



Published in final edited form as:

*J Affect Disord.* 2011 October ; 133(3): 537–545. doi:10.1016/j.jad.2011.04.056.

## The 2-Week Duration Criterion and Severity and Course of Early Childhood Depression: Implications for Nosology

Michael S. Gaffrey, Ph.D., Andy C. Belden, Ph.D., and Joan L. Luby, M.D.

Washington University School of Medicine, Departments of Psychiatry (Child Division), Washington University in St. Louis, Saint Louis, Missouri

### Abstract

**Background**—Although validity for DSM-IV MDD symptom criteria in preschoolers has been demonstrated, whether the 2-week duration criterion is an appropriate threshold of clinical significance at this age remains unclear. The current study aimed to begin addressing this question.

**Method**—Three hundred and six preschoolers were recruited from community sites and followed longitudinally for 2 years. A subsample including healthy preschoolers (N=77) and those with MDD (N=74) were examined. The MDD group was further divided based upon meeting (DSM, N=24) or failing to meet (<DSM, N=50) the DSM-IV 2-week duration criterion. Groups were compared on parent and teacher report measures of symptom severity and functional impairment at baseline and 2-year follow-up.

**Limitations**—A larger sample of depressed preschoolers and refined measures of duration are needed to replicate the current study.

**Results**—Preschoolers with MDD differed significantly from controls on the majority of measures examined regardless of duration status and time of assessment. Further, the DSM group significantly differed from the <DSM group at baseline on measures of MDD symptom severity and impairment. No differences in the risk of a MDD diagnosis at follow-up were found on the basis of duration group status.

**Conclusions**—DSM-IV duration criterion failed to capture all clinically affected preschoolers at baseline or confer greater predictive validity for a depression diagnosis 2 years later. Findings suggest that preschoolers meeting all DSM-IV MDD criteria except for episode duration exhibit a clinically significant form of depression and experience a 2-year MDD outcome similar to those meeting full criterion.

### Keywords

Preschool; Depression; Depressive Episode; Mood Disorder; Pediatric; Developmental Psychopathology

---

© 2011 Elsevier B.V. All rights reserved.

Correspondence to: Michael S. Gaffrey, Ph.D., Washington University School of Medicine, Department of Psychiatry (Child Division), 660 South Euclid Avenue, Campus Box: 8134, St. Louis, MO 63110, gaffrey@wustl.edu, 314-747-5641 (phone), 314-286-2732 (fax).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Introduction

In the identification of very early onset depression, Luby and Belden have suggested there is a critical need for a more developmentally nuanced consideration of depression related symptomatology given the unique and variable patterns of emotional competence during this developmental period (Luby and Belden, 2006). That is, the flexibility, modulation, and complexity of emotion expression and regulation during this time can vary greatly based on setting, situational demands, as well as the emotion being experienced (Denham, 1998, Saarni, 2000, Cole, 1986b). Luby and colleagues were the first to directly investigate the validity of major depressive disorder (MDD) in preschoolers using a developmentally informed assessment and developmentally modified criteria (Luby et al., 2003, Luby et al., 2002a). Their findings indicated that the clinical picture of MDD in preschoolers was similar to that found in older children and adolescents. Specifically, the manifestation of core depression symptoms in older age groups (e.g. anhedonia, difficulty concentrating, irritability, etc.) were also those that characterized preschoolers with MDD. However, they emphasized that the key to identifying these symptoms was a developmentally informed 'translation' of symptom manifestations in the assessment process. For example, items pertaining to preoccupation with death and suicidality were modified to account for their possible manifestation as recurrent themes in play rather than verbal expressions alone. Research by this group has also suggested that the strict 2-week duration criteria for a depressive episode be 'set aside' in Preschool Onset Major Depressive Disorder (PO-MDD; Luby et al., 2003). This was in large part due to the significant number of highly impaired preschoolers who did not qualify for a depression diagnosis on the basis of failing to meet this duration requirement alone. Specifically, of the 49 children meeting the developmentally modified criteria for depression in Luby and colleagues 2002 study (Luby et al., 2002a), 76% (37) of them did not meet standard DSM-IV duration criteria when assessed with the Diagnostic Interview Schedule for Children-Version 4 (Schaffer et al., 2000). As a result, the authors suggested that although the validity of core DSM-IV symptom constructs for depression were supported in preschoolers, the duration of these symptoms (including depressed mood/irritability and anhedonia) in young children warranted further investigation.

It should be noted that questions surrounding the required diagnostic criteria for MDD, including episode duration, while salient and controversial in this age group are not unique to it. In fact, a large body of research investigating 'subthreshold depressive disorders' in children, adolescents, and adults (i.e., Minor Depression, subsyndromal depression) has suggested that individuals failing to meet the full DSM-IV MDD diagnostic requirements still demonstrate significant functional and psychosocial impairment (Fergusson et al., 2005; Kendler and Gardner, 1998; Karlsson et al., 2009; Klein, 2008; Birmaher et al., 2009; Lewinsohn et al., 2000). Additionally, the question of whether subthreshold depressive disorders represent unique diagnostic entities unto themselves or are better conceptualized as a prodromal or residual phase of MDD has also been raised with no clear answer apparent. This issue has been especially relevant in childhood depression where a number of studies have examined the presence of depressive symptoms (both kind and quantity) as predictive of a future MDD diagnosis. Notably, this body of literature has suggested that the presence of depressive symptoms in children and adolescents without a history of depression is highly predictive of later MDD in an additive manner (i.e., 2-4 fold increase in risk per symptom; Ialongo et al., 2001; Keenan et al., 2008) and that any combination of several depressive symptoms (rather than a specific symptom or cluster of symptoms) present for at least one week is important in this regard (Kovacs and Lopez-Duran, 2010). While the presence and number of symptoms has received much attention in pediatric depression, there is a dearth of research examining the influence of episode duration in children who meet the full symptom requirements of MDD. However, there is some support for individuals at younger ages

exhibiting a less persistent pattern of depressed mood and that major depression may be best viewed as a continuum of symptoms varying in degree of severity and duration rather than a categorical entity (Kendler and Gardner, 1998; Karlsson et al., 2009; Klein, 2008; Birmaher et al., 2009; Carta et al., 2003). Nevertheless, the influence of episode duration in childhood depression remains an area in need of further empirical exploration.

Although still ambiguous in many domains, the body of literature concerning diagnosis in older children is far more substantial than what is currently available in early childhood. And, while empirical findings on the diagnostic boundaries in MDD using older age groups offers some guidance, consideration of developmental differences in emotional competency as noted above is of fundamental importance when considering psychopathology in very early childhood. This has been of particular benefit in the identification of depressive symptomatology in preschoolers as well as in the conceptualization of other relevant psychiatric disorders at this age (e.g., reduced number of C criterion for Post Traumatic Stress Disorder; Meiser-Stedman et al., 2008). Following from this, we sought to investigate whether the current conceptualization of a minimum duration of 2 weeks for an episode of depression was appropriate for preschool age children. Based on reports of developmental variation in the durational features of childhood mood disorders, and the greater variability in expression and regulation of emotion in younger children, we hypothesized that the 2-week duration may not be a developmentally appropriate threshold (Birmaher et al., 2009; Cole et al., 1996). To date, the influence of episode duration has not been directly addressed in preschool age children with MDD and, more specifically, whether or not the required 2-week duration criteria as specified in the DSM-IV captures all clinically relevant manifestations of depression at this age remains unclear. Additionally, whether or not episode duration impacts symptom severity and/or impairment at long term follow up, an issue pertinent to the importance of duration in setting the diagnostic threshold, is also unknown.

The purpose of the current study was to build upon previous findings (Luby and Belden, 2006; Luby et al., 2003; Luby et al., 2009b) generated from a federally funded program of research examining the onset and course of preschool onset major depression by addressing the above noted gaps concerning episode duration in preschool onset MDD using parent and teacher data. It was hypothesized that similar to studies examining episode duration in older children and adults, DSM-IV duration criteria would fail to capture all clinically significant groups. Rather, when comparing PO-MDD groups based on duration of depressive episode (i.e., 2 weeks vs. a portion of the day for several days) with each other and a group of similarly aged healthy control children, a 'stepwise' hierarchy would be found. Specifically, children with depressive episodes meeting the DSM-IV duration criteria would demonstrate greater impairment and level of symptomatology (i.e., number of symptoms) than children meeting all DSM-IV MDD criteria except for duration and both groups would demonstrate significantly more symptoms of depression, higher severity and impairment than healthy control children (Luby et al., 2003). Additionally, it was hypothesized that these between group differences would persist when the same children were examined 2 years later and that episode duration (i.e., DSM vs. <DSM) would not influence the relative risk for a diagnosis of MDD at this later time.

## Methods

### Participants

The Validation of Preschool Depression Study (PDS) is an ongoing longitudinal investigation of 306 preschoolers and their families funded by the National Institute of Mental Health. Details regarding recruitment and assessment of this sample have been previously described (Luby et al., 2009b). Children between the ages of 3 to 5.11 years and

their families were recruited from pediatricians' offices, daycare centers, and preschools in the greater St. Louis metropolitan area using the Preschool Feelings Checklist (PFC), a sensitive screening measure for depression in young children (Luby et al., 2004a). Of the 6,000 PFCs distributed, 1,474 were returned. Caregivers endorsing no symptoms on the PFC as well as those endorsing 3 or more internalizing and/or externalizing symptoms were contacted for further screening (n=899). Children who were reported to have chronic illnesses, neurological problems, suspected of having a pervasive developmental disorder and/or a significant suspected cognitive and/or speech delay, and those outside the study age range were excluded. Of the 416 eligible families, 306 agreed to participate. Upon enrollment into the study, each family participated in 3 annual comprehensive age appropriate mental health and developmental assessments (spanning 2 years for each subject) at the Early Emotional Development Program at the Washington University School of Medicine in St. Louis, St. Louis, Missouri. After complete description of the study, written informed consent was obtained from the children's guardian(s) prior to study participation. Given their young age, child assent was not obtained. The current study reports on a subsample chosen from this group based on their meeting symptom criteria for PO-MDD (n=74) at the baseline assessment or failing to qualify for any psychiatric diagnosis across all study waves (n=77; deemed a "Healthy" comparison group). The Washington University Institutional Review Board approved all study procedures.

### Diagnostic Assessment

Diagnostic assessments were conducted at each annual wave using the Preschool Age Psychiatric Assessment (PAPA; Egger et al., 1999, 2003). The PAPA is an interviewer-based instrument with established test-retest reliability for use with caregivers of children between 2.0 to 6.0 years of age (Egger et al., 2006). It was developed to assess for DSM-IV Axis I disorders in preschoolers using developmentally appropriate symptom manifestations. Ratings of symptom severity, frequency, duration and resulting impairment are also collected during the PAPA. Trained research assistants blind to diagnostic group membership and individual performance on study measures administered the PAPAs. Weekly calibration meetings with master coders as well as coding checks on 20% of PAPAs, where any coding discrepancies were reviewed with a senior clinician (JL), were conducted in order to maintain calibration and avoid rater drift over the course of the Preschool Depression Study. Computer algorithms applying relevant symptom and duration criteria gathered during the PAPA were used to generate DSM-IV diagnoses. It should be noted that only the duration criteria for MDD was set aside and serves as the focus of the current report.

### Duration Groups

Once the MDD symptom items of the PAPA were completed, interviewers reviewed/ summarized all depression symptoms that had been endorsed at clinically significant levels by caregivers. Caregivers were instructed to answer duration questions based on their observations of their children's expressions of these specific symptoms. To detail how duration is assessed on the PAPA, after discussing all MDD symptoms endorsed at a clinically relevant intensity, caregivers were asked whether they had observed their child expressing these symptoms for 2 or more consecutive hours during any given day in the past 6 months. Caregivers who reported no went on to the next PAPA module. Caregivers with children who had expressed MDD symptoms for 2 or more consecutive hours on a single day were then asked whether these episodes had ever occurred on 4 or more days in a single week. Caregivers with children who had symptom durations lasting 2 consecutive hours on 4 out of 7 days a week were then asked whether this type of episode had ever occurred for 2 or more consecutive weeks. Based on parental response children were divided into 2 groups. 1) DSM duration: defined by parental endorsement of depressive symptoms episodes that

lasted for 2 or more hours, for at least 4 out of 7 days a week, and that had at least once occurred for 2 consecutive weeks, and 2) <DSM duration: defined by parental endorsement of depressive symptoms episodes that lasted for 2 or more hours, for at least 4 out of 7 days a week, and that had *not* occurred for 2 consecutive weeks at any point in the last 6 months. As previously reported, of the children included in the <DSM duration group, 12 (24%) were reported as exhibiting MDD symptoms 2 or more hours in a day, 4 or more days a week but not for 2 consecutive weeks; 19 (38%) were reported as exhibiting MDD symptoms 2 or more hours a day but for fewer than 4 consecutive days in any given week; and 19 (38%) were reported as exhibiting MDD symptoms for less than 2 hours in any given day (Luby et al., 2009b).

### Co-morbidity

Co-morbid disorders were also assessed using the PAPA. A ‘yes/no’ variable was created based on the presence or absence of a co-occurring disorder (i.e., in addition to MDD). For the purpose of the analyses that follow co-morbid disorders were grouped into “internalizing” or “externalizing” categories. This variable was of interest in order to control for any effects of co-morbidity on the key outcome variables if group differences were found. Disorders included in the internalizing group were generalized anxiety disorder, separation anxiety, social phobia, and post traumatic stress disorder. Those included in the externalizing group were attention deficit-hyperactivity disorder, oppositional defiant disorder, and conduct disorder.

### MDD Severity Scores

MDD symptom severity scores were calculated for each participant using information gathered during the PAPA. Specifically, the total number out of 31 possible MDD symptom related items endorsed on the PAPA were summed and used to create severity scores at both the baseline and 2 year follow up assessments. Given that previous research has indicated that dimensional approaches incorporating such scores capture the developmental variation of MDD symptom manifestations in young children and related impairment and are sensitive indicators of depression severity, we chose to use all 31 items rather than the traditional 9 ‘core’ symptoms only (Luby et al., 2004b). This also allowed for a more fine grained analysis of potential differences between the clinical groups of interest.

### Parent and Teacher Report

The MacArthur Health and Behavior Questionnaire (HBQ) is a caregiver report measure that assesses the mental and physical health of young children in multiple contexts using information provided by caregivers (i.e., parent or teacher; Essex et al., 2002). Responses are scored dichotomously (e.g., yes or no) or on 3- or 4-point Likert scales ranging, for example, from 0 (“never or not true”) to 2 (“often or very true”). The HBQ provides dimensional ratings of function in a number of domains, including emotional and behavioral symptomatology, physical health, social adaptation, and school adaptation. Each domain is further divided into more content specific subscales (e.g., Depression). The current study focuses on the domain of emotional symptomatology (i.e., Internalizing) and its relevant subscales. The parent version of the HBQ evaluates internalizing symptoms using three subscales, Overanxious, Depression, and Separation Anxiety, while the teacher version only incorporates the Overanxious and Depression subscales. In addition, the HBQ provides a more global measure of impairment relevant to each context in which the child functions. For example, child specific difficulties such as not getting along with their parents or peers is measured on the Impairment-Child subscale while difficulties imposed on family functioning (e.g., interfering with family activities) as a result of the child’s behavior are rated on the Impairment-Family subscale. Similar difficulties at school, such as fighting with friends or diminishing performance, are measured by the teacher rated HBQ impairment



subscale. Higher scores indicate greater severity or impairment. The internal consistency, test-retest reliability, and discriminant validity of this measure have been previously established (Essex et al., 2002).

### Statistical Analyses

Chi-square ( $\chi^2$ ) tests and univariate analyses of variance (ANOVA) were used in order to examine between group differences in age, gender, caregiver marital status and ethnicity at baseline. When statistically significant between-group differences were found, follow up  $\chi^2$  and ANOVA tests were used to determine which specific groups differed.

The main effects of group on depression severity and HBQ scores at both time points (i.e., baseline and 2-year follow-up) were tested using ANOVAs. All between-group pairwise comparisons were examined using post hoc tests when the omnibus ANOVA was significant. Specifically, post hoc analyses were conducted using Tukey's Honestly Significant Difference procedure and, when the homogeneity of variance assumption was violated within the omnibus ANOVA, a Games-Howell correction for multiple comparisons was substituted. (Games and Howell, 1976) Additional analyses of covariance (ANCOVA) were conducted in order to control for the potential effects of age and co-morbid disorders and are noted where appropriate below.

In order to examine the influence of episode duration on the relative risk for an MDD diagnosis at the 2 year follow-up assessment, a  $\chi^2$  test of association was conducted.

## Results

### Demographics

No significant between group differences were found for gender. However, groups were found to differ in age. Specifically, the <DSM duration group was significantly older than the Healthy group at baseline. Therefore, age was controlled in analyses specifically involving the <DSM duration and Healthy groups. Additionally, differences in ethnicity, family income, and parental marital status were also found. That is, both the DSM and <DSM duration groups were found to have a greater number of children from ethnicities other than Caucasian or African American (<DSM N=8; DSM N=5; Healthy N=2), the <DSM duration group was found to have a higher percentage of lower income families than healthy controls, and the Healthy group had a higher percentage of married parents than either of the two clinical groups. It should be noted that similar demographic differences have been noted in a previous study of PO-MDD using an independent sample of preschoolers (Luby et al., 2002b; see Table 1).

### Participant Attrition

Children in the DSM duration group who did not participate in the 2 year follow up assessment (N=8) were compared to those who did (N=16) on all demographic and clinical variables collected at baseline. The group missing from the two year follow up assessment was found to have significantly higher levels of functional impairment at baseline as reported by their teacher on the HBQ (present  $\bar{x} = .37$ ; missing  $\bar{x} = .76$ ,  $p < .05$ ). No other differences were found. An identical comparison was carried out for the <DSM duration group. Children missing at the 2 year follow up (N=7) in the <DSM group did not differ from those who participated (N=43) on any variables relevant to the current study. By design of the planned analyses, no children were missing in the Healthy group at the 2 year follow up assessment.

### Co-morbidity

Children in the DSM duration group were more likely to have a co-morbid internalizing disorder when compared to the <DSM duration group. As a result, the presence of an internalizing disorder was controlled for in analyses specifically involving the DSM and <DSM duration groups. No between group differences were found for co-morbid externalizing disorders. (see Table 1)

### MDD Severity

Examination of MDD severity at baseline revealed that children meeting DSM-IV duration criteria exhibited significantly more symptoms than children in the <DSM duration group. Additionally, as expected based on study design, both groups were found to exhibit significantly more symptoms than the Healthy group at baseline. Follow up examination of MDD severity 2 years later indicated that the duration groups no longer differed from each other. However, both groups continued to be significantly more impaired than the Healthy group. (see Table 2 and Figure 1)

### Relative Risk for MDD at 2 year Follow-Up

Findings indicated that children in the DSM duration group were not significantly more likely than those in the <DSM duration group to receive a diagnosis of MDD using either adjusted (2-week criteria set aside) or formal DSM-IV criteria at the 2 year follow-up assessment (Adjusted criteria:  $\chi^2 = 2.2$ ,  $p = .137$ ; Risk Ratio=1.61, 95% CI [.89–2.9]; <DSM 15/43, DSM 9/16) (DSM criteria:  $\chi^2 = .99$ ,  $p = .318$ ; Risk Ratio=2.02, 95% CI [.51–8]; <DSM 4/43, DSM 3/16).

### Health and Behavior Questionnaire

Parent report on the HBQ at baseline revealed significantly more impairment for both duration groups on all areas assessed when compared to the Healthy group. In addition, parent report on the HBQ indicated that the DSM duration group was significantly more impaired on the family functioning subscale than the <DSM duration group at baseline as well.

Parent report data from the HBQ two years later revealed similar findings. That is, both duration groups continued to be significantly more impaired than the Healthy group in all areas assessed. However, the DSM duration group was no longer significantly more impaired than the <DSM duration group on the family functioning subscale. Further, while controlling for the presence of a co-morbid internalizing disorder, the DSM duration group demonstrated significantly higher scores than the <DSM duration group on the Overanxious subscale at this time as well. (see Table 3 and Figure 1)

Teacher report on the HBQ at baseline revealed significantly more impairment for the 2 duration groups on the Overanxious and Impairment subscales when compared to the Healthy group. In addition, the <DSM duration group scored significantly higher on the Internalizing scale than the Healthy group at baseline as well.

Similar findings emerged when teacher responses on the HBQ were compared 2 years later. As at baseline, both duration groups continued to score significantly higher than the Healthy group on the Overanxious and Impairment subscales and the <DSM group duration continued to be significantly higher than the Healthy group on the Internalizing scale. However, only the DSM duration group was significantly higher than the Healthy group on the Depression subscale at this time. It should be noted that parent- and teacher report HBQ scores were not available for some children at each time point. (see Tables 3 and 4 for the numbers included in each group)

## Discussion

The goal of the current study was to examine the influence of the strict 2-week episode duration criterion on depression symptom severity and impairment (teacher and parent rated) at baseline and 2 year follow up in a group of depressed and healthy preschool age children. Specifically, we investigated whether preschool children who met all criteria for MDD, including the 2-week duration criteria, differed in clinically meaningful ways from those who met all MDD symptom but not duration criteria. As hypothesized, a 'stepwise' relationship between the groups was found at the baseline assessment (see Figure 1). That is, based on parent report, both duration groups were found to be significantly more impaired in all areas when compared to the Healthy group at baseline. Further, the DSM duration group demonstrated significantly higher levels of MDD symptom severity and impairment in family functioning when compared to the <DSM duration group. Teacher report at baseline indicated that both duration groups were significantly more impaired than healthy controls in a number of areas as well, such as mood related symptomatology and classroom impairment. Results from the 2 year follow up assessment indicated a similar pattern of findings. As at baseline, both duration groups continued to demonstrate significantly greater levels of mood symptomatology and impairment at home and school when compared to the healthy control group. However, at 2 year follow up the duration groups no longer significantly differed in terms of depression symptom severity or impaired family functioning. Additionally, and of particular interest, the relative risk for a diagnosis of MDD using adjusted (2-week episode requirement set aside) or strict DSM-IV criteria at the 2 year follow up assessment did not differ based on episode duration at baseline (as defined by categorical group membership). Consistent with previous findings in PO-MDD of increased risk for depression later in childhood, this finding suggests that failing to meet 2-week episode duration during the preschool period does not diminish the risk for clinically significant manifestations of depressive symptomatology 2 years later (Luby et al., 2009b).

The current findings indicate that children meeting all DSM-IV MDD criteria except for 2-week episode duration exhibit a similar pattern of disorder related symptomatology and impairment to those who meet strict DSM-IV MDD criteria including 2-week duration. Additionally, these findings were stable across both rater (i.e., caregiver/teacher) and time (2 years), indicating that this was not a transient or context specific finding. As such, the current study suggests that adherence to the strict DSM-IV 2-week duration criteria fails to capture the entire group of preschoolers with clinically significant depressive symptoms. The finding of equivalent risk for a MDD diagnosis, using either formal or adjusted DSM (i.e. two week criteria set aside) criteria, in both duration groups two years later also suggests that the DSM-IV duration criteria did not confer greater predictive validity. However, given the small size of the DSM group at follow up and the possibility of a type II error, comparative studies including larger samples of children meeting the full DSM-IV criteria for MDD will be needed to explore this issue further.

Based on the current DSM taxonomy, studies of older children, adolescents, and adults have previously characterized differences similar to those identified in the current study as 'subthreshold' versus 'full syndrome' depressive disorders (Lewinsohn et al., 2000; Sadek and Bona, 2000; Pincus et al., 1999). While there continues to be debate surrounding this issue, the current findings suggest that such a dichotomy may not be clinically useful or meaningful when considering depressive episode duration in preschool onset major depressive disorder. Support for this conclusion comes from the highly similar pattern of significant differences that the duration groups demonstrated from healthy controls at both time points, with a few significant differences between the groups identified only at baseline. As such, while the 'stepwise' pattern of results at baseline and follow up may indicate meaningful differences between the duration groups, we feel that this is best



interpreted as differences in course rather than ‘kind’. That is, for children meeting the necessary symptom requirements for MDD, a categorical distinction, we suggest that duration may be more appropriately viewed dimensionally as an indicator of severity rather than dichotomously when defining ‘caseness’ at this early age. However, the limits of the current study data and developmental literature to date do not allow for any firm conclusions about identifying appropriate distinctions within episode duration (e.g., normative, mildly, moderately, or severely ill) and the potential clinical, rather than statistical, significance of relative differences in severity. Nevertheless, and while keeping these limitations in mind, study findings suggest that the 2 week duration criterion fails to identify a group of children who would likely benefit from further clinical attention and are at risk for continued mood related impairment in the future.

It should be noted that while the current findings could have implications for defining episode durations in MDD more generally (e.g. across the lifespan), we feel that any conjecture on this topic would be premature and instead intend for our comments to be constrained specifically to preschool onset major depression. This is of particular importance given the necessary consideration of emotional development at this age and the issues relevant to assessing early childhood psychopathology. For example, research with typically developing preschoolers has indicated that the demands of a given situation (e.g., expected positive response to a gift) may result in spontaneous expressive control, lessening their expression of negative emotions when an anticipated outcome does not occur (receipt of desired gift; Cole, 1986a). Additional work in typically developing preschoolers has also suggested that behavior following the experience of sadness remains context appropriate and similar (e.g., behavioral distraction), rather than distinctly different (e.g., giving up), to that seen for anger and happiness but at a reduced frequency (Dennis et al., 2009). Similarly, while observational work in depressed preschoolers has noted reduced enthusiasm and increased expressions of negative emotions during mildly stressful activities when compared to their typically developing peers, these differences were found to be most apparent in a severely affected subgroup of depressed preschoolers exhibiting anhedonia. As such, it may be that for some depressed preschoolers the frequency of observable symptomatology is more variable than what would be anticipated given our understanding of this disorder in older children and adults. Additionally, given the known limitations of preschoolers self-report on their internal experience and the possible variation in emotion regulation and expression at this age, parental report may be unable to provide a comprehensive narrative for the daily emotional experience of these children (Luby et al., 2007). Alternatively, though not mutually exclusive, parental expectations of increased variation in emotional expression during the preschool years, and a reluctance to recognize clinically negative mood states in such young children may lead to a greater tendency to ‘normalize’ symptoms that would be considered concerning behaviors at older ages.

Several limitations of the current study should be noted. While our sample represents the largest group of preschoolers with MDD studied to date to our knowledge, the sample size was not sufficient for examining depressive episode duration in finer increments beyond ‘less than 2 weeks’ (e.g., 10 days). However, in order to address whether those with the shortest durations were equally affected, we examined the subgroup of preschoolers who reported symptoms for <2 hours per day (38% of MDD sample). Notably, we found significant differences nearly identical to those reported for the <DSM group as a whole (all  $p < .01$ ; available upon request). Future studies including larger numbers of preschoolers with MDD are needed to more fully elaborate the potential implications of varying depressive episode criteria (i.e., dimensionally) at this age. Additionally, future studies examining episodicity in PO-MDD would likely benefit from incorporating a detailed dimensional approach to symptom duration (e.g., 0–12 hours) and episode duration (e.g., 0–14 days) measurement, rather than present or absent distinctions (e.g.,  $\geq 2$  hours) as used in the

current study. Incorporating such measures would likely also help to clarify the time course of symptoms and episodes of the disorder in early childhood. Given the relatively small number of DSM duration group children available at follow-up, and the short period between baseline and follow-up, it will be necessary to further examine the recurrence of MDD at longer intervals with larger groups to fully address the question of whether duration confers predictive validity. Nevertheless, to our knowledge this study represents the largest sample of depressed preschoolers yet available and thus we believe it provides unique information critical for informing this issue.

As noted in the introduction, our findings confirm the need for a more developmentally informed diagnostic conceptualization of MDD, including considerations of symptom durations, among preschool children. While this point has been demonstrated in previous research examining age appropriate manifestations of MDD related symptomatology (Luby et al., 2009a; Luby et al., 2003), the current study extends these findings into the area of depressive episode duration and suggests that the 2-week duration criteria may not be optimal for identifying all clinically relevant manifestations of depression in preschoolers. However, more research is required before firmly concluding that episode duration should be given less “weight” in diagnostic decisions at this age. Though beyond the scope of the current report, we believe that continued research into new and meaningful ways of assessing and conceptualizing early childhood psychopathology is of the utmost importance as we move towards a more developmentally informed DSM-V.

## Acknowledgments

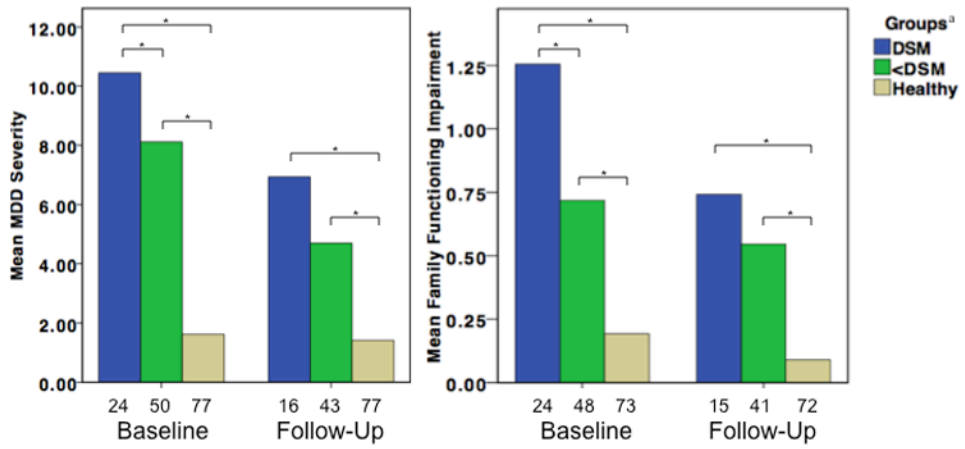
We wish to acknowledge the Early Emotional Development Program Staff, our preschool participants and their parents, and community recruiting sites whose participation and cooperation made this research possible.

## References

- Birmaher B, Axelson D, Strober M, Gill MK, Yang M, Ryan N, Goldstein B, Hunt J, Esposito-Smythers C, Iyengar S, Goldstein T, Chiapetta L, Keller M, Leonard H. Comparison of manic and depressive symptoms between children and adolescents with bipolar spectrum disorders. *Bipolar Disord.* 2009; 11:52–62. [PubMed: 19133966]
- Carta MG, Altamura AC, Hardoy MC, Pinna F, Medda S, Dell’Osso L, Carpiniello B, Angst J. Is recurrent brief depression an expression of mood spectrum disorders in young people? Results of a large community sample. *Eur Arch Psychiatry Clin Neurosci.* 2003; 253:149–53. [PubMed: 12904979]
- Cole PM. Children’s Spontaneous Control of Facial Expressions. *Child Dev.* 1986a; 57:1309–1321.
- Cole PM. Children’s Spontaneous Control of Facial Expression. *Child Dev.* 1986b; 57:1309–1321.
- Cole PM, Zahn-Waxler C, Fox NA, Usher BA, Welsh JD. Individual differences in emotion regulation and behavior problems in preschool children. *J Abnorm Psychol.* 1996; 105:518–29. [PubMed: 8952185]
- Denham, SA. Emotional development in young children. New York, NY: Guilford Press; 1998.
- Dennis TA, Cole PM, Wiggins CN, Cohen LH, Zalewski M. The Functional Organization of Preschool-Age Children’s Emotion Expressions and Actions in Challenging Situations. *Emotion.* 2009; 9:520–530. [PubMed: 19653775]
- Egger, HL.; Ascher, B.; Angold, A. The Preschool Age Psychiatric Assessment: Version 1.4. Duke University Medical Center; 1999, 2003.
- Egger HL, Erkanli A, Keeler G, Potts E, Walter B, Angold A. Test-Retest Reliability of the Preschool Age Psychiatric Assessment (PAPA). *J Am Acad Child Adolesc Psychiatry.* 2006; 45:538–549. [PubMed: 16601400]
- 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. *J Am Acad Child Adolesc Psychiatry*. 2002; 41:588–603. [PubMed: 12014792]
- Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Subthreshold depression in adolescence and mental health outcomes in adulthood. *Arch Gen Psychiatry*. 2005; 62:66–72. [PubMed: 15630074]
- Games PA, Howell JF. Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study. *Journal of Educational Statistics*. 1976; 1:113–125.
- Ialongo NS, Edelsohn G, Kellam SG. A further look at the prognostic power of young children's reports of depressed mood and feelings. *Child Dev*. 2001; 72:736–47. [PubMed: 11405579]
- Karlsson L, Marttunen M, Karlsson H, Kaprio J, Hillevi A. Minor change in the diagnostic threshold leads into major alteration in the prevalence estimate of depression. *J Affect Disord*. 2009
- Keenan K, Hipwell A, Feng X, Babinski D, Hinze A, Rischall M, Henneberger A. Subthreshold symptoms of depression in preadolescent girls are stable and predictive of depressive disorders. *J Am Acad Child Adolesc Psychiatry*. 2008; 47:1433–42. [PubMed: 19034189]
- Kendler KS, Gardner CO Jr. Boundaries of major depression: an evaluation of DSM-IV criteria. *Am J Psychiatry*. 1998; 155:172–7. [PubMed: 9464194]
- Klein DN. Classification of depressive disorders in the DSM-V: proposal for a two-dimension system. *J Abnorm Psychol*. 2008; 117:552–60. [PubMed: 18729608]
- Kovacs M, Lopez-Duran N. Prodromal symptoms and atypical affectivity as predictors of major depression in juveniles: implications for prevention. *J Child Psychol Psychiatry*. 2010
- Lewinsohn PM, Solomon A, Seeley JR, Zeiss A. Clinical implications of “subthreshold” depressive symptoms. *J Abnorm Psychol*. 2000; 109:345–51. [PubMed: 10895574]
- Luby J, Heffelfinger A, Koenig-McNaught A, Brown K, Spitznagel E. The preschool feelings checklist: A brief and sensitive screening measure for depression in young children. *J Am Acad Child Adolesc Psychiatry*. 2004a; 43:708–717. [PubMed: 15167087]
- Luby J, Heffelfinger A, Mrakotsky C, Hessler M, Brown K, Hildebrand T. Preschool major depressive disorder: preliminary validation for developmentally modified DSM-IV criteria. *J Am Acad Child Adolesc Psychiatry*. 2002a; 41:928–37. [PubMed: 12162628]
- Luby, JL.; Belden, AC. Mood disorders: phenomenology and a developmental emotion reactivity model. In: Luby, JL., editor. *Handbook of Preschool Mental Health: Development, Disorders, and Treatment*. New York: Guilford Press; 2006.
- Luby JL, Belden AC, Pautsch J, Si X, Spitznagel E. The clinical significance of preschool depression: Impairment in functioning and clinical markers of the disorder. *J Affect Disord*. 2009a; 112:111–119. [PubMed: 18486234]
- Luby JL, Belden AC, Sullivan J, Spitznagel E. Preschoolers' contribution to their diagnosis of depression and anxiety: Uses and limitations of young child self-report of symptoms. *Child Psychiatry Hum Dev*. 2007
- Luby JL, Heffelfinger AK, Mrakotsky C, Hessler MJ, Brown KM, Hildebrand T. Preschool major depressive disorder: preliminary validation for developmentally modified DSM-IV criteria. *J Am Acad Child Adolesc Psychiatry*. 2002b; 41:928–37. [PubMed: 12162628]
- Luby JL, Mrakotsky C, Heffelfinger A, Brown K, Hessler M, Spitznagel E. Modification of DSM-IV criteria for depressed preschool children. *Am J Psychiatry*. 2003; 160:1169–72. [PubMed: 12777277]
- Luby JL, Mrakotsky C, Heffelfinger A, Brown K, Spitznagel E. Characteristics of depressed preschoolers with and without anhedonia: evidence for a melancholic depressive subtype in young children. *Am J Psychiatry*. 2004b; 161:1998–2004. [PubMed: 15514399]
- Luby JL, Si X, Belden AC, Tandon M, Spitznagel E. Preschool depression: homotypic continuity and course over 24 months. *Arch Gen Psychiatry*. 2009b; 66:897–905. [PubMed: 19652129]
- Meiser-Stedman R, Smith P, Glucksman E, Yule W, Dalgleish T. The posttraumatic stress disorder diagnosis in preschool- and elementary school-age children exposed to motor vehicle accidents. *Am J Psychiatry*. 2008; 165:1326–37. [PubMed: 18676592]
- Pincus HA, Davis WW, McQueen LE. ‘Subthreshold’ mental disorders. A review and synthesis of studies on minor depression and other ‘brand names’. *Br J Psychiatry*. 1999; 174:288–96. [PubMed: 10533546]

- Saarni, C. Emotional competence: A developmental perspective. In: Bar-On, R.; Parker, JDA., editors. The handbook of emotional intelligence. San Francisco: Jossey-Bass; 2000.
- Sadek N, Bona J. Subsyndromal symptomatic depression: a new concept. *Depress Anxiety*. 2000; 12:30–9. [PubMed: 10999243]
- Schaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J Am Acad Child Adolesc Psychiatry*. 2000; 39:28–38. [PubMed: 10638065]



**Figure 1.** Group Differences In MDD Severity and HBQ Family Functioning Impairment at Baseline and 2 Year Follow Up.

Note. HBQ = Health Behavior Questionnaire; MDD severity is based on scores from the Preschool Age Psychiatric Assessment interview

\*  $p < .01$

<sup>a</sup> Number of participants for each group are noted under their representative bar



Table 1

## Demographic Characteristics of the Study Sample

Characteristic	Patients, Number (%)			$\chi^2$	P Value
	<DSM (N=50)	DSM (N=24)	Healthy (N=77)		
Ethnicity					
White	22(44)	14(58)	52(67)		
African American	20(40)	5(21)	23(30)	13.83	.01
Other <sup>a</sup>	8(16)	5(21)	2(3)		
Gender					
Male	28(56)	16(67)	34(44)		
Female	22(44)	8(33)	43(56)	4.28	.12
Family Income, \$ per year					
20,000	12(25)	5(23)	5(7)		
20,001–40,000	13(28)	6(27)	13(19)		
40,001–60,000	10(21)	2(9)	14(20)	14.82	.02
≥60,001	12(26)	9(41)	38(54)		
Caregiver marital status					
Married	22(44)	10(45)	53(69)		
Other <sup>b</sup>	28(56)	12(55)	24(31)	9.04	.01
Internalizing Disorders					
Present	17(34)	14(58)	na		
Absent	33(66)	10(42)	na	3.9	.04
Externalizing Disorders					
Present	30(61)	15(62)	na		
Absent	19(39)	9(38)	na	.01	.92
Age, mean(SD), years	4.7(.76)	4.6(.98)	4.3(.78)	$F_{2,148}=5.2$	.01

<sup>a</sup> Asian American, Latino, Biracial<sup>b</sup> Separated, divorced, widowed, single

**Table 2**  
Between Group Comparison of Depression Symptom Severity at Baseline and Two Year Follow Up.

Severity	Duration Group <sup>a,b</sup>				F Value	P Value	$\eta^2$ Value
	<DSM	DSM	Healthy	Healthy			
Baseline	<i>8.1(3)</i>	<b>10.5(4.2)</b>	1.6(1.5)	77	<i>F</i> <sub>2, 148</sub> = 150.9	.01	.67
Follow-Up	50	24	77	77	<i>F</i> <sub>2, 132</sub> = 36.8	.01	.36
	43	16	77				

Note. Data presented as mean (standard deviation) number in group

<sup>a</sup> Bold numbers indicate a significant difference with the Healthy group

<sup>b</sup> Within the duration groups all analyses were covaried for age and co-morbid internalizing disorder. Italicized numbers indicate a significant difference from the DSM duration group

**Table 3**  
Comparison Using the Health Behavior Questionnaire-Parent Version at Baseline and Follow Up

Scale	Duration Group <sup>a,b</sup>				F	P	$\eta^2$
	<DSM	DSM	Healthy	Baseline			
Internalizing	<b>.56(.25)</b>	<b>.71(.36)</b>	.26(.22)	73	$F_{2, 140} = 36.38$	.01	.34
	47	23					
<b>Subscales</b>							
Depression	<b>.48(.28)</b>	<b>.63(.39)</b>	.15(.23)	73	$F_{2, 140} = 37.06$	.01	.35
	47	23					
Overanxious	<b>.59(.3)</b>	<b>.8(.39)</b>	.32(.26)	72	$F_{2, 138} = 28.21$	.01	.3
	46	23					
Separation Anxiety	<b>.58(.35)</b>	<b>.7(.52)</b>	.31(.28)	73	$F_{2, 136} = 14.36$	.01	.17
	45	21					
Impairment Child	<b>.46(.35)</b>	<b>.67(.35)</b>	.11(.18)	73	$F_{2, 140} = 46.14$	.01	.39
	48	22					
Impairment Family	<b>.72(.59)</b>	<b>1.3(.73)</b>	.19(.24)	73	$F_{2, 142} = 49.17$	.01	.41
	48	24					
<b>Scale</b>							
Internalizing	<b>.46(.26)</b>	<b>.64(.25)</b>	.18(.17)	73	$F_{2, 125} = 41.57$	.01	.4
	41	14					
<b>Subscales</b>							
Depression	<b>.38(.31)</b>	<b>.5(.31)</b>	.09(.13)	73	$F_{2, 125} = 32.89$	.01	.35
	41	14					
Overanxious	<b>.56(.28)</b>	<b>.83(.32)</b>	.28(.24)	72	$F_{2, 124} = 34.1$	.01	.36
	41	14					
Separation Anxiety	<b>.45(.35)</b>	<b>.58(.32)</b>	.18(.22)	73	$F_{2, 124} = 18.49$	.01	.23
	41	13					
Impairment Child	<b>.38(.42)</b>	<b>.52(.31)</b>	.08(.15)	72	$F_{2, 125} = 24.49$	.01	.28
	41	15					
Impairment Family	<b>.55(.59)</b>	<b>.74(.53)</b>	.09(.19)	73	$F_{2, 125} = 26.45$	.01	.3

<b>Duration Group<sup>a,b</sup></b>		<b>F</b>	<b>P</b>	<b><math>\eta^2</math></b>
<b>&lt;DSM</b>	<b>Healthy</b>			
41	72	15		

Note. Data presented as mean (standard deviation) number in group

<sup>a</sup> Bold numbers indicate a significant difference with the Healthy group

<sup>b</sup> Within the duration groups all analyses were covaried for age and co-morbid internalizing disorder. Italicized numbers indicate a significant difference from the DSM duration group

**Table 4**  
Between Group Comparison Using the Health Behavior Questionnaire-Teacher Version at Baseline and 2 Year Follow Up

Scale	Teacher				
	Duration Group <sup>a,b</sup>		F	P	$\eta^2$
	<DSM	Healthy			
	<b>Baseline</b>				
Internalizing	<b>.41(.27)</b>	.24(.29) <sup>c</sup>	F <sub>2, 101</sub> = 4.38	.01	.1
	36	17	51		
	<b>Subscales</b>				
Depression	.36(.32)	.38(.38)	F <sub>2, 104</sub> = 2.62	.07	.05
	37	17	53		
Overanxious	<b>.43(.31)</b>	.26(.27)	F <sub>2, 101</sub> = 4.74	.01	.08
	36	17	51		
Impairment	<b>.49(.52)</b>	.17(.28)	F <sub>2, 102</sub> = 9.4	.01	.16
	37	15	53		
	<b>Follow-Up</b>				
Internalizing	<b>.44(.38)</b>	.55(.4) <sup>c</sup>	F <sub>2, 99</sub> = 6.7	.01	.12
	32	12	58		
	<b>Subscales</b>				
Depression	.33(.37) <sup>c</sup>	<b>.44(.39)</b>	F <sub>2, 101</sub> = 4.11	.02	.08
	33	12	59		
Overanxious	<b>.55(.43)</b>	.32(.26)	F <sub>2, 99</sub> = 5.91	.01	.12
	32	12	58		
Impairment	<b>.37(.39)</b>	.13(.22)	F <sub>2, 101</sub> = 10.5	.01	.17
	33	12	59		

Note. Data presented as mean (standard deviation) number in group

<sup>a</sup> Bold numbers indicate a significant difference with the Healthy group

<sup>b</sup> Within the duration groups all analyses were covaried for age and co-morbid internalizing disorder.

<sup>c</sup>  $p < .05$  prior to correction for multiple comparisons