



## ORIGINAL ARTICLE

# The general movement checklist: A guide to the assessment of general movements during preterm and term age



Carolina Yuri Panvequio Aizawa <sup>a,\*</sup>, Christa Einspieler <sup>b</sup>,  
Fernanda Franoso Genovesi <sup>a</sup>, Silvia Maria Ibidi <sup>c</sup>, Renata Hydee Hasue <sup>a</sup>

<sup>a</sup> Universidade de So Paulo (USP), Faculdade de Medicina, Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional, So Paulo, SP, Brazil

<sup>b</sup> Medical University of Graz, Division of Phoniatrics, Graz, Austria

<sup>c</sup> Universidade de So Paulo (USP), Hospital Universitrio, Unidade de Neonatologia, Diviso de Clnica Pediatrca, So Paulo, SP, Brazil

Received 30 May 2020; accepted 1 September 2020

Available online 2 November 2020

### KEYWORDS

Cerebral palsy;  
Child development;  
Early intervention  
(education);  
General movements;  
Neurologic  
examination;  
Premature infant

### Abstract

**Objectives:** To develop a checklist describing features of normal and abnormal general movements in order to guide General Movement Assessment novices through the assessment procedure, to provide a quantification of General Movement Assessment; and to demonstrate that normal and abnormal GMs can be distinguished on the basis of a metric checklist score.

**Methods:** Three examiners used General Movement Assessment and the newly developed GM checklist to assess 20 videos of 16 infants (seven males) recorded at 31–45 weeks postmenstrual age (writhing general movements). Inter- and intra-scorer agreement was determined for General Movement Assessment (nominal data; Kappa values) and the checklist score (metric scale ranging from 0 to 26; Intraclass Correlation values). The scorers' satisfaction with the usefulness of the checklist was assessed by means of a short questionnaire (score 10 for maximum satisfaction).

**Results:** The scorers' satisfaction ranged from 8.44 to 9.14, which indicates high satisfaction. The median checklist score of the nine videos showing normal general movements was significantly higher than that of the eleven videos showing abnormal general movements (26 vs. 11,  $p < 0.001$ ). The checklist score also differentiated between poor-repertoire (median = 13) and cramped-synchronized general movements (median = 7;  $p = 0.002$ ). Inter- and intra-scorer agreement on (i) normal vs. abnormal general movements was good to excellent (Kappa = 0.68–1.00); (ii) the distinction between the four general movement categories was considerable to excellent (Kappa = 0.56–0.93); (iii) the checklist was good to excellent (ICC = 0.77–0.96).

\* Corresponding author.

E-mail: [carol.aizawa@gmail.com](mailto:carol.aizawa@gmail.com) (C.Y. Aizawa).

*Conclusion:* The general movement checklist proved an important tool for the evaluation of normal and abnormal general movements; its score may potentially document individual trajectories and the effect of therapeutic intervention.

© 2020 Sociedade Brasileira de Pediatria. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

The Prechtl General Movement Assessment (GMA) has a high predictive power for the neurodevelopmental outcome in preterm and term infants with risk factors. The GMA enables early identification of infants who are at increased risk for cerebral palsy,<sup>1–7</sup> minor neurological deficits,<sup>8</sup> cognitive impairments,<sup>9</sup> or autism spectrum disorders.<sup>10</sup> The assessment is based on visual Gestalt perception of videoed age-specific normal and abnormal movement patterns, and is non-invasive, cost-effective, and highly reliable, with inter-scoring reliability values of 89–93% and an average Kappa = 0.88.<sup>3,11</sup>

General Movements (GMs) are spontaneous movements with a rich and complex repertoire and a specific spatial-temporal organization. GMs can be detected from early fetal life onwards until the predominance of intentional and anti-gravity movements at 4–5 months of post-term age.<sup>2</sup>

GMs comprise the entire body and manifest themselves in a variable sequence of arm, leg, neck and trunk movements. They come and go gradually, with variable intensity and speed. Rotations and frequent slight variations of the direction of motion make them appear complex and smooth.<sup>12</sup> From term age until the second month post-term, GMs assume the form of “writhing movements”. They are very similar to that of preterm-age GMs, but have a reduced amplitude and speed.<sup>2,3,12</sup>

Preterm GMs and writhing movements are classified as normal or abnormal, i.e. poor-repertoire (PR), cramped-synchronized (CS) or chaotic (Ch) GMs.<sup>1,12</sup> CS GMs have a particularly high predictive value (70% sensitivity and 97% specificity) for spastic cerebral palsy.<sup>13,14</sup> They are characterized by rigidity and lack of fluency and elegance, and include almost simultaneous contraction and relaxation of limbs and trunk muscles.<sup>13,15</sup> Infants with Ch GMs typically develop CS GMs around term and have a high risk for spastic cerebral palsy.<sup>16</sup> PR GMs, on the other hand, are less predictive and rather unspecific.<sup>1,9</sup>

Sensitivity (97–98%) and specificity (89–91%)<sup>5,14</sup> of GMA are superior to cranial ultrasound (74% and 92%, respectively) and neurological examination (88% and 87%, respectively), but similar to magnetic resonance image performed at term age (86–100% and 89–97%, respectively).<sup>5,6</sup> Although writhing movements’ sensitivity (93%) and specificity (59%)<sup>14</sup> are slightly lower than fidgety movements’ (FM), it is no less important the identification of abnormal GMs, especially CS, which can persist until FM’s phase and indicate a worse prognosis.<sup>7,13</sup> This makes GMA a very important tool for early detection of neurodevelopmental impairment, which provides the rationale for early intervention and the aspired minimization of sequelae.

Although preventive health care and early diagnosis of neurodevelopmental disorders represent important actions of health programs, they are not sufficiently implemented in low- and middle-income countries (LMIC). This includes Brazil, where at least 37% of Caucasian and 61% of Indigenous and Afro-Brazilian infants are born into vulnerable families.<sup>17</sup> Because of its high reliability and cost- and time-efficiency, GMA can be the most appropriate evaluation tool for LMIC. Qualitative assessment by GMA provides important data on risk factors for alterations in neurodevelopment that are of major importance especially in these countries.<sup>18</sup> Nevertheless, quantification of GMA would improve a wider spectrum of data analysis and information about risks for altered neurodevelopment. A “GMA checklist” could also guide GMA beginners in applying the method.

The Motor Optimality Score (MOS) is a detailed GMA that evaluates age-specific motor repertoire including fidgety movements and other movements and postural patterns expected to be present in this period.<sup>7,12</sup> It is applied to several populations and has high reliability (0.80 to 0.94 intra-class correlation coefficients of inter-observer reliability).<sup>12</sup> However, it covers only infants at the fidgety movements phase, i.e. between three and five months of age. Therefore, we focus this work on a semi-quantitative assessment of GMs at preterm, term and early post-term age.

We therefore developed a checklist of features of normal and abnormal GMs to guide GMA novices through the assessment procedure and to provide a quantification of GMA. Our specific aims were to assess (a) the scorers’ satisfaction with a checklist provided as supplement to the classic GMA; (b) inter- and intra-scoring agreement for both classic GMA and the newly developed checklist; (c) the possibility to differentiate between normal and abnormal GMs based on a checklist score.

## Method

In this exploratory study three observers analyzed 20 videotapes of 16 infants applying the newly developed checklist and following a standardized assessment procedure.<sup>12</sup>

## Observers

The three observers, referred to as A, B, and C, are highly qualified neuropediatric physiotherapists. Observers A and B were certified after basic and advanced training courses and had used the method in clinical practice and ongoing research projects. Observer C successfully attended a basic training course but hardly used GMA. All observers received written instructions through a manual developed by the first author on how to apply the checklist, which were reinforced

orally before the assessment. None of them were familiar with the infants' medical histories. The scoring took place in the Neurofunctional Evaluation Laboratory of the Faculty of Medicine of the University of São Paulo (USP).

## Subjects

Nineteen videotapes of 15 infants, recorded between 31 weeks postmenstrual age and 5 weeks post-term age, were carried out between May and October 2014 at the Neonatology Intermediate Care Unit and at the Clinic of Early Intervention at the University Hospital of USP. Four preterm infants were recorded twice, at preterm and term age; the other 11 infants were recorded only once, totalizing 19 videos. Six recordings were performed <37 weeks, and 13 recordings at term and early post-term age. The intention was to include a diversified group of infants regarding both gestational age and risks for later neurological impairments, consecutively selected, and representative of all movement patterns (normal, PR and CS). Infants with osteoarticular disorders or congenital malformations were not included. As none of the infants showed chaotic GMs, one video recording was recruited from the GM Trust Medical Guide "Spontaneous Motor Activity as a Diagnostic Tool"<sup>19</sup> (case S, born at 30 weeks and recorded at 34 weeks). Hence, the final sample consisted of 20 videos of 16 infants (9 females).

Birth weight ranged from 1250 to 3450 g, and 8 infants were born preterm (gestational age: 28–35 weeks). Two infants had intraventricular hemorrhage; moderate or severe perinatal asphyxia was diagnosed in three infants. Eight infants presented risk factors for later neurodevelopmental impairment such as meningitis, intrauterine exposure to drug abuse, and/or syphilis.

The Ethics Committee of the University of São Paulo approved the study (Protocol CAPPEQ 091/14; 283/15) and all parents gave their written consent for the video recordings to be used for research purposes.

## Video recordings

In accordance with the Prechtl GMA,<sup>12,20</sup> infants were videotaped for 2–4 min in supine position, in active sleep or active wakefulness. Sequences that included fussing or crying were discarded. Infants were only wearing diapers or short-sleeved bodysuits to ensure optimal visualization of trunk and limbs' movements. Any interaction with the infants during the recording was avoided.

## The GM checklist

The current GMA mentor and senior tutor, Professor C. E., contributed to the elaboration of the checklist and authorized its publication in this study.

The first part of the checklist refers to core data such as the infant's name, identification (ID) number, date of birth, gestational age (GA), due date (for preterm infants), the mother's name, and recording age (Fig.1).

Part 2 (Fig. 1) starts with features related to normal GMs in preterm infants (<37 weeks postmenstrual age) and includes statements to be answered in the affirmative or

negative ("yes" or "no"). This part of the checklist also ends with a summary question to conclude if the GMs assessed would be considered normal or abnormal.

Part 3 (Fig. 2) of the checklist focuses on abnormal GMs in preterm infants and consists of other statements, which had to be answered with "yes" or "no". The final question consists of what pattern of abnormal GM would the scorer ascribe to the infant assessed: PR, CS, or Chaotic GMs.

Part 4 of the checklist addresses the assessment of infants at term and early post-term age (37 weeks postmenstrual to 5 weeks post-term age), starting with features of normal writhing GMs. Statements (1), (2), (3), (4), (8), and (9), as well as the final question are the same as for Part 2; new statements include: (5), (6), and (7) (Fig.3).

Part 5, about abnormal writhing GMs, is identical to Part 3 (Fig.2).

Each statement related to normal GMs (Parts 2 and 4) that is answered in the *affirmative* scores 1 point; each statement related to abnormal GMs (Parts 3 and 5) that is answered in the *negative* also scores 1 point. No point is given for statements related to normal features answered in the negative and for statements related to abnormal features answered in the affirmative. That makes a maximum checklist score of 26 for the best possible performance; the minimum score is 0, indicating the worst performance.

At the end of the checklist, a key for abbreviations is given. In addition, there are instructions as to how the Prechtl GMA<sup>12,20</sup> should ideally be performed.

## The assessment procedure

The scorers assessed the 20 videos in the same room on a large screen. They did not communicate during the scoring procedure. Upon request, they were allowed to see video sequences repeatedly. Each observer saw the video recordings the same number of times and for the same period of time. The average time taken to assess a video was 4 min.

To determine intra-observer reliability, each scorer reassessed the videos (presented in a different order) after 4 weeks.

Upon completing the second assessment, the scorers were given a form with five questions regarding their satisfaction with the checklist. They were asked whether the checklist facilitated (i) visualization of normal and abnormal GMs, (ii) the assessment procedure itself, and (iii) choosing a GM category, both for (iv) preterm GMs and (v) writhing GMs. The scorers replied on an 11-point semantic differential scale, with 0 indicating "not satisfied at all" and 10 indicating "highly satisfied".

## Statistics

Statistical analysis was performed using the SPSS package for Windows, version 22.0;  $p < 0.05$  indicated statistical significance. Descriptive analysis was performed to assess satisfaction with the checklist, using averages of the observers' answers for each question. Inter-observer and intra-observer agreement were determined for the GMA (nominal data) and the checklist score (metric scale). Intra-class correlation coefficient (ICC) statistics were applied for the checklist score to examine pairwise agreement

**THE GENERAL MOVEMENT CHECKLIST – PRETERM AND TERM AGE**

Name: \_\_\_\_\_ Register (ID Number): \_\_\_\_\_

Date of Birth: \_\_\_\_/\_\_\_\_/\_\_\_\_ Age: \_\_\_\_\_ Gestational Age: \_\_\_\_\_

Date of complete 40 weeks postmenstrual age: \_\_\_\_/\_\_\_\_/\_\_\_\_ Mother's name: \_\_\_\_\_

<b>Infants to be scored until 37 weeks postmenstrual age</b>	Date:		Date:		Date:	
	Age:		Age:		Age:	
	YES	NO	YES	NO	YES	NO
1. Movements consist of rotations superimposed on extensions and flexions of upper and lower limbs						
2. Movements occur in the whole space available						
3. The direction of movements varies						
4. The sequence of movements varies constantly and is unpredictable						
5. Especially the lower limbs move with a large amplitude						
6. Movements of the lower limbs are more obvious than those of the upper limbs						
7. Movement speed is higher than it will be around term						
8. Movements are initiated in one part of the body and spread to other parts like a wave						
9. Sequences set in and end gradually						
<b>In conclusion, do you consider these GMs as normal?</b>						

**Figure 1** Part 1 of the General Movement Assessment checklist, referring to core data of the infant, and Part 2 (in blue) referring to normal features of preterm GMs.

among the scorers A, B, and C and the overall agreement among all scorers. Values of ICC < 0.75 indicate poor agreement; ICC > 0.75 indicates good agreement; and ICC > 0.90 indicates excellent agreement.<sup>21</sup> To calculate the inter- and intra-observer agreement for the GMA, we performed Cohen’s Kappa, regarding Kappa < 0.20 as poor agreement, Kappa = 0.21–0.40 as slight agreement, Kappa = 0.41–0.60 as considerable agreement, Kappa = 0.61–0.80 as good agreement, and Kappa > 0.81 as excellent agreement.<sup>22</sup> The internal consistency of the checklist was analyzed by means of Cronbach’s alpha, considering alpha > 0.90 as excellent, alpha = 0.70–0.90 as good, and alpha = 0.6–0.7 as acceptable.<sup>23</sup> The Mann-Whitney *U* test was applied to compare two groups with nominal variables (e.g. GM categories) on one dependent variable (i.e. the checklist score).

## Results

The scorers’ satisfaction ranged from 8.44 to 9.14 on the 11-point semantic differential scale ranging from 0 to 10, which indicates high satisfaction.

The first author, who conceived the study, found the GMs in 9/20 videos (45%) to be normal and the GMs in 11 videos (55%) to be abnormal; 6 of the latter were assessed as PR, 4 as CS, 1 as Ch. Inter-scorer agreement with scorer A was excellent (Kappa = 0.92), and good agreement was reached

between the first author and scorers B (Kappa = 0.78) and C (Kappa = 0.77).

Inter-scorer agreement between A, B, and C regarding the differentiation between normal and abnormal GMs revealed a Kappa ranging from 0.68 to 0.80, which is a good result (Table 1); it is slightly higher than their agreement regarding the differentiation between normal GMs and the three abnormal categories, PR, CS, and Ch GMs (Table 1). As for the checklist, the three scorers reached an average agreement of ICC = 0.80 (95% CI: 0.63–0.91; Table 1).

Table 1 also demonstrates that the reassessment of the videos after 4 weeks resulted in good to excellent intra-scorer agreement, both in case of the checklist and the qualitative GMA.

The internal consistency of the checklist was excellent, with Cronbach’s alpha values ranging from 0.93 to 0.97.

Because inter-scorer agreement was not 100%, we compared A, B, and C’s checklist scores with the first author’s GM assessment. Five videos assessed as showing normal GMs reached the maximum score of 26 on the checklist (median = 26; P25 = 25, P75 = 26; range: 12–26). The 11 videos presenting abnormal GMs revealed a median checklist score of 10 (P25 = 8, P75 = 13; range: 3–25). Although there was a certain overlap in the outliers, the difference between normal and abnormal GMs was highly significant (Mann-Whitney *U* test, *p* < 0.001). Splitting the abnormal GMs

<b>Abnormal movements from preterm period to 37 weeks postmenstrual age</b>	Date:		Date:		Date:	
	Age:		Age:		Age:	
	YES	NO	YES	NO	YES	NO
10. Fluency is absent						
11. The sequence is monotonous						
12. Movements appear to be repetitive						
13. The sequence is suddenly interrupted						
14. The sequence is too short						
15. There is a lack of rotations						
16. The speed appears to be too low						
17. The quantity of GMs seems to be reduced						
18. Movements are too rigid						
19. Movements are abrupt						
20. Limb and trunk muscles appear to contract simultaneously and then relax simultaneously (on/off pattern)						
21. The posture appears to be stiff						
22. The child appears to be holding the breath due to longer-lasting contractions of the trunk, followed by an increase in the respiratory rate						
23. Cramped-synchronized movements are most evident in the lower limbs						
24. The movement amplitude is significantly high						
25. Movements are very fast and abrupt						
26. Tremulous movements are frequent						
<b>In conclusion, what pattern of abnormal GM would you ascribe to this infant: PR, CS, or Chaotic GMs?</b>						

Figure 2 Part 3 and Part 5 of the General Movement Assessment checklist, referring to abnormal features of preterm and writhing GMs.

<b>Infants to be scored from term age (37 to 40 weeks of postmenstrual age) until 5 weeks post-term age (so-called writhing movements)</b>	Date:		Date:		Date:	
	Age:		Age:		Age:	
	YES	NO	YES	NO	YES	NO
1. Movements consist of rotations superimposed on extensions and flexions of upper and lower limbs						
2. Movements occur in the whole space available						
3. The direction of movements varies						
4. The sequence of movements varies constantly and is unpredictable						
5. Compared to preterm age, the amplitude is smaller						
6. The writhing characteristic is obvious, especially in upper limbs						
7. The speed is moderate						
8. Movements are initiated in one part of the body and spread to other parts like a wave						
9. Sequences set in and end gradually						
<b>In conclusion, do you consider these GMs as normal?</b>						

Figure 3 Part 4 of the General Movement Assessment checklist, referring to normal features of writhing movements.

into PR, CS, and Ch GMs showed the median of PR GMs to be significantly higher (median = 13; P25 = 9, P75 = 14) than that of CS and Ch GMs (median = 7, P25 = 6, P75 = 10; Mann-Whitney *U* test, *p* = 0.002).

The results did not differ between GMs recorded at preterm age and writhing GMs recorded at term or early post-term age (Mann-Whitney *U* test, *p* = 0.600).

### Discussion

We developed a checklist of features of normal and abnormal GMs during preterm, term and early post-term age. The scorers' satisfaction with the checklist was high. Inter- and intra-scorer agreement on (i) normal vs. abnormal GMs was good to excellent; (ii) the distinction between the four GM

**Table 1** Inter-scorer agreement between scorers A, B, and C and intra-scorer agreement (after a 4-week interval) for GM categories (Kappa values) and for the checklist (ICC).

	Kappa (normal vs. abnormal)	Kappa (categories)	ICC (95% confidence interval)
Inter-scorer agreement			
Scorer A – Scorer B	K = 0.70 (good <sup>a</sup> )	K = 0.78 (good <sup>a</sup> )	ICC = 0.77 (0.51–0.90) (good <sup>b</sup> )
Scorer A – Scorer C	K = 0.80 (good <sup>a</sup> )	K = 0.69 (good <sup>a</sup> )	ICC = 0.85 (0.66–0.94) (good <sup>b</sup> )
Scorer B – Scorer C	K = 0.68 (good <sup>a</sup> )	K = 0.56 (considerable <sup>a</sup> )	ICC = 0.78 (0.52–0.91) (good <sup>b</sup> )
All scorers	K = (0.68–0.80) (good <sup>a</sup> )	K = (0.56–0.78) (considerable-good <sup>a</sup> )	ICC = 0.80 (0.63–0.91) (good <sup>b</sup> )
Intra-scorer agreement			
Scorer A	K = 0.90 (excellent <sup>a</sup> )	K = 0.85 (excellent <sup>a</sup> )	ICC = 0.89 (0.76–0.96) (good <sup>b</sup> )
Scorer B	K = 1.00 (excellent <sup>a</sup> )	K = 0.93 (excellent <sup>a</sup> )	ICC = 0.96 (0.91–0.99) (excellent <sup>b</sup> )
Scorer C	K = 0.89 (excellent <sup>a</sup> )	K = 0.69 (good <sup>a</sup> )	ICC = 0.89 (0.75–0.96) (good <sup>b</sup> )

ICC, Intra-class correlation coefficient; K, Cohen Kappa.

<sup>a</sup> According to Landis et al.<sup>22</sup>.

<sup>b</sup> According to Portney et al.<sup>21</sup>.

categories was considerable to excellent; (iii) the checklist score was good to excellent. The median checklist score of the nine videos showing normal GMs was significantly higher than that of the eleven videos showing abnormal GMs, and the checklist score provided differentiated scoring between PR and CS GMs.

The GMA was introduced by Prechtl in 1990<sup>24</sup> and saw its breakthrough as a systematic, valuable and reliable tool for the assessment of the integrity and function of the young nervous system seven years later.<sup>1</sup> Since then, regular training courses have mainly been provided in Europe, but have also spread to the Americas, Australia, Asia and South Africa in recent years. Valentin et al. evaluated the first 18 training courses held between 1997 and 2002 and reported that 83% of more than 8000 assessments were performed correctly after a mere 4 days of training.<sup>25</sup> Although training has proved to be effective, novices often struggle when left alone, especially with difficult cases. Bernhardt et al. reported that inter-scorer agreement was only fair to substantial (Kappa < 0.60) if a scorer was not experienced enough, regardless of their GMA certificate.<sup>26</sup> We thought that a checklist specifying the most important features of normal and abnormal GMs might be helpful, especially for GMA performed in LMIC, where training is still rare and the GMA could prove a very helpful alternative diagnostic tool when sophisticated imaging techniques are unavailable or unaffordable.<sup>5,27</sup>

Einspieler and colleagues have introduced an optimality score for GMs recorded at preterm, term and early post-term age that covers criteria such as amplitude, speed, range in space, onset and offset, etc.<sup>16</sup> So far, this detailed GMA is only used in research and by experts, as its implementation requires the successful and certified attendance of an advanced GM training course. By contrast, our checklist has proved a help to GMA novices in their clinical practice with the three scorers' high satisfaction with its application. These seem like a good enough reason to recommend the method. Furthermore, as detailed GMA (MOS) for assessment at a fidgety age, our checklist might be useful as an instrument with good predictive value to assessment at preterm and writhing movements age.

As GMA is based on visual Gestalt perception (pattern recognition), which is a powerful tool when it comes to the

analysis of complex motor phenomena, it is of the utmost importance that inter-scorer agreements are high (89–93% among experienced observers).<sup>3</sup> Only few studies reported on the reliability for preterm and/or term age. Mutlu et al.<sup>28</sup> observed excellent agreement among three observers who rated the GMs of 25 preterm infants (Kappa = 0.85; 95% CI: 0.46–1.00), and of 31 writhing GMs (Kappa = 0.94; 95% CI: 0.55–1.00).<sup>28</sup> However, these values only reflect the comparison of normal vs. abnormal GMs and do not apply to the subcategories of abnormal GMs. Our Kappa values for normal vs. abnormal GMs reach 0.68–0.80 (good agreement), but our scorers were far less experienced than those of the Mutlu study, who were well-versed in GMA, including even one instructor of the method.<sup>28</sup> Our scorers were probably about as experienced as those in the study by Bernhardt et al.,<sup>26</sup> but achieved much higher Kappa values than the latter. Their Kappa values ranged from 0.20 to 0.63 for the normal vs. abnormal ratings and from 0.16 to 0.60 for the four response category ratings (normal, PR, CS, Ch), which indicates only poor to considerable agreement,<sup>26</sup> while our results reflect good to excellent agreement for normal vs. abnormal GMs and considerable to excellent agreement for the four categories.

Our Kappa values for intra-scorer agreement after the 4-week interval were excellent for normal vs. abnormal ratings and good to excellent for the four response category ratings. Again, the intra-scorer agreement results published by Bernhardt et al. were partly lower, demonstrating only slight to good agreement (Kappa = 0.30–0.78) for normal vs. abnormal ratings and slight to excellent agreement (Kappa = 0.25–0.82) for the four subcategories.<sup>26</sup> There's no telling whether these differences are due to a longer intra-scoring interval (9-week interval) or the scorers' varying experience.

The reliability studies on detailed GMA revealed good to excellent agreement, with ICCs ranging from 0.87<sup>29</sup> to 0.98.<sup>30</sup> However, these studies focused on fidgety movements, and consequently, on different features than our checklist. The Motor Optimality Score considers the presence (or absence) of fidgety movements (period between three and five months of age) and also evaluates the concurrent motor repertoire at this age.<sup>7,12</sup> Our GMA checklist focuses on main features of normal and abnormal GMs at

preterm and writhing movements age (period until two months), mainly the presence of variability, complexity and fluency of movement, described in items, important to guide GMA novices through the assessment procedure. With an overall ICC = 0.80, we reached good agreement applying our checklist, which—along with a high satisfaction score—encourages us to further promote its implementation, especially by less experienced GM assessors. Besides, the possibility of converting the nominal GMA into numeric data allows more robust statistical analysis on future researches.

Another advantage is that the checklist score clearly distinguishes between normal and abnormal GMs, and between PR and CS GMs. Especially the latter might be clinically relevant, as CS GMs have a high predictive power for bilateral<sup>7</sup> spastic cerebral palsy<sup>1,7,13</sup> and are strongly associated with Gross Motor Function Classification System (GMFCS) outcomes III-V;<sup>7</sup> whereas PR GMs, albeit abnormal, are less predictive.<sup>1,3,9,10,12</sup> PR GMs can either normalize<sup>10</sup> or deteriorate into CS GMs.<sup>1</sup>

The present study has some limitations. First, our sample size is small. Although our checklist score has produced some significant results for the distinction between normal and abnormal GMs, these results need to be confirmed in a larger group of infants. A second limitation is that the scorers' satisfaction with the usefulness of the checklist was based on a mere five questions. Yet we consider our scorers representative of colleagues who just finished their GMA training and have different experiences.

In conclusion, we developed a checklist of 26 statements that describe features of normal and abnormal GMs during preterm, term and early post-term age. It adds to qualitative assessments by guiding scorers through the evaluation procedure and helping them to make decisions, as evidenced by the scorers' high satisfaction with the applicability of this new tool. Inter- and intra-scorer agreement was as high as in similar studies, or higher. As a supplement to the complex categorical GMA, the GM checklist clearly facilitates assessment for novices. Further research on a larger sample is needed to demonstrate whether our checklist score can capture small changes in the quality of GMs related to therapeutic interventions. In any case, a metric scale facilitates comparison with other quantitative assessments, which is extremely important for monitoring individual neurodevelopmental trajectories.

## Funding

This study was financed in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – 2013/21041-7 to C.Y.P.A.).

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

This study was financed in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – 2013/21041-7 to C.Y.P.A.). We also would like to thank all physiothera-

pists who helped to collect relevant data; Alexandra Siqueira Colombo, chief of the Division of Physiotherapy, Speech Therapy and Occupational Therapy, and the nurses of the Neonatology Division at the University Hospital of the University of São Paulo; the parents, who permitted us to record their infants; and Miha Tavcar (scriptophil), who assisted in copyediting the paper.

## References

1. Prechtl HF, Einspieler C, Cioni G, Bos AF, Ferrari F, Sontheimer D. An early marker for neurological deficits after perinatal brain lesions. *Lancet*. 1997;349:1361–3.
2. Prechtl HF. General movement assessment as a method of developmental neurology: new paradigms and their consequences. The 1999 Ronnie MacKeith lecture. *Dev Med Child Neurol*. 2001;43:836–42.
3. Einspieler C, Prechtl HF. Prechtl's assessment of general movements: a diagnostic tool for the functional assessment of the young nervous system. *Ment Retard Dev Disabil Res Rev*. 2005;11:61–7.
4. Noble Y, Boyd R. Neonatal assessments for the preterm infant up to 4 months corrected age: a systematic review. *Dev Med Child Neurol*. 2012;54:129–39.
5. Bosanquet M, Copeland L, Ware R, Boyd R. A systematic review of tests to predict cerebral palsy in young children. *Dev Med Child Neurol*. 2013;55:418–26.
6. Novak I, Morgan C, Adde 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.
7. Einspieler C, Bos AF, Krieger-Tomantschger M, Alvarado E, Barbosa VM, Bertocelli N, et al. Cerebral palsy: early markers of clinical phenotype and functional outcome. *J Clin Med*. 2019;8:1616.
8. Bruggink JL, Einspieler C, Butcher PR, Van Braeckel KN, Prechtl HF, Bos AF. The quality of the early motor repertoire in preterm infants predicts minor neurologic dysfunction at school age. *J Pediatr*. 2008;153:32–9.
9. Einspieler C, Bos AF, Libertus ME, Marschik PB. The General Movement Assessment Helps Us to Identify Preterm Infants at Risk for Cognitive Dysfunction. *Front Psychol*. 2016;7:406.
10. Einspieler C, Sigafoos J, Bölte S, Bratl-Pokorny KD, Landa R, Marschik PB. Highlighting the first 5 months of life: general movements in infants later diagnosed with autism spectrum disorder or Rett Syndrome. *Res Autism Spectr Disord*. 2014;8:286–91.
11. Einspieler C, Utsch F, Brasil P, Panvequio Aizawa CY, Peyton C, Hyde Hasue R, et al. Association of infants exposed to prenatal Zika virus infection with their clinical, neurologic, and developmental status evaluated via the general movement assessment tool. *JAMA Netw Open*. 2019;2:e187235.
12. Einspieler C, Prechtl HF, Bos AF, Ferrari F, Cioni G. Prechtl's method on the qualitative assessment of general movements in preterm, term and young infants. *Clin Dev Med*. 2004;167:1–91.
13. Ferrari F, Cioni G, Einspieler C, Roversi MF, Bos AF, Paolicelli PB, et al. Cramped synchronized general movements in preterm infants as an early marker for cerebral palsy. *Arch Pediatr Adolesc Med*. 2002;156:460–7.
14. Kwong AK, Fitzgerald TL, Doyle LW, Cheong JL, Spittle AJ. Predictive validity of spontaneous early infant movement for later cerebral palsy: a systematic review. *Dev Med Child Neurol*. 2018;60:480–9.
15. Yang H, Einspieler C, Shi W, Marschik PB, Wang Y, Cao Y, et al. Cerebral palsy in children: movements and postures during early

- infancy, dependent on preterm vs. full term birth. *Early Hum Dev.* 2012;88:837–43.
16. Einspieler C, Marschik PB, Pansy J, Scheuchenegger A, Kriebler M, Yang H, et al. The general movement optimality score: a detailed assessment of general movements during preterm and term age. *Dev Med Child Neurol.* 2016;58:361–8.
  17. United Nations International Children's Emergency Fund (UNICEF) [<https://www.unicef.org/brazil>]. Infância e adolescência no Brasil - Censo Demográfico 2010 [cited 2020 Apr 13]. Available from: <https://www.unicef.org/brazil/situaca-das-criancas-e-dos-adolescentes-no-brasil>.
  18. Tomantschger I, Herrero D, Einspieler C, Hamamura C, Voos MC, Marschik PB. The general movement assessment in non-European low- and middle-income countries. *Rev Saude Publica.* 2018;52:6.
  19. Prechtl HF, Einspieler C, Cioni G, Bos AF, Ferrari F. Spontaneous motor activity as a diagnostic tool. Functional assessment of the young nervous system. A scientific illustration of Prechtl's method. Demonstration video. London, Graz: The GM Trust; 1997.
  20. Einspieler C, Prechtl HF, Ferrari F, Cioni G, Bos AF. The qualitative assessment of general movements in preterm, term and young infants-review of the methodology. *Early Hum Dev.* 1997;50:47–60.
  21. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 3rd ed. Upper Sadle River, NJ: Prentice Hall; 2008.
  22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–74.
  23. George D, Mallery P. *SPSS for windows step by step: a simple guide and reference, 11.0 update.* 4th edition Boston: Allyn and Bacon; 2003.
  24. Prechtl HF. Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Hum Dev.* 1990;23:151–8.
  25. Valentin T, Uhl K, Einspieler C. The effectiveness of training in Prechtl's method on the qualitative assessment of general movements. *Early Hum Dev.* 2005;81:623–7.
  26. Bernhardt I, Marbacher M, Hilfiker R, Radlinger L. Inter- and intra-observer agreement of Prechtl's method on the qualitative assessment of general movements in preterm, term and young infants. *Early Hum Dev.* 2011;87:633–9.
  27. Soleimani F, Badv RS, Momayezi A, Biglarian A, Marzban A. General movements as a predictive tool of the neurological outcome in term born infants with hypoxic ischemic encephalopathy. *Early Hum Dev.* 2015;91:479–82.
  28. Mutlu A, Einspieler C, Marschik PB, Livanelioglu A. Intra-individual consistency in the quality of neonatal general movements. *Neonatology.* 2008;93:213–6.
  29. Fjørtoft T, Einspieler C, Adde L, Strand LI. Inter-observer reliability of the "Assessment of Motor Repertoire-3 to 5 Months" based on video recordings of infants. *Early Hum Dev.* 2009;85:297–302.
  30. Zang FF, Yang H, Han Q, Cao JY, Tomantschger I, Kriebler M, et al. Very low birth weight infants in China: the predictive value of the motor repertoire at 3 to 5 months for the motor performance at 12 months. *Early Hum Dev.* 2016;100:27–32.