

# Primary Care Autism Screening and Later Autism Diagnosis

Paul S. Carbone, MD,<sup>a</sup> Kathleen Campbell, MD, MHSc,<sup>a</sup> Jacob Wilkes, BS,<sup>d</sup> Gregory J. Stoddard, MPH, MBA,<sup>b</sup> Kelly Huynh, MStat,<sup>d</sup> Paul C. Young,<sup>a</sup> Terisa P. Gabrielsen, PhD<sup>c</sup>

**OBJECTIVES:** To describe the proportion of children screened by the Modified Checklist for Autism in Toddlers (M-CHAT), identify characteristics associated with screen completion, and examine associations between autism spectrum disorder (ASD) screening and later ASD diagnosis.

## abstract

**METHODS:** We examined data from children attending 18- and 24-month visits between 2013 and 2016 from 20 clinics within a health care system for evidence of screening with the M-CHAT and subsequent coding of ASD diagnosis at age >4.75 years. We interviewed providers for information about usual methods of M-CHAT scoring and ASD referral.

**RESULTS:** Of 36 233 toddlers, 73% were screened and 1.4% were later diagnosed with ASD. Hispanic children were less likely to be screened (adjusted prevalence ratio [APR]: 0.95, 95% confidence interval [CI]: 0.92–0.98), and family physicians were less likely to screen (APR: 0.12, 95% CI: 0.09–0.15). Compared with unscreened children, screen-positive children were more likely to be diagnosed with ASD (APR: 10.3, 95% CI: 7.6–14.1) and were diagnosed younger (38.5 vs 48.5 months,  $P < .001$ ). The M-CHAT's sensitivity for ASD diagnosis was 33.1%, and the positive predictive value was 17.8%. Providers routinely omitted the M-CHAT follow-up interview and had uneven referral patterns.

**CONCLUSIONS:** A majority of children were screened for ASD, but disparities exist among those screened. Benefits for screen-positive children are improved detection and younger age of diagnosis. Performance of the M-CHAT can be improved in real-world health care settings by administering screens with fidelity and facilitating timely ASD evaluations for screen-positive children. Providers should continue to monitor for signs of ASD in screen-negative children.



Departments of <sup>a</sup>Pediatrics and <sup>b</sup>Internal Medicine, University of Utah, Salt Lake City, Utah; <sup>c</sup>Department of Counseling, Psychology, and Special Education, Brigham Young University, Provo, Utah; and <sup>d</sup>Pediatric Analytics, Intermountain Healthcare, Salt Lake City, Utah

Dr. Carbone conceptualized and designed the study, analyzed data, and drafted the initial manuscript; Mr Wilkes supervised data collection and assisted in data analyses; Dr Campbell critically reviewed the manuscript and approved the manuscript as submitted; Mr Stoddard supervised data analyses and approved manuscript as submitted; Ms Huynh assisted with data collection and analyses; Dr Young assisted with the conceptualization and design of the study and critically reviewed the manuscript; Dr Gabrielsen assisted with the conceptualization of the study and critically reviewed the manuscript; and all authors approved the final manuscript as submitted.

**DOI:** <https://doi.org/10.1542/peds.2019-2314>

Accepted for publication May 20, 2020

Address correspondence to Paul S. Carbone, MD, Department of Pediatrics, University of Utah, 295 Chipeta Way, Salt Lake City, UT 84109. E-mail: paul.carbone@hsc.utah.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2020 by the American Academy of Pediatrics

**WHAT'S KNOWN ON THIS SUBJECT:** Universal autism screening in toddlers is recommended, but it is unknown how frequently this occurs, what factors are associated with screening, and the performance characteristics of the most commonly used screening instrument in real-world health care settings.

**WHAT THIS STUDY ADDS:** Autism screening was completed in the majority of toddlers but was less likely to occur in Hispanic children. Children who screened positive were more likely to be diagnosed with autism and were diagnosed earlier, but false-negative screens were common.

**To cite:** Carbone PS, Campbell K, Wilkes J, et al. Primary Care Autism Screening and Later Autism Diagnosis. *Pediatrics*. 2020;146(2):e20192314

Early identification of autism spectrum disorder (ASD), through developmental surveillance and screening, allows children access to ASD-specific behavioral interventions that improve long-term outcomes.<sup>1-3</sup> Developmental surveillance that entails history-taking and observation for signs of ASD lacks sensitivity, especially during brief visits.<sup>4</sup> In 2007, to improve sensitivity and lower age of ASD diagnosis, the American Academy of Pediatrics recommended ASD screening at 18- and 24-month visits.<sup>2</sup> Despite this, there has not been a significant decrease in age of ASD diagnosis over the ensuing decade.<sup>5</sup>

Delayed identification may be related to several factors, including low rates of ASD screening, lower performance characteristics of ASD screening instruments in “real-world” practice, and barriers to screen-positive children being evaluated for ASD in a timely fashion. Among pediatricians, 81% report routinely administering ASD screening tools,<sup>6</sup> most commonly the Modified Checklist for Autism in Toddlers (M-CHAT) or the Modified Checklist for Autism in Toddlers, Revised.<sup>3,7</sup> Studies have shown that use of the M-CHAT leads to identification of ASD at younger ages. However, these studies were performed in practices with research support.<sup>3,7</sup> In a recent longitudinal population-based study of screening at 18 months, authors reported a low sensitivity of the M-CHAT in identifying children with ASD due to a high proportion of false-negative screens.<sup>8</sup> Additionally, the recommended M-CHAT follow-up interview for screen-positive children may be difficult to complete and is often omitted in community practices, affecting performance characteristics.<sup>7,9-11</sup> Therefore, studies in which researchers examine the M-CHAT in practices without support for scoring or completion of the screen and cost-free ASD evaluations are needed to identify

barriers to early identification of children with ASD.

With our study, we had 6 objectives: (1) estimate the rate of ASD screening at 18- and 24-month visits, (2) identify factors associated with screen completion, (3) evaluate whether screening is associated with improved identification of ASD, (4) determine if ASD screening was associated with a younger age of ASD diagnosis, (5) report performance characteristics of the M-CHAT, and (6) obtain qualitative information about M-CHAT administration from participating practices.

## METHODS

We analyzed electronic health record (EHR) data from children aged 16 to 30 months attending 18- and 24-month health supervision visits between 2013 and 2016 at Intermountain Healthcare (IHC) clinics in Utah. We included visits to clinics whose providers used an EHR M-CHAT result field that had been added in 2013, allowing providers to document whether the M-CHAT was completed (yes or no) and the result (negative [pass] or positive [fail]). To identify factors associated with M-CHAT completion, we analyzed visit-level variables (18- or 24-month visit, year of visit, provider type [pediatrician, family physician, advanced practice provider]) and patient and family variables (sex, race and ethnicity, insurance type, family location, and Area Deprivation Index [proxy measure of socioeconomic status (SES) that utilizes US Census Bureau data on poverty, housing, employment, and education<sup>12</sup>]). We identified children in our cohort diagnosed with ASD by *International Classification of Diseases* (ICD) codes (ICD-9 299.xx, ICD-10 F84.x) entered in visits up to May 31, 2019 (aged 57–107 months at time of data pull). In Utah, ASD diagnostic services are provided by IHC specialty clinics, in university-based clinics, in Title

V-supported clinics administered by the Utah Department of Health, and by providers in private practice. We estimated the age of ASD diagnosis by identifying the age at which the first ASD code was entered into the EHR, whether by the diagnosing professional or another provider within IHC on the date when seeing the child and learning of the diagnosis. To examine associations between ASD screening and ASD identification, we compared the prevalence of ASD at the time of our data pull in screened children with the prevalence of children who were not screened. We also compared the prevalence of ASD in screen-positive children with the prevalence of ASD in children who screened negative. To assess for associations between ASD screening and age of ASD diagnosis, we compared the age of first ASD diagnosis among screened children with those not screened as well as in those who screened positive with those who screened negative. By examining the screening result (positive or negative) and later diagnosis of ASD among all children included in our cohort, we were able to determine the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the M-CHAT. Finally, we obtained qualitative data by contacting physicians in participating practices. We asked how the M-CHAT was administered, how the first stage of the M-CHAT was typically scored, if the follow-up interview was administered when indicated, and the referral pattern for screen-positive children.

Our data included 3 levels of nesting: provider visits (1 or 2 visits) nested within patient, patients nested within provider, and providers nested within clinics. There was no interest in a within-patient analysis (visit 1 compared with visit 2), so regression models were fitted for either a visit-specific result or a combined-visit result, eliminating the need to include a visits within patient level of nesting in the models. We used multilevel

regression models to account for the remaining 2 levels of nesting: patients nested with provider, and providers nested within clinics. For describing patient and family and visit characteristics, we used ordinary descriptive statistics, rather than estimates derived from multilevel models. We modeled age of ASD diagnosis by receipt of screening (yes or no) and result of screen (positive or negative) using multilevel linear regression. We modeled ASD screening prevalence (visit-level data) and the association of screening and ASD diagnosis (child-level data) using multilevel binary Poisson regression with robust SEs. We choose binary Poisson regression because, for cohort or cross-sectional study designs, prevalence ratios can be accurately estimated directly by exponentiating the regression coefficient.<sup>13,14</sup> We added visit and patient and family variables to these models to obtain adjusted mean ages and adjusted prevalence ratios (APRs). We calculated test characteristics (sensitivity, specificity, PPV, NPV) of the M-CHAT in predicting a later diagnosis of ASD using simple 2 × 2 tables, with no adjustment for the multilevel structure in the data. Given the small amount of missing data (2.2%), all analyses are complete case analyses, in which patients with missing data were dropped from the analyses. When <5% of data are missing, particularly with large sample sizes, the complete case analysis approach provides sufficiently unbiased estimates.<sup>15</sup> All reported *P* values are for 2-sided comparisons using the Stata-15 statistical software (StataCorp LLC, College Station, TX).

The University of Utah Institutional Review Board determined the study was exempt from review.

## RESULTS

### Demographics

There were 48 307 children with 18- and 24-month visits during the study period. We excluded visits by children

seen at clinics not routinely using the M-CHAT result field, leaving data from 36 223 children seen during 59 139 visits at 20 clinics by 172 providers (Fig 1). Patient and family demographic variables and visit-level characteristics are shown in Tables 1 and 2. Our cohort consisted of predominantly non-Hispanic white children. The majority lived in urban or suburban areas, were privately insured, and were from families from higher socioeconomic backgrounds. Most visits were completed by pediatricians.

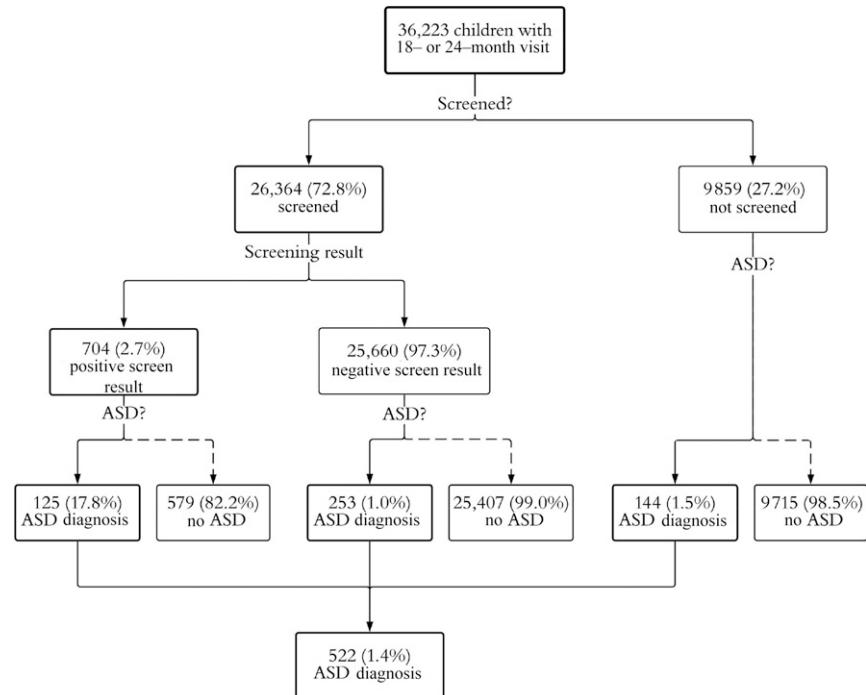
### M-CHAT Screening

Among 36 223 children, 72.8% were screened at either the 18- or the 24-month visit; among those who attended 18- and 24-month visits (*n* = 20 072), 54.4% were screened twice (Table 3). More children were screened at the 24-month visit than at the 18-month visit (66% and 62%; *P* < .001), but fewer had positive screens (1.6% vs 2.4%; *P* < .001). A total of 378 (72%) of the 522 children diagnosed with ASD had

been screened. Of these, 228 (60.3%) were screened once; 150 (39.7%) were screened twice. A total of 165 (72%) of the 228 children with ASD who had been screened once (at either 18- or 24-month visits) screened negative. Eighty-eight (59%) of the 150 children screened twice screened negative at both visits. Thus, of the 378 children later diagnosed with ASD, 253 (67%) had screened negative (Fig 1). Table 4 shows the factors associated with M-CHAT completion. In univariable analysis, Hispanic children and those from lowest socioeconomic backgrounds were least likely to be screened, and family physicians were less likely to screen. In multivariable analysis, visits with Hispanic children (APR: 0.96, 95% confidence interval [CI]: 0.93–0.99) and by family physicians (APR: 0.12, 95% CI: 0.09–0.15) remained associated with a lower likelihood of M-CHAT completion.

### M-CHAT Screening and ASD Diagnosis

Among the entire cohort (*n* = 36 223), 522 children (1.4%, 1 in 69) had been



**FIGURE 1**

Flowchart of ASD screening and diagnosis for all patients studied.

**TABLE 1** Patient and Family Characteristics

	Entire Cohort, n = 36 223	ASD Cohort, n = 522 (1.4%, 1 in 69)
Sex (n = 36 222), n (%)		
Male	18 741 (51.7)	403 (77.2)
Female	17 481 (48.3)	119 (22.8)
Race and ethnicity, n (%)		
Non-Hispanic white	29 676 (81.9)	399 (76.4)
Hispanic	3935 (10.9)	70 (13.4)
Other	976 (2.7)	19 (3.6)
Asian or Pacific Islander	848 (2.3)	19 (3.6)
Non-Hispanic black	422 (1.2)	11 (2.1)
Missing	366 (0.9)	4 (0.8)
Location, n (%)		
Urban or suburban	24 419 (67.4)	365 (69.9)
Rural	713 (2.0)	10 (1.9)
Mixed	10 308 (28.5)	143 (27.4)
Missing	783 (2.2)	4 (0.8)
Insurance, n (%)		
Private	21 037 (58.1)	302 (57.9)
Government	10 375 (28.6)	146 (28.0)
Other	3608 (10.0)	59 (11.3)
Self-pay	1164 (3.2)	15 (2.9)
Missing	39 (0.1)	0 (0.0)
SES, n (%)		
1 (highest)	8471 (23.4)	77 (14.8)
2	9832 (27.1)	132 (25.3)
3	7293 (20.1)	118 (22.6)
4	5225 (14.4)	93 (17.8)
5 (lowest)	4619 (12.8)	98 (18.8)
Missing	783 (2.2)	4 (0.8)
ASD diagnosis		
Yes, n (%)	—	522 (1.4)
No, n (%)	35 701 (98.6)	—
Average age of diagnosis (n = 522), mo (SD)	—	46.8 (17.7)
Time to diagnosis after positive screen, d (SD)	—	646 (476)

—, not applicable.

diagnosed with ASD at the time of our data pull (age range 57–107 months). The prevalence of ASD in children screened and not screened in both univariable and multivariable analyses was similar. In multivariable analysis, children who screened

positive were 17 times more likely to be diagnosed with ASD than screen-negative children (APR: 17.0, 95% CI: 13.1–22.0) and 10 times more likely than children not screened (APR: 10.3, 95% CI: 7.6–14.1) (Table 5).

**TABLE 2** Visit Characteristics

All Visits	N = 59 139, n (%)
Visit type	
18-mo	30 695 (51.9)
24-mo	28 444 (48.1)
Provider type	
Pediatrician	55 815 (94.4)
Family physician	2557 (4.3)
Advanced practice clinician	723 (1.2)
Other	44 (0.0)
Visit year	
2013	10 329 (17.5)
2014	16 464 (27.8)
2015	16 849 (28.5)
2016	15 497 (26.2)

## M-CHAT Screening and Age of ASD Diagnosis

There was no difference in the age of ASD diagnosis in children who were screened and not screened (46.5 months [95% CI: 43.5–49.6] versus 48.5 months [95% CI: 44.6–52.3], respectively,  $P = .34$ ). However, among screened children later diagnosed with ASD, those who screened positive were diagnosed at a younger age compared with those who screened negative (38.5 months [95% CI: 34.2–42.7] compared with 50.6 months [95% CI: 47.0–54.1];  $P < .001$ ) (Table 6). Screen-positive children were also diagnosed 10 months earlier than children not screened (38.5 months [95% CI: 34.2–42.7] vs 48.5 months [95% CI: 44.8–52.2];  $P < .001$ ).

## Performance Characteristics of the M-CHAT

We assessed performance characteristics of the M-CHAT in predicting a diagnosis of ASD among the 26 364 children who had been screened and found the following: sensitivity of 33.1% (95% CI: 28.3–38.1), specificity of 97.8% (95% CI: 97.6–97.9), PPV of 17.8% (95% CI: 15.0–20.8), and NPV of 99.0% (95% CI: 98.9–99.1). There was a significantly higher sensitivity in nonwhite compared with white children and in children screened twice compared with once. The PPV in boys was higher compared with girls and in children from lower SES families compared with those from higher SES families (Table 7).

## M-CHAT Administration in Practices

After contacting all 20 participating practices, we were able to interview physicians from 12 practices who described how the M-CHAT was typically administered. All practices administered M-CHATS on paper because there were no electronic or online means to screen through the EHR. With regards to scoring, 10 practices used the recommended

**TABLE 3** M-CHAT Screening Rate and Results, by Patients and Visits

	n (%)
Children seen for 18- or 24-mo visits	36 223
One M-CHAT at either 18- or 24-mo visit	
Completed	26 364 (72.8)
Not completed	9 859 (27.2)
M-CHAT at 18- and 24-mo visits (if attended both visits, n = 20 072)	
Completed	10 916 (54.4)
Not completed	9 156 (45.6)
M-CHAT result (n = 26 364)	
Positive (on at least 1 screen)	704 (2.7)
Negative	25 660 (97.3)
Visits at either 18- or 24-mo	59 139
M-CHAT	
Completed	37 426 (63.3)
Not completed	21 713 (36.7)
18-mo visit M-CHAT (n = 30 695)	
Completed	18 878 (61.5)
Not completed	11 817 (38.5)
24-mo visit M-CHAT (n = 28 444)	
Completed	18 548 (65.2)
Not completed	9 896 (34.8)
18-mo M-CHAT result (n = 18 878)	
Positive	460 (2.4)
Negative	18 418 (97.6)
24-mo M-CHAT result (n = 18 548)	
Positive	303 (1.6)
Negative	18 245 (98.4)
Children with ASD	522
At least 1 M-CHAT	
Completed	378 (72.4)
Not completed	144 (27.6)
Completed M-CHATS (n = 378)	
1 M-CHAT at either 18- or 24-mo visit	228 (60.3)
18- and 24-mo visit	150 (39.7)
M-CHAT result (n = 378)	
Positive	125 (33.1)
Negative	253 (66.9)
M-CHAT negative (n = 253)	
18-mo visit but not 24-mo visit	83 (32.8)
24-mo visit but not 18-mo visit	82 (32.4)
18- and 24-mo visits	88 (34.8)

scoring, 1 used a mix of “eyeballing” completed screens and scoring, and 1 did not score M-CHATS. In 7 of 10 practices scoring the M-CHAT, medical assistants typically administered and scored the first step of the M-CHAT. One practice typically administered follow-up interviews, whereas 9 practices did not, and 1 practice used a case manager to “informally” clarify missed items using the follow-up interview as a guide. Referral patterns after positive screens were mixed. Three practices referred simultaneously for ASD evaluations and to Early Intervention Program (EIP). Three

practices typically referred only to EIP, and 4 practices referred for ASD evaluations but not to EIP. Two practices had no set referral patterns.

## DISCUSSION

In this study of ASD screening in real-world practices within a large health care system, we found approximately three-fourths of children were screened for ASD at least once and half were screened at both recommended ages. Screening occurred less often in Hispanic children and at visits with family physicians. Children who screened

positive were more likely to later be identified with ASD and were diagnosed at a younger age. Our ability to follow-up on outcomes of screen-negative children allowed for broader estimates of performance characteristics than previous studies. In doing so, we found a lower-than-expected sensitivity and PPV than previously estimated (33% and 18%, respectively). Qualitative data regarding lack of fidelity in screen administration and uneven referral patterns after positive screens provides context into why our findings may differ from studies of the M-CHAT with research support.

Previous estimates of ASD screening have varied significantly (between 17% and 81%) and relied on physician report of “usual practice.”<sup>6,16</sup> Our study, along with 2 recent studies, adds to understanding of ASD screening because estimates are based on data from actual visits rather than physician recall and indicates that, although ASD screening is far from universal, a high proportion of children are screened at least once.<sup>10,11</sup>

With our study, we provide important insight into disparities in ASD screening; researchers in previous studies suggested that ASD screening in primary care reduces racial and ethnic disparities in ASD identification and age of diagnosis.<sup>17</sup> Hispanic children in our cohort were less likely to be screened compared with non-Hispanic white children, a finding supported by one previous study.<sup>11</sup> We also found family physicians rarely administered ASD screening tools. The American Academy of Family Physicians does not recommend universal ASD screening, which may explain differences in screening between pediatricians and family physicians.<sup>18</sup> No previous studies have had estimated rates of ASD screening among family physicians, although researchers of one qualitative study of a small group of family physicians

**TABLE 4** Factors Associated With M-CHAT Completion

Characteristic	Completed M-CHAT, n (%)	Univariable Analysis			Multivariable Analysis <sup>a</sup>		
		PR	95% CI	P	APR	95% CI	P
<b>Sex</b>							
Male	23 432 (76.6)	Reference	—	—	—	—	—
Female	21 993 (77.0)	1.00	0.98–1.02	.58	—	—	—
<b>Race and ethnicity</b>							
Non-Hispanic white	37 553 (77.2)	Reference	—	—	—	—	—
Hispanic	4602 (73.1)	0.95	0.92–0.98	.003	0.96	0.93–0.99	.009
Non-Hispanic black	508 (76.7)	0.96	0.87–1.04	.31	0.96	0.88–1.05	.369
Asian or Pacific Islander	1158 (79.9)	0.97	0.92–1.03	.34	0.97	0.91–1.03	.332
<b>Location</b>							
Urban	31 459 (78.3)	Reference	0.99	—	—	—	—
Rural	669 (59.2)	0.99	0.92–1.08	.95	1.01	0.93–1.09	.794
Mix	12 527 (75.0)	1.01	0.99–1.04	.18	1.02	0.94–1.10	.654
<b>SES</b>							
1 (highest)	11 797 (83.4)	Reference	—	—	—	—	—
2	12 874 (79.4)	1.00	0.98–1.03	.67	1.01	0.98–1.03	.574
3	8915 (74.8)	1.00	0.97–1.02	.78	1.00	0.97–1.03	.895
4	6164 (73.3)	1.00	0.97–1.03	.99	1.01	0.98–1.05	.476
5 (lowest)	4905 (67.0)	0.96	0.92–0.99	.02	0.97	0.94–1.01	.159
<b>Insurance</b>							
Medicaid	8613 (74.4)	Reference	—	—	—	—	—
Private	25 304 (76.0)	1.01	0.98–1.03	.63	—	—	—
Self-pay	2196 (78.2)	1.01	0.97–1.04	.71	—	—	—
<b>Provider type</b>							
Pediatrician	44 600 (79.9)	Reference	—	—	—	—	—
Family physician	240 (9.4)	0.12	0.09–0.15	<.001	0.12	0.09–0.15	<.001
Advanced practice clinician	558 (77.2)	0.76	0.60–1.02	.07	0.78	0.56–1.03	.077
<b>Visit type</b>							
18-mo	23 569 (76.8)	Reference	—	—	—	—	—
24-mo	21 857 (76.8)	0.99	0.98–1.02	.10	—	—	—

PR, prevalence ratio; —, not applicable.

<sup>a</sup> Covariates included in multivariable model: sex, race and ethnicity, location, SES, and insurance.

found that, rather than screening, participants relied on developmental surveillance to identify children with ASD.<sup>19</sup> Given that family physicians provide 16% to 21% of pediatric care, this new finding uncovers an opportunity to further increase ASD screening rates in the United States.<sup>20</sup>

A strength of our study was the ability to longitudinally follow children seen for 18- and 24-month visits, allowing an estimate of performance characteristics of the M-CHAT not possible in previous validation studies.<sup>3</sup> Even after controlling for potential confounding

factors, we found a strong association between a positive ASD screen and a later diagnosis of ASD. This new finding supports the policy of ASD screening as a means of improving identification of children with ASD. We found that the age of diagnosis did not differ between screened and

**TABLE 5** Associations Between ASD Diagnosis and M-CHAT Completion or M-CHAT Screen-Positive

Characteristic	ASD n (%)	No ASD n (%)	Univariable analyses			Multivariable Analyses <sup>a</sup>		
			PR	95% CI	P	APR	95% CI	P
<b>M-CHAT completion</b>								
Screened	378 (1.4)	25 986 (98.6)	0.90	0.73–1.10	.29	0.87	0.69–1.10	.25
Not screened	144 (1.5)	9715 (98.5)	Reference	—	—	—	—	—
<b>M-CHAT result</b>								
Screen-positive	125 (17.8)	579 (82.2)	19.0	15.2–23.8	<.001	17.0	13.1–22.0	<.001
Screen-negative	253 (1.0)	25 407 (99.0)	Reference	—	—	—	—	—
<b>M-CHAT</b>								
Screen-positive	125 (17.8)	579 (82.2)	11.4	8.8–14.9	<.001	10.3	7.6–14.1	<.001
Not screened	144 (1.5)	9715 (98.5)	Reference	—	—	—	—	—

PR, prevalence ratio; —, not applicable.

<sup>a</sup> Covariates included in multivariable model: sex, race and ethnicity, location, SES, insurance.

**TABLE 6** Age of ASD Diagnosis by M-CHAT Completion and Result

ASD Screening Test	Age of First ASD Diagnosis, mo <sup>a</sup>	95% CI	P
M-CHAT			
Screened	46.5	43.5–49.6	.34
Not screened	48.5	44.6–52.3	Reference
M-CHAT result			
Screen-positive	38.5	34.2–42.7	<.001
Screen-negative	50.6	47.0–54.1	Reference

<sup>a</sup> Covariates included in multivariable model: sex, race and ethnicity, location, SES, and insurance.

unscreened children. Although this could be interpreted as a lack of impact of screening, we did find a clinically important difference when we separated the screened group into those who screened negative and those who screened positive. Specifically, we found that children who screened positive had a diagnosis of ASD in the EHR 12 months earlier than those who screened negative and 10 months earlier than those not screened. Therefore, these findings add new evidence that when a child screens positive, screening may lower the age of diagnosis. We also conclude from these results that the lack of difference between the screened and unscreened groups is being driven by

later diagnosis in the subgroup with false-negative results on screening. This result highlights the importance of developing a system or a screening instrument that can reduce false negatives. Our uncontrolled, real-world setting does not allow us to draw conclusions about whether false-negative screens are driven by children who were missed by the screening instrument or who were missed by a flawed system of screening and referrals. Our findings support the American Academy of Pediatrics recommendation for universal screening in that more children and younger children with ASD will have opportunities to access evidence-based therapies that improve outcomes.<sup>2</sup>

We identified a large number of children later diagnosed with ASD who screened negative (false negatives), leading to unexpectedly low sensitivity of the M-CHAT. This finding is consistent with previous studies in which researchers identified groups of children with ASD who passed a single M-CHAT despite having deficits in social communication, suggesting that providers should closely monitor for signs of ASD with developmental surveillance and broadband developmental screening tests and refer for concerns even after a negative ASD screen.<sup>2,21–23</sup> Similar to another study conducted in a real-world setting, we found that the sensitivity of the M-CHAT improved among children screened twice, suggesting that providers should screen at both visits.<sup>11</sup>

We additionally found children who screened positive but were later identified as not having ASD in the EHR. This resulted in a lower PPV, 17.8%, than reported in the most recent validation study of the M-CHAT (47%) but closer to that in another study (14.6%) done in a similar real-world setting.<sup>3,11</sup> Our

**TABLE 7** Performance Characteristics of the M-CHAT Within the Entire Cohort and Subgroups in Predicting Later Diagnosis of ASD

Group	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
Entire cohort (n = 26 364)	33.1 (28.3–38.1)	97.8 (97.6–97.9)	17.8 (15.0–20.8)	99.0 (98.9–99.1)
Race and ethnicity				
White (n = 21 762)	30.1 (24.8–35.7)	98.1 (97.9–98.2)	17.1 (13.9–20.7)	99.1 (98.9–99.2)
Nonwhite (n = 3638)	43.1 (31.4–55.3)	96.4 (95.8–97.0)	19.6 (13.7–26.7)	98.8 (13.7–26.7)
Comparison	P = .04	P < .001	P = .47	P = .19
Sex				
Male (n = 13 609)	34.8 (29.4–40.5)	97.4 (97.1–97.6)	22.9 (19.1–27.0)	98.5 (98.3–98.7)
Female (n = 12 754)	26.6 (17.3–37.7)	98.2 (98.0–98.4)	8.4 (5.3–12.6)	99.5 (99.4–99.6)
Comparison	P = .17	P < .001	P < .001	P < .001
M-CHAT completion				
18- or 24-mo (n = 15 448)	27.6 (21.9–33.9)	97.9 (97.7–98.2)	16.8 (13.1–20.9)	98.9 (98.7–99.1)
18- and 24-mo (n = 10 916)	41.3 (33.4–49.7)	97.5 (97.2–97.8)	18.9 (14.8–23.6)	99.2 (99.0–99.3)
Comparison	P = .006	P = .03	P = .46	P = .03
Insurance				
Medicaid (n = 5022)	27.5 (17.5–39.6)	97.8 (97.3–98.2)	14.6 (9.0–21.9)	99.0 (98.7–99.2)
Private (n = 14 685)	37.0 (30.4–43.9)	97.9 (97.6–98.1)	20.1 (16.2–24.4)	99.1 (98.9–99.2)
Comparison	P = .15	P = .70	P = .17	P = .57
SES				
High SES (n = 14 165)	34.6 (27.2–42.5)	97.8 (97.5–98.0)	14.9 (11.5–19.0)	99.2 (99.1–99.4)
Low SES (n = 11 700)	32.6 (26.3–39.3)	97.7 (97.5–98.0)	21.3 (17.0–26.1)	98.7 (98.5–98.9)
Comparison	P = .68	P = .91	P = .03	P < .001

qualitative data revealed low usage of the recommended M-CHAT follow-up interview, which is intended to reduce false positives. Previous studies of the M-CHAT have used trained research assistants to administer the follow-up interview by phone; pediatricians might have difficulty completing follow-up interviews in the context of brief visits.<sup>3,7</sup> Interventions that have been shown to improve pediatrician completion of follow-up interviews include online or digital decision support tools.<sup>24,25</sup> Additionally, what we defined as “false positives” may in reality be children with ASD who were not referred for ASD evaluations after screening positive and thus not formally diagnosed, a practice that has been reported by providers in this and other studies.<sup>10,16,22</sup> In contrast, families of children who screened positive in validation studies of the M-CHAT were offered free and timely ASD evaluations, which facilitated identification and a young age of diagnosis. Families of children in our study, like those across the country, may struggle to access affordable and timely ASD diagnostic evaluations.<sup>26,27</sup> For example, there was an average of 646 days between a positive screening test and the first ASD codes entered into the EHR. This suggests that families were not referred after positive screens, had difficulty accessing evaluations once they were referred, or chose not to pursue ASD evaluations in a timely fashion. A similar study showed that, among children who screened positive, only 31% were referred for specialty evaluation and, among those, 65% were completed, suggesting that further interventions are needed to facilitate referral completion.<sup>10</sup> Taken together, the lower-than-expected performance characteristics of the M-CHAT found in this study may have been due to system-level issues seen in real-world health care settings (not administering the screen twice, not administering the follow-up

interview, not referring after positive screens, difficulty accessing ASD evaluations) and issues related to child characteristics or the screen itself, all of which warrant further study. Various strategies to address system-level barriers at the practice level, such as electronic decision support, clinical informatics reporting, and follow-up phone calls, may allow for tracking of positive screens and referral completion.<sup>28</sup>

We found differences in performance characteristics of the M-CHAT in several subgroups. The M-CHAT had a lower sensitivity and lower PPV in non-Hispanic white children and children from higher SES backgrounds, respectively. Researchers of 2 previous studies assessed the accuracy of screening and found mixed results with regards to PPV in nonwhite children compared with white children.<sup>9,11</sup> We also identified a lower PPV in girls compared with boys, a finding that differs from the most recent validation of the M-CHAT but aligns with an M-CHAT study that systematically followed-up with all children screened.<sup>3,11</sup> Given that girls with ASD are diagnosed less often and later, improving the accuracy of ASD screening in girls should be prioritized. Disparities in availability of screening and diagnostic services across subgroups has implications for psychometric properties of screeners; therefore, these results deserve more in-depth investigation in future studies.

The limitations of our study are related to reliance on EHR data to identify ASD screening and diagnoses. Without being able to clinically confirm ASD diagnoses, results regarding performance characteristics of the M-CHAT should be considered preliminary. However, the accuracy of ASD diagnoses in our study is supported by the similar prevalence of ASD in our cohort compared with national prevalence estimates of same-aged children.<sup>5,29</sup> Additionally, our estimates of time to diagnosis are limited by our use of time when the

diagnosis entered the EHR and, therefore, likely overestimate the time by some unknown amount and should be interpreted with caution. Findings from a previous M-CHAT validation study suggest that we should have expected a screen-positive rate of 7.4% if not using the follow-up interview, and yet our screen-positive rate was only 2.7%.<sup>3</sup> In a more recent study, however, done within a real-world health care setting that did not use the follow-up interview, researchers found a screen-positive rate similar to ours (3%).<sup>10</sup> The reasons for a lower-than-expected screen-positive rate in real-world settings in which the M-CHAT follow-up interview is not used is one that deserves further study. In our study, we hypothesize that the lower-than-expected proportion of positive screens could stem from several possible causes. One may be that physicians informally asked follow-up questions, although they reported not formally administering the follow-up interview. Another could be that errors in scoring of paper screeners resulted in positive M-CHATS recorded as negative, as was seen in a previous study that rescored paper forms of the M-CHAT.<sup>25</sup> If incorrect scoring of paper screeners was responsible for the lower-than-expected screen-positive rate in our study, health care systems could invest in electronic administration and scoring to improve the accuracy of ASD screening. Although the EHR data we had access to did not allow us to review the scoring of the M-CHAT, we are in the process of performing chart reviews of screen-negative children with ASD to investigate how M-CHAT scoring was performed and whether there were concerns related to ASD documented during visits, as reported in a study finding delays in social communication in children with ASD who screened negative at 18 months.<sup>21</sup> We are also limited by reliance on data and processes from one large health system in Utah, which may not be representative of other health care

systems or patient populations. Finally, there was a lack of information about Spanish language availability of the M-CHAT among the practices, although administering the M-CHAT in Spanish did not eliminate disparities in screen completion between English- and Spanish-speaking families in a previous study.<sup>11</sup> A closer investigation of the factors that might enhance efforts to increase the parity of screening rates in Hispanic families is needed. Despite these limitations, our study provides new information and perspective about ASD screening in real-world primary care practices.

## CONCLUSIONS

Our results suggest progress toward universal ASD screening in primary

care practices. Still, only half of children in our study were screened at both 18- and 24-month visits, suggesting further work to be done to encourage providers to complete screens as recommended, with more efforts to screen Hispanic children. Further advocacy and education are needed to encourage family physicians to screen for ASD. More resources are needed for implementing ASD screening at the practice and health care system level, with particular attention to administering ASD screening tools with fidelity, prompt referral of children who are found to be at-risk, and increasing the availability of ASD diagnostic providers to facilitate prompt ASD evaluations.

## ACKNOWLEDGMENT

We thank Catherine Jolma, MD for her review of the article.

## ABBREVIATIONS

- APR: adjusted prevalence ratio  
ASD: autism spectrum disorder  
CI: confidence interval  
EHR: electronic health record  
EIP: Early Intervention Program  
ICD: *International Classification of Diseases*  
IHC: Intermountain Healthcare  
M-CHAT: Modified Checklist for Autism in Toddlers  
NPV: negative predictive value  
PPV: positive predictive value  
SES: socioeconomic status

**FINANCIAL DISCLOSURE:** Dr Campbell discloses that she is an inventor on a patent related to screening for autism spectrum disorder; the other authors have indicated they have no financial relationships relevant to this article to disclose.

**FUNDING:** Supported by the University of Utah Population Health Research Foundation, with funding in part from the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1TR002538 (formerly 5UL1TR001067-05, 8UL1TR000105, and UL1RR025764). The research reported in this publication was supported in part by the Utah Stimulating Access to Research in Residency Transition Scholar under award number 1R38HL143605-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funded by the National Institutes of Health (NIH).

**POTENTIAL CONFLICT OF INTEREST:** Dr Campbell discloses that she is an inventor on a patent related to screening for autism spectrum disorder; the other authors have indicated they have no potential conflicts of interest to disclose.

**COMPANION PAPER:** A companion to this article can be found online at [www.pediatrics.org/cgi/doi/10.1542/peds.2020-1467](http://www.pediatrics.org/cgi/doi/10.1542/peds.2020-1467).

## REFERENCES

- 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*. 2015; 136(suppl 1):S60–S81
- Johnson CP, Myers SM; American Academy of Pediatrics Council on Children With Disabilities. Identification and evaluation of children with autism spectrum disorders. *Pediatrics*. 2007; 120(5):1183–1215
- Robins DL, Casagrande K, Barton M, Chen CM, Dumont-Mathieu T, Fein D. Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*. 2014; 133(1):37–45
- Gabrielsen TP, Farley M, Speer L, Villalobos M, Baker CN, Miller J. Identifying autism in a brief observation. *Pediatrics*. 2015;135(2). Available at: [www.pediatrics.org/cgi/content/full/135/2/e330](http://www.pediatrics.org/cgi/content/full/135/2/e330)
- Baio J, Wiggins L, Christensen DL, et al. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR Surveill Summ*. 2018;67(6):1–23
- Coury D, Wolfe A, Lipkin PH, Hyman S. Screening of young children for autism spectrum disorders: results from a national survey of pediatricians. In: Proceedings from the Pediatric Academic Societies Annual Meeting; May 6–9, 2017; San Francisco, CA.
- Chlebowski C, Robins DL, Barton ML, Fein D. Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*. 2013;131(4). Available at: [www.pediatrics.org/cgi/content/full/131/4/e1121](http://www.pediatrics.org/cgi/content/full/131/4/e1121)
- Stenberg N, Bresnahan M, Gunnes N, et al. Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatr Perinat Epidemiol*. 2014;28(3):255–262
- Khawaja MK, Hazzard AP, Robins DL. Sociodemographic barriers to early detection of autism: screening and evaluation using the M-CHAT, M-CHAT-R, and M-CHAT-R/F. *J Dev Behav Pediatr*. 2017;18(1):1–7

- and follow-up. *J Autism Dev Disord.* 2015;45(6):1797–1808
10. 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):e20183326
  11. Guthrie W, Wallis K, Bennett A, et al. Accuracy of autism screening in a large pediatric network. *Pediatrics.* 2019; 144(4):e20183963
  12. Knighton AJ, Savitz L, Belnap T, Stephenson B, VanDerslice J. Introduction of an area deprivation Index measuring patient socioeconomic status in an integrated health system: implications for population health. *EGEMS (Wash DC).* 2016;4(3):1238
  13. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. *Am J Epidemiol.* 2004;160(4):301–305
  14. Zou GY, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res.* 2013;22(6):661–670
  15. Harrell FE Jr. *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis.* New York, NY: Springer-Verlag; 2001
  16. Self TL, Parham DF, Rajagopalan J. Autism spectrum disorder early screening practices: a survey of physicians. *Comm Disord Q.* 2015;36(4): 195–207
  17. Herlihy LE, Brooks B, Dumont-Mathieu T, et al. Standardized screening facilitates timely diagnosis of autism spectrum disorders in a diverse sample of low-risk toddlers. *J Dev Behav Pediatr.* 2014; 35(2):85–92
  18. Sanchack KE, Thomas CA. Autism spectrum disorder: primary care principles. *Am Fam Physician.* 2016; 94(12):972–979
  19. Fenikilé TS, Ellerbeck K, Filippi MK, Daley CM. Barriers to autism screening in family medicine practice: a qualitative study. *Prim Health Care Res Dev.* 2015; 16(4):356–366
  20. Phillips RL Jr., Bazemore AW, Dodoo MS, Shipman SA, Green LA. Family physicians in the child health care workforce: opportunities for collaboration in improving the health of children. *Pediatrics.* 2006;118(3): 1200–1206
  21. Øien RA, Schjølberg S, Volkmar FR, et al. Clinical features of children with autism who passed 18-month screening. *Pediatrics.* 2018;141(6): e20173596
  22. Beacham C, Reid M, Bradshaw J, et al. Screening for autism spectrum disorder: profiles of children who are missed. *J Dev Behav Pediatr.* 2018;39(9): 673–682
  23. Beuker KT, Schjølberg S, Lie KK, Swinkels S, Rommelse NN, Buitelaar JK. ESAT and M-CHAT as screening instruments for autism spectrum disorders at 18 months in the general population: issues of overlap and association with clinical referrals. *Eur Child Adolesc Psychiatry.* 2014;23(11): 1081–1091
  24. Sturner R, Howard B, Bergmann P, et al. Autism screening with online decision support by primary care pediatricians aided by M-CHAT/F. *Pediatrics.* 2016; 138(3):e20153036
  25. Campbell K, Carpenter KLH, Espinosa S, et al. Use of a digital modified checklist for autism in toddlers - revised with follow-up to improve quality of screening for autism. *J Pediatr.* 2017; 183:133–139.e1
  26. 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
  27. McKenzie K, Forsyth K, O'Hare A, et al. Factors influencing waiting times for diagnosis of Autism Spectrum Disorder in children and adults. *Res Dev Disabil.* 2015;45–46:300–306
  28. Talmi A, Bunik M, Asherin R, et al. Improving developmental screening documentation and referral completion. *Pediatrics.* 2014;134(4). Available at: [www.pediatrics.org/cgi/content/full/134/4/e1181](http://www.pediatrics.org/cgi/content/full/134/4/e1181)
  29. Christensen DL, Bilder DA, Zahorodny W, et al. Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *J Dev Behav Pediatr.* 2016; 37(1):1–8