Autism Screening at 18 Months of Age: a Comparison of the Q-CHAT and the M-CHAT

Raymond Sturner (rsturner@childhealthcare.org)
Johns Hopkins School of Medicine: Johns Hopkins University School of Medicine
https://orcid.org/0000-0002-4029-5048
Barbara Howard
Johns Hopkins Medical Institutions: Johns Hopkins Medicine
Paul Bergmann
Foresightlogic
Shana Attar
University of Washington Seattle Campus: University of Washington
Lydia Stewart
Center for Promotion of Child Development through Primary Care
Kerry Bet
Center for Promotion of Child Development through Primary Care
Carrie Allison
Cambridge University: University of Cambridge
Simon Baron-Cohen
Cambridge University: University of Cambridge

Research

Keywords: autism screening, developmental screening, M-CHAT, Q-CHAT

Posted Date: March 31st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-359293/v1

Abstract

Background: Autism screening is recommended at 18 and 24-month pediatric well visits. The M-CHAT-R screen requires a follow-up interview (M-CHAT-R/F) when positive. M-CHAT-R/F may be less accurate for 18- than 24-month-olds and accuracy for identification prior to two years is not known in samples including children passing screens. Since autism symptoms may emerge gradually, items with ordinal responses, such as Q-CHAT-10, might better capture autism signs than the dichotomous (i.e., yes/no) items in M-CHAT-R. The objective of this study was to determine the accuracy of the M-CHAT-R/F and the Q-CHAT-10 for predicting autism in a sample of children who screen positive or negative at 18 months old.

Methods: This is a community pediatrics validation study with screen positive (96) and age and practice-matched screen negative children (314) recruited for diagnostic evaluations completed prior to two years old.

Clinical diagnosis of autism based on results of in-person diagnostic autism evaluations by research reliable testers blind to screening results and using the ADOS-2 Toddler module, Mullen Scales of Early Learning (MSEL) per standard guidelines, and ADOS-T algorithm.

Results: While the M-CHAT-R/F had higher specificity and PPV compared to M-CHAT-R as expected, performance characteristics, including improved sensitivity with similar PPV, favored Q-CHAT-10 (scoring adjusted to capture the full range of frequency items).

Limitations: Many parents declined participation and the sample skewed toward higher educated parents. Results cannot be extended to older ages.

Conclusions: The limitations of the M-CHAT-R/F at the 18-month-visit include low sensitivity with minimal balancing benefit of improved PPV from the follow-up interview. The Q-CHAT-10 shows advantages over M-CHAT-R/F including requiring half the number of items and no follow-up interview, yet it did not improve the low PPV at 18 months. Ordinal, rather than dichotomous, scoring of autism screening items appears to be beneficial at this age. The Q-CHAT-10 with adjusted scoring is thus a better alternative to the M-CHAT at the 18-month-visit, providing better detection of children who can be diagnosed at this age even though psychometrics of both approaches still fall below recommended standards, consistent with growing recognition that screening needs to also occur repeatedly beyond this age.

Background

Autism Spectrum Disorder (henceforth autism) is a prevalent and life-long condition, with a rate of 1 in 54[1] by 8 years of age. A strong association between early evidence-based intervention with improved long-term outcomes for children with autism[2, 3, 4, 5] is one rationale for the recommendation by the American Academy of Pediatrics (AAP)[6] and the Centers for Disease Control (CDC)[7] for screening of autism in all children at 18- and 24-months. However, the 2015 US Preventive Services Task Force

(USPSTF) asserted that additional data are needed, in part due to a lack of adequate validation of the recommended tools in community samples.[8]

The CHAT, one of the first validated autism screening tests,[9] showed initial promise for screening at 18-months with a high concurrent positive predictive value (PPV). However, at a six-year follow-up, the 18-month CHAT had only identified 38% of autistic children with an autism diagnosis.[10] The low sensitivity of the CHAT, and a desire to eliminate its child observation items, led to modifications of the screen, e.g. the Modified CHAT (M-CHAT),[11] that also added parent report items. Another modification, the Quantitative CHAT(Q-CHAT),[12] changed the dichotomous responses of the CHAT and M-CHAT (yes/no) to ordinal responses (how much/often), to acknowledge that autistic traits lie on a dimension.[13]

While the M-CHAT is the most widely used autism screening test, it may not exceed the CHAT in long-term sensitivity, in part due to the late diagnosis of autism as reviewed below. Neither the M-CHAT nor the Q-CHAT have been studied in a representative community population of 18-month-olds with validation testing that includes both screen negatives and positives as needed to estimate sensitivity. Furthermore, the different item response approaches of the ordinal Q-CHAT-10 and the dichotomous M-CHAT-R have never been directly compared.

Additionally, the revised M-CHAT (M-CHAT-R) now requires the use of a standardized follow-up clinician-administered interview for most positive screens.[14] Of note, during validation studies, the required follow-up interview was conducted by telephone as part of a “two stage screener” process known as the M-CHAT-R/F,[14] which increased PPV from 0.14 to 0.48 in a sample of M-CHAT-R screen-positive children at 18- and 24-month well-child visits.[15] However, the follow-up interview has an extremely low rate of utilization in primary care settings.[16] Even with the follow-up interview, the PPV was lower (0.28) in younger compared to (0.61) older toddlers in one community sample[17] and similarly lower (0.36) in younger than in older (0.69) in another.[18]

Studies also suggest that, over subsequent years, for toddlers who screened both positive and negative on the initial M-CHAT, estimates for M-CHAT sensitivity and PPV are lower than reported in the M-CHAT validation study, with predictive indices lowest for the youngest toddlers. For example, a follow up study in Norway showed that a positive M-CHAT (without follow-up) at 18-months identified only 34% of children with an autism diagnosis by 9-years-old.[19] A recent report of screening with the M-CHAT at both the 18- and 24-month well-child visits, whose medical records were reviewed for autism diagnoses as outcomes at 4 to 8 years of age, reported a similar sensitivity of 0.35 for 18-month screening; lower than sensitivity at 24 months (0.49).[20] Guthrie, et al. (2019) found that for the 41.2 % of children whose score triggered the follow up portion of the M-CHAT-R/F, the PPV was also higher at the 24-month visit (0.25) than when the same child was screened at 18 months (0.18).[20]

An obstacle to estimates of prediction of autism diagnoses made years later is that some children may not have had any clinical manifestations at the earlier age and thus negative screens were ambiguous. In

addition, longitudinal studies show that an average of 32% of toddlers appear typical at 18 months and then regress between 18- and 24-months;[21] one reason the AAP recommends rescreening at 24 months.[6] Additionally, recent data from prospective studies of high-risk infant siblings reveal that as many as 45% of those diagnosed with autism at three years were not diagnosed at two years despite comprehensive assessments at that time.[22] Prevalence is also reported as 30% higher at ages 8–12 than at 3–7-years old.[23] Children identified later with autism tend to have milder symptoms and higher cognitive functioning.[24]

Screening strategies that are age-relevant and capture the natural emergence of autism are needed to address the screening challenges at 18-months. One solution may be ordinal polytomous (> 2 response options) scoring ("ordinal scoring") such as the Q-CHAT's "how much" item responses.[12] A 10 item version[26] of the original 25 item Quantitative Checklist for Autism in Toddlers (Q-CHAT)[12] is particularly suited for primary care because of its brevity and reported sensitivity of 0.91 and specificity 0.89 in a case comparison study.[26] However, since data from unselected primary care populations are lacking, we cannot consider this to be a true estimate of sensitivity. Also, while the Q-CHAT uses a five-point frequency response, its standard scoring instruction utilizes a pass/fail cut-off rather than ordinal scoring using the full range of all items. In this study, we compare the predictive utility of the M-CHAT-R, the M-CHAT-R/F, and the Q-CHAT-10 in a representative community sample that includes both toddlers who pass and fail initial screening measures. We also test the Q-CHAT-10 with its original pass/fail scoring and an experimental ordinal scoring to better understand the contribution of ordinal scoring to accurate 18-month screening.

Methods

Sampling and screening procedures:

Parents completed the M-CHAT-R before 18-month pediatric visits (16–20 mo.) via an online clinical process support system called CHADIS.[27, 28] M-CHAT-R positive screens prompted completion of M-CHAT-R/F by the PCP during the visit or later (if the PCP did not) by phone by a research assistant using online prompts in CHADIS via a previously validated method.[29] The follow up interview was completed for all M-CHAT-R positive parent reports except 23 which were not done (17) or begun but not completed (6). In addition, parents completed the Q-CHAT-10 and Ages & Stages Questionnaire (ASQ-3).[30] The order of presentation of Q-CHAT-10 to parents alternated with M-CHAT-R every month. A total of 11,878 parents of children age 16–20 months from pediatric offices using the same online system in Maryland, Massachusetts, and North Carolina completed the M-CHAT-R and Q-CHAT-10 screens. The version of the Q-CHAT-10 used that has been recommended for clinical usage by the authors included pictures illustrating each of the items. Of 787 children with any positive screen result (Q-CHAT-10 or M-CHAT-R even if follow-up was negative), it was possible to contact parents of 308 by phone or email for enrollment. Age and gender matched controls with both screens negative (n = 301) were then successfully contacted from the same practice or a practice in the same area with similar demographics. Children

were excluded if their parents reported that they were exposed to English at home less than 50% of the time (N = 43) or if they were not yet walking or scooting (N = 2) as required to complete the ADOS-2 Toddler module[31] for autism diagnostic testing. See Fig. 1 for a flowchart of recruitment. The study enrolled 472 children during the study period and lost 56 to follow up. The final sample includes 406 children with available data on key items and final autism case status determination. Of the final sample, 165 children were screen positives on M-CHAT-R (including 51 M-CHAT-R/F positives); 98 who were Q-CHAT-10 positive (of which 65 were also M-CHAT-R positive and 33 who were M-CHAT-R negative); and 238 were screen negatives. The initial screening component was deemed exempt by the IRB; parents of recruited children provided written consent.

Diagnostic procedures:

All children completed in-person diagnostic autism evaluations using the ADOS-2 Toddler (ADOS-T) module[31]and Mullen Scales of Early Learning (MSEL)[32]. Diagnostic testers were all experienced autism evaluators either certified as research reliable on the toddler module prior to the evaluations (3) or in one case attaining reliability through video review by a certified research reliable tester before finalized scoring. All diagnostic testers were blinded to screening results. Results of the ADOS-T algorithm and information from the MSEL informed a clinical judgment of whether a child met criteria for autism as required.[31, 32] Developmental disorder was defined as the typical criteria for early intervention services (score > 1½ SD below the mean on two or more subscales or > 2 SD on a single subscale of the MSEL)[34].

Results

Respondents were primary caregivers, almost all mothers who tended to be well educated and privately insured (see demographics Table 1). The range of ADOS scores shown in Table 2. As intended, the follow-up interview (M-CHAT-R/F) procedure increased PPV and specificity over the initial screen (M-CHAT-R) but with compromise in sensitivity and NPV (Table 3). Six of 10 children with M-CHAT-R scores > 7 were diagnosed with autism. A follow-up interview for these 10 children was falsely negative for 3 of the 6 with an autism diagnosis. Q-CHAT-10 with standard scoring, compared with M-CHAT-R, had higher PPV and specificity, but lower sensitivity. This is a similar pattern as the M-CHAT-R/F compared with M-CHAT-R. Q-CHAT-10 had both higher specificity and PPV than M-CHAT-R/F but did not improve sensitivity. However, unlike M-CHAT-R/F, the Q-CHAT-10 had a higher likelihood ratio for a positive (LR + = sensitivity/1-specificity) than M-CHAT-R.

To explore the potential of scoring Q-CHAT-10 differently than using the standard pass/fail decision for each item, we used ordinal scoring of the items. Ordinal scoring resulted in an area under the ROC curve (AUC) of 0.74 [0.70, 0.78] and a cut-point of > = 15 that optimized the balance of sensitivity (0.63) vs. specificity (0.79) via Youden's J (0.42). While the Q-CHAT-10 scored ordinally showed a lower sensitivity than M-CHAT-R, it was substantially greater than that of M-CHAT-R/F and Q-CHAT-10. The Q-CHAT-10's specificity (0.79) is generally accepted as adequate, and was much higher than that of M-CHAT-R, but lower than MCHAT-R/F. Overall, Q-CHAT-10-O had similar results to adding the follow up to M-CHAT-R in

having higher specificity and PPV than the M-CHAT-R with less compromise in sensitivity than M-CHAT-R/F or Q-CHAT-10, although Q-CHAT-10 scored ordinally loses the advantage of higher positive likelihood ratio shown by Q-CHAT-10.

Discussion

The current study results suggest that, when toddlers are screened in primary care at the 18-month visit, one cannot assume that children passing the M-CHAT-R/F are unlikely to have an autism diagnosis. In fact, most children identified as having autism in this sample passed the currently required MCHAT-R/F. Two community screening follow-up studies showed most children diagnosed with autism 2 ½ to 7 years later had passed the M-CHAT screen when 18-months-old.[19, 20] However, as noted earlier, several prospective studies[21,22,23] now point to the natural history of autism including evolution in subsequent years suggesting most cases of autism that are diagnosable years later would not have been evident by diagnostic testing of children presenting for 18-month well childcare. Unlike for typical clinical care when diagnostic testing is often delayed due to waiting lists or other issues, in this study we were often able to accommodate families through home testing and thereby completed all diagnostic testing prior to age two. The sensitivity estimates therefore should be more representative of children prior to the increases in prevalence expected by the natural history of progression of this condition. The sensitivity estimate in this study was higher for the M-CHAT-R than in either long-term outcome study cited.[19,20]

M-CHAT-R scoring without the follow-up was, as expected, more sensitive to autism diagnoses than the recommended two-stage procedure (M-CHAT-R/F) but at the cost of lower Positive Predictive Value (PPV), consistent with prior M-CHAT studies.[14] However, even with the follow-up interview, the PPV was relatively low and similar to reports from other studies showing PPV lower at 18 than at 24 months.[17,18] It should be noted that these M-CHAT-R/F estimates are limited by omission of the follow-up interview in some M-CHAT-R positive cases. Even with access to electronic support for completing the M-CHAT-F/U, the follow-up interview was inadvertently omitted 14% of the time in the current study and 59% of the time in the above noted follow-up study when a similar application of electronic decision support was available.[20] On the other hand, studies of practices without any decision supports reveal that the follow-up interview is only very rarely completed.[16] This study's results are consistent with the recommendation to omit the follow-up interview in cases with fail scores above 7 on the M-CHAT, although for a very small number, rather than risking false negatives by employing the interview for those few scoring in that range.

In a prior study, prediction of an autism diagnosis when pediatricians used the online decision support for conducting the follow-up interview was equivalent to when used by autism center personnel and also provided similar results at 24 months as in prior M-CHAT data.[18] This suggested that the differing results by age were not due to inaccurate follow-up interviews. There are a number of possible explanations for differing results across this age range that have potential implications for autism screening test development. In another prior study we found that when older toddlers (20 + months) were compared with younger (< 20 months), the younger toddlers had higher rates of item failure, with items that reflected

more advanced developmental milestones having the highest failure rates[38] suggesting that autism screening tests may need age-related scoring cut-offs. Prospective studies suggest that autism symptoms emerge gradually,[13] which may be reflected in a lower number of endorsed autism-specific items in younger children.

Q-CHAT-10 with standard scoring had the same advantages over M-CHAT-R as the M-CHAT-R/F in greater specificity and PPV, and both the specificity and PPV of Q-CHAT-10 was statistically superior to M-CHAT-R/F. Yet Q-CHAT-10 did not have greater sensitivity than M-CHAT-R/F. However, scoring using the full range of each Q-CHAT-10 item's ordinal scale resulted in performance characteristics that were more favorable than either the M-CHAT-R or two-stage M-CHAT-R/F because of improvement in specificity and PPV, compared to M-CHAT-R. While the Q-CHAT-10 with ordinal scoring resulted in some loss in positive likelihood ratio compared the Q-CHAT-10, sensitivity was improved compared to either the Q-CHAT-10 or M-CHAT-R/F. The Q-CHAT-10 scored ordinally has other advantages since it requires half the number of items of M-CHAT-R and does not require a follow-up clinician interview.

The limitations of the M-CHAT-R identified in this study are not necessarily an indictment of the potential of screening using parent report for autism in this age group. However, the attempts of the M-CHAT-R to simplify scoring with yes/no responses and of the Q-CHAT by using dichotomous cut-offs of frequency-based items with the same simple scoring for all ages may be unnecessary since computers can assist scoring. This study represents one of the largest groups of toddlers with diagnostic testing before age two from a community sample. However, as in prior autism screening studies, complete sampling of the community populations was limited because of the reluctance of previously unconcerned parents to submit their children to in-person testing as well difficulty contacting parents. While these sampling limitations preclude an absolute estimate of screening sensitivity, we present the typical public health indices to allow comparison to other studies. We also present likelihood indices and diagnostic odds ratios as metrics that can help compare tests independent of condition prevalence.

Conclusions

When children passing autism screening at the recommended 18-month pediatric visit are considered along with screen failures the recommended two-stage, the M-CHAT-R/F autism screen shows lower sensitivity to autism diagnoses than was previously estimated from studies that included both 18 and 24-month-olds and excluded children passing screens. Since the outcome comparison in this study involved timely completion of diagnostic testing it provides a better estimate of what is possible from using these screens at an age prior to known developmental shifts in the natural history[22, 23, 24, 25] of this condition. The lower sensitivity occurred with less balancing benefit of increase in PPV as in studies including both 18 and 24-month visits. However, when the ten-item Q-CHAT is scored using the full range of responses for each item, the resultant performance shows advantages over both the M-CHAT-R and M-CHAT-R/F, requires half the number of items, and eliminates the need for a follow-up interview. The Q-CHAT-10 is freely available from the authors and can be administered and scored without a computer. The Q-CHAT-10 can therefore be recommended for autism screening at 18 months. A caveat is that, even this solution,

similarly to M-CHAT-R, it falls below generally recommended psychometrics for screening[39] with the challenge of a relatively low PPV, meaning most children screening positive will not be confirmed by diagnostic testing as having autism. When making an autism referral for young toddlers, clinicians might also consider the possibility that the child may have a developmental problem other than autism and thereby benefit from an evaluation even if the result is not an autism diagnosis, as false positive children for autism have been shown to have a high rate of “developmental concerns”.[14] The overlap between autism and developmental delay outcomes in this sample is beyond the scope of this report. Another strategy to consider is one of tracking and rescreening after 20 months, when M-CHAT-R/F screen results appear to be more predictive[18]. Further research is needed for greater accuracy of screening at the 18-month visit including approaches our group is developing involving promising parent report computer-based strategies utilizing language items[18, 40] and more fully integrating autism screening with screening for developmental delay.[40] In a separate study we have shown that no screen at any age group identifies all or even most autism cases suggesting that autism screening should be conducted continuously at different ages during childhood, adolescence, and adulthood.[41]

Limitations

Limitations to generalizing from this study include the sample not being fully representative, since an entire population was not studied as many parents recruited declined participation, more so in the screen pass group than in the screen fail group, and parent participants were more highly educated than national rates. The version of the Q-CHAT-10 used had illustrations of items as suggested by the authors, with unknown impact compared to the original. Results from these studies of 16-20-month-olds cannot be extended to older ages.

Abbreviations

PCP: Primary Care Provider;

M-CHAT-R : Modified Checklist for Autism in Toddlers-Revised;

MCHAT-R/F: Modified Checklist for Autism in Toddlers Follow-up Interview;

ADOS-2: The Autism Diagnostic Observation Schedule, Second Edition;

MSEL: Mullen Scales of Early Learning;

ROC: Receiver Operating Characteristics;

Q-CHAT-10: Quantitative Checklist for Autism in Toddlers-10 item version;

Q-CHAT-10-O: Q-CHAT-10 scored through full range of ordinal (frequency) items;

AUC: Area Under the Curve;

ASD: Autism Spectrum Disorder;

CHADIS: Comprehensive Health and Decision Information System;

PPV: Positive Predictive Value;

AAP: American Academy of Pediatrics;

CDC: Center for Disease Control

Declarations

Availability of data and materials:

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate:

Ethics approval for this study was obtained from the Medstar Health Research Institute Office of Research (Reference number 2014-069).

Consent for Publication:

Not applicable.

Availability of data and materials:

Data is available from the corresponding author for reasonable requests

Competing Interests:

This study was conducted by the Center for Promotion of Child Development through Primary Care and its for-profit subsidiary, Total Child Health (TCH), Inc. CHADIS, the web-tool used in the study was developed by Dr. Sturner and his spouse, Dr. Howard. Dr. Sturner is Director of the Center and Dr. Howard is President of TCH. Both are members of the Board of Directors of Center and are paid employees or consultants to both entities. The other authors have indicated they have no financial relationships relevant to this article to disclose.

Funding:

All phases of this study were supported by NIMH grant no. R44MH085399. SBC received funding from the Wellcome Trust 214322\Z\18\Z. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Further to this SBC received funding from Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant

agreement No 777394. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. SBC also received funding from the Autism Research Trust, Autistica, the MRC and the NIHR Cambridge Biomedical Research Centre. The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR or Department of Health and Social Care.

Authors Contributions:

Raymond Sturner and Barbara Howard conceptualized, designed, and oversaw data collection during the study, drafted the initial manuscript, and approved the final manuscript as submitted.

Lydia Stewart and Kerry Bet managed data collection and data entry, coordinated with offices, reviewed the manuscript and approved the final manuscript as submitted. Shana Attar performed autism and developmental evaluations, entered autism diagnostic conclusions, helped prepare the data for analysis, reviewed the manuscript and approved the final manuscript as submitted. Paul Bergmann helped design the project, carried out the analyses, reviewed and revised the manuscript, and approved the final manuscript as submitted. Simon Baron-Cohen and Carrie Allison conceptualized and assisted in design, reviewed and revised the manuscript, and approved the final manuscript as submitted. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Acknowledgments:

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health Award Number R44MH085399. The content is solely the responsibility of the authors and does not necessarily represent the views of the National Institutes of Health. The participating primary care pediatricians and their office staffs are acknowledged for their cooperation during the project and continuing implementation. We acknowledge testers including Ruth Williams, MA, Katherine Campe, MS, Jaime Allison, MEd, Margaret DeRamus, MS, and Tiffany Garner, PsyD with Trellis Services. We acknowledge Linda Lee who created illustrations for Q-CHAT items.

References

1. Maenner, Matthew J., Kelly A. Shaw, Jon Baio, EdS1, Anita Washington, Mary Patrick, Monica DiRienzo, et al. "Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016." *MMWR. Surveillance Summaries* 69, no. 4 (March 27, 2020): 1–12. https://doi.org/10.15585/mmwr.ss6904a1
2. *National Research Council (2001). Educating Children with Autism. Washington, DC: National Academies Press*.

3. Howlin P, Magiati I, Charman T. Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities*. 2009;114 (1): 23–41.
4. Dawson, G., Rogers, S., Munson, J., Milani Smith, Jamie Winter, J., Greenson, J., Donaldson, A., Varley, J., Randomized, Controlled Trial of an Intervention for Toddlers with Autism: The Early Start Denver Model. Pediatrics. 2010 Jan; 125(1).
5. Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-Vanderweele J. A systematic review of early intensive intervention for autism spectrum disorders Pediatrics. 2011 May;127(5):e1303-11. doi: 10.1542/peds.2011-0426. Epub 2011 Apr 4.
6. American Academy of Pediatrics (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. Pediatrics, 118(1), 405–420.
7. https://www.cdc.gov/ncbddd/autism/screening.html accessed 1/17/2021.
8. Siu, A. L., & The US Preventive Services Task Force (USPSTF). (2016). Screening for autism spectrum disorder in young children. US preventive services task force recommendation statement. JAMA: The Journal of the American Medical Association, 315(7), 691–696. doi:10.1001/jama.2016.0018.
9. Baron-Cohen S, Allen J, Giliberg C. Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry* 1992; 161, 839–843.
10. Baron–Cohen S, Cox A, Baird G, et al. Psychological markers in the detection of autism in infancy in a large population. *Br J Psychiatry*. 1996; 168 (2): 158–63.
11. Robins D L, Fein D, Barton, M L, Green, J. A. The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. Journal of Autism and Developmental Disorders 2001; 31(2), 131–144. DOI:10.1177/1362361308094502.
12. Allison C, Baron–Cohen S, Wheelwright S, et al. The Q–CHAT (Quantitative CHecklist for Autism in Toddlers): a normally distributed quantitative measure of autistic traits at 18–24 months of age: preliminary report. J Autism Dev Disord. 2008; 38(8): 1414–25.
13. Ozonoff S, Heung K, Byrd R, Hansen R, and Hertz-Picciotto I. The onset of autism: Patterns of symptom emergence in the first years of life. *Autism Research* 2008; 1(6), 320–328. doi:10.1002/aur.53.
14. Robins, D. L., Casagrande, K., Barton, M., Chen, C-MA., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). Pediatrics, 133(1), 37–45; p43; doi:10.1542/peds.2013-1813.
15. Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., & Barton, M. (2008). The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. Journal of Autism and Developmental Disorders, 38(5), 827–839. doi:10.1007/s10803-007-0450-9.
16. Swanson A R Warren Z E, Stone W L, Vehorn A C, Dohrmann E, and Humberd Q. The diagnosis of autism in community pediatric settings: Does advanced training facilitate practice change? Autism,

2014; 18(5), 555–561. https://doi.org/10.1177/1362361313481507.

17. Pandey J, Verbalis A, Robins D L, Boorstein H, Klin A, Babitz T, and Fein D. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. Autism: The International Journal of Research and Practice 2008: 12(5), 513–535. doi:10.1177/1362361308094503.
18. Sturner, R. Howard, B. J., Bergmann, P., Morrel, T.. Landa, R., Walton, K., Marks, D., Accurate Autism Screening at the 18-Month Well-Child Visit Requires Different Strategies Than at 24 Months J. Autism and Related Disorders 2017; p 1- 15, 10.1007/s10803-017-3231-0.
19. Stenberg N, Bresnahan M, Gunnes N, Hirtz D, Hornig M, Lie K K, and Schjlberg, S. Identifying children with autism spectrum disorder at 18 months in a general population sample. Paediatric and Perinatal Epidemiology 2014; 28(3), 255–262. doi:10.1111/ppe.12114.
20. Guthrie W, Wallis K, Bennett A, et al. Accuracy of Autism Screening in a Large Pediatric Network. Pediatrics. 2019;144(4).
21. Barger, B. D., Campbell, J. M., & McDonough, J. D. (2013). Prevalence and onset of regression within autism spectrum disorders: A meta-analytic review. Journal of Autism and Developmental Disorders, 43(4), 817–828. doi:10.1007/s10803-012-1621-x.
22. 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
23. Zablotsky B, Black LI, Blumberg SJ. Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014–2016. NCHS Data Brief , No. 291. Hyattsville, MD: CDC, National Center for Health Statistics; 2017;291:1–8. PubMedexternal icon
24. Zwaigenbaum L, Duku E, Fombonne E, et al. Developmental functioning and symptom severity influence age of diagnosis in Canadian preschool children with autism. *Paediatr Child Health.* 2019;**24***(1):*e57–e65
25. Allison C, Auyeung B, & Baron-Cohen, S. Toward brief “red flags” for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. J American Academy of Child & Adolescent Psychiatry, 2012; 51(2), 202-212.
26. CHADIS (2019). CHADIS (Comprehensive Health and Decision Information System). http://www.chadis.com/site/. Accessed 8 November 2020.
27. Howard BJ and Sturner R. Use of an Online Clinical Process Support as an Aid to Identification and Management of Developmental and Mental Health Problems. Curr Dev Disord. Rep. 2017; 4(4):108–117.
28. Sturner, R., Howard, B., Bergmann, P., Morrel, T., Andon, A., Marks, D., Rao, P., & Landa, R. (2016). Autism screening with online decision support by primary care pediatricians aided by M-CHAT/F. Pediatrics, 138(3), e20153036. doi:10.1542/peds.2015-3036
29. Squires,J. Bricker, D. Potter,L., Revision of a Parent-Completed Developmental Screening Tool: Ages and Stages Questionnaires, Journal of Pediatric Psychology, Volume 22, Issue 3, June 1997, Pages 313–328, https://doi.org/10.1093/jpepsy/22.3.313.

30. Lyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., Richler, J., Risi, S., & Lord, C. (2009). The Autism Diagnostic Observation Schedule—Toddler Module: A new module of a standardized diagnostic measure for autism spectrum disorders. Journal of Autism and Developmental Disorders, 39(9), 1305–1320. .
31. *Mullen, E. M. (1995). The Mullen Scales of Early Learning: AGS Edition. Circle Pines, MN: American Guidance Service, Inc*.
32. Filipek PA, Accardo PJ, Baranek GT, Cook EH, Jr, Dawson G, Gordon B, et al. The screening and diagnosis of autistic spectrum disorders. Journal of Autism and Developmental Disorders. 1999;29(6):439–484.
33. Ringwalt, S. (Comp.). (2015, March). Summary table of states' and territories' definitions of/criteria for IDEA Part C eligibility. Retrieved from http://ectacenter.org/~pdfs/topics/earlyid/partc_elig_table.pdf.
34. Sturner R, Howard B J, Bergmann P Stewart L, Afarian T. Comparison of Autism Screening in Younger and Older Toddlers. *J. Autism and Related Disorders* 2017; P 1- 9. 10.1007/s10803-017-3230-1
35. *American Academy of Pediatrics, Addressing Mental Health Concerns in Primary Care: A Clinician's Toolkit. 2010, Revised January 2012*.
36. Sturner, R., Bergman, P., Howard, B., Bet, K., Stewart, L. Attar, S., A Toddler Adaptive Development and Autism Screen (TADAS) efficiently improves Identification at the 18-month Visit presented at the Society for Developmental and Behavioral Pediatrics, Sept., 2019, Washington, DC.
37. Allison, C, Matthews, F.E., Ruta, L., Pasco, G., Soufer, R., Brayne, C., Charman, T., & Baron–Cohen, S. The Quantitative Checklist for Autism in Toddlers (Q–CHAT). A population screening study with follow-up: the case for multiple time–point screening for autism. Submitted for publication, available on request, 2021.

Tables

Table 1.

Child Demographics

Age	**Mean Months (S.D.)**
At screening	18.0 months (0.53)
At diagnostic testing overall	20.5 months (1.9)
At diagnostic testing: screen positives	20.5 months (1.9)
At diagnostic testing: screen negatives	20.3 months (1.8)
Sex	**N (%)**
Male	294 (72.1)
Female	114 (27.9)
Race	**N (%)**
White	281 (68.9)
African American	39 (9.6)
Asian	19 (4.7)
Multi-racial	63 (15.4)
Unknown	6 (1.5)
Hispanic Ethnicity[2]	**N (%)**
Yes	27 (6.6)
No	375 (91.9)
Not Sure	1 (0.2)
Exposure to English Language[3]	**N (%)**
50% of the time	13 (3.2)
75% of the time	45 (11.0)
100% of the time	300 (73.5)
Missing	50 (12.2)
Primary Language in Home	**N (%)**
English	391 (95.8)
Other	12 (2.9)
Missing	5 (1.2)

Health Insurance	**N (%)**
Private Payer	325 (79.7)
Public (Medicaid, Medicare, SCHIP)	42 (10.3)
Other	41 (10.0)
Missing	5 (1.2)
Household Income	**N (%)**
< $25,000	2 (0.5)
$25,000 - $49,999	16 (3.9)
$50,000 - $99,999	58 (14.2)
$100,000 - $149,999	73 (17.9)
$150,000 - $199,999	63 (15.4)
$200,000 - $249,999	29 (7.1)
>= $250,000	32 (7.8)
Missing	135 (33.1)

Table 2.

Respondent Characteristics

Relationship to Child	**N (%)**
Mother	3742 (91.2)
Father	28 (6.9)
Other	3 (0.7)
Education Level[4]	**N (%)**
High School Diploma/GED or Less	14 (3.4)
Some College	61 (15.0)
Bachelor Degree	1487 (36.0)
Advanced Degree	181 (44.4)
Employment	**N (%)**
Professional/Major Business	236 (57.8)
Medium Business/Minor Professional	26 (6.4)
Technical/Craftsperson	16 (3.9)
Clerical/Sales	2 (0.5)
Semiskilled	1 (0.2)
Service/Unskilled	2 (0.5)
At Home/Manage Family	113 (27.7)
Unemployed	7 (1.7)

Table 3.

Diagnostic Score Results

ADOS Scores (Overall)	**n**	**Mean**	**SD**	**Min**	**Max**
Social Affect Total	406	5.0	4.9	0	20
Restricted/Repetitive Behavior Total	405	1.7	1.9	0	8
Overall Total (SA + RRB)	405	6.6	6.1	0	24
Range of Concerns	401	1.4	0.8	1	3
ADOS Scores (ASD Negative)	**n**	**Mean**	**SD**	**Min**	**Max**
Social Affect Total	344	3.5	3.2	0	14
Restricted/Repetitive Behavior Total	343	1.2	1.5	0	6
Overall Total (SA + RRB)	343	4.7	4.0	0	17
ADOS Scores (ASD Positive)	**n**	**Mean**	**SD**	**Min**	**Max**
Social Affect Total	62	13.4	3.6	6	20
Restricted/Repetitive Behavior Total	62	4.1	1.9	1	8
Overall Total (SA + RRB)	62	17.5	3.7	9	24
Mullen T-Scores	**n**	**Mean**	**SD**	**Min**	**Max**
Gross Motor	405	50.8	9.6	20	80
Visual Reception	404	54.5	12.0	20	80
Fine Motor	405	50.6	9.2	20	80
Receptive Language	404	50.3	9.2	20	80
Expressive Language	405	45.9	13.7	20	80

Table 4.

ASD Screening Performance – Sensitivity, Specificity, PPV & NPV

Row	Screen	Sensitivity	Specificity	PPV	NPV	LR+	LR-
1	M-CHAT-R n=406	0.73 [0.60, 0.82][1] 2D | 3D | 4D[2]	0.65 [0.60, 0.70] 2D | 3D | 4D	0.27 [0.21, 0.34] 2D | 3D | 4D	0.93 [0.89, 0.96] 2I | 3D | 4E	2.1 [1.7, 2.6]	0.4 [0.3, 0.6]
2	M-CHAT-R/F n=368	0.34 [0.23, 0.47] 1D | 3I | 4D	0.90 [0.86, 0.92] 1D | 3D | 4D	0.35 [0.24, 0.49] 1D | 3D | 4I	0.89 [0.85, 0.92] 1I | 3E | 4I	3.2 [2.0, 5.3]	0.7 [0.5, 0.9]
3	Q-CHAT-10 n=406	0.31 [0.21, 0.43] 1D | 2I | 4D	0.96 [0.93, 0.98] 1D | 2D 12D	0.58 [0.41, 0.73] 1D | 2D | 4D	0.88 [0.85, 0.91] 1D | 2E | 4I	7.5 [4.0, 14.2]	0.7 [0.6, 0.9]
4	Q-CHAT-10-O n=406	0.63 [0.50, 0.74] 1D | 2D | 3D	0.79 [0.74, 0.83] 1D | 2D | 3D	0.35 [0.27, 0.44] 1D | 2I | 3D	0.92 [0.89, 0.95] 1E | 2I | 3I	3.0 [2.3, 4.0]	0.5 [0.3, 0.6]

[1] Cells contain value, [95% CI] and TOST results.

[2] Number refers to the row that was compared to the current row; letter refers to the result of the TOST comparison: D – Statistically significant difference, E – Statistically significant equivalence, I – Indeterminate finding.

Figures

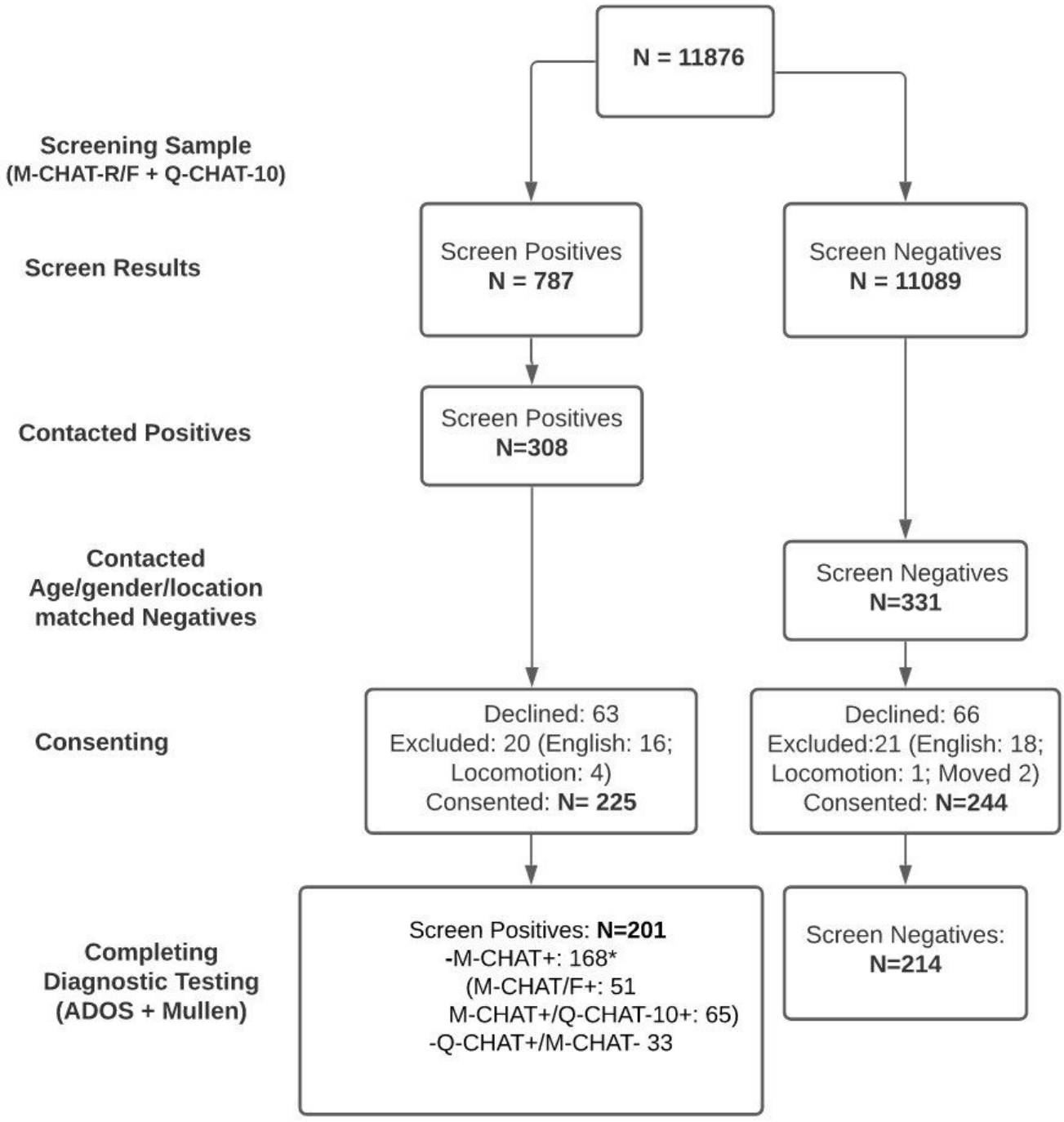

Figure 1

Recruitment Flow. * M-CHAT-R F/U interview completed on all but 23 with incomplete or omitted

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementQCHAT10items.docx