Psychological Medicine

cambridge.org/psm

Developmental pathway for first onset of depressive disorders in females: from adolescence to emerging adulthood

Wenting Mu[1,*], Chuncheng Huang[1,*], Nisha Yao[2], Jiaju Miao[3], Greg Perlman[4], David Watson[5], Daniel N. Klein[6] and Roman Kotov[3,4,6]

[1]Department of Psychology, Tsinghua University, Beijing, China; [2]School of Kinesiology and Health, Capital University of Physical Education and Sports, Beijing, China; [3]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA; [4]Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA; [5]Department of Psychology, University of Notre Dame, Notre Dame, Indiana, USA and [6]Department of Psychology, Stony Brook University, Stony Brook, NY, USA

Original Article

*Wenting Mu and Chuncheng Huang share first authorship of this work.

Cite this article: Mu W, Huang C, Yao N, Miao J, Perlman G, Watson D, Klein DN, Kotov R (2024). Developmental pathway for first onset of depressive disorders in females: from adolescence to emerging adulthood. *Psychological Medicine* **54**, 753–762. https://doi.org/10.1017/S0033291723002441

Received: 5 February 2023
Revised: 29 June 2023
Accepted: 27 July 2023
First published online: 29 August 2023

Keywords:
chronic risk; dynamic risk; first-onset depression; prediction; risk change

Corresponding author:
Wenting Mu;
Email: wmu@mail.tsinghua.edu.cn; mwttwm@gmail.com

Abstract

Background. Although risk markers for depressive disorders (DD) are dynamic, especially during adolescence, few studies have examined how change in risk levels during adolescence predict DD onset during transition to adulthood. We compared two competing hypotheses of the dynamic effects of risk. The risk escalation hypothesis posits that worsening of risk predicts DD onset beyond risk level. The chronic risk hypothesis posits that persistently elevated risk level, rather than risk change, predicts DD onset.

Methods. Our sample included 393 girls (baseline age 13.5–15.5 years) from the adolescent development of emotions and personality traits project. Participants underwent five diagnostic interviews and assessments of risk markers for DD at 9-month intervals and were re-interviewed at a 6-year follow-up. We focused on 17 well-established risk markers. For each risk marker, we examined the prospective effects of risk level and change on first DD onset at wave six, estimated by growth curve modeling using data from the first five waves.

Results. For 13 of the 17 depression risk markers, elevated levels of risk during adolescence, but not change in risk, predicted first DD onset during transition to adulthood, supporting the chronic risk hypothesis. Minimal evidence was found for the risk escalation hypothesis.

Conclusions. Participants who had a first DD onset during transition to adulthood have exhibited elevated levels of risk throughout adolescence. Researchers and practitioners should administer multiple assessments and focus on persistently elevated levels of risk to identify individuals who are most likely to develop DD and to provide targeted DD prevention.

Depressive disorders (DD) are highly prevalent and debilitating. A leading cause of global disease burden, the prevalence of DD increases dramatically in adolescence and emerging adulthood (Rohde, Lewinsohn, Klein, Seeley, & Gau, 2013), especially for females, and is associated with impaired functioning and increased mortality (Malhi & Mann, 2018). Although numerous risk markers have been identified as predictors of DD onset, most studies assume risk markers are static and assess them only once, without considering how risk changes over time (Mu et al., 2021; Nelson, McGorry, Wichers, Wigman, & Hartmann, 2017). In fact, many risk markers have been shown to be highly dynamic in nature and change substantially over time (e.g. Goldstein, Perlman, Eaton, Kotov, & Klein, 2020; Klein, Kotov, & Bufferd, 2011; Roberts, Walton, & Viechtbauer, 2006), especially during adolescence (Roberts & DelVecchio, 2000; Trzesniewski, Donnellan, & Robins, 2003).

It is largely unknown how the patterning of changes in risk factors is related to risk for DD onset. Two competing hypotheses have been proposed (Mu et al., 2021). The risk escalation hypothesis posits that DD onset is most likely following an escalation in levels of risk factors. Several studies have provided support for this hypothesis, demonstrating that increases in risk factors predicted subsequent increases in the severity of depressive symptoms (Mu, Luo, Nickel, & Roberts, 2016; Steiger, Allemand, Robins, & Fend, 2014) and probability of DD onset (Laceulle, Ormel, Vollebergh, Van Aken, & Nederhof, 2014). Alternatively, the chronic risk hypothesis posits that sustained elevation of risk over time predicts DD onset and that change in risk over time contributes little additional information in predicting DD onset. There has been some indirect support for this hypothesis. For example, chronic stressors predict increases in depressive symptoms (Jenness, Peverill, King, Hankin, & McLaughlin, 2019) and subsequent DD onset (Hammen, Kim, Eberhart, & Brennan, 2009). Furthermore, using a trait–state decomposition method, some studies have found that the stable component, rather than the state component, of personality traits predicted DD onset (Kendall et al., 2015) and the course of depression (Naragon-Gainey, Gallagher, & Brown, 2013). Finally, adolescents

who exhibited subthreshold depression across two consecutive assessment waves had a higher likelihood of developing full-criteria depression at later assessment waves than those with subthreshold depression only at one assessment wave (Klein, Shankman, Lewinsohn, & Seeley, 2009).

Although empirical studies have provided evidence that is consistent with both of these hypotheses, only one study, to our knowledge, has directly compared the risk escalation and chronic risk hypotheses (Mu et al., 2021). In a sample of never-depressed early adolescent girls who were assessed prospectively across five waves, they evaluated the predictive effect of mean level and change in levels of risk factors over time on the development of a first DD onset. Results showed that mean level of risk factors outperformed increases in levels of risk factors in predicting DD onset, supporting the chronic risk hypothesis. This pattern of findings was consistent across 16 well-established risk markers in four different domains.

The present study extends Mu et al.'s (2021) work in several ways. First, while Mu et al. (2021) examined risk for DD onset during adolescence, the current study focused on transition to adulthood. DD incidence peaks in transition to adulthood (Rohde et al., 2013), which is an important but less studied developmental phase for understanding the development of DD. Furthermore, adolescence is characterized with rapid changes in many psychological aspects, and trajectories of risk factors in adolescence can have long-term consequences on mental health in adulthood (Steiger et al., 2014). However, it is unclear how dynamic risk during adolescence predicts DD onset in transition to adulthood. Second, unlike Mu et al. (2021), the current study adopted a more stringent definition of DD by excluding DD NOS and defined DD as major depressive disorder (MDD) or dysthymic disorder, which improves the internal validity of the current study. Third, we adopted an entirely different analytic approach. While Mu et al. (2021) assessed change using difference scores between adjacent assessments separated by 9-month intervals, we examined whether more lasting change (i.e. 3 years) in risk factors predict subsequent DD onset. After all, 9 months may not have been long enough to capture meaningful change, as it may fail to distinguish transient fluctuations from lasting change. Hence, we employed latent growth curve modeling instead of cox regression as used in Mu et al. (2021). This approach models changes in levels of risk factors adjusting for measurement error and distinguishing it from transient fluctuations, providing a more precise estimate of change.

The current study aims to provide a rigorous test of the risk escalation and chronic risk hypotheses while addressing the above-mentioned issues. To our knowledge, no research to date has ever examined how dynamic risk during adolescence predicts DD onset in transition to adulthood. Well-established risk markers were examined, including subclinical depression and anxiety symptoms (e.g. Klein et al., 2009, 2013), personality traits (e.g. neuroticism, extraversion, and conscientiousness; Bagby, Quilty, & Ryder, 2008; Goldstein, Kotov, Perlman, Watson, & Klein, 2018; Klein et al. 2011), depressogenic cognitive or personality styles (e.g. rumination, self-criticism, dependency; Burkhouse, Uhrlass, Stone, Knopik, & Gibb, 2012; Kopala-Sibley, Klein, Perlman, & Kotov, 2017; Nolen-Hoeksema, 2000), and social risk markers (e.g. social support, peer/parent–child relationship, school engagement; Shochet, Homel, Cockshaw, & Montgomery, 2008; Smith, Nelson, & Adelson, 2019; Starr & Davila, 2008). This study included six assessment waves over 6 years, during which a sample of girls was followed from adolescence into emerging adulthood. We examined how levels and changes in risk factors in adolescence predicted DD onset in transition to adulthood.

Methods

Participants

Data included in the current study were collected as part of the Adolescent Development of Emotions and Personality Traits (ADEPT) project. The original ADEPT project followed 550 adolescent girls from the start of adolescence for 3 years with five waves of assessments, with 9 months between two waves. The extended ADEPT project attempted to re-interview all 550 participants when they entered early adulthood. For the current study, we focused on girls who did not have a first onset of MDD or dysthymic disorder until wave 6, so the 72 participants who had MDD or dysthymic disorder during the first five waves were excluded (see online Supplementary Fig. S1 for a comparison of risk factors between participants who had a first onset between waves 5 and 6 and those who had an onset between waves 1 and 5). Participants who had medically caused DD ($n = 1$), bipolar disorder ($n = 6$), and those who were missing at wave 6 ($n = 78$) were also excluded from the analyses.

The current sample consisted of 393 females. Participants were 13.14–15.63 years old (mean = 14.38, S.D. = 0.62) at enrollment (Table 1; see online Supplementary Table S1 for the age range of participants at each wave). Most were non-Hispanic White (81.9%), and 34.5% came from families where both parents had a bachelor's or higher degree. None of these participants had intellectual disability or lifetime history of MDD or dysthymic disorder before enrollment. Details of the recruitment and inclusion/exclusion criteria can be found in Michelini et al. (2021).

Participants completed five diagnostic interviews for DD and assessments of relevant risk factors at 9-month intervals for 3 years. Then, at a 6-year follow-up, participants completed a diagnostic interview for DD. Participants used life events calendar to aid recall of DD symptoms in the 3-year interval. Among the 393 participants, 315 participants never developed DD, and 78 participants had their first DD onsets during the wave 6 interval (MDD, $n = 76$; dysthymic disorder, $n = 2$). In onset group, there were no new onsets after 15 March 2020,[†1] indicating that the new onset is not the result of the pandemic. Informed assent and consent were obtained from all participants and their parents in waves 1–5 and informed consent was obtained from the participants in wave 6. The study was approved by the Stony Brook University Institutional Review Board.

Assessments

Depression diagnosis

DD (i.e. MDD and dysthymic disorder) were assessed using the Kiddie Schedule for Affective Disorders and Schizophrenia for School Aged Children Present and Lifetime Version (K-SADS-PL; Kaufman et al., 1997). The K-SADS-PL is a semi-structured diagnostic interview designed to assess current and lifetime psychiatric diagnoses in children and adolescents based on the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV). The K-SADS-PL interview was conducted by two well-trained research staff who were supervised by licensed

†The notes appear after the main text.

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

Table 1. Baseline demographic characteristics of participants with first DD onset at wave 6 *v.* participants with no DD onset across all waves

	First DD onset		Never DD onset		Test		
Characteristic	Mean	S.D.	Mean	S.D.	*t*	*p*	Cohen's *d*
Age	14.30	0.61	14.40	0.62	1.34	0.18	0.16
	N	%	*N*	%	χ^2	*p*	Cramer's *V*
White ethnicity	68	87.18%	281	89.21%	0.26	0.61	0.026
Hispanic ethnicity	8	10.26%	32	10.16%	0.00	0.98	0.001
Parents' education					1.07	0.30	0.052
Neither parent completed college	28	36.36%	95	30.25%			
One or both parent(s) completed college	49	63.64%	219	69.75%			
Household income					6.02	0.05	0.13
<$60 000	8	10.53%	28	10.00%			
$60 000–$120 000	38	50.00%	99	35.36%			
>$120 000	30	39.47%	153	54.64%			

Note. First DD onset, $n = 78$; never DD onset, $n = 315$.

clinical psychologists. An independent rater derived diagnoses from 48K-SADS-PL interview recordings to assess interrater reliability. Kappas for MDD and dysthymic disorder were 0.73 and 0.85, respectively (Mu et al., 2021).

Depression and anxiety symptoms

The Inventory of Depression and Anxiety Symptoms (IDAS-II; Watson et al., 2012) was used to measure symptoms of depression and anxiety. The IDAS-II contains 18 non-overlapping subscales and a General Depression scale composed of items from six subscales measuring depressive symptoms. Participants were asked to indicate to what extent they experienced a symptom over the past 2 weeks on a Likert scale ranging from 1 (not at all) to 5 (extremely). Based on the relevance to risk for depression onset, the current study included the General Depression scale (20 items) and six subscales measuring ill temper (5 items), panic symptoms (8 items), social anxiety (6 items), traumatic intrusions (4 items), traumatic avoidance (i.e. avoiding reminders of past traumas; 4 items), and mania (5 items). The seven specific depressive symptom subscales were not included to avoid redundancy with the General Depression scale. The claustrophobia subscale, the euphoria subscale, and the three obsessive compulsive subscales were not included due to low relevance to risk for depression onset.

Personality traits

The Big Five Inventory (BFI; John & Srivastava, 1999) was used to measure the personality traits of neuroticism, conscientiousness, and extraversion. The neuroticism and the extraversion subscales each contain eight items, while the conscientiousness subscale contains nine items. Participants were asked to indicate the extent to which they agreed that a characteristic (e.g. moody; a reliable worker; sociable) applied to them on a Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly).

Rumination

The Ruminative Responses Scale (RRS) of the Response Styles Questionnaire (Nolen-Hoeksema, 1987) contains 22 items. Participants were asked to indicate what they generally thought (e.g. think 'Why can't I get going?') when they felt depressed on a four-point Likert scale ranging from 1 (almost never) to 4 (almost always).

Self-criticism

Bagby, Parker, Joffe, and Buis's (1994) revision of the self-criticism subscale of the Depressive Experiences Questionnaire (DEQ; Blatt, D'Afflitti, & Quinlan, 1976) was used to measure self-criticism. The subscale contains nine items (e.g. 'I often find that I don't live up to my own standards or ideals'). Participants were asked to rate each item on a five-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly).

Dependency

The emotional reliance subscale of the Interpersonal Dependency Inventory (IDI; Hirschfeld et al., 1977) has 10 items (e.g. 'I would be completely lost if I didn't have someone special'). Each item was rated on a four-point Likert scale ranging from 1 (not characteristic of me) to 4 (very characteristic of me).

Social support

The Multidimensional Scale of Perceived Social Support (MSPSS; Zimet, Dahlem, Zimet, & Farley, 1988) has 12 items. Participants indicated the perceived adequacy of social support received from their family, friends, and significant others on a seven-point Likert scale ranging from 1 (very strongly disagree) to 7 (very strongly agree).

School engagement

Three subscales of the School Attitude Assessment Survey – Revised (SAAS-R; McCoach & Siegle, 2003) – attitudes toward school, attitude toward teachers, and self-motivation/regulation – were used to measure school engagement. The three subscales were aggregated, totaling 14 items. A sample item was 'I am glad that I go to this school'. Each item was rated on a seven-point Likert-type agreement scale (1 = very strongly disagree, 7 = very strongly agree).

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

Bullying

The Victim Version of the Revised Peer Experiences Questionnaire (RPEQ; De Los Reyes & Prinstein, 2004) has three subscales: overt victimization, relational victimization, and reputational victimization. The scales were aggregated, totaling 12 items. Participants indicated how often they were treated in a certain way (e.g. A teen excluded me from his/her group of friends) over the past 9 months on a five-point Likert scale (1 = never, 5 = a few times a week).

Parental criticism

The criticism subscale of the Network of Relationships Inventory (NRI; Furman & Buhrmester, 2009) includes three items each for the participant's mother and father. Participants rated how often each parent interacted with them in a certain way (e.g. How often does this person criticize you?) on a five-point Likert scale ranging from 1 (little or none) to 5 (the most), respectively. These six items were aggregated into a total score.

Psychometric properties of these measures (i.e. Cronbach's alphas and stability coefficients) can be found in Mu et al. (2021).

Statistical analysis

The outcome was whether or not DD onset occurred during the 3-year interval between waves 5 and 6. Predictors were intercepts and slopes estimated by latent growth curve modeling using five waves of assessments for each of 17 risk factors, in turn. We conducted the tests in two steps. First, we specified latent growth curve models for each of the 17 risk factors to estimate the mean and interindividual difference in the level and change of risk. Raw scores of each risk marker were standardized using means and standard deviations across all participants and assessment waves available. Grand mean standardized scores of a risk marker at each of the five assessment waves were used as indicators. For each risk factor, the intercept factor was set at either the first or the fifth assessment wave respectively to estimate both the initial and the proximal level of risk relative to time of onset. We therefore examined two univariate growth curve models for each of the 17 risk factors with the intercept setting at the first or the fifth assessment wave separately. Maximum likelihood estimation with robust standard error was performed.

Next, we address dichotomous wave 6 onset of DD in these models. Specifically, DD onset was regressed on the intercept and linear slope, which were latent factors with random effects estimated using latent growth curve modelling and were allowed to correlate (Fig. 1). The intercept factor was set at the first and the fifth assessment occasions respectively. No covariates were included in the structural model.

Missing data were addressed using full-information maximum likelihood estimation. All analyses were conducted using Mplus 8.7 (Muthén & Muthén, Los Angeles, CA, USA). We used the following indices to evaluate the fit of the latent growth curve models: the comparative fit index (CFI), the Tucker–Lewis index (TLI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA). Conventional guidelines (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004; Hooper, Coughlan, & Mullen, 2008) suggest that TLI and CFI values of 0.90 or greater indicate adequate fit and values of 0.95 or greater indicate excellent fit; for SRMR and RMSEA, values of 0.08 or less indicate acceptable and 0.05 or less indicate excellent fit.

Results

Level and change of risk before onset

Descriptive statistics

Among the 393 participants, 78 had a first DD onset between waves 5 and 6 (i.e. onset group), while the remaining participants had no DD onset in their life (i.e. no onset group). Standardized mean scores of each risk marker at each wave for the two groups are presented in Fig. 2, and the raw scores (i.e. means and standard deviations) are presented in online Supplementary Table S2. In general, the onset group exhibited a chronically higher level of risk compared to the no onset group for most risk markers throughout the first five assessment waves.

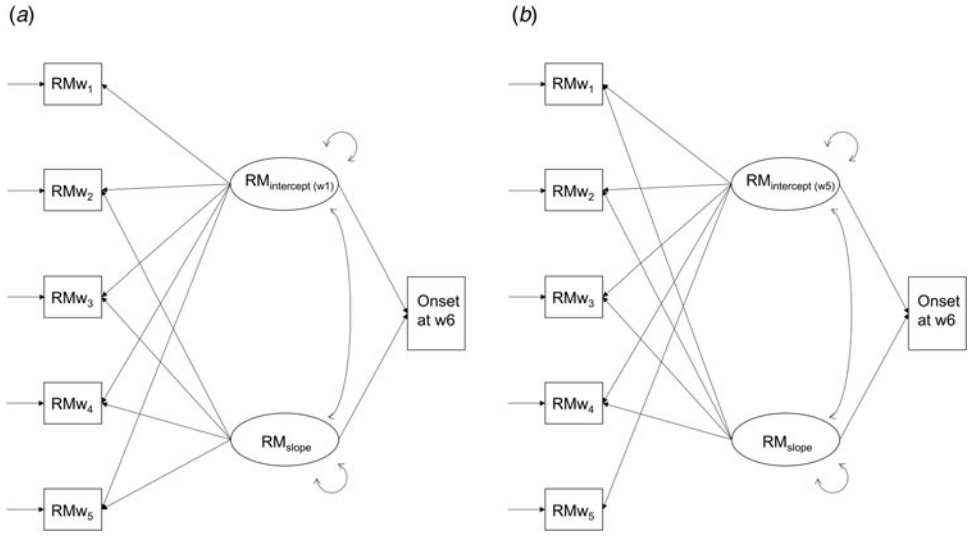

Figure 1. Structural equation model predicting first DD onset at wave 6 using (a) intercept (w1) and slope and (b) intercept (w5) and slope for each risk marker estimated from wave 1 to 5.
Note. RM = risk marker; W1 = wave 1; W2 = wave 2; W3 = wave 3; W4 = wave 4; W5 = wave 5; W6 = wave 6.

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

Figure 2. Standardized scores of each risk marker from wave 1 to 5 by onset group at wave 6.

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

Table 2. Fit indices of growth curve model with intercept set at wave 5 for each risk marker

	N	χ^2	RMSEA	90% CI of RMSEA	CFI	TLI	SRMR	AIC	BIC
Depression and anxiety symptoms									
1. General depression	393	33.62**	0.08	[0.05–0.11]	0.94	0.94	0.06	4666.75	4706.49
2. Ill temper	393	19.45*	0.05	[0.01–0.08]	0.95	0.95	0.06	4787.98	4827.72
3. Social anxiety	392	14.57	0.03	[0.00–0.07]	0.98	0.98	0.05	4684.60	4724.32
4. Panic	393	14.16	0.03	[0.00–0.07]	0.97	0.97	0.07	4843.73	4883.47
5.Traumatic intrusions	393	14.57	0.03	[0.00–0.07]	0.96	0.96	0.06	5019.79	5059.53
6.Traumatic avoidance	393	12.76	0.03	[0.00–0.06]	0.99	0.99	0.05	4840.74	4880.48
7. Mania	393	14.09	0.03	[0.00–0.07]	0.99	0.99	0.05	4631.62	4671.36
Personality traits									
8. Neuroticism	392	57.79**	0.11	[0.08–0.14]	0.95	0.95	0.05	4294.33	4334.04
9. Conscientiousness	393	41.62**	0.09	[0.06–0.12]	0.98	0.98	0.05	3932.44	3972.18
10. Extraversion	393	26.14**	0.06	[0.04–0.10]	0.99	0.99	0.03	3763.32	3803.05
Depressogenic cognitive/personality styles									
11. Rumination	392	8.94	0.00	[0.00–0.05]	1.00	1.00	0.02	4513.31	4553.02
12. Self-criticism	392	22.83*	0.06	[0.03–0.09]	0.98	0.98	0.04	4365.95	4405.66
13. Dependency	392	18.61*	0.05	[0.01–0.08]	0.98	0.98	0.03	4542.85	4582.56
Social risk factors									
14. Social support	393	5.66	0.00	[0.00–0.03]	1.00	1.00	0.03	4912.06	4951.79
15. School engagement	393	18.81*	0.05	[0.01–0.08]	0.96	0.96	0.04	4756.43	4796.17
16. Bullying	393	24.79**	0.06	[0.03–0.09]	0.92	0.92	0.06	4810.43	4850.17
17. Parental criticism	383	14.99	0.04	[0.00–0.07]	0.99	0.99	0.03	4163.58	4203.06

Note. RMSEA, the root mean square error of approximation; CFI, comparative fit index; TLI, Tucker–Lewis index; SRMR, standardized root mean square residual; AIC, Akaike information criteria; BIC, Bayesian information criteria. We did not present fit indices with intercept set at wave 1 because no differences in fit indices were observed with intercept set at wave 1 *v.* wave 5.
*$p < 0.05$; **$p < 0.01$.

Estimates from latent growth curve models

To estimate level and change of risk, growth curve models were constructed for each risk marker with the intercept set at the first or the fifth assessment wave. Fit indices for the growth curve models are presented in Table 2. Growth curve models generally had adequate fit.

Level and change estimates from the growth curve models are presented in online Supplementary Table S3. The mean of the slope for each risk marker indicates the change rate of risk from wave 1 to 5. In terms of magnitude of change, 11 out of 17 risk markers showed significant change over time based on the means of the slopes ($ps < 0.05$). In terms of their change direction, the trajectories of 10 risk markers (i.e. general depression, ill temper, social anxiety, panic, traumatic intrusion, traumatic avoidance, mania, self-criticism, dependency, bullying) indicated a significant decrease in risk from wave 1 to 5, whereas the slope of one risk marker (i.e. parental criticism) indicated a significant increase in risk across waves.

The variance of the slope indicates interindividual difference in the intraindividual change of the risk factors over time. The variances of the slopes for all risk markers reached statistical significance ($ps < 0.01$), except for panic ($p = 0.56$) and traumatic intrusions ($p = 0.33$), indicating significant interindividual difference in the change of risk (see online Supplementary Fig. S2 for a spaghetti plot of each participant's risk trajectory in a random subsample of 10% of the current sample).

Predictive utility of level and change in risk

We created a structural model for each risk marker to examine the predictive utility of the level and change of risk factors on first DD onset. The latent intercept and slope of each risk marker estimated using data from wave 1 to 5 were used as predictors of DD onset at wave 6. Model fit information (i.e. log-likelihood value, the Akaike information criteria, and the Bayesian information criteria) is presented in online Supplementary Table S4. Table 3 displays the standardized results of structural models.

A higher level of risk at wave 5 significantly predicted first DD onset for all risk markers except for traumatic avoidance, traumatic intrusions, bullying, and parental criticism. In contrast, across all risk markers, the change in risk from wave 1 to 5 did not significantly predict DD onset after controlling for the wave 5 level of risk. A similar pattern was observed when controlling for wave 1 risk assessment as baseline. Across all risk factors, risk assessment at wave 1 significantly predicted first DD onset for all risk markers except for traumatic avoidance, traumatic intrusions, bullying, and parental criticism. For only two out of 17 risk markers (i.e. self-criticism; social support), an increase in risk significantly predicted first DD onset.

Discussion

This study extends our prior study (Mu et al., 2021) – which was the first to directly test the chronic risk and risk escalation

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

Table 3. Model results of structural equation model predicting first DD onset at wave 6 using latent intercept and slope estimated using standardized scores of each risk marker from wave 1 to 5

		Predicting onset using intercepts			Predicting onset using slopes		
		Estimate	S.E.	*p*	Estimate	S.E.	*p*
Depression and anxiety symptoms							
1. General depression	Wave 5	0.39**	0.07	0.00	−0.04	0.13	0.77
	Wave 1	0.42**	0.08	0.00	0.25	0.14	0.07
2. Ill temper	Wave 5	0.22**	0.08	0.00	−0.13	0.12	0.29
	Wave 1	0.30*	0.12	0.01	0.10	0.18	0.59
3. Social anxiety	Wave 5	0.27**	0.08	0.00	−0.06	0.11	0.60
	Wave 1	0.38**	0.11	0.00	0.25	0.16	0.11
4. Panic	Wave 5	0.24**	0.07	0.00	−0.12	0.19	0.53
	Wave 1	0.25**	0.07	0.00	0.02	0.22	0.93
5. Traumatic intrusions	Wave 5	0.14	0.08	0.06	−0.17	0.18	0.32
	Wave 1	0.15	0.08	0.05	−0.07	0.20	0.72
6. Traumatic avoidance	Wave 5	0.06	0.08	0.48	−0.13	0.13	0.32
	Wave 1	0.08	0.11	0.48	−0.07	0.18	0.70
7. Mania	Wave 5	0.18**	0.07	0.01	−0.11	0.10	0.29
	Wave 1	0.20**	0.08	0.01	0.06	0.10	0.60
Personality traits							
8. Neuroticism	Wave 5	0.37**	0.08	0.00	−0.16	0.11	0.16
	Wave 1	0.33**	0.07	0.00	0.11	0.10	0.28
9. Conscientiousness	Wave 5	−0.35**	0.07	0.00	0.06	0.10	0.56
	Wave 1	−0.35**	0.07	0.00	−0.16	0.10	0.12
10. Extraversion	Wave 5	−0.22**	0.08	0.00	−0.03	0.11	0.82
	Wave 1	−0.20**	0.07	0.01	−0.16	0.09	0.09
Depressogenic cognitive/personality styles							
11. Rumination	Wave 5	0.45**	0.08	0.00	−0.13	0.13	0.31
	Wave 1	0.39**	0.07	0.00	0.16	0.11	0.14
12. Self-criticism	Wave 5	0.43**	0.07	0.00	−0.04	0.11	0.69
	Wave 1	0.43**	0.07	0.00	0.32**	0.09	0.00
13. Dependency	Wave 5	0.23**	0.08	0.00	−0.11	0.12	0.39
	Wave 1	0.20**	0.07	0.00	0.08	0.11	0.46
Social risk factors							
14. Social support	Wave 5	−0.21**	0.07	0.01	−0.04	0.11	0.71
	Wave 1	−0.22**	0.08	0.01	−0.23*	0.10	0.03
15. School engagement	Wave 5	−0.31**	0.07	0.00	0.07	0.12	0.55
	Wave 1	−0.32**	0.08	0.00	−0.17	0.13	0.17
16. Bullying	Wave 5	0.05	0.07	0.53	−0.18	0.10	0.06
	Wave 1	0.06	0.10	0.54	−0.13	0.14	0.34
17. Parental criticism	Wave 5	0.12	0.08	0.11	0.08	0.11	0.46
	Wave 1	0.12	0.08	0.11	0.18	0.10	0.09

Note. Standardized results of structural models were presented. Two-sided statistical tests were performed at a level of significance of 5%. Wave 1 = intercepts set at wave 1, wave 5 = intercepts set at wave 5. Red indicates significant and positive prediction; green indicates significant and negative prediction. Lighter color indicates significance test coefficients lower than 0.05, whereas darker color indicates significance test coefficients lower than 0.01.
*$p < 0.05$; **$p < 0.01$.

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

hypotheses for depression – by doubling the follow-up period, using latent growth curve models instead of subtraction-based difference scores, and examining the transition from adolescence to emerging adulthood. For 13 out of 17 well-established risk markers for DD, chronic risk elevation during adolescence, but not change in risk, predicted first DD onset in transition to adulthood. Girls who had a first lifetime onset of DD in transition to adulthood have exhibited persistent and elevated levels of risk since entering adolescence. Our sample showed significant changes on almost all of the risk markers from early to late adolescence, but for 15 of the 17 risk factors tested, the change pattern did not significantly differ for those girls who had DD onset *v.* those who remained healthy. These results provide strong support for the chronic risk hypothesis, and little support for the risk escalation hypothesis.

The current findings advance our understanding of the developmental pathways of DD from adolescence to emerging adulthood. Our results demonstrated that chronic risk elevation, rather than an increase in risk over time, confers susceptibility to DD onset, supporting the chronic risk hypothesis. This pattern of results aligned with Mu et al. (2021), where the mean level of risk across multiple assessment waves outperformed risk escalation when predicting first DD onset in adolescence. Similarly, previous research observed strong links between persisting high levels of risk, such as sustained subthreshold depression, chronic stress, and the stable component of risk factors, and a greater likelihood of DD onset (e.g. Kendall et al., 2015; Klein et al., 2009). It is important to note that the current study focused on girls who had their first DD onset during transition to adulthood, whereas Mu et al. (2021) focused on girls who had their first DD onset in adolescence.

One natural question that follows is why these girls did not have DD onset earlier, like some other girls who had onset during the first five waves of assessment (Mu et al., 2021). Random exposure to life events may be an important mechanism for the timing of DD onset, as twin studies show that unique environmental variance explains over 60% of the risk of MDD (Sullivan, Neale, & Kendler, 2000). One related possibility is that some individuals may be more vulnerable to a specific set of stressors than others (Lazarus & Folkman, 1984). The type of stressors adolescents face (e.g. family conflicts, dramatic biological changes brought on by puberty) are different from those faced by emerging adults (e.g. independence, careers). Individual differences in vulnerability to stressors may explain the difference we observed in timing of DD onset. Another possibility pertains to a dose–response association between elevated risk for DD and health outcomes, with the impact of elevated risk accumulating until it reaches a threshold to trigger DD onset (Dunn et al., 2018). From this perspective, individuals with higher sustained levels of risk cross the threshold earlier than those with lower sustained levels of risk.

Unexpectedly, for four of the 17 risk markers (i.e. traumatic avoidance, traumatic intrusions, bullying, parental criticism), neither risk level nor change predicted DD onset, which is inconsistent with previous work supporting the predictive role of these risk markers (Klein et al., 2013; Mu et al., 2021). One possible explanation is that the depressogenic effects of these risk markers are time-limited (Shanahan, Copeland, Costello, & Angold, 2011). Thus, levels and changes in these risk markers in adolescence may have a weaker influence on mental health in the transition to adulthood. For example, Shanahan et al. (2011) found that the associations between psychosocial risk factors and DD onset were stronger when they were measured concurrently than at different ages. Similarly, Jaffee et al. (2002) found that early childhood levels of risk had a stronger predictive effect on juvenile-onset DD than on adult-onset DD. Future research is needed to clarify whether certain risk factors have time-limited depressogenic effects. An alternative possibility is that the influence of some risk factors differs as a function of developmental stage (Jaffee et al., 2002; Shanahan et al., 2011). However, this is less likely given that previous research has supported the predictive effect of trauma-related symptoms and interpersonal factors on adult-onset DD.

According to the risk escalation hypothesis, risk should increase throughout adolescence for girls who developed DD. However, for all 17 risk markers examined, change of risk from early to late adolescence did not predict DD onset when the effect of risk level was controlled (worsening in self-criticism and social support no longer predicted DD onset when its level at wave 5 was controlled). This pattern is consistent with Mu et al. (2021), but inconsistent with other past evidence supporting the risk escalation hypothesis (e.g. Laceulle et al., 2014; Mu et al., 2016; Steiger et al., 2014). The current study estimates change using data from five assessment waves, thereby allowing us to separate slope and intercept more precisely as compared to previous studies. With a more precise estimation of change, we still found no evidence supporting the predictive role of change in risk levels. The current study has several limitations. First, the window for DD onsets was wide, that is, up to 3 years after wave 5, allowing the possibility that our model could have missed an escalation shortly before onset. However, this possibility is not likely, given that Mu et al. (2021) found no evidence supporting the predictive effect of risk change 9-month priors to onset when risk level was controlled. Second, participants were mainly white girls from a particular region of the USA, which limits the generalizability of our findings to other cultures and populations. Replication of the current study in more diverse samples is warranted. Third, the assessment of risk factors relied solely on self-report. Although self-report has been shown to be the best assessment approach for many risk factors (Babor, Brown, & Del Boca, 1990), multiple methods (e.g. behavioral observations, physiological and neurocognitive assessments) should be employed in future studies to provide a more comprehensive picture of risk profiles. Fourth, the current study did not examine an exhaustive list of risk markers, as some other well-established risk markers, such as romantic relationship, which has been shown to be effective for adolescent girls, (Starr et al., 2012) were not included. More studies could be conducted on whether other risk markers fit the same pattern in the future.

Notwithstanding these limitations, the current study has the following strengths. First, we explored the developmental period spanning adolescence to emerging adulthood, which is when DD incidence peaks (Rohde et al., 2013). More importantly, we examined the long-term effects of risk level and change on first DD onset. Second, we estimated the latent intercepts and slopes of risk factors using five assessments over 3 years, whereas most previous studies estimated risk change using data from two assessment waves, rendering it difficult to control for measurement errors or to disentangle lasting change from temporal fluctuations (e.g. Laceulle et al., 2014; Mu et al., 2021).

In conclusion, we directly compared the chronic risk and risk escalation hypotheses. Our findings demonstrate that for the established depression risk factors, elevated level of risk, but not risk change, predicts DD onset. They also argue that individual

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press

differences in depression risk emerge in early adolescence or even childhood and are maintained in the absence of clinical depression. This suggests that to predict and prevent DD, researchers and practitioners should focus on persistently elevated levels of risk to identify individuals who are most likely to develop DD and to provide targeted DD prevention. In addition, adolescents with high levels of risk factors can be targeted for programs to prevent the development of DD. School-based monitoring of risk level among adolescents would be informative for identifying those who are especially vulnerable for DD.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291723002441.

Financial support. This work was supported by the National Institute of Mental Health of the National Institutes of Health (R01MH093479 to R. K.).

Competing interests. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Note

[1] These date are critical because on 20 March 2020, New York State declared a state-wide stay-at-home order, with all non-essential businesses closed and all non-essential gatherings cancelled/postponed (Francescani, 2020).

References

Babor, T. F., Brown, J., & Del Boca, F. K. (1990). Validity of self-reports in applied research on addictive behaviors: Fact or fiction? *Behavioral Assessment*, *12*(1), 5–31. Retrieved from https://www.proquest.com/scholarly-journals/validity-self-reports-applied-research-on/docview/617758254/se-2

Bagby, R. M., Parker, J. D., Joffe, R. T., & Buis, T. (1994). Reconstruction and validation of the depressive experiences questionnaire. *Assessment*, *1*(1), 59–68. https://doi.org/10.1177/1073191194001001009

Bagby, R. M., Quilty, L. C., & Ryder, A. C. (2008). Personality and depression. *The Canadian Journal of Psychiatry*, *53*(1), 14–25. https://doi.org/10.1177/070674370805300104

Blatt, S. J., D'Afflitti, J. P., & Quinlan, D. M. (1976). Experiences of depression in normal young adults. *Journal of Abnormal Psychology*, *85*(4), 383. https://doi.org/10.1037/0021-843X.85.4.383

Burkhouse, K. L., Uhrlass, D. J., Stone, L. B., Knopik, V. S., & Gibb, B. E. (2012). Expressed emotion-criticism and risk of depression onset in children. *Journal of Clinical Child & Adolescent Psychology*, *41*(6), 771–777. https://doi.org/10.1080/15374416.2012.703122

De Los Reyes, A., & Prinstein, M. J. (2004). Applying depression-distortion hypotheses to the assessment of peer victimization in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *33*(2), 325–335. https://doi.org/10.1207/s15374424jccp3302_14

Dunn, E. C., Soare, T. W., Raffeld, M. R., Busso, D. S., Crawford, K. M., Davis, K. A., … Susser, E. S. (2018). What life course theoretical models best explain the relationship between exposure to childhood adversity and psychopathology symptoms: Recency, accumulation, or sensitive periods? *Psychological Medicine*, *48*(15), 2562–2572. https://doi.org/10.1017/S0033291718000181

Francescani, C. (2020, June 17). Timeline: The first 100 days of New York Gov. Andrew Cuomo's COVID-19 response. Retrieved from https://abcnews.go.com/US/News/timeline-100-days-york-gov-andrew-cuomos-covid/story?id=71292880

Furman, W., & Buhrmester, D. (2009). Methods and measures: The network of relationships inventory: Behavioral systems version. *International Journal of Behavioral Development*, *33*(5), 470–478. https://doi.org/10.1177/0165025409342634

Goldstein, B. L., Kotov, R., Perlman, G., Watson, D., & Klein, D. N. (2018). Trait and facet-level predictors of first-onset depressive and anxiety disorders in a community sample of adolescent girls. *Psychological Medicine*, *48*(8), 1282–1290. https://doi.org/10.1017/S0033291717002719

Goldstein, B. L., Perlman, G., Eaton, N. R., Kotov, R., & Klein, D. N. (2020). Testing explanatory models of the interplay between depression, neuroticism, and stressful life events: A dynamic trait-stress generation approach. *Psychological Medicine*, *50*(16), 2780–2789. https://doi.org/10.1017/S0033291719002927

Hammen, C., Kim, E. Y., Eberhart, N. K., & Brennan, P. A. (2009). Chronic and acute stress and the prediction of major depression in women. *Depression and Anxiety*, *26*(8), 718–723. https://doi.org/10.1002/da.20571

Hirschfeld, R. M., Klerman, G. L., Gouch, H. G., Barrett, J., Korchin, S. J., & Chodoff, P. (1977). A measure of interpersonal dependency. *Journal of Personality Assessment*, *41*(6), 610–618. https://doi.org/10.1207/s15327752jpa4106_6

Hooper, D., Coughlan, J., & Mullen, M. (2008). Evaluating model fit: a synthesis of the structural equation modelling literature. In 7th European Conference on research methodology for business and management studies (pp. 195–200).

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Jaffee, S. R., Moffitt, T. E., Caspi, A., Fombonne, E., Poulton, R., & Martin, J. (2002). Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry*, *59*(3), 215–222. https://doi.org/10.1001/archpsyc.59.3.215

Jenness, J. L., Peverill, M., King, K. M., Hankin, B. L., & McLaughlin, K. A. (2019). Dynamic associations between stressful life events and adolescent internalizing psychopathology in a multiwave longitudinal study. *Journal of Abnormal Psychology* (1965), *128*(6), 596–609. https://doi.org/10.1037/abn0000450

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, *2*(1999), 102–138. Retrieved from https://www.proquest.com/books/big-five-trait-taxonomy-history-measurement/docview/619404914/se-2

Kaufman, J., Birmaher, B., Brent, D., Rao, U. M. A., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. https://doi.org/10.1097/00004583-199707000-00021

Kendall, A. D., Zinbarg, R. E., Mineka, S., Bobova, L., Prenoveau, J. M., Revelle, W., & Craske, M. G. (2015). Prospective associations of low positive emotionality with first onsets of depressive and anxiety disorders: Results from a 10-wave latent trait-state modeling study. *Journal of Abnormal Psychology*, *124*(4), 933. https://doi.org/10.1037/abn0000105

Klein, D. N., Glenn, C. R., Kosty, D. B., Seeley, J. R., Rohde, P., & Lewinsohn, P. M. (2013). Predictors of first lifetime onset of major depressive disorder in young adulthood. *Journal of abnormal psychology*, *122*(1), 1. https://doi.org/10.1037/a0029567

Klein, D. N., Kotov, R., & Bufferd, S. J. (2011). Personality and depression: Explanatory models and review of the evidence. *Annual Review of Clinical Psychology*, *7*(1), 269–295. https://doi.org/10.1146/annurev-clinpsy-032210-104540

Klein, D. N., Shankman, S. A., Lewinsohn, P. M., & Seeley, J. R. (2009). Subthreshold depressive disorder in adolescents: Predictors of escalation to full-syndrome depressive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(7), 703–710. https://doi.org/10.1097/CHI.0b013e3181a56606

Kopala-Sibley, D. C., Klein, D. N., Perlman, G., & Kotov, R. (2017). Self-criticism and dependency in female adolescents: Prediction of first onsets and disentangling the relationships between personality, stressful life events, and internalizing psychopathology. *Journal of Abnormal Psychology*, *126*(8), 1029. https://doi.org/10.1037/abn0000297

Laceulle, O. M., Ormel, J., Vollebergh, W. A., Van Aken, M. A., & Nederhof, E. (2014). A test of the vulnerability model: Temperament and temperament change as predictors of future mental disorders – the TRAILS study.

Journal of Child Psychology and Psychiatry, *55*(3), 227–236. https://doi.org/10.1111/jcpp.12141

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Malhi, G. S., & Mann, J. J. (2018). Depression. *The Lancet* (*British Edition*), *392*(10161), 2299–2312. https://doi.org/10.1016/S0140-6736(18)31948-2

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341.

McCoach, D. B., & Siegle, D. (2003). The school attitude assessment survey-revised: A new instrument to identify academically able students who underachieve. *Educational and Psychological Measurement*, *63*(3), 414–429. https://doi.org/10.1177/0013164403063003005

Michelini, G., Perlman, G., Tian, Y., Mackin, D. M., Nelson, B. D., Klein, D. N., & Kotov, R. (2021). Multiple domains of risk factors for first onset of depression in adolescent girls. *Journal of Affective Disorders*, *283*, 20–29. https://doi.org/10.1016/j.jad.2021.01.036

Mu, W., Li, K., Tian, Y., Perlman, G., Michelini, G., Watson, D., … Kotov, R. (2021). Dynamic risk for first onset of depressive disorders in adolescence: Does change matter?. *Psychological Medicine*, *53*(6), 2352–2360. https://doi.org/10.1017/S0033291721004190

Mu, W., Luo, J., Nickel, L., & Roberts, B. W. (2016). Generality or specificity? Examining the relation between personality traits and mental health outcomes using a bivariate bi-factor latent change model. *European Journal of Personality*, *30*(5), 467–483. https://doi.org/10.1002/per.2052

Naragon-Gainey, K., Gallagher, M. W., & Brown, T. A. (2013). Stable 'trait' variance of temperament as a predictor of the temporal course of depression and social phobia. *Journal of Abnormal Psychology*, *122*(3), 611. https://doi.org/10.1037/a0032997

Nelson, B., McGorry, P. D., Wichers, M., Wigman, J. T. W., & Hartmann, J. A. (2017). Moving from static to dynamic models of the onset of mental disorder: A review. *JAMA Psychiatry*, *74*(5), 528–534. https://doi.org/10.1001/jamapsychiatry.2017.0001

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, *101*(2), 259. https://doi.org/10.1037/0033-2909.101.2.259

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*(3), 504. https://doi.org/10.1037/0021-843X.109.3.504

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*(1), 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1. https://doi.org/10.1037/0033-2909.132.1.1

Rohde, P., Lewinsohn, P. M., Klein, D. N., Seeley, J. R., & Gau, J. M. (2013). Key characteristics of major depressive disorder occurring in childhood, adolescence, emerging adulthood, and adulthood. *Clinical Psychological Science*, *1*(1), 41–53. https://doi.org/10.1177/2167702612457599

Shanahan, L., Copeland, W. E., Costello, E. J., & Angold, A. (2011). Child-, adolescent- and young adult-onset depressions: Differential risk factors in development? *Psychological Medicine*, *41*(11), 2265–2274. https://doi.org/10.1017/S0033291711000675

Shochet, I. M., Homel, R., Cockshaw, W. D., & Montgomery, D. T. (2008). How do school connectedness and attachment to parents interrelate in predicting adolescent depressive symptoms? *Journal of Clinical Child and Adolescent Psychology*, *37*(3), 676–681. https://doi.org/10.1080/15374410802148053

Smith, O. A., Nelson, J. A., & Adelson, M. J. (2019). Interparental and parent–child conflict predicting adolescent depressive symptoms. *Journal of Child and Family Studies*, *28*(7), 1965–1976. https://doi.org/10.1007/s10826-019-01424-6

Starr, L. R., & Davila, J. (2008). Differentiating interpersonal correlates of depressive symptoms and social anxiety in adolescence: Implications for models of comorbidity. *Journal of Clinical Child & Adolescent Psychology*, *37*(2), 337–349. https://doi.org/10.1080/15374410801955854

Starr, L. R., Davila, J., Stroud, C. B., Clara Li, P. C., Yoneda, A., Hershenberg, R., & Ramsay Miller, M. (2012). Love hurts (in more ways than one): Specificity of psychological symptoms as predictors and consequences of romantic activity among early adolescent girls. *Journal of Clinical Psychology*, *68*(4), 373–381. https://doi.org/10.1002/jclp.20862

Steiger, A. E., Allemand, M., Robins, R. W., & Fend, H. A. (2014). Low and decreasing self-esteem during adolescence predict adult depression two decades later. *Journal of Personality and Social Psychology*, *106*(2), 325. https://doi.org/10.1037/a0035133

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, *157*(10), 1552–1562.

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology*, *84*(1), 205–220. https://doi.org/10.1037/0022-3514.84.1.205

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., … Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment*, *19*(4), 399–420. https://doi.org/10.1177/1073191112449857

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, *52*(1), 30–41. https://doi.org/10.1207/s15327752jpa5201_2

https://doi.org/10.1017/S0033291723002441 Published online by Cambridge University Press