Behaviour Research and Therapy 87 (2016) 40–47

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Testing for interactive and non-linear effects of risk factors for binge eating and purging eating disorders

Karina L. Allen [a, b, c, *], Susan M. Byrne [a], Ross D. Crosby [d, e], Eric Stice [f]

[a] School of Psychology, The University of Western Australia, Australia
[b] Eating Disorders Service, Maudsley Hospital, South London and Maudsley NHS Foundation Trust, UK
[c] Institute of Psychiatry, Psychology and Neuroscience, Kings College London, UK
[d] Department of Psychiatry and Behavioral Science, University of North Dakota School of Medicine and Health Sciences, USA
[e] Department of Biostatistics, Neuropsychiatric Research Institute, USA
[f] Oregon Research Institute, USA

ARTICLE INFO

Article history:
Received 24 November 2015
Received in revised form
16 July 2016
Accepted 26 August 2016
Available online 27 August 2016

Keywords:
Eating disorders
Risk factors
Sex differences
Classification tree analysis
Prospective
Adolescents
Raine study

ABSTRACT

Almost no research has tested whether risk factors interact in the prediction of future eating disorder onset, which might suggest qualitatively distinct etiologic pathways. Accordingly, this prospective study tested for possible interactions between risk factors in the prediction of binge eating and purging eating disorders in adolescents. It also examined sex differences in pathways to risk. Two analytical approaches were used: (1) classification tree analysis (CTA), which is ideally suited to identifying non-linear interactions and the optimal cut-points for defining risk, with follow-up random forest analyses; and (2) two-way interaction terms in a series of logistic regression models. Data were drawn from the Western Australian Pregnancy Cohort (Raine) Study, a population-based study that followed participants from pre-birth to young adulthood. This study involved 1297 adolescents (49% male), 146 (11%) of whom developed bulimia nervosa, binge eating disorder or purging disorder in late adolescence. In CTA, sex was the first and most potent predictor of eating disorder risk with females showing a 5-fold increase in risk relative to males. For males and females, weight and eating concerns were the next most potent predictor of risk and three risk groups emerged, reflecting non-linear risk. For females with intermediate weight and eating concerns, externalizing problems emerged as an additional predictor. Interaction terms in logistic regression models did not produce significant results after correcting for multiple testing. Findings advance knowledge on risk pathways to eating disorder onset, highlight non-linear risk processes, and provide cut-points for prospectively identifying high-risk youth for prevention programs.

Eating disorders are serious mental illnesses that affect up to 15% of adolescent females and 3% of adolescent males (Allen, Byrne, Oddy, & Crosby, 2013a; Stice, Marti, & Rohde, 2013). These disorders are associated with long-term psychosocial impairment (Johnson, Cohen, Kasen, & Brook, 2002; Solmi et al., 2015), carry increased mortality (Berkman, Lohr, & Bulik, 2007), and are difficult and expensive to treat (Begg et al., 2007; Simon, Schmidt, & Pilling, 2005). An improved understanding of the risk factors that predict future onset of eating disorders is vital for developing optimally effective prevention programs and for identifying the youth most in need of these programs. This is important because even the most effective prevention programs could produce larger effects.

Factors shown to predict eating disorder onset in at least two prospective studies include perceived pressure to be thin, body dissatisfaction, weight and eating concerns, negative affect or depressive symptoms, dietary restraint, and social support deficits (e.g., Allen, Byrne, Oddy, Schmidt, & Crosby, 2014; Beato-Fernandez, Rodriguez-Cano, Belmonte-Llario, & Martinez-Delgado, 2004; Ghaderi & Scott, 2001; Martinez-Gonzalez et al., 2003; The McKnight Investigators, 2003). Weight-related constructs have also been found to predict disorder onset, in the form of Body Mass Index (BMI), childhood overweight, and/or parent-perceived childhood overweight (e.g., Allen et al., 2014; Stice, 2016). These findings primarily relate to eating disorders characterized by binge eating and purging (subsequently referred to as binge eating and purging disorders: bulimia nervosa [BN], binge eating disorder [BED] and purging disorder [PD]), rather than

* Corresponding author. Maudsley Hospital, Denmark Hill, London, SE5 8AZ, UK.
E-mail address: karina.allen@uwa.edu.au (K.L. Allen).

http://dx.doi.org/10.1016/j.brat.2016.08.019
0005-7967/

anorexia nervosa (AN). However, very few studies have tested for interactions between risk factors, which is an important focus because we know that risk factors do not operate independently (Jacobi, Hayward, de Zwaan, Kraemer, & Agras, 2004). The current study sought to extend work on risk factors for binge eating and purging disorders, with a particular focus on how risk factors may interact to increase the likelihood of these disorders developing.

Interaction effects can be investigated using cross-product terms in traditional regression analyses. However, classification tree analysis (CTA), a recursive partitioning analytic approach, is ideally suited to this work because it can identify non-linear associations between risk factors and the outcome of interest, and determine specific and optimal cut-points for defining high risk groups. The latter feature is particularly useful in psychological research because many proposed risk factors occur on a continuum. Identifying specific cut-points for risk would guide the implementation of prevention programs to at-risk sub-populations after screening a full population.

Only three studies have used CTA to identify predictors of eating disorder onset. The first (Stice, Marti, & Durant, 2011) identified a three-way interaction between body dissatisfaction, depression, and dieting in the prediction of eating disorder onset in a community sample of adolescent girls (n = 496). Girls in the top 24% of body dissatisfaction showed an incidence of eating disorder onset of 24% versus 6% for those with lower body dissatisfaction. In the high dissatisfaction group, girls in the top 32% of depressive symptoms showed an eating disorder incidence of 43% versus 15% for those with lower depressive symptoms. In the low body dissatisfaction group, girls in the top 12% of dieting showed an eating disorder incidence of 18% versus 5% for those lower in dieting. These results suggest that prevention programs may be beneficial for girls with high body dissatisfaction and high depressive symptoms, and girls with low body dissatisfaction but high dieting frequency.

The second study (Jacobi et al., 2011) involved young women with weight concerns from the control condition of a prevention trial (n = 236). It found that negative comments by a coach or teacher about eating showed the strongest relation to future eating disorder onset. Women reporting negative comments showed an eating disorder incidence of 39% versus 8% for those who did not receive such comments. For women who did not receive comments, those with a history of major depression showed an eating disorder incidence of 30% relative to 4% for those without a history of depression. These results suggest targeted prevention may be helpful for women with weight concerns who have received negative comments about weight, as well as those who have not received such comments but who have a history of major depression.

The third study (Stice, Marti, Spoor, Presnell, & Shaw, 2008) used data from a randomized prevention trial with adolescent girls who reported body dissatisfaction (n = 481) and found that denial of the costs of pursuing the thin ideal was the most potent predictor of eating disorder onset. Girls in the top 17% of denial showed eating disorder incidence of 23% versus 6% for girls with lower denial scores. Among girls with high denial, eating disorder incidence was 0% for those who completed a dissonance-based prevention program versus 18% for those who completed alternative prevention programs and 50% for those in the control condition. Among girls with low denial, those in the upper 18% of emotional eating showed a 16% eating disorder incidence versus 4% for those with lower emotional eating. Among girls with low denial and low emotional eating, those in the upper 17% of externalizing symptoms showed an eating disorder incidence of 11% versus 2% for those with lower externalizing symptoms. These results suggest that prevention programs should be targeted towards young women who deny the costs of pursuing the thin ideal, as well as those who are low in denial but high on emotional eating, and those who are low in denial and emotional eating but who report high externalizing problems.

Although these CTA studies have elucidated interactions between risk factors, only one used a general community recruited sample (Stice et al., 2011) and none included boys, essential to identify sex-specific risk processes. Indeed, almost no prospective risk factor studies have been conducted with males, despite over 30% reporting at least occasional binge eating, purging or overeating (e.g., Field et al., 2014) and up to 3% meeting criteria for a binge eating or purging disorder (Allen et al., 2013a; Field et al., 2014). One recent study found that dieting and extreme weight control behaviors predicted later onset of binge eating and eating disorders for females but not males (Liechty & Lee, 2013), which highlights that risk pathways may differ by sex. Identifying pathways to male eating disorders is important if prevention initiatives are to accommodate young men at risk of eating pathology.

We have previously reported on risk factors for binge eating and purging eating disorders in the Western Australian Pregnancy Cohort (Raine) Study. In these analyses, female sex, parent-perceived childhood overweight, and weight and eating concerns emerged as the most potent predictors of early (14-year) and later (17/20-year) onset binge eating and purging disorders (Allen, Byrne, Forbes, & Oddy, 2009; Allen et al., 2014). However, analyses to date have not considered sex differences. Interactions between risk factors in the prediction of eating disorder onset are also unexplored.

The current report sought to identify interactions between risk factors in early-middle adolescence, and the specific cut-points on these factors, that could best predict onset of binge eating and purging disorders in later adolescence. Sex differences were also investigated. We used both CTA and cross-product terms in traditional logistic regression models to address this aim, to determine whether the former analytic approach is more sensitive and whether there is evidence that traditional cross-product terms miss non-linear relations. We focused on adolescence because most eating disorder symptoms appear to emerge during this period (Stice et al., 2013) and most prevention trials target adolescents (Stice & Shaw, 2004). Identifying the adolescent variables that best predict eating disorder onset, and specific risk cut-points for these variables, will facilitate selection of youth most in need of prevention work. Again, this is important because prevention programs are most effective when targeting high risk groups and, to date, even the most effective programs have produced only moderate effects (Stice & Shaw, 2004).

Given the small number of studies examining interactions between eating disorder risk factors, no specific predictions were made regarding the results. However, predictor variables were selected a priori based on findings from previous eating disorder studies with CTA (Jacobi et al., 2011; Stice, Rohde, Gau, & Shaw, 2012; Stice et al., 2011) and with the Raine Study (Allen et al., 2009, 2014). Thus, our predictors of interest included dietary restraint, weight and eating concerns, depressive symptoms, internalizing problems (capturing negative affect, anxiety, social problems and somatic complaints), externalizing problems (capturing inattention, rule-breaking behavior and anti-social/aggressive behavior) and body mass index (BMI), as well as participant sex. We were not in a position to assess thin-ideal internalization, perceived pressure to be thin or social support as these variables were not measured. We chose to include a specific measure of depressive symptoms as well as the broader internalizing scale, as previous studies have assessed negative affect in different ways (Jacobi et al., 2004; Stice, 2016) and we were interested in whether one measure would provide better predictive utility.

1. Method

1.1. Design and participants

Full details of the Raine Study are available elsewhere (Allen et al., 2009; Allen et al., 2013a; Newnham, Evans, Michael, Stanley, & Landau, 1993). In brief, 2900 women were recruited between 16 and 20 weeks gestation from the major public maternity hospital in Perth, Western Australia. Recruitment occurred between May 1989 and November 1991. Of the 2900 women enrolled in the study, 2868 delivered live birth children. Assessments were conducted with mothers and their offspring at birth and ages 1, 2, 3, 5, 8, 10, 14, 17 and 20 years.

Eating disorder symptoms were assessed in offspring at ages 14, 17 and 20. There were 1383 adolescents (49% male) who provided eating disorder data at age 14 and at least one of the subsequent two assessments. This sample represented 59% of the sample eligible to participate in the 14 through 20-year assessments (i.e., of those not deceased and not previously withdrawn from the study; n = 2344) and 48% of the original cohort. Families lost to follow-up were more socially disadvantaged than those who remained in the study but the Raine Study initially over-sampled disadvantaged families (see Allen, Crosby, Oddy, & Byrne, 2013 for full details).

Missing eating disorder data were imputed for adolescents who completed the 14-year assessment but missed either the 17 or 20-year assessment, using expectation maximization (EM) imputation with maximum likelihood. This applied to 141 (10.2%) cases at age 17 and 140 (10.1%) cases at age 20. No evidence was found to suggest that data were not missing at random (Little's MCAR X^2 [1383] = 1376, $p = 0.664$) and results did not differ when using the imputed and non-imputed datasets. Results from the imputed dataset are reported.

The dependent variable was first onset of BN, BED or PD at age 17 or 20. We have previously found evidence for shared risk pathways to BN, PD and BED in the Raine Study sample (Allen, Byrne, & Crosby, 2015), which supports the use of a single binge eating and purging disorder outcome group. Participants with an eating disorder at age 14 (n = 74) and those with AN or atypical AN in later adolescence (n = 12) were excluded (there were too few AN cases to reliably analyze, and findings from prospective studies suggest that risk factors for AN may be qualitatively distinct from risk factors for binge eating and purging disorders) (Stice, 2016). This resulted in an effective sample of 1297 adolescents (49% male), with 146 participants (18% male; n = 26) developing an eating disorder in later adolescence (81 BN; 43 BED; 22 PD). Consistent with the ethnic make-up of Western Australia, 82% of participants were Caucasian. One-fifth of families (n = 279; 21%) had an annual family income in the lowest two Australian income quintiles. In 30% of cases (n = 906) the child's biological father did not live in the family home.

1.2. Measures

1.2.1. Eating disorder diagnoses

Raine Study eating disorder assessment items, and diagnostic algorithms for determining DSM-5 eating disorders, have been described in full (Allen et al., 2013a; Allen et al., 2013). In sum, eating disorder symptoms were assessed using 24 self-report items adapted from the Eating Disorder Examination-Questionnaire (EDE-Q) (Fairburn & Beglin, 1994). The main adaptation was use of a simplified 4-point response scale, intended to facilitate accurate reporting by adolescents. The scale ranged from 0 ("Not at all") to 3 ("Most of the time [every day or nearly every day]) (Allen et al., 2013). Diagnoses were based on responses to the EDE-Q diagnostic items plus measured height and weight. Convergence between questionnaire and interview assessment of eating disorder symptoms is acceptable based on reported kappa values (e.g., range = 0.57–0.83) (Berg et al., 2012; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Stice, Fisher, & Martinez, 2004; Stice, Telch, & Rizvi, 2000).

Diagnoses of BN and PD were made according to DSM-5 criteria (Allen et al., 2013a). Questionnaire items did not allow for identification of sub-threshold BN or collect information on criterion B for BED (i.e., whether three symptoms relating to dysregulated eating or distress over eating are present). Over-evaluation of weight or shape was used instead of this BED criterion (Allen et al., 2013a). Others have shown that over-evaluation of weight and shape reliably distinguishes individuals with BED from individuals who binge eat without clinical impairment (e.g., Hrabosky, Masheb, White, & Grilo, 2007). Nonetheless, our BED definition does differ from strictly defined DSM-5 BED.

1.2.2. Potential predictor variables

Continuous predictors included *BMI* (calculated using measured height and weight); *eating, weight and shape concerns* on the adapted EDE-Q (referred to as *weight and eating concerns* for simplicity and described below); *dietary restraint* on the adapted EDE-Q; *depression* on the Beck Depression Inventory-Youth (BDI-Y) (Beck, Beck, & Jolly, 2001); *internalizing problems* on the Youth Self-Report (YSR; negative affect, anxiety, social problems and somatic complaints) (Achenbach, 1991); and *externalizing problems* on the YSR (inattention, rule-breaking behavior and anti-social/aggressive behavior). Participant *sex* was also used as a categorical predictor.

The EDE-Q items relating to eating, weight and shape concerns were combined on a single scale as previous analyses have shown the eating concern, weight concern and shape concern items to be highly correlated (Allen et al., 2015). Further, this index has previously been identified as a potent predictor of eating disorder onset within the Raine Study sample (Allen et al., 2014, 2015). It may be seen as capturing body dissatisfaction as well as specific concerns about eating (e.g., eating in front of others, guilt after eating). In this sample, correlations between the separate Eating Concern, Weight Concern and Shape Concern subscales ranged from 0.73 to 0.90 ($ps < 0.001$) across ages 14 and 17. The EDE-Q items used to calculate the weight and eating concern score are separate from those used to determine eating disorder diagnoses, with the exception of two items assessing over-evaluation of weight and shape, which contribute to the weight and eating concern scale and the diagnostic algorithms for BN and BED. As in previous studies (Allen et al., 2015), running analyses with these items omitted from the weight and eating concern scale did not change the pattern of results. As such, we retained the over-evaluation of weight and shape items in line with traditional EDE-Q scoring.

The EDE-Q, BDI-Y and YSR all have well-established psychometric properties (Achenbach, 1991; Beck et al., 2001; Fairburn & Beglin, 1994). In this sample, alpha coefficients at age 14 were 0.93 for weight and eating concerns, 0.78 for dietary restraint, 0.97 for depression, 0.74 for internalizing problems and 0.79 for externalizing problems.

1.3. Procedure

Self-report questionnaires were posted to adolescents at ages 14, 17 and 20 years, for at-home completion prior to attendance at a face-to-face assessment. Height and weight were measured during the assessment by a trained research assistant. Data collection was approved by the ethics committees of Princess Margaret Hospital for Children and the University of Western Australia.

1.4. Statistical analysis

Consistent with past research (Stice et al., 2011), continuous predictors were lagged to reflect the value at the assessment prior to eating disorder onset. Thus, for participants who developed a disorder at age 17, values at age 14 were entered as predictors and 20-year data were not used. For participants who developed a disorder at age 20, and for participants who did not develop an eating disorder by age 20, values at age 17 were entered as predictors. This means that predictors were lagged by 3 years for all participants. The same assessment measures were used at 14 and 17. Previous analyses suggest that risk factors for eating disorders that develop in early adolescence are similar to those for eating disorders that develop in later adolescence (Allen et al., 2014). All continuous predictor variables were normalized prior to analysis (using Blom's method, a rank-based transformation) to facilitate interpretation of parameter estimates for the various predictors.

Two sets of analyses were conducted. The first used CTA with follow-up random forest analysis. Classification tree analysis is an exploratory, data-driven, recursive partitioning analytic approach that seeks to explain outcomes on a single categorical variable (Loh, 2011). It is hierarchical in nature (unlike traditional regression analyses) and identifies optimal cut-points on predictor variables. After the first and most potent predictor in the full sample is identified, the model then searches for the most potent predictor in the successive spits of the sample in a recursive fashion. Analyses were conducted in SPSS Statistics Version 22 and used the CHAID method (Chi-Square Automatic Interaction Detection), which allows for non-binary splits and is particularly well suited to large data sets (Ripley, 1996). Since each split point (decision point) in CTA corresponds to a statistical test, Bonferroni corrections are applied automatically. The smallest possible group size was set to 20, to set limits on over-fitting the model to the data and increase the likelihood that results will replicate (Stice et al., 2011). As child sex has previously emerged as the most potent predictor of eating disorder risk within the Raine sample (Allen et al., 2009, 2014), and we were specifically interested in sex differences, sex was specified as the first split within the CTA model. Beyond this specification, CTA determined which predictors should enter the model and in which combinations. Any time that different variables emerged as the next most potent predictor of eating disorder onset from the same node, it signified that the effects of those second level predictors were different depending on the variable used to make the first split. This is the form interaction effects take in CTA (Camp & Slattery, 2002; Lemon, Roy, Clark, Friedmann, & Rakowski, 2003).

Whereas CTA generates a single classification tree predicting the outcome of interest, random forest analysis builds an ensemble ('forest') of classification trees using bootstrap estimates of the data. Further, whilst CTA produces decision splits using the best of all possible predictor variables, random forest analysis produces splits using the best of a randomly selected subset of variables (Breiman, 2001). By combining bootstrap aggregation ("bagging") and random variable selection, this approach produces a forest of classification trees that are aggregated to produce overall estimates of predictor importance and of prediction error. The error estimate is known as the aggregated "out of bag" or OOB error, and refers to how well the data *not* in the bootstrap samples are predicted by the classification trees generated *with* the bootstrap samples (Liaw & Wiener, 2002).

Random forest analysis does not produce a single classification tree or specific cut-points for the identified predictor variables. However, when used with CTA, it can help to guard against the overfitting of CTA models. We used it for this purpose, and ran random forest analyses using all possible predictor variables and then only the predictors obtained in CTA. Support for the CTA model would come from (i) a lower aggregated OOB for the random forest analyses using the CTA predictors, compared to the analyses using all possible predictors, and (ii) random forest analyses ranking the predictors included in the CTA model as important. Random forest analyses were run using the random forest package in R, which can be operated from SPSS (Liaw & Wiener, 2002). The default setting of 500 trees was retained.

The second set of analyses involved manual testing of all possible two-way interactions between predictor variables, in a series of logistic regression models with eating disorder onset as the categorical outcome variable. Each model included the normalized scores for the two predictors of interest (e.g., BMI and Restraint) and the interaction term between these predictors (e.g., BMI x Restraint). These analyses were conducted to compare findings from CTA with those from traditional interaction testing within a regression framework. Analyses were run separately for male and female participants, but where differences in results emerged, 3-way interactions with sex (e.g., Sex x BMI x Restraint) were specified to test for the statistical significance of any sex differences. In total, 15 two-way interactions were tested. A Bonferroni correction was used, giving an alpha level of 0.0033. Power calculations were conducted using the procedure outlined by Demidenko (2008), which allows for an estimate of power taking into account proposed interactions. The sample provided 80% power to detect interactions with an odds ratio of at least 1.67. This calculation was based on each variable contributing to the interaction itself predicting outcome with an odds ratio of at least 1.50, which, based on previous Raine Study analyses would be reasonable to expect (Allen et al., 2014).

It is worth nothing that interaction effects in CTA take a different form to those in regression models (Lemon et al., 2003; Loh, 2011). An interaction in CTA is testing the strongest predictor of a dichotomous outcome (i.e., eating disorder development) in separate subsamples (such as males and females, or those with high weight and eating concerns vs. low concerns) using recursive partioning and taking into account non-linear effects. If different predictors, or different cut-points for predictors, emerge for these subsamples then the predictive effects are significantly different for each sample (a significant interaction). In contrast, interactions in regression models assess differences in the effects of one predictor variable according to levels of a second predictor variable. Any splits in the sample are pre-determined rather than identified as part of the analyses (Camp & Slattery, 2002; Lemon et al., 2003).

2. Results

2.1. Descriptive statistics

Means and standard deviations for predictor variables are shown in Table 1. Correlations between predictor variables are shown in Table 2. Most correlations were statistically significant, positive, and moderate in magnitude.

2.2. Classification tree and random forest analyses

2.2.1. Classification tree analyses

Results from CTA are shown in Fig. 1. The initial split by participant sex was significant ($X^2[1] = 59.18$, $p < 0.001$) and eating disorder incidence over follow-up was 18.1% for females versus 4.0% for males. Among males, the only subsequent split was for weight and eating concerns ($X^2[2] = 50.28$, $p < 0.001$) and three non-linear risk categories were identified. Of the male participants with normalized weight and eating concern scores in the bottom 12% of participants (at least 1.16 SD below the mean), 1.2% subsequently developed an eating disorder. Of the male participants with normalized scores between the 12th and 59th percentiles (47% of the sample, normal score between −1.15 and 0.21), 1.5% developed an eating disorder. Of the

Table 1
Means and standard deviations for predictor variables, by sex and eating disorder group. Raw scores are presented first, followed by normal scores.

	Males		Females	
	No ED (n = 609)	ED at age 17 or 20 (n = 26)	No ED (n = 542)	ED at age 17 or 20 (n = 120)
Body mass index	20.74 (3.86)	23.98 (3.55)	21.03 (3.86)	22.71 (4.25)
Normalized	−0.13 (1.00)	0.77 (0.68)	−0.02 (0.95)	0.41 (0.89)
Weight & eating concern	0.30 (0.30)	0.73 (0.43)	0.50 (0.41)	0.81 (0.49)
Normalized	−0.37 (0.79)	0.54 (0.79)	0.08 (0.84)	0.65 (0.76)
Restraint	0.30 (0.37)	0.75 (0.56)	0.42 (0.44)	0.71 (0.54)
Normalized	−0.22 (0.73)	0.57 (0.88)	0.01 (0.80)	0.49 (0.86)
Depression	5.67 (6.89)	12.61 (12.77)	8.43 (7.82)	15.79 (9.88)
Normalized	−0.27 (0.90)	0.45 (1.17)	0.13 (0.88)	0.87 (0.88)
Externalizing Problems	48.28 (9.42)	51.48 (12.29)	49.55 (9.41)	55.42 (8.78)
Normalized	−0.13 (0.98)	0.20 (1.28)	0.01 (0.97)	0.61 (0.89)
Internalizing Problems	46.35 (9.59)	53.30 (11.17)	46.78 (8.50)	41.50 (8.94)
Normalized	−0.08 (1.03)	0.64 (1.15)	−0.03 (0.90)	0.47 (0.90)

Note. ED = Binge eating or purging eating disorder. Participants with an eating disorder at age 14 (n = 74) and those with anorexia nervosa or atypical anorexia nervosa in later adolescence (n = 12) were excluded.
Consistent with the use of lagged predictor variables, values for the eating disorder group are for the assessment immediately prior to disorder onset. Values for the non-eating disorder group are for age 17. Differences between the eating disorder and non-eating disorder groups were significant for all variables ($p < 0.01$), with the exception of externalizing problems in males.

Table 2
Pearson bivariate correlations between normalized predictor variables. Results for males (n = 635) are shown above the diagonal, and those for females (n = 662) below the diagonal and in italics.

	BMI	Weight & eating concern	Restraint	Depression	Externalizing problems	Internalizing problems
BMI	–	0.41**	0.47**	−0.04	0.07	0.01
Weight & eating concern	*0.45***	–	0.62**	0.32**	0.21**	0.35**
Restraint	*0.50***	*0.70***	–	0.15**	0.06	0.15**
Depression	*0.14***	*0.44***	*0.26***	-	0.46**	0.70**
Externalizing problems	*0.18***	*0.33***	*0.22***	*0.50***	–	0.48**
Internalizing problems	*0.11***	*0.44***	*0.26***	*0.70***	*0.47***	–

*$p < 0.05$ **$p < 0.01$.

male participants with normal scores in the top 41% of the sample (at least 0.21 SD above the mean), 17.1% developed an eating disorder.

Weight and eating concerns also formed the first split for females ($X^2[2] = 102.42$, $p < 0.001$), and again, three non-linear risk categories were identified. The lowest cut-point was equivalent to that obtained for male participants (bottom 12% of participants) and 1.8% of the females in this group subsequently develop an eating disorder. For girls with concern scores between the 12th and 77th percentiles (65% of the sample, normal score between −1.16 and 0.74), 8.2% developed an eating disorder. For girls with a concern score in the top 23% of the sample (at least 0.74 SD above the mean), 45.1% developed an eating disorder.

For females in the middle category of weight and eating concerns, an additional split was observed for externalizing problems ($X^2[1] = 8.46$, $p = 0.033$). There was a protective relationship between moderate concern and low externalizing problems, with girls in the bottom 39% of externalizing problems (at least −0.28 SD below the mean) having an eating disorder incidence of 2.8% over follow-up (vs. 8.2% for the moderate weight and eating concern group as a whole). For females with average or above average externalizing problems (top 61% of participants/above −0.28 SD from the mean), eating disorder incidence was 12.3%.

Fig. 2 summarizes eating disorder incidence rates by sex and weight and eating concern group.

2.2.2. Random forest analyses

For boys, random forest analyses with all possible predictor variables generated an aggregated error rate (OOB) of 0.04% for classifying eating disorder onset. Error was much higher for predicting eating disorder cases (89.3%) than non-cases (0.01%). Rankings of predictor importance placed weight and eating concerns (importance rating 10.3) and BMI (10.3) as equally important, and these variables were distinguishable from other predictors (where importance ratings ranged from 6.0 to 6.6). When analyses were repeated with weight and eating concern as the only predictor (as per CTA findings), the aggregated error rate was again 0.04% and the error rate for eating disorder cases reduced to 78.6%. Error for non-cases remained at 0.01%. Thus, CTA findings were partially supported by random forest analyses for boys.

For girls, random forest analyses with all possible predictor variables generated an aggregated error rate of 13.8%, with error for eating disorder cases being 56.8% and that for non-cases 0.04%. Weight and eating concern was ranked as most important (=40.9) and this variable was distinguishable from other predictors, which had importance ratings between 23.7 and 27.7. Body mass index was ranked as more important than externalizing problems (27.7 vs. 24.7). When analyses were repeated with weight and eating concern and externalizing problems as the two predictors (as per CTA findings), the aggregated error rate increased slightly to 16.4%, with error for eating disorder cases being 63.6% and that for non-cases being 0.06%. Thus, CTA findings were not fully supported for girls.

2.3. Two-way interaction tests

In logistic regression analyses, no two-way interaction terms were significant in predicting eating disorder onset when applying the corrected alpha level of $p < 0.0033$ ($ps = 0.019$ to 0.971 for males and 0.035 to 0.912 for females[1]).

[1] Interaction effects with $p < 0.05$ included Restraint x Externalizing Problems for boys ($p = 0.024$), Depression x Externalizing Problems for girls ($p = 0.044$) and Internalizing Problems x Externalizing Problems for boys ($p = 0.019$) and girls ($p = 0.035$).

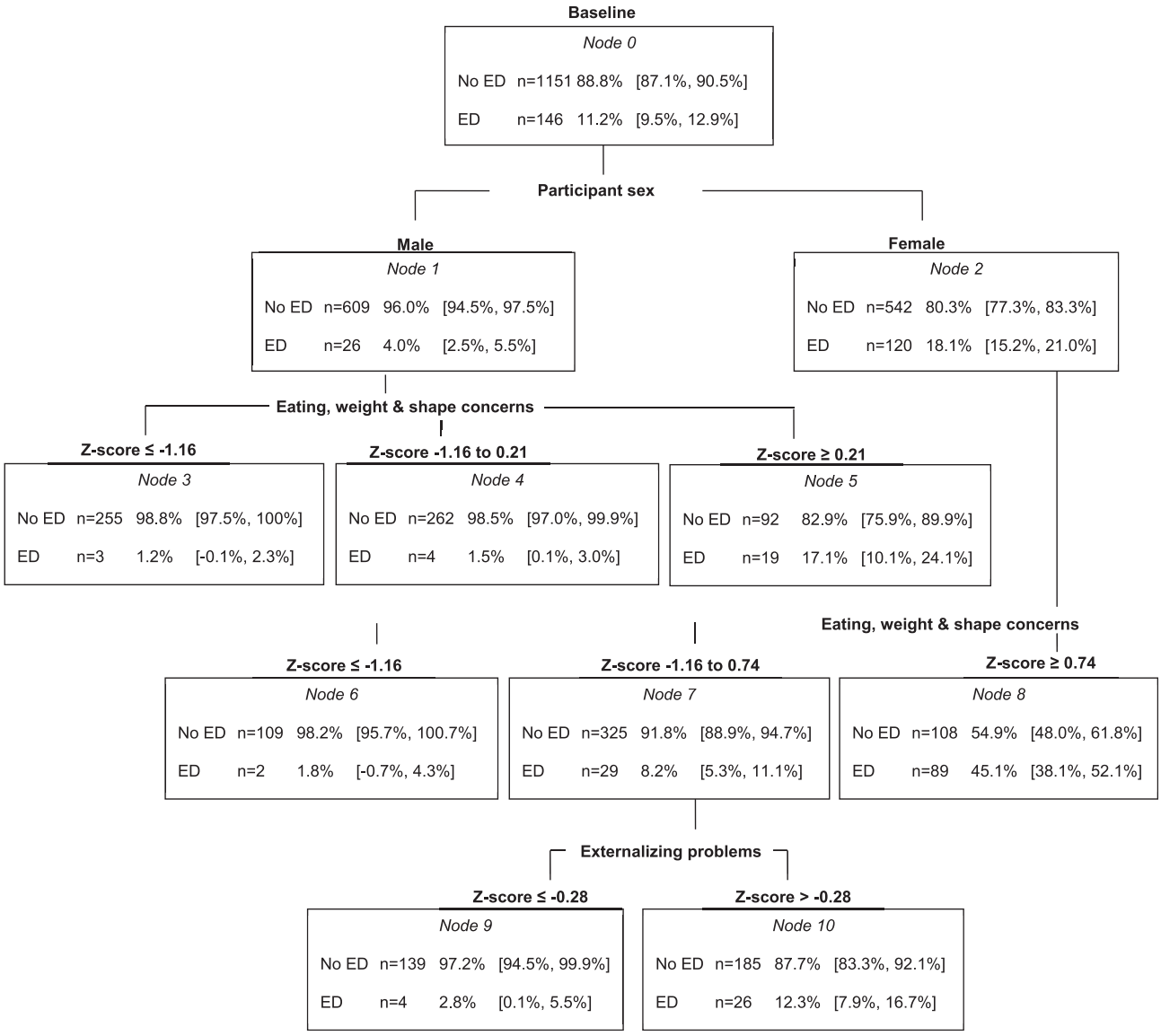

Fig. 1. The empirically derived classification tree predicting eating disorder onset. Sample size and eating disorder incidence is shown for each branch and node. Cut-points are for normalized scores. Square brackets show 95% CI.

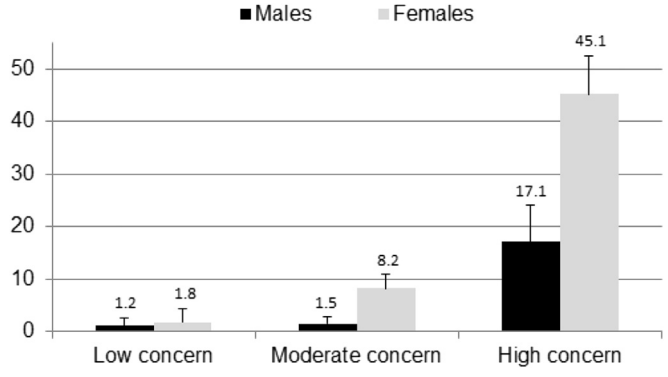

Fig. 2. Eating disorder incidence rates (%) and 95% confidence intervals for males and females, by weight and eating concern group.

3. Discussion

This study sought to identify the combination/s of risk factors in early-middle adolescence that could best account for the development of binge eating and purging eating disorders in later adolescence, with attention to sex differences and specific cut-points for risk. In CTA, participant sex interacted with weight and eating concern scores to predict eating disorder onset. Thus, cut-points on weight and eating concern differed by sex but this was still the most potent predictor of eating disorder onset for both sexes. An interaction was also found between moderate weight and eating concern scores and externalizing problems in girls. These results were partially supported by random forest analyses, which emphasized the importance of weight and eating concerns in the prediction of eating disorders in males and females. However, the importance of externalizing problems was less clear using this approach and BMI was suggested to be important. In regression models, no two-way interaction terms were significant in predicting risk after correcting for multiple testing.

The different results across CTA and regression-based analyses are noteworthy because very few studies have used CTA to investigate eating disorder risk factors. There are a number of advantages of this analytical approach over traditional regression, including the capacity to test for non-linear effects and interactions and to identify the cut-points that best characterize low versus high risk groups (Loh, 2011). Given the non-linear relationship found between weight and eating concerns and eating disorder risk in CTA (as demonstrated in Fig. 2), it is not surprising that regression models did not detect this interaction. Interaction terms in regression models assume a linear relation between the predictors and outcome (Lemon et al., 2003).

For females, eating disorder incidence rates were greatest (45.1%) when weight and eating concern scores were in the top 23% of the sample. This is an important finding because the cut-point for high risk converges almost exactly with that found by Stice and colleagues in their CTA study in 2011. In Stice et al. (2011) sample, body dissatisfaction in the top 24% (normal score >0.69) was the strongest predictor of eating disorder risk. Taken together, these results provide strong support for a risk cut-point of approximately 0.70 SD above the mean (top 23–24% of girls) on measures of body dissatisfaction or weight and eating concerns. Girls who fall above this cut-point are likely to be particularly in need of targeted eating disorder prevention programs.

Whilst girls needed to score in the top 23% for concern to be classified as high risk, a more conservative cut-point, the top 41%, emerged for boys. With weight and eating concern scores being lower in males than females overall, it may be that even slightly elevated concern will place adolescent males at risk for binge eating and purging disorders. It is also noteworthy that the effect of weight and eating concern on eating disorder onset was more pronounced for females than for males. Males in the high concern group had an incidence of eating disorders that was 11.4 times that of the moderate concern group (17.1% vs. 1.5%) and 14.2 times that of the low concern group (1.2%). Females in the high concern group had an incidence of eating disorders that was 5.5 times that of the moderate concern group (45.1% vs. 8.2%) but 37.6 times that of the low concern group (1.2%). These results suggest that screening for weight and eating concern may be important for adolescent boys and girls, but that interpretation of concern scores requires sex-specific norms and risk cut-points.

To our knowledge, this is the first study to find an interaction between moderate weight and eating concerns and externalizing symptoms in the prediction of binge eating and purging disorders in females. However, externalizing problems have been found to predict *persistent* eating pathology in Raine Study participants in the past (Allen, Byrne, Oddy, & Crosby, 2013b) and were linked to eating disorder onset in one of the three prior CTA studies in this area (Stice et al., 2011). In Stice et al. (2011) study, externalizing problems only predicted eating disorders when other identified risk factors (denial of the costs of pursuing the thin-ideal and emotional eating) were absent. This converges with results from the current study, where externalizing symptoms only predicted risk when the more potent risk factor of weight and eating concern was not strongly endorsed. The results also suggested a protective relationship for girls with low-moderate weight and eating concerns and low-moderate externalizing problems. Externalizing problems capture a range of symptoms relating to emotional and behavioral under-control and can also be thought of as a proxy for trait impulsivity, which has been linked to binge eating and purging disorders in the past (e.g., Wonderlich, Connolly, & Stice, 2004). At the same time, random forest analyses did not rate externalizing problems as more important than other psychosocial variables in the prediction of eating disorder onset across bootstrapped data. Further research is thus needed to assess the generalizability of this particular interaction.

Contrary to previous findings, depression did not predict eating disorder onset in male or female Raine Study participants. There are a number of possible reasons for this. One is that depression was not a strong predictor of outcome in this sample once other key variables were taken into account. Another is the timing of assessment. Weight and eating concerns are known to increase across adolescence, on average (Neumark-Sztainer, Wall, Larson, Eisenberg, & Loth, 2011; Rohde, Stice, & Marti, 2015), and so assessing risk factors earlier in adolescence (<14 years) may allow for the effects of other variables, including depression, to be identified more readily. In the Stice et al. (2011) study where body dissatisfaction and depressive symptoms interacted to predict disorder risk, the mean age of participants at baseline was 13.5 years. However, as these participants were followed over 8 years, there was still overlap with the developmental period of the current study. A third possible reason for the null effect of depression is that the BDI-Y, a clinical measure of depression, may not be sensitive enough to capture low levels of negative affect. In all, it seems that further research is warranted to determine if and how depression interacts with other variables to predict binge eating and purging disorders.

One limitation of this research is that the number of male eating disorder cases was small ($n = 26$). This may account for the lack of risk interaction effects with this group, and makes ongoing attention to risk factors for male eating disorders important. Moreover, we were unable to assess for possible male-specific risk factors, such as muscularity concerns or sexuality, due to not having measures to assess these variables. A second limitation is that eating disorder symptoms were assessed via questionnaires. Eating disorder symptoms assessed with questionnaire versus interview have shown reasonable agreement (Stice et al., 2000, 2004) but the latter is still considered the gold standard. Further to this, eating disorder prevalence rates in this study are somewhat higher than those reported elsewhere (e.g., Stice et al., 2013), although there are still very few reports of prevalence rates according to DSM-5 criteria. Third, we examined a relatively narrow set of predictors, chosen to extend previous studies. Finally, our results relate to a largely Caucasian sample. We can not comment on the generalizability of our findings to other ethnic or racial groups and more research is needed to test whether ethnicity moderates the effects of risk factors on eating disorder onset. A key strength of the research is the application of CTA to prospective risk factor data, something that has been done on relatively few occasions in the past, and the use of random forest analyses to evaluate the generalizability of CTA findings. Additional strengths include the use of a large, prospective cohort followed over adolescence; well-validated questionnaire measures; and modelling of sex differences.

In sum, this study provides new data on the combinations of risk factors that may best predict binge eating and purging eating disorders in male and female adolescents. Results highlight a non-linear weight and eating concern pathway to eating pathology in both sexes. This is an important finding that will allow for more accurate screening of youth at risk for eating disorders, by providing clearer cut-points for increased risk. Results also revealed an additional risk pathway for females, involving moderate weight and eating concerns combined with externalizing symptoms. The high risk cut-point for female weight and eating concerns converges almost exactly with that of Stice et al. (2011). Theoretically, these results extend knowledge on risk pathways to eating disorders and provide support for the use of CTA (or other analyses catering to non-linear risk processes) when conducting research in this area. Practically, these results suggest that targeted eating disorder prevention should be provided to girls with weight and eating concern scores in the top quartile of their peers and males

with weight and eating concern scores in the top 40% of their peers. We also recommend replication studies to assess the generalizability of an interaction between moderate weight and eating concerns and average/above-average externalizing symptoms in the prediction of eating disorders in females.

Conflicts of interest

None.

Acknowledgements

We are extremely grateful to the Raine Study participants and their families, and to the Raine Study team for cohort management and data collection. The first author was supported by an early career research fellowship from the National Health and Medical Research Council (NHMRC) of Australia. Core funding for the Raine Study is provided by the Telethon Kids Institute (previously the Telethon Institute for Child Health Research), Raine Medical Research Foundation, University of Western Australia, Faculty of Medicine, Dentistry and Health Sciences at UWA, Women's and Infant's Research Foundation, Curtin University, and Edith Cowen University. Funding for the 17 and 20-year follow-ups was provided by the NHMRC, Canadian Institutes of Health Research, and Lions Eye Institute.

References

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Allen, K. L., Byrne, S. M., & Crosby, R. D. (2015). Distinguishing between risk factors for bulimia nevosa, binge eating disorder, and purging disorder. *Journal of Youth and Adolescence, 44*, 1580–1591.

Allen, K. L., Byrne, S. M., Forbes, D., & Oddy, W. H. (2009). Risk factors for full- and partial-syndrome early adolescent eating disorders: A population-based pregnancy cohort study. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*, 800–809.

Allen, K. L., Byrne, S. M., Oddy, W. H., & Crosby, R. D. (2013a). DSM-IV-TR and DSM-5 eating disorders in adolescents: Prevalence, stability, and psychosocial correlates in a population-based sample of male and female adolescents. *Journal of Abnormal Psychology, 122*, 720–732.

Allen, K. L., Byrne, S. M., Oddy, W. H., & Crosby, R. D. (2013b). Early onset binge eating and purging eating disorders: Course and outcome in a population-based study of adolescents. *Journal of Abnormal Child Psychology, 41*, 1083–1096.

Allen, K. L., Byrne, S. M., Oddy, W. H., Schmidt, U., & Crosby, R. D. (2014). Risk factors for binge eating and purging eating disorders: Differences based on age of onset. *International Journal of Eating Disorders, 47*, 802–812.

Allen, K. L., Crosby, R. D., Oddy, W. H., & Byrne, S. M. (2013). Eating disorder symptom trajectories in adolescence: Effects of time, participant sex, and early adolescent depressive symptoms. *Journal of Eating Disorders, 1*(32). http://dx.doi.org/10.1186/2050-2974-1181-1132.

Beato-Fernandez, L., Rodriguez-Cano, T., Belmonte-Llario, A., & Martinez-Delgado, C. (2004). Risk factors for eating disorders in adolescents: A spanish community-based longitudinal study. *European Child & Adolescent Psychiatry, 13*, 287–294.

Beck, S., Beck, A. T., & Jolly, J. (2001). *Beck youth inventory*. New York: Psychological Corporation.

Begg, S., Vos, T., Barker, B., Stevenson, C., Stanley, L., & Lopez, A. D. (2007). *The burden of disease and injury in Australia 2003*. Canberra: Australian Institute of Health and Welfare (AIHW).

Berg, K. C., Stiles-Shields, E. C., Swanson, S. A., Peterson, C. B., Lebow, J., & Le Grange, D. (2012). Diagnostic concordance of the interview and questionnaire versions of the Eating Disorder Examination. *International Journal of Eating Disorders, 45*, 850–855.

Berkman, N. D., Lohr, K. N., & Bulik, C. M. (2007). Outcomes of eating disorders: A systematic review of the literature. *International Journal of Eating Disorders, 40*, 293–309.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

Camp, N. J., & Slattery, M. L. (2002). Classification tree analysis: A statistical tool to investigate risk factor interactions with an example for colon cancer (United States). *Cancer Causes and Control, 13*, 813–823.

Demidenko, E. (2008). Sample size and optimal design for logistic regression with binary interaction. *Statistics in Medicine, 27*, 36–46.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*, 363–370.

Field, A. E., Sonneville, K. R., Crosby, R., Swanson, S. A., Eddy, K. T., Camargo, C. A., … Micali, N. (2014). Prospective associations of concerns about physique and the development of obesity, binge drinking, and drug use among adolescent boys and young adult men. *JAMA Pediatrics, 168*, 34–39.

Ghaderi, A., & Scott, B. (2001). Prevalence, incidence and prospective risk factors for eating disorders. *Acta Psychiatrica Scandinavica, 104*, 122–130.

Hrabosky, J. I., Masheb, R. M., White, M. A., & Grilo, C. M. (2007). Overvaluation of shape and weight in binge eating disorder. *Journal of Consulting and Clinical Psychology, 75*, 175–180.

Jacobi, C., Fittig, E., Bryson, S. W., Wilfley, D. E., Kraemer, H. C., & Taylor, C. B. (2011). Who is really at risk? Identifying risk factors for subthreshold and full syndrome eating disorders in a high-risk sample. *Psychological Medicine, 41*, 1939–1949.

Jacobi, C., Hayward, C., de Zwaan, M., Kraemer, H. C., & Agras, W. S. (2004). Coming to terms with risk factors for eating disorders: Application of risk terminology and suggestions for a general taxonomy. *Psychological Bulletin, 130*, 19–65.

Johnson, J. G., Cohen, P., Kasen, S., & Brook, J. S. (2002). Eating disorders during adolescence and the risk for physical and mental disorders during early adulthood. *Archives of General Psychiatry, 59*, 545–552.

Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analyses in public health: Methodological review and comparison with logistic regression. *Annals of Behavioral Medicine, 26*, 172–181.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*, 18–22.

Liechty, J. M., & Lee, M. (2013). Longitudinal predictors of dieting and disordered eating among young adults in the U.S. *International Journal of Eating Disorders, 46*, 790–800.

Loh, W. Y. (2011). Classification and regression trees. *Data Mining and Knowledge Discovery, 1*, 14–23.

Martinez-Gonzalez, M. A., Gual, P., Lahortiga, F., Alonso, Y., Irala-Estevez, J. d., & Cervera, S. (2003). Parental factors, mass media influences, and the onset of eating disorders in a prospective population-based cohort. *Pediatrics, 111*, 315–320.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy, 42*, 551–567.

Neumark-Sztainer, D., Wall, M., Larson, N. I., Eisenberg, M. E., & Loth, K. (2011). Dieting and disordered eating behaviors from adolescence to young adulthood: Findings from a 10-year longitudinal study. *Journal of the American Dietetic Association, 111*, 1004–1011.

Newnham, J. P., Evans, S. F., Michael, C. A., Stanley, F. J., & Landau, L. I. (1993). Effects of frequent ultrasound during pregnancy: A randomised controlled trial. *Lancet, 342*, 887–891.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge, UK: Cambridge University Press.

Rohde, P., Stice, E., & Marti, C. (2015). Developmental and predictive effects of eating disorder risk factors during adolescence: Implications for prevention efforts. *International Journal of Eating Disorders, 48*, 187–198.

Simon, J., Schmidt, U., & Pilling, S. (2005). The health service use and cost of eating disorders. *Psychological Medicine, 35*, 1543–1551.

Solmi, F., Sonneville, K. R., Easter, A., Horton, N. J., Crosby, R. D., Treasure, J., … Micali, N. (2015). Prevalence of purging at age 16 and associations with negative outcomes among girls in three community-based cohorts. *Journal of Child Psychology and Psychiatry, 56*, 87–96.

Stice, E. (2016). Interactive and mediational etiological mdoels of eating disorder onset: Evidence from prospective studies. *Annual Review of Clinical Psychology, 12*, 359–381.

Stice, E., Fisher, M., & Martinez, E. (2004). Eating disorder diagnostic scale: Additional evidence of reliability and stability. *Psychological Assessment, 16*, 60–71.

Stice, E., Marti, C. N., & Durant, S. (2011). Risk factors for onset of eating disorders: Evidence of multiple risk pathways from an 8-year prospective study. *Behaviour Research and Therapy, 49*, 622–627.

Stice, E., Marti, C., & Rohde, P. (2013). Prevalence, incidence, impairment and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *Journal of Abnormal Psychology, 122*, 445–457.

Stice, E., Marti, C. N., Spoor, S. T. P., Presnell, K., & Shaw, H. E. (2008). Dissonance and healthy weight eating disorder prevention programs: Long-term effects from a randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 76*, 329–340.

Stice, E., Rohde, P., Gau, J., & Shaw, H. E. (2012). *Effect of a dissonance-based prevention program on risk for eating disorder onset in the context of eating disorder risk factors, 13*(Prevention Science), 129–139.

Stice, E., & Shaw, H. E. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin, 130*, 206–227.

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the eating disorder diagnostic scale: A brief self-report measure for anorexia, bulimia, and binge eating disorder. *Psychological Assessment, 12*, 123–131.

The McKnight Investigators. (2003). Risk factors for the onset of eating disorders in adolescent girls: Results of the McKnight longitudinal risk factor study. *American Journal of Psychiatry, 160*, 248–254.

Wonderlich, S. A., Connolly, K. M., & Stice, E. (2004). Impulsivity as a risk factor for eating disorder behavior: Assessment implications with adolescents. *International Journal of Eating Disorders, 36*, 172–182.