EUR J PHYS REHABIL MED 2008;44:213-20

The term diplegia should be enhanced. Part III: inter-observer reliability of the new rehabilitation oriented classification

R. PASCALE [1], S. PERAZZA [2], G. BORELLI [4], E. BIANCHINI [2], S. ALBORESI [3], P. B. PAOLICELLI [2], A. FERRARI [3, 4], G.CIONI [1, 2]

[1]Division of Child Neurology and Psychiatry
University of Pisa, Pisa, Italy
[2]Department of Developmental Neuroscience
IRCCS Stella Maris, Pisa, Italy
[3]Children Rehabilitation Unit
Santa Maria Nuova Hospital, Reggio Emilia, Italy
[4]Department of Neuroscience
University of Modena and Reggio Emilia
Modena and Reggio Emilia, Italy

Aim. **The aim of this study was to validate a recent classification of gait in children with the spastic diplegic form of cerebral palsy (CP) by checking the reliability of different scorers in assigning subject walking performance to one of the four specific patterns described in the classification.**

Methods. **The gait patterns of 50 children and adolescents with CP (23 males, 27 females; age range 3-17 years) were selected among patients whose videos were stored in the archives of the Pisa and Reggio Emilia Hospitals. Only video recordings of gait with homogeneous features (duration of at least 90 s, simultaneous recordings on sagittal and frontal views, and other criteria) were taken for examination. The videos were blindly scored using an observational gait scale, at first by two of the authors of the classification system (defined as "maximum experts"), then by ten expert observers, and finally by 206 professionals of rehabilitation after a one-day training on the classification. Cohen's kappa statistics (k) and intra class correlations (ICC) were calculated.**

Results. **Kappa and ICC indicate an almost perfect agreement both between the two maximum experts and among the ten expert observers. Good results were also obtained in the group of one-day trained scorers. Only a few cases were assigned to the "unclassified" category. The profession of the observer (doctor or therapist) and previous knowledge of the classification had no significant influence on reliability scores.**

Conclusion. **The results suggest that the proposed classification can be reliably applied, even utilizing short video recordings, to arrange diplegic children into different patterns. Further studies are needed to validate the use of this classification system for clinical and research aims.**

KEY WORDS: Cerebral palsy, classification - Cerebral palsy, diagnosis - Cerebral palsy, therapy.

Over the last few decades, the interpretation of cerebral palsy (CP) has changed, and a great effort has been made to develop newer nosographic classifications.[1, 2] After traditional classifications, mainly based on muscle tone and reflexes, steps have been made towards more modern approaches based on the study of the functional competence of the nervous system.[3-6] Some models are useful for a systematic reorganization of clinical findings, for eas-

Acknowledgements.—The Authors would like to thank M Coluccini, S Lanciani, R Di Pietro, S Frosini (Pisa), AR Maoret, S Muzzini, M Lodesani, and A Ovi (Reggio Emilia) for their help in designing the study and preparing the video recordings.

Fundings.—This research was supported by grant RC 1/2006 of the Italian Ministry of Health.

Received on July 18, 2006.
Accepted for publication on February 19, 2008.

Address reprint requests to: Prof. G. Cioni, Department of Developmental Neuroscience, IRCCS Stella Maris, Via dei Giacinti 2, 56128 Calambrone Pisa (Italy). E-mail: cioni@inpe.unipi.it

ier data collection and for epidemiological comparison.[7, 8] They can contribute to the acquisition of a common language providing a reference guide useful in different countries. Unfortunately, however, they offer a limited contribution to prognosis in children with CP.[3-6]

The traditional classifications are generic and based on signs and symptoms of pathology that cannot be modified by therapy.[9, 10] Moreover, they only classify CP on the basis of the type and severity of motor impairment and associated movement disorders without considering that the overall motor performances also encompass perceptual and cognitive deficits (vision, sense of movement, interest, attention, tolerance, etc.). These classifications only distinguish broad categories (tetraplegia, diplegia, hemiplegia) without making any discrimination within each group and without considering the natural history of each child, including the development of individual strategies and the changes occurring during growth.

The continuous evolution of CP concept and the large variety of clinical findings underline the need for newer approaches to classification. These must be clearly articulated and functionally-based to better understand child organization and therapeutic needs, in order to tailor an appropriate treatment with customized strategies carried out at the appropriate time.

Ferrari [4, 11] suggested a classification for diplegic children based on gait features, suitable to distinguish homogeneous groups. This classification takes advantage of many descriptions of gait patterns in CP children reported in the literature.[12-16] It includes four patterns of walking, whose principal items, specific for each pattern, are reported elsewhere.[11, 17, 18]

The proposed classification is derived from clinical observations of large series of patients; however, it still requires several studies to be validated. So far, this model seems to be useful to assess and understand many aspects of CP motor patterns. It is a simple and quick "gestalt" approach to the assessment, helpful for prognosis and serving as a guide for therapeutic programs.

The study described in this paper aims at validating this classification of gait in children with diplegia by checking the inter-rater reliability of judgements provided by either expert or trained observers through the evaluation of gait patterns taken from short video recordings.

Materials and methods

Subjects

The sample was composed of 50 CP children and adolescents, 23 male and 27 female, with ages ranging from 3 to 17 years (mean 9.4 years). Their video recordings, available from the archives of the two Centres (Pisa and Reggio Emilia Hospitals) involved in this study, show a suitable sequence of walking patterns.

The subjects were randomly selected among those fulfilling the following inclusion criteria: affected by CP, diagnosed according to the international definition,[2, 8, 9] and with non-progressive brain lesions documented by brain magnetic resonance imaging (MRI). All subjects had to be able to walk for at least 6 months. Children were excluded if they had been submitted to orthopaedic surgery in the 12 months preceding the videorecording or to botulinum toxin injection in the previous six months (*i.e.*, treatments which modify walking patterns).

Methods

The videos had to be performed according to the protocol proposed by the Italian Group for Cerebral Palsy (GIPCI)[17] with clear frontal and lateral views. The minimum acceptable duration was set at 90 s. Children were suitably dressed so that upper limb, trunk, pelvis and lower limb behaviour was easily observable. Videos of gait, either with or without devices (canes, walker), were accepted.

The raters were asked to observe a DVD displaying gait recordings: each child was shown twice at normal speed. During a short pause between each videoclip, they had to assign the child to one of the four gait pattern categories.

In the first part of the study two observers (AF and PBP), taken from the authors who contributed to the construction of this classification of gait in diplegia, independently evaluated all video recordings by means of a reference observation chart.

The agreed scoring by the maximum experts was considered the "gold standard" for comparison by successive observers.

As a second step, ten expert observers (four medical doctors and six physical therapists) scored the tapes; all of them were experienced in child rehabilitation and familiar with Ferrari's classification of gait patterns.[11, 18]

In the second part of the study, the 40 cases where the highest agreement (*i.e.*, ≥8 out of 10, 80%) was reached by the ten experts relative to the gold standard (see below), were later classified by numerous observers (medical doctors, physical therapists, students, orthopaedic technicians) during a one-day training on the use of the classification. The training included both lectures and video interpretation. Clinical cases shown as examples during training were not used during the reliability assessment. The scoring of the videos was performed in two separate sessions on the same day, with twenty clinical cases being presented in each session. The first session was completed by 206 observers: 113 physical therapists (70 self-declared as having a pre-training preliminary information about the classification, 43 not informed), 50 physiatrists (27 self-declared as informed, 23 not informed), 20 orthopaedic technicians (13 informed, 7 not informed), 11 child neurologists (3 informed, 8 not informed), one paediatric surgeon (not informed), one orthopaedic surgeon (not informed), 8 students (not informed) and two people who did not declare their profession. The second session was only attended by 165 observers out of the former 206: 89 physical therapists (56 self-declared informed of the classification, 33 not informed), 36 physiatrists (24 informed, 12 not informed), 9 orthopaedic technicians (6 informed, 3 not informed), 9 child neurologists (3 informed, 6 not informed), one orthopaedic surgeon (not informed), 6 students (not informed) and 15 unclassified people.

Statistical analysis

For each case, the perfect agreement among the observers was assessed as a percentage.[19]

Chance-corrected agreement between observers was measured using Cohen's kappa statistic (k) [20] and ICC (Intraclass Correlation). Kappa values were interpreted according to conventional groups (0.0-0.20=slight agreement; 0.21-0.40=fair; 0.41-0.60=moderate; 0.61-0.80=substantial; 0.81-1=almost perfect agreement).[21] The ICC was interpreted according to recommendations by Fleiss: an ICC greater than 0.75 indicates excellent reliability and from 0.4 to 0.75 indicates fair to good reliability.[22]

TABLE I.—*Scoring by the two maximum experts.*

Case number	Form according to observer 1	Form according to observer 2
1	IV	IV
2*	IV/IV*	III/IV*
3	II	II
4	IV	IV
5	III	III
6	IV	IV
7	III	III
8	II	II
9	I	I
10	IV	IV
11	II	II
12	I	I
13	II	II
14	I	I
15	III	III
16	III	III
17	II	II
18	I	I
19	III	III
20*	IV/IV*	II/IV*
21	IV	IV
22	IV	IV
23	III	III
24	IV	IV
25	III	III
26	IV	IV
27	IV	IV
28*	II/II*	IV/II*
29	I	I
30	III	III
31	III	III
32	IV	IV
33	III	III
34	IV	IV
35	I	I
36	III	III
37	II	II
38	IV	IV
39	I	I
40	IV	IV
41	III	III
42	I	I
43	IV	IV
44	IV	IV
45	II	II
46	III	III
47*	IV/IV*	III/IV*
48	IV	IV
49	IV	IV
50	IV	IV

*Cases of disagreement and final judgement after discussion and rescoring.

TABLE II.—*Results of the scoring of the 50 cases by 10 expert observers, compared to gold standard.*

Case number	Gold standard	Judgement of 10 expert observers				
		Form I	Form II	Form III	Form IV	Unclassified
1	IV			20%	80%	
2	IV			10%	90%	
3	II		100%			
4	IV				100%	
5	III			100%		
6	IV				100%	
7	III		20%	80%		
8	II		90%	10%		
9	I	100%				
10	IV				100%	
11*	II		70%	30%		
12	I	100%				
13	II		100%			
14	I	100%				
15	III			100%		
16	III			100%		
17	II		100%			
18	I	100%				
19	III			90%	10%	
20*	IV		50%	10%	40%	
21	IV				100%	
22	IV			10%	90%	
23	III	10%		80%	10%	
24	IV			10%	90%	
25*	III	30%		60%	10%	
26	IV				100%	
27	IV				100%	
28*	II		60%		40%	
29	I	80%	10%			10%
30	III			90%	10%	
31	III			100%		
32	IV				100%	
33*	III			50%	50%	
34	IV				100%	
35	I	90%	10%			
36	III			90%		10%
37*	II		60%	40%		
38	IV			10%	90%	
39*	I	40%	60%			
40	IV				100%	
41	III		20%	80%		
42	I	90%	10%			
43	IV				100%	
44	IV				100%	
45*	II	30%	70%			
46*	III	30%		70%		
47*	IV		10%	60%	30%	
48	IV				100%	
49	IV				100%	
50	IV				100%	

*Cases with agreement lower than 80%, not used for the one-day training.

TABLE III.—*k values of agreement between 10 expert observers.*

	Obs.2	Obs.3	Obs.4	Obs.5	Obs.6	Obs.7	Obs.8	Obs.9	Obs.10
Obs.1	0.889	0.696	0.889	0.694	0.805	0.657	0.710	0.745	0.795
Obs.2		0.751	0.89	0.695	0.806	0.658	0.665	0.747	0.797
Obs.3			0.751	0.692	0.777	0.62	0.634	0.715	0.709
Obs.4				0.751	0.862	0.715	0.721	0.747	0.797
Obs.5					0.804	0.681	0.605	0.686	0.707
Obs.6						0.855	0.746	0.772	0.824
Obs.7							0.766	0.732	0.756
Obs.8								0.798	0.791
Obs.9									0.94

TABLE IV.—*ICC values for groups with different levels of experience and different professions among the 10 expert observers.*

	Level of experience		Profession	
	>5 years (5 observers)	<5 years (5 observers)	Doctors (4 observers)	Therapists (6 observers)
ICC (CI 95%)	0.982 (0.973-0.989)	0.975 (0.962-0.985)	0.968 (0.951-981)	0.975 (0.963-0.985)

Results

Maximum experts

The judgments provided by the two maximum experts differed in 4/50 cases: two were scored as form III instead of IV, two as II instead of IV or vice versa (Table I). The inter-observer agreement evaluated by k statistics was 0.859 (*i.e.*, "almost perfect"), and ICC (95% IC) was 0.951 (0.914-0.972). After discussion and re-scoring of the four discordant videos, a common judgment was reached for all cases. The final sample consisted of 8 cases with Form I, 8 with Form II, 13 with Form III and 21 with Form IV.

Expert observers

The inter-(expert) observer agreement was firstly computed for each case as a percentage in comparison with the gold standard (Table II).

The ten observers agreed with the gold standard 100% in 24 cases (14 belonging to Form IV, 4 to III, 4 to I and 2 to II); 90% in 10 cases (3 Form III, 4 IV, 2 I and 1 II); and 80% in 6 cases (3 Form III, 1 IV, 1 I and 1 II).

The agreement among the ten expert observers was similar to that of the maximum experts. All values were between "moderate" (k=0.605 in the case of observer 5 *vs.* observer 8) and "almost perfect" (k=0.94 in the case of observer 9 *vs.* observer 10) (Table III). The ICC value (95% IC) of this group was 0.985 (0.978-0.991).

The influence of the level of experience in the specific area (more or less than 5 years) and the profession (medical doctor or physical therapist) on inter-rater reliability was also evaluated. The ICC values between groups with different levels of experience or profession were similar (Table IV). Only two observers used the "unclassified" option for two cases.

One-day trained observers

During the first session of 20 cases, the one-day trained observers agreed with the gold standard 100% for no cases; 90% in 9 cases: 4 belonging to Form IV, 3 to I and 2 to III; and 80% in 7 cases: 3 belonging to Form II, 2 to III, 1 to I and 1 to IV (Table V).

The mean k value of agreement was 0.82, ranging from 0.66 for orthopaedic technicians to 0.93 for the orthopaedic surgeon. The different levels of knowledge of the classification did not have any influence (subjects declaring some knowledge: mean k value of 0.81; subjects without any knowledge: mean k value of 0.80; and undeclared: mean k value of 0.89).

During the second session for the other 20 cases, the observers agreed with the gold standard 100% in one case (Form IV); 90% in 7 cases (6 belonging to Form IV and 1 to III); and 80% in 5 cases (3 belonging to Form IV and 2 to III) (Table V).

TABLE V.—*Results of scoring for 40 cases by the one-day trained observers compared to gold standard.*

Case number	Gold standard	Judgement of one-day observers				
		Form I (%)	Form II (%)	Form III (%)	Form IV (%)	Unclassified (%)
Session I						
1	IV	8	1.5	33	56	1.5
2	IV		4.5	9	86	0.5
3	II		80	2	13.5	4.5
4	IV		1	1.5	97	0.5
5	III	4.5	3	91.5	0.5	0.5
6	IV		2	1	95	2
7	III	6	13	75	1	5
8	II	0.5	82.5	1.5	11	4.5
9	I	98	0.5		1	0.5
10	IV	0.5	1	1.5	96	1
12	I	84	14.5	0.55		1
13	II	1	37.5	7	44	10.5
14	I	95	4			1
15	III	4.5	2.5	91.5	2	0.5
16	III	6.5	9	82		2.5
17	II		87		12.5	0.5
18	I	96.5	3			0.5
19	III	4.5	13	82		0.5
21	IV	1.5	16.5	6.5	71.5	4
22	IV		1		93.5	5.5
Session II						
23	III	28	5	64		3
24	IV				100	
26	IV	0.6	0.6	16.3	81.3	1.2
27	IV		0.6		92.8	1.2
29	I	45	47.2	5.4		2.4
30	III	1.2	0.6	91	5.4	1.8
31	III		1.2	46.6	49.2	3
32	IV		9.6	1.2	82.6	6.6
34	IV		0.6		98.2	1.2
35	I	71.6	26.6	0.6	0.6	0.6
36	III	9.6	4.8	80.2	1.2	4.2
38	IV	1.2			98.2	0.6
40	IV			0.6	97	2.4
41	III	4.2	4.8	86.8	0.6	3.6
42	I	64.3	31.5	1.8		2.4
43	IV	1.8	1.8	16.2	80.2	
44	IV		24.8	10.3	45.5	19.3
48	IV		0.6		94.6	4.8
49	IV				97.6	2.4
50	IV		29.6	2.4	64.4	3.6

The mean k values of agreement with the gold standard was 0.73, ranging from 0.55 for orthopedic technicians to 0.74 for the orthopaedic surgeons and the child neurologists; the different levels of knowledge had no influence (subjects with some knowledge: mean k value of 0.74; with no knowledge: mean k value of 0.76; undeclared: mean k value of 0.70) (Table VI, upper part). There were no significant differences resulting from previous knowledge of the classification (Table VII).

Values of agreement among one-day trained observers, measured by ICC for the most represented professional groups, were excellent: they ranged from 0.950 to 0.999 (Table VII, lower part). The "unclassi-

TABLE VI.—*Mean k values, compared to gold standard, and ICC for different professional groups among the one-day trained observers.*

	Session I		Session II	
	N	Mean k value	N	Mean k value
Physiotherapists	113	0.863	89	0.73
Physiatrists	50	0.811	36	0.55
Orthopedic technicians	20	0.662	9	0.68
Child neurologists	11	0.812	9	0.74
Paediatric surgeon	1	0.779	—	—
Orthopedic surgeon	1	0.932	1	0.74
Students	8	0.894	6	0.69
Undeclared	2	0.819	15	0.55

	Session I		Session II	
	N	ICC (95%IC)	N	ICC (95%IC)
Physiotherapists	113	0.998 (0.996-1)	89	0.999 (0.996-1)
Physiatrists	50	0.996 (0.992-0.999)	36	0.992 (0.982-0.998)
Orthopedic technicians	20	0.976 (0.953-0.990)	9	0.954 (0.901-0.985)
Child neurologists	11	0.988 (0.975-0.996)	9	0.950 (0.898-0.981)
Students	8	0.967 (0.945-0.985)	6	0.963 (0.925-0.986)

TABLE VII.—*Mean k values, compared to gold standard, of the one-day trained observers with different knowledge of the classification system.*

	Session I		Session II	
	SS	Mean k value	SS	Mean k value
Physiotherapists	113	0.863	89	0.73
Some knowledge	109	0.816	89	0.74
No knowledge	95	0.807	61	0.76
Undeclared	2	0.894	15	0.70

fied" option was used by 53 of the 206 one-day trained observers of session I (mean 1.7 times, range 1-5) and by 65 of the 156 one-day trained observers of session II (mean 1.6 times, range 1-4).

Discussion

Despite a large interest in different types of gait in CP, so far no well defined and generally accepted classification has been produced, and many proposals available in the literature do not completely fulfil required criteria of both internal and external validity.[15] Thus further efforts are needed to confirm the existing classification system or to produce and validate new systems.

The scale of walking classification in spastic diplegia presented in this paper has obtained several positive indices in terms of validity and reliability. To be valid, any scale or classification system must meet the needs for which it has been created. Moreover, to be reliable, the proposed system has to give the same results when used by different observers. In this study the validity is indirectly indicated by the frequency of using the "unclassified" option by the observers. Among the ten "expert" observers group, only two observers used this option for two cases The use of this option was more frequent in one-day training group. The mild increase of the number of "unclassified" options from session I to II (used one or two times by about 1/4 of the scorers of session I and more than 1/3 of session II) could be due to observer tiredness (session II was scheduled in the afternoon of the training day) or to the presence of more difficult cases in the second group.

The reliability, measured by k statistics and ICC, was good. Data analysis shows that the agreement among expert observers was not influenced by different levels of clinical experience (years of work). Although the younger observers showed lower statistic values, they had "almost perfect" mean values. Similar results were obtained by different professions: medical doctors and physical therapists all had similar ICC. Different professions had no influence on learning ability or on the use of the classification among the one-day trained observers.

The two "maximum" experts agreed on the classification of 46 out of 50 subjects. The percentage of each form of these 46 cases is in agreement with those recently found on large hospital based population of children with diplegia.[19] Forms IV and I are the most frequently occurring forms, followed by III and II.

The concordance of the ten expert observers was also compared with that of the maximum experts (gold standard). After arbitrarily fixing a cut-off of 80% (agreement of at least 8 observers over 10 *versus* the gold standard), an agreement was found for 40 of the 50 cases.

The most easily identifiable forms were Form IV (19/21 cases identified with at least an agreement of

80% among the scorers) and Form I (7/8 cases), followed by Form III (10/13 cases) and Form II (4/8 cases).

Forms IV and I are the extremes of classification in terms of motor impairment. They also significantly differ in terms of Gross Motor Function Classification System.[23] Likely for this reason, they are the most easily identifiable forms at gestalt. Equally, Form III is easily identifiable because its main core consists of the frontal swing of trunk, an evident element at a gestalt evaluation. The main core of Form II is a knee flexion in mid stance, an element only evident in sagittal view and less so in the frontal view.

Conclusions

The results obtained in this study confirm the possibility to use this classification in the clinical practice. Preliminary findings reported above suggest that even rehabilitation professionals not familiar with this classification can learn to use it quite rapidly. This tool is fast and quick to apply and does not require expensive technologies. In the future, a study with a more homogenous sample in terms of number of subjects for each form will be needed. This will allow for more precise statistical analyses. When fully validated, this classification system could be useful in clinical and research settings.

References

1. Bax MC. Terminology and classification of cerebral palsy. Dev Med Child Neurol 1964;6:295-7.
2. Bax M, Goldestein M, Roseubaum P, Leviton A, Paneth N, Dan B et al. Executive Committee for the definition of cerebral Palsy. Proposed definition and classification of cerebral palsy. Dev Med Child Neurol 2005;47:571-7.
3. Palisano RJ, Rosenbaum PL, Walter S, Russel S, 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.
4. Ferrari A. Around the problem of classification of cerebral palsy (In Italian). Giornale di Neuropsichiatria dell'Età Evolutiva 1995;15:191-205.
5. Ferrari A. Proposte riabilitative nelle paralisi cerebrali infantili. Pisa: Edizioni Del Cerro; 1997.
6. Dan B, Cheron G. Reconstructing Cerebral Palsy. J Ped Neurol 2004;2:57-64.
7. Colver AF, Sethumanhavan T. The term diplegia should be abandoned. Arch Dis Child 2003;88:286-90.
8. Surveillance of Cerebral Palsy Europe (SCPE). Dev Med Child Neurol 2000;42:816-24.
9. Hagberg B. Nosology and classification of Cerebral Palsy. Giornale di Neuropsichiatria dell'età evolutiva 1989;4:12-7.
10. Hagberg B, Hagberg G. The changing panorama of cerebral palsy - bilateral spastic forms in particular. Acta Paediatr Suppl 1996;416:48-52.
11. Ferrari A, Cioni G. Le forme spastiche della paralisi cerebrale infantile: guida all'esplorazione delle funzioni adattive. Milano: Edizioni Springer; 2005.
12. Bobath B, Bobath K. Motor development in the different types of cerebral palsy. London: William Heinemann Medical Books; 1975.
13. 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;11:224-34.
14. Yokochi K. Gait patterns in children with spastic diplegia and periventricular leucomalacia. Brain Dev 2001;23:34-7.
15. Dobson F, Morris ME, Baker R, Kerr Graham H. Gait classification in children with cerebral palsy: a systemic review. Gait Posture 2007;25:140-52.
16. Rodda J, Graham HK. Classification of gait patterns in spastic hemiplegia and spastic diplegia: a basis for a management algorithm. Eur J Neur 2001;5:98-108.
17. Gruppo Italiano Paralisi Cerebrale Infantile. Protocol of videorecording (In Italian). Giornale di Neurops Età Evolutiva 1998;2:52-7.
18. Ferrari A., Alboresi S, Muzzini S, Pascale R, Perazza S, Cioni G. The terms diplegia should be enhanced (I): around the problem of classification of cerebral palsy. Eur Medicoph (submitted).
19. Morris C, Kurinczuk JJ, Fitzpatrick R, Rosenbaum PL. Reliability of the Manual Ability Classification System for children with cerebral palsy. Dev Med Child Neurol 2006;48:950-3.
20. Cohen J. A coefficient of agreement for nominal scales. Educ Psycol Meas 1960;20:37-46.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
22. Shourt PE, Feiss LJ. Intraclass Correlation: uses in assessing rater reliability. Psycol Bull 1979;86:420-8.
23. Cioni G, Lodesani M, Coluccini M, Sassi S, Paolicelli PB, Pascale R et al. The term diplegia should be enhanced (II): contribution to validation of a new classification system. Eur Medicoph (submitted).