

# Reliability of the Gross Motor Function Measure-66 scale in the evaluation of children with cerebral palsy: validation for Colombia

Diana M. Rivera-Rujana<sup>1\*</sup>, Diana I. Muñoz-Rodríguez<sup>2</sup>, and Maite C. Agudelo-Cifuentes<sup>2</sup>

<sup>1</sup>Departamento de Fisioterapia, Facultad de Ciencias de la Salud, Universidad del Cauca, Popayán, Cauca; <sup>2</sup>Facultad de Fisioterapia, Universidad CES, Medellín, Antioquia. Colombia

## Abstract

**Background:** Infantile cerebral palsy is the leading cause of physical disability in childhood and generates different alterations in motor development that prevent the child's independence. The Gross Motor Function Measure (GMFM) scale is considered the gold standard for this measurement in children with infantile cerebral palsy. In Colombia, its use is delayed due to its original language (English) and no studies on its validity in this specific field. This study aimed to determine whether cultural equivalence allows maintaining the reliability characteristics of the instrument to favor its use in the clinical setting.

**Methods:** We conducted a cross-sectional study that included 330 children with infantile cerebral palsy from three departments of Colombia, to whom the GMFM-66 scale was applied. Reliability was evaluated from interobserver consistency by estimating intraclass correlation coefficients and internal consistency with the omega coefficient ( $\omega$ ) or McDonald's test.

**Results:** The scale demonstrates consistency and stability in its measurements in terms of reliability. The internal consistency was satisfactory only for the first dimension, Lying and rolling ( $\omega = 0.91$ ). For the other dimensions, the  $\omega$ -value was always  $> 0.95$ . Good agreement was found among the experts in 83.3% of the items and dimensions evaluated. **Conclusions:** The GMFM-66 scale in Spanish and for the Colombian context demonstrates good psychometric properties and provides a better understanding of the motor development of children with infantile cerebral palsy so that it can be recommended for use in the Colombian context.

**Keywords:** Reliability. Cerebral palsy. Motor skills. Physical therapy modalities. Pediatrics. Children with disability.

## Confiability de la escala de Medición de la Función Motora Gruesa-66 en la evaluación de niños con parálisis cerebral: validación para Colombia

## Resumen

**Introducción:** La parálisis cerebral infantil es la principal causa de discapacidad física en la infancia y genera diferentes alteraciones en el desarrollo motor que impiden la independencia del niño. La escala de Medición de la Función Motora Gruesa (GMFM) se considera el método de referencia para esta medición en niños con parálisis cerebral. En Colombia, su uso está rezagado por el lenguaje original (inglés) y la carencia de estudios sobre su validez en este contexto. El objetivo de este estudio fue determinar si la equivalencia cultural permite mantener las características de confiabilidad del instru-

### Correspondence:

\*Diana M. Rivera-Rujana

E-mail: dianarivera@unicauca.edu.co

1665-1146/© 2021 Hospital Infantil de México Federico Gómez. Published by Permanyer. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Date of reception: 11-05-2021

Date of acceptance: 18-08-2021

DOI: 10.24875/BMHIM.21000094

Available online: 24-01-2022

Bol Med Hosp Infant Mex. 2022;79(1):33-43

[www.bmhim.com](http://www.bmhim.com)

mento para favorecer su uso en el ámbito clínico. **Métodos:** Se llevó a cabo un estudio transversal que incluyó 330 niños con parálisis cerebral infantil de tres Departamentos de Colombia, a quienes se aplicó la escala GMFM-66. Se evaluó la confiabilidad desde la consistencia interobservador mediante la estimación de coeficientes de correlación intraclasé y la consistencia interna con el coeficiente omega ( $\omega$ ) o prueba de McDonald. **Resultados:** En términos de confiabilidad, la escala demuestra consistencia y estabilidad en sus mediciones. La consistencia interna fue satisfactoria únicamente para la primera dimensión, Decúbito y rolado ( $\omega = 0.91$ ). Para las demás dimensiones, el valor de  $\omega$  siempre fue  $> 0.95$ . Se encontró un buen acuerdo entre los jueces en el 83.3% de los ítems y las dimensiones evaluadas. **Conclusiones:** La GMFM-66 en español y para el contexto colombiano demuestra buenas propiedades psicométricas y proporciona una mejor comprensión del desarrollo motor de los niños con parálisis cerebral infantil, por lo que se puede recomendar su uso en el contexto colombiano.

**Palabras clave:** Confiabilidad. Parálisis cerebral. Logros motores. Modalidades terapéuticas. Pediatría. Niños con discapacidad.

## Introduction

Infantile cerebral palsy (ICP) is the clinical description that shares the typical characteristics of a non-progressive brain injury acquired during the prenatal, perinatal, or early postnatal period<sup>1</sup>. It primarily affects the development of movement, tone, and posture, causing limitations in physical activity and different degrees of permanent disability, associated with multiple comorbidities and sensory, musculoskeletal, neurological, and behavioral deficits<sup>1-3</sup>. Management is, therefore, multidisciplinary and involves the physician working with a team of healthcare professionals from rehabilitation, orthopedic, psychology, and social work<sup>3</sup>.

ICP is currently the leading cause of physical disability in childhood, with an estimated 17 million people affected worldwide. Australia and Europe have reported a decrease in ICP cases from 1.99 and 2.5 cases in the last decade, respectively, to 1.77 cases per 1000 live births<sup>4,5</sup>. Conversely, recent studies in the United States, Taiwan, and Egypt have found rates above 3 per 1000<sup>1,2</sup>. Rates in low- and middle-income countries are uncertain; they appear to be higher, with projections of 4.4 to 10 cases per 1000<sup>1</sup>, probably due to higher burden of infectious diseases and deficiencies in prenatal and perinatal care<sup>6,7</sup>.

Meeting the needs of people with cerebral palsy in low-resource settings is one of the most significant challenges. Although the sequelae of this disorder affect individuals throughout their lives, most research efforts and management strategies focus on the needs of the pediatric population, their clinical management, and the prevention of secondary problems<sup>1,3,4</sup>. Resources must be justified in assessing needs, among which are those derived from developmental motor disorders, as they acquire relevance due to the negative impact on the quality of life of children, their families,

and the community, and the disability burden they generate for health systems<sup>1,5-7</sup>.

Therefore, gross motor function assessment instruments play a key role in identifying, diagnosing, and evaluating motor difficulties in childhood<sup>8</sup>. The Gross Motor Function Measure (GMFM) scale, in all its versions, has been the most researched measure, with the best results, the highest evidence of validity and response properties, making it recognized worldwide as the gold standard for quantitatively assessing changes in gross motor function in children with ICP<sup>9</sup>.

From the original version<sup>10</sup> of the GMFM-88, developed in Canada in English, later versions emerged, such as the GMFM-66 items<sup>11</sup> and, from this, the GMFM-66 Item Sets (GMFM-66-IS) and the GMFM-66 Basal & Ceiling (GMFM-66-B&C)<sup>12</sup>; both abbreviated versions demonstrated high levels of validity and reliability for clinical and research use. Finally, the Gross Motor Performance Measure (GMPM)<sup>13</sup>, developed for use together with the GMFM, assesses the quality of motor patterns rather than their quantity. The extended 66-item version, the subject of this research, proved to be equally sensitive in measuring both motor function<sup>10,14,15</sup> (even with reports of greater sensitivity than the 88-item version) and changes in motor function in response to treatment<sup>12</sup>. This feature allows determining the effectiveness of intervention strategies, planning them, and monitoring the child's motor development with ICP. However, although the high frequency of live births with a diagnosis of ICP is estimated for Colombia, the application of this scale has not been widely adopted<sup>11</sup>. Furthermore, it has not been documented whether cultural equivalence allows maintaining the instrument's reliability to facilitate its use in clinical settings in the Colombian context.

## Methods

### Study design

In this cross-sectional study, we evaluated the reliability of the 66-item version of the GMFM scale to determine the accuracy of the results obtained with its application when assessing the motor characteristics of children with ICP in the Colombian context. The Spanish version, resulting from the appearance validation focused on the cross-cultural component by Cobo et al.<sup>16</sup>, was used with the previous authorization of these authors.

### Subjects

The reference population was children with a medical diagnosis of ICP attended on an outpatient basis in different health care institutions, foundations, and different programs of the Colombian Institute of Family Welfare (Instituto Colombiano de Bienestar Familiar) in the departments of Nariño, Cauca, and Antioquia, in Colombia. An open invitation was made for children's selection through the communication channels of the participating universities during 2018 and 2019. Non-probabilistic techniques were used to include those who met the selection criteria until the estimated sample size was completed. The selection criteria included children between 5 months and 16 years of age, with a confirmed medical diagnosis of ICP, informed consent form signed by the parents or legal representative of the child, and acceptance of a responsible adult to accompany the minor during the evaluation process. Cases in which the information provided by the parent or guardian could not be confirmed or validated with the clinical history and cases with associated pathologies or disorders that prevented the application of the scale were excluded.

The sample size was calculated based on the factorial sampling criterion, which suggests having at least five subjects with the syndrome for each item contained in the instrument<sup>17</sup>. Therefore, for this study on the GMFM-66 scale, the number of patients included was 330 children who met the selection criteria previously described. This type of sampling has been suggested for validation studies since the representativeness of the participants is not necessary, but rather the sufficiency of the sample size to allow estimates of each item, among them, and the factors that contain them.

### Data collection and analysis

For the evaluation of the population, the researchers designed a questionnaire to collect the most relevant demographic and clinical characteristics (sex, age, medical diagnosis, type of ICP, risk factors, and motor development history), which were asked to parents or guardians and confirmed in the medical records. Subsequently, the 66-item version of the GMFM scale, considered the gold standard for obtaining an objective estimate of gross motor function in children with ICP, was applied. The GMFM-66 items are grouped into five dimensions as follows:

- Dimension A. Lying (supine and prone) and rolling: contains four items with tasks involving the upper limbs, head, and midline control.
- Dimension B. Sitting: contains 15 items that seek to demonstrate axial control of the head and trunk in an intermediate position both dynamically and statically.
- Dimension C. Crawling and kneeling: contains ten items that assess the ability to adopt the crawling and kneeling positions, maintain them, and move in these positions.
- Dimension D. Standing: contains 13 items that assess the transition to bipedal position, its maintenance with one and two feet, with and without the support of the upper limbs, and the control of the center of gravity in this position.
- Dimension E. Walking, running, and jumping: contains 24 items that assess the child's ability to perform anterior, lateral, and posterior gait, with different amplitude of the base of support, climb up and down steps, run a short distance, and jump with one or two feet<sup>16</sup>.

The scale was applied in the presence of a child's family member or caregiver and videotaped upon signed consent.

Each item of the GMFM-66 is scored on a four-category scale as follows: 0 indicates that the child is unable to initiate the activity; 1 indicates that the child initiates the task but performs less than 10% of the task completion; 2 indicates that the child completes the task, but partially, greater than 10% but does not achieve task completion; 3 indicates that the child has completed the task. We also included the qualification "Not evaluable," which refers to the fact that the item was not evaluated due to impossibility of being performed or due to the child's refusal—even though they showed skills that would allow at least a partial

execution. In this type of case, a score of 0 will correspond<sup>18</sup>.

### **Reliability of the scale**

For this study, clinical experts assessed the reliability of the GMFM scale and the accuracy of the results obtained with its application. With these measurements, the interobserver reliability of the scale was analyzed by estimating intraclass correlation coefficients (ICC) with a 95% confidence interval (CI) for each scale item, considering that the scores corresponded to categorical scales. In the scale proposed by Bland and Altman<sup>19,20</sup>, the degree of interobserver reliability is classified as poor or null (ICC < 0.20), mediocre (ICC 0.21-0.40), moderate (ICC 0.41-0.60), good (ICC 0.61-0.80), or very good (ICC 0.81-1.00).

The following criteria were established to select judges as clinical experts. Judges were required to have a background in neuropediatrics, previous and current (at the time of the study) experience with pediatric ICP patients (theoretical and clinical), and experience in applying ICP scales and semiology of at least one year. In order to link them, each university participating in the study was asked to suggest experts (professors and graduates) with these characteristics who were interested and could participate in the study. Subsequently, individual invitations were sent out explaining the intention of the study, the evaluation methodology, and the reporting of the results. Of four judges, two were external, and two were internal (linked to clinical teaching in neuropediatrics at the participating universities), but they were masked (not known to each other) to ensure the independence of the evaluation and avoid bias in their agreement.

It should be noted that the four observers independently analyzed the measurements performed on the children, and from them, assigned the score for each of the 66 items for the five dimensions of the scale.

Additionally, the internal consistency analysis was performed, for which the omega or McDonald's coefficient ( $\omega$ )<sup>21</sup> was used, considering that the level of measurement of the items was ordinal with a response scale from 0 to 3. The  $\omega$  coefficient has shown a better performance than Cronbach's alpha for the calculation of internal consistency when there is low variance in the responses to the items when they are not measured continuously and when the number of response alternatives is small, as in this case<sup>22</sup>. Values between

0.7 and 0.9 in the  $\omega$  coefficient are considered satisfactory for adequate internal consistency.

With the results of each evaluator, a summary table was designed in Excel® and then exported to the Statistical Package for the Social Sciences (SPSS) V 2.0, licensed by Universidad CES, where the statistical analysis was conducted.

This research was of minimal risk, and the ethics committee of each of the participating universities approved its execution.

## **Results**

### **Analysis of demographic and clinical variables**

This study included 330 participants with ICP, primarily males (62.1%), with a median age of 9 years (interquartile range (IQR) = 6-12.2 years). The most frequent clinical form of ICP was spastic (60%). The risk factors identified as significant were prolonged labor (24.2%), perinatal asphyxia (17.3%), and prenatal infection (12.8%).

Regarding the motor development of the included patients, it was found that cephalic control was initiated in most subjects at 12 months (IQR = 5.2-24.0 months), rolling at 17 months (IQR = 8.0-36.0), independent sitting together with crawling at 18 months (IQR = 9.0-36.0 and IQR = 8.0-36.0, respectively). Standing was achieved at 24 months (IQR = 11-48) and walking at 33.5 months (IQR = 13.5-54.2).

### **Descriptive analysis of the items**

For dimension A, *Lying (supine and prone) and rolling*, four motor tasks were evaluated, and the prevalent score was 3. Of the 330 children, 14% (n = 46) were unable to initiate the tasks, so the dimension score was zero (0); 270 children achieved all the tasks of the dimension. The remaining 14 children scored zero on three consecutive tasks, so the evaluation was concluded for this dimension.

For dimension B, *Sitting*, 15 motor tasks were assessed and the prevalent score was 3. Three children failed to perform any task of the dimension, making a statistical contribution of 267 children, of whom only 42% (n = 113) scored up to the last task.

For dimension C, *Crawling and kneeling*, ten motor tasks were assessed and the prevalent score was 3. One child failed to complete any task; of the remaining

112, 78% (n = 87) managed to complete the dimension with a score > 0.

For dimension D, *Standing*, 13 motor tasks were assessed and the prevalent score was 3, except for item *D57: While standing, lifts left foot without arm support for 10 seconds*, where the most prevalent score was 2. One child scored zero for the dimension. Of the remaining 86, 83% (n = 71) scored up to the last task.

For dimension E, *Walking, running, and jumping*, 24 motor tasks were evaluated. Due to the complexity of the activities, only 71 children (21% of the sample) were able to initiate the dimension, suggesting less neuro-motor compromise in these children. Although the prevalent score was 3, item *E74: While standing, walks ten consecutive steps forward on a straight line 2 cm wide* had a prevalent score of 2, and *E82: While standing, hops on the right foot 10 times within a 60 cm circle*, a prevalent score of 1. Further analysis is required to determine the behavior of these items within the sub-scale. Forty-nine children completed the dimension, corresponding to 69% of those who started it and 15% of the total study sample.

### Reliability analysis of the GMFM scale

Four observers independently evaluated the results of the GMFM scale in the 330 children to analyze the reliability of the instrument. Of the five dimensions evaluated by the scale, dimension E (*Walking, running, and jumping*) had a higher mean agreement in the evaluators' responses for the different items (0.90). However, it should be considered that this value was given mainly by the number of children who were unable to initiate each of the tasks and therefore had a score of zero.

The next dimension with the highest mean agreement among the judges was dimension C (*Crawling and kneeling*), with a value of 0.92. It was followed by dimension D (*Standing*), with a mean ICC of 0.89. Finally, dimensions B (*Sitting*) and A (*Lying and rolling*) obtained ICC values of 0.84 and 0.80, respectively. It is worth noting that the *Lying and rolling* dimension obtained the most significant variability in the initial state of the children; that is, most of them were able to initiate the tasks of this dimension, unlike dimensions C, D, and E, where most of them had a score of 0 for not initiating.

Regarding the *Lying and rolling* dimension (dimension A), the item with the highest agreement among the judges was *Turns head to both sides with symmetrical*

*limbs*, while the item *Raises head upright* showed low consistency among the responses of the four observers (Table 1).

Regarding dimension B, *Sitting*, it was observed that there were items with low agreement among the judges, such as *Remains seated without resting the arms for 3 seconds* with a correlation coefficient of 0.63, and *Descends to the prone position with control*, with an ICC of 0.67. In contrast, it was observed that in questions such as *Can sit on a low bench*, the agreement among the judges was satisfactory (ICC 0.93) (Table 1).

Meanwhile, the *Crawling and kneeling* dimension showed good agreement among judges' responses on all items, ranging from 0.84 (*Climbs four steps crawling on hands and knees/feet*) to 0.95 (*Supports on four points—crawling position—with weight on hands and knees*) (Table 1).

Agreement among the judges on the questions of the *Standing* dimension was variable. On items such as *Descends with control to sit on the floor without supporting the arms*, the consistency of the judges' responses was low, of 0.6. In contrast, on items such as *Achieves squat without supporting arms* and *Picks up an object from the floor and returns to standing without supporting arms*, the consistency in the responses was > 0.94 (Table 2).

Finally, in the dimension *Walking, running, and jumping*, few participants initiated the activities of the different items. For most of the items, the agreement among the judges was > 0.9, although it should be emphasized that, for most of the items, the child did not initiate the motor task (Table 2).

The internal consistency analysis was satisfactory only for the first dimension, *Lying and rolling*, composed of four items, whose internal consistency was 0.91 according to the McDonald  $\omega$  test, noting that it is not necessary to eliminate any of the items to improve its internal consistency. As for the correlations between the items, all were > 0.6, with the highest correlations observed between items A2–A6 (0.80) and A6 - A7 (0.79) (Table 1, Figure 1). Moreover, the item that correlated best with the total dimension was A6 (0.85), followed by A7 (0.82).

The second dimension (*Sitting*) comprises 15 items that showed a high internal consistency (McDonald's test  $\omega = 0.97$ ). This high consistency (close to 1) indicates that some items may measure almost the same aspect. More precisely, the correlations among the items allowed showing which of them may be making similar measurements. The highest correlations were observed between items B35–B36 (Pearson's

**Table 1.** Analysis of interobserver consistency and internal consistency between items in dimensions A, B, and C

Item	Analysis of agreement between judges		Internal consistency analysis	
	Intraclass correlation coefficient	Confidence Interval (CI)	Total item correlation	McDonald's $\omega$ test if the item is eliminated
<i>Dimension A: Lying (supine and prone) and rolling</i>				
A2	0.82	0.78-0.84	0.78	0.890
A6	0.80	0.77-0.83	0.85	0.875
A7	0.79	0.75-0.81	0.82	0.885
A10	0.77	0.73-0.80	0.75	0.908
<i>Dimension B: Sitting</i>				
B18	0.82	0.78-0.84	0.65	0.974
B21	0.77	0.73-0.80	0.59	0.974
B22	0.80	0.76-0.83	0.67	0.974
B23	0.76	0.72-0.79	0.79	0.972
B24	0.63	0.58-0.68	0.88	0.970
B25	0.90	0.88-0.91	0.89	0.970
B26	0.90	0.88-0.91	0.91	0.970
B27	0.89	0.86-0.90	0.90	0.970
B30	0.67	0.62-0.71	0.89	0.970
B31	0.89	0.86-0.90	0.89	0.970
B32	0.88	0.85-0.90	0.89	0.970
B34	0.91	0.89-0.92	0.87	0.970
B35	0.91	0.89-0.92	0.87	0.970
B36	0.93	0.92-0.94	0.84	0.971
B37	0.92	0.90-0.93	0.85	0.971
<i>Dimension C: Crawling and kneeling</i>				
C39	0.94	0.92-0.94	0.92	0.985
C40	0.93	0.91-0.93	0.95	0.984
C41	0.95	0.94-0.95	0.95	0.984
C42	0.94	0.93-0.95	0.96	0.984
C43	0.93	0.92-0.94	0.95	0.984
C44	0.92	0.90-0.93	0.94	0.984
C45	0.92	0.90-0.93	0.92	0.985
C46	0.84	0.80-0.86	0.89	0.986
C48	0.94	0.92-0.94	0.92	0.985
C51	0.92	0.90-0.93	0.89	0.986

Source: SPSS results viewer of data obtained from applying the Gross Motor Function-66 (GMFM-66) scale to each participant.

correlation coefficient = 0.95), B31–B32 (Pearson's correlation coefficient = 0.94), and B26-B27 (Pearson's correlation coefficient = 0.93) (Table 1, Figure 2). We also observed that the internal consistency did not decrease or improve with eliminating any item. However, the correlation with the total dimension was higher in items B26, B27, and B31.

The third dimension (*Crawling and kneeling*) comprises ten items, which also showed high internal consistency (McDonald test  $\omega$  = 0.99), reflecting that they may be measuring similar or the same aspects. High correlations were identified between items C42–C43 (Pearson's correlation coefficient = 0.96), C40–C41 (Pearson's correlation coefficient = 0.95), C43–C44 (Pearson's correlation coefficient = 0.94),

and C39–C40 (Pearson's correlation coefficient = 0.94) (Table 1, Figure 3). We did not identify that there could be an improvement in internal consistency by eliminating some items. However, the items with the highest correlations with the total dimension were C40 (total item correlation = 0.95) and C42 (total item correlation = 0.96).

The fourth dimension (*Standing*), composed of 13 items, showed an internal consistency of 0.99 according to the McDonald  $\omega$  test. Moreover, no improvement in the consistency of the total dimension was observed with the elimination of some items. The items with the highest correlations with each other were D63–D64 (Pearson's correlation coefficient = 0.97), D62–D63 (Pearson's correlation coefficient = 0.96),

**Table 2.** Analysis of interobserver consistency and internal consistency between items, for dimensions D and E

Item	Analysis of agreement between judges		Internal consistency analysis	
	Intraclass correlation coefficient	Confidence Interval (CI)	Total item correlation	McDonald's $\omega$ test if the item is eliminated
<i>Dimension D: Standing</i>				
D52	0.94	0.93-0.95	0.91	0.987
D53	0.93	0.92-0.94	0.95	0.987
D54	0.93	0.91-0.93	0.92	0.987
D55	0.92	0.90-0.93	0.94	0.987
D56	0.94	0.93-0.95	0.92	0.987
D57	0.86	0.83-0.87	0.85	0.988
D58	0.83	0.80-0.86	0.88	0.988
D59	0.91	0.89-0.92	0.95	0.987
D60	0.89	0.87-0.91	0.90	0.987
D61	0.92	0.90-0.93	0.92	0.987
D62	0.61	0.55-0.66	0.95	0.986
D63	0.95	0.93-0.95	0.96	0.986
D64	0.96	0.94-0.96	0.96	0.986
<i>Dimension E: Walking, running and jumping</i>				
E65	0.95	0.94-0.96	0.93	0.991
E66	0.95	0.93-0.95	0.91	0.991
E67	0.97	0.96-0.97	0.91	0.991
E68	0.94	0.93-0.95	0.94	0.991
E69	0.96	0.95-0.97	0.95	0.991
E70	0.95	0.94-0.96	0.96	0.991
E71	0.94	0.93-0.95	0.94	0.991
E72	0.95	0.94-0.96	0.95	0.991
E73	0.91	0.89-0.92	0.93	0.991
E74	0.90	0.88-0.91	0.94	0.991
E75	0.94	0.93-0.95	0.94	0.991
E76	0.93	0.91-0.94	0.92	0.991
E77	0.96	0.94-0.96	0.97	0.991
E78	0.95	0.94-0.96	0.95	0.991
E79	0.95	0.93-0.95	0.95	0.991
E80	0.94	0.92-0.95	0.94	0.991
E81	0.92	0.90-0.93	0.89	0.991
E82	0.84	0.81-0.86	0.79	0.991
E83	0.86	0.83-0.88	0.79	0.991
E84	0.95	0.94-0.96	0.92	0.991
E85	0.94	0.93-0.95	0.91	0.991
E86	0.92	0.91-0.93	0.83	0.991
E87	0.92	0.90-0.93	0.80	0.991
E88	0.86	0.83-0.88	0.72	0.992

Source: SPSS results viewer of data obtained from applying the Gross Motor Function-66 (GMFM-66) scale to each participant.

and D54–D55 (Pearson's correlation coefficient = 0.95) (Table 2, Figure 4). Regarding the correlation of each item with the total dimension, the highest correlation was observed for item D63 (total item correlation = 0.96) and the lowest for item D57 (total item correlation = 0.85).

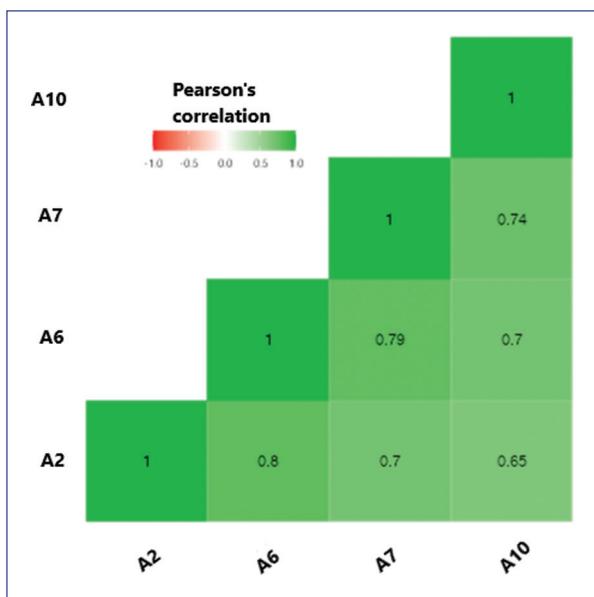
The fifth and last dimension comprises 24 items that showed an internal consistency of 0.99 according to the McDonald  $\omega$  test. This high consistency does not improve with the elimination of aWny item. When analyzing the correlations of the items with the total dimension, the highest correlations were observed for items E70 (total item correlation = 0.95) and E77 (total

item correlation = 0.97), while the lowest correlation was found for item E88 (total item correlation = 0.72) (Table 2).

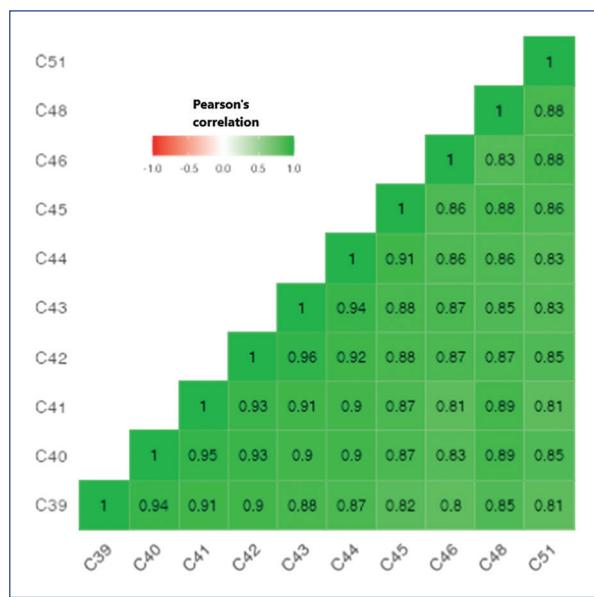
## Discussion

In this study, we found that the GMFM-66 scale shows good reliability when applied to the population of children with cerebral palsy in three departments of Colombia.

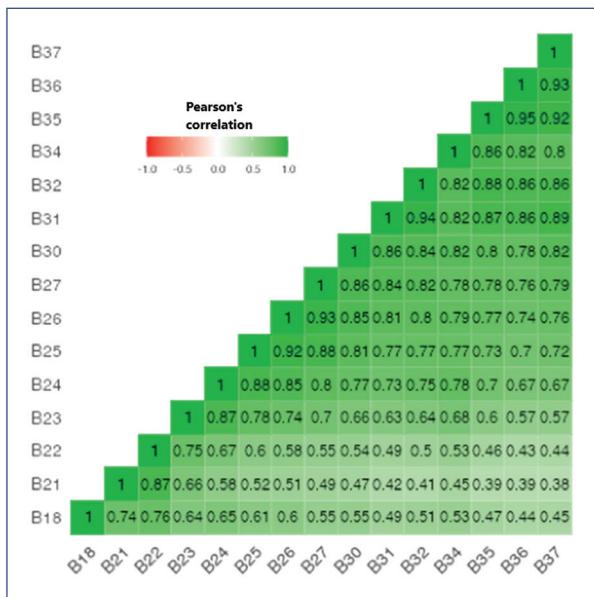
The clinical profile of Colombian children with ICP in this study is consistent with other reports in the



**Figure 1.** Analysis of the correlations between the items of dimension A, *Lying (supine and prone) and rolling*. Source: Created by the authors from SPSS results with Jamovi solid software version 1.6.23.

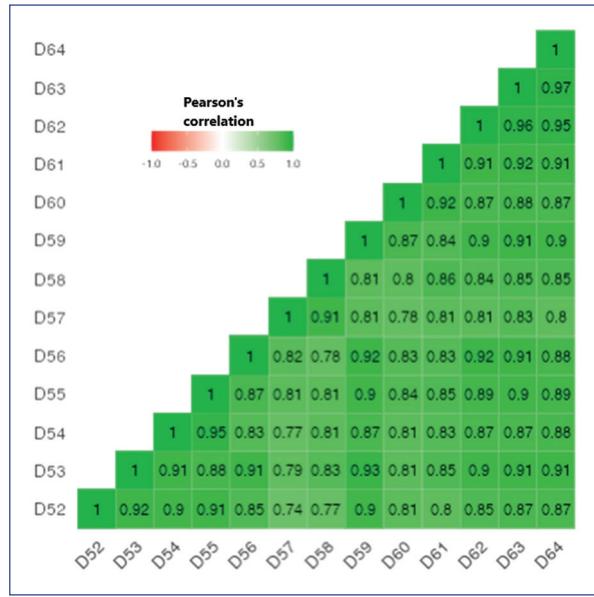


**Figure 3.** Analysis of the correlations among the items of dimension C, *Crawling and kneeling*. Source: Created by the authors from SPSS results using Jamovi 1.6.23 solid software.



**Figure 2.** Analysis of the correlations among the items of dimension B, *Sitting*. Source: Created by the authors from SPSS results using Jamovi 1.6.23 solid software.

international literature: higher frequency in males and spastic type according to tone deviation<sup>23</sup>. Furthermore, for the population in this study, neither low birth weight nor prematurity was different from worldwide reports



**Figure 4.** Analysis of the correlations among the items of dimension D, *Standing*. Source: Created by the authors from SPSS results using Jamovi 1.6.23 solid software.

that describe close and strong relationships of these factors with the development of ICP<sup>1,6,24-26</sup>.

According to international reports, all prenatal, perinatal, and postnatal antecedents found in our study are

related to the leading causes of ICP, which remains a crucial aspect in the prevention of this injury and its consequent disability<sup>1,2,26-29</sup>. Furthermore, secondary prevention plays a fundamental role in reducing the prevalence of ICP by preventing the development of brain injury<sup>1,6,25-28</sup> through the care of the pregnant woman, the neonate at risk, and early referral, which is one of the challenges of public health in developing countries such as Colombia.

Knowledge of the risk factors in the Colombian context would allow the establishment of promotion and prevention strategies for early identification and timely care, seeking to control and reduce the frequency of cases. In this context, the documented experience in Europe and Australia is based on efficient epidemiological surveillance and registration. Consequently, emerging strategies and policies were implemented to address and mitigate the impact of the determinants of health on pregnant women and newborns. Through this, they have managed to reduce the prevalence of ICP compared to reports from other areas of the world, becoming a current role model<sup>4,5</sup>.

Once the delay in motor development has been identified through early signs suggestive of neuromotor injury or through the follow-up of children with risk factors, it is essential to establish an accurate and timely diagnosis to initiate actions and strategies to mitigate the impact of the disability. In this process, evaluation is the pillar that represents the basis for identifying, diagnosing, planning, rethinking, and anticipating rehabilitation. The use of standardized, validated, and objective assessment tools enables accurate clinical information to be obtained and interventions to be adequately monitored<sup>30</sup>. The GMFM scale has been widely used in research to test the efficacy of therapeutic or surgical interventions performed on children with cerebral palsy, such as selective dorsal rhizotomy<sup>31</sup>, virtual reality<sup>32</sup>, robotics<sup>33</sup>, or equine therapies<sup>34</sup>. The severity of the damage, its relationship with the functional level, and the limitation for the development of gross motor activities (such as head control, rolling, sitting, assuming the quadruped position, and crawling to achieve the bipedal position and activities at this level) considerably affect the possibility of interacting with their environment in these children<sup>35</sup>. In this study, this situation was evidenced in the considerably delayed acquisition of motor skills for those who achieve them, and the low scores and poor task performance on the scale for those children with more significant compromise, and the inability to

perform tasks as the level of complexity on the scale increases for each subsequent dimension.

For all these reasons, assessment instruments are an essential and determining resource for collecting information and objectively measuring the achievement of the goals and objectives of therapeutic interventions. Instruments such as the Alberta Infant Motor Scale<sup>36</sup>, the Bruininks-Oseretsky Test of Motor Proficiency<sup>37</sup>, and the Peabody Developmental Motor Scale-2<sup>38</sup> are available. However, linguistic and cultural gaps between the countries in which these scales have been developed limit their application in the Colombian population since the validation process involves translation, semantic adaptation, and application to measure their validity and reliability in another context<sup>39</sup>.

In Colombia, there are no standardized scales that evaluate the motor development of children with special needs, and research development in this area is scarce. Currently, the only scale that has been validated and implemented by the Ministry of Health is the Abbreviated Developmental Scale Third Version (Escala Abreviada de Desarrollo, versión 3)<sup>40</sup>. This scale is used as a diagnostic screening for detecting developmental problems in childhood; however, it does not help to diagnose children with developmental disorders, such as ICP, leaving the evaluation, reevaluation, and diagnosis of motor function to medical and physiotherapy professionals. The above interposes biases in the monitoring and follow-up of the strategies applied to this population, generating increasingly higher costs to the health system due to the impact on the quality of family life derived from the permanent disability of those affected.

The result of the reliability analysis process of the scale for the Colombian context reiterated what has been found in multiple international publications of the original version: the GMFM is a valid, reliable, and sensitive observation instrument, widely used in different countries and serving as a reference for the development of other tests and classification systems, standing out as the most applicable in both research and clinical practice. This instrument has been adapted to different languages and cultures, with versions published in Korean, Dutch, Portuguese (Brazil), Norwegian, and Spanish<sup>41</sup>. In its 66-item version, the scale is valid for detecting changes in gross motor skills in children with ICP<sup>12,14</sup>. In addition, it showed excellent inter-observer reliability, reporting in our study an ICC between 0.92 and 0.97, being the lowest for dimension C (*Crawling and kneeling*) and the highest for the dimensions *Lying and rolling* and *Running and jumping* (subscale A and E, respectively). Russell et al. obtained an ICC between 0.87 and 0.99, being the

lowest for dimension A and the highest for dimension E, in addition to test-retest reliability, demonstrating high internal consistency<sup>9,10,13</sup>. This psychometric characteristic was not reported in this study since the children were assessed at a single time point. Other studies reported strong construct validity<sup>42,43</sup>. Given that the theoretical basis of the items represents this validity, this study did not aim to reevaluate it.

The degree of agreement of the expert evaluators was very good for 83.3% of the items. Only item D62: *While standing, squats to sit on the floor with control without supporting the arms* obtained a mediocre score. From the analysis of this item by the researchers, in consensus with the experts, it was concluded that the definition of the word “control” was confusing; for this reason, as it was understood from different perspectives, it was not scored in the same way. However, Russell et al. established the definitions of this term in the scale’s user manual<sup>44</sup>, so it is recommended to be defined and standardized in the Colombian context for use among clinicians.

The scale demonstrates consistency and stability in the measurements in terms of reliability. A good inter-judge agreement was found for the items and dimensions evaluated, showing that the GMFM-66 in Spanish and for the Colombian context demonstrates good psychometric properties and provides a better understanding of the motor development of children with ICP<sup>18</sup>.

We suggest deepening the validity since the variability of the scale was found only for the initial dimensions. This finding may be because, in patients with ICP, the other dimensions involve greater complexity in the execution of movements, so there is a significant proportion of missing data from children who failed to initiate any task in subscales D and E.

Finally, it should be noted that the scale is easy to use, the description of the items is clear, concrete, and the score is sensitive to the different degrees of functionality and types of ICP regardless of age. We recommend its application in a comfortable environment, prior training in the use of the scale, and having the necessary materials for its application, in addition to videotaping the evaluation to adequately assess and have the possibility of repeating the observation of each item. The application time will depend on the degree of commitment of the subject to be evaluated, which can vary from 30 to 90 minutes.

One of the limitations of this study was the decrease in sample size as the complexity of the items and dimensions increased. This decrease was because of the participant’s clinical conditions, as the ICP itself

limits the execution of complex tasks and those related to the latter dimensions. However, the sample size of each dimension was documented, from which the analyses were performed, and the estimation of the psychometric properties considered these losses in the interpretation of the findings, mainly in the last dimensions. In addition, the original scale considers that there are items that could not be evaluated due to the motor condition of the participants (the case of this study) and not because of missing data.

In conclusion, we showed the reliability of the GMF-66 in the assessment of motor function in children with ICP for the Colombian context, using (with authorization) a prior appearance validation performed by Cobo et al.<sup>16</sup> focused on the cross-cultural component. We found that cultural equivalence allows maintaining inter-rater consistency and internal consistency, and consequently, the instrument’s reliability. Therefore, we recommend using this scale in the evaluation and reevaluation of children with ICP and its implementation as a standardized scale in the national context.

To avoid problems of loss of sample size due to the clinical conditions of the participants, we suggest considering sample sizes that exceed the minimum number of participants for each item (more than five) in future studies.

## Ethical disclosures

**Protection of human and animal subjects.** The authors declare that no experiments were performed on humans or animals for this study.

**Confidentiality of data.** The authors declare that they have followed the protocols of their work center on the publication of patient data.

**Right to privacy and informed consent.** The authors have obtained the written informed consent of the patients or subjects mentioned in the article. The corresponding author has this document.

## Conflicts of interest

The authors declare no conflict of interest.

## Funding

For this study, funding was received from the Internal Development Project, Universidad del Cauca (VRI ID4688) and Convocatoria de Mínima Cuantía, Universidad CES (Imputation INV.022019.014).

## Acknowledgments

We thank the physiotherapists Angie Mesa, Catherine Ante, Elizabeth Roldan, Mabel Salazar, Mónica Osorio, and Rosmery Imbachi for participating as expert judges in this validation process. We also thank the physiotherapist Natalia Mejía and the physiotherapy students of the Universidad del Cauca and Universidad CES for their contributions in the labor field. Finally, we thank the Colombian institutions, parents, and children with cerebral palsy.

## References

- Graham HK, Rosenbaum P, Paneth N, Dan B, Lin JP, Damiano DL, et al. Cerebral palsy. *Nat Rev Dis Primers.* 2016;2:15082.
- Wimalasundera N, Stevenson VL. Cerebral palsy. *Pract Neurol.* 2016;16:184-94.
- Gulati S, Sondhi V. Cerebral palsy: an overview. *Indian J Pediatr.* 2018;85:1006-16.
- Sellier E, Platt MJ, Andersen GL, Krägeloh-Mann I, De La Cruz J, Cans C, et al. Decreasing prevalence in cerebral palsy: a multi-site European population-based study, 1980 to 2003. *Dev Med Child Neurol.* 2016;58:85-92.
- Sellier E, McIntyre S, Smithers-Sheedy H, Platt MJ, SCPE and ACPR Groups. European and Australian Cerebral Palsy Surveillance networks working together for collaborative research. *Neuropediatrics.* 2020;51:105-12.
- Novak I, Morgan C, Addle L, Blackman J, Boyd RN, Brunstrom-Hernandez J, et al. Early, accurate diagnosis and early intervention in cerebral palsy: advances in diagnosis and treatment. *JAMA Pediatr.* 2017;171:897-907.
- Afzali M, Etemad K, Kazemi A, Rabiei R. Cerebral palsy information system with an approach to information architecture: a systematic review. *BMJ Health Care Inform.* 2019;26:e100055.
- Griffiths A, Toovey R, Morgan PE, Spittle AJ. Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open.* 2018;8:e021734.
- Adair B, Said CM, Rodda J, Morris ME. Psychometric properties of functional mobility tools in hereditary spastic paraparesis and other childhood neurological conditions. *Dev Med Child Neurol.* 2012;54:596-605.
- Russell DJ, Rosenbaum PL, Cadman DT, Cowland C, Hardy S, Jarvis S. The Gross Motor Function Measure: a means to evaluate the effects of the physical therapy. *Dev Med Child Neurol.* 1989;31:341-52.
- Alotaibi M, Long T, Kennedy E, Bavishi S. The efficacy of GMFM-88 and GMFM-66 to detect changes in gross motor function in children with cerebral palsy (CP): a literature review. *Disabil Rehabil.* 2014;36:617-27.
- Brunton LK, Bartlett DJ. Validity and reliability of two abbreviated versions of the Gross Motor Function Measure. *Phys Ther.* 2011;91:577-88.
- Gowland C, Boyce WF, Wright V, Russell DJ, Goldsmith CH, Rosenbaum PL. Reliability of the Gross Motor Performance Measure. *Phys Ther.* 1995;75:597-602.
- Russell DJ, Avery LM, Walter SD, Hanna SE, Bartlett DJ, Rosenbaum PL, et al. Development and validation of item sets to improve efficiency of administration of the 66-item Gross Motor Function Measure in children with cerebral palsy. *Dev Med Child Neurol.* 2010;52:e48-54.
- Robles-Pérez DA, Rodríguez PDM, Zarco PMJ, Rendón FB, Mesa LC, Echevarría RDC. Versión española de la Gross Motor Function Measure (GMFM): fase inicial de su adaptación transcultural. *Rehabilitación.* 2009;43:197-203.
- Cobo-Mejía EA, Quino-Avila AC, Díaz-Vidal DM, Chacón-Serna MJ. Validez de apariencia del Gross Motor Function Measure-88. *Univ Salud.* 2014;16:47-57.
- Carretero-Díos H, Pérez C. Normas para el desarrollo y revisión de estudios instrumentales: consideraciones sobre la selección de tests en la investigación psicológica. *Int J Clin Heal Psychol.* 2005;5:521-51.
- Russell DJ, Avery LM, Rosenbaum PL, Raina PS, Walter SD, Palisano RJ. Improved scaling of the Gross Motor Function Measure for children with cerebral palsy: evidence of reliability and validity. *Phys Ther.* 2000;80:873-85.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;8:307-10.
- Delgado-Noguera MF, Merchán-Galvís ÁM, Mera-Mamián AY, Muñoz-Manquillo DM, Calvache JA. Evaluación de la calidad metodológica de las Guías Colombianas de Práctica Clínica en Pediatría. *Pediatría.* 2015;48:87-93.
- McDonald RP. *Test theory: a unified treatment.* New York: Psychology Press; 1999.
- Ventura-León J, Caycho-Rodríguez T. El coeficiente omega: un método alternativo para la estimación de la confiabilidad. *Rev Latinoam Cienc Soc Niñez Juv.* 2017;15:625-7.
- Gómez-Vega JC, Ocampo-Navia MI, Acevedo-González JC. Espasticidad. *Univ Med.* 2021;62:1-13.
- Downes KL, Grantz KL, Shenassa ED. Maternal, labor, delivery, and perinatal outcomes associated with placental abruption: a systematic review. *Am J Perinatol.* 2017;34:935-57.
- Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Prognostic factors for cerebral palsy and motor impairment in children born very preterm or very low birthweight: a systematic review. *Dev Med Child Neurol.* 2016;58:554-69.
- Spittle AJ, Morgan C, Olsen JE, Novak I, Cheong JLY. Early diagnosis and treatment of cerebral palsy in children with a history of preterm birth. *Clin Perinatol.* 2018;45:409-20.
- Korzeniewski SJ, Slaughter J, Lenski M, Haak P, Paneth N. The complex aetiology of cerebral palsy. *Nat Rev Neurol.* 2018;14:528-43.
- Shepherd E, Salam RA, Middleton P, Makrides M, McIntyre S, Badawi N, et al. Antenatal and intrapartum interventions for preventing cerebral palsy: an overview of Cochrane systematic reviews. *Cochrane Database Syst Rev.* 2017;8:CD012077.
- Eden RD, Evans MI, Britt DW, Evans SM, Schifrin BS. Safely lowering the emergency Cesarean and operative vaginal delivery rates using the Fetal Reserve Index. *J Matern Fetal Neonatal Med.* 2020;33:1473-9.
- Ponomarenko GN. [Physical therapy: prospects for the systematic development]. *Vopr Kurortol Fizioter Lech Fiz Kult.* 2017;94:59-64. Article in Russian.
- Morota N. Clinically practical formula for preoperatively estimating the cutting rate of the spinal nerve root in a functional posterior rhizotomy. *Childs Nerv Syst.* 2019;35:665-72.
- Arnoni JLB, Pavão SL, Dos Santos Silva FP, Rocha NACF. Effects of virtual reality in body oscillation and motor performance of children with cerebral palsy: a preliminary randomized controlled clinical trial. *Complement Ther Clin Pract.* 2019;35:189-94.
- Matsuda M, Iwasaki N, Mataki Y, Mutsuzaki H, Yoshikawa K, Takahashi K, et al. Robot-assisted training using Hybrid Assistive Limb® for cerebral palsy. *Brain Dev.* 2018;40:642-8.
- Yeo SM, Lee JY, Shin HY, Seo YS, Kwon JY. Factors influencing motor outcome of hippotherapy in children with cerebral palsy. *Neuropediatrics.* 2019;50:170-7.
- Schiariti V, Longo E, Shoshmin A, Kozhushko L, Besstrashnova Y, Król M, et al. Implementation of the International Classification of Functioning, Disability, and Health (ICF) Core Sets for Children and Youth with Cerebral Palsy: global initiatives promoting optimal functioning. *Int J Environ Res Public Health.* 2018;15:1899.
- Albuquerque PL, Guerra MQF, Lima MC, Eickmann SH. Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor development in preterm infants: a comparative study with the Bayley III. *Dev Neurorehabil.* 2018;21:408-14.
- Brown T. Structural validity of the Bruininks-Oseretsky test of motor proficiency—Second edition brief form (BOT-2-BF). *Res Dev Disabil.* 2019;85:92-103.
- Morgan C, Novak I, Dale RC, Badawi N. Optimising motor learning in infants at high risk of cerebral palsy: a pilot study. *BMC Pediatr.* 2015;15:30.
- García Correa HR, Sánchez DP, Ordoñez Mora LT. Escalas validadas para neurorrehabilitación en Hispanoamérica: revisión exploratoria. *Rehabilitación.* 2021;55:301-11.
- Pontificia Universidad Javeriana. Escala abreviada de desarrollo-3. Gobierno de Colombia; 2016. Available from: <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/Escala-abreviada-de-desarrollo-3.pdf>
- Ferre-Fernández M, Murcia-González MA, Ríos-Díaz J. Traducción y adaptación transcultural del Gross Motor Function Measure a la población española de niños con parálisis cerebral. *Rev Neurol.* 2020;71:177-85.
- Ferre-Fernández M, Murcia-González MA, Barnuevo Espinosa MD, Ríos-Díaz J. Measures of motor and functional skills for children with cerebral palsy: a systematic review. *Pediatr Phys Ther.* 2020;32:12-25.
- Josenby AL, Jarnlo G, Gummesson C, Nordmark E. Longitudinal construct validity of the GMFM-88 total score and goal total score and the GMFM-66 score in a 5-year follow-up study. *Phys Ther.* 2009;89:342-50.
- Russell DJ, Rosenbaum P, Wright M, Avery LM. Gross Motor Function Measure (GMFM-66 and GMFM-88) User's Manual. CanChild Centre for Childhood Disability Research. Hamilton: Cambridge University Press; 2013.