



VCU

Virginia Commonwealth University
VCU Scholars Compass

Psychiatry Publications

Dept. of Psychiatry

2017

Symptoms of major depression: Their stability, familiarity, and prediction by genetic, temperamental, and childhood environmental risk factors

Kenneth S. Kendler

Virginia Commonwealth University, kenneth.kendler@vcuhealth.org

Steven H. Aggen

Virginia Commonwealth University

Follow this and additional works at: http://scholarscompass.vcu.edu/psych_pubs

 Part of the [Psychiatry and Psychology Commons](#)

(C) 2017 Wiley Periodicals, Inc.

Downloaded from

http://scholarscompass.vcu.edu/psych_pubs/76

This Article is brought to you for free and open access by the Dept. of Psychiatry at VCU Scholars Compass. It has been accepted for inclusion in Psychiatry Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

RESEARCH ARTICLE

Symptoms of major depression: Their stability, familiarity, and prediction by genetic, temperamental, and childhood environmental risk factors

 Kenneth S. Kendler, M.D.^{1,2} | Steven H. Aggen, Ph.D.¹

¹Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, VA, USA

²Department of Human and Molecular Genetics, Virginia Commonwealth University School of Medicine, Richmond, VA, USA

Correspondence

Kenneth S. Kendler, Virginia Institute for Psychiatric and Behavioral Genetics of VCU, Box 980126, Richmond, VA 23298-0126. Email: kenneth.kendler@vcuhealth.org.

Contract grant numbers: MH40828 and UL1RR031990.

Background: Psychiatry has long sought to develop biological diagnostic subtypes based on symptomatic differences. This effort assumes that symptoms reflect, with good fidelity, underlying etiological processes. We address this question for major depression (MD).

Methods: We examine, in twins from a population-based registry, similarity in symptom endorsement in individuals meeting criteria for last-year MD at separate interview waves and in concordant twin pairs. Among individuals with MD, we explore the impact of genetic-temperamental and child adversity risk factors on individual reported symptoms. Aggregated criteria do not separate insomnia from hypersomnia, weight gain from loss, etc. while disaggregated criteria do.

Results: In twins with MD at two different waves, the mean tetrachoric correlations (\pm SEM) for aggregated and disaggregated DSM-IV A criteria were, respectively, $+0.31 \pm 0.06$ and $+0.34 \pm 0.03$. In monozygotic (MZ) and dizygotic (DZ) twin pairs concordant for last-year MD, the mean tetrachoric correlations for aggregated and disaggregated criteria were, respectively, $+0.33 \pm 0.07$ and $+0.43 \pm 0.04$, and $+0.05 \pm 0.08$ and $+0.07 \pm 0.04$. In individuals meeting MD criteria, neuroticism predicted the most MD symptoms (10), followed by childhood sexual abuse (8), low parental warmth (6), and genetic risk (4).

Conclusions: The correlations for individual depressive symptoms over multiple episodes and within MZ twins concordant for MD are modest suggesting the important role of transient influences. The multidetermination of individual symptoms was further evidenced by their prediction by personality and exposure to early life adversities. The multiple factors influencing symptomatic presentation in MD may contribute to our difficulties in isolating clinical depressive subtypes with distinct pathophysiologies.

KEYWORDS

genetics, major depression, stability, symptoms, twins

1 | INTRODUCTION

Nature, in the production of disease, is uniform and consistent; so much so, that for the same disease in different persons the symptoms are for the most part the same; and the selfsame phenomena that you would observe in the sickness of a Socrates you would observe in the sickness of a simpleton. Sydenham (1676/1981) ((Sydenham, 1981), p. 148).

The essential clinical phenomena which occur in any illness are, first, those which refer to the actual cause of the illness, which endow it with a specific character; second those that may be said to share the disorder in that they give content, colouring, and contour to the individual illnesses whose basic form and character have already been biologically established. The former group of phenomena I would call pathogenic, and the latter *pathoplastic* Birnbaum (1923/1974) ((Birnbaum, 1974), p. 203).

Since the beginning of modern neuropsychiatry in 19th century in Europe, efforts have been made to define clinically distinct and increasingly homogeneous syndromes with the hope of better

understanding their etiology (Kendler & Engstrom, 2016). This approach assumes that individual symptoms reflect, with some fidelity, underlying etiologic processes. Sydenham above reflects the optimistic view that symptoms are highly indicative of the underlying disease state and stable across sufferers (Sydenham, 1981). Birnbaum articulates a more complex view that psychiatric symptoms can reflect *pathogenic* processes—arising from stable underlying pathophysiology—and *pathoplastic* processes that result from a wide variety of different factors including age, sex, character, and environmental experiences.

Despite the conceptual importance of this question, empirical evaluations of the sources of symptoms of psychiatric disorders have been infrequent. One approach has been to examine the correlation of symptoms in pairs of affected relatives with the goal of clarifying the importance of familial influences on symptom occurrence. This has been studied most extensively for schizophrenia where results in concordant twin and siblings pairs suggest modest to moderate correlations for most symptoms (Bleuler, 1978; Cardno et al., 1999; Cardno, Sham, Murray, & McGuffin, 2001; DeLisi, Goldin, Maxwell, Kazuba, & Gershon, 1987; Kendler et al., 1997). A meta-analysis of the three schizophrenic symptom dimensions estimated sibling correlations at +0.18 to +0.28 (Rietkerk et al., 2008). A different approach found that specific types of stressful life events were related to distinct depressive symptom profiles (Keller, Neale, & Kendler, 2007), suggesting that such events function as pathoplastic influences.

This report seeks to clarify further the causes of symptoms of major depression (MD). We pose three questions about individuals who met DSM-IV criteria for MD in the last year in multiple waves of a longitudinal population-based twin study (Kendler & Prescott, 2006). First, in twins who met criteria for MD on at least two different occasions, how stably are individual MD symptoms reported across episodes? Second, what are the correlations in symptoms of depressive episodes in the last year in monozygotic (MZ) and dizygotic (DZ) twin pairs concordant for MD? Third, among affected individuals, how well can the occurrence of individual depressive symptoms be predicted by biological/temperamental and childhood environmental risk factors?

2 | MATERIALS AND METHODS

2.1 | Sample

Participants were from two cohorts that formed the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (Kendler, & Prescott, 2006). Caucasian twins were ascertained from the birth-certificate based Virginia Twin Registry. Female-female (FF) twin pairs, born 1934–1974, were eligible if both members responded to a mailed questionnaire in 1987–1988. This initial sample ($N = 2,163$, mean age \pm (SD) 30.1 (7.6)) was interviewed in person or by phone four times from January 1987 to April 1997 with cooperation rates ranging from 85 to 92% (Kendler et al., 2006). The male-male/male-female pairs (MMM/MFM, $N = 6,812$ mean age \pm (SD) 35.5 (9.1)), birth years 1940–1974, were ascertained with a 72% cooperation rate, directly from registry

records. The first interview was completed largely by phone between 1993 and 1996. A second interview conducted between 1994 and 1998, had a response rate of 83%.

Interviews were conducted at least 1 year apart by trained mental health professionals blind to knowledge about the cotwin. Informed consent was obtained for all personal interviews and assent for all phone interviews. Zygosity was determined by discriminate function analyses using standard questions validated against DNA genotyping (Kendler & Prescott, 1999).

2.2 | Measures

In all interviews, twins were asked about the occurrence in the last year of 17 symptoms reflecting all DSM-IV A criteria for MD (American Psychiatric Association, 1994). As detailed elsewhere (Kendler et al., 2006), questions for this section were adapted from the SCID interview (Spitzer & Williams, 1985). Follow-up inquiries asked whether these symptoms occurred at the same time, reported total number of episodes and dates, to the month, for the onset and offset of each episode. Test-retest reliability for last year MD, based on two blinded interviews 1–2 months apart, was good: $\kappa = 0.74$ (SE = 0.08) tetrachoric $r = 0.96 \pm 0.03$. Neuroticism was assessed by the Eysenck Personality Questionnaire short form (Eysenck, Eysenck, & Barrett, 1985). A percent correct score was calculated for each twin from their pattern of responses to the 12 Neuroticism items from the EPQ short-scale obtained at the first waves (Eysenck et al., 1985). Genetic risk was assessed by a composite measure of the lifetime history of MD in the co-twin, and in the mother and father as assessed by both twins using the Family History Research Diagnostic Criteria (Endicott, Andreasen, & Spitzer, 1975) accounting for the varying genetic correlation with the proband (1.00 for MZ cotwins and 0.50 for DZ co-twins and parents). Parental coldness as reported by the twin on their mother and father was measured by a modified version of Parker's Parental Bonding Instrument (Parker, Tupling, & Brown, 1979). Childhood sexual abuse before the age of 16 was the only risk factor utilized here which was assessed differently in the two cohorts. In the FF sample, it was assessed by questions developed by Mullen and colleagues (Martin, Anderson, Romans, Mullen, & O'Shea, 1993). For this report, we counted as positive reports of fondling and sexual touching and attempted or completed intercourse. In the MMM/MFM sample, it was assessed by a single item asked about being "sexually abused or molested" prior to age 16. These two assessment methods are similarly predictive of MD in women (Kendler, Gardner, & Prescott, 2002) and men (Kendler, Gardner, & Prescott, 2006). All nonbinary predictor variables were standardized to a mean of 0 and standard deviation of 1 to facilitate within analysis comparisons.

2.3 | Statistical modeling

To empirically evaluate our substantive questions, we organized our data in three different ways. First, for the within-person across waves correlational analyses, an individual record dataset was constructed by identifying all twins that algorithmically met MD criteria in at least two different waves ($N = 448$). One set collected item responses for the first time the MD criterion was met and an independent set for the

second time. For the FF sample, these data could come from any of the possible pairings of the four waves. Second, a twin pair dataset was assembled to examine individual criteria correlations for pairs of twins concordant for MD ($N = 224$). The twin pair dataset was constructed in the same manner as was the within-person across waves dataset. For the FF and MMMF samples, both twins within a pair had to have a positive MD diagnosis at one of the four or two waves, respectively. Third, for the regression analyses, a different individual record dataset was constructed consisting of new binary criteria variable sets that contained within-person symptom endorsements from the first time, in which an individual met the MD diagnostic cutoff ($N = 1,716$).

To obtain estimates of the within-person across time correlations, 2×2 contingency tables were generated for each of the compiled time 1 and time 2 binary MD symptoms. Tetrachoric correlations were estimated using the SAS version 9.4 procedure CORR with the option POLYCHORIC (SAS Institute, 2012). Tetrachoric correlations, first proposed by Karl Pearson (1901), examine dichotomous (i.e., “yes-no”) variables. But instead of assessing the correlation in the overt measure, they estimate the correlation in an underlying normal distributed liability. They can vary from -1.00 to $+1.00$ and can be interpreted in the same way as a standard Pearson product-moment correlation.

The version 9.4 SAS statistical procedure GENMOD (SAS Institute, 2012) was used to fit general estimating equation (GEE) logistic regression models using the logit link and binary distribution options to obtain parameter estimates for the set of selected predictors of the binary MD symptom criteria. GEE was implemented to account for the observation nesting due to concordant twins for MD ($N = 225$). An unstructured correlation form was specified for the purpose of adjusting standard errors for the nonindependent observations.

3 | RESULTS

The frequency of endorsement for the nine aggregated and eight disaggregated DSM-IV depressive symptoms among those meeting diagnostic criteria for MD are seen in the first column of Table 1. They ranged from 18.0% for weight gain to 98.4% for depressed mood. Because of insufficient variance in response to depressed mood, this criterion had to be eliminated from further analysis.

In the 448 individuals who reported MD episodes in the last year at different interview waves, correlation in the aggregated and disaggregated criteria are seen in Table 1 which had mean (\pm SEM) correlations of $+0.31 \pm 0.06$ [SD -15.6] and $+0.34 \pm 0.03$ [SD 7.5], respectively. While all significant, individual correlations for the aggregated criteria varied widely from $+0.14$ and $+0.17$ for sleep change and tiredness to $+0.40$ and $+0.63$ for worthlessness and thoughts of death. Correlations were moderately higher for disaggregated than for aggregated criteria for appetite, sleep, and psychomotor changes with a notably high value of $+0.49$ for hypersomnia.

In the 86 pairs of MZ twins concordant for last year MD, significant correlations were seen for all criterion except sleep change (Table 1). Correlations in the aggregated and disaggregated criteria had a mean of $+0.33 \pm 0.07$ and $+0.43 \pm 0.04$, respectively.

Results were very different in the 138 pairs of DZ twins concordant for last year MD. In this sample, we had insufficient variance for the anhedonia criterion to yield meaningful results. Of the remaining 16 criteria, statistically significant resemblance was seen in only 3 of them: sleep change, trouble sleeping and difficulty concentrating. Correlations in the aggregated and disaggregated criteria had a mean of $+0.05 \pm 0.08$ and $+0.07 \pm 0.04$, respectively.

Our aggregate genetic risk factor index was significantly associated with reporting four symptoms all in a positive direction: appetite gain, psychomotor change, psychomotor agitation, and feelings of worthlessness. By contrast, endorsement of 10 of the 16 available criteria was significantly predicted by neuroticism scores, all but one (appetite loss), in the positive direction. Notably, feelings of worthlessness and thoughts of death—among the most stably reported criteria within individuals—were the most strongly predicted by neuroticism. A number of the other relatively stable symptoms—particularly hypersomnia, appetite gain, and psychomotor agitation—were also strongly predicted by neuroticism.

Endorsement of eight of the 16 available criteria was significantly predicted by a history of childhood sexual abuse all in the positive direction. The most strongly predicted symptoms were loss of interest, weight gain, and thoughts of death. Other relatively stable symptoms predicted by a history of abuse included hypersomnia, psychomotor retardation, and feelings of worthlessness.

Parental coldness was significantly associated with reporting of six depressive symptoms, five in the expected positive direction. The effect sizes of these five (appetite, and weight gain, hypersomnia, feelings of worthlessness, and thoughts of death) were all relatively similar.

4 | DISCUSSION

The goal of this report was to further clarify, through addressing three questions, the nature of the influences on depressive symptoms among individuals meeting criteria for MD. First, we asked: to what degree are the occurrences of individual symptoms of MD stable over multiple depressive episodes? Individual DSM depressive criteria were only modestly correlated across episodes occurring at least 1 year apart with a mean correlation slightly exceeding $+0.30$. Furthermore, cross-episode stability varied widely among the individual criteria. If most symptoms of MD were strongly pathogenic—that is, reflect closely the pathophysiological substrate activated when individuals enter into depressive episodes—the interepisode correlation among the symptoms would be likely to be higher. Instead, these results suggest that episode-specific factors strongly influence depressive symptomatology.

These results are complementary to the findings obtained by Keller et al. (2007) in this same sample. In that study, the symptomatic presentation of depression was substantially influenced by the kind of stressful life event that precipitated the depressive syndrome. Our results—modest temporal stability of individual symptoms—likely reflect the same process. Personal experiences around episode onset may provide important pathoplastic influences on the symptoms of MD.

TABLE 1 Frequency, tetrachoric correlations for endorsement of disaggregated DSM-IV criteria for last-year major depression within individuals over time and within monozygotic and dizygotic twin pairs concordant for major depression and the prediction of these criteria by genetic risk for major depression, neuroticism, childhood sexual abuse, and parental coldness

DSM-IV A criterion number and description	Frequency % Endorsed N = 1,716	Within-individual		Between MZ twin pairs		Between DZ twin pairs		Prediction by genetic risk		Prediction by neuroticism		Prediction by child sexual abuse		Prediction by parental coldness	
		Tetrachoric correlation (\pm SE)	NE	Tetrachoric correlation (\pm SE)	NE	Tetrachoric correlation (\pm SE)	NE	Linear regression beta (\pm 95% CIs)	NE	Linear regression beta (\pm 95% CIs)	NE	Logistic regression Odds ratio (\pm 95% CIs)	NE	Linear regression beta (\pm 95% CIs)	NE
1. Depressed mood	98.4	0.23 (\pm 0.14)	NE	0.49 (\pm 0.25)	NE	0.16 (0.00–0.32)	NE	0.16 (0.00–0.32)	NE	–0.05 (–0.21 to 0.11)	NE	1.94 (1.17–3.22)	NE	–0.06 (–0.22 to 0.11)	NE
2. Anhedonia	88.6	0.35 (\pm 0.08)	NE	0.43 (\pm 0.19)	NE	–0.33 (\pm 0.18)	NE	0.12 (–0.01–0.25)	NE	–0.03 (–0.15 to 0.09)	NE	1.19 (0.84–1.68)	NE	0.02 (–0.10 to 0.14)	NE
3. Appetite/weight Change	77.3	0.37 (\pm 0.07)	NE	0.30 (\pm 0.16)	NE	–0.01 (\pm 0.14)	NE	0.00 (–0.10 to 0.11)	NE	–0.15 (–0.25 to –0.04)	NE	0.92 (0.70–1.21)	NE	–0.08 (–0.18 to 0.02)	NE
3a. Appetite loss	54.3	0.36 (\pm 0.08)	NE	0.62 (\pm 0.15)	NE	0.06 (\pm 0.16)	NE	0.05 (–0.09 to 0.18)	NE	0.19 (0.07–0.31)	NE	1.36 (1.00–1.85)	NE	0.15 (0.02–0.27)	NE
3b. Appetite gain	21.6	0.29 (\pm 0.08)	NE	0.37 (\pm 0.17)	NE	0.15 (\pm 0.14)	NE	0.06 (–0.05 to 0.17)	NE	–0.10 (–0.21 to 0.01)	NE	1.34 (1.00–1.78)	NE	0.01 (–0.10 to 0.12)	NE
3c. Weight loss	31.2	0.31 (\pm 0.09)	NE	0.53 (\pm 0.17)	NE	0.16 (\pm 0.17)	NE	0.15 (0.02–0.29)	NE	0.14 (0.01–0.27)	NE	1.93 (1.40–2.65)	NE	0.21 (0.08–0.33)	NE
3d. Weight gain	18.0	0.14 (\pm 0.10)	NE	–0.17 (\pm 0.23)	NE	0.38 (\pm 0.17)	NE	0.08 (–0.06 to 0.22)	NE	0.00 (–0.12 to 0.13)	NE	1.23 (0.85–1.76)	NE	–0.01 (–0.04 to 0.02)	NE
4. Sleep change	80.4	0.38 (\pm 0.07)	NE	0.41 (\pm 0.17)	NE	0.22 (\pm 0.14)	NE	0.06 (–0.06 to 0.17)	NE	–0.03 (–0.14 to 0.07)	NE	0.82 (0.62–1.10)	NE	–0.04 (–0.06 to –0.01)	NE
4a. Trouble sleeping	67.7	0.49 (\pm 0.07)	NE	0.37 (\pm 0.17)	NE	0.01 (\pm 0.15)	NE	0.07 (–0.04 to 0.19)	NE	0.14 (0.03–0.26)	NE	1.70 (1.26–2.29)	NE	0.04 (0.01–0.07)	NE
4b. Sleep more than usual	24.2	0.25 (\pm 0.09)	NE	0.50 (\pm 0.16)	NE	–0.13 (\pm 0.17)	NE	0.13 (0.01–0.26)	NE	0.14 (0.02–0.26)	NE	1.42 (1.02–1.98)	NE	0.05 (–0.07 to 0.17)	NE
5. Psychomotor change	77.4	0.37 (\pm 0.07)	NE	0.43 (\pm 0.15)	NE	0.09 (\pm 0.14)	NE	0.14 (0.03–0.25)	NE	0.17 (0.06–0.27)	NE	1.11 (0.84–1.47)	NE	–0.04 (–0.15 to 0.06)	NE
5a. Psychomotor agitation	60.8	0.28 (\pm 0.07)	NE	0.31 (\pm 0.16)	NE	–0.16 (\pm 0.14)	NE	0.03 (–0.07 to 0.13)	NE	0.14 (0.04–0.25)	NE	1.45 (1.10–1.92)	NE	0.05 (–0.05 to 0.16)	NE
5b. Psychomotor retardation	37.2	0.17 (\pm 0.09)	NE	0.35 (\pm 0.17)	NE	0.10 (\pm 0.16)	NE	0.06 (–0.06 to 0.19)	NE	0.13 (0.01–0.25)	NE	0.99 (0.72–1.36)	NE	–0.01 (–0.12 to 0.10)	NE
6. Feeling tired/fatigued	74.4	0.40 (\pm 0.07)	NE	0.19 (\pm 0.17)	NE	0.09 (\pm 0.14)	NE	0.15 (0.04–0.26)	NE	0.50 (0.39–0.61)	NE	1.34 (1.02–1.77)	NE	0.16 (0.09–0.29)	NE
7. Feeling worthless	57.7	0.32 (\pm 0.07)	NE	0.44 (\pm 0.15)	NE	0.30 (\pm 0.13)	NE	0.05 (–0.05 to 0.16)	NE	0.09 (–0.02 to 0.19)	NE	1.06 (0.80–1.41)	NE	–0.03 (–0.13 to 0.08)	NE
8. Difficulty concentrating	58.8	0.63 (\pm 0.06)	NE	0.48 (\pm 0.16)	NE	–0.10 (\pm 0.17)	NE	0.10 (–0.02 to 0.22)	NE	0.36 (0.24–0.48)	NE	1.84 (1.35–2.51)	NE	0.17 (0.11–0.39)	NE
9. Thoughts of death	23.2		NE		NE		NE		NE		NE		NE		NE

Note: Covariates included in all logistic regression analyses—age, sex, and interview wave. Logistic regression predictor variables are standardized. Statistical estimates at nominal significance level ($P < .05$, two-tailed) are in **bold**. NE, not estimable; \pm SE, asymptotic standard errors; \pm 95% CIs, symmetric Wald-based 95% confidence intervals for GEE regression estimates. Aggregated symptoms are in normal type while disaggregated symptoms are in italics.

Second, we examined individual symptoms in MZ and DZ twin pairs concordant for last year MD to understand the impact of aggregate genetic and familial-environmental effects on endorsement. Symptomatic resemblance in MZ pairs was modest with mean correlations around +0.35 and very similar to that seen within individuals across time. Symptom resemblance was much weaker in concordant DZ pairs. When they get depressed, individuals who share at birth their entire genome and who were raised together in the same home modestly resemble each other in the specific symptoms they report. Consistent with our within-individual results, these findings suggest that the majority of variation in symptom presentation in MD results from environmental experiences unique to individuals and not to shared genetic and family-environmental background.

Third, to gain further insight into other possible etiologic influences that impact symptoms displayed by those affected with MD, we predicted symptom endorsement from four diverse risk factors, two of which were “genetic/temperamental” (genetic risk for MD and neuroticism) and two of which reflect childhood adversities (childhood sexual abuse and parental coldness). Of our 16 depressive symptoms, three were predicted by neither class of risk factors (appetite/weight, sleep change, and difficulty concentrating), seven by both classes of risk factors (appetite and weight gain, hypersomnia, psychomotor change, psychomotor retardation, feeling worthless, and thoughts of death), three only by genetic/temperamental factors (appetite loss, psychomotor agitation, and feeling tired), and three only by childhood environmental risk (anhedonia, weight loss, and trouble sleeping). As expected, the seven criteria significantly predicted by both classes of risk factors were more stable over time 0.39 ± 0.05 than those which were not 0.29 ± 0.03 . These results suggest that multiple widely diverse factors influence the probability of endorsement of specific symptoms in those suffering from MD. These pathoplastic factors reflect biological/genetic processes, personality, traumatic early environmental experiences, and the quality of parent-child relationships.

Our results for childhood sexual abuse—whose impact on risk for MD is likely to be largely causal (Kendler et al., 2000; Nelson et al., 2002)—are congruent with findings from a large clinically ascertained sample of depressed women in China (Li et al., 2014). In that sample, childhood sexual abuse was commonest among women with atypical depression that was characterized by many of the same symptoms predicted by sexual abuse in our twins: hypersomnia, weight gain, psychomotor retardation, suicidal ideation, and feelings of worthlessness.

Several of the symptoms—including thoughts of death, worthlessness and hypersomnia—were moderately correlated over time and in MZ pairs and thus might be considered useful indices of core depressive pathophysiology. However, these symptoms were also among those most strongly predicted by childhood adversities, suggesting a more complex etiology.

4.1 | Limitations

Our results should be viewed in the context of six potentially important methodological limitations. First, sample sizes of concordant twin pairs were insufficient for more sophisticated modeling which could examine the role of genetic versus environmental influences and the relationship between liability to MD and liability to specific symptoms

given MD. Indeed, correlations in DZ twins were on average lower than expected for a genetically influenced trait (where they should equal at least half that found for MZ twins). However, given the large confidence intervals seen in these correlations, these results could easily have arisen from chance effects. Second, our sample of concordant DZ twins contained both same- and opposite-sex pairs which we combined to maximize power. An examination of their results separately did not suggest meaningful differences.

Third, we had no good way to correct for the unreliability of assessment of depressive symptoms and to therefore quantify the contribution of measurement error to the observed correlations observed over time and between twin pairs. The accuracy of recall for lifetime psychiatric disorders in general and depression in particular has long been questioned (Takayanagi et al., 2014; Wells & Horwood, 2004). However, our recall interval was limited to 12 months likely reducing the impact of memory problems. Our interview emphasized the dating of specific time periods in the last year for assessment both of environmental adversities and depressive episodes, a method shown to improve recall accuracy (Belli, 1998). The interrater reliability of our assessment of depressive symptoms was very high (Kendler, Neale, Kessler, Heath, & Eaves, 1992). We have previously examined twin resemblance for symptoms of *lifetime* MD in the FF twins from our sample (Kendler et al., 1992). They were considerably lower than observed for last year MD (MZ pairs $+0.17 \pm 0.04$ and DZ pairs $+0.01 \pm 0.06$) suggesting that our shorter recall period indeed improved recall accuracy (Kendler et al., 1992).

Fourth, in both within-individual and twin analyses, the time elapsed between the reported MD episodes varied widely. To determine if these differences influenced our results, we examined whether predicting later from earlier symptoms within individuals or across twins interacted with the months separating the two reports. Of the 32 comparisons (16 symptoms within and across twins), only one was significant, consistent with chance expectations.

Fifth, some of our twins might have been in a depressive episode when they completed their neuroticism questionnaire. Given prior evidence from this sample that neuroticism scores increase during such episodes (Kendler, Neale, Kessler, Heath, & Eaves, 1993), we re-ran the relevant analyses eliminating these twins, around 11% of the sample. The results changed little and none of the 17 estimates of the reduced sample were outside the 95% CIs of the results reported in Table 1. We conclude that inclusion of these cases produced minimal biases in our findings.

Sixth, the presence or absence of co-occurring disorders might influence the degree to which individual depressive symptoms are reported across different depressive episodes. Our analyses did not take this into account largely because our information on the occurrence of most potential comorbid disorders did not have sufficient temporal resolution to make this possible.

5 | CONCLUSION

Historically, an important task in psychiatry has been to find etiologically distinct psychiatric disorders, often by subdividing existing broad

syndromes on the basis of distinctive symptoms. This effort would be substantially facilitated if Sydenham's view about the relationship between diseases and symptoms were true—that symptoms are highly stable and strong indices of underlying pathophysiology. Our results do not support this optimistic vision. Rather, we find evidence that the individual symptoms reported by individuals suffering from MD are themselves multifactorial in etiology. They are not highly stable over time suggesting, consistent with prior work (Keller et al., 2007), that they are subject to important transient influences. While clearly influenced by familial factors, resemblance for most symptoms is only modest among concordant MZ twins suggesting a major influence on them from unique environmental experiences (that is, unshared with their twin). We also see that endorsement of many of these symptoms is influenced by personality and childhood environmental adversities.

Our conclusion is that symptoms assessed at structured interviews—at least for MD—are likely quite imperfect indices of the pathophysiological processes underlying the depressive syndrome. These results are congruent with the relatively low yield of efforts over the last century to use symptoms to develop useful subtypes of psychiatric illness that would guide us to etiology. These results are also compatible with the decision to eliminate from DSM-5 one of the longest standing efforts at clinical subtyping—the subtypes of schizophrenia—because they lacked stability, validity, and clinical utility (Braff, Ryan, Rissling, & Carpenter, 2013). Finally, they are consistent with the existence of a complex pathway from the brain state that predisposes to depression and the speech act of a patient articulating depressive symptoms such as feelings of worthlessness (Markova & Berrios, 1995). Such a pathway is likely far more susceptible to external influences than that from cystitis to burning on urination or from urticaria to itching. Birnbaum's understanding of the etiology of the symptoms of psychiatric disorders is likely more accurate than Sydenham's. In Birnbaum's terminology, our results suggest that most of the symptoms of MD are strongly influenced by pathoplastic processes.

ACKNOWLEDGMENTS

This project was supported by the grant MH40828. The Virginia Twin Registry is supported by grant UL1RR031990. Dr. Carol Prescott played a central role in the design and implementation of this twin study.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

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.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

- Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6, 383–406.
- Birnbaum, K. (1974). The making of a psychosis: The principles of structural analysis in psychiatry. In S. R. Hirsch & M. Shepherd (Eds.), *Themes and variations in European psychiatry: An anthology* (pp. 199–238). Charlottesville: University Press of Virginia.
- Bleuler, M. (1978). *The Schizophrenic disorders: Long-term patient and family studies*. New Haven: Yale University Press.
- Braff, D. L., Ryan, J., Rissling, A. J., & Carpenter, W. T. (2013). Lack of use in the literature from the last 20 years supports dropping traditional schizophrenia subtypes from DSM-5 and ICD-11. *Schizophrenia Bulletin*, 39, 751–753.
- Cardno, A. G., Jones, L. A., Murphy, K. C., Sanders, R. D., Asherson, P., Owen, M. J., & McGuffin, P. (1999). Dimensions of psychosis in affected sibling pairs. *Schizophrenia Bulletin*, 25, 841–850.
- Cardno, A. G., Sham, P. C., Murray, R. M., & McGuffin, P. (2001). Twin study of symptom dimensions in psychoses. *British Journal of Psychiatry*, 179, 39–45.
- DeLisi, L. E., Goldin, L. R., Maxwell, M. E., Kazuba, D. M., & Gershon, E. S. (1987). Clinical features of illness in siblings with schizophrenia or schizoaffective disorder. *Archives of General Psychiatry*, 44, 891–896.
- Endicott, J., Andreasen, N., & Spitzer, R. L. (1975). *Family history-research diagnostic criteria*. New York: New York State Psychiatric Institute, Biometrics Research Department.
- Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, 6, 21–29.
- Keller, M. C., Neale, M. C., & Kendler, K. S. (2007). Association of different adverse life events with distinct patterns of depressive symptoms. *American Journal of Psychiatry*, 164, 1521–1529.
- Kendler, K. S., Bulik, C. M., Silberg, J., Hettema, J. M., Myers, J., & Prescott, C. A. (2000). Childhood sexual abuse and adult psychiatric and substance use disorders in women: An epidemiological and cotwin control analysis. *Archives of General Psychiatry*, 57, 953–959.
- Kendler, K. S., & Engstrom, E. J. (2016). Kahlbaum, Hecker, and Kraepelin and the transition from psychiatric symptom complexes to empirical disease forms. *American Journal of Psychiatry*.
- Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry*, 159, 1133–1145.
- Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2006). Toward a comprehensive developmental model for major depression in men. *American Journal of Psychiatry*, 163, 115–124.
- Kendler, K. S., Karkowski-Shuman, L., O'Neill, F. A., Straub, R. E., MacLean, C. J., & Walsh, D. (1997). Resemblance of psychotic symptoms and syndromes in affected sibling pairs from the Irish Study of high-density schizophrenia families: Evidence for possible etiologic heterogeneity. *American Journal of Psychiatry*, 154, 191–198.
- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Familial influences on the clinical characteristics of major depression: A twin study. *Acta Psychiatrica Scandinavica*, 86, 371–378.
- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A longitudinal twin study of personality and major depression in women. *Archives of General Psychiatry*, 50, 853–862.
- Kendler, K. S., & Prescott, C. A. (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry*, 56, 39–44.
- Kendler, K. S., & Prescott, C. A. (2006). *Genes, environment, and psychopathology: Understanding the causes of psychiatric and substance use disorders* (1st ed.). New York: Guilford Press.

- Li, Y., Aggen, S., Shi, S., Gao, J., Li, Y., Tao, M., ... Kendler, K. S. (2014). Subtypes of major depression: Latent class analysis in depressed Han Chinese women. *Psychological Medicine*, *44*, 327–388.
- Markova, I. S., & Berrios, G. E. (1995). Mental symptoms: Are they similar phenomena? The problem of symptom heterogeneity. *Psychopathology*, *28*, 147–157.
- Martin, J., Anderson, J., Romans, S., Mullen, P., & O'Shea, M. (1993). Asking about child sexual abuse: Methodological implications of a two stage survey. *Child Abuse and Neglect*, *17*, 383–392.
- Nelson, E. C., Heath, A. C., Madden, P. A., Cooper, M. L., Dinwiddie, S. H., Bucholz, K. K., ... Martin, N. G. (2002). Association between self-reported childhood sexual abuse and adverse psychosocial outcomes: Results from a twin study. *Archives of General Psychiatry*, *59*, 139–145.
- Parker, G., Tupling, H., & Brown, L. (1979). A parental bonding instrument. *British Journal of Medical Psychology*, *52*, 1–10.
- Pearson, K. (1901). Mathematical contributions to the theory of evolution. VII. On the correlation of characters not quantitatively measurable. *Proceedings of the Royal Society*, *66*, 241–244.
- Rietkerk, T., Boks, M. P., Sommer, I. E., Liddle, P. F., Ophoff, R. A., & Kahn, R. S. (2008). The genetics of symptom dimensions of schizophrenia: Review and meta-analysis. *Schizophrenia Research*, *102*, 197–205.
- SAS Institute, I. (2012). *SAS/STAT*® online documentation, version 9.4. Cary, N.C.: SAS Institute, Inc.
- Spitzer, R. L., & Williams, J. B. W. (1985). *Structured clinical interview for DSM-III-R (SCID)*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Sydenham, T. (1981). Preface to the third edition, *Observations Medicae* (R.G. Latham, Trans.) (1676/1981). In A. L. Caplan, H. T. Engelhardt Jr., & J. J. McCartney (Eds.), *Concepts of health and disease* (pp. 145–155). Reading, MA: Addison-Wesley Publishing Company.
- Takayanagi, Y., Spira, A. P., Roth, K. B., Gallo, J. J., Eaton, W. W., & Mojtabai, R. (2014). Accuracy of reports of lifetime mental and physical disorders: Results from the Baltimore epidemiological catchment area study. *JAMA Psychiatry*, *71*, 273–280.
- Wells, J. E., & Horwood, L. J. (2004). How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychological Medicine*, *34*, 1001–1011.

How to cite this article: Kendler KS, and Aggen SH. Symptoms of major depression: Their stability, familiarity and prediction by genetic, temperamental and childhood environmental risk factors. *Depression and Anxiety*. 2017;34:171–177. DOI: 10.1002/da.22591