Published in final edited form as:

Ambul Pediatr. 2006 ; 6(6): 347–351.

Parent and teacher mental health ratings of children using primary care services: Inter-rater agreement and implications for mental health screening

Jonathan D. Brown, PhD, MHS and Lawrence S. Wissow, MD, MPH

Anne Gadomski, MD, MPH

Bassett Research Institute, Bassett Healthcare, Cooperstown, NY

Ciara Zachary, BS

Department of Health, Behavior, and Society, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Edward Bartlett, MD

Johns Hopkins Community Physicians, Baltimore, MD

Ivor Horn, MD, MPH

Children's National Medical Center, Washington, DC

Abstract

Objective—To examine inter-rater agreement when screening for child mental health problems during primary care visits.

Design and Methods—Children age 5-10 ($n = 227$) and one of their parents were systematically recruited from the waiting rooms of 15 primary care sites located in Baltimore, MD, Washington, DC, and rural New York from 2002-2005. The parent and teacher of the child completed the Strengths and Difficulties Questionnaire to measure the child's emotions, behaviors, and functional impairment.

Results—Parents and teachers identified a similar proportion of children as having high symptoms (25% vs. 23%) and high impairment (27% vs. 32%) but rarely agreed in their assessments of specific children. Parent ratings alone missed 52% of children rated by teachers as having both high symptoms and high impairment ($\kappa = 0.15$). Only 6% of these discrepant visits were for mental health problems, making it unlikely that teacher reports would have been solicited.

Conclusion—Parent reports failed to detect half of school-aged children considered to be seriously disturbed by their teachers. Efforts to improve detection of mental health problems using screening tools in primary care may require algorithms that help providers judge when to solicit teacher reports and how to interpret conflicting information from parents and teachers.

Keywords

mental health screening; primary care; agreement

Correspondence to: Lawrence S. Wissow.

Corresponding author: Jonathan D. Brown, PhD, MHS, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway Room 753 Baltimore, MD 21205 Phone: 410-274-2323, Fax 410-955-7241; Email: jobrown@jhsph.edu. **Request for reprints:** Lawrence S. Wissow, MD, MPH, Professor, Department of Health, Behavior, and Society, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway Room 703 Baltimore, MD 21205 Phone: 410-614-1243, Fax: 410-955-7241 Email: lwissow@jhsph.edu.

Sources of support: This work was supported by NIMH grants R01MH062469 (Lawrence Wissow) and F31MH75531 (Jonathan Brown).

Conflict of interest: The authors have no conflict of interest.

Introduction

Researchers and clinicians have proposed that routine screening during primary care visits may improve the identification of children's mental health problems.¹⁻⁴ One potential problem with screening school-aged children in primary care is that ratings are often only gathered from the parent present at the visit.⁴ Teachers are usually consulted only if the parent expresses a concern about behavior or achievement in school.⁵

Obtaining mental health ratings from a single informant is inconsistent with findings from community-based studies, which have demonstrated that raters often do not agree in their assessments of child mental health.⁶⁻¹⁴ These discrepancies are not necessarily errors in observation. Rather, when parent and teacher reports have been compared to independent psychiatric assessments, the combination of reports is more sensitive than either alone.¹⁵ Thus, discrepancies may reflect that children's symptoms, or the opportunities to observe them, vary across contexts.¹⁶ Obtaining teacher reports may be particularly important for children, given the high correlation of school difficulties with future mental health problems and social dysfunction.¹⁷

The implications of inter-rater agreement for screening in primary care have not been extensively explored. Inter-rater agreement in primary care might be better than in community samples because the prevalence of mental health problems among children in primary care is higher than in the general population, and detection generally improves with severity.¹⁸⁻²⁰ Further, visits may be initiated because an adult suspects that the child has a mental health problem.²¹

This investigation takes advantage of data collected for another study to examine parent-teacher agreement on mental health screening reports about children seen in primary care. The data represent what would result if PCPs systematically obtained mental health screening reports from both the teacher and parent. We hypothesized that parent-teacher agreement would be better than in community samples, but that, consistent with community studies, teachers would report more externalizing symptoms and impairment compared to parents.¹⁵

Methods

Participants and Procedures

The data were originally collected to assess whether the youth was eligible to participate in an intervention to improve the psychosocial communications skills of PCPs. Data were collected from 15 primary care offices in Baltimore, MD, Washington, DC, and rural New York which were chosen to represent the range of specialties, payers, locations, and practice structures that provide pediatric primary care in the United States.²²⁻²⁵ All sites served patient populations with a mix of insurance types and who were within the age of infant to 18 years.

Fifty-four PCPs participated; 81.4% ($n=44$) medical doctors, 16.6% ($n=9$) nurse practitioners, and 1.8% ($n=1$) physician assistant. Forty percent were male, 64.8% were pediatricians, and 35.2% were family practitioners.

Recruitment of families took place from December 2002 to August 2005. Research interviewers systematically approached all parents in the waiting area with the goal of recruiting 10 children for each PCP. Children were eligible if they were between 5 -16 years old, scheduled to see a participating PCP, and reported pain as 4 or less on an analog scale of 1-10, in which 1 was no pain. The researcher randomly selected one child per family to participate if more than one child was eligible.

The parent provided written consent and was asked if a researcher could contact their child's primary school teacher. The parent provided the name of the school and teacher. While in the waiting area, the parent completed a questionnaire about the child's mental health (described below). The parent also reported whether the visit was for a mental health problem, well-child visit, acute illness, or follow-up of a medical problem. The parent was compensated \$15 for participation.

The teacher identified by the parent was informed by mail that the child was participating in a research project to teach doctors skills and was asked to complete a questionnaire that would not be included in the child's medical record (described below). No information about the teacher was collected. All procedures were approved by the Johns Hopkins Bloomberg School of Public Health Committee on Human Research and the research review committees of participating sites.

Measure

The parent and teacher completed the Strengths and Difficulties Questionnaire (SDQ), a measure of children's emotions, behaviors, and functional impairment that has been widely used to screen for mental health problems in research and clinical settings.²⁶⁻²⁷ The SDQ is available in over 40 languages and has been validated as a screening tool for children age 4-17 among community-based and clinic populations in the United States, United Kingdom, Australia, South America, Asia and many European countries.²⁸⁻³³ Versions are available for parent, teacher, and teen self-report (www.sdqinfo.com).

The SDQ consist of 5 scales (emotion, conduct, inattention-hyperactivity, peer problems, and prosocial behavior), each composed of 5 items which measure the extent to which the item describes the child during the past 6 months using a 3-point Likert scale. Items are scored 0 = "not true," 1 = "somewhat true," and 2 = "certainly true" so that each scale ranges from 0-10. The emotion, conduct, inattention-hyperactivity, and peer problems scales are summed to create the Total Difficulties score, which ranges from 0-40.

Five items inquire whether the child has difficulties in four areas (emotions, concentration, behavior or getting along with others) and whether such difficulties interfere with home life, friendships, classroom learning, and leisure activities. These items are scored 0 = "not at all," 1 = "little," 2 = "medium amount," and 3 = "great deal," to yield an Impairment scale score of 0-10.³⁴

U.S. normative data are based on parent reports of 9,878 children age 4-17 years, who completed the SDQ during the 2001 National Health Interview Survey (NHIS).³⁴ In the NHIS, scores in what was designated as the high range of each scale identified children as belonging to the worse 7-10% of the normative population depending on the domain.³⁴ In the NHIS, children who scored within the high range of the Total Difficulties scale were 12.2 times (95% CI, 9.6 -15.4) more likely to have received mental health services from the general medical sector during the past year.³⁴

The Total Difficulties and Impairment scales demonstrated excellent internal consistency in the NHIS investigation (Cronbach's $\alpha = 0.83$ and 0.80 respectively).³⁴ Internal consistency was also good for the conduct, emotions, inattention-hyperactivity, and prosocial behavior scales (Cronbach's $\alpha = 0.63$ to 0.77) but not peer problems (Cronbach's $\alpha = 0.46$).³⁴ Similar internal consistency has been found among European samples.³⁵⁻³⁶ In our sample, all parent-reported symptom scales demonstrated good to excellent internal consistency (Cronbach's $\alpha = 0.63$ to 0.82), as did the Impairment scale (Cronbach's $\alpha = 0.85$). Teacher-reported symptom scales also demonstrated good internal consistency (Cronbach's $\alpha = 0.72$ to 0.85), as did the Impairment scale (Cronbach's $\alpha = 0.79$).

Low levels of agreement between parent and teacher ratings (Pearson's $r = 0.39$) have been found among samples in the United Kingdom, which was similar to findings among populations in Germany and the Netherlands.³⁶⁻³⁸

Statistical Analysis

SDQ scales were summed using the published algorithm (www.sdqinfo.com). SDQ scores for each domain were categorized as high according to the NHIS normative cut-points.³⁴ As in the NHIS investigation, we created a dichotomous variable to categorize children who were rated as having both a high Total Difficulties score and a high Impairment score. These would be children likely to meet diagnostic criteria for a psychiatric disorder in our sample.

We examined agreement between parent and teacher ratings using proportions and kappa. All analyses were conducted using Stata 9.³⁹

Results

Children and Teachers

This report focuses on children age 5-10 because the teacher completed the SDQ only for this age group. Of the 871 children initially screened for the original study, 291 were between the ages of 5-10, met eligibility requirements, and were recruited during the school year. Eighty-six percent ($n = 253$) of those parents consented to have a researcher contact their child's teacher. There were no differences in race, gender, or insurance status between the children of parents who consented to have a researcher contact the teacher and those who did not (results not shown). Teachers returned the SDQ for 227 children (89.7% of teachers contacted) for an overall response rate of 78.0%. Of the 227 children, 44.9% were female, 71.8% Caucasian, 19.3% African-American, 7.0% Other, and 1.7% American Indian. Ten parents (4.4%) identified as Hispanic and completed the Spanish language SDQ. There were no differences in race, gender, or insurance status between the children with returned teacher SDQ ratings compared with those without a teacher rating (result not shown). Half of visits (51.1%) were well-child appointments, 29.1% acute illness, 11.8% follow-up of a medical problem, and 7.9% were for a mental health problem.

Agreement between parent and teacher ratings

The group of parents and teachers both identified 17.6% of children as having a high Total Difficulties score and a high Impairment score, that is, children who would be most likely to meet diagnostic criteria in our sample (Table 1). However, parents and teachers did not identify the same children. Twenty-six percent of children were rated as having high scores on both the Total Difficulties and Impairment scales by *either* the parent or teacher, but only 9.3% were rated as such by *both* the parent and teacher. Among children rated as having both a high Total Difficulties score and high Impairment score by at least one observer, parents did not identify 52.5% of children who were identified by the teacher ($k = 0.15$). Of the children who the teacher, but not the parent, rated as having high Total Difficulties and high Impairment, only 2 parents reported that the visit was for their child's mental health problem, making it unlikely that teacher reports would have been solicited.

Agreement was best for children with attention or conduct problems ($k = 0.42$ and 0.35 , respectively). Teachers reported a higher proportion of children with attention problems while parents reported a higher proportion with conduct problems. The worst agreement was for emotional problems and low prosocial behavior ($k = 0.25$ and 0.18 , respectively). Teachers reported a higher proportion of children as having low prosocial behavior and parents reported a higher proportion as having high emotional symptoms.

Discussion

In our primary care sample, the prevalence of mental health symptoms and impairment was consistent with previous investigations in primary care settings, but parent-teacher agreement was no better than has been observed in community samples that used the SDQ.^{28,36,38,40-42} Our results parallel those from a British population-based study, which found that combining parent and teacher reports on the SDQ detected 62% of psychiatric disorders among children age 5-10, compared with 30% for parent report alone.²⁸

If PCPs were to routinely gather information from both the parent and teacher, the prevalence of children requiring further evaluation would be much higher than if only the parent completed the screening instrument. One alternative would be to ask routinely for teacher reports when the parent brings the child for evaluation of a mental health concern. However, in our sample, this would have led to getting additional information in only 2 of the cases where teachers suspected disorder-level problems but the parent did not.

Limitations

Our study is limited in that we do not have a way of assessing the child's "true" mental health status. However, prior studies with the SDQ and the Ontario Child Health Study Scales have established that parent reports alone underestimate the prevalence of disorder as measured independently with diagnostic instruments.^{6,14} Although the SDQ has not been used extensively in clinical settings in the United States, it has been adopted as a component of the National Health Interview Survey, where it has shown strong associations with the use of mental health services.³⁴

Conclusion

Routine screening using only parent reports is likely to under-identify symptoms and functional problems that would be uncovered if reports were also solicited from teachers. Gathering a mental health rating from the teacher, however, greatly increases the prevalence of children needing further evaluation. Efforts to improve detection of mental health problems in primary care may require development of algorithms that help providers judge when to solicit teacher reports and how to interpret contrasting information from parents and teachers.

Acknowledgements

We wish to thank the families and the staff of the clinics participating clinics. Our research team included: Carmen Ivette Diaz, Mark Celio, O'Neil Costley, Xianghua Luo, Lucia Martinez, Nancy Tallman, Alexandra Suchman, Mei-Chen Wang, Nancy Weissflog. This work was supported by NIMH grants R01MH062469 and F31MH75531.

References

1. Frankenfield DL, Keyl PM, Gielen A, Wissow LS, Werthamer L, Baker SP. Adolescent patients--healthy or hurting? Missed opportunities to screen for suicide risk in the primary care setting. *Arch Pediatr Adolesc Med* 2000;154:162-168. [PubMed: 10665603]
2. Asarnow JR, Jaycox LH, Anderson M. Depression among children in primary care: Models for delivering mental health services. *Child Adolesc Psychiatr Clin N Am* 2002;11:477-497. [PubMed: 12222079]
3. Brugman E, Reijneveld SA, Verhulst FC, Verloove-Vanhorick SP. Identification and management of psychosocial problems by preventive child health care. *Arch Pediatr Adolesc Med* 2001;155:462-469. [PubMed: 11296073]
4. Borowsky IW, Mozayeny S, Ireland M. Brief psychosocial screening at health supervision and acute care visits. *Pediatrics* 2003;112:129-133. [PubMed: 12837878]

5. Reiff MI, Stein MT. Attention-deficit/hyperactivity disorder evaluation and diagnosis: A practical approach in office practice. *Pediatr Clin North Am* 2003;50:1019–1048. [PubMed: 14558680]
6. Offord DR, Boyle MH, Racine Y, Szatmari P, Fleming JE, Sanford M, et al. Integrating assessment data from multiple informants. *J Am Acad Child Adolesc Psychiatry* 1996;35:1078–1085. [PubMed: 8755805]
7. Holmbeck GN, Li ST, Schurman JV, Friedman D, Coakley RM. Collecting and managing multisource and multimethod data in studies of pediatric populations. *J Pediatr Psychol* 2002;27:5–18. [PubMed: 11726675]
8. Jensen PS, Rubio-Stipec M, Canino G, Bird HR, Dulcan MK, Schwab-Stone ME, et al. Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *J Am Acad Child Adolesc Psychiatry* 1999;38:1569–1579. [PubMed: 10596258]
9. Jellinek MS, Murphy JM, Little M, Pagano ME, Comer DM, Kelleher KJ. Use of the pediatric symptom checklist to screen for psychosocial problems in pediatric primary care: A national feasibility study. *Arch Pediatr Adolesc Med* 1999;153:254–260. [PubMed: 10086402]
10. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987;101:213–232. [PubMed: 3562706]
11. Foley DL, Rutter M, Angold A, Pickles A, Maes HM, Silberg JL, et al. Making sense of informant disagreement for overanxious disorder. *J Anxiety Disord* 2005;19:193–210. [PubMed: 15533704]
12. MacLeod RJ, McNamee JE, Boyle MH, Offord DR, Friedrich M. Identification of childhood psychiatric disorder by informant: Comparisons of clinic and community samples. *Can J Psychiatry* 1999;44:144–150. [PubMed: 10097834]
13. Shemesh E, Newcorn JH, Rockmore L, Shneider BL, Emre S, Gelb BD, et al. Comparison of parent and child reports of emotional trauma symptoms in pediatric outpatient settings. *Pediatrics* 2005;115:582–590.
14. Weissman MM, Wickramaratne P, Warner V, John K, Prusoff BA, Merikangas KR, et al. Assessing psychiatric disorders in children. Discrepancies between mothers' and children's reports. *Arch Gen Psychiatry* 1987;44:747–753. [PubMed: 3632247]
15. Goodman R, Renfrew D, Mullick M. Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *Eur Child Adolesc Psychiatry* 2000;9:129–134. [PubMed: 10926063]
16. Stein MT. When parents and teachers disagree about a child's behavior: An opportunity for further evaluations. *J Dev Behav Pediatr* 2004;25:50–51. [PubMed: 14767357]
17. DeSocio J, Hootman J. Children's mental health and school success. *J Sch Nurs* 2004;20:189–196. [PubMed: 15283617]
18. Briggs-Gowan MJ, Horwitz SM, Schwab-Stone ME, Leventhal JM, Leaf PJ. Mental health in pediatric settings: distribution of disorders and factors related to service use. *J Am Acad Child Adolesc Psychiatry* 2000;39:841–849. [PubMed: 10892225]
19. Costello EJ, Costello AJ, Edelbrock C, Burns BJ, Dulcan MK, Brent D, Janiszewski S. Psychiatric disorders in pediatric primary care. Prevalence and risk factors. *Arch Gen Psychiatry* 1988;45:1107–1116. [PubMed: 3264146]
20. Robinson JW, Roter DL. Psychosocial problem disclosure by primary care patients. *Soc Sci Med* 1999;48:1353–1362. [PubMed: 10369436]
21. Costello EJ, Pescosolido BA, Angold A, Burns BJ. A family network-based model of access to child mental health services. *Research in Community and Mental Health* 1998;9:165–190.
22. Freed GL, Nahra TA, Wheeler JR. Which physicians are providing health care to America's children? trends and changes during the past 20 years. *Arch Pediatr Adolesc Med* 2004;158:22–26. [PubMed: 14706953]
23. Cherry DK, Woodwell DA. National Ambulatory Medical Care Survey: 2000 summary. *Adv Data* 2002;328:1–32. [PubMed: 12661586]
24. Rushton JL, Clark SJ, Freed GL. Pediatrician and family physician prescription of selective serotonin reuptake inhibitors. *Pediatrics* 2000;105:957–962. [PubMed: 10742354]

25. Simpson GA, Bloom B, Cohen RA, Blumberg S, Bourdon KH. U.S. children with emotional and behavioral difficulties: Data from the 2001, 2002, and 2003 National Health Interview Surveys. *Adv Data* 2005;360:1–13. [PubMed: 16004071]
26. Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiatry* 1999;40:791–799. [PubMed: 10433412]
27. Vostanis P. Strengths and Difficulties Questionnaire: Research and clinical applications. *Curr Opin Psychiatry* 2006;19:367–372. [PubMed: 16721165]
28. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001;40:1337–1345. [PubMed: 11699809]
29. Marzocchi GM, Capron C, Di Pietro M, Duran Tauleria E, Duyme M, Frigerio A, et al. The use of the Strengths and Difficulties Questionnaire (SDQ) in southern European countries. *Eur Child Adolesc Psychiatry* 2004;13:40–46.
30. Garcia-Cortazar P, Mazaira JA, Goodman R. The initial validation study of the Gallego version of the Strengths and Difficulties Questionnaire. *Revista de Psiquiatria Infanto-Juvenil* 2000;2:95–100.
31. Hawes DJ, Dadds MR. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Aust N Z J Psychiatry* 2004;38:644–651. [PubMed: 15298588]
32. Rothenberger A, Woerner W. Strengths and Difficulties Questionnaire (SDQ)-evaluations and applications. *Eur Child Adolesc Psychiatry* 2004;13(Suppl 2):II1–II2. [PubMed: 15243779]
33. Woerner W, Fleitlich-Bilyk B, Martinussen R, Fletcher J, Cucchiaro G, Dalgalarondo P, et al. The Strengths and Difficulties Questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. *Eur Child Adolesc Psychiatry* 2004;13(Suppl 2):II47–II54. [PubMed: 15243786]
34. Bourdon KH, Goodman R, Rae DS, Simpson G, Koretz DS. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *J Am Acad Child Adolesc Psychiatry* 2005;44:557–564. [PubMed: 15908838]
35. Ronning JA, Handegaard BH, Sourander A, Mørch WT. The Strengths and Difficulties Self-report Questionnaire as a screening instrument in Norwegian community samples. *Eur Child Adolesc Psychiatry* 2004;13:73–82. [PubMed: 15103532]
36. van Widenfelt BM, Goedhart AW, Treffers PD, Goodman R. Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *Eur Child Adolesc Psychiatry* 2003;12:281–289. [PubMed: 14689260]
37. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry* 1998;7:125–130. [PubMed: 9826298]
38. Muris P, Meesters C, van den Berg F. The Strengths and Difficulties questionnaire (SDQ)--further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *Eur Child Adolesc Psychiatry* 2003;12:1–8. [PubMed: 12601558]
39. StataCorp.. *Stata Statistical Software: Release 9*. StataCorp LP; College Station, TX: 2005.
40. Bergeron L, Valla JP, Breton JJ. Pilot study for the Quebec child mental health survey: Part II. Correlates of DSM-III-R criteria among six to 14 year olds. *Can J Psychiatry* 1992;37:381–386. [PubMed: 1394013]
41. Costello EJ, Angold A, Burns BJ, Erkanli A, Stangl DK, Tweed DL. The Great Smoky Mountains study of youth. Functional impairment and serious emotional disturbance. *Arch Gen Psychiatry* 1996;53:1137–1143. [PubMed: 8956680]
42. Jensen PS, Watanabe HK, Richters JE, Cortes R, Roper M, Liu S. Prevalence of mental disorder in military children and adolescents: Findings from a two-stage community survey. *J Am Acad Child Adolesc Psychiatry* 1995;34:1514–1524. [PubMed: 8543520]

Table 1

Parent and teacher ratings of children age 5-10 years ($n = 227$)

Score	Parent ratings	Teacher ratings	Either parent or teacher	Parent and teacher agree	Teacher concerns missed if only parent screened	kappa (κ)
High Total difficulties	25.4% 58/227	22.9% 52/227	34.3% 78/227	14.0% 32/227	38.4% 20/52	0.44
High Impairment	27.3% 62/227	30.4% 69/227	42.2% 96/227	14.9% 34/227	50.7% 35/69	0.31
High Total difficulties and high Impairment	17.6% 40/227	17.6% 40/227	26.4% 60/227	9.3% 21/227	52.5% 21/40	0.15
High emotional	19.3% 44/227	11.8% 27/227	25.5% 58/227	5.7% 13/227	51.8% 14/27	0.25
High conduct	22.4% 51/227	15.8% 36/227	29.0% 66/227	9.2% 21/227	41.6% 15/36	0.35
High inattention-Hyperactivity	22.0% 50/227	29.9% 68/227	37.0% 84/227	14.9% 34/227	50.0% 34/68	0.42
High peer problems	17.6% 40/227	21.1% 48/227	29.9% 68/227	8.8% 20/227	58.3% 28/48	0.31
Low prosocial behavior	11.8% 27/227	22.0% 50/227	28.6% 65/227	5.2% 12/227	76.0% 38/50	0.18