

NOTICE: This is the author's version of a work that was accepted for publication in *Gait & Posture*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms, may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Gait & Posture* Volume 36, Issue 2, June 2012, Pages 307-311, doi:10.1016/j.gaitpost.2012.03.017.

Kinematic Assessment of Stereotypy in Spontaneous Movements in Infants

Dominik Karch¹, Keun-Sun Kang², Katarzyna Wochner², Heike Philippi³,
Mijna Hadders-Algra⁴, Joachim Pietz², Hartmut Dickhaus¹

¹ Institut für Medizinische Biometrie und Informatik, Heidelberg University, Germany

² Center for Child and Adolescent Medicine, University Hospital Heidelberg, Germany

³ Center of Developmental Medicine, Frankfurt, Germany

⁴ Beatrix Children's Hospital - Institute of Developmental Neurology, University Medical Center, Groningen, Netherlands

Movement variation constitutes a crucial feature of infant motor development. Reduced variation of spontaneous infant movements, i.e. stereotyped movements, may indicate severe neurological deficit at an early stage. Hitherto evaluation of movement variation has been mainly restricted to subjective assessment based on observation. This article introduces a method for quantitative assessment yielding an objective definition of stereotyped movements which may be used for the prognosis of neurological deficits such as cerebral palsy (CP).

Movements of 3-months-old infants were recorded with an electromagnetic tracking system facilitating the analysis of joint angles of the upper and lower limb. A stereotypy score based on dynamic time warping has been developed describing movements which are self-similar in multiple degrees of freedom. For clinical evaluation, this measure was calculated in a group of infants at risk for neurological disorders (n=54) and a control group of typically developing children (n=21) on the basis of spontaneous movements at the age of three months. The stereotypy score was related to outcome at the age of 24 months in terms of CP (n=10) or no-CP (n=53). Using the stereotypy score of upper limb movements CP cases could be identified with a sensitivity of 90% and a specificity of 96%. The corresponding score of the leg movements did not allow for valid discrimination of the groups.

The presented stereotypy feature is a promising candidate for a marker that may be used as a simple and noninvasive quantitative measure in the prediction of CP. The method can be adopted for the assessment of infant movement variation in research and clinical applications.

1 Introduction

Movement variation is a hallmark of infant motor development. Beginning in early fetal life movements are characterized by an abundant variation of movement patterns which vary temporally and spatially [1]. Variation is also attributed to be a key aspect of general movements [2, 3, 4] which are the most frequently occurring movement patterns of the fetus and young infant.

It has been put forward that a lack of movement variation, i.e. monotonous movements of a stereotype quality, indicates an atypical motor development and can be used as an early marker of developmental disabilities [2], [3], [5, 6, 7]. On the other hand it has been noted that less variation of goal directed or intended movements represents good neurological function [8]. However, these less variable movements are not stereotyped but precise

or "consistent" as Piek named it. In motor development the so called "adaptive variability" means that among variable movement options the most suitable movement for task solution is selected.

Dusing and Harbourne [9] remark that multiple definition and quantification systems have limited the clinical interpretation of variation. This indecisiveness has hitherto hampered the comparison of research results regarding the meaning of low movement variation. Hence, an objective definition of stereotyped movements which facilitates quantitative measurements would be of great value for the description of infant motor development.

Previous attempts to describe movement variation were mainly based on subjective observation of overall movement patterns which can have questionable accuracy [10]. Furthermore, muscle activation has been examined using electromyography during spontaneous infant

movements [11]. However, this investigation focused on the developmental course of the spontaneous movements of healthy children and does not give an objective description of movement variation. Objective approaches studying the variation of kinematics were hitherto limited to the quantification of specific movement features such as joint angles during predefined movements of the lower limb such as kicking movements or walking [12], [13]. A kinematic description of variation of infant motility, especially of arbitrary movement patterns, is largely missing.

The novel method presented in this article aims at representing movements of the lower as well as the upper limb in all degrees of freedom of the corresponding joints. It allows for the quantification of arbitrary stereotyped movement patterns. The contribution of this article is twofold:

1. An objective definition of movement stereotypy in terms of kinematic parameters is provided.
2. The relationship between kinematically defined movement stereotypies at three months and the development of cerebral palsy is assessed.

2 Patients

Movements of a group of infants at risk for neurological disorders ($n=54$) and a control group ($n=21$) were recorded at the age of three months (calculated from the expected date of delivery) according to the following protocol: The infants were lying on a mattress in supine position and had ample space for spontaneous motility. Movements were recorded with an electromagnetic tracking system (3D Guidance medSAFE™, Ascension Technology) with a sampling rate of 50Hz. Eight sensors were tracked in two different configurations: At first, the sensors were attached to the right upper and lower limb. Motility was recorded for at least five minutes. This was possible in all infants. Next, the sensors were attached to both upper limbs and motility was recorded for another five minutes. This was possible in 1/3 of the children (14 of the risk group, 11 of the control group) because some of the children started crying or technical problems occurred. A specifically designed biomechanical model was used in order to calculate anatomically meaningful, interpretable movement parameters, hereby enabling repeatable measurements which are independent of the exact location of the attached sensors. Angles of the elbow, shoulder, knee and hip joints were derived from recorded data. These angle time-series constitute the degrees of freedom of the model (according to [14, 15, 16]). At the age of 24 months the infants attended a follow-up exam-

ination in order to determine if they had developed cp. They were examined by pediatric neurologists according to the definition given in [17]. According to the clinical diagnosis [18] ten children had developed spastic cp and 53 had not. 12 infants were lost to follow-up because parents had either opted to withdraw from the study or moved away from the study site. The study was approved by the Ethics Committee. Before the study, the experimental protocol was explained to all parents and their written informed consent was obtained.

3 Methods

The term “variation of movements” implies an abundance of different patterns which vary in speed, amplitude, involvement of participating joints, etc. Hence, movements lack variation if they possess a stereotype, monotonous quality. Such movements exhibit patterns which occur several times during a recording and show similar time-series trajectories in multiple degrees of freedom (e.g. elbow flexion-extension). They might differ to some extent in speed, amplitude or onset, but the basic shapes of the trajectories remain the same. These shapes depend on the kind of movement and are highly subject-specific. Fig. 1 shows an example of arm swipes whose basic shape is repeated several times.

In order to automatically quantify the degree of self-similarity segments of the time-series have to be identified which are to be assessed. Start and endpoints of segments which show monotonous characteristics are a priori unknown. Identifying them is a daunting task since it is not computationally feasible to compare all possible subsegments [19, 20, 21]. We tackle this problem heuristically by making a sensible guess: time-series are partitioned into segments of movement and rest. A movement segment in the interval $[t_1, t_2]$ is an element of the set of movement segments \mathbb{M} if the speed $v(t)$ in this segment exceeds some basic threshold $v_1 = 0.25 \text{ rad/s}$ at all times and exceeds at least once a higher threshold $v_2 = 1 \text{ rad/s}$ in both directions

$$\forall_{t_1 < t < t_2} v(t) > v_1 \wedge \max v(t) > v_2 \wedge \min v(t) < -v_2 \Rightarrow [t_1, t_2] \in \mathbb{M} \quad (1)$$

The thresholds were chosen empirically so that only segments which show distinct movements were included. All such segments are then assessed for similarity.

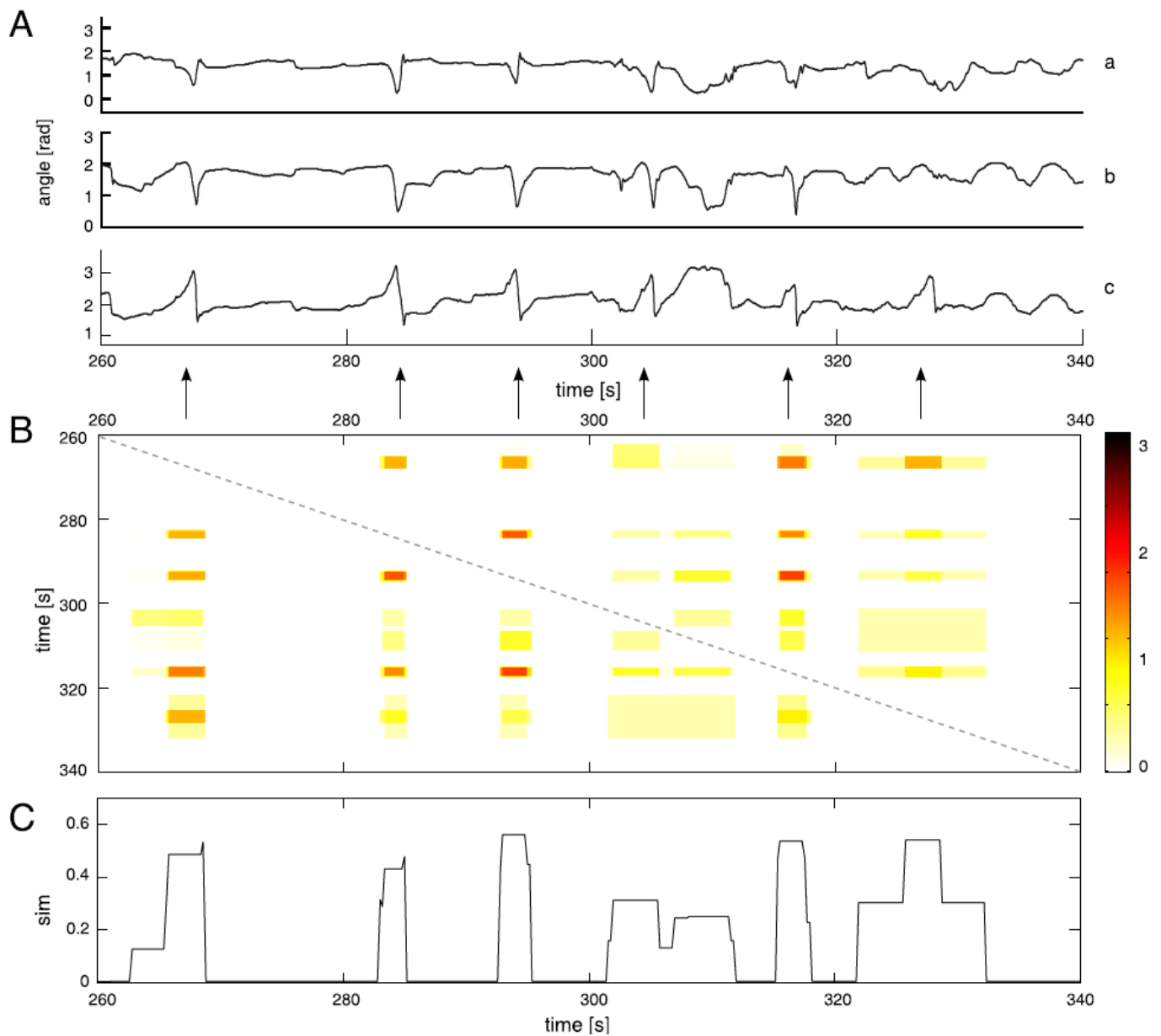


Figure 1: A) 80 s of a recording of arm movements (time-series represent flexion-extension of the elbow joint [a], elevation [b] and plane of elevation [c] of the shoulder joint) of a child who developed CP showing repetitive movements (see arrows). B) The diagram visualizes the similarity function (cf. Eq. 3) $sim(t_1, t_2)$. Each rectangle depicts the start and endpoints of two segments of the time-series which are similar to a certain extent. The color value represents the degree of similarity (see color bar). C) Plot of $sim(t)$ (cf. Eq. 4) representing the mean degree of similarity of each date t . This function is summarized with the stereotypy score $ster$ (cf. Eq. 6).

3.1 Assessment of stereotypy

A movement segment is considered to be stereotyped if its time-series trajectory shows a high degree of similarity to other segments of the same recording. Hence, movement segments have to be compared to each other in order to assess whether they share a common shape. We use the Dynamic Time-Warping (DTW) distance as a measure which complies with this property [22]. This similarity evaluation is performed on each time-series (e.g. flexion-extension of the knee joint). Results of all time-series of one limb can then be combined to yield a general assessment of self-similarity.

1. For every pair of movement segments (I_1, I_2) the DTW-distance $\text{dtw}(I_1, I_2)$ can be calculated. Hence, for each pair of dates (t_1, t_2) where t_1 resides in the interval of a movement segment I_1 and t_2 resides in the interval I_2 the distance is determined by calculating the DTW-distance of these segments. A distance of zero corresponds to a perfect accordance; with growing dissimilarity the distance grows. In order to obtain a similarity function which can be aggregated for several time-series, the distance is mapped to the range $[1, 0)$ using the exponential function, where 1 corresponds to maximal similarity. If one of the dates does not belong to a movement segment, the similarity is zero.

$$s_i(t_1, t_2) = \begin{cases} e^{-\text{dtw}(I_1, I_2)} & | \exists_{I_1, I_2 \in \mathbb{M}} t_1 \in I_1 \wedge t_2 \in I_2 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The function s_i evaluates the similarity between two dates of a time-series i .

2. The similarity values of all time-series of one limb can now be summed.

$$\text{sim}(t_1, t_2) = \sum_i s_i \quad (3)$$

Hence, the two-dimensional function sim evaluates the similarity of movements from one limb between each date t_1 and all possible other dates t_2 . Fig. 1B illustrates this function with an example.

3. A mean similarity $\text{sim}(t)$ is calculated by averaging the similarity values of all movement segments for each t

$$\text{sim}(t) = \frac{1}{|T_M|} \sum_{t_1 \in T_M} \text{sim}(t, t_1) \quad (4)$$

where T_M is the unification of all movement intervals, i.e. it is the set of all dates residing in a move-

ment interval

$$T_M = \left\{ \bigcup_{i=1}^{|M|} I_i | (I_i \in M) \right\} \quad (5)$$

Fig. 1C illustrates the function $\text{sim}(t)$ with an example.

4. Infants who exhibit stereotyped movements do not show them all of the time. In order to define a statistic which represents a sequence of these typical movements a moving average is applied to $\text{sim}(t)$ and the maximum value is chosen, constituting the stereotypy score ster

$$\text{ster} = \max_{0 < t < T-w} \frac{1}{w} \sum_{i=t}^{t+w} \text{sim}(t_i) \quad (6)$$

For clinical evaluation we calculated the stereotypy score based on the movements of the upper and the lower limb, respectively. If the scores could be calculated for both arms, the maximum value was taken. The distributions of scores were then compared to see if there is a relation between stereotyped movements and the outcome CP and no-CP. Then the receiver operating characteristic (ROC) was examined in order to evaluate the use of the stereotypy score for the prognosis of CP.

4 Results

The algorithm for the calculation of the stereotypy score is illustrated in Fig. 1: In the angle time-series similar shapes reappear several times synchronously in all degrees of freedom. The time-series are evaluated with the function $\text{sim}(t_1, t_2)$ (see Eq. 3), which is illustrated by the graph in Fig. 1B. With this graph it can be seen which segments of the recording are very similar to each other. This two-dimensional function is condensed to the one-dimensional function $\text{sim}(t)$ in Fig. 1C which gives the mean similarity of each date to all other movement segments. Comparing this function with the original time-series one can see that the algorithm is able to detect movement segments which exhibit stereotyped behavior. Observing stereotypy scores of both groups we observed that only the stereotypy score of the arm is suitable for their discrimination. We evaluated the influence of the parameter 'window size' in Eq. 6 regarding its influence on the discriminative ability measured by the Youden index [23] which is the maximum value of 'sensitivity + specificity - 1'. For sizes smaller than 210 s the index varied between 0.7 and 0.8; for sizes between 210 s and

420 s the index remained nearly constant around 0.85. For longer window sizes the index lowered again below 0.8. We chose a window size of 300 s. Table 1 shows the stereotypy scores. Children who did not develop CP showed similar values for arm and leg. In contrast, the CP group showed a lower stereotypy score for leg movements and high values for the upper limb. The distribution of the stereotypy score of the upper limb can be seen in Fig. 2. The suitability of this feature for the prognosis of CP can be evaluated with the ROC depicted in Fig. 3. Using the optimal threshold 0.036 a sensitivity of 90% can be obtained having a specificity of 96%.

5 Discussion

The kinematic definition of stereotyped movements as given in Eq. 1–6 and summarized in the stereotypy score concisely represents the phenomena of low variation in infant movements and allows for an objective assessment. The application to a group of children at risk for neurologic disorders and a control group revealed a strong correlation between the stereotypy score of the upper limb and the outcome CP. This novel approach differs from prior approaches because it defines a lack of variation as the self-similarity of arbitrary movements in various joints. This closely resembles the visual, subjective notion of stereotypy.

In contrast, prior approaches regarding statistical variation of movements are based on the measurement of signals which are supposed to be related to movement variation. For instance, the center of pressure which can be measured when the infant is sitting or moving in supine position [24], [25] can be considered as a signal that can capture some aspect of movement variation of the whole body. Similarly, Adde et al. [26] calculated one ‘centroid of motion’ time-series from video recordings which is supposed to represent the distribution of motility regarding the four limbs. This kind of signals is commonly analyzed by the means of linear statistics, such as the standard deviation or the frequency content of the signal’s distribution [27], or nonlinear techniques such as entropy measures or the embedding dimension of a signal [28]. Summary statistics like standard deviation do not consider intra-individual variation of non-stationary signals. Non-linear techniques have the potential to capture these variations, but it might be very difficult to interpret them. E.g., Ohgi et al. analyzed the acceleration signal of a sensor attached to the wrist of three healthy infants and three infants with brain injury and found embedding dimensions with different degrees of freedom. But as Fetters et al. [29] pointed out, it is difficult to say what a difference

between embedding dimensions 7 and 8 actually means. To be useful non-linear techniques should either allow for the discrimination of groups or allow for an interpretation; otherwise they cannot foster a better understanding of the nature of movement variation.

A different class of approaches for the quantitative description of stereotype movements is based on the classification of predefined, labeled patterns which often possess periodic properties. E.g., Westeyn et al. [30] used activity recognition based on acceleration signals for the identification of seven stereotype patterns in autistic children such as drumming or rocking. In a similar fashion, Goodwin et al. [31] performed recognition of predefined, periodic movements in autistic children. These approaches do not fit our application because we strived for the quantitative description of a qualitative impression of a ‘lack of variation’. In our case, neither are there predefined patterns nor are the movements periodic (see e.g. Fig. 4 for an example of movements of the CP group which are not periodic and were attributed a high stereotypy score). Hence, our approach can be applied usefully if there are no predefined patterns.

While children of the no-CP group showed similar stereotypy values for arm and leg, most children of the CP group showed a low stereotypy score for the lower limb, but high values for the upper limb. Indeed, many of those children showed few distinct leg movements but quite repetitive arm movements. Hence, the two groups could be distinguished using the stereotypy score for the upper limb, not the lower limb. There were two cases from the no-CP group which showed high stereotypy scores. These infants did not develop CP, but both showed neurological dysfunction at the age of 24 months. One child with outcome CP did not show a high stereotypy score. This recording was characterized by movements of low amplitude (see Fig. 4). Our algorithm is based on shape recognition, and there are few shapes to detect in this case. Obviously the movement behavior of this child differs from the rest of the group. It might be necessary to describe this kind of movements with a different feature.

Three properties turned out to be crucial for the ability of the stereotypy score of the upper arm to differentiate between the groups. Firstly, the score is based on the gross movements of the elbow and shoulder joints. The time-series of the hand joint do not contribute to the differentiation: if the hand is included, the Youden index is lowered to 0.78. Secondly, atypical stereotypy is expressed during a long enough movement phase, not during the whole recording. Thirdly, the score implicitly gives higher priority to shorter segments, since longer ones tend to have lower similarity due to their more com-

Table 1: Robust measures of location (median) and dispersion (interquartile range, IQR) of the stereotypy scores of arm and leg for the groups CP and no-CP. In the children who did not develop CP values for leg and arm movements are similar. In contrast, the CP-group shows a low stereotypy score for leg movements and very high values for the upper limb. For discrimination of both groups the stereotypy score of the upper limb is suitable due to the different medians and the low dispersion. Fig. 2 shows the corresponding distributions.

	ARM	LEG
no-CP: MEDIAN (IQR)	14.3 (19.7)	18.4 (23.0)
CP: MEDIAN (IQR)	41.7 (16.9)	17.3 (19.6)

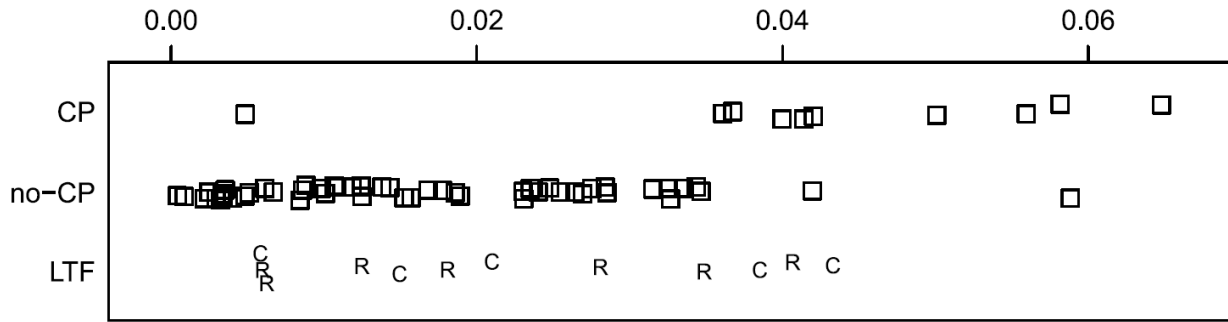


Figure 2: Distribution of the stereotypy score of the upper limb for the groups no-CP and CP (values for infants from the risk (R) and control (C) group that were lost to follow up (LTF) are given in the last row). Nine of the ten recordings of the CP group show very high values.

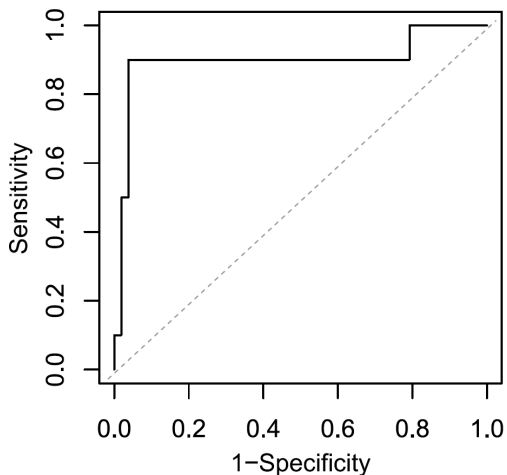


Figure 3: Receiver operating characteristic of the relationship between the stereotypy score of the arm and the outcome CP. The threshold corresponding to the maximal Youden index results in a sensitivity of 90% with a specificity of 96%.

plex structure. Hence, long and continuous movements are not attributed high scores. The distribution of segment lengths is similar in both groups (median values 6.8 s and 5.5 s) with few very long segments in the no-CP group. In order to explore the relation between this kind of movements and the outcome, we split all long segments to smaller segments of a maximum duration of 5 s. As a consequence, the differences between the groups became smaller. That means that it is the short, stereotyped movement segments (as seen in Fig. 1) that are pivotal for the characterization of the infants with outcome CP.

There are limitations to our approach. Some children show movements preferably on one side, in this case the stereotypy score might be underestimated if only the opposite side of the body is tracked. However, this is a weakness due to the technical limitations, not the method. Since not all of the infants possess recordings of both arms, one could argue that this biases the results regarding the differences between the groups CP and no-CP. However, we do not expect there to be a bias, since the presence of a left arm recording depends on neither outcome nor group membership.

We applied the method to a sample of infants of which ten developed spastic CP which is the most common form

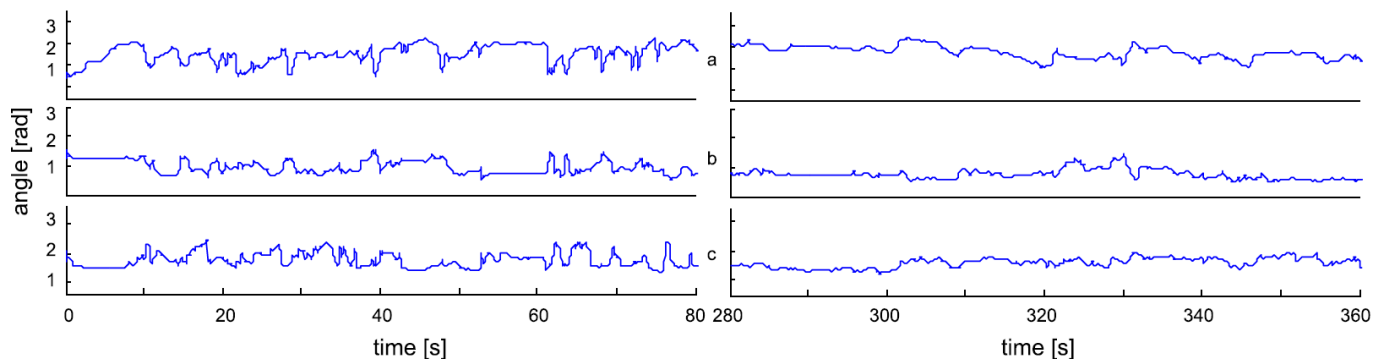


Figure 4: Arm movements (time-series according to Fig. 1) of two children who developed cp. Left side: this infant exhibited a high stereotypy score. There are no notable periodic movements. Right side: this infant mainly shows movements of low amplitude. The stereotypy score is not adequate for the description of such movements since it is based on the recognition of shapes.

of cp. Since there were no cases of dyskinetic and atactic cp, this sample is not representative for all forms of cp. However, due to the inclusion of a control group, we think that the stereotypy score is a good candidate for a prognostic marker for cp. Further studies are needed to replicate clinical validity and explore generalizability.

References

- [1] M Hadders-Algra. The Neuronal Group Selection Theory: a framework to explain variation in normal motor development. *Developmental Medicine & Child Neurology*, 42:566–572, 2000.
- [2] F Ferrari, G Cioni, and HFR Prechtl. Qualitative changes of general movements in preterm infants with brain lesions. *Early Human Development*, 23(3):193–231, 1990.
- [3] M Hadders-Algra. General Movements: A Window for Early Identification of Children at High Risk for Developmental Disorders. *Journal of Pediatrics*, 145:12–18, 2004.
- [4] M Hadders-Algra. Variation and variability: key words in human motor development. *Physical therapy*, 90(12):1823–37, December 2010.
- [5] BCL Touwen. How normal is variable, or how variable is normal? *Early Human Development*, 34(1-2): 1–12, September 1993.
- [6] C Einspieler and HFR Prechtl. Prechtl’s assessment of general movements: a diagnostic tool for the functional assessment of the young nervous system. *Mental retardation and developmental disabilities research reviews*, 11(1):61–7, January 2005.
- [7] HFR Prechtl, C Einspieler, G Cioni, AF Bos, F Ferrari, and D Sontheimer. An early marker for neurological deficits after perinatal brain lesions. *The Lancet*, 349:1361–1363, 1997.
- [8] J Piek. The role of variability in early motor development. *Infant Behavior and Development*, 25(4): 452–465, 2002.
- [9] SC Dusing and RT Harbourne. Variability in postural control during infancy: implications for development, assessment, and intervention. *Physical therapy*, 90(12):1838–49, December 2010.
- [10] DAM Pyles, MM Riordan, and JS Bailey. The stereotypy analysis: An instrument for examining environmental variables associated with differential rates of stereotypic behavior. *Research in Developmental Disabilities*, 18(1):11–38, February 1997.
- [11] M Hadders-Algra, LA Van Eykern, AWJ Klip-Van Den Nieuwendijk, and HFR Prechtl. Developmental course of general movements in early infancy. II. EMG correlates. *Early Human Development*, 28(3): 231–251, March 1992.
- [12] S Jeng, L Chen, and K Yau. Kinematic analysis of kicking movements in preterm infants with very low birth weight and full-term infants. *Physical Therapy*, 82(2):148–59, March 2002.
- [13] L Fetters, Y Chen, J Jonsdottir, and EZ Tronick. Kicking coordination captures differences between full-term and premature infants with white matter disorder. *Human Movement Science*, 22(6):729–748, 2004.

- [14] D Karch, K S Kim, K Wochner, J Pietz, H Dickhaus, and H Philippi. Quantification of the segmental kinematics of spontaneous infant movements. *Journal of Biomechanics*, 41(13):2860–2867, 2008.
- [15] D Karch, K Kim, K Wochner, H Philippi, J Pietz, and H Dickhaus. Compensation of large motion sensor displacements during long recordings of limb movements. *Journal of Biomechanics*, 43:1844–48, 2010.
- [16] D Karch. *Quantitative Analyse der Spontanmotorik von Säuglingen für die Prognose der infantilen Cerebralparese*. PhD thesis, Universität Heidelberg, 2011.
- [17] SCPE. Prevalence and characteristics of children with cerebral palsy in Europe. *Developmental Medicine & Child Neurology*, 44:633–640, 2002.
- [18] P Baxter, C Morris, M Goldstein, M Bax, and A Colver. The Definition and Classification of Cerebral Palsy. *Developmental Medicine & Child Neurology*, 49(s109):1–44, 2007.
- [19] F Duchêne, C Garbay, and V Rialle. Learning recurrent behaviors from heterogeneous multivariate time-series. *Artificial intelligence in medicine*, 39(1):25–47, January 2007.
- [20] TS Huang. *Learning to extract temporal signal patterns from temporal signal sequence*. IEEE Comput. Soc, September 2000. ISBN 0-7695-0750-6.
- [21] J Buhler and M Tompa. Finding motifs using random projections. *Journal of computational biology*, 9(2):225–42, January 2002.
- [22] H Sakoe and S Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(1):43–49, 1978.
- [23] W Youden. Index rating for Diagnostic Test. *Cancer*, 3:32–35, 1950.
- [24] SC Dusing, A Kyvelidou, VS Mercer, and N Stergiou. Infants born preterm exhibit different patterns of center-of-pressure movement than infants born at full term. *Physical therapy*, 89(12):1354–62, December 2009.
- [25] A Kyvelidou, RT Harbourne, WA Stuberg, J Sun, and N Stergiou. Reliability of center of pressure measures for assessing the development of sitting postural control. *Archives of physical medicine and rehabilitation*, 90(7):1176–84, July 2009.
- [26] L Adde, JL Helbostad, AR Jensenius, G Taraldsen, and R Støen. Using computer-based video analysis in the study of fidgety movements. *Early Human Development*, 85(9):541–7, 2009.
- [27] D Gilden, T Thornton, and M Mallon. 1/f noise in human cognition. *Science*, 267(5205):1837–1839, March 1995.
- [28] S Ohgi, S Morita, K Khee Loo, and C Mizuike. Time Series Analysis of Spontaneous Upper-Extremity Movements of Premature Infants With Brain Injuries. *Physical Therapy*, 88(9):1022–1033, 2008.
- [29] L Fethers and JP Scholz. Commentary on Time series analysis of spontaneous upper-extremity movements of premature infants with brain injuries. *Physical Therapy*, 88(9):1034–1036, 2008.
- [30] T Westeyn, K Vadas, X Bian, T Starner, and GD Abowd. Recognizing mimicked autistic self-stimulatory behaviors using hmms. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, pages 164–167. IEEE, 2005.
- [31] MS Goodwin, SS Intille, F Albinali, and WF Velicer. Automated Detection of Stereotypical Motor Movements. *Journal of autism and developmental disorders*, September 2010.