Association between long-term PM$_{2.5}$ exposure and depression among Chinese adults in the context of population aging: a quasi-experimental study

Tao Xue$^{1}$, Tianjia Guan$^{2}$, Yixuan Zheng$^{3}$, Guannan Geng$^{3}$, Qiang Zhang$^{3}$, and Tong Zhu$^{4,*}$

$^{1}$ Institute of Reproductive and Child Health / Ministry of Health Key Laboratory of Reproductive Health and Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China, 100191

$^{2}$ School of Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, 100730

$^{3}$ Department of Earth System Science, Tsinghua University, Beijing, China, 100084

$^{4}$ BIC-ESAT and SKL-ESPC, College of Environmental Science and Engineering, Peking University, Beijing, China, 100871

* Corresponding authors: Tong Zhu (Room 501, College of Environmental Science and Engineering, Peking University, Beijing 100871, China; Telephone: 010-62754789; E-mail: tzhu@pku.edu.cn).

The authors declare no competing interests.
Abstract

**Background:** Recent studies suggest an association between mental disorders such as depression and air pollution. However, few studies examined the association between air pollution reduction and mental health improvement. Since 2013, China has carried out a series of clean air actions that have rapidly improved air quality, and provided a quasi-experimental scenario to examine the association.

**Method:** Based on nationwide surveys of Chinese adults from 2011 to 2015, we evaluated the association between long-term PM$_{2.5}$ exposure and a widely-used depression score (C-ESD score), using a mixed-effects model with multivariate adjustment. The association between PM$_{2.5}$ reduction and the score change was further explored using a difference-in-difference analysis of the temporal contrast between 2011 (before the actions) and 2015 (after the actions). To increase interpretability of the association, the estimated impact of PM$_{2.5}$ levels was compared to that of aging, a well-known risk factor for depression.

**Results:** A 10-μg/m$^3$ increase in PM$_{2.5}$ concentration was associated with a 3.63% (95% confidence interval [CI]: 2.00–5.27%) increase in the C-ESD score (a higher score indicates larger probability of depression). Aging of 1 year was associated with a score increase of 0.76% (0.45–1.07%), equivalent to the effect of a 2.1-μg/m$^3$ (95% CI: 1.1–4.2 μg/m$^3$) increase in PM$_{2.5}$ concentration. Difference-in-difference analysis confirmed the significant association between PM$_{2.5}$ reduction and the score reduction. We also found improved air quality during 2011–2015 offset the negative impact from 5-years’ aging.

**Conclusions:** This study added the epidemiological evidence on the association between depression and long-term exposure to PM$_{2.5}$. Our findings also suggested the mental health benefits from China’s recent efforts to reduce air pollution.

**Keywords:** Air quality policy; PM$_{2.5}$ pollution; Depression; Mental health; Quasi-experiment
Mental disorders have contributed to a large proportion of the global burden of diseases by causing either non-fatal (e.g., disability) or fatal outcomes (e.g., suicide). Mental illness is the leading cause of years lived with disability (YLD), and the related disease burden increased by 37.6% from 1990 to 2010 (Ferrari et al. 2014). The increased burden of mental illness may be caused by both demographic dynamics, including population growth and aging (Ferrari et al. 2014), and changes in relevant risk factors, such as physical activity and social stress. Recent epidemiological evidence suggests that air pollution may adversely affect mental health (H Chen et al. 2017; Generaal et al. 2019; Li et al. 2018; Pun et al. 2017; Roberts et al. 2019; Xue et al. 2019b; Zhang et al. 2018; Zheng et al. 2019). Due to the ubiquitous and prolonged exposure to ambient pollutants, air quality can be a major contributor to mental illness, such as depression. However, the level of long-term exposure to air pollution usually varies slowly; thus, for an individual, its impact on mental health cannot be easily observed due to the long timespan needed to observe any changes, which impedes establishment of a longitudinal study (e.g., a cohort study) to examine the association between air quality and mental disorders.

Since 2013, air quality in China has improved rapidly due to a series of clean air actions (Cheng et al. 2019; Huang et al. 2018; Xue et al. 2019a; Zheng et al. 2017), including optimization of the industrial structure, improvements in end-of-pipe control, and reductions in residential usage of unclean fuels. The long-term concentrations of fine particulate matter less than 2.5 μm in diameter (PM$_{2.5}$), the major air pollutant in China, decreased significantly nationwide from 2013 to 2017. The population-weighted concentration of PM$_{2.5}$ decreased rapidly from 67.4 μg/m$^3$ in 2013 to 45.5 μg/m$^3$ in 2017 (Xue et al. 2019a). This dramatic change in air quality provides an opportunity to study the mental effect of PM$_{2.5}$ under a quasi-experimental scenario. A quasi-experiment can result in a shape contrast of air pollution levels during a relatively short period, which is advantageous to control confounders (Chen et al. 2013; Rich et al. 2012).

In this study, we made use of the China Health and Retirement Longitudinal Study (CHARLS), which repeatedly measured the depression score (C-ESD score) of a representative sample of Chinese adults in 2011, 2013, and 2015 using a standard questionnaire (Zhao et al. 2014). Due to the overlap between the CHARLS study period and the duration of China’s clean air actions, we were able to associate the reduced concentrations of PM$_{2.5}$ to changes in the C-ESD score at the individual level, and to quantify the impact of improved air quality on adult mental health. Because the C-ESD is not clinically indicative, interpreting the impact of air quality is difficult. To increase interpretability, we further compared the effect of PM$_{2.5}$ concentrations to that of population aging, a well-known risk factor for poor mental health (Ferrari et al. 2014).
Methods

Study population

The studied adults were obtained from the open-access CHARLS database (Zhao et al. 2013; Zhao et al. 2014). CHARLS is an ongoing nationwide longitudinal survey on the health and socioeconomic status of older Chinese adults. Using a four-stage and well-established sample design, CHARLS researchers surveyed a representative sample of ~20,000 Chinese adults from ~150 county-level units in 2011 (Figure 1). Subsequently, researchers performing the CHARLS regularly visited the respondents every 2 years (Figure S1), and thus surveyed a national-scale cohort of Chinese adults. In each survey wave, all subjects were visited by well-trained interviewers in a face-to-face computer-assisted personal interview, which gathered data on demographic characteristics, behavioral risk factors, the residential city, and housing conditions. The mental health status of the subjects was measured using the 10-question Center for Epidemiologic Studies Depression scale (CESD-10). The validity of the CESD-10 has been examined among Chinese adults (Boey 1999). Each question measures the frequency of a specific type of negative mood (e.g., fearful or depressed) using a score of 0 (rarely or none), 1 (some days), 2 (occasionally), or 3 (most of the time). We calculated the sum of all question-specific scores as the C-ESD score (0–30) to indicate the general depression status of each subject. Details of the study design and the purpose of the CHARLS are provided elsewhere (Zhao et al. 2014). The CHARLS was approved by the Ethics Review Committee of Peking University (IRB00001052–11015). CHARLS has supported studies on mental health risk factors (Hu et al. 2019) and the effects of air pollution on other outcomes (Liu et al. 2016). This study utilized the open-accessed CHARLS data, which is publicly available from the website: http://opendata.pku.edu.cn/. All analyses in this study adhered to the data usage guidelines.

This study was based on currently available data from CHARLS surveys in 2011, 2013, and 2015, which involved a sample of 24,805 individual adults. To conduct a longitudinal analysis, we focused on the valid records of older adults (≥ 40 years old), who were visited at least twice. A total of 16,151 adults remained after excluding subjects who were only observed once. We further excluded the adults without age data (which was derived from the time of the survey and the birthdate) and adults <40 years old at survey time. Finally, this study involved 41,031 observations of 15,954 individual adults distributed across 447 communities in 126 cities (Figure 1).

Environmental exposure

The ambient exposure assessment was based on the PM_{2.5} Hindcast Database for China (2000–2016) introduced in our previous work (Xue et al. 2019a). The database can be accessed from the website:
http://www.meicmodel.org/dataset-phd.html. Due to the lack of nationwide monitoring data on PM$_{2.5}$ concentrations in China before 2013, historical PM$_{2.5}$ concentrations were estimated from satellite remote-sensing measurements and outputs from a chemical transport model (CTM) of air pollution emission inventories. Satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer, retrieve the column concentrations of aerosols from the earth’s surface to the top of the atmosphere by measuring electromagnetic signals. Satellite-based annual estimates of PM$_{2.5}$ have been applied in many health-related studies at the national and global scales (Goyal et al. 2019; Zheng et al. 2017). The CTM simulations were based on the multi-resolution emission inventory for China (http://www.meicmodel.org/index.html). The results provide a complete characterization of the spatiotemporal variations in PM$_{2.5}$ concentrations and have been applied to support studies on health risk assessments and relevant policy analyses in China (Cheng et al. 2019; Q Zhang et al. 2017). In a previous study (Xue et al. 2019a), we developed a machine-learning model to bring the satellite measurements and CTM simulations together by relating them to the nationwide monitoring concentrations of PM$_{2.5}$ from 2013 to 2016, and then applied the model to hindcast the PM$_{2.5}$ values before 2013. The estimator was in good agreement with the *in situ* observations on monthly ($R^2 = 0.71$) and yearly scales ($R^2 = 0.77$) and has been utilized in other epidemiological studies. Please refer to our previous study for more details on the PM$_{2.5}$ estimator (Xue et al. 2019a).

The original PM$_{2.5}$ data have a spatial resolution of $0.1^\circ \times 0.1^\circ$ and daily concentrations across the mainland of China, during 2000–2016. For consistency, exposure assessments before and after 2013 were based on the estimated PM$_{2.5}$ concentrations. The subjects in the open-access CHARLS data could only be geo-coded at the city level for confidentiality reasons, and detailed survey dates were not available. Therefore, we first pooled the PM$_{2.5}$ data into city-level averages by matching the pixels of a regular grid with a map of China’s prefectures (Figure 1), and further calculated the monthly averages. We utilized the PM$_{2.5}$ concentration averages during the 12 months preceding the surveyed months as the exposure values (Figure S1).

We also obtained gridded estimates of temperature with an original resolution of $0.1^\circ \times 0.1^\circ$ by fusing the satellite measurements of land surface temperatures, *in situ* observations, and simulations from a weather-forecast research model (Xue et al. 2019b). The three types of temperature values were assembled by day using a universal kriging approach. Random cross-validations indicated that the fused estimates were in good agreement with the monitored values ($R^2 = 0.96$). Details of the temperature data assembly are documented in our previous study (Xue et
City-level monthly averages for temperature data were also calculated for each record before the regression analyses.

**Difference-in-difference analyses**

In quasi-experiments, the difference-in-difference (DD) approach has been typically utilized to evaluate the impact of policy interventions (Deschenes and Meng 2018), including in studies of the effect of air pollutants on health (Pope III et al. 2009; Xue et al. 2019b; Xue et al. 2019b). We utilized the DD model to examine the association between PM$_{2.5}$ concentration reductions on the C-ESD score change.

The reductions in PM$_{2.5}$ from 2011 to 2015 were driven by the clean air actions (Zhang et al. 2019); the magnitude of the changes can be interpreted in the context of different emission-control measures, and might be unrelated to conventional risk factors for depression, such as aging. In conventional DD analysis of a binary exposure, subjects residing in areas subject to more stringent policies (which resulted in a larger PM$_{2.5}$ reduction) were included in the treatment group, and living in areas subject to less stringent policies were the controls (Figure S2). The between-group difference in the change in C-ESD score from 2011 (i.e., before intervention) to 2015 (i.e., after intervention) could be taken as a proxy of the effect of policy. Because PM$_{2.5}$ is a continuous variable, a regression analysis of the association between changes in PM$_{2.5}$ and C-ESD score was performed. The DD analysis requires treatments and controls to be randomly assigned. Therefore, we utilized the inverse probability of changes in PM$_{2.5}$ as the regression weights. Because the subjects were clustered by household and sampling site (i.e., community), we also incorporated the corresponding random effects into the regression model. DD analysis is advantageous by controlling for some unmeasured confounders (e.g., aging) by the study design itself (Deschenes and Meng 2018). Because the CHARLS 2013 samples were partially influenced by the air quality intervention (Figure S1), the DD model included only samples from the 2011 and 2015 waves. Because the distribution of the absolute changes in C-ESD score was skewed and leptokurtic, we modeled the relative changes. To do this, we added 1 to the original C-ESD score and calculated the changes on a logarithmic scale such that the skewness of the distribution was reduced (skewness = 0.03), as was the leptokurtosis (kurtosis = 0.65). We calculated the excess risk (ER) for each 10 μg/m$^3$ change in PM$_{2.5}$ concentration using the following equation: $ER = (\exp(10\beta) - 1) \times 100\%$, where $\beta$ denotes the estimated regression coefficient.

In sensitivity analyses of the DD model, we first examined whether our results were altered by adjustment for covariates. The covariates included constant characteristics (i.e., urban/rural
residency, sex, educational level, and age in 2011) and changes in longitudinal variables (i.e., ambient temperature, marriage status, smoking, drinking, cooking energy type, building type, residential rent payment, presence of an in-house telephone, and indoor temperature maintenance). The missing covariate values were first imputed using the chained-equation approach (Buuren and Groothuis-Oudshoorn 2010). Next, we presented the results of the re-estimated DD models without regression weights, and evaluated the importance of randomly assigning the samples across different levels of exposure change. Additionally, we removed the random effects to control for spatial clustering of the samples from the original DD models or changed them through replacing the community term by the city term, and recalculated the effect estimates. Finally, since the DD model presumed the outcome variable processed in the same pattern between treatments and controls, we examined whether the C-ESD score variations during the pre-treatment period (2011-2013) were associated to the treatment-control assignment.

**Longitudinal association analyses**

To make full use of the data, we applied a mixed-effects model to evaluate the association between the C-ESD score and long-term PM$_{2.5}$ exposure, including a few risk factors for depression, such as age. The aim of the longitudinal analysis was to clarify the impact of PM$_{2.5}$ exposure by comparing it with the impact of other risk factors.

The mixed-effects model was specified as follows:

$$\log(\text{score}_{i,j}) = \beta_0 + \beta_1 \text{PM}_{2.5,ij} + \beta_2 \text{age}_{i,j} + \beta_3 (\text{PM}_{2.5,ij} \times \text{age}_{i,j}) + \mathbf{z}_{i,j}^T \mathbf{\gamma} + \text{city}_i + \eta(\text{community}_i) + \lambda, \quad \ldots \quad (1)$$

where $i$ denotes the subject index; $j$ denotes the visit index; $\beta_0$ denotes the intercept; $\beta_1$-$\beta_3$ denote the regression coefficients for PM$_{2.5}$, age, and their interaction term, respectively; $\mathbf{z}_{i,j}$ denotes a set of adjusted covariates and $\mathbf{\gamma}$ denotes the corresponding coefficients; city, denotes a fixed effect to control the unmeasured city-specific risk factors of depression, such as traditional culture (Ng 1997); and $\lambda$ and $\eta$ denote two random slopes to model the correlations between records from the same subject or the same community, respectively. The adjusted covariates ($\mathbf{z}_{i,j}$) included (1) annual temperature; (2) demographic characteristics (urban/rural residency, sex, education level, and marriage status); (3) lifestyle risk factors (smoking and drinking); (4) housing conditions (cooking energy type, building type, residential rent payment, presence of an in-house telephone, and indoor temperature maintenance). The regression incorporating all the above covariates was termed the fully adjusted model. The model might be too complex to produce a stable estimator, and a large fraction of the regressed samples involved the imputed missing values, which increased the uncertainty of the model. To examine whether the estimated associations were sensitive to these
limitations, besides the full adjustment, we also applied a standard adjustment, which involved only the first three sets of covariates, demographic characteristics, and lifestyle risk factors. The association was also evaluated by ER, similar to the DD model.

In sensitivity analyses of the longitudinal model, we explored how the estimated association between PM$_{2.5}$ concentration and the C-ESD score varied with (1) sub-regions of the study domain (Figure 1), (2) demographic sub-groups, and (3) exposure levels. Considering the balanced sample sizes (Figure 1), we divided the study domain into three sub-regions: midwest ($n = 13,886$), north ($n = 11,488$) and southeast ($n = 15,657$). We utilized interaction analyses to explore the modifications on the effect of PM$_{2.5}$ concentrations by a sub-region indicator or demographic characteristic (e.g., sex), and replaced the linear term of PM$_{2.5}$ concentration in the regression model with a penalized spline term to examine the linearity of the effect. We also conducted parallel analyses for the association between age and the C-ESD score to compare the effect of PM$_{2.5}$ concentration to that of population aging. Finally, we re-fitted the standard model using subject-specific fixed effects instead of the random effects in equation (1). The fixed-effects model is a generalized version of the DD method that can incorporate more than two repeated measurements (Deschenes and Meng 2018).

A longitudinal regression with fixed effects is considered less biased, but also less efficient, than a comparable regression with random effects (Gunasekara et al. 2014). The refitted model was utilized to examine whether the estimated association was sensitive to the different approaches used to model subject-specific effects.

To illustrate the impact of the changes in air quality on depression, we conducted a post-hoc analysis based on the 9,123 adults who participated in all three CHARLS waves. We first calculated the change in one risk factor ($\Delta x = x_{2015} - x_{2011}$; $x =$ PM$_{2.5}$, age, or z) from 2011 to 2015, and quantified its impact on C-ESD score as $[\exp(\Delta x \beta) - 1] \times 100\%$. We compared the impact of the risk factors in the association model with standard adjustment, and focused on the combined impact of PM$_{2.5}$ reduction and population aging in a group of adults.

All statistical analyses were performed using R (version 3.3.2; R Foundation for Statistical Computing, Vienna, Austria). The linear mixed-effects and fixed-effects models were inferred using the lme4 package (Bates et al. 2014) and the plm package (Croissant and Millo 2008), respectively. Imputation was performed using the mice package (Buuren and Groothuis-Oudshoorn 2010). Inverse probability weights were calculated using the ipw package. Penalized spline functions were parameterized using the mgcv package (Wood 2011), and inference of the nonlinear mixed-effects models was done using the gamm4 package (Wood and Scheipl 2014). The relevant R codes are documented in the Supplementary materials.
Results

Descriptive summary

This study involved 15,954 adults, and each was visited an average of 2.6 times. In 2011, the mean age of the adults was 58.4 years (standard deviation: 9.4 years). Because of the different response rates (X Chen et al. 2017), more rural adults (61.21%) were screened than urban adults (38.79%). There were slightly more female (52.22%) than male subjects (47.77%) due to differences in longevity. A summary of the demographic information of the adults in our study is presented in Table 1.

The longitudinal variables, including population characteristics and environmental exposure, are summarized in Table 1. Although the mean C-ESD score in the 2013 (score = 7.8) or 2015 CHARLS (8.1) was smaller than that in the 2011 CHARLS (8.3), the comparison is not indicative of the trend in depression, given the between-wave difference in the surveyed population. The long-term mean PM$_{2.5}$ concentration for 2015 CHARLS subjects was 53.1 μg/m$^3$, which was considerably lower than the value for the 2013 (60.3 μg/m$^3$) or 2011 subjects (61.6 μg/m$^3$). The changes in PM$_{2.5}$ concentrations were within our expectations and were attributed to the clean air actions in China.

However, there were many other potential depression risk factors that changed in opposite directions during the study period. For example, we observed an increase in the fraction of unmarried adults from 15.52% in 2011 to 15.86% in 2013, and then to 17.74% in 2015. Marriage results in mental health benefits (Stack and Eshleman 1998); thus, this trend increased the C-ESD score. Living conditions, such as in-house bathing facilities and the ability to maintain the indoor temperature, have improved during the study period, which may have reduced the C-ESD score.

Additional information on the longitudinal variables can be found in Table 2. Because of the complexity of the drivers of depression, the lack of similarity in the trends of C-ESD score and PM$_{2.5}$ should not be interpreted as evidence against our hypothesis.

Associating PM$_{2.5}$ reductions with depression risk changes using difference-in-difference models

Among the 10,725 adults who participated in CHARLS 2011 and 2015, 741 had an increased concentration of PM$_{2.5}$ after the clean air actions. To illustrate the DD methodology, we conducted a preliminary analysis of binary PM$_{2.5}$ exposure (Figure S2), and set those subjects as controls (i.e., were not affected by the policy) and the others as the treatments. From 2011 to 2015, the mean C-ESD score of the controls increased by 3.7%, but that of the treatment group decreased by 0.9%. The between-group difference in the C-ESD score change was significant (P = 0.026), likely due to the intervention. In contrast, between-group difference before the completely implying the actions (2011-2013) was not statistically significant (P = 0.48; Table S1). The results (Table S1) suggest that the treatments and controls had a similar pattern of C-ESD score before the intervention.
Figure 2 presents the estimated effect of PM$_{2.5}$ on the C-ESD score, with the adjusted covariates of the DD models. The results of the different models were statistically comparable, considering uncertainties. The estimates consistently suggested that long-term exposure to PM$_{2.5}$ was significantly related to the C-ESD score. Based on the estimated ER from the standard model (Figure 2), a 10 μg/m$^3$ increase in PM$_{2.5}$ concentration was associated with a 4.14% (95% confidence interval [CI]: 0.41–8.00%) increase in the C-ESD score. In sensitivity analyses, the results were not significantly changed if we changed the settings to control for the effect of clustering of the samples, or did not weight the samples according to the probability of the reduction in PM$_{2.5}$ (Figure S3). However, the estimated effects in the unweighted models were slightly weaker than the weighted results. For instance, after removing the inverse probability weights, the standard DD model yielded an ER of 2.89% (95% [CI]: -0.59%, 6.48%). Because the samples were not optimally randomized (Figure S4), the unweighted estimates (Figure S3) could be slightly biased.
Estimating the exposure-response association between PM$_{2.5}$ with depression risk using longitudinal models

Based on the 41,031 samples from the three CHARLS waves, we quantified the association between PM$_{2.5}$ and C-ESD score using a longitudinal model (Figure 2). Although the point-estimates from the longitudinal models were slightly lower than those from the DD models, the two types of models reported comparable effects of PM$_{2.5}$ on the C-ESD score. Additionally, because the longitudinal models incorporated more samples, they had narrower CIs than the DD models. Adjusting for confounders, except age, did not significantly change the longitudinal results. As the studied population ages, a model that does not control for the effect of aging could yield a biased result.

According to the model with standard adjustments, a 10 μg/m$^3$ increase in PM$_{2.5}$ concentration was associated with a 3.63% (95% CI: 2.00–5.27%) increase in the C-ESD score. The results of the longitudinal models are listed in Table S2.

In sensitivity analyses, we showed that the association between PM$_{2.5}$ concentration and the C-ESD score did not vary significantly between sub-populations (Figure S5). The sub-region analysis, which displayed a weak association in the north but strong associations in the midwest and southeast (P = 0.04; likelihood ratio test). The weak association may have been caused by the small sample size in the north (Table 1) or a higher fraction of natural particles (e.g., dust in the northwest), which are known to be less toxic than anthropogenic particles. The nonlinear analysis showed an approximately linear exposure-response function of PM$_{2.5}$ concentration without a no-effect threshold (Figure 3a), which was consistent with estimates from the other models (Table S2).

Additionally, we examined the interaction between PM$_{2.5}$ concentration and age, and found that the effect of PM$_{2.5}$ concentration was weaker in older adults (Table S2). Finally, the estimated effect of PM$_{2.5}$ was not affected by use of random- or fixed-effects models (Table S3). However, the mixed-effects model, which could be more efficient, but also more biased, than the fixed-effects model (Gunasekara et al. 2014) was used only to enhance the interpretability of the estimates.

Evaluating the impact of PM$_{2.5}$

This study also examined the associations between age and the C-ESD score, so as to better interpret the mental health effects of PM$_{2.5}$ concentrations. The standard model indicated that the C-ESD score increased by 0.76% (95% CI: 0.45–1.07%) for an adult who aged by 1 year (Table S2). The association was robust, given different sets of adjusted covariates (Table S2). Subgroup analyses (Figure S5) suggested that the effect of age was not considerably modified by sub-region or sub-population indicators. The nonlinear analysis (Figure 4b) showed a sublinear exposure-response function for age, suggesting a weak marginal effect for adults older than ~70 years. Additionally, considering the collinearity between age and the temporal trend, we re-estimated the effect within
each CHARLS wave, based on the cross-sectional comparison of C-ESD score among adults of different ages. The results of the cross-sectional analysis (Table S4) were consistent with the estimates from the longitudinal models (Table S2).

A comparison between the effect of PM$_{2.5}$ concentration and that of age showed that they were statistically comparable (Figure 3b). Based on the standard model, the effect of 1 year was of equivalent magnitude to the effect of a PM$_{2.5}$ concentration increase of 2.1 μg/m$^3$ (95% CI: 1.1–4.2 μg/m$^3$).

The impact of a risk factor on the C-ESD score is determined by its effect magnitude and temporal changes. Based on the 9,123 adults who participated in all three waves of the CHARLS, we quantified the impacts of air quality improvement and aging on their mental health status from 2011 to 2015 (Figure 4). According to the results, aging and PM$_{2.5}$ reduction were the major drivers of changes in C-ESD score. Using the C-ESD score obtained before implementation of the clean air actions (by CHARLS 2011) as the reference, we found that a decreased concentration of PM$_{2.5}$ resulted in a relative reduction in the score of 0.27% (95% CI: 0.15–0.40%) and 2.87% (95% CI: 1.61–4.21%) in 2013 and 2015, respectively; in contrast, aging resulted in a relative increase of 1.53% (95% CI: 0.87–2.13%) and 3.08% (95% CI: 1.74–4.31%), respectively. Analysis of the combined impact of age and PM$_{2.5}$ (including their interactions), which together resulted in a small reduction in the C-ESD score of 0.25% (95% CI: -0.78–1.30%) from 2011 to 2015, showed that the benefit of reducing the PM$_{2.5}$ concentration offset the negative effect of aging. Because the impact evaluation was based on a fixed population, we quantified the impact of individual-level aging, which was faster than the aging of the population (Figure S1).

**Discussion**

Based on the CHARLS surveys, we conducted a longitudinal study of Chinese adults to associate mental health with long-term exposure to ambient PM$_{2.5}$ during a quasi-experimental scenario, in which air quality was rapidly improved by a series of stringent emission-control policies. We found a robust exposure-response function between air pollution and a commonly used indicator of depression, which was estimated to increase by 3.63% (95% CI: 2.00–5.27%) after a 10 μg/m$^3$ increase in long-term PM$_{2.5}$ concentration. Our results also indicate that the effect of PM$_{2.5}$ concentration reduction was comparable with that of aging. Although aging partially offset the benefits from reducing PM$_{2.5}$ concentrations, the rapidly improved air quality since 2013 could still bring a net positive impact.

An increasing number of studies have reported on the association between air pollution and mental disorders, including depression (Generaal et al. 2019; Lin et al. 2017; Pun et al. 2017; Pun et al. 2019;
Roberts et al. 2019; Tian et al. 2015; Vert et al. 2017; Wang et al. 2018; Wang et al. 2019; Zijlema et al. 2016), cognitive functions (Li et al. 2018; Tallon et al. 2017; Zhang et al. 2018), and other indicators of mental health (Klompmaker et al. 2019; Oudin et al. 2016; Shin et al. 2018; Sui et al. 2018; Sun et al. 2019; X Zhang et al. 2017; Zheng et al. 2019). For instance, Gu et al. (2019) conducted a meta-analysis of 684,859 subjects and reported that a 10 µg/m$^3$ increase in PM$_{2.5}$ concentration is strongly associated with a 19% (95% CI: 7–33%) ER of depression, and weakly associated with a 5% ER of suicide (95% CI: −1–11%) (Gu et al. 2019). Many of these studies examined the association between air pollution and mental illness among residents in China (Li et al. 2018; Sui et al. 2018; Sun et al. 2019; Tian et al. 2015; Wang et al. 2018; Wang et al. 2019; Xue et al. 2019b; X Zhang et al. 2017; Zhang et al. 2018). However, most of them were cross-sectional studies (Sui et al. 2018; Sun et al. 2019; Tian et al. 2015; Wang et al. 2018; Wang et al. 2019). Mental health can be affected by a number of complex factors, and comprehensively controlling these factors is critical to minimize bias. However, this may be difficult to achieve in a cross-sectional study. For example, Tian et al. conducted a cross-sectional study of 6,630 older adults (≥60 years old) from the 2013 CHARLS to associate the C-ESD score with sulfur dioxide emissions, and reported a U-shaped exposure-response function (Tian et al. 2015), which is biologically implausible and may have been caused by a failure to adjust for potential confounding factors. In the DD models and longitudinal models, the C-ESD score of a subject was compared to another observation of the same subject. Thus, the intra-individual comparison controlled for many unmeasured mental health risk factors (e.g., genetics), which varied only inter-individually. Therefore, our longitudinal study is more powerful than the previous cross-sectional analysis (Tian et al. 2015).

The quasi-experimental design based on the policy-driven rapid changes in air quality also added power to this study. The intervention exerted a marked effect on PM$_{2.5}$ exposure, which may have had a significant effect on the health outcomes. For instance, Pun et al. (2017) conducted a longitudinal study of the two waves of the National Social Life, Health and Aging project conducted in the United States (US), and examined the effect of PM$_{2.5}$ on depression based on an exposure contrast of 2.3 µg/m$^3$ from measurements separated by 5 years (annual PM$_{2.5}$ concentration = 11.1 and 8.8 µg/m$^3$ for the 2005–2006 and 2010–2011 waves, respectively). Compared to the US study, the mean difference of PM$_{2.5}$ in our study is 8.5 µg/m$^3$ during a 5-year period (Table 2), and may have an easily-distinguishable impact on health.

The analysis of the interaction between age and PM$_{2.5}$ suggested that the PM$_{2.5}$-depression association is weaker among older adults (Table S2). Although few similar depression findings have been reported, the relative risk of PM$_{2.5}$ is assumed to decline with age for the exposure-response
functions of other outcomes, such as mortality (Burnett et al. 2014). Because the risk factors other than air pollution increase with age, such an effect-modification is feasible.

The central government carried out an action plan of air pollution control and prevention from 2013 to 2017 to tackle severe air pollution in China (Xue et al. 2019a). Guided by the plan, provincial governments committed to a targeted reduction in annual concentration of particulate matter over 5 years and drew up emission-control pathways towards the targets. For example, the Beijing government planned to reduce the annual concentration of PM$_{2.5}$ to 60 $\mu g/m^3$ by 2017 and enacted a series of clean air actions, including (1) coal-fired boiler control, (2) promotion of clean fuels in the residential sector, (3) optimization of the industrial structure, (4) fugitive dust control, (5) vehicle-emission control, (6) improved end-of-pipe control, and (7) integrated treatment of volatile organic compounds. Based on a chemical-transport model analysis (Cheng et al. 2019), these seven measures respectively led to PM$_{2.5}$ concentration reductions of 5.9, 5.3, 3.2, 2.3, 1.9, 1.8, and 0.2 $\mu g/m^3$, which resulted in a considerable change in the mean monitored PM$_{2.5}$ concentration from 89.5 $\mu g/m^3$ in 2013 to 58.0 $\mu g/m^3$ in 2017. The clean air actions were expected to protect the public health by reducing cardiorespiratory mortalities associated with PM$_{2.5}$ exposure (Xue et al. 2019a).

The present study showed that these actions can also bring benefits to adult mental health, which further reduces the disease burden by decreasing the YLDs attributable to mental disorders.

The present study also examined the effect of aging on depression and compared it to the effect of PM$_{2.5}$ concentration. Such an analysis improved the interpretability of the association between PM$_{2.5}$ concentration and the depression and has policy implications. The Chinese population is expected to age rapidly because of a low fertility rate. According to the world population projection for the total population of China (Figure S6) (http://dataexplorer.wittgensteincentre.org/wcde-v2/), the mean age of all Chinese adults ≥ 40 years old will increase from 55 years in 2010 to 62–65 years in 2050. Based on our model (Table S2), population aging could offset the benefits from a 14.9–21.3 $\mu g/m^3$ reduction in PM$_{2.5}$ concentration. Assuming that air quality meets the current national standard (annual mean PM$_{2.5} < 35 $ $\mu g/m^3$) in 2050, the population-weighted concentration of PM$_{2.5}$ would decrease by 26 $\mu g/m^3$ (from 61 $\mu g/m^3$ in 2010 to 35 $\mu g/m^3$ in 2017). Therefore, the positive impact of PM$_{2.5}$ concentration reduction on depression is comparable to the negative impact of population aging. This rough analysis suggests that the older population requires more stringent air quality standards to protect their mental health. However, notably, the estimated association between age and depression not only reflects the underlying physiological effects (e.g., increased oxidative stress) but also the socioeconomic effects (e.g., decreased income) attributable to aging, and thus it may not be representative of the future population. Therefore, exactly quantifying the combined impacts of a change in air quality and population aging is beyond the capacity of this study. Further
prospective studies of the effects of changes in societal and cultural factors on the mental health of an aging population, and air quality, are warranted. This study was subjected to the following limitations. First, exposure to ambient PM$_{2.5}$ was assessed at the city-level due to the lack of specific addresses, which resulted in misclassification of exposure by ignoring the within-city variation in PM$_{2.5}$ concentrations. The misclassification might also be caused by falsely specifying the exposure time-window. Because detailed survey dates were unavailable in the open-access CHARLS data, the long-term exposure to PM$_{2.5}$ was evaluated using annual means based on monthly scale time-series. Additionally, uncertainties in the PM$_{2.5}$ concentration hindcast estimator are another source of exposure misclassification, which usually leads to an underestimated association. Second, although this study controlled for temporally invariant confounders (e.g., bullying victimization and childhood sexual abuse) and a few longitudinal covariates (including marriage, aging, and alcohol drinking), it may have missed some risk factors for depression. Although the estimates are unlikely to be confounded by some of them (e.g., drug abuse), failure to adjust for all potential confounders could have biased our results. Third, this study utilized PM$_{2.5}$ mass concentration as a general indicator of ambient air quality, which may have underestimated the complexities of air pollution toxicity. For instance, PM$_{2.5}$ is a mixture of particles of different chemical species and sizes, and between-component variations in toxicity have been shown for PM$_{2.5}$ (Han and Zhu, 2015, Han et al., 2016). Additionally, the estimated association between PM$_{2.5}$ and depression might not be attributable to only itself, but also to other air pollutants, such as ozone (Zhao et al. 2018), that co-vary with PM$_{2.5}$. Therefore, further studies on the effects on mental health of other air pollutants are warranted. Finally, some CHARLS subjects dropped out of the study for various reasons, which may have biased the results (Chen et al. 2017). Also, the limited information on missing data from the publicly available CHARLS datasets hampered the derivation of sampling weights for the statistical models, reducing the national representativeness of the findings. Given the above limitations, the causality of our findings should be interpreted cautiously.

Conclusions

Based on longitudinal measurements from 2011 to 2015 of a nationwide representative sample of 15,954 subjects, we found a robust association between an increase in PM$_{2.5}$ concentration and depression risk among the older in China. Due to clean air actions, the exposure concentrations of PM$_{2.5}$ decreased rapidly from 61.1 μg/m$^3$ to 53.3 μg/m$^3$, which improved mental health, and had a stronger impact than the negative effect from 5 years of aging. This study not only enriched the epidemiological evidence on the adverse effects of air pollution on mental health, but also indicated the mental health benefits from the clean air policies in China.
Declarations

Ethics approval and consent to participate
The CHARLS and its data usage was approved by the Ethics Review Committee of Peking University (IRB00001052–11015).

Consent for publication
Not applicable

Availability of data and materials
The data that support the findings of this study are available from China Center for Economic Research, Peking University but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of China Center for Economic Research, Peking University.

Competing interests
The authors declare that they have no competing interests.

Funding
National Natural Science Foundation of China (41701591, 81571130100, and 41421064) and Ministry of Science and Technology of China (2015CB553401).

Authors’ contributions
T.X. & T.Z. designed this study; T.X., Y.Z., G.G. & Q.Z. prepared and analysed the data; T.X., T.G. & T.Z. drafted the manuscript; all co-authors revised the manuscript together.

Acknowledgements
Thanks to the China Center for Economic Research, National School of Development, Peking University for providing the CHARLS data. This work was supported by National Natural Science Foundation of China (41701591, 81571130100, and 41421064) and Ministry of Science and Technology of China (2015CB553401). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.
References


Figure Legends

Figure 1. Map of study region with long-term averages of PM$_{2.5}$ concentrations (2010–2015).
Figure 2. **Estimated associations between the depression score and PM$_{2.5}$ concentration.** In the difference-in-difference (DD) analysis, the model adjusted for age incorporated age at 2011 as a covariate; the standard adjustment also involved the fixed variables of urban/rural residence, sex, and education, as well as the longitudinal variables of ambient temperature, marriage status, smoking, and drinking. The full adjustment also considered longitudinal changes in cooking energy type, building type, residential rent payment, presence of an in-house telephone, and indoor temperature maintenance. The longitudinal models adjusted for the constant variables in a similar manner as in the DD models; however, adjustment of the longitudinal variables was based on the values recorded in each survey wave, rather than their between-wave changes therein.
Figure 3. Nonlinear effects of (a) PM$_{2.5}$ concentration and (b) age on the C-ESD score. Dashed lines, pointwise 95% confidence intervals. The histograms show the distribution of PM$_{2.5}$ exposure and age.
Figure 4. Impacts of reducing PM$_{2.5}$ concentration, aging, and changes in other factors on the C-ESD score. The changes in risk factors ($\Delta x$) were evaluated using a fixed group of subjects who participated in all three China Health and Retirement Longitudinal Study (CHARLS) waves; the corresponding coefficients were estimated from the longitudinal model with standard adjustment (Figure 2) and presented in Table S3.
Tables

Table 1. Constant variables of the studied subjects.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subgroup</th>
<th>Number of subjects (percentage of the total)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All subjects</td>
</tr>
<tr>
<td>Total number of subjects</td>
<td></td>
<td>15,954 (100%)</td>
</tr>
<tr>
<td>Education</td>
<td>Below elementary</td>
<td>6,133 (38.44%)</td>
</tr>
<tr>
<td></td>
<td>Elementary and middle</td>
<td>6,099 (38.23%)</td>
</tr>
<tr>
<td></td>
<td>Above middle</td>
<td>1,712 (10.73%)</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>2,010 (12.60%)</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>8,331 (52.22%)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>7,621 (47.77%)</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>2 (0.01%)</td>
</tr>
<tr>
<td>Place of residence</td>
<td>Rural</td>
<td>9,765 (61.21%)</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>6,189 (38.79%)</td>
</tr>
<tr>
<td>Region</td>
<td>Midwest</td>
<td>5,367 (33.64%)</td>
</tr>
<tr>
<td></td>
<td>North</td>
<td>4,451 (27.90%)</td>
</tr>
<tr>
<td></td>
<td>Southeast</td>
<td>6,136 (38.46%)</td>
</tr>
</tbody>
</table>

Table 2. Longitudinal variables of the studied subjects.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>2011 CHARLS*</th>
<th>2013 CHARLS</th>
<th>2015 CHARLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression score</td>
<td>8.1 (6.2)</td>
<td>8.3 (6.3)</td>
<td>7.8 (5.8)</td>
<td>8.1 (6.4)</td>
</tr>
<tr>
<td>PM\textsubscript{2.5} concentration ((\mu g/m^3))</td>
<td>58.2 (19.8)</td>
<td>61.6 (18.9)</td>
<td>60.3 (22.2)</td>
<td>53.1 (16.7)</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>14.0 (5.3)</td>
<td>13.6 (5.2)</td>
<td>13.9 (5.6)</td>
<td>14.4 (5.0)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>60.5 (9.3)</td>
<td>59.2 (9.2)</td>
<td>60.3 (9.4)</td>
<td>61.9 (9.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subgroup</th>
<th>Number of visits (percentage of the total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of visits</td>
<td></td>
<td>41,031 (100%)</td>
</tr>
<tr>
<td>Married</td>
<td>No</td>
<td>6727 (16.39%)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>34,303 (83.60%)</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>1 (0.00%)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------</td>
<td>------------------</td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>28422 (69.27%)</td>
<td>12608 (30.73%)</td>
</tr>
<tr>
<td>Yes</td>
<td>8797 (69.50%)</td>
<td>3861 (30.50%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>9970 (69.47%)</td>
<td>4381 (30.53%)</td>
</tr>
<tr>
<td><strong>Drinking</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequent</td>
<td>10944 (26.67%)</td>
<td>3238 (25.58%)</td>
</tr>
<tr>
<td>Rare</td>
<td>3312 (8.07%)</td>
<td>977 (7.72%)</td>
</tr>
<tr>
<td>Never</td>
<td>26758 (65.21%)</td>
<td>8443 (66.70%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>3772 (26.90%)</td>
<td>3934 (27.41%)</td>
</tr>
<tr>
<td><strong>Cooking energy type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clean</td>
<td>21257 (51.81%)</td>
<td>5548 (43.83%)</td>
</tr>
<tr>
<td>Unclean</td>
<td>19292 (47.02%)</td>
<td>6965 (55.02%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>3934 (27.41%)</td>
<td>6512 (45.37%)</td>
</tr>
<tr>
<td><strong>Building type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One story</td>
<td>23683 (57.72%)</td>
<td>7884 (62.28%)</td>
</tr>
<tr>
<td>Multi-storey</td>
<td>17159 (41.82%)</td>
<td>4722 (37.30%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>7115 (50.75%)</td>
<td>5614 (39.12%)</td>
</tr>
<tr>
<td><strong>Rent payment for residence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>39353 (95.91%)</td>
<td>12289 (97.08%)</td>
</tr>
<tr>
<td>Yes</td>
<td>1200 (2.92%)</td>
<td>285 (2.25%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>13330 (95.07%)</td>
<td>457 (3.18%)</td>
</tr>
<tr>
<td><strong>In-house telephone</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24798 (60.44%)</td>
<td>6350 (50.17%)</td>
</tr>
<tr>
<td>Yes</td>
<td>16122 (39.29%)</td>
<td>6262 (49.47%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>9966 (71.08%)</td>
<td>5823 (40.57%)</td>
</tr>
<tr>
<td><strong>Indoor temperature maintenance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very hot</td>
<td>412 (1.00%)</td>
<td>254 (2.01%)</td>
</tr>
<tr>
<td>Hot</td>
<td>3573 (8.71%)</td>
<td>1345 (10.63%)</td>
</tr>
<tr>
<td>Bearable</td>
<td>34561 (84.23%)</td>
<td>10491 (82.88%)</td>
</tr>
<tr>
<td>Cold</td>
<td>1342 (3.27%)</td>
<td>436 (3.44%)</td>
</tr>
<tr>
<td>Very cold</td>
<td>102 (0.25%)</td>
<td>63 (0.50%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1041 (2.54%)</td>
<td>69 (0.55%)</td>
</tr>
</tbody>
</table>

* CHARLS, China Health and Retirement Longitudinal Study*