Supplementary Materials for
Prediction of autistic tendencies at 18 months of age via markerless video analysis of spontaneous body movements in 4-month-old infants

Hirokazu Doi\(^1\), Naoya Iijima\(^2\), Akira Furui\(^3\), Zu Soh\(^3\),
Rikuya Yonei\(^4\), Kazuyuki Shinohara\(^4\), Mayuko Iriuchi\(^4\),
Koji Shimatani\(^5\), and Toshio Tsuji\(^3\)\(^*\)

\(^{1}\)Department of Science and Engineering, Kokushikan University, Japan.
\(^{2}\)Graduate School of Engineering, Hiroshima University, Japan.
\(^{3}\)Graduate School of Advanced Science and Engineering, Hiroshima University, Japan.
\(^{4}\)School of Engineering, Hiroshima University, Japan.
\(^{5}\)Graduate School of Biomedical Sciences, Nagasaki University, Japan.
\(^{6}\)Faculty of Health and Welfare, Prefectural University of Hiroshima, Japan.

\(^{*}\) tsuji-c@bsys.hiroshima-u.ac.jp

The PDF file includes:

Supplementary Texts
Supplementary Tables S1–S3
Supplementary Figure S1
Supplementary Table S1: Mean scores of questionnaires sent to mother-infant pairs, excluding M-CHAT. Standard deviations are in parentheses; statistics $W$ and $p$ values obtained from between-group comparisons by the Brunner-Munzel test are also shown.

| Questionnaire | 4 months | 18 months |
|---------------|----------|-----------|
|               | High risk | Low risk  | $W$    | $p$-value |
| CESD          | 9.429 (3.505) | 5.029 (5.595) | 3.766 | $1.611 \times 10^{-3}$ |
| EPDS          | 3.429 (3.207) | 2.765 (3.358) | 0.493 | 0.636 |
| Bonding scale | 3.714 (2.812) | 1.147 (1.987) | 1.972 | 0.090 |
| KIDS          | 46.571 (7.480) | 48.824 (8.193) | -0.805 | 0.440 |
|               | 14.429 (3.919) | 7.441 (7.287) | 4.706 | $3.417 \times 10^{-5}$ |
| Bonding scale | 4.143 (4.018) | 2.147 (2.311) | 0.801 | 0.451 |
| KIDS          | 48.571 (14.444) | 64.059 (11.040) | -3.537 | $6.459 \times 10^{-3}$ |
Supplementary Table S2: List of items in the Japanese version of M-CHAT

| Item no. | M-CHAT item                     | Critical |
|----------|--------------------------------|----------|
| 1        | Enjoys being swung              |          |
| 2        | Interest in other children      | ✓        |
| 3        | Climbs up stairs                |          |
| 4        | Enjoys peek-a-boo               |          |
| 5        | Pretend play                    |          |
| 6        | Imperative pointing             | ✓        |
| 7        | Declarative pointing            | ✓        |
| 8        | Functional play                 |          |
| 9        | Brings objects to show          | ✓        |
| 10       | Eye contact                     |          |
| 11       | Oversensitive to noise          |          |
| 12       | Responds to smile               |          |
| 13       | Imitation of action             | ✓        |
| 14       | Responds to name                | ✓        |
| 15       | Point following                 | ✓        |
| 16       | Walking                         |          |
| 17       | Gaze-following                  |          |
| 18       | Unusual finger movement         |          |
| 19       | Gaining parent’s attention      |          |
| 20       | Wondering hearing               | ✓        |
| 21       | Understands what is said        | ✓        |
| 22       | Stares at nothing               |          |
| 23       | Social reference                | ✓        |
Supplementary Texts:
Definitions of Bodily Movement Features

Movement Magnitude

Movement frequency \((I_1)\)

This feature is defined as the proportion of the number of frames with bodily movement between the first and \(S\)-th sub-segments using the following equations:

\[
(A_k) I_1 = \frac{100}{L} \sum_{s=1}^{S} \sum_{l=1}^{L_s} (A_k)_{kl},
\]

\[
(A_k)_{kl} = \begin{cases} 
1 & ((A_k) M_l \geq M_{th}) \\
0 & ((A_k) M_l < M_{th}) 
\end{cases}
\]

where \(L\) is the total number of frames in each sub-segment, and \(M_{th}\) is the threshold for determining whether movement has occurred. In this paper, we set \(M_{th} = 0.005\).

Movement strength \((I_2)\)

This feature is defined as the magnitude of movement per unit time between the first and \(S\)-th segments using the following equations:

\[
(A_k) I_2 = \frac{1}{L'} \sum_{s=1}^{S} \sum_{l=1}^{L_s} (A_k)_{vl},
\]

\[
(A_k)_{vl} = \begin{cases} 
(A_k) M_l & ((A_k) M_l \geq M_{th}) \\
0 & ((A_k) M_l < M_{th}) 
\end{cases}
\]

where \(L'\) is the total number of frames in which \((A_k) M_l\) is \(M_{th}\) or more between the first and \(S\)-th segments.

Movement count \((I_3)\)

This feature is defined as the number of continuous movement detections, divided by the total number of frames \(L\):

\[
(A_k) I_3 = \frac{1}{L} \sum_{s=1}^{S} (A_k) Q_s,
\]

where \((A_k) Q_s\) represents the number of movements detected in the \(s\)-th video sub-segment. A movement is detected when \((A_k) M_l\) becomes \((A_k) M_l \geq M_{th}\) once and then falls below \(M_{th}\) again.
Movement Balance

Ratio of Movement Frequency ($I_4$)

This feature is defined as the ratio of the movement frequency between the bodily regions $A_{k_1}$ and $A_{k_2}$:

$$\frac{(A_{k_1}, A_{k_2})}{(A_{k_2})} I_4 = \frac{(A_{k_1})}{(A_{k_2})} I_1.$$  \hspace{1cm} (6)

Note that if $(A_{k_2}) I_1 = 0$, then $(A_{k_1}, A_{k_2}) I_4 = 0$.

Ratio of Movement Strength ($I_5$)

This feature is defined as the ratio of the movement strength between the bodily regions $A_{k_1}$ and $A_{k_2}$:

$$\frac{(A_{k_1}, A_{k_2})}{(A_{k_2})} I_5 = \frac{(A_{k_1})}{(A_{k_2})} I_2.$$  \hspace{1cm} (7)

Note that if $(A_{k_2}) I_2 = 0$, then $(A_{k_1}, A_{k_2}) I_5 = 0$.

Movement Coordination ($I_6$)

Movement coordination between the bodily regions $A_{k_1}$ and $A_{k_2}$ is calculated as the correlation coefficient between the time-series data of $(A_{k_1}) M_l$ and $(A_{k_2}) M_l$. Correlation coefficients were computed within sliding temporal windows with a length of 300 frames. The stride of the sliding window was one frame. Correlation coefficients were then averaged to yield a single feature of the movement coordination.

Movement Rhythm

Central Frequency ($I_7$) and Second Moment around Central Frequency ($I_8$) of Motor Alteration

The time-series data of $(A_{k_i}) M_l$ of each video sub-segment were analyzed in the frequency domain. Within each video sub-segment, the time-series data within a sliding window with the length of 128 frames were subjected to the fast Fourier transformation. The stride of the moving window was one frame. Then, the average power spectrum density (PSD), $P(f)$, was computed by grand-averaging PSDs across all the sliding windows in all the video sub-segments. Based on $P(f)$, central frequency ($F_{\text{centr}}$) and second moment around the central frequency
(D$_{cntr}$) were computed according to the following equations:

$$F_{cntr} = \frac{\sum_{f=0}^{f_{max}} f P(f)}{\sum_{f=0}^{f_{max}} P(f)},$$  \hspace{1cm} (8)$$

$$D_{cntr} = \sqrt{\frac{\sum_{f=0}^{f_{max}} P(f)(f - F_{cntr})^2}{\sum_{f=0}^{f_{max}} P(f)}},$$  \hspace{1cm} (9)$$

where $f_{max} = 15$ Hz is the maximal frequency range for analysis. $F_{cntr}$ and $D_{cntr}$, computed on the basis of $P(f)$, correspond to ($A_k$)$_I^7$ and ($A_k$)$_I^8$, respectively.

Central Frequency ($I_{9x}$, $I_{9y}$) and Second Moment around Central Frequency ($I_{10x}$, $I_{10y}$) of Body Center Velocity

The time-series data of the body center velocity along the $x$- (top-bottom) and $y$- (left-right) axes were subjected to the fast Fourier transformation. Based on the resulting PSD, the central frequency and the second moment around the central frequency were computed in essentially the same manner as ($A_k$)$_I^7$ and ($A_k$)$_I^8$.

Central Frequency ($I_{11x}$, $I_{11y}$) and Second Moment around Central Frequency ($I_{12x}$, $I_{12y}$) of Body Center Fluctuation

The time-series data of body center fluctuation along the $x$- (top-bottom) and $y$- (left-right) axes were subjected to the fast Fourier transformation. Based on the resulting PSD, the central frequency and the second moment around the central frequency were computed in essentially the same manner as ($A_k$)$_I^7$ and ($A_k$)$_I^8$.

Movement of the Body Center

Variation in the Body Center Velocity ($I_{13}$)

The average absolute values of the body center velocity along the $x$- ($G_{v,l,x}^v$) and $y$-axes ($G_{v,l,y}^v$) were computed using the following equations:

$$\langle (A_r) I_{13x}, (A_r) I_{13y} \rangle = \left( \frac{1}{T} \sum_{s=1}^{S} \sum_{l=1}^{L_x} |G_{v,l,x}^v|, \frac{1}{T} \sum_{s=1}^{S} \sum_{l=1}^{L_y} |G_{v,l,y}^v| \right).$$ \hspace{1cm} (10)$$

Standard Deviation of Body Center Fluctuation ($I_{14}$)

The standard deviations of the body center fluctuations along the $x$- ($G_{l,x}^d$) and $y$- axis ($G_{l,y}^d$) within the sliding temporal window of length $L_g$ are calculated as
follows:

\[
\sigma_j = \sqrt{\frac{1}{L_g} \sum_{l=1}^{L_g} \left( G_{l,j}^d - \overline{G}_j^d \right)^2},
\]

where \( j \in \{x, y\} \), \( L_g = 300 \) frames, and \( \overline{G}_j^d \) represents the average value of \( G_{l,j}^d \) within the window. The \( \sigma_j \) is calculated in all video sub-segments, in the same way; then, \( I_{14} \) is calculated by averaging all \( \sigma_j \).

**Area of Body Center Excursion (I_{15})**

This feature represents the area surrounded by the outermost circumference of the trajectory of the body center excursion. First, in the sliding temporal window of length \( L_g = 300 \) frames, the outermost circumference of the trajectory of \( (G_{x}^d, G_{y}^d) \) are arranged in a clockwise direction based on the method proposed by Kim *et al.* [1] and defined as \( (G_{b,x}^d, G_{b,y}^d) \) \( (b = 1, 2, \ldots, B; B \) is the total number of points of the outermost circumference). The area surrounded by the outermost circumference is then calculated using the following equation:

\[
\frac{1}{2} \sum_{b=1}^{B} \left| G_{b,x}^d G_{b+1,y}^d - G_{b,y}^d G_{b+1,x}^d \right|.
\]

Note that if \( b = B \), then \( b + 1 = 1 \). This calculation is performed continuously, sliding the window by one frame. The areas are calculated in all video sub-segments, and the average of all obtained areas is defined as the feature \( (A_u)_{I_{15}} \).

**References**

[1] Kim, G., Ferdjallah, M., Harris, G. F., Fast computational analysis of sway area using center of pressure data in normal children and children with cerebral palsy, *Am. J. Biomed. Sci.*, 1, 364–372, (2009).
Supplementary Texts:
Analysis of the relationship between bodily movement features and ordinal scores of M-CHAT items

Methods

To investigate the relationship between motor features at 4 months and ordinal scores of M-CHAT items at 18 months, regression analysis based on a generalized linear model (GLM) was performed. The subjects were 41 infants (seven in ASD high-risk group and 34 in ASD low-risk group) as in the main text.

The objective variable of the GLM was the ordinal score of the M-CHAT items at 18 months (the number of failed items on the M-CHAT), and the explanatory variable was the 26 motor features obtained from the video at 4 months. Because the ordinal scores of the M-CHAT items can be regarded as count data with an upper limit, a binomial distribution was set for the GLM error structure, and a logit function was selected for the link function. Forward-backward stepwise model selection based on Akaike’s information criterion (AIC) was performed to search for the best model for predicting the ordinal scores. The prediction accuracy of the selected model was evaluated based on leave-one-out cross-validation.

Results and discussion

Supplementary Table S3 shows the information on the best regression model obtained from the stepwise analysis. In total, ten features were selected as the best features for predicting the ordinal scores of the M-CHAT items. The results of the model accuracy evaluation by cross-validation showed a mean absolute error between the predicted and true scores of 0.459. The relationship between the true and predicted scores is shown in Supplementary Figure S1. A significant positive correlation was found between both scores ($r = 0.6214; p < 0.0001$)

Effective movement features were almost consistent whether the ASD high- and low-risk groups were compared dichotomously based on a cutoff criterion or the M-CHAT score itself was regressed. Specifically, among ten features selected by this analysis, the following eight features showed significant differences between the ASD high- and low-risk groups (see main text): $(A_6)I_2$, $(A_7)I_{10y}$, $(A_7)I_{11x}$, $(A_7)I_{11y}$, $(A_7)I_{12y}$, $(A_7)I_{13y}$, $(A_7)I_{14x}$, and $(A_7)I_{15}$. These results support the validity of the proposed video-based movement analysis method for assessing ASD-like behaviors in infants.
Supplementary Table S3: Features selected by the stepwise analysis based on a generalized linear model.

| Features | Estimate | Std. Error | z-value | p-value |
|----------|----------|------------|---------|---------|
| (Intercept) | -0.654 | 4.944 | -0.132 | 0.894 |
| \((A_6)I_2\) | -264.424 | 94.0 | -2.813 | 4.908 \times 10^{-3} |
| \((A_7)I_{11x}\) | 20.277 | 7.297 | 2.779 | 5.457 \times 10^{-3} |
| \((A_7)I_{11y}\) | -37.753 | 12.295 | -3.071 | 2.136 \times 10^{-3} |
| \((A_7)I_{12y}\) | 36.693 | 10.859 | 3.379 | 7.270 \times 10^{-4} |
| \((A_7)I_{13y}\) | 221.799 | 66.228 | 3.349 | 8.110 \times 10^{-4} |
| \((A_7)I_{14x}\) | 210.40 | 136.119 | 1.546 | 0.122 |
| \((A_7)I_{15}\) | -3015.140 | 974.782 | -3.093 | 1.980 \times 10^{-3} |
| \((A_7)I_{9x}\) | -18.689 | 6.284 | -2.974 | 2.939 \times 10^{-3} |
| \((A_7,A_6)I_6\) | 23.991 | 7.717 | 3.109 | 1.878 \times 10^{-3} |
| \((A_7)I_{10y}\) | 12.994 | 5.944 | 2.186 | 0.0288 |

\((A_6)I_2\): Movement strength (lower body). \((A_7)I_{11x}\): Center frequency of body center fluctuations (x axis). \((A_7)I_{11y}\): Center frequency of body center fluctuations (y axis). \((A_7)I_{12y}\): Second moment around \((A_7)I_{11y}\). \((A_7)I_{13y}\): Variations in body center velocity (y axis). \((A_7)I_{14x}\): Variations in body center fluctuations (x axis). \((A_7)I_{15}\): Area of body center excursion. \((A_7)I_{9x}\): Center frequency of body center velocity (x axis). \((A_5,A_6)I_6\): Symmetry in upper and lower body. \((A_7)I_{10y}\): Second moment around \((A_7)I_{9x}\).
Supplementary Figure S1: Relationship between the true and predicted values of ordinal scoring of M-CHAT items. The solid black line and the equation show the results of fitting a linear regression model to the relationship. Pearson’s correlation coefficient (Pearson’s $r$) and the results of the test of no correlation are also shown.