

BMJ Open 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

Lindsay Millar,^{1,2} Alex McConnachie,³ Helen Minnis,⁴ Philip Wilson,⁵ Lucy Thompson,^{5,6} Anna Anzulewicz,⁷ Krzysztof Sobota,⁷ Philip Rowe,^{1,2} Christopher Gillberg,^{4,6} Jonathan Delafield-Butt¹

To cite: Millar L, McConnachie A, Minnis H, *et al.* 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* 2019;**9**:e026226. doi:10.1136/bmjopen-2018-026226

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-026226>).

Received 6 September 2018
Revised 26 March 2019
Accepted 1 May 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Jonathan Delafield-Butt;
jonathan.delafield-butt@strath.ac.uk

ABSTRACT

Introduction Recent evidence suggests an underlying movement disruption may be a core component of autism spectrum disorder (ASD) and a new, accessible early biomarker. Mobile smart technologies such as iPads contain inertial movement and touch screen sensors capable of recording subsecond movement patterns during gameplay. A previous pilot study employed machine learning analysis of motor patterns recorded from children 3–5 years old. It identified those with ASD from age-matched and gender-matched controls with 93% accuracy, presenting an attractive assessment method suitable for use in the home, clinic or classroom.

Methods and analysis This is a phase III prospective, diagnostic classification study designed according to the Standards for Reporting Diagnostic Accuracy Studies guidelines. Three cohorts are investigated: children typically developing (TD); children with a clinical diagnosis of ASD and children with a diagnosis of another neurodevelopmental disorder (OND) that is not ASD. The study will be completed in Glasgow, UK and Gothenburg, Sweden. The recruitment target is 760 children (280 TD, 280 ASD and 200 OND). Children play two games on the iPad then a third party data acquisition and analysis algorithm (Play.Care, Harimata) will classify the data as positively or negatively associated with ASD. The results are blind until data collection is complete, when the algorithm's classification will be compared against medical diagnosis. Furthermore, parents of participants in the ASD and OND groups will complete three questionnaires: Strengths and Difficulties Questionnaire; Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations Questionnaire and the Adaptive Behavioural Assessment System-3 or Vineland Adaptive Behavior Scales-II. The primary outcome measure is sensitivity and specificity of Play.Care to differentiate ASD children from TD children. Secondary outcomes measures include the accuracy of Play.Care to differentiate ASD children from OND children.

Ethics and dissemination This study was approved by the West of Scotland Research Ethics Service Committee 3 and the University of Strathclyde Ethics Committee. Results will be disseminated in peer-reviewed publications and at international scientific conferences.

Trial registration number NCT03438994; Pre-results.

Strengths and limitations of this study

- This study tests the accuracy of a new computational serious game assessment for the early identification of autism in preschool children.
- It is the first phase 3 diagnostic study to evaluate an iPad-based serious game assessment for autism spectrum disorder.
- A computational analysis of motor pattern variables, rather than social and emotional variables, is employed.
- Novel data collection methods exploit the touch screen and inertial sensors for accurate, reliable movement measurement of children during gameplay.
- Children with clinical diagnoses of autism are compared with typically developing controls and children with other neurodevelopmental disorders that are not autism.

INTRODUCTION

Autism spectrum disorder (ASD) is a childhood neurodevelopmental disorder with a prevalence estimated as high as 1 in 59 children in the USA.¹ In the UK, ca. 700 000 individuals live with autism² and the aggregate annual cost of healthcare and support is £28 billion.³ Early identification and consequent early therapeutic intervention may afford family and caregivers opportunity to adjust, and can trigger early healthcare intervention and children's services support. Such provision can produce significant, lifelong health and economic benefit.^{4–7}

Early diagnosis of children with autism remains complex and can be difficult to obtain. Diagnosis currently relies on specialist medical expertise with diagnostic instrumentation dependent on subjectively rated scores during child observations, parent interviews and testing. These instruments are time

consuming, clinically demanding and can be poorly validated against population controls. Medical diagnosis can be withheld for many years due to wait-list times or uncertainty in clinical diagnostic fit, and schools and families can struggle not knowing whether or not a child formally requires additional support needs.^{8 9}

Recent identification of motor disturbance in young children who develop ASD presents a new target for early assessment.^{10 11} ASD is typically considered a social and emotional disorder. Therefore, current diagnostic instruments directly address social and emotional aspects of the syndrome. However, motor control underpins social engagement, emotional expression, linguistic and cognitive development,¹²⁻¹⁷ and its precise subsecond form is now accessible to non-invasive, ecologically valid assessment methods from the neonatal stage onward. These methods of motor analysis can detect risk for neurodevelopmental disorder well before current methods allow,¹⁸⁻²¹ but more work is needed to better characterise these motor signatures. In particular, characterisation of an autism-specific motor signature may serve as an accessible, non-invasive biomarker of the disorder for its early identification, but more work is needed to better define its specificity within the broad clinical population.²²

Children with ASD exhibit a motor signature with a disruption in the development of purposeful movement,¹⁰ evident in, for example, delays in the attainment of motor milestones,^{23 24} poor coordination,^{25 26} unusual gait patterns²⁷⁻²⁹ and difficulties in gross and fine motor skills.³⁰ In a meta-analysis of all evidence, Fournier *et al*¹¹ concluded motor disruption to be a core feature of ASD. This motor perspective is beginning to gain significant clinical and research interest³¹ and although movement differences in autism have not yet entered into the diagnostic criteria, the Diagnostic and Statistical Manual of Mental Disorders (DSM-5)³² now includes 'awkward or clumsy gait' as an associated symptom. More work is now required to characterise how these motor differences manifest across the spectrum²² and to identify their aetiology.^{33 34} Motor signs identified in autism require differentiation from those that extend across the wider spectrum of Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examination (ESSENCE).³⁵

Recently, technological developments have miniaturised inertial movement unit (IMU) sensors that now include triaxial accelerometers, gyroscopes and magnetometers, together with touch-sensitive screens integrated into mobile consumer smart devices such as phones, tablets and wearable devices such as wristbands and watches. These new devices provide unprecedented access to motor information with high levels of accuracy and reliability previously restricted to high-end research or clinical laboratories.

At the same time, an emerging field of 'serious games' has established a new paradigm for engaging children with autism in psychological assessment or intervention in fun and playful manners while retaining a 'serious' scientific agenda.³⁶⁻³⁹ Smart tablet serious games are

an especially attractive format, because children with ASD do not readily share the same social intentions as the researcher. On the other hand, children with ASD and typically developing (TD) children alike typically enjoy screen-based games and will engage these freely of their own volition, presenting an attractive and ecologically valid paradigm for assessment.

In previous work, Anzulewicz *et al*¹⁸ tested a novel, serious smart tablet game for early identification of young children with autism. Two attractive educational cartoon games (Duckie Deck, <http://duckiedeck.com>) were engineered with bespoke code to collect sensor data (touch screen and triaxial accelerometer and gyroscope) as the children engaged for 5 min with each game. Machine learning analysis identified those children diagnosed with autism with 93% accuracy, proving the principle that smart game identification of ASD in children was a viable solution for assessment. Furthermore, this methodology required no professional training to deploy, no verbal instructions were required to engage the child, and only limited supervision was necessary. The result of the assessment was purely statistical; no subjective coding was required. Thus, we proposed this serious game digital health instrument could serve as an attractive addition to the diagnostician's tool box, as a paediatric public health screening tool, or for teachers and psychologists in schools.

This game is now developed into a commercial product called Play.Care that is entering its final stages of preparation by Harimata sp. z o.o. (<https://harimata.pl/>) with support from a European Union Horizon 2020 Small to Medium Enterprise Instrument (grant No. 756079) to maximise its development and availability. This diagnostic study is a part of that grant to test with gold-standard methodology the classification accuracy of the serious game assessment, and therefore its value in educational, clinical and population screening contexts, or as a tool in clinical diagnosis.

A current best performance for a validated subcomponent of the gold standard in autism diagnosis, the Autism Diagnostic Observation Schedule (ADOS) V.2, yields diagnostic performance of, at most, 91% sensitivity and 71% specificity, or 82% sensitivity and 88% specificity when the algorithm is adjusted to increase specificity.⁴⁰ In our previous pilot work, the Play.Care iPad serious game assessment demonstrated a sensitivity and specificity of 83% and 85%, respectively.¹⁸ Serious game assessment by smart device appears to be an attractive resource that could improve sensitivity and specificity ratings, while at the same time improving accessibility. Moreover, the device offers a pure statistical computation without the disadvantages of subjectively rated codes.

The aim of this study is to validate the predictive value of the Play.Care for ASD diagnosis by conducting a multisite, controlled diagnostic study. Further aims include assessing the ability of Play.Care to differentiate ASD from other childhood developmental disorders. The

Table 1 Participant numbers across groups and sites

Site	ASD	OND	TD	Total
Glasgow	100	140	180	420
Gothenburg	180	60	100	340
Total	280	200	280	760

ASD, Autism Spectrum Disorder; OND, Other neurodevelopmental disorder; TD, typically developing.

primary outcome measures will be sensitivity and specificity of Play.Care to detect ASD in the study population.

Limitations inherent in the earlier pilot study are overcome in this diagnostic trial. The pilot study was based on a sample of 82 children recruited from a small number of specialist day-care centres in Poland. Such limited subject recruitment is prone to selection bias and site-specific effects and precludes assessment of predictive value in population-based settings. Furthermore, the pilot study employed the same dataset to train and test the machine learning algorithm. It did so with 10 repetitions of a 10-fold cross-validation procedure that is not prone to overfitting. Nevertheless, it only trained and tested patterns on that particular dataset. To alleviate both concerns, this study sets out to test an adaptation of the algorithm previously developed and published¹⁸ on new, blinded data obtained independently of the Play.Care software engineers and commercial concerns in an international, multisite diagnostic evaluation with a general population and clinical cohort and with phase III clinical study structure and size. This trial is designed to test the Play.Care assessment to determine its generalised,

real-world prediction rates across the range of possible ASD expression, as well as against TD and other neurodevelopmental disorder (OND) variability. Should the sensitivity and specificity of the assessment fall within clinically useful levels comparable to or exceeding current observer-rated tests (ADOS-2, Autism Diagnostic Interview-Revised, ADI-R), we expect this instrument to be an attractive and useful addition to the assessment and diagnostic tool box, accessible to a wide range of professionals in psychiatry, paediatrics and specialist children's services, and those in education. This solution is attractive and timely, and benefits from a surge in machine learning analysis of smart tech data. In our case, this is a fun, useful and mobile serious game for autism assessment.

METHODS AND ANALYSIS

Study design

This study is a phase III prospective, diagnostic classification study. The protocol follows the Standards for Reporting Diagnostic Accuracy Studies guidelines.⁴¹ The study is a multicentre, international trial based in Glasgow, Scotland, UK and Gothenburg, Sweden. The Glasgow cohort comprises a clinical sample and the Gothenburg cohort comprises a general population sample. The study has three groups of participants recruited across two sites (table 1): children who have received a clinical diagnosis of ASD; children who have received a clinical diagnosis of an (OND; see table 2 for inclusion/exclusion criteria); and children TD. The recruitment phase of the study runs from 1 January 2018 to 30 June 2019 with data analysis for

Table 2 Participant inclusion and exclusion criteria

Group	Inclusion criteria	Exclusion criteria
All participants	Aged 30 months to 5 years 11 months	
Participants with a diagnosis of ASD	Diagnosis of autism spectrum disorder on the basis of DSM-5 criteria or equivalent ICD-10 framework	Uncorrected sensory (visual, hearing) impairments Presence of any motor impairments or behavioural impairment that may obstruct testing
Participants with a diagnosis of OND	Other childhood developmental disorders including, but not limited to intellectual disability, non-verbal disability, communication disability, attention deficit hyperactivity disorder and developmental coordination disorder, Down's syndrome and cerebral palsy.	Subclinical or secondary expressions of ASD Uncorrected hearing or vision impairments
TD participants	Within age range	Diagnosis or suspicion of ▶ Neuropathology ▶ Psychopathology ▶ Sensory or motor disorder including mild tremors

ASD, Autism Spectrum Disorder; OND, Other neurodevelopmental disorder; TD, typically developing.

primary and secondary outcome measures concluding and reports submitted by 31 October 2019.

Patient and public involvement

The research questions addressed in this study were the result of previous dialogue between research, digital health, commercialisation and clinical, educational, and parental need. Demand for early, accessible and computationally accurate assessment of ASD was found to be growing within the context of a rise in awareness and incidence of diagnosed autism. Yet, wait-list times remained high and demand on clinician time increased. An accessible instrument was needed that could identify autism and differentiate it from ONDs—the primary and secondary outcome measures of this study—with ease and reduced expert involvement. Patient involvement helped inform recruitment patterns of this diagnostic study, particularly to fit around family care and education patterns, and especially the children's daily routine. The latter was an important consideration for timing subject engagement. Educators and clinicians assisted with the early stages of pilot protocol development, ensuring the Play.Care test phase would fit alongside the normal day routine in the nursery, school, clinic, or in some cases, in the home.

Participants

A total of 760 children aged 30 months to 5 years 11 months will be recruited. Table 1 illustrates the participant numbers between sites and groups and table 2 details inclusion and exclusion criteria for each group.

Group sizes were decided on by assuming Play.Care will achieve a moderate 85% sensitivity. In this case, a study with 200 children with ASD will yield a confidence interval (CI) with a width of $\pm 4.9\%$, that is, the final sensitivity calculation will be accurate to within this range. The target recruitment size for this group is 280 children to afford some reduction in sensitivity or a significant non-completion or under-recruitment rate while preserving a narrow CI. The wider population recruited in this study in comparison to the earlier study may reduce sensitivity of the algorithm's prediction, inclusion of low functioning children with ASD increases the likelihood of disengagement with or disinterest in the Play.Care assessment, and the new recruitment strategies employed in this study may present recruitment challenges. Assuming a similar level of specificity, the same precision will be obtained with 200 TD children. Furthermore, if the sensitivity (or specificity) is as high as 90%, there will be 80% power to show that the sensitivity (or specificity) exceeds 83.0%, and 90% power to show it exceeds 81.9%. On the other hand, if the sensitivity (or specificity) is only 70%, there will be 91% power to show that it is below 80%. That is, if the Play.Care assessment is underperforming, which is the major concern of healthcare services, this study has a high probability of detecting it.

For the Glasgow cohort, recruitment of children with ASD or OND will be facilitated through the National

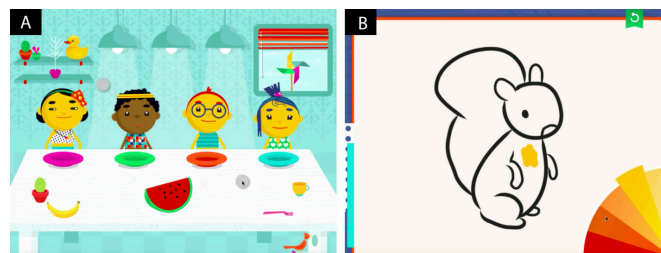


Figure 1 Play.Care assessment serious games. (A) 'Sharing'. The child taps the fruit to divide it into four pieces for sharing among the game characters. When all four characters have a slice of fruit, they express happiness for three seconds before the fruit is replaced with another food and the characters return to their original positions. (B) 'Creativity'. The child is free to choose an object or animal they wish to trace and then colour in freely by choosing a colour from the colour wheel. If the child is satisfied with their colouring, a new shape can be chosen by selecting the green button at the top of the screen, and the process repeats. Reproduced from Anzulewicz *et al.* (2016).

Health Service (NHS) Greater Glasgow and Clyde Specialist Children's Services, other specialist children's centres, nurseries and schools. ASD and OND recruitment will occur via two streams. The first will recruit children who receive a positive diagnosis of ASD or OND following an assessment at an NHS clinic. The second will recruit children already diagnosed with ASD or OND from specialist nurseries, preschools and development centres. TD children will be recruited from local private nurseries in the Greater Glasgow area.

The Gothenburg cohort is a general population sample as all children are screened for ASD within the Gothenburg area at 30 months. Families of children screened positively for ASD are referred to the Gillberg Neuropsychiatry Centre (GNC) in Gothenburg for neuropsychiatric assessment and it is at this stage that families will be made aware of the study and asked if they would like to participate. OND children in Gothenburg will be recruited through speech and language therapists in collaboration with the GNC.

Diagnostic device

The diagnostic tool under assessment in this study is the Play.Care iPad mini application, developed by Harimata sp. z o.o., Kraków, Poland. The assessment consists of two serious games played on the iPad. The first is called 'sharing' and requires the child to tap a piece of food to split it into four pieces, available for sharing with four cartoon game characters waiting patiently (figure 1A). The second game is called 'creativity' and requires the child to trace the outline of a drawing and then colour it in (figure 1B). Each gameplay session consists of a 2 min training phase where the researcher and/or parent can assist the child, and a 5 min assessment phase where the child must play unassisted. Data are collected for analysis only from the 5 min assessment phase. Following completion of the assessment, no further participation is required.

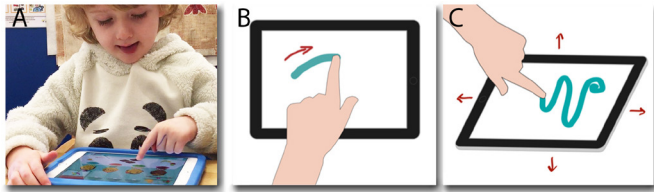


Figure 2 Movement data acquisition. (A) A child engages freely with the serious game. The tablet is protected by a bumper and placed firmly on a table. Movement data are acquired from (B) the touch screen and (C) the inertial movement unit sensor that detect the kinematics and contact forces of a gesture, respectively. Adapted from Anzulewicz *et al.* (2016).

Anzulewicz *et al.*¹⁸ details the data analytics employed, including identification of movement patterns by calculation of more than 200 ‘features’ that are assessed by an algorithm developed by machine learning. This algorithm statistically determines whether or not a particular child’s gameplay fits a diagnostic classification. It delivers a statistical prediction.

The gameplay data are collected by two sets of sensors within the iPad (figure 2): (1) the touch screen sensor records the Cartesian coordinate of the touch and its displacement as a gesture travels across the screen with a sampling rate of ca. 60 times per second; and (2) a triaxial accelerometer and gyroscope IMU sensor that detects the small accelerations and rotations of the iPad device with a sampling rate of ca. 100 times per second as the child’s fingers impact on the screen and push into it giving subtle, but significant displacive forces during a gesture. More than 200 movement ‘features’ are calculated from the raw sensor signals to characterise the child’s gameplay. These include, for example, for the touch screen sensors, the duration of a gesture, its maximum velocity, deviation from a straight line, its peak acceleration and the variance of these parameters across a gameplay session. Features from the raw IMU sensors are similarly extracted by calculating, for example, peak acceleration and rotation for each axis, their means and SDs. These features are described in full in Anzulewicz *et al.*¹⁸ Altogether, 10 repetitions of a 10-fold cross-validation procedure were carried out with Regularized Greedy Forest⁴² machine learning enhanced with additional data collected by Harimata in collaboration with the University of Strathclyde prior to this study to produce the algorithm employed in this study. The algorithm is fixed for the duration of the study.

The iPads used in this study are commercially available Apple products currently widely used for leisure, business and research. The iPad specifications employed here are as follows: white or black bevel, silver back iPad mini 4, 128GB. The iOS version will be the latest version available on purchase of the iPads. These specifications will remain fixed for the duration of the study. Furthermore, software updates will be switched off to prevent automatic updates of iOS which may alter the sampling rate of the internal gyroscope and accelerometers.

Procedure

Potential participants in the ASD and OND assessment stream will first be approached by the clinician undertaking the assessment. The family will be presented with the patient information sheet (PIS) following a positive diagnosis and if the clinician judges them to be suitable participants. The family will complete an interest slip indicating that they are willing to be further contacted by the research team regarding the study. This method was adopted to allow participation to be strictly ‘opt-in’ and to prevent the research team contacting families who had not expressed an interest in participating. After a minimum period of 24 hours, the research team will contact interested families to ask if they would like to participate and to arrange a data collection appointment.

Potential participants in the ASD and OND already diagnosed stream will be first approached at their child’s preschool where the research team will introduce the study during a parent’s evening or similar event. At this point, the PIS will be made available and families who are interested will complete an interest slip, as above, and after a minimum of 24 hours the research team will contact them.

Potential participants who are TD will be recruited through local private nurseries. Nursery staff will circulate the PIS among parents and those who wish their child to participate will return a completed consent form to the nursery. Prior to completing the Play.Care assessment, informed consent will be given by the child’s parents or caregivers. The researcher will then complete a paper-based case report form (CRF) for participants in the ASD and OND groups which details study ID, Scottish Community Health Index (‘CHI’) number, name, address, postal code, phone number and general practitioner details. Participants’ CHI numbers will be recorded to allow follow-up for up to 10 years following participation. This will provide potential future insight into whether diagnoses remain consistent into adolescence and whether any other diagnoses are received. Any further diagnoses can then be compared with the iPad data to determine any correlations between new diagnoses and movement characteristics in childhood.

An additional electronic CRF (eCRF) will be completed detailing the child’s state of arousal and interest on the day of data collection, and for the ASD and OND groups, data from a number of additional psychometric instruments including the Strengths and Difficulties Questionnaire (SDQ), the ESSENCE Questionnaire (ESSENCE-Q) and the Adaptive Behaviour Assessment System (ABAS) at Glasgow and the Vineland Adaptive Behavior Scale in Gothenburg will be recorded. In some cases, additional data will be recorded from instruments employed as part of the child’s neuropsychiatric assessment: Wechsler Preschool and Primary Scale of Intelligence ; ADI-R; and ADOS-2. The diagnostic criteria that determine a child’s diagnosis (International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) or DSM 5) will be recorded.

These psychometric instruments will be completed at an appropriate time during or following the data collection appointment. Additionally, the Glasgow TD group will also carry out the ESSENCE-Q, SDQ and ABAS where possible.

Following completion of the CRF and eCRF, the child will complete the Play.Care assessment. During the assessment, the iPad will be protected within a spongy back cover and bumper (iPad Mini2 Air Protect, Belkin, USA) and will be placed flat on the table top, where it will remain for the duration of the assessment. Parents/caregivers will have the option to consent to their child being videotaped during the assessment. If consent is given, a video camera will be setup and recording will begin prior to the child starting the assessment.

If a participant is unable to complete the assessment (eg, they become distracted or distressed), eCRF and CRF data will still be recorded and the participant will remain a subject in the study. This group will be classified as 'Play.Care incompletes' and will provide insight into the clinical applicability of the assessment. This group will be monitored at monthly intervals. If the rate of Play.Care incompletes reaches a level where it will impact the statistical power of the study, administration protocols will be reviewed and adjusted to ensure best possible compliance. This criterion will apply to the Glasgow cohort of participants only. The Gothenburg cohort is a general population sample and therefore participation cannot be controlled dependent on failure rates. If a participant does not have a confirmed clinical diagnosis by the end of the study, they will fall into the category of 'diagnosis failures'. The acceptable level of diagnosis failures is 7.5%, or, up to 21 subjects in the ASD group and 15 in the OND group.

Throughout the data collection phase, all members of the research team, the trial consortium and Harimata will be blinded to the Play.Care diagnostic prediction until all data have been collected. Following data collection, clinical diagnoses will be compared with Play.Care diagnostic predictions. The primary outcome measure is the ability of Play.Care to identify ASD cases within the general population (Gothenburg) and clinical population (Glasgow) by assessing classification efficacy against TD children. The secondary outcome measures are the ability of Play.Care to differentiate these ASD children from OND children.

CONCLUSIONS

This is the first phase 3 equivalent diagnostic study to test the predictive diagnostic accuracy of a smart tablet serious game for the early detection of autism in preschool children. The system is based on computational analysis of motor patterns inherent in the smart tablet sensor gameplay data, with pilot data suggesting sensitivity and specificity comparable or exceeding the current gold standard. This new system offers the benefit of rapid computational analysis with no lengthy subjectively rated

clinician-led elements to it, rather the system engages the child's interest in an attractive tablet-based serious game with raw sensor data analysed by classification algorithm. The instrument's computational accuracy is tested in this multisite diagnostic study of 760 children.

DISSEMINATION

Following completion of data collection, results will be disseminated in the peer-reviewed scientific literature, at national and international scientific conferences, on the Laboratory for Innovation in Autism website (<https://www.strath.ac.uk/research/innovationinautism/>) and by final report to Harimata sp. z o.o. and the European Commission. A lay summary of the study results will be available for participating families.

Author affiliations

¹Laboratory for Innovation in Autism, University of Strathclyde, Glasgow, UK

²Biomedical Engineering, University of Strathclyde, Glasgow, UK

³Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK

⁴Mental Health and Wellbeing, University of Glasgow, Glasgow, UK

⁵Centre for Rural Health, University of Aberdeen, Aberdeen, UK

⁶Gillberg Neuropsychiatry Centre, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

⁷Harimata sp. z o.o., Kraków, Poland

Acknowledgements We are grateful to the children, their parents, teachers and clinicians who have worked so hard to make this study possible, both in its design and in its implementation.

Contributors JD-B, CG, HM, PR, PW, LT, AM, KS and AA produced the study design. LM, JD-B, CG, HM, PR, AM, PW and LT developed the study protocol with input from KS and AA. Sample size calculations were carried out by AM. KS, AA and JD-B were responsible for technical development and pilot work. This paper was written by LM and JD-B with input from all coauthors.

Funding This work was subcontracted to the University of Strathclyde by Harimata sp. z o.o. as an integral part of a Horizon 2020 SME Instrument, grant number 756079.

Competing interests The academic authors LM, HM, PW, AM, PR, CG and JD-B are members of the trial steering and management committees and declare no financial interest in this product or the funding company, Harimata sp. z o.o. Coauthors AA and KS are board members of the Harimata sp. z o.o. that intends to commercialise the Play.Care assessment technology. AA and KS have options vesting in the company. AA is a voting member of the trial steering committee.

Patient consent for publication Not required.

Ethics approval This study was granted approval by the West of Scotland Research Ethics Service Committee 3 (reference number: 17-WS-0223 231435) for the National Health Service and by the University of Strathclyde Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Baio J, Wiggins L, Christensen DL, *et al*. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveill Summ* 2018;67:1-23.

2. Brugha TS, McManus S, Bankart J, *et al.* Epidemiology of autism spectrum disorders in adults in the community in England. *Arch Gen Psychiatry* 2011;68:459–65.
3. Knapp M, Romeo R, Beecham J. Economic cost of autism in the UK. *Autism* 2009;13:317–26.
4. Bradshaw J, Steiner AM, Gengoux G, *et al.* Feasibility and effectiveness of very early intervention for infants at-risk for autism spectrum disorder: a systematic review. *J Autism Dev Disord* 2015;45:778–94.
5. Pickles A, Le Couteur A, Leadbitter K, *et al.* Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *Lancet* 2016;388:2501–9.
6. Peters-Scheffer N, Didden R, Korzilius H, *et al.* Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in The Netherlands. *Res Dev Disabil* 2012;33:1763–72.
7. Chasson GS, Harris GE, Neely WJ. Cost Comparison of Early Intensive Behavioral Intervention and Special Education for Children with Autism. *J Child Fam Stud* 2007;16:401–13.
8. Liptrot K. Concern as children in York face long waits for autism diagnosis, The Press. 2019 <https://www.yorkpress.co.uk/news/17369842.concern-as-children-in-york-face-long-waits-for-autism-diagnosis/>.
9. BBC Wales. Autism assessment delay concerns across Wales: BBC News. 2018 <https://www.bbc.co.uk/news/uk-wales-42791769>.
10. Trevarthen C, Delafield-Butt JT. Autism as a developmental disorder in intentional movement and affective engagement. *Front Integr Neurosci* 2013;7:49.
11. Fournier KA, Hass CJ, Naik SK, *et al.* Motor coordination in autism spectrum disorders: a synthesis and meta-analysis. *J Autism Dev Disord* 2010;40:1227–40.
12. Delafield-Butt J. The emotional and embodied nature of human understanding: sharing narratives of meaning. In: Trevarthen C, Delafield-Butt J, Dunlop A-W, eds. *The child's curriculum: working with the natural voices of young children*. Oxford: Oxford University Press, 2018.
13. Delafield-Butt JT, Gangopadhyay N. Sensorimotor intentionality: The origins of intentionality in prospective agent action. *Developmental Review* 2013;33:399–425.
14. Delafield-Butt JT, Trevarthen C. The ontogenesis of narrative: from moving to meaning. *Front Psychol* 2015;6.
15. Iverson JM. Developing language in a developing body: the relationship between motor development and language development. *J Child Lang* 2010;37:229–61.
16. Trevarthen C, Delafield-Butt JT. Biology of shared meaning and language development: regulating the life of narratives. In: Legerstee M, Haley D, Bornstein M, eds. *The infant mind: origins of the social brain*. New York: Guildford Press, 2013:167–99.
17. Trevarthen C, Delafield-Butt JT. Development of consciousness. In: Hopkins B, Geangu E, Linkenauger S, eds. *Cambridge encyclopedia of child development*. Cambridge: Cambridge University Press, 2017:821–35.
18. Anzulewicz A, Sobota K, Delafield-Butt JT. Toward the Autism Motor Signature: Gesture patterns during smart tablet gameplay identify children with autism. *Sci Rep* 2016;6:31107.
19. Delafield-Butt JT, Freer Y, Perkins J, *et al.* Prospective organization of neonatal arm movements: a motor foundation of embodied agency, disrupted in premature birth. *Dev Sci* 2018;21:e12693.
20. Craig CM, Grealy MA, Lee DN. Detecting motor abnormalities in preterm infants. *Exp Brain Res* 2000;131:359–65.
21. Taffoni F, Tamilia E, Palminteri MR, *et al.* Ecological sucking monitoring of newborns. *IEEE Sens J* 2013;13:4561–8.
22. Hocking DR, Caeyenberghs K. What is the nature of motor impairments in autism, are they diagnostically useful, and what are the implications for intervention? *Current Developmental Disorders Reports* 2017;4:19–27.
23. Teitelbaum O, Teitelbaum P. *Does your baby have autism?: Detecting the earliest signs of autism*. Garden City Park, NY: Square One Publishers, 2008.
24. Teitelbaum P, Teitelbaum O, Nye J, *et al.* Movement analysis in infancy may be useful for early diagnosis of autism. *Proc Natl Acad Sci U S A* 1998;95:13982–7.
25. Mari M, Castiello U, Marks D, *et al.* The reach-to-grasp movement in children with autism spectrum disorder. *Philos Trans R Soc Lond B Biol Sci* 2003;358:393–403.
26. Ghaziuddin M, Butler E, autism Cin. and Asperger syndrome: a further report. *J Intellect Disabil Res* 1998;42(Pt 1):43–8.
27. Weiss MJ, Moran MF, Parker ME, *et al.* Gait analysis of teenagers and young adults diagnosed with autism and severe verbal communication disorders. *Front Integr Neurosci* 2013;7:33.
28. Nayate A, Tonge BJ, Bradshaw JL, *et al.* Differentiation of high-functioning autism and Asperger's disorder based on neuromotor behaviour. *J Autism Dev Disord* 2012;42:707–17.
29. Kindregan D, Gallagher L, Gormley J. Gait deviations in children with autism spectrum disorders: a review. *Autism Res Treat* 2015;2015:1–8.
30. Lloyd M, MacDonald M, Lord C. Motor skills of toddlers with autism spectrum disorders. *Autism* 2013;17:133–46.
31. Torres EB, Donnellan AM, eds. *Autism: the movement perspective*: Nature Frontiers Group, 2013. <https://www.frontiersin.org/research-topics/801/autism-the-movement-perspective..>
32. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington DC: American Psychiatric Association, 2013.
33. Bosco P, Giuliano A, Delafield-Butt J, *et al.* Brainstem enlargement in preschool children with autism: Results from an intermethod agreement study of segmentation algorithms. *Hum Brain Mapp* 2019;40:7–19.
34. Delafield-Butt J, Trevarthen C. On the brainstem origin of autism: disruption to movements of the primary self. In: Torres E, Whyatt C, eds. *Autism: the movement sensing perspective*: Taylor & Francis CRC Press, 2017.
35. Hatakenaka Y, Kotani H, Yasumitsu-Lovell K, *et al.* Infant Motor Delay and Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations in Japan. *Pediatr Neurol* 2016;54:55–63.
36. Chen J, Wang G, Zhang K, *et al.* A pilot study on evaluating children with autism spectrum disorder using computer games. *Comput Human Behav* 2019;90:204–14.
37. Noor HAM, Shahbodin F, Ananta GP, *et al.* A prototype development of visual perception diagnosis games for autism children. *International Journal of Advanced and Applied Sciences* 2017;4:68–72.
38. Zakari HM, Ma M, Simmons D. A Review of Serious Games for Children with Autism Spectrum Disorders (ASD). In: M M, MF O, BH J, eds. *Serious Games Development and Applications SGDA 2014 Lecture Notes in Computer Science*. Cham: Springer, 2014:93–106.
39. Fleming TM, Bavin L, Stasiak K, *et al.* Serious Games and Gamification for Mental Health: Current Status and Promising Directions. *Front Psychiatry* 2016;7.
40. de Bildt A, Sytema S, van Lang ND, *et al.* Evaluation of the ADOS revised algorithm: the applicability in 558 Dutch children and adolescents. *J Autism Dev Disord* 2009;39:1350–8.
41. Bossuyt PM, Reitsma JB, Bruns DE, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology* 2015;277:826–32.
42. Johnson R, Tong Z. Learning Nonlinear Functions Using Regularized Greedy Forest. *IEEE Trans Pattern Anal Mach Intell* 2014;36:942–54.