Gene–environment interplay in depressive symptoms: moderation by age, sex, and physical illness

A. J. Petkus[1], C. R. Beam[2], W. Johnson[3], J. Kaprio[4,5], T. Korhonen[4,5,6], M. McGue[7,8], J. M. Neiderhiser[9], N. L. Pedersen[10,11], C. A. Reynolds[12] and M. Gatz[10,11]*, for the IGEMS Consortium†

[1] *Department of Neurology, University of Southern California, Los Angeles, CA, USA;* [2] *Department of Psychology & Davis School of Gerontology, University of Southern California, Los Angeles, CA USA;* [3] *Department of Psychology, Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK;* [4] *Department of Public Health, University of Helsinki, Helsinki, Finland;* [5] *Institute for Molecular Medicine (FIMM), University of Helsinki, Helsinki, Finland;* [6] *Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland;* [7] *Department of Psychology, University of Minnesota, Minneapolis, MN, USA;* [8] *The Danish Twin Registry, University of Southern Denmark, Institute of Public Health, Epidemiology, Odense C, Denmark;* [9] *Department of Psychology, Penn State University, University Park, PA, USA;* [10] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden;* [11] *Department of Psychology, University of Southern California, Los Angeles, CA, USA;* [12] *Department of Psychology, University of California Riverside, Riverside, CA, USA*

Background. Numerous factors influence late-life depressive symptoms in adults, many not thoroughly characterized. We addressed whether genetic and environmental influences on depressive symptoms differed by age, sex, and physical illness.

Method. The analysis sample included 24 436 twins aged 40–90 years drawn from the Interplay of Genes and Environment across Multiple Studies (IGEMS) Consortium. Biometric analyses tested age, sex, and physical illness moderation of genetic and environmental variance in depressive symptoms.

Results. Women reported greater depressive symptoms than men. After age 60, there was an accelerating increase in depressive symptom scores with age, but this did not appreciably affect genetic and environmental variances. Overlap in genetic influences between physical illness and depressive symptoms was greater in men than in women. Additionally, in men extent of overlap was greater with worse physical illness (the genetic correlation ranged from near 0.00 for the least physical illness to nearly 0.60 with physical illness 2 S.D. above the mean). For men and women, the same environmental factors that influenced depressive symptoms also influenced physical illness.

Conclusions. Findings suggested that genetic factors play a larger part in the association between depressive symptoms and physical illness for men than for women. For both sexes, across all ages, physical illness may similarly trigger social and health limitations that contribute to depressive symptoms.

Introduction

Depressive symptoms are common throughout adulthood. These symptoms often impair daily function, especially in older adults, even when their total number does not meet criteria for clinical diagnosis (Judd *et al.* 2002). Both cross-sectional and longitudinal reports have observed U-shaped relations between depressive symptom scores and age (Kessler *et al.* 1992; Sutin *et al.* 2013). Additionally, at most ages, women report more depressive symptoms than men, although the difference may diminish in very old age (e.g. Berkman *et al.* 1986; Forlani *et al.* 2014).

In older twins, genetic influences generally account for up to 30% of the variance in symptom scores (Johnson *et al.* 2002; Schnittker, 2014). Possible age and sex differences in heritability of depressive symptoms have been considered using twin designs. Most find similar heritability estimates of depressive symptoms for men and women (Gillespie *et al.* 2004). Some studies have observed greater heritability of depressive symptoms with higher age (e.g. Carmelli *et al.* 2000), although others did not find significant age differences in heritability (McGue & Christensen, 1997; Johnson *et al.* 2002). Because heritability estimates represent the relative proportions of genetic and

* Address for correspondence: M. Gatz, Ph.D., Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA.

(Email: gatz@usc.edu)

† Members of the IGEMS Consortium are given in the Appendix.

environmental variance components, greater heritability may be a function of reduced environmental influences, increased genetic influences, or both.

Physical illness is a well-established risk factor for depressive symptoms, particularly in later life (e.g. Meeks *et al.* 2000; Djernes, 2006), with these observations not simply due to overlap with somatic symptoms used in diagnosing depression (Sutin *et al.* 2013). Depressive symptoms and depressive disorders are elevated among medical inpatients and primary-care outpatients (Blazer, 2003). The association between depression and physical illness may lie in vascular, neuroendocrine, or inflammatory manifestations of physical illness that also contribute to risk for explicit depressive symptoms (Steptoe, 2007; Fiske *et al.* 2009). There may be common factors such as low education or high neuroticism that predispose to both physical illness and depression (Steptoe, 2007). Physical illness may occasion emotional distress and existential struggles that lead to elevated depressive symptoms (Steptoe, 2007). Moreover, physical illness may impose functional limitations and disruptions to daily living that potentiate depression, representing a behavioral pathway (Fiske *et al.* 2009).

Twin studies can be valuable in illuminating the extents to which genetic and environmental sources of influence explain observed covariation between depressive symptoms and physical illness. Vulnerability factors in common to physical illness and depressive symptoms, as well as proposed physiological pathways linking the two, suggest the possibility that genetic factors account for some of the association. Behavioral explanations support roles for environmental factors. Physical illness may also moderate genetic and environmental vulnerabilities to depressive symptoms, either uniquely or via shared pathways (Johnson, 2007).

In this paper, we explored the cross-sectional interplay among depressive symptoms, physical illness, sex, and age. Prior twin studies of depressive symptoms have not considered the role of physical illness, and research pointing to co-occurring physical illness and depressive symptoms in older adults has not addressed how genetic and environmental influences interplay. We thus examined whether genetic and environmental variance in depressive symptoms varies with sex, age, and different levels of physical illness.

Method

Participants

Summarized in Table 1, the sample was drawn from ten independent studies from the Interplay of Genes and Environment across Multiple Studies (IGEMS) Consortium (Pedersen *et al.* 2013). A total of 24 436

individual twins (53.1% women) had relevant data. Depressive symptom data were taken from the first occasion at which each participant answered self-report depressive symptom questionnaires. Incomplete twin pairs were included, as they are informative about age differences in means and variances. Demographic characteristics in Table 1 include numbers of monozygotic (MZ), same-sex dizygotic (SSDZ), and opposite-sex dizygotic (OSDZ) pairs. Demographics by study are given in Supplementary Table S1.

Twins in the Swedish Adoption Twin Study of Aging (SATSA; Finkel & Pedersen, 2004), Origins of Variance in the Oldest-Old (OCTO-Twin; McClearn *et al.* 1997), Ageing in Women and Men: A Longitudinal Study of Gender Differences in Health Behaviour and Health among Elderly (GENDER; Gold *et al.* 2002), and Twin and Offspring Study in Sweden (TOSS; Neiderhiser & Lichtenstein, 2008) were ascertained from the population-based Swedish Twin Registry. SATSA includes all same-sex twins from the Swedish Twin Registry who were reared apart and a matched sample who were reared together, with baseline depressive symptoms assessed at the first questionnaire follow-up in 1987. OCTO-Twin includes same-sex pairs who both survived until age 80, with baseline in 1991. GENDER consists of OSDZ twin pairs born 1906–1925, with baseline in 1994. Baseline for TOSS was 1997 or 2004.

Longitudinal Study of Aging Danish Twins (LSADT; Christensen *et al.* 1999) and Middle-Age Danish Twins (MADT; Skytthe *et al.* 2013) were ascertained from the Danish Twin Registry. LSADT includes same-sex twin pairs aged ⩾70 years, with baseline in 1995. MADT includes both same- and opposite-sex twin pairs born 1931–1952, with baseline in 1998.

The Minnesota Twin Study of Adult Development and Aging (MTSADA; Finkel & McGue, 1993) includes same-sex twin pairs aged ⩾60 years ascertained through Minnesota state birth records, with baseline data collected 1984–1994. The Vietnam Era Twin Study of Aging (VETSA; Kremen *et al.* 2006) includes male twins aged 51–60 at baseline in 2003–2007. The twin subsample of Midlife in the United States (MIDUS; Kendler *et al.* 2000) includes same- and opposite-sex twin pairs, with depressive symptoms collected at follow-up in 2007.

The older Finnish Twin Cohort (FTC; Kaprio & Koskenvuo, 2002) is a registry of same-sex Finnish twin pairs with age range comparable to the Swedish and Danish twin registries. Depressive symptom data reported here come from wave 4 in 2011 (Kaprio, 2013).

Previous publications from SATSA (Gatz *et al.* 1992), LSADT (McGue & Christensen, 1997), MADT (Johnson *et al.* 2002), VETSA (Franz *et al.* 2011), and FTC

Table 1. *Mean scores on depressive symptoms and physical illness (with standard deviations), phenotypic correlations between depressive symptoms and physical illness, intra-pair correlations, and cross-twin cross-trait correlations (with 95% confidence intervals)*

	N pairs	Depressive symptoms mean (s.d.)	I-CIRS mean (s.d.)	Phenotypic Correlation (95% CI)	Depressive symptoms intra-pair correlation (95% CI)	I-CIRS Intra-pair correlation (95% CI)	Cross-twin cross-trait correlation (95% CI)
Men	3605	19.43 (4.17)	1.59 (1.66)	0.20 (0.18–0.21)	0.25 (0.22–0.28)	0.27 (0.24–0.30)	0.09 (0.06–0.12)
Women	5331	20.42 (4.80)	1.97 (1.80)	0.17 (0.15–0.18)	0.24 (0.21–0.26)	0.29 (0.26–0.31)	0.08 (0.05–0.11)
MZM	1621	19.45 (4.34)	1.55 (1.61)	0.19 (0.15–0.22)	0.33 (0.28–0.37)	0.35 (0.31–0.39)	0.08 (0.03–0.13)
MZW	1762	20.30 (4.77)	1.91 (1.80)	0.16 (0.13–0.19)	0.36 (0.32–0.40)	0.38 (0.34–0.42)	0.11 (0.06–0.15)
DZM	1984	19.60 (4.15)	1.63 (1.71)	0.19 (0.17–0.22)	0.18 (0.14–0.22)	0.20 (0.16–0.24)	0.10 (0.05–0.14)
DZW	2416	20.69 (4.90)	2.05 (1.84)	0.15 (0.13–0.18)	0.16 (0.12–0.20)	0.24 (0.20–0.27)	0.04 (0.01–0.08)
DZOS	1153	19.11 (3.98)	1.66 (1.58)	0.25 (0.22–0.29)	0.15 (0.09–0.21)	0.24 (0.19–0.29)	0.12 (0.06–0.17)
Total	8936	19.95 (4.54)	1.79 (1.75)	0.19 (0.18–0.20)	0.25 (0.23–0.27)	0.29 (0.27–0.31)	0.09 (0.07–0.11)

I-CIRS, IGEMS revision of Cumulative Illness Rating Scale; MZM, monozygotic male pairs; MZW, monozygotic female pairs; DZM, same-sex dizygotic male pairs; DZW, same-sex dizygotic female pairs; DZOS, opposite-sex dizygotic pairs.

(Takkinen *et al.* 2004; Korhonen *et al.* 2007) have presented results from behavior genetic analyses of depressive symptoms. The benefits of the IGEMS Consortium are a large sample of twins offering increased power to detect potentially small moderation effects of physical illness, age, and sex and inclusion of a large number of opposite-sex twin pairs to explore potential sex differences in genetic and environmental influences on depressive symptoms.

Zygosity

Specific methods for zygosity determination varied across studies, but in most cases involved use of standard questionnaires about physical similarity between twins and DNA analysis to resolve uncertain cases.

Measures

Depressive symptoms

Two measures of depressive symptoms were used in IGEMS studies: the Center for Epidemiologic Studies – Depression (CES-D) scale (Radloff, 1977) and a modified version of the Cambridge Mental Disorders of the Elderly Examination (CAMDEX; Roth *et al.* 1986). The CES-D score is the sum of 20 items, each scored from 0 to 3 to indicate frequency of experiencing the symptom during the past week. Four positively worded items were reverse-scored so that higher scores indicated more severe depressive symptoms. Previous research supports a four-factor structure representing depressed mood, psychomotor retardation/somatic symptoms, lack of well-being (the four reverse-scored items), and interpersonal difficulties (Radloff, 1977; Gatz *et al.* 1992).

The CAMDEX was administered only to Danish twins and consists of 16 items assessing frequency of depressive symptoms. Items were scored from 1 to 3 to indicate frequency of experiencing the symptom, except for two items rated on a 'yes/no' scale. Items were scored so that higher scores indicate more severe depressive symptoms. The CAMDEX consists of two factors: affective and somatic symptoms (McGue & Christensen, 1997).

To create a common depressive symptom metric, CES-D and CAMDEX scales were administered to a separate data-harmonization sample. Item-response theory methods were applied to compare items from the two measures and create a conversion table between the scales (Gatz *et al.* 2015b). For substantive comparability, we retained only the 14 CES-D items that loaded onto the depressed mood and psychomotor retardation/somatic symptoms subscales. The co-calibrated score was expressed in CAMDEX units, such that the total score could range from 16 for someone who endorsed no symptoms of depression to a maximum of 46 for someone who endorsed the most severe frequency of depressive symptoms. Because the harmonized measure did not include the same items for those who completed different scales, we calculated separate reliabilities for each scale's harmonization-relevant items. For those who completed the CES-D, $\alpha = 0.90$ ($N = 16\,177$), while for those who completed the CAMDEX, $\alpha = 0.84$ ($N = 8718$).

Physical illness

To create a harmonized index of physical illness (I-CIRS; Gatz *et al.* 2015a), we created a revised version of the Modified Cumulative Illness Rating Scale (CIRS;

Salvi *et al.* 2008). I-CIRS is a list of 13 body systems, with a score that represents the number of systems with moderate disability or morbidity. All studies had self-reported lists of physical illnesses that we recoded to be consistent across IGEMS studies. I-CIRS categories included cardiac; hypertension; vascular (circulatory); respiratory; eye, ear, nose and throat; upper gastrointestinal; hepatic; renal; musculoskeletal; neurological (excluding stroke which was rated in a separate category); endocrine/metabolic; cancer; and stroke. Each participant received a score from 0 to 13, with higher scores indicating more systems affected by illness.

Statistical analysis

Depressive symptoms and I-CIRS were positively skewed, so scores were log-transformed to approximate assumptions of multivariate normality. Depressive symptoms and I-CIRS further were standardized (mean = 0, S.D. = 10) to ease computational burden. Twin pairs younger than 40 (240 complete pairs) or older than 90 (10 complete pairs) were removed from analyses due to sparse coverage.

We examined the phenotypic means of depressive symptoms by sex and linear, quadratic, and cubic continuous age using SAS Proc Mixed (SAS Institute Inc., 2013) to account for dependencies between members of twin pairs. Intra-pair correlations for depressive symptoms and cross-twin cross-trait correlations for depressive symptoms and physical illness were calculated by sex, age (by decade), and zygosity (see Supplementary Table S2). To evaluate cohort effects, we formed a sample of 50- to 59-year-olds tested between 1987 and 1991 (comprised of SATSA and MTSADA) and a sample of 50- to 59-year-olds tested between 2003 and 2007 (comprised of TOSS twins from 2004, MIDUS twins, and a random sample from VETSA to make the two samples similar in size).

Twin studies allow variance in a measure to be decomposed into additive genetic variance (A), shared environmental variance (C), dominant genetic variance (D), and non-shared environmental variance including error (E). Shared environment refers to environmental influences that contribute to twin pair similarity, such as rearing environment or contact as adults; non-shared environment encompasses influences contributing to differences within a pair. Conventional twin methods possess too few degrees of freedom to estimate all ACDE components, but preliminary models suggested no evidence of dominant genetic influences. Variance decomposition relies on assumptions that SSDZ and OSDZ twins share 50% of their segregating genes, on average, while MZ twins share 100% of their genes; and that genetic and environmental biometric

components are independent. Based on these assumptions, twin ('ACE') models specify three sets of equations based on degree of genetic relationship:

$$\text{MZ} = \text{cov MZ} = V_a + V_c;$$
$$\text{SSDZ} = \text{cov SSDZ} = 0.5 \times V_a + V_c;$$
$$\text{OSDZ} = \text{cov OSDZ} = 0.5 \times V_a + V_c.$$

All multivariate twin models were estimated in OpenMx 2.2.6 (Neale *et al.* 2016).

Qualitative sex differences (i.e. different sources of influences) were tested in univariate biometric models by freely estimating the genetic correlation in OSDZ pairs instead of constraining it to be 0.50 as well as freely estimating the shared environmental correlation instead of constraining it to be 1.00. An estimated genetic correlation between OSDZ twins significantly <0.50 suggests qualitative sex differences. Due to documented problems estimating quantitative sex-limitation in bivariate models (Neale *et al.* 2006), bivariate moderation models were estimated separately for males and females using only SSDZ twins. To test significance of sex differences we constrained parameter estimates for women to equal those for men. We compared this constrained model to the model allowing women's parameter estimates to be freely estimated.

We fitted age-moderated univariate biometric models (van der Sluis *et al.* 2012) to address whether and how continuous age moderated genetic and environmental variance components underlying depressive symptoms. Preliminary results suggested nonlinear age differences in depressive symptom scores. These were best characterized as two slopes with a turning point at age 75 (see Supplementary Fig. S1). Therefore, we estimated two linear age effects centered at age 75 – with age recoded into two variables: aged 40–75 (i.e. AgeA = age 75; if age ⩾ 75 then AgeA = 0) and aged 75–90 (AgeB = age 75; if age < 75 then AgeB = 0).

Under the model, additive genetic, shared environmental, and non-shared environmental variances are expressed as:

$$A = a + (\beta_{a40} \times \text{AgeA}) + (\beta_{a75} \times \text{AgeB}),$$
$$C = c + (\beta_{c40} \times \text{AgeA}) + (\beta_{c75} \times \text{AgeB}),$$
$$E = e + (\beta_{e40} \times \text{AgeA}) + (\beta_{e75} \times \text{AgeB}).$$

The 'a' parameter is the estimated additive genetic variance at the age-75 turning point, β_{a40} is the estimate of the linear slope on additive genetic variance for participants aged 40–75, and the β_{a75} is the second linear slope for participants aged 75–90. The C and E variance components are computed in the same way. Likelihood-ratio tests were used to compare nested models.
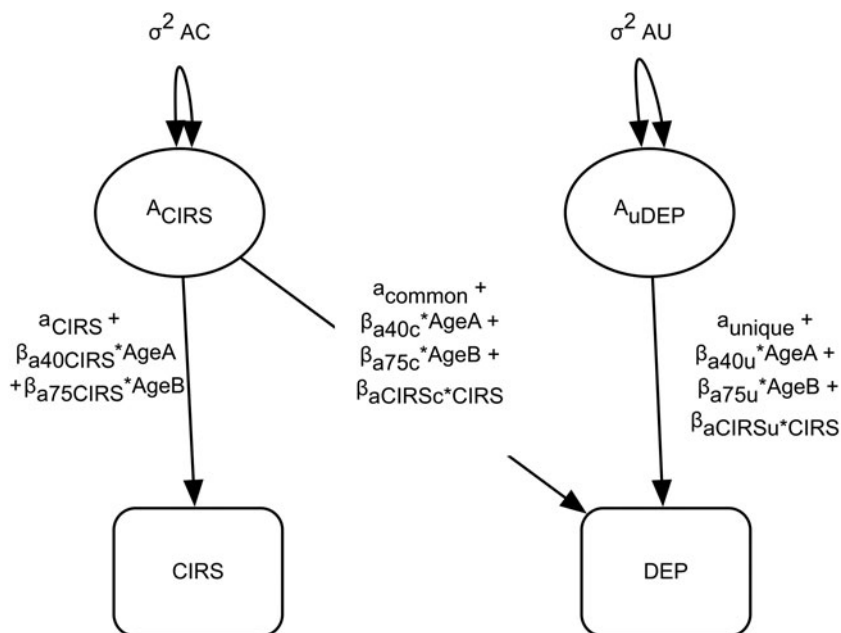
Fig. 1. Diagram of the additive genetic components from the age and I-CIRS moderated bivariate biometric model examining physical health (CIRS) and depressive symptoms (DEP). Note that slope a40 pertains to individuals age 40–75 while slope a75 pertains to individuals aged $\geqslant$75. The equation for calculating the genetic covariance between I-CIRS and depressive symptoms is the product of the pathway shared between CIRS and depressive symptoms and the genetic pathway on CIRS. The equation can be written as: $(a_{common} + \beta_{a40c} \times AgeA + \beta_{as75c} \times AgeB + \beta_{aCIRSc} \times CIRS) \times (a_{CIRS} + \beta_{a40CIRS} \times AgeA + \beta_{a40CIRS} \times AgeB)$.

We fitted continuous age and I-CIRS-moderated bivariate biometric models to test whether I-CIRS and age moderated genetic and environmental variance between I-CIRS and depressive symptoms (Johnson, 2007). Fig. 1 depicts the additive genetic variance components of this model, which were expressed analogously for the shared and non-shared environmental components (not presented). The model estimated the genetic variance of I-CIRS (a_{CIRS}), the genetic covariance between I-CIRS and depressive symptoms (a_{common}), and a genetic variance component unique to depressive symptoms (a_{unique}). Age-moderating terms for 40–75 and 75–90 years age groups on depressive symptoms and I-CIRS were included in the I-CIRS, common, and unique depressive symptoms paths. Genetic (r_G) and environmental (r_C and r_E) correlations between I-CIRS and depressive symptoms were calculated to aid in interpreting the covariance. The equation for genetic correlation, for example, is:

$$rG = \frac{(a_{common} + (\beta_{a40c} \times AgeA) + (\beta_{a75c} \times AgeB) + (\beta_{aCIRSc} \times CIRS))}{\sqrt{\begin{array}{l}((a_c + (\beta_{a40c} \times AgeA) + (\beta_{a75c} \times AgeB) \\ + (\beta_{aCIRSc} \times CIRS))^2 + (a_{unique} + (\beta_{a40u} \times AgeA) \\ + (\beta_{a75u} \times AgeB) + (\beta_{aCIRSu} \times CIRS))^2)\end{array}}}$$

A possible alternative explanation to the variance moderation model is the uniform nonlinear main effects model. To address this possible explanation we ran a nonlinear main effects model of I-CIRS on depressive symptoms (Van Hulle *et al.* 2013) and compared this model to the variance moderation model.

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.

Results

Descriptive statistics

Harmonized depressive symptom scores and I-CIRS scores by age group and sex, pooled across studies are shown in Table 1, and by age group in Supplementary Table S2. Depressive symptom scores were positively correlated with physical illness. Results from mixed-effects regression models indicated significant age and sex associations with depressive symptoms. Women scored significantly higher than men, and both sexes showed accelerating (cubic) score increases with age

Table 2. *Model fitting results for I-CIRS and age-moderated bivariate biometric models of depressive symptoms*

Model	Men				Women			
	−2LL	df	Comp	Δ−2LL (Δdf)	−2LL	df	Comp	Δ−2LL (Δdf)
1. Full ACE	108 024	14 850	–	–	126 727	17 240	–	–
2. AE	108 032	14 861	1	8.14 (11)	126 742	17 251	1	15.24 (11)
3. Drop all CIRS moderation	108 317	14 865	2	285.03 (4)**	126 904	17 255	2	161.91 (4)**
4. Drop age 40–75 covariance moderation	108 034	14 863	2	2.29 (2)	126 744	17 253	2	2.23 (2)
5. Drop age 75–90 covariance moderation	**108 036**	**14 865**	**4**	**1.19 (2)**	**126 744**	**17 255**	**4**	**0.02 (2)**
6. Drop age 40–75 unique depression moderation	108 072	14 867	5	35.86 (2)**	126 759	17 257	5	14.11 (2)**
7. Drop age 75–90 unique depression moderation	108 056	14 867	5	20.17 (2)**	126 754	17 257	5	9.02 (2)**

More complete model-fitting results are given in Supplementary Tables S7 and S8.

I-CIRS, IGEMS revision of Cumulative Illness Rating Scale; −2LL, negative log likelihood multiplied by 2; df, degrees of freedom; Comp, model to which compared, Δ−2LL (Δdf), the difference in log likelihood by difference in degrees of freedom between the model being fit and its comparison model; A, additive genetic variance; C, shared environmental variance; E, non-shared environmental variance; ACE, full model estimating additive genetic, shared environmental, and nonshared environmental variance components; AE, constrained model estimating only additive genetic and nonshared environmental variance components.

** Significant deterioration in fit compared to comparison model. Best-fitting model is shown in bold.

(Supplementary Fig. S1). The variability of depressive symptoms was greater in women than in men and was larger at older ages than at younger ages for both men and women (Supplementary Table S2). There was no significant interaction between age and sex. I-CIRS scores increased quadratically with age in both sexes. The phenotypic association between depressive symptoms and physical illness was linear, and correlations showed no systematic differences with age.

There were differences in mean depressive symptom scores across participating studies (Supplementary Table S3). However, differences were not systematic by country or by original measure of depressive symptoms, and adding study to the biometric models did not affect findings.

Next, we evaluated whether there were cohort differences in depressive symptoms (Supplementary Table S4). We compared 50- to 59-year-olds tested between 1987 and 1991 with 50- to 59-year-olds tested between 2003 and 2007 and found no significant differences between the earlier cohort (mean = 20.09, s.d. = 4.76, N = 402) and the later cohort (mean = 20.05, s.d. = 4.56, N = 400). There was no evidence of interaction between sex and cohort.

Twin correlations (shown in Table 1) suggested genetic and non-shared environmental influences on both depressive symptoms and I-CIRS in male and female twins. For most ages there was little evidence of shared environmental influences; however, there was some evidence for them (i.e. DZ correlation greater than half the MZ correlation) on depressive symptoms at

ages 80–90 (Supplementary Table S2). The small but significant MZ and DZ cross-twin cross-trait correlations were similar for men but not for women. Similar correlations indicate overlapping shared environmental influences, while larger correlations for MZ than for DZ twins suggest that additive genetic effects likely mediate the association between the two traits.

Biometric model-fitting results

Age-moderated univariate biometric models with qualitative sex-limitation suggested no qualitative sex differences in genetic influences or shared environmental influences on depressive symptoms [$r_G = 0.50$, 95% confidence interval (CI) 0.39–0.50; $r_C = 1.00$, 95% CI 0.91–1.00]. This result was expected given the similar correlations for OSDZ and SSDZ pairs. Consistent with greater variance in women, magnitudes of genetic and environmental influences could not be constrained equal without significant deterioration of model fit (see Supplementary Table S5, with parameter estimates for the full univariate model in Supplementary Table S6). Therefore, separate age-moderated bivariate biometric models were calculated for men and for women. Bivariate model-fitting results are shown in Table 2 (with additional details in Supplementary Tables S8 and S9 for males and females, respectively). For both sexes, shared environmental influences were very small, not significantly different from 0.00, and could be dropped without significant reduction in model fit (see Table 2, model 2). Table 3 presents the

Table 3. *Unstandardized parameter estimates (95% confidence intervals) from the best fitting bivariate biometric model from Table 2 (model number 5) of depressive symptoms*

	Men			Women		
		95% CI			95% CI	
Parameter	Estimate	Lower	Upper	Estimate	Lower	Upper
Additive genetic						
a_{CIRS}	6.00	5.14	6.84	5.18	4.53	5.81
a_{common}	0.28	0.22	0.32	0.26	0.18	0.32
a_{unique}	3.96	3.12	4.74	5.41	4.63	6.16
β_{aCIRSc}	0.08	0.01	0.13	0.01	−0.04	0.07
β_{a40u}	−0.06	−0.10	0.01	−0.01	−0.06	0.03
β_{a75u}	0.20	−0.08	0.45	0.16	−0.03	0.34
β_{aCIRSu}	0.03	−0.02	0.08	0.02	−0.02	0.07
Non-shared environmental						
e_{CIRS}	7.90	7.35	8.46	7.12	6.69	7.55
e_{common}	0.14	0.10	0.17	0.15	0.12	0.19
e_{unique}	6.80	6.30	7.32	7.59	7.10	8.10
β_{eCIRSc}	0.09	0.05	0.13	0.11	0.07	0.15
β_{e40u}	−0.04	−0.01	−0.07	−0.03	−0.06	−0.01
β_{e75u}	0.16	0.01	0.33	0.03	−0.09	0.16
β_{eCIRSu}	0.10	0.07	0.13	0.06	0.03	0.09

a_{CIRS}, Additive genetic genetic path estimate to the additive genetic I-CIRS factor at age 75; a_{common}, common pathway estimate of the additive genetic contributions of I-CIRS on depressive symptoms; a_{unique}, estimated variance unique to depressive symptoms at age 75; β_{aCIRSc}, estimate of the moderating effect of I-CIRS on the additive genetic covariance. β_{a40u}, estimate of the linear slope on the additive genetic variance unique to depressive symptoms for participants aged 40–75; β_{a75u}, estimate of the linear slope on the additive genetic variance unique to depressive symptoms for participants aged 75–90; β_{aCIRSu}, estimate of the moderating effect of I-CIRS on the additive genetic variance unique to depressive symptoms. The parameters denoted with an e represent corresponding parameter estimates for the non-shared environmental contributions.

unstandardized parameter estimates from the best-fitting model, which was model 5 (parameter estimates for the full bivariate model are given in Supplementary Table S9).

Genetic and environmental influences on depressive symptoms

Results for sex and age moderation of depressive symptoms were similar in the univariate analyses of depressive symptoms and for influences unique to depressive symptoms in the bivariate model. In the full univariate model (Supplementary Table S6), additive genetic influences were greater for women than for men; in the best-fitting bivariate model (Table 3), this pattern held but with the confidence intervals somewhat overlapping.

Fig. 2(a,b) presents total (common plus unique) unstandardized A and E variance of depressive symptoms as a function of age. Although age moderation of influences unique to depression could not be dropped, the effects were quite small. Specifically, for both men and women, non-shared environmental variances were

smaller with greater age until age 75. After age 75, for men only, non-shared environmental contributions accounted for larger amounts of variance at older than at younger ages. Heritability estimates, which represent relative proportions of genetic and environmental variance, were not significantly different for men and women: At age 75, h^2 for men was 28.6% (95% CI 21.7–33.7) and for women 34.7% (95% CI 30.6–37.8).

Gene–environment interplay in association between depressive symptoms and physical illness

Physical illness significantly moderated genetic and environmental contributions to depressive symptoms. There was significant overlap in both genetic and non-shared environmental influences between I-CIRS and depressive symptoms. From the information in Table 3, we calculated the proportion of variance in depressive symptoms that was shared with physical illness. For men, at age 75, with a mean I-CIRS score, 15% (95% CI 12–18) of the genetic variance and 2% (95% CI 1–4) of the non-shared environmental variance in depressive symptoms was shared with physical
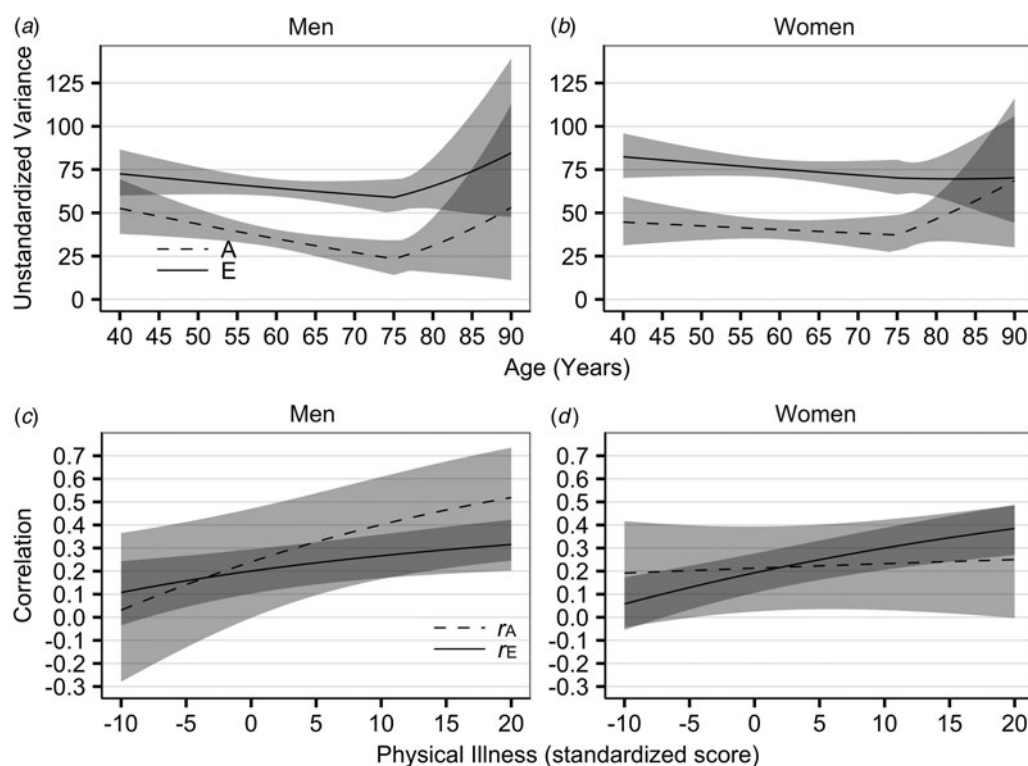
Fig. 2. Graphs of the estimated unstandardized genetic and environmental variance components for depressive symptoms for men (*a*) and women (*b*) by age from combining the common and unique raw variance estimates from the AE age and I-CIRS moderated bivariate biometric model. Graphs of the estimated genetic, and non-shared environmental correlations between I-CIRS and depressive symptoms by I-CIRS score for men (*c*) and women (*d*) shown at age 75 for purposes of illustration. A, Additive genetic variance; E, non-shared environmental variance; r_G, genetic correlation between depressive symptoms and I-CIRS; r_E, non-shared environmental correlation between depressive symptoms and I-CIRS. Shaded area represents 95% confidence interval of the estimate. The I-CIRS score was standardized so the mean CIRS score equals zero and with standard deviation equal ±10.

illness. The analogous figures for women were 6% of the genetic variance (95% CI 3–8) and 2% (95% CI 1–3) of the non-shared environmental variance. These figures reflect both the greater phenotypic correlation in men and greater I-CIRS variance moderation.

Consistent with variance moderation, genetic correlations varied with level of physical illness for men but not women (Fig. 2*c,d*). That is, for men there was greater overlap in additive genetic influences underlying I-CIRS and depressive symptoms at higher I-CIRS scores than at lower I-CIRS scores, regardless of age. The genetic correlation ranged from near 0.00 for the least physical illness to nearly 0.60 with physical illness 2 s.d. above the mean. Also referring to Fig. 2(*c,d*), environmental factors contributing to twin differences in I-CIRS scores accounted for more of the environmental variation in depressive symptoms at higher I-CIRS scores than at lower I-CIRS scores in both sexes. Moreover, referring to β_{eCIRSu} in Table 3, for both men and women, greater physical illness was associated with greater non-shared environmental contributions unique to depressive symptoms.

Nonlinear and linear means moderation did not provide better model fit in either sex, suggesting that non-linear uniform main effects of I-CIRS on depressive symptoms were not explaining the moderation detected in these models.

Discussion

We explored the cross-sectional interplay among depressive symptoms, physical illness, sex, and age in a combined sample of 24 436 twins aged 40–90. We highlight five findings. First, although additive genetic variance in depressive symptoms was greater in women than men, heritability of depressive symptoms was not significantly different for men and women. Finding no sex differences in heritability is consistent with past work with depressive symptoms (e.g. Johnson *et al.* 2002; Gillespie *et al.* 2004) although some research has reported greater heritability of depressive disorders in women than men (e.g. Bierut *et al.* 1999; Kendler *et al.* 2001, 2006). The sex-limitation model indicated that the same genes influenced

depression in women and men, although others have reported that different genes influenced depressive illness in men and women (Kendler & Gardner, 2014).

Second, for men but not for women, non-shared environmental variance in depressive symptoms (i.e. influences on depressive symptoms unique to one individual in a twin pair) was greater at higher ages. Some have suggested that men, in particular, may be increasingly sensitive to stressful life events – including health issues – with age (Sonnenberg *et al.* 2000; Kendler & Gardner, 2014; Forlani *et al.* 2014). This account might explain the patterns of non-shared environmental variances after age 75.

Third, physical illness significantly moderated non-shared environmental influences unique to depressive symptoms. Regardless of sex, greater self-reported physical illness was associated with greater non-shared environmental influences on the unique variance in depressive symptoms.

Fourth, overlap in genetic influences between physical illness and depressive symptoms was greater for men than for women. In men, as well, the genetic correlation between depressive symptoms and physical illness was higher with greater physical illness. These effects did not differ with age. Analyses of genetic correlations derived from summary statistics of common disorders from genome-wide association studies indicate shared genetic effects of major depressive disorder with multiple psychiatric, neurological and cardiovascular traits and risk factors (Anttila *et al.* 2016).

Fifth, non-shared environmental influences mediated the association between physical illness and depressive symptoms, suggesting small but potentially directly causal influences (Heath *et al.* 1993; Turkheimer & Harden, 2014). These environmental influences did not vary with age; however, they were stronger with greater physical illness. With poorer health, environmental influences that make people ill (e.g. poor diet, lack of physical exercise) may also contribute to experiencing more depressive symptoms.

We note several limitations. First, the data were cross-sectional so intra-individual changes in depressive symptoms over time could not be observed. Cohort differences, however, were not statistically significant. Additionally, previous research does not suggest secular diminution of CES-D scores in younger cohorts (Radloff, 1977; Berkman *et al.* 1986; Ehrlich & Isaacowitz, 2002; Twenge, 2015).

Second, reverse causality remains an alternative possibility in this study, including reciprocal processes between depressive symptoms and physical illness (Steptoe, 2007; Scott *et al.* 2016). Here, longitudinal data are preferable to cross-sectional data to disentangle direction of causation (Heath *et al.* 1993) and dual change mechanisms (McArdle & Hamagami, 2003).

Third, twins' histories of depression were unknown. Onset of depressive symptoms in late life might have different etiologies than chronic or recurring depressive symptoms (Fiske *et al.* 2009). Fourth, other potential risk factors for depression such as personality characteristics, smoking, use of alcohol, and stressful life events unrelated to illness were not included in the models. Fifth, results may have been influenced by differences in measurement or recruiting procedures among the IGEMS studies, although we were not able to identify any. Sixth, while the sample was large and diverse in many respects, population-based, and not selected for depressive symptoms or physical illness, ethnic and racial compositions were fairly homogenous.

Overall, results showed that co-occurring physical illness and depressive symptoms in older adults reflects both genetic and environmental mechanisms. Results suggesting shared genes may indicate common pathogenic pathways. Those for non-shared environmental influences suggest that physically ill twins may face experiences (e.g. functional limitations, pain, disruptions to daily living) that lead to decreased physical and social engagement, elevating depressive symptoms compared to their physically healthy co-twins. One implication is that alterations in physically ill older adults' environments, such as better chronic-pain treatment, sleep therapy, assistive devices, and better overall disease management may offset these effects. Interventions directed at the physical symptoms might be paired with behavioral therapies for depression that emphasize learning new skills to adapt to physical limitations and challenges.

Appendix. Consortium on Interplay of Genes and Environment across Multiple Studies (IGEMS)

Members of IGEMS include: Nancy L. Pedersen (Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, and Department of Psychology, University of Southern California, Los Angeles, CA), Kaare Christensen (Department of Epidemiology, University of Southern Denmark, Odense, Denmark), Anna Dahl Aslan (Institute of Gerontology, School of Health Sciences, Jönköping University, Jönköping, Sweden), Deborah Finkel (Department of Psychology, Indiana University Southeast, New Albany, IN), Carol E. Franz (Department of Psychiatry, University of California, San Diego, La Jolla, CA), Margaret Gatz (Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, and Department of Psychology, University of Southern California, Los Angeles, CA), Briana N. Horwitz (Department of Psychology, California

State University, Fullerton, CA), Boo Johansson (Department of Psychology, University of Gothenburg, Gothenburg, Sweden), Wendy Johnson (Department of Psychology and Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK), Jaakko Kaprio, Department of Public Health, University of Helsinki, Helsinki, Finland), William S. Kremen (Department of Psychiatry, University of California, San Diego, and VA Center of Excellence for Stress and Mental Health, La Jolla, CA), Robert Krueger (Department of Psychology, University of Minnesota, Minneapolis, MN), Michael J. Lyons (Department of Psychological and Brain Sciences, Boston University, Boston, MA), Matt McGue (Department of Psychology, University of Minnesota, Minneapolis, MN), Miriam A. Mosing (Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden), (Jenae M. Neiderhiser (Department of Psychology, The Pennsylvania State University, University Park, PA), Matthew S. Panizzon (Department of Psychiatry, University of California, San Diego, La Jolla, CA), Inge Petersen (Department of Epidemiology, University of Southern Denmark, Odense, Denmark), and Chandra A. Reynolds (Department of Psychology, University of California-Riverside, Riverside, CA).

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0033291717000290.

Acknowledgements

Declaration of Interest

None.

References

Anttila V *et al*. (2016). Analysis of shared heritability in common disorders of the brain. *bioRxiv*. doi:10.1101/048991.

Berkman LF, Berkman CS, Kasl S, Freeman Jr. DH, Leo L, Ostfeld AD, Cornoni-Huntley J, Brody J (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology* **124**, 372–388.

Bierut LJ, Heath AC, Bucholz KK, Dinwiddie SH, Madden PA, Statham DJ, Dunne MP, Martin NG (1999). Major depressive disorder in a community-based twin sample: are there different genetic and environmental contributions for men and women? *Archives of General Psychiatry* **56**, 557–563.

Blazer DG (2003). Depression in late life: review and commentary. *Journal of Gerontology Series A: Biological Sciences and Medical Sciences* **58A**, 249–265.

Carmelli D, Swan GE, Kelly-Hayes M, Wolf PA, Reed T, Miller B (2000). Longitudinal changes in the contribution of genetic and environmental influences to symptoms of depression in older male twins. *Psychology and Aging* **15**, 505–510.

Christensen K, Holm NV, McGue M, Corder L, Vaupel JW (1999). A Danish population-based twin study on general health in the elderly. *Journal of Aging and Health* **11**, 49–64.

Djernes JK (2006). Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatrica Scandinavia* **113**, 372–387.

Ehrlich BS, Isaacowitz DM (2002). Does subjective well-being increase with age? *Perspectives in Psychology* **5**, 20–26.

Finkel D, McGue M (1993). The origins of individual differences in memory among the elderly: a behavior genetic analysis. *Psychology and Aging* **8**, 527–537.

Finkel D, Pedersen NL (2004). Processing speed and longitudinal trajectories of change for cognitive abilities:

The Swedish Adoption/Twin Study of Aging. *Aging, Neuropsychology, and Cognition. Special Issue: Longitudinal Studies of Cognitive Aging* **11**, 325–345.

Fiske A, Wetherell JL, Gatz M (2009). Depression in older adults. *Annual Review of Clinical Psychology* **5**, 363–389.

Forlani C, Morri M, Ferrari B, Dalmonte E, Menchetti M, De Ronchi D, Atti AR (2014). Prevalence and gender differences in late-life depression: a population-based study. *American Journal of Geriatric Psychiatry* **22**, 370–380.

Franz CE, Lyons MJ, O'Brien R, Panizzon MS, Kim K, Bhat R, Grant MD, Toomey R, Eisen S, Xian H, Kremen WS (2011). A 35-year longitudinal assessment of cognition and midlife depression symptoms: the Vietnam Era Twin study of aging. *The American Journal of Geriatric Psychiatry* **19**, 559–570.

Gatz M, Pedersen NL, Plomin R, Nesselroade JR, McClearn GE (1992). The importance of shared genes and shared environments for symptoms of depression in older adults. *Journal of Abnormal Psychology* **101**, 701–708.

Gatz M, Petkus A, Reynolds CA, Franz C, Kaprio J, Christensen K, for the IGEMS Consortium (2015a). Age moderation of individual differences in chronic medical illness burden [Abstract]. *Behavior Genetics* **45**, 657.

Gatz M, Reynolds CR, Finkel D, Hahn CJ, Zhou Y, Zavala C (2015b). Data harmonization in aging research: not so fast. *Journal of Experimental Aging Research* **41**, 475–495.

Gillespie NA, Kirk KM, Evans DM, Heath AC, Hickie IB, Martin NG (2004). Do the genetic or environmental determinants of anxiety and depression change with age? A longitudinal study of Australian twins. *Twin Research* **7**, 39–53.

Gold CH, Malmberg B, McClearn GE, Pedersen NL, Berg S (2002). Gender and health: a study of older unlike-sex twins. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences* **57B**, S168–S176.

Heath AC, Kessler RC, Neale MC, Hewitt JK, Eaves LJ, Kendler KS (1993). Testing hypotheses about direction of causation using cross sectional family data. *Behavior Genetics* **23**, 29–50.

Johnson W (2007). Genetic and environmental influences on behavior: capturing all the interplay. *Psychological Review* **114**, 423–440.

Johnson W, McGue M, Gaist D, Vaupel JW, Christensen K (2002). Frequency and heritability of depression symptomatology in the second half of life: evidence from Danish twins over 45. *Psychological Medicine* **32**, 1175–1185.

Judd LL, Schettler PJ, Akiskal HS (2002). The prevalence, clinical relevance, and public health significance of subthreshold depressions. *Psychiatric Clinics of North America* **25**, 685–698.

Kaprio J (2013). The Finnish Twin Cohort Study: an update. *Twin Research and Human Genetics* **16**, 157–162.

Kaprio J, Koskenvuo M (2002). Genetic and environmental factors in complex diseases: the older Finnish Twin Cohort. *Twin Research and Human Genetics* **5**, 358–365.

Kendler KS, Gardner CO (2014). Sex differences in the pathways to Major Depression: a study of opposite-sex twin pairs. *The American Journal of Psychiatry* **171**, 426–435.

Kendler KS, Gardner CO, Neale MC, Prescott CA (2001). Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychological Medicine* **4**, 605–616.

Kendler KS, Gatz M, Gardner CO, Pedersen NL (2006). A Swedish national twin study of lifetime major depression. *American Journal of Psychiatry* **163**, 109–114.

Kendler KS, Thornton LM, Gilman SE, Kessler RC (2000). Sexual orientation in a U.S. national sample of twin and nontwin sibling pairs. *American Journal of Psychiatry* **157**, 1843–1846.

Kessler RC, Foster C, Webster PS, House JS (1992). The relationship between age and depressive symptoms in two national surveys. *Psychology and Aging* **7**, 119–126.

Korhonen T, Broms U, Varjonen J, Romanov K, Koskenvuo M, Kinnunen T, Kaprio J (2007). Smoking behaviour as a predictor for depression among Finnish men and women – a prospective study of adult twins. *Psychological Medicine* **37**, 705–715.

Kremen WS, Thompson-Brenner H, Leung YJ, Grant MD, Franz CE, Eisen SA, Jacobson KC, Boake C, Lyons MJ (2006). Genes, environment, and time: the Vietnam Era Twin Study of Aging (VETSA). *Twin Research and Human Genetics* **9**, 1009–1022.

McArdle JJ, Hamagami F (2003). Structural equation models for evaluating dynamic concepts within longitudinal twin analyses. *Behavioral Genetics* **33**, 137–159.

McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, Plomin R (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science* **276**, 1560–1563.

McGue M, Christensen K (1997). Genetic and environmental contributions to depression symptomatology: evidence from Danish twins 75 years of age and older. *Journal of Abnormal Psychology* **106**, 439–448.

Meeks S, Murrell SA, Mehl RC (2000). Longitudinal relationships between depressive symptoms and health in normal older and middle-aged adults. *Psychology and Aging* **15**, 100–109.

Neale M, Hunter M, Pritikin JN, Zahery M, Brick TR, Kickpatrick RM, Estabrook R, Bates TC, Maes H, Boker SM (2016). Open Mx 2.0: extended structural equation and statistical modeling. *Psychometrika* **81**, 535–549.

Neale M, Roysamb E, Jacobson K (2006). Multivariate genetic analysis of sex limitation and GxE interaction. *Twin Research and Human Genetics* **9**, 481–489.

Neiderhiser JM, Lichtenstein P (2008). The Twin and Offspring Study in Sweden: advancing our understanding of genotype-environment interplay by studying twins and their families. *Acta Psychologica Sinica* **40**, 1116–1123.

Pedersen NL, Christensen K, Dahl A, Finkel D, Franz C, Gatz M, Horwitz BN, Johansson B, Johnson W, Kremen WS, Lyons MJ, Malmberg B, McGue M, Neiderhiser JM, Peterson I, Reynolds CA (2013). IGEMS: The Consortium on Interplay of Genes and Environment across Multiple Studies. *Twin Research and Human Genetics* **16**, 481–489.

Radloff LS (1977). The CES – D scale: a self – report depressive mood scale for research in the general population. *Applied Psychological Measurement* **3**, 385–401.

Roth M, Tym E, Mountjoy CQ, Huppert FA, Hendrie H, Verma S, Goodard R (1986). CAMDEX: a standardised

instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry* **149**, 698–709.

Salvi F, Miller MD, Grilli A, Giorgi R, Towers AL, Morichi V, Spazzafumo L, Mancinelli L, Espinosa E, Rappelli A, Dessì-Fulgheri P (2008). A manual of guidelines to score the modified cumulative illness rating scale and its validation in acute hospitalized elderly patients. *Journal of the American Geriatrics Society* **56**, 1926–1931.

SAS Institute Inc. (2013). *SAS 9.4*. SAS Institute Inc.: Cary, NC.

Schnittker J (2014). *Natural Symptoms? The Intersection of Social, Biological, and Genetic Determinants of Depression in Later Life*. Working paper: University of Pennsylvania Population, Aging Research Center. PARC Working Paper Series, WPS 14–01.

Scott KM, Lim C, Al-Hamzawi A, Alonso J, Bruffaerts R, Caldas-de-Almeida JM, Florescu S, de Girolamo G, Hu C, de Jonge P, Kawakami N, Medina-Mora ME, Moskalewicz J, Navarro-Mateu F, O'Neill S, Piazza M, Pasada-Villa J, Torres Y, Kessler RC (2016). Association of mental disorders with subsequent chronic physical conditions: world mental health surveys from 17 countries. *JAMA Psychiatry* **73**, 150–158.

Skytthe A, Christiansen L, Kyvik KO, Bødker FL, Hvidberg L, Petersen I, Nielsen MMF, Bingley P, Hjelmborg J, Tan Q, Holm NV, Vaupel JW, McGue M, Christensen K (2013). The Danish Twin Registry: linking surveys, national registers, and biological information. *Twin Research and Human Genetics* **16**, 104–111.

Sonnenberg CM, Beekman ATF, Deeg DJH, van Tilburg W (2000). Sex differences in late life depression. *Acta Psychiatrica Scandinavica* **101**, 286–292.

Steptoe A (2007). *Depression and Physical Illness*. Cambridge University Press: New York.

Sutin AR, Terracciano A, Milaneschi Y, An Y, Ferrucci L, Zonderman AB (2013). The trajectory of depressive symptoms across the adult lifespan. *JAMA Psychiatry* **70**, 803–811.

Takkinen S, Tolvanen A, Kaprio J, Berg S, Koskenvuo M, Rantanen T (2004). The genetic and environmental effects on depressive symptoms among older female twins. *Twin Research* **7**, 626–636.

Turkheimer E, Harden KP (2014). Behavior genetic research methods: testing quasi-causal hypotheses using multivariate twin data. In *Handbook of Research Methods in Personality and Social Psychology* (ed. H. T. Reis and C. M. Judd), pp. 159–187. John Wiley: New York.

Twenge JM (2015). Time period and birth cohort differences in depressive symptoms in the U.S., 1982–2013. *Social Indicators Research* **121**, 437–454.

van der Sluis S, Posthuma D, Dolan CV (2012). A note on false positives and power in GxE modelling of twin data. *Behavior Genetics* **42**, 170–186.

Van Hulle C, Lahey B, Rathouz P (2013). Operating characteristics of alternative statistical methods for detecting gene-by-measured environment interaction in the presence of gene-environment correlation in twin and sibling studies. *Behavior Genetics* **43**, 71–84.