Electrophysiological Abnormalities in Angelman Syndrome Correlate With Symptom Severity

Supplementary Information

Supplementary Methods

Study Population

The reported data were obtained as part of the AS Natural History Study (ASNHS) (ClinicalTrials.gov Identifier: NCT00296764), a longitudinal multi-center study of AS. Consent was obtained according to the Declaration of Helsinki and was approved by the institutional review boards of the participating sites. Subsets of these data have been analyzed previously (1–6).

For our analyses, we used data from participants with deletion AS genotype (the common class I and II deletions only (7)) and visits where both EEG and clinical scale assessments were performed. In a few instances, EEGs and clinical scales were not assessed on the same day but were separated by no more than two weeks. Data presented here are from children and adolescents (1 – 18 years). We focused on this age range since there were very limited data available for individuals younger than 1 year (n = 3 visits) or older than 18 years (n = 1 visit).

In total, we analyzed 72 visits from 45 individuals with deletion AS (class I: 23 visits from 16 individuals; class II: 49 visits from 29 individuals). 16 participants had more than 1 and up to 5 visits. 42 individuals had at least one visit after the onset of epilepsy. The Clinical Severity Scale (CSS) data were missing for 4 visits from 2 participants (see Supplementary Table S1 for a detailed breakdown).

The age of individuals was 59 ± 40.1 months (mean ± SD) with a sex ratio of 15:30 (F:M), see Supplementary Tables S2 and S3 for detailed breakdowns of age and sex by subpopulations entering the different analyses. The sex ratio deviated from the expected approximate equality. We tested if the clinical scales, the age of epilepsy onset and the delta-band EEG differ with sex (see Supplementary Table S4). This was not the case and we did not further account for sex in this study.

EEG

EEGs were performed in the clinical setting with an international 10-20 system (19 electrodes) montage and comprised wake sections as well as sleep for some participants (1). Data were acquired using BioLogic and Xltek systems (Natus Medical Incorporated, Pleasanton, California; sampling rates: 200 Hz, 256 Hz, and 512 Hz) with gold disc electrodes or gold plated silver silver-chloride electrodes and impedances were below 10kΩ. The analyses reported herein are restricted to sections where participants were awake.

Offline, EEG data were bandpass filtered (0.5 – 45 Hz, FIR filter) and cleaned of artifacts (interpolation of artifactual channels; exclusion of sections with strong artifacts; removal of independent components that capture EOG, EMG or technical artifacts) and re-referenced to average across all channels for analysis. Supplementary Table S5 provides details on the artifact rejection. This careful pre-processing was done to ensure that signals analyzed reflect brain activity rather than technical artifacts such as eye movements.

Power spectral estimates were derived for logarithmically-scaled frequencies ranging from 1 to 32 Hz (f/σ_f = 8.7) using Morlet Wavelets (8). Center frequencies were spaced logarithmically according to the exponentiation of the base 2 with exponents ranging from 0 (1 Hz) to 5 (32 Hz) in steps of 1/8. Power spectral estimates were derived as average power values of successive 3/4-overlapping temporal windows. Absolute power values were then averaged across electrodes (where applicable) and log-transformed to have units $10*\log_{10}(\mu V^2/\log_2(Hz))$. Consequently, differences between signals have the unit decibel (dB). We used dB as a unit as well in graphs where no changes are shown. The main EEG parameter was power averaged across all electrodes at the delta peak frequency of 2.8 Hz that we identified in our prior work (1).

Clinical Scales

We quantified AS symptoms using three clinical scales:

1. Growth-scale scores (GS) from the Bayley Scales of Infant and Toddler Development, Third edition (Bayley-3; scores included: cognitive, expressive communication, receptive communication, fine motor, gross motor) (9). We used GS rather than raw scores, which transform the data into an interval scale with defined mean and standard deviation (10). We analyzed the mean of the five growth scale values ("Bayley-3 Mean") as a proxy for global development before analyzing individual domains. This is not a commonly used metric but we found it helpful as an initial single test, which was then followed by testing individual scores.
2. The raw scores from the Vineland Adaptive Behavior Scales, Second edition (VABS-II-II) (11). The following 8 sub-domains were analyzed: communication (receptive, expressive), daily living skills (personal), socialization (interpersonal relationships, play and leisure time, coping skills), motor skills (gross and fine). We excluded the following three sub-domains because our prior work showed they exhibited floor effects in the AS population (4): communication (written), daily living skills (domestic, community).
3. The Clinical Severity Scale (CSS), a scale developed for the ASNHS (4). On the CSS, higher scores indicate greater impairment. To simplify the interpretation of correlation values in tables and figures, we report the correlation with the negative of the CSS score (i.e. "-CSS").

All scores analyzed (Bayley-3 scores, VABS-II sub-domains, CSS) have a non-linear developmental trajectory that we account for by using linear mixed-effects regression models with 3rd order polynomials as described in (4).

Supplementary Discussion

Estimating true correlations

The true correlation between two variables is reduced by the reliability with which each can be measured (12,13). In fact, if the test-retest correlations for the EEG metric and the clinical scales were known, the true underlying association could be estimated as:

$$r_{EEG,Scale} = \frac{\hat{r}_{EEG,Scale}}{\sqrt{r_{EEG-trt}}\sqrt{r_{Scale-trt}}}$$

, where $\hat{r}_{EEG,Scale}$ is the measured correlation between an EEG metric and a clinical parameter (as reported in the main text), $r_{EEG-trt}$ and $r_{Scale-trt}$ are the test-retest correlations of the EEG metric and the clinical parameter, respectively, and $r_{EEG,Scale}$ is the "true correlation" between EEG and clinical scales.

We currently do not know the test-retest correlations for EEG metrics and clinical parameters in AS that we analyzed here. However, we can derive a rough estimate. For the clinical scales investigated here, the re-test correlation over repeated assessments, which were at least one year apart, is ~0.5 – 0.7 (4). Test-retest correlation should be evaluated on a much shorter time-scale and are expected to be higher. We do not know the retest correlation of the AS EEG metrics, but generally EEG metrics have relatively high test-retest reliability.

Assuming 0.8 for both the clinical parameter and the EEG metrics would lead to a correction factor of 1.25; i.e. the true correlation would be 25% higher than the estimate reported in the Discussion section in the main text. As an example, the correlations derived for the "Bayley-3 Mean" as a proxy for global development of r = -0.38 (cross-sectional) and r = -0.55 (longitudinal), would have true underlying correlations of about r = -0.48 (cross-sectional) and r = -0.69 (longitudinal) under these assumptions.

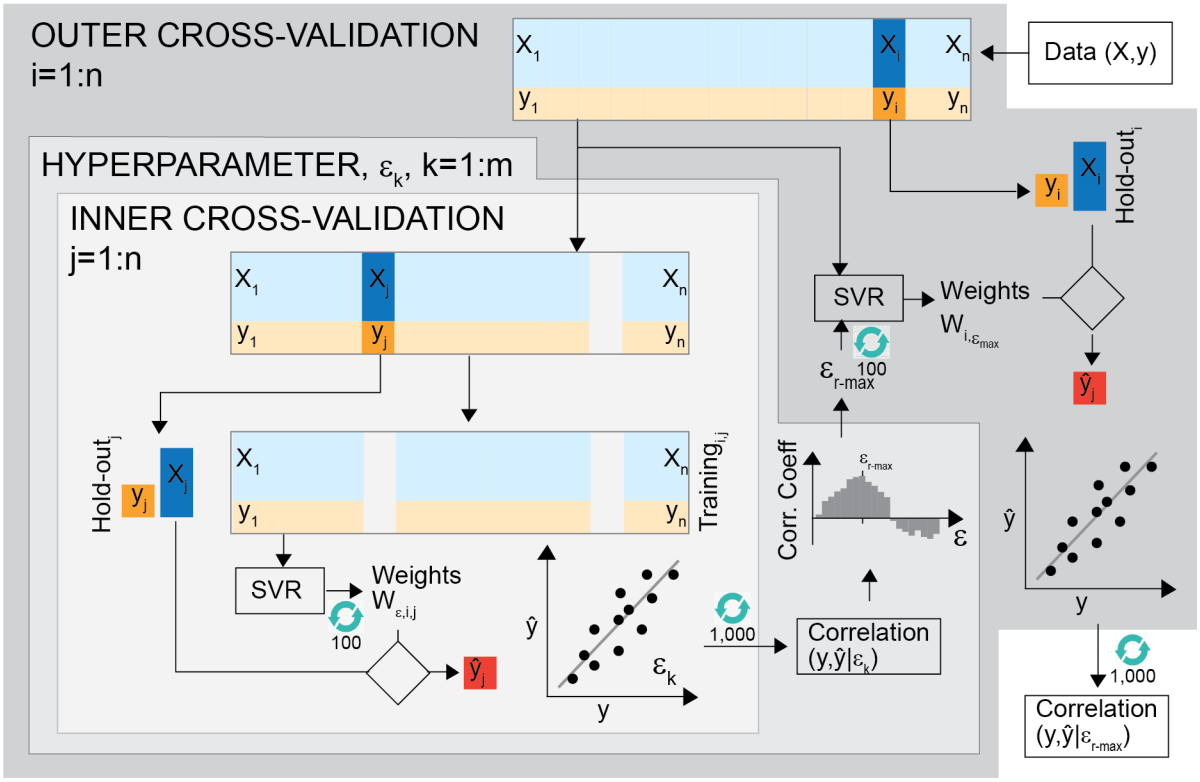
Notably, the test-retest reliably of any measured quantity is finite and there is nothing one can do about this. The "true correlation" estimates may be helpful when trying to infer from the correlation of the actual measures to the general question of high tightly electrophysiology and clinical severity are related.

Characteristics and Limitations of the ML-based analyses

ML approaches are being increasingly deployed in neuroscience research. While potentially very valuable, there are numerous different approaches deployed in various different contexts, which are subject to potential pitfalls. This section aims at briefly classifying the context of our approach, discussing how it relates to best practices (14), and pointing out limitations:

- We use the ML approach to investigate if spatio-spectral information can predict clinical scales better than scalp average delta-band EEG power (see Discussion section in the main paper). This is different from other use cases, such as aiming to create a biomarker that predicts a clinical outcome.
- We used a regression approach (as opposed to classification) to predict the clinical scores. The metric we are interested in is the correlation coefficient between the predicted clinical score and the actual clinical score, since this can be compared to the correlation found for EEG delta-band power (univariate analysis).
- We used a nested cross-validation approach, where all operations are performed on the test data and the model is then applied to the hold-out data (see Supplementary Figure S1). This approach is the gold-standard and should avoid over-fitting.
- It has been pointed out that k-fold cross-validation should be used instead of leave-one-out cross-validation (14). Notably, this is in the context of using R^2 as metric, in which case large spurious negative correlations can be mistaken as predictive. Here we used the signed correlation coefficient, and therefore do not suffer from that risk.
- A clear limitation of our study is the relatively low number of participants from the ML perspective (n=45). Additional studies with more data are needed to confirm our results.

Supplementary Figures



Supplementary Figure S1. Illustration of ML learning approach for the multivariate cross-sectional correlation between EEG and clinical scales. We employed a nested cross-validation approach with an inner cross-validation to determine the hyperparameter ε and an outer cross-validation loop to generate a prediction from data that were not used to generate the model. Note, both EEG data (X, spatio-spectral power) and clinical scales were corrected for age (regression of log-age) before being subjected to the ML procedure (see Methods). Nomenclature: i,j: Participant indices; n: Number of participants; X: EEG feature vector; y: Clinical scale value; ŷ: Prediction of y from X; ε_k.. k^{th} hyperparameter from set; ε_{max}: Selected hyperparameter; SVR: Support vector regression; W: Weights derived from SVR; Green circulating arrows indicate 100 and 1,000 resamples where each time only a single trial from each participant was used. The results are then averaged across resamples.

Supplementary Tables

Clinical	Cross-sectional		Longitudinal	
	n_{Subj}	n_{Visit}	n_{Subj}	n_{Visit}
Bayley-3 Mean	45	72	16	43
Bayley-3 Cognitive	45	72	16	43
Bayley-3 Expressive Comm.	45	72	16	43
Bayley-3 Receptive Comm.	45	72	16	43
Bayley-3 Fine Motor	45	72	16	43
Bayley-3 Gross Motor	45	72	16	43
CSS	43	68	n/a	n/a
Age Onset Epilepsy	42	42	n/a	n/a
VABS-II Soc. Coping Skills	45	72	16	43
VABS-II Soc. Interp. Relation	45	72	16	43
VABS-II Soc. Play & Leisure	45	72	16	43
VABS-II Comm. Expressive	45	72	16	43
VABS-II Comm. Receptive	45	72	16	43
VABS-II DLS Personal	45	72	16	43
VABS-II Motor Fine	45	72	16	43
VABS-II Motor Gross	45	72	16	43

Supplementary Table S1. Overview of individuals and visits used for the different analyses.

Dataset	Mean	SD	Min	Max
Cross-sectional, Bayley-3, VABS-II	58.58	40.1	13	168
Cross-sectional, CSS	57.42	37.91	13	168
Cross-sectional, age onset epilepsy	61.12	40.31	13	168
Longitudinal	50.81	34.26	15	168

Supplementary Table S2. Age in months broken down by analyses.

Dataset	Female	Male
Cross-sectional, Bayley-3, VABS-II	15	30
Cross-sectional, CSS	15	28
Cross-sectional, age onset epilepsy	13	29
Longitudinal	3	13

Supplementary Table S3. Sex broken down by analyses.

Clinical	Cross-sectional p	Longitudinal p
Bayley-3 Mean	0.811	0.811
Bayley-3 Cognitive	0.548	0.548
Bayley-3 Expressive Comm.	0.646	0.646
Bayley-3 Receptive Comm.	0.177	0.177
Bayley-3 Fine Motor	0.519	0.519
Bayley-3 Gross Motor	0.237	0.237
CSS	0.162	-
Age Onset Epilepsy	0.265	-
VABS-II Soc. Coping Skills	0.644	0.644
VABS-II Soc. Interp. Relation	0.990	0.990
VABS-II Soc. Play & Leisure	0.676	0.676
VABS-II Comm. Expressive	0.282	0.282
VABS-II Comm. Receptive	0.429	0.429
VABS-II DLS Personal	0.564	0.564
VABS-II Motor Fine	0.292	0.292
VABS-II Motor Gross	0.208	0.208

Supplementary Table S4. Dependence of clinical scales and age of epilepsy onset on sex. We performed as sensitivity analysis testing for effect of sex. To this end, we fitted two linear mixed models, one with sex as a categorical variable and one without (Model with sex as categorical variable: $Y \sim 1 + SEX + (1|SUBJ)$; Reference model: $Y \sim 1 + (1|SUBJ)$; with Y being a clinical scale, age of epilepsy onset, or EEG delta-band power), to the clinical data after removal of age effects, and compared them using log-likelihood ratio tests. The number of individuals and visits for these analyses are summarized in Supplementary Table S2. The results of these tests, for both the cross-sectional and the longitudinal datasets, are reported in this table. The analysis showed no sign of sex-dependence of the clinical scale. Similarly, the EEG was not dependent on age for all data used in the analyses ($p > 0.492$). The sensitivity analyses support not adjusting for sex in the analyses reported in the main text.

Parameter	Mean	SD	Min	Max
Interpolated electrodes	0.44	0.71	0	2
Rejected ICs	4.54	2.4	0	9
Artifact sections (%)	45.82	21.02	7.19	90.38
Analyzed signal (sec)	1,044.55	534.27	172.98	2,226.47

Supplementary Table S5. Details on EEG pre-processing.

Clinical	ρ	ρ_{LB}	ρ_{UB}	p
Bayley-3 Mean	**-0.34**	-0.55	-0.09	0.0133
Bayley-3 Cognitive	**-0.32**	-0.54	-0.08	0.0169
Bayley-3 Expressive Comm.	*-0.23*	-0.46	0.03	0.0714
Bayley-3 Receptive Comm.	**-0.26**	-0.48	0.00	0.0479
Bayley-3 Fine Motor	**-0.26**	-0.48	0.00	0.0481
Bayley-3 Gross Motor	**-0.34**	-0.55	-0.09	0.0124
-CSS	**-0.39**	-0.59	-0.14	0.0060
Age Onset Epilepsy	**-0.33**	-0.55	-0.07	0.0193
VABS-II Soc. Coping Skills	-0.01	-0.27	0.25	0.4737
VABS-II Soc. Interp. Relation	-0.04	-0.29	0.22	0.4108
VABS-II Soc. Play & Leisure	-0.07	-0.32	0.19	0.3402
VABS-II Comm. Expressive	*-0.24*	-0.47	0.02	0.0618
VABS-II Comm. Receptive	*-0.23*	-0.46	0.03	0.0693
VABS-II DLS Personal	**-0.34**	-0.55	-0.09	0.0127
VABS-II Motor Fine	**-0.30**	-0.51	-0.05	0.0268
VABS-II Motor Gross	*-0.24*	-0.46	0.02	0.0641

Supplementary Table S6. Cross-sectional non-parametric correlation between EEG delta-band power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset. This table reports Spearman rank correlation coefficients (see Table 1 for Pearson correlation coefficients). Significant correlations are highlighted in bold font ($p < 0.05$, uncorrected); trends are highlighted in italics ($p < 0.1$, uncorrected). ρ_{LB} and ρ_{UB} are lower and upper bound of the 90% confidence interval, respectively. P-values are uncorrected for multiple testing (see Methods). The number of individuals (degrees of freedom) can be found in Supplementary Table S2. For abbreviations see Figure 1.

Clinical	r	r_{LB}	r_{UB}	p	ρ	$ρ_{LB}$	$ρ_{UB}$	$p_ρ$	n Subj	n Visit
Bayley-3 Mean	-0.22	-0.44	0.03	0.0758	-0.19	-0.43	0.07	0.1106	45	72
Bayley-3 Cognitive	**-0.25**	-0.47	-0.01	0.0467	-0.23	*-0.46*	0.03	0.0720	45	72
Bayley-3 Expr. Comm.	-0.05	-0.29	0.20	0.3724	-0.04	-0.29	0.22	0.4110	45	72
Bayley-3 Receptive Comm.	*-0.21*	-0.44	0.04	0.0841	-0.22	*-0.45*	0.04	0.0845	45	72
Bayley-3 Fine Motor	-0.11	-0.35	0.14	0.2319	-0.09	-0.34	0.17	0.2830	45	72
Bayley-3 Gross Motor	*-0.24*	-0.46	0.01	0.0540	-0.26	**-0.49**	-0.01	0.0448	45	72
-CSS	-0.17	-0.41	0.09	0.1365	-0.20	-0.44	0.06	0.1035	43	68
Age Onset Epilepsy	-0.17	-0.41	0.09	0.1384	-0.25	*-0.48*	0.02	0.0616	42	42
VABS-II Soc. Coping Skills	-0.10	-0.34	0.15	0.2594	-0.08	-0.33	0.18	0.3129	45	72
VABS-II Soc. Interp. Relation	0.02	-0.23	0.27	0.5629	0.00	-0.26	0.25	0.4890	45	72
VABS-II Soc. Play & Leisure	-0.13	-0.37	0.12	0.1896	-0.11	-0.35	0.15	0.2514	45	72
VABS-II Comm. Expressive	-0.17	-0.40	0.09	0.1376	-0.19	-0.42	0.07	0.1191	45	72
VABS-II Comm. Receptive	-0.09	-0.33	0.16	0.2716	-0.12	-0.36	0.14	0.2236	45	72
VABS-II DLS Personal	*-0.23*	-0.46	0.02	0.0616	-0.22	*-0.45*	0.04	0.0790	45	72
VABS-II Motor Fine	**-0.31**	-0.52	-0.07	0.0174	-0.30	**-0.51**	-0.04	0.0270	45	72
VABS-II Motor Gross	**-0.27**	-0.49	-0.03	0.0339	-0.32	**-0.53**	-0.07	0.0184	45	72

Supplementary Table S7. Cross-sectional nonparametric correlation between relative EEG delta-band power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset. In the literature "relative delta power", i.e. the fraction of delta power of the total power, has been suggested as an alternative metric. This table reports Pearson correlation coefficients (r, r_{LB}, r_{UB}, p) and Spearman rank correlation coefficients (ρ, $ρ_{LB}$, $ρ_{UB}$, $p_ρ$) for relative EEG delta power (see Table 1 and Supplementary Table S1 for analyses with absolute EEG delta power) . Significant correlations are highlighted in bold font ($p < 0.05$, uncorrected); trends are highlighted in italics ($p < 0.1$, uncorrected). r_{LB}/$ρ_{LB}$ and r_{UB}/$ρ_{UB}$ are lower and upper bound of the 90% confidence interval, respectively. P-values are uncorrected for multiple testing (see Methods). n_{Subj} and n_{Data} report the number of participants and visits that were used for a respective analysis. For abbreviations see Figure 1. These results show that relative delta power also correlates negative with most clinical scores (average: $r = -0.17 \pm 0.09$, mean ± SD) but associations were weaker compared to total power (3/16 significant, $p < 0.05$; 4/16 showed a trend, $p < 0.1$; Figure 1A-C, Supplementary Table S1). This is in line with relative delta power also providing less separation to controls compared to absolute delta power (1).

	All	"Younger half"			"Older half"			Difference	
Clinical	**r**	**r**	**p**	**n_{Subj}**	**r**	**p**	**n_{Subj}**	**r**	**p**
Bayley-3 Mean	**-0.38**	**-0.43**	0.0082	30	**-0.42**	0.0182	25	-0.01	0.9661
Bayley-3 Cognitive	**-0.36**	**-0.42**	0.0095	30	**-0.45**	0.0111	25	0.03	0.8975
Bayley-3 Expressive Com.	**-0.24**	*-0.28*	0.0712	30	*-0.19*	0.1807	25	-0.08	0.7399
Bayley-3 Receptive Com.	**-0.28**	-0.20	0.1480	30	**-0.42**	0.0174	25	0.22	0.3937
Bayley-3 Fine Motor	**-0.30**	**-0.42**	0.0093	30	-0.22	0.1454	25	-0.20	0.4354
Bayley-3 Gross Motor	**-0.37**	**-0.41**	0.0112	30	**-0.44**	0.0137	25	0.03	0.8985
-CSS	**-0.32**	**-0.41**	0.0127	29	-0.24	0.1306	24	-0.17	0.5154
Age Onset Epilepsy	**-0.28**	-0.20	0.1657	27	-0.14	0.2582	24	-0.06	0.8361
VABS-II Soc. Coping Skills	**0.06**	0.01	0.5230	30	0.02	0.5362	25	-0.01	0.9722
VABS-II Soc. Interp. Relation	**-0.03**	-0.06	0.3799	30	-0.01	0.4880	25	-0.05	0.8616
VABS-II Soc. Play & Leisure	**-0.11**	-0.02	0.4633	30	-0.17	0.2069	25	0.15	0.5975
VABS-II Com. Expressive	**-0.22**	-0.24	0.1045	30	-0.14	0.2527	25	-0.10	0.7177
VABS-II Com. Receptive	**-0.21**	*-0.29*	0.0624	30	-0.20	0.1749	25	-0.09	0.7386
VABS-II DLS Personal	**-0.31**	**-0.40**	0.0147	30	**-0.34**	0.0482	25	-0.06	0.8086
VABS-II Motor Fine	**-0.33**	**-0.51**	0.0017	30	*-0.29*	0.0812	25	-0.22	0.3577
VABS-II Motor Gross	**-0.25**	**-0.41**	0.0116	30	-0.12	0.2842	25	-0.29	0.2727

Supplementary Table S8. Age sensitivity analysis for cross-sectional correlation between EEG delta-band power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset. This table reports Pearson correlation coefficients separately for younger and older individuals (half-split) as well as a comparison between those correlation coefficients. The age split was at 58 months, the age for the "younger half" was 36.2 ± 13.8 months (mean ± SD), the age for the "older half" was 98.8 ± 31.9 months (mean ± SD). Significant correlations (Null-hypothesis no correlation, one-tailed and equal correlation values, respectively) are highlighted in bold font ($p < 0.05$, uncorrected); trends are highlighted in italics ($p < 0.1$, uncorrected). P-values are uncorrected for multiple testing (see Methods). For abbreviations see Figure 1. The results show a similar pattern of correlation values across clinical scales for the "younger" and "older half" of the data (correlation between z-transformed correlation values: $r = 0.6264$; $p = 0.0094$). Overall the magnitude of correlation values was higher for the correlations derived from the "younger" compared to the "older half" but differences were not significant. In sum, this analysis suggest that the correlations between EEG delta-band power and clinical scales exist across the broad age range studied.

Clinical	r_{ML}	n_{Subj}	n_{Visit}	$\varepsilon_{r\text{-}max}$
Bayley-3 Mean	0.46	45	72	$2.1 \cdot 10^{-5}$
Bayley-3 Cognitive	0.47	45	72	$1.9 \cdot 10^{-5}$
Bayley-3 Expressive Com.	0.28	45	72	$2.2 \cdot 10^{-5}$
Bayley-3 Receptive Com.	0.14	45	72	$3.5 \cdot 10^{-5}$
Bayley-3 Fine Motor	0.35	45	72	$2.8 \cdot 10^{-5}$
Bayley-3 Gross Motor	0.63	45	72	$3.3 \cdot 10^{-5}$
-CSS	0.30	43	68	$3.3 \cdot 10^{-6}$
Age Onset Epilepsy	0.37	42	n/a	$2.0 \cdot 10^{-1}$
VABS-II Soc. Coping Skills	0.32	45	72	$1.9 \cdot 10^{-6}$
VABS-II Soc. Interp. Relation	0.25	45	72	$3.0 \cdot 10^{-6}$
VABS-II Soc. Play & Leisure	-0.04	45	72	$4.5 \cdot 10^{-6}$
VABS-II Com. Expressive	0.43	45	72	$3.8 \cdot 10^{-6}$
VABS-II Com. Receptive	0.49	45	72	$4.2 \cdot 10^{-6}$
VABS-II DLS Personal	0.32	45	72	$4.9 \cdot 10^{-6}$
VABS-II Motor Fine	0.37	45	72	3.7
VABS-II Motor Gross	0.49	45	72	$8.5 \cdot 10^{-6}$

Supplementary Table S9. ML: Multivariate cross-sectional correlation between EEG spatio-spectral power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset – additional information. This table reports additional information on the multivariate prediction of the clinical scores using ML (see Methods) and the clinical scores (r_{ML}) (see Table 2 for main results). n_{Subj} and n_{Data} report the number of participants and visits that were used for a respective analysis. $\varepsilon_{r\text{-}max}$ is the hyperparameter identified in the inner cross-validation loop, see Methods and Supplementary Figure S1. For abbreviations see Figure 1.

Clinical	$-\rho_\delta$	ρ_{ML}	ρ_{LB}	ρ_{UB}	p	Inc. δ (%)
Bayley-3 Mean	**0.34**	**0.48**	0.26	0.66	0.0004	43
Bayley-3 Cognitive	**0.32**	**0.44**	0.21	0.62	0.0015	35
Bayley-3 Expressive Com.	*0.23*	**0.33**	0.08	0.54	0.0154	44
Bayley-3 Receptive Com.	**0.26**	0.10	-0.16	0.34	0.2680	-
Bayley-3 Fine Motor	**0.26**	**0.32**	0.07	0.53	0.0174	25
Bayley-3 Gross Motor	**0.34**	**0.57**	0.37	0.72	0.0000	68
-CSS	**0.39**	**0.29**	0.03	0.51	0.0345	-26
Age Onset Epilepsy	**0.33**	**0.36**	0.10	0.57	0.0120	8
VABS-II Soc. Coping Skills	0.01	**0.28**	0.02	0.50	0.0372	-
VABS-II Soc. Interp. Relation	0.04	*0.22*	-0.04	0.45	0.0798	-
VABS-II Soc. Play & Leisure	0.07	-0.01	-0.26	0.25	0.5162	-
VABS-II Com. Expressive	*0.24*	**0.39**	0.15	0.59	0.0050	62
VABS-II Com. Receptive	*0.23*	**0.46**	0.23	0.64	0.0009	98
VABS-II DLS Personal	**0.34**	**0.27**	0.02	0.49	0.0392	-20
VABS-II Motor Fine	**0.30**	**0.36**	0.11	0.56	0.0089	21
VABS-II Motor Gross	*0.24*	**0.49**	0.27	0.67	0.0003	108

Supplementary Table S10. Multivariate non-parametric cross-sectional correlation between EEG spatio-spectral power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset. This table reports Spearman rank correlation coefficients (see Table 1 for Pearson correlation coefficients) between the multivariate prediction of the clinical scores using ML (see Methods) and the clinical scores (ρ_{ML}). $-\rho_\delta$ shows the negative of the Spearman rank correlation coefficient for the univariate analysis with EEG delta power that is reported in Supplementary Table S1 to allow easy comparison with ρ_{ML}. Note, for ρ_{ML} correlations are expected to be positive (ML is trained to predict the scores), while for r_δ correlations are expected to be negative (greater EEG abnormality, lower clinical scores). Significant correlations are highlighted in bold font ($p < 0.05$, uncorrected); trends are highlighted in italics ($p < 0.1$, uncorrected). ρ_{LB} and ρ_{UB} are lower and upper bound of the 90% confidence interval for ρ_{ML}, respectively. P-values are uncorrected for multiple testing (see Methods). n_{Subj} and n_{Data} report the number of participants and visits that were used for a respective analysis. Inc. δ (%) reports the increase in magnitude of the rank correlation coefficient ρ_{ML} compared to $-\rho_\delta$ for those where both correlation coefficients were significant or showed a trend ($p < 0.1$, uncorrected; increases for non-significant correlations would not be well-defined mathematically and not meaningful). Details on the number of individuals and the ML hyperparameter ε can be found in Supplementary Table S9. For abbreviations see Figure 1.

Clinical	ρ	ρ_{LB}	ρ_{UB}	p	n_{Subj}	n_{Data}
Bayley-3 Mean	**-0.60**	-0.80	-0.15	0.0154	16	43
Bayley-3 Cognitive	**-0.52**	-0.77	-0.08	0.0278	16	43
Bayley-3 Expressive Com.	-0.22	-0.58	0.27	0.2439	16	43
Bayley-3 Receptive Com.	*-0.42*	-0.70	0.08	0.0838	16	43
Bayley-3 Fine Motor	*-0.40*	-0.72	0.02	0.0591	16	43
Bayley-3 Gross Motor	-0.41	-0.67	0.13	0.1198	16	43
VABS-II Soc. Coping Skills	**-0.50**	-0.79	-0.13	0.0173	16	43
VABS-II Soc. Interp. Relation	-0.02	-0.50	0.37	0.3918	16	43
VABS-II Soc. Play & Leisure	-0.16	-0.61	0.22	0.1984	16	43
VABS-II Com. Expressive	0.11	-0.49	0.39	0.4139	16	43
VABS-II Com. Receptive	-0.12	-0.59	0.26	0.2387	16	43
VABS-II DLS Personal	**-0.56**	-0.77	-0.07	0.0295	16	43
VABS-II Motor Fine	-0.26	-0.64	0.19	0.1631	16	43
VABS-II Motor Gross	0.03	-0.35	0.52	0.6374	16	43

Supplementary Table S11. Longitudinal nonparametric correlations between EEG delta-band power and clinical severity as measured with Bayley-3, VABS-II, CSS and the age of epilepsy onset. This table reports Spearman rank correlation coefficients (see Table 3 for Pearson correlation coefficients). Significant correlations are highlighted in bold font ($p < 0.05$, uncorrected); trends are highlighted in italics ($p < 0.1$, uncorrected). r_{LB} and r_{UB} are lower and upper bound of the 90% confidence interval, respectively. P-values are uncorrected for multiple testing (see Methods). n_{Subj} and n_{Data} report the number of participants and visits that were used for a respective analysis. For abbreviations see Figure 1.

Supplementary References

1. Frohlich J, Miller M, Bird LM, Garces P, Purtell H, Hoener MC, *et al.* (2019): Electrophysiological phenotype in Angelman syndrome differs between genotypes. *Biol Psychiatry*. https://doi.org/10.1016/j.biopsych.2019.01.008
2. Frohlich J, Bird LM, Dell'Italia J, Johnson MA, Hipp JF, Monti MM (2020): High-voltage, diffuse delta rhythms coincide with wakeful consciousness and complexity in Angelman syndrome. *Neurosci Conscious* 2020. https://doi.org/10.1093/nc/niaa005
3. Gentile JK, Tan W-H, Horowitz LT, Bacino CA, Skinner SA, Barbieri-Welge R, *et al.* (2010): A neurodevelopmental survey of Angelman syndrome with genotype-phenotype correlations. *J Dev Behav Pediatr JDBP* 31: 592.
4. Keute M, Miller MT, Krishnan ML, Sadhwani A, Chamberlain S, Thibert RL, *et al.* (2020): Angelman syndrome genotypes manifest varying degrees of clinical severity and developmental impairment. *Mol Psychiatry* 1–9.
5. Sidorov MS, Deck GM, Dolatshahi M, Thibert RL, Bird LM, Chu CJ, Philpot BD (2017): Delta rhythmicity is a reliable EEG biomarker in Angelman syndrome: a parallel mouse and human analysis. *J Neurodev Disord* 9: 17.
6. Vendrame M, Loddenkemper T, Zarowski M, Gregas M, Shuhaiber H, Sarco DP, *et al.* (2012): Analysis of EEG patterns and genotypes in patients with Angelman syndrome. *Epilepsy Behav* 23: 261–265.
7. Clayton-Smith J, Laan L (2003): Angelman syndrome: a review of the clinical and genetic aspects. *J Med Genet* 40: 87–95.
8. Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1997): Oscillatory γ-band (30–70 Hz) activity induced by a visual search task in humans. *J Neurosci* 17: 722–734.
9. Bayley N (2006): *Bayley Scales of Infant and Toddler Development. San Antonio, TX: The Psychological Corporation*. Harcourt Assessment.
10. Farmer CA, Kaat AJ, Thurm A, Anselm I, Akshoomoff N, Bennett A, *et al.* (2020): Person Ability Scores as an Alternative to Norm-Referenced Scores as Outcome Measures in Studies of Neurodevelopmental Disorders. *Am J Intellect Dev Disabil* 125: 475–480.
11. Sparrow SS, Cicchetti DV, Balla DA (2005): Vineland adaptive behavior scales:(Vineland II), survey interview form/caregiver rating form. *Livonia MN Pearson Assess*.

12. Bergholm F, Adler J, Parmryd I (2010): Analysis of Bias in the Apparent Correlation Coefficient Between Image Pairs Corrupted by Severe Noise. *J Math Imaging Vis* 37: 204–219.
13. Spearman C (1904): The Proof and Measurement of Association between Two Things. *Am J Psychol* 15: 72–101.
14. Poldrack RA, Huckins G, Varoquaux G (2020): Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psychiatry* 77: 534–540.