Canvas Dx as a Diagnostic Aid for Autism in Young Children

Health Technology Assessment

November 2024

New York State Department of Health
Office of Health Insurance Programs
99 Washington Ave, One Commerce Plaza, Suite 720
Albany, NY 12210
(518) 473-2160
https://www.health.ny.gov/health_care/medicaid/ebbrac
EBBRAC@health.ny.gov

Canvas Dx as a Diagnostic Aid for Autism in Young Children

Health Technology Assessment

November 2024

Prepared by:

Center for Evidence-based Policy
Oregon Health & Science University
3030 S Moody, Suite 250
Portland, OR 97201
Phone: 503-494-2182
Fax: 503-494-3807
http://centerforevidencebasedpolicy.org/

Authors

Susan Connor, PhD; Elizabeth Brown, MD, MSHP; Laura Pavlech, DVM, MSLS; Valerie King, MD, MPH; Beth Shaw, MSc

About the Center for Evidence-based Policy

The Center for Evidence-based Policy (Center) is recognized as a national leader in evidence-based decision making and policy design. The Center understands the needs of policymakers and supports public organizations by providing reliable information to guide decisions, maximize existing resources, improve health outcomes, and reduce unnecessary costs. The Center specializes in ensuring that diverse and relevant perspectives are considered and appropriate resources are leveraged to strategically address complex policy issues with high-quality evidence and collaboration. The Center is based at Oregon Health & Science University in Portland, Oregon.

Conflict of Interest Disclosures: No authors have conflicts of interest to disclose. All authors have completed and submitted the Oregon Health & Science University form for Disclosure of Potential Conflicts of Interest, and none were reported.

Funding and Support: This research was funded by the New York State Department of Health through a contract with the Center for Evidence-based Policy at Oregon Health & Science University (Contract number C038339).

Table of Contents

Executive Summary

Background

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition that affects approximately 1 in 36 children and is often associated with impaired communication and social interaction, restricted interests, and repetitive behaviors.[1] Comprehensive assessment by a specialist with expertise in ASD, such as a developmental-behavioral pediatrician, psychologist, neurologist, or psychiatrist, is recommended for diagnosis of ASD.[2] While ASD can be reliably identified as early as 18 months of age, the median age of ASD diagnosis in the US is 49 months.[3] The delay between initial concern and official diagnosis is due in part to regional variability in access to specialist services, a limited specialty workforce leading to long wait lists for diagnostic evaluations, and cultural and language barriers in some communities.[2,4-6]

Pediatric primary care providers often cite as barriers to diagnosis a lack of training in recognizing and diagnosing ASD, discomfort with the perceived subjectivity of diagnostic criteria, and lack of knowledge and training in use of long and involved diagnostic instruments.[7,8] ASD is frequently comorbid with attention-deficit/hyperactivity disorder (ADHD), and often co-occurs with other developmental or behavioral diagnoses.[2,9,10] Pediatric primary care providers must carefully consider whether conditions other than ASD, such as ADHD, anxiety, global developmental delays, or intellectual disability, may better explain a child's behavior.[2] Emerging strategies to aid in ASD diagnosis in primary care are being developed and tested, including mobile health applications like Canvas Dx.[11]

The Cognoa ASD Diagnosis Aid, marketed as Canvas Dx by Cognoa, Inc., is a software as a medical device (SaMD) designed to assist physicians in evaluating a child with suspected ASD.[12,13] Canvas Dx is not a stand-alone diagnostic instrument.[14] As envisioned by Cognoa, when a concern for ASD or other developmental disability has been identified through initial screening, clinicians can prescribe Canvas Dx to help aid in decision making, either ruling out a diagnosis, making a diagnosis, or referring a patient to specialists for further evaluation.[14] Machine learning algorithms process the information from questionnaires filled out by a parent and a health care provider, and from videos of the child uploaded by caregivers.[12,13] Canvas Dx then reports a positive or negative diagnosis for ASD if its algorithm has sufficient information, or gives an indeterminate response if there is insufficient information to determine a diagnosis.[12,13] The US Food and Drug Administration (FDA) reviewed the Cognoa ASD Diagnosis Aid through the De Novo premarket review pathway and granted the device class II status restricted to prescription use on June 2, 2021.[14]

Key Questions

KQ1. What is the effectiveness (accuracy and utility) of Canvas Dx as a diagnostic tool for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

KQ2. What are the adverse effects of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

KQ3. What are the costs or cost-effectiveness studies related to Canvas Dx in children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

KQ4. What are clinical practice guideline recommendations for the use of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or autism spectrum disorder?

KQ5. What are relevant Medicaid program coverage policies and private payer policies for the use of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

Methods

Researchers from the Center for Evidence-based Policy (Center) searched Ovid MEDLINE, Cochrane Database of Systematic Reviews via the Cochrane Library, and other databases and information sources for diagnostic accuracy studies, randomized controlled trials, nonrandomized comparative trials, prospective cohort studies, interrupted time series with comparison groups, controlled before-after studies, cost and cost-effectiveness studies, and clinical practice guidelines. We also searched trial registries for relevant ongoing trials. We searched 10 state Medicaid program websites, 13 private payer websites, and the Centers for Medicare & Medicaid Services for local and national coverage determinations on the use of Canvas Dx to aid in the diagnosis of ASD.

Summary of Clinical Evidence and Recommendation Findings

We identified 3 publications from 3 eligible studies with diagnostic accuracy outcomes.[13,15,16] We did not identify any studies that captured information on use of Canvas Dx in a primary care setting or measured how or if Canvas Dx altered the care pathway between identification of possible developmental delay and diagnosis. Table 1 presents a summary of findings for the diagnostic ability of Canvas Dx for children with suspected ASD compared with decisions of clinical specialists. Full details of the GRADE assessment are in Appendix H.

Table 1. Summary of Findings (GRADE)

Outcome	Number of Participants and Studies	Median (Range)	Test Accuracy CoE	Rationale[a]
Sensitivity	1 diagnostic accuracy study[16] N = 425	Canvas DX had a high sensitivity when compared with a reference standard of diagnosis made by a specialist clinician, based on DSM-5 criteria and validated by 1 or more blinded reviewing specialist clinician(s) (sensitivity of 98.4%; 95% CI, 91.6% to 100%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias and 1 level for indirectness[a,b]
Specificity	1 diagnostic accuracy study[16] N = 425	78.9% (95% CI, 67.6% to 87.7%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias, 1 level for imprecision (i.e., wide confidence intervals), 1 level for indirectness.[a,b]
PPV	2 diagnostic accuracy studies[13,16] N = 1,147	Presumably using the same algorithm and an overlapping data set, 1 study (Wall et al., 2023) reports PPV at 89.7% (95% CI, 83.9% to 90.3%)[13] and the other (Megerian et al., 2022) at 80.8% (95% CI, 70.3% to 88.8%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias, 1 level for inconsistency, 1 level for indirectness, and 1 level for imprecision (large confidence interval in 1 study).
NPV	2 diagnostic studies[13,16] N = 1,147	Presumably using the same algorithm and an overlapping data set, 1 study (Wall) reports NPV at 96.1% (95% CI 93.4% to 98.6%)[13] and the other (Megerian) at 98.3% (90.6% to 100%)[16]	●○○○ VERY LOW	Downgraded 2 levels for risk of bias and 1 level for indirectness.[a]
Time to diagnosis	No evidence	--	--	--
Time to service initiation	No evidence	--	--	--

Notes. [a] *Indirectness refers to the degree to which children in the diagnostic study are representative of the broader population of children whose parents or clinicians have concern for developmental disability. Imprecision refers to the level of certainty of the true effect.*[17,18]

[b] *Unable to rate for inconsistency as only 1 study. For methods and interpretation of GRADE ratings, see Appendix C.*

Abbreviations. AUC: area under the curve; CoE: certainty of evidence; CI: confidence interval; GRADE: Grading of Recommendations, Assessment, Development, and Evaluations approach; NPV: negative predictive value; PPV: positive predictive value; ROC receiver operating characteristic.

Cost and Cost-Effectiveness

No studies exploring costs or cost-effectiveness of Canvas Dx were identified.

Clinical Practice Recommendations

We did not identify any clinical practice guidelines that specifically referenced Cognoa or Canvas Dx, or had guidance on the role of machine learning or SaMD in the ASD diagnostic process.

Key Policy Findings

We identified coverage policies related to Canvas Dx from 3 private payers: Aetna, Anthem Blue Cross and Blue Shield (formerly Empire BlueCross BlueShield), and Highmark Blue Shield of Northeastern New York.[19-21] None of the 10 states on our list of states to search for Medicaid policies had a policy for Canvas Dx. Two of the 3 private payers did not cover Canvas Dx, describing the SaMD as experimental and investigational with insufficient data to demonstrate either its efficacy or clinical utility.[19,20] Highmark Blue Shield of Northeastern New York covered Canvas Dx.[21]

Conclusions

Evidence about Canvas Dx was limited to diagnostic accuracy studies.[13,15,16] Canvas Dx's classification algorithm is sensitive to extreme cases (very high or very low risk of autism spectrum disorder) but has more difficulty in identifying less extreme cases, with relatively high sensitivity and negative predictive value, but lower specificity and positive predictive value.[13,15,16] In the study that informed Canvas Dx's FDA authorization, determinate output (either ASD positive or negative) was 31.8%.[16] In a 2023 publication that described efforts to optimize Canvas Dx through an algorithmic threshold-optimization procedure using a series of train-test validation procedures with repeated sampling of the same dataset, determinate results were increased to 66.5% without any significant change in accuracy.[13]

All identified studies were sponsored by the manufacturer of the system, Cognoa.[13,15,16] While mobile health applications are promoted as a means to decrease disparities in diagnosis and treatment initiation, only 1 of the 3 Canvas Dx diagnostic accuracy studies included in this review addressed differential performance of the application by gender, race, ethnic group, or family socioeconomic status.[16] The study was not powered to detect differences in these covariates, and confidence intervals were too wide to lead to any meaningful conclusions.[16] Cognoa research of its video assessment algorithms reported significantly worse performance on children of African American (61% accuracy) and South Asian (57% accuracy) ethnicities than on European American children (69% accuracy).[22]

The developers of Canvas Dx claim the tool provides a means to increase diagnostic rates by health care providers in primary care while relieving the strain on specialty services, with the goal of reducing age at diagnosis and initiation of interventions.[13,16,23] There is no evidence that use of Canvas Dx would increase diagnosis rates in primary care settings or lead to more rapid referral to early intervention services. Comparative studies could gain insight into acceptability and usability of the digital health technology in real-world settings, shed light on Canvas Dx's ability to affect time to diagnosis or to ASD-specific service initiation, determine whether its use can eliminate or substantially decrease the need for specialist diagnostic services, and gauge the tool's ability to create meaningful change in clinical management of the diagnostic and referral processes.

Background

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition often associated with impaired communication and social interaction, restricted interests, and repetitive behaviors.[24] According to the Centers for Disease Control and Prevention (CDC), approximately 1 in 36 children in the US has ASD.[1] American Academy of Pediatrics (AAP) guidelines recommend screening for ASD during primary care visits at 18 months and 24 months of age.[2] Results of a screening test are not diagnostic but can help identify children at risk for a diagnosis of ASD and who may require additional evaluation.[2] Comprehensive assessment by a specialist with expertise in ASD, such as a developmental-behavioral pediatrician, psychologist, neurologist, or psychiatrist, is recommended for diagnosis of ASD.[2] The assessment includes individual child history, physical examination, diagnostic tools, and ancillary testing (e.g., speech, language, and communication assessment; intelligence testing; sensorimotor or occupational therapy evaluation).[2] The members of the multidisciplinary team use diagnostic tools in conjunction with clinical judgment and input from the child's caregivers.[2] The AAP cautions that diagnostic tools should not be used in isolation.[25,26]

National guidelines recommend these tools[2,27,28]:

- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule, Second edition (ADOS-2)
- Childhood Autism Rating Scale, Second edition (CARS-2)
- Gilliam Autism Rating Scale, Third edition (GARS-3)
- Social Responsiveness Scale, Second edition (SRS-2)

With the 2013 release of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), definitions of autistic disorder and Asperger syndrome were merged with childhood disintegrative disorder and pervasive developmental disorder, not otherwise specified, which created greater heterogeneity in the ASD spectrum and contributed to the increase in ASD diagnoses over the past decade.[29] Autistic traits such as language deficits or repetitive behaviors, and the social deficits associated with ASD are common across developmental disorders and also present along a continuum in the general population.[30] As a result, drawing the line between typical and atypical development is challenging for specialists working with individuals who are neurodiverse, and even more so for clinicians without particular specialization or training in ASD and other developmental disorders.[31]

The differential diagnosis for children with suspicion of ASD is broad, which makes the diagnostic process complex.[2,9,10] ASD is often comorbid with attention-deficit/hyperactivity disorder (ADHD) and both have numerous associated symptom patterns, such as sensory integration, that are not part of diagnostic criteria.[9,10] ASD frequently co-occurs with other developmental or behavioral diagnoses, such as anxiety or mood disorders.[2] Multiple other conditions diagnosed in childhood may include social communication impairment and repetitive movements (e.g., global delay, tic disorder), and clinicians must assess whether any of those conditions better explain the child's behavior.[2,32] The diagnostic process is further affected by parental attitudes and expectations, both positive and negative.[33-38] In studies of caregivers' views and experiences with the developmental evaluation process, parents are eager for a sense of certainty and access to early intervention services,[34,38] but struggle with denial, shame, guilt, and fear of stigma that could make them resistant to diagnosis.[39,40]

Although most children will need to see a developmental-behavioral or neurodevelopmental pediatrician, psychologist, or child and adolescent psychiatrist for a diagnostic evaluation, pediatric primary care clinicians who are comfortable with application of the DSM-5 criteria can make an initial clinical diagnosis,[2] which can start appropriate services. An accurate diagnosis in general pediatrics requires more information than can be gained in a short clinical visit.[2] Longitudinal history can be obtained through family interviews and by use of questionnaires such as the Social Communication Questionnaire or Social Responsiveness Scale, but the AAP cautions questionnaires alone are not enough to make a diagnosis of ASD.[2] The AAP further cautions that clinician training in diagnosis of autism and related conditions beyond the basic overview in residency programs is essential for diagnosis in a pediatric primary care setting.[2]

Importance of Early Diagnosis

ASD can be reliably identified as early as 18 months of age, although children with milder symptoms may not be identified until school age.[2] Diagnostic stability is high for children diagnosed with ASD at 18 to 36 months of age, with greater than 80% of diagnoses confirmed a year or more later.[2,41] Although most parents of children with an ASD diagnosis first reported concerns from 12 to 18 months of age,[42,43] CDC analysis of 2020 data from the Autism and Developmental Disabilities Monitoring (ADDM) Network found the median age of earliest known ASD diagnosis in the US was 49 months across 11 states analyzed, ranging from 36 months in California to 59 months in Minnesota (New York was not in the network).[3]

The first years of a life are a time of substantial neuroplasticity and a vital stage in synapse formation and refinement.[44] Early intensive intervention is associated with significant effects on social communication, adaptive skills, and cognitive abilities,[45-48] although systematic reviews noted small sample sizes, lack of blinding, limited follow-up data, and incomplete outcome data introduce high risk of bias.[46,49] There is evidence the effects of early intervention are sustained over time,[50-52] although wide variation across studies in participants' phenotypic profiles and ages, and setting, timing, frequency, and delivery of interventions contribute to inconsistent findings.[53] While lack of long-term follow-up data constrains the ability to conduct cost-effectiveness analysis, evidence suggests children who receive early intensive intervention may require fewer services later on, resulting in long-term cost savings.[54-56] The reported benefits of early intervention led efforts to lower the age of diagnosis and spurred development of innovative models for early identification,[44] although barriers to high-quality early intervention services, particularly for children from marginalized groups, can limit the value of early identification.[2,57]

ASD Prevalence and Access to Services in New York State

In New York state, 58,655 children aged 4 to 21 years were receiving special education services for ASD in the 2022 school year, representing 12% of all children receiving special education services in the state.[58] This represents a 168% change from 2010, before the release of DSM-5, when 22,284 New York children were receiving special education services for autism.[58] These numbers likely represent an underestimation of ASD prevalence in the state, as not all children with an ASD diagnosis access special education services and specific numbers for younger children (less than 4 years of age) are not given.[59] As with other research over the past decade,[60] rates of ASD in New York were highest among school districts with higher proportions of Black

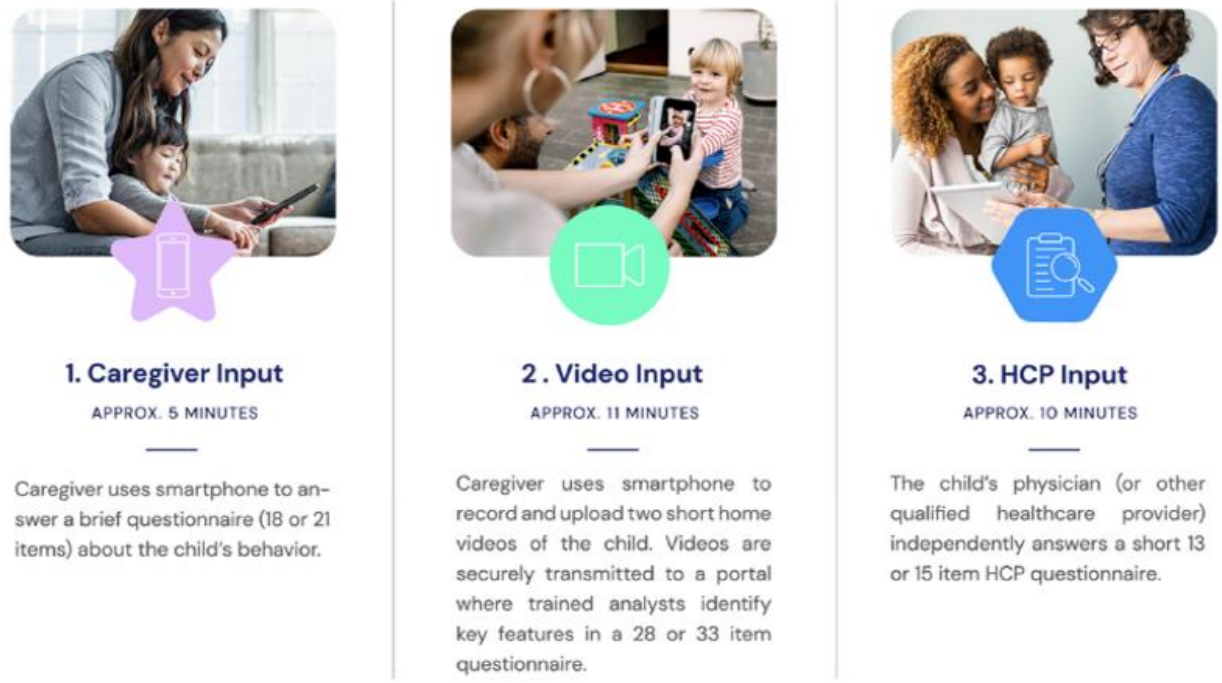
and Hispanic students, with significant clustering of school districts with high ASD rates in New York City and Albany.[61]

Residents of urban areas such as New York City, Rochester, Buffalo, Albany, and Syracuse are most likely to have diagnostic resources (developmental pediatricians, pediatric neurologists, child and adolescent psychiatrists, or a specialized autism center or autism evaluation center) available within 25 miles, with the greatest concentration of resources in New York City.[62] Families not residing in an urban area travel longer distances to access resources.[62] A resident of Potsdam, for example, has to travel 28 miles to reach the closest specialist in New York state (developmental pediatrician, pediatric neurologist, child and adolescent psychiatrist) and 110 miles to reach the nearest specialized autism center, while a resident of Jamestown has to travel more than 50 miles to reach the nearest specialist or autism center in the state.[62]

Canvas Dx

The Cognoa ASD Diagnosis Aid, marketed as Canvas Dx by Cognoa, Inc., is a software as a medical device (SaMD) designed to assist qualified health care providers in evaluating children with suspected ASD (Figure 1).[12,13] After a physician prescribes Canvas Dx, caregivers download a mobile app to answer questions about behavior problems and upload videos of their child.[12,13] Manufacturer-trained viewers analyze uploaded videos of patients through a video portal.[12,13] Another portal for health care providers lets the prescriber enter answers to preloaded questions about behavior problems, track the information supplied by parents or caregivers, and review a report of the results.[12,13] Machine-learning algorithms process the information from the parent questionnaire, videos, and health care provider questionnaire, then Canvas Dx reports a positive or negative ASD result if there is sufficient information to make a diagnosis.[13] Canvas Dx will report that no result can be generated if there is insufficient information.[13]

Figure 1. Components of Cognoa's Canvas Dx[2]



Source: Wall et al., 2023[13]

The parent and provider questionnaires used in Canvas Dx are based on research using artificial intelligence (AI) to shorten lengthy standard diagnostic questionnaires into what the algorithm judged to be key questions. Canvas Dx questionnaires are informed by both the ADOS-2[63-66] and ADI-R.[67,68] There is some criticism of the Cognoa teams' methodology and results from the earlier versions of the app.[69,70] A 2016 paper[70] reports on an unsuccessful attempt to reproduce results from 2 early studies led by Dennis Wall, founder of Cognoa, using larger datasets.[66,68]

Parents are asked to upload 2 short videos (1.5 to 5 minutes in length) of their child from a mobile device.[13,67] The parents are told to show the child's face and hands and opportunities for social engagement.[13,67] For review and scoring of parent-uploaded videos, Cognoa recruits nonexpert workers from paid platforms (such as Amazon Mechanical Turk and Microworkers.com) who receive minimal training before rating videos.[71-73] The crowdsourced workers answer a series of multiple choice questions about the child's behavior in each video, and a machine-learning classifier then predicts the diagnosis from the viewer responses.[71-73] Cognoa-sponsored research states the use of crowdsourcing for annotation of behavioral features is reliable,[63,67,71-76] although clinicians had higher sensitivity in their recognition of autism-related symptoms and were less likely to miss positive cases.[71] In a test of 30 videos of children with ASD and 30 neurotypical controls, both sensitivity and specificity of crowdsourced workers ranged from 80% to 90%, depending on level of blurring or pixilation used to preserve patient privacy.[71] No independent analyses of accuracy of crowdsourcing for rating of behavioral elements in home videos or other tasks related to autism diagnosis were identified. A systematic review by Wazny (2018) that examined use of crowdsourcing for health-related applications (which did not included Canvas Dx) found diagnosis was the most common use of crowdsourcing, primarily the rating or grading of laboratory samples.[77] Wazny found mixed results in the accuracy of crowdsourced workers, with better performance on easier and more straightforward cases, but lower accuracy on more complex cases.[77] Wazny also found gamification of tasks appeared to improve accuracy.[77]

FDA Authorization of Canvas Dx

The Food and Drug Administration (FDA) reviewed the Cognoa ASD Diagnosis Aid through the De Novo premarket review pathway, a regulatory pathway for low- (class I) to moderate- (class II) risk devices of a new type, and granted the device class II status on June 2, 2021.[78] The FDA's De Novo classification is an alternate pathway to classify novel medical devices without a substantial equivalent and which is outside existing product codes.[79] Devices determined to be class I or class II through a De Novo risk-based classification request may be marketed without submitting a 510(k) application and can also be used as predicates for future 510(k) submissions.[79] A De Novo application proposes a product risk classification (class I or class II), shares a written benefit and risk analysis starting with the clinical benefits of the device, and recommends special controls for the new product code based upon the risks to health and the mitigation measures for each risk.[79]

The FDA's De Novo application decision restricted the Cognoa ASD Diagnosis Aid to prescription use not intended as a stand-alone diagnostic device.[78] Rather, the Cognoa ASD Diagnostic Aid is an adjunct to the diagnostic process.[78] Acceptance of the De Novo application was based on a single clinical validation study by Megerian and colleagues in 2022.[16,78] The validation study was a prospective, double-blinded, single-arm study conducted at 14 sites in the

US that used data from 425 participants to evaluate the safety and diagnostic accuracy of the device to aid in the diagnosis of ASD, with the comparator being the clinical reference standard.[16,78] As part of the benefit and risk analysis in the De Novo application, the FDA cautions the device may give unreliable results if used in patients with other conditions that would have been excluded from the clinical study by Megerian and colleagues in 2022,[16] such as known deafness or blindness, major dysmorphic features (e.g., fetal alcohol syndrome), diagnosis of genetic conditions (i.e., Rhett syndrome, fragile X), and diagnosis of epilepsy.[78] Cognoa's De Novo application identified a risk of misdiagnosis and delayed diagnosis of ASD associated with use of the Canvas DX app, which can result in delayed treatment of ASD and delivery of treatment not appropriate for ASD; however, the FDA decision determined that benefits outweighed risks.[78]

Clinical Evidence Required by the De Novo and 510(k) Pathways

The De Novo classification process was created as part of the FDA Modernization Act of 1997 and is now the preferred regulatory pathway for novel diagnostics without predicate devices.[80] The De Novo pathway can promote innovation and remove a regulatory burden for software as a medical device (SaMD) with low to moderate risk, increasing the availability of up-to-date device predicates upon which subsequent 510(k) device applications can be based,[80] although questions about the amount of clinical evidence required in support of De Novo applications remain.[81-84]

Johnson and colleagues (2020) conducted a retrospective cross-sectional analysis of 63 devices cleared through the De Novo pathway between January 1, 2011, and December 31, 2019, to explore the strength of clinical evidence supporting FDA clearance through this pathway.[82] They found 43% of devices were cleared without supporting premarket pivotal studies or based on pivotal studies that failed to meet at least 1 primary effectiveness endpoint.[82] The FDA required postmarket study in only 1 case. In 32 cases (47.8%), devices cleared through the De Novo pathway served as the basis for new models and competitor products subsequently cleared via the 510(k) process.[82]

By using the De Novo pathway, Canvas Dx received FDA authorization on the basis of a single diagnostic efficacy study judging the performance of the underlying algorithm that did not measure use, harms, or risks in clinical diagnostic settings.[16] FDA judgment on balance of benefits and risks is based on the risk and benefit analysis reported by Cognoa in its application. Postmarket studies were not required for Canvas Dx and seldom are required for SaMD authorized through the De Novo pathway.[82] Subsequent evolutions of the product or similar products from competitors can now be cleared through the easier and faster 510(k) pathway.[80]

Coverage Policy Issues

At present, there are no federal requirements for coverage of digital therapeutics or digital diagnostics by Medicaid, although multiple bills have been put before Congress.[85-88] The most recent, the Access to Prescription Digital Therapeutics Act of 2023 (HR 1458 and S 723), was introduced in March 2023.[86,87] The bill proposes Medicare and Medicaid coverage of prescription digital therapeutics and would require the Centers for Medicare & Medicaid Services (CMS) to establish a payment methodology for prescription digital therapeutics, assign product-specific Healthcare Common Procedure Coding System (HCPCS) codes for these products, and create a

system for manufacturer reporting. The latest action occurred on September 19, 2023, when HR 1458 was on the agenda for a subcommittee meeting.[86]

The first HCPCS code for a prescription digital therapeutic became effective in April 2022.[89] This Level II code, A9291, was created as a request from the manufacturer of reSET, a prescription digital therapeutic for the management of substance use disorder. The code describes "prescription digital behavioral therapy, fda [*sic*] cleared, per course of treatment."[89(p46)] This language was later modified to "prescription digital cognitive and/or behavioral therapy."[90(p75)] At the public meeting where this new Current Procedural Terminology (CPT) code was discussed, the primary speaker for the application noted the American Medical Association included digital cognitive behavioral therapy in the existing codes for remote therapeutic monitoring.[89] The speaker observed this CPT code change did not create a billing pathway for prescription digital therapeutics because the cognitive behavioral therapy remote therapeutic monitoring supply code is for FDA-cleared devices prescribed by a physician who has incurred a cost, taken ownership or title of the device, and billed for the device, qualifications that did not apply to prescription digital therapeutics.[89] After code A9291 was established, 2 additional HCPCS Level II codes for prescription digital therapeutics have been created: A9292, for a prescription digital therapeutic to improve visual acuity in patients with amblyopia (October 2023),[90] and A9293, for a prescription digital therapeutic to monitor fertility (April 2024).[91]

An application to establish a new HCPCS Level II code for Canvas Dx was discussed at CMS's HCPCS public meeting in November 2022.[92] The proposed language for this new code was "prescription digital diagnostic device for neurodevelopmental/behavioral disorders, FDA authorized, per diagnostic assessment."[92(p26)] The applicant, Cognoa, Inc., noted the existing HCPCS Level II code A9291 was not appropriate for Canvas Dx because it was a diagnostic product used at a single point in time, not a therapeutic product for ongoing or episodic management of a chronic condition.[92] CMS decided not to create a new HCPCS Level II code for Canvas Dx because the product typically would be used in a procedure reported with a HCPCS Level I (CPT) code. The average wholesale price for Canvas Dx is $2,280 per evaluation.[93]

Contextual Questions

CQ1. *What is the place of Canvas Dx as a diagnostic tool in clinical care for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?*

The AAP recommends routine autism-specific screening at 18- and 24-month well-child visits,[2] but primary care providers may not strictly adhere to these guidelines.[44] Some may screen at one visit but not the other; others screen at both ages but rely on their clinical judgment rather than using validated screening tools, which are more accurate than clinical judgment alone.[44] Still others screen only when a concern is raised by the parent.[44,94] In a study of 94 primary care providers across 13 states, chart review revealed 51% administered autism screening tools at all 18-month well-child visits, while 41% screened at all 24-month visits.[95]

Screening for and diagnosis of ASD is covered under Early and Periodic Screening, Diagnostic, and Treatment (EPSDT), the Medicaid program benefit for children and adolescents younger than 21 years of age.[96] In New York state, EPSDT is implemented through the Child/Teen Health Program.[97] The Child/Teen Health Program follows the AAP recommendations for developmental screening of children at 9, 18, and 30 months of age, and ASD screening at 18

and 24 months of age.[97] While the AAP recommends children be screened for developmental delay and ASD with a standardized tool, it does not endorse or approve specific tools for screening.[98] In New York state, licensed physicians, including psychiatrists and developmental and behavioral pediatricians, and licensed psychologists can diagnose ASD.[99]

Under New York State Public Health Law (Section 2500-J), the Commissioner of Health is required to establish best practice protocols for early screening of children for ASD.[100] This requirement is fulfilled in the Best Practice Protocol for Early Screening of Young Children for Autism Spectrum Disorders (ASDs) by Pediatric Primary Care Providers,[101] which follows recommendations from the AAP and the New York State Department of Health Clinical Practice Guideline on Assessment and Intervention Services for Young Children with Autism Spectrum Disorders.[99,102] This protocol and guideline describe 2 levels of screening tools for ASD: Level 1 tools are autism-specific and can be used to screen all children for ASD, and Level 2 tools are used to screen children for whom there is already a concern for ASD.[99,101] Level 2 tools generally have more items and require specialized training to administer.[101] The protocol strongly recommends use of Level 1 tools, specifically the Modified Checklist for Autism in Toddlers-Revised with Follow-up (M-CHAT-R/F), Infant-Toddler Checklist (ITC), and Parent Observation of Social Interactions (POSI), in pediatric primary care settings.[101] The guideline recommends the use of the Level 2 tools: Screening Tool for Autism in Toddlers and Young Children (STAT), Autism Detection in Early Childhood (ADEC), and Parent Observation of Early Milestones (POEMS).[99] The New York State Department of Health Bureau of Early Intervention keeps a list of approved developmental assessment instruments.[103]

When a primary care provider or other early childhood professional suspects a child younger than 3 years of age might have a developmental delay or ASD, they must refer the child to the Early Intervention Program (EIP) for further evaluation.[99] The EIP is part of the Individuals with Disabilities Education Act (IDEA), a federal law that guarantees an appropriate public education for children and adolescents aged 3 to 21 years. EIPs provide therapeutic and support services to children younger than 3 years of age with disabilities or developmental delay and their families.[99] In New York state, EIP was established in Article 25 of the Public Health Law and through Subpart 69-4 of the Department of Health's administrative rules and regulations (Title 10, Chapter II).[104,105] When a primary care provider or other early childhood professional suspects a child aged 3 to 5 years might have a developmental delay or disability that could affect school performance, they should refer the child to the school district for evaluation by the Committee on Preschool Special Education (CPSE).[97]

New York State Medicaid reimburses for developmental screening of children younger than 3 years of age in addition to payment for an Evaluation and Management service such as a well-child exam.[106] In a child's first 3 years of life, primary care providers may be reimbursed up to 2 times for using a validated screening tool for ASD, and 1 time per year for using a validated screening tool for global development.[106] Providers are to use the CPT code 96110 (developmental screening, with interpretation and report, per standardized form) and the International Classification of Diseases Tenth Revision Clinical Modification (ICD-10-CM) codes Z13.41 (encounter for autism screening) or Z13.42 (encounter for screening for global developmental delays-milestones) when billing for these services.[106] The nonfacility and facility global fees for CPT code 96110 were listed as $15.76.[107]

Families in the US may wait as long as 18 months between initial screening of their child by a primary care clinician and diagnosis by a specialist.[108] One proposed solution to expedite diagnosis is to equip primary care providers with the tools and expertise needed to diagnose ASD within their practices.[109] In a review of National Survey of Children's Health (NSCH) data from 2016 through 2019, an average of only 13% of children were first diagnosed by their primary care provider as having ASD.[109] Children whose autism was first diagnosed by a primary care provider were approximately 1 year younger at the time of diagnosis than those whose autism was first diagnosed in nonprimary care settings, highlighting the important role of pediatric providers in increasing early diagnosis and opportunities for early intervention.[109] While the AAP encourages primary care clinicians to make the diagnosis based on DSM-5 criteria when they feel comfortable doing so,[2] a study of screening practices at 18- and 24-month well-child visits in a hospital-owned network of 290 pediatricians within 54 primary care pediatric practices in Texas found screening rates exceeded 90% at 18 months, but primary care pediatricians made a diagnosis in only 1% of cases in which a child was at risk on 1 or both (18 and 24 months) M-CHAT (Modified Checklist for Autism in Toddlers) screenings.[110]

Use of Canvas Dx in pediatric practice does not solve the bottleneck for diagnostic or therapeutic specialist services, but may have a role in increasing diagnosis at the primary care level.[108] In the protocol for a Cognoa-sponsored observational study exploring feasibility of Canvas Dx use in a primary care setting, Sohl and colleagues (2022) posit the tool can help primary care physicians who may not have ASD-specific training, lack self-efficacy in making the diagnosis, or do not have the time or skills necessary to use gold standard ASD diagnostic instruments.[111] The diagnostic instruments are time-intensive, must be done in-person, and require specialized training for administration.[111] The tool may increase primary care capacity to diagnose or rule out ASD; however, it is unknown how many primary care providers will still refer patients to specialty clinics for confirmatory diagnosis even if a positive result is produced by Canvas Dx.[111,112] An inconclusive result would require additional decision making from a primary care provider or referral to a specialist.[111]

In a podcast interview describing the development of Canvas Dx and its potential role in increasing early diagnosis, Cognoa founder Dennis Wall argues that even if it yields an inconclusive result, the pediatrician receives a detailed report to make a decision themselves, or refer the child for further testing.[23] Canvas Dx would only be successful in reducing time to diagnosis if the primary care clinician made the diagnosis, since a family referred for further testing would still face the barriers of long wait times and limited access to diagnostic services.[108]

In a 2022 commentary on emerging approaches to address barriers to ASD diagnosis in primary care settings, Wieckowski and colleagues identify Canvas Dx as a mobile health application with the potential to be applied to ASD diagnosis, although the authors make no judgments or claims regarding its efficacy and do not explicitly recommend use of Canvas Dx or any of the other tools or strategies reviewed.[11]

A prospective observational study (NCT05223374) completed in February 2024 that has not yet published results may shed some light on the place of Cognoa in the diagnostic pathway, although the methodology will limit applicability of results for primary care practices.[111] The Cognoa-funded study tested use of the Canvas Dx diagnostic tool in practices taking part in the Extension for Community Health Outcomes (ECHO) autism primary care training model.[111] The

ECHO Autism project trains clinicians to screen, diagnose, and care for children with ASD in primary care settings.[113] The ECHO Autism project is an adaptation of the larger Project ECHO movement to bring best practices in health care to underserved patients by creating information exchanges that use telemonitoring to share knowledge between specialists (the hub) and outlying primary care providers (the spokes).[113]

ECHO Autism clinicians from participating primary care practices in Missouri were assigned to dispense routine clinical care and conduct best practice ECHO Autism diagnostic evaluations with the addition of Canvas Dx device prescriptions for patients 18 to 72 months of age suspected of having ASD.[111] The primary objective was to assess the time from the initial concern by the ECHO Autism clinician to diagnosis when using the device as part of the diagnostic pathway.[111] Secondary outcomes included time to initiating therapy following a positive ASD diagnosis, caregiver and clinician satisfaction with the diagnostic process, and assessment of insurance coverage or reimbursement.[111]

While the study results will give some insight on the usability and performance of Canvas Dx in clinical practice, it involves clinicians who already have specialized training and mentorship in ASD diagnosis and treatment, more so than the typical pediatric primary care provider (particularly those in under-resourced areas).[111] As a result, the study will not have information on how the Canvas Dx tool might be integrated into practice by clinicians without such specialized training, who have not been instructed to prescribe the Canvas Dx tool for screening, or have not been specifically assigned to provide diagnostic workups in primary care.[2,111,112] Additionally, lack of a control group or comparative cohort will limit conclusions drawn regarding any relationship between Canvas Dx and time to diagnosis or to initiate therapy.[111]

CQ2. *What are the health equity considerations for Canvas Dx or other SaMD with machine-learning algorithms to diagnose children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?*

While most experts believe the expected harms of digital health technologies like Canvas Dx are limited, these may include data issues (e.g., not keeping patient data secure or selling patient data without notification),[114-117] offering inaccurate or ineffective information or treatment,[114,118] problems related to software defects,[115,119,120] lack of transparency in AI algorithms,[83] and concerns about patient consent and autonomy.[83] Researchers raised additional health equity concerns for digital health technologies, such as whether product content is available in multiple languages and is culturally appropriate for diverse populations,[117,121-124] if there will be equitable access to internet services and required devices,[117,122,124,125] and if the technologies have been adequately studied in diverse populations.[126] Commentators also note potential for conflicts of interest in research on commercialized products.[127] Because enthusiasm for use of digital technologies may outpace the evidence, there is also potential for harm due to reduced use of existing evidence-based practices.[127]

Biases incorporated into machine-learning algorithms and unequal access to AI technologies in health care have the potential to increase health disparities.[83,128] AI-driven SaMD uses machine learning based on a large amount of data to predict a specific outcome.[128] Because AI is only as good as the data on which it is trained, AI-driven SaMD could amplify health care bias and exacerbate health disparities.[128,129] These biases may result from underrepresentation of minority groups, unbalanced gender enrollment, variable amounts of missing or inaccurate data, or

inadequate consideration of potential confounders.[128,129] The encoded biases can lead to less accurate or less beneficial predictions for medically underserved patients, perpetuating disparities due to gender, race, or ethnicity.[128,130] Systemic factors such as clinical bias, geographic barriers, and insurance coverage may limit patients' access to SaMD.[131]

Because many SaMD use the FDA's De Novo classification process or 510(k) application pathway and are considered to be of low-to-moderate risk, these devices are not required to undergo the kinds of premarket testing that could detect population-specific differences in performance.[82,83] Analyzation by demographic subgroups often lacks power to produce clinically significant insights.[132] While the FDA recommends AI in health care be regulated to ensure equity, no standards exist to accomplish this and current FDA regulations do not address bias in AI SaMD outputs.[128] Additionally, the potential of SaMD to address barriers to entry to diagnostic and therapeutic pathways can only be maximized if app developers assess the usability, acceptability, and validity of these tools in underserved communities and diverse populations.[133]

CQ3. *What are the recommendations or guidelines from national organizations about using Canvas Dx or other SaMD that rely on machine-learning algorithms to diagnose developmental or health conditions?*

A search for guidelines and recommendation statements from the US Preventive Services Task Force (USPSTF), AAP, ECRI Guideline Trust, OVID Medline, and Google Scholar identified no professional guidelines or recommendation statements that addressed Canvas Dx or prescription digital diagnostic aids for ASD in general.

American Academy of Pediatrics

The AAP recommends screening all children for symptoms of ASD through a combination of developmental surveillance at all visits and standardized autism-specific screening tests at 18 and 24 months of age in their primary care visits.[2] For children identified as at-risk for a diagnosis of ASD, primary care providers should give a timely referral to a specialist for clinical diagnostic evaluation, although general pediatricians and child psychologists trained in autism diagnosis and comfortable with application of the DSM-5 criteria can make an initial clinical diagnosis and refer to early intervention services.[2]

While the developers of Canvas Dx identify the need to more quickly refer children with an ASD diagnosis to early intervention services,[13] the AAP notes a definitive diagnosis is not necessary to institute services for documented delays through early intervention or school services.[2] Children should be referred for intervention for all developmental delays at the time of identification and not wait for an ASD diagnostic evaluation.[2] The AAP does not approve or endorse any specific screening tools.[2] AAP recommendations do not offer any guidance on use of mobile applications, machine-learning algorithms, SaMD, or digital health technologies in the diagnostic process.[2]

USPSTF

The USPSTF published recommendations for ASD screening in young children in 2016,[134] and an update is underway.[135] The USPSTF in 2016 concluded the current evidence was insufficient to assess the balance of benefits and harms of screening for ASD in children between 18 and 30 months of age for whom no concerns of ASD have been raised by their parents or a clinician.[134]

However, the USPSTF recommendations do not address the diagnostic process and make no reference to SaMD or digital health diagnostic tools.[134]

Methods

This review is based on key questions (KQs) and contextual questions (CQs) identified by the New York State Department of Health. Search parameters, KQs, CQs, and methodologies for identifying, assessing, and reporting evidence are described below.

Key Questions

These key questions are addressed in the clinical evidence review and payer policies section:

KQ1. What is the effectiveness (accuracy and utility) of Canvas Dx as a diagnostic tool for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

   a. Does effectiveness vary by patient characteristics (e.g., age, sex), characteristics or symptoms of developmental delay, co-occurring conditions (e.g., comorbid diagnoses), provider characteristics, or setting?

KQ2. What are the adverse effects of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

   a. Do the adverse effects vary by patient characteristics (e.g., age, sex), provider characteristics, symptoms of developmental delay, or setting?

KQ3. What are the costs or cost-effectiveness studies about Canvas Dx in children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

KQ4. What are clinical practice guideline recommendations for the use of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

KQ5. What are relevant Medicaid program coverage policies and private payer policies for the use of Canvas Dx for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

Contextual Questions

Information to answer the following contextual questions is summarized in the Background section:

CQ1. What is the place of Canvas Dx as a diagnostic tool in clinical care for children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

CQ2. What are the health equity considerations for Canvas Dx or other SaMD with machine-learning algorithms to diagnose children aged 18 to 72 months whose caregivers or health care providers suspect developmental delay or ASD?

CQ3. What are the recommendations or guidelines from national organizations about use of Canvas Dx or other SaMD that use machine-learning algorithms to diagnose developmental or health conditions?

Eligible Studies

Table 2 summarizes the study inclusion and exclusion criteria.

Table 2. Key Study Inclusion and Exclusion Criteria

Study components	Inclusion	Exclusion
Population	Children between the ages of 18 and 72 months with suspected developmental delay	Children younger than 18 months of age or older than 72 months of age Children without concern for developmental delay
Intervention	Canvas Dx, as an adjunct to the diagnostic process	Any other diagnostic tool or process not using Canvas Dx
Comparators	Diagnostic process with evaluation by a clinician, alone or in conjunction with other professionals (e.g., speech and language pathologist, educator) that includes history, physical examination, diagnostic tools, and ancillary testing, but not Canvas Dx	Other SaMD intended to diagnose ASD in young children
Outcomes	Critical outcomes: • Sensitivity, specificity, positive predictive value, and negative predictive value; time to diagnosis; and time to ASD-specific service initiation Important outcomes: • Change in clinical management of diagnostic and referral processes	Other outcomes

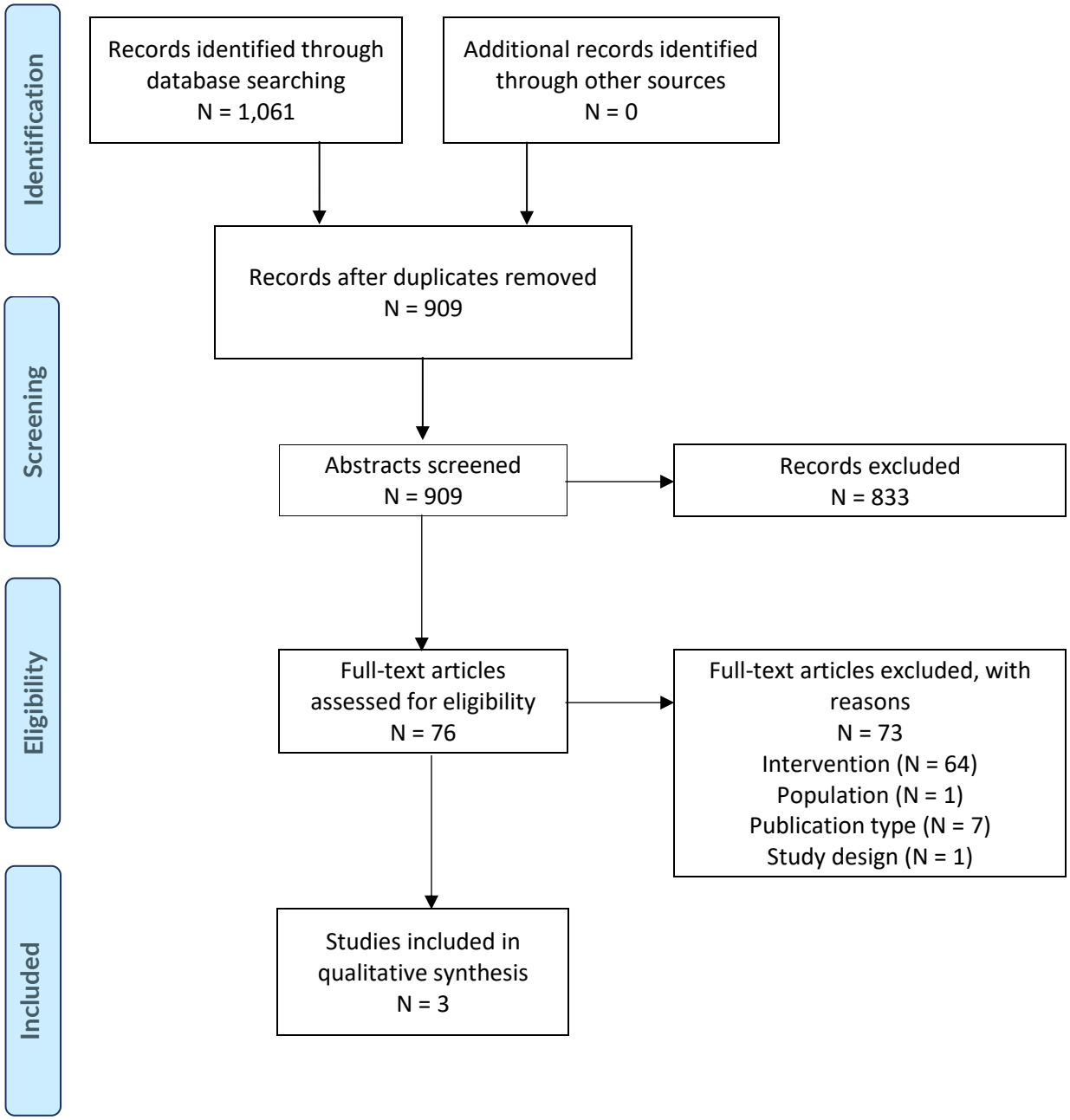
Abbreviations. ASD: autism spectrum disorder; SaMD: software as a medical device.

Evidence and Policy Searches

Researchers from the Center for Evidence-based Policy (Center) searched Ovid MEDLINE, Cochrane Database of Systematic Reviews via the Cochrane Library, and other databases and information sources for diagnostic accuracy studies, randomized controlled trials, nonrandomized comparative trials, prospective cohort studies, interrupted time series with comparison groups, controlled before-after studies, cost and cost-effectiveness studies, and clinical practice guidelines. We identified 909 potentially relevant publications for the key questions for clinical evidence and clinical practice guidelines (Figure 2). We also searched trial registries for relevant ongoing trials. A full list of sources we searched and the search strategies are listed in Appendix A. We did not conduct systematic searches to identify publications to answer contextual questions.

We also searched 10 state Medicaid program websites, 13 private payer websites, and the Centers for Medicare & Medicaid Services for local and national coverage determinations on the use of Canvas Dx to aid in the diagnosis of ASD. Three relevant coverage policies were identified. Appendix A lists search terms used to identify relevant policies.

Figure 2. PRISMA Diagram With Details



Screening and Inclusion

Two researchers used the systematic review software platform DistillerSR to screen publications identified in the searches through the inclusion and exclusion criteria listed in Appendix B. Disagreement about inclusion was resolved through discussion. Appendix F lists included studies, and Appendix J lists studies excluded during full text screening with the primary reason for exclusion. Figure 2 shows the numbers of studies screened and included or excluded at each step.

Risk-of-Bias Assessment

Two researchers assessed each included diagnostic accuracy study for risk of bias using standard forms. Appendix C has detailed tables with criteria considered for assessing risk of bias or methodological quality. Disagreement between the researchers was resolved through discussion.

Data Abstraction

One researcher used a standard form to extract all data presented in this report, and a second researcher checked each data point against the information in the publication from which it was abstracted to ensure accuracy.

Synthesis

We report a narrative synthesis of the clinical evidence and policy findings. We applied the GRADE system (Appendix C) to rate the overall quality of evidence for key outcomes and test performance.[17,18] The GRADE system[17,18] defines the overall quality of a body of evidence for an outcome in the following manner:

- **High**: Raters are very confident the estimate of the efficacy of the device studied for this outcome measure lies close to the true effect. Typical sets of studies are RCTs with few or no limitations, and the effect estimate is likely stable.
- **Moderate**: Raters are moderately confident in the estimate of the efficacy of the device studied for this outcome measure. The true effect is likely to be close to the estimate, but there is a possibility that it is different. Typical sets of studies include RCTs with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.
- **Low**: Raters have little confidence in the estimate of the efficacy of the device for this outcome measure. The true value of the outcome measure may be substantially different from the estimate. Typical sets of studies include RCTs with serious limitations or nonrandomized studies without special strengths.
- **Very low**: Raters have no confidence in the estimate of the efficacy of the device studied for this outcome measure. The true value is likely to be substantially different from the estimate. Typical sets of studies include nonrandomized studies with serious limitations or inconsistent results across studies.
- **Not applicable**: Researchers did not identify any eligible articles.

Evidence Summary

We identified 3 publications from 3 eligible studies with diagnostic accuracy outcomes. We did not identify any relevant ongoing clinical trials, diagnostic accuracy studies, or clinical practice guidelines that referenced Cognoa or Canvas Dx or discussed the role of machine learning or SaMD in the ASD diagnostic process. The clinical evidence review is organized by key question.

Additionally, we identified 20 articles published between 2012 and 2023 regarding development and testing of individual components of what would become the Canvas Dx program. These publications are not tabled in the evidence review because not all 3 components of the product authorized by the FDA were included: parent survey, video assessment, and health care provider assessment. Details regarding these pre-Canvas articles are in Appendix E.

Measures of Test Performance

Diagnostic accuracy studies, which traditionally use a cohort or cross-sectional study design, assess the ability of a diagnostic test to correctly classify a patient as having a condition, such as ASD, or not.[136] An ideal diagnostic test would not result in any false positive (patient wrongly diagnosed with the condition) or false negative (patient wrongly diagnosed without the condition) results; however, in reality diagnostic tests are rarely perfect and clinicians must have an understanding of accuracy and error rate.[136] Measures can be used to assess how well a diagnostic test identifies people with a given condition (details are in Appendix I):

- Sensitivity (or detection rate): probability a test result will be positive when the disease is present (true positive rate)[137]
- Specificity (or true negative rate): probability a test result will be negative when the disease is not present (true negative rate)[137]
- Positive predictive value (PPV): probability the disease is present when the test is positive[137]
- Negative predictive value (NPV): probability the disease is not present when the test is negative[137]

Generally, a good diagnostic test has high degrees of sensitivity and specificity. The predictive value is determined by the sensitivity and specificity of the test and the prevalence of the condition in the population being tested.[137] The more sensitive a test, the less likely an individual with a negative test actually has the condition and thus the greater the NPV.[137] The more specific the test, the less likely an individual with a positive test does not have the condition and the greater the PPV.[137]

While sensitivity and specificity are characteristics of the test itself, PPV is a function of both the test characteristics and the underlying risk of the condition (prevalence) in the population being tested. The less prevalent a condition, the more likely a positive test is a false positive. For example, if the prevalence of a condition is 0.1%, a test with a sensitivity of 100% and a specificity of 99.5% will have a PPV of 16.7%. The same test will have a PPV of 66.7% if the prevalence of the condition is 1.0%.

KQ1: Effectiveness

All 3 publications in this evidence review were diagnostic accuracy studies. Table 3 presents the characteristics of the included diagnostic accuracy studies, all of which recruited children aged 18 to 72 months. In the study that formed the basis of the FDA's June 2021 authorization decision, Megerian and colleagues used data from a cohort study (NCT04151290) that initially recruited 711 participants at 14 sites in 6 states.[16] However, the COVID-19 emergency was declared 7 months after enrollment began and significantly affected study operations.[16] Analysis is limited to the 425 participants who completed both device input and specialist evaluation.[16]

Wall and colleagues described a study designed to validate the modification of the Canvas Dx diagnostic algorithm through use of a predetermined change control protocol, a mechanism proposed by the FDA to allow manufacturers of SaMD to intermittently unlock devices under certain circumstances to apply planned modifications to the device's algorithms and assess the impact.[13] Wall and colleagues use data from the cohort study described by Megerian and colleagues (NCT04151290), limited to the same sample of 425 individuals who completed both Cognoa screening and specialist evaluation, combining it with data from another cohort study

(NCT03871179) to create with a total sample size of 722.[13] While ClinicalTrials.gov reports that NCT03871179 was withdrawn with no subjects recruited,[138] Wall and colleagues used data from 297 children with concern for developmental delay who were enrolled through the study.[13] The research team used a machine-learning training methodology that randomly sorted data into training sets (n = 504) and test sets (n = 218); however, it is unclear whether performance values presented for the optimized algorithm (specificity, sensitivity, etc.) were based on the full dataset of 722 participants or the test data set of 218 participants.[13] As a result, any overlap of individual subjects in diagnostic accuracy calculations between the Wall and Megerian studies is unclear.[13,16]

Abbas and colleagues (2020) described a study which recruited 375 children aged 18 to 72 months at 3 sites in the US which was designed to demonstrate the assessment reliability of the Canvas tool (Western IRB project number 2202803) in relation to common screening tools,.[15] However, screening tools are not intended for diagnosis and are used to identify cases that warrant further diagnostic assessment. Because the studies by Wall and Abbas report numbers of cases included in analysis, but do not report on enrollment numbers, there is insufficient information to understand the impact of missing data.[13,15]

The publication by Megerian and colleagues has details on age, gender, and racial or ethnic distribution of enrolled children, and education and income level of parents, indicating study participants were broadly representative of patients who would use the diagnostic device in a real-world setting.[13] Wall and colleagues reported on mean age and gender distribution of both datasets. While the authors pointed readers to the publication by Megerian and colleagues[16] for further demographic details about 1 of the included datasets, they do not share any additional demographic detail about the other.[13] Abbas and colleagues reported enrollment of children aged 18 to 72 months, but do not have details on age, gender, or racial and ethnic distributions.[15]

We assessed all 3 studies as being at high risk of bias[13,15,16] (details in Appendix G). Standard clinical diagnosis based on DSM-5 criteria by specialist clinicians was the reference standard in all 3 studies and both participants and clinicians were blinded to diagnostic device results.[13,15,16] Data from NCT04151290 were used in 2 publications.[13,16] Changes to study operation due to the COVID-19 pandemic resulted in a smaller-than-planned pool of participants for analysis, which was not considered as a risk of bias in our assessment because of the unusual nature of the situation and demographic similarity between the recruited cohort and the subset of completers. Association of authors in the 3 studies[13,15,16] with Cognoa did contribute to risk of bias.[139] The publication by Abbas and colleagues reported the model was revised in the middle of the study, so the version of the diagnostic device used for the earlier group (N = 162) was different than that used for the later group (N = 213), and not all participants received all 3 portions of the Canvas assessment.[15] Two studies had incomplete reporting of recruitment methods and participant inclusion and exclusion criteria.[13,15] The publication by Wall and colleagues was sponsored by Cognoa, and though the sponsor was blinded, no details are given on the specific components of the trial, or how that was accomplished, since all authors were affiliated with Cognoa.[16] The study by Megerian and colleagues reported Cognoa sponsorship, while the remaining article identified author affiliations with Cognoa, but did not explicitly report Cognoa funding.[13,15]

Table 3. Characteristics of Included Studies

Publication Author, Year Trial Identifier Study Design Study Location	Population Description N Analyzed Participants Participant demographics	Intervention Comparator	Risk of Bias Outcomes of Interest Reported
Wall et al., 2023[13] NCT04151290 (425 children) NCT03871179 (297 children) US *Note: NCT03871179 is categorized as withdrawn in ClinicalTrials.gov, with no participants enrolled*	Dataset 1 (NCT04151290): children aged 18 to 72 months old with identified concerns for developmental delay by a health care provider or caregiver Dataset 2 (NCT03871179): children aged 18 to 72 months old with concern for developmental delay (source of concern undefined) Analyzed N = 722 (Enrollment figures not reported) NCT04151290: Mean age 3.33 years (SD 1.15) Gender 36.4% female NCT03871179: Mean age 3.97 years (no SD) Gender 42.9% female	Canvas Dx output compared with a clinical reference standard. For NCT04151290, the reference standard was a diagnosis made by a specialist clinician, based on DSM-5 criteria and validated by 1 or more blinded reviewing specialist clinicians. For NCT03871179, the reference standard was diagnosis based on DSM-5 criteria made by board certified pediatric psychiatrists, pediatric neurologists, developmental behavioral pediatricians, or psychologists with at least 5 years of experience diagnosing autism.	High risk of bias Outcomes • PPV • NPV • Indeterminate rate (no definitive categorization as ASD positive or ASD negative)
Megerian et al., 2022[16] NCT04151290 (425 children) US	Children 18 to 72 months old with identified concerns for developmental delay by a health care provider or caregiver Analyzed N = 425 (722 enrolled) Mean age 3.33 years (SD 1.15) Gender, 36.2% female White non-Hispanic, 53.9% Black, 13.2% Hispanic or Latino, 11.5%	Canvas Dx output compared with a diagnosis made by a specialist clinician, based on DSM-5 criteria and validated by 1 or more blinded reviewing specialist clinicians.	High risk of bias Outcomes • PPV • NPV • Sensitivity • Specificity • Indeterminate rate (no definitive categorization as ASD positive or ASD negative)

Publication Author, Year Trial Identifier Study Design Study Location	Population Description N Analyzed Participants Participant demographics	Intervention Comparator	Risk of Bias Outcomes of Interest Reported
Abbas et al., 2020[15] Western IRB #2202803 (375 children) US	Children 18 to 72 months of age referred through the typical referral process for suspicion of autism Analyzed N = 375 (Enrollment figures not reported) Mean age, not reported Gender, not reported	Cognoa diagnostic tool output compared with: • Outcomes of autism screening instruments (such as ADOS-2, M-CHAT-R, and/or CBCL) as appropriate for child's age and administered by the referring provider and • Clinical diagnosis by a licensed health care provider at a center specialized in neurodevelopmental disorders.	High risk of bias Outcomes • Difference in AUC for sensitivity and specificity in comparison with expert clinical diagnosis for 3 subsets: ○ Parent questionnaire only, ○ Parent questionnaire and video review ○ Parent questionnaire, video review, and clinician questionnaire • Differences in specificity at 90% sensitivity in comparison with clinical screening models (M-CHAT-R for children aged 1.5 to 3 years, CBCL, SRS-2) for 3 subsets described above

Abbreviations. ADOS-2: Autism Diagnostic Observation Schedule, Second Edition; ASD: autism spectrum disorder; AUC: area under the curve; CBCL: Child Behavior Checklist; M-CHAT-R: Modified Checklist for Autism in Toddlers, Revised; NPV: negative predictive value; PPV: positive predictive value; SD: standard deviation; SRS-2: Social Responsiveness Scale, Second Edition.

Accuracy

Three publications examined accuracy of the Canvas Dx diagnostic tool, as described in Tables 4 through 6.[13,15,16] Additional information on measures of test performance is in Appendix I, while details on study outcomes are found in Appendix D. Wall and colleagues report on results of an algorithmic threshold optimization procedure designed to decrease the number of indeterminate results and improve Canvas Dx's ability to detect or rule out ASD.[13] Among the subset of participants with a determinant result, both PPV and NPV decreased in the optimized model, but not significantly so, while the percent of children receiving a definitive ASD positive or negative result increased significantly from 45.4% (95% CI 41.3% to 48.6%) to 66.5% (95% CI 62.5% to 71.0%).[13]

In the study that informed Canvas Dx's De Novo application to the FDA,[78] Megerian and colleagues compared Canvas Dx outputs with diagnostic agreement by 2 or more independent specialists.[16] Canvas Dx produced a determinant positive or negative ASD result in only 31.8% of cases.[16] Among the subset of participants with a determinant result, sensitivity (98.4%) and NPV (98.3%) were high, while specificity (78.9%) and PPV (80.8%) were lower.[16] Children younger than 3 years of age were significantly more likely to receive a determinate result (38.9% vs.

26.2% for children aged 3 years and older), but Canvas Dx had significantly lower specificity in younger children (67% in children younger than age 3, and 88% in children aged 3 years and older).[16] Sensitivity, specificity, PPV, and NPV were all lower for girls in comparison with boys, while the indeterminate rate was higher for girls (71% for girls and 67% for boys), but confidence intervals were wide and differences were not statistically significant.[16] The authors performed a covariate analysis of Canvas Dx by racial category and Hispanic ethnicity, but categories were nonexclusive and a single child could be counted in more than 1 category.[16] The ability of the study to aid understanding of device performance in children of different genders and races is limited, and the authors pointed out the study was not powered for such analyses.[16]

Abbas and colleagues compared results of the Cognoa tool in assessment of 375 children who had been referred to 1 of 3 specialty autism centers for comprehensive ASD evaluation.[15] The majority of the sample completed only the parent questionnaire and video screening, while a subset (n = 213) also completed the clinician questionnaire.[15] For the subset of participants who completed all 3 portions of the Canvas Dx diagnostic tool, area under the receiver operating characteristics (ROC) curve (AUC) was 0.916 for children aged 18 to 72 months and 0.837 for the subset of children younger than age 4, using clinical diagnosis at a center specializing in neurodevelopmental disorders as the reference standard and allowing for inconclusive determination in up to 30% of cases.[15] Confidence intervals are not included and no details are given on number of inconclusive results, sensitivity, specificity, PPV, NPV, or accuracy.[15] When the ability to abstain from decision making was removed and the Cognoa tool had to make a choice between ASD and no ASD, AUC dropped to 0.846 for children aged 18 to 84 months and to 0.759 for children younger than age 4.[140] It is unknown whether the machine-learning algorithm supporting testing in the study by Abbas and colleagues[15] is the same as that used by Megerian and colleagues[16] in support of FDA authorization.

Only the study by Megerian and colleagues has data for construction of a confusion matrix, which summarizes the performance of a machine-learning model on a set of test data.[16] Wall and colleagues report determinate rate, as well as PPV and NPV, but do not give detailed information on number of true positives, false positives, or other metrics.[13] Abbas and colleagues only share data on area under the ROC curve in their analysis of module performance in relation to the clinical reference standard.[15] The Megerian study calculations of positive and negative predictive value, sensitivity, and specificity exclude all cases for which the algorithm did not return a result (indeterminate cases)[16] and both Wall and Abbas also exclude indeterminate cases from calculation of accuracy metrics.[13,15] While the abstention is presented as a safeguard to increase test accuracy,[13] calculating test performance as if those inconclusive results do not exist does not accurately reflect test performance and can produce biased estimates of accuracy.[141] There are multiple approaches to handling indeterminate results that would avoid overstated summary statistics and create greater transparency, none of which are used in these studies.[141-143]

Table 4. Positive and Negative Predictive Values for Canvas Dx

Study ID	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)
Wall et al. 2023[13] (some overlap with Megerian)[16]	Revised algorithm 87.5% (82.5-96.7) Original algorithm 89.7 (83.9-90.3)	Revised algorithm 95.6% (93.7-97.9) Original algorithm 96.1 (93.4-98.6)
Megerian et al. 2022[16]	80.8% (70.3-88.8)	98.3 (90.6-100)

Abbreviations. CI: confidence interval.

Table 5. Sensitivity and Specificity for Canvas Dx

Study ID	Sensitivity (95% CI)	Specificity (95% CI)
Megerian et al. 2022[16]	98.44% (91.60% to 99.96%)	78.87% (67.56% to 87.67%)
Abbas et al. 2020[15]	0.7 coverage, N = 112, AUC = 0.837	0.7 coverage, N = 204, AUC = 0.916

Abbreviations. AUC: area under the curve; CI: confidence interval.

Table 6. Conclusive Results Produced by Canvas Dx (positive or negative for ASD)

Study ID	Returned conclusive result (ASD positive or ASD negative)
Wall et al. 2023[13] (some overlap of data with Megerian)[16]	Revised algorithm 66.5% (62.5-71.0) Original algorithm 45.4 (41.3-48.6)
Megerian et al. 2022[16]	31.8% (135/425)
Abbas et al. 2020[15]	Not specified (minimum 70%)

Abbreviations. ASD: autism spectrum disorder.

Strength of Evidence

We applied the GRADE system (Appendix C) to rate the overall quality of evidence for key outcomes and test performance.[17,18] Table 7 presents a summary of findings for the diagnostic ability of Canvas Dx for children with suspected ASD compared with decisions of clinical specialists. Strength of evidence was very low across all outcomes. Full details of the GRADE assessment are in Appendix H.

Table 7. Summary of Findings (GRADE)

Outcome	Number of Participants and Studies	Median (Range)	Test Accuracy CoE	Rationale[a]
Sensitivity	1 diagnostic accuracy study[16] N = 425	Canvas DX had a high sensitivity when compared with a reference standard of diagnosis made by a specialist clinician, based on DSM-5 criteria and validated by 1 or more blinded reviewing specialist clinicians (sensitivity of 98.4%; 95% CI, 91.6% to 100%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias and 1 level for indirectness[a,b]
Specificity	1 diagnostic accuracy study[16] N = 425	78.9% (95% CI, 67.6% to 87.7%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias, 1 level for imprecision (i.e., wide confidence intervals), 1 level for indirectness.[a,b]
PPV	2 diagnostic accuracy studies[13,16] N = 1,147	Presumably using the same algorithm and an overlapping data set, 1 study (Wall et al., 2023) reports PPV at 89.7% (95% CI, 83.9% to 90.3%)[13] and the other (Megerian et al., 2022) at 80.8% (95% CI, 70.3% to 88.8%)	●○○○ VERY LOW	Downgraded 2 levels for risk of bias, 1 level for inconsistency, 1 level for indirectness, and 1 level for imprecision (large confidence interval in 1 study).

Outcome	Number of Participants and Studies	Median (Range)	Test Accuracy CoE	Rationale[a]
NPV	2 diagnostic studies[13,16] N = 1,147	Presumably using the same algorithm and an overlapping data set, 1 study (Wall) reports NPV at 96.1% (95% CI 93.4% to 98.6%)[13] and the other (Megerian) at 98.3% (90.6% to 100%)[16]	●○○○ VERY LOW	Downgraded 2 levels for risk of bias and 1 level for indirectness.[a]
Time to diagnosis	No evidence	--	--	--
Time to service initiation	No evidence	--	--	--

Notes. [a] *Indirectness refers to the degree to which children in the diagnostic study are representative of the broader population of children who have concern for developmental disability identified by parents or clinicians. Imprecision refers to the level of certainty of the true effect.*[17,18]

[b] *Unable to rate for inconsistency as only 1 study. For methods and interpretation of GRADE ratings, see Appendix C.*

Abbreviations. AUC: area under the curve; CoE: certainty of evidence; CI: confidence interval; GRADE: Grading of Recommendations, Assessment, Development, and Evaluations approach; NPV: negative predictive value; PPV: positive predictive value; ROC receiver operating characteristic.

Health Equity and Canvas Dx Accuracy

The AAP has identified gender and race as potential barriers to identifying risk for ASD.[2] The study by Megerian and colleagues is the only article on Canvas Dx to explicitly report on device performance by gender or racial or ethnic subgroup, although the authors caution the study was not powered to detect differences in these covariates.[16] They report no difference in device performance (PPV or NPV, indeterminate rate, sensitivity, or specificity) across participants' race, ethnicity, or gender, by examining the overlap of corresponding 95% confidence intervals, although confidence intervals were very wide (Table 7).[16] Participants who reported multiple races or ethnicities (n = 57) were included in each category they identified, with the result that an individual might be compared with themselves.[16] The publication by Megerian and colleagues is also the only study to report data on education and income level of participant families.[16] They report no difference in device performance (PPV or NPV, indeterminate rate, sensitivity, or specificity) across parent education level or income level, although numbers were small and confidence intervals were very wide.

Articles about the training of the video screening portion of Canvas Dx frequently report on the gender mix of videos used, but do not report on results by gender or any differential results in performance of the classifier.[71-73] Publications on the video assessment portion of the Canvas Dx app do not report on the racial or ethnic mix of included children and do not provide data on differential performance of the device in different racial or ethnic groups.[67,71-74,144]

Table 7. Gender Differences in Canvas Dx Performance Described by Megerian et al. (2022)[1]

Gender	n	PPV (n/N) 95% CI	NPV (n/N) 95% CI	Indeterminate rate (n/N) 95% CI	Sensitivity (n/N) 95% CI	Specificity (n/N) 95% CI
Female	154	60% (12/20) 36%, 81%	96% (24/25) 80%, 100%	71% (109/154) 61%, 78%	92% (12/13) 64%, 100%	75% (24/32) 57%, 89%
Male	271	88% (51/58) 77%, 95%	100% (32/32) 89%, 100%	67% (181/271) 61%, 72%	100% (51/51) 93%, 100%	82% (32/39) 66%, 92%
Black	77	95% (19/20) 75%, 100%	100% (7/7) 59%, 100%	65% (50/77) 53%, 75%	100% (19/19) 82%, 100%	88% (7/8) 47%, 100%
White	259	73% (27/37) 56%, 86%	97% (37/38) 86%, 100%	71% (184/259) 65%, 76%	96% (27/28) 82%, 100%	79% (37/47) 64%, 89%
Hispanic (any race)	75	74% (14/19) 49%, 91%	100% (10/10) 69%, 100%	61% (46/75) 49%, 72%	100% (14/14) 77%, 100%	67% (10/15) 38%, 88%

Notes. The study was not powered to detect differences in these covariates.
Abbreviations. CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

Effectiveness

We did not identify any studies that captured information on use of Canvas Dx in a primary care setting or measured how or if Canvas Dx altered the care pathway between identification of concern about the possibility of ASD or developmental delay and diagnosis or referral to early intervention services.

Relevant Ongoing Studies

We identified no ongoing studies registered on ClinicalTrials.gov that fit the inclusion criteria for this review.[13] As noted in the Background section under Contextual Question 1, a prospective observational study (NCT05223374) designed to assess the feasibility and impact of integrating Canvas Dx into the ECHO Autism STAT diagnosis model was completed in February 2024 but has not yet published results.[111] In this Cognoa-funded study, ECHO Autism clinicians from participating primary care practices in Missouri were assigned to initiate Canvas Dx device prescriptions during diagnostic evaluations for patients 18 to 72 months of age suspected of having ASD.[111] While the study results will provide some insight on the usability and performance of Canvas Dx in clinical practice, it involves clinicians who already have specialized training and mentorship in ASD diagnosis and treatment and are not representative of the typical pediatric primary care provider (particularly those in under-resourced areas).[111] The lack of a control group or comparative cohort will limit conclusions drawn regarding any relationship between Canvas Dx and time to diagnosis or time to initiating therapy.[111]

KQ2. Adverse Effects

No studies were identified that addressed the safety of the Canvas Dx diagnostic device. Both Wall and colleagues[13] and Megerian and colleagues[16] describe the algorithm's ability to abstain from making a determination when presented with insufficient information as a safety feature that minimizes false negatives. None of the 3 studies addressed adverse events in the studies or the potential for adverse events during use of the diagnostic device in real-world applications.[13,15,16]

As part of the benefit and risk analysis in the De Novo application, the FDA cautions that the device may give unreliable results if used in patients with other conditions that would have been excluded from the clinical study by Megerian and colleagues,[16] such as known deafness or blindness, major dysmorphic features (e.g., fetal alcohol syndrome), history or diagnosis of genetic conditions (i.e., Rhett syndrome, fragile X), and history or diagnosis of epilepsy.[78] Identified risks associated with the use of Canvas Dx include misdiagnosis and delayed diagnosis of ASD, which can result in delayed treatment of ASD and delivery of treatment not appropriate for ASD; however, the FDA decision determined that benefits outweighed risks.[78]

KQ3. Costs or Cost-Effectiveness

No studies related to cost or cost-effectiveness of Canvas Dx were identified.

KQ4. Clinical Practice Guideline Recommendations

No clinical practice guidelines were identified that addressed use of Canvas Dx for children whose caregivers or health care providers suspect developmental delay or ASD.

KQ5. Medicaid Program Coverage Policies and Private Payer Policies

We identified coverage policies for Canvas Dx from 3 of the 13 private payers searched for this report: Aetna, Anthem Blue Cross and Blue Shield (formerly Empire BlueCross BlueShield), and Highmark Blue Shield of Northeastern New York.[19-21] There was no national coverage determination for Canvas Dx. None of the 10 states on our predetermined list of states to search had a Medicaid policy for Canvas Dx. See Table 8 for a summary of the policies and Appendix K for a full description of the policies. See Appendix A for the policy sources searched and terms used to identify relevant coverage policies.

One (Highmark Blue Shield of Northeastern New York) of the 3 policies for Canvas Dx was specific to digital diagnostics and Canvas Dx.[21] The remaining payers mentioned Canvas Dx in the policy for ASDs (Aetna),[145] prescription digital therapeutics (Aetna),[19] or mobile device-based health management applications (Anthem Blue Cross and Blue Shield).[20]

Two of the 3 payers with policies for Canvas Dx did not cover this product (Table 8).[19,20] These payers considered Canvas Dx experimental and investigational with insufficient data to demonstrate either its efficacy or clinical utility.[19,20,146] Highmark Blue Shield of Northeastern New York was the only payer to cover Canvas Dx.[21] Their policy stated that Canvas Dx may be considered medically necessary to aid in the diagnosis of ASD in individuals aged 18 to 72 months exhibiting persistent deficits in social communication and social interaction across multiple contexts or restricted, repetitive patterns of behaviors, interests, or activities, or both, as noted by a parent, caregiver, or provider.[21] Highmark, Inc. is a commercial payer partner of Cognoa, the maker of Canvas Dx.[147,148]

Table 8. Summary of Coverage for Canvas Dx

Payer	Canvas Dx Covered Last Review Date
Private payers	
Aetna	**No** July 17, 2024[19] September 20, 2024[145]
Anthem Blue Cross and Blue Shield	**No** August 8, 2024[20]
Capital District Physicians' Health Plan	No policy identified
Cigna Healthcare	No policy identified
EmblemHealth	No policy identified
Excellus BlueCross BlueShield	No policy identified
Fidelis Care	No policy identified
Healthfirst	No policy identified
Highmark Blue Shield of Northeastern New York	**Yes** May 2024[21]
MetroPlusHealth	No policy identified
Molina Healthcare	No policy identified[a]
Tufts Health Plan	No policy identified
United Healthcare	No policy identified
Medicare	
Centers for Medicare & Medicaid Services	No policy identified
State Medicaid programs	
California (Medi-Cal)	No policy identified
Florida	No policy identified
Massachusetts (MassHealth)	No policy identified
New Jersey (NJ FamilyCare)	No policy identified
New York	No policy identified
North Carolina	No policy identified
Oregon (Oregon Health Plan)	No policy identified
Pennsylvania	No policy identified
Texas	No policy identified
Washington (Apple Health)	No policy identified

Note. [a] *Canvas Dx was reviewed in Molina Healthcare's 2023 policy on prescription digital therapeutics, but not mentioned in the 2024 version of this policy.*

Discussion

Canvas Dx is authorized by the FDA as a diagnostic aid to assist clinicians in evaluating a child with suspected ASD.[14] As envisioned by Cognoa, when a concern for ASD or other developmental disability has been identified through initial screening, clinicians can prescribe Canvas Dx to help aid in decision making, either ruling out a diagnosis, making a diagnosis, or referring a patient to specialists for further evaluation.[14] We identified 3 publications from 3 eligible studies with diagnostic accuracy outcomes for Canvas Dx, all of which were sponsored by Cognoa or were led by the Cognoa research team.[13,15,16] We did not identify any independent evaluations of the Canvas Dx diagnostic tool, no clinical practice guidelines that addressed Canvas Dx specifically or digital health diagnostic technologies generally, no relevant ongoing trials related to Canvas Dx, and no studies related to cost or cost-effectiveness of Canvas Dx. Critical identified outcomes of interest for this review were sensitivity, specificity, PPV, NPV, time to diagnosis (critical), and time to ASD-specific service initiation. Change in clinical management of diagnostic and referral processes was identified as an important outcome. Two of 3 included publications had data on PPV and NPV of the Canvas Dx tool,[13,16] while 2 also had information on sensitivity and specificity.[15,16]

For all outcomes, strength of evidence was very low (see Appendix H) primarily due to risk of bias (see Appendix G), indirectness (lack of information to determine how well the analyzed sample represents the broader population of toddlers and preschoolers with concern for ASD), imprecision springing from wide confidence intervals, and failure to use a method of calculating accuracy that accounted for indeterminate results. Because accuracy measures were limited to the subset of cases for which Canvas Dx produced a definitive result (ASD or no ASD), summary statistics are likely overstated in all 3 studies.[13,15,16] None of the identified studies measured the critical outcomes of time to diagnosis or time to ASD-specific service initiation, or the important outcome of change in clinical management of the diagnostic and referral processes.

Health Equity Considerations

Of the 3 publications in this evidence review, the study by Megerian and colleagues is the only 1 to report on device performance by gender or racial or ethnic subgroup.[16] While the study reports no difference in device performance (PPV, NPV, indeterminate rate, sensitivity, or specificity) across participants' race, ethnicity, or gender, by examining the overlap of corresponding 95% confidence intervals, the authors caution the study was not powered to detect differences in these covariates and confidence intervals were extremely wide.[16] The study by Megerian and colleagues is the only 1 of the 3 to give demographics on the study sample, reporting that 64% of children were male, 54% were Caucasian, 13% were Black, and 13% were mixed race (11.5% were categorized as Hispanic only).[16] Without further demographic information, it is impossible to know how the population of the studies reflects the New York Medicaid population.

Publications about the video assessment portion of the Canvas Dx app do not report on the racial or ethnic mix of included children and do not share data on performance of the device in different racial or ethnic groups.[67,71-74,144] A Cognoa team led by Washington and colleagues (2022) concludes that future work should evaluate the methods across races, ethnicities, and other sensitive attributes to ensure fair and unbiased AI.[71] Reporting on an automated facial

expression classifier being developed by the Cognoa team for use in diagnostic and therapeutic tools, Banerjee and colleagues (2023) reported the model performed significantly worse on African American and South Asian ethnicities than on European American children, with as much as 11.6% lower accuracy.[22] While accuracy was 68.7% in European American children and 66.1% in Asian children, it dropped to 61.3% for African American children and 57.1% for South Asian children.[22]

Tariq and colleagues (2019), members of the Cognoa research team, sought to validate the machine learning models used in Canvas Dx using parent-submitted home videos of Bangladeshi children: 55 videos of children with ASD, 50 of children with another speech or language condition, and 54 of children with neurotypical development.[76] Videos were rated by 9 crowdsourced non-Bengali-speaking US-based raters without clinical training who received a 1-hour training session with an analyst before rating videos.[76] Questionnaires completed by raters were used to assemble machine-learning classifiers.[76] Relative accuracy was low.[76] Classifiers trained on test data from US videos performed with 94.2% accuracy when tested on a US video data set, but dropped to 54% when used on the Bangladeshi data.[76]

Evidence Related to Canvas Dx's Influence on Diagnosis Rates and Care Pathways

The developers of Canvas Dx promote the tool as a means of increasing diagnostic rates by health care providers in primary care and decreasing strain on specialty services.[13,16,23] The stated goal of the Canvas Dx tool is to reduce age at diagnosis and initiation of interventions.[13,16,23] At this time there is no evidence Canvas Dx would increase diagnosis in primary care settings, which could aid in early diagnosis, help families avoid long wait times and potentially long travel times for access to limited diagnostic services, or potentially speed access to early intervention services. If families are referred for expert consultation rather than receiving a diagnosis in primary care, it is not likely the Canvas Dx tool would decrease age at diagnosis or reduce wait times for limited services.

No comparative studies have been completed and no studies in progress were identified that quantify what role Canvas Dx may have in increasing ASD diagnosis rates by primary care practitioners. According to the New York State Department of Health Clinical Practice Guideline on Assessment and Intervention Services for Young Children with Autism Spectrum Disorders, a primary care provider or other early childhood professional who suspects a child may have a developmental delay or ASD must refer the child to the Early Intervention Program (EIP) for further evaluation.[99]

One recently completed but as yet unpublished single-arm observational study was designed to test use of the Canvas Dx diagnostic tool in pediatric practices simultaneously enrolled in an ASD training and mentorship program, a population not representative of typical primary care practices.[111] Study results will not shed light on how use of Canvas Dx might affect time to diagnosis or time to initiation of intervention services in most pediatric practices. Similarly, the study cannot add any objective insight on how use of the diagnostic tool might affect clinical management pathways in pediatric practices that are not receiving special training and mentorship in ASD recognition and management.[111]

There are currently no federal requirements for coverage of digital diagnostics by Medicaid and no HCPCS code for Canvas Dx. Highmark Blue Shield of Northeastern New York was the only

payer amongst the 3 payers with policies related to Canvas Dx to cover this product.[21] The Highmark Blue Shield of Northeastern New York policy states that Canvas Dx may be considered medically necessary to aid in the diagnosis of ASD in children aged 18 to 72 months who exhibit 1 or more signs of developmental delay, as noted by a parent, caregiver, or provider.[21] Highmark, Inc. is a commercial payer partner of Cognoa.[147,148]

Summary

Canvas Dx's classification algorithm is sensitive to extreme cases (very high or very low risk of autism), but has more difficulty in identifying less extreme cases.[133] While the inclusion of an indeterminate category is described as a safety feature that minimizes false negatives,[13,16] the result is the diagnostic tool is most likely to identify children with uncomplicated presentations and relatively severe symptoms, while not identifying children with mild to moderate presentations, groups who could directly benefit from early, intensive intervention.[149]

Performance of the Canvas Dx device across genders and varying racial and ethnic groups and socioeconomic strata, and its ability to affect disparities in diagnosis of ASD and promotion of early intervention, is unestablished. The developers of Canvas Dx acknowledge the need for more research to evaluate the diagnostic method across races, ethnicities, and other attributes, such as caregiver education level or primary language.[71] Additionally, research on the usability, acceptability, and validity of Canvas Dx in underserved communities and diverse populations could help quantify the ability of the diagnostic tool to address barriers to entry to diagnostic and therapeutic pathways.[133] Independent research on use of Canvas Dx is needed to corroborate results of manufacturer-led studies, give insight into acceptability and usability of the digital health technology in real-world settings, shed light on Canvas Dx's ability to affect time to diagnosis or time to ASD-specific service initiation, and gauge if the tool can create change in clinical management of the diagnostic and referral processes.

References

1. Centers for Disease Control and Prevention. What is autism spectrum disorder? 2022; https://www.cdc.gov/ncbddd/autism/facts.html. Accessed August 5, 2023.

2. Hyman SL, Levy SE, Myers SM, Council On Children With Disabilities Section On Developmental Behavioral Pediatrics. Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics.* 2020;145(1). doi: 10.1542/peds.2019-3447.

3. Meanner MJ, Warren Z, A RW, Amoakohene E, Bakian AV BD. Prevalence and characteristics of autism spectrum disorder among children aged 8 years--Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020. *Morb Mortal Weekly Rep.* 2023;72(2):1-14.

4. Liu BM, Paskov K, Kent J, et al. Racial and Ethnic Disparities in Geographic Access to Autism Resources Across the US. *JAMA Netw Open.* 2023;6(1):e2251182. doi: 10.1001/jamanetworkopen.2022.51182.

5. Yingling ME, Ruther MH, Dubuque EM, Mandell DS. County-level variation in geographic access to Board Certified Behavior Analysts among children with Autism Spectrum Disorder in the United States. *Autism.* 2021;25(6):1734-1745. doi: 10.1177/13623613211002051.

6. Singh JS, Bunyak G. Autism Disparities: A Systematic Review and Meta-Ethnography of Qualitative Research. *Qual Health Res.* 2019;29(6):796-808. doi: 10.1177/1049732318808245.

7. Finke EH, Drager KDR, Ash S. Pediatricians' perspectives on identification and diagnosis of autism spectrum disorders. *J Early Child Res.* 2010;8(3):254-268. doi: 10.1177/1476718X10366773.

8. Hamp N, DeHaan SL, Cerf CM, Radesky JS. Primary Care Pediatricians' Perspectives on Autism Care. *Pediatrics.* 2023;151(1). doi: 10.1542/peds.2022-057712.

9. Gascon A, Gamache D, St-Laurent D, Stipanicic A. Do we over-diagnose ADHD in North America? a critical review and clinical recommendations. *J Clin Psychol.* 2022;78(12):2363-2380. doi: 10.1002/jclp.23348.

10. Hus Y, Segal O. Challenges surrounding the diagnosis of autism in children. *Neuropsychiatr Dis Treat.* 2021;17:3509-3529. doi: 10.2147/ndt.S282569.

11. Wieckowski AT, Zuckerman KE, Broder-Fingert S, Robins DL. Addressing current barriers to autism diagnoses through a tiered diagnostic approach involving pediatric primary care providers. *Autism Res.* 2022;15(12):2216-2222. doi: 10.1002/aur.2832.

12. Cognoa Inc. Canvas Dx. https://cognoa.com. Accessed May 22, 2024.

13. Wall DP, Liu-Mayo S, Salomon C, Shannon J, Taraman S. Optimizing a de novo artificial intelligence-based medical device under a predetermined change control plan: improved ability to detect or rule out pediatric autism. *Intell Based Med.* 2023;8. doi: 10.1016/j.ibmed.2023.100102.

14. US Food and Drug Administration. FDA authorizes marketing of diagnostic aid for autism spectrum disorder. 2021; https://www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-diagnostic-aid-autism-spectrum-disorder. Accessed April 28, 2024.

15. Abbas H, Garberson F, Liu-Mayo S, Glover E, Wall DP. Multi-modular AI approach to streamline autism diagnosis in young children. *Sci Rep.* 2020;10(1):5014. doi: 10.1038/s41598-020-61213-w.

16. Megerian JT, Dey S, Melmed RD, et al. Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. *NPJ Digit Med.* 2022;5(1):57. doi: 10.1038/s41746-022-00598-6.

17. Guyatt G, Oxman A, Vist G, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008;336(7650):924-926. doi: 10.1136/bmj.39489.470347.AD.

18. Schünemann H, Brozek J, Guyatt G, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendations. 2013; http://gdt.guidelinedevelopment.org/app/handbook/handbook.html. Accessed December 15, 2015.

19. Aetna. Prescription digital therapeutics. 2024; https://www.aetna.com/cpb/medical/data/900_999/0999.html. Accessed October 22, 2024.

20. Anthem Blue Cross Blue Shield. Mobile device-based health management applications. 2024; https://www.anthem.com/dam/medpolicies/abcbs/active/guidelines/gl_pw_e000367.html. Accessed October 22, 2024.

21. Highmark Inc. Digital diagnostics. 2024; https://securecms.highmark.com/content/medpolicy/en/highmark/ny/commercial/policies/Diagnostic_Medical/M-86/M-86-002.html. Accessed October 22, 2024.

22. Banerjee A, Mutlu OC, Kline A, Surabhi S, Washington P, Wall DP. Training and profiling a pediatric facial expression classifier for children on mobile devices: machine learning study. *JMIR Form Res.* 2023;7:e39917. doi: 10.2196/39917.

23. Autism Science Foundation Weekly Science Podcasts. Machine learning in autism, explained (an interview with Dennis Wall). 2024; https://asfpodcast.org/archives/1682. Accessed May 14, 2024.

24. National Institute of Environmental Health Sciences. Autism. 2023. https://www.niehs.nih.gov/health/topics/conditions/autism. Accessed February 15, 2024.

25. Centers for Disease Control and Prevention. Autism spectrum disorder (ASD): screening and diagnosis of autism spectrum disorder. 2022; https://www.cdc.gov/ncbddd/autism/screening.html. Accessed January 19, 2024.

26. National Institute of Mental Health. Autism spectrum disorder. 2023; https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd. Accessed January 19, 2024.

27. American Psychiatric Association. Autism spectrum disorder. 2022; https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787. Accessed September 30, 2024.

28. Volkmar F, Siegel M, Woodbury-Smith M, et al. Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):237-257. doi: 10.1016/j.jaac.2013.10.013.

29. Mottron L, Bzdok D. Autism spectrum heterogeneity: fact or artifact? *Mol Psychiatry*. 2020;25(12):3178-3185. doi: 10.1038/s41380-020-0748-y.

30. Constantino JN, Todd RD. Autistic traits in the general population: a twin study. *Arch Gen Psychiatry*. 2003;60(5):524-530. doi: 10.1001/archpsyc.60.5.524.

31. Cumin J, Pelaez S, Mottron L. Positive and differential diagnosis of autism in verbal women of typical intelligence: a delphi study. *Autism*. 2022;26(5):1153-1164. doi: 10.1177/13623613211042719.

32. First MB. Mutually exclusive versus co-occurring diagnostic categories: the challenge of diagnostic comorbidity. *Psychopathology*. 2005;38(4):206-210. doi: 10.1159/000086093.

33. Jacobs D, Steyaert J, Dierickx K, Hens K. Parents' views and experiences of the autism spectrum disorder diagnosis of their young child: a longitudinal interview study. *Eur Child Adolesc Psychiatry*. 2020;29(8):1143-1154. doi: 10.1007/s00787-019-01431-4.

34. Jacobs D, Steyaert J, Dierickx K, Hens K. Parents' multi-layered expectations when requesting an Autism Spectrum Disorder assessment of their young child: an in-depth interview study. *BMC Psychiatry*. 2020;20(1):440. doi: 10.1186/s12888-020-02806-7.

35. Lappé M, Lau L, Dudovitz RN, Nelson BB, Karp EA, Kuo AA. The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics.* 2018;141(Suppl 4):S272-s279. doi: 10.1542/peds.2016-4300C.

36. Mackie TI, Schaefer AJ, Ramella L, et al. Understanding How Parents Make Meaning of Their Child's Behaviors During Screening for Autism Spectrum Disorders: A Longitudinal Qualitative Investigation. *J Autism Dev Disord.* 2021;51(3):906-921. doi: 10.1007/s10803-020-04502-7.

37. Carlsson E, Miniscalco C, Kadesjö B, Laakso K. Negotiating knowledge: parents' experience of the neuropsychiatric diagnostic process for children with autism. *Int J Lang Commun Disord.* 2016;51(3):328-338. doi: 10.1111/1460-6984.12210.

38. Dounavi K, Koldas M. Parental Perspectives on Early Life Screening and Genetic Testing for ASD: A Systematic Review. *J Autism Dev Disord.* 2024. doi: 10.1007/s10803-023-06231-z.

39. Barak-Levy Y, Paryente B. Diving into the Resolution Process: Parent's Reactions to Child's Diagnosis. *Int J Environ Res Public Health.* 2023;20(4). doi: 10.3390/ijerph20043295.

40. Naicker VV, Bury SM, Hedley D. Factors associated with parental resolution of a child's autism diagnosis: A systematic review. *Front Psychiatry.* 2022;13:1079371. doi: 10.3389/fpsyt.2022.1079371.

41. Zwaigenbaum L, Bryson SE, Brian J, et al. Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. *Autism Res.* 2016;9(7):790-800. doi: 10.1002/aur.1585.

42. Becerra-Culqui TA, Lynch FL, Owen-Smith AA, Spitzer J, Croen LA. Parental first concerns and timing of autism spectrum disorder diagnosis. *J Autism Dev Disord.* 2018;48(10):3367-3376. doi: 10.1007/s10803-018-3598-6.

43. Zablotsky B, Colpe LJ, Pringle BA, Kogan MD, Rice C, Blumberg SJ. Age of parental concern, diagnosis, and service initiation among children with autism spectrum disorder. *Am J Intellect Dev Disabil.* 2017;122(1):49-61. doi: 10.1352/1944-7558-122.1.49.

44. James SN, Smith CJ. Early autism diagnosis in the primary care setting. *Semin Pediatr Neurol.* 2020;35:100827. doi: 10.1016/j.spen.2020.100827.

45. Fuller EA, Kaiser AP. The effects of early intervention on social communication outcomes for children with autism spectrum disorder: a meta-analysis. *J Autism Dev Disord.* 2020;50(5):1683-1700. doi: 10.1007/s10803-019-03927-z.

46. Reichow B, Hume K, Barton EE, Boyd BA. Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD) 2018; https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD009260.pub3/full?highlightAbstract=autism. Accessed April 28, 2024.

47. Rodgers M, Marshall D, Simmonds M, et al. Interventions based on early intensive applied behaviour analysis for autistic children: a systematic review and cost-effectiveness analysis. *Health Technol Assess.* 2020;24(35):1-306. doi: 10.3310/hta24350.

48. Okoye C, Obialo-Ibeawuchi CM, Obajeun OA, et al. Early diagnosis of autism spectrum disorder: a review and analysis of the risks and benefits. *Cureus.* 2023;15(8):e43226. doi: 10.7759/cureus.43226.

49. French L, Kennedy EMM. Annual Research Review: Early intervention for infants and young children with, or at-risk of, autism spectrum disorder: a systematic review. *J Child Psychol Psychiatry.* 2018;59(4):444-456. doi: 10.1111/jcpp.12828.

50. Estes A, Munson J, Rogers SJ, Greenson J, Winter J, Dawson G. Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry.* 2015;54(7):580-587. doi: 10.1016/j.jaac.2015.04.005.

51. Green J, Pickles A, Pasco G, et al. Randomised trial of a parent-mediated intervention for infants at high risk for autism: longitudinal outcomes to age 3 years. *J Child Psychol Psychiatry.* 2017;58(12):1330-1340. doi: 10.1111/jcpp.12728.

52. Landa RJ, Kalb LG. Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention. *Pediatrics.* 2012;130 Suppl 2:S186-190. doi: 10.1542/peds.2012-0900Q.

53. Landa RJ. Efficacy of early interventions for infants and young children with, and at risk for, autism spectrum disorders. *Int Rev Psychiatry.* 2018;30(1):25-39. doi: 10.1080/09540261.2018.1432574.

54. Cidav Z, Munson J, Estes A, Dawson G, Rogers S, Mandell D. Cost offset associated with Early Start Denver Model for children with autism. *J Am Acad Child Adolesc Psychiatry.* 2017;56(9):777-783. doi: 10.1016/j.jaac.2017.06.007.

55. Sampaio F, Feldman I, Lavelle TA, Skokauskas N. The cost-effectiveness of treatments for attention deficit-hyperactivity disorder and autism spectrum disorder in children and adolescents: a systematic review. *Eur Child Adolesc Psychiatry.* 2022;31(11):1655-1670. doi: 10.1007/s00787-021-01748-z.

56. Segal L, Green J, Twizeyemariya A, et al. Estimated therapy costs and downstream cost consequences of iBASIS-Video Interaction to Promote Positive Parenting Intervention vs

usual care among children displaying early behavioral signs of autism in Australia. *JAMA Netw Open*. 2023;6(4):e235847. doi: 10.1001/jamanetworkopen.2023.5847.

57. Mendez AI, McQueen E, Gillespie S, Klin A, Klaiman C, Pickard K. Access to Part C, Early Intervention for children younger than 4 years evaluated for autism spectrum disorder. *Autism*. 2024;28(6):1431-1440. doi: 10.1177/13623613241229150.

58. New York State Education Department. Number of New York State children and youth with disabilities receiving special education programs and services. 2022-2023; https://www.p12.nysed.gov/sedcar/archived/2223pdrpts.html. Accessed 7/25/2024.

59. DeMasi ME. Children with autism spectrum disorders. New York Council on Children and Families; 2020; https://www.ccf.ny.gov/files/3613/8262/2276/Autism20brief.pdf. Accessed September 6, 2024.

60. Gallin Z, Kolevzon AM, Reichenberg A, Hankerson SH, Kolevzon A. Racial Differences in the Prevalence of Autism Spectrum Disorder: A Systematic Review. *J Autism Dev Disord*. 2024. doi: 10.1007/s10803-024-06403-5.

61. McGrath K, Bonuck K, Mann M. Exploratory spatial analysis of autism rates in New York school districts: role of sociodemographic and language differences. *J Neurodev Disord*. 2020;12(1):35. doi: 10.1186/s11689-020-09338-x.

62. Autism Speaks. Resource Guide. 2024; https://www.autismspeaks.org/resource-guide. Accessed September 6, 2024.

63. Duda M, Kosmicki JA, Wall DP. Testing the accuracy of an observation-based classifier for rapid detection of autism risk. *Transl Psychiatry*. 2014;4(8):e424. doi: 10.1038/tp.2014.65.

64. Kosmicki JA, Sochat V, Duda M, Wall DP. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. *Transl Psychiatry*. 2015;5(2):e514. doi: 10.1038/tp.2015.7.

65. Levy S, Duda M, Haber N, Wall DP. Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. *Mol Autism*. 2017;8:65. doi: 10.1186/s13229-017-0180-6.

66. Wall DP, Kosmicki J, Deluca TF, Harstad E, Fusaro VA. Use of machine learning to shorten observation-based screening and diagnosis of autism. *Transl Psychiatry*. 2012;2(4):e100. doi: 10.1038/tp.2012.10.

67. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP. Mobile detection of autism through machine learning on home video: A development and prospective validation study. *PLoS Med*. 2018;15(11):e1002705. doi: 10.1371/journal.pmed.1002705.

68. Wall DP, Dally R, Luyster R, Jung JY, Deluca TF. Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PLoS One.* 2012;7(8):e43855. doi: 10.1371/journal.pone.0043855.

69. Bone D, Bishop SL, Black MP, Goodwin MS, Lord C, Narayanan SS. Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *J Child Psychol Psychiatry.* 2016;57(8):927-937. doi: 10.1111/jcpp.12559.

70. Bone D, Goodwin MS, Black MP, Lee CC, Audhkhasi K, Narayanan S. Applying machine learning to facilitate autism diagnostics: pitfalls and promises. *J Autism Dev Disord.* 2015;45(5):1121-1136. doi: 10.1007/s10803-014-2268-6.

71. Washington P, Chrisman B, Leblanc E, et al. Crowd annotations can approximate clinical autism impressions from short home videos with privacy protections. *Intell Based Med.* 2022;6:None. doi: 10.1016/j.ibmed.2022.100056.

72. Washington P, Leblanc E, Dunlap K, et al. Precision telemedicine through crowdsourced machine learning: testing variability of crowd workers for video-based autism feature recognition. *J Pers Med.* 2020;10(3):13. doi: 10.3390/jpm10030086.

73. Washington P, Tariq Q, Leblanc E, et al. Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection. *Sci Rep.* 2021;11(1):7620. doi: 10.1038/s41598-021-87059-4.

74. Duda M, Haber N, Daniels J, Wall DP. Crowdsourced validation of a machine-learning classification system for autism and ADHD. *Transl Psychiatry.* 2017;7(5):e1133. doi: 10.1038/tp.2017.86.

75. Duda M, Ma R, Haber N, Wall DP. Use of machine learning for behavioral distinction of autism and ADHD. *Transl Psychiatry.* 2016;6(2):e732. doi: 10.1038/tp.2015.221.

76. Tariq Q, Fleming SL, Schwartz JN, et al. Detecting developmental delay and autism through machine learning models using home videos of Bangladeshi children: development and validation study. *J Med Internet Res.* 2019;21(4):e13822. doi: 10.2196/13822.

77. Wazny K. Applications of crowdsourcing in health: an overview. *J Glob Health.* 2018;8(1):010502. doi: 10.7189/jogh.08.010502.

78. US Food and Drug Administration. De novo classification request for Cognoa ASD diagnosis aid. 2020; https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200069.pdf. Accessed April 28, 2024.

79. US Food and Drug Administration. Medical device De Novo classification process. 2021; https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/medical-device-de-novo-classification-process. Accessed April 28, 2024.

80. Aboy M, Crespo C, Stern A. Beyond the 510(k): the regulation of novel moderate-risk medical devices, intellectual property considerations, and innovation incentives in the FDA's De Novo pathway. *NPJ Digit Med.* 2024;7(1):29. doi: 10.1038/s41746-024-01021-y.

81. Fargen KM, Frei D, Fiorella D, et al. The FDA approval process for medical devices: an inherently flawed system or a valuable pathway for innovation? *J Neurointerv Surg.* 2013;5(4):269-275. doi: 10.1136/neurintsurg-2012-010400.

82. Johnston JL, Dhruva SS, Ross JS, Rathi VK. Clinical evidence supporting US Food and Drug Administration clearance of novel therapeutic devices via the de novo pathway between 2011 and 2019. *JAMA Intern Med.* 2020;180(12):1701-1703. doi: 10.1001/jamainternmed.2020.3214.

83. Joshi G, Aditi J, Reddy Araveeti S, Adhikari S, Garg H, M B. FDA-approved artificial intelligence and machine learning (AI/ML)-enabled medical devices: an updated landscape. *Electronics.* 2023;13:498. doi: 10.3390/electronics13030498.

84. Van Norman GA. Drugs, devices, and the FDA: Part 2: an overview of approval processes: FDA approval of medical devices. *JACC Basic Transl Sci.* 2016;1(4):277-287. doi: 10.1016/j.jacbts.2016.03.009.

85. US Senate. Medicaid and CHIP Access to Prescription Digital Therapeutics Act, S 5238, 117th Congress. 2022; https://www.congress.gov/bill/117th-congress/senate-bill/5238/text. Accessed March 25, 2024.

86. US House. Access to Prescription Digital Therapeutics Act of 2023, HR 1458, 118th Congress. 2023; https://www.congress.gov/bill/118th-congress/house-bill/1458. Accessed March 25, 2024.

87. US Senate. Access to Prescription Digital Therapeutics Act of 2023, S 723, 118th Congress. 2023; https://www.congress.gov/bill/118th-congress/senate-bill/723. Accessed March 25, 2024.

88. Anonymous. Access to Prescription Digital Therapeutics Act of 2022, HR 7051, 117th Congress. 2022; https://www.congress.gov/bill/117th-congress/house-bill/7051. Accessed April 1, 2024.

89. Centers for Medicare & Medicaid Services. Centers for Medicare & Medicaid Services (CMS) Healthcare Common Procedure Coding System (HCPCS) application summaries and coding recommendations second biannual, 2021 HCPCS coding cycle. 2021;

https://www.cms.gov/files/document/2021-hcpcs-application-summary-biannual-2-2021-non-drug-and-non-biological-items-and-services.pdf. Accessed May 22, 2024.

90. Centers for Medicare & Medicaid Services. Centers for Medicare & Medicaid Services' (CMS') Healthcare Common Procedure Coding System (HCPCS) Level II final coding, benefit category and payment determinations first biannual (B1), 2023 HCPCS coding cycle. 2023; https://www.cms.gov/files/document/2023-hcpcs-application-summary-biannual-1-2023-non-drug-and-non-biological-items-and-services.pdf. Accessed May 22, 2024.

91. Centers for Medicare & Medicaid Services. Centers for Medicare & Medicaid Services' (CMS') Healthcare Common Procedure Coding System (HCPCS) Level II final coding, benefit category and payment determinations. 2023; https://www.cms.gov/files/document/2023-hcpcs-application-summary-biannual-2-2023-non-drug-and-non-biological-items-and-services-posted.pdf. Accessed May 22, 2024.

92. Centers for Medicare & Medicaid Services. Centers for Medicare & Medicaid Services' (CMS') first biannual 2022 Healthcare Common Prcedure Coding System (HCPCS) public meeting agenda. 2022; https://www.cms.gov/files/document/hcpcs-public-meeting-agenda-non-drug-and-non-biological-items-and-services-november-29-2022.pdf. Accessed April 8, 2024.

93. Micromedex Red Book. Canvas Dx. Merative; 2024; https://www.micromedexsolutions.com/home/dispatch/. Accessed October 22, 2024.

94. Self T, Parham D, Rajagopalan J. Autism spectrum disorder early screening practices: a survey of physicians. *Comm Disord Q.* 2015;36(4):195-207. doi: 10.1177/1525740114560060.

95. Mazurek MO, Kuhlthau K, Parker RA, Chan J, Sohl K. Autism and general developmental screening practices among primary care providers. *J Dev Behav Pediatr.* 2021;42(5):355-362. doi: 10.1097/dbp.0000000000000909.

96. Department of Health and Human Services. EPSDT - A guide for states: coverage in the Medicaid benefit for children and adolescents. 2014; https://www.medicaid.gov/sites/default/files/2019-12/epsdt_coverage_guide.pdf. Accessed March 27, 2024.

97. New York State Department of Health. New York Medicaid Child/Teen Health Program (C/THP) provider manual. 2019; https://www.emedny.org/ProviderManuals/EPSDTCTHP/PDFS/EPSDT-CTHP.pdf. Accessed March 29, 2024.

98. American Academy of Pediatrics. Autism spectrum disorder: links to commonly used screening instruments and tools. 2021; https://publications.aap.org/toolkits/pages/asd-screening-tools. Accessed September 13, 2024.

99. New York State Department of Health. Clinical practice guideline on assessment and intervention services for young children with autism spectrum disorders (ASD) 2017 update. 2017; https://www.health.ny.gov/publications/20152.pdf. Accessed March 29, 2024.

100. New York State. Public Health Law Chapter 45, Article 25, Title 1, Section 2500-J. Autism spectrum disorders; screening of children. 2018; https://www.nysenate.gov/legislation/laws/PBH/2500-J. Accessed September 13, 2024.

101. New York State Department of Health. Best practice protocol for early screening of young children for autism spectrum disorders (ASDs) by pediatric primary care providers. 2019; https://www.health.ny.gov/community/infants_children/early_intervention/autism/docs/best_practice_protocol.pdf. Accessed March 29, 2024.

102. New York State Department of Health. New York state clinical practice guideline on assessment and intervention services for young children (ages 0-3) with autism spectrum disorders: update - 2017 report of the research of evidence. 2017; https://www.health.ny.gov/community/infants_children/early_intervention/autism/docs/report_of_the_research.pdf. Accessed March 29, 2024.

103. New York State Department of Health. List of developmental assessment instruments. 2023; https://www.health.ny.gov/community/infants_children/early_intervention/docs/2023-05_list_developmental_assessment_instruments.pdf. Accessed March 29, 2024.

104. New York State. Public Health Law Chapter 45, Article 25, Title 2-A. Early Intervention Program for infants and toddlers with disabilities and their families. 2022; https://www.nysenate.gov/legislation/laws/PBH/A25T2-A. Accessed September 13, 2024.

105. New York State. New York codes, rules, and regulations: section 69-4.8 - evaluation and screening of the child and assessment of the child and family. 2024; https://regs.health.ny.gov/content/section-69-48-evaluatorsscreening-evaluation-and-assessment-responsibilities. Accessed March 29, 2024.

106. New York State Department of Health. Coverage for developmental screening, including autism spectrum disorder, in the "first three years of life". 2021; https://www.health.ny.gov/health_care/medicaid/program/update/2021/no14_2021-12.htm. Accessed 14.

107. New York State Department of Health. NYS Medicaid physician medicine services fee schedule. 2024; https://www.emedny.org/ProviderManuals/Physician/PDFS/Physician_Manual_Fee_Schedule_Sect2.xls. Accessed September 13, 2024.

108. 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.* 2016;63(5):851-859. doi: 10.1016/j.pcl.2016.06.007.

109. Smith JV, Menezes M, Brunt S, Pappagianopoulos J, Sadikova E, M OM. Understanding autism diagnosis in primary care: rates of diagnosis from 2004 to 2019 and child age at diagnosis. *Autism.* 2024:13623613241236112. doi: 10.1177/13623613241236112.

110. Monteiro SA, Dempsey J, Berry LN, Voigt RG, Goin-Kochel RP. Screening and referral practices for autism spectrum disorder in primary pediatric care. *Pediatrics.* 2019;144(4). doi: 10.1542/peds.2018-3326.

111. Sohl K, Kilian R, Brewer Curran A, et al. Feasibility and impact of integrating an artificial intelligence-based diagnosis aid for autism into the Extension for Community Health Outcomes Autism Primary Care Model: protocol for a prospective observational study. *JMIR Res Protoc.* 2022;11(7):e37576. doi: 10.2196/37576.

112. American Academy of Pediatrics. Autism diagnosis in primary care. 2023; https://www.aap.org/en/patient-care/autism/autism-diagnosis-in-primary-care/. Accessed March 5, 2024.

113. Mazurek MO, Curran A, Burnette C, Sohl K. ECHO Autism STAT: accelerating early access to autism diagnosis. *J Autism Dev Disord.* 2019;49(1):127-137. doi: 10.1007/s10803-018-3696-5.

114. American Psychiatric Association. App Advisor: an American Psychiatric Association initiative. https://www.psychiatry.org/psychiatrists/practice/mental-health-apps. Accessed March 20, 2024.

115. Ambrose M. Ensuring public trust in digital therapeutics: a pharmacopeial perspective. *J Manag Care Spec Pharm.* 2021;27(4):533-535. doi: 10.18553/jmcp.2021.27.4.533.

116. Academy of Managed Care Pharmacy. AMCP partnership forum: digital therapeutics-what are they and where do they fit in pharmacy and medical benefits? *J Manag Care Spec Pharm.* 2020;26(5):674-681. doi: 10.18553/jmcp.2020.1941.

117. Academy of Managed Care Pharmacy. AMCP partnership forum: the evolving role of digital therapeutics. *J Manag Care Spec Pharm.* 2022;28(7):804-810. doi: 10.18553/jmcp.2022.22093.

118. Ambrose M, Seiler D, Barrett B, Yu C, Poldolosky D, Levy M. The role of public standards in assuring quality of digital therapeutics. US Pharmacopeia; 2020; https://qualitymatters.usp.org/sites/default/files/user-uploaded-files/USP_Digital_Therapeutics_Paper_2020-06-11.pdf. Accessed March 27, 2024.

119. Patel NA, Butte AJ. Characteristics and challenges of the clinical pipeline of digital therapeutics. *NPJ Digit Med.* 2020;3(1):159. doi: 10.1038/s41746-020-00370-8.

120. Rassi-Cruz M, Valente F, Caniza MV. Digital therapeutics and the need for regulation: how to develop products that are innovative, patient-centric and safe. *Diabetol Metab Syndr.* 2022;14(1):48. doi: 10.1186/s13098-022-00818-9.

121. Parcher B, Coder M. Decision makers need an approach to determine digital therapeutic product quality, access, and appropriate use. *J Manag Care Spec Pharm.* 2021;27(4):536-538. doi: 10.18553/jmcp.2021.27.4.536.

122. Salsabili M, Tesell M, Alcusky M, et al. Prescription digital therapeutics: applying Medicaid experience to value assessment and formulary management. *J Manag Care Spec Pharm.* 2023;29(6):685-691. doi: 10.18553/jmcp.2023.29.6.685.

123. Rodriguez-Villa E, Torous J. Regulating digital health technologies with transparency: the case for dynamic and multi-stakeholder evaluation. *BMC Med.* 2019;17(1):226. doi: 10.1186/s12916-019-1447-x.

124. Substance Abuse and Mental Health Services Administration. Digital therapeutics for management and treatment in behavioral health publication no. PEP23-06-00-001. 2023; https://store.samhsa.gov/sites/default/files/pep23-06-00-001.pdf. Accessed March 13, 2024.

125. Brezing CA, Brixner DI. The rise of prescription digital therapeutics in behavioral health. *Adv Ther.* 2022;39(12):5301-5306. doi: 10.1007/s12325-022-02320-0.

126. Kumar A, Ross JS, Patel NA, Rathi V, Redberg RF, Dhruva SS. Studies of prescription digital therapeutics often lack rigor and inclusivity. *Health Aff (Millwood).* 2023;42(11):1559-1567. doi: 10.1377/hlthaff.2023.00384.

127. Nuske HJ, Mandell DS. Digital health should augment (not replace) autism treatment providers. *Autism.* 2021;25(7):1825-1827. doi: 10.1177/13623613211043368.

128. Dortche K, McCarthy G, Banbury S, Yannatos I. Promoting health equity through improved regulation of artificial intelligence medical devices. *J Sci Policy Gov.* 2023;21(3). doi: 20.38126/JSPG210302.

129. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med.* 2018;178(11):1544-1547. doi: 10.1001/jamainternmed.2018.3763.

130. Hoffman S. The emerging hazard of AI-related health care discrimination. *Hastings Cent Rep.* 2021;51(1):8-9. doi: 10.1002/hast.1203.

131. Kadakia KT, Rathi VK, Ramachandran R, Johnston JL, Ross JS, Dhruva SS. Challenges and solutions to advancing health equity with medical devices. *Nat Biotechnol.* 2023;41(5):607-609. doi: 10.1038/s41587-023-01746-3.

132. Fox-Rawlings SR, Gottschalk LB, Doamekpor LA, Zuckerman DM. Diversity in medical device clinical trials: do we know what works for which patients? *Milbank Q.* 2018;96(3):499-529. doi: 10.1111/1468-0009.12344.

133. Ponzo S, May M, Tamayo-Elizalde M, et al. App characteristics and accuracy metrics of available digital biomarkers for autism: scoping review. *JMIR Mhealth Uhealth.* 2023;11(1). doi: 10.2196/52377.

134. US Preventive Services Task Force. Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *JAMA.* 2016;315(7):691-696. doi: 10.1001/jama.2016.0018.

135. US Preventive Services Task Force. Recommendations in progress. 2024; https://www.uspreventiveservicestaskforce.org/uspstf/topic_search_results?category%5B%5D=17&searchterm=. Accessed May 28, 2024.

136. Chassé M, Fergusson DA. Diagnostic Accuracy Studies. *Seminars in Nuclear Medicine.* 2019;49(2):87-93. doi: https://doi.org/10.1053/j.semnuclmed.2018.11.005.

137. Shreffler J, Huecker MR. Diagnostic testing accuracy: sensitivity, specificity, predictive values and likelihood ratios. *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2024.

138. NCT03871179. Cognoa ASD Diagnostic Device - Data Collection Study. ClinicalTrials.gov; 2024; https://clinicaltrials.gov/study/NCT03871179?tab=table. Accessed July 16, 2024.

139. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev.* 2017;2(2):Mr000033. doi: 10.1002/14651858.MR000033.pub3.

140. Kanne SM, Carpenter LA, Warren Z. Screening in toddlers and preschoolers at risk for autism spectrum disorder: Evaluating a novel mobile-health screening tool. *Autism Res.* 2018;11(7):1038-1049. doi: 10.1002/aur.1959.

141. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ*. 2013;346:f2778. doi: 10.1136/bmj.f2778.

142. Makhtar M, Neagu DC, Ridley MJ. Comparing multi-class classifiers: on the similarity of confusion matrices for predictive toxicology applications. Springer Berlin Heidelberg; 2011; https://link.springer.com/chapter/10.1007/978-3-642-23878-9_31.

143. Stahlmann K, Reitsma JB, Zapf A. Missing values and inconclusive results in diagnostic studies - A scoping review of methods. *Stat Methods Med Res*. 2023;32(9):1842-1855. doi: 10.1177/09622802231192954.

144. Leblanc E, Washington P, Varma M, et al. Feature replacement methods enable reliable home video analysis for machine learning detection of autism. *Sci Rep*. 2020;10(1):21245. doi: 10.1038/s41598-020-76874-w.

145. Aetna. Autism spectrum disorders. 2024; https://www.aetna.com/cpb/medical/data/600_699/0648.html. Accessed October 22, 2024.

146. Molina Healthcare. Prescription digital therapeutics: policy no. 412. 2023; https://www.molinahealthcare.com/-/media/Molina/PublicWebsite/PDF/Common/Molina-Clinical-Policy/Policies-to-be-deleted/MCP-APR-2022/Prescription-Digital-Therapeutics_R.pdf. Accessed September 6, 2024.

147. Gliadkovskaya A. Cognoa's FDA-approved autism diagnostic tool now in network for Highmark members. Fierce Healthcare; 2024; https://www.fiercehealthcare.com/payers/cognoas-autism-diagnostic-tool-now-network-highmark-members. Accessed March 26, 2024.

148. Cognoa. Canvas Dx for Highmark families. 2024; https://cognoa.com/canvas-dx-for-highmark-families/. Accessed September 6, 2024.

149. Kim SH, Macari S, Koller J, Chawarska K. Examining the phenotypic heterogeneity of early autism spectrum disorder: subtypes and short-term outcomes. *J Child Psychol Psychiatry*. 2016;57(1):93-102. doi: 10.1111/jcpp.12448.

150. Washington P, Park N, Srivastava P, et al. Data-driven diagnostics and the potential of mobile artificial intelligence for digital therapeutic phenotyping in computational psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2020;5(8):759-769. doi: 10.1016/j.bpsc.2019.11.015.

151. Abbas H, Garberson F, Glover E, Wall DP. Machine learning approach for early detection of autism by combining questionnaire and home video screening. *J Am Med Inform Assoc*. 2018;25(8):1000-1007. doi: 10.1093/jamia/ocy039.

152. Duda M, Daniels J, Wall DP. Clinical Evaluation of a Novel and Mobile Autism Risk Assessment. *J Autism Dev Disord*. 2016;46(6):1953-1961. doi: 10.1007/s10803-016-2718-4.

153. Highmark Inc. Prescription digital therapeutics. 2024; https://securecms.highmark.com/content/medpolicy/en/highmark/ny/commercial/policies/Miscellaneous/Z-105/Z-105-007.html. Accessed October 22, 2024.

Appendix A. Search Methods

Clinical Evidence Sources and Search Strategies

We searched selected bibliographic databases and grey literature sources using key words such as *autism*, *developmental delay*, *neurodevelopmental disorder*, *diagnosis*, *mobile app*, *software as a medical device*, and *Canvas Dx* to identify diagnostic accuracy studies, randomized controlled trials, nonrandomized comparative trials, prospective cohort studies, cost-effectiveness studies, and clinical practice guidelines. We did not use date limits, but we did limit search results to publications available in the English language. Searches were conducted March 3, 2024, through March 6, 2024.

Bibliographic Database Sources

- Cochrane Central Register of Controlled Trials (CENTRAL)
- Cochrane Database of Systematic Reviews (CDSR)
- EBSCO Cumulated Index to Nursing and Allied Health Literature (CINAHL)
- Ovid MEDLINE
- Ovid PsycInfo

Evidence Synthesis Sources

- Agency for Healthcare Research and Quality (AHRQ)
- Canada's Drug Agency
- Epistemonikos
- Health Quality Ontario
- Institute for Clinical and Economic Review (ICER)
- Institute for Health Quality and Efficiency in Health Care
- International HTA Database
- National Institute for Health and Care Excellence (NICE)
- Oregon Health Evidence Review Commission (HERC)
- Veterans Administration Evidence Synthesis Program (ESP)
- Washington Health Technology Assessment

Clinical Practice Guideline Sources

- American Academy of Child and Adolescent Psychiatry
- American Academy of Family Physicians (AAFP)
- American Academy of Pediatrics (AAP)
- American Medical Association (AMA)
- American Psychological Association
- American Psychiatric Association
- Canadian Pediatric Society
- Guidelines International Network (GIN) International Guidelines Library
- Scottish Intercollegiate Guidelines Network (SIGN)
- US Preventive Services Task Force (USPSTF)
- Veterans Administration/Department of Defense Clinical Practice Guidelines

Clinical Trial Sources

- ClinicalTrials.gov
- ScanMedicine

Regulatory Body and Manufacturer Sources

- Cognoa
- US Food and Drug Administration Manufacturer and User Facility Device Experience (MAUDE)

Ovid MEDLINE ALL Search Strategy

1 exp child development disorders, pervasive/
2 neurodevelopmental disorders/
3 developmental disabilities/
4 autis*.ti,ab,kf.
5 asperger*.ti,ab,kf.
6 kanner*.ti,ab,kf.
7 ((development* or neurodevelopment*) adj (delay* or disorder*)).ti,ab,kw.
8 (pervasive adj3 child*).ti,ab,kw.
9 (pdd-nos or pdd nos).ti,ab,kw.
10 or/1-9
11 algorithms/
12 artificial intelligence/
13 exp machine learning/
14 software/
15 software design/
16 software validation/
17 mobile applications/
18 cell phone/
19 exp computers, handheld/
20 (artificial adj1 intelligen*).ti,ab,kw.
21 (machine adj2 learn*).ti,ab,kw.
22 (software adj3 medical device?).ti,ab,kw.
23 samd.ti,ab,kf.
24 ((cell* or mobile* or smart*) adj1 (device* or phone* or telephone* or technolog*)).ti,ab,kw.
25 (cellphone* or cell-phone* or smartphone* or smart-phone*).ti,ab,kf.
26 mobile app*.ti,ab,kw.
27 (canvasdx or canvas-dx or canvas dx).ti,ab,kf.
28 cognoa.ti,ab,ia.

29 or/11-28
30 diagnosis/
31 diagnosis.fs.
32 exp clinical decision making/
33 delayed diagnosis/
34 diagnosis, computer-assisted/
35 diagnosis, differential/
36 exp diagnostic errors/
37 "direct-to-consumer screening and testing"/
38 exp mass screening/
39 early diagnosis/
40 overdiagnosis/
41 exp "sensitivity and specificity"/
42 diagnos*.ti,kf.
43 (overdiagnos* or over-diagnos*).ti,ab,kf.
44 (misdiagnos* or mis-diagnos*).ti,ab,kf.
45 screening.ti,kf.
46 ((device or test) adj3 (sensitiv* or specific*)).ti,ab,kw.
47 ((negative or positive) adj3 predict* value?).ti,ab,kw.
48 (npv or ppv).ti,ab,kf.
49 or/30-48
50 clinical decision rules/
51 exp clinical protocols/
52 consensus/
53 exp consensus development conferences as topic/
54 critical pathways/
55 decision making, shared/
56 exp guidelines as topic/
57 health planning guidelines/
58 consensus development conference.pt.
59 consensus development conference, NIH.pt.
60 guideline.pt.
61 practice guideline.pt.
62 guideline?.ti,kf.
63 ((committee or executive) adj2 (recommendation* or statement* or summar*)).ti,kw.
64 (consensus adj2 (document* or paper* or recommendation* or report* or statement)).ti,kw.

65 (joint adj2 (document* or recommendation* or statement*)).ti,kw.
66 ((policy or position) adj2 (paper* or statement*)).ti,kw.
67 ((clinical or critical or practice) adj2 (pathway? or standard?)).ti,kw.
68 or/50-67
69 infant/
70 exp child/
71 pediatrics/
72 infan*.ti,ab,kf.
73 child*.ti,ab,kf.
74 toddler?.ti,ab,kf.
75 (preschool* or pre-school* or pre school*).ti,ab,kf.
76 (nursery adj2 school*).ti,ab,kw.
77 (kindergarten* or kinder-garten* or kinder garten*).ti,ab,kf.
78 p?ediatric*.ti,ab,kf.
79 or/69-78
80 and/10,29,49,79
81 and/10,49,68,79
82 limit 81 to yr="2019 -Current"
83 or/80,82
84 (exp animals/ not humans/) or (baboon? or bovine? or canine? or cat? or chimpanzee? or cow? or dog? or feline? or fish or goat? or hens or macque? or mice or monkey? or mouse or murine? or ovine or pig? or porcine or primate? or sheep or rabbit? or rat or rats or rattus or rhesus or rodent? or zebrafish).ti.
85 83 not 84
86 limit 85 to english language

Cochrane Database of Systematic Reviews (CDSR) and Cochrane Central Register of Controlled Trials (CENTRAL) via the Cochrane Library Search Strategy

1 [mh "child development disorders, pervasive"]
2 [mh ^"neurodevelopmental disorders"]
3 [mh ^"developmental disabilities"]
4 autis*:ti,ab,kw
5 asperger*:ti,ab,kw
6 kanner*:ti,ab,kw
7 ((development* OR neurodevelopment*) NEXT (delay* OR disorder*)):ti,ab,kw
8 (pervasive NEAR/3 child*):ti,ab,kw
9 (pdd-nos OR "pdd nos"):ti,ab,kw
10 [OR #1-#9]

11 [mh ^algorithms]
12 [mh ^"artificial intelligence"]
13 [mh "machine learning"]
14 [mh ^software]
15 [mh ^"software design"]
16 [mh ^"software validation"]
17 [mh ^"mobile applications"]
18 [mh ^"cell phone"]
19 [mh "computers, handheld"]
20 (artificial NEAR/1 intelligen*):ti,ab,kw
21 (machine NEAR/2 learn*):ti,ab,kw
22 (software NEAR/3 "medical device"):ti,ab,kw
23 (samd):ti,ab,kw
24 ((cell* OR mobile* OR smart*) NEAR/1 (device* OR phone* OR telephone* OR technolog*)):ti,ab,kw
25 (cellphone* OR cell-phone* OR smartphone* OR smart-phone*):ti,ab,kw
26 (mobile NEXT app*):ti,ab,kw
27 (canvasdx OR canvas-dx OR "canvas dx"):ti,ab,kw
28 (cognoa):ti,ab,kw
29 [OR #11-#28]
30 [mh ^diagnosis]
31 [mh /DI]
32 [mh "clinical decision making"]
33 [mh ^"delayed diagnosis"]
34 [mh ^"diagnosis, computer-assisted"]
35 [mh ^"diagnosis, differential"]
36 [mh "diagnostic errors"]
37 [mh ^"direct-to-consumer screening and testing"]
38 [mh "mass screening"]
39 [mh ^"early diagnosis"]
40 [mh ^overdiagnosis]
41 [mh "sensitivity and specificity"]
42 diagnos*:ti,kw
43 (overdiagnos* OR over-diagnos*):ti,ab,kw
44 (misdiagnos* OR mis-diagnos*):ti,ab,kw
45 screen*:ti,kw
46 ((device OR test) NEAR/3 (sensitiv* OR specific*)):ti,ab,kw

47 ((negative OR positive) NEAR/3 (predict* NEXT value?)):ti,ab,kw

48 (npv OR ppv):ti,ab,kw

49 [OR #30-#48]

50 [AND #10, #29, #49] in Trials

51 [AND #10, #49] in Cochrane Reviews

CINAHL Plus with Full Text

S1 (MH "Child Development Disorders, Pervasive+")

S2 (MH "Child Development Disorders+")

S3 (MH "Developmental Disabilities")

S4 TI(autis*) OR AB(autis*)

S5 TI(asperger*) OR AB(asperger*)

S6 TI(kanner*) OR AB(kanner*)

S7 TI((development* OR neurodevelopment*) N1 (delay* OR disorder*)) OR AB((development* OR neurodevelopment*) N1 (delay* OR disorder*))

S8 TI(pervasive N3 child*) OR AB(pervasive N3 child*)

S9 TI(pdd-nos OR "pdd nos") OR AB(pdd-nos OR "pdd nos")

S10 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9

S11 (MH "Algorithms")

S12 (MH "Artificial Intelligence")

S13 (MH "Machine Learning+")

S14 (MH "Software")

S15 (MH "Software Design")

S16 (MH "Mobile Applications")

S17 (MH "Cellular Phone")

S18 (MH "Computers, Hand-Held+")

S19 TI(artificial N1 intelligen*) OR AB(artificial N1 intelligen*)

S20 TI(machine N2 learn*) OR AB(machine N2 learn*)

S21 TI(software N3 "medical device*") OR AB(software N3 "medical device*")

S22 TI(samd) OR AB(samd)

S23 TI((cell* OR mobile* OR smart*) N1 (device* OR phone* OR telephone* OR technolog*)) OR AB((cell* OR mobile* OR smart*) N1 (device* OR phone* OR telephone* OR technolog*))

S24 TI(cellphone* OR cell-phone* OR smartphone* OR smart-phone*) OR AB(cellphone* OR cell-phone* OR smartphone* OR smart-phone*)

S25 TI("mobile app*") OR AB("mobile app*")

S26 TI(canvasdx OR canvas-dx OR "canvas dx") OR AB(canvasdx OR canvas-dx OR "canvas dx")

S27 TI(cognoa) OR AB(cognoa) OR AF(cognoa)

S28 S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27

S29 (MH "Diagnosis")

S30 (MH "Decision Making, Clinical+")

S31 (MH "Diagnosis, Computer Assisted")

S32 (MH "Diagnosis, Delayed")

S33 (MH "Diagnosis, Developmental+")

S34 (MH "Diagnosis, Differential")

S35 (MH "Diagnostic Errors+")

S36 (MH "Early Diagnosis")

S37 (MH "Home Diagnostic Tests")

S38 (MH "Overdiagnosis")

S39 (MH "Health Screening")

S40 (MH "Sensitivity and Specificity")

S41 (MH "Predictive Value of Tests")

S42 TI(diagnos*)

S43 TI(overdiagnos* OR over-diagnos*) OR AB(overdiagnos* OR over-diagnos*)

S44 TI(misdiagnos* OR mis-diagnos*) OR AB(misdiagnos* OR mis-diagnos*)

S45 TI(screening)

S46 TI((device OR test) N3 (sensitiv* OR specific*)) OR AB((device OR test) N3 (sensitiv* OR specific*))

S47 TI((negative OR positive) N3 "predict* value*") OR AB((negative OR positive) N3 "predict* value*")

S48 TI(npv OR ppv) OR AB(npv OR ppv)

S49 S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48

S50 (MH "Child")

S51 (MH "Child, Preschool")

S52 (MH "Infant")

S53 TI(child*) OR AB(child*)

S54 TI(toddler*) OR AB(toddler*)

S55 TI(nursery N2 school*) OR AB(nursery N2 school*)

S56 TI(preschool* OR pre-school* OR "pre school*") OR AB(preschool* OR pre-school* OR "pre school*")

S57 TI(kindergarten* OR kinder-garten* OR "kinder garten*) OR AB(kindergarten* OR kinder-garten* OR "kinder garten*)

S58 TI(infan*) OR AB(infan*)

S59 TI(p#ediatric*) OR AB(p#ediatric*)

S60 S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59

S61 S10 AND S28 AND S49 AND S60

S62 S61
Limiters – English Language
Source Types – Academic Journals

PsycInfo

1 exp autism spectrum disorders/

2 neurodevelopmental disorders/

3 developmental disabilities/

4 autis*.ti,ab,id.

5 asperger*.ti,ab,id.

6 kanner*.ti,ab,id.

7 ((development* or neurodevelopment*) adj2 (delay* or disorder*)).ti,ab,id.

8 (pervasive adj3 child*).ti,ab,id.

9 (pdd-nos or pdd nos).ti,ab,id.

10 or/1-9

11 artificial intelligence/

12 machine learning/

13 exp machine learning algorithms/

14 computer software/

15 mobile applications/

16 exp mobile devices/

17 (artificial adj1 intelligen*).ti,ab,id.

18 (machine adj2 learn*).ti,ab,id.

19 (software adj3 medical device?).ti,ab,id.

20 samd.ti,ab,id.

21 ((cell* or mobile* or smart*) adj1 (device* or phone* or telephone* or technolog*)).ti,ab,id.

22 (cellphone* or cell-phone* or smartphone* or smart-phone*).ti,ab,id.

23 mobile app*.ti,ab,id.

24 (canvasdx or canvas-dx or canvas dx).ti,ab,id.

25 cognoa.ti,ab,in.

26 or/11-25

27 diagnosis/

28 differential diagnosis/

29 medical diagnosis/

30 exp computer assisted diagnosis/

31 misdiagnosis/
32 exp health screening/
33 test sensitivity/
34 test specificity/
35 diagnos*.ti,id.
36 (overdiagnos* or over-diagnos*).ti,ab,id.
37 (misdiagnos* or mis-diagnos*).ti,ab,id.
38 screening.ti,id.
39 ((device or test) adj3 (sensitiv* or specific*)).ti,ab,id.
40 ((negative or positive) adj3 predict* value?).ti,ab,id.
41 (npv or ppv).ti,ab,id.
42 or/27-41
43 and/10,26,42
44 limit 43 to english language
45 limit 44 to ("0100 journal" or "0110 peer-reviewed journal" or "0120 non-peer-reviewed journal" or "0130 peer-reviewed status unknown" or "0500 electronic collection")

Policy Sources and Search Terms

We searched websites for the state Medicaid programs and private payers listed below using terms such as *digital diagnostic*, *mobile app*, *software as a medical device*, and *Canvas Dx*.

State Medicaid Programs

- California Medicaid
- Florida Medicaid
- Massachusetts Medicaid
- New Jersey Medicaid
- New York Medicaid
- North Carolina Medicaid
- Oregon Medicaid and the Health Evidence Review Commission (HERC) coverage guidance (including topics under consideration)
- Pennsylvania Medicaid
- Texas Medicaid
- Washington Medicaid and the Washington Health Technology Assessment Program coverage determinations (including topics under consideration)

Private Payers

- Aetna
- Anthem Blue Cross and Blue Shield (formerly Empire BlueCross BlueShield)
- Capital District Physicians' Health Plan
- Cigna
- EmblemHealth
- Excellus BlueCross BlueShield

- Fidelis Care
- Healthfirst
- Highmark Blue Shield of Northeastern New York
- MetroPlusHealth
- Molina Healthcare
- Tufts Health Plan
- UnitedHealthcare

Appendix B. Detailed Inclusion and Exclusion Criteria

Table B. Detailed Inclusion and Exclusion Criteria

Study Component	Inclusion	Exclusion
Populations	• Children aged 18 to 72 months suspected of having developmental delay based on concerns of a caregiver or health care provider	• Children younger than 18 months of age • Children older than 72 months of age • Children whose caregivers or health care providers do not have a concern about developmental delay
Interventions	• Use of Canvas Dx as an adjunct to the diagnostic process	• Any other diagnostic tool or process not using Canvas Dx • Other SaMD intended to diagnose ASD in young children
Comparators	• Evaluation by a qualified clinician, alone or with other professionals (e.g., speech and language pathologist, educator) that involves history, diagnostic tool, physical examination, and ancillary testing • Diagnostic tools, such as: ○ Autism Diagnostic Interview-Revised (ADI-R) ○ Autism Diagnostic Observation Schedule, Second edition (ADOS-2) ○ Childhood Autism Rating Scale, Second edition (CARS-2) ○ Developmental, Dimensional and Diagnostic Interview (3di) ○ Diagnostic Interview for Social and Communication Disorders (DISCO) ○ Gilliam Autism Rating Scale, Third edition (GARS-3) ○ Social Responsiveness Scale, Second edition (SRS-2)	• None listed
Outcomes	• Critical ○ Sensitivity ○ Specificity ○ Positive predictive value ○ Negative predictive value ○ Time to diagnosis ○ Time to ASD-specific service initiation • Important ○ Change in clinical management of diagnostic and referral processes	• Other outcomes
Timing and follow-up	• No minimum	• None listed
Setting	• Studies conducted in countries categorized as *very high* on the United Nations Human Development Index	• Studies conducted in countries not categorized as *very high* on the United Nations Human Development Index

Study Component	Inclusion	Exclusion
Study design	KQ1-KQ2 • Diagnostic accuracy studies • Randomized controlled trials • Nonrandomized comparative trials • Prospective cohort studies • Interrupted time series with comparison group • Controlled before-after studies KQ3 • Comparative studies and economic evaluations • Cost-effectiveness analyses • Economic modeling studies KQ4 • Evidence-based clinical practice guidelines with specific recommendations	• Studies without extractable data • Retrospective studies unless otherwise noted • Uncontrolled studies • Proof-of-principle studies, such as algorithm development
Sample size	• No minimum	• None listed
Publication type	• Peer-reviewed publication of primary study results • Published in the English language • Ancillary publications with additional comparative follow-up	• Abstracts, conference proceedings, posters, editorials, letters • Studies not formally peer reviewed (i.e., preprint publications) • Studies published in languages other than English • Studies that cannot be located • Duplicate publications of the same study that do not report different outcomes or follow-up times, or single-site reports from published multicenter studies

Abbreviations. ASD: autism spectrum disorder; KQ: key question; SaMD: software as a medical device.

Appendix C. Additional Methods

Table C1. Risk of Bias Assessment: Diagnostic Test Accuracy Studies

Domain	Domain Elements[a]
Patient Representation	• Spectrum of patients is representative of the patients who will receive the test in practice • Index test, its use, and interpretation are similar to the review question
Patient Selection	• Selection criteria are clearly described • Consecutive or random sample of patients were enrolled • Case-control design was not used • Study avoided inappropriate exclusions
Reference Standard	• Reference standard is likely to classify the condition correctly
Test Timing	• Period between the reference standard and index test is short enough to be reasonably sure the target condition did not change between the tests
Verification	• Whole sample, or a random selection of the sample, received verification using the same diagnostic reference standard
Use of Reference Standard	• All patients received the same reference standard, regardless of the index test result
Test Independence	• Reference standard was independent of the index test (i.e., the index test did not form part of the reference standard)
Interpretation of the Index Test	• Index test results were interpreted without knowledge of the results of the reference standard • If a threshold was used, it was prespecified
Interpretation of the Reference Standard	• Reference standard results were interpreted without knowledge of the results of the index test
Uninterpretable or Intermediate Test Results	• Uninterpretable or intermediate test results are reported
Withdrawals	• All patients enrolled were included in the analysis • Explanation is given for all withdrawals or losses from the study
Interest Disclosure	• Disclosures of interest are given for authors/funders/commissioners of the study • Interests are unlikely to significantly affect study validity
Funding Source	• There is a description of source(s) of funding • Funding source is unlikely to have a significant impact on study validity

Note. [a] The elements included in each domain are assessed and rated as yes, no, unclear, or not applicable based on performance and documentation of the individual elements in each domain. The overall risk of bias for the study is assessed as high, moderate, or low based on assessment of how well the overall study methods and processes were performed to limit bias and ensure validity.

Table C2. GRADE System for Rating the Certainty of Evidence for Outcomes for Diagnostic Accuracy Studies

GRADE Rating	Plain Language Description	Detailed Category Description
High	New research is very unlikely to change our understanding of the relationship between this outcome and the health technology.	Center researchers are very confident the estimate of the efficacy of the diagnostic tool for the outcome of interest lies close to the true effect. Typical sets of studies are randomized controlled trials with few or no limitations, and the estimate of effect is likely stable.
Moderate	New research may change our understanding of the relationship between this outcome and the health technology.	Center researchers are moderately confident in the estimate of the efficacy of the diagnostic tool for the outcome of interest. The true value is likely to be close to the estimate of the value, but there is a possibility it is different. Typical sets of studies are randomized controlled trials with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.
Low	New research is likely to change our understanding of the relationship between this outcome and the health technology.	Center researchers have little confidence in the estimate of the efficacy of the diagnostic tool related to the reported outcome. The true value may be substantially different from the estimated value. Typical sets of studies are randomized controlled trials with serious limitations or nonrandomized studies without special strengths.
Very low	New research is very likely to change our understanding of the relationship between this outcome and the health technology.	Center researchers have no confidence in the estimate of the efficacy of the diagnostic tool for the outcome of interest. The true value is likely to be substantially different from the estimated value. Typical sets of studies are nonrandomized studies with serious limitations or inconsistent results across studies.
Not applicable	There is no research to report.	Center researchers did not identify any eligible articles.

Source. Adapted from 2 publications about GRADE.[17,18]
Abbreviation: GRADE: Grading of Recommendations, Assessment, Development, and Evaluations.

Appendix D. Evidence Tables

Table D. Accuracy of the Canvas Dx Diagnostic Tool

Publication Author, Year Risk of Bias Study Objective Comparison	Input Participants analyzed	Outcomes of Interest Reported
Wall et al., 2023[13] Risk of bias: High Study objective: To conduct an algorithmic threshold optimization procedure to improve the device's ability to detect or rule out autism in children aged 18 to 72 months without changing the accuracy or intended use of the device Comparison: Performance of Canvas Dx version 1 vs. Canvas Dx version 2 vs. the clinical reference standard	Enrolled: N not disclosed Analyzed: N = 722 (425 of which come from the same dataset used by Megerian et al.)[16] • In algorithm refinement, 70% of cases were used for algorithm training and 30% for algorithm testing • It is unclear if the diagnostic accuracy outcomes were based on the test subset or the full sample	PPV (95% CI) for subset with determinate results: • V1, 89.7% (83.9% to 90.3%) • V2, 87.5% (82.5% to 96.7%) • Performance difference, -2.0% (-4.4% to 0.9%) NPV (95% CI) for subset with determinate results: • V1, 96.1% (93.4% to 98.6%) • V2, 95.6% (93.7% to 97.9%) • Performance difference, -0.9% (-3.2% to 1.7%) Determinate result (ASD positive or ASD negative): • V1, 45.4% (41.3% to 48.6%) • V2, 66.5% (62.5% to 71.0%) • Performance difference, +21.2% (17.4% to 25.2%)
Megerian et al., 2022[16] Risk of bias: High Study objective: To evaluate the ability of the device to aid in the diagnosis of ASD in children aged 18 to 72 months Comparison: Canvas Dx output compared with the clinical reference standard	Enrolled N = 711 Analyzed N = 425	Determinate results (ASD positive or ASD negative) of Canvas Dx device: • 31.8% (135/425) For 135 participants with determinate output (95% CI): PPV: 80.8% (70.3% to 88.8%) NPV: 98.3% (90.6% to 100%) Sensitivity: 98.4% (91.6% to 100%) Specificity: 78.9% (67.6% to 87.7%) Difference in indeterminate rate by age, (95% CI) (N = 135)[a]: • < 3 years: 61% (54% to 68%) • ≥ 3 years: 74% (68% to 79%) • *P* = 0.006 Difference in specificity by age, (95% CI) (N = 135)[a]: • < 3 years: 67% (47% to 83%) • ≥ 3 years: 88% (74% to 96%) • *P* = 0.03 Differences by gender (95% CI)[a]: • Female PPV, 60% (36% to 81%) • Male PPV, 88% (77% to 95%) • Female NPV, 96% (80% to 100%) • Male NPV, 100% (89% to 100%) • Female sensitivity, 92% (64% to 100%) • Male sensitivity, 100% (93% to 100%)

Publication Author, Year Risk of Bias Study Objective Comparison	Input Participants analyzed	Outcomes of Interest Reported
		• Female specificity, 75% (57% to 89%) • Male specificity, 82% (66% to 92%) Differences by race/ethnicity (95% CI)[a]: • White PPV, 73% (56% to 86%) • Black PPV, 95% (75% to 100%) • Hispanic PPV, 74% (49% to 91%)) • White NPV, 97% (86% to 100%) • Black NPV, 100% (59% to 100%) • Hispanic NPV, 100% (69% to 100%) • White sensitivity, 96% (82% to 100%) • Black sensitivity, 100% (82% to 100%) • Hispanic sensitivity, 100% (77% to 100%) • White specificity, 79% (64% to 89%) • Black specificity, 88% (47% to 100%) • Hispanic specificity, 67% (38% to 88%)
Abbas et al., 2020[15] Risk of bias: High Study objective: To compare results of Canvas Dx with screening instruments used in clinical evaluation of ASD risk Comparison: Canvas Dx output compared with the clinical reference standard and results of screening instruments used in primary care	Enrolled N not disclosed Analyzed N = 375 (162 from first wave of enrollment; 213 from second wave of enrollment, following changes to the algorithm) Only children in the 2nd wave (N = 213) received all 3 components of Canvas Dx Data from a separate, unspecified number of de-identified ADOS and ADI-R score sheets of children between 18 and 84 months of age used for training the algorithm	ROC curves for the subset of the clinical sample with parent, video, and clinician modules, comparing predictive ability with conventional autism screeners. An inconclusive determination was allowed for up to 30% of cases and only conclusive results are included in accuracy calculations Confidence intervals, number of inconclusive results, and details on sensitivity, specificity, PPV, and NPV not included[22] Children 18 to 72 months: N = 204, AUC = 0.916 Subset of children < 4 years of age: N = 112, AUC = 0.837 Median time to completion Parent module, ~4 minutes Video scoring, 20 minutes Clinician module, 1.2 minutes

Notes. [a] *Study not powered for covariate analysis.*

Abbreviations. ADI-R: Autism Diagnostic Interview, Revised. ADOS: Autism Diagnostic Observation Schedule. ASD: autism spectrum disorder. AUC: area under the curve. CBCL: Child Behavior Checklist. CI: confidence interval. M-CHAT-R: Modified Checklist for Autism in Toddlers, Revised. NPV: negative predictive value. PPV: positive predictive value. SD: standard deviation. SRS-2: Social Responsiveness Scale, Second Edition.

Appendix E. Additional Evidence Regarding Canvas Development

We identified 20 articles published between 2012 and 2023 regarding development and testing of individual components of what would become the Canvas Dx program. One publication by Sohl and colleagues involves the FDA-authorized version of Canvas Dx, but is not tabled because it is a study protocol and does not give any results or outcomes.[111] The remaining publications are not tabled in the evidence review because they do not include all 3 components of the product that was authorized by the FDA: parent survey, video assessment, and health care provider assessment, as described below.

Table E. Other Articles Related to Varying Pieces of the Cognoa Diagnostic Device

Lead Author, Year	Title	Related Canvas module(s): *parent questionnaire, video assessment, clinician questionnaire*
Banerjee, 2023[22]	Training and profiling a pediatric facial expression classifier for children on mobile devices: machine learning study	Video[a]
Sohl, 2022[111]	Feasibility and impact of integrating an artificial intelligence-based diagnosis aid for autism into the extension for community health outcomes autism primary care model: protocol for a prospective observational study	Parent, video, clinician[b]
Washington, 2022[71]	Crowd annotations can approximate clinical autism impressions from short home videos with privacy protections	Video
Washington, 2021[73]	Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection	Video
Washington, 2020[72]	Precision telemedicine through crowdsourced machine learning: testing variability of crowd workers for video-based autism feature recognition	Video
Leblanc, 2020[144]	Feature replacement methods enable reliable home video analysis for machine learning detection of autism	Video
Washington, 2020[150]	Data-driven diagnostics and the potential of mobile artificial intelligence for digital therapeutic phenotyping in computational psychiatry	Video
Tariq, 2019[76]	Detecting developmental delay and autism through machine learning models using home videos of Bangladeshi children: development and validation study	Video[c]
Abbas, 2018[151]	Machine learning approach for early detection of autism by combining questionnaire and home video screening	Parent, video
Tariq, 2018[67]	Mobile detection of autism through machine learning on home video: a development and prospective validation study	Video

Lead Author, Year	Title	Related Canvas module(s): *parent questionnaire, video assessment, clinician questionnaire*
Kanne, 2018[140]	Screening in toddlers and preschoolers at risk for autism spectrum disorder: evaluating a novel mobile-health screening tool	Parent, video
Levy, 2017[65]	Sparsifying machine-learning models identify stable subsets of predictive features for behavioral detection of autism	Parent
Duda, 2017[74]	Crowdsourced validation of a machine-learning classification system for autism and ADHD	Video
Duda, 2016[152]	Clinical evaluation of a novel and mobile autism risk assessment	Mobile clinical screening tool (not Canvas Dx)[d]
Duda, 2016[75]	Use of machine learning for behavioral distinction of autism and ADHD	Parent
Bone, 2015[70]	Applying machine learning to autism diagnostics: pitfalls and promises	Parent[e]
Kosmicki, 2015[64]	Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning	Video
Duda, 2014[63]	Testing the accuracy of an observation-based classifier for rapid detection of autism risk	Parent
Wall, 2012[68]	Use of artificial intelligence to shorten the behavioral diagnosis of autism	Parent
Wall, 2012[66]	Use of machine learning to shorten observation-based screening and diagnosis of autism	Video

Notes. [a] Discusses both diagnostic and therapeutic uses of machine learning modules. [b] Protocol for NCT05223374, which was completed in February 2024 but has not yet been published. [c] Testing of the Cognoa video algorithm with Bangladeshi children. [d] Testing of a mobile clinical screening tool which predates the Cognoa clinician questionnaire tool. [e] Research not completed by the Cognoa team; another research team attempted to replicate Cognoa's findings with a larger dataset and identified what they perceived as conceptual and methodological problems with the Cognoa studies. This study is not included in evidence review because it did not include all 3 elements of the FDA-authorized SaMD.

Abbreviations. ADHD: attention deficit hyperactivity disorder. ASD: autism spectrum disorder. SaMD: software as a medical device.

Appendix F. Included Studies

Table F. Included Studies

Primary Publication from Included Trial	Publications Reporting Additional Results
Abbas et al. 2020[15] Western IRB #2202803	None
Megerian et al. 2022[16] NCT04151290	Wall et al. (2023) uses data from this clinical trial, but for other purposes and does not report additional trial results
Wall et al. 2023[13] NCT04151290 NCT03871179	None

Appendix G. Risk of Bias Assessment

Table G1. Risk of Bias: Diagnostic Test Accuracy Studies, Part 1

Study	Patient Selection				Index Test		Reference Standard	
	Consecutive or random sample of patients	Case-control design avoided	Inappropriate exclusions avoided	Applicability	Interpreted without knowledge of reference standard results	Used pre-specified threshold	Likely to correctly classify target condition	Results interpreted without knowledge of index test results
Wall et al. 2023	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Megerian et al. 2022	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Abbas et al. 2020	Unclear	Yes	No	Yes	Yes	No	Yes	Yes

Table G2. Risk of Bias: Diagnostic Test Accuracy Studies, Part 2

Study	Flow and Timing			Disclosures and Funding		Overall Risk of Bias
	Appropriate interval between index test and reference standard	All patients received the same reference standard	All patients included in analysis	Disclosures given for authors	Description of funding source	
Wall et al. 2023	Yes	Yes	Yes	Yes	No	**High** Lack of clarity regarding inclusion/exclusion criteria and how subjects were recruited; combination of data from 2 separate studies (1 of which was classified as withdrawn in clinicaltrials.gov); no distinction between authors and Cognoa. Did not use a method of calculating diagnostic test accuracy that accounted for indeterminate results.
Megerian et al. 2022	Yes	Yes	No	Yes	Yes	**High** Loss to follow up was high; although understandable due to COVID-19, this limits data used in analysis. Some concerns regarding conflict of interest. Did not use a method of calculating diagnostic test accuracy that accounted for indeterminate results.
Abbas et al. 2020	Yes	Yes	Unclear	Yes	No	**High** Unclear if a consecutive sample was used, undefined inclusion/exclusion criteria, algorithms changed/refined in mid-study, only half of sample included all 3 portions of Canvas app, number of uninterpretable cases not reported, unclear if all participants were included in analysis.

Appendix H. GRADE

Table H. GRADE Profile:
Test Accuracy (Sensitivity and Specificity) and Test Performance (PPV and NPV) of Canvas Dx for Diagnosis of Autism

Outcome	Number of Participants and Studies	Study Design	Factors That May Decrease Certainty of Evidence					Mean (95% CI)	Test Accuracy CoE
			Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias		
Sensitivity	1 study (Megerian) 425 participants	Prospective cohort study	Very serious (-2) See Risk of Bias assessment	Unknown Single study[a]	Serious (-1) 40% of enrolled subjects were not included in analysis, casting doubt on representative-ness of sample	No serious	Not assessed	Single study: 98.4% (95% CI, 91.6 to 100)	●○○○ VERY LOW
Specificity	1 study (Megerian) 425 participants	Prospective cohort study	Very serious (-2) See Risk of Bias assessment	Unknown Single study[a]	Serious (-1) 40% of enrolled subjects were not included in analysis, casting doubt on representative-ness of sample	Serious (-1) Wide CI	Not assessed	Single study: 78.9% (95% CI, 67.6 to 87.7)	●○○○ VERY LOW

Outcome	Number of Participants and Studies	Study Design	Factors That May Decrease Certainty of Evidence					Mean (95% CI)	Test Accuracy CoE
			Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias		
PPV	2 studies (Wall, Megerian) 1,147 participants	Prospective and retrospective cohort studies	Very serious (-2) See Risk of Bias assessment	Serious (-1) Difference in PPV between versions using the same algorithm	No serious	Serious (-1) Wide CI	Not assessed	Evaluations of the same algorithm: 89.7% (83.9 to 90.3) in Wall et al.[13] 80.8% (70.3 to 88.8) in Megerian et al.[16]	●○○○ VERY LOW
NPV	2 studies (Wall, Megerian) 1,147 participants	Prospective and retrospective cohort studies	Very serious (-2) See Risk of Bias assessment	No serious	Serious (-1) 40% of enrolled subjects in one input data set were not included in analysis, casting doubt on representative-ness of sample	No serious	Not assessed	Evaluations of the same algorithm: 96.1 (93.4 to 98.6) in Wall et al.[13] 98.3 (90.6 to 100) in Megerian et al.[16]	●○○○ VERY LOW
Time to diagnosis	0 studies	--	--	--	--	--	--	--	--
Time to service initiation	0 studies	--	--	--	--	--	--	--	--

Notes. [a] *Unable to rate for inconsistency.*

Abbreviations: CoE: certainty of evidence; NPV: negative predictive value; PPV: positive predictive value.

Appendix I. Measures of Test Performance

True positive (TP): A child with ASD diagnosed by the clinical reference standard who is correctly identified as having ASD with the diagnostic tool.

False positive (FP): A child who is identified as not having an ASD diagnosis by the clinical reference standard who is incorrectly identified as having ASD by the diagnostic tool.

True negative (TN): A child who is identified as not having an ASD diagnosis by the clinical reference standard who is correctly identified as not having ASD by the diagnostic tool.

False negative (FN): A child who is identified as having an ASD diagnosis by the clinical reference standard who is incorrectly identified as not having ASD by the diagnostic tool.

Sensitivity (true positive rate): The proportion of children who test positive among all those who have ASD. A highly sensitive test is effective at ruling out ASD when negative.

Specificity (true negative rate): The proportion of children who test negative among all those who do not have ASD. A highly specific test is effective at ruling in ASD when positive.

Positive predictive value (PPV): Probability that following a positive diagnostic test result, a child will have ASD as confirmed by the clinical reference standard.

Negative predictive value (NPV): Probability that following a negative diagnostic test result, a child truly not have ASD as confirmed by the clinical reference standard.

$$\textbf{Sensitivity} = \frac{(TP)}{(TP+FN)} \; ; \; \textbf{Specificity} = \frac{(TN)}{(TN+FP)}$$

$$\textbf{PPV} = \frac{(TP)}{(TP+FP)} \; ; \; \textbf{NPV} = \frac{(TN)}{(TN+FN)}$$

Table I. Measures of Test Performance

Screening Test Result	Condition Status, Confirmed by Clinical Reference Standard (diagnosis made by a specialist clinician, based on DSM-5 criteria)		Total
	Positive	Negative	
Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
Total	TP + FN	FP + TN	TP + FP + FN + TN

Abbreviations: FN: false negatives; FP: false positives; TN: true negatives; TP: true positives.

Appendix J. Excluded Studies with Primary Reason for Exclusion

Table J lists the publications that were excluded during full text review and the primary reason for exclusion. There may be multiple reasons for exclusion for any given publication, and the table lists only the most influential reason for exclusion.

Table J. Excluded Studies With Primary Reason for Exclusion

Reference Information	Primary Reason for Exclusion
Abbasi J. Mobile device app helps distinguish toddlers with autism. *JAMA*. 2021; 325(22):2243.	Publication Type
Achenie LEK, Scarpa A, Factor RS, et al. A machine learning strategy for autism screening in toddlers. *J Dev Behav Pediatr*. 2019; 40(5):369-376.	Intervention
Al Mamun KA, Bardhan S, Ullah MA, et al. Smart autism - a mobile, interactive and integrated framework for screening and confirmation of autism. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016; 2016:5989-5992.	Publication Type
Alqaysi ME, Albahri AS, Hamid RA. Diagnosis-based hybridization of multimedical tests and sociodemographic characteristics of autism spectrum disorder using artificial intelligence and machine learning techniques: a systematic review. *Int J Telemed Appl*. 2022:3551528.	Intervention
Anwar A, Abruzzo PM, Pasha S, et al. Advanced glycation end products, dityrosine and arginine transporter dysfunction in autism - a source of biomarkers for clinical diagnosis. *Mol Autism*. 2018; 9:3.	Intervention
Barbaro J, Yaari M. Study protocol for an evaluation of ASDetect - a mobile application for the early detection of autism. *BMC Pediatr*. 2020; 20(1):21.	Intervention
Blank R, Smits-Engelsman B, Polatajko H, et al. European Academy for Childhood Disability (EACD): recommendations on the definition, diagnosis and intervention of developmental coordination disorder (long version). *Dev Med Child Neurol*. 2012; 54(1):54-93.	Intervention
Bone D, Bishop SL, Black MP, et al. Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *J Child Psychol Psychiatry*. 2016; 57(8):927-37.	Study Design
Briguglio M, Turriziani L, Curro A, et al. A machine learning approach to the diagnosis of autism spectrum disorder and multi-systemic developmental disorder based on retrospective data and ADOS-2 score. *Brain Sci*. 2023; 13(6):31.	Intervention
Cavus N, Lawan AA, Ibrahim Z, et al. A systematic literature review on the application of machine-learning models in behavioral assessment of autism spectrum disorder. *J Pers Med*. 2021; 11(4):14.	Intervention
Dauchez T, Camelot G, Levy C, Rajerison T, et al. Diagnostic process for autism spectrum disorder: a meta-analysis of worldwide clinical practice guidelines for the initial somatic assessment. *Children*. 2022; 9(12):1886.	Intervention
Denis F, Maurier L, Carillo K, et al. Early detection of neurodevelopmental disorders of toddlers and postnatal depression by mobile health app: observational cross-sectional study. *JMIR Mhealth Uhealth*. 2022; 10(5):e38181.	Intervention
Desideri L, Perez-Fuster P, Herrera G. Information and communication technologies to support early screening of autism spectrum disorder: a systematic review. *Children*. 2021; 8(2):93.	Intervention
Ehteshami S, Mirzakhani Araghi N, Pashmdarfard M. Psychometric properties of autism spectrum disorders screening assessment tools: systematic review. *Med J Islam Repub Iran*. 2023; 37:117.	Intervention

Reference Information	Primary Reason for Exclusion
Gotham K, Risi S, Dawson G, Tager-Flusberg H, et al. A replication of the Autism Diagnostic Observation Schedule (ADOS) revised algorithms. *J Am Acad Child Adolesc Psychiatry*. 2008; 47(6):642-651.	Intervention
Grazioli S, Crippa A, Rosi E, et al. Exploring telediagnostic procedures in child neuropsychiatry: addressing ADHD diagnosis and autism symptoms through supervised machine learning. *Eur Child Adolesc Psychiatry*. 2024; 33(1):139-149.	Intervention
Hoffmann W, Konig U, Heinzel-Gutenbrunner M, et al. Early identification of Asperger syndrome in young children. *Res Dev Disabil*. 2013; 34(1):640-9.	Intervention
Hsu CF, Chien TW, Chow JC, et al. An app for identifying children at risk for developmental problems using multidimensional computerized adaptive testing: development and usability study. *JMIR Pediatr Parent*. 2020; 3(1):e14632.	Intervention
Joudar SS, Albahri AS, Hamid RA, et al. Artificial intelligence-based approaches for improving the diagnosis, triage, and prioritization of autism spectrum disorder: a systematic review of current trends and open issues. *Artif Intell Rev*. 2023; 56(S1):53-117.	Intervention
Joudar SS, Albahri AS, Hamid RA. Triage and priority-based health care diagnosis using artificial intelligence for autism spectrum disorder and gene contribution: a systematic review. *Comput Biol Med*. 2022; 146:105553.	Intervention
Knopf A. FDA authorizes marketing of diagnostic aid for autism spectrum disorder. *The Brown University Child & Adolescent Psychopharmacology Update*. 2021; 23(8):7-7.	Publication Type
Ko C, Lim JH, Hong J, et al. Development and validation of a joint attention-based deep learning system for detection and symptom severity assessment of autism spectrum disorder. *JAMA Netw Open*. 2023; 6(5):e2315174.	Intervention
Lavi R. The OlaMind screening tool for autism spectrum disorder in children aged 4-18 years. *Int J Child Health Hum Dev*. 2020; 13(4):395-405.	Intervention
Liu X, Zhao W, Qi Q, et al. A survey on autism care, diagnosis, and intervention based on mobile apps focusing on usability and software design. *Sensors*. 2023; 23(14):09.	Publication Type
Marciano F, Venutolo G, Ingenito CM, et al. Artificial intelligence: the "trait d'union" in different analysis approaches of autism spectrum disorder studies. *Curr Med Chem*. 2021; 28(32):6591-6618.	Publication Type
Marks KP, Page Glascoe F, Macias MM. Enhancing the algorithm for developmental-behavioral surveillance and screening in children 0 to 5 years. *Clin Pediatr*. 2011; 50(9):853-68.	Intervention
McPheeters ML, Weitlauf A, Vehorn A, et al. Screening for autism spectrum disorder in young children: a systematic evidence review for the U.S. Preventive Services Task Force. Agency for Healthcare Research and Quality (AHRQ); 2016; https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/autism-spectrum-disorder-in-young-children-screening. Accessed March 4, 2024.	Intervention
Metcalfe D, McKenzie K, McCarty K, et al. Screening tools for autism spectrum disorder, used with people with an intellectual disability: a systematic review. *Res Autism Spectr Disord*. 2020; 74:101549.	Intervention
Moon SJ, Hwang J, Kana R, et al. Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder: systematic review and meta-analysis of brain magnetic resonance imaging studies. *JMIR Ment Health*. 2019; 6(12):e14108.	Intervention
Mukherjee D, Bhavnani S, Lockwood Estrin G, et al. Digital tools for direct assessment of autism risk during early childhood: a systematic review. *Autism*. 2024; 28(1):6-31.	Intervention
Paolucci C, Federica G, Scheda R, et al. Early prediction of autism spectrum disorders through interaction analysis in home videos and explainable artificial intelligence. *Comput Human Behav*. 2023; 148(1):107877.	Intervention

Reference Information	Primary Reason for Exclusion
Rahman MKK, Subashini MM. A deep neural network-based model for screening autism spectrum disorder using the Quantitative Checklist for Autism in Toddlers (QCHAT). *J Autism Dev Disord*. 2022; 52(6):2732-2746.	Intervention
Randall M, Egberts KJ, Samtani A, et al. Diagnostic tests for autism spectrum disorder (ASD) in preschool children. *Cochrane Database Syst Rev*. 2018; 7(7):CD009044.	Intervention
Romaszko MZ, Ochal MR, Januszko-Giergielewicz B. Possibilities of early diagnosis of autism spectrum disorder, with a special attention to Asperger syndrome: a systematic literature review. *Polish Annals of Medicine*. 2020; 28(1):99-105.	Intervention
Romero-Garcia R, Martinez-Tomas R, Pozo P, et al. Q-CHAT-NAO: a robotic approach to autism screening in toddlers. *J Biomed Inform*. 2021; 118:103797.	Intervention
Sanchez-Garcia AB, Galindo-Villardon P, Nieto-Librero AB, et al. Toddler screening for autism spectrum disorder: a meta-analysis of diagnostic accuracy. *J Autism Dev Disord*. 2019; 49(5):1837-1852.	Intervention
Sanders BW, Bedrick S, Broder-Fingert S, et al. Mobile and online consumer tools to screen for autism do not promote equity. *Autism*. 2023; 27(3):714-722.	Intervention
Schrader E, Delehanty AD, Casler A, et al. integrating a new online autism screening tool in primary care to lower the age of referral. *Clin Pediatr*. 2020; 59(3):305-309.	Intervention
Shahamiri SR, Thabtah F. Autism AI: a new autism screening system based on artificial intelligence. *Cogn Comput*. 2020; 12(4):766-777.	Intervention
Shannon J, Taraman S, Wall DP, et al. 3.46 Optimizing a de novo artificial intelligence–based medical device under a predetermined change control plan: improved ability to detect or rule out ASD in general pediatric settings. *J Am Acad Child Adolesc Psychiatry*. 2022; 61(10):S242-S243.	Publication Type
Siu AL, Bibbins-Domingo K, Grossman DC, et al. Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *JAMA*. 2016; 315(7):691-6.	Intervention
Song DY, Kim SY, Bong G, et al. The use of artificial intelligence in screening and diagnosis of autism spectrum disorder: a literature review. *Soa Chongsonyon Chongsin Uihak*. 2019; 30(4):145-152.	Intervention
Souza PVC, Guimaraes AJ, Araujo VS, et al. An intelligent Bayesian hybrid approach to help autism diagnosis. *Soft* Comput. 2021; 25(14):9163-9183.	Intervention
Tartarisco G, Cicceri G, Di Pietro D, et al. Use of machine learning to investigate the quantitative checklist for autism in toddlers (Q-CHAT) towards early autism screening. *Diagnostics*. 2021; 11(3):22.	Intervention
Thabtah F, Kamalov F, Rajab K. A new computational intelligence approach to detect autistic features for autism screening. *Int J Med Inform*. 2018; 117():112-124.	Intervention
Thabtah F, Peebles D. A new machine learning model based on induction of rules for autism detection. *Health Informatics J*. 2020; 26(1):264-286.	Intervention
Thabtah F. An accessible and efficient autism screening method for behavioural data and predictive analyses. *Health Informatics J*. 2019; 25(4):1739-1755.	Intervention
Tunc B, Pandey J, St John T, et al. Diagnostic shifts in autism spectrum disorder can be linked to the fuzzy nature of the diagnostic boundary: a data-driven approach. *J Child Psychol Psychiatry*. 2021; 62(10):1236-1245.	Intervention
Valizadeh A, Moassefi M, Nakhostin-Ansari A, et al. Automated diagnosis of autism with artificial intelligence: state of the art. *Rev Neurosci*. 2024; 35(2):141-163.	Intervention
Wei Q, Xu X, Xu X, et al. Early identification of autism spectrum disorder by multi-instrument fusion: a clinically applicable machine learning approach. *Psychiatry Res*. 2023; 320:115050.	Intervention

Reference Information	Primary Reason for Exclusion
Wells C, Hill S, Argaez C. Artificial intelligence and machine learning in mental health services: a literature review. Canadian Agency for Drugs and Technology in Health (CADTH), Mental Health Commission of Canada (MHCC); 2021; https://www.cadth.ca/artificial-intelligence-and-machine-learning-mental-health-services-literature-review. Accessed March 4, 2024.	Population
Wells C, Hill S, Argaez C. Artificial intelligence and machine learning in mental health services: an environmental scan. Canadian Agency for Drugs and Technology in Health (CADTH), Mental Health Commission of Canada (MHCC); 2021; https://www.cadth.ca/artificial-intelligence-and-machine-learning-mental-health-services-environmental-scan. Accessed March 4, 2024.	Publication Type
Xipolitopoulos C, Nikiforos MN, Exarchos T. machine learning for autistic spectrum disorder risk screening. *Adv Exp Med Biol*. 2021; 1338:81-87.	Intervention

Appendix K. Description of Payer Policies

Table K. Overview of Coverage Criteria for Canvas Dx

Policy Author Last Review Date	Policy Terminology and Definition of Digital Health Technologies	Medical Necessity Criteria for Digital Health Technologies	Canvas Dx Coverage
Private payers			
Aetna July 17, 2024[19] September 20, 2024[145]	Prescription digital therapeutics	Only covers FDA-approved or cleared mobile apps for contraception based on fertility awareness, when prescribed by a treating provider, per federal preventive care mandates	Not covered • Considered experimental and investigational • Insufficient evidence in published peer-reviewed literature on effectiveness • Peer-reviewed literature does not support the use of artificial intelligence-based devices for diagnosis of ASD
Anthem Blue Cross and Blue Shield (formerly Empire BlueCross BlueShield)[20] August 8, 2024	Mobile device-based health management applications Practitioner-prescribed software applications for health management purposes when used on a mobile device to evaluate, diagnose, or treat an illness, injury, disease, or its symptoms	When the following criteria have been met: • Criteria to evaluate the MSA: ○ Approved or cleared by FDA ○ Credible scientific evidence that permits reasonable conclusions regarding the impact of the MSA on health outcomes ○ MSA proven materially to improve the net health outcome or be as beneficial as any established alternative • Criteria to evaluate the appropriateness of the MSA for the individual: ○ MSA prescribed by a health care practitioner ○ Documentation supporting that MSA ordered for a covered purpose, and	Not covered Insufficient data to understand whether use of Canvas Dx decreases time to diagnosis in a real-world setting, and if it is likely to improve clinically relevant ASD outcomes

Policy Author Last Review Date	Policy Terminology and Definition of Digital Health Technologies	Medical Necessity Criteria for Digital Health Technologies	Canvas Dx Coverage
		according to accepted standard of medical practice ○ Requested MSA is not primarily for the convenience of the individual, prescribing clinician, caregiver, or other health care provider	
Highmark Blue Shield of Northeastern New York May 2024[21] June 2024[153]	Digital diagnostics: • Technology-based health care diagnosis aids for medical and behavioral conditions • Intended to be used as adjunct to diagnostic process, not a stand-alone diagnostic device Prescription digital therapeutics: • Technology-based health care therapeutic interventions for the treatment of medical and behavioral conditions • Intended to be used as a part or whole of a treatment plan for appropriate health diagnoses	FDA approved digital therapeutics prescribed by a licensed health care professional for therapeutic intervention may be considered medically necessary when the following criteria are met: • Used within approved indications • Prescribed by a provider for whom the condition is within the scope of their practice • Individual must be at least 18 years of age, unless the digital therapeutic is designed and approved for pediatric use and the individual is within the recommended age range • Only used for outpatient care • Individual must be able to reasonably interact with the software to receive prescription for any digital therapeutic intervention • Approved by Highmark New Technology Advisory Committee	Covered • May be considered medically necessary to aid in the diagnosis of ASD when the following indications are met: ○ Individual aged between 18 to 72 months ○ Individual exhibiting 1 or more of the following developmental delays: persistent deficits in social communication and social interaction across multiple contexts; restricted, repetitive patterns of behaviors, interests, or activities • Digital diagnostic device not intended for use as a stand-alone diagnostic device, but as an adjunct to the diagnostic process • HCPCS code A9291 is used

Abbreviations. ASD: autism spectrum disorder; FDA: US Food and Drug Administration; MSA: mobile-based software application.

Appendix L. Applicable Codes

Table L. Applicable Codes for Canvas Dx

Code	Description
ICD-10-CM Codes	
F84.0	Autistic disorder
Z13.4	Encounter for screening for certain developmental disorders in childhood
Z13.40	Encounter for screening for unspecified developmental delays
Z13.41	Encounter for autism screening
Z13.42	Encounter for screening for global developmental delays (milestones)
Z13.49	Encounter for screening for other developmental delays
CPT Codes	
96110	Developmental screening (e.g., developmental milestone survey, speech and language delay screen), with scoring and documentation, per standardized instrument
96112	Developmental test administration (including assessment of fine and gross motor, language, cognitive level, social, memory and executive functions by standardized developmental instruments when performed), by physician or other qualified health care professional, with interpretation and report; first hour
96113	Developmental test administration (including assessment of fine and gross motor, language, cognitive level, social, memory and executive functions by standardized developmental instruments when performed), by physician or other qualified health care professional, with interpretation and report; each additional 30 minutes
91999	Unlisted special service, procedure or report [when specified as a mobile-based health management software application]
HCPCS Codes	
A9291[a]	Prescription digital cognitive and/or behavioral therapy, Food and Drug Administration cleared, per course of treatment
G0451	Development testing; with interpretation and report, per standardized instrument form

Notes [a] Highmark Blue Shield of Northeastern New York uses this code for Canvas Dx.

Abbreviations: CPT: Current Procedural Terminology; HCPCS: Healthcare Common Procedure Coding System; ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification.