



HAL
open science

General Movement Assessment from videos of computed 3D infant body models is equally effective compared to conventional RGB Video rating

A. Sebastian Schroeder, Nikolas Hesse, Raphael Weinberger, Uta Tacke, Lucia Gerstl, Anne Hilgendorff, Florian Heinen, Michael Arens, Linze Dijkstra, Sergi Pujades, et al.

► To cite this version:

A. Sebastian Schroeder, Nikolas Hesse, Raphael Weinberger, Uta Tacke, Lucia Gerstl, et al.. General Movement Assessment from videos of computed 3D infant body models is equally effective compared to conventional RGB Video rating. *Early Human Development*, 2020, 144, pp.104967. 10.1016/j.earlhumdev.2020.104967 . hal-02988419

HAL Id: hal-02988419

<https://inria.hal.science/hal-02988419v1>

Submitted on 4 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **General Movement Assessment from videos of computed 3D infant body models is**
2 **equally effective compared to conventional RGB Video rating**

3

4 Schroeder, A. Sebastian^{1*} ; Hesse, Nikolas^{2*,#} ; Weinberger, Raphael^{1,5} ; Tacke, Uta^{1,3} ;
5 Gerstl, Lucia¹ ; Hilgendorff, Anne¹ ; Heinen, Florian¹ ; Arens, Michael² ; Dijkstra, Linze J.⁶
6 Pujades Rocamora, Sergi⁷ ; Black, Michael⁷ ; Bodensteiner, Christoph² ; Hadders-Algra,
7 Mijna⁶

8

9 ¹ Ludwig Maximilian University of Munich (LMU), Hauner Children's Hospital, Department of
10 Paediatric Neurology and Developmental Medicine, Munich, Germany

11 ² Swiss Children's Rehab, University Children's Hospital Zurich, Affoltern am Albis, Switzerland

12 ³ University Children's Hospital, Basel, Switzerland

13 ⁴ Ludwig Maximilian University of Munich (LMU), Comprehensive Developmental
14 Care (CDeC), Munich, Germany

15 ⁵ Ludwig Maximilian University of Munich (LMU), Department of Epidemiology, Institute of
16 Social Paediatrics and Adolescent Medicine, Munich, Germany

17 ⁶ University of Groningen, University Medical Center Groningen, Department of
18 Paediatrics, Groningen, Netherlands

19 ⁷ Max Planck Institute for Intelligent Systems, Tübingen, Germany

20

21 *Authors contributed equally

22 # work was done while NH was with Fraunhofer Institute of Optronics, System Technologies
23 and Image Exploitation, Ettlingen, Germany.

24 All authors have approved the final article

25 Declarations of interest: none

26 **Corresponding author:**

27 A. Sebastian Schroeder

28 Ludwig Maximilian University of Munich (LMU), Hauner Children's Hospital, Paediatric
29 Neurology, Developmental Medicine, Lindwurmstr. 4, 80337 Munich

30 Phone: +49 89 4400 55110

31 Fax: +49 89 4400 55111

32 Email: Sebastian.Schroeder@med.lmu.de

33

1 **General Movement Assessment from videos of computed 3D infant body models is**
2 **equally effective compared to conventional RGB Video rating**

3

4 **Highlights (3-5 bullet points in an extra document):** maximum 85 characters, including
5 spaces, per bullet point

- 6 • Presentation of markerless 3D capture and modelling of general movements
7 • High agreement between conventional and model-based General Movement Assessment
8 • Prediction of CP of conventional and model-based GMA is equally good.

9

Preprint Manuscript Accepted

1 **Abstract:**

2 **Background:** General Movement Assessment (GMA) is a powerful tool to predict Cerebral
3 Palsy (CP). Yet, GMA requires substantial training challenging its broad implementation in
4 clinical routine. This inspired a world-wide quest for automated GMA.

5 **Aim:** To test whether a low-cost, marker-less system for three-dimensional motion capture
6 from RGB depth sequences using a whole body infant model may serve as the basis for
7 automated GMA.

8 **Study design:** Clinical case study at an academic neurodevelopmental outpatient clinic.

9 **Subjects:** Twenty-nine high risk infants were assessed at their clinical follow-up at 2-4 month
10 corrected age (CA). Their neurodevelopmental outcome was assessed regularly up to 12-31
11 months CA.

12 **Outcome measures:** GMA according to Hadders-Algra by a masked GMA-expert of
13 conventional and computed 3D body model ("SMIL motion") videos of the same GMs.
14 Agreement between both GMAs was tested using dichotomous and graded scaling with Kappa
15 and intraclass correlations, respectively. Sensitivity and specificity to predict CP at ≥ 12 months
16 CA were assessed.

17 **Results:** Agreement of the two GMA ratings was moderate-good for GM-complexity ($\kappa=0.58$;
18 ICC=0.874 [95%CI 0.730;0.941]) and substantial-good for fidgety movements (FMs;
19 Kappa=0.78, ICC=0.926 [95%CI 0.843;0.965]). Five children were diagnosed with CP (four
20 bilateral, one unilateral CP). The GMs of the child with unilateral CP were twice rated as mildly
21 abnormal with FMs. GM-complexity and somewhat less FMs, of both conventional and SMIL
22 motion videos predicted bilateral CP comparably to published literature.

23 **Conclusions:** Our computed infant 3D full body model is an attractive starting point for
24 automated GMA in infants at risk of CP.

25

26 **Keywords:** General Movement Assessment (GMA), high risk infants, early diagnosis, fidgety,
27 automated motion analysis, cerebral palsy, Kinect

28

29

30

31

1 **Abbreviations:**

2 BSID: Bayley Scales of Infant Development

3 CA: Corrected Age

4 CP: Cerebral Palsy

5 DA: Definitely Abnormal GMA rating

6 FMs: Fidgety Movements

7 GA: Gestational Age

8 GMA: General Movement Assessment

9 GMs: General Movements

10 MA: Mildly Abnormal GMA rating

11 NS: Normal Suboptimal GMA rating

12 NO: Normal Optimal GMA rating

13 SMIL: Skinned Multi-Infant Linear Model

14 SMPL: Skinned Multi-Person Linear Model

15 RGB-D: Red Green Blue-Depth

16

17 This research did not receive any specific grant from funding agencies in the public,
18 commercial, or not-for-profit sectors.

19

20

1 Cerebral palsy (CP) describes the most common physical disability in children. It occurs in
2 approximately 2/1.000 live births [1]. The average age at which the diagnosis is made and
3 communicated with the families has been shown to be approximately 11 months. For less
4 severely affected children diagnosis may even be made as late as 24 months of age [2]. Yet,
5 early and accurate detection of infants at very high risk of CP is a prerequisite for timely
6 initiation of early intervention, i.e. intervention starting before the age of 6 months [3]. Early
7 intervention programs use caregiver coaching and address the infant's sensorimotor
8 development, attention, self-regulation, early communication skills and parental mental well-
9 being [4, 5].

10 To detect very high risk of CP, it is recommended to use a combination of standardized tools
11 in conjunction with clinical history. Before 5 months corrected age (CA), the most predictive
12 tools to detect risk of CP are term-age cerebral magnetic resonance imaging (MRI) and the
13 video-based Qualitative Assessment of General Movements (GMA) between 2 and 5 months
14 CA [6-8]). Term-age MRI may, however, not serve as a screening tool in clinical practice,
15 leaving GMA as the most predictive tool in everyday practice.

16 General movements (GMs) are spontaneous movements, which involve all body parts [9, 10].
17 They emerge during early fetal life and disappear when goal-directed motor behaviour
18 emerges around 4-5 months CA. The form of typical GMs changes as a result of developmental
19 transformations of the nervous system. In the last phase, at 2 to 5 months CA, GMs have a
20 'fidgety' character. Fidgety GMs (FMs) occur irregularly all over the body and consist of a
21 continuous stream of tiny elegant movements. During each phase, typical GMs are primarily
22 characterized by complexity and variation (in short: complexity) [10]. GMs are considered as
23 abnormal when their complexity is reduced. [9, 10]. At fidgety age – the age at which GMA
24 best predicts developmental outcome [6, 11] – atypical GMs are also characterized by a
25 reduction or disappearance of the age-specific FMs. Therefore Hadders-Algra suggested that
26 GMA at 2-5 months CA (fidgety age) is performed with a two-step procedure: (1) grading
27 movement complexity and (2) assessing FMs [11, 12].

28 Unfortunately, GMA requires a high level of expertise in the raters, which impedes the
29 application in broad clinical practice [13-17]. Therefore, it is desirable to develop an easy to
30 use, automated screening tool that can be applied in general paediatric practice as an
31 alternative for the expert dependent video ratings of GMs. To be generally applicable, this
32 diagnostic tool should be sensitive and specific and should use low-cost, commercially
33 available equipment that does not influence the infants' motions.

34 A primary prerequisite for automatic analysis of motor development is to capture infant
35 movements with high precision. Motions are represented as continuous time-series of the
36 kinematics of limb and trunk movements, e.g., 2-dimensional (2D) or 3-dimensional (3D) body

1 joint positions or joint angles. In the subsequent analysis stage, motion features have to be
2 extracted from the captured movements to serve as input to machine learning classifiers, which
3 can be used to identify infants with abnormal motor activity or to predict the risk of CP.

4 The first computerized approach towards automated detection of CP used a marker-based
5 Vicon system to track movements in 3D [18]. Relevant characteristics of the captured
6 movements were encoded in motion features based on limb velocities and accelerations. A
7 machine learning classifier was trained to predict the CP risk based on these features [18]. Up
8 to date, a wide variety of sensors with different properties have been used for capturing infant
9 movements, e.g., body-worn miniaturized movement sensors [19, 20] or magnetic tracking
10 systems [21]. Video-based motion tracking has the advantage that it does not require the
11 attachment of sensors [22, 23].

12 Marschik et al. [24] developed the 'iDN Fingerprint Model' that combines multiple biological
13 signals of the young infant, including a 3D video of GMs with a Red-Green-Blue Depth (RGB-
14 D) camera. The Italian group of Orlandi and colleagues proposed a system based on the
15 combination of audio and video data, i.e., the infant's spontaneous vocalizations and its GMs
16 [25]. However, currently both systems' psychometric properties have not been reported.

17 Purely computer-based 2D video analysis for the assessment of GMs was first published by
18 Adde and colleagues [26]. They developed a General Movement Toolbox (GMT), which uses
19 the displacement of pixels from one video frame to the next for quantitative analyses of GMs
20 occurring at 2-5 months CA, i.e., the fidgety GMs. Visual representations from GMT resulted
21 in recognisable patterns of the fidgety movements. They reported that the application of the
22 automatic movement classification in their mixed study group of 82 high and low risk infants
23 born at term or preterm would reduce the need for further referral by 70%. Their study showed
24 that video recordings can be used for qualitative and quantitative analyses of GMs [26]. More
25 recently, the same group suggested on the basis of a group of 150 high risk infants that the
26 analysis could be used to reduce the rate of infants needing intensive developmental
27 monitoring [27]. However, two things should be noted. First, the system misclassified some
28 children, e.g., missing those who needed monitoring (10%) or referring those who did not
29 require intensive follow-up (20%). Second, the authors had to exclude infants with a specific
30 type of moderately abnormal GMs (abnormal FMs) from the analysis as the algorithm could
31 not deal with these movements.

32 More recently, two other groups reported on computer-based analysis of 2D GM-videos. The
33 group of Orlandi used an approach based on a skin model for segmentation and large
34 displacement optical flow (LDOF) [28]. The kinematic features extracted from the model were
35 used to classify fidgety GMs as typical or atypical. Their retrospective study based on 127
36 videos of preterm infants showed a specificity of 99%, but a sensitivity of only 44% in predicting

1 CP [28]. The group of Marchi and colleagues reported an approach based on automated pose
2 estimation to capture key aspects of GMs [29]. They used OpenPose - an open source
3 computer vision software - developed by Cao et al. [30] to analyse the 2D videos of 21
4 preselected infants at fidgety GM-age. They considered their approach as feasible for
5 automatizing infant motion analysis. However, they indicated that pose tracking errors
6 occurred in 14 out of 21 recorded videos, which hindered automatized analysis. They
7 concluded, that their approach would improve significantly by the use of three dimensional
8 (3D) instead of 2D camera technology [29].

9 To summarize, up to now no results are available on the application of 3D systems for
10 automated assessment of movements in infancy that do not use sensors/markers that have to
11 be attached to the infants' limbs, which may affect their behaviour.

12 Since 2015, we have developed a motion tracking system that estimates the full body pose of
13 infants in 3D [31-34] (Kinematic Motion Analysis Tool; KineMAT). The system is based on a
14 commercially available, low-cost RGB-depth sensor (Kinect 1.0, Microsoft, USA) and can be
15 applied without attaching markers to the infant's body (see Figure 1) [31-33]. Recently, we
16 developed the Skinned Multi-Infant Linear (SMIL) body model [34], which is based on the adult
17 body model SMPL (Skinned Multi-Person Linear Model) [35]. The pose and shape of this virtual
18 body can be adjusted with the model parameters. This allows us to use SMIL for capturing the
19 3D shape and 3D pose of infants from RGB-D sequences. We create a video of the resulting
20 SMIL body reproducing the movements of the recorded infant, which we denote "SMIL motion
21 video". Figure 2 illustrates the resemblance of the RGB and SMIL motion videos.

22 The aim of this study was to evaluate how reliably our developed 3D infant body model (SMIL
23 motion video) represents true infant spontaneous movements to a human observer. To this
24 end, we assessed the agreement on the classifications of a human expert of GM-quality of two
25 videos of the same GMs of 29 consecutively recruited high risk infants at fidgety age. The data
26 of both videos originated from the same camera recording. One video consisted of the
27 conventional GM-video used in clinical practice, the other of the SMIL motion video (Fig. 1 and
28 Fig. 2). Both videos were assessed according to Hadders-Algra, i.e., assessing GM-complexity
29 (in the present study with three degrees of fine-grading) and assessing FMs with two degrees
30 of grading [12, 36, 37].

31

32 **Material and Methods**

33 **Participants**

1 Participants were high risk infants who received standardized follow up at the Centre of
2 Developmental Care of the integrated Social Paediatric Centre at the University Hospital in
3 Munich, Germany. High risk infants were infants meeting at least one of the following criteria:
4 born \leq 32 weeks of gestation (GA), birth weight below 1500 grams, or other perinatal
5 complications such as term asphyxia or early onset meningitis. Exclusion criteria were
6 additional, potentially confounding diagnoses (e.g. genetic disorders, brachial plexus palsy,
7 neuromuscular disorders). All infants meeting the inclusion criteria were examined by the first
8 author between November 2016 and October 2017. The video recording was performed prior
9 to the general developmental neurological examination. The infants' prenatal, perinatal and
10 neonatal history was documented on standardized charts. After fidgety GM-age, the infants
11 generally were assessed at a three to six months interval depending on the clinical need. Each
12 follow-up visit consisted of a detailed neuropaediatric clinical examination. In children \geq 12
13 months CA the diagnosis of CP was based on clinical findings in line with international
14 recommendations as published by Boychuk et al. [38]. At 24 months CA, children were invited
15 to participate in a structured neurodevelopmental assessment using the Bayley Scales of
16 Infant and Toddler Development (BSITD, Bayley III Scales, full version).

17 The local ethics committee gave clearance for this study prior to the inclusion of the first patient
18 (Ethikkommission LMU, Project Nr. 454-16). Parents gave written informed consent prior to
19 study entry. No family withdrew consent during the study period.

20

21 **Methods**

22 The infant lay in supine position and an RGB-D camera (Microsoft Kinect V1, USA) was
23 positioned directly above the infant at a distance of 100 cm. This camera allows simultaneous
24 recordings of regular RGB videos (typically used for clinical GMA ratings) and depth videos,
25 which provide the 3D distance to the camera for each pixel of the image. The infant's GMs
26 were recorded for 3 minutes in a peaceful, quiet environment during active wakefulness.

27 To capture the infant's motion, we use SMIL (Skinned Multi-Infant Linear body model; see
28 introduction for additional details on the development of the model [34]), which models the 3D
29 body surface together with an underlying skeleton. The pose and shape of this virtual body
30 can be adjusted with the model parameters. This allows the "registration" of the model to RGB-
31 D data. Registration denotes the process of (automatically) adjusting model parameters so that
32 the model best describes the input data, i.e. the model takes the same shape and pose as the
33 recorded infant. We create a video of the SMIL registration results, i.e., the SMIL body
34 reproducing the recorded infant movements, which we denote as "SMIL motion video" (see
35 Fig. 1 and Fig. 2). SMIL allows access to 3D positions of all points of the body and the skeleton,

1 as well as body joint angles in all degrees of freedom across time. The SMIL Model achieves
2 a metric accuracy, i.e., the average distance of RGB-D scan points to the model surface, of
3 2.51 mm (SD 0.21 mm) and represents the original (correct) infant body pose in 98.8% of the
4 time [34]. The model is adapted to the specifics of infantile body proportions. Our registration
5 method can handle fast movements and self-occlusions, which are a common problem when
6 capturing motions of freely moving humans with only one camera. Further details on the
7 registration of the SMIL Model to RGB-D sequences can be found in previous publication [34].

8 The conventional video and the SMIL motion video of each infant were used for blinded GMA
9 rating according to Hadders-Algra. For GMA at fidgety age this means a two-step procedure.
10 First, the degree of GM-complexity is assessed – this is the parameter that can be assessed
11 at any GM-age [10]. It supplies information on the integrity of the subcortical-cortical networks
12 [11]. We graded GM-complexity in three ways: (a) by using a classification into four categories:
13 normal-optimal movements (NO; abundant complexity and variation), normal-suboptimal
14 movements (NS; sufficient complexity and variation), mildly abnormal movements (MA;
15 insufficient variation and complexity; reflecting normal but non-optimal brain function) and
16 definitely abnormal movements (DA; very limited complexity and variation including cramped
17 synchronized movements; reflecting brain dysfunction) [10] (b) by applying a fine-graded 10-
18 point Likert-scale, in which score 1 denotes a virtual absence of GM-complexity and 10 a very
19 abundant GM-complexity [37]. Scores of NO movements ranged from 10 to 8, those of NS
20 movements from 7 to 6, those of MA movements from 5 to 4 and those of DA movements from
21 3 to 1. The Likert-scale fine grading has been applied before; it helps to describe movement
22 quality but it does not assist improvement of prediction of developmental outcome compared
23 to that based on the four classes [37]. (c) GMA ratings were dichotomized into the clinically
24 relevant DA category (1-3 on the Likert scale; high predictive value for CP) versus a non-DA
25 category (MA, NS, NO, i.e. 4-10 on the Likert scale) as previously described by Hadders-Algra
26 et al. [39]. Second, the degree of FMs is assessed: continuously present, intermittently present,
27 sporadic and absent [40]. Continuously and intermittently present FMs were considered as
28 typical FMs. In the calculation of predictive values FMs were categorized as absent or present,
29 i.e., sporadic or typical. [12].

30 The senior author assessed all SMIL motion videos in random order (both in terms of GM-
31 complexity and in terms of presence of FMs), followed by the assessment of all RGB videos in
32 random order. The senior author is a GMA expert with more than 25 years of experience in
33 clinical and scientific GMA rating (MH-A). She was masked for the infant's medical history and
34 outcome.

35

36 **Statistical analysis**

1 Descriptive statistics were used to describe the infants' clinical information on single subject
2 level (gestational age at birth, APGAR, umbilical cord pH and base excess, age at GMA
3 recording, neonatal morbidity diagnoses, brain ultrasound findings, results of
4 neurodevelopmental outcome at ≥ 12 months, GMA scores).

5 For the assessment of both videos by the same GMA rater, statistical tests representing test-
6 retest and intra-rater reliability measures were used. The Intraclass Correlation Coefficient
7 (ICC; two way mixed, absolute agreement) was used to evaluate the degree of agreement
8 between the two GM-ratings of GM-complexity using the 1-10 Likert scale and the FM-ratings.
9 Reliability is considered high when ICC is 0.70 or higher. The ICC values are presented with
10 95% confidence intervals (CI). In addition, we used Bland-Altman plots to report the agreement
11 between the GM-complexity on the basis of both videos.

12 Kappa statistics were used to describe the agreement of DA and Not-DA as well as for the
13 classification of absent and present FMs of both types of video. Kappa-values of 0.41-0.60 are
14 considered a "moderate agreement", 0.61-0.80 a "substantial agreement" and values of 0.81-
15 1.00 as "almost perfect agreement". Two-by-two tables were used to describe the consistency
16 of GMA for the conventional videos compared to those based on the SMIL motion video with
17 respect to neurodevelopmental long-term outcome (CP present or absent). In addition,
18 specificity, sensitivity, negative predictive value, and positive predictive value were calculated.
19 Statistics were performed with IBM SPSS, version 25.0.0.1 (IBM Corporation USA). ROC
20 curves were used to graphically illustrate the relationship between sensitivity and the
21 occurrence of false positive values for DA versus non-DA GM-complexity ratings of both types
22 of video.

23

24 **Results**

25 Twenty-nine infants met the inclusion criteria. Thus, 2x29 sets of videos were assessed. The
26 infants' mean corrected age at recording was 14.8 weeks (95%CI 14.1;15.5). More than half
27 (18/29, 62%) of the infants was born before 32 weeks of gestation. All infants born at higher
28 gestational ages were highrisk due to perinatal or postnatal complications (e.g. term asphyxia
29 or meningitis) requiring standardised neurodevelopmental follow up examinations. At
30 neurodevelopmental follow-up between 12-31 months CA 10 children (34%) showed a typical
31 development; 14 had a global developmental delay (48%) and five (17%) had an ascertained
32 diagnosis of CP (four a bilateral spastic CP, one a unilateral spastic CP). Detailed clinical
33 characteristics of the infants are provided in **Table 1 and Supplementary Table S1**.

34

35 *GMA rating: conventional video versus SMIL motion video*

1 On the basis of GM-complexity assessments six infants were rated as DA, both in the
2 conventional videos and in the SMIL motion video rating, 18 (13 SMIL motion video) as MA,
3 and 6 (10) as NS. In both video ratings no infant was rated as NO (**Table 1**).

4 Using the dichotomy DA versus non-DA GM-complexity, 25/29 (86%) of both ratings were
5 identical. The corresponding inter-rating agreement measured with the Cohen's Kappa was
6 0.58 ($p = 0.002$), meaning "moderate" agreement. For FMs present versus absent Cohen's
7 Kappa was 0.78 ($p = 0.001$), meaning "substantial" agreement.

8 The GM-complexity ratings of both video methods using the 1-10 Likert scale were identical
9 for 16/29 (55%) infants. An additional 10/29 (34%) ratings showed only a +/-1 difference
10 between both video-methods. The ratings of three infants (10%) differed more than one point.
11 In two infants the ratings differed by two points (#6: conventional GMA rating of 5, corresponding
12 to MA GM-complexity, and a SMIL motion video GMA rating of 3, corresponding to DA GM-
13 complexity; #15 conventional GMA rating of 3 and a SMIL motion video GMA rating of 5). In
14 the third infant the ratings differed by three points (#3, conventional GM-complexity 5 versus
15 SMIL motion video GM-complexity 2) (**Table 2, Bland-Altman-Plot see Supplementary**
16 **Figure S1**). The mean Intraclass Correlation Coefficient (ICC) of both GM-complexity ratings
17 was excellent 0.874 (95% CI 0.730;0.941).

18 In the FM rating of the conventional video FM's were absent in three infants (#1, #2, #6) of
19 which two infants had DA GM-complexity and one MA GM-complexity. All three were later
20 diagnosed with CP. In the SMIL motion video also three infants were rated as absent FM (#1,
21 #6, #8) of which again two infants were rated having DA GM-complexity and one as having
22 MA GM-complexity. Two of these were diagnosed with CP and one with global developmental
23 delay at follow up. ICC of FM ratings was 0.926 (95% CI 0.843;0.965).

24 One infant who was diagnosed with bilateral CP (#5) and the infant with unilateral CP (#20)
25 had presented either with sporadic or typical FMs in both videos (**Table 1**).

26

27 *Prediction of developmental outcome*

28 **Table 2** summarizes the predictive values of the various GMA ratings including positive and
29 negative predictive values. Sensitivity and specificity of DA vs non-DA rating of GM-complexity
30 on the basis of the conventional videos to predict CP were 0.600 and 0.875, respectively; those
31 of the corresponding SMIL motion videos 0.800 and 0.917, respectively. The difference
32 between the two is illustrated by ROC curves (Supplementary Figure S2). Sensitivity and
33 specificity of absent FMs of the conventional video were 0.600 and 1.000, those of the SMIL
34 motion videos 0.400 and 1.000, respectively. Finally, the combination of GM-complexity and
35 FM ratings resulted in a sensitivity and specificity of both types of video of 0.400 and 1.000,

1 respectively. Child #20, later diagnosed with unilateral CP, was consistently missed: she
2 showed MA GMs with FMs.

5 **Discussion**

6 Our clinical case-study indicated that GMA based on computer generated virtual 3D infant
7 body models (SMIL motion videos) closely corresponds to the established gold standard based
8 on conventional RGB videos. The SMIL motion videos were able to catch movement
9 complexity and FMs. In addition, our data suggest that GMA ratings based on SMIL motion
10 videos result in a similar prediction of CP as GMA based on conventional videos.

11 *Comparison of GMA based on conventional versus SMIL motion videos*

12 Several studies have reported excellent inter-observer agreement of skilled observers for
13 conventional GMA [10, 41, 42]. Our comparison of the 29 simultaneously recorded video pairs
14 could be considered a similar type of test-retest analysis. Corresponding to literature data, we
15 found substantial inter-observer reliability (dichotomous FMs and GM-complexity Cohen's
16 Kappa 0.78 and 0.58 respectively, finer graded ratings ICC 0.926 and 0.874 respectively).

17
18 The predictive values of both ratings of GM-complexity were comparable to those described in
19 the literature; those of FMs were somewhat lower than the values reported in the literature,
20 where summary estimates of sensitivity and specificity of GMA based on FM assessment have
21 been reported to be 98% [95% CI: 74;100] and 91% [95% CI: 83;93%] respectively [6].
22 Interestingly, the predictive values based on the SMIL motion video were not inferior to those
23 based on the conventional videos. While GM-complexity rating seemed to be superior using
24 the SMIL motion video (Sensitivity 0.8 versus 0.6; PPV 0.667 versus 0.5, NPV 0.957 versus
25 0.913), the conventional video seemed to be slightly superior regarding the assessment of
26 FMs (Sensitivity 0.6 versus 0.4; NPV 0.923 versus 0.889). It is conceivable that the SMIL
27 motion video is superior to conventional GMA rating of GM-complexity, as the SMIL motion
28 video forces the observer to focus on the Gestalt perception of GM-quality, whereas in the
29 conventional video the assessor may be distracted and/or biased by additional visible details,
30 such as naso-gastric tube, oxygen supply, monitoring cables or details of facial expression.
31 Regarding the assessment of FMs, however, the SMIL motion video at times contains slight
32 pose inaccuracies due to noisy input data, which may be misinterpreted as very small
33 amplitude FMs. This could explain that it was occasionally hard to distinguish sporadic and
34 absent FMs on the basis of the SMIL videos due to insufficient data quality. This underpins the
35 conclusion, that assessing both parameters (GM-complexity and FMs) is relevant for an
36 optimal clinical prediction.

1 To understand the differences in the GMA ratings based on the conventional and SMIL motion
2 videos of infants #2, #3, #6, and #15 MH-A reviewed these videos. She noticed the following
3 for the infants (#3 and #6) whose GM-complexity was rated better on the conventional video
4 (MA) than on the SMIL video (DA): Infant #3 (developmental delay at follow-up) showed
5 especially some complex movements in the distal joints (with FMs), apparently these
6 movements made a less prominent impression on the SMIL video. Infant # 6 who was later
7 diagnosed with bilateral CP had shown on the conventional video a discrepancy in movement
8 complexity between the upper and the lower part of the body, with the upper being better than
9 the lower. This discrepancy was less striking on the SMIL video. Both videos were rated as
10 absent FMs. Infant #15 (typical outcome at follow-up) had been assessed at 9 weeks CA. GM-
11 complexity of the conventional video was rated as DA, due to prominent stereotyped
12 movement sequences in the proximal joints; the distal joints showed some complexity. The
13 latter was more prominently visible during the SMIL video, which was therefore rated as MA.
14 Both videos were rated as sporadic FMs. It has been shown by others, that movement quality
15 at 9 weeks CA may improve to better qualities during the following weeks [43]. Presumably
16 this happened also in this infant. Finally, the GMs of infant #2 (bilateral CP) were rated on both
17 videos as DA. However, the FMs were scored as absent on the conventional video, but as
18 sporadic on the SMIL video – the few FMs had been observed at the wrists only. The latter
19 may also be due to the slight pose inaccuracies due to noisy input data of the SMIL Model
20 suggesting the presence of sporadic FM. The differences in scores between conventional and
21 SMIL videos could be due to such technical matters, but they also underline that GMA based
22 on Gestalt perception may remain difficult even after years of experience; it stresses the need
23 of automatic movement analysis following strict algorithms.

24 GMA was especially adequate in the prediction of bilateral CP. Yet, it missed the child who
25 was diagnosed with unilateral CP. It has been reported before that GMA does not detect all
26 children with unilateral CP [44-46]. The differential ability of GMA to predict bilateral and
27 unilateral CP corresponds to the putative neural substrate of the quality of GMs. GM-quality is
28 considered to reflect the integrity of extensive neural networks involving not only cortical areas,
29 but also their connectivity with subcortical relay stations [7]. Indeed, evidence is accumulating
30 that DA GMs are especially related to damage of the periventricular white matter [7, 47]. In
31 part of the children with unilateral CP the lesion is restricted to the cortical grey or does only
32 affect part of the cross-roads running through the periventricular white matter (as in infant #20)
33 and does not massively disrupt white matter connectivity, therewith not producing DA GMs.
34 The detection of children with unilateral CP is also challenging due to the finding that fidgety
35 movements are frequently present.

1 Two children showed DA GMs (#4 and #7) but were not diagnosed with CP. They both
2 presented with “global developmental delay” at follow-up. Recent data indicate that the
3 limitations in the infant’s early motor repertoire do not only predict an increased risk of CP, but
4 also an increased risk of cognitive impairment irrespective of the GMA method used for
5 classification [48-50].

6 *Strengths and limitations:*

7 The strength of the current study is that it is the first that compares GMA based on videos of
8 computed 3D infant body models to GMA rating of conventional videos. Another strength is
9 that the videos were rated by one masked and highly experienced GMA assessor using
10 multiple grading strategies including explicit assessment of FMs.

11 The study also has limitations: the sample size was small and consisted of high risk infants
12 only. However, this convenience sample might be considered as representative of many infant
13 follow-up clinics and represented infants with normal, MA and DA GMs. It may also be
14 considered that the duration of follow-up was limited, which may mean that some children with
15 a developmental diagnosis may have been missed [37].

16 *Current status and future perspective of GMA based on SMIL modelling*

17 In contrast to previous approaches that capture a limited number of body joint positions based
18 on 2D-video [27-29, 51, 52] or relying on marker based sensors attached to the baby [18, 19,
19 53, 54], our method allows the extraction of angles of all body joints (3 degrees of freedom per
20 joint), 3D positions of body joints and of arbitrary points on the body surface without the need
21 of potentially distracting markers. Our method is robust to self-occlusions (e.g. hands and feet
22 occluding large parts of the body) and fast motion. It has been proven to display the true whole
23 body movements of the recorded infant in 98.8% of recorded video sequences, and to have a
24 metric accuracy, i.e. the average distance of RGB-D scan points to the model surface, of 2.51
25 mm (SD 0.21 mm) [34].

26 In order to further improve the SMIL tracking, we plan to integrate means to detect and fix the
27 rare occurrence of failure cases, as recently published by Aristidou et al. [55]. In addition, we
28 aim to apply the SMIL motion video in a larger patient population, not only to improve the
29 detection of predetermined motion parameters (e.g. fidgety general movements, centroid of
30 motion, cramped synchronized activity, or a reduction of segmental movements), but also to
31 use machine learning algorithms to quantify the motion parameters which are assumed to drive
32 the “Gestalt perception”.

33

34 *Concluding remarks*

1 Our case study of 29 high risk infants demonstrated that the amount of motion details captured
2 by the SMIL motion video - based on a low-cost Kinect recording and the KineMAT tool -
3 enables accurate GMA at fidgety age. This implies that the SMIL motion video adequately
4 catches the movement characteristics needed for GMA of infants with movements ranging
5 from a normal to a definitely abnormal quality, turning it into an attractive tool for automatic
6 GMA.
7

Preprint Manuscript Accepted

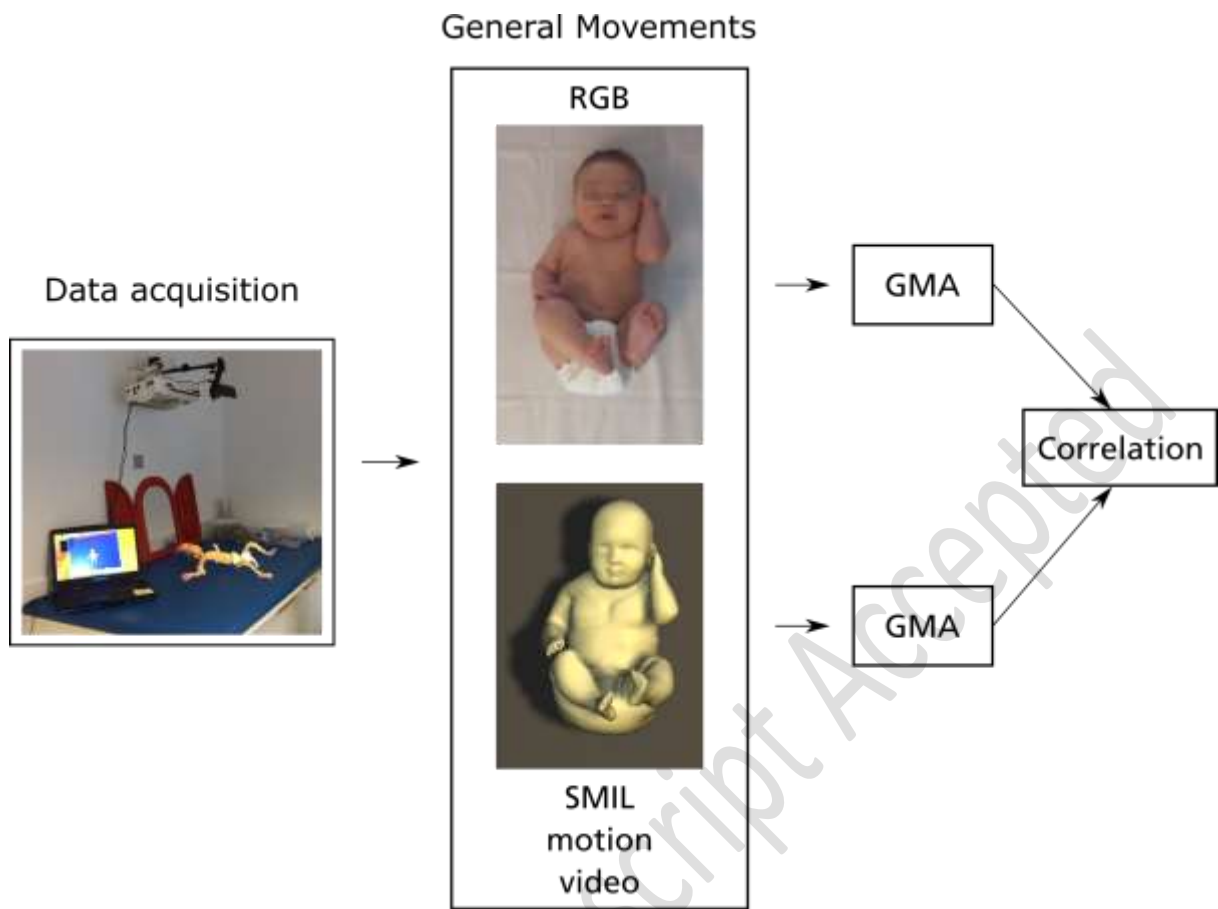
1 References

- 2 1. Himmelmann K, Uvebrant P. The panorama of cerebral palsy in Sweden part XII shows that
3 patterns changed in the birth years 2007-2010. *Acta Paediatr.* 2018;107(3):462-8.
- 4 2. Granild-Jensen JB, Rackauskaite G, Flachs EM, Uldall P. Predictors for early diagnosis of cerebral
5 palsy from national registry data. *Dev Med Child Neurol.* 2015;57(10):931-5.
- 6 3. Novak I, Morgan C, Adde L, Blackman J, Boyd RN, Brunstrom-Hernandez J, et al. Early, Accurate
7 Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment. *JAMA*
8 *Pediatr.* 2017;171(9):897-907.
- 9 4. Akhbari Ziegler S, Dirks T, Hadders-Algra M. Coaching in early physical therapy intervention: the
10 COPCA program as an example of translation of theory into practice. *Disabil Rehabil.*
11 2019;41(15):1846-54.
- 12 5. Hutchon B, Gibbs D, Harniess P, Jary S, Crossley SL, Moffat JV, et al. Early intervention
13 programmes for infants at high risk of atypical neurodevelopmental outcome. *Dev Med Child*
14 *Neurol.* 2019;61(12):1362-7.
- 15 6. Bosanquet M, Copeland L, Ware R, Boyd R. A systematic review of tests to predict cerebral palsy
16 in young children. *Dev Med Child Neurol.* 2013;55(5):418-26.
- 17 7. Hadders-Algra M. Early human brain development: Starring the subplate. *Neurosci Biobehav Rev.*
18 2018;92:276-90.
- 19 8. Kwong AKL, Fitzgerald TL, Doyle LW, Cheong JLY, Spittle AJ. Predictive validity of spontaneous
20 early infant movement for later cerebral palsy: a systematic review. *Dev Med Child Neurol.*
21 2018;60(5):480-9.
- 22 9. Prechtl HF. Qualitative changes of spontaneous movements in fetus and preterm infant are a
23 marker of neurological dysfunction. *Early Hum Dev.* 1990;23(3):151-8.
- 24 10. Hadders-Algra M, Mavinkurve-Groothuis AM, Groen SE, Stremmelaar EF, Martijn A, Butcher PR.
25 Quality of general movements and the development of minor neurological dysfunction at toddler
26 and school age. *Clin Rehabil.* 2004;18(3):287-99.
- 27 11. Hadders-Algra M. Neural substrate and clinical significance of general movements: an update. *Dev*
28 *Med Child Neurol.* 2018;60(1):39-46.
- 29 12. Hamer EG, Bos AF, Hadders-Algra M. Assessment of specific characteristics of abnormal general
30 movements: does it enhance the prediction of cerebral palsy? *Dev Med Child Neurol.*
31 2011;53(8):751-6.
- 32 13. Adde L, Rygg M, Lossius K, Oberg GK, Stoen R. General movement assessment: predicting cerebral
33 palsy in clinical practise. *Early Hum Dev.* 2007;83(1):13-8.
- 34 14. Ricci E, Einspieler C, Craig AK. Feasibility of Using the General Movements Assessment of Infants in
35 the United States. *Phys Occup Ther Pediatr.* 2018;38(3):269-79.
- 36 15. Brown AK, Greisen G, Haugsted U, Jonsbo F. Formal training in general movement assessment is
37 required to effectively evaluate infants with perinatal asphyxia in outpatient settings. *Acta*
38 *Paediatr.* 2016;105(9):1056-60.
- 39 16. Bernhardt I, Marbacher M, Hilfiker R, Radlinger L. Inter- and intra-observer agreement of Prechtl's
40 method on the qualitative assessment of general movements in preterm, term and young infants.
41 *Early Hum Dev.* 2011;87(9):633-9.
- 42 17. Datta AN, Furrer MA, Bernhardt I, Huppi PS, Borradori-Tolsa C, Bucher HU, et al. Fidgety
43 movements in infants born very preterm: predictive value for cerebral palsy in a clinical
44 multicentre setting. *Dev Med Child Neurol.* 2017;59(6):618-24.
- 45 18. Meinecke L, Breitbach-Faller N, Bartz C, Damen R, Rau G, Disselhorst-Klug C. Movement analysis in
46 the early detection of newborns at risk for developing spasticity due to infantile cerebral palsy.
47 *Hum Mov Sci.* 2006;25(2):125-44.
- 48 19. Heinze F, Hesels K, Breitbach-Faller N, Schmitz-Rode T, Disselhorst-Klug C. Movement analysis by
49 accelerometry of newborns and infants for the early detection of movement disorders due to
50 infantile cerebral palsy. *Med Biol Eng Comput.* 2010;48(8):765-72.

- 1 20. Gima H, Shimatani K, Nakano H, Watanabe H, Taga G. Evaluation of Fidgety Movements of Infants
2 Based on Gestalt Perception Reflects Differences in Limb Movement Trajectory Curvature. *Phys*
3 *Ther.* 2019;99(6):701-10.
- 4 21. Philippi H, Karch D, Kang KS, Wochner K, Pietz J, Dickhaus H, et al. Computer-based analysis of
5 general movements reveals stereotypies predicting cerebral palsy. *Dev Med Child Neurol.*
6 2014;56(10):960-7.
- 7 22. Rahmati H, Martens H, Aamo OM, Stavadahl O, Stoen R, Adde L. Frequency-based features for
8 early cerebral palsy prediction. *Conf Proc IEEE Eng Med Biol Soc.* 2015;2015:5187-90.
- 9 23. Stahl A, Schellewald C, Stavadahl O, Aamo OM, Adde L, Kirkerod H. An optical flow-based method
10 to predict infantile cerebral palsy. *IEEE Trans Neural Syst Rehabil Eng.* 2012;20(4):605-14.
- 11 24. Marschik PB, Pokorny FB, Peharz R, Zhang D, O'Muircheartaigh J, Roeyers H, et al. A Novel Way to
12 Measure and Predict Development: A Heuristic Approach to Facilitate the Early Detection of
13 Neurodevelopmental Disorders. *Curr Neurol Neurosci Rep.* 2017;17(5):43.
- 14 25. Orlandi S, Guzzetta A, Bandini A, Belmonti V, Barbagallo SC, Tealdi C, et al. AVIM—A contactless
15 system for infant data acquisition and analysis: Software architecture and first results. *Biomedical*
16 *Signal Processing and Control.* 2015;20:85-99.
- 17 26. Adde L, Helbostad JL, Jensenius AR, Taraldsen G, Stoen R. Using computer-based video analysis in
18 the study of fidgety movements. *Early Hum Dev.* 2009;85(9):541-7.
- 19 27. Stoen R, Songstad NT, Silberg IE, Fjortoft T, Jensenius AR, Adde L. Computer-based video analysis
20 identifies infants with absence of fidgety movements. *Pediatr Res.* 2017;82(4):665-70.
- 21 28. Orlandi S, Raghuram K, Smith CR, Mansueto D, Church P, Shah V, et al. Detection of Atypical and
22 Typical Infant Movements using Computer-based Video Analysis. *Conf Proc IEEE Eng Med Biol Soc.*
23 2018; 3598-3601.
- 24 29. Marchi V, Hakala A, Knight A, D'Acunto F, Scattoni ML, Guzzetta A, et al. Automated pose
25 estimation captures key aspects of General Movements at eight to 17 weeks from conventional
26 videos. *Acta Paediatr.* 2019;108(10): 1817-1824.
- 27 30. Cao Z, Simon T, Wei SE, Sheikh Y. Realtime multi-person 2D pose estimation using part affinity
28 fields. *CVPR, IEEE.* 2017:1302-10.
- 29 31. Hesse N, Stachowiak G, Breuer T, Arens M. Estimating Body Pose of Infants in Depth Images using
30 Random Ferns. *IEEE International Conference on Computer Vision Workshops (ICCVW).* 2015:35–
31 43.
- 32 32. Hesse N, Schroeder A, Mueller-Felber W, Bodensteiner C, Arens M, Hoffmann U. Markerless
33 motion analysis for early detection of infantile movement disorders. *EMBEC & NBC 2017, Joint*
34 *Conference of the European Medical and Biological Engineering Conference (EMBEC) and the*
35 *Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC): Tampere, Finland,*
36 *June 2017. DOI: 10.1007/978-981-10-5122-7_50.*
- 37 33. Hesse N, Schroeder A, Mueller-Felber W, Bodensteiner C, M. A, Hofmann U. Body Pose Estimation
38 in Depth Images for Infant Motion Analysis. *EMBC 2017, 39th Annual International Conference of*
39 *the IEEE Engineering in Medicine and Biology Society : July 11 to 15, 2017, Jeju Island, Korea. DOI:*
40 *10.1109/EMBC.2017.8037221.*
- 41 34. Hesse N, Pujades S, Black M, Arens M, Hofmann U, Schroeder S. Learning and Tracking the 3D
42 Body Shape of Freely Moving Infants from RGB-D sequences. *IEEE Trans Pattern Anal Mach Intell.*
43 2019. DOI: 10.1109/TPAMI.2019.2917908.
- 44 35. Loper M, Mahmood N, Romero J, Pons-Moll G, Black MJ. Smpl: A skinned multi-person linear
45 model. *ACM Transactions on Graphics (TOG).* 2015;34(6):248:1-:16.
- 46 36. Hadders-Algra M. Evaluation of motor function in young infants by means of the assessment of
47 general movements: a review. *Pediatr Phys Ther.* 2001;13(1):27-36.
- 48 37. Groen SE, de Blecourt AC, Postema K, Hadders-Algra M. General movements in early infancy
49 predict neuromotor development at 9 to 12 years of age. *Dev Med Child Neurol.* 2005;47(11):731-
50 8.

- 1 38. Boychuck Z, Andersen J, Bussieres A, Fehlings D, Kirton A, Li P, et al. International expert
2 recommendations of clinical features to prompt referral for diagnostic assessment of cerebral
3 palsy. *Dev Med Child Neurol.* 2020; 62(1): 89-96.
- 4 39. Hadders-Algra M. Putative neural substrate of normal and abnormal general movements.
5 *Neurosci Biobehav Rev.* 2007;31(8):1181-90.
- 6 40. Einspieler C, Prechtl HF, Bos A, Ferrari F, Cioni G. Prechtl's Method on the Qualitative Assessment
7 of General Movements in Preterm, Term and Young Infants. First ed: Mac Keith Press; 2008
8 September 2008.
- 9 41. Hadders-Algra M, Groothuis AM. Quality of general movements in infancy is related to
10 neurological dysfunction, ADHD, and aggressive behaviour. *Dev Med Child Neurol.*
11 1999;41(6):381-91.
- 12 42. Fjortoft T, Einspieler C, Adde L, Strand LI. Inter-observer reliability of the "Assessment of Motor
13 Repertoire--3 to 5 Months" based on video recordings of infants. *Early Hum Dev.* 2009;85(5):297-
14 302.
- 15 43. Saether R, Stoen R, Vik T, Fjortoft T, Vagen RT, Silberg IE, et al. A change in temporal organization
16 of fidgety movements during the fidgety movement period is common among high risk infants.
17 *Eur J Paediatr Neurol.* 2016;20(4):512-7.
- 18 44. Cioni G, Bos AF, Einspieler C, Ferrari F, Martijn A, Paolicelli PB, et al. Early neurological signs in
19 preterm infants with unilateral intraparenchymal echodensity. *Neuropediatrics.* 2000;31(5):240-
20 51.
- 21 45. Guzzetta A, Mercuri E, Rapisardi G, Ferrari F, Roversi MF, Cowan F, et al. General movements
22 detect early signs of hemiplegia in term infants with neonatal cerebral infarction. *Neuropediatrics.*
23 2003;34(2):61-6.
- 24 46. Bouwstra H, Dijk-Stigter GR, Grooten HM, Janssen-Plas FE, Koopmans AJ, Mulder CD, et al.
25 Predictive value of definitely abnormal general movements in the general population. *Dev Med*
26 *Child Neurol.* 2010;52(5):456-61.
- 27 47. Peyton C, Yang E, Msall ME, Adde L, Stoen R, Fjortoft T, et al. White Matter Injury and General
28 Movements in High-Risk Preterm Infants. *AJNR Am J Neuroradiol.* 2017;38(1):162-9.
- 29 48. Butcher PR, van Braeckel K, Bouma A, Einspieler C, Stremmelaar EF, Bos AF. The quality of
30 preterm infants' spontaneous movements: an early indicator of intelligence and behaviour at
31 school age. *J Child Psychol Psychiatry.* 2009;50(8):920-30.
- 32 49. Heineman KR, Schendelaar P, Van den Heuvel ER, Hadders-Algra M. Motor development in
33 infancy is related to cognitive function at 4 years of age. *Dev Med Child Neurol.* 2018;60(11):1149-
34 55.
- 35 50. Einspieler C, Bos AF, Libertus ME, Marschik PB. The General Movement Assessment Helps Us to
36 Identify Preterm Infants at Risk for Cognitive Dysfunction. *Front Psychol.* 2016;7:406.
- 37 51. Rahmati H, Aamo OM, Stavadahl O, Dragon R, Adde L. Video-based early cerebral palsy prediction
38 using motion segmentation. *Conf Proc IEEE Eng Med Biol Soc.* 2014;2014:3779-83.
- 39 52. Rahmati H, Martens H, Aamo OM, Stavadahl O, Stoen R, Adde L. Frequency Analysis and Feature
40 Reduction Method for Prediction of Cerebral Palsy in Young Infants. *IEEE Trans Neural Syst*
41 *Rehabil Eng.* 2016;24(11):1225-34.
- 42 53. Gao Y, Long Y, Guan Y, Basu A, Baggaley J, Plötz T. Towards Reliable, Automated General
43 Movement Assessment for Perinatal Stroke Screening in Infants Using Wearable Accelerometers.
44 *ACM.* 2019. arXiv:1902.08068 [cs.HC]
- 45 54. Karch D, Kang KS, Wochner K, Philipp H, Hadders-Algra M, Pietz J, et al. Kinematic assessment of
46 stereotypy in spontaneous movements in infants. *Gait Posture.* 2012;36(2):307-11.
- 47 55. Aristidou A, Cohen-Or D, Hodgins JK, A S. Self-similarity Analysis for Motion Capture Cleaning.
48 *Computer Graphics-Forum.* 2018;37(2):297-309.
- 49 56. SCPE. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and
50 registers. *Surveillance of Cerebral Palsy in Europe (SCPE).* *Dev Med Child Neurol.* 2000;42(12):816-
51 24.

1 **Figure 1: Setup**



2

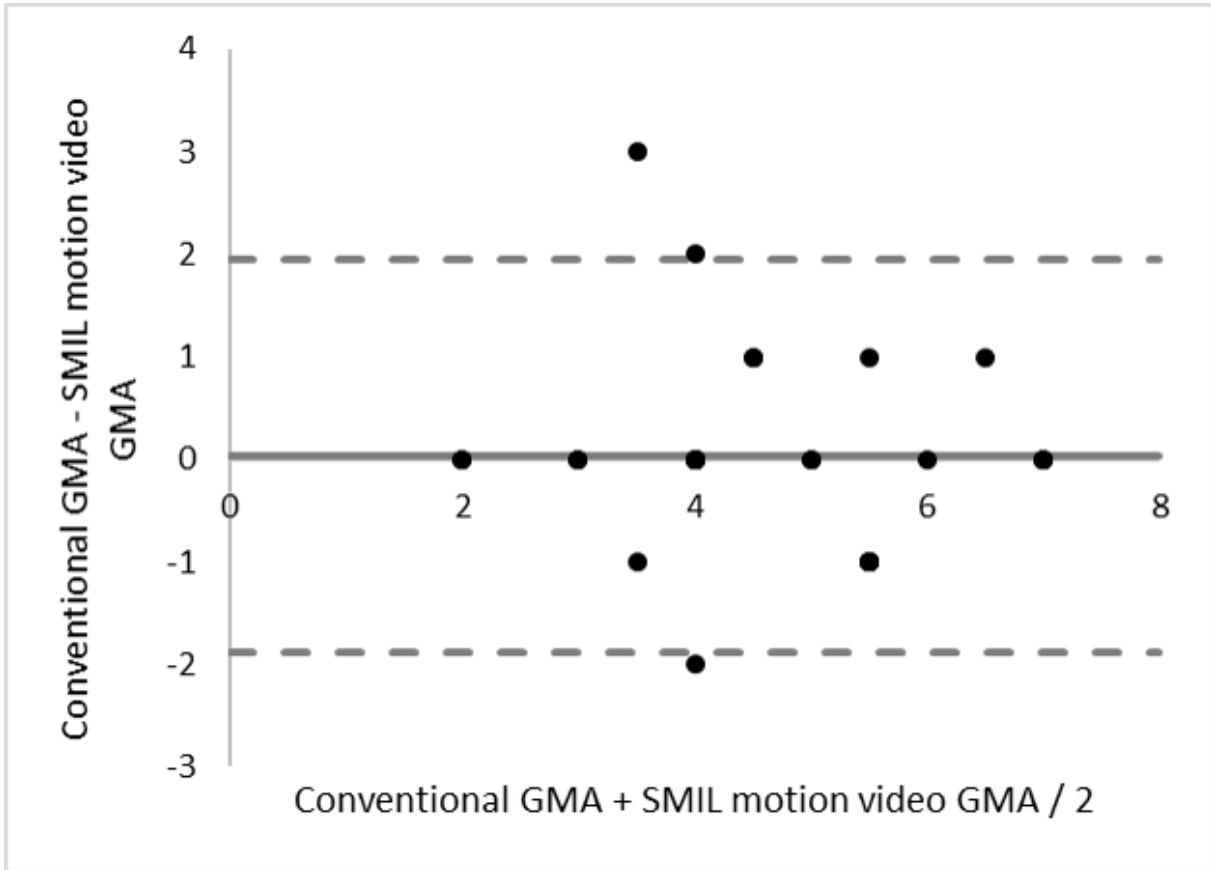
Preprint Manuscript Accepted

1 **Figure 2: Conventional video recording and its corresponding SMIL motion video pose.**



Preprint Manuscript Accepted

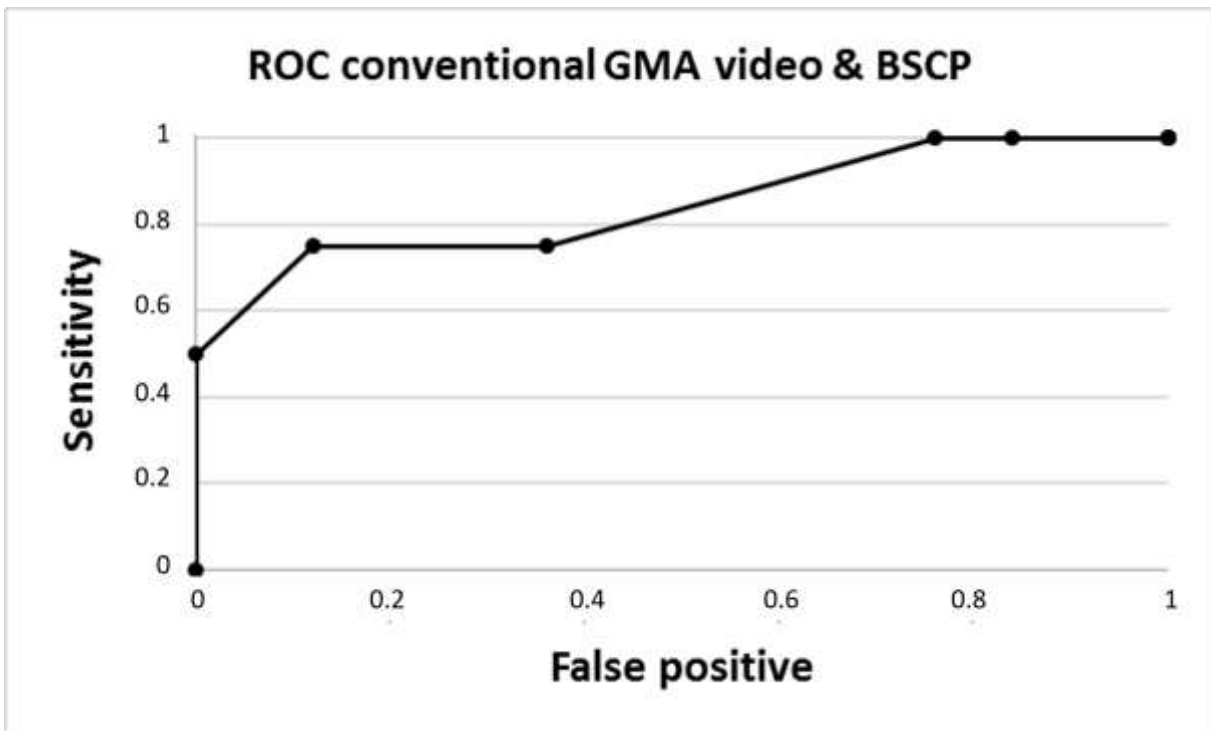
- 1 **Supplementary Figure S1: Bland-Altman Plot:** 1-10 Likert scale rating conventional GMA
- 2 versus SMIL motion video rating.



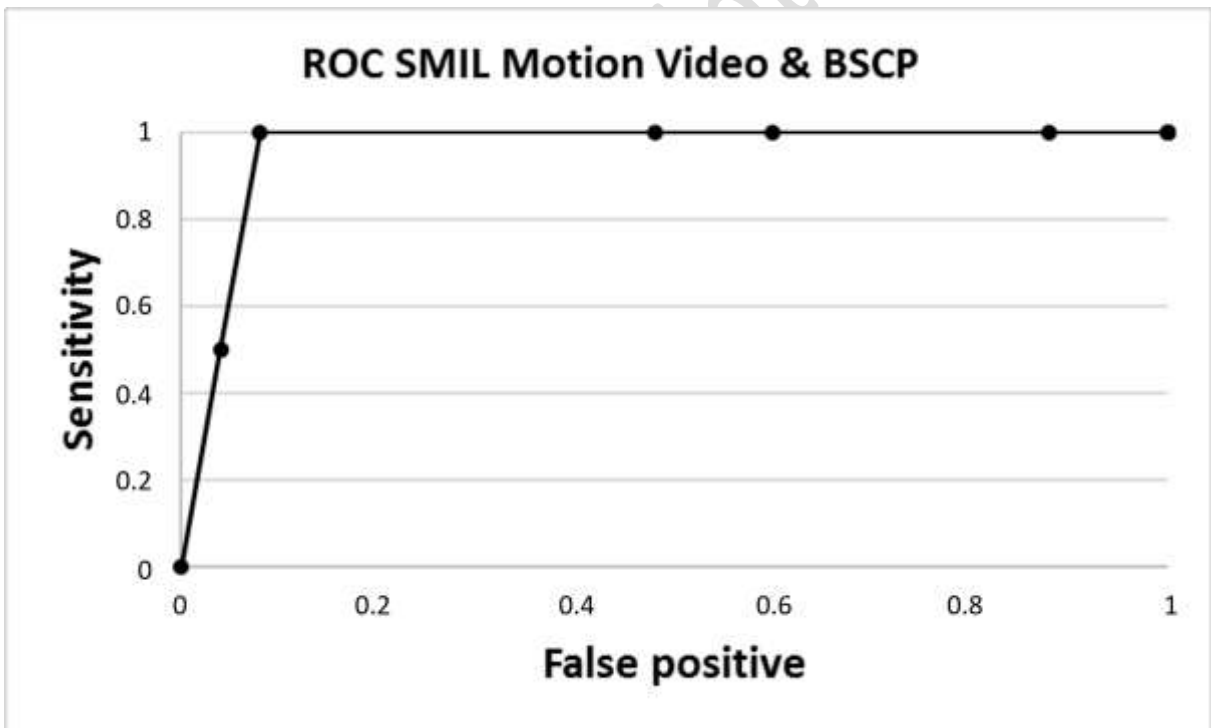
3

Preprint Manuscript

1 **Supplementary Figure S2a,b: ROC**



2



3

1 **Table 1: GMA Conventional video versus SMIL motion video**

Pat ID	sex	GA	CA	Clinical Diagnosis at 12-30 mo CA [38]	GM-complexity Likert score Conventional video	Fidgety movements Conventional video	GM-complexity Likert score SMIL motion video	Fidgety movements SMIL motion video
1	M	29+4	12	Bilateral CP	2	0	2	0
2	M	40+0	17	Bilateral CP	2	0	2	1
3	F	28+4	12	Dev. Delay	5	2	2	1
4	M	29+1	18	Dev. Delay	3	2	3	2
5	M	40+2	9	Bilateral CP	3	1	3	2
6	F	32+2	14	Bilateral CP	5	0	3	0
7	F	24+5	12	Dev. Delay	3	2	4	2
8	M	24+0	13	Dev. Delay	4	2	4	2
9	M	35+2	18	Age adequate	4	2	4	2
10	F	28+0	15	Dev. Delay	4	2	4	2
11	M	40+6	17	Age adequate	4	2	4	2
12	M	25+5	14	Dev. Delay	4	2	4	2
13	F	28+5	16	Age adequate	4	2	4	2
14	F	28+0	15	Dev. Delay	5	2	4	2
15	M	41+5	9	Age adequate	3	1	5	1
16	M	27+1	14	Dev. Delay	5	2	4	2
17	M	31+1	17	Age adequate	5	2	5	2
18	M	29+1	13	Age adequate	5	2	5	2
19	F	33+4	12	Age adequate	5	2	6	2
20	F	28+0	15	Unilateral CP	5	2	6	2
21	F	28+5	16	Age adequate	5	2	6	2
22	M	29+1	13	Age adequate	5	2	6	2
23	M	23+6	15	Dev. Delay	5	2	6	2
24	F	29+4	16	Dev. Delay	6	2	5	2
25	M	29+4	16	Dev. Delay	6	2	6	2
26	F	36+3	12	Dev. Delay	7	2	6	2
27	F	26+6	17	Age adequate	7	2	7	2
28	M	25+4	15	Dev. Delay	7	2	7	2
29	M	41+1	16	Dev. Delay	7	2	7	2

2 **Legend:** GM-complexity Likert scale ratings correspond to the following classifications 1-3 “definitely
3 abnormal”, 4-5 “mildly abnormal”, 6-7 “normal suboptimal”, 8-10 “ normal optimal. Fidgety movements
4 (FMs): 0 = absent; 1 = sporadically present; 2 = intermittently or continuously present. Abbreviations:
5 CA = corrected age; convent = conventional; CP = cerebral palsy; definition of CP in accordance with
6 international expert recommendations.[38]; terminology of the Surveillance of Cerebral Palsy in Europe
7 (SCPE; [56]. DD: developmental delay, either based on performance on the Bayley Scales of Infant and
8 Toddler Development (BSITD, Version 3) or on the clinical examination (see Supplementary Table 1);
9 GA: gestational age; F: female; M: male;

10

Table 2: Prediction of CP at follow up by conventional video GMA rating and GMA rating based on SMIL motion videos.

a) based on GM-complexity

		Cerebral palsy	No CP	Total	Q criteria
Conventional video					Sensitivity 0.6 Specificity 0.875 PPV 0.5 NPV 0.913
DA-compl.		3	3	6	
Not-DA		2*	21	23	
	Total	5	24	29	
SMIL motion video					Sensitivity 0.8 Specificity 0.917 PPV 0.667 NPV 0.957
DA compl.		4	2	6	
Not-DA		1*	22	23	
	Total	5	24	29	

b) based on FMs

		Cerebral palsy	No CP	Total	Q criteria
Conventional video					Sensitivity 0.6 Specificity 1 PPV 1 NPV 0.923
FMs absent		3	0	3	
FMs present		2*	24	26	
	Total	5	24	29	
SMIL motion video					Sensitivity 0.4 Specificity 1 PPV 1 NPV 0.889
FMs absent		2	0	2	
FMs present		3*	24	27	
	Total	5	24	29	

c) based on GM-complexity and FMs

		Cerebral palsy	No CP	Total	Q criteria
Conventional video					Sensitivity 0.4 Specificity 1 PPV 1 NPV 0.889
DA-compl., FMs absent		2	0	2	
Not-DA, FMs absent		1	0	1	
DA-compl., FMs present		1	3	4	
Not-DA, FMs present		1*	21	22	
	Total	5	24	29	
SMIL motion video					Sensitivity 0.4 Specificity 1 PPV 1 NPV 0.889
DA-compl, FMs absent		2	0	2	
Not-DA, FMs absent		0	0	0	
DA-compl, FMs present		2	2	4	
Not-DA, FMs present		1*	22	23	
	Total	5	24	29	

Legend Tabel 2a-c: * including infant #20 later diagnosed with unilateral spastic CP. Abbreviations: DA compl. = definitely abnormal complexity GM rating (Likert Scale 1-3); Not-DA = all other GM ratings (mildly abnormal, normal suboptimal; Likert scale 4-7). FMs absent = fidgety movements scored "0", FMs present = all other fidgety movements (sporadic, intermittend, continuously). [36, 37], SMIL Skinned Multi-Infant Linear Body Model [34].

Supplementary Table S1: Clinical characteristics of the study population

Pat ID	sex	GA	CA	BW	APGAR 1/5/10 minutes	Umbilical cord pH	Umbilical Base-excess	Cerebral ultrasound	Clinical Diagnosis at 12-30 months
1	M	29+4	12	1370	8/9/9	7.41	-2	Abnormal (PVL)	Bilateral CP
2	M	40+0	17	3500	1/2/6	7.14	-12	Abnormal (HIE term asphyxia)	Bilateral CP
3	F	28+4	12	800	8/9/9	7.3	-8	Unspecific	Bayley III: DD at 23 mo CA
4	M	29+1	18	960	2/8/9	7.36	1	Normal	Clinically DD at 18 mo CA
5	M	40+2	9	2500	6/7/7	7.08	-11	Abnormal (HIE term asphyxia)	Bilateral CP
6	F	32+2	14	1810	6/8/-	7.36	0	Abnormal (PVL)	Bilateral CP
7	F	24+5	12	820	7/9/9	7.42	2	Normal	Bayley III: DD at 24 mo CA
8	M	24+0	13	620	7/8/9	7.39	1	Unspecific	Clinically DD at 12 mo CA
9	M	35+2	18	2590	9/9/10	7.4	-2	Abnormal (ICH I right, II° left)	Typical development at 12 mo CA
10	F	28+0	15	480	7/8/10	7.19	-5.5	Normal	Bayley III: DD at 26 mo CA
11	M	40+6	17	4526	-/9/10	7.32	0	Unspecific	Typical development at 12 mo CA
12	M	25+5	14	750	1/2/4	7.31	-7	Unspecific	Clinically DD at 12 mo CA
13	F	28+5	16	1300	9/9/10	7.47	-4.5	Unspecific	Bayley III: typical development at 28 mo CA
14	F	28+0	15	650	8/10/10	7.38	-6	Unspecific	Bayley III: DD at 24 mo CA
15	M	41+5	9	3450	3/6/10	7.23	-8.4	Normal	Bayley III: typical development at 24 mo CA
16	M	27+1	14	1100	4/9/9	7.31	-4.3	Normal	Bayley III: DD at 30 mo CA

17	M	31+1	17	-	8/7/9	7.33	-1	Unspecific	Bayley III: typical development at 26 mo CA
18	M	29+1	13	1190	8/8/9	7.34	-3.2	Normal	Bayley III: typical development at 24 mo CA
19	F	33+4	12	1850	9/10/10	7.47	-2.5	Normal	Typical development at 12 mo CA
20	F	28+0	15	900	9/9/9	7.35	-5.5	Abnormal (PVC left)	Unilateral CP at 6 mo CA
21	F	28+5	16	1170	5/8/9	7.08	-10.3	Normal	Bayley III: typical development at 28 mo CA
22	M	29+1	13	1320	7/8/9	7.36	-3	Unspecific	Bayley III: typical development at 24 mo CA
23	M	23+6	15	-	-	-	-	Unspecific	Bayley III: DD at 31 mo CA
24	F	29+4	16	1756	5/6/7	7.36	-4.9	Normal	Clinically DD at 12 mo CA
25	M	29+4	16	850	2/2/2	0	0	Normal	Clinically DD at 12 mo CA
26	F	36+3	12	840	1/5/8	7.4	0	Normal	Clinically DD 12 mo CA
27	F	26+6	17	890	6/8/10	7.42	-0.7	Unspecific	Bayley III: typical development at 25 mo CA
28	M	25+4	15	655	4/9/9	7.4	-2.1	Normal	Clinically DD at 12 mo CA
29	M	41+1	16	4280	1/3/5	6.79	-27.2	Unspecific	Clinically DD at 12 mo CA

Legend: Abbreviations: sex = male (m), female (f), GA = gestational age, CA = corrected age; BW = bodyweight, HIE = Hypoxic ischaemic encephalopathy, PVL = periventricular leucomalasia, PVC = periventricular cyst, CP = cerebral palsy, DD = developmental delay, either based on performance on the Bayley Scales of Infant and Toddler Development (BSITD, Version 3) or on the clinical examination mo = months. Definition of CP in accordance with international expert recommendations.[38]; terminology of the Surveillance of Cerebral Palsy in Europe (SCPE; [56].

Legends to the figures

Figure 1 Study set-up

We recorded General Movements of Infants in the clinical setting of an outpatient clinic. We correlated expert GMA ratings of standard RGB videos with GMA ratings on SMIL motion videos of the same sequence. SMIL motion videos are the result of capturing 3D shape and motion with the SMIL model [34]. (further explanation see: <https://www.youtube.com/watch?v=aahF1xGurmM&feature=youtu.be>)



Figure 2: Conventional video recording and its corresponding SMIL motion video pose.

The upper panel shows 8 representative poses of a three minute video recording of an infant at fidgety age. The infant presents with typical motion features of complexity and variation during active wakefulness; his GMA complexity was rated as normal suboptimal (Likert-score 7). The lower panel are the corresponding SMIL motion video frames; also in these panels GM-complexity can be clearly observed. Note that fidgety movements can not be shown in the static poses of the figure.

Supplementary Figure S1: Bland-Altman Plot of correspondence of Likert-scale rating of GM-complexity comparing conventional RGB video and SMIL motion video.

The Bland-Altman plot depicts the agreement between the ratings of both videos. The Y axis shows the difference between the two paired measurements, the X-axis represents the average of these measures. Mean value of the 29 GMA rating differences = 0.034 (grey solid line). Standard deviation of mean value: ± 0.981 . Upper dotted line = upper reference limit (Mean value + $1.96 \times \text{Stddev}$) = 1.958. Lower dotted line = lower reference limit (Mean value - $1.96 \times \text{Stddev}$) = -1.888. Regarding the three outliers infant #3 (developmental delay) infant #6 (bilateral spastic CP) and infant #15 refer to the discussion section.

Supplementary Figure S2: ROC of sensitivity of GMA DA complexity rating based on the two types of video

Fig S2a: Receiver operating characteristic curve (ROC curve) of DA vs non-DA complexity rating based on conventional RGB video; S2b: similar curve but now based on rating of SMIL motion video. The lower panel suggests a superior diagnostic ability of the SMIL motion video GM-complexity rating of the binary classification system.