

Predictive validity of the Infant Toddler Checklist in primary care at the 18-month visit and developmental diagnosis at 3–5 years: a prospective cohort study

Cornelia M Borkhoff,¹ Marina Atalla,² Imaan Bayoumi,³ Catherine S Birken,¹ Jonathon L Maguire,² Patricia C Parkin ¹

To cite: Borkhoff CM, Atalla M, Bayoumi I, *et al.* Predictive validity of the Infant Toddler Checklist in primary care at the 18-month visit and developmental diagnosis at 3–5 years: a prospective cohort study. *BMJ Paediatrics Open* 2022;**6**:e001524. doi:10.1136/bmjpo-2022-001524

Received 25 April 2022
Accepted 3 June 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Paediatrics, Hospital for Sick Children, Toronto, Ontario, Canada

²Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

³Department of Family Medicine and Centre for Studies in Primary Care, Queen's University, Kingston, Ontario, Canada

Correspondence to

Dr Patricia C Parkin; patricia.parkin@sickkids.ca

ABSTRACT

Objective There is international variation in recommendations regarding developmental screening and growing recognition of the low sensitivity of commonly used developmental screening tools. Our objective was to examine the predictive validity of the Infant Toddler Checklist (ITC) at 18 months to predict a developmental diagnosis at 3–5 years, in a primary care setting.

Methods We designed a prospective cohort study, recruiting in primary care in Toronto, Canada. Parents completed the ITC at the 18-month visit and reported developmental diagnosis at 3–5 years (developmental delay, autism spectrum disorder (ASD), attention deficit hyperactivity disorder (ADHD), learning problem). We calculated screening test properties with 95% CIs. We used multivariable logistic regression analyses adjusted for important covariates.

Results In the final sample (n=488), mean age at screening was 18.5 (SD 1.1) months, and at follow-up was 46.6 (SD 10.0) months. At screening, 46 (9.4%) had a positive ITC. At follow-up, 26 (5.3%) had a developmental diagnosis, including: developmental delay (n=22), ASD (n=4), ADHD (n=1), learning problem (n=1); parents of two children each reported two diagnoses (total of 28 diagnoses). Of four children with a diagnosis of ASD at follow-up, three had a positive ITC at 18 months. The ITC specificity (92%, 95% CI: 89% to 94%) and negative predictive value (96%, 95% CI: 95% to 97%) were high; false positive rate was low (8%, 95% CI: 6% to 11%); sensitivity was low (31%, 95% CI: 14% to 52%). There was a strong association between a positive ITC at 18 months and later developmental diagnosis (adjusted OR 4.48, 95% CI: 1.72 to 11.64; p=0.002).

Conclusion The ITC had high specificity, high negative predictive value, low false positive rate, and identified children with later developmental delay and ASD. The ITC had low sensitivity, similar to other screening tools underscoring the importance of continuous developmental surveillance at all health supervision visits.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There is international variation in recommendations regarding developmental screening and growing recognition of the low sensitivity of commonly used developmental screening tools. The Infant Toddler Checklist (ITC) was developed for early identification of children, 6–24 months, who have, or are at risk of developing, a communication impairment. There is little research on the predictive validity of the ITC when used to screen children at the 18-month visit in primary care.

WHAT THIS STUDY ADDS

⇒ The ITC at 18 months had high specificity and negative predictive value, and a low false positive rate for a developmental diagnosis at 3–5 years, including developmental delay and autism spectrum disorder. Children with a positive ITC at the 18-month visit had higher odds of a developmental diagnosis at follow-up. The low sensitivity of the ITC suggests that a positive ITC at the 18-month visit cannot accurately identify those who will have a developmental diagnosis at 3–5 years.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Practitioners and policymakers may consider the ITC as a developmental screening tool at the 18-month health supervision visit. The low sensitivity of the ITC is similar to other screening tools and underscores the importance of continuous developmental surveillance at all health supervision visits.

INTRODUCTION

Early identification of young children with developmental disorders is recommended in many countries. While primary care practitioners often perform developmental surveillance, the addition of standardised developmental screening tools may lead to earlier identification and referral for intervention, as shown in a randomised trial.¹ However, there

is international variation in recommendations regarding screening, including type of tool (general/broadband or domain/disorder-specific tool), age of screening and one-time versus repeat screening.^{2–11}

In Canada, the Canadian Paediatric Society recommends an enhanced 18-month visit in primary care, including the use of a developmental screening tool to 'stimulate discussion with parents about their child's development'.² However, there is no consensus on which screening tool is best suited for one-time screening at this visit, and there is a growing recognition of the low sensitivity of commonly used developmental screening tools.^{12–17}

Considering the importance of speech, language and social communication at 18 months, we examined the Infant Toddler Checklist (ITC). The ITC was developed by Wetherby and Prizant for early identification of children, 6–24 months, who have, or are at risk of developing, a communication impairment.¹⁸ Scoring provides recommendations for monitoring and referral.¹⁹ The developers also generated evidence supporting the concurrent and predictive criterion validity of the ITC for detection of a range of developmental concerns, including language delay, global developmental delay and autism spectrum disorder (ASD).^{20–24} However, an instrument should be validated for the target population and setting in which it will be used; we are interested in the criterion validity of the ITC in routine primary care.

Pierce *et al* selected the ITC for their screen–evaluate–treat model for early identification of children with ASD and have screened more than 44 000 children at 12, 18 and 24 months in a network of 203 primary care paediatricians in California, USA, demonstrating feasibility and physician satisfaction.^{25 26} Diagnostic evaluation was completed in those referred (39%), leading to developmental diagnoses including ASD and other delays.²⁶ However, since not all children had a diagnostic evaluation, criterion validity of the ITC could not be examined.^{25 26}

Concurrent criterion validity of many developmental screening tools has been evaluated, with concurrent measurement of screening and criterion measures. However, there is great interest in the predictive validity of developmental tools which 'offers a critical illustration of whether a screening test measures dimensions of development that are enduring and have a meaningful impact on children's long-term outcomes'.²⁷ Three recent systematic reviews addressed the predictive criterion validity of developmental screening tools; however, of included studies, few examined screening at 18 months or younger, and none included the ITC.^{28–30}

While the ITC may be a promising developmental screening tool for the 18-month visit in primary care, we have identified gaps in evidence regarding its predictive validity. Our objective was to examine the predictive validity of the ITC at the 18-month visit to predict a developmental diagnosis at 3–5 years of age.

METHODS

Design, setting, population

Using a prospective cohort design, we recruited healthy children between 2011 and 2018 during a scheduled 18-month health supervision visit at primary care practices participating in a longitudinal cohort in Toronto, Canada called *TARGET Kids!* (www.targetkids.ca). The cohort profile has been published and registered at ClinicalTrials.gov (NCT01869530).³¹ Current Canadian recommendations define screening as use of standardised tools to search for developmental delay in asymptomatic populations.³² To ensure a healthy, asymptomatic cohort, we excluded children with chronic health conditions (except asthma), established diagnosis of developmental delay, gestational age <32 weeks, birth weight <1.25 kg and unscheduled visit. To optimise questionnaire completion, parents unable to communicate in English were not included.

ITC at 18 months

At the 18-month visit, parents completed the ITC, a one-page, 24-item checklist, which takes 5 min to complete and 2 min to score.^{18 19} The checklist produces three composite scores (social, speech, symbolic) and their sum produces a total score. Age-based cut-offs (at or below the 10th percentile) indicate concern/no concern for each of the four scores (three composite scores and total score). The developers recommend that a child with concern on the speech composite should be monitored with repeat ITC in 3 months and referred for evaluation if there is concern on a second ITC (monitor/refer); and a child with concern on the social composite, symbolic composite or total score should be referred for evaluation (refer).

For this study, the ITC was only measured once at the 18-month visit. We examined three components of the ITC: concern for speech delay (defined as concern on the speech composite); concern for other communication delays (defined as concern on the social composite, symbolic composite or total score); positive ITC (defined as concern for speech delay and/or other communication delays).

Developmental diagnosis at 3–5 years

At 3–5 years, using a standardised questionnaire, parents responded to the question: 'Has your child been diagnosed with any of the following conditions?' (response options: developmental delay, ASD, attention deficit hyperactivity disorder (ADHD), learning problem, none). Parent response was dichotomised as a developmental diagnosis (yes/no). Parent report of clinician diagnosis, using similarly worded questions, has been used extensively in national surveys of developmental disorders including ASD.^{33–36}

Statistical analysis

We used all available data from children in the *TARGET Kids!* cohort meeting the eligibility criteria. Descriptive

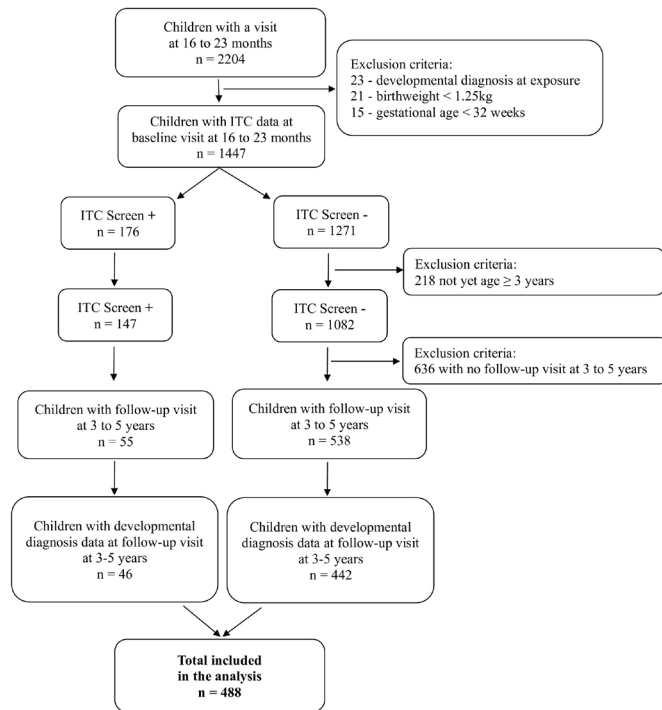


Figure 1 Study participant flow chart. ITC, Infant Toddler Checklist.

statistics were used to characterise the study participants. To assess the predictive validity of the ITC, we used two approaches.

First, we calculated the screening test properties (sensitivity, specificity, false positive rate, positive predictive value (PPV), negative predictive value (NPV)), with 95% CIs, for each of the three components of the ITC at 18 months, with a developmental diagnosis at 3–5 years as the criterion measure.

Second, we further evaluated the strength of the relationship between a positive ITC at the 18-month visit and the criterion measure, using multivariable logistic regression analyses. Three models were created corresponding to the three components of the ITC: positive ITC; concern for speech delay; concern for other communication delays. Potential confounding variables selected a priori included child age at follow-up, sex, birth weight, maternal ethnicity and family income. Models were adjusted for all covariates regardless of statistical significance.³⁷ All potential confounders had <13% missing data. Missing covariate data were handled by multiple imputation using the fully conditional specification method.³⁸ To reduce the potential for bias, models were run on 20 imputed data sets.³⁹ Results of the 20 imputed data sets were combined, and the parameter estimates (95% CI) for the adjusted pooled models were reported. Statistical significance was defined as $p < 0.05$; all statistical tests were two sided. Statistical analysis was conducted using SAS V.9.4 statistical software (SAS Institute).

Patient and public involvement

Parents and members of the public were not involved in the design, analysis or interpretation of the research presented here. *TARGet Kids!* has now developed a Parent and Clinician Team (<https://www.targetkids.ca/pact>).

RESULTS

Participants

Of 593 children with an ITC at baseline and follow-up at 3–5 years, 488 (82%) had outcomes on developmental diagnosis and were included in the analysis (figure 1). The mean age at screening was 18.5 (SD 1.1) months and at follow-up was 46.6 (SD 10.0) months (table 1). At screening, 46 (9.4%) children had a positive ITC: concern for speech delay (n=28, 5.7%); concern for other communication delays (n=30, 6.2%); concern for both speech delay and other communication delays (n=12, 2.5%). At follow-up, 26 (5.3%) children had a developmental diagnosis, with parents of two children each reporting two diagnoses, for a total of 28 diagnoses: developmental delay (n=22); ASD (n=4); ADHD (n=1); learning problem (n=1). Of the four children with a diagnosis of ASD at follow-up (0.8% of the total sample), three had a positive ITC at 18 months, all with concern for other communication delays.

Predictive validity

Sensitivity was 23%–31%, specificity 92%–95%, false positive rates 5%–8%, PPV 17%–21% and NPV 96% (table 2). Children with a positive ITC at 18 months had higher odds of a developmental diagnosis at 3–5 years (adjusted OR (aOR) 4.48, 95% CI: 1.72 to 11.64; $p=0.002$), as did male sex (aOR 3.05, 95% CI: 1.17 to 7.97; $p=0.02$) (table 3). There was a strong association between ITC concern for speech delay (aOR 4.78, 95% CI: 1.65 to 13.81; $p=0.004$) and ITC concern for other communication delays (aOR 4.46, 95% CI 1.71 to 11.64; $p=0.002$) and a developmental diagnosis at 3–5 years (table 4).

DISCUSSION

In this study, parents of 488 children completed the ITC at their child's 18-month health supervision visit. On the basis of a 10th percentile cut-off, 48 children were expected to have a positive ITC screen. In our sample, 46 children (9.4%) had a positive ITC screen, including 5.7% with concern for speech delay, 6.2% with concern for other communication delays and 2.5% with concern for both. At follow-up, at a mean age of 4 years, approximately 5% had a developmental diagnosis, including developmental delay, ASD, ADHD and learning problems. Of four children with a diagnosis of ASD at 3–5 years, three had a positive ITC screen, all with concern for other communication delays, which is notable as this ITC component is thought to capture ASD.^{24–26} Children with a positive ITC at the 18-month visit had 4.48 higher odds of a developmental diagnosis at follow-up. The

Table 1 Characteristics of study participants (n=488)

Characteristic	All participants	Infant Toddler Checklist (ITC)	
		Positive	Negative
N	488	46	442
Child and family characteristics	n		
Female sex	488	223 (45.7)	19 (41.3)
Birth weight, kg	456	3.2 (0.6)	3.1 (0.6)
Body mass index, z-score	484	0.13 (1.1)	0.17 (0.9)
Only child	486	220 (45.3)	18 (39.1)
Maternal age at birth, years	452	34.4 (4.0)	34.2 (3.8)
Maternal ethnicity*	425		
European		286 (67.3)	22 (55.0)
Non-European		139 (32.7)	18 (45.0)
Maternal education	480		
High school or less		24 (5.0)	2 (4.4)
College/university		456 (95.0)	43 (95.6)
Family income (\$C)	480		
Less than \$40 000		33 (6.9)	9 (20.0)
\$40 000–\$79 999		54 (11.3)	8 (17.8)
\$80 000–\$149 999		164 (34.2)	13 (28.9)
\$150 000+		229 (47.7)	15 (33.3)
Family immigration status	465		
Canadian born		269 (57.9)	24 (54.6)
Immigrant, industrialised		60 (12.9)	2 (4.6)
Immigrant, non-industrialised		136 (29.3)	18 (40.9)
Family history of developmental concern†	356	29 (8.2)	4 (10.8)
Baseline			
Age at baseline, months	488	18.5 (1.1)	18.3 (0.8)
Positive ITC screen	488	46 (9.4)	46 (100)
Concern for speech delay	488	28 (5.7)	28 (60.9)
Concern for other communication delays	488	30 (6.2)	30 (65.2)
Follow-up			
Age at follow-up, months	488	46.6 (10.0)	47.0 (9.5)
Developmental diagnosis	488	26 (5.3)	18 (4.1)

Data regarding baseline characteristics are presented as mean (SD) or N (%).

*Non-European consists of 37 mixed=2 or more ethnic groups (8.2%), 33 South Asian (6.8%), 31 East Asian (6.4%), 14 African and Caribbean (2.9%), 13 Latin American (2.7%), 7 Southeast Asian (1.4%), 3 West Asian/North African (0.6%), 1 Indigenous (0.2%).

†Family history of developmental concerns includes history of ASD, ADHD, or learning disability in mother, father or siblings.

ADHD, attention deficit hyperactivity disorder; ASD, autism spectrum disorder.

high specificity (92%–95%) and NPV (96%) suggest that most children with a negative ITC screen will not have a later developmental diagnosis. The low false positive rate (5%–8%) suggests that use of the ITC will result in few unintended harms related to overdiagnosis and over-referral.

We also identified low sensitivity of the ITC suggesting that a positive ITC at the 18-month visit cannot accurately identify those who will have a developmental diagnosis at 3–5 years. Low sensitivity of developmental screening

tools to predict later outcomes has been recognised as a challenge, due to the dynamic nature of children's development.^{27 30} Factors associated with lower sensitivity include younger age at screening and longer latency from screening to outcome.³⁰ Therefore, it is not surprising that screening at 18 months resulted in a low sensitivity to predict developmental diagnosis at 3–5 years. This underscores the importance of continuous developmental surveillance at all health supervision visits, as recommended by professional organisations.^{2 5}

Table 2 Screening test properties of the Infant Toddler Checklist (ITC) screen at 18 months compared with developmental diagnosis at 3–5 years (n=488)

ITC screen*	Developmental diagnosis		
	Yes	No	
Positive	8	38	46
Negative	18	424	442
	26	462	488
Sensitivity (95% CI), %	30.8 (14.3 to 51.8)		
Specificity (95% CI), %	91.8 (88.9 to 94.1)		
False positive rate (95% CI), %	8.2 (5.8 to 11.3)		
Positive predictive value (95% CI), %	17.4 (9.9 to 28.8)		
Negative predictive value (95% CI), %	95.9 (94.8 to 96.8)		
Concern for speech delay	Yes	No	
Positive	6	22	28
Negative	20	440	460
	26	462	488
Sensitivity (95% CI), %	23.1 (9.0 to 43.7)		
Specificity (95% CI), %	95.2 (92.9 to 97.0)		
False positive rate (95% CI), %	4.8 (3.0 to 7.2)		
Positive predictive value (95% CI), %	21.4 (10.8 to 38.1)		
Negative predictive value (95% CI), %	95.7 (94.7 to 96.5)		
Concern for other communication delays	Yes	No	
Positive	6	24	30
Negative	20	438	458
	26	462	488
Sensitivity (95% CI), %	23.1 (9.0 to 43.7)		
Specificity (95% CI), %	94.8 (92.4 to 96.7)		
False positive rate (95% CI), %	5.2 (3.3 to 7.7)		
Positive predictive value (95% CI), %	20.0 (10.1 to 35.8)		
Negative predictive value (95% CI), %	95.6 (94.7 to 96.4)		

*An ITC screen is positive if there is concern for speech delay and/or other communication delays.

There have been three recent systematic reviews of predictive validity of developmental screening in early childhood, highlighting the interest in understanding which tools best identify children who may benefit from early identification and intervention.^{28–30} Of included studies, few examined screening at ≤ 18 months, and none examined the ITC.

Sim *et al* examined the predictive validity of language screening tools at 2–6 years in six studies.²⁸ Studies of children < 18 months were not eligible, and only two studies had a time to follow-up of ≥ 12 months, with a mean sensitivity of 54%. Cairney *et al* examined the predictive value of developmental assessment at 1–5 years in 13 studies and found a positive association between poor early child development and later educational difficulties, high specificity and NPV, and low sensitivity.³⁰ Three studies included children < 18 months, which assessed association but not screening test properties.

Schönhaut *et al* examined the predictive validity of the Ages & Stages Questionnaire (ASQ), which is of great importance given its common use in many countries.²⁹ Of five included studies, three screened children at 36 or 48 months; and while the remaining two studies screened some children at 18 months, they included children born preterm. Lamsal *et al* (not included in any systematic review) examined the predictive validity of the ASQ for parent report of a developmental diagnosis at 4–5 years.¹⁵ At 24 months (the youngest age examined), using the 1 SD cut-off for the ASQ, sensitivity was 84% and specificity was 69%; using the 2 SD cut-off, sensitivity was 32% and specificity was 91%. In summary, little is known about the predictive validity of the ASQ in healthy term infants screened at 18 months.

Low to moderate sensitivity has been found in concurrent validity studies of the ASQ.^{12 13 16} Warren *et al* found a sensitivity of 55% and specificity of 86% in a meta-analysis of four studies.¹⁶ Sheldrick *et al* found a sensitivity of 35% and specificity of 89% in children 9 months–5.5 years (mean 2.6 years).¹³ Wilson *et al* found a sensitivity of 55% and specificity of 95% on the ASQ Communication Scale in children 23–30 months (mean 25 months).¹²

Low sensitivity of ASD-specific tools has also been found. In a predictive validity study of the Modified Checklist for Autism in Toddlers with Follow-Up in children 16–26 months in primary care, Guthrie *et al* found a sensitivity of 39% and specificity of 95% for ASD; and a sensitivity of 12% and specificity of 97% for any delay.¹⁴

The ITC was developed by Wetherby and Prizant in US children.¹⁸ For identification of communication disorders, concurrent validity was assessed in children 12–17 and 18–24 months: sensitivity 86%–89% and specificity 74%–77%²¹; and predictive validity in children 12–24 months, with language assessed at 3 years: sensitivity 83% and specificity 70%.²¹ For identification of ASD, in children 9–24 months (n=5385), 60 who were ≥ 4 years received a diagnosis of ASD; of these, the ITC identified 56, for a sensitivity of 93%.²⁴

Pierce *et al* selected the ITC in their screen–evaluate–treat model embedded in primary care, aiming for early ASD intervention before 2 years.^{25 26} A network of 203 paediatricians screened more than 44 000 children at 12, 18 and 24 months using the ITC. Approximately 39% of children with a positive ITC were referred for diagnostic evaluation. Of these, almost half received a diagnosis of ASD and about one-third received a diagnosis



Table 3 Logistic regression models for the association between screening with the Infant Toddler Checklist (ITC) at 18 months and developmental diagnosis at 3–5 years (n=488)

Predictor	Unadjusted analysis		Adjusted analysis*	
	OR (95% CI)	P value	OR (95% CI)	P value
Positive ITC screen	4.96 (2.02 to 12.16)	<0.001	4.48 (1.72 to 11.64)	0.002
Age at follow-up, months	0.98 (0.94 to 1.02)	0.29	0.97 (0.93 to 1.02)	0.19
Sex, male	2.95 (1.16 to 7.49)	0.02	3.05 (1.17 to 7.97)	0.02
Birth weight, kg	0.98 (0.50 to 1.95)	0.96	0.95 (0.47 to 1.93)	0.88
Maternal ethnicity				
Non-European	2.36 (1.05 to 5.31)	0.04	2.10 (0.83 to 5.31)	0.12
European	1.00 (ref)	—	1.00 (ref)	—
Family income (\$C)				
Less than \$40 000	2.73 (0.82 to 9.15)	0.1	1.32 (0.33 to 5.30)	0.69
\$40 000–\$79 999	1.17 (0.31 to 4.33)	0.82	0.69 (0.17 to 2.89)	0.62
\$80 000–\$149 999	0.75 (0.27 to 2.08)	0.58	0.65 (0.23 to 1.85)	0.42
\$150 000+	1.00 (ref)	—	1.00 (ref)	—

Bold=statistically significant findings at p<0.05.

*Adjusted for child age at follow-up in months, child sex, birth weight, maternal ethnicity, family income. All covariates were measured at baseline except child age which was at follow-up.

of developmental delay, language delay or other delays.²⁶ Pierce *et al* found that a diagnosis of ASD becomes stable starting at 14 months, that the most common diagnostic transition was from language or developmental delay to ASD, and that almost 24% of children with an ASD diagnosis at 3–4 years were late identified.⁴⁰ In our study, three of four children with an ASD diagnosis at 3–5 years had a positive ITC at 18 months, supporting the potential for early identification of ASD using the ITC.

Strengths of the ITC include its focus on communication, an important developmental domain at 18 months for which interventions are available. The ITC distinguishes concern for speech delay (monitor/refer) and other communication delays such as ASD (refer). Overall, the ITC screen positive rate is about 10% and false positives are low, minimising overdiagnosis and over-referral. Additional advantages include its one-page format, ease of completion and availability free of charge.

Limitations of this study include parent report of a developmental diagnosis rather than a standardised clinician assessment. However, parent report of physician diagnosis of developmental disorders including ASD has been used extensively in national surveys such as the US National Survey of Children's Health and the National Health Interview Survey.^{33–36} Kogan *et al* have summarised the evidence supporting the validity of parent report compared with physician diagnosis.³³ In our study, maternal education and family income were high, which may limit the generalisability of our findings. However, family income was included as a covariate in our analysis and the overall prevalence of a positive ITC screen in our sample was close to the expected 10th percentile cut-off score.¹⁹

CONCLUSION

Developmental screening at 18 months may lead to early identification of communication impairments (including

Table 4 Multivariable logistic regression models for the association between screening with the Infant Toddler Checklist (ITC) at 18 months and developmental diagnosis at 3–5 years (n=488)

Predictor†	Developmental diagnosis			
	Unadjusted analysis		Adjusted analysis*	
	OR (95% CI)	P value	OR (95% CI)	P value
Positive ITC screen‡	4.96 (2.02 to 12.16)	<0.001	4.48 (1.72 to 11.64)	0.002
Concern for speech delay	6.00 (2.19 to 16.44)	<0.001	4.78 (1.65 to 13.81)	0.004
Concern for other communication delays	5.48 (2.01 to 14.89)	<0.001	4.46 (1.71 to 11.64)	0.002

Bold=statistically significant findings at p<0.05.

*Adjusted for child age at follow-up in months, child sex, birth weight, maternal ethnicity, family income. All covariates were measured at baseline except child age which was at follow-up.

†Each row shows results from three separate models.

‡An ITC screen is positive if there is concern for speech delay and/or other communication delays.

speech delay and other communication disorders such as ASD) for which interventions are available. This study provides evidence on the predictive validity of the ITC in primary care at the 18-month visit, suggesting that it performs as well as other currently available developmental screening tools.

Acknowledgements We thank all participating children and families for their time and involvement in this study. We thank the *TARGet Kids!* Collaboration for supporting this study (details may be found on our website: www.targetkids.ca). The *TARGet Kids!* Collaboration is a primary care practice-based research network and includes practice site physicians, research staff, collaborating investigators, trainees, methodologists, biostatisticians, data management personnel, laboratory management personnel and advisory committee members.

Contributors CB and PP drafted the manuscript. CB and MA analysed the data. All authors contributed to the design of the study, edited the manuscript and approved the final submission. CB and PP contributed equally, are joint senior authors and act as guarantors for the paper.

Funding Funding to support *TARGet Kids!* was provided by multiple sources including the Canadian Institutes for Health Research (CIHR), namely the Institute of Human Development, Child and Youth Health and the Institute of Nutrition, Metabolism and Diabetes, as well as the St Michael's Hospital Foundation, Toronto. The Pediatric Outcomes Research Team is supported by a grant from the Hospital for Sick Children Foundation, Toronto.

Disclaimer Funding agencies had no role in the design, collection, analyses or interpretation of the results of this study or in the preparation, review or approval of the manuscript.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval Consent was obtained from parents. Ethics approval was granted by the Hospital for Sick Children, Toronto Ethics Board (REB# 1000012436). The ITC was completed for research purposes; practitioners and parents were blind to the results.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data available upon reasonable request according to the *TARGet Kids!* data access policy.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Patricia C Parkin <http://orcid.org/0000-0003-2935-7145>

REFERENCES

- Guevara JP, Gerdes M, Localio R, *et al.* Effectiveness of developmental screening in an urban setting. *Pediatrics* 2013;131:30–7.
- Williams R, Clinton J, Canadian Paediatric Society, Early Years Task Force. Getting it right at 18 months: in support of an enhanced well-baby visit. *Paediatr Child Health* 2011;16:647–54.
- Zwaigenbaum L, Brian JA, Ip A. Early detection for autism spectrum disorder in young children. *Paediatr Child Health* 2019;24:424–32.
- Guttmann A, Saunders NR, Kumar M, *et al.* Implementation of a physician incentive program for 18-month developmental screening in Ontario, Canada. *J Pediatr* 2020;226:213–20.
- Lipkin PH, Macias MM, Council On Children With DISABILITIES, SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS. Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics* 2020;145:e20193449.
- Hyman SL, Levy SE, Myers SM, *et al.* Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics* 2020;145:e20193447.
- Garg P, Ha MT, Eastwood J, *et al.* Health professional perceptions regarding screening tools for developmental surveillance for children in a multicultural part of Sydney, Australia. *BMC Fam Pract* 2018;19:42.
- McLean K, Goldfeld S, Molloy C. *Screening and surveillance in early childhood health: rapid review of evidence for effectiveness and efficiency of models. An evidence check review brokered by the.* Australia: Sax Institute for NSW Kids and Families, 2014.
- Bellman M, Byrne O, Sege R. Developmental assessment of children. *BMJ* 2013;346:e8687.
- UK Department of Health. The healthy child programme 2 year review. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377800/dh_108329.pdf [Accessed 05 Apr 2022].
- Wilson P, Wood R, Lykke K, *et al.* International variation in programmes for assessment of children's neurodevelopment in the community: understanding disparate approaches to evaluation of motor, social, emotional, behavioural and cognitive function. *Scand J Public Health* 2018;46:805–16.
- Wilson P, Rush R, Charlton J, *et al.* Universal language development screening: comparative performance of two questionnaires. *BMJ Paediatr Open* 2022;6:e001324.
- Sheldrick RC, Marakovitz S, Garfinkel D, *et al.* Comparative accuracy of developmental screening questionnaires. *JAMA Pediatr* 2020;174:366–74.
- Guthrie W, Wallis K, Bennett A, *et al.* Accuracy of autism screening in a large pediatric network. *Pediatrics* 2019;144:e20183963.
- Lamsal R, Dutton DJ, Zwicker JD. Using the ages and stages questionnaire in the general population as a measure for identifying children not at risk of a neurodevelopmental disorder. *BMC Pediatr* 2018;18:122.
- Warren R, Kenny M, Fitzpatrick-Lewis D. Screening and treatment for developmental delay in early childhood (ages 1–4): systematic review Hamilton, Ontario: McMaster University; 2014. screening-and-treatment-for-dev-delay-in-early-childhood-ages-1-4-plus-add-finaljan-4-2016-.pdf (canadiantaskforce.ca) [Accessed 05 Apr 2022].
- Wetherby AM, Guthrie W, Hooker JL, *et al.* The early screening for autism and communication disorders: Field-testing an autism-specific screening tool for children 12 to 36 months of age. *Autism* 2021;25:2112–23.
- Wetherby AM, Prizant G. *CSBS DP manual*. First Normed Edition. Baltimore: Brookes Publishing, 2008.
- Checklist IT. First words project – checklist and scoring. Available: <https://firstwords.fsu.edu/pdf/checklist.pdf> and https://firstwords.fsu.edu/pdf/Checklist_Scoring_Cutoffs.pdf [Accessed 05 Apr 2022].
- Wetherby AM, Allen L, Cleary J, *et al.* Validity and reliability of the communication and symbolic behavior scales developmental profile with very young children. *J Speech Lang Hear Res* 2002;45:1202–18.
- Wetherby A, Goldstein H, Cleary J. Early identification of children with communication disorders: concurrent and predictive validity of the CSBS developmental profile. *Infants Young Child* 2003;16:161–74.
- Wetherby AM, Woods J, Allen L, *et al.* Early indicators of autism spectrum disorders in the second year of life. *J Autism Dev Disord* 2004;34:473–93.
- Wetherby AM, Watt N, Morgan L, *et al.* Social communication profiles of children with autism spectrum disorders late in the second year of life. *J Autism Dev Disord* 2007;37:960–75.
- Wetherby AM, Brosnan-Maddox S, Peace V, *et al.* Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism* 2008;12:487–511.
- Pierce K, Carter C, Weinfeld M, *et al.* Detecting, studying, and treating autism early: the one-year well-baby check-up approach. *J Pediatr* 2011;159:458–65.
- Pierce K, Gazestani V, Bacon E, *et al.* Get SET early to identify and treatment refer autism spectrum disorder at 1 year and discover factors that influence early diagnosis. *J Pediatr* 2021;236:179–88.
- Marks K, Glascoe FP, Aylward GP, *et al.* The thorny nature of predictive validity studies on screening tests for developmental-behavioral problems. *Pediatrics* 2008;122:866–8.
- Sim F, Thompson L, Marryat L, *et al.* Predictive validity of preschool screening tools for language and behavioural difficulties: a PRISMA systematic review. *PLoS One* 2019;14:e0211409.
- Schonhaut L, Maturana A, Cepeda O, *et al.* Predictive validity of developmental screening questionnaires for identifying children with later cognitive or educational difficulties: a systematic review. *Front Pediatr* 2021;9:698549.



- 30 Cairney DG, Kazmi A, Delahunty L, *et al.* The predictive value of universal preschool developmental assessment in identifying children with later educational difficulties: a systematic review. *PLoS One* 2021;16:e0247299.
- 31 Carsley S, Borkhoff CM, Maguire JL, *et al.* Cohort Profile: The Applied Research Group for Kids (TARGet Kids!). *Int J Epidemiol* 2015;44:776–88.
- 32 Canadian Task Force on Preventive Health Care. Recommendations on screening for developmental delay. *CMAJ* 2016;188:579–87.
- 33 Kogan MD, Vladutiu CJ, Schieve LA, *et al.* The prevalence of parent-reported autism spectrum disorder among US children. *Pediatrics* 2018;142:e20174161.
- 34 Xu G, Strathearn L, Liu B, *et al.* Prevalence of autism spectrum disorder among US children and adolescents, 2014–2016. *JAMA* 2018;319:81–2.
- 35 Yuan J, Li M, Lu ZK. Racial/ethnic disparities in the prevalence and trends of autism spectrum disorder in US children and adolescents. *JAMA Netw Open* 2021;4:e210771.
- 36 Gurney JG, McPheeters ML, Davis MM. Parental report of health conditions and health care use among children with and without autism: national survey of children's health. *Arch Pediatr Adolesc Med* 2006;160:825–30.
- 37 Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Wiley, 2001.
- 38 Buuren Svan, Groothuis-Oudshoorn K. mice : Multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
- 39 Little R, Rubin DB. *Statistical analysis with missing data*. Wiley 2002.
- 40 Pierce K, Gazestani VH, Bacon E, *et al.* Evaluation of the diagnostic stability of the early autism spectrum disorder phenotype in the general population starting at 12 months. *JAMA Pediatr* 2019;173:578–87.