3.03.01	Digital Health Technologies: Diagnostic Applications		
Original Policy Date:	September 1, 2022	**Effective Date:**	October 1, 2023
Section:		**Page:**	Page 1 of 15

Policy Statement

I. Prescription digital health technologies for diagnostic application that have received clearance for marketing by the U.S. Food and Drug Administration as a diagnostic aid for autism spectrum disorder (Canvas Dx) are considered **investigational**.

NOTE: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

Policy Guidelines

- N/A

Description

Digital health technologies is a broad term that includes categories such as mobile health, health information technology, wearable devices, telehealth and telemedicine, and personalized medicine. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device, and include technologies intended for use as a medical product, in a medical product, as companion diagnostics, or as an adjunct to other medical products (devices, drugs, and biologics). The scope of this review includes only those digital technologies that are intended to be used for diagnostic application (detecting the presence or absence of a condition, the risk of developing a condition in the future, or treatment response [beneficial or adverse]) and meet the following 3 criterion- 1) Must meet the definition of "Software as a medical device" which states that software is intended to be used for a medical purpose, without being part of a hardware medical device or software that stores or transmits medical information. 2) Must have received marketing clearance or approval by the U.S. Food and Drug Administration either through the *de novo premarket* process or 510(k) process or pre-market approval and 3) Must be prescribed by a healthcare provider.

Related Policies

- N/A

Benefit Application

Benefit determinations should be based in all cases on the applicable contract language. To the extent there are any conflicts between these guidelines and the contract language, the contract language will control. Please refer to the member's contract benefits in effect at the time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal mandates (e.g., Federal Employee Program [FEP]) prohibits plans from denying Food and Drug Administration (FDA)-approved technologies as investigational. In these instances, plans may have to consider the coverage eligibility of FDA-approved technologies on the basis of medical necessity alone.

Regulatory Status

Digital health technologies that meet the current scope of review are shown in Table 1.

Table 1. Digital Health Technology for Diagnostic Applications

Application	Manufacturer	FDA Cleared Indication	Description	FDA Product Code	FDA Marketing Clearance	Year
Canvas DX (formerly known as Coagnoa App)	Cognoa	"Canvas Dx is intended for use by healthcare providers as an aid in the diagnosis of Autism Spectrum Disorder (ASD) for patients ages 18 months through 72 months who are at risk for developmental delay based on concerns of a parent, caregiver, or healthcare provider. The device is not intended for use as a stand-alone diagnostic device but as an adjunct to the diagnostic process. The device is for prescription use only (Rx only)."	Artificial intelligence app for use by health care providers as an adjunct in the diagnosis of autism spectrum disorder for patients ages 18 to 72 months. Canvas DX includes 3 questionnaires: parent/caregiver, a video analyst, and a health care provider, with an algorithm that synthesizes the 3 inputs for use by the primary care provider.	QPF	DEN200069	2021

FDA: U.S. Food and Drug Administration;

Rationale

Background

Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a biologically based neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction and restricted, repetitive patterns of behavior, interests, and activities. ASD can range from mild social impairment to severely impaired functioning; as many as half of individuals with autism are non-verbal and have symptoms that may include debilitating intellectual disabilities, inability to change routines, and severe sensory reactions. The American Psychiatric Association's Diagnostic and Statistical Manual, Fifth Edition (DSM-5) provides standardized criteria to help diagnose ASD.[1,]

Diagnosis of ASD in the United States generally occurs in two steps: developmental screening followed by comprehensive diagnostic evaluation if screened positive. American Academy of Pediatrics (AAP) recommends general developmental screening at 9, 18 and 30 months of age and ASD specific screening at 18 and 24 months of age.[2,3,] Diagnosis and treatment in the first few years of life can have a strong impact on functioning as it allows for treatment during a key window of developmental plasticity.[4,5,]However, early diagnosis in US remains an unmet need even though studies have demonstrated a temporal trend of decreasing mean ages at diagnosis over time.[6,7,]. According to a 2020 study by Autism and Developmental Disabilities Monitoring (ADDM) Network, an active surveillance system that provides estimates of ASD in the US, reported median age of earliest known ASD diagnosis ranged from 36 months in California to 63 months in Minnesota.[8,]

Scope of Review

Software has become an important part of product development and is integrated widely into digital platforms that serve both medical and non-medical purposes. Three broad categories of software use in medical device are

1. Software used in the manufacture or maintenance of a medical device (example software that monitors x-ray tube performance to anticipate the need for replacement),
2. Software that is integral to a medical device or software in a medical device (example software used to "drive or control" the motors and the pumping of medication in an infusion pump)
3. Software, which on its own is a medical device referred to as "Software as a Medical Device" (SaMD) (example, software that can track the size of a mole over time and determine the risk of melanoma)

The International Medical Device Regulators Forum, a consortium of medical device regulators from around the world led by the U.S. Food and Drug Administration (FDA) defines SaMD as "software that is intended to be used for one or more medical purposes that perform those purposes without being part of a hardware medical device".[9,] Such software was previously referred to by industry, international regulators, and health care providers as "standalone software," "medical device software," and/or "health software," and can sometimes be confused with other types of software. The scope of this review includes only those digital technologies that are intended to be used for diagnostic application (detecting presence or absence of a condition, the risk of developing a condition in the future, or treatment response [beneficial or adverse]) and meet the following 3 criterion-

1. Must meet the definition of "Software as a medical device" which states that software is intended to be used for a medical purpose, without being part of a hardware medical device or software that stores or transmits medical information.
2. Must have received marketing clearance or approval by the U.S. Food and Drug Administration either through the *de novo* premarket process or 510(k) process or pre-market approval and
3. Must be prescribed by a healthcare provider.

BCBSA Evaluation Framework for Digital Health Technologies

SaMDs, as defined by FDA, are subject to the same evaluation standards as other devices; the Blue Cross and Blue Shield Association Technology Evaluation Criterion are as follows:

1. The technology must have final approval from the appropriate governmental regulatory bodies.
2. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.
3. The technology must improve the net health outcome.[a]
4. The technology must be as beneficial as any established alternatives.
5. The improvement must be attainable outside the investigational settings.[b]

[a] The technology must assure protection of sensitive patient health information as per the requirements of The Health Insurance Portability and Accountability Act of 1996 (HIPAA)
[b] The technology must demonstrate usability in a real-world setting
Other regulatory authorities such as the United Kingdom's National Institute for Health and Care Excellence (NICE) have proposed standards to evaluate SaMD.[10,]

Literature Review

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

Promotion of greater diversity and inclusion in clinical research of historically marginalized groups (e.g., People of Color [African-American, Asian, Black, Latino and Native American]; LGBTQIA (Lesbian, Gay, Bisexual, Transgender, Queer, Intersex, Asexual); Women; and People with Disabilities [Physical and Invisible]) allows policy populations to be more reflective of and findings more applicable to our diverse members. While we also strive to use inclusive language related to these groups in our policies, use of gender-specific nouns (e.g., women, men, sisters, etc.) will continue when reflective of language used in publications describing study populations.

Autism Spectrum Disorder

Clinical Context and Test Purpose

The American Academy of Pediatrics provides details on the screening and diagnosis for autism spectrum disorder (ASD).[2,3,] Children with ASD can be identified as toddlers, and early intervention can and does influence outcomes.[11,] The Academy recommends screening all children for symptoms of ASD through a combination of developmental surveillance at 9, 18, and 30 months of age and standardized autism-specific screening tests at 18 and 24 months of age.

Screening tools typically use questionnaires that are answered by a parent, teacher, or clinician and are designed to help caregivers identify and report symptoms observed in children at high risk for ASD. While they are generally easy and inexpensive to administer, they have limited sensitivity (ability to identify young children with ASD) and specificity (ability to discriminate ASD from other developmental disorders, such as language disorders and global developmental delay).[12,] Results of a screening test are not diagnostic. Due to the variability in the natural course of early social and language development, some children who have initial positive screens (suggesting that they are at risk for ASD) ultimately will not meet diagnostic criteria for ASD.[13,] Other children who pass early screens for ASD may present with atypical concerns later in the second year of life and eventually be diagnosed with ASD. In the context of early identification and diagnosis of ASD, sensitivity is more important than specificity for a screening test as the potential over-referral of children with positive screens is preferable to missing children at risk for ASD. Once a child is determined to be at risk for a diagnosis of ASD, either by screening or surveillance, a timely referral for a comprehensive clinical diagnostic evaluation is warranted. Structured observation of symptoms of ASD during clinical evaluation is helpful to inform the diagnostic application of the DSM-5 criteria. These tools require long and expensive interactions with highly trained clinicians. To meet diagnostic criteria, the symptoms must impair function.

Cognoa, the manufacturer of Canvas Dx, has stated on its website that the test "is intended for use by healthcare providers as an aid in the diagnosis of ASD for patients ages 18 months through 72 months who are at risk for developmental delay based on concerns of a parent, caregiver, or healthcare provider."[14,] The device is not intended for use as a stand-alone diagnostic device but as an adjunct to the diagnostic process. Further the manufacturer states, "Canvas Dx can aid primary care physician in diagnosing ASD in children starting at 18 months of age during a critical period when interventions are shown to provide/lead to optimal long-term outcomes." The manufacturer also makes indirect and direct assertions that the use of Canvas Dx may allow children with ASD to be diagnosed earlier than the current average age of diagnosis and that the use of this test fulfills an unmet need for a delayed formal diagnosis of ASD after parenteral concern.[14,] Some of the reasons cited for the unmet need of a delayed diagnosis is shortage of specialists, time-intensive evaluations, lack of access to care for children from ethnic/racial minorities and/or disadvantaged socioeconomic backgrounds and in rural areas, lack of standard diagnostic process for ASD and use of multiple types of specialists for referral with no clear pathway for primary care physicians.

To evaluate the utility of the test, an explication of how the test would be integrated into the current AAP-recommended screening and diagnostic pathway is needed. The FDA authorized indication is for children who are at risk of developmental delay. It is unclear how Canvas Dx should be used as adiagnostic aid. The diagnostic accuracy of Canvas Dx was evaluated in a community setting by physicians who completed residency training in either general pediatrics or family medicine. However, the referral pathway after completion of Canvas Dx test lacks clarity. Two potential scenarios are possible and summarized in Table 2. Note that each of these hypothetical scenarios have a unique PICO formulation and require a different metric to understand test accuracy. For example, positive predictive value (PPV) answers the question, "How likely is it that the patient with a positive test actually has the condition?" and is the more important measure for a rule-in test. On the other hand, a negative predictive value (NPV) answers the question, "How likely is it that a patient with a negative test is actually free of the condition?" and is the more important measure for rule-out test.

Table 2. Potential Referral Strategies with Canvas Dx

Canvas Dx Test Referral Strategy	Implications
Assumption 1: For a negative test, further testing by a specialist is not required. For all others results (positive/indeterminate), further testing by a specialist for confirmatory diagnosis is required.	Under these assumptions, Canvas Dx is a "rule out test".
Assumption 2: For a negative or positive test, further testing by a specialist is not required. For indeterminate results, further testing by a specialist for confirmatory diagnosis.	Under these assumptions, Canvas Dx is both a "rule out test" and "rule in" test.

The purpose of Canvas DX in individuals who are in the age range of 18 to 72 months and in whom there is a suspicion of ASD by a parent, caregiver, or healthcare provider is unclear.

The following PICO was used to select literature to inform this review.

Populations
The relevant population of interest is children who are in the age range of 18 to 72 months and who are at risk of developmental delay.

Interventions
The test being considered is Canvas DX (formerly known as Cognoa App). According to the manufacturer, Canvas Dx is a prescription diagnostic aid to healthcare professionals considering the diagnosis of ASD in patients 18 months through 72 months of age at risk for developmental delay.[14,] Canvas Dx incorporates 3 separate inputs. The patient's caregiver uses a smartphone application ("App") to fill out a caregiver questionnaire (4-minute) that asks about the child's behavior and development. The patient's caregiver also uses the smartphone application to make video recordings of behavior at home. A lightly trained video analyst reviews these videos of the child recorded by the parent/caregiver and completes a questionnaire (2-minute). Finally, a health care professional meets with the child and a parent/caregiver and completes an online questionnaire (2-minute) via a health care provider portal. Canvas Dx utilizes a machine-learning algorithm that receives the 3 independent inputs and produces one of the 3 outputs listed in Table 3.

Canvas DX uses a machine learning-based assessment of autism comprising the above-mentioned modules for a unified outcome of diagnostic-grade reliability. The parent and the clinician questionnaire modules are based on behavioral patterns probed by a Autism Diagnostic Interview-Revised (ADI-R) while the video assessment module is based on behavioral patterns probed by the Autism Diagnostic Observation Schedule (ADOS).[15,]

Abbas et al (2020) states that the responses from the 3 modules are each considered to be a 'probability and combined mathematically'.[15,] Upper and lower thresholds are applied to produce the

categories in Table 3. The paper states that 'thresholds can be tuned independently to optimize the sensitivity, specificity, and model coverage'.

Table 3. Outputs of Canvas Dx[14,]

Canvas Dx Output	Interpretation
Positive for ASD	The patient has ASD if the healthcare professional confirms the clinical presentation of the patient is consistent with and meets diagnostic criteria for ASD.
Negative for ASD	The patient does NOT have ASD if the healthcare professional confirms the clinical presentation of the patient is consistent with ruling out ASD and does NOT meet diagnostic criteria for ASD. A negative result does not necessarily mean that the patient will not develop ASD in the future and continued monitoring or evaluation for non-ASD conditions may be warranted.
No result	The available information does not allow the algorithm to render a reliable result. This does not mean that the patient either has or does not have ASD.

ASD: autism spectrum disorder

Comparators

The comparator would be comprehensive diagnostic evaluation tests for confirmatory diagnosis of ASD that are commonly used in the United States.

Diagnostic tools commonly used in the US are summarized in Table 4. The accuracy of many of these tools has not been well studied.[16,] Tools that are recommended in national guidelines and used in the U.S. include Autism Diagnostic Interview-Revised (ADI-R), Autism Diagnostic Observation Schedule-2nd edition (ADOS-2), and Childhood Autism Rating Scale 2nd edition (CARS-2). According to a 2018 Cochrane systematic review and meta-analyses, authors observed substantial variation in sensitivity and specificity of all tests. According to summary statistics for ADOS, CARS, and ADI-R, ADOS was found to be the most sensitive. All tools performed similarly for specificity.[16,]

Table 4. Commonly Used Diagnostic Instruments and Tools for Autism Spectrum Disorder in the United States[a]

Tool	Age	Description	Comments
ADI-R	Mental age ≥18 months	• 2- to 3-hour 93-point semi-structured clinical interview that probes for ASD symptoms	• Not practical for clinical settings • Usually used in research settings, often combined with the ADOS-2
ADOS-2nd edition	Age 12 months through adulthood	• Semi-structured assessment by trained clinician of social interaction, play/imaginative use of materials, communication and atypical behaviors • 5 modules based on child's expressive language abilities (including one for toddlers) • Takes 40 to 60 minutes to administer	• Reference standard for diagnosis of ASD in research studies and clinical settings • The information obtained from the ADOS-2 is used by the clinician in conjunction with the history of peer interactions, social relationships, and functional impairment from symptoms to determine if the DSM-5 criteria are met

Tool	Age	Description	Comments
CARS-2	Children ≥2 years of age	• 15 items directly observed by a trained clinician and a parent unscored questionnaire • Takes 20 to 30 minutes to administer	• 15 items are correlated with DSM-5

[a] This table is not exhaustive, and other tests are available such as Developmental Dimensional and Diagnostic Interview (3di), Diagnostic Interview for Social and Communication Disorder (DISCO), Gilliam Autism Rating Scale (GARS) and Social Responsiveness Scale, Second edition (SRS). According to AAP, validated observation tools include the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) and the Childhood Autism Rating Scale, Second Edition (CARS-2). No single observation tool is appropriate for all clinical settings.[3,]
ADI-R: Autism Diagnostic Interview-Revised; ADOS-2: Autism Diagnostic Observation Schedule-2nd edition (ADOS-2); ASD: autism spectrum disorder; CARS-2: Childhood Autism Rating Scale 2nd edition; DSM-5: The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

Outcomes

The general outcomes of interest are test validity, symptoms, functional outcomes, quality of life.

- Beneficial outcomes resulting from a true negative test result are avoiding unnecessary subsequent testing.
- Beneficial outcomes resulting from a true positive test result are early referral for comprehensive evaluation and identification of ASD leading to early intervention and improved health outcomes.
- Harmful outcomes resulting from a false-positive test result are unnecessary testing or treatment, potential stigmatization and other ethical, legal, and social implications such as educational and employment discrimination.
- Harmful outcomes resulting from a false-negative test result are diagnostic delay and possibility of missing treatment during the key window of developmental plasticity.

A fuller explanation of appropriate outcomes is not possible until the position of the test in the screening and diagnostic pathway is clarified.

Study Selection Criteria

For the evaluation of clinical validity of Canvas Dx, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores);
- Included a suitable reference standard;
- Patient/sample clinical characteristics were described;
- Patient/sample selection criteria were described.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Diagnostic Performance

Two studies on diagnostic performance of Canvas Dx have been published. The first study by Abbas et al (2020) reported on the technical development and performance of the Canvas Dx (formerly known as Cognoa App) for diagnosing ASD and is not reviewed in detail.[15,] The second study by Megerian (2022) was a double-blind, multicenter, prospective, comparator cohort study testing the diagnostic accuracy of Canvas Dx in a primary care setting.[17,]The study compared Canvas Dx output to diagnostic agreement by 2 or more independent specialists in a cohort of 18 to 72-month-olds with developmental delay concerns. Characteristics and results are summarized in Tables 5 and 6. A total of 711 participants were enrolled and 425 completed both the device input and specialist evaluation component of the study between August 2019 and June 2020. The majority of study participants (68% or 290/425) were classified as "indeterminates" by Canvas DX. For the 32% of participants who received a determinate output (ASD positive or negative), sensitivity was 98.4% (95% CI, 91.6% to

100%), specificity was 78.9% (95% CI, 67.6% to 87.7%), PPV was 80.8% (95% CI, 70.3% to 88.8%) and NPV was 98.3% (95% CI, 90.6% to 100%).

Relevance, design, and conduct gaps in the studies are described in Tables 7 and 8. Major limitations in study relevance are the lack of clarity on how the test fits into the current pathway and the appropriate referral process subsequent to testing. It is unclear if Canvas Dx is a "rule-out" or "rule-in" test or perhaps both. Major limitations in the design and conduct of the study include missing data and lack of generalizability. As per the protocol, the study planned to enroll 725 participants between the ages of ≥18 months and <72 months of age from 30 clinical sites within the United States.

However, 711 participants were enrolled from 14 sites across 6 states. Of these, 425 completed both the device input and specialist evaluation component of the study and were included in the final analysis. The estimated drop out rate was 40%. Authors reported that COVID-19 control measures led to changes in study visit schedules, missed visits, patient discontinuations, and site closures (9 out of 14 sites). No clear description of reasons for discrepancy in the number of clinical sites (30 proposed sites versus 14 actual sites), characteristics of missing observations, or sensitivity analyses of missing data assumptions were provided. Issues related to the generalizability of the study findings were also noted. Data on participants stratified by enrollment sites/states and origin of primary concern for developmental delay (whether it was patient/caregiver or healthcare professional) were not reported. More clarity on these issues is needed to understand generalizability of this study.

Table 5. Characteristics of Studies of Clinical Validity of a Diagnostic Test

Study	Study Population	Design	Reference Standard for ASD	Threshold for Positive Canvas Dx	Timing of Reference and Canvas Dx	Blinding of Assessors	Comment
Megerian et al (2022)[17], [NCT04151290]	Children 18 to 72 months of age Functional English capability in the home environment Caregivers must have smartphone capabilities	Double-blind, multicenter, prospective, active comparator Study participants were recruited from 14 trial sites across 6 states if the primary HCP identifies a child at risk for developmental delay or learns of caregiver concern about developmental delay.	After completion of assessment by Canvas Dx, trialists contacted caregiver to schedule an appointment for a diagnostic evaluation by the specialist clinician. Diagnostic process aligned with best practice recommendations for ASD evaluation and specialist assessments were conducted. by board-certified child and adolescent psychiatrists, child neurologists, developmental-behavioral pediatricians, or child psychologists with more than 5 years of experience diagnosing ASD. To ascertain diagnostic certainty, all cases were independently assessed by a second reviewing	Proprietary algorithm across 3 inputs uses 64 questions to identify behavioral, executive functioning, and language and communication features that are maximally predictive of an ASD diagnosis. Thresholds not defined.		Yes	None

Study	Study Population	Design	Reference Standard for ASD	Threshold for Positive Canvas Dx	Timing of Reference and Canvas Dx	Blinding of Assessors	Comment
			specialist. If the 2 specialists disagreed about the ASD diagnosis, then the case was referred to a third specialist.				

ASD: autism spectrum disorder; HCP: healthcare professionaled

Table 6. Clinical Validity Results of Canvas Dx

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Sensitivity	Specificity	Predictive Value
Megerian et al (2022)[17],	Consented: 711 Completed Canvas Dx: 585	Completed Canvas Dx and specialist evaluation: 425	Did not complete Canvas Dx: 126 Did not complete specialist evaluation: 160	29% (122 of 425 with developmental concern diagnosed with ASD via specialist evaluation)	98.4 (91.6 to 99.9)	78.9 (67.6 to 87.7)	PPV: 80.8 (70.3 to 88.8) NPV: 98.3 (90.6 to 99.9)

ASD: autism spectrum disorder; PPV: positive predictive value; NPV: negative predictive value

Table 7. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Megerian et al (2022)[17],	2. Test use in current diagnostic pathway unclear (lack of clarity on how the test fits into the current pathway and the appropriate referral process subsequent to testing)				

[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use; 5 Enrolled study populations do not reflect relevant diversity; 6. Other
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest (e.g., older version of test, not applied as intended); 4. Other.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose; 4. Other.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported; 4. Reclassification of diagnostic or prognostic risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests); 6. Other.
e Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined); 2: Other.

Table 8. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Megerian et al (2022)[17],					1. Inadequate description of indeterminate and missing samples 3. High loss to follow-up or missing data (approximately 40%)	

[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience); 3. Other.
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests; 2. Other.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described; 5. Other.
d Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication; 4. Other.
e Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data; 4. Other.
f Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported; 3: Insufficient consideration of potential confounding; 4. Other.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility. There are no studies comparing clinical outcomes for patients diagnosed using Canvas Dx with alternative methods for testing for ASD (ie, no direct evidence that the test is clinically useful). Currently, it is unclear whether a chain of evidence can be constructed because of the lack of clarity on how the test results would be used to change management practices.

Section Summary: Autism Spectrum Disorder

The evidence on Canvas Dx includes a single double-blinded, multi-site, prospective, comparator cohort study which reported on the diagnostic accuracy of Canvas Dx in a primary care setting (enrolled 711, completed 425). The study compared Canvas Dx output to diagnostic agreement by two or more independent specialists in a cohort of 18 to 72-month-olds with developmental delay concerns. Majority of study participants (68% or 290/425) were classified as "indeterminates" by Canvas DX. For the 32% of participants who received a determinate output (ASD positive or negative), sensitivity was 98.4% (95% CI: 91.6% to 100%), specificity was 78.9% (95% CI: 67.6% to 87.7%), PPV was 80.8% (95% CI: 70.3 to 88.8%) and NPV was 98.3% (95% CI: 90.6% to 100%). Major limitation in study relevance is the lack of clarity on how the test fits into the current pathway and the appropriate referral process subsequent to testing. It is unclear if Canvas Dx is a "rule-out" or "rule-in" test or perhaps both. Major limitation in the design and conduct of the study included missing data and lack of generalizability. The estimated drop out rate was 40%. Authors reported that COVID-19

control measures led to changes in study visit schedules, missed visits, patient discontinuations, and site closures (9 out of 14 sites). No clear description of reasons for discrepancy in the number of clinical sites (30 proposed sites versus 14 actual sites), missingness, characteristics of missing observations, or sensitivity analyses of missing data assumptions were provided. Issues related to the generalizability of the study findings were also noted. Data on participants stratified by enrollment sites/states and origin of primary concern for developmental delay (whether it was patient/caregiver or healthcare professional) was not reported. More clarity on these issues is needed to understand generalizability of this study. Other limitations include differences that may occur between the testing environments of a structured clinical trial setting versus the home setting and lack of data on usability outside of a clinical trial.

Supplemental Information

The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

Practice Guidelines and Position Statements

Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

American Academy of Pediatrics

The American Academy of Pediatrics (AAP) guidelines recommend autism spectrum disorder (ASD) - specific universal screening in all children at ages 18 and 24 months in addition to developmental surveillance and monitoring.[2,] Toddlers and children should be referred for diagnostic evaluation when increased risk for developmental disorders (including ASD) is identified through screening and/or surveillance. Children should be referred for intervention for all identified developmental delays at the time of identification and not wait for an ASD diagnostic evaluation to take place. The AAP does not approve nor endorse any specific tool for screening purposes. The AAP has published a toolkit that provides a list of links to tools for developmental surveillance and screening for use at the discretion of the health care professional.[18,]

The American Academy of Child and Adolescent Psychiatry

The American Academy of Child and Adolescent Psychiatry recommends that the developmental assessment of young children and the psychiatric assessment of all children should routinely include questions about ASD symptomatology.[19,]

U.S. Preventive Services Task Force Recommendations

The U.S. Preventive Services Task Force (USPSTF) published recommendations for ASD in young children in 2016.[20,] The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for ASD in young children (children 18 to 30 months of age) for whom no concerns of ASD have been raised by their parents or a clinician.

Medicare National Coverage

There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Ongoing and Unpublished Clinical Trials

Some currently unpublished trials that might influence this review are listed in Table 9.

Table 9. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT05223374	Extension for Community Healthcare Outcomes (ECHO) Autism Diagnostic Study in Primary Care Setting	100	Jun , 2023
Unpublished			
NCT04326231[a]	Cognoa ASD Digital Therapeutic Engagement and Usability Study	30	Jul 2020

NCT: national clinical trial.
[a] Denotes industry-sponsored or cosponsored trial.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5), 5th ed. Washington, DC: American Psychiatric Association; 2013
2. Lipkin PH, Macias MM, Norwood KW, et al. Promoting Optimal Development: Identifying Infants and Young Children With Developmental Disorders Through Developmental Surveillance and Screening. Pediatrics. Jan 2020; 145(1). PMID 31843861
3. Hyman SL, Levy SE, Myers SM, et al. Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. Pediatrics. Jan 2020; 145(1). PMID 31843864
4. Dawson G, Bernier R. A quarter century of progress on the early detection and treatment of autism spectrum disorder. Dev Psychopathol. Nov 2013; 25(4 Pt 2): 1455-72. PMID 24342850
5. Dawson G, Rogers S, Munson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics. Jan 2010; 125(1): e17-23. PMID 19948568
6. Hertz-Picciotto I, Delwiche L. The rise in autism and the role of age at diagnosis. Epidemiology. Jan 2009; 20(1): 84-90. PMID 19234401
7. Leigh JP, Grosse SD, Cassady D, et al. Spending by California's Department of Developmental Services for Persons with Autism across Demographic and Expenditure Categories. PLoS One. 2016; 11(3): e0151970. PMID 27015098
8. Maenner MJ, Shaw KA, Bakian AV, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. MMWR Surveill Summ. Dec 03 2021; 70(11): 1-16. PMID 34855725
9. International Medical Device Regulators Forum. Software as a Medical Device (SaMD): Key Definitions. 2013. http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf. Accessed June 6, 2023.
10. National Institute for Health and Care Excellence (NICE). Evidence standards framework for digital health technologies. 2021. nice.org.uk/corporate/ecd7/chapter/section-a-evidence-for-effectiveness-standards. Accessed June 6, 2023.
11. Zwaigenbaum L, Bauman ML, Choueiri R, et al. Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. Pediatrics. Oct 2015; 136 Suppl 1(Suppl 1): S60-81. PMID 26430170
12. Zwaigenbaum L, Bryson S, Lord C, et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. Pediatrics. May 2009; 123(5): 1383-91. PMID 19403506
13. Kleinman JM, Ventola PE, Pandey J, et al. Diagnostic stability in very young children with autism spectrum disorders. J Autism Dev Disord. Apr 2008; 38(4): 606-15. PMID 17924183
14. Canvas Dx Website. Available at https://canvasdx.com/ Accessed on June 6, 2023.
15. Abbas H, Garberson F, Liu-Mayo S, et al. Multi-modular AI Approach to Streamline Autism Diagnosis in Young Children. Sci Rep. Mar 19 2020; 10(1): 5014. PMID 32193406

16. Randall M, Egberts KJ, Samtani A, et al. Diagnostic tests for autism spectrum disorder (ASD) in preschool children. Cochrane Database Syst Rev. Jul 24 2018; 7(7): CD009044. PMID 30075057
17. 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. May 05 2022; 5(1): 57. PMID 35513550
18. Autism Spectrum Disorder: Links to Commonly Used Screening Instruments and Tools (AAP Toolkits). American Academy of Pediatrics. Available at https://publications.aap.org/toolkits/pages/asd-screening-tools. Accessed on June 6, 2023.
19. 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. Feb 2014; 53(2): 237-57. PMID 24472258
20. 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. Feb 16 2016; 315(7): 691-6. PMID 26881372

Documentation for Clinical Review

- No records required

Coding

This Policy relates only to the services or supplies described herein. Benefits may vary according to product design; therefore, contract language should be reviewed before applying the terms of the Policy.

The following codes are included below for informational purposes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy. Policy Statements are intended to provide member coverage information and may include the use of some codes for clarity. The Policy Guidelines section may also provide additional information for how to interpret the Policy Statements and to provide coding guidance in some cases.

Type	Code	Description
CPT®	None	
HCPCS	E1399	Durable medical equipment, miscellaneous

Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

Effective Date	Action
09/01/2022	New policy.
10/01/2023	Annual review. No change to policy statement. Literature review updated

Definitions of Decision Determinations

Medically Necessary: Services that are Medically Necessary include only those which have been established as safe and effective, are furnished under generally accepted professional standards to treat illness, injury or medical condition, and which, as determined by Blue Shield, are: (a) consistent

with Blue Shield medical policy; (b) consistent with the symptoms or diagnosis; (c) not furnished primarily for the convenience of the patient, the attending Physician or other provider; (d) furnished at the most appropriate level which can be provided safely and effectively to the patient; and (e) not more costly than an alternative service or sequence of services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of the Member's illness, injury, or disease.

Investigational/Experimental: A treatment, procedure, or drug is investigational when it has not been recognized as safe and effective for use in treating the particular condition in accordance with generally accepted professional medical standards. This includes services where approval by the federal or state governmental is required prior to use, but has not yet been granted.

Split Evaluation: Blue Shield of California/Blue Shield of California Life & Health Insurance Company (Blue Shield) policy review can result in a split evaluation, where a treatment, procedure, or drug will be considered to be investigational for certain indications or conditions, but will be deemed safe and effective for other indications or conditions, and therefore potentially medically necessary in those instances.

Prior Authorization Requirements and Feedback (as applicable to your plan)

Within five days before the actual date of service, the provider must confirm with Blue Shield that the member's health plan coverage is still in effect. Blue Shield reserves the right to revoke an authorization prior to services being rendered based on cancellation of the member's eligibility. Final determination of benefits will be made after review of the claim for limitations or exclusions.

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department at (800) 541-6652, or the Transplant Case Management Department at (800) 637-2066 ext. 3507708 or visit the provider portal at www.blueshieldca.com/provider.

We are interested in receiving feedback relative to developing, adopting, and reviewing criteria for medical policy. Any licensed practitioner who is contracted with Blue Shield of California or Blue Shield of California Promise Health Plan is welcome to provide comments, suggestions, or concerns. Our internal policy committees will receive and take your comments into consideration.

For utilization and medical policy feedback, please send comments to: MedPolicy@blueshieldca.com

Disclaimer: This medical policy is a guide in evaluating the medical necessity of a particular service or treatment. Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.

Appendix A

POLICY STATEMENT (No changes)	
BEFORE	**AFTER**
Digital Health Technologies: Diagnostic Applications **3.03.01** **Policy Statement:** I. Prescription digital health technologies for diagnostic application that have received clearance for marketing by the U.S. Food and Drug Administration as a diagnostic aid for autism spectrum disorder (Canvas Dx) are considered **investigational**.	**Digital Health Technologies: Diagnostic Applications** **3.03.01** **Policy Statement:** I. Prescription digital health technologies for diagnostic application that have received clearance for marketing by the U.S. Food and Drug Administration as a diagnostic aid for autism spectrum disorder (Canvas Dx) are considered **investigational**.