Original Paper

Integration of an FDA-Authorized Digital Autism Diagnostic Tool Into an Outpatient Developmental Pediatrics Workflow: Preliminary Quality Improvement Study

Jenna B Resnik[1], BS; Rebecca T Lane[1,2], MD; Sinan O Turnacioglu[1,2], MD

[1]School of Medicine and Health Sciences, George Washington University, Washington, DC, United States

[2]Department of Neurodevelopmental Pediatrics, Children's National Hospital, Washington, DC, United States

Corresponding Author:
Jenna B Resnik, BS
School of Medicine and Health Sciences
George Washington University
2300 I St. NW
Washington, DC, 20052
United States
Phone: 1 2029942295
Email: jresnik267@gwu.edu

Abstract

Background: Long wait times and specialist shortages delay autism spectrum disorder (ASD) diagnosis, often beyond the period when early intervention is most effective. Digital diagnostic tools may help increase diagnostic capacity by supporting clinicians in delivering timely autism evaluations within routine outpatient care.

Objective: This study aimed to conduct a preliminary assessment of the integration of a digital autism diagnostic device (Canvas Dx) into an outpatient developmental pediatrics workflow, including a descriptive evaluation of agreement between device outputs and specialist clinical diagnoses and an exploration of clinician and caregiver experience.

Methods: A single-center, prospective quality improvement study evaluated the implementation and performance of Canvas Dx in a specialty developmental pediatrics clinic. Canvas Dx is a Food and Drug Administration (FDA)–authorized digital ASD diagnostic device approved for use in children aged 18 to 72 months with caregiver or health care provider concern for developmental delay. Eligible children were consecutively prescribed Canvas Dx during the study period and completed the assessment as part of routine clinical care. Device diagnostic outputs were compared with the final specialist clinical diagnoses. Clinician and caregiver feedback was collected to explore their experience with the assessment process and perceived workflow impact via poststudy surveys that captured both structured quantitative questions and open-ended qualitative questions.

Results: A total of 28 children were included (mean age 52.2, SD 13.8 months; 7/28, 25% female), with an autism prevalence of 67.9% (19/28) based on specialist diagnosis. The device generated a determinate result for 22 (78.6%) participants. Among participants with determinate device results, sensitivity was 94.1% (95% CI 73%-99%) and specificity was 80% (95% CI 37.6%-96.4%). Positive predictive value was 94.1% (95% CI 73%-99%), and negative predictive value was 80% (95% CI 37.6-96.4%) when compared with specialist diagnosis. These estimates were imprecise, particularly for specificity and negative predictive value, because of the small number of participants without an ASD diagnosis. Clinicians reported improved insights and ease of use during evaluation visits, and caregivers reported high satisfaction with the assessment process.

Conclusions: This preliminary quality improvement assessment demonstrates the feasibility of integrating Canvas Dx into a routine outpatient developmental pediatrics workflow and provides initial descriptive data on agreement with specialist diagnostic determinations and clinician and caregiver experience. Given the small sample size and resulting imprecision of diagnostic performance estimates, these findings should be interpreted as preliminary and used to inform larger prospective evaluations of diagnostic performance and workflow impact.

(JMIR Form Res 2026;10:e98697) doi: 10.2196/98697

KEYWORDS

autism spectrum disorder; digital health; artificial intelligence; AI; pediatrics; quality improvement

XSL•FO
RenderX

Introduction

Background

The prevalence of autism spectrum disorder (ASD) has steadily increased over the last 2 decades [1], while there is simultaneously a shortage of specialists trained to diagnose ASD [2-4] This has contributed to an average age of diagnosis that is later than the optimal period [5] for young children to receive the intensive behavioral intervention that contributes to optimal long-term outcomes [6,7]. Long wait times for ASD diagnostic evaluations remain a significant challenge across the United States, with recent studies reporting waits of several months to more than 1 year [8-10]. At the time this project was conducted, internal clinic waitlist data maintained by the Division of Neurodevelopmental Pediatrics at Children's National Hospital indicated that wait times for autism diagnostic evaluations exceeded 12 months. There is thus an urgent need for improved processes to establish ASD diagnoses in a timely fashion so that families of young children with ASD can access appropriate intervention services [11].

Canvas Dx is a Food and Drug Administration (FDA)–authorized diagnostic device that can support the diagnosis or ruling out of ASD in children with developmental delay or other ASD-specific concerns between 18 and 72 months of age [12]. This project assessed the impact of incorporating the Canvas Dx system into the workflow of outpatient developmental pediatrics clinics at Children's National. The system was used with 28 consecutive patients who met the criteria for the Canvas Dx system's indications for use, as determined by the division's intake coordinator on initial contact with referred families.

Canvas Dx and Existing Evidence

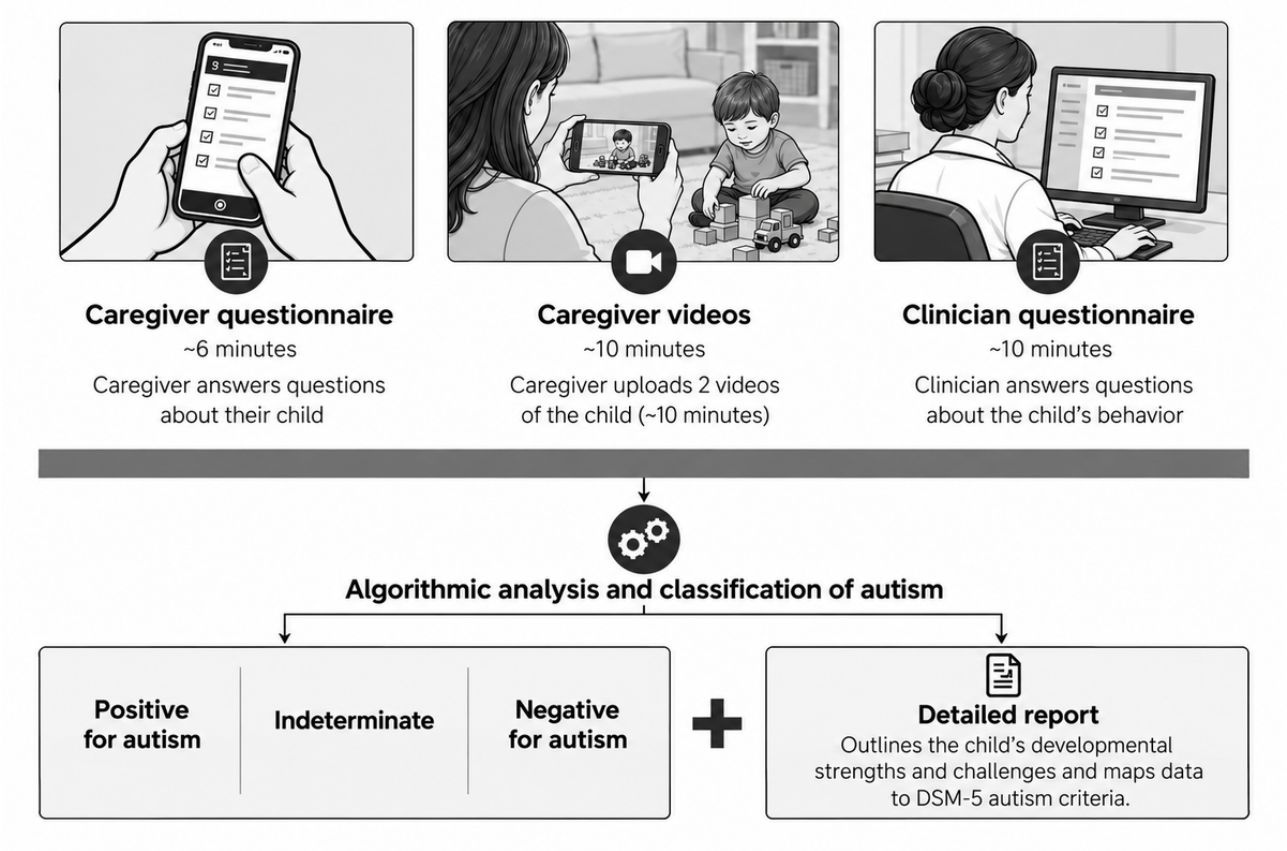
Canvas Dx is a prescription-only, FDA-authorized software-based diagnostic aid intended to support clinicians in diagnosing or ruling out ASD in children aged 18 to 72 months who have concerns about developmental delay [13]. The device combines information from 3 independent sources, including a caregiver questionnaire, a treating clinician questionnaire, and a standardized clinical analysis of caregiver-uploaded home videos, within a locked machine learning algorithm to generate a positive, negative, or indeterminate result for autism. To support clinical interpretation, contextual information, including a detailed developmental profile that maps observed behaviors to the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]) diagnostic criteria for ASD and summarizes strengths and challenges across key developmental domains, is also generated to help inform clinical decision-making and intervention planning. For participants receiving an indeterminate result, the report also includes a predefined indeterminate category (low, moderate, or high). The low, moderate, and high categories are defined by prespecified algorithmic thresholds designed to stratify ASD likelihood among children for whom the device cannot generate a definitive positive or negative result. The thresholds were selected during device development to provide the clearest separation in ASD likelihood within the indeterminate population, with higher categories representing progressively greater likelihood of ASD. These evidence-based categories were established using an independent development dataset (n=261; autism prevalence 26.8%; 34% female). Within this development dataset, 5% of children classified as low indeterminate, 25% classified as moderate indeterminate, and 35% classified as high indeterminate received a specialist-confirmed ASD diagnosis, supporting the intended risk stratification across categories [12]. Data inputs and outputs are summarized in Figure 1.

Evidence supporting Canvas Dx includes results from a prospective, double-blinded, multisite pivotal clinical trial that demonstrated high agreement between Canvas Dx outputs and blinded specialist diagnoses, with consistent performance across sex, race and ethnicity, and socioeconomic status [14]. Subsequent real-world studies have demonstrated that Canvas Dx maintains accuracy outside of investigational settings and can substantially reduce the time to ASD diagnosis compared with traditional specialty referral pathways [12,15].

In a prospective observational study integrating Canvas Dx into a primary care diagnostic workflow, for example, use of the system was reported to be feasible for clinicians and acceptable to families, while compressing the diagnostic timeline by several months relative to nearby specialty centers [16]. A larger real-world case series further demonstrated strong diagnostic performance across diverse practice settings, with determinate outputs produced for most cases and median ages for positive results occurring well before the national average age of ASD diagnosis in the United States [12]. To date, most published evaluations of Canvas Dx supporting its clinical validity, real-world effectiveness, and potential system-level benefits have focused on primary care setting integration [12,16]. In contrast, the present study provides a preliminary examination of its potential applicability within a specialty diagnostic pathway.

XSL•FO
RenderX

Figure 1. The Canvas Dx system integrates data from a caregiver-completed questionnaire, features extracted from short videos depicting the child's play and social interactions, and a clinically focused questionnaire assessing the child's behavior and development. These inputs are combined using machine learning methods to generate an autism spectrum disorder (ASD) prediction (positive or negative) and a detailed developmental report that maps observed features and behavioral data to DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]) ASD diagnostic criteria, in support of clinicians' independent clinical judgment.



Objectives

This quality improvement project was designed as a preliminary assessment to characterize real-world use of Canvas Dx and generate descriptive data to inform future larger-scale evaluation.

Primary Objective

The primary objective was to preliminarily characterize agreement between Canvas Dx determinate outputs and standard clinical autism diagnoses made by developmental pediatrics specialists in a routine outpatient clinical setting and to generate descriptive diagnostic performance estimates to inform future larger-scale evaluation.

Secondary Objectives

The secondary objective was to explore clinician and caregiver experience with Canvas Dx and perceived workflow impact associated with incorporating the system into outpatient developmental pediatrics practice.

Methods

Study Design

This single-center, prospective quality improvement study preliminarily evaluated the implementation of Canvas Dx in a specialty developmental pediatrics clinic and descriptively characterized its diagnostic performance compared with specialist clinical diagnosis.

Ethical Considerations

This preliminary quality improvement project was reviewed by the Children's National Hospital Institutional Review Board (IRB), which determined that the project was exempt from full IRB review because it met the criteria for a quality improvement initiative intended to improve clinical processes and patient care. The project did not involve experimental interventions. Because the project was determined to be exempt, the IRB determined that formal informed consent was not required. All data used for analysis were deidentified prior to analysis and reporting. Participant privacy and confidentiality were protected in accordance with institutional policies and applicable ethical standards, and no individually identifiable information was included in the study results. Participants did not receive financial compensation for participation in this quality improvement project.

Setting

The project took place within an outpatient developmental pediatrics clinic affiliated with Children's National Hospital in Washington, DC.

Participants

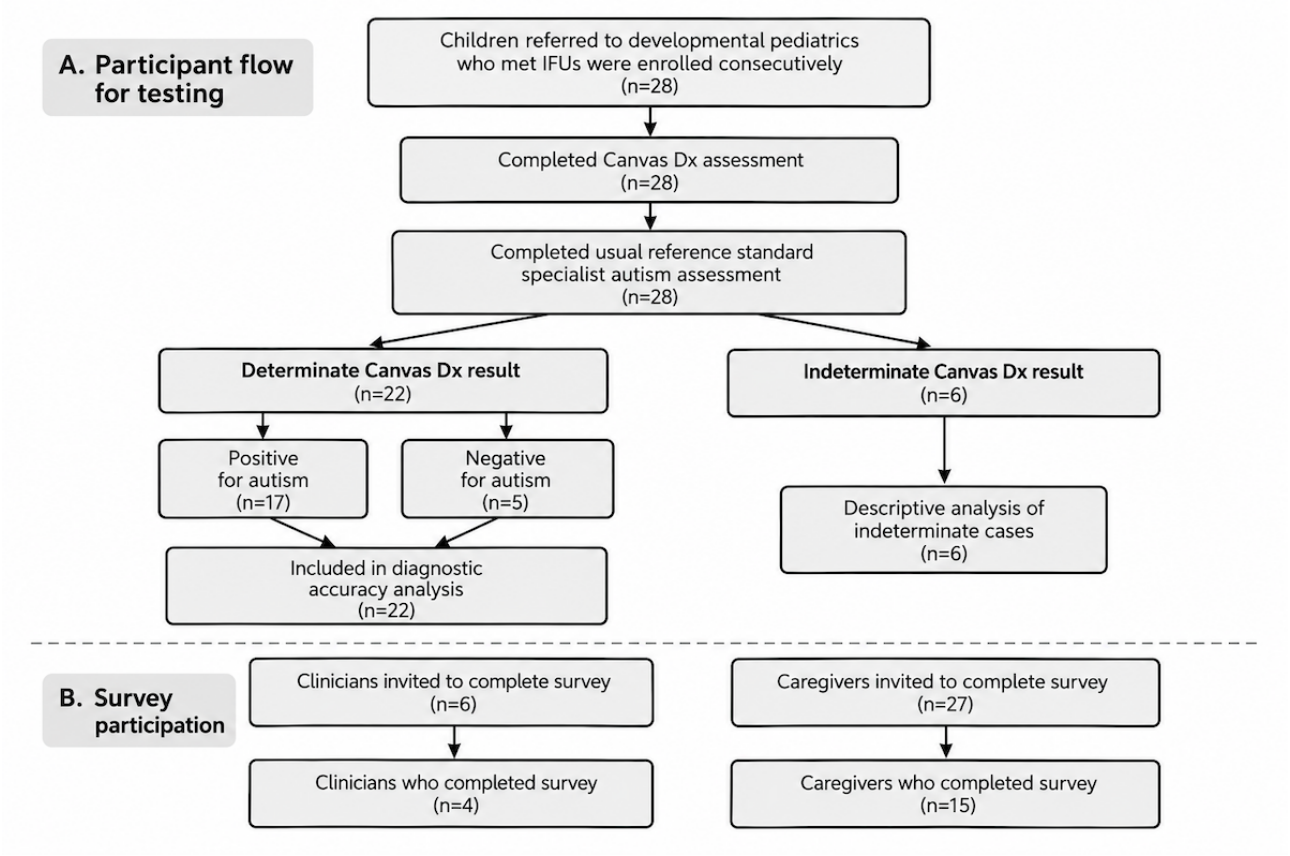
Participants were enrolled and data were collected between May and September 2025 as part of routine clinical care. Participants

XSL•FO
RenderX

included children aged 18 to 72 months with caregiver or health care provider concern for developmental delay who met the device's indications for use. The complete Canvas Dx indications for use are included in Multimedia Appendix 1. Children were referred from primary care to the developmental pediatrics clinic for evaluation of a range of developmental concerns, including but not limited to suspected ASD. Consistent with the Canvas Dx indications for use, eligibility was based on the presence of broad developmental concerns rather than an ASD-specific referral; therefore, children were not excluded if suspicion of ASD was not the primary reason for referral. Eligibility was determined during intake prior to clinic visits.

No formal sample size target or power calculation was established because this was a time-limited quality improvement project rather than a prospectively powered diagnostic accuracy study. Enrollment was determined pragmatically by the number of eligible patients meeting the device indications for use who presented during the implementation period and by the limited number of device licenses available for the project. Accordingly, the resulting sample was intended to provide preliminary descriptive data to inform future larger-scale evaluation. Participant flow through the study is summarized in Figure 2.

Figure 2. Participant assessment flow and survey participation. (A) Of 28 enrolled children, 22 received determinate Canvas Dx results and were included in the diagnostic accuracy analysis, while 6 received indeterminate results and were analyzed separately. Diagnostic accuracy was calculated using determinate results only. (B) Survey completion among invited clinicians (4/6) and caregivers (15/27). IFU: indications for use.



Intervention

Caregivers completed the Canvas Dx app-based questionnaire and uploaded short home videos of their child prior to the clinic visit. Clinicians completed the Canvas Dx clinical questionnaire. The reference standard clinic visit consisted of a comprehensive, approximately 1-hour, in-person developmental evaluation conducted by a board-certified developmental pediatrician. The evaluation included a caregiver interview, review of developmental and medical history, direct clinical observation of the child, review of available records, and any additional developmental assessments deemed clinically appropriate. Following completion of the evaluation, the developmental pediatrician integrated all available clinical information to determine whether the child met *DSM-5* criteria for ASD or another neurodevelopmental condition.

Outcome Measures

The primary outcome measured was diagnostic agreement between the digital diagnostic device and specialist clinical diagnosis. Secondary outcomes included clinician- and caregiver-reported experience with the device.

Diagnostic Agreement Measurement

The specialist clinical diagnosis was categorized as ASD or non-ASD based on the developmental pediatrician's final clinical determination. Canvas Dx generated 1 of 3 possible outputs: positive for ASD, negative for ASD, or indeterminate. For participants receiving a determinate Canvas Dx result, agreement was defined as a positive Canvas Dx result with a specialist diagnosis of ASD or a negative Canvas Dx result with a specialist diagnosis of non-ASD. Discordance was defined as a positive Canvas Dx result with a specialist diagnosis of

XSL•FO
RenderX

non-ASD or a negative Canvas Dx result with a specialist diagnosis of ASD.

Clinician and Caregiver Experience Measures

Overview

Clinician and caregiver experience with Canvas Dx was assessed using brief surveys administered following use of the device. Surveys measured secondary outcomes including satisfaction and comfort with the assessment process, perceived ease of use, acceptability, time burden, confidence in the diagnostic result, and willingness to recommend the approach. The complete caregiver and clinician survey instruments are provided in Multimedia Appendices 2 and 3.

Caregiver Experience

Caregivers were asked whether they were comfortable using the Canvas Dx device for their child's evaluation (yes or no). Five additional items assessed whether the device worked without problems, whether instructions for recording the child were easy to follow, whether the caregiver could use the device independently, willingness to recommend the device to other families, and trust in the diagnosis received from the health care provider. These items used a 5-point Likert response scale ranging from strongly disagree to strongly agree. Time burden was assessed by asking caregivers how much time they spent completing Canvas Dx tasks, with categorical response options. Optional open-ended questions invited caregivers to describe what they liked and did not like about the device and provide additional feedback.

Clinician Experience

Clinicians were asked whether they were comfortable using Canvas Dx as part of their patient's evaluation (yes or no). Five additional items assessed whether the platform worked without difficulties, ability to use it without assistance from the Canvas Dx team, confidence in the diagnosis received from Canvas Dx, perceived usefulness of the assessment summary, and willingness to recommend this type of evaluation to other health care providers. These items used a 5-point Likert response scale ranging from strongly disagree to strongly agree. Clinicians were also asked whether they reviewed the full Canvas Dx report and whether they would consider continuing to use Canvas Dx in clinical practice (yes or no). Optional open-ended questions solicited feedback regarding the report, reasons for continued use, comparison with other diagnostic tools, and additional experiences with Canvas Dx.

Data Analysis

Diagnostic Agreement Analysis

Descriptive analyses and diagnostic performance calculations were conducted using Python (version 3.12; Python Software Foundation). Diagnostic agreement and determinate versus indeterminate outcomes were summarized using counts and percentages. Given the preliminary nature of this quality improvement study and the small sample size, no inferential statistical testing was performed.

Sensitivity, specificity, positive predictive value, and negative predictive value were calculated among determinate cases, with 95% CIs calculated using the Wilson score method. Indeterminate results were excluded from these calculations because they do not represent positive or negative diagnostic classifications, but rather cases in which the algorithm abstains from making a diagnostic call because of uncertainty.

The abstention mechanism allows the system to acknowledge uncertainty and avoid overconfident classification when evidence is insufficient [17,18]. Analyzing abstention cases separately is consistent with prior evaluations of Canvas Dx [12,16] and reflects the intended behavior of an abstaining classifier, for which diagnostic accuracy metrics describe performance only when the algorithm issues a classification, while the determinate rate provides complementary information regarding the frequency of abstention.

Survey Analysis

Responses to quantitative clinician and caregiver survey items were analyzed descriptively. Categorical responses were summarized using frequencies and percentages, and Likert-scale responses were summarized by the distribution of responses across response categories. Given the small sample size and exploratory nature of these secondary outcomes, no inferential statistical testing was performed.

Open-ended clinician and caregiver survey responses were analyzed using an inductive qualitative content analysis approach [19,20]. Responses were reviewed in their entirety, and meaningful units of text were identified and assigned descriptive codes. Related codes were then grouped into broader descriptive categories reflecting recurring aspects of clinician and caregiver experience with Canvas Dx. Caregiver and clinician responses were analyzed separately because the survey questions and perspectives differed between groups. Responses provided in Spanish were translated into English for analysis, with the original responses retained for reference. Given the small number and brief nature of the open-ended responses, the analysis was intended to provide descriptive context regarding participant experience rather than to achieve thematic saturation. One researcher performed the initial coding and category development, and a second researcher independently assigned the source responses to the established categories, with no discrepancies identified.

Results

Demographics

A total of 28 children meeting the indications for use were consecutively enrolled in the quality improvement project (mean age 52.2, SD 13.8 months). Seven (25%) were female, and 21 (75%) were male. Autism prevalence according to the reference standard specialist diagnosis was 67.9% (19/28). There were no missing data for the primary diagnostic accuracy analyses. Participants represented a racially and ethnically diverse cohort, with the largest groups identifying as Black and not Hispanic and as other and Hispanic. The distribution of participant race and ethnicity is presented in Table 1.

XSL•FO
RenderX

Table 1. Self-reported race and ethnicity of participants enrolled in the developmental pediatrics quality improvement study (N=28)[a].

Caregiver-reported race and ethnicity categories	Overall, n (%)
Black and not Hispanic	9 (32.1)
Other and Hispanic	8 (28.6)
White and not Hispanic	3 (10.7)
Declined to include	2 (7.1)
Asian and not Hispanic	1 (3.6)
Black and Hispanic	1 (3.6)
White and Hispanic	1 (3.6)
Race declined and Hispanic	1 (3.6)
Race declined and not Hispanic	1 (3.6)
American Indian or Alaska Native or ethnicity not reported	1 (3.6)

[a]Percentages are based on total sample size (N=28) and may not sum to 100% due to rounding. Race and ethnicity categories are reported as provided by participants or caregivers.

Agreement Between Determinate Device Outputs and Specialist Clinical Diagnoses

The device generated a determinate result (positive or negative for autism) for 22 (78.6%) of the 28 participants. Among participants with determinate results, the device demonstrated strong agreement with specialist clinical diagnoses. Of the 17 participants who received a specialist-confirmed autism diagnosis, 16 also received a positive autism result from the device, corresponding to a sensitivity of 94.1% (16/17; 95% CI 73%-99%). Of the 5 participants who were determined by specialists not to meet the criteria for autism, 4 also received a negative result from the device, yielding a specificity of 80% (4/5; 95% CI 37.6%-96.4%).

The positive predictive value of the device was 94.1% (16/17; 95% CI 73%-99%), with 16 of 17 participants who screened positive receiving a specialist-confirmed ASD diagnosis. The negative predictive value was 80% (4/5; 95% CI 37.6%-96.4%), with 4 of 5 participants who screened negative confirmed as not having ASD.

Indeterminate Analysis

A total of 6 indeterminate results were received (6/28, 21.4%). Four (66.7%) were high indeterminate, 2 (33.3%) were moderate indeterminate, and none were low indeterminate. Autism prevalence within the indeterminate group was 33.3% (2/6). Both high indeterminate cases were confirmed as having autism by a clinician. More than half (4/6, 66.7%) of the participants with indeterminate results had reported co-occurring conditions. For indeterminate cases, global developmental delay was reported for 33.3% (2/6), language disorder was reported for 16.7% (1/6), and attention-deficit/hyperactivity disorder was reported for 16.7% (1/6).

Clinician Experience

Of the 6 clinicians who participated in the pilot, 5 were board-certified developmental-behavioral pediatricians, and 1 was a nurse practitioner. Four (66.7%) of the participating clinicians completed poststudy feedback surveys regarding their experience using Canvas Dx as part of the diagnostic process. All 4 indicated that they were comfortable using Canvas Dx as part of their patients' evaluations. All 4 agreed or strongly agreed that the site worked without difficulties, that they were able to use the site without assistance from the Canvas Dx team, that they were confident in the diagnoses their patients received from Canvas Dx, and that they would recommend this type of evaluation to other health care providers. Responses regarding the usefulness of the assessment summary were more variable: 2 clinicians strongly agreed that it was useful, 1 agreed, and 1 strongly disagreed. All 4 clinicians reported that they had read the full Canvas Dx report and that they would consider continuing to use Canvas Dx in their clinical practice.

Inductive qualitative content analysis of clinician free-text responses identified 4 primary descriptive categories: value of structured clinical information, value of naturalistic behavioral information, perceived clinical and access value, and limitations. Clinicians described the Canvas Dx report as organized and clinically informative, including its structured presentation of behaviors and symptoms and organization according to *DSM-5* criteria. One clinician described the report as "well organized," noting that it "highlighted behaviors and symptoms appropriately" and "provided relevant information for care management." Another noted, "I appreciated the breakdown of response by *DSM-5* criteria."

Clinicians also described the value of naturalistic behavioral information, particularly the ability to review videos of children outside the clinical environment. One clinician reported, "I really liked having videos of the patients in their home or other comfortable setting to see before I met the family." Perceived clinical and access value included convenience for families, parent involvement, confirmation of clinical impressions, and the potential to improve access to autism diagnosis and intervention. One clinician noted that Canvas Dx "allows the parents to be involved in the diagnostic process, which often helps with acceptance of the diagnosis."

The qualitative analysis also identified several individual limitations or concerns. One clinician reported difficulty with 2-factor authentication due to inconsistent cellular service. The

XSL•FO
RenderX

same clinician described uncertainty regarding the incremental impact on their own diagnostic decision-making, noting that although Canvas Dx provided confirmation of their clinical impression, it did not necessarily change their diagnostic decision. The clinician also suggested that the tool may have greater value for less specialized health care providers seeking to provide an earlier diagnosis. Additionally, one clinician expressed concern regarding the report template's requirement to assign a level of support need, indicating that this feature may be limiting in certain clinical contexts.

Caregiver Experience

In total, 15 (55.6%) of the 27 participating caregivers returned their postassessment feedback survey regarding their experience using Canvas Dx as part of their child's diagnostic evaluation. One caregiver was the parent of twins who completed the survey once for her 2 children. All respondents (15/15, 100%) reported that they were comfortable using the Canvas Dx device, as indicated by affirmative responses to a binary yes or no item assessing overall comfort with the platform.

Overall ease of use and functionality were rated highly. For the statement "The Canvas Dx app worked fine without problems," 86.7% (13/15) of caregivers selected "strongly agree," while the remaining 13.3% (2/15) selected "agree." Similarly, 93.3% (14/15) strongly agreed that the directions for recording their child were easy to follow, with one caregiver (6.7%) selecting "agree." With respect to independent use, 86.7% (13/15) of respondents strongly agreed that they could use the app without assistance, and 13.3% (2/15) agreed. When asked whether they would recommend the app to other families, 86.7% (13/15) of caregivers strongly agreed and 13.3% (2/15) agreed. Using an adapted net promoter score (NPS) calculation in which "strongly agree" responses were categorized as promoters and "agree" responses as passives, the resulting NPS was 87, reflecting a high willingness to recommend the device among respondents. Caregivers also reported favorable perceptions of efficiency. Ten (66.7%) respondents indicated that completing the questionnaire and recording and uploading 2 home videos took less than 15 minutes, while 5 (33.3%) caregivers reported completion times of 15 to 30 minutes. No participants reported requiring more than 30 minutes to complete the process, suggesting low time burden for most families. Caregivers also reported high levels of trust in the results their child received from the health care provider at the clinic visit, with 86.7% (13/15) strongly agreeing and 13.3% (2/15) agreeing with this statement.

Inductive qualitative content analysis of caregiver free-text responses identified ease and simplicity of completing the assessment as the primary descriptive category. All 6 caregivers who responded to the question regarding what they liked about Canvas Dx described aspects of ease of use, including the simplicity and practicality of the device, clarity of instructions, or ease of recording and uploading videos. Representative comments included, "It is very easy and practical," "I liked that the instructions are very well explained," and "Videos were easy to record or upload." Additional open-ended comments reflected trust and appreciation for the assessment process, including "Overall, there is great confidence in how it works" and "I am very grateful for being shown the device." No negative aspects of the Canvas Dx device were identified in caregiver open-ended responses.

Discussion

Principal Findings

In this preliminary quality improvement project evaluating the integration of Canvas Dx into a specialist autism diagnostic workflow, Canvas Dx determinate results showed a high level of agreement with specialist clinical diagnoses. However, the small sample size, particularly the small number of participants without an ASD diagnosis, resulted in wide CIs and limited the precision of diagnostic performance estimates. Accordingly, these findings should be considered preliminary and descriptive rather than confirmatory. They provide initial evidence to inform larger studies evaluating Canvas Dx performance within our specialist developmental pediatrics workflows. While subgroup sample sizes were too small to support meaningful comparisons of diagnostic performance across sex, race, and ethnicity, the study cohort was demographically diverse and included substantial representation of participants identifying as Black or African American (non-Hispanic), a population historically underrepresented in clinical research [21,22] and disproportionately affected by delayed diagnosis. Equitable representation across demographic groups is a critical consideration for digital diagnostic tools intended for broad clinical deployment [23], as it supports their potential to contribute to more consistent and timely identification across populations that have traditionally faced barriers to care.

Of the 6 participants who received an indeterminate result, autism prevalence was lower than in the positive subset, and a majority had documented co-occurring developmental conditions, including global developmental delay, language disorder, or attention-deficit/hyperactivity disorder. Although the small number of indeterminate cases precludes firm conclusions, the observed co-occurrence of developmental conditions among these children warrants further investigation into whether phenotypic complexity may be associated with indeterminate outputs. The ability to return an indeterminate result rather than forcing classification in ambiguous cases is consistent with the device's abstention-based responsible AI framework [17,18].

Beyond diagnostic performance, clinician and caregiver feedback indicated generally positive experiences with Canvas Dx, including satisfaction, perceived ease of use, and acceptability within the clinical workflow. Qualitative content analysis of clinician feedback identified perceived value of structured clinical information and naturalistic behavioral information as key categories, including the structured presentation of *DSM-5* criteria and the ability to review videos of children in naturalistic settings prior to in-person visits. Caregivers also reported high levels of comfort with the assessment process and found the platform easy to use and time efficient. Consistent with these quantitative findings, qualitative content analysis of caregiver feedback identified ease and simplicity of completing the assessment as the primary descriptive category. These findings suggest that the assessment

XSL•FO
RenderX

can be completed with minimal burden on families and is acceptable for use within routine clinical care. Caregivers also expressed a strong willingness to recommend the tool to others, further supporting its overall acceptability.

Comparison With Prior Work

The largest real-world Canvas Dx study to date did not report on caregiver or clinician experience [12]. However, findings from our study related to clinician and caregiver acceptability mirror those reported in the study by Sohl et al [16], in which Canvas Dx was integrated into a primary care workflow. In that study, both health care providers and families reported ease of use, and device results showed agreement with the final clinical diagnosis. The present project extends those findings to a specialist diagnostic context and demonstrates similarly favorable perceptions of device workflow integration and clinical accuracy.

Notably, the proportion of children receiving a determinate result in the current project was higher than that reported in both the Sohl et al [16] and Salomon et al [12] studies. This difference may reflect variation in population characteristics, referral patterns, clinical setting, or implementation workflow. The higher determinate rate observed here suggests that in certain clinical contexts, the device may be able to provide actionable diagnostic support for a substantial majority of patients. Further comparative evaluation across care settings would be valuable to better understand the factors that influence device performance in different contexts.

More broadly, the present findings contribute to a growing literature evaluating alternative approaches to traditional autism diagnostic pathways. Primary care–based diagnostic models, telehealth assessment, remote behavioral observation, and digital approaches have increasingly been evaluated as strategies to expand diagnostic capacity and facilitate more timely autism evaluation [24-27]. A 2023 scoping review of ASD diagnosis in US primary care settings identified studies evaluating diagnostic accuracy, time to diagnosis, specialty wait times, and clinician comfort with diagnosis, supporting the feasibility of expanding diagnostic evaluation beyond traditional pathways [24]. Within specialty settings, investigators have evaluated fast-track and stratified diagnostic models designed to better allocate specialist resources and reduce time to diagnosis [28]. Collectively, these studies reflect increasing efforts to move beyond a uniform, resource-intensive diagnostic pathway toward approaches that match the intensity of assessment to individual clinical needs [29]. The present study contributes to this literature by providing preliminary evidence regarding the integration of an FDA-authorized digital diagnostic device into an existing specialty developmental pediatrics workflow.

Clinical and Implementation Implications

These findings have several implications for future clinical evaluation and implementation. First, the preliminary diagnostic performance estimates observed among determinate cases, particularly for sensitivity and positive predictive value, suggest that the device may be particularly useful in supporting accurate identification of ASD, with the potential to facilitate more timely access to appropriate intervention. However, given the small sample size and resulting imprecision of these estimates, particularly for specificity and negative predictive value, larger studies are needed to more precisely characterize diagnostic performance and potential clinical impact. Second, the abstention mechanism appears to appropriately flag cases with phenotypic complexity, reinforcing rather than replacing clinician expertise. Third, the integration of structured caregiver-reported data and home-based video provides clinicians with insight into child behavior in naturalistic settings, which may enhance diagnostic confidence and family engagement.

From a workflow perspective, clinician feedback from this preliminary quality improvement project suggests that the device can be incorporated into routine evaluation processes without substantial technical barriers, although these findings should be confirmed in larger implementation studies. Caregiver feedback suggests low time burden and high user satisfaction, perceived ease of use, and acceptability, which are critical components for sustained adoption.

Limitations

This project has several limitations. Although all eligible patients were enrolled consecutively, the small sample size resulted in wide CIs around estimates of diagnostic accuracy, limiting the precision of these estimates. Accordingly, the diagnostic performance findings should be interpreted as preliminary and hypothesis-generating rather than confirmatory. Because specialists were not blinded to the Canvas Dx result, diagnostic independence between the index test and reference standard was not maintained. This may have introduced incorporation bias [30] and should be considered when interpreting the diagnostic accuracy estimates. However, this workflow reflects the intended real-world clinical use of Canvas Dx as an adjunct to, rather than a replacement for specialist clinical judgment. Only 4 of the 6 clinicians participated in the survey, which limits the generalizability of health care provider experience findings. Additionally, caregiver survey completion was 56%, raising the possibility of response bias. It is plausible that caregivers with more polarized experiences may be most likely to complete the survey [31], potentially inflating or deflating satisfaction metrics. Given the uniformly positive caregiver responses observed, positivity bias cannot be excluded, and the survey findings should therefore be interpreted as descriptive of respondents rather than representative of all participants. Future studies with larger samples and higher response rates will be important to validate these findings. The study was conducted within a single implementation context, and findings may not generalize to other geographic regions, practice types, or patient populations.

Because diagnostic accuracy analyses were limited to participants receiving determinate results, overall system performance must be interpreted in light of the abstention rate. While the abstention mechanism reflects responsible AI design, future research should continue to examine the characteristics of children who receive an indeterminate output and best-practice protocols for making accurate diagnostic decisions in this group.

XSL•FO
RenderX

Conclusions

In this preliminary quality improvement project, Canvas Dx demonstrated high observed concordance with specialist diagnostic determinations among children receiving determinate results and was well accepted by both clinician and caregiver users. The device's ability to provide structured behavioral insights, incorporate naturalistic videos, and integrate into routine workflow suggests potential for meaningful clinical utility. However, given the preliminary nature of this study, these findings should be interpreted as exploratory and require confirmation in larger-scale evaluations. Collectively, these findings suggest that Canvas Dx has the potential to support routine ASD diagnostic workflows and provide a foundation for larger implementation studies. In the context of persistent specialist shortages and prolonged wait times for ASD evaluation, further evaluation of technologies that may help streamline diagnostic pathways is warranted.

Acknowledgments

The authors would like to thank the participating families and clinicians, including Dr Nicol Awadalla Bacon, Dr Cathy Scheiner, Dr Naterra Walker, Dr Britany Weissman, and Ms Lindsay Damon, for their participation in and contributions to this quality improvement project. The authors are also grateful to Dr Carmela Salomon and her team at Cognoa Inc for their support and guidance over the duration of the project.

Funding

The authors declared that no financial support was received for this work.

Data Availability

Data are not publicly available because they contain sensitive patient information. Individual deidentified participant data that underlie the results reported in this article and the study protocol may be made available from the corresponding author to qualified researchers upon reasonable request. Proposals related to the Cognoa Canvas Dx device can be directed to the chief executive officer of Cognoa (sharief@cognoa.com) to gain access. Data requesters will be required to sign a data-sharing agreement prior to access.

Generative AI Use

Generative AI (ChatGPT model GPT-6; OpenAI) was used to assist with language editing, improving the clarity of the manuscript, and preparing illustrative figures and diagrams. All study data, statistical analyses, interpretation of results, and scientific conclusions were generated and verified by the authors. The authors reviewed, edited, and take full responsibility for all manuscript content.

Authors' Contributions

All authors contributed equally to this work. All authors contributed to conceptualization, methodology, investigation, formal analysis, interpretation of the data, writing of the original draft, and review and editing of the manuscript. All authors approved the final manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

JBR and RTL have no conflicts of interest to disclose. SOT is a paid consultant for Floreo Inc and Arcadia Pharmaceuticals.

Multimedia Appendix 1

Canvas Dx complete indications for use.
[PDF File (Adobe PDF File), 51 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Provider survey instrument.
[PDF File (Adobe PDF File), 37 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Family survey instrument.
[PDF File (Adobe PDF File), 37 KB-Multimedia Appendix 3]

References

1. Maenner MJ, Warren Z, Williams AR, Amoakohene E, Bakian AV, Bilder DA, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11

XSL•FO
RenderX

sites, United States, 2020. MMWR Surveill Summ. Mar 24, 2023;72(2):1-14. [FREE Full text] [doi: 10.15585/mmwr.ss7202a1] [Medline: 36952288]
2. Gordon-Lipkin E, Foster J, Peacock G. Whittling down the wait time: exploring models to minimize the delay from initial concern to diagnosis and treatment of autism spectrum disorder. Pediatr Clin North Am. Oct 2016;63(5):851-859. [FREE Full text] [doi: 10.1016/j.pcl.2016.06.007] [Medline: 27565363]
3. Kraft C, Badesch S, Shannon J, Salomon C, Seal M, Chettiath T, et al. Wait times and processes for autism diagnostic evaluations: a first report survey of autism centers in the U.S. CMS Health Equity Conference. 2023. URL: https://www.cms.gov/files/document/wait-times-and-processes-autism-diagnostic-evaluations-first-report-survey-autism-centers-us.pdf [accessed 2026-10-13]
4. Baum RA, Berman BD, Fussell JJ, Patel R, Roizen NJ, Voigt RG, et al. Child health needs and the developmental-behavioral pediatrics workforce supply: 2020-2040. Pediatrics. Feb 01, 2024;153(Suppl 2):e2023063678H. [doi: 10.1542/peds.2023-063678H] [Medline: 38300001]
5. Okoye C, Obialo-Ibeawuchi CM, Obajeun OA, Sarwar S, Tawfik C, Waleed MS, et al. Early diagnosis of autism spectrum disorder: a review and analysis of the risks and benefits. Cureus. Aug 09, 2023;15(8):e43226. [FREE Full text] [doi: 10.7759/cureus.43226] [Medline: 37692637]
6. Zuckerman K, Lindly OJ, Chavez AE. Timeliness of autism spectrum disorder diagnosis and use of services among U.S. elementary school-aged children. Psychiatr Serv. Jan 01, 2017;68(1):33-40. [FREE Full text] [doi: 10.1176/appi.ps.201500549] [Medline: 27476809]
7. About autism spectrum disorder. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/autism/about/index.html [accessed 2026-02-24]
8. Nasir AK, Strong-Bak W, Bernard M. Diagnostic evaluation of autism spectrum disorder in pediatric primary care. J Prim Care Community Health. 2024;15:21501319241247997. [FREE Full text] [doi: 10.1177/21501319241247997] [Medline: 38650542]
9. Marsh K, Nguyen NT, Burton T, Hicks S. Reliability, accuracy, and timeliness of autism telehealth evaluations by pediatricians. Telemed J E Health. Jul 2026;32(7):739-745. [doi: 10.1177/15305627261440763] [Medline: 41954026]
10. Braddock B, Turner K, Kumarason K, Bock L, Stumpe K, Wallace M, et al. Improving access to autism evaluation: a coordinated developmental-behavioral pediatrics and pediatric neurology diagnostic pathway. Mo Med. 2024;121(3):225-230. [FREE Full text] [Medline: 38854594]
11. Chen YH, Drye M, Chen Q, Fecher M, Liu G, Guthrie W. Delay from screening to diagnosis in autism spectrum disorder: results from a large national health research network. J Pediatr. Sep 2023;260:113514. [doi: 10.1016/j.jpeds.2023.113514] [Medline: 37244580]
12. Salomon C, Heinz K, Aronson-Ramos J, Wall DP. An analysis of the real world performance of an artificial intelligence based autism diagnostic. Sci Rep. Aug 12, 2025;15(1):29503. [FREE Full text] [doi: 10.1038/s41598-025-15575-8] [Medline: 40796609]
13. 510(k) Premarket notification. US Food & Drug Administration. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K243558 [accessed 2026-02-24]
14. Megerian JT, Dey S, Melmed RD, Coury DL, Lerner M, Nicholls CJ, et al. Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. NPJ Digit Med. May 05, 2022;5(1):57. [FREE Full text] [doi: 10.1038/s41746-022-00598-6] [Medline: 35513550]
15. Taraman S, Shannon J, Liu-Mayo S, Kraft CA, Chettiath T, Salomon M, et al. 6.27 Exploring the real-world performance of an artificial intelligence-based diagnostic device for ASD: an aggregate analysis of early Canvas Dx prescription and output data. J Am Acad Child Adolesc Psychiatry. Oct 2023;62(10):S294. [doi: 10.1016/j.jaac.2023.09.434]
16. Sohl K, Linstead E, Heinz K, Lledo EE, Brewer Curran A, Mahurin M, et al. Integration of an artificial intelligence-based autism diagnostic device into the ECHO autism primary care workflow: prospective observational study. JMIR Form Res. Oct 21, 2025;9:e80733. [FREE Full text] [doi: 10.2196/80733] [Medline: 41124399]
17. Kompa B, Snoek J, Beam AL. Second opinion needed: communicating uncertainty in medical machine learning. NPJ Digit Med. Jan 05, 2021;4(1):4. [FREE Full text] [doi: 10.1038/s41746-020-00367-3] [Medline: 33402680]
18. Landsheer JA. The clinical relevance of methods for handling inconclusive medical test results: quantification of uncertainty in medical decision-making and screening. Diagnostics (Basel). May 09, 2018;8(2):32. [FREE Full text] [doi: 10.3390/diagnostics8020032] [Medline: 29747402]
19. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs. Apr 2008;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
20. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. Feb 2004;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
21. Grant SR, Lin TA, Miller AB, Mainwaring W, Espinoza AF, Jethanandani A, et al. Racial and ethnic disparities among participants in US-based phase 3 randomized cancer clinical trials. JNCI Cancer Spectr. Jul 07, 2020;4(5):pkaa060. [FREE Full text] [doi: 10.1093/jncics/pkaa060] [Medline: 33225207]

XSL•FO
RenderX

22. Aylward BS, Gal-Szabo DE, Taraman S. Racial, ethnic, and sociodemographic disparities in diagnosis of children with autism spectrum disorder. J Dev Behav Pediatr. 2021;42(8):682-689. [FREE Full text] [doi: 10.1097/DBP.0000000000000996] [Medline: 34510108]
23. Sharma Y, Saha A, Goldsack JC. Defining the dimensions of diversity to promote inclusion in the digital era of health care: a lexicon. JMIR Public Health Surveill. Feb 09, 2024;10:e51980. [FREE Full text] [doi: 10.2196/51980] [Medline: 38335013]
24. Johnson NL, Fial A, Van Hecke AV, Whitmore K, Meyer K, Pena S, et al. A scoping review of diagnosis of autism spectrum disorder in primary care. J Pediatr Health Care. 2023;37(5):519-527. [doi: 10.1016/j.pedhc.2023.04.003] [Medline: 37178094]
25. Onishchenko D, Huang Y, van Horne J, Smith PJ, Msall ME, Chattopadhyay I. Reduced false positives in autism screening via digital biomarkers inferred from deep comorbidity patterns. Sci Adv. Oct 08, 2021;7(41):eabf0354. [FREE Full text] [doi: 10.1126/sciadv.abf0354] [Medline: 34613766]
26. Lingren T, Chen P, Bochenek J, Doshi-Velez F, Manning-Courtney P, Bickel J, et al. Electronic health record based algorithm to identify patients with autism spectrum disorder. PLoS One. Jul 29, 2016;11(7):e0159621. [FREE Full text] [doi: 10.1371/journal.pone.0159621] [Medline: 27472449]
27. Chang Z, Di Martino JM, Aiello R, Baker J, Carpenter K, Compton S, et al. Computational methods to measure patterns of gaze in toddlers with autism spectrum disorder. JAMA Pediatr. Aug 01, 2021;175(8):827-836. [FREE Full text] [doi: 10.1001/jamapediatrics.2021.0530] [Medline: 33900383]
28. Davis JM, Harrington MB, Howie FR, Mohammed KS, Gunderson JA. Reducing time to diagnosis of autism spectrum disorder using an integrated community specialty care model: a retrospective study. J Pediatr. Jul 2024;270:114009. [doi: 10.1016/j.jpeds.2024.114009] [Medline: 38492915]
29. Fox E, Gerdts J, Ahlers K, Kotchick B. Age-based tracks: a method to tailor autism diagnostic evaluation in large-scale autism specialty centers. J Clin Med. Oct 27, 2022;11(21):6332. [FREE Full text] [doi: 10.3390/jcm11216332] [Medline: 36362561]
30. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. Ann Intern Med. Dec 16, 2008;149(12):889-897. [FREE Full text] [doi: 10.7326/0003-4819-149-12-200812160-00008] [Medline: 19075208]
31. Perneger TV, Peytremann-Bridevaux I, Combescure C. Patient satisfaction and survey response in 717 hospital surveys in Switzerland: a cross-sectional study. BMC Health Serv Res. Mar 02, 2020;20(1):158. [FREE Full text] [doi: 10.1186/s12913-020-5012-2] [Medline: 32122346]

Abbreviations

ASD: autism spectrum disorder
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
FDA: Food and Drug Administration
IRB: Institutional Review Board
NPS: net promoter score

Edited by L MacNeill; submitted 26.Apr.2026; peer-reviewed by R Beaumont, C Klaiman; comments to author 19.Jun.2026; revised version received 26.Sep.2026; accepted 08.Oct.2026; published 03.Nov.2026

Please cite as:
Resnik JB, Lane RT, Turnacioglu SO
Integration of an FDA-Authorized Digital Autism Diagnostic Tool Into an Outpatient Developmental Pediatrics Workflow: Preliminary Quality Improvement Study
JMIR Form Res 2026;10:e98697
URL: https://formative.jmir.org/2026/1/e98697
doi: 10.2196/98697
PMID:

©Jenna B Resnik, Rebecca T Lane, Sinan O Turnacioglu. Originally published in JMIR Formative Research (https://formative.jmir.org), 03.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX