ORIGINAL PAPER

CrossMark

A Comparison Between Two Screening Approaches for ASD Among Toddlers in Israel

Orly Kerub[1,2] · Eric J. Haas[2] · Gal Meiri[3] · Nadav Davidovitch[1] · Idan Menashe[4,5]

Abstract

Systematic screening of autism spectrum disorder (ASD) can improve early diagnosis of ASD. We compared the efficacy of two ASD screening methods, the Global Developmental Screening (GDS), and the Modified Checklist for Autism in Toddlers-Revised, with Follow-Up (M-CHAT/F) in 1591 toddlers of ages 18–36 months from 35 government-funded clinics in south Israel. The M-CHAT/F performed better than the GDS in detecting toddlers with ASD (sensitivity: 70.0% vs. 50.0%, and specificity: 98.2% vs. 96.6% respectively). Both methods had an equivalent performance in detecting other forms of developmental delays (sensitivity = 63%; and specificity ~ 98%). In addition, remarkable inter-nurse variation was observed in the GDS referral decisions. Thus, employment of the M-CHAT/F in the Israeli health system may improve early detection of ASD among toddlers.

Keywords Autism spectrum disorder · Screening · Toddlers · M-CHAT

Introduction

Autism spectrum disorders (ASD) are a range of neurodevelopmental disorders characterized by deficits in social communication and the existence of restricted and repetitive patterns of behavior, interest or activity (American Psychiatric Association 2013; Levy et al. 2009). At present, there is no cure to ASD, however, significant improvement in ASD symptoms can be achieved through a range of intensive behavioral interventions (Eaves and Ho 1996; Estes et al.

✉ Idan Menashe
  idanmen@bgu.ac.il

1 Department of Health Systems Management, Ben-Gurion University of the Negev, Beer Sheva, Israel

2 Ministry of Health, Jerusalem, Israel

3 Pre-School Psychiatry Unit, Soroka University Medical Center, Beer Sheva, Israel

4 Public Health Department, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel

5 Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer Sheva, Israel

2015). Substantial evidence suggest that earlier application of such interventions achieve better outcomes (Dawson et al. 2010; Eaves and Ho 1996; Zwaigenbaum et al. 2015). Thus, early diagnosis of ASD on a population basis is warranted to reduce the burden associated with the disorder among affected children and their families.

It is accepted that ASD can be reliably diagnosed before the second birthday (Dawson et al. 2010; Worley et al. 2011; Zwaigenbaum et al. 2015). Thus, the American Academy of Pediatrics recommends that all children should undergo standardized screening testing to detect autism between the ages of 18–24 months (Johnson and Myers 2007). There are multiple screening tools for ASD; some are based on parental reports, while others require direct observation and engagement by clinicians. Parent-report tools often have the advantage of being brief, inexpensive, and practical in the office setting.

One of the most commonly used ASD-screening tools worldwide is the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al. 2001). This two-stage screening tool uses 23 yes/no items that can be completed by parents of toddlers between 16 and 30 months of age. Both the M-CHAT and its recent revision, the Modified Checklist for Autism in Toddlers, Revised, with Follow-Up (M-CHAT-R/F) (Robins et al. 2014), demonstrated high sensitivity and specificity in detecting toddlers with

ASD in the US (Kleinman et al. 2008; Robins 2008), as well as in several other countries (Mohamed et al. 2016; Nygren et al. 2012; Samadi and McConkey 2015). In addition, it performed better in detecting children with autism than various other more general screening tools (Pinto-Martin et al. 2008; Wiggins et al. 2014). Yet, there is still insufficient evidence regarding the benefits and harms of utilization of the M-CHAT as a universal screening tool for ASD (Robins et al. 2016; Siu et al. 2016). In particular, concerns have been raised about the fate of children with ASD who pass the M-CHAT at 18 months (false negative) (Øien et al. 2018; Stenberg et al. 2014). Thus, further evaluation of the M-CHAT and other ASD screening tools in diverse populations is warranted.

Toddlers in Israel undergo global developmental screening (GDS) by nurses at government-funded maternal child health centers (MCHCs) (Israel Ministry of Health 2016). The GDS is performed at each visit in the MCHC (i.e. at ages 3, 6, 9, 12, 18, 24, 36, 48, and 60 months) and includes 4–9 items that assess different age appropriate developmental skills in the areas of gross motor, fine motor, language and communication, and emotional-social domain which are examined through interview of the parent or direct observation of the child. If the child shows risk on one or more of the GDS items, he/she will be invited for a follow-up examination or will be referred to a child developmental clinic for further evaluation. Since the GDS has been designed to detect general developmental problems, we hypothesized it might have a poorer performance in detecting toddlers with ASD compared to ASD-specific screening tools such as the M-CHAT. To test this hypothesis we compared the screening performance of the M-CHAT and GDS in a sample of 1591 toddlers at ages of 18–36 months at 35 randomly selected MCHCs in the Negev.

Methods

Population

There are approximately 16,000 live births in the southern part of Israel, the Negev, annually of which ~50% are of Bedouin Arabs, ~50% are Jewish, and less than 1% are immigrants of other ethnicities (The Central Bureau of Statistics of Israel 2016). The development of these newborns is routinely monitored from birth until age 6 years in 47 government-funded MCHCs distributed across the region. The attendance at these MCHCs is extremely high with 95%–99% of all newborns attending MCHC for developmental assessments, vaccinations, and other services during infancy (Bin Nun et al. 2010).

ASD Screening

Screening for ASD or other DD was conducted between March 2015 and December 2016 in 35 out of the 47 MCHCs from the Negev where the nurses agreed to administer the M-CHAT questionnaire in addition to their routine training in administration of the GDS. We used the original version of the M-CHAT (the one with the 23 yes/no items; Robins et al. 2001), because at the time of the study there was no Hebrew translation of the more recent M-CHAT-R/F version. Nurses who participated in the study had a 1-day workshop about ASD screening and diagnosis as well as specific training in M-CHAT administration. All these nurses had previous experiences in the administration of the GDS test.

Subjects were toddlers at ages of 18–36 months who came for regular developmental assessments in these clinics. Both M-CHAT and GDS screening were conducted by the nurses at the clinics because many parents, especially in the Bedouin population, have poor reading and writing capacities. Items that could not be evaluated at the clinic were scored according to the parents experience with their child. Toddlers who screened positive at the M-CHAT questionnaire were followed-up by an ASD specialist nurse using the M-CHAT follow-up questionnaire (M-CHAT/F). Toddlers who screened positive at either the M-CHAT/F or the GDS, were referred to further developmental evaluation at the Soroka University Medical Center (SUMC). Diagnoses of ASD or other developmental problems were determined by a child psychiatrist or child neurologist according to the DSM-5 criteria (American Psychiatric Association 2013) as described before (Meiri et al. 2017).

Screening Evaluation

False negatives were determined 10 months after study completion by reviewing the medical records of all toddlers who were screen negative at both the M-CHAT/F and the GDS, and identifying those with reported ASD or DD diagnosis. We used standard screening validity measures [i.e. sensitivity, specificity, positive & negative predicted values (PPV & NPV)] to compare the ability of both M-CHAT/F and GDS to detect toddlers with ASD or other developmental problems.

Assessment of the GDS Tool

Currently, nurses at the MCHCs refer toddler for further developmental evaluation based on the GDS results and their own impression of the child development. To have an objective assessment of GDS tool we collected the full history of the GDS tests performed at 9, 12, 18, 24 and

36 months of the 50 toddlers who failed at the M-CHAT/F test, and another 150 randomly selected toddlers with normal M-CHAT/F scores. Then, we asked eleven experienced MCHC nurses (2–30 years working at the MCHC) to review these data and give a referral decision for each of these toddlers. Importantly, the nurses had no other information about these toddlers except of their sex and age. We used Fleiss-Kappa statistics to assess inter-rater agreement between these eleven nurses. The developmental skills that are evaluated by the GDS at 9, 12, 18, 24 and 36 months are described in the supplementary materials.

Results

M-CHAT/F—GDS Comparison

A total of 1,591 toddlers, between 18 and 36 months of age (Mean = 21.30 ± 3.45 months) were screened by both GDS and M-CHAT/F at the selected MCHCs in this study (Table 1). There was a small but significant overrepresentation of Bedouins and males in this sample (60% and 54% respectively; *p < 0.001*).

Figure 1 shows the screening results of this study. Overall, 24 toddlers were detected by both the M-CHAT/F and the GDS as needing further developmental evaluation (Kappa = 42%, *p < 0.001*). Another 26 and 35 toddlers were detected by only the M-CHAT/F or the GDS respectively. Of these 84 toddlers, seven received a diagnosis of ASD, 30 received a diagnosis of other forms of developmental delays (DD) (e.g. Motor, or Speech and language delay etc.), 11 did not complete their diagnostic process at the time of this analysis, and four were lost to follow up. Notably, the M-CHAT/F was more sensitive than the GDS in detecting toddlers with ASD (sensitivity of 70.0% vs. 50.0% respectively), and slightly more specific (specificity of 98.2% vs. 96.6% respectively) (Table 2). These differences translated into a 2.3-fold difference between the PPV of these two screening tools (20.0% vs. 8.6% for the

M-CHAT/F and GDS respectively). Notably, all the five toddlers that were screened positive in the GDS, were positive also in the M-CHAT/F, thus combining the results of the GDS + M-CHAT/F did not produce better screening results than the M-CHAT/F alone.

We also assessed the efficacy of these two screening tools to detect toddlers with other forms of DD (Table 3). Both M-CHAT/F and GDS had similar specificity and sensitivity in detecting toddlers with DD (Sensitivity = 63.3%, Specificity = 97.5–99%). However, the combination of these two tools, resulted in a remarkable increase in the sensitivity of the screening (Sensitivity = 93.7%) without changing its specificity (Specificity = 98%).

GDS Results Evaluation

Results of the GDS evaluation by a sample of eleven MCHC nurses are displayed in Fig. 2. There was a moderate inter-rater agreement between nurses regarding the referral decision of toddlers with ASD or other DD (Fleiss Kappa = 0.435), with only four of these toddlers (13%) detected by all eleven nurses (Fig. 2a). There was a better consensus among the nurses regarding the 170 typically developing toddlers in this sample (Fleiss Kappa = 0.721) (Fig. 2b). Interestingly, there were four typically developing toddlers whom all nurses thought they needed further developmental evaluation, and another five toddlers who were referred by 10/11 of the nurses. Further examination of the factors associated with nurse's referral decisions revealed a significant association with the number of failed GDS items (Spearman r = 0.65; *p < 0.001*), with ≥ 5 failed items serving as the referral cutoff for the majority (≥ 6) of the nurses (Fig. 2c). Notably, seven toddlers who did not fail at any of the GDS items were still referred for further evaluation by a few nurses. In addition, there was one toddler who failed in only two GDS items and was referred for further developmental evaluation by seven nurses. In contrast, there was one toddler who failed in five GDS items and was referred for further developmental evaluation by only one nurse. Using the ≥ 5 failed GDS items as a referral cutoff obtained 63% sensitivity

Table 1 Study population

	ASD (N = 10)	DD (N = 30)	ND (N = 1551)	*p* Values
Males, N (%)	7 (70%)	20 (66.7%)	863 (53.9%)	0.314[a]; 0.216[b]
Bedouins, N (%)	3 (30%)	16 (53.3%)	927 (59.8%)	0.057[a]; 0.635[b]
Age (Months), Mean ± SD	23.18 ± 4.38	23.01 ± 3.44	21.28 ± 3.45	0.082[c]; **0.007**[d]

Statistically significant difference (*p* < 0.05) is highlighted in bold

ASD autism spectrum disorder, *DD* (Other) developmental delays, *ND* normal development

[a]Chi-square comparing ASD vs. ND

[b]Chi-square comparing DD vs. ND

[c]*t* test comparing ASD vs. ND

[d]*t* test comparing DD vs. ND

1,591 toddlers screened with both GDS and M-CHAT/F

Screen Positive

Screen Negative

24 Positive by both M-CHAT/F and GDS	26 Positive by M-CHAT/F only	35 Positive by GDS only	1506 Negative by both M-CHAT/F and GDS
5 ASD 10 DD 1 LF 8 ND	2 ASD 9 DD 14 LF 1 ND	0 ASD 9 DD 0 LF 26 ND	3 ASD 2 DD 0 LF 1501 ND

Fig. 1 A flow chart of the screening results of this study. Participants were assigned diagnoses of autism spectrum disorder (ASD), other types of developmental delay (DD), or normal development (ND). Toddlers who did not complete their diagnostic evaluation by the time of this study were considered as lost to follow-up (LF)

Table 2 Screening efficacy of the M-CHAT/F and GDS in detecting toddlers with ASD

	M-CHAT/F[a]		GDS[b]		M-CHAT/F + GDS[c]		Total
	+	−	+	−	++/−+/+−	−−	
ASD (+)	7	3	5	5	7	3	10
ASD (−)	28	1538	53	1513	51	1515	1566
Total	35	1541	58	1518	58	1518	1576
Sensitivity	70.0%		50.0%		70.00%		Prevalence = 0.63%
Specificity	98.2%		96.6%		96.76%		
PPV	20.0%		8.6%		12.07%		
NPV	99.8%		99.7%		99.80%		

[a]Screening results of the M-CHAT/F: (+) screen positive; (−) screen negative

[b]Screening results of the M-CHAT/F: (+) screen positive; (−) screen negative

[c]Screening results of both the M-CHAT/F and GDS: (++/+−/−+) screen positive in either of the tests; (−−) screen negative in both tests

and 92% specificity in identifying toddlers with ASD or other DD. Nurse's referral decisions were not associated with the age, gender or race of the toddlers.

Discussion

The purpose of this study was to compare the screening efficacy of GDS vs. M-CHAT/F in detecting ASD or other DD in a sample of 1591 toddlers from southern Israel. The prevalence of toddlers with ASD in our sample (0.68%) was slightly higher than the prevalence of ASD in Israel reported in two other recent studies (Davidovitch et al. 2013; Raz et al. 2014). This difference in ASD prevalence can be mostly attributed to the active screening for ASD with both M-CHAT/F and GDS that was used in this study. Such active screening for developmental problems is particularly relevant to the Bedouin population that traditionally has been characterized with low prevalence of ASD (Mahajnah et al. 2015; Meiri et al. 2017). Thus,

Table 3 Screening efficacy of the M-CHAT/F and GDS in detecting toddlers with non-ASD developmental delay (DD)

	M-CHAT/F [a]		GDS[b]		M-CHAT/F + GDS[c]		Total
	+	−	+	−	++/−+/+−	−−	
DD (+)	19	11	19	11	28	2	30
DD (−)	16	1530	39	1507	30	1516	1546
Total	35	1541	58	1518	58	1518	1576
Sensitivity	63.33%		63.33%		93.67%		Prevalence = 1.90%
Specificity	98.97%		97.48%		98.05%		
PPV	54.29%		32.76%		48.28%		
NPV	99.29%		98.28%		99.87%		

[a]Screening results of the M-CHAT/F: (+) screen positive; (−) screen negative

[b]Screening results of the M-CHAT/F: (+) screen positive; (−) screen negative

[c]Screening results of both the M-CHAT/F and GDS: (++/+−/−+) screen positive in either of the tests; (−−) screen negative in both tests

incorporating routine, standardized screening for ASD in this population will help reduce existing ethnic disparities in ASD diagnosis and possibly also in ASD treatments as implied by other studies (Herlihy et al. 2014).

Notably, our ability to review the medical records of all 1,591 toddlers in this study and consequently assess the false negative rates of both M-CHAT/F and GDS is a major strength of this study. The relatively high rates (up-to 30%) of false negatives in this study are consistent with the reports of Stenberg et al. (2014), and Øien et al. (2018), and likely due to children with milder symptoms of ASD that are hardly noticeable at ages of 18–36 months. While somewhat better performance could be achieved if the more recent version of the M-CHAT, the M-CHAT-R/F, were used (Robins et al. 2014), additional false negatives are also expected at later ages. Nevertheless, the primary goal of screening tools for ASD is to detect children with the disorder as early as possible. In this regard, the better performance of the M-CHAT/F over the GDS in detecting ASD in these ages suggest that its implementation at the MCHC clinics could help reducing the age of ASD diagnosis in this population.

The better performance of the M-CHAT/F in detecting ASD cases compared to the GDS is consistent with results of other studies demonstrating the value of the M-CHAT/F in detecting toddlers with ASD (Chlebowski et al. 2013; Toh et al. 2017). On the other hand, the GDS is applicable to a wider age range, and therefore more flexible. Thus, toddlers tested by the GDS might be given several opportunities to achieve certain milestones before failure is determined. Given the differences between the GDS and the M-CHAT/F, a combination of these two tests in toddler's developmental screening might be optimal (Kamio et al. 2014; Nygren et al. 2012). Our results suggest that screening toddlers with both M-CHAT/F and GDS will have a similar efficacy as the M-CHAT/F alone in detecting cases with ASD and will improve the efficacy in detecting toddlers with other types of DD compared to application of either the M-CHAT/F or the GDS screening approaches alone.

The GDS test is designed to identify various developmental issues. However, the decision whether to refer toddlers to further developmental assessment currently relies on the examiner's judgement, and therefore variable (Mendonca et al. 2016). The moderate inter-rater agreement between the eleven nurses in our study indicates that nurses often disagree regarding a referral of the same toddler. One possible source for this variability is that sometimes failure or achievement of specific milestones are reported by parents without the ability to test them in the clinic (e.g. toddlers who refuse to cooperate). This may result in examples such as the toddler in our sample who failed in five items and still was not referred for further developmental evaluation by the majority of the nurses. In addition, when conducting the GDS test, some nurses record the parents' concerns regarding the development of their child. This may be the reason for the few toddlers who apparently did not fail at any of the items but still were referred for further evaluation by some of the nurses. Another factor that may affect the referral decision of the nurses is the cultural or ethnic suitability of developmental milestones as suggested by Magalhães et al. (Magalhães et al. 2015). Although we didn't observe significant ethnic differences in the nurses' referral decision in our study, such ethnic disparities should be considered in the development of future developmental tests. These results emphasize the need for standardized referral guidelines for the GDS along with complementary training of the nurses at the MCHCs.

We noticed that an important factor in the referral decision of the nurses in this study was the number of failed GDS items with failure in five items or more serving as a referral cutoff for the majority of the nurses. Interestingly, a similar cutoff was used for both toddlers with developmental

Fig. 2 Referral decision of 11 nurses based on blinded GDS results. Gray and White boxes indicate referral and no referral decisions respectively. **a** Referral decision for the 30 toddlers with ASD or other developmental problems. **b** Referral decision for the 170 toddlers with normal development. **c** The number of referral nurses are plotted vs. the number of failed GDS items. Gray circles are for typically developed toddlers, and empty black circles are for toddlers diagnosed with ASD/DD. Circle sizes are proportional to the number of toddler in each group. Vertical dashed line indicates a putative referral cutoff of five failed GDS items, and horizontal dashed line indicates the median of referring nurses
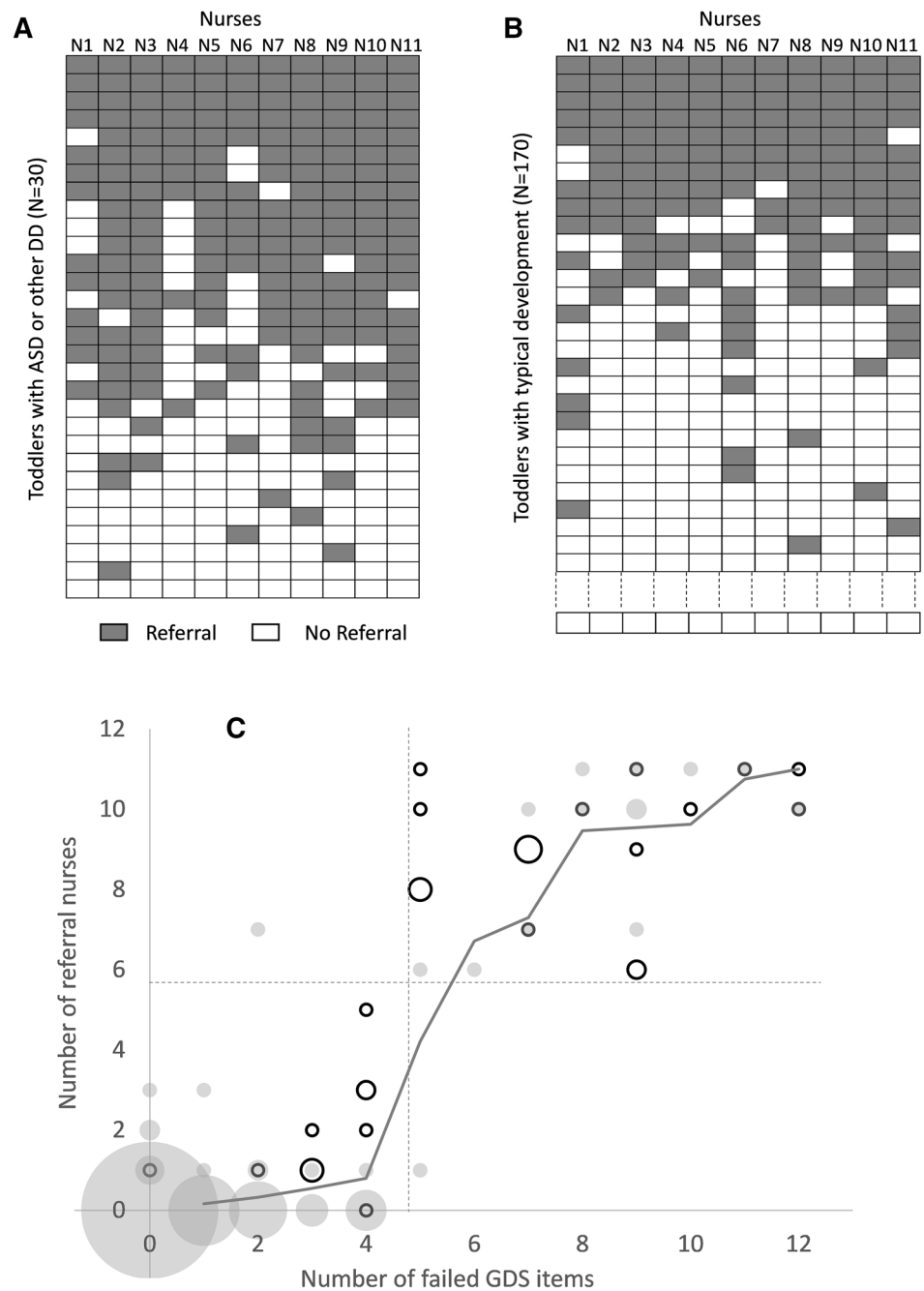
problems and toddlers with normal development without significant differences between the two groups. The small size of our sample didn't allow us to examine if any of the GDS items or specific combinations of them are better predictors of ASD or other developmental problems. Results of such analysis will help to develop standardized referral guidelines for the GDS, which will reduce the remarkable variability in nurses referral as observed in this study.

Our study has several noticeable limitations. (1) Despite the large initial sample size of 1591 toddlers, the low prevalence of ASD and DD in the population resulted in a

small number of cases that limit our ability to have a stable assessment of the screening sensitivity. (2) 15 of the 50 toddlers (30%) who were positive at the M-CHAT/F did not complete their diagnosis. This relatively large fraction of lost to follow-up may result in under estimation of the sensitivity of this screening tool. We currently investigates the factors that are associated with referral compliance and time-to-diagnosis of toddlers that have been referred to SUMC with suspicion of ASD. (3) The referral assessment by the eleven nurses was done retrospectively without a direct examination of the children by these nurses. It is

possible that better inter-rater agreement between nurses would have been achieved if this evaluation was done in a prospective manner with all nursed actually observing these toddlers at the clinic.

Conclusions

Our findings indicate that employment of the M-CHAT together with the GDS in the routine developmental screening at the MCHCs, would improve detection of toddlers with ASD and other DD in this population.

Author Contributions OK conceived of the study, participated in its design and coordination, performed the statistical analysis, and drafted the manuscript; EJH helped in the interpretation of the data and drafted the manuscript; GM participated in the design and coordination of the study and made all patient diagnoses; ND conceived of the study and participated in the design and coordination of the study; IM conceived of the study, participated in its design and coordination, oversaw the statistical analysis and draft the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, fifth edition (DSM-V)*. Arlington: American Psychiatric Publishing.

Bin Nun, G., Berlovitz, Y., Shani, M. (2010). *The Health Care System in Israel*, (2nd ed.), Tel Aviv: Am Oved Publishers Ltd.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics, 131*, e1121–e1127. https://doi.org/10.1542/peds.2012-1525.

Davidovitch, M., Hemo, B., Manning-Courtney, P., & Fombonne, E. (2013). Prevalence and incidence of autism spectrum disorder in an Israeli population. *Journal of Autism and Developmental Disorders, 43*, 785–793. https://doi.org/10.1007/s10803-012-1611-z.

Dawson, G., et al. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver Model. *Pediatrics, 125*, e17–e23. https://doi.org/10.1542/peds.2009-0958.

Eaves, L. C., & Ho, H. H. (1996). Brief report: Stability and change in cognitive and behavioral characteristics of autism through childhood. *Journal of Autism and Developmental Disorders, 26*, 557–569.

Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., & Dawson, G. (2015). Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*, 580–587. https://doi.org/10.1016/j.jaac.2015.04.005.

Herlihy, L. E., et al. (2014). Standardized screening facilitates timely diagnosis of autism spectrum disorders in a diverse sample of low-risk toddlers. *Journal of Developmental and Behavioral Pediatrics: JDBP, 35*, 85–92. https://doi.org/10.1097/dbp.0000000000000014.

Israel Ministry of Health. (2016). Guidance for the assessing of developmental assessments for infants and toddlers up to the age of six years.

Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics 120*, 1183–1215. https://doi.org/10.1542/peds.2007-2361.

Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the Modified Checklist for Autism in toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of Autism and Developmental Disorders, 44*, 194–203. https://doi.org/10.1007/s10803-013-1864-1.

Kleinman, J. M., et al. (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*, 827–839. https://doi.org/10.1007/s10803-007-0450-9.

Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *Lancet, 374*, 1627–1638. https://doi.org/10.1016/s0140-6736(09)61376-3.

Magalhães, L. D. C., Cardoso, A. A., Guimarães, M. A. (2015). How can we make our assessment of motor ability relevant cross-culturally. *Current Developmental Disorders Reports, 2*, 157–164.

Mahajnah, M., Sharkia, R., Shalabe, H., Terkel-Dawer, R., Akawi, A., & Zelnik, N. (2015). Clinical characteristics of autism spectrum disorder in Israel: Impact of ethnic and social diversities. *BioMed Research International*. https://doi.org/10.1155/2015/962093.

Meiri, G., et al. (2017). Brief report: The Negev Hospital-University-Based (HUB) Autism Database. *Journal of Autism and Developmental Disorders, 47*, 2918–2926. https://doi.org/10.1007/s10803-017-3207-0.

Mendonca, B., Sargent, B., & Fetters, L. (2016). Cross-cultural validity of standardized motor development screening and assessment tools: A systematic review. *Developmental Medicine & Child Neurology, 58*, 1213–1222. https://doi.org/10.1111/dmcn.13263.

Mohamed, F. E., et al. (2016). Screening of Egyptian toddlers for autism spectrum disorder using an Arabic validated version of M-CHAT; Report of a community-based study (Stage I). *European Psychiatry, 34*, 43–48. https://doi.org/10.1016/j.eurpsy.2016.01.2421.

Nygren, G., Sandberg, E., Gillstedt, F., Ekeroth, G., Arvidsson, T., & Gillberg, C. (2012). A new screening programme for autism in a general population of Swedish toddlers. *Research in Developmental Disabilities, 33*, 1200–1210. https://doi.org/10.1016/j.ridd.2012.02.018.

Øien, R. A., et al. (2018). Clinical features of children with autism who passed 18-month screening. *Pediatrics*. https://doi.org/10.1542/peds.2017-3596.

Pinto-Martin, J. A., Young, L. M., Mandell, D. S., Poghosyan, L., Giarelli, E., & Levy, S. E. (2008). Screening strategies for autism spectrum disorders in pediatric primary care. *Journal of Developmental and Behavioral Pediatrics, 29*, 345–350. https://doi.org/10.1097/DBP.0b013e31818914cf.

Raz, R., Weisskopf, M. G., Davidovitch, M., Pinto, O., & Levine, H. (2014). Differences in autism spectrum disorders incidence by sub-populations in Israel 1992–2009: A total population study. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-014-2262-z.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism, 12*, 537–556. https://doi.org/10.1177/1362361308094502.

Robins, D. L., et al. (2016). Universal autism screening for toddlers: Recommendations at odds. *Journal of Autism and Developmental Disorders, 46*, 1880–1882. https://doi.org/10.1007/s10803-016-2697-5.

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M., 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*, 37–45. https://doi.org/10.1542/peds.2013-1813.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). 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, 31*, 131–144.

Samadi, S. A., & McConkey, R. (2015). Screening for autism in Iranian preschoolers: Contrasting M-CHAT and a scale developed in Iran. *Journal of Autism and Developmental Disorders, 45*, 2908–2916. https://doi.org/10.1007/s10803-015-2454-1.

Siu, A. L., et al. (2016). Screening for autism spectrum disorder in young children: US Preventive Services task force recommendation statement. *JAMA, 315*, 691–696. https://doi.org/10.1001/jama.2016.0018.

Stenberg, N., et al. (2014). Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatric and Perinatal Epidemiology 28*, 255–262. https://doi.org/10.1111/ppe.12114.

The Central Bureau of Statistics of Israel. (2016). Live birth statistics. Statistical abstract of Israel.

Toh, T. H., Tan, V. W., Lau, P. S., & Kiyu, A. (2017). Accuracy of modified checklist for autism in toddlers (M-CHAT) in detecting autism and other developmental disorders in community clinics. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-017-3287-x.

Wiggins, L. D., Piazza, V., & Robins, D. L. (2014). Comparison of a broad-based screen versus disorder-specific screen in detecting young children with an autism spectrum disorder. *Autism, 18*, 76–84. https://doi.org/10.1177/1362361312466962.

Worley, J. A., Matson, J. L., Mahan, S., Kozlowski, A. M., & Neal, D. (2011). Stability of symptoms of autism spectrum disorders in Toddlers: An examination using the Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT). *Developmental Neurorehabilitation, 14*, 36–40. https://doi.org/10.3109/17518423.2010.530638.

Zwaigenbaum, L., et al. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics, 136*(Suppl 1), S60–S81. https://doi.org/10.1542/peds.2014-3667E.