Interrater reliability of the *Saint-Anne Dargassies* Scale in assessing the neurological patterns of healthy preterm newborns

Carla Ismirna Santos Alves [1]
João Gonçalves de Medeiros [2]
Luciana Arruda Carriço [3]
Nívia Maria Arraes [4]
Roselene Ferreira de Alencar [5]
Aurea Nogueira de Melo [6]

[1] Programa de Pós-graduação em Ciências da Saúde. Universidade Federal do Rio Grande do Norte. Natal, RN, Brasil.
[2] Departamento de Imagem da Maternidade Escola Januário Cicco. Universidade Federal do Rio Grande do Norte. Natal, RN, Brasil.
[3] Centro de Reabilitação Infantil, Natal, RN, Brasil.
[4,6] Departamento de Pediatria. Centro de Ciências da Saúde. Universidade Federal do Rio Grande do Norte. Av. Gal. Gustavo Cordeiro de Faria, s.n. Petrópolis. Natal, RN, Brasil. CEP: 59.078-970. Email: aureanmelo@yahoo.com.br
[5] Departamento de Fisioterapia. Centro de Ciências da Saúde. Universidade Federal do Rio Grande do Norte. Natal, RN, Brasil.

Abstract

Objectives: to assess the interrater reliability of the Saint-Anne Dargassies Scale in assessing neurological patterns of healthy preterm newborns.

Methods: twenty preterm newborns met the inclusion criteria for participation in this prospective study. The neurologic examination was performed using the Saint-Anne Dargassies Scale, showing normal serial cranial ultrasound examination. In order to test the reliability, the study was structured as follows: group I (rater 1/physiotherapist; rater 2/neonatologist); group II (rater 3/physiotherapist; rater 4/child neurologist) and the gold standard (expert and professor in pediatric neurology).

Results: high interrater agreement was observed between groups I – II compared with the gold standard in assessing postural pattern (p<0.01). Regarding the assessment of primitive reflexes, greater agreement was observed in the evaluation of palmar grasp reflex and Moro reflex (p<0.01) for group I compared with the gold standard. An analysis of tone demonstrated heterogeneous agreement, without compromising the reliability of the scale. The probability of equality between measurements of head circumference in the two groups, compared with the gold standard, was observed.

Conclusions: the Saint-Anne Dargassies Scale demonstrated high reliability and homogeneity with significant power of reproducibility and may be capable to identify preterm newborns suspected of having neurological deficits.

Key words *Neonatal screening, Reliability, Infant, premature, Preterm, Evaluation*

Introduction

Scientific and technical advances have produced satisfactory results in newborn care, including significant improvements in intensive care, especially for preterm infants, which resulted in higher survival rates, making neurological assessment indispensable for these newborns at risk.[1-3]

In the last two centuries, the advent of more advanced complementary exams has not reduced the value of the clinical assessment of newborns with and without either clinical or neurological complications. The use of these resources has enabled clinicians to predict the development of neurological lesions with greater accuracy, as clinical assessments remain the most accessible patent care resource in most hospitals.[4-6]

Volpe[7] argued that the identification of isolated neurological signs during the neonatal period should not be considered predictive. However, he underscored that the predictive value increases when a set of altered signs are identified during a neurological examination, as the presence of several abnormalities on such an assessment is suggestive of the presence of a severe neurological disorder, which increases the predictive ability of the clinical exam.[7]

Neonatal studies on the reliability of scales that assess preterm newborns are related primarily to the assessment instrument for acute pain.[8-10] With respect to scale reliability in assessing neurological pattern in preterm, Deschênes *et al.*[11] underscored the interrater reliability of the *Amiel-Tison* Neurological Assessment at term, showing its importance and exhibiting excellent reliability, while Simard *et al.*[12] demonstrated good validity and reliability in preterms and terms up to 6 years of age. However, Gagnon, in his doctoral thesis, showed that Premie-Neuro raw scores had acceptable reliability and validity for use by clinicians to identify at-risk preterm infants, however, its classifications should be interpreted cautiously.[13] On the other hand, Fernãndez *et al.*[14] reported acceptable validity and reliability using the Spanish version of the Premie-Neuro scale for preterm children. Leroux *et al.*[15] showed that interrater reliability of the *Amiel-Tison* assessment tool is very good, and when performed by a highly trained examiner, the results correlate with developmental performance at 2 years of corrected age.

Thus, it is important to evaluate the reliability of assessment instruments, determining the skill of examiners in measuring or identifying subject or object differences, thereby decreasing errors inherent to diagnosis, scores or measurements.[16]

The *Saint-Anne Dargassies* Scale (SDS) is reference for the follow-up of preterm infants, particularly because it defines the maturational evolution of these newborns every two weeks until term.[17] However, its assessment properties must be verified, as satisfactory levels of both validity and reliability of an instrument are not guaranteed when used in a population different from that for which its use was intended.[18-20]

Although the SDS is considered the gold standard in assessing preterm newborns (PTNBs), it can be verified that the scale has been underused in PTNBs researches.

The aims of this study were to evaluate the reliability of the SDS in assessing the neurological patterns of healthy PTNBs, and to analyze the level of agreement between interraters and the interrater considered the gold standard.

Methods

This was a prospective study, performed using the SDS[17] in PTNBs who were born and treated in the Intensive Care Unit of Januário Cicco Maternity School (MEJC) in Natal, Brazil. The sample was calculated based on the number of complicated and non-complicated preterm newborns born at the MEJC over a one-year period. The PTNBs were assessed by a neonatologist for clinical, physical and biochemical aspects following birth and defined as optimal to be part of the sample, for not exhibiting any clinical or neurological complications of prematurity.

This project was approved by the Research Ethics Committee of the Federal University of Rio Grande do Norte, under protocol number 423/2010. Written informed consent was provided by the parents of all PTNBs.

The inclusion criteria for participation in this study were as follows: a) being preterm newborns treated at the Intensive Therapy Unit of MEJC: b) having a gestational age (GA) between 32 and 37 weeks; c) exhibiting no abnormalities and d) having undergone serial cranial ultrasonography evaluations without abnormalities. PTNBs were not included if they had any of the following: malformations of the central nervous system (myelomeningocele, hydrocephaly, anencephaly, and others), neurological or clinical complications requiring either intubation or sedation, maternal sedation during the first 48 hours of life and abnormal cranial ultrasound, as well as failure to comply with the assessments as established via research.

To test the reliability of the SDS and interrater

agreement, the collection was structured as follows: group I consisted of rater/1 (physiotherapist) and rater/2 (neonatologist); group II consisted of rater/3 (physiotherapist) and rater/4 (child neurologist); these groups were compared with the gold standard (GS). The GS is the childhood neurology professor, with a medical residency in pediatric neurology and exercising this specialty for at least 35 years. Each group independently assessed 10 PTNBs for a final total of 20 PTNBs. The GS assessed the total n of the sample under the same conditions and on the same days as groups I and II. The PTNBs were well fed, and a 1-hour interval was allowed between each observer's exam, with an application time of approximately 10 minutes.

On the SDS, the following items were assessed: the cardinal points reflexes and the palmar grasp, Moro, crossed extension and gait reflexes; passive muscle tone was examined by measuring the popliteal, foot-leg and heel-to-ear articular angles, and active tone was assessed via the observation of spontaneous movements, lower limb straightening, head straightening and an examination of the neck flexors (traction maneuver). The state of awareness was proposed by SDS, who classified them as sleepiness, provoked wakefulness, spontaneous wakefulness, wakefulness, altered wakefulness or asleep, and sedated.[17] The first neurological examination was conducted within 72 hours following birth and repeated every two weeks until the infant reached term.

To assess the infants' articular angles (A), a goniometer was specially adapted, per Alves and Melo.[21] At the end of each item assessed by the SDS, the score values between 0 and 2 were estimated, where 0 was an absent response, 1 an altered response and 2 an expected response.

The cranial ultrasound exams were performed and interpreted by an examiner blinded to the neurological patterns of the PTNBs, using a GE-LOGIQ P6 ultrasound machine with a convex transducer of 6 to 10 MHz and diameter of 3 cm, which was applied using the anterior and posterior transfontanellar technique. The exams were conducted in the coronal (anterior, medial and posterior), sagittal (median and paramedian) and axial planes. The cranial ultrasound was performed before the neurological exam and every two weeks thereafter until each infant reached term.

Before the application of the SDS, all four observers attended a preparatory workshop taught by an experienced examiner (GS). The workshop was divided into two stages. Stage 1 addressed the items assessed by the scale, and stage 2 involved practical training to standardize the assessment of the scale's items. The first 4 meetings consisted of the GS applying the scale in PTNBs hospitalized in the MEJC to evaluate the 4 raters. During the remaining 8 meetings, the four raters attempted the SDS under the supervision of the GS. The preparatory workshop concluded after 12 meetings, when the raters were considered qualified to utilize the SDS.

The data were tabulated and stored using Microsoft Excel 2010. The databank was exported to SPSS 20.0 software, the primary tool with which the statistical analyses were performed, using R software, version 3.3.1. An analysis of the interrater agreement between the two groups and the GS was completed using the binomial test. The significance level was set at 2%, and a critical value of 3 disagreements was established to reject the disagreement hypothesis; that is, in the observation of seven interrater agreements, the hypothesis that the observer disagreed with the assessment was rejected. The Wilcoxon test was used to compare continuous measurements, with a significance level of 5%. If the *p*-value of the interrater test was less than 0.05, there was considerable evidence to reject the equality between the distributions of the values used between the observers and the GS.

Results

A total of 26 PTNBs without either clinical or neurological complications were selected. However, the study sample was ultimately composed of 20 non-complicated PTNBs because six were excluded from study due to the following: the detection of intracranial hemorrhage during the first ultrasound exam (1 case), signs of alcohol and drug withdrawal (1 case), signs of lethargy (1 case) and failure to complete each of the assessment stages of the study (3 cases). Of the 20 PTNBs selected, 13 (65%) were female, and 7 (35%) were male. With respect to gestational age, 1 (5%) was 32 weeks, 3 (15%) were 33 weeks, 9 (45%) were 35 weeks, and 7 (35%) were 34 weeks. Regarding weight for gestational age, 15 (75%) were considered appropriate for gestational age and 5 (25%) were considered small for gestational age. Regarding resuscitation in the delivery room, 13 (65%) did not require resuscitation, whereas 7 did. Regarding the use of resuscitation resources in the delivery room were 6 (30%) required an oxygen hood or helmet and 1 (5%) newborn required continuous positive airway pressure (CPAP).

On the APGAR scores at 1st and 5th minutes, the PTNBs were stable, and those that required

resuscitation in the delivery room progressed satisfactorily following resuscitation. During the first minute of APGAR screening, the following was observed: 7 (35%) had APGAR scores of 9, 9 (45%) had APGAR scores of 8, 2 (10%) had APGAR scores of 7, and 2 had no record of their first minute APGAR scores. Regarding the 5-minute APGAR score, the following was noted: 16 (80%) had a score of 9, 1 (5%) had a score of 8, 1 (5%) had a score of 6 and 2 (10%) had no record of their 5-minute APGAR scores.

The likely etiologies of premature deliveries were as follows: 4 (20%) with pre-eclampsia, 12 (60%) with premature labor, 1 (5%) with premature rupture of membranes, 1 (5%) with specific hypertensive disease of pregnancy, 1 (5%) with amniorrhexis and 1 (5%) with respiratory discomfort.

The maternal ages were grouped as follows: 14-20 years (9/45%); 20-30 years (7/35%); and 30-39 years (4/20%). With respect to the current pregnancy, 9 (45%) had a single pregnancy, 8 (40%) had twins, and 3 (15%) had triplet. Regarding prenatal care, 11 (55%) had between 3 and 5 appointments, 5 (25%) had between 6-8 appointments, and 4 (20%) had between 9-10 appointments.

When asked about complications during pregnancy, 1 (5%) had urinary tract infection (UTI) and high blood pressure (HBP), 5 (25%) had anemia, 3 (15%) experienced only HBP, 3 (15%) exhibited both HBP and anemia; 2 (10%) had UTI, and 6 (30%) women reported having no complications during pregnancy. Regarding the trimester in which the aforementioned complications occurred, 2 (10%) developed complications during the first trimester, 9 (45%) developed complications during the second trimester, 3 (15%) developed complications during the third trimester, and 6 (30%) denied having any complications.

Regarding the type of delivery, 11 (55%) underwent a cesarean delivery and 9 (45%) a vaginal delivery. Regarding their presentations at delivery, 15 (75%) were born with a cephalic presentation, 3 (15%) with a pelvic presentation, and 2 (10%) with no record of their presentation at the time of delivery.

Table 1 shows the interrater agreement in the assessment of postural patterns between groups I and II; 100% agreement was noted between rater/1 and the GS ($p<0.01$) for the four assessments of postural pattern; however, between rater/2 and the GS, the first assessment had a lower index of agreement (90%). However, the P-value remained significant ($p<0.01$) and 100% agreement was noted in the postural assessment between rater/3 and rater/4 compared with the GS ($p<0.01$).

Table 2 demonstrates the likelihood of significant interrater agreement in the evaluation of head circumference between groups I and II compared with the GS. The group I P-values for the head circumference measurements in the four SDS assessments were 0.944, 1.000, 0.905 and 0.915, and the P-values for group II were 0.634, 0.833, 0.259 and 0.191.

The assessments of the infants' primitive

Table 1

Agreement between group I and group II raters and the gold standard in assessing postural patterns of 20 preterm newborns using the *Saint-Anne Dargassies* Scale.

Type of assessment	Interrater agreement			
	R1 and GS		R2 and GS	
	Agreement (%)	*p*	Agreement (%)	*p*
PP 1SDS	100	<0.01	90	<0.01
PP 2SDS	100	<0.01	100	<0.01
PP 3SDS	100	<0.01	100	<0.01
PP 4SDS	100	<0.01	100	<0.01
	R3 with GS		R4 with GS	
	Agreement (%)	*p*	Agreement (%)	*p*
PP 1SDS	100	<0.01	100	<0.01
PP 2SDS	100	<0.01	100	<0.01
PP 3SDS	100	<0.01	90	<0.01
PP 4SDS	100	<0.01	100	<0.01

PP = postural pattern; R1 = rater/1; R2 = rater/2; R3 = rater/3; R4 = rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

reflexes demonstrated that the highest rate of interrater agreement between group I and the GS was noted in the assessment of the palmar grasp and Moro reflexes, with a P-value of 0.01 for the four assessments (Table 3). Significant agreement was noted between raters 3 and 4 and the GS with respect to the assessment of the infants' primitive reflexes. The interrater agreement was not homogeneous as observed in the assessment of the palmar grasp and Moro reflexes in group I. Group II exhibited greater heterogeneity between the observers and the GS (Table 4).

Table 5 displays the level of agreement between the group I and group II raters and the GS regarding articular angles and demonstrates that all measurements of the popliteal angle (PA) were significant for group I; however, the strongest agreement was noted between rater/2 and the GS. The measurements of only the foot-leg angle (FLA) demonstrated significant agreement for group I in the second, third and fourth assessments. All measurements of the heel-ear angles (HEA) in group 1 were considered significant.

Significant interrater agreement in the assessment of the PA was noted only in the fourth assessment, with a P-value of 0.01 for both raters/3 and 4 compared with the GS, as demonstrated in Table 5. The measurements of the HEA exhibited significant agreement between rater/3 and the GS only in the third and fourth assessments (p=0.01). Rater 4 and the GS agreed significantly in all four

assessments of the foot-leg angle. With respect to the HEA assessment in group II, significant agreement was noted between rater 3 and the GS only in the first and third assessments, and between rater 4 and the GS in the first three assessments.

The assessment of active tone in the PTNBs demonstrated significant homogeneity between the group I and II raters and the GS on the upper limb rebound test, with a P-value of 0.01 for the four group assessments. During the spontaneous movements exam, all group I assessments obtained significant agreement, with a P-value of 0.01. In group II, this agreement was significant only for the assessments of this item between rater/4 and the GS, as rater/3 and the GS agreed significantly on the evaluation of spontaneous movement only in assessments 2th, 3th and 4th (p=0.01).

Regarding the traction maneuver, which assessed muscle flexor strength, significant agreement was noted among all of group I assessments, with a $p<0.01$ for all of the assessments in this group. Significant agreement in the four group II assessments was noted only between rater/3 and the GS. Interrater agreement was considered significant between rater/4 and the GS only for the first, second and fourth assessments, with $p<0.01$ for the first and second assessments and a P-value of 0.02 for the 4th assessment.

Table 2

Significance between the group I and II raters and the gold standard in measurements of head circumferences, and biauricular and anteroposterior parameters of 20 preterm newborns.

Type of assessment	Interrater agreement	
	R1 and GS	R2 and GS
	p	*p*
HC 1SDS	0.944	0.634
HC 2SDS	1.000	0.833
HC 3SDS	0.905	0.259
HC 4SDS	0.915	0.191
BA 1SDS	0.198	0.015
BA 2SDS	0.010	0.030
BA 3SDS	0.016	0.009
BA 4SDS	0.030	0.058
AP 1SDS	0.028	0.032
AP 2SDS	0.009	0.014
AP 3SDS	0.014	0.013
AP 4SDS	0.030	0.022

HC = head circumference; BA = biauricular; AP = anteroposterior; R1=rater/1; R2 = rater/2; R3 = rater/3;R4 = rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Table 2   **conclusion**

Significance between the group I and II raters and the gold standard in measurements of head circumferences, and biauricular and anteroposterior parameters of 20 preterm newborns.

Type of assessment	Interrater agreement	
	R3 and GS	R4 and GS
	p	*p*
HC 1SDS	1.000	0.048
HC 2SDS	0.013	0.035
HC 3SDS	0.581	0.179
HC 4SDS	0.077	0.122
BA 1SDS	0.233	0.462
BA 2SDS	0.156	0.035
BA 3SDS	0.362	0.076
BA 4SDS	0.227	0.309
AP 1SDS	0.498	0.622
AP 2SDS	0.931	0.671
AP 3SDS	0.160	0.049
AP 4SDS	0.172	0.097

HC = head circumference; BA = biauricular; AP = anteroposterior; R1=rater/1; R2 = rater/2; R3 = rater/3;R4 = rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Table 3

Agreement between the group I raters and the gold standard regarding the assessment of primitive reflexes of 20 preterm newborns using the *Saint-Anne Dargassies* Scale.

Type of assessment	Interrater agreement			
	R1 and GS		R2 and GS	
	Agreement (%)	*p*	Agreement (%)	*p*
CPR 1SDS	80	<0.01	80	<0.01
CPR 2SDS	100	<0.01	90	<0.01
CPR 3SDS	100	<0.01	100	<0.01
CPR 4SDS	90	<0.01	90	<0.01
PGR 1SDS	100	<0.01	100	<0.01
PGR 2SDS	100	<0.01	100	<0.01
PGR 3SDS	100	<0.01	100	<0.01
PGR 4SDS	100	<0.01	100	<0.01
MR 1SDS	100	<0.01	100	<0.01
MR 2SDS	100	<0.01	100	<0.01
MR 3SDS	100	<0.01	100	<0.01
MR 4SDS	100	<0.01	100	<0.01
CER 1SDS	70	0.02	80	<0.01
CER 2SDS	90	<0.01	90	<0.01
CER 3SDS	100	<0.01	100	<0.01
CER 4SDS	100	<0.01	100	<0.01
AGR 1SDS	90	<0.01	90	<0.01
AGR 2SDS	90	<0.01	100	<0.01
AGR 3SDS	100	<0.01	100	<0.01
AGR 4SDS	100	<0.01	100	<0.01

CPR = cardinal points reflex; PGR = palmar grasp index; RM = Moro reflex; CER = crossed extension reflex; AGR=automatic gait reflex; R1 = rater/1; R2 = rater/2; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Table 4

Agreement between group II raters and the gold standard regarding the assessment of primitive reflexes of 20 preterm newborns using the *Saint-Anne Dargassies* Scale.

Type of assessment	Interrater agreement			
	R3 and GS		R4 and GS	
	Agreement (%)	*p*	Agreement (%)	*p*
CPR 1SDS	50	0.213	60	0.076
CPR 2SDS	80	<0.01	90	<0.01
CPR 3SDS	90	<0.01	100	<0.01
CPR 4SDS	80	<0.01	100	<0.01
PGR 1SDS	100	<0.01	90	<0.01
PGR 2SDS	90	<0.01	80	<0.01
PGR 3SDS	100	<0.01	100	<0.01
PGR 4SDS	90	<0.01	90	<0.01
MR 1SDS	50	0.213	60	0.076
MR 2SDS	90	<0.01	100	<0.01
MR 3SDS	100	<0.01	100	<0.01
MR 4SDS	100	<0.01	100	<0.01
CER1SDS	40	0.787	70	0.02
CER 2SDS	50	0.213	40	0.787
CER 3SDS	70	0.02	60	0.076
CER 4SDS	60	0.076	70	0.02
AGR 1SDS	80	<0.01	80	<0.01
AGR 2SDS	90	<0.01	90	<0.01
AGR 3SS	70	0.02	90	<0.01
AGR 4SDS	80	<0.01	80	<0.01

RPC = cardinal points reflex; PGR = palmar grasp; RM = Moro reflex; CER = crossed extension reflex; AGR = automatic gait reflex; R3 = rater/3; R4 = rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Table 5

Agreement between the group I and II raters and the gold standard in the assessment of articular angles of 20 preterm newborns using the *Saint-Anne Dargassies* Scale.

Type of assessment	Interrater agreement			
	R1 and GS		R2 and GS	
	Agreement (%)	*p*	Agreement (%)	*p*
PA 1SDS	70	0.02	80	<0.01
PA 2SDS	80	<0.01	80	<0.01
PA 3SDS	100	<0.01	100	<0.01
PA 4SDS	100	<0.01	100	<0.01
FLA 1SDS	50	0.213	60	0.787
FLA 2SDS	70	0.02	90	<0.01
FLA 3SDS	80	<0.01	70	0.02
FLA 4SDS	80	<0.01	80	<0.01
HEA 1SDS	90	<0.01	100	<0.01
HEA 2SDS	90	<0.01	80	<0.01
HEA 3SDS	70	0.02	70	0.02
HEA 4SDS	80	<0.01	80	<0.01

PA = popliteal angle; FLA = foot-leg angle; HEA = heel-ear angle; R1 = rater/1; R2 = rater/2; R3 = rater/3; R4 =rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Table 5

Agreement between the group I and II raters and the gold standard in the assessment of articular angles of 20 preterm newborns using the *Saint-Anne Dargassies* Scale.

Type of assessment	Interrater agreement			
	R3 and GS		R4 and GS	
	Agreement (%)	*p*	Agreement (%)	*p*
PA 1SDS	20	0.299	50	0.213
PA 2SDS	50	0.213	50	0.213
PA 3SDS	60	0.076	60	0.076
PA 4SDS	80	<0.01	90	<0.01
FLA 1SDS	60	0.076	70	0.02
FLA 2SDS	60	0.076	70	0.02
FLA 3SDS	80	<0.01	80	<0.01
FLA 4SDS	100	<0.01	100	<0.01
HEA 1SDS	80	<0.01	70	0.02
HEA 2SDS	60	0.076	70	0.02
HEA 3SDS	90	<0.01	70	0.02
HEA 4SDS	60	0.076	30	0.599

PA = popliteal angle; FLA = foot-leg angle; HEA = heel-ear angle; R1 = rater/1; R2 = rater/2; R3 = rater/3; R4 =rater/4; GS = gold standard; SDS = *Saint-Anne Dargassies* Scale.

Discussion

In this study, three parameters exhibited an important relationship in defining healthy preterm infants, parameters such as birth weight for gestational age, APGAR score and premature birth etiology. In our study, birth weight for gestational age was considered adequate in 75% of the sample, suggesting good maturational evolution among these PTNBs.[22-24]

According to the literature, when the APGAR score is below seven following the 5th minute, special attention must be paid to these newborn infants, even in the absence of altered laboratory exams.[25] The PTNBs of this study exhibited 1 and 5 minute APGAR scores greater than or equal to 7, confirming that these newborns evolved satisfactorily.

Bittar and Zugab[26] reported that spontaneous prematurity accounts for 75% of cases and results from premature labor. Moreover, the etiology is considered complex or multifactorial or unknown, which hinders the implementation of preventive measures. The data from our study are similar to those obtained by these authors, as 12 of the participants (60% of the sample) went into premature labor without a secondary etiology.

The neurological examination is an important part of the newborn assessment, whether the infant in question is term or preterm, and is a useful tool for identifying newborns that require follow-up due to the risk of neurodevelopmental abnormalities.[27] It is important that the scale and the research method are used correctly to identify newborns suspected of having abnormalities and to differentiate these patients from those with normal development.

According to Noble and Boyd, there is increasing evidence regarding the impact of prematurity on brain development.[28] The authors report that approximately 10 to 15% of extreme PTNBs are diagnosed with cerebral palsy and that there is growing evidence of premature birth effects persisting into school age, adolescence and adulthood, which categorize premature birth as a severe public health problem.

Sampath *et al*.[29] report that even preterm newborns with normal cranial ultrasound results are susceptible to neurodevelopmental alterations, reinforcing the importance of performing the neurological exam in conjunction with valid and reliable instruments such as the SDS to identify infants suspected of having developmental disorders. The use of standardized assessments has become increasingly necessary in clinical practice. Studies have demonstrated that this need exists in all areas of medicine.[30] The choice of an adequate instrument to study PTNBs should be a concern for researchers, as these instruments must be both valid and reliable for this population.

The results of interrater agreement in the present

study demonstrated significant indexes in most of the items assessed by the SDS. An analysis of the interrater agreement noted in determining the postural patterns of the PTNBs demonstrates significant agreement between both the group I and II raters and the GS, indicating that the SDS is reliable in evaluating postural patterns.

The primitive reflexes of the 20 PTNBs also exhibited strong agreement between the group I and II raters and the GS. This significant agreement in assessing primitive reflexes further strengthened the reliability of this scale for the follow-up of these high-risk newborns.

Although the statistical analysis of the articular angles also demonstrated heterogeneity between the group I and II raters and the GS, this heterogeneity did not compromise the significant level of agreement, which also reinforced the reliability of this assessment instrument.

The interrater agreement observed in the utilization of the SDS may have resulted from differences in both the experience and the performance of the raters, in addition to previous training in the preparatory workshop that preceded the data collection.

The present study had some limitations. The first regards the sample of 20 preterm newborns; however, it is important to relate that we studied only preterm with no clinical or neurological complications, a factor that restricted the sample n

of the study. Thus, the homogeneous nature of the groups analyzed to determine reliability minimized this problem. Other factor should be considered the number of examiners (n=5) and only one of whom was considered GS. However, in spite of these limitations, SDS can be considered for future validations in studies, using more examiners to confirm our previously unpublished findings.

The strength of this study was to demonstrate that the SDS exhibits both high reliability and homogeneity and may be considered a scale with accurate reproducibility. Our results allow us to recommend it as an assessment instrument for PTNBs without either clinical or neurological complications. We also believe that the use of the SDS is a good reference for the follow-up of preterm infants, in addition to demonstrating high reliability.

In conclusion, our data pointed strong agreement among examiners using the SDS to screen pre-term newborn with high neurologic risk. We suggest the use of this SDS for the follow-up of PTNBs during prematurity period. It should be utilized every 15 days by professionals who work with preterm newborns. Given its reliability and practical feasibility, it will be useful to health professionals: pediatric neurologists, neonatologists, pediatricians, physiotherapists and neonatal nurses to identify premature newborns suspected of having neurological deficits.

References

1. Zomignani AP, Zambelli HJL, Antonio MRGM. Cerebral development in preterm newborn infants. Rev Paul Pediatr. 2009; 27 (2): 198-203.

2. Golin MO, de Souza FIS, Sarni ROS. Neonatal neurological assessment by the Dubowitz method in preterm infants at term corrected age compared with term infants. Rev Paul Pediatr. 2009; 27 (4): 402-9.

3. Procianoy RS, Guinsburg R. Avanços no manejo do recém-nascido prematuro extremo. J Pediatr. 2005; 8: S1-2.

4. Wang S, Liao C, Liang S, Zhong D, Liu J, Li Z. Ultrasound findings of mild neonatal periventricular-intraventricular hemorrhage after different treatments. Int J Clin Exp Med. 2015; 8(4): 5085-93.

5. Whyte HEA, Blaser S. Limitations of routine neuroimaging in predicting outcomes of preterm infants. Neuroradiology 2013; 55 (2): S3-S11.

6. De Vries LS, Benders MJNL, Groenendaal F. Imaging the premature brain: ultrasound or MRI? Neuroradiology 2013; 55(2): S13-S22.

7. Volpe J. Intrcranialhemorrage. In: W.B.Saunders Company (ed) Neurology of the Newborn. Pliladelphia. 2001; 397-427.

8. Guinsburg R,Peres CA, Shinzato AR, Kopelman BI. Reliability of two behavioral tools to assess pain in preterm neonates. São Paulo Med J. 2003; 121 (2): 72-6.

9. Hummel P, Lawlor-Klean P, Weiss MG. Validity and reliability of the N-PASS assessment tool with acute pain. J Perinatol. 2010; 30: 474-8.

10. Taddio A, Hogan ME, Moyer P, Girgis A, Gerges S, Wang L, Ipp M. Eavaluation of the reliability, validity and practicality of 3 measures of acute pain in infants undergoing immunization injections. Vaccine. 2011; 29 (7): 1390-4.

11. Deschênes G, Gosselin J, Couture M, Lachance C. Interobserver Reliability of the Amiel-Tison Neurological Assessment at Term. Pediatr Neurol. 2004. 30 (3): 190-4.

12. Simard MN, Lambert J, Lachance C, Audibert F, Gosselin J. Interexaminer Reliability of Amiel-Tison Neurological Assessments. Pediatr Neurol. 2009; 41 (5): 347-52.

13. Gagnon KL. Reliability and validity of the premie-neuro in detecting early neurodevelopmental delay and disability in preterm infants [Tese]. Kansas: Faculty of the University of Kansas; 2009.

14. Fernãndez D, Alvarez MJ, Rodríguez D, Rodríguez M, Fernãndez E, Urdiales P. Spanish Validation of the Premie-Neuro Scale in Premature Infants. J Pediatr Nurs. 2015; 30 (4): 560-7.

15. Leroux BG, Tich SN, Branger B, Gascoin G, Rouger V, Berlie I, Montcho Y, Ancel PY, Rozé JC, Flamant C. Neurological assessment of preterm infants for predicting neuromotor status at 2 years: results from the LIFT cohort. BMJ Open. 2013; 3 (2): 1-5.

16. Kottner J, Audigé, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, Roberts C, Shoukri M, Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011; 64 (1): 96-106.

17. Saint-AnneDargassies. Desarrollo Neurologico del Recien Nacido de Termino y Prematuro. Panamericana, Buenos Aires 1997; 305p.

18. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press; 1995.

19. Wusthoff C. How to use: the neonatal neurological examination. Arch Dis Child Educ Pract Ed. 2013; 98 (4): 148-53.

20. Almeida KM,Dutra MVP, de Mello RR, Reis ABR,Martins PS. Concurrent validity and reliability of the Alberta Infant Motor Scale in premature infants. J Pediatr. 2008; 84 (5): 442-8.

21. Alves CIS, Melo AN. The neurological examination of non-complicated preterm newborns using the Saint-Anne Dargassies Scale from birth to term. Arq Neuro-Psiquiatr. 2010; 68 (6): 893-7.

22. da Silva ES, Nunes ML. The influence of gestational age and birth weight in the clinical assessment of the muscle tone of healthy term and preterm newborns. Arq Neuro-Psiquiatr. 2005; 63: 956-62.

23. Van Kranen- Mastenewbornroeck VH, Folmer KB, Kingma H, Caberg H, Blanco CE, Hasaart TH et al. Postural behaviour of term SGA and AGA newborn infants. Dev Med Child Neurol. 1993; 35: 516-24.

24. Lagatta J, Yan K, Hoffmann R. The association between 5-min Apgar score and mortality disappears after 24 h at the borderline of viability. Acta Paediatr. 2012; 101 (6): 243-7.

25. Takazono PS, Golin MO. Asfixia Perinatal: Repercussões Neurológicas e Detecção Precoce. Rev Neurocienc. 2013; 21 (1): 108-17.

26. Bittar RE, Zugaib M. Indicadores de risco para o parto prematuro. Rev Bras Ginecol Obstet. 2009; 31 (4): 203-9.

27. Wusthoff CJ. How to use: the neonatal neurological examination. Arch Dis Child Educ Pract. 2013; 98: 148-53.

28. Noble Y, Boyd R. Neonatal assessments for the preterm infant up to 4 months correctedvelopmental Medicine & age: a systematic review. Dev MedChid Neurol. 2012; 54: 129-39.

29. Sampath V, Bowen J, Gibson F.Risk factors for adverse neurodevelopment in extremely low birth weight infants with normal neonatal cranial ultrasound. J Perinatol. 2005; 25: 210-5.

30. Gartner FR, de Miranda E, Rijnders ME, Freeman LM, Middeldorp JM, Bloemenkamp KWM, Stiggelbout AM, Marle MEA. Good reliability and validity for a new utility instrument measuring the birth experience, the labor and delivery index. J Clin Epidemiol. 2015; 68: 1-11.