THE LANCET
Psychiatry

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Lee SL, Pearce E, Ajnakina O, et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: a 12-year population-based cohort study. *Lancet Psychiatry* 2020; published online Nov 9. http://dx.doi.org/10.1016/S2215-0366(20)30383-7.

Table of contents

Supplementary Table 14	Primary outcome analysis of separate timepoints	21
STROBE Statement	Completed checklist	22

Measures of social experiences related to but distinct from loneliness

Social support: Questions on social support covered four relationship types (partner, children, family members, and friends) and for each relationship, participants responded to three items (e.g. "How much can you rely on them if you have a serious problem?") on a 4-point Likert scale ranging from a lot to not at all. Participants without the relevant relationship scored zero. Total scores ranged from 0-36, higher scores indicating better support.

Social network size: We used three questions about the number of children, other family members and friends that participants had a close relationship with. Participants could list between 0 and 10 relationships per category. Total scores ranged from 0 to 30, higher scores indicating a larger network size.

Social contact frequency: We rescaled and combined questions on how often respondents contacted their children, other family members and friends on a 6-point Likert scale ranging from less than once a year to three or more times a week. Total scores ranged from 0 to 18, higher scores indicating greater contact frequency.

Participation in social groups: We used a binary variable, asking respondents if they participated in groups such as social clubs, residents' groups, or religious groups.

Measurement of confounders

Genetic variables: We examined potential genetic confounding of associations between loneliness and depressive symptoms using polygenic risk scores (PRS) calculated for genotyped ELSA participants. The genome-wide genotyping in ELSA was performed at University College London Genomics in 2013-2014 using the Illumina HumanOmni2.5 BeadChips (HumanOmni2.5-4v1, HumanOmni2.5-8v1.3), which measures ~2.5 million markers that capture the genomic variation down to 2.5% minor allele frequency (MAF). Genotyping was performed in two batches. Allele frequencies were compared between the batches after filtering for 5% of missingness. The correlation was calculated between the batches for a number of chromosomes and exceeded 99%. The two batches were merged as one data set. After quality reassurance, which entailed excluding ethnic outliers (self-reported) and duplicates, the genome-wide data were available for total 7412 ELSA participants of European ancestry and 2230767 SNPs.

Using PLINK,[1] R studio and VCFtools,[2] single-nucleotide polymorphism (SNPs) were excluded if they were non-autosomal, the minor allele frequency was <0.01%, if more than 2% of genotype data were missing and if the Hardy-Weinberg Equilibrium *P*-value<10^{-4}. Samples were removed based on call rate (<0.99), suspected non-European ancestry, sex difference in allelic frequency of ≥0.2, heterozygosity and relatedness. We employed the principal components analysis 4 to identify those individuals who deviated from the ethnic population they self-reported to be (i.e., ethnic outliers).[3,4] This set of analyses demonstrated the presence of ancestral admixture in the 65 individuals, who were subsequently removed. Concordant genetic ancestry and self-reported ethnicity participants were retained for further analyses. After these QC steps 7183 (96.9% n=7412) individuals and 1372240 (61.5% of n=2230767 SNPs) directly genotyped SNPs remained for further analyses.

Polygenic Risk Scores (PRS). The PRS for depressive symptoms was created using results from 2016 GWAS conducted by the Social Science Genetic Association Consortium (SSGAC) as part of their subjective wellbeing GWAS.[5] SSGAC summary statistics contained 6,524,474 SNPs; of these,

1,187,563 SNPs overlapped with the ELSA genetic database and were included in the PRS for depressive symptoms. The PRS for loneliness was created using results from a 2016 GWAS conducted by the Psychiatric Genomics Consortium utilising genotypic and phenotypic data from 10 760 individuals aged ⩾50 years that were collected by the Health and Retirement Study (HRS) to perform the first genome-wide association study of loneliness.[6] GWAS summary statistics contained 5,768,558 SNPs and, of these, 1,055,906 overlapped with the ELSA genetic database and were included in the PRS for loneliness. To calculate PRSs depressive symptoms and loneliness, SNPs associated with each of these outcomes, weighted by their effect size derived from the SSGAC and PGC respectively, and were summed in a continuous score using PRSice. As previous research highlighted that PGSs built from directly genotyped data either had more predictive power or did not differ significantly from PGSs calculated using imputed data, we calculated PGSs based on genotyped data at different *P*-value cut-offs.[7,8] Because PGSs including all available SNPs either explain the most amount of variation in a trait or are not significantly different than PGSs based on different *P*-value thresholds,[8] we utilised PGS that was based on threshold of *P*-value of 1.

Socio-demographics: Age in years was a continuous variable. Consistent with other ELSA studies, ethnicity was dichotomised into white and ethnic minority (due to small numbers in ethnic minority groups).[9] Marital status was dichotomised into married and other (single/divorced/separated/widowed). Educational attainment was a three-category variable indicating highest qualification (higher education, school-level, and no formal qualifications). Net non-pension wealth in quintiles was used as an indicator of socio-economic status. Working status was classified into employed and not employed (due to retirement, unemployment, illness or family reasons).

Physical health: Respondents were asked whether they suffered from any long-standing physical illness, yes or no. For mobility impairment, respondents were asked if they had difficulties with ten everyday activities (e.g. walking 100 yards, carrying over ten pounds). We analysed mobility impairment as binary. A binary variable was also created for pain, with respondents reporting whether

they were often troubled with pain (yes, no). BMI and waist circumference were continuous. We adjusted for waist circumference in addition to BMI because BMI may be a poor indicator for obesity-related health risk in older people due to its reduced ability to predict body fat.[10] Cognitive function was assessed using neuropsychological tests of immediate and delayed verbal memory, prospective memory, verbal fluency, cognitive speed and attention, and time orientation. As the scoring of each neuropsychological test varied, consistent with a prior study we generated z-scores and summed and re-standardised individual cognitive domain z-scores to obtain global cognitive function z-score.[11]

Supplementary Table 1. Specification for linear multilevel regression models.

Model	Specification
Model 1	Univariable: loneliness (exposure) and depressive symptoms (outcome)
Model 2	Model 1 plus continuous linear time variable
Model 3	Model 2 plus continuous quadratic time variable
Model 4	Model 3 plus social experiences (social support, social network size, social contact frequency, and participation in social groups)
Model 5	Model 4 plus polygenic risk scores
Model 6	Model 4 plus sociodemographics (age, sex, ethnicity, marital status, education, working status, and wealth)
Model 7	Model 6 plus health indicators (physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function)
Model 8	Model 7 plus depressive symptoms from wave two
Model 9	Model 8 plus interaction between loneliness and time[a]
Model 10	Model 9 plus interaction between loneliness and quadratic time[b]

[a]To test whether the association between loneliness and depressive symptoms differed across time-points.

[b]To test for non-linearity in the association between loneliness and depressive symptoms according to time-point.

Supplementary Table 2. Characteristics of the sample used for analyses compared with the sample excluded because of missing data.

Characteristic	Analytic sample (n=4211) [a]	Excluded sample (n=4960)[a]	*p* value
Sex			
Female	2310 (54.9%)	2778 (56.0%)	
Male	1901 (45.1%)	2182 (44.0%)	0.27
Ethnicity			
White	4211 (100%)	4736 (95.7%)	
Ethnic minority	0 (0%)	214 (4.3%)	<0.0001
Marital Status			
Married	2954 (70.2%)	3106 (62.6%)	
Other (single/divorced/separated/widowed)	1257 (29.9%)	1853 (37.4%)	<0.0001
Level of education			
Higher education (degree level and higher)	603 (14.3%)	512 (10.5%)	
Intermediate (school level qualifications)	2272 (54.0%)	2158 (44.4%)	
No qualification	1336 (31.7%)	2189 (45.1%)	<0.0001
Working status			
Employed	1426 (32.4%)	1367 (28.0%)	
Not employed	2785 (66.1%)	3515 (72.0%)	<0.0001
Wealth quintiles			
1 (Lowest)	490 (11.6%)	998 (22.5%)	
2	742 (17.6%)	915 (20.7%)	
3	896 (21.3%)	857 (19.4%)	
4	1014 (24.1%)	829 (18.7%)	
5 (Highest)	1069 (25.4%)	828 (18.7%)	<0.0001
Participation in social groups			
Yes	960 (22.8%)	989 (29.6%)	
No	3251 (77.2%)	2352 (70.4%)	<0.0001
Long-standing illness			
Yes	2297 (54.6%)	2930 (59.1%)	
No	1914 (45.5%)	2024 (40.9%)	<0.0001
Mobility impairment			
Present	2313 (54.9%)	2313 (54.9%)	

Absent	1847 (37.3%)	1898 (45.1%)	<0.0001
Often troubled with pain			
Yes	1490 (35.4%)	1947 (40.3%)	
No	2721 (64.6%)	2889 (59.7%)	<0.0001
Age (years)	65.1 (8.9)	67.5 (11.2)	<0.0001
Loneliness score at Wave 2 (range 3-9)	4.0 (1.4)	4.2 (1.6)	<0.0001
Depressive symptoms score at Wave 2 (range 0-8)	1.3 (1.2)	1.6 (1.9)	<0.0001
Social support score (range 0-36)	23.2 (7.2)	22.2 (7.6)	<0.0001
Social network size score (range 0-30)	7.3 (4.2)	6.9 (4.3)	<0.0001
Social contact frequency score (range 0-18)	9.0 (3.0)	8.5 (3.3)	<0.0001
Body mass index (BMI)	27.8 (4.8)	28.1 (5.0)	<0.0001
Waist circumference in cm	95.2 (12.9)	96.2 (13.5)	<0.0001
Overall cognitive function	0.0 (1.0)	-0.2 (1.1)	<0.001

Note. Data are mean (SD) or n (%).

[a]Sample with complete data are those with no missing data on exposure or confounders and at least one depressive symptoms outcome at any time-point (4211). The sample with complete data and those with missing (4960) add up to the total eligible sample at wave two (9171; see Figure 1).

Supplementary Table 3. Odds ratios for depression for each point-increase in loneliness, using a repeated measures depression outcome from six waves of follow-up (N=4211).

Model	Odds ratio (95% CI)	p value
Model 1: Univariable association	1.93 (1.82 to 2.04)	<.0001
Model 2: Model 1 adjusted for time	1.93 (1.84 to 2.05)	<.0001
Model 3: Model 2 adjusted for other social variables[a]	1.86 (1.75 to 1.99)	<.0001
Model 4: Model 3 adjusted for polygenic risk scores[b]	1.84 (1.73 to 1.95)	<.0001
Model 5: Model 4 adjusted for sociodemographics[c]	1.72 (1.62 to 1.82)	<.0001
Model 6: Model 5 adjusted for health indicators[d]	1.58 (1.49 to 1.68)	<.0001
Model 7: Model 6 adjusted for depressive symptoms at wave two	1.28 (1.21 to 1.35)	<.0001

[a]Other social variables: social network size, social contact frequency, social support, and participation in social groups.

[b]PRS for depressive symptoms and loneliness.

[c]Sociodemographic factors: age, sex, ethnicity, marital status, education, working status, and wealth.

[d]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

Supplementary Table 4. Unadjusted and adjusted odds ratios for depression according to a one-point increase in loneliness, at each individual timepoint (N=4211).

Timepoint	Odds ratio (95% CI) p value	
	Unadjusted	Fully adjusted [a]
Wave 3	2.12 (1.97 to 2.29) p<0. 0001	1.38 (1.27 to 1.51) p<0. 0001
Wave 4	1.99 (1.84 to 2.16) p<0. 0001	1.31 (1.21 to 1.43) p<0. 0001
Wave 5	1.92 (1.75 to 2.10) p<0. 0001	1.26 (1.61 to 1.35) p<0. 0001
Wave 6	1.82 (1.67 to 1.99) p<0. 0001	1.19 (1.08 to 1.30) p<0. 0001
Wave 7	1.92 (1.73 to 2.10) p<0. 0001	1.25 (1.13 to 1.36) p<0. 0001
Wave 8	1.82 (1.65 to 2.01) p<0. 0001	1.20 (1.08 to 1.32) p<0. 0001

[a]Adjusted for: social network size, social contact frequency, social support, participation in social groups, age, sex, ethnicity, marital status, education, working status, wealth, long-standing illness, mobility impairment, pain, body mass index, waist circumference, cognitive function and depressive symptoms at wave two.

Supplementary Table 5. Effect estimates, and 95% confidence intervals, for social experiences related to loneliness and polygenic risk scores (on depressive symptoms outcome, from primary model; N=4211).

Timepoint	Change in depressive symptom score (95% CI) p value	
	Unadjusted	Fully adjusted [b]
Social network size	0.0072 (-.0024 to 0.17) p=.14	0.0023 (-.0057 to 0.010) p=.57
Social contact frequency	0.012 (-.0020 to 0.026) p=.094	0.082 (-.0038 to 0.020) p=.18
Participation in social groups	0.24 (0.16 to 0.33) p<0.00001	0.019 (-.056 to 0.093) p=.62
Social support	-0.0094 (-.016 to -.0030) p=.0036	-0.0075 (-.013 to -.0015) p=.015
PRS for depressive symptoms	0.0050 (.0034 to .0065) p<0.00001	0.0033 (.0020 to .0046) p<0.00001
PRS for loneliness	0.00029 (-.00022 to .00081) p=.26	0.00016 (-.00027 to .00059) p=.46

[a]Includes loneliness, social network size, social contact frequency, social groups, social support and polygenic risk scores.

[b]Adjusted for loneliness, age, sex, ethnicity, marital status, education, working status, wealth, long-standing illness, mobility impairment, pain, body mass index, waist circumference, cognitive function, and depressive symptoms at baseline (wave two).

Supplementary Table 6. Change in depressive symptom score for each point increase in loneliness, using repeated measures of depressive symptoms from six waves of follow-up (N=4211) and multilevel negative binomial regression.

Model	Change in depressive symptoms (95% CI)	*p* value
Model 1: Univariable association	0.31 (0.28 to 0.33)	<0.00001
Model 2: Model 1 plus continuous linear time variable	0.31 (0.28 to 0.33)	<0.00001
Model 3: Model 2 plus continuous quadratic time variable[a]	0.31 (0.28 to 0.33)	<0.00001
Model 4: Model 2 adjusted for social experiences related to loneliness[b]	0.29 (0.26 to 0.31)	<0.00001
Model 5: Model 4 adjusted for polygenic risk scores[c]	0.28 (0.26 to 0.31)	<0.00001
Model 6: Model 5 adjusted for socio-demographic factors[d]	0.25 (0.23 to 0.28)	<0.00001
Model 7: Model 6 adjusted for health indicators[e]	0.21 (0.19 to 0.24)	<0.00001
Model 8: Model 7 adjusted for depressive symptoms at baseline[f]	0.12 (0.09 to 0.14)	<0.00001

[a]Time squared was subsequently excluded from models because of no evidence of departure from linearity.

[b]Social network size, social contact frequency, participation in social groups and perceived social support.

[c]PRS for depressive symptoms and loneliness.

[d]Confounders: age, sex, ethnicity, marital status, education, working status, and wealth.

[e]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

[f]Interactions between loneliness and time and loneliness and time squared were added to model 8; results for interaction terms are reported in the text.

Supplementary Table 7. Unadjusted and adjusted change in depressive symptoms for each point-increase in loneliness, according to timepoint (associations drawn from multilevel negative binomial models; N=4211)

Timepoint	Change in depressive symptom score (95% CI) p value	
	Unadjusted	Fully adjusted[a]
Wave 3	0.33 (0.30 to 0.36) p<0.00001	0.14 (0.11 to 0.17) p<0.00001
Wave 4	0.32 (0.29 to 0.35) p<0.00001	0.13 (0.10 to 0.16) p<0.00001
Wave 5	0.29 (0.26 to 0.32) p<0.00001	0.10 (0.07 to 0.13) p<0.00001
Wave 6	0.30 (0.27 to 0.33) p<0.00001	0.11 (0.08 to 0.14) p<0.00001
Wave 7	0.29 (0.26 to 0.33) p<0.00001	0.10 (0.07 to 0.13) p<0.00001
Wave 8	0.28 (0.25 to 0.32) p<0.00001	0.09 (0.06 to 0.13) p<0.00001

[a]PRS for depressive symptoms and loneliness were not imputed; genetic data cannot be imputed using standard imputation due to linkage disequilibrium.

[b]Adjusted for: social network size, frequency of social contacts, social support, participation in social groups, age, sex, ethnicity, marital status, education, working status, wealth, long-standing physical illness, mobility impairment, pain, body mass index, waist circumference, cognitive function and depressive symptoms at baseline (wave two).

Supplementary Table 8. Change in depressive symptom score for each point increase in loneliness, using panel data from waves 2-8 (N=4211).

Model	Change in depressive symptoms (95% CI)	p value
Model 1: Univariable association	0.36 (0.34 to 0.37)	<0.00001
Model 2: Model 1 plus continuous linear time variable	0.36 (0.34 to 0.37)	<0.00001
Model 3: Model 2 plus continuous quadratic time variable[a]	0.38 (0.36 to 0.41)	<0.00001
Model 4: Model 2 adjusted for social experiences related to loneliness[b]	0.35 (0.33 to 0.36)	<0.00001
Model 5: Model 4 adjusted for polygenic risk scores[c]	0.34 (0.33 to 0.36)	<0.00001
Model 6: Model 5 adjusted for socio-demographic factors[d]	0.33 (0.32 to 0.35)	<0.00001
Model 7: Model 6 adjusted for health indicators[e]	0.32 (0.31 to 0.34)	<0.00001
Fixed effects model (within-person change in loneliness)	0.23 (0.22 to 0.24)	<0.00001

[a]Time squared was subsequently excluded from models because of no evidence of departure from linearity.

[b]Social network size, social contact frequency, participation in social groups and perceived social support.

[c]PRS for depressive symptoms and loneliness.

[d]Confounders: age, sex, ethnicity, marital status, education, working status, and wealth.

[e]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

[f]Interactions between loneliness and time and loneliness and time squared were added to model 8.

Supplementary Table 9. Change in depressive symptom score for each point increase in loneliness, using repeated measures of depressive symptoms from six waves of follow-up (N =4211), including loneliness item in CES-D.

Model	Change in depressive symptoms (95% CI)	p value
Model 1: Univariable association	0.46 (0.44 to 0.49)	<0.00001
Model 2: Model 1 plus continuous linear time variable	0.46 (0.44 to 0.49)	<0.00001
Model 3: Model 2 plus continuous quadratic time variable[a]	0.46 (0.44 to 0.49)	<0.00001
Model 4: Model 2 adjusted for social experiences related to loneliness[b]	0.44 (0.41 to 0.47)	<0.00001
Model 5: Model 4 adjusted for polygenic risk scores[c]	0.44 (0.41 to 0.47)	<0.00001
Model 6: Model 5 adjusted for socio-demographic factors[d]	0.40 (0.37 to 0.43)	<0.00001
Model 7: Model 6 adjusted for health indicators[e]	0.36 (0.33 to 0.39)	<0.00001
Model 8: Model 7 adjusted for depressive symptoms at baseline[f]	0.20 (0.17 to 0.23)	<0.00001

[a]Time squared was subsequently excluded from models because of no evidence of departure from linearity.

[b]Social network size, social contact frequency, participation in social groups and perceived social support.

[c]PRS for depressive symptoms and loneliness.

[d]Confounders: age, sex, ethnicity, marital status, education, working status, and wealth.

[e]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

[f]Interactions between loneliness and time and loneliness and time squared were added to model 8; results for interaction terms are reported in the text.

Supplementary Table 10. Unadjusted and adjusted change in depressive symptoms for each point-increase in loneliness, according to timepoint (associations drawn from multilevel models), including loneliness item in CES-D.

Timepoint	Number of participants[a]	Change in depressive symptom score (95% CI) p value	
		Unadjusted	Fully adjusted [b]
Wave three	4060	0.41 (0.38 to 0.45) p<0.0001	0.25 (0.21 to 0.28) p<0. 0001
Wave four	3651	0.37 (0.34 to 0.41) p<0.0001	0.18 (0.14 to 0.22) p<0. 0001
Wave five	3387	0.38 (0.34 to 0.41) p<0.0001	0.20 (0.16 to 0.24) p<0. 0001
Wave six	3142	0.36 (0.32 to 0.40) p<0.0001	0.16 (0.12 to 0.21) p<0. 0001
Wave seven	2751	0.37 (0.33 to 0.41) p<0.0001	0.16 (0.12 to 0.21) p<0. 0001
Wave eight	2449	0.35 (0.31 to 0.40) p<0.0001	0.15 (0.10 to 0.19) p<0. 0001

[a]The number of participants with complete data on all variables at each Wave. This number varies because there are a different number of people with depression data at each wave (due to attrition).

[b]Adjusted for social network size, social contact frequency, participation in social groups, social support, age, sex, ethnicity, marital status, education, working status, wealth, long-standing illness, mobility impairment, pain, body mass index, waist circumference, cognitive function, and depressive symptoms at baseline (wave two).

Supplementary Table 11. Odds ratios for depression for each point-increase in loneliness, using a repeated measures depression outcome from six waves of follow-up (N=4211), including loneliness item in CES-D.

Model	Odds ratio (95% CI)	p value
Model 1: Univariable association	2.11 (1.82 to 2.04)	<.00001
Model 2: Model 1 adjusted for time	2.11 (1.82 to 2.04)	<.00001
Model 3: Model 2 adjusted for other social variables[a]	2.00 (1.88 to 2.14)	<.00001
Model 4: Model 3 adjusted for polygenic risk scores[b]	1.98 (1.85 to 2.11)	<.00001
Model 5: Model 4 adjusted for sociodemographics[c]	1.83 (1.72 to 1.95)	<.00001
Model 6: Model 5 adjusted for health indicators[d]	1.69 (1.60 to 1.79)	<.00001
Model 7: Model 6 adjusted for depressive symptoms at wave two	1.33 (1.26 to 1.41)	<.00001

[a]Other social variables: social network size, social contact frequency, social support, and participation in social groups.

[b]PRS for depressive symptoms and loneliness.

[c]Sociodemographic factors: age, sex, ethnicity, marital status, education, working status, and wealth.

[d]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

Supplementary Table 12. Unadjusted and adjusted odds ratios for depression according to a one-point increase in loneliness, at each individual timepoint (N=4211), including loneliness item in CES-D.

Timepoint	Odds ratio (95% CI) p value	
	Unadjusted	Fully adjusted [a]
Wave 3	2.38 (2.18 to 2.60) p<0. 00001	1.51 (1.38 to 1.64) p<0. 00001
Wave 4	2.07 (1.72 to 2.51) p<0. 00001	1.34 (1.22 to 1.45) p<0. 00001
Wave 5	2.17 (1.79 to 2.59) p<0. 00001	1.38 (1.26 to 1.51) p<0. 00001
Wave 6	1.93 (1.59 to 2.33) p<0. 00001	1.21 (1.11 to 1.37) p<0. 00001
Wave 7	1.98 (1.64 to 2.41) p<0. 00001	1.24 (1.13 to 1.38) p<0. 00001
Wave 8	1.95 (1.59 to 2.38) p<0. 00001	1.23 (1.17 to 1.38) p<0. 00001

[a]Adjusted for: social network size, social contact frequency, social support, participation in social groups, age, sex, ethnicity, marital status, education, working status, wealth, long-standing illness, mobility impairment, pain, body mass index, waist circumference, cognitive function and depressive symptoms at wave two.

Supplementary Table 13. Change in depressive symptom score for each point increase in loneliness, using repeated measures of depressive symptoms from six waves of follow-up (imputed sample; N=7974[a]).

Model	Change in depressive symptoms (95% CI)	*p* value
Model 1: Univariable association	0.37 (0.35 to 0.39)	<0.00001
Model 2: Model 1 adjusted for time	0.37 (0.35 to 0.39)	<0.00001
Model 3: Model 2 adjusted for quadratic time	0.37 (0.35 to 0.39)	<0.00001
Model 4: Model 2 adjusted for other social variables[b]	0.33 (0.31 to 0.35)	<0.00001
Model 5: Model 4 adjusted for sociodemographics[c]	0.30 (0.28 to 0.32)	<0.00001
Model 6: Model 5 adjusted for health indicators[d]	0.25 (0.23 to 0.27)	<0.00001
Model 7: Model 6 adjusted for depressive symptoms at Wave 2	0.12 (0.10 to 0.14)	<0.00001

[a]The imputed sample comprises all people with complete exposure data. PRS for depressive symptoms and loneliness were not imputed; genetic data cannot be imputed using standard imputation due to linkage disequilibrium.
[b]Other social variables: social network size, social contact frequency, social support, and participation in social groups.
[c]Sociodemographic factors: age, sex, ethnicity, marital status, education, working status, and wealth.
[d]Health indicators: physical illness, mobility impairment, pain, BMI, waist circumference, and cognitive function.

Supplementary Table 14. Unadjusted and adjusted change in depressive symptoms for each point-increase in loneliness, according to timepoint (associations drawn from multilevel models; imputed sample; N=7974[a]).

Timepoint	Change in depressive symptom score (95% CI) p value	
	Unadjusted	Fully adjusted[a]
Wave 3	0.41 (0.38 to 0.45) p<0.0001	0.18 (0.16 to 0.21) p<0. 0001
Wave 4	0.37 (0.34 to 0.41) p<0.0001	0.13 (0.11 to 0.16) p<0. 0001
Wave 5	0.38 (0.34 to 0.41) p<0.0001	0.13 (0.11 to 0.16) p<0. 0001
Wave 6	0.36 (0.32 to 0.40) p<0.0001	0.10 (0.08 to 0.13) p<0. 0001
Wave 7	0.37 (0.33 to 0.41) p<0.0001	0.09 (0.07 to 0.12) p<0. 0001
Wave 8	0.35 (0.31 to 0.40) p<0.0001	0.08 (0.06 to 0.11) p<0. 0001

[a]PRS for depressive symptoms and loneliness were not imputed; genetic data cannot be imputed using standard imputation due to linkage disequilibrium.

[b]Adjusted for: social network size, frequency of social contacts, social support, participation in social groups, age, sex, ethnicity, marital status, education, working status, wealth, long-standing physical illness, mobility impairment, pain, body mass index, waist circumference, cognitive function and depressive symptoms at baseline (wave two).

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract **(page 1)**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found **(page 3)**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **(page 6)**
Objectives	3	State specific objectives, including any prespecified hypotheses **(page 8)**
Methods		
Study design	4	Present key elements of study design early in the paper **(page 9)**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **(page 9)**
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up **(page 6)** *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed **N/A** *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable **(pages 10-11 and Appendix pages 4-6)**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **(pages 10-11 and Appendix pages 4-6)**
Bias	9	Describe any efforts to address potential sources of bias **(page 9)**
Study size	10	Explain how the study size was arrived at **(pages 9, 14 and 22)**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **(pages 12-14)**

Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **(pages 12-14)**
		(*b*) Describe any methods used to examine subgroups and interactions **(pages 12-14)**
		(*c*) Explain how missing data were addressed **(pages 13-14)**
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed **(pages 13-14)** *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses **(page 13)**

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed **(sample flowchart page 22)**
		(b) Give reasons for non-participation at each stage **(sample flowchart page 22)**
		(c) Consider use of a flow diagram **(sample flowchart page 22)**
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **(page 14)**
		(b) Indicate number of participants with missing data for each variable of interest **(sample flowchart page 22)**
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) **(page 14 and sample flowchart page 22)**
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time **(pages 14 and 25)**
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included **(pages 15-16 and 26-27)**
		(*b*) Report category boundaries when continuous variables were categorized **(page 10)**
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period **N/A**
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses **(page 16)**
Discussion		
Key results	18	Summarise key results with reference to study objectives **(page 17)**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **(pages 17-19)**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **(page 17 and pages 19-20)**
Generalisability	21	Discuss the generalisability (external validity) of the study results **(page 18)**
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based **(pages 1-2 and page 28)**

References

1. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795

2. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156-2158. doi:10.1093/bioinformatics/btr330

3. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909. doi:10.1038/ng1847

4. Wang D, Sun Y, Stang P, Berlin JA, Wilcox MA, Li Q. Comparison of methods for correcting population stratification in a genome-wide association study of rheumatoid arthritis: principal-component analysis versus multidimensional scaling. *BMC Proc*. 2009;3(S7). doi:10.1186/1753-6561-3-s7-s109

5. Okbay A, Baselmans BML, De Neve JE, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet*. 2016;48(6):624-633. doi:10.1038/ng.3552

6. Gao J, Davis LK, Hart AB, et al. Genome-Wide Association Study of Loneliness Demonstrates a Role for Common Variation. *Neuropsychopharmacology*. 2017;42(4):811-821. doi:10.1038/npp.2016.197

7. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*. 2016;533(7604):539-542. doi:10.1038/nature17671

8. Ware EB, Schmitz LL, Faul J, et al. Heterogeneity in polygenic scores for common human traits. *bioRxiv*. 2017;(5):106062. doi:10.1101/106062

9. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A*. 2013;110(15):5797-5801. doi:10.1073/pnas.1219686110

10. Janssen I, Katzmarzyk PT, Ross R. Waist circumference and not body mass index explains obesity-related health risk. *Am J Clin Nutr*. 2004;79(3):379-384. doi:10.1093/ajcn/79.3.379

11. Llewellyn DJ, Lang IA, Langa KM, Huppert FA. Cognitive function and psychological well-being: Findings from a population-based cohort. *Age Ageing*. 2008;37(6):685-689. doi:10.1093/ageing/afn194