Malnutrition in early life and adult mental health: Evidence from a natural experiment

Cheng Huang a,*, Michael R. Phillips b,c,d, Yali Zhang e, jingxuan Zhang f, Qichang Shi g, Zhiqiang Song h, Zhijie Ding i, Shutao Pang j, Reynaldo Martorell k

a Department of Global Health, School of Public Health and Health Services, George Washington University, Washington, DC, USA
b Suicide Research and Prevention Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
c Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
d Department of Psychiatry, Emory University School of Medicine, Atlanta, USA
e WHQ Collaborating Center for Research and Training in Suicide Prevention, Beijing Huai Long Gouan Hospital, Beijing, China
f Shandong Provincial Mental Health Center, Jinan City, Shandong Province, China
g Li Tong De Hospital, Hangzhou City, Zhejiang Province, China
h Tianshui City Psychiatric Hospital, Tianshui City, Gansu Province, China
i Qingdao Mental Health Center, Qingdao City, Shandong Province, China
j Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
k Hubert Department of Global Health, Rollins School of Public Health, Emory University, USA

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Famine
Mental health
Selection effect
Natural selection in utero
China
Life course

ABSTRACT

As natural experiments, famines provide a unique opportunity to test the health consequences of nutritional deprivation during the critical period of early life. Using data on 4972 Chinese born between 1956 and 1963 who participated in a large mental health epidemiology survey conducted between 2001 and 2005, we investigated the potential impact of exposure to the 1959–1961 Chinese Famine in utero and during the early postnatal life on adult mental illness. The risk of mental illness was assessed with the 12-item General Health Questionnaire (GHQ-12) and eight other risk factors, and the famine impact on adult mental illness was estimated by difference-in-difference models. Results show that compared to unexposed women born in 1963, women born during the famine years (1959–1961) had higher GHQ scores (increased by 0.95 points; CI: 0.26, 1.65) and increased risk of mental illness (OR = 2.80; CI: 1.23, 6.39); those born in 1959 were the most affected and had GHQ scores 1.52 points higher (CI: 0.42, 2.63) and an OR for mental illness of 4.99 (CI: 1.68, 14.84). Compared to men in the 1963 birth cohort, men born during the famine had lower GHQ scores (decreased by 0.89 points; CI: −1.59, −0.20) and a nonsignificant decrease in the risk of mental illness (OR = 0.60; CI: 0.26, 1.40). We speculate that the long-term consequences of early-life famine exposure include both the selection of the hardest and the enduring deleterious effects of famine on those who survive. The greater biological vulnerability and stronger natural selection in utero of male versus female fetuses during severe famine may result in a stronger selection effect among men than women, obscuring the deleterious impact of famine exposure on the risk of mental illness in men later in life.

© 2012 Elsevier Ltd. All rights reserved.

Introduction

Child undernutrition remains highly prevalent across the world, in particular in low and middle income countries, where 32% (178 million) of children younger than 5 years had weight-for-age Z scores of less than −2 in 2005 according to the new WHO Growth Standards (Black et al., 2008; de Onis & Blossner, 2003), which accounted for a third of child deaths and more than 10% of the total global disease burden (Black et al., 2008). Child undernutrition has been increasingly implicated in long-term health and development problems (Barker, 2003; Harper, Susser, St. Clair, & He, 2010; Lumey, Stein, & Susser, 2011; Stein et al., 2007; Victora et al., 2008). In particular, malnutrition during the critical period of brain development of the first 1000 days (pregnancy and the first two years) may also cause permanent deficits in brain and behavioral function (Galler & Barrett, 2001; Galler, Weber, Harrison, & Ramsey, 2005; Levitsky & Strupp, 1995; Venables & Raine, 2012).

* Corresponding author. 2175 K Street, NW, Suite 200, Washington, DC 20037, USA. Tel.: +1 202 994 1015.
E-mail addresses: chenghuang@gwu.edu, pkuteach@gmail.com (C. Huang).

0277-9536/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.socscimed.2012.09.051

Please cite this article in press as: Huang, C., et al., Malnutrition in early life and adult mental health: Evidence from a natural experiment, Social Science & Medicine (2012), http://dx.doi.org/10.1016/j.socscimed.2012.09.051
As a natural experiment, famine provides a unique opportunity to test the effect of early-life severe nutritional deprivation on a series of adult outcomes including mental (Brown & Susser, 2008; Lumey et al., 2011; Stein, 1975). The earliest evidence suggested that prenatal exposure to the Dutch famine caused by a Nazi blockade of western Holland from 1944 until liberation in May 1945 was associated with an excess of congenital nervous system abnormalities, mainly neural tube defects (Lumey et al., 2011; Stein, 1975). Subsequent studies also suggested that famine exposure in early gestation likely increases adult schizophrenia risk (see a review by Susser, Hoek, & Brown, 1998). Similar associations were observed in the Chinese 1959–1961 famine. Two recent studies based on psychiatric hospital records from different sites reported a two-fold increased risk of schizophrenia among those conceived or in early gestation at the height of famine (St. Clair et al., 2005; Xu et al., 2009). Another study based on the 1987 China National Survey on Disability showed that in urban areas Chinese who were conceived, born, and raised during the famine have a higher risk of schizophrenia than both prefamine and postfamine cohorts. In rural areas, however, the famine cohorts and prefamine cohorts had a lower risk of developing schizophrenia than the postfamine cohorts (Song, Wang, & Hu, 2009), which was interpreted as the result of a more intense survival selection weeding out the frail in rural areas due to the greater severity of the famine in these areas (Song et al., 2009). Research on the Dutch famine has also suggested that early life famine exposure may increase risk of other psychiatric disorders in addition to schizophrenia in adulthood, including antisocial personality disorder (Neugebauer, Hoek, & Susser, 1995; Susser et al., 1998) and major affective disorders (Brown, Susser, Lin, Neugebauer, & Gorman, 1995; Brown, van Os, Driessens, Hoek, & Susser, 2000) as well as general symptomatic measures of mental illness that are not diagnosis-specific among offspring of famine survivors (Stein, Pierik, Verrrips, Susser, & Lumey, 2009).

In the present study, we used data from a recent epidemiological survey on mental disorders in China to investigate the potential impact of famine exposure in utero and/or in early postnatal life on the risk of a broad spectrum of mental illnesses in late adulthood. The purposes of this study were twofold. First, we estimated the potential long-term effect of the famine on mental illness among the famine survivors. Second, we tested a hypothesis that the effect of famine on mental disease is not sex-neutral; that is, it is different between men and women. Our study contributes to the literature in several ways. First, previous Chinese studies replicated the Dutch famine results for schizophrenia, leaving the potential effect of the famine on other mental disorders unexamined in the Chinese context. Famine research in various contexts is crucial for a better isolation of the independent health consequences of famine exposure and consequent nutritional deficiency from alternative mechanisms such as toxic food substitutes, which are more difficult to rule out in any single context (Lumey et al., 2011). Second, extensive literature has suggested a male vulnerability under environmental stress (Catalano et al., 2012; Trivers & Willard, 1973), in particular that famine effects might be different between men and women survivors (Mu & Zhang, 2011). We were especially interested in whether the mental health consequence of famine exposure is sex-specific.

### Background

**The Chinese Great Leap Forward Famine**

The Chinese famine occurring between 1959 and 1961 was one of the worst catastrophes in Chinese history (Ashton, Hill, Piazza, & Zeitz, 1984; Smil, 1999). Although the causes of the famine remain under debate, the weather, specifically a drought, has traditionally been blamed; this notion has been perpetuated by the Chinese, who continue to refer to it as “three years of natural disasters (san nian zi ran zai hai)” despite more evidence suggesting that policy failures played a major role (Lin, 1990; Lin & Yang, 2000). In 1958, Mao launched the Great Leap Forward Campaign, during which millions of peasants were mobilized to assist in heavy industry and to promote iron production in particular. In rural areas, mess hall communes were built, and private kitchens were prohibited. Farmers were provided free meals, and tremendous food waste was recorded during the harvest year of 1958 right before the famine (Chang & Wen, 1997; Johnson, 1998; Li & Yang, 2005; Lin & Yang, 2000; Peng, 1987; Yang, 1996; Yang & Su, 1998). Famine suddenly hit China in 1959, and grain output dropped sharply during the next three years. During the worst year—1961—only 70% of the amount produced in 1958 was reached (Chen & Zhou, 2007). Daily consumption per capita decreased dramatically to 1500 calories, far below average energy requirements of 2100 calories per day (Ashton et al., 1984; Chen & Zhou, 2007). Because of the inflexible nature of the centrally planned procurement system, the Chinese government failed to adjust the procurement and transfer of food to thousands of heterogeneous counties as a necessary response to the food shortage. Overprocurement occurred in many regions where food stores ran out, and large numbers of people subsequenly died of starvation (Meng & Qian, 2009). As a result, mortality dramatically increased and fertility rates dropped sharply during 1959–1961. The famine led to an estimated 30 million excess deaths and another 30 million lost births (Ashton et al., 1984; Chen & Zhou, 2007; Peng, 1987).

The famine affected rural areas disproportionately (Chen & Zhou, 2007; Huang, Li, Wang, & Martorell, 2010). Urban residents were less impacted by the famine because of preferential treatment by governments through a grain ration system even though rations were cut modestly during the famine years in some regions (Chen & Zhou, 2007). Dramatic regional variations in famine intensity also occurred, evident in the substantial difference in mortality rates across regions in the year 1960. The highest rates were recorded in Anhui (68.6 per thousand) and Sichuan (54.0 per thousand); the lowest rates were in provinces in Northeast China (Liaoning, Heilongjiang, and Jilin, about 11 per thousand). The variation in famine severity was more likely to be determined by the degree of willingness of local leaders to follow the radical central government policies, not local grain production. In fact, high productivity areas were more likely to experience severe famine (Meng, Qian, & Yared, 2010).

Although famines are normally accompanied by migration away from famine areas, the Chinese famine led only to mild mobility, primarily caused by the restriction imposed by the residence registration (hukou) system (Chen & Zhou, 2007; Meng & Qian, 2009). Hukou regulation was initiated in 1951, formalized, and then strictly reinforced in both cities and rural areas by 1958 (Chan & Zhang, 1999). Under the hukou regulation, every citizen was required to register one and only one permanent residency. Regulation on migration was firmly controlled by a public security system, which monitored and controlled not only the rural influx to the cities but also all intrarural and intrarural movement (Chan & Zhang, 1999).

Sex-specific famine effect and natural selection

Recent studies have shown that early-life exposure to the Chinese famine was associated with stunted growth (Chen & Zhou, 2007; Gorgens, Meng, & Vaithianathan, 2012; Huang et al., 2010; Meng & Qian, 2009) and higher risk of adult schizophrenia (St. Clair et al., 2012).
et al., 2005). Furthermore, adult female survivors with famine exposure in utero or during infancy were more likely to experience stillbirths and miscarriages than other women (Cai & Feng, 2005).

Most strikingly, famine effects were found to differ by sex, with larger long-term impacts among women. For example, exposure to famine at the early stage of life is associated with higher risk of overweight among survivors, in particular for women (Luo, Mu, & Zhang, 2006). These results replicate findings that insult in early life exhibits more harmful impacts on women (Maccini & Yang, 2009; Ravelli, van der Meulen, Osmond, Barker, & Bleker, 1999). This gender difference is usually interpreted as a result of sex-specific differences in the “scarring” effect and “selection” effect of early-life traumas (Mu & Zhang, 2011). Briefly, scarring effects can permanently impair the health of shock survivors, but robust individuals are more likely to survive a shock than their fragile counterparts (Preston, Hill, & Dreonstedt, 1998), so the scarring effects may be attenuated or canceled in cohorts that are less likely to survive the initial shock (Bozzoli, Deaton, & Quintana-Domeque, 2009; Mu & Zhang, 2011; Song et al., 2009). In addition, survival selection in response to health shocks or environmental stress is usually more severe among men than women (Noymer & Gareen, 2000). For example, higher excess mortality in males compared to females was observed in dozens of historical famines, including the 1959–1961 Chinese famine (Macintyre, 2002; Mu & Zhang, 2011); therefore, the long-term scarring effect of such shocks or stress is usually less likely to be fully observed among men compared to women (Mu & Zhang, 2011).

This stress-induced selection, in particular, sex-specific natural selection in utero, has been well discussed, but its biological underpinnings remain unclear in the literature (Catalano et al., 2012). It has been hypothesized that (a) women spontaneously terminate gestations least likely to yield grandchildren (Catalano et al., 2012; Trivers & Willard, 1973; Wells, 2000); (b) fragile or small fetuses below some critical rank on hardiness, in particular, male fetuses, may be “culled” in stressful environments (Catalano, Saxton, Bruckner, Goldman, & Anderson, 2009; Forbes, 1997) because small fetuses compared to larger fetuses of the same sex are less likely to survive harsh environment if born; and (c) small male fetuses are less likely to survive compared to female fetuses of equivalent size (Catalano et al., 2012; Wells, 2000). Such a hypothesis is supported by evidence that the ratio of males to females at birth declined among birth cohorts in gestation in the presence of ambient stressors such as economic depressions (Catalano, 2003; Catalano & Bruckner, 2005), wars (Kemkes, 2006), terrorist events (Catalano, Bruckner, Gould, Eskenza, & Anderson, 2005; Catalano, Bruckner, Marks, & Eskenza, 2006), and natural and human-made disasters (Fukuda, Fukuda, Shimizu, & Moller, 1998; Lyster, 1974) including the Chinese Great Leap Forward famine (Song, 2012).

Methods

Data

The data derived from a series of epidemiological surveys on mental disorders conducted in four provinces in China—Zhejiang, Shandong, Qinghai, and Gansu—from 2001 to 2005 (Phillips et al., 2009). Multistage stratified random sampling methods were used to identify 96 urban and 267 rural primary sampling sites from 24 counties or districts in these four provinces; the sampling frame of 113 million individuals aged 18 years or older included 12% of the adult population in China. A total of 63,004 individuals were screened with an expanded version of the General Health Questionnaire. The screening interviews, which took a mean (SD) of 25 (9) minutes to complete, were done face to face in respondents’ homes by 170 psychiatric nurses who had been trained for seven to 10 days before the main study commenced. The local institutional review board at each research site provided ethics approval for the study. Extensively pilot-tested at the research sites, the screening instrument included the 12-item General Health Questionnaire (GHQ-12) as well as eight additional items that assessed the presence of other risk factors for mental illness, including current very poor physical health, very poor psychological health, frequent obsessive thoughts or compulsive behaviors, frequent restriction of behavior because of phobias, frequent feelings of extreme nervousness or anxiety, social problems resulting from drinking in the prior year, any previous treatment for psychological problems, and any previous suicidal ideation or suicidal behavior. The 12-item General Health Questionnaire was developed to screen for nonspecific psychiatric morbidity and has been widely used and translated into many languages (Hankins, 2008). Its validity and reliability have been found to be acceptable as a screening instrument for psychiatric disorders in various nonclinical settings (Goldberg & Williams, 1988; Hu, Stewart-Brown, Twigg, & Weich, 2007; Yang, Huang, & Wu, 2003), but its utility in clinical psychometrics is questioned as limited (Hankins, 2008). The literature suggests that the GHQ combined with the assessment of additional risk factors was a good proxy for risk of current mental illness (Phillips et al., 2009; Yang & Link, 2009).

We included all screened individuals in rural areas born from 1956 to 1963 (see Table 1, n = 4972), among whom the 1956–1958 birth cohorts were exposed to the famine during postnatal life; the 1959–1961 birth cohorts were exposed in utero as well as during early postnatal period. For example, at the mid-year point, the 1957 cohort would have been exposed from 1.5 to 4.5 years and the 1959 cohort late in pregnancy and from birth to 2.5 years. The majority of the 1962 birth cohort, that is, those who were conceived during 1961, was exposed in utero but not postnatal life; and the 1963 birth cohort had no famine exposure and was used as reference
group in the main analysis (The 1964 birth cohort and the combined 1963–1964 birth cohorts were used as alternative reference group in the additional robustness analysis). We restricted the main analysis to a relatively short period to avoid other major natural disasters before the famine years, such as the extremely cold winter of 1954–1955 (Ding, Jia, Wang, Chen, & Chen, 2009) and the political chaos of the Chinese Cultural Revolution of 1966–1976. We also restricted the main analysis to rural subjects born in these years because food security in urban areas was guaranteed by the rationing system even during the famine years, and urban residents were therefore much less likely to have experienced the famine compared to rural residents (Chen & Zhou, 2007; Huang et al., 2010). The final sample size was large enough for sufficient power to detect even a “small” (effect size = 0.2) famine effect (Cohen, 1988).

Assessment of famine severity

We assessed famine severity in the 24 counties included in this survey by comparing the size of the famine cohorts relative to the surrounding non-famine cohorts in the 1% sample of China’s 1990 population census (https://international.ipums.org/international/), following the strategy employed by a previous study (Huang et al., 2010). The underlying assumption is that the smaller the size of the famine cohorts relative to the surrounding cohorts because of reduced fertility and/or increased mortality, the greater the severity of the famine (Huang et al., 2010). We calculated the mean size of cohorts born during the three years immediately before the famine (1956–1958) and the three years immediately after the famine (1962–1964), labeled \( N_{\text{nonfam}} \); we also calculated the mean size of cohorts born during the famine years (1959–1961), labeled \( N_{\text{famine}} \). We then calculated the cohort size shrinkage index (CSSI) as the difference between \( N_{\text{nonfam}} \) and \( N_{\text{famine}} \), divided by \( N_{\text{famine}} \). Among the 24 counties, CSSI ranged from 0.27 to 0.65 with a mean of 0.51 and standard deviation of 0.12, indicating that on average the cohorts alive in 1990 who were born during the three years of the famine were 51% smaller than the cohorts born in the three years immediately preceding or following the famine in these counties. Further analysis suggested that this cohort size shrinkage index generated at the province level (Table 2) is highly correlated \((r = 0.85; p < 0.001)\) with provincial excess mortality-based index, which contrasts the excess mortality in famine years to nonfamine years, a famine intensity indicator employed in a previous study (Chen & Zhou, 2007). In addition, the 1988 National Survey of Fertility and Contraception, in which about a half million ever-married women were interviewed on complete pregnancy/birth history, suggested that live births during the three famine years were significantly lower than the surrounding years and that the fertility-reduction index is also highly correlated with the cohort size shrinkage index at the province level \((r = 0.90; p < 0.001)\). Because the 1988 National Survey of Fertility and Contraception is representative at the province level and the excess mortality rate was available at the province level, we were unable to further test the correlation of these three indices at more disaggregated level, neither could we include the other two indices to the present analyses that examined ecological famine exposure at the county level and its impact on mental disorders. Nevertheless, the high correlations across indices at the province level based on different data sources suggest an acceptable validity and reliability of the cohort size shrinkage index as a single famine intensity indicator.

Statistical model

Only individuals conceived during or before the 1959–1961 famine were exposed to the famine, and the intensity of famine exposure depended on the famine severity in the region where the individual was born. We used the variation of famine exposure across cohorts and regions to construct a Difference-in-Difference (DID) estimator (Chen & Zhou, 2007; Huang et al., 2010). We estimated separate models for the two different mental health outcomes.

The first outcome was based on the overall GHQ-12 score. Each GHQ item was scored 1–4 with scores of 3 or 4, indicating that the symptom was present to a moderate to severe degree. The GHQ-12 was used as a unidimensional measure because our exploratory factor analyses (EFA) results indicated a one-factor solution that explained 65% of the variance with each item loading at above 0.4. When computing the overall score, item scores of 1 or 2 were coded absent (scored 0) and item scores of 3 or 4 were coded as present (scored 1), so the total score ranges from 0 to 12, which had a negative binomial distribution. We fit a negative binomial regression with the DID estimator as shown below.

\[
\log (E(Y_{tk} | CSSI_r, Cohort_k)) = B_0 + a_r \text{County}_r + \delta_k \text{Cohort}_k \\
+ \gamma \text{CSSI}_r + \sum_{k=1956}^{1982} \beta_k (\text{CSSI}_r \times \text{Cohort}_k)
\]

where \( Y_{tk} \) is the GHQ-12 score for an individual who was born in county \( r \) and year \( k (k = 1956–1962, \text{and } 1963 \text{ is the reference year}) \), \( a_r \) represents the county fixed effect, \( \delta_k \) represents the cohort fixed effect. \( \text{CSSI}_r \) refers to the cohort size shrinkage index in county \( r \), and \( \beta_k \) represents the coefficient of the interaction between the cohort size shrinkage index and the cohort dummy variables (CSSI \( \times \) COHORT).

### Table 2

<table>
<thead>
<tr>
<th>Province</th>
<th>Cohort size shrinkage index</th>
<th>Excess mortality index</th>
<th>Fertility reduction index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td>0.22</td>
<td>1.67</td>
<td>0.30</td>
</tr>
<tr>
<td>Tianjin</td>
<td>0.33</td>
<td>1.80</td>
<td>0.26</td>
</tr>
<tr>
<td>Hebei</td>
<td>0.34</td>
<td>3.12</td>
<td>0.36</td>
</tr>
<tr>
<td>Shanxi</td>
<td>0.26</td>
<td>0.95</td>
<td>0.23</td>
</tr>
<tr>
<td>Nei Mongol</td>
<td>0.2</td>
<td>0.47</td>
<td>0.22</td>
</tr>
<tr>
<td>Liaoning</td>
<td>0.32</td>
<td>5.55</td>
<td>0.31</td>
</tr>
<tr>
<td>Jilin</td>
<td>0.23</td>
<td>2.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>0.19</td>
<td>1.75</td>
<td>0.14</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.22</td>
<td>0.62</td>
<td>0.25</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>0.42</td>
<td>5.10</td>
<td>0.36</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>0.35</td>
<td>1.88</td>
<td>0.36</td>
</tr>
<tr>
<td>Anhui</td>
<td>0.63</td>
<td>21.97</td>
<td>0.62</td>
</tr>
<tr>
<td>Fujian</td>
<td>0.37</td>
<td>3.68</td>
<td>0.33</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>0.35</td>
<td>2.37</td>
<td>0.30</td>
</tr>
<tr>
<td>Shandong</td>
<td>0.4</td>
<td>7.87</td>
<td>0.37</td>
</tr>
<tr>
<td>Henan</td>
<td>0.47</td>
<td>10.22</td>
<td>0.46</td>
</tr>
<tr>
<td>Hubei</td>
<td>0.42</td>
<td>5.02</td>
<td>0.38</td>
</tr>
<tr>
<td>Hunan</td>
<td>0.54</td>
<td>8.80</td>
<td>0.48</td>
</tr>
<tr>
<td>Guangdong</td>
<td>0.35</td>
<td>3.37</td>
<td>0.29</td>
</tr>
<tr>
<td>Guangxi</td>
<td>0.46</td>
<td>10.90</td>
<td>0.31</td>
</tr>
<tr>
<td>Sichuan</td>
<td>0.61</td>
<td>28.63</td>
<td>0.52</td>
</tr>
<tr>
<td>Guizhou</td>
<td>0.52</td>
<td>16.38</td>
<td>0.42</td>
</tr>
<tr>
<td>Yunnan</td>
<td>0.38</td>
<td>3.15</td>
<td>0.37</td>
</tr>
<tr>
<td>Shanxi</td>
<td>0.25</td>
<td>0.13</td>
<td>0.26</td>
</tr>
<tr>
<td>Gansu</td>
<td>0.45</td>
<td>10.48</td>
<td>0.44</td>
</tr>
<tr>
<td>Qinghai</td>
<td>0.52</td>
<td>12.65</td>
<td>0.55</td>
</tr>
<tr>
<td>Ningxia</td>
<td>0.41</td>
<td>2.60</td>
<td>0.52</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>0.22</td>
<td>2.63</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Note: Each of these indices was generated in a similar way, which was discussed in context under the section entitled Assessment of Famine Severity. Source: The excess mortality rate is from Lin and Yang (2000), fertility information is from the 1988 National Survey of Fertility and Contraception, and the cohort size has been generated from the 1990 Census of Chinese Population.
and is a measure of the impact of famine exposure on GHQ scores. A detailed discussion of estimating “treatment” effect using the interaction term in nonlinear difference-in-difference models has been presented elsewhere (Athey & Imbens, 2006; Puhani & Sonderhof, 2010). To estimate the average effect across all 24 counties, we multiplied the interaction coefficient by 0.51 (the mean CSSI across all counties).

The second outcome variable, a more comprehensive measure of the risk of mental illness, combined GHQ scores and eight risk factors. Additional risk for mental illness (i.e., beyond that measured by GHQ) was considered present if the respondent reported any of the eight additional risk behaviors. All respondents were classified into three strata for mental illness: low risk, moderate risk, and high risk, following the screening practice in the survey (Phillips et al., 2009). The high-risk stratum included a GHQ score of at least 4 (on a 0–12 point scale) or any of the eight risk factors. Individuals with moderate risk had no risk factors and had a GHQ score of 1–3. Those with low risk had no risk factors and had a GHQ score of 0. We fitted ordered logit models for the risk of mental illness with the difference-in-difference estimator presented above and with the risk categorized as low, moderate, and high. The proportional odds assumption was met by the score test ($p > 0.70$).

For all the models, males and females were analyzed separately. Confidence intervals were adjusted for clustering by county. All analyses were conducted using SAS, Version 9.

### Results

The characteristics of our study sample (rural survey participants born from 1956 to 1963) are presented by birth year and sex in Table 1. Both women and men reported relatively low (i.e., healthy) GHQ scores across birth years. In general women had higher mean GHQ scores than men ($P < 0.05$), ranging from 0.77 (1963 birth cohort) to 1.15 (1960 birth cohort) among women and from 0.58 (1963 birth cohort) to 0.86 (1959 birth cohort) among men. Among those born from 1956 to 1963, about 68.4% of the women and 72.7% of the men scored zero for the GHQ. About two thirds of the sample was classified as being at low risk of current mental illness.

Estimates of average famine impact by birth year on GHQ scores are shown in Table 3. With the exception of women in the 1957 birth cohort, famine exposure was associated with increased mean GHQ scores in women and with decreased mean GHQ scores in men. In women, the famine had a significant impact on the GHQ scores among those in the 1959 birth cohort; these women would have had a significantly lower mean GHQ score (lower by 1.52 points; 95% CI: 0.42, 2.63) in the absence of famine. The effect size of this estimated impact is 0.84 (1.52 divided by 1.79, the standard deviation of GHQ scores among the 1959 cohort women), which is considered a large effect according to Cohen’s effect size criteria (Cohen, 1988). In contrast, famine exposure among men is significantly associated with lower mean GHQ scores among the 1960 birth cohort ($-0.99$; 95% CI: $-1.60$, $-0.37$; effect size = 0.70) and 1961 birth cohort ($-1.14$; 95% CI: $-2.21$, $-0.06$; effect size = 0.63). The differences in the coefficients between men and women are statistically significant.

The odds ratios for risk of mental illness based on the expanded three-level measure of risk in the famine-exposed birth cohorts compared to the unexposed 1963 birth cohorts are shown in Table 4. Consistent with findings for the GHQ scores, with the exception of the women in the 1957 cohort, all the ORs were greater than one for female cohorts and less than one for male cohorts; but only a few of the odds ratios were statistically significant. Women in the 1959 birth cohort were 4 times more likely to be at risk of mental disorder in adult life than women born in 1963 ($P < 0.01$). The only male birth cohort with statistically significant decreased risk of mental illness compared to the 1963 cohort was the 1962 cohort (OR = 0.34; 95% CI: 0.14–0.80; $P = 0.01$).

We did a post-hoc secondary analysis where individuals born from 1956 to 1958 were combined into a prefamine cohort, individuals born from 1959 to 1961 were combined into the famine cohort, and the 1962 cohort was considered the postfamine cohort; these three groups were then compared with the unexposed 1963 birth cohort. Results from these analyses with GHQ scores and the expanded three-level measure of risk as outcomes variables suggest that famine exposure increased the mean GHQ score by 0.95 points (95% CI: 0.26, 1.65; $P = 0.01$; effect size = 0.51) and increased the risk of mental disorder (OR = 2.88; 95% CI: 1.23, 6.39; $P = 0.01$) among women in the famine cohorts (i.e., women born during the famine), but it decreased the mean score of GHQ by 0.89 points (95% CI: $-1.59$, $-0.20$; $P = 0.01$; effect size = 0.54). In men, the risk of mental disorder also tended to decrease but the association was not statistically significant (OR = 0.60; 95% CI: 0.26, 1.24; $P = 0.24$).

In addition, the robustness analyses suggested that when the 1964 birth cohort or the combined 1963–1964 birth cohort was used as the reference group, the results were highly similar. Further robustness analysis on the urban sample did not reveal any significant association between famine exposure and mental disorder outcomes.

### Discussion

We studied relationships between early-life exposure to the Chinese famine and risk of mental illness in mid- to late-adulthood among Chinese born from 1956 to 1963 in four provinces, and we

Table 3 Estimated coefficients of famine effect on scores of GHQ-12, based on negative binomial regression with difference-in-difference estimator.

<table>
<thead>
<tr>
<th>Year</th>
<th>Women Estimated change in mean score*</th>
<th>95% CI</th>
<th>Men Estimated change in mean score*</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>0.81</td>
<td>(−0.08, 1.69)</td>
<td>−0.50</td>
<td>(−1.42, 0.42)</td>
</tr>
<tr>
<td>1957</td>
<td>−0.35</td>
<td>(−1.47, 0.77)</td>
<td>−0.13</td>
<td>(−1.07, 0.82)</td>
</tr>
<tr>
<td>1958</td>
<td>0.55</td>
<td>(−0.36, 1.47)</td>
<td>−0.56</td>
<td>(−1.32, 0.20)</td>
</tr>
<tr>
<td>1959</td>
<td>1.52</td>
<td>(0.42, 2.63)</td>
<td>−0.50</td>
<td>(−1.95, 0.96)</td>
</tr>
<tr>
<td>1960</td>
<td>0.68</td>
<td>(−0.20, 1.56)</td>
<td>−0.99</td>
<td>(−1.60, −0.37)</td>
</tr>
<tr>
<td>1961</td>
<td>0.69</td>
<td>(−0.52, 1.90)</td>
<td>−1.14</td>
<td>(−2.21, −0.06)</td>
</tr>
<tr>
<td>1962</td>
<td>0.63</td>
<td>(−0.22, 1.48)</td>
<td>−0.84</td>
<td>(−1.70, 0.03)</td>
</tr>
</tbody>
</table>

Reference group is the 1963 birth cohort. Method of computing the estimated effect of famine exposure on mean GHQ scores is described in the Methods Section. Higher GHQ-12 scores represent greater risk of mental illness.

Table 4 Estimated odds ratios of risk of mental illness predicted by famine exposure, based on ordered logit regression with difference-in-difference estimator.

<table>
<thead>
<tr>
<th>Year</th>
<th>Women OR</th>
<th>95% CI</th>
<th>Men OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>1.95</td>
<td>(0.80, 4.76)</td>
<td>0.81</td>
<td>(0.25, 2.60)</td>
</tr>
<tr>
<td>1957</td>
<td>0.86</td>
<td>(0.26, 2.84)</td>
<td>0.68</td>
<td>(0.23, 2.02)</td>
</tr>
<tr>
<td>1958</td>
<td>1.48</td>
<td>(0.52, 4.22)</td>
<td>0.54</td>
<td>(0.18, 1.57)</td>
</tr>
<tr>
<td>1959</td>
<td>4.99</td>
<td>(1.68, 14.84)</td>
<td>0.69</td>
<td>(0.19, 2.53)</td>
</tr>
<tr>
<td>1960</td>
<td>2.24</td>
<td>(0.71, 7.05)</td>
<td>0.55</td>
<td>(0.22, 1.37)</td>
</tr>
<tr>
<td>1961</td>
<td>1.62</td>
<td>(0.54, 4.11)</td>
<td>0.65</td>
<td>(0.18, 2.32)</td>
</tr>
<tr>
<td>1962</td>
<td>2.34</td>
<td>(0.98, 5.59)</td>
<td>0.34</td>
<td>(0.14, 0.80)</td>
</tr>
</tbody>
</table>

Reference group is the 1963 birth cohort. Method of computing the estimated effect of famine exposure on the expanded three-level measure of risk of mental disorders is described in the Methods Section.
also tested the hypotheses that famine effects are different between men and women survivors based on existing evidence. We found significant and strong associations between famine exposure and increased risk of mental disorders among women born during the famine. In contrast, famine exposure came with an apparent reduced risk of mental disorder among men later in their lives. These findings are consistent with previous studies that reported enduring effects of early-life shocks in women but not men (Luo et al., 2006; Maccini & Yang, 2009; Mu & Zhang, 2011; Ravelli et al., 1999; Yang et al., 2008). We speculate that the observed sex-specific famine effect is related to “culled” cohort argument, which posits that environment stress may in particular cull weaker males in utero and leave more robust male births, who may exhibit some health advantages later in life despite of insult in early life (Catalano & Bruckner, 2006).

Early-life famine effects on adult females may also be attenuated by selection but not to the point of completely masking the scarring (Stein et al., 1998) as well as specific mental disorder examined (Brown et al., 2000). For example, early gestational famine may increase risk of schizophrenia, and later gestational famine may give rise to affective disorder in particular (Brown et al., 2000).

The long-term effects of early-life nutrition on physical health are well recognized, and this study is an important addition to the accumulating body of evidence that early-life nutrition also has a profound, life-long effect on neurological development and psychological health (Galler & Barrett, 2001; Galler et al., 2005; Levitsky & Strupp, 1995; Lumey et al., 2011; Roth et al., 2011; Venables & Raine, 2012), but our findings are limited to women. Previous studies suggested that neurodevelopmental disorders may be “programmed” by early life stress exposures, which may alter brain development and influence fetal brain growth through epigenetic modifications (Bale et al., 2010), a way similar to early life programming of adult chronic diseases including diabetes (Barker, 2003). With respect to famine effects, famine exposure may result in nutritional deficiency in both micro and macronutrients, which may in turn have direct effects on the risk of developing mental disorders. Severe prenatal micronutrient deficiency in tolate and iodine results in neural tube defects and cretinism, respectively, but effects from moderate deficiency on growth and development are also evident. Prenatal protein—energy malnutrition influences the development of brain structure, such as the hippocampus, and impairs the function of neurotransmitters, such as dopamine and serotonin (Brown et al., 2000). These findings should be interpreted with caution. A major drawback this study shares with all famine studies is that it sheds little light on the mechanisms underlying the associations between famine exposure and mental illness. Nutritional deficiency caused by famine is most likely the mechanism underlying the observed association between famine exposure and increased risk of mental disorders. Although alternative mechanisms, such as the psychogenic effect of experiencing famine and toxic food substitute (Lumey et al., 2011; Xu et al., 2009), may not be entirely excluded, they are less likely to play an important role given the consistency of finding across various contexts (Lumey et al., 2011). Second, unlike the Dutch famine with a short and clear starting and ending point, which therefore allows a linkage between famine exposure at different stages of gestation to adult mental disorders (Susser, St. Clair, & He, 2008), the Chinese famine lasted three years and its duration varies slightly across regions, which makes it impossible to define birth cohorts exposure to famine at specific period of gestation but allows an examination of consequences of famine exposure and malnourishment in a broadly defined “critical period,” that is, pregnancy and the first two years of life. In addition, in a natural selection argument for the sex-specific famine effect, some mechanism was assumed in this study that women spontaneously terminate fetuses, in particular male fetuses with some traits that may be related to adult risk of mental disorders if born and survive to a certain age. A more careful test of such an assumption merits further research. Finally, we are unsure whether our results represent famine effects in places other than the four studied provinces—Zhejiang, Shandong, Qinghai, and Gansu; neither can we directly assess the potential effect of selective out-migration on our findings, but evidence suggests that this effect is not likely to be important. Out-migration was low during the famine and up to the 1990s (Chen & Zhou, 2007), and since then, the migrants have primarily been labor migrants characterized as single and young with a mean age of mid-20s, according to a national survey conducted around 2002, when the survey for the present study was conducted (Tao, 2008). Moreover, our findings about the effect of the Chinese Great Leap Forward famine may not necessarily suggest similar psychological consequences of other famines and ambient stressors. Nevertheless, our findings about the deleterious effect of famine exposure in utero and in infancy along with similar findings from other famine studies underscore the importance of early-life nutrition in improving physical and psychological health over the entire lifetime. Our findings about the larger long-term effects of early-life insult among women relative to men also merit further research, in particular on potential mechanisms, including sex-specific natural selection in utero.

Contributors

CH and RM designed the study. MRP, YZ, JZ, QS, ZS, ZD, and SP contributed research data to the study. CH, RM, RZ, and MRP contributed to data analysis and interpretation. CH drafted the report, and all co-authors critically revised for significant scientific content, and have seen and approved the final version of the report.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

This research was funded by a grant from the Global Health Institute of Emory University and a grant from National Institutes of Health (1R03HD072104-01). The funders had no role in the conduct of the study.

References


