Available online at www.sciencedirect.com
ScienceDirect
Gait & Posture 25 (2007) 140–152
GAIT & POSTURE
www.elsevier.com/locate/gaitpost

Review

Gait classification in children with cerebral palsy: A systematic review

Fiona Dobson [a,b,*], Meg E. Morris [b,c], Richard Baker [a], H. Kerr Graham [a,b,d]

[a] Hugh Williamson Gait Laboratory & Murdoch Childrens Research Institute, Royal Children's Hospital, Flemington Road, Parkville, Vic. 3052, Australia
[b] University of Melbourne, Vic. 3010, Australia
[c] RASP Program, Southernhealth, Vic., Australia
[d] Department of Orthopaedics, Royal Children's Hospital, Vic. 3052, Australia

Received 26 August 2005; received in revised form 20 December 2005; accepted 11 January 2006

Abstract

This systematic review of the literature evaluates the validity of existing classifications of gait deviations in children with cerebral palsy (CP). Numerous efforts have been made to develop classification systems for gait in CP to assist in diagnosis, clinical decision-making and communication. The internal and external validity of gait classifications in 18 studies were examined, including their sampling methods, content validity, construct validity, reliability and clinical utility. Half of the studies used qualitative pattern recognition to construct the gait classification and the remainder used statistical techniques such as cluster analysis. Few adequately defined their samples or sampling methods. Most classifications were constructed using only sagittal plane gait data. Many did not provide adequate guidelines or evidence of reliability and validity of the classification system. No single classification addressed the full magnitude or range of gait deviations in children with CP. Although gait classification in CP can be useful in clinical and research settings, the methodological limitations of many classifications restrict their clinical and research applicability.

Keywords: Cerebral palsy; Gait patterns; Classification system; Validity; Systematic review

Contents

1. Introduction . . . 141
2. Method . . . 141
   2.1. Search strategy . . . 141
   2.2. Inclusion criteria . . . 141
   2.3. Data extraction . . . 141
   2.4. Quality assessment . . . 142
3. Results . . . 142
   3.1. Search strategy yield . . . 142
   3.2. Descriptive aspects of reviewed studies . . . 142
   3.3. Quality aspects of reviewed studies . . . 146
   3.4. Sample definition and selection . . . 146
   3.5. Classification construction . . . 146
   3.6. Psychometric properties . . . 146

* Corresponding author. Tel.: +61 3 9345 5354; fax: +61 3 9345 5447.
E-mail address: fionadobson@aanet.com.au (F. Dobson).

0966-6362/$ – see front matter
doi:10.1016/j.gaitpost.2006.01.003

4. Discussion . . . 146
4.1. Sample definition . . . 146
4.2. Sampling method . . . 147
4.3. Classification construction . . . 147
4.4. Psychometric properties . . . 149
5. Conclusions. . . . 150
Acknowledgements . . . 151
References . . . 151

1. Introduction

One of the most striking features of cerebral palsy (CP) is the variability of its clinical presentation [1]. The diversity of gait deviations observed in children with CP has led to repeated efforts to develop gait classification systems to assist in diagnosis, clinical decision-making and communication [2–19]. Gait classifications may enable clinicians to differentiate gait patterns into clinically significant categories that assist with clinical decision-making. This is possible when the underlying variables that define the classification groups can be modified by intervention. Gait classifications can also provide clinicians and researchers with a common language that quickly conveys a "clinical snapshot" of an individual's gait impairments. The term "gait classification" refers to a system that allows the allocation of gait patterns into groups that can be differentiated from one another based on a set of defined variables. This is distinct from gait indexes, assessment scores and scales, which score individual gait variables or provide an overall index to quantify deviations from normal gait without group allocation.

In order to understand the range and quality of existing gait classifications for children with CP, a systematic review of the available literature was performed. This was considered to be an essential step before the development of a new gait classification. The review was conducted to highlight both the positive qualities and limitations of previous classification systems and to provide evidence as to whether existing systems were adequate, required modification or required re-conceptualisation. The main aim was to critically evaluate the internal and external validity of existing gait classifications by critically appraising their design, sampling methods, content, construct, psychometric properties and clinical utility.

2. Method

2.1. Search strategy

An electronic search of the following publication databases was performed in March 2005: MEDLINE (1966–March 2005), Embase (1988–Week 10, 2005), Current Contents (1993–Week 10, 2005), AMED (1985–March 2005), CINAHL (1982–March 2005), Health and Psychosocial Instruments (1985–December 2004), PsycINFO (1967–March 2005), Inspec (1987–Week 9, 2005) and Recal Bibliographic Database (1991–2005). Keywords used in the search strategy included cerebral palsy, gait, gait analysis, classification, gait pattern, observation and cluster analysis. Key terms were matched to MeSH subject headings and exploded. Relevant truncation or wildcard symbols were used to retrieve all possible suffix variations of a root word. Targeted searching was conducted to help identify literature that could be missed by electronic database searching. This included an online search of journals that were likely to contain target articles and a manual search of the cited references from each relevant study identified for the review.

2.2. Inclusion criteria

The title and abstract of each study identified from the search strategies were assessed. They were included in the review when they satisfied the following criteria: (1) studies that involved the classification of gait impairment(s)/deviation(s) where they described or allocated gait variables into categories, groups or clusters. This could include kinematic, temporal–spatial kinetic or electromyographic data; (2) full papers or abstracts (not later published as full papers) published between 1966 and March 2005; (3) the participants were predominantly aged 0–18 years with a diagnosis of CP. If a comparatively small part of the sample contained a few older subjects outside the inclusion age range, then these studies were considered eligible. Excluded from the review were studies of outcome measurement scales or indices (for example, the Normalcy Index [20,21]) that score gait parameters without allocating them into categories. Classifications of functional mobility, such as the Functional Mobility Scale [22] or gross motor function, such as the GMFCS [23], where gait is only a component of the functional activity were also excluded. This was because the focus of this review was identifying classifications of gait patterns rather than functional activities.

2.3. Data extraction

Key details of each study were extracted using a customised data extraction form. Two independent reviewers (FD and MM) piloted the form to check the form content and to ensure reliable data extraction. The major themes for data extraction were: subject characteristics; subject selection and recruitment; measurement tool and type of gait variables measured; classification construction including method, scale type and number of classifications levels; and psychometric properties including reliability, validity, sensitivity, specificity and clinical utility. These themes were chosen to gain a comprehensive account of each study in order to synthesise the studies and to assess the quality of each study.

2.4. Quality assessment

In contemporary systematic reviews, quality assessments are provided in addition to the standard qualitative review [24]. Quality assessment was performed on a subset of the data extracted to assess both internal and external validity of the studies. As no standardised or validated checklists were available for this type of review, a customised quality checklist was developed using a number of sources with similar quality themes assessed in this review. These included generic systematic review principles [24–26], diagnostic guidelines [27], other systematic reviews of gait measures [28–30], other systematic reviews of rating scales or questionnaires [31–34] and the International Classification of Function, Disability and Health (ICF) guidelines on classification construction [35]. A scored checklist was not used because the validity of using a scoring system in a customised quality checklist has not been determined. The major themes for the quality checklist were external validity including sampling method and subject characteristics and internal validity including classification construction and scale psychometric properties. Two reviewers (FD and MM) used the checklist independently to evaluate the quality of each study. Any discrepancies on quality criteria were checked using the original article to ascertain the correct coding based on objective a priori decision rules (Table 1), according to previous recommendations [36–38]. For example, if one reviewer found 'Limited' inclusion/exclusion criteria (i.e. only one or two points listed) and the other found 'Not stated', then the article was re-read to determine the correct coding. Table 1 outlines the questions and the a priori decision rules used to complete the quality checklist.

Table 1
Assessment Checklist Questions

Questions	Response/decision rule criteria
Sample characteristics	
Are the subject characteristics defined adequately? (include: motor type, number, age, gender)	Adequate—all details provide Partial—all details except gender Inadequate—missing details
What sampling method is used?	List type: convenience, case consecutive Population-based Not stated
Are clear inclusion/exclusion criteria stated?	Stated—clear list of both provided Limited—one or two points only Not stated—not provided
Classification construction	
What planes of motion are used in the classification? (content validity)	List: sagittal, coronal, transverse
Is the classification construction method clearly defined?	Adequate—clear definition provided Partial—limited/unclear Inadequate—not defined
Were the set of variables defining the levels/groups on the classification clearly defined?	Adequate—clear details provided Partially—limited, unclear details Inadequate—not defined
Psychometric properties	
Was the reliability of t he classification tested?	Not reported List: inter-rater, intra-rater
Was internal validity of the classification groups demonstrated?	List: construct, criterion, discriminant
Was the utility of the tool described?	Described Limited description Not described

3. Results

3.1. Search strategy yield

The electronic search of selected databases identified 333 published studies. A further five studies were identified from the targeted journal/proceedings search and a further two studies were found from the hand-search giving a yield of 340 publications. Following the application of a priori inclusion/exclusion criteria (Section 2.2), two reviewers (FD and MM) identified 18 studies for inclusion in the systematic review. The details of these studies are summarised in Table 2.

3.2. Descriptive aspects of reviewed studies

Subject characteristics varied across the studies. Seven analysed children with diplegic CP, five with hemiplegic CP and six analysed a mixture of CP types. The sample sizes ranged from 15 to 588 subjects with a median of 46 subjects. Half the studies reported the age range of subjects, which varied between 2 and 30.5 years of age. In three studies [8,15,16], the age range of the sample exceeded 18 years of age, however the majority of participants were within the inclusion age range of 0–18 years.

Three-dimensional gait analysis was used in 15 studies and video observational analysis was used in one study. The majority of studies (15/18) analysed kinematic data to form

Table 2
Identified studies of classification of gait in children with cerebral palsy

Study	Subjects type; number; age range (years)	Measurement tool used	Type of gait variables used in classification	Scale type	Number of groups	Anatomical level(s) classified	Method of classification construction	Reliability test results
Simon et al. [13]	SH, SD; 15; 4–16	3-DGA (Vanguard) Surface EMG	T/S, kinematic, kinetic, EMG	Nominal	3	Knee	Qualitative pattern recognition using quantitative data	Not reported
Wong et al. [16]	SH, SD, SQ; 86; 3–24	3-DGA (3 camera)	T/S, kinematic	Nominal	4	Ankle, knee, hip	Quantitative Cluster analysis *k*th nearest neighbour	Not reported
Winters et al. [15]	SH, other; 46; 3.5–30.5	3-DGA Surface EMG	T/S speed, kinematic	Ordinal	4	Ankle, knee, hip, pelvic	Qualitative pattern recognition using quantitative data	Not reported
Kadaba et al. [4][a]	SD; 30; –	3-DGA	Kinematic, kinetics	Ordinal	10	Ankle, knee, hip	Quantitative with arbitrary selection Principal component analysis Cluster analysis (unspecified)	Not reported
Sutherland and Davids [14]	SD; 588; –	3-DGA	Kinematic, EMG, force plate	Ordinal	4	Knee	Qualitative pattern recognition using quantitative data	Not reported
Stout et al. [12][a]	SH; 83; –	3-DGA	Kinematic, kinetic, EMG, energy (O_2 cost)	Ordinal	5	Ankle, knee, hip	Quantitative and subjective nterpretation Cluster analysis (unspecified) One-way ANOVA	Not reported
Hullin et al. [3]	SH; 26; –	3-DGA	Kinematic, kinetic, EMG	Ordinal	5	Ankle, knee, hip, pelvis	Qualitative pattern recognition using quantitative data	Not reported
O'Malley et al. [8]	SD; 88; 2–20	3-DGA (Vicon)	Stride length, cadence, leg length, age	Ratio	5	N/A	Quantitative Fuzzy *k*-means cluster analysis	Not reported
O'Byrne et al. [7]	SH, SD; 146; –	3-DGA (CODA-3)	Kinematic	Nominal	8	Ankle, knee, hip	Quantitative using subjective interpretation *k*-Means cluster analysis	Not reported
Wong et al. [17]	SD, SQ; 42; 3–17	DynoGraph load sensors	Temporal–spatial, foot loading forces	Ordinal	2 × 4	Foot	Qualitative pattern recognition using quantitative data	Not reported

Table 2 (*Continued*)

Study	Subjects type; number; age range (years)	Measurement tool used	Type of gait variables used in classification	Scale type	Number of groups	Anatomical level(s) classified	Method of classification construction	Reliability test results
Kienast et al. [5][a]	SD; 24; –	3-DGA (Vicon)	Kinematic	Nominal	3	Ankle, knee, hip	Quantitative *k*th nearest neighbour cluster analysis	Not reported
Lin et al. [6]	SD; 23; mean 9	3-DGA Expert Vision	Kinematic, kinetic, EMG	Nominal	4	Knee	Qualitative pattern recognition using quantitative data	Not reported
Yokochi [18]	SD; 20; 6–16	Video observations	Kinematic stability	Nominal	5	Ankle, knee	Qualitative pattern recognition	Not reported
Stebbins et al. [11][a]	SH; 39; –	Surface EMG	EMG of Tib Ant	Ordinal	3	Tib Ant	Qualitative pattern recognition using quantitative data	Inter-rater
			Gasroc muscles		5	Gastroc		Three raters 67–100% agreement
Carollo et al. [2][a]	Not stated; 155; 4.4–18.2	3-DGA Vicon	Kinematic	Nominal	5	Ankle, knee	Quantitative using subjective interpretation Hidden Markov Models	Not reported
Rodda et al. [9]	SD; 187; 4–18	3-DGA Vicon	Kinematic	Ordinal	6	Ankle, knee, hip, pelvis	Qualitative pattern ecognition using uantitative data	Intra-rater (wκ 0.89) and inter-rater (wκ 0.6–0.74) Six raters
Zwick et al. [19]	SD, SH; 66; 5–17	3-DGA Vicon	Kinematic, kinetic, EMG	Nominal	2	Calf muscle	Quantitative using subjective interpretation Generalised dynamic neural network	Not reported
Salazar et al. [10]	SH; 206; –	3-DGA Vicon	Kinematic	Ordinal	4	Ankle, knee, hip, pelvis	Quantitative using subjective interpretation Support Vector Machine	Not reported

Symbol (–) denotes missing data; SH: spastic hemiplegia; SD: spastic diplegia; SQ: spastic quadriplegia; 3-DGA: three-dimensional gait analysis; EMG: electromyography; G: gastrocnemius muscle; TA: tibialis anterior muscle; T/S: temporal–spatial; ‘other’ includes traumatic brain injury and juvenile cerebrovascular accident.

[a] Abstract only

Table 3
Quality assessment of studies of classification of gait in children with cerebral palsy

Study	Sample characteristics			Classification construction			Psychometric properties		
	Subject definition	Inclusion criteria	Sampling method	Planes of motion	Construction method	Classification groups	Reliability	Validity	Utility
Simon et al. [13]	Adequate	Stated	Not stated	Sagittal, coronal	Not stated	Adequately defined	Not reported		Limited description
Wong et al. [16]	Partially	Not stated	Not stated	Sagittal, coronal, transverse	Adequately defined	Adequately defined	Not reported	Discriminant, construct	Limited description
Winters et al. [15]	Adequate	Stated	Not stated	Sagittal	Not stated	Adequately defined	Not reported		Described
Kadaba et al. [4][a]	Inadequate	Stated	Not stated	Sagittal	Partially defined	Inadequately defined	Not reported		Described
Sutherland and Davids [14]	Inadequate	Not stated	Not stated	Sagittal	Not stated	Inadequately defined	Not reported		Limited description
Stout et al. [12][a]	Inadequate	Limited	Not stated	Sagittal	Partially defined	Partially defined	Not reported	Construct	Limited description
Hullin et al. [3]	Inadequate	Not stated	Not stated	Sagittal	Not stated	Adequately defined	Not reported		Not described
O'Malley et al. [8]	Inadequate	Limited	Not stated	N/A	Adequately defined	Adequately defined	Not reported	Discriminant, construct, cross-validation	Described
O'Byrne et al. [7]	Inadequate	Stated	Convenience retrospective 2-year period	Sagittal	Adequately defined	Inadequately defined	Not reported	Construct	Described
Wong et al. [17]	Adequate	Stated	Not stated	N/A	Inadequately defined	Partially defined	Not reported	Concurrent	Limited description
Kienast et al. [5][a]	Inadequate	Not stated	Not stated	Sagittal	Adequately defined	Adequately defined	Not reported		Limited description
Lin et al. [6]	Adequate	Limited	Not stated	Sagittal	Inadequately defined	Inadequately defined	Not reported		Limited description
Yokochi [18]	Adequate	Stated	Convenience	Sagittal, coronal	Inadequately defined	Partially defined	Not reported		Not described
Stebbins et al. [11][a]	Inadequate	Not stated	Not stated	N/A	Inadequately defined	Inadequately defined	Inter-rater		Not described
Carollo et al. [2][a]	Inadequate	Not stated	Convenience retrospective	Sagittal	Adequately defined	Inadequately defined	Not reported	Discriminant, construct, cross-validation	Limited description
Rodda et al. [9]	Partially	Stated	Not stated	Sagittal	Inadequately defined	Adequately defined	Inter-rater, intra-rater		Described
Zwick et al. [19]	Partially	Stated	Convenience retrospective 3-year period	Sagittal	Adequately defined	Adequately defined	Not reported	Construct, criterion, discriminant	Described
Salazar et al. [10]	Inadequate	Not stated	Convenience retrospective	Sagittal	Adequately defined	Adequately defined	Not reported	Construct, discriminant, cross-validation	Described

[a] Abstract only.

the classification and five studies used kinetic information to supplement kinematics. EMG data were incorporated in six studies and two studies used temporospatial data.

Classifications were constructed using a nominal scale in eight studies, an ordinal scale in nine studies and a ratio scale in one study. The median number of classification groups was 5 and ranged from 2 to 10. Four studies incorporated four anatomical levels (ankle, knee, hip and pelvis) into the classification, five studies incorporated three anatomical levels (ankle, knee and hip), four studies were based on two anatomical levels (ankle and knee) and a further four studies were based on only one anatomical level. In one study, the number of anatomical levels was not applicable as the classification was constructed using stride length and cadence [8].

A qualitative method of construction for the classification was used in half the studies, with 8/9 studies incorporating quantitative data into subjective pattern recognition to form the classification. The remaining studies used quantitative statistical construction techniques with seven studies using a variation of cluster analysis techniques including *k*th nearest neighbour, *k*-means, fuzzy *k*-means and generalised dynamic neural networks. The remaining two studies incorporated less common quantitative approaches including Hidden Markov Models and the Support Vector Machine.

3.3. Quality aspects of reviewed studies

As this review did not examine clinical interventions, meta-analysis was not appropriate. Evidence was therefore limited to a qualitative process of ‘‘best evidence synthesis’’. The quality assessment decisions for each study are outlined in Table 3 under three major themes: study sample (external validity), classification construction and psychometric properties (internal validity).

3.4. Sample definition and selection

Adequate sample definition, including the motor type, topographical distribution, number, age range and gender of participants was provided in only 5/18 studies. Nine studies did not report the age range of subjects and one study did not report the type of CP. One study included a small number of subjects with a diagnosis of traumatic brain injury and juvenile cerebrovascular accident as part of the sample [15]. The study inclusion/exclusion criteria were clearly stated in fewer than half of the studies. The sampling method was not stated in the majority of studies (13/18). The five studies that stated a sampling method used a retrospective sample of convenience.

3.5. Classification construction

Classifications of gait were based on analyses in the sagittal plane in 12 studies. Only one study used gait variables in all three planes of motion (sagittal, coronal and transverse). A clear definition of the construction method and process was provided in only seven studies and half of the studies did not state the method and process. Likewise, details on the set of variables defining the levels/groups on the classification were clearly provided in only half of the studies.

3.6. Psychometric properties

The reliability of the classification constructed in each study was assessed in only two studies [9,11]. Construct validity was demonstrated in 7/18 studies and discriminant validity was demonstrated in five studies. All used a quantitative method for the construction of the classification. Criterion validity was addressed in one study and concurrent validity in one other. The utility or description of how the classification was being used or could be used in clinical practice or research were described in seven studies.

4. Discussion

4.1. Sample definition

Adequate sample definitions and inclusion criteria were provided in only 4 of the 18 publications included in the evaluation [13,15,17,18]. None of the studies provided information about the functional severity of participants. It was difficult to determine if the classifications from studies without adequate definitions applied to a wide spread of individuals of varying severity across the childhood and adolescent years or only to a select age group and severity.

Adequate documentation of the clinical and demographic characteristics of the sample for each study, including the inclusion criteria, is necessary to enable the generalisation of the study findings to other populations. Information concerning previous spasticity management (such as botulinum toxin, selective dorsal rhizotomy and intrathecal baclofen) and orthopaedic surgical interventions of the lower limbs is required for the interpretation of the classification. For example, a participant with a history of previous orthopaedic surgery, such as a tendo-achilles lengthening, may acquire a variant gait pattern, such as a crouch gait, due to the effects of surgery. This variant pattern may not be typical of the naturally occurring gait pattern range and therefore may not be relevant when applying the classification to a population without previous surgery. The majority of studies assessed in this review did not provide adequate sample definition, limiting the generalisability of the findings to other populations. This is consistent with the limited sample descriptions found in other systematic reviews of observational gait analysis used in clinical practice [29].

4.2. Sampling method

For the development of a meaningful gait classification, a representative sample is required. A representative sample ensures that the range of characteristics of interest in a population will most likely be present in the sample [39]. The sampling method was not stated in over two-thirds of the studies (see Table 3). It is likely that these studies used a sample of convenience because this type of sampling involves the selection of the most accessible members of the target population [40]. Five studies [2,7,10,18,19] reported using a retrospective sample of convenience chosen from children who had already attended a health service such as a gait laboratory or orthopaedic outpatient unit. The main limitation of this sampling method is that the sample may be biased and not representative of the entire population of interest [40]. These samples were more likely to contain a skewed representation towards the more extreme gait deviations and an underestimation of milder impairments. They were also less likely to contain the range of possible gait deviations that is present in a certain population, which limits the population validity of these studies.

As summarised in Table 3, several studies used small sample sizes with six having samples of 30 or less [3–6,13,18]. Using small samples of convenience also limits the population validity of these studies because it is unlikely that the complete range of gait characteristics from the entire population of interest will be present in these samples. This selection bias leads to difficulties in the quantification of the incidence and magnitude of gait deviations in the population of interest. These small samples become even more problematic when they are classified into a large number of groups. For example, Simon et al. [13] classified only 15 subjects into three groups. Kadaba et al. [4] used a slightly larger sample of 30 subjects, however then classified these subjects into 10 groups.

4.3. Classification construction

Despite most studies using sophisticated three-dimensional measurement systems to collect gait data in three planes of motion, most classifications were constructed using only sagittal plane data. This is consistent with many classifications of CP gait that are based on observational gait analysis [41–45]. This may be due to the increased ease at which deviations can be reliably visualised in the sagittal plane [45] and because deviations in the coronal and transverse plane are more difficult to interpret visually [46]. Classifications based on sagittal plane data may limit their content validity and restrict their application because typical deviations in other planes of motion were not captured or considered. Although many gait deviations observed and treated in children with CP occur in the sagittal plane [47], deviations in the transverse and coronal planes are also considered important in clinical decision-making and intervention planning [48–50]. Decisions to proceed to some of the more invasive surgical interventions, such as osteotomies of the femur and tibia, rely heavily on gait information in the transverse plane [47,50]. For example, an in-toeing gait pattern may result from internal rotation of the femur, tibia or both. If the pelvis is retracted (externally rotated) on the same side the femur is internally rotated, the femur may appear to be neutrally rotated (i.e. coming straight at the observer when viewed from the front). Correct identification and quantification of these transverse plane deviations can influence decisions on the amount and type of surgical intervention.

The method in which the classification was constructed was coded as either *qualitative* or *quantitative*. *Qualitative* construction methods included those where decisions to group members relied on the judgment and experience of those making the decisions. These judgements could vary in different clinicians leading to different interpretations of the classification groups. The types of gait variables chosen to make the decision could also rely on individual judgment and could vary from study to study. This could increase the likelihood of different judgements and decisions being made by different clinicians. A potential advantage of qualitative methods is enhanced clinical relevance as groups are partitioned using clinical reasoning and rely on identifiable patterns of movement observed by clinicians. Therefore, they may be more meaningful in terms of clinical diagnosis and treatment planning.

For the nine studies that used a qualitative method, none provided an adequate description or justification of the process that was undertaken to determine the groups defined in the classification. They also lacked information on which specific gait variables were chosen for analysis, why they were chosen and who chose them. It was not clear if any method, such as an expert consensus group, was used to guide the classification construction process, or if the decisions were the responsibility of the investigators. For example, in one study [14], there was no indication of how or why subjects with diplegic CP were allocated into the four groups of knee patterns. There was no indication of either mean scores or range scores for each gait variable used to describe the groups. Although this study contained the largest cohort of subjects, it was not stated how many of the 588 subjects were represented in each group making it impossible to judge the likely prevalence of different knee patterns in this large population of children with diplegic CP. The lack of transparency of the construction process in the qualitative studies limits the reproducibility and clinical interpretation of these construction techniques.

Inadequate descriptions of classification construction methods used in the qualitative studies also makes it is difficult to determine the most likely number and type of gait patterns to be represented in a population. This is highlighted in children with diplegic CP. Sutherland and Davids [14] and Lin et al. [6] both reported four groups of sagittal plane knee patterns using samples of 588 and 23 subjects, respectively. Three of the four groups described were similar for both studies. The differences found in these two studies may be due to different study populations, including different sample sizes, sampling methods (not stated in either study) or subject characteristics (inadequately defined by Sutherland and Davids). It may also have been due to the additional information obtained by kinetic data included in the study by Lin et al. [6] or the failure by both studies to adequately describe the classification construction method.

More recently, Rodda et al. [9] described six groups of gait patterns in a sample of 187 children with spastic diplegia. This classification was based on a hierarchical analysis of the ankle followed by that of the knee, hip and pelvis and was described on an ordinal scale. The groups described in this study had characteristics in common with the groups described by Sutherland and Davids [14] and Lin et al. [6], however unlike these studies, no 'recurvatum gait' category was described by Rodda et al. [9]. This type of pattern would have required previous lengthening of the hamstring muscles, which was an a priori exclusion in the study by Rodda and coworkers. Differences in the number of groups and type of groups found in these three studies may have been due to different study populations and classification construction methods. Because sample definitions were limited, sampling methods were not stated and construction methods were inadequately defined in all these studies, it is difficult to determine the impact that these factors had on the discrepancies found in the number and type of groups.

Determination of the most likely number and type of gait patterns represented in children with hemiplegic CP was also limited due to inadequate definitions of classification constructions. Winters et al. [15] found four groups of gait patterns based on the sagittal plane kinematics of the ankle, knee, hip and pelvis in a sample of 46 children with spastic hemiplegic CP. They suggested progressive pathological involvement of anatomical levels with each group and pathology in a specific muscle–tendon unit or joint at each level is responsible for each gait pattern. Although appealing for clinical decision-making, no statistical evidence was provided that the groups were distinct from each other. Based on this same classification, Hullin et al. [3] subsequently found five groups of gait patterns from 26 children with spastic hemiplegia. Additional information from kinetic data was used to determine the groups in this classification and may explain some differences found. Using electromyographic data from the gastrocnemius and tibialis anterior muscles, Stebbins et al. [11] reported eight different gait patterns in children with spastic hemiplegia. Lack of transparency of the construction process, sampling method and sample characteristics makes it difficult to determine the real number of naturally occurring gait groups in children with hemiplegic CP.

When *quantitative* methods for classification construction were used, all except one study [2] provided adequate definition of the construction method and process. The majority of the studies that used quantitative methods used a variation of cluster analysis. Cluster analysis is a generic term for a wide range of statistical techniques used to examine multivariate data in order to uncover groups or clusters of homogenous observations [51]. The purpose of cluster analysis is to explore data sets to assess whether they contain a number of subgroups of objects that resemble each other and which are different from objects in other subgroups [51]. The advantage of this quantitative method is that a systematic and structured approach based on objective data from the sample can be conducted. It is useful for detecting groups or clusters that cannot be easily identified by visual inspection. It can be applied to data where no a priori structure of the data or clustering solution is known.

The limitations of cluster analysis include that the process may uncover "artificial" groups that have no clinical relevance or meaning. Cluster solutions may impose a structure on the data set rather than finding naturally occurring groups. Therefore, it is possible to artificially divide a homogenous group or fail to partition a heterogenous group using this method. Cluster analysis also requires that a number of a priori rules or choices are set. These choices are set by the user, require subjective judgment and have a direct influence on the outcome of the clustering solution. For gait classifications, the investigator needs to choose the gait variables for the statistical analysis, determine when one group of subjects with a type of gait pattern become sufficiently similar to call them a homogenous group, and alternatively, determine when two groups of subjects with different types of gait patterns are sufficiently dissimilar to call them distinct groups. It is at the discretion of the investigator to choose the ultimate number of clusters. Although this has the potential to enhance clinical relevance of the final solution, the answers to clustering solutions are not entirely objective. The final solution, including the number of groups and the structure of the groups can be influenced by the rules set. Unless the investigator is clear about the rules set in the construction phase, it may be impossible to replicate the findings or use the findings with confidence.

Nearest neighbour cluster techniques were employed by two studies in the review [5,16]. These hierarchical type clustering methods are designed to detect high-density clusters [52]. Using a sample of 86 children with spastic diplegia, hemiplegia and quadriplegia, Wong et al. [16] chose a four-cluster solution despite confirming that one cluster was most likely a group of outliers. Use of a four-cluster gait classification to classify many types of CP severely limits the applicability of this classification to be used in a meaningful way within the clinical environment. Additionally, only gait variables of the right leg of each participant were used to construct the classification. Considering that the sample included children with hemiplegia, the classification may have been based on either the affected or unaffected limb of these children.

This undermines the validity of this classification. In the study by Kienast et al. [5], three clusters of gait patterns were found using high-density clusters. Because nearest neighbour cluster analysis relies on the use of a representative sample from the population of interest, a sample of convenience of 24 subjects with spastic diplegia is unlikely to provide a true representative sample from which subpopulations can be detected. This sampling strategy not only undermines the external validity of the findings but it also limits the construct and discriminant validity of the classification system.

k-Means cluster analysis was utilised in one study [7]. This type of cluster analysis relies on an optimisation algorithm to find the optimal partition of the data set into a pre-determined number of groups. In a sample of 146 children with either hemiplegic or diplegic CP, O'Byrne et al. [7] chose an eight-group cluster solution to classify 237 limbs based on sagittal plane kinematics. Kinematic values including means, range and standard deviations were provided for six gait variables featured in the descriptions of each pattern but there was no explanation of who or why these variables were chosen. As *k*-means cluster analysis requires that the number of final groups in the classification is pre-set, a clear rationale for how this is done is essential to determine the validity of this method. No clear rationale for choosing the final number of groups was provided.

Fuzzy clustering analysis was used in one study [8]. In this type of analysis, subjects are not assigned to a single cluster but have differing degrees of strength of membership in some or all clusters [53]. This allows for the classification of impairments that fall along a continuum rather than into well-defined groups. None of the qualitative studies reviewed had considered this issue. Fuzzy cluster analysis requires the investigator to pre-select the number of subsets there are likely in a given data set. In the study by O'Malley et al. [8], which used a sample of 88 children with spastic diplegic cerebral palsy, a clear and valid explanation was given for the final five subgroups chosen. Although this study primarily considered gait function rather than gait patterns, it was the only study to provide transparent information on the decision rules that were set in the analysis. Unfortunately, inadequate reporting of the sample characteristics may limit the generalisation of this classification to other populations of interest.

A number of less common approaches were also used to construct classifications. These included a generalised dynamic neural network [19], Hidden Markov Models [2] and the 'Support Vector Machine' [10]. These types of techniques or models use computing algorithms that attempt to imitate and automate complex pattern recognition tasks much like the human brain [51]. One of the major limitations of these types of models is that their accuracy is dependent on large sets of training data. The more training data, the more responsive, sensitive and specific they become. Sufficient training data may not always be available for the development of this accuracy in some populations. The utility of these approaches is yet to be embraced in the clinical environment.

Following the construction of a classification system, an adequate definition or description of the resultant groups is required so they can be used in a meaningful and useful way within the clinical setting. It is helpful if a summary of the identifying features of each group is clearly described to assist with the applicability of the classification. Only half of the studies provided such detail (refer to Table 3). Many studies did not report descriptive statistics, such as means, ranges or variances for each of the gait variables used to define the classification groups. Some studies [3,13,15,17,18,54] provided an illustrative representation of the distinguishing characteristics of each gait pattern, which may assist in the user-friendliness of the classification system.

4.4. Psychometric properties

Essential to the validity of any measurement tool is that it be reliable. The reliability of the classifications was assessed in only two studies. Using a weighted kappa (wκ) statistic, Rodda et al. [9] reported high intra-rater reliability (wκ 0.86) and moderate inter-rater reliability (wκ 0.60–0.74) for six experienced clinicians over two testing sessions. Inter-rater agreement was reported to be substantially lower on the second testing occasion (wκ 0.60, 95% CI 0.36, 0.83) when compared to the first (wκ 0.74, 95% CI 0.58, 0.95) for all groups on the classification except Group IV—''crouch gait''. The authors reported that on each occasion the raters were provided with written instructions and illustrations for each gait pattern on the classification however provided no explanation for the difference found between testing occasions.

Stebbins et al. [11] reported that inter-rater agreement between three clinicians ranged between 67% and 100%. The limitation of reporting raw percent agreement alone is that chance agreement is not taken into consideration. Stebbins and coworkers also reported the inter-rater reliability of another gait classification included in the current review. Using the Winters et al. [15] classification, 39 gait patterns from children with hemiplegic CP were rated by two independent assessors. The agreement between the two assessors was reported as 55% raw agreement. This raw agreement is very low when taking into consideration the high possibility of chance agreement of the two raters on the four-point scale of the Winters classification system.

Although reliability is fundamental to the clinical usefulness of gait classifications, no other studies reported reliability as part of the initial classification development or as a follow-up step. Further testing of the reliability of other classifications reported in this review is required before definitive conclusions can be made.

High reliability of a measurement tool does not guarantee high validity of that tool. Using gait classifications, clinicians can have good intra-rater and inter-rater agreement but be consistently wrong about their observations or assignment of gait patterns to a particular group. Broadly, a classification system can be described as valid if it consistently and accurately measures what it is intended to measure [55].

Studies that demonstrated construct and discriminant validity (refer to Table 3) all used a quantitative method for the construction of the classification. Due to inadequate reporting and low level of methodological quality of studies using qualitative construction methods, construct and discriminant validity was either limited or difficult to determine. Increased transparency and methodological rigour is required in order for definitive conclusions to be made.

In one study using qualitative construction methods, Wong et al. [17] examined the correlation of their classification systems with Minear's classification of daily activity limitation, which were all measured on a 4-point ordinal scale. Pearson correlations were selected to examine this concurrent validity, however justification for using parametric techniques for ordinal data was not provided.

Cross-validation attempts for the classifications were conducted by four studies [2,8,10,16]. The purpose of this cross-validation process was to assess the precision of the classification system by using a test set of data that was not from the original set of data used to construct the classification. Wong et al. [16] used 42 of the 128 subjects as test cases to cross-validate the classification system. Because the testing was applied to a subsequent classification scheme based on a partitioning decision rule from the clustering method, evaluation of the sensitivity of the original classification clusters was not possible. O'Malley et al. [8] performed a cross-validation process on their classification system with just four cases and found that two neurologically intact subjects of different ages and genders both fell very close to the control group cluster centre featured in their fuzzy clustering system. Additionally, improved membership changes were noted for two children with spastic diplegia who had improved postoperative functional mobility. The small number of cross-validation cases in this study limits the strength of this evidence and further cross-validation with a greater number of test cases is warranted. Carollo et al. [2] retained 122 of the original 261 subjects as test cases and reported a classification recognition rate of 81.1%. Recognition for two groups described in this classification ('apparent equinus' and 'crouch') were 5–20% lower than the other three groups and not all limbs from all 155 subjects could be classified using the five classification groups. This limits the clinical usefulness of this classification system. Salazar et al. [10] reserved 52 cases from 206 cases for testing and found high sensitivity (82.8% and 100%) and high specificity (84.6% and 98.7%). A limitation of this model was that gait patterns used for training data were pre-classified using the Winters et al. [15] classification. Details on how this was conducted, including who pre-classified the gait patterns, were not provided.

The impact of the classification on clinical practice, research or education was not always reported. Winters et al. [15] discussed how their classification could guide treatment for each group on the classification. They also speculated about interventions that would not be suitable for certain classification levels [15]. The Winters classification system is one of the most cited gait classifications in cerebral palsy and has been used in many outcome studies to define the study population [50,56–60]. It has also been used by other researchers as the basis for developing new classifications [3,10–12,54]. As such, the limitations of construct validity and internal validation of these classification systems needs to be considered.

Kadaba et al. [4] described how their classification could be used to separate primary from secondary gait deviations and how this could be useful in the surgical management of children with diplegia. Similarly, Rodda et al. [9] described how their classification could be used in clinical decision-making process, including surgical interventions and physiotherapy practice although no follow-up studies have yet been reported. O'Byrne et al. [7] and Zwick et al. [19] outlined how their classification system could be used as a diagnostic tool and Salazar et al. [10] described how a Support Vector Machine (SVM) might assist computer aided gait pattern recognition in children with spastic hemiplegia. The nature of the SVM equipment and analysis used in these systems limits the availability and accessibility for everyday use by clinicians.

5. Conclusions

Although CP gait classifications are used for diagnostic purposes, to streamline communication and to facilitate clinical decision-making, the overall methodological quality of the studies evaluated in this systematic review was low. No single classification system appeared to reliably and validly describe the full range of gait deviations in children with CP. Many classifications appeared to use arbitrary decisions to allocate patients into groups as opposed to using clinical decision-making principles to allocate them to clinically meaningful categories. Furthermore, evidence on how classification groups were related to specific impairments, such as muscle contracture, was limited. Future CP gait classifications need to be constructed around clinically meaningful categories and may benefit from incorporation of large, prospective population-based cohorts. Analyses in coronal and transverse planes of motion in addition to the sagittal plane could improve content validity. It is recommended that new gait classifications demonstrate evidence of reliability and validity, which could generate greater confidence in the clinical and research applicability of these tools.

Acknowledgments

This work was supported by the CCRE Clinical Gait Analysis and Gait Rehabilitation NHMRC grant. Fiona Dobson received support from the Sir Robert Menzies Allied Health Sciences Scholarship and the Faculty of Health Sciences, La Trobe University.

References

[1] Liptak GS, Accardo PJ. Health and social outcomes of children with cerebral palsy. J Pediatr 2004;145:S36–41.
[2] Carollo JJ, He Q, Debrunner C. Gait pattern classification in children with cerebral palsy: results from a 12-state hidden Markov model created with a large training set. Gait Posture 2004;S20:S4 [abstract].
[3] Hullin MG, Robb JE, Loudon IR. Gait patterns in children with hemiplegic spastic cerebral palsy. J Pediatr Orthop Part B 1996;5: 247–51.
[4] Kadaba MP, Ramakrishnan HK, Jacobs D, Chambers C, Scarborough N, Goode B. Gait pattern recognition in spastic diplegia. Dev Med Child Neurol 1991;S33:28 [abstract].
[5] Kienast G, Bachmann D, Steinwender G, Zwick EB, Saraph V. Determination of gait patterns in children with cerebral palsy using cluster analysis. Gait Posture 1999;10:57 [abstract].
[6] Lin CJ, Guo LY, Su FC, Chou YL, Cherng RJ. Common abnormal kinetic patterns of the knee in gait in spastic diplegia of cerebral palsy. Gait Posture 2000;11:224–32.
[7] O'Byrne JM, Jenkinson A, O'Brien TM. Quantitative analysis and classification of gait patterns in cerebral palsy using a three-dimensional motion analyzer. J Child Neurol 1998;13:101–8.
[8] O'Malley MJ, Abel MF, Damiano DL, Vaughan CL. Fuzzy clustering of children with cerebral palsy based on temporal-distance gait parameters. IEEE Trans Rehabil Eng 1997;5:300–9.
[9] Rodda JM, Carson L, Graham HK, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. J Bone Joint Surg [Br] 2004;86B:251–8.
[10] Salazar AJ, De Castro OC, Bravo RJ. Novel approach for spastic hemiplegia classification through the use of the support vector machines. In: Proceedings of the 26th annual international conferences of the IEEE EMBS; 2004. p. 466–9.
[11] Stebbins J, Harrington M, Thompson N, Wainwright A, Forster H, Theologis TN. Gait classification in hemiplegic cerebral palsy based on EMG. Gait Posture 2004;20:S82–3.
[12] Stout J, Bruce B, Gage JR, Schutte L. Joint kinetic patterns in children with spastic hemiplegia cerebral palsy. Gait Posture 1995;3:274 [abstract].
[13] Simon SR, Deutsch SD, Nuzzo RM, Mansour MJ, Jackson JL, Koskinen M, et al. Genu recurvatum in spastic cerebral palsy. Report on findings by gait analysis. J Bone Joint Surg [Am] 1978;60: 882–94.
[14] Sutherland DH, Davids JR. Common gait abnormalities of the knee in cerebral palsy. Clin Orthop Relat Res 1993;139–47.
[15] Winters T, Gage J, Hicks R. Gait patterns in spastic hemiplegia in children and adults. J Bone Joint Surg [Am] 1987;69A:437–41.
[16] Wong MA, Simon S, Olshen RA. Statistical analysis of gait patterns of persons with cerebral palsy. Stat Med 1983;2:345–54.
[17] Wong AM, Chen CL, Hong WH, Chiou WK, Chen HC, Tang FT. Gait analysis through foot pattern recognition for children with cerebral palsy. J Musculoskelet Res 1999;3:71–81.
[18] Yokochi K. Gait patterns in children with spastic diplegia and periventricular leukomalacia. Brain Dev 2001;23:34–7.
[19] Zwick EB, Leistritz L, Milleit BVS, Zwick G, Galicki M, Witte HGS. Classification of equinus in ambulatory children with cerebral palsy—discrimination between dynamic tightness and fixed contractures. Gait Posture 2004;20:273–9.
[20] Romei M, Galli M, Motta F, Schwartz M, Crivellini M. Use of the normalcy index for the evaluation of gait pathology. Gait Posture 2004;19:85–90.
[21] Schutte LM, Narayanan U, Stout JL, Selber P, Gage JR, Schwartz MH. An index for quantifying deviations from normal gait. Gait Posture 2000;11:25–31.
[22] Graham HK, Harvey A, Rodda J, Nattrass G, Pirpiris M. The functional mobility scale (FMS). J Pediatr Orthop 2004;24:514–20.
[23] Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. Dev Med Child Neurol 1997;39:214–23.
[24] National Health and Medical Research Council. How to review the evidence: systematic identification and review of the scientific literature. Canberra, Australia: Biotext; 1999.
[25] Oxman A. Checklist for review articles. BMJ 1994;309:648–51.
[26] Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997;126:376–80.
[27] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Standards for Reporting of Diagnostic Accuracy. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med 2003;138:W1–2.
[28] McGinley JL. Observation of push off during gait following stroke: criterion-related validity and reliability of physiotherapy ratings, School of Physiotherapy, Faculty of Health Sciences Bundoora, Vic., Australia: La Trobe University; 2004.
[29] Toro B, Nester C, Farren P. A review of observational gait assessment in clinical practice. Physiother Theory Pract 2003;19:137–49.
[30] Woods G. The high-level mobility assessment tool for traumatic brain injury, School of Physiotherapy, Faculty of Health Sciences Bundoora, Vic., Australia: La Trobe University; 2005.
[31] Solway S, Brooks D, Lacasse Y, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. Chest 2001;119:256–70.
[32] Pietrobon R, Coeytaux RR, Carey TS, Richardson WJ, DeVellis RF. Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review. Spine 2002;27: 515–22.
[33] Ramaker C, Marinus J, Stiggelbout AM, Van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov Disord 2002;17:867–76.
[34] de Boer MR, Moll AC, de Vet HC, Terwee CB, Volker-Dieben HJ, van Rens GH. Psychometric properties of vision-related quality of life questionnaires: a systematic review. Ophthalmic Physiol Opt 2004;24: 257–73.
[35] World Health Organization. Towards a common language for functioning, disability and health: the international classification of functioning, disability and health. Geneva; 2002, http://www.who.int/classification/icf [retrieved January 24, 2005].
[36] Dodd KJ, Taylor NF, Damiano DL. A systematic review of the effectiveness of strength-training programs for people with cerebral palsy. Arch Phys Med Rehabil 2002;83:1157–64.
[37] Morris SL, Dodd KJ, Morris ME. Outcomes of progressive resistance strength training following stroke: a systematic review. Clin Rehabil 2004;18:27–39.
[38] Moseley AM, Herbert RD, Sherrington C, Maher CG. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). Aust J Physiother 2002;48:43–9.
[39] Fritz JM, Wainner R. Examining diagnostic tests: an evidence-based perspective. Phys Ther 2001;81:1546–64.
[40] Polgar T, Thomas SA. Introduction to research in the health sciences, 3rd ed., Melbourne: Churchill Livingstone; 1995.
[41] Koman LA, Mooney 3rd JF, Smith BP, Goodman A, Mulvaney T. Management of spasticity in cerebral palsy with botulinum-A toxin: report of a preliminary, randomized, double-blind trial. J Pediatr Orthop 1994;14:299–303.

[42] Boyd RN, Graham HK. Objective measurement of clinical findings in the use of Botulinum toxin A for the management of children with cerebral palsy. Eur J Neurol 1999;6:S23–35.
[43] Corry IS, Cosgrove AP, Duffy CM, McNeill S, Taylor TC, Graham HK. Botulinum toxin A compared with stretching casts in the treatment of spastic equinus: a randomised prospective trial. J Pediatr Orthop 1998;18:304–11.
[44] Koman LA, Brashear A, Rosenfeld S, Chambers H, Russman B, Rang M, et al. Botulinum toxin type a neuromuscular blockade in the treatment of equinus foot deformity in cerebral palsy: a multicenter, open-label clinical trial. Pediatrics 2001;108:1062–71.
[45] Mackey AH, Lobb GL, Walt SE, Stott NS. Reliability and validity of the Observational Gait Scale in children with spastic diplegia. Dev Med Child Neurol 2003;45:4–11.
[46] Gage JR. The treatment of gait problems in cerebral palsy. Clinics in Developmental Medicine No. 164–165 London: Mac Keith Press; 2004
[47] Graham HK, Selber P. Musculoskeletal aspects of cerebral palsy. J Bone Joint Surg [Br] 2003;85:157–66.
[48] Gage JR. Gait analysis in cerebral palsy. Clinics in Developmental Medicine No. 121. London: MacKeith Press; 1991.
[49] Ounpuu S, Deluca P, Davis RB. Gait analysis. In: Neville B, Goodman R, editors. Congenital hemiplegia. Clinics in Developmental Medicine No. 150. London: MacKeith Press; 2000. p. 81–97.
[50] Dobson F, Graham HK, Baker R, Morris ME. Multilevel orthopaedic surgery in group IV spastic hemiplegia. J Bone Joint Surg [Br] 2005;87-B:548–55.
[51] Everitt BS, Landau S, Leese M. Cluster analysis New York: Oxford University Press; 2001.
[52] Hartigan JA. Clustering algorithms New York: Wiley; 1975.
[53] Bezdek JC, Pal SK. Fuzzy models for pattern recognition New York: Institute for Electrical and Electronics Engineers Press; 1992.
[54] Rodda JM, Graham HK. Classification of gait patterns in spastic hemiplegia and spastic diplegia: a basis for a management algorithm. Eur J Neurol 2001;8:98–108.
[55] Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use, 2nd ed., New York: Oxford University Press Inc.; 2002.
[56] Aminian A, Vankoski SJ, Dias L, Novak RA. Spastic hemiplegic cerebral palsy and the femoral derotation osteotomy: effect at the pelvis and hip in the transverse plane during gait. J Pediatr Orthop 2003;23:314–20.
[57] Dobson F, Boyd RN, Parrott J, Nattrass GR, Graham HK. Hip surveillance in children with cerebral palsy. Impact on the surgical management of spastic hip disease. J Bone Joint Surg [Br] 2002;84: 720–6.
[58] Graham HK. Botulinum toxin type A management of spasticity in the context of orthopaedic surgery for children with spastic cerebral palsy. Eur J Neurol 2001;8:30–9.
[59] Õunpuu S, Michalak R, Romness M, Bell K, Deluca P. An evaluation of pelvic rotation pre- and post unilateral femoral derotation osteotomy in children with spastic hemiplegic cerebral palsy. Dev Med Child Neurol 2002.
[60] Saraph V, Zwick EB, Zwick G, Dreier M, Steinwender G, Linhart W. Effect of derotation osteotomy of the femur on hip and pelvis rotations in hemiplegic and diplegic children. J Pediatr Orthop B 2002;11: 159–66.