Screening for Childhood Mental Health Problems: Outcomes and Early Identification

Marilyn J. Essex, PhD[1], **Helena C. Kraemer, PhD**[2], **Marcia J. Slattery, MD, MHS**[1], **Linnea R. Burk, PhD**[1], **W. Thomas Boyce, MD**[3], **Hermi R. Woodward, BS**[4], and **David J. Kupfer, MD**[4]

1*Department of Psychiatry, University of Wisconsin-Madison*

2*Department of Psychiatry and Behavioral Sciences, Stanford University*

3*College for Interdisciplinary Studies and Faculty of Medicine, University of British Columbia*

4*Department of Psychiatry, University of Pittsburgh*

Abstract

Background—Many childhood psychiatric problems are transient. Consequently, screening procedures to accurately identify children with problems unlikely to remit and thus, in need of intervention, are of major public health concern. This study aimed to develop a universal school-based screening procedure based on the answers to three questions: (1) What are the broad patterns of mental health problems from kindergarten to grade 5? (2) What are the grade 5 outcomes of these patterns? (3) How early in school can children likely to develop the most impairing patterns be identified accurately?

Methods—Mothers and teachers reported on a community sample (N=328) of children's internalizing and externalizing symptoms in kindergarten and grades 1, 3, and 5. In grade 5, teachers reported on children's school-based functional impairments, physical health problems, and service use; mothers reported on children's specialty mental health care.

Results—Four patterns distinguished children who (1) never evidenced symptoms; (2) evidenced only isolated symptoms; or evidenced recurrent symptoms, either (3) without or (4) with comorbid internalizing and externalizing. By grade 5, children with recurrent comorbid symptoms had the greatest impairments, physical health problems, and service use. These children can be identified quite accurately by grade 1.

Conclusions—Universal screening at school entry can effectively identify children likely to develop recurrent comorbid symptoms, and would provide a basis for developing optimal targeted intervention programs.

Keywords

universal screening; childhood; mental health problems; impairments; longitudinal

---

Childhood mental health problems, including internalizing and externalizing, have high point prevalence. Although rates vary by definitional criteria (e.g., categories v. dimensions), assessment method (e.g., semi-structured interview v. questionnaire), and informant (e.g., parent v. teacher), at any given time, approximately 20% of children evidence mental health

---

Corresponding Author Dr. Marilyn J. Essex, University of Wisconsin Department of Psychiatry, 6001 Research Park Boulevard, Madison, WI, 53719. Phone: 608-263-7887. Fax: 608-265-4008. Email: E-mail: mjessex@wisc.edu.

problems with at least minimal functional impairment, 10% to 15% evidence more severely impairing psychiatric disorders (U. S. Department of Health and Human Services, 1999), and by adolescence, approximately 40% will have met criteria for a psychiatric diagnosis at least once (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Kim-Cohen et al., 2003). These rates increase further when they include children with sub-threshold diagnostic symptoms, many of whom are significantly impaired (Angold, Costello, Farmer, Burns, & Erkanli, 1999; Leaf et al., 1996). Such high rates are of concern because childhood problems are often the origin of impairing adolescent and adult psychiatric disorders (Hofstra, van der Ende, & Verhulst, 2002; McGee, Feehan, Williams, & Anderson, 1992). Despite this, for multiple reasons primarily related to inadequacies in insurance coverage, screening programs, numerous other aspects of the mental health care delivery system (e.g., lack of facilities in geographic areas of greatest need), and parental knowledge of available resources, only 20% of children in need of services receive specialty mental health care in the USA (Kataoka, Zhang, & Wells, 2002; U. S. Department of Health and Human Services, 1999).

Although these statistics highlight widespread childhood mental health problems and unmet need for treatment, the limited resources of the mental health care delivery system in the USA preclude providing treatment for all symptomatic children. Further, there is some evidence that childhood mental health problems exhibit substantial discontinuity (Costello et al., 2003; Prior, Smart, Sanson, & Oberklaid, 2001; Verhulst & van der Ende, 1992). Thus, the best use of scarce resources may be to develop cost-effective universal screening procedures to identify early those children most likely to evidence persistent impairing mental health problems (New Freedom Commission on Mental Health, 2003).

Extensive research has documented the broad patterns of childhood mental health problems and impairments. Externalizing problems are more stable than internalizing problems; and, comorbid internalizing and externalizing problems exhibit the greatest stability and highest levels of impairment (Eisenberg et al., 2001; Esser, Schmidt, & Woerner, 1990; Hofstra et al., 2002; Prior et al., 2001; Verhulst & van der Ende, 1992). Further, early onset and childhood trajectories characterized by persistent problems, further increase risk for later psychiatric disorders and impairments (Campbell, Spieker, Burchinal, & Poe, 2006; Fontaine et al., 2008; Kearney, Sims, Pursell, & Tillotson, 2003). These and other studies clearly provide the basis for consideration of early mental health screening to prevent progression to more toxic long-term outcomes. However, studies have not identified the longitudinal patterns of childhood mental health problems associated with the worst outcomes and, most importantly, the development of a specific algorithm for early identification with sufficient sensitivity and specificity to serve as a basis of a screening program.

The present study explores the broad patterns of children's mental health problems beginning at the transition into primary school and outcomes at grade 5, with the overarching aim of developing a universal school-based screening procedure. Such a screening procedure is one important component of a broader public health effort aimed at improving children's mental health (Figure 1), which includes a prior step of identifying risk factors for early mental health problems (necessary for the development of early prevention strategies), and a later step of referring the children identified by the screener for evaluation by mental health specialists for diagnosis and treatment (i.e., targeted intervention).

In the present study, kindergarten was chosen as the initial assessment time, and the broad-band categories of internalizing and externalizing problems assessed by short adult-report questionnaires were selected as appropriate for a universal school-based screening. Although mental health problems can emerge earlier in childhood (e.g., (Tremblay et al., 2004)), school-based screening is advantageous because of the natural aggregation of most children at the transition to primary school and the accessibility of both teachers and parents as informants,

which may increase the accuracy of identification of children with problems (Offord et al., 1996). We address three questions: (1) What are the broad patterns of mental health problems from kindergarten to grade 5? (2) What are the grade 5 outcomes of these patterns? (3) How early in school can children likely to develop the most impairing patterns be identified accurately?

Methods

Participants

The children represent a subset of the Wisconsin Study of Families and Work, a longitudinal study of child development. The original sample comprised 570 women and their partners recruited during pregnancy from obstetric/gynecology clinics and a low income clinic in two Midwestern cities (Hyde, Klein, Essex, & Clark, 1995); 560 of the women had live births and were eligible to continue in the study. Attrition resulted in response rates of 86% (n=479) at kindergarten and 71% (n=400) at grade 5. Because the present analyses focused on longitudinal patterns of mental health symptoms, analyses included only the 328 children (59% of the original eligible sample; 167 girls) of the 400 participating at grade 5 who had complete mother and teacher reports of children's symptoms in kindergarten and grades 1, 3, and 5. There were no statistically significant differences between these 328 children and the remaining 72 with incomplete data in the levels of symptoms at any assessment. Informed consent was obtained from subjects after complete description of the study at each assessment.

At recruitment (1990 – 1991), 24% of the 328 mothers and 28% of the fathers had a high school degree or less, 19% and 21% had some college, and 57% and 51% were college graduates. The median annual family income was $48,000 ($7,500 to >$200,000). Mother's average age was 29.7 (SD = 4.2); father's was 31.7 (SD = 5.1). Most couples were married (95%) and Caucasian (90%). There were no statistically significant differences in these characteristics between the 328 participants and the remaining families in the original sample with two minor exceptions: compared with participants, non-participating parents were one year younger (mothers, $M = 28.9$ v 29.7 years, $t(568) = -2.34$, $p = .020$; fathers, $M = 30.7$ v 31.7 years, $t(548) = -2.31$, $p = .021$), and non-participating fathers had a half year less education ($M = 14.8$ v 15.3 years of education, $t(548) = -2.01$, $p = .045$).

Measures

Measures included child mental health symptoms at kindergarten and grades 1, 3, and 5, and grade 5 outcomes, all included in the suite of measures comprising the adult-report MacArthur Health and Behavior Questionnaire (HBQ) (Boyce et al., 2002; Essex et al., 2002). The mental health scales, which take approximately 15 minutes to complete, are derived from the DSM-based Ontario Child Health Study (Boyle, Offord, Racine, Szatmari, & Sanford, 1993). All HBQ scales have strong psychometric properties (Essex et al., 2002) and its mental health scales have been shown to discriminate groups of children with and without signs of early psychopathology (Luby et al., 2002).

Mothers and teachers reported on the frequency (0=never, 1=sometimes, 2=often) of children's mental health symptoms in the past 6 months. *Internalizing Symptoms* included subscales for depression, overanxious, and for mothers only, separation anxiety (mother-report=29 items, teacher-report=14 items; α-coefficients = .64 – .90; all but two > .70). *Externalizing Symptoms* included subscales for oppositional defiance, conduct problems, inattention, impulsivity, overt aggression and relational aggression (mother-report= 46 items, teacher-report = 45 items; α-coefficients = .58 – .93; all but four > .70). All scale scores were computed as means. Mother and teacher reports, which overlapped modestly at each assessment (rs = . 20 – .46; all ps < .05), were averaged since using both informants best conforms to clinical

practice. The resulting Internalizing and Externalizing scores at each assessment ranged from 0 to maximums of .94 to 1.36. Cut-points defining children as high v. low Internalizing (high > .54) and Externalizing (high > .64) were obtained from an independent multi-site case-control study of the same-age children (kindergarten and grade 1) using Receiver Operator Characteristic analyses of averaged mother- and teacher-report HBQ scale scores to distinguish clinic-referred children from community controls (Ablow et al., 1999). No sex differences in these cut-points were found. In the present analyses, these cut-points corresponded to the upper 8% to 15% of the distributions of Internalizing and Externalizing at each assessment.

In grade 5, school-based assessments of children's functional impairments, physical health, and service use were obtained from teachers as outcome measures. In addition to being consistent with the focus on school-based screening, because children have different teachers each year, the use of only grade 5 teacher reports, and not mother reports, minimizes any method bias resulting from the use of the same informants for both prediction and outcome. Teachers used 3- to 5-point scales to report on children's *Academic Impairment* (13 items, e.g., How would you evaluate this child's current school performance in math-related skills), including subscales for academic competence (α=.95) and school engagement (α=.87); *Social Impairment* (12 items, e.g., Is teased and ridiculed by other children), including subscales for peer acceptance/rejection (α=.92) and bullied by peers (α=.72); *Global Functional Impairment* (7 items, e.g., How much have child's grades gone down as a result of these emotional/behavioral problems; α=.79); and *Global Physical Health Problems* (5 items, e.g., How often in average month does child stay or go home because of illness; α=.88). Teachers also reported on whether or not children had used any *School-based Services* (i.e., resource room, occupational/physical therapy, counseling/therapy); mothers reported on whether or not children had received *Specialty Mental Health Care* (i.e., evaluation or treatment by a psychiatrist or psychologist) outside of the school setting.

Data Analysis

Descriptive statistics were compiled for the prevalence of high symptoms at each assessment, gender differences in the rates of internalizing and externalizing symptoms, and the incidence of high symptoms from kindergarten to grade 5.

Transitions between states defined by low/high Internalizing and/or Externalizing symptoms from one assessment to the next were described using one-step transition matrices. These transition patterns were then used to define the most common longitudinal symptom patterns.

To compare the longitudinal symptom groups on grade 5 outcomes, Chi Square Contingency Analysis (with categorical dependent variables) and One-Way Analysis of Variance (with continuous dependent variables) with a 5% significance level were used. When statistical significance occurred, pair-wise comparison of groups indicated differences. For estimating effect size, Number Needed to Take (NNT) was chosen because it is strongly recommended for its clinical interpretability (Kraemer & Kupfer, 2006). NNT indicates the number of high-risk children one would need to take to find one more child with the outcome of interest than if one sampled the same number of low-risk children. Thus, the smaller the NNT, the greater the clinical importance of the difference in outcome between the two risk groups. If NNT = 1, every child in the high-risk group would have the outcome and none of the children in the low-risk group would have it, i.e., perfect discrimination. As NNT gets larger, it indicates increasingly weaker discrimination. Corresponding to Cohen's standards (Cohen, 1988), a small effect size (d = .2) would be NNT of > 8.9, a medium effect size (d = .5) would be NNT of 3.6, and a large effect size (d = .8) would be NNT < 2.3 (Kraemer & Kupfer, 2006).

Receiver Operating Characteristic (ROC) methods (Kiernan, Kraemer, Winkleby, King, & Taylor, 2001; Kraemer, 1992) were applied to distinguish as early in school as possible the

children who developed the most problematic pattern of mental health symptoms vs. all others. Six variables defining children with high (1) vs. low (0) internalizing, externalizing, or comorbid symptoms in kindergarten and grade 1, and child gender, were included as predictors. In ROC, each predictor variable is used to split the sample, and the success of each split is evaluated using a weighted kappa coefficient with the weight determined by the relative clinical importance of false positive and false negative classifications (here, both were weighted equally). The split with the maximal kappa is selected as optimal and accepted if it passes a pre-set stopping rule (here, $p < .01$). The sample is then split into the two subgroups, and the process is repeated separately in each. This process continues until either the subgroup size is too small for adequate evaluation (here, fewer than 10 subjects) or until no other significantly distinguishing predictors are found, i.e., the split fails the stopping rule. This results in a "decision tree" defining subgroups, or classifications, at varying levels of risk for the outcome. Those classifications in which the proportion of subjects with the outcome exceeds the base rate of the outcome are defined as high risk and included in the final classification; the rest are not. A quick visual assessment of the final classification and its major competitors can be obtained from a ROC plane, using sensitivity (Se; probability of a positive classification among those with the outcome) and specificity (Sp; probability of a negative classification among those without the outcome) to locate the position of each classification between random (Se=1-Sp) and ideal (Se=Sp=1).

Results

Descriptive Statistics

Children evidencing high symptoms included 19% at kindergarten, 21% at grade 1, 27% at grade 3, and 16% at grade 5, with approximately equal percentages showing high externalizing and high internalizing symptoms. For all but the grade 5 assessment, there were significantly more boys with high externalizing symptoms [kindergarten to grade 3, range of X^2 (1, N=328) = 3.41 to 5.70, ps < .05; grade 5, p = .057]; there were no gender differences in high internalizing symptoms (all ps > .15). By grade 5, 41% of children had evidenced high symptoms at some point since kindergarten; of these children, 39% had high symptoms once, 36% twice, 15% three times, and 10% all four times. These prevalence and incidence rates are consistent with those of previous epidemiological research (Costello et al., 2003).

Longitudinal Patterns of Mental Health Symptoms

The patterns of transitions in symptoms from one assessment time to the next were quite similar. Thus, Table 1 highlights the average pattern.

As expected, the most common transition from one assessment time to the next was from Low Symptoms to Low Symptoms (average 88%). Consequently the most common longitudinal pattern was one in which high symptoms were never seen (i.e., *Never Symptoms*; 58.8%).

Among children with only Internalizing or Externalizing symptoms, the most common transition was a return to Low Symptoms (average 69% of Internalizing Only and 43% of Externalizing Only). Thus the next most common pattern is one in which there are no two consecutive assessments with symptoms (i.e., *Isolated Symptoms*; 21.6%). In addition to those with only Internalizing or Externalizing symptoms, most children who transition from Low Symptoms to Internalizing Only will likely be in this group since the probability of transitioning from Internalizing Only to either of the other two high symptom groups is very low.

Finally, the remaining children were those with recurrent symptoms (i.e., high symptoms at consecutive assessments), with Comorbid Symptoms the most stable over time (average 42%). Thus, the remaining patterns were divided into two groups with recurrent symptoms, either

without comorbid internalizing and externalizing symptoms (i.e., *Recurrent Non-comorbid Symptoms*; 10.1%) or with comorbid symptoms at least once (i.e., *Recurrent Comorbid Symptoms*; 9.5%).

Notably, children with Recurrent Comorbid Symptoms were most likely to have high symptoms at least three times (77%) and almost half had high symptoms all four times (45%). In contrast, children with Recurrent Non-comorbid Symptoms were most likely to have high symptoms only twice (70%) and none had symptoms all four times. Children with Isolated Symptoms generally had high symptoms only once (73%).

Comparisons of the longitudinal symptom groups showed that there were no statistically significant differences in child gender (χ^2 [3, N=328] = 5.45, *p* = .142).

Grade 5 Outcomes

Overall, children with Recurrent Comorbid Symptoms had the highest levels of impairment and service use by grade 5 (Table 2). Teachers reported significantly higher levels of academic, social, and global impairment, and global physical health problems for this group compared with the other three groups. Teachers also reported the highest rates of service use for this group, with almost half (48.4%) using school-based services in grade 5. Mothers reported that, compared with the Never and Isolated Symptoms groups, children in the Recurrent Comorbid Symptoms group had higher rates of specialty mental health treatment, with over a third (36.7%) having received treatment from a psychologist or psychiatrist.

Children in the Isolated and Recurrent Non-comorbid Symptoms groups also differed significantly from the Never Symptoms group, showing higher levels of academic, social, and global impairment, and higher rates of school-based service use and specialty mental health care.

There were no statistically significant gender differences in the grade 5 outcomes or the associations of the outcomes with the longitudinal symptom groups (main and interactive effects, all *p*s < .19).

The clinical significance of the symptom group differences in grade 5 outcomes is highlighted in Figure 2. NNT comparing the Recurrent Comorbid Symptoms versus the Never Symptoms groups indicates generally strong discrimination (all NNTs ≤ 3.2), especially in Academic, Social, and Global Impairment and Specialty Mental Health Treatment [NNTs ≤ 2.3, corresponding to Cohen's standard for a large effect size, *d* > .8 (Cohen, 1988)]. These NNT values indicate much stronger discrimination than that between the Never Symptoms group and Recurrent Non-comorbid Symptoms (NNTs = 2.0 to 9.5) or Isolated Symptoms (NNTs = 3.4 to 12.0) groups.

Early Identification of Children with Recurrent Comorbid Symptoms

Children's trajectories of mental health problems are quite evident as early as kindergarten. Kindergarteners with Low Symptoms were not likely to develop later symptoms (72% Never Symptoms). Kindergarteners with Internalizing Only symptoms were most likely to develop a pattern of Isolated Symptoms (65%), or less likely, Recurrent Non-comorbid Symptoms (30%). Kindergarteners with Externalizing Only symptoms were less likely to develop a pattern of Isolated Symptoms (39%), and most likely to develop recurrent symptoms, either Recurrent Comorbid Symptoms (32%) or Recurrent Non-comorbid Symptoms (29%). Notably, although only 10 kindergarteners evidenced Comorbid Symptoms, 100% of them developed a pattern of Recurrent Comorbid Symptoms.

Finally, an ROC analysis, as described in the Data Analysis section, was applied to identify as early in school as possible the children who developed Recurrent Comorbid Symptoms (n = 31 of 328; 9.5% base rate). The most optimal predictor identified the high-risk subgroup of 10 children with Comorbid Symptoms in kindergarten (3% of the population), described above, 100% (10/10) of whom developed Recurrent Comorbid Symptoms compared with only 6.6% (21/318) of the remaining children. Next, within the remaining subgroup of 318 children (those without comorbid symptoms in kindergarten), the most optimal predictor identified the subgroup of 29 children with externalizing symptoms in grade 1 (9% of the population), 51.7% (15/29) of whom developed the Recurrent Comorbid pattern. When both high-risk groups were considered together, 64% (25/39) developed Recurrent Comorbid Symptoms compared with only 2.1% (6/289) of the children in neither high-risk group. Moreover, this final classification correctly identified 81% (25/31) of the children who developed Recurrent Comorbid Symptoms (sensitivity) and 95% (283/297) of those who did not (specificity). Figure 3 (#1), compared with the sensitivity and specificity of the other classifications considered, is the most optimal based on its relative proximity to ideal and distance from random.

Discussion

This study investigated patterns of childhood mental health symptoms from kindergarten to grade 5, and grade 5 outcomes, with the overarching goal of developing a universal school-based screening procedure to identify early the children likely to develop the most impairing symptom patterns and thus, in need of early intervention. The findings highlight the distinctions between children who develop a pattern of recurrent comorbid internalizing and externalizing symptoms and those with patterns of recurrent non-comorbid symptoms, isolated symptoms, or who never evidence high symptoms. Previous studies have found higher levels of impairment and service use among children with recurring problems (Esser et al., 1990; Verhulst & van der Ende, 1992) or with comorbid internalizing and externalizing problems (Eisenberg et al., 2001; Hofstra et al., 2002). The present study emphasizes that it is the combination of recurring and comorbid symptoms that strongly distinguishes children likely to suffer pervasive impairment by early adolescence, and proposes a universal school-based screening strategy to identify them.

The proposed screening strategy is part of a larger public health approach to improving children's mental health (Figure 1), which aims to minimize the costs and maximize the benefits of universal, targeted, and clinical strategies (Offord, Kraemer, Kazdin, Jensen, & Harrington, 1998). Overall, the findings suggest that universal school-based screening can greatly improve detection of children likely to benefit from early mental health intervention. Specifically, the findings indicate that children most likely to develop recurrent comorbid symptoms can be identified quite accurately by the end of grade 1 using a relatively low-cost universal screening procedure based on mother and teacher questionnaire reports of children's mental health symptoms during the transition to primary school. Only the small subgroup of children who screen positive would then be referred for more costly expert diagnosis and intervention, alleviating referral burden for mental health specialists and thus, permitting more timely and specialized evaluations and interventions for children truly in need. Clinical follow-up would also provide opportunities for assessment of risk factors, symptom specificity and clustering, and differential treatment responses in this high-risk group of children.

Such benefits must be weighed against the costs of any screening procedure (Offord et al., 1998). Screening procedures, which typically target only a small portion of the population (here, 12%), may be considered too costly; and, participation rates may be lowest for those at highest risk for the outcome. The proposed strategy is based on short (15-minute), self-administered adult-report questionnaires that could be integrated with other routine school-based screenings, thus minimizing costs and enhancing participation rates. Problems with

accuracy are also of concern. In the present study, there were no false positives in the kindergarten high-risk group, but almost half of the children in the grade 1 high-risk group were false positives. Importantly, however, all children identified by high externalizing symptoms in grade 1 developed a pattern of isolated or, more likely, recurrent non-comorbid symptoms; and both of these groups, especially the recurrent group, suffered significantly greater impairments by grade 5 than children who never evidenced high symptoms. This suggests that these children may also benefit from early referral to mental health specialists. In summary, the advantages of universal school-based screening and the high levels of screening accuracy shown here, coupled with the substantial morbidity associated with recurrent comorbid internalizing and externalizing symptoms, suggest that the potential benefits of early screening and identification significantly outweigh the potential costs.

These findings must be considered in light of potential limitations. The non-representativeness of this population, and sample attrition, raise issues of generalizability. There were, however, only minor differences in demographics and no differences in levels of children's mental health symptoms between the participants and non-participants. Further, although the study considered gender differences, it did not include other psychosocial factors that might influence the link between the longitudinal patterns of mental health symptoms and the outcomes. Finally, the lack of a diagnostic criterion, and the necessary use of symptom cut-points from an independent study, raises additional concerns. However, this was an exploratory study with the goal of generating pilot information necessary for future hypothesis-testing studies. The next steps would be an independent validation followed by studies addressing important questions regarding the specific problems of children who screen positive, and optimal referral and treatment strategies.

Conclusion

The success of any screening program ultimately depends not only on its accuracy, but on whether the referrals to mental health specialists result in a decrease of poor outcomes. In the present study, only half of the children with recurrent comorbid symptoms were receiving any school-based services, and only a third had received specialty mental health care, underscoring the fact that children most in need of treatment are not receiving it (Kataoka et al., 2002). Further, studies of mental health screening in primary care settings indicate that even when screening is completed and children referred, many families do not follow through (Hacker et al., 2006). It is our hope that universal screening at school entry, where services are already in place, will improve follow-up evaluation and care, with subsequent increases in positive outcomes for children most in need of early intervention.

Acknowledgements

References

Ablow JC, Measelle JR, Kraemer HC, Harrington R, Luby J, Smider N, et al. The MacArthur Three-City Outcome study: Evaluating multi-informant measures of young children's symptomatology. Journal of the American Academy of Child and Adolescent Psychiatry 1999;38(12):1580–1590. [PubMed: 10596259]

Angold A, Costello EJ, Farmer EMZ, Burns BJ, Erkanli A. Impaired but undiagnosed. Journal of the American Academy of Child and Adolescent Psychiatry 1999;38(2):129–138. [PubMed: 9951211]

Boyce WT, Essex MJ, Goldstein LH, Armstrong JM, Kraemer HC, Kupfer DJ. The confluence of mental, physical, social, and academic difficulties in middle childhood. I: Exploring the "headwaters" of early

life morbidities. Journal of the American Academy of Child and Adolescent Psychiatry 2002;41(5): 580–587. [PubMed: 12014791]

Boyle MH, Offord DR, Racine Y, Szatmari P, Sanford M. Evaluation of the revised Ontario Health Study Scales. Journal of Child Psychology and Psychiatry 1993;34:189–213. [PubMed: 8444992]

Campbell SB, Spieker S, Burchinal M, Poe MD. Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. Journal of Child Psychology and Psychiatry 2006;47(8):791–800. [PubMed: 16898993]

Cohen, J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Archives of General Psychiatry 2003;60:837–844. [PubMed: 12912767]

Eisenberg N, Cumberland A, Spinrad TL, Fabes RA, Shepard SA, Reiser M, et al. The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. Child Development 2001;72(4):1112–1134. [PubMed: 11480937]

Esser G, Schmidt MH, Woerner W. Epidemiology and course of psychiatric disorders in school-age children: Results of a longitudinal study. Journal Child Psychology and Psychiatry 1990;31(2):243–263.

Essex MJ, Boyce WT, Goldstein LH, Armstrong JM, Kraemer HC, Kupfer DJ. The confluence of mental, physical, social, and academic difficulties in middle childhood: II. Developing the MacArthur Health and Behavior Questionnaire. Journal of the American Academy of Child and Adolescent Psychiatry 2002;41(5):588–603. [PubMed: 12014792]

Fontaine N, Carbonneau R, Barker ED, Vitaro F, Hebert M, Cote SM, et al. Girls' hyperactivity and physical aggression during childhood and adjustment problems in early adulthood. Archives of General Psychiatry 2008;65(3):320–328. [PubMed: 18316678]

Hacker KA, Myagmarjav E, Harris V, Suglia SF, Weidner D, Link D. Mental health screening in pediatric practice: Factors related to positive screens and the contribution of parental/personal concern. Pediatrics 2006;118:1896–1906. [PubMed: 17079560]

Hofstra MB, van der Ende J, Verhulst FC. Child and adolescent problems predict DSM-IV disorders in adulthood: A 14-year follow-up of a Dutch epidemiological sample. Journal of the American Academy of Child and Adolescent Psychiatry 2002;41(2):182–189. [PubMed: 11837408]

Hyde JS, Klein MH, Essex MJ, Clark R. Maternity leave and women's mental health. Psychology of Women Quarterly 1995;19:257–285.

Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: Variation by ethnicity and insurance status. American Journal of Psychiatry 2002;159(9):1548–1555. [PubMed: 12202276]

Kearney CA, Sims KE, Pursell CR, Tillotson CA. Separation anxiety disorder in young children: A longitudinal and family analysis. Journal of Clinical Child and Adolescent Psychiatry 2003;32(4):593–598.

Kiernan M, Kraemer HC, Winkleby MA, King AC, Taylor CB. Do logistic regression and signal detection identify different subgroups at risk? Implications for the design of tailored interventions. Psychological Methods 2001;6:35–48. [PubMed: 11285811]

Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne B, Poulton R. Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective longitudinal cohort. Archives of General Psychiatry 2003;60:709–717. [PubMed: 12860775]

Kraemer, HC. Evaluating medical tests: Objective and quantitative guidelines. Newbury Park, CA: Sage Publications; 1992.

Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biological Psychiatry 2006;59(11):990–996. [PubMed: 16368078]

Leaf PJ, Alegria M, Cohen P, Goodman SH, Horwitz SM, Hoven CW, et al. Mental health service use in the community and schools: Results from the four-community MECA study. Journal of the American Academy of Child and Adolescent Psychiatry 1996;35(7):889–897. [PubMed: 8768348]

Luby JL, Heffelfinger A, Measelle JR, Ablow JC, Essex MJ, Dierker L, et al. Differential performance of the MacArthur HBQ and DISC-IV in identifying DSM-IV internalizing psychopathology in young

children. Journal of the American Academy of Child and Adolescent Psychiatry 2002;41(4):458–466. [PubMed: 11931603]

McGee R, Feehan M, Williams S, Anderson J. DSM-III disorders from age 11 to age 15 years. Journal of the American Academy of Child and Adolescent Psychiatry 1992;31(1):50–59. [PubMed: 1537781]

New Freedom Commission on Mental Health. Achieving the promise: Transforming mental health care in America: Final report (No. SMA-03-3832). Rockville, MD: Department of Health and Human Services; 2003.

Offord DR, Boyle MH, Racine Y, Szatmari P, Fleming JE, Sanford M, et al. Integrating assessment data from multiple informants. Journal of the American Academy of Child and Adolescent Psychiatry 1996;35:1078–1085. [PubMed: 8755805]

Offord DR, Kraemer HC, Kazdin AE, Jensen PS, Harrington R. Lowering the burden of suffering from child psychiatric disorder: Trade-offs among clinical, targeted, and universal interventions. Journal of the American Academy of Child and Adolescent Psychiatry 1998;37(7):686–694. [PubMed: 9666623]

Prior M, Smart D, Sanson A, Oberklaid F. Longitudinal predictors of behavioral adjustment in pre-adolescent children. Australian & New Zealand Journal of Psychiatry 2001;35:297–307. [PubMed: 11437802]

Tremblay RE, Nagin DS, Seguin JR, Zoccolillo M, Zelazo PD, Boivin M, et al. Physical aggression during early childhood: Trajectories and predictors. Pediatrics 2004;114(1):43–50.

U. S. Department of Health and Human Services. Mental Health: A report of the surgeon general. Rockville, MD: Substance Abuse and Mental Health Services Administration, National Institutes of Health; 1999.

Verhulst FC, van der Ende J. Six-year developmental course of internalizing and externalizing problem behaviors. Journal of the American Academy of Child and Adolescent Psychiatry 1992;31(5):924–931. [PubMed: 1400127]
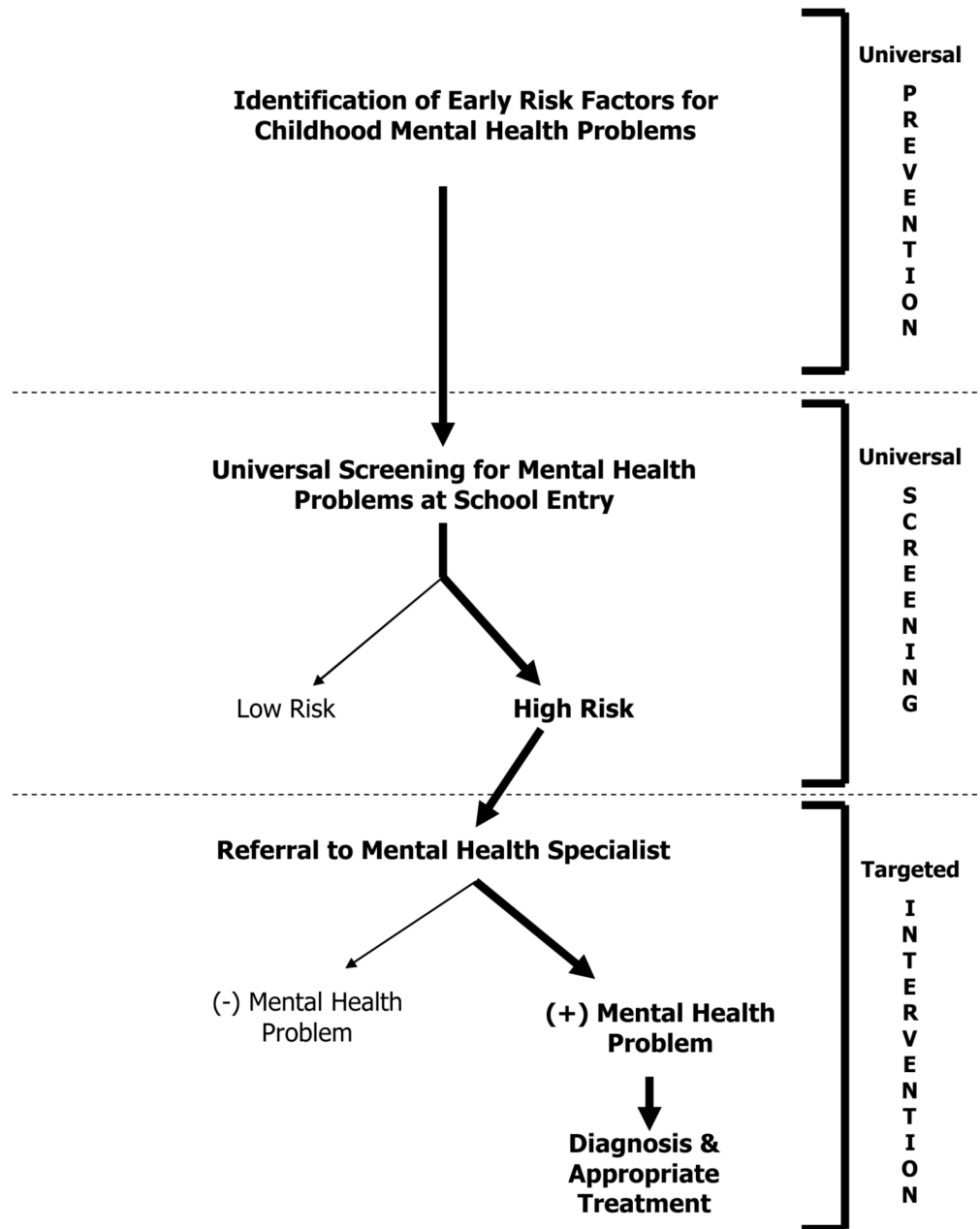
Identification of Early Risk Factors for Childhood Mental Health Problems

Universal
P
R
E
V
E
N
T
I
O
N

Universal Screening for Mental Health Problems at School Entry

Low Risk   **High Risk**

Universal
S
C
R
E
E
N
I
N
G

Referral to Mental Health Specialist

(-) Mental Health Problem   **(+) Mental Health Problem**

Diagnosis & Appropriate Treatment

Targeted
I
N
T
E
R
V
E
N
T
I
O
N

Figure 1.
Universal screening at the transition to primary school as one component of a larger public health effort to improve children's mental health.
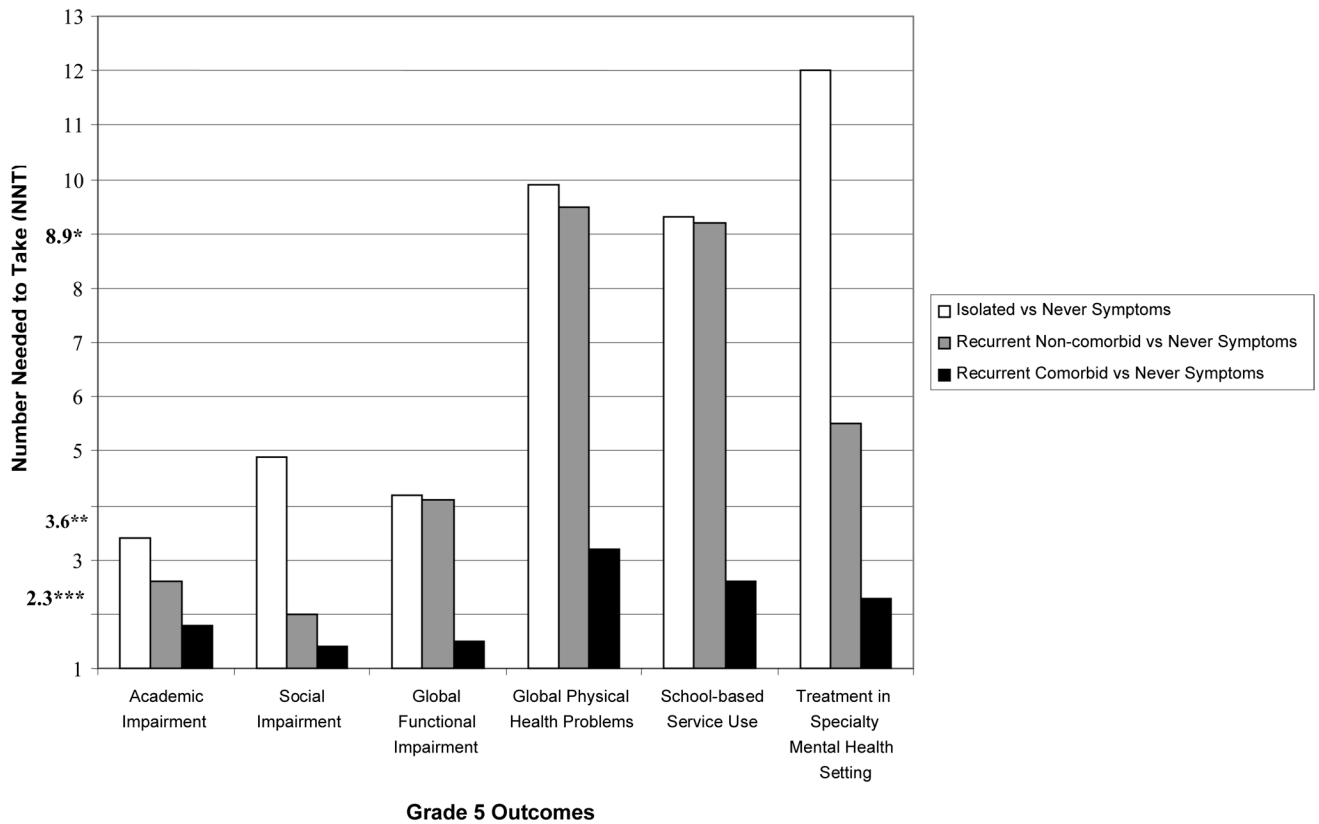
Figure 2.
Number Needed to Take (NNT), an effect size indicating the clinical significance of differences in each grade 5 outcome between children in the Never Symptoms group (n = 193) and those in the Isolated (n = 71), Recurrent Non-comorbid (n = 33), and Recurrent Comorbid (n = 31) groups. Asterisks indicate NNT values corresponding to Cohen's standards for small (*d = . 2), medium (**d = .5), and large (***d = .8) effect sizes (Kraemer & Kupfer, 2006).
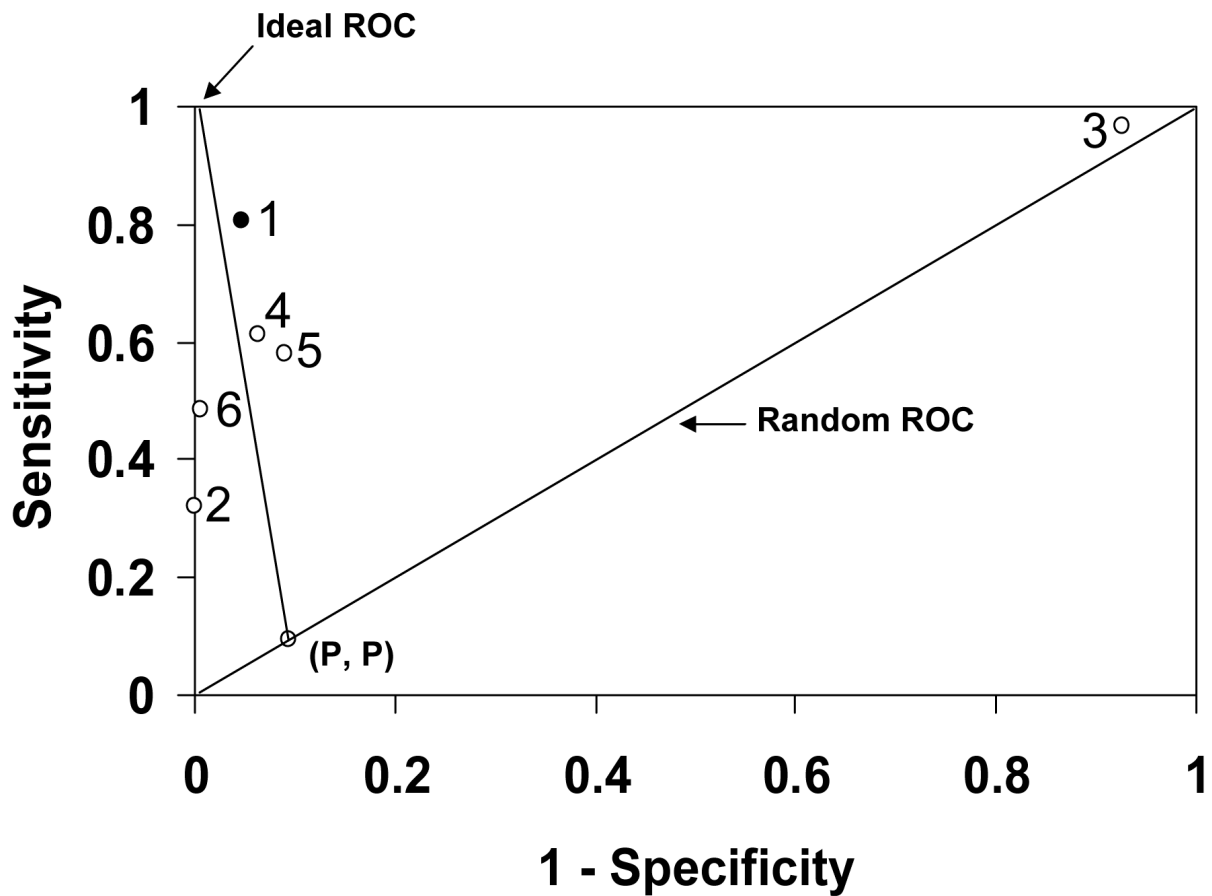
Figure 3.
Receiver Operating Characteristic (ROC) plane using sensitivity and specificity to compare the proposed screening algorithm based on kindergarten comorbidity and grade 1 externalizing (#1) with alternatives based on kindergarten comorbidity alone (#2), or in combination with kindergarten internalizing (#3), kindergarten externalizing (#4), grade 1 internalizing (#5), or grade 1 comorbidity (#6). P = Proportion with Recurrent Comorbid Symptoms.

Table 1

Average Transition Matrix from One Assessment Time to the Next, Kindergarten to Grade 5. Three matrices were defined: kindergarten to grade 1, grade 1 to grade 3, and grade 3 to grade 5. Percentages are averages of the three matrices, with those for the separate matrices in parentheses. N = 328

Symptom Status at One Time	Symptom Status at the Next Time				
	Low Symptoms	Internalizing Only	Externalizing Only	Comorbid Symptoms	Total
Low Symptoms	88% (87, 83, 92)	8% (8, 10, 7)	3% (3, 5, 1)	1% (2, 2, 0)	100%
Internalizing Only	69% (70, 57, 81)	24% (26, 30 15)	3% (4, 3, 2)	4% (0, 10, 2)	100%
Externalizing Only	43% (43, 16, 69)	8% (7, 12, 4)	32% (46, 32, 18)	17% (4, 40, 9)	100%
Comorbid Symptoms	17% (0, 25, 26)	17% (20, 16, 13)	24% (30, 17, 26)	42% (50, 42, 35)	100%
Total	79% (79, 74, 84)	10% (9, 12, 8)	6% (8, 7, 4)	5% (4, 7, 4)	100%

Table 2

Grade 5 Outcomes of Longitudinal Symptom Groups.

	Longitudinal Symptom Groups				
	Never Symptoms (n = 193)	Isolated Symptoms (n = 71)	Recurrent Non-comorbid Symptoms (n = 33)	Recurrent Comorbid Symptoms (n = 31)	Test Statistic
Impairment at Grade 5 (all standardized; $M = 0$, $SD = 1$)	M (SD)	M (SD)	M (SD)	M (SD)	
Academic impairment	−.30 (.83)^{abc}	.17 (.79)^{ad}	.24 (1.06)^{ad}	1.05 (1.16)^{bcd}	$F(3,321) = 23.62$**
Social impairment	−.34 (.71)^{abc}	.04 (.87)^{ad}	.50 (1.15)^{ad}	1.37 (1.15)^{bcd}	$F(3,321) = 41.46$**
Global functional impairment	.10 (.17)^{abc}	.20 (.23)^{ad}	.23 (.31)^{ad}	.55 (.48)^{bcd}	$F(3,324) = 30.07$**
Global physical health problems	.16 (.34)^{a}	.23 (.39)^{a}	.07 (.15)^{a}	.45 (.57)^{bcd}	$F(3,324) = 7.06$*
Use of Services at Grade 5	N (%)	N (%)	N (%)	N (%)	
School-based services	20 (10.4%)^{abc}	15 (21.1%)^{ad}	7 (21.2%)^{ad}	15 (48.4%)^{bcd}	χ^2 (3, N=328) = 28.4**
Specialty mental health care	6 (3.1%)^{abc}	8 (11.4%)^{ad}	7 (21.2%)^{d}	11 (36.7%)^{cd}	χ^2 (3, N=326) = 39.3**

Within rows, superscripts indicate that a mean differs significantly ($p < .05$) from the mean for

^{a}Recurrent Comorbid

^{b}Recurrent Non-comorbid

^{c}Isolated , and/or

^{d}Never symptoms.

*
$p < .01$

**
$p < .001$