Rates of detection of developmental problems at the 18-month well-baby visit by family physicians’ using four evidence-based screening tools compared to usual care: a randomized controlled trial

R. E. Thomas,* W. Spragins,† G. Mazloum,‡ M. Cronkhite§ and G. Maru¶
*Family Medicine, Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada
†Independent Research Consultant, Calgary, Alberta, Canada
‡Foothills Family Practice, Alberta, Canada
§Family Physician, Calgary, Alberta, Canada, and
¶Children’s Hospital, Calgary, Alberta and Independent Research Consultant, Calgary, Alberta, Canada
Accepted for publication 4 February 2016

Abstract

Background Early and regular developmental screening can improve children’s development through early intervention but is insufficiently used. Most developmental problems are readily evident at the 18-month well-baby visit. This trial’s purpose is to: (1) compare identification rates of developmental problems by GPs/family physicians using four evidence-based tools with non-evidence based screening, and (2) ascertain whether the four tools can be completed in 10-min pre-visit on a computer.

Methods We compared two approaches to early identification via random assignment of 54 families to either: ‘usual care’ (informal judgment including ad-hoc milestones, n = 25); or (2) ‘Evidence-based’ care (use of four validated, accurate screening tools, n = 29), including: the Parents’ Evaluation of Developmental Status (PEDS), the PEDS-Developmental Milestones (PEDS-DM), the Modified Checklist for Autism in Toddlers (M-CHAT) and PHQ9 (maternal depression).

Results In the ‘usual care’ group four (16%) and in the evidence-based tools group 18 (62%) were identified as having a possible developmental problem. In the evidence-based tools group three infants were to be recalled at 24 months for language checks (no specialist referrals made). In the ‘usual care’ group four problems were identified: one child was referred for speech therapy, two to return to check language at 24 months and a mother to discuss depression. All forms were completed on-line within 10 min.

Conclusions Despite higher early detection rates in the evidence-based care group, there were no differences in referral rates between evidence-based and usual-care groups. This suggests that clinicians: (1) override evidence-based screening results with informal judgment; and/or (2) need assistance understanding test results and making referrals. Possible solutions are improve the quality of information obtained from the screening process, improved training of physicians, improved support for individual practices and acceptance by the regional health authority for overall responsibility for screening and creation of a comprehensive network.
Introduction

Developmental problems

The US 2003 National Survey of Children’s Health identified that of children 6–17 years 11% had learning disabilities, 8.8% ADHD and 6.3% behavioral problems (Blanchard et al., 2006). A 2014 publication assessed that in the US about 17% of children have behavioral or developmental problems (e.g. motor, language, pre-academic or academic skills, behavioral, social–emotional, mental health or autism spectrum disorders). Twenty to 25% have psychosocial risks (parental mental health, education, poverty, housing or failure to encourage their child’s language or academic skills) (Glascoe 2014).

One method of comparing child development across countries is to compare results on a screening instrument that has been translated into several languages (The Ages and Stages Questionnaire). Kerstens (Kerstjens et al., 2009) compared the Ages and Stages for 2508 Dutch children to Norwegian, US and Korean samples and concluded: ‘The mean scores of Dutch children for all domains except for the fine motor domain differed significantly from the US mean scores. Moreover, Dutch mean scores were statistically significantly lower than the Norwegian scores in all domains. The Dutch and Korean children differed significantly with regard to the fine and gross motor domains. Differences were generally small, being only clinically relevant (effect sizes (Cohen’s delta) > 0.5, or differences in raw scores > 5 points, the smallest possible increase in domain scores) in the problem solving domain (US) and the fine motor domain (Norway and Korea).’ Janson & Squires (2004) compared 1341 Norwegian children to the original US norming study of over 8000 children (Squires Bricker & Potter, 1999). Out of 50 comparisons only five were either larger than 0.5 d or 5 raw score points (1/2 item), which is the smallest increment on an ASQ domain score. In these samples from a few countries there are differences but overall a pattern of similarity.

In the US only 20% or 30% of children with developmental problems are identified before school age (Brothers, Glascoe & Robertshaw 2008). The children who tend to be identified early are those who have marked symptoms, other obvious medical needs or have severe developmental and/or behavioral problems.

Structured screening tools vs. informal surveillance

There is significant variability between primary care physicians in their screening practices. The benefits of more structured information exchanges are shown in the study by Glascoe (2014), in which the number and range of concerns discussed during a routine well-baby visit were markedly improved by using parent-completed brief questionnaires to help structure the visits.

Because of their frequent contact with and observation of their children parents are likely to be the best observers of their children, and the literature shows that if a parent is concerned, then there is a high probability of a problem (Glascoe 2014) (however, if a parent is not concerned it does not mean there are no problems).

Parent-completed questionnaires tend to produce more accurate and reliable responses than a few open-ended general ‘surveillance’ questions, and have the advantages that they can be completed rapidly by parents, are easily incorporated into clinical care and tend to initiate a structured dialogue between the parents and the health professionals (Glascoe 2014). Guevara (Guevara et al., 2013) in a CDC-funded trial randomized 2103 children to either screening with the Ages and Stages-II and M-CHAT, or developmental surveillance using milestones. In the evidence-based group 25% were assessed as developmental delay and 19% of all infants were referred, whereas in the surveillance groups 13% were detected and 10% referred. Logistic regression found that the multivariate adjusted odds ratio of referral depended on parental factors (maternal age ≥ 30, AOR = 0.81; African–American 0.68; single parent 0.75), child factors (two or more developmental domains 3.15; special health needs 3.16) and practitioner (developmental screening without additional staff support 0.44; ‘usual care’ 0.45) (Jimenez et al., 2014). A study of 95 children at the 9, 18 and 24-month health supervision visits found that 11 were identified as delayed using surveillance, 15 failed with the ASQ (of whom 10 were not identified by surveillance) and 28 were scored to be monitored (Thomas et al., 2012). The use of standardized screening tools has increased. A survey of American Academy of Pediatrics members found that the proportion reporting using ≥1 screening tool increased from 23% in 2002 to 47.7% in 2009 (but this is still <50% of all pediatricians!) (Radecki et al., 2011).

Goals and objectives

The purpose of this trial is to: (1) compare the rate of identification of developmental problems by GPs/Family physicians using ‘usual care’ – meaning informal judgment including ad-hoc milestones; or ‘Evidence-based’ care – meaning © 2016 John Wiley & Sons Ltd, Child: care, health and development, 42, 3, 382–393
use of four validated, accurate screening tools \((n = 29)\), including: the Parents Evaluation of Developmental Status (PEDS), the PEDS-Developmental Milestones (PEDS-DM), the Modified Checklist for Autism in Toddlers (M-CHAT) and PHQ9 (maternal depression). To provide a common point of comparison both groups were also asked to use the non-evidence based screening tool most family physicians in Canada use, the Rourke Baby Record. (2) ascertain whether the four tools can be completed in a 10 minute pre-visit on a computer without disrupting the office routine. These tools have been used in some pediatricians’ offices in Canada and the US and we wished to perform a ‘proof-of-principle’ test whether this could be accomplished in the offices of GPs/family physicians who usually have much shorter visits. The PEDS and PEDS-DM had not previously been tested in family physician’s offices in Canada.

**Methods**

**Selection of screening tools: instrument sensitivities and specificities**

We reviewed screening forms which use parental input to assess child development (Table 1) and parental input to assess behavioral and emotional concerns (Table 2). We needed forms that were short enough to be used in busy family physicians’ offices yet had high sensitivities and specificities. We chose the PEDS, PEDS-DM, M-CHAT and PHQ9 (to measure maternal depression).

**Parents’ evaluation of developmental status**

We chose this for two reasons (Tables 1 and 2): it has excellent sensitivity across ages (86%) and specificity (83%) for parental

<table>
<thead>
<tr>
<th>Developmental screening form and time to complete</th>
<th>Age range</th>
<th>Description</th>
<th>Scoring</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents’ Evaluation of Developmental Status (2 min)</td>
<td>Birth to 9 years</td>
<td>10 questions, based on 300 item Brigance screen</td>
<td>Identifies three paths: low moderate or high risk for disabilities or delays. Whether to monitor development, provide education to parents, re-screen or refer</td>
<td>74–79% according to age level</td>
<td>70–80% according to age level</td>
</tr>
<tr>
<td>Parents’ Evaluation of Developmental Status: Milestones (5 min)</td>
<td>Birth to 8 years</td>
<td>6–8 items per visit (fine motor, gross motor, social-emotional, self-help, expressive language, receptive language; and reading and math for older children)</td>
<td>Scoring template builds bar graph of development in each domain</td>
<td>70%</td>
<td>77–93%</td>
</tr>
<tr>
<td>Ages and Stages Questionnaire (15 min if interview needed)</td>
<td>4–60 months</td>
<td>19 separate 30 item questionnaires for each age group (communication, gross motor, fine motor, problem solving, personal-social skills)</td>
<td>Pass or fail on developmental status</td>
<td>70–90% according to age level</td>
<td>71–91% according to age level</td>
</tr>
<tr>
<td>Infant Development Inventory (10 min if interview needed)</td>
<td>3–18 months</td>
<td>60 yes/no descriptions of development in five domains. A 300 item assessment-level version can be used by professionals and produces age equivalent and cutoff scores for each domain</td>
<td>Cutoffs for each domain for either 30% delay or 30% advanced level relative to chronological age</td>
<td>75%</td>
<td>70%</td>
</tr>
<tr>
<td>Infant–Toddler Checklist for language and Communication (5–10 min)</td>
<td>6–24 months</td>
<td>24 MCQs. Screens for language delays as the first symptom the child is not developing typically.</td>
<td>Manual table of cutoff scores (1.25 standard deviations below mean)</td>
<td>78%</td>
<td>84%</td>
</tr>
</tbody>
</table>
concerns for which their child is at high/moderate developmental risk (Glascoe 2013), and it is the most time efficient tool for asking parents to identify their child’s problems. A systematic review identified 37 studies of the PEDS (n=210 242) and based on parental concerns 13.8% (95%CI 10.9, 16.8%) were at high and 19.8% (95%CI 16.7, 22.9%) at moderate developmental risk. Parental concerns associated with high developmental risk were related to low birth weight, male gender, poor or fair child health rating, minority ethnicity, lower socioeconomic status and limited health care access/insurance. Parents had more concerns about older children irrespective of SES, perhaps indicating increased developmental demands with age (Wooffenden et al. 2014).

The pooled estimate of combined high and moderate developmental risk PEDS (27.0–36.0%) is higher than for the Australian Early Development Index (AEDI) (Goldfeld et al. 2012) and the Ages and Stages Questionnaire (Sices et al. 2009; Limbos & Joyce 2011; Restall & Borton 2010) but similar to the Denver Developmental Screening Test (DDST) (Miller 1998; Nair et al. 2009).

### Table 2.
Forms which use parental input to assess behavioral and emotional concerns

<table>
<thead>
<tr>
<th>Developmental screening form and time to complete</th>
<th>Age range</th>
<th>Description</th>
<th>Scoring</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents’ Evaluation of Developmental Status (2 min)</td>
<td>Birth to 9 years</td>
<td>10 questions, based on 300 item Brigance screen</td>
<td>Identifies three paths: low moderate or high risk for disabilities or delays. Whether to monitor development, provide education to parents, re-screen or refer</td>
<td>74–79% according to age level</td>
<td>70–80% according to age level</td>
</tr>
<tr>
<td>Parents’ Evaluation of Developmental Status: Milestones (5 min)</td>
<td>Birth to 8 years</td>
<td>6–8 items per visit (fine motor, gross motor, social–emotional, self-help, expressive language, receptive language; and reading and math for older children)</td>
<td>Scoring template builds bar graph of development in each domain</td>
<td>70%</td>
<td>77–93%</td>
</tr>
<tr>
<td>Ages and Stages Questionnaire: Social Emotional (ASQSE) (10–15 min if interview needed)</td>
<td>6–60 months</td>
<td>A separate 30 item form for each of eight visits between 6 and 60 months. Items focus on self-regulation, compliance, communication, adaptive functioning, autonomy, affect and interaction with people</td>
<td>Single cutoff score indicating if referral indicated</td>
<td>71–85% according to age level</td>
<td>90–98% according to age level</td>
</tr>
<tr>
<td>Brief infant toddler social emotional assessment (BIT-SEA) (5–7 min)</td>
<td>12–36 months</td>
<td>42 items</td>
<td>Cutoffs based on child’s age and gender indicate presence or absence of problems and competence</td>
<td>80–85%</td>
<td>75–80%</td>
</tr>
<tr>
<td>Eyberg Child behavior Inventory/Sutter-Eyberg Student behavior Inventory (7 min if interview needed)</td>
<td>2–16 years</td>
<td>36–38 short statements. More than 16 suggests referral for behavioral interventions. If less than 16, can serve as guide for in-office counselling</td>
<td>A refer/nonrefer score for each behavioral domain (e.g. aggressivity, conduct)</td>
<td>80%</td>
<td>86%</td>
</tr>
<tr>
<td>Pediatric Symptom Checklist (7 min if interview needed)</td>
<td>4–16 years</td>
<td>35 short statements including externalizing (e.g. aggression) and internalizing (e.g. depression) behaviors. Scores ≥ 28 suggest referral</td>
<td>A single refer/nonrefer score</td>
<td>80–95%</td>
<td>68–100%</td>
</tr>
<tr>
<td>Modified Checklist of Autism in Toddlers (M-CHAT) (5 min)</td>
<td>16–52 months</td>
<td>24 items</td>
<td>If answer yes to two critical items or any three items, suggests referral</td>
<td>96%</td>
<td>94%</td>
</tr>
</tbody>
</table>
PEDS-Developmental Milestones

The PEDS-DM used the samples for the Brigance Inventory of Early Development–II (IED–II) and the Brigance Comprehensive Inventory of Basic Skills–Revised (CIBS–R) \((n = 1619)\) and was normed on the number of failing items below the 16th percentile \((\text{across ages sensitivity} = 83\% \text{ and specificity} = 84\%)\) (Brothers, Glascoe & Robertshaw 2008). [An example of a PEDS and PEDS:DM case is in Supplemental Table 1].

Modified Checklist for Autism In Toddlers

Robins studied 4797 children during toddler checkups with the Modified Checklist for Autism in Toddlers (M-CHAT) and the M-CHAT Follow-Up Interview. The difference between screening and actual diagnosis is shown by the fact that of the 4797 children, 466 screened positive on the M-CHAT, and of the 362 who completed the follow-up interview to reduce false positives, 61 continued to show risk factors for autism spectrum disorders. However, of the 41 children who underwent further evaluation only 21 were diagnosed with ASD, and 17 with non-autism spectrum disorder delays (Robins 2008). This can represent a large output of screening energy.

Rourke Baby Record

The Canadian College of Family Physicians advises its members to use the Rourke Baby Record, and most Canadian family physicians may either use it or some general surveillance questions depending on their preferences and time availability. The authors of the Rourke searched Medline, PubMed and the American Academy of Pediatrics and by consensus assigned a strength of recommendation \((\text{good, fair, consensus})\) to each of the assessments or interventions suggested in the tool (http://www.rourkebabyrecord.ca). There are no recorded sensitivities, specificities or reliabilities for the Rourke Baby Record (Rourke et al. 2013). An analysis of the coverage of developmental domains by age shows that the coverage of domains is uneven and uneven across age groups. A survey suggests that it is used by 79% of Canadian family physicians (Rourke et al. 2009). To provide a common point of comparison between the two groups all family physicians in both groups were invited to use the Rourke Baby Record.

Patient Health Questionnaire 9 (PHQ-9)

This a validated general depression assessment tool for adolescents and adults. In adult populations the PHQ9 has a sensitivity > 88% and specificity > 88% as diagnosed by a mental health interview and the Structured Clinical Interview for DSM-IV (Berkule et al. 2014).

Recruitment

We chose the 18-month visit as all developmental problems should then be fully evident. At three family medicine clinics in or near Calgary, Alberta, parents of infants 18 months old were invited via letters and posters in the offices to participate in a study at their 18-month visit, and six family physicians participated. All physicians were in private practice within the Alberta Health Care system. The University of Calgary provided Ethics approval (Ethics ID 24165).

Randomization

For each practice, lists of infants due for 18-month check-ups were prepared, and the research assistant randomized them by computer to ‘usual care’ or evidence-based screening. A developmental pediatrician advised us that children aged 16–20 months could take the 18-month screen. Because of the differences in the interventions neither the participants nor family physicians could be blinded once the participants had been assigned.

Administration

The parents attended 10 min before the visit (one parent for each child attended and they were on time). The tools were completed on a computer in a private booth and were immediately wirelessly marked from the PEDStest.com site and either printed and given to the family physician or electronically forwarded to the physician on the electronic medical record. After the tools had been completed on the computer and marked, the physician and parent prioritized and discussed any concerns raised. The research associate was present throughout each visit and facilitated parental use of the screening forms as needed. All mothers except one were very comfortable with computers with minimal help. We did not ascertain how much time the regular office staff would have needed to devote to helping parents if the research assistant were not present. The charts were reviewed three months later to ascertain if any referrals had resulted from the screening visit. No incentives were given to parents.

Results

The infants were evenly balanced on gender \((52\% \text{ female})\) and 91% were term births. The usual care and evidence
based care groups were very similar in gestational age at birth and age at screening (17.84 and 17.59 months). They differed markedly in female gender (40%, 62%); the research assistant did all randomization by computer ahead of the sessions and no reason was found for this imbalance (Table 3).

Evidence-based tools group: PEDS

One child was identified as at high risk of developmental disabilities, 8 as moderate risk of disability and 7 as low risk of developmental disability and increased risk of mental health problems (total 62% of infants identified at some level of risk).

PEDS-Developmental Milestones

Twenty-two milestone failures were noted: gross motor (nine infants) receptive language (four), fine motor (six), social emotional (two) and expressive language (one) (there was >1 problem for some infants).

M-CHAT

Five children failed one question, two 2 questions and one 3 questions. Only five questions were failed: 10 (looks you in the eye for more than a second or two); 11 (oversensitive to noise); 13 (imitates you); 18 (unusual finger movements near face) and 22 (sometimes stares at nothing or wanders with no purpose). None of these infants failed the M-CHAT when the research assistant tested them with the longer screen. For several infants sound sensitivity was to high-pitched appliances (e.g. one child’s chair in the kitchen was next to a blender).

PHQ9

No mother scored as depressed on the PHQ9, but some mothers scored on the PHQ fatigue item if they had several children.

Physicians’ decisions

For three infants language was to be checked by the physician at 24 months, and no referrals were made.

Physicians’ decisions for the ‘usual care’ group

Four (16%) problems were identified: one child was referred for speech therapy, two to return to the physician to check language at 24 months and a mother to discuss her depression. Overall, parents in the evidence-based tools screening group found significantly more problems for discussion (Table 4, \(p = 0.002\)), but the physicians in both groups after discussion with the parents decided that most problems were not serious, and that only six needed follow-up with the family physician within three months (three in each group and one referral from the ‘usual care’ group to a speech therapist). For some children with indicated language problems who spent all day with non-English speaking grandparents or siblings usually spoke for them, the physician encouraged more opportunities for speaking, with a check-up in 3 months.

Discussion

The results of this study

The answer to the first research question is that evidence-based screening identified significantly more problems than usual care (Table 4, \(p = 0.002\)). However, there were no differences in

Table 4. Results

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Control group</th>
<th>Enhanced screening group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No problem identified</td>
<td>21 (84%)</td>
<td>11 (38%)</td>
</tr>
<tr>
<td>Problems identified</td>
<td>4 (16%)</td>
<td>18 (62%)</td>
</tr>
<tr>
<td>Total</td>
<td>25 (100%)</td>
<td>29 (100%)</td>
</tr>
</tbody>
</table>

Chi-Squared = 11.803, \(p = 0.001\) (two sided), df = 1; with continuity correction for 2 × 2 table Chi-Squared = 9.971, \(p = 0.002\) (two sided).

Table 3. Results: Gestational age at birth, age at screening and gender

<table>
<thead>
<tr>
<th></th>
<th>Usual care (N = 25)</th>
<th>Evidence-based screening (N = 29)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gestational age at birth</td>
<td>Gender (% female)</td>
</tr>
<tr>
<td></td>
<td>Age at screening (months)</td>
<td></td>
</tr>
<tr>
<td>32 weeks (n = 1)</td>
<td>8 at 17, 13 at</td>
<td>10(40%)</td>
</tr>
<tr>
<td>34 weeks (n = 2)</td>
<td>18, 4 at 19</td>
<td></td>
</tr>
<tr>
<td>term (n = 23)</td>
<td>Avg = 17.84</td>
<td></td>
</tr>
</tbody>
</table>
referral rates between the evidence-based and usual care groups, and rates were low from both groups. We have very few clues. The physicians’ comments on the children in the evidence-based care group are brief and do not comment on the problems identified by the screening tools. The comments by the research assistant (a researcher and former school secretary who had worked with many children’s problems) may provide more insight: she commented on characteristics of the children that may have reassured her they were doing well, particularly busy infants, and she may have anchored on these characteristics. Cox (Cox et al., 2010) examined 752 PEDS forms and found that 94.6% of parents identified at least one concern on the English form and 69.7% on the Spanish form. Twenty-per cent of parents’ written comments concerned behavioral problems, 14% medical and 9% expressive language, and 23.9% of the comments were not matched to the intent of the questions (e.g. a behavioral concern was entered for questions about motor skills). Cox concluded that: ‘The inappropriate developmental expectations, limited health literacy, and culturally distinct comments on the PEDS form reinforce the importance of using screening tools to enhance the care provided during visits but not to replace patient-provider communication.’ The results of the current research project suggest that clinicians: (1) override evidence-based screening results with informal judgment; and/or (2) need assistance understanding test results and making referrals when needed. Care coordination and improved information sharing to and from referral sources should help improve referral and enrolment rates.

The answer to the second research question is that this proof-of-principle study shows that parents can complete all four evidence-based forms on a computer in a family physician’s office (average 10 minutes completion time) before the appointment, with no disruption to office routines. The research assistant was present whenever infants were screened so the effect on office routines when the office staff initially would be responsible for setting up screening and communicating with parents is not known. The tools were accessed and immediately wirelessly marked from PEDStest.com. The 10 PEDS questions ask about all developmental areas and permitted mothers to voice concerns and queries, which the mother and family physician prioritized and discussed.

In the context of societies which do not have organized screening organizations

Education and health in Canada are the responsibility of the provinces and not the federal government, and there are no complete national or provincial screening programs. The experience of this researcher was that it was very hard to get any practices to participate in screening, even physicians well-known to the researcher, and the projects if initially accepted for a presentation stalled with the clinic director or office staff anticipating substantial more work. Solutions are national and provincial comprehensive initiatives, and the Canadian Family Medicine residency organization mandating a country-wide screening program taught in residency.

Analysing reasons why comprehensive screening programs are not implemented and why once implemented they do not result in appropriate referral rates (Table 5)

(a) Do screening tools add too much time to the interview?

Schonwald (Schonwald et al., 2009a) found no change in the interview length when PEDS was included. Hix-Small (Hix-Small et al., 2007) estimated that 6 min of staff time were required for the Ages and Stages, but four of these minutes were required for data entry into the EMR, and this could be eliminated if done initially on a computer.

(b) Do physicians take on board all the information from the screening tools and implement it?

Rates of referral are usually low and were also found in this project. Training in the understanding of evidence-based screening tools and expected rates of developmental and behavioral problems by age and from the social context of the child (ethnicity, poverty, family disorganization, parental addiction) and in-time monitoring of screening and referral rates are required to detect which physicians are insufficiently screening or not referring when required. Hix-Small (Hix-Small et al., 2007) found only a 54% return of Ages and Stages questionnaires by parents, and it will be important to provide several avenues for questionnaire completion (Internet, paper and on-line before the office visit).

(c) Do physicians understand that surveillance is less effective than screening?

Rydz (Rydz et al., 2006) found for 288 children in Québec that incorporating the physicians’ opinion regarding the developmental status of the child did not improve the accuracy of the ASQ and Child Development Inventory assessments.
For overworked and underfunded practices the best solution may be for the Regional health authority to train and deploy screening by dedicated teams.

### Table 5. Reasons why children may not be referred for further screening and possible solutions

<table>
<thead>
<tr>
<th>Reasons</th>
<th>Solutions</th>
</tr>
</thead>
</table>
| (a) Improve the quality of information obtained from the screening process | 1. Medical office assistant to assemble and review all screening results ahead of time  
2. Lengthen interview to 30 min  
3. Invite to follow-up interview if insufficient information obtained. |
| For dysfunctional families especially where detailed assessment of parenting skills is required, 15 min is too short a time to review screening test results, examine child, discuss with parent, diagnose and make a management plan | Conduct one section of the interview to obtain all significant negatives and positives (similarly to the completeness of a pre-surgical medical). After discussing the parental comments on the PEDS, one option is to repeat the 10 PEDS questions with the focus of comprehensively asking for all the skills the child has:  
Q2. Tell me all the words and sounds your child makes of any kind  
Q3. Tell me all the ways your child understands what you say  
Q4. Tell me all the ways your child uses his hands and fingers to do things  
Q5. Tell me all the ways your child uses his arms and legs  
Q6. Tell me all the ways your child behaves that you like or concern you  
Q7. Tell me all the ways how your child gets on with others  
Q8. Tell me all the ways your child has learned to do things for her/himself  
Q9. Tell me what preschool skills your child has learned. |
| Insufficient diagnostic information is obtained. Cox (2010) for 740 written parental PEDS comments found: (1) 20% were about behavior, 14% medical and 9% expressive language (each of which require detailed questioning), and (2) 23% of comments did not match the PEDS question (indicating parental misunderstanding requiring clarification). | Lack of knowledge about typical problem rates detected in screening. Unaware false positives are different from true negatives and have worse results when checked with diagnostic tests. Unaware casual surveillance questions detect far fewer problems than organized screening. |
| (b) Improved training of physicians and improved support for individual practices | Qualitative research is needed observing physicians as they ask questions about the child’s developmental process and then debrief the physicians and parents afterwards about their perceptions and decision-making. |
| Knowledge of screening. Rydz et al. (2006) for 288 children found that the accuracy of the ASQ and Child Development Inventory assessments was not improved by incorporating the physicians’ opinion about developmental status | The practices’ concerns were that screening would affect clinic flow, the need to distribute responsibilities among staff at multiple levels, capturing visits on computers, maintaining screening during busy times and staff turnover, tracking referrals and parental failure to follow-up on referrals. |
| There is insufficient understanding how physicians interpret developmental screens and why some decide to ignore the results and prefer their own observations | Lack of knowledge about typical problem rates detected in screening. Unaware false positives are different from true negatives and have worse results when checked with diagnostic tests. Unaware casual surveillance questions detect far fewer problems than organized screening. |
| Attitudes to screening. King 2009 found that 54 practices which were early enthusiastic joiners of the American Academy of Pediatrics quality improvement project had multiple concerns | Qualitative research is needed observing physicians as they ask questions about the child’s developmental process and then debrief the physicians and parents afterwards about their perceptions and decision-making. |
| (c) Acceptance by the regional health authority for overall responsibility for screening | The Regional health authority needs to take screening as a key priority, motivate staff, require high standard screening, provide training and monitor that rates of diagnoses and referral for individual health practitioners are similar to those achieved by the highest standard practitioners. |
| Rates of screening and referral are affected by the complexity of the child and family and require the help of an organized well-funded network. (Jimenez et al., 2014) in a logistic regression found that the multivariate adjusted odds ratio of referral depended on parental factors (maternal age ≥ 30, AOR = 0.81; African-American 0.68; single parent 0.75), child factors (two or more developmental domains 3.15; special health needs 3.16) and practitioner (developmental screening without additional staff support 0.44; ‘usual care’ 0.45) | Seamless referrals such that (1) physicians know exactly the progress and outcome of each referral and the next steps proposed, and (2) the referral is made so that parents do not have to make any phone calls, and the results of the referral and proposed therapies are discussed regularly with the parents to motivate them. |
| It may be necessary to Increase payment for screening to enable practices to meets the highest standards | Increased payment requires that random screenings are observed and independently rated by experts, and feedback provided by Skype, and indicated areas for improvement planned for monitoring on next screenings. |
| For overworked and underfunded practices the best solution may be for the Regional health authority to train and deploy screening by dedicated teams | The regional health authority would mandate such teams to identify each infant requiring screening and conduct screening of the highest standard. |
(d) Do physicians understand that false positives are different than true negatives and more likely to perform adversely on diagnostic screens?

Tarini (Tarini et al., 2011) found 818 false positive newborn screens in a cohort of 49,959 births: the pretermers had more acute outpatient visits than the term births, and chronic illness or parental anxiety was queried. Glascoe (2013) in a national sample of 512 infants found that the false positives when tested with diagnostic measures had a 6.7 relative risk of lower scores for intelligence, 4.9 academic achievement, 3.1 language and 2.6 adaptive behavior. When criteria from the US Individuals with Disabilities Education Act were applied to predict which children would have difficulty benefitting from usual classroom instruction, 70% of those with false positive scores and 29% of the true negatives were predicted to have such difficulties.

(e) Do physicians understand that assessing parenting behaviors is a key component of assessing the child’s environment?

Glascoe (Glascoe & Leew, 2010) for 382 children from the Brigance Infant and Toddler Screens standardization and validation study identified that six positive parenting behaviors predicted average to above-average development on the Brigance screens but <2 positive parenting behaviors and negative perceptions of children indicated performance 2 SDs below the mean. Psychosocial risk factors included >3 children in the home, multiple moves, limited English and parental depression.

(f) Are explanation and persuasion enough?

Schoenwald (Kerstjens et al., 2009) at two clinics associated with the Boston Children’s Hospital found only 61% of children received a PEDS evaluation despite detailed explanations and training (Schonwald et al., 2009b). Much higher rates would be expected at an elite academic clinic. Fifty-four US pediatric practices applied to participate in an American Academy of Pediatrics quality improvement pilot to improve developmental surveillance and screening. The author concluded: ‘At the project’s conclusion practices reported screening more than 85% of patients presenting at recommended screening ages. They achieved this by dividing responsibilities among staff and actively monitoring implementation. Despite these efforts, many practices struggled during busy periods and times of staff turnover. Most practices were unable or unwilling to adhere to 3 specific AAP recommendations: to implement a 30-month visit; to administer a screen after surveillance suggested concerns; and to submit simultaneous referrals both to medical subspecialists and local early-intervention programs. Overall practices reported referring only 61% of children with failed screens. Many practices struggled to track their referrals. Those that did found that many families did not follow through with recommended referrals.’ (King et al., 2010) Presumably these 54 enthusiastic practices had carefully discussed benefits of the project and the workload, yet were not able to achieve their goals.

(g) Does the referral system work seamlessly?

Major problems in referral systems that are not working optimally are that (1) clinicians may not know where to refer children when problems are found, (2) referral sources may not keep primary providers informed of children’s status, whether the parents attend the offered resources, and whether the problems have been solved.

Possible solutions (Table 5)

(a) Improve the quality of information obtained from the screening process

With dysfunctional families with inadequate parenting skills it will be necessary to lengthen the interview. Moreover, Cox (Cox et al., 2010) showed that 23% of parental comments on the PEDS were mismatched to the question so one solution would be to go through each PEDS question with the parents and adopt a skills approach and ask the parent to tell the physician the child’s skills in relationship to each question.

(b) Improved training of physicians and improved support for individual practices

Improved training of physicians to understand screening results and the benefits of referral are needed, and detailed administrative help from and regional health authorities to individual practices to help staff implement screening programs so that the burden on the practices is relieved. Qualitative research watching physicians as they ask questions about the child’s developmental process and then debriefing them afterwards about their perceptions and decision-making would be invaluable.
(c) Acceptance by the regional health authority for overall responsibility for screening

The ideal solution would be for the regional health authority to use its budget and staff and create an organized network for helping practices implement screening, implement a seamless referral and referral tracking system, provide increased finance to reward superior screening and referral achievements and, if necessary, train teams to screen in those practices which are so overburdened that they cannot facilitate screening.

(d) In summary

The problems noted over a wide range of screening projects are: (1) the need for uniform system-wide referral methods, (2) strongly involving parents so they understand what the referral is expected to achieve, (3) how to cope with the chaos in referrals caused by overwork and staff turnover, (4) seamless completed referrals so that (a) the physician knows the progress and outcome of each referral, and (b) parents are not simply given a phone number and have to negotiate the bureaucracy and call in a language that they do not understand. This leads to the conclusions that there is no substitute for mandatory participation in screening, detailed monitoring and consequences. These can only be expected if appropriate financing and support are supplied by a central authority. High-risk areas of medical practice tend to have mandatory safety routines, and the overall medical authorities have not yet stepped up to completing mandatory monitoring of child screening as a key factor affecting children’s future welfare and life chances.

Directions for future research

(1) Qualitative studies are a top priority because we have so little description of direct observations how physicians interact with children and parents during screening, how they discuss their findings with the parents and how they decide there are problems and whether referrals are needed; (2) identify the most complete training programs that teach how to use and interpret screening tools, and improve and implement those training programs; and (3) the ASQ and PEDS give different results. Sices (Sices et al., 2009) found that 37% of children failed the PEDS and 27% the ASQ, 52% passed both and 15% failed both. Limbos & Joyce (2011) found the PEDS had 74% sensitivity and 64% specificity, and the ASQ 82% and 78%. Sensitivity of the PEDS was 89% for >2 predictive concerns (specificity 41%) and for the ASQ 94% for >2 abnormal domains (specificity 47%). Large studies comparing the ASQ and PEDS with optimum training and qualitative studies of the interview process are required to identify the optimum tools for screening. (4) Identify the most successful integrated screening and referral systems and their results.

Key messages

- Detection rates of developmental problems at the 18-month well-baby visit by family physicians using four evidence-based screening tools are much higher than with ‘usual care.’
- Mothers offered the opportunity to state in open text concerns about 10 developmental domains were enabled to identify many more concerns.
- Four evidence based tools (the Parents Evaluation of Developmental Status (Peds), the Peds-Developmental Milestones (Peds-Dm), the Modified Checklist for Autism in Toddlers (M-Chat) and PHQ9 (maternal depression)) can be answered on a computer by parents within 10 min before their child’s appointment.
- Although many more problems were identified using evidence-based tools, the family physicians after discussion with the parents overrode the screening tools and identified few children in each group who needed follow-up within 3 months.
- Possible solutions are (1) to improve the quality of information obtained from the screening process by training, (2) improved training of physicians in understanding evidence-based screening tools, (3) qualitative studies on how physicians understand evidence-based screens, and the discussion between physicians and patients about which problems have been identified and need referral, (4) improved support for individual practices and (5) acceptance by the regional health authority for overall responsibility for screening and creation of a comprehensive network.

Acknowledgments

Dr. Beverly Gold, Dr. Ian Kendall and Dr. Trevor Mailey (Heritage Pointe Family Practice, De Winton, Alberta); Dr. Sandy Lahey and Dr. Jennifer Corales (Lahey Medical Clinic, Cochrane, Alberta); and Dr Gloria Mazloum (Foothills Family Practice, Okotoks, Alberta) for participating in screening their
18 month olds. Dr. Ben Gibbard, MD, MCS, MSc., FRCPC, Les Cunning MD and Shirley Leew, Ph.D, SLP(C), RSLP for much advice and discussion how to screen children in this project.

References


Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Supplemental Table 1: Using Parents’ Evaluation of Developmental Status (PEDS) and PEDS: Developmental Milestones (PEDS:DM): A Case Example