Artificial Intelligence Identification of Autism

Using a Smart Tablet Serious Game for Preschool Children

Results from a Phase 3 diagnostic trial of 779 Children in Sweden and the United Kingdom

Jonathan Delafield-Butt[1,2], Szu-Ching Lu[1,2], Krzysztof Sobota[3], Robin Young[4], Lucy Thompson[5,6], Max Thorsson[6], Christos Tachtatzis[1,7], Ivan Andonovic[1,7], Philip Rowe[1,8], Phil Wilson[5], Helen Minnis[9], Alex McConnachie[4], Christopher Gillberg[6]

(1) Laboratory for Innovation in Autism, University of Strathclyde; (2) School of Education, University of Strathclyde; (3) Department of Psychology, University of Warsaw; (4) Robertson Biostatistics Centre, University of Glasgow; (5) Institute of Applied Health Sciences, University of Aberdeen; (6) Gillberg Neuropsychiatry Centre, University of Gothenburg; (7) Electronic and Electrical Engineering, University of Strathclyde; (8) Department of Biomedical Engineering, University of Strathclyde; (9) Institute for Health and Wellbeing, University of Glasgow

1. Introduction

Early detection of autism spectrum disorder (ASD) allows early intervention and potentially the best lifelong health outcomes (1). However, ASD's complex symptomatology makes diagnosis complex, time-consuming and requires specialist clinical input. Waiting times for diagnosis can be many months or years. Recent evidence suggests the motor system is linked with autism aetiology (2,3), providing an accessible modality for computational assessment (4). Further, most children with or without autism are attracted to smart tablet gameplay. Its touch screen and inertial sensors enable collection of reliable motor kinematic and behavioural data, suggesting a promising new route for accessible, scalable early assessment. This study set out to test promising pilot results of an iPad serious game assessment paradigm (4) with a gold-standard blinded, multi-site phase 3 diagnostic trial (5).

2. Objectives

To determine the predictive accuracy of a serious smart tablet game for the early identification of autism using pre-trained algorithms naïve to trial data collected in two sites using blinded comparison against clinical diagnosis.

3. Methods

Two serious games (Figure 1) running on iPad mini tablets (Apple Inc.) set within a bespoke app to organise the display of the games sequentially for a 2-minute training phase followed by a single 5-minute test phase with code for collecting inertial sensor and touch screen data (Play.Care, Harimata) was employed. Previous machine learning analysis demonstrated 93% classification accuracy based on motor kinematic features (4).

A registered phase 3 prospective, diagnostic classification study (NCT03438994; Full protocol: Ref. 5) tested the predictive accuracy of a smart tablet serious game with artificial intelligence data analytics to identify autism. Three cohorts aged 3-6 years participated: children typically developing (TD); children with a clinical diagnosis of autism (ASD); and children with diagnoses of other non-autism neurodevelopmental disorders (OND). 779 children were recruited from Scotland (Glasgow) and Sweden (Gothenburg). Children played two 5-minute games on an iPad. One commercial algorithm and four research algorithms were trained on a *previous* cohort of children collected prior to this trial (*n*=767).

Algorithms were tested naïvely on new, blinded trial data to classify gameplay patterns as positively or negatively associated with an ASD diagnosis. Classification results were then compared against medical diagnosis by a clinical trial unit. Social-Emotional Questionnaire (SEQ) and adaptive function scores were collected for a subset of participants. Sensitivity and specificity of the algorithms to differentiate ASD children from TD children are reported.

Contact details

* For more information contact,
jonathan.delafield-butt@strath.ac.uk

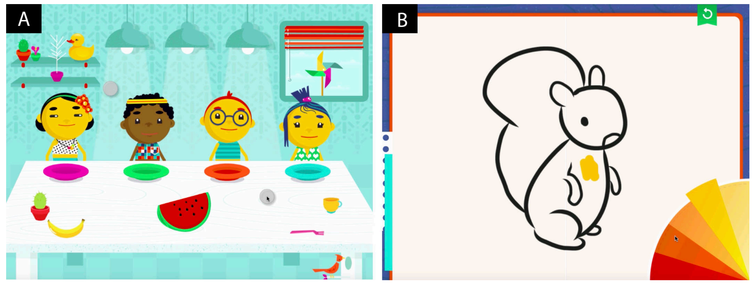

Figure 1. Two serious games were employed. (A) Sharing consisted of dividing a piece of food and distributing it evenly among four cartoon children present on the screen. When the food was distributed, all children exclaimed, "Yipee!" and proceeded to munch the food in a delightful manner for 3 seconds. Then, the trial repeated. **(B) Creativity** was a colouring game with no rules of engagement. An object outline appeared for tracing, then a colouring wheel appeared and the child could select a colour. The toy or animal outline always remained unobstructed.

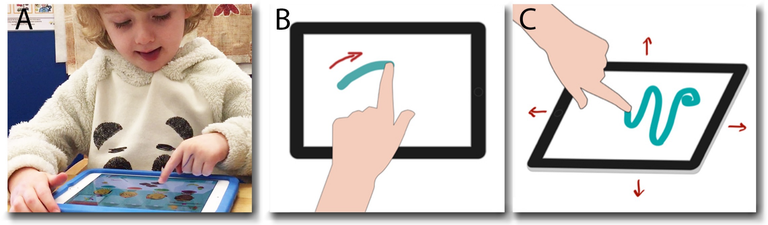

Figure 2. Computational gameplay data were collected during (A) children's gameplay at a table that included two main sources: (B) gesture kinematics from the touch screen and (B) impact and and pressure from the iPad's inertial sensors.

4. Results

Table 1. Participant recruitment was carried out in Sweden and Scotland. 694 participants (331 TD, 185 ASD, and 178 OND) were included in the final analysis, derived from a total recruitment of 779 participants; 85 were excluded due to incomplete data transfer or backup.

Variable	Statistics	All participants (*n*=694)	TD group (*n*=331)	ASD group (*n*=185)	OND group (*n*=178)
Age (months)	n_{obs} (n_{miss})	694 (0)	331 (0)	185 (0)	178 (0)
	Mean (SD)	52.43 (10.59)	51.69 (11.07)	53.54 (9.51)	52.63 (10.67)
	Median (IQR)	53.0 [44.00, 61.00] (30.00, 72.00)	53.00 [43.00, 61.00] (30.00, 71.00)	54.00 [45.75, 61.00] (32.00, 72.00)	52.50 [45.00, 61.00] (30.00, 72.00)
	Range				
Sex	n_{obs} (n_{miss})	694 (0)	331 (0)	185 (0)	178 (0)
Female	*n* (%)	275 (39.63%)	162 (48.94%)	42 (22.70%)	71 (39.89%)
Male	*n* (%)	419 (60.37%0	169 (51.06%)	143 (77.30%)	107 (60.11%)
Severity level of ASD	n_{obs} (n_{miss})	n/a	n/a	171 (14)	n/a
Level 1: Requiring support	*n* (%)			70 (40.94%)	
Level 2: Requiring substantial support	*n* (%)			84 (49.12%)	
Level 3: Requiring very substantial support	*n* (%)			17 (9.94%)	

Table 2. Algorithm performance was calculated on **(i)** the iPad data alone and **(ii)** iPad data plus additional data from a small multiple-choice questionnaire on the social and emotional aspects of the child (25 questions) completed by the parent.

Index test	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC
Commercial algorithm	0.287 (0.22, 0.36)	0.468 (0.41, 0.52)	0.005	0.985	0.403	0.670
Full-feature ML	0.605 (0.53, 0.68)	0.840 (0.80, 0.88)	0.037	0.995	0.756	0.811
Reduced-feature ML	0.697 (0.63, 0.76)	0.776 (0.73, 0.82)	0.031	0.996	0.748	0.817
Kinematic-feature ML	0.751 (0.68, 0.81)	0.713 (0.66, 0.76)	0.026	0.997	0.727	0.776
CNN	0.800 (0.74, 0.86)	0.595 (0.54, 0.65)	0.020	0.997	0.669	0.763
SEQ	0.625 (0.52, 0.72)	1.000 (0.98, 1.00)	1.000	0.996	0.848	0.963
SEQ+ full-feature ML	0.904 (0.83, 0.95)	0.934 (0.88, 0.97)	0.122	0.999	0.922	0.977
SEQ+ reduced-feature ML	0.904 (0.83, 0.95)	0.947 (0.90, 0.98)	0.148	0.999	0.930	0.979
SEQ+ kinematic-feature ML	0.952 (0.89, 0.98)	0.882 (0.82, 0.93)	0.075	0.999	0.910	0.973
SEQ+ CNN	0.942 (0.88, 0.98)	0.901 (0.84, 0.94)	0.088	0.999	0.918	0.968

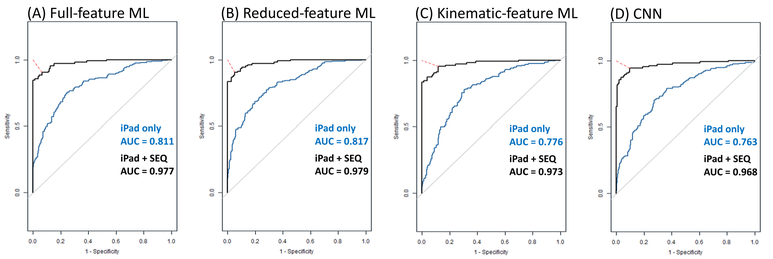

Figure 3. Area Under the Curve measures of diagnostic accuracy were computed for 24 algorithmic models. Four models are presented here (blue curves) represented four classes of algorithm type: (A) Full features algorithms based on the original model features extraction; (B) a reduced features model based on the previous; (C) a model derived solely from motor kinematic features; and (D) a model derived from convolution neural networks.

5. Conclusions

A world-first phase 3 diagnostic accuracy study of a digital health smart tablet assessment of autism.

Clinically useful sensitivities and specificities for screening or diagnostic pathways.

Computational analysis of motor patterns indicates strong (>70%), but not total predictive value for early identification of autism.

Best performance is achieved in combination with a brief, supplementary questionnaire on socio-emotional health.

Future work is required to

(i) improve the algorithms further (AI development), including for OND differentiation.
(ii) develop this into a clinic- or school-ready tool (commercialisation or not-for-profit).
(iii) integrate with screening, assessment or diagnostic pathways (translation to practice).

References

1) Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., Gammer, I., Lowry, J., Vamvakas, G., Byford, S., Aldred, C., Slonims, V., McConachie, H., Howlin, P., Parr, J. R., Charman, T., & Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. The Lancet, 388(10059), 2501-2509. https://doi.org/10.1016/S0140-6736(16)31229-6
2) Trevarthen, C., & Delafield-Butt, J. T. (2013). Autism as a developmental disorder in intentional movement and affective engagement. Frontiers in Integrative Neuroscience, 7, 49. https://doi.org/10.3389/fnint.2013.00049
3) Bhat, A. N. (2020). Is Motor Impairment in Autism Spectrum Disorder Distinct From Developmental Coordination Disorder? A Report From the SPARK Study. Physical Therapy, 100(4), 633-644. https://doi.org/10.1093/ptj/pzz190
4) Anzulewicz, A., Sobota, K., & Delafield-Butt, J. T. (2016). Toward the autism motor signature: Gesture patterns during smart tablet gameplay identify children with autism. Scientific Reports, 6. https://doi.org/10.1038/srep31107
5) Millar, L., McConnachie, A., Minnis, H., Wilson, P., Thompson, L., Anzulewicz, A., Sobota, K., Rowe, P., Gillberg, C., & Delafield-Butt, J. (2019). Phase 3 diagnostic evaluation of a smart tablet serious game to identify autism in 760 children 3–5 years old in Sweden and the United Kingdom. BMJ Open, 9(7), e026226. https://doi.org/10.1136/bmjopen-2018-026226