This is a repository copy of *Longitudinal Profiles of Girls' Irritable, Defiant and Antagonistic Oppositional Symptoms: Evidence for Group Based Differences in Symptom Severity*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/108677/

Version: Accepted Version

The final publication is available at Springer via http://doi.org/10.1007/s10802-016-0231-z.

Abstract

Objective: Three subdimensions of ODD symptoms have been proposed –angry/irritable (IR), argumentative/defiant (DF) and antagonism (AN). This study tested whether longitudinal symptom trajectories could be identified by these subdimensions. **Method:** Group-based trajectory analysis was used to identify developmental trajectories of IR, DF and AN symptoms. Multi-group trajectory analysis was then used to identify how subdimension trajectories were linked together over time. Data were drawn from the Pittsburgh Girls Study (PGS; N=2450), an urban community sample of girls between the ages of five--eight at baseline. We included five waves of annual data across ages five-13 to model trajectories. **Results:** Three trajectories were identified for each ODD subdimension: DF and AN were characterized by high, medium and low severity groups; IR was characterized by low, medium stable, and high increasing groups. Multi-trajectory analysis confirmed these subdimensions were best linked together based on symptom severity. We did not identify girls' trajectory groups that were characterized predominantly by a particular subdimension of ODD symptoms. Membership in more severe symptom groups was significantly associated with worse outcomes five years later. **Conclusions:** In childhood and early adolescence girls with high levels of ODD symptoms can be identified, and these youth are characterized by a persistently elevated profile of IR, DF and AN symptoms. Further studies in clinical samples are required to examine the ICD-10 proposal that ODD with irritability is a distinct or more severe form of ODD.

Key words: Oppositional, trajectory, girls, longitudinal

Longitudinal profiles of girls' irritable, defiant and antagonistic oppositional symptoms:

Evidence for group based differences in symptom severity

There is a growing interest in characterizing the dimensionality of oppositional defiant disorder (ODD) symptoms. There is exceptionally strong and consistent support for the distinction of two dimensions distinguishing an affectively-oriented irritable dimension from at least one other clearly behaviorally-oriented dimension (Burke, Rowe, & Boylan, 2014; Herzhoff &Tackett, 2015; Lavigne, Bryant, Hopkins & Gouze, 2014).The ODD items which are best assigned to each dimension differ across studies. Some evidence also exists for a three-dimensional (irritability/anger, defiant/oppositional and vindictive/antagonistic) model of symptoms, as proposed in DSM-5 (American Psychiatric Association, 2013; Burke & Loeber, 2010; Stringaris & Goodman, 2009).

This paper is concerned with determining whether a dimension of irritability can be distinguished from one or more behavioral dimensions of ODD at the person level. If so, this would have bearing on the description of diagnostic subtypes of ODD. For the present review and analyses, we will use the term "irritability" to refer to a dimension of ODD symptoms that is characterizedby problems with being angry and touchy and that is distinct from behavioral problems such as arguing and defying which may also include vindictive behaviours.   We recognize that important questions remain about varying alternative models of these symptom dimensions.

Developmental trajectories of ODD symptoms

Several studies have examined developmental trajectories of ODD symptoms.

Thesstudies across ages five to 18 consistently find evidence forthree or four distinct trajectories of oppositional symptoms across childhood into adolescence (, Bongers, Koot, van der Ende & Verhulstet al, 2004; Boylan, Vaillancourt & Szatmari 2012; Leadbeater, Thompson & Gruppson, 2012; Nagin & Tremblay,1999). Across all studies, there is substantial replication of the trajectories observed where between 5 and 10% of the sample presents with high and persisting oppositional symptoms, 25% to 35% have moderate stable or decreasing oppositional symptoms, approximately 45% have low levels of symptoms which decrease over time and the remainder have no symptoms. Two studies identified no significant differences in either ODD symptom prevalence or trajectory shapes by sex (Bongers et al., 2004; Boylan et al., 2012).

If the distinct dimensions within ODD symptoms are useful to identify clinically, it might be expected that the course of each symptom cluster over time would be dissociable at the level of the individual, such that groups of individuals could be identified as having relative elevations of irritability, defiant or vindictive symptoms relative to the other dimensions.   No studies have yet examined ODD symptom dimension course over time using person centered methods. Irritable mood versus defiant ODD: Diagnostic subtypes of ODD?

Given that a diagnostic subtype must be identifiable at the level of the person to be clinically relevant, we next review studies which employ person-level data to discern what is known about how ODD symptom dimensions present across groups of youth. Five cross-sectional studies have used various measures and person-oriented techniques to identify whether youth can be identified as having a profile specific to irritability or defiant symptoms of ODD. At the time of writing, no study has examined whether there may be a profile (or possible subtype) of vindictive symptoms in youth.

There have been three publications about latent classes of ODD symptoms in youth

involving five community samples (Althoff, Kuny-Slock, Verhulst, Hudziak & van der Ende, 2014; Herzhoff & Tackett, 2016; Kuny et al., 2013). One of these studies investigated multiple age groups (Kuny et al., 2013) and one included eight ODD symptoms (Herzhoff & Tackett, 2015) allowing the most robust descriptive analysis of potential subgroups. Across these studies, several findings are replicated. First, in all, a large "no symptom" group is identified comprising approximately 70% of the sample. Three (confirmed by three of the studies) or four latent classes (confirmed by two of the studies) are identified. While all studies identify an oppositional/defiant symptoms only group, two studies identify a group with irritability symptoms only and three identify a group with mixed irritability and defiant symptoms each comprising about 10% of the sample. Of further interest, the study which compared multiple age groups suggested similarity in class membership from age seven to 12. Membership in the class distinguished by both symptom dimensions was least consistent across the age groups suggesting some fluctuation (to either the behavioral or the affective only classes) across development.

In all analyses from these studies, children in the irritability only class, and in some cases those in the mixed symptom class were more likely to show problems with mood and anxiety symptoms, or symptom indicators of anxious temperament than those in the defiant behavior class. This suggests that there are psychological differences between the groups based on ODD symptom profile.

A limitation in most of these latent class analyses is that the items at the core of the characterization of the groups had a low probability of assignment to their groups. Most notably, in Kuny et al. (2013) the symptom of "touchy or easily annoyed" had only a moderate probability (50%) of assignment to an "irritability" group in one sample. In Althoff et al. (2014) the groups distinguished as "Irritable Class" and "Defiant Class" showed very modest

differentiation in the probability of endorsement of "temper tantrums or hot temper" and "argues." While evidence from community samples using person-oriented analyses supports the distinction of affective from behavioral symptoms of ODD, there are groups of youth with concurrent irritable and defiant symptoms and class membership may shift over time.

Two studies have also conducted cross-sectional person-oriented analyses with clinical samples (Burke, 2012; Drabick & Gadow, 2012). In a sample of 177 boys between agesseven and 12, Burke (2012) found that 35% belonged in a latent class with both irritability and defiant symptoms along with a class distinguished by defiant behavior without irritability and a class low in all eight DSM-IV ODD symptoms. Similar to the preceding studies, the classes showing irritability displayed higher rates of anxiety and depression scores into young adulthood. Drabick and Gadow (2012) studied clustering of ODD symptoms in their clinical sample of 1160 youths aged six to 18. They found that, when using a cut-score of three irritability symptoms, 26% of the boys and girls in their study were categorized as irritable (along with defiant behavior); the remainder were categorized as having predominantly defiant behavior (along with two or fewer symptoms of irritability) or having fewer than four symptoms of ODD overall (a reference group). In this study, the group with irritability and defiant symptoms had significantly higher rates of symptoms of Generalized Anxiety Disorder, Major Depressive Disorder, and manic symptoms concurrently than all other groups.

Although these are only a small number of studies, key differences are evident between the clinical and community samples. Notably, thus far, clinical samples do not identify a group based on being irritable without also having oppositional behavioral features as well. It may be that the presence of both irritability and defiant behavior differentially influences whether a child engages in services, leading to a low rate of children in clinic samples with irritability alone.

There is some support for an irritability only group in community samples, but other studies are needed which have a full complement of ODD symptoms available to test the robustness of the potential sub-groups. While other studies are needed, these four suggest that ODD with irritability is a distinct form of the disorder, and the presence of irritability symptoms may imply increased risk for mood and anxiety problems relative to ODD defiant symptoms alone.

Current Study

The aim of the study was to test whether irritability could be distinguished from one or more behavioural dimensions of ODD at the person level. Testing this aim requires two component analyses that build on each other. The first step was to estimate trajectories of girls' irritable, defiant and antagonistic ODD dimensions independently of each other across childhood and early adolescence using latent class growth analysis (Nagin 2005). We chose to use this method as opposed to other person centered methods of data analysis such as growth mixture modeling (Muthén & Muthéen,2000) because it models population heterogeneity as distinct homogeneous groupings and is appropriate when there is an implicit assumption that groups are present. It is a good approach prior to testing the fallibility of an hypothesis with structural or growth modeling where group assignment is directly challenged (Jung & Wickrama, 2008). We chose to include a test of the antagonistic dimension at the person level because this dimension was identified when testing the factorial structure of ODD in the Pittsburgh Girls Study in a previous analysis (Burke, Hipwell & Loeber, 2010) and it is also identified as a pattern of ODD symptoms in DSM-5.

We used data from the Pittsburgh Girls' Study (PGS) for a female specific analysis for several reasons. Significantly less is known about the phenomenology and course of disruptive behaviours in girls, and specific focus on course of behaviour in girls is needed to identify group

differences within sex. This study which focuses on girls alone will have increased power to

identify between group differences such as a group with increasing irritability, or a group with

specifically high irritability. The PGS data contains the full complement of DSM-5 ODD

symptoms at all waves of the study to enable subdimension identification. Of particular

relevance for the study of ODD subdimensions in girls is that the prevalence of depression and

anxiety disorders increases exponentially in girls at, or post puberty (Angold, Costello &

Worthman, 1998; Hankin et al., 1998). This fact, in concert with the association of ODD

irritability with later depression in general, suggest that girls' trajectories of irritability may also

increase or plateau around the transition to adolescence. As no study has specifically examined

the question of subdimension trajectories, we do not know whether this profile exists. Other

trajectory studies of four to seven ODD symptoms grouped together (Boylan et al., 2012;

Leadbeater et al., 2012) have not identified sex differences in the trajectories and it may be the

result of not specifying factors specific to irritability and behavioural symptoms.

       Based on findings from studies of general ODD trajectories, we hypothesized that fewer

than 10% of girls belong to a group with persistently high irritability, defiant and antagonistic

ODD symptoms consistent with the prevalence of ODD in the general population. We also

expected groups of girls with trajectories of very low or decreasing levels of each ODD symptom

dimension (>50% of the sample). Given the significant increase in the incidence of new onset

internalizing disorders in girls in adolescence, and previous reports of association between ODD

irritability symptoms and depression and anxiety disorders, we similarly expected a group of

girls with an increasing trajectory of the irritability symptom dimension over time.

       The second part of the primary objective was to estimate how these group-based

symptom trajectories develop concurrently over time within individuals by modeling trajectory

groups identified across the three dimensions concurrently (i.e., multi-trajectory group analysis). The observed patterns of subdimension trajectory linkage within the multi-dimension trajectory groups will suggest whether there is subdimension-specificity in the presentation of ODD symptoms within individuals over time, providing robust evidence for the independence of the proposed symptom subdimensions.

We hypothesized that the ODD subdimensions would covary strongly at the person level; therefore, the majority of girls would belong to trajectory groups that show a similar intercept and slope for each of irritability, oppositional and antagonistic symptoms jointly over time. In particular, we hypothesized that we would identify a group with high level symptom trajectories with each of irritability, defiant and antagonistic behavior. In addition, we also anticipated finding a group high in defiant behavior without irritability, but not a group high in irritability without defiant behavior, as found by others (Burke, 2012; Drabick & Gadow, 2012; Herzhoff & Tackett, 2015). The relationship of antagonistic trajectories to the other classes was unclear, although it was thought to be more likely that antagonistic symptoms would covary developmentally with defiant as opposed to irritability symptoms.

Our secondary objective was to test the association between trajectory group membership and symptom measures including depressive, anxiety, conduct and ADHD symptom severity as a means of externally validating the trajectory groups. Consistent with the extant literature, we hypothesized that girls with trajectories of elevated irritability would have stronger associations with depression and anxiety outcomes than behavioral outcomes, and those with elevated oppositional or antagonistic symptom trajectories will have elevated behavioural problems as outcomes. We tested how multi-trajectory group membership predicts the various symptom measures, and hypothesized that the multi-trajectory group with high levels of all ODD

symptoms would have the highest level of symptom outcomes of all types compared to the no ODD symptom and medium level ODD symptom group. Were there to be a multi-trajectory group characterized by high levels of only one subdimension, we would expect significant association of that trajectory group with elevated symptoms as outlined above (irritability with depression and anxiety symptoms and other groups with externalizing (CD and ADHD) symptoms.

Methods

Participants

Data were drawn from the Pittsburgh Girls Study (PGS; N=2450), an urban community sample of four age-based cohorts of girls between the ages of five to eight at baseline assessment. These families have been followed annually; the present study includes data from the first five assessment waves spanning a period from five to13 years in the accelerated longitudinal design. Low income families were oversampled such that neighbourhoods in which at least 25% of families who were living at or below poverty level were fully enumerated and a random selection of 50% of all households in all other neighborhoods were enumerated (Hipwell et al., 2002 and Keenan et al., 2010provide details on study design and recruitment). African American girls comprised 52.9% of the sample (41.2% Caucasian, 4.9% multiracial, 0.9% Asian). At baseline, the majority of caretakers were female (92.3%), were cohabitating with a spouse or domestic partner (58.7%), and had completed at least 12 years of education (83.2%). Retention in the study was high, with follow up retention rates that ranged from 97.2% to 93.1%. Two demographic variables, race (1=Caucasian, 2=minority race) and receipt of public assistance at the first wave of the study (1=yes, 0=no), were included as covariates in the estimation of the trajectories.

Data were collected in the homes of the girls and interviews were conducted separately for the girl and the caregiver. A combination of computerized and paper data collection methods were used. Approval for all study procedures was obtained from the University of Pittsburgh Institutional Review Board, and participants received financial compensation for their participation.

Measures

Symptom trajectories. ODD symptom severity was assessed using caretaker reports on the Child Symptom Inventory – 4th Edition (CSI-4; Gadow & Sprafkin, 1994) when girls were ages 5-12 years. The CSI-4 includes DSM-IV symptoms scored on a 4-point rating scale (0=never to 3=very often). Symptom severity scores were used to create the ODD dimension ratings. The ODD items were assigned to the irritable (IR), defiant (DF) and antagonistic (AN) dimensions, consistent with Burke et al. (2010) and are shown in Table 1 with comparisons to the current DSM-5 dimensional structure. The irritable dimension consisted of "been angry and resentful," "been touchy or easily annoyed by others" and "taken anger out on others or tried to get even" (α (age 5) = .71 to α (age 12)=.80). The oppositional dimension included "argued with adults," "defied you or refused to do what you told her to do," and "lost her temper" (α (age 5) = .64 to α (age 12) =.80). The antagonistic dimension consisted of "blamed others for her own misbehaviour or mistakes" and "done things to deliberately annoy others" (α (age 5) = .56 to α (age 12) = .71)

Outcomes at Wave 6 (ages 10-13). ADHD, CD and Depression. Girls reported on symptoms of ADHD (alpha = .91), CD (alpha = .70), and depression (alpha = .74) at wave 6 using the Child Symptom Inventory (CSI-IV, Gadow & Sprafkin, 1994). The measure has good internal consistency and test-retest reliability for these symptom constructs . Anxiety was measured by

the total score on the parent-reported Screen for Child Anxiety and Related Emotional Disorders (SCARED, Birmaher et al., 1997). The 41 items in the scale measure generalized anxiety, separation anxiety, panic, school anxiety, and social phobia. Items are scored on a 3-point scale (0=not/hardly ever true to 2=true to very true). The SCARED has good reliability and has been shown to significantly discriminate between anxiety depression and disruptive disorders, as well as within individual anxiety disorders (Birmaher et al., 1997). Internal consistency was.91 in our sample.

Data Analytic Plan

Semi-parametric group-based trajectory modeling, or GBTM (Nagin, 2005), was used to identify the numbers and shapes of distinct trajectories of irritability (IR), defiant (DF) and antagonistic (AN) symptoms across ages fiveto 13 using five annual prospective waves of data. The ODD item assignments used in the study are contrasted to the proposed DSM-5 item structure in Table 1. Previous studies have shown that the ODD subdimension structure of the PGS data is best organized according to factors of irritability, defiant behaviour and antagonisticness using exploratory factor analysis (Burke et al.., 2010). The superiority of this item assignment compared to alternate models has been confirmed in recent studies testing two and three factor CFA models in unique large datasets (Herzhoff & Tackett, 2015; Lavigne et al., 2014). Each of these studies, as well as Burke et al. (2014), have also shown the superiority of two or three factor ODD models with this item assignment to be superior to a single ODD factor including all 8 items.With group-based trajectory modeling, individual variation over time is considered to be normally distributed within groups which each have distinct growth patterns. This assumption is useful for conceptualization and prognostication of clinical problem behaviour where the presence of separate groups with differing patterns of behaviour over time is

assumed. Within each trajectory group, the intercept and slope terms are assumed to be invariant

when they are modelled. Therefore, it is not practice to vary these terms when doing GBTM.

Invariance of intercepts and slopes across the trajectory groups was tested using the Wald test,

which is an extension of the PROC TRAJ SAS macro (Jones & Nagin, 2007). Models were

estimated using the PROC TRAJ macro in SAS 9.0 (Jones et al., 2001). Estimation of the

trajectories followed two steps: (1) selecting the number of trajectory groups, followed by (2)

estimating the slope (shape) of each trajectory. To estimate the optimal number of groups, we

used a censored normal model and several practical fit indices (Nagin, 2005). Maximization of

the Bayes Information Criterion (BIC) using any of one to five classes was first used to identify

the ideal number of groups described by the data (D'Unger et al., 1998). Very strong differences

between BIC scores were considered in the range of 10 or more units (Raftery, 1995) as

supported by the Bayes' criterion calculation (delta BIC between comparator models). To select

the shape of each group, quadratic, followed by linear and zero growth factor options were

specified and tested. The best fitting model corresponded to the one with significant growth and

intercept terms for each trajectory in the model, and BIC was maximized. In addition, entropy

(E) a measure of classification uncertainty, was included. Higher classification certainty is

supported by a lower entropy value.

When multiple alternate models presented with similar BIC scores, the selection of the

best model was based on a combination of (1) the highest (-) BIC score, (2) statistically

significant growth coefficients for all trajectory parameters, (3) posterior probabilities of correct

trajectory group assignment being greater than 70% or 0.7, and finally (4) the most parsimonious

model was achieved (Roeder, Lynch & Nagin, 1999; Nagin, 2005). Considering the accelerated

design of the study, the effect of baseline age on trajectory group membership was tested treating

cohort as a covariate during the trajectory estimation procedure.

Following estimation of the individual trajectories for each dimension of ODD, multi-trajectory group analysis proceeded to test how trajectories within each dimension are linked across dimensions within each group of girls. For example, if each oppositional behavior dimension (IR, DF, AN) is found to consist of 3 distinct trajectories (e.g., high, medium, low level of symptoms in each case), it is possible there could potentially be 27 groups of girls with distinct multi-trajectory profiles ($3^3 = 27$; each dimension with 3 trajectories combined with the other dimensions: e.g., girls characterized by high irritability, high defiant and high antagonistic behavior, and other girls characterized by membership in the moderate classes of all of these dimensions, etc.). The syntax for the analysis can be found in Jones and Nagin (2007). To test for the optimum number of multi-trajectory groups from the data, the researcher specifies models with varying numbers of expected multi-trajectory groups (three or more, as there are three oppositional dimensions in this case) and then identifies the best fitting model according to the approach reported previously. Each output identifies the specified number of groups (i.e., three) where the first group would be characterized by the trajectory shapes described by the first set of parameters for each model (i.e., low level trajectory for each of IR DF and AN, for example), the second group by the second sets (i.e., medium level trajectory for each of IR, DF and AN) and third group by the third sets (i.e., high level for each), although the data may identify other trajectory linkages (i.e., low IR, medium DF and medium AN, for example). The idea is that there could be resultant multi-trajectory groups that are best described as consisting of one trajectory from a particular dimension, but differing trajectories from other dimensions. The impact of minority race and household receipt of public assistance on trajectory group membership was examined by entering these variables as covariates in supplemental analyses.

ANOVA was used to compare mean scores of various outcomes measured at wave six across the multi trajectory groups. The approach used was to examine first for the presence of overall group effect for a particular covariate.   If present, the nature of significant differences in post hoc tests across the covariates was examined.

Results

Descriptive Statistics

The correlations between and within the dimensions were tested by assessment wave rather than age because multiple ages were assessed at each assessment wave (Table 2). Between wave correlations within the IR dimension ranged between r = .41 (wave 1 to 5) and r = .67 (wave 4 to 5), within the DF dimension ranged between r = .51 (wave 1 to 5) to r= .67 (wave 4 to 5) and within the AN dimension ranged between .41 (wave 1 to 5) to .56 (wave 4 to 5). Cross-sectional correlations between IR and DF dimensions ranged from r = .56 (wave 1, lowest) to .71(wave 2, highest), for IR and AN dimensions from r = .60 (wave 1, lowest) to r = .73 (wave 2, highest) and between DF and AN from r=.51 (wave 1, lowest) to r = .65 (wave 2, highest).

Mean scores for the IR dimension ranged between 1.66 (1.09) at age five to 1.76 (1.29) at age 13, for the AN dimension ranged between 1.59 (1.08) at age 5 to 1.41 (1.11) at age 13, and for the DF dimension ranged between 2.52 (1.63) at age 5 to 1.98 (1.61) and age 10, and to as high as 2.12 (1.70) at age 13.

Objective 1: Estimation of single dimension trajectories

Each data wave in PGS consisted of four age cohorts. Prior to modeling the trajectories we tested whether age cohort was significantly predictive of trajectory group membership. We found that in all models it was not, thus we modelled the data by wave of assessment and not age.

The developmental trajectories of IR were best characterized by three trajectory groups, BIC = -19137.57, Entropy ,,E,=1.25. Compared with a four group model (Table 2), BIC was substantially improved (a greater than 10 unit difference is considered significant (Raftery, 1995) and alternative slope selections for the three group model had significantly higher entropy and lower posterior group membership probabilities. These are displayed in Figure 1 and correspond to (1) a group with no symptoms (severity count of one or less) with a negative quadratic slope (39% of the sample), (2) a group with one IR symptom (56% of the sample) with no significant slope and (3) a moderate level (one or two symptoms) of IR with quadratic slope (5% of the sample). Intercepts and (linear and quadratic) slopes differed significantly across trajectory groups (data not shown). Comparison of fit indices between the chosen and next best fitting models is shown in Table 3, along with the posterior probabilities of correct assignment. These data support very high model specification and fit.

The best fitting model describing DF symptom trajectories consisted of three groups ,BIC = -18442.16, E = 1.28, as displayed in Figure 2. No solutions were possible for several specified 4-group models, and the 2-group model had a BIC fit of < -20000. The alternative next best fitting model was a 3-group model (group slopes 222), which also had substantially poorer fit than the chosen model (Table 2). There was (1) a no DF symptoms (severity score less than 2) trajectory group with linear negative slope (37% of the sample), (2) a low level DF symptoms trajectory group with a quadratic slope (55%) and (3) moderate to high DF symptoms group with a positive linear slope (8%). Intercepts and (linear and quadratic) slopes differed significantly across trajectory groups (data not shown). Alternative model fit indices and posterior probabilities of correct assignment for each trajectory group are displayed in Table 2. These data support a very high level of assignment accuracy for the oppositional behaviour trajectory

groups.

A three-group trajectories model (BIC = -16975.44, E = 1.00) maximized model fit and parsimony for the AN symptoms. Several two-, three- and four-group models were tested. The fit indices for two-group models were notably inferior to three- and four-group models. The three-group model was chosen over the next best fitting four -group model based on parsimony and because posterior probabilities of group assignment in two out of the four groups were significantly discrepant (60 and 68%) compared to the other groups, suggesting a three group model provides a better fit to the data. The three groups are displayed (Figure 3) and correspond to (1) a no AN symptoms (severity count less than two) group with quadratic slope (25% of the sample), (2) a low level AN symptoms group with a negative linear slope (68%) and (3) moderate to high AN symptoms group with a linear slope (7%). Intercepts and (linear and quadratic) slopes differed significantly across trajectory groups (data not shown). Posterior probabilities of correct assignment for the high level trajectory group in each trajectory group are displayed in Table 2. These data support a very high level of assignment accuracy for the AN trajectory groups.

The impact of minority race and household receipt of public assistance on trajectory group membership were examined by entering these variables as covariates in supplemental analyses. Specifically, we repeated all trajectory analyses adjusting for each of these covariates simultaneously. The trajectory estimates did not differ significantly with the covariates included and minority race was not predictive of trajectory membership. Receipt of public assistance was significantly predictive of membership in the medium and high versus low level antagonistic trajectory groups but was not related to membership in the IR or DF trajectory groups (data available from authors).

Objective 2: Multi-trajectory group estimation

Multi-trajectory group modelling identified only one model that had acceptable fit indices, as displayed in Table 4. This model identified three groups, a high multi-trajectory group comprised of high level trajectories of each symptom dimension (9.5% of the sample), a "medium" multi-trajectory group comprised of the middle severity trajectory of each symptom dimension (55.5% of the sample) and a low multi-trajectory group with the low level symptom trajectories of each dimension (35.0% of the sample). Posterior probability of correct assignment for the high multi-trajectory group was 95%, the medium multi-trajectory group was 95% and the low multi-trajectory group was 95%. As with the single trajectory models described above, these data support a very high level of assignment accuracy for the multi-trajectory group model.

Objective 3: Association of trajectory group membership with psychological outcomes (Table 5)

The high multi-dimension trajectory group (high IR, high BD and high AN) had significantly higher levels of internalizing (depressive and anxiety) and externalizing (CD and ADHD) symptoms compared to the medium and low level multi-dimension groups, which did not differ from each other on these outcomes.

We compared the mean levels of ODD symptoms endorsed in each multi-dimension trajectory group in relation to CSI-IV clinical cut points (T-scores) for low/non clinical (T-score <30), moderate (T-score range 59-70) or high (T-score $>$ 70) symptom severity. From this we can infer that girls in the low and middle level multi-dimension trajectory groups have very low levels of oppositionality (i.e., CSI oppositionality T-score $<$ 59) and likelihood of concurrent ODD diagnosis, whereas girls in the highest level trajectory group would have high severity of oppositionality and likelihood of ODD diagnosis (Tscore >70). Considering this, these data

estimate a prevalence of six-eight% of girls having high severity of ODD problem behaviours with a stable course across ages five to 12 in the sample.

Discussion

Ours is the first study to identify longitudinal trajectories of ODD subdimensions across childhood. We identified developmental trajectories of irritable (IR), defiant (DF) and antagonistic (AN) behaviour dimensions in girls in the Pittsburgh Girls Study. Each dimension was characterized by three trajectory groups corresponding to distinctly different severities of symptoms as well as symptom courses over time. Trajectory models had excellent fit and a high accuracy of group classification for each dimension tested, attesting to the robustness of the finding across dimensions. We did not identify a group of girls with specific elevations in DF symptoms only as initially hypothesized, nor did we find a particular association between trajectories of AN symptoms with DF as opposed to IR symptoms.

While stability was evident for the trajectories of most dimensions across ages five-13, a notable exception was the positive growth observed in the high IR trajectory group over this age period. In the low level multi-symptom trajectory group, symptom decline was observed for all three dimensions. Several other trajectory studies have reported reductions in ODD symptoms over time in girls of this age group (Boylan et al., 2012; Leadbeater et al., 2012), as well as persistence/stability of ODD symptoms in youth with high levels of oppositional symptoms at baseline (Bongers et al., 2004). Stability of irritability symptoms relative to desistance of defiance symptoms over time has been reported across ages 12-25 years (Leadbeater & Homel, 2015). The current study of late childhood and early adolescence observes comparatively more developmental change in irritability compared with defiance and AN symptom trajectories, most notably with increasing levels of irritability with age in girls. This finding should be considered

in relation to the robust increases in internalizing symptoms in girls in early adolescence, suggesting that irritability may be an indicator of risk for internalizing difficulties particularly in female children with disruptive behaviour. Other types of disruptive behaviours may be less relevant to risk for internalizing disorder in this age group.

Additional developmental trajectory studies are needed to reliably describe the course of irritability, DF and AN behaviours in adolescence, as this may be the period of greatest symptom change.

Multi-trajectory groups were also robustly identified. The IR, defiant and AN trajectories in each group were linked by severity. Girls classified as having a high score on one ODD subdimension were found to have high scores on all other subdimensions. Although the subdimensions were linked by severity, they were not necessarily linked by developmental course as the slopes of the trajectories differed across subdimension within the multi-trajectory groups.   For example, in the high multi-trajectory group, the slope of the IR trajectory was quadratic, the slope for DF was positive linear and the slope for AN was positive linear, with differing slopes across subdimension in the low and moderate level multi-dimension groups. This finding provides support that the person-level linkage of the subdimensions by severity is not the result of measurement artifact. It was not possible to formally test for differences in the slopes across the subdimensions within each multi-trajectory group, as the group membership is not independent.

Our findings suggest that the trajectories of subdimensions of ODD overlap substantially within individual girls. There was not evidence for a developmental course with predominantly IR, DF or AN features independent from the other dimensions. As with the single trajectory analyses, the high level multi-trajectory group was associated with greater levels of

psychopathology compared to the mid and low level multi-trajectory groups. Thus, unlike Kuny et al., (2013), who found evidence for stability within distinct dimensional classifications when examining transitions between time points, the present results suggest no such distinction when considering common developmental trajectories using group-based trajectory modeling.

Contrary to our expectations, these subdimensions did not appear to emerge as distinct trajectories, but covaried highly within the multi-trajectory groups. The present findings are concordant with the person-oriented analyses of Herzhoff and Tackett (2015), Drabick and Gadow (2012) and Burke (2012), who found that classes of children with elevated irritability also had elevated defiant symptoms, although in those studies, groups with defiant symptoms only were identified, suggesting some subdimension separation at the person level.

The most parsimonious explanation for the absence of IR, DF or AN only groups in this study may be the longitudinal nature of the modelling strategy. All published person-level studies of ODD subdimensions have used cross-sectional data, whereas we model the subdimensions across five annual waves of data. As such, unless membership in a specific IR, DF or AN predominant group is notably stable over time, it may not emerge as a distinct group in the trajectory analysis. It may be that a subtype of ODD with irritability can be distinguished and does suggest elevated risk for depression, but is not sufficiently stable over time to emerge as a distinct trajectory classification. This is of concern since a lack of developmental stability may call into question the clinical utility of a subtyping framework. Additional population based studies are needed to replicate this finding, and to address the issue of possible sex differences so as to better describe ODD subdimension-specific phenomenology at the person level.

Additional explanations for the absence of subdimension-specific (pure) groups may relate to the age or exclusively female nature of the sample. As the severity of depression and

anxiety symptoms increases exponentially in girls at, or post puberty (Angold, Costello & Worthman, 1998; Hankin et al., 1998), it may be that the post-pubertal years may be associated with the observation of a group with IR prominence in girls. In support of this, we did note that while the slope of the other trajectories decreased over time, the slope of the high irritability trajectory increased. At older ages, irritability may present in the relative absence of behavioural symptoms and the shape and membership proportion of group based trajectory models may change significantly in this situation.

The observed high developmental overlap observed across the ODD subdimensions in girls is reminiscent of other studies which similarly note substantially higher rates of co-occurrence or comorbidity of conduct and depressive symptoms in girls compared to boys (Keenan, Loeber & Green, 1999; Wiesner & Kim, 2006) and externalizing and internalizing comorbidity in general (Rudolph & Clarke, 2001). These differences may result from higher rates of externalizing behaviours in boys compared to girls (and thus externalizing behaviours can present in boys without internalizing behaviours) or that externalizing behaviour being less common in girls, is associated with more severe psychopathology and comorbidity when present (Keenan, Loeber & Green, 1999). Overall our findings suggest that while ODD is a heterogeneous disorder in terms of comorbidity and etiology, the presentation of symptoms is more consistent with severity as opposed to phenotypic heterogeneity – at least in pre and peri-pubertal girls. This finding does contrast with analyses of Herzhoff and Tackett (2015) who did not identify significant sex differences in the latent classes of ODD subfactors they examined.

It is important to consider the function of the present analytic strategy when interpreting these results. This modeling strategy is intended to be an approximation of the complex underlying reality of the developmental course of these subdimensions over time. The three

latent subdimensions in this paper are derived from a single construct (ODD), thus, when

modeled over time, the dimensions may overlap, or covary, substantially more than they differ.

These groupings of girls' levels of symptoms over time reflect the best balance of high within-

group similarly with high between-group distinctions. Had the data came from a restricted

sample, such as a clinical population, that balance might very well yield different groupings with

different taxonomical implications (i.e., smaller between-group distinctions). Group based

trajectory analysis - as compared with growth mixture modelling, another analytic strategy for

identifying longitudinal latent groups - assumes within group homogeneity, whereas the latter

does not. In growth mixture modelling, variance in within-group slope and intercept can be

formally tested. The use of GMM may serve as a rigorous follow up approach to ensure that the

severity based pattern of trajectory linkage we observed is robust, and important to consider as

our observed pattern of results has not been consistently reported in previously published latent

class studies (Althoff et al., 2014; Herzhoff and Tackett, 2015; Kuny et al., 2013).

Is ODD irritability important to identify in girls?

        Despite the strong developmental association between IR, DF and AN symptoms

observed in this study, the presence of irritability remains important to identify. The extant

literature about ODD subdimensions strongly suggests that irritability is associated with

depression and anxiety outcomes as opposed to behavioural outcomes. In this study, however,

because IR was not distinct from DF and AN in the multi-trajectory analysis, girls in the high

level multi-trajectory group also showed increases in ADHD and CD symptoms compared to

their peers. Other person-centered studies (Gadow & Drabick, 2012; Kuny et al., 2013) have

identified that the presence of high irritability in a group is associated with greater severity of

outcomes of all types, and Burke (2012) and Gadow & Drabick (2012) also suggest that

irritability does not present in the absence of oppositional symptoms. This is consistent with the

proposed revision to the International Classification of Diseases (ICD-11; Lochman et al., 2015),

which includes a specifier for ODD with irritability, rather than incorporating the new Disruptive

Mood Dysregulation Disorder to represent a classification of children with problems of elevated

irritability independent of oppositional or DF behavior.

Further, using bifactor modeling Burke et al. (2014) showed that the relationship of the

IR and oppositional subdimensions to each other may be best understood in the context of a

general ODD syndrome, despite the fact that each factor may have distinct causal origins.

Additional studies are needed to explore which biological predictors are specifically or generally

associated with the syndrome of ODD as opposed to its subdimensions.

Future directions

An important implication of this work is that ODD, or high levels of ODD symptoms, is

important to identify in girls as it appears to persist into adolescence and is predictive of negative

mental health outcomes. The ominous impact of childhood ODD in general has been shown by

others, indexing risk for relationship dysfunction in areas of school, work and family (Burke et

al., 2014) and income and educational attainment (Stringaris & Goodman, 2009). Our study

suggests that girls with ODD are at increased risk for depressive symptoms in adolescence and

ODD is likely to persist at clinical levels of severity in the presence of these depressive

symptoms. Our results indicate that both irritability and ODD are important to identify for

depression prevention purposes, possibly because ODD and irritability are highly associated

phenomena in girls. Second, while irritability predicts depression, in this study it is also

predictive of other mental health outcomes. In summary, irritable girls have a very high

likelihood of ODD diagnosis and also appear to be at increased risk for persisting impairment

across multiple domains as early adolescents. Irritability and the diagnosis of ODD may be more likely to occur independently in boys and this is an important question to examine prior to establishing that irritability is a specifier of severity in ODD.

From a measurement perspective, this study, as well as others (Burke & Loeber, 2010; Herzhoff & Tackett, 2015; Lavigne et al., 2014), have identified an item assignment for ODD subdimensions that is statistically robust, but is not the model proposed in DSM-5. This presents a problem for researchers in this area as it is not clear which is preferable to use when testing alternative models. There are many similarities between this configuration and the DSM-5 approach, therefore it is likely that results would have been similar is the DSM-5 approach was used to model the trajectories. Group based trajectories modelled using DSM-5 item assignments would help to address this issue.

Limitations

This study used parent report data on ODD symptoms because youth report was not available. For report of behavioural symptoms, parent report alone has been shown to be more valid that youth alone (Jensen et al., 1999). This approach may have limited the ability to disentangle the subdimensions, particularly of IR and AN, where the child may have a differing opinion as to the presence or severity of their symptoms (Angold & Costello, 1996). As a counterpoint, the external correlates upon which the trajectories were compared were reported by the youth (except for anxiety) whereas the trajectories were reported by the parents. This reduced the impact of shared method variance and increased the confidence of their value as independent correlates.

It is also important to recognize that these findings may not generalize to clinic-referred youth.

Conclusion

This is the first study to use GBTM to test whether unique subgroups of oppositional girls could be identified based on their longitudinal profile on three ODD subdimensions representing IR, DF and AN. We did not identify girls' trajectory groups that were characterized by a particular subdimension of ODD symptoms; subdimensions were linked by symptom severity. Severity of symptoms of all types was significantly associated with worse long term outcomes, although significant post-hoc associations were observed between high irritability trajectory membership with later anxiety and depressive symptoms consistent with previous literature. Findings suggest that in childhood and early adolescence girls with high levels of ODD symptoms can be identified, and these youth are characterized by a persistent elevated profile of IR, DF and AN symptoms. This study does not support the proposal that ODD with irritability is a distinct or more severe form of ODD in girls.

Compliance with ethical standards

Ethical approval: All procedures performed in this study were in accordance with the ethical standards of the instutitional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in this study.

References

Althoff, R.R., Kuny-Slock, A.V., Verhulst, F.C., Hudziak, J.J., & van der Ende, J.(2014). Classes of oppositional-defiant behavior: concurrent and predictive validity. Journal of Child Psychology and Psychiatry, 55, 1162-1171. American Psychiatric Association: Diagnostic and statistical manual of mental disorders. (2013). (Fifth ed.). Arlington, VA: American Psychiatric Association.

Angold, A., & Costello, E.J. (1996). Toward establishing an empirical basis for the diagnosis of oppositional defiant disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 35, 1205-1212.

Angold, A., Costello, E.J., & Worthman, C.M. (1998). Puberty and depression: The roles of age, pubertal status and pubertal timing. Psychological Medicine, 28, 51-61.

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & McKenzie Neer, S. (1997). The screen for child anxiety related emotional disorders (SCARED): scale construction and psychometric characteristics. Journal of American Academy of Child and Adolescent Psychiatry, 36, 545-553.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. Child Development, 75, 1523-1537.

Boylan, K. R., Vaillancourt, T., & Szatmari, P. (2012). Linking oppositional behaviour trajectories to the development of depressive symptoms in childhood. Child Psychiatry and Human Development, 43, 484-497.

Burke, J. D. (2012). An affective dimension within oppositional defiant disorder symptoms

   among boys: Personality and psychopathology outcomes into early childhood. Journal of

   Child Psychology and Psychiatry, 53, 1176-1183.

Burke, J. D., Hipwell, A. E., & Loeber, R. (2010). Dimensions of oppositional defiant disorder as

   predictors of depression and conduct disorder in preadolescent girls. Journal of the

   American Academy of Child & Adolescent Psychiatry, 49, 484-492.

Burke, J. D., Rowe, R., & Boylan, K. (2014). Functional outcomes of child and adolescent

   oppositional defiant disorder symptoms in young adult men. Journal of Child Psychology

   and Psychiatry, 55, 264-272.

Drabick, D. A. G., & Gadow, K. D., (2012). Deconstructing oppositional defiant disorder:

   Clinic-based evidence for an Anger/Irritability phenotype. Journal of American Academy of

   Child & Adolescent Psychiatry, 51, 384-393.

D'Unger,A., Land, K.C., McCall, P.L., & Nagin, D.S. (1998). How many latent classes of

   delinquent/criminal careers? results from mixed Poisson regression analyses of the London,

   Philadelphia and Racine cohort studies. American Journal of Sociology, 103, 1593-1960.

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E. & Domenech, J.M. (2012). Dimensions of

   oppositional defiant disorder in 3-year-old preschoolers. Journal of Child Psychology and

   Psychiatry, 53,1128-1138.Gadow, K. D., & Sprafkin, J. (1994). Child Symptom Inventories

   Manual. Stony Brook, NY: Checkmate Plus. 53, 1128-1138.

Hankin, B.L., Abramson, L.Y., Moffitt, T.E., Silva, P.A., McGee, R., & Angell, K.E. (1998).

   Development of depression from preadolescence to young adulthood: Emerging gender

   differences in a 10-year longitudinal study. Journal of Abnormal Psychology, 107, 128-140.

Herzhoff, K., & Tackett, J.L. (2016). Subfactors of oppositional defiant disorder: converging evidence from structural and latent class analyses. Journal of Child Psychology and Psychiatry,57, 18-29.

Hipwell, A. E., Loeber, R., Stouthamer-Loeber, M., Keenan, K., White, H.R., & Kroneman, L.(2002). Characteristics of girls with early onset disruptive and antisocial behaviour. Criminal Behavioural Mental Health, 12, 99-118.

Jensen, P.S., Rubio-Stipec, M., Canino, G., Bird, H.R., Dulcan, M.K., Schwab-Stone, M.E., & Lahey, B.B. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? Journal of the American Academy of Child & Adolescent Psychiatry, 38, 1569-1579.

Jones, B.L., Nagin, D.S. & Roeder, K. (2001). A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. Sociological Methods & Research, 29:373-394.

Jones, B.L., & Nagin, D.S. (2007). Advances in group-based trajectory modeling and an SAS procedure for estimating them. Sociological Methods & Research, 35, 542-571.Jung, T., & Wickrama, T.A.S. (2008). An introduction to latent class growth analysis and growth mixture modelling. Social and Personality Psychology, 2, 302-317.

Keenan, K., Hipwell, A., Chung, T., Stepp, S., Stouthamer-Loeber, M., Loeber, R., & McTigue, K. (2010) The Pittsburgh Girls Study: Overview and initial findings. Journal of Clinical Child and Adolescent Psychology, 39, 506-521.

Keenan, K., Loeber, R., & Green, S. (1999). Conduct disorder in girls: A review of the literature. Clinical Child & Family Psychological Review, 2, 3-19.

Krieger, F. V., Polanczyk, G. V., Goodman, R., Rohde, L. A., Graeff-Martins, A. S., Salum, G.,

.Stringaris, A. (2013). Dimensions of oppositionality in a brazilian community

      sample:Testing the DSM-5 proposal and etiological links. Journal of American Academy

      of Child & Adolescent Psychiatry, 52, 389-400.

Kuny, A. V., Althoff, R. R., Copeland, W., Bartels, M., Van Beijsterveldt, C. E. M., Baer, J., &

      Hudziak, J. J. (2013). Separating the domains of oppositional behavior: Comparing latent

      models of the conners' oppositional subscale. Journal of the American Academy of Child &

      Adolescent Psychiatry, 52, 172.

Lavigne, J.V., Bryant F., Hopkins, J., Gouze, K. (2014). Dimensions of oppositional defiant

      disorder in young children: Model comparisons, gender and longitudinal invariance. Journal

      of Abnormal Child Psychology, 43, 423-439.

Leadbeater, B. J., & Homel, J. (2015). Irritable and defiant sub-dimensions of ODD: Their

      stability and prediction of internalizing symptoms and conduct problems from adolescence

      to young adulthood. Journal of Abnormal Child Psychology, 43, 407-421.

Leadbeater, B. J., Thompson, K., & Gruppuso, V. (2012). Co-occurring trajectories of symptoms

      of anxiety, depression, and oppositional defiance from adolescence to young adulthood.

      Journal of Clinical Child and Adolescent Psychology, 41, 719-730.

Lochman, J.E., Evans, S.C., Burke, J.D., Roberts, M.C., Fite, P.J., Reed, G.M., de la la Peña,

      F.R., Matthys, W., Ezpeleta, L., Siddiqui, S., & Elena Garralda, M. (2015). An empirically

      based alternative to DSM-5's disruptive mood dysregulation disorder for ICD-11. World

      Psychiatry, 14, 30-33.

Maughan, B., Rowe., R., Messer, J., Goodman, R., & Meltzer, H. (2004). Conduct disorder and

    oppositional defiant disorder in a national sample: Developmental epidemiology. Journal of

    Child Psychology and Psychiatry, and Allied Disciplines, 45, 609-621.

Muthén B, Muthén LK. (2000). Integrating person-centered and variable-centered analyses:

    growth mixture modeling with latent trajectory classes. Alcohol Clinical Experimental

    Research, 24, 882-91.

Nagin, D. S. (2005). Group-Based Modeling of Development. Cambridge: Harvard University

    Press.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and

    hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child

    Development, 70, 1181-1196.

Raftery, A.E. (1995). Bayesian model selection in social research. Sociological Methodology

    25, 111-164.

Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its

    application to schizophrenia. The American Journal of Psychiatry, 126, 983-987.

Roeder, K., Lynch, K., & Nagin, D. (1999). Modeling uncertainty in latent class membership: A

    case study in criminology. Journal of the American Statistical Association, 94, 766-776.

Rowe, R., Costello, E., Angold, A., Copeland, W. E., & Maughan, B. (2010). Developmental

    pathways in oppositional defiant disorder and conduct disorder. Journal of Abnormal Child

    Psychology, 119, 726-738.

Rudolph, K. D., & Clark, A. G. (2001). Conceptions of relationships in children with depressive
       and aggressive symptoms: Social-cognitive distortion or reality? Journal of Abnormal Child
       Psychology, 29, 41-56.

Stringaris, A., Cohen, P., Pine, D. S., & Leibenluft, E. (2009). Adult outcomes of youth
       irritability: A 20-year Prospective Community-based study. American Journal of
       Psychiatry, 166, 1048-1054.

Stringaris, A., & Goodman, R. (2009). Three dimensions of oppositionality in youth. Journal of
       Child Psychology & Psychiatry, 50, 216-223.

Wiesner, M., & Kim, H.K. Co-occurring delinquency and depressive symptoms of adolescent
       boys and girls: A dual trajectory modeling approach. Developmental Psychology, 42, 122-
       135.