ACCSCIENCE PUBLISHING

ISSN: 2424-8150 (Print)
ISSN: 2424-8606 (Online)
Volume 8 · Issue 1
June 2022

INTERNATIONAL JOURNAL OF POPULATION STUDIES

International Journal of Population Studies

Print ISSN: 2424-8150

Online ISSN: 2424-8606

International Journal of Population Studies (IJPS) is an open access, multidisciplinary journal that publishes high quality original research and timely reviews of recent advances and emerging issues in population processes; dynamics of fertility, mortality, and migration; and linkages with socioeconomic and environmental change across times, spaces, and cultures. The journal aims to provide a platform for researchers worldwide to promote and share cutting-edge knowledge and advances in different areas of population research. Article formats include editorials, research articles, review articles, letters to editors, commentaries, perspectives, reports, and book reviews that address demography and population-related issues. The journal also offers special issues arising from conferences and other meetings.

About the Publisher

AccScience Publishing is a publishing company based in Singapore. We publish a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of disciplines.

Contact Us

Managing Editor
ijps.office@accscience.sg

AccScience Publishing
9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619.

Volume 8 • Issue 1 • June 2022
ISSN 2424-8150 (print) ISSN 2424-8606 (online)

INTERNATIONAL JOURNAL OF POPULATION STUDIES

Editor-in-Chief
Danan Gu
United Nations, New York, United States

ACCSCIENCE PUBLISHING
Access Science Without Barriers

Full issue copyright © 2022 AccScience Publishing
All rights reserved. Without permission in writing from the publisher, this full issue publication in its entirety may not be reproduced or transmitted for commercial purposes in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system. Permissions may be sought from ijps.office@accscience.sg.

Article copyright © Respective Author(s)
See articles for copyright year. All articles in this full issue publication are open-access. There are no restrictions in the distribution and reproduction of individual articles, provided the original work is properly cited. However, permission to reuse copyrighted materials of an article for commercial purposes is applicable if the article is licensed under Creative Commons Attribution-NonCommercial License. Check the specific license before reusing.

International Journal of Population Studies

ISSN: 2424-8150 (print)

ISSN: 2424-8606 (online)

Editorial and Production Credits

Publisher: AccScience Publishing
Managing Editor: Alicia Tian
Production Editor: Sharmila Velapasamy
Special Issue Commissioning Editors: Alicia Tian, Eva Liu
Article Layout and Typeset: Sinjore Technologies (India)
Cover Design: ProPub (China)

For all advertising queries, contact ijps.office@accscience.sg.

Supplementary file

Supplementary files of articles can be obtained at https://accscience.com/journal/IJPS/8/1.

Disclaimer

AccScience Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. AccScience Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Population Studies

Editorial Board

Editor-in-Chief

Danan Gu
Population Division, Department of Economic and Social Affairs, United Nations, New York, United States

Deputy Editor

Qiushi Feng
Department of Sociology and Anthropology, National University of Singapore, Singapore

Associate Editor

Hans-Peter Blossfeld
Graduate Centre Trimberg Research Academy (TRAc), Otto-Friedrich-Universität Bamberg, Bamberg, Germany

Editorial Board Members*

Huda Alkitkat, *Egypt*
Luciana Correia Alves, *Brazil*
Elena Ambrosetti, *Italy*
Xue Bai, *China*
Pau Baizan, *Spain*
Federico Benassi, *Italy*
Gabriel Mendes Borges, *Brazil*
Tianji Cai, *Macau*
Cassandra D. Chaney, *USA*
Huashuai Chen, *China*
Wei Chen, *China*
Zhenxiang Chen, *Canada*
Kailash Chandra Das, *India*
Gustavo De Santis, *Italy*
Dusan Drbohlav, *Czech Republic*
Sonja Drobnič, *Germany*
Matthew E. Dupre, *USA*
Viviana Egidi, *Italy*
Ann Evans, *Australia*
Zhixin Feng, *China*
Fabiane Ribeiro Ferreira, *Brazil*
Mirjam Fischer, *Germany*
Angel M. Foster, *Canada*
Yuanyuan Fu, *China*
Elizabeth Fussell, *USA*
Víctor Manuel García-Guerrero, *Mexico*
Vasilis Gavalas, *Greece*
Cecilia Gayet, *Mexico*
Ashley Larsen Gibby, *USA*
Cristina Giudici, *Italy*
Raphael Mendonça Guimarães, *Brazil*
Monica Das Gupta, *USA*
Masa Higo, *Japan*
Quanbao Jiang, *China*
Aramide Kazeem, *USA*
Caroline Krafft, *USA*
David P. Lindstrom, *USA*
Daniel Lois, *Germany*
Rikiya Matsukura, *Japan*
Goran Miladinov, *Macedonia*
Komanduri S. Murty, *USA*
Rangasamy Nagarajan, *India*
Lorretta Ntoimo, *Nigeria*
Livia Olah, *Sweden*
José Antonio Ortega, *Spain*
John Lekan Oyefara, *Nigeria*
Neir Antunes Paes, *Brazil*
Sangram Kishor Patel, *India*
Yaolin Pei, *USA*
Gina Potarca, *Switzerland*
Chiara Daniela Pronzato, *Italy*
Amany Refaat, *Egypt*
Rosa María Aisa Rived, *Spain*
Gabriele Ruiu, *Italy*
Luule Sakkeus, *Estonia*
Ivett Szalma, *Hungary*
Md. Ismail Tareque, *Bangladesh*
David B. Ugal, *Nigeria*
Eunice Danitza Vargas Valle, *Mexico*
Kun Wang, *USA*
Ning Wang, *China*
Senhu Wang, *Singapore*
Philippe Wanner, *Switzerland*
Tom Wilson, *Australia*
Hongwei Xu, *USA*
Fang Yang, *China*
Na Yin, *USA*
Haiyan Zhu, *USA*

ACCSCIENCE PUBLISHING

Volume 8 | Issue 1 | June 2022 | Page 1 – 101 |

International Journal of Population Studies

https://accscience.com/journal/IJPS

CONTENTS

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Validity and reliability of Mini-Mental State Examination in Older Adults in China: Inline Mini-Mental State Examination with cognitive functions

Teck Kiang Tan[1]*, and Qiushi Feng[2]

[1]Institute for Applied Learning Sciences and Educational Technology, National University of Singapore, Singapore

[2]Department of Sociology, Centre for Family and Population Research, National University of Singapore, Singapore

Abstract

The main aim of the study is to validate the factor structure of the Mini-Mental State Examination (MMSE) of China's older population using the Chinese Longitudinal Healthy Longevity Survey. The validation process used the exploratory factor analysis (EFA) to determine the number of dimensions of MMSE, the confirmatory factor analysis (CFA) to confirm the factorial structure of MMSE, and the factorial invariance to conclude the factor structure does not differ between the young-old (aged 65 – 79) and old-old (aged 80 or older). The results of the EFAs suggested two possible factor structures: A six-factor and a seven-factor solution. The seven-factor confirmatory factor model turned out as the best fit by comparison to the four competing confirmatory models. Strict factorial invariance was attained for the two age groups, indicating a high level of measurement equality, a property of invariance was seldom achieved in the literature of factorial invariance studies. In comparison to the MMSE literature that focused solely on EFA that aims to establish a single summated score, the present study suggests using EFA, CFA, and factorial invariance that takes into consideration of measurement errors as the preferred procedure since it establishes the appropriate MMSE dimensionality that is in line with their respective cognitive functions.

Keywords: Mini-Mental State Examination; Validation; Exploratory factor analysis; Confirmatory factor analysis; Factorial invariance

***Corresponding author:**
Teck Kiang Tan
(Singapore.alsttk@nus.edu.sg)

Citation: Tan TK, Feng Q. (2022). Validity and reliability of Mini-Mental State Examination in Older Adults in China: Inline Mini-Mental State Examination with cognitive functions. *International Journal of Population Studies*, 8(1):1-15. https://doi.org/10.36922/ijps.v8i1.1285

Received: February 23, 2022

Accepted: May 17, 2022

Published Online: June 08, 2022

Copyright: © 2022 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The cognitive function represents an important health dimension of older adults, which significantly affects their physical health, well-being, and mortality (Hsieh, Wu, Wang, *et al.*, 2021; Langa, Llewellyn, Lang, *et al.*, 2009; Park, Kwon, Jung, *et al.*, 2012; van der Meulen, Irven, Bakunina, *et al.*, 2021). The Mini-Mental State Examination (MMSE) was first introduced by Folstein *et al.* (1975) more than 40 years ago to measure cognitive performance. Since then, MMSE was widely used and has been accepted as an overall measure of cognitive impairment in clinical, research, and community settings. The

MMSE is nowadays a widely used test of cognitive function among older adults. Its usability and applicability were well noted in social and medical research as a short screening tool, especially for older adults at risk of mild cognitive impairment (Folstein, Folstein, and McHugh, 1975; Lezak, Howieson, and Loring, 2004), screening the risk of dementia and Alzheimer's disease (Burke, Grudzien, Burgess, *et al.*, 2021; Mitchell, 2009; Arevalo-Rodriguez, Smailagic, Roqué-Figuls, *et al.*, 2021; Shigemori, Ohgi, Okuyama, *et al.*, 2010).

The MMSE specifies a list of cognitive domain functions as an inventory that contains items to measure them which include orientation, registration, attention, calculation, spelling, recall, delayed recall, naming, repetition, verbal and written comprehension, and visuospatial capability (Davey and Jamieson, 2004; Folstein, Folstein, and McHugh, 1975; Reilly, Challis, Burns *et al.*, 2004; Shulman, Herrmann, Brodaty, *et al.*, 2006). The past studies that investigated the factorial structure of MMSE remained unsettled about its dimensionality. These studies did not show the alignment between the domain functions to the MMSE items about the dimensionality it is supposed to form and normally ended up with a lower dimension, not in line with the theoretically expected dimension (Folstein, Folstein, and McHugh, 1975; Park, Kwon, Jung, *et al.*, 2012; Shigemori, Ohgi, Okuyama, *et al.*, 2010; Tinklenberg, Brooks, Tanke, *et al.*, 1990). The first study on MMSE factorial structure in the literature indicated a two-factor structure with the first factor including attention/concentration, language, and constructional praxis, and the second comprising time-space orientation and delayed recall (Fillenbaum, Heyman, Wilkinson, *et al.*, 1987). Three years later, similar results of a two-factor solution were established with the first factor including writing, naming, immediate memory, reading a sentence, and verbal comprehension, and the second factor including constructional praxis, delayed recall, temporal orientation, attention/concentration, and spatial orientation (Tinklenberg, Brooks, Tanke, *et al.*, 1990). The non-alignment of the MMSE items to the cognitive functions was clearly shown by these two studies that the MMSE items within the two factors differed. The number of dimensions increased in the later year publication to three factors (Shigemori, Ohgi, Okuyama, *et al.*, 2010). While the number of factors was not a straightforward solution empirically, the number of dimensions also differed within a study between the control and experimental groups (Baek, Kim, Park, *et al.*, 2016), and some ignored the determination of dimensionality procedure and directly used the domain MMSE summated subscores (e.g., Park, Kwon, Jung, *et al.*, 2012).

The exploratory factor analysis (EFA) was the main approach in the MMSE validation literature to determine the number of factors for the MMSE inventory and the main statistics used was the percentage of variance explained (e.g., Shigemori, Ohgi, Okuyama, *et al.*, 2010). This statistic explains the level of variances that could be extracted from an MMSE inventory, however, it does not point out whether the number of dimensions is statistically and theoretically appropriate. Choosing a higher level of variance explained will end in a higher dimension and vice versa. This statistic easily allows an arbitrary decision in determining the dimension. Theoretically, a prior expectation about the dimensionality should be decided before starting a study such as recall and visuospatial capability which are most probably two separate factors. Using the appropriate and developed statistical procedures that have been established in the measurement literature to determine the dimensionality of MMSE were also absent from the MMSE validation literature. These developments include the Velicer's MAP criteria (Lim, Li, Xie, *et al.*, 2019; Velicer and Jackson, 1990), the Horn's parallel analysis (Dinno, 2009; Garrodp, Abad, and Ponsoda, 2013; Hayton, Allen, and Scarpello, 2004; Horn, 1965), the very simple structure (VSS) criterion (Revelle and Rocklin, 1979), various information criteria, and approaches such as BIC and sample adjusted BIC (Schwartz, 1978; Sclove, 1987).

EFA to establish the number of dimensions is the preliminary step in a validation process. Further incorporating confirmatory factor analysis (CFA) and the factorial invariance test into the validation procedure were already established in the fields of measurement, psychology, and education (Bollen and Lennox, 1991; Brown, 2006; Liau, Chow, Tan, *et al.*, 2010; Lane, Anderson, Ponce, *et al.*, 2012) and acknowledged as part of the validation process. That is, after EFA, CFA should be carried out to confirm the structure of an inventory. Providing the fit of a hypothesized CFA model and comparing it with competing CFA models are part of the validation procedure to ensure that the best fitted CFA is chosen (Brown, 2006; Jackson, Gillaspy, and Purc-Steogebsib, 2009; Liau, Chow, Tan, *et al.*, 2010). The factorial invariance further confirms if the chosen CFA model is applicable in various controlled conditions. The reliability and validity of measurements on the health of older adults were constantly a concern (Gu, 2005). Due to the major health disparity between young-old (aged 65 – 79 years old) and oldest-old (aged over 80 years old), it is desirable to check the factorial invariance of these two major age groups. The factorial invariance testing was carried out in this study to make sure the MMSE psychometric properties for both the young-old and old-old age groups were invariant. After the best fitted confirmatory factor model was chosen, four factorial invariance tests were carried out to ensure there is no inequality in their factor structure and form despite

their declining health conditions (Putnick and Bornstein, 2016; Revuelta, Franco-Martinez, and Ximénez, 2021; Schürer, van Ophuysen, and Behrmann, 2021).

The main purpose of this study is to find out the MMSE factorial structure of China's older population using the Chinese Longitudinal Healthy Longevity Survey (CLHLS). While the reliability and validity of MMSE have been carried out for each wave of the CLHLS using EFA (Gu, 2005; CLHLS, 2020), the present paper is to reveal the factor structure of MMSE from a more refined methodological base. Using the most recent 2018 wave of the CLHLS, this study carried out EFA and CFA, further examined the factorial invariance for the young-old and old-old, and reported reliability levels with the consideration of the binary nature of the data, using an appropriate estimation method.

2. Data and methods

2.1. Data sources

The data were extracted from the CLHLS, which is a national longitudinal survey of Chinese older adults. The CLHLS started in the year 1998 and accomplished a total of eight waves in 1998, 2000, 2002, 2005, 2008/2009, 2011/2012, 2014, and 2018/2019. More details about the survey were published in the literature (Gu *et al*., 2021). The present study used the most recent wave in the year 2018/2019. The 2018/2019 wave had 15,896 older adults. For the purposes of this study, 7603 respondents were dropped for missing values in any of the MMSE items, which led to a sample size of 8293 older adults. As the sample size is large after deletion, it has little effect on the validation results. All participants involved in the survey consented to take part in the survey, including answering the MMSE items. The sample was further divided into the young-old aged 65 – 79 and the oldest-old aged 80+, with the former consisting of 4751 and the latter having 4118 persons.

2.2. Measures for MMSE inventory

The MMSE inventory contained 23 items (Folstein, Folstein, and McHugh, 1975). The CLHLS developed a Chinese version of the MMSE, similar to that of Yi and Vaupel (2002), with revised items that take into account the Chinese cultural and socioeconomic context making the items easily understandable and practically answerable. These items comprise the components to measure orientation to time, place, and season, short registration, delayed registration, calculation, instruction, visualization, attention, and language. The content, construct, and concurrent validities were covered in the various waves of CLHLS which were not repeated in the present paper (e.g., Gu, 2005). The usual practice of CLHLS users tended to sum up these items for a score ranging from 0 to 30. In the present paper, all the 23 items were categorized as binary.

2.3. Procedure of validation

The validation process is broadly breakdown into three major stages: First, EFA to determine the number of dimensions; second, CFA to obtain the best factor structure of MMSE; and last, factorial invariance testing to examine the degree of invariance concerning age. These three steps are elaborated on in the following three subsections. The various functions from the software package R were used for analyses.

2.3.1. EFA

EFAs were carried out to determine the number of dimensions and the factor structure of MMSE. First, factorability was carried out to assess the adequacy of carrying out factor analysis. Bartlett test of sphericity (Bartlett, 1951) and Kaiser–Mayer–Olkin (KMO) measures of sampling adequacy (MSA) (Kaiser, 1970; Kaiser and Rice, 1974) were generated to provide the evidence of factorability of factor analysis, indicating the adequacy for carrying out the factor analysis and providing evidence about the sufficient numbers of significant correlations among the 23 MMSE indicators to justify undertaking factor analysis (Pett, Lackey, and Sullivan, 2003). For individual MMSE items, individual MSA was reported. Function KMO and cortest.bartlett from the package psych were used to generate these three factorability indices (Revelle, 2021).

Next, the heterogeneous correlation matrix for the 23 binary coded MMSE items was generated using the function hector from the package polycor (Fox, 2019) as an input to carrying out EFA using the package psych, function fa (Revelle, 2021). The present study used several methods that were commonly recommended in the literature (Finch, 2020) to determine the dimension of the MMSE inventory. These methods included the Kaiser's greater than 1 rule (Kaiser, 1960), scree test (Cattell, 1966), Velicer's minimum average partial procedure criteria (map; Velicer, 1976; Velicer and Jackson, 1990), Horn's parallel analysis (Dinno, 2009; Glorfeld, 1995; Hayton, Allen, and Scarpello, 2004; Horn, 1965), VSS criterion (Revelle and Rocklin, 1979), BIC (Schwartz, 1978), sample size adjusted BIC (SABIC; Sclove, 1987), root mean square error of approximation (RMSEA; Browne and Cudeck, 1992), and standardized root means square of the residuals (RMSR; Bentler, 1995). The package psych, function VSS. scree, and VSS generated the scree plot, the VSS, RMSEA, RMSR, and the various information criteria indicators. The package paran and function paran produced the parallel

analysis. The ultimate decision to determine the number of factors was based on the outcomes of the above-mentioned methods and the interpretability of the EFA results.

2.3.2. CFA

After determining the number of factors, CFAs were carried out using the package lavaan function cfa with the specification of ordered = TRUE to indicate binary CFA which was performed. The weighted least square mean and variance adjusted estimator was used for the estimation. It used diagonally weighted least squares to estimate the model parameters. Model fit for CFAs was assessed using multiple indices, consisting of the χ^2 statistic (Bollen, 1989; Jöreskog, 1993), comparative fit index (CFI; Bentler, 1990), Tucker and Lewis index (TLI; Tucker and Lewis, 1973), (RMSEA; Browne and Cudeck, 1992), and standardized root mean square residual (SRMR; Brown, 2006). The χ^2, RMSEA, and SRMR assess how well the covariances predicted from the estimates reproduced the sample covariances (Pomplun and Omar, 2001). The CFI and TLI assess the degree of fit of the proposed model accounted for the sample covariances. RMSEA values approximating 0.06 demonstrate a close fit of the model (Browne and Cudeck, 1992; Hu and Bentler, 1999). A value <0.08 is generally considered a good fit for SRMR (Hu and Bentler, 1999). CFI and TLI values of 0.90 (Bentler, 1990) and 0.95 (Hu and Bentler, 1999), respectively, indicate an acceptable and good fit for the model.

2.3.3. Factorial invariance

After determining the CFA model that best described the structure of the MMSE, factorial invariance testing was carried out to examine the degree of invariance concerning age. Factorial invariance is a concept applied in the context of psychometric analysis of an inventory (Revuelta, Franco-Martinez, and Ximénez, 2021; Schürer, van Ophuysen, and Behrmann, 2021; Putnick and Bornstein, 2016) to measure the degree of invariance assurance for a categorical variable about its applicability level. In the current context, this concept postulates the psychometric properties of MMSE, and its applicability for the two age groups, young-old and old-old. The first invariance testing is configural invariance, commonly referred to as pattern invariance. This testing procedure aims to examine whether both the young-old and old-old have the same MMSE factor structure (Cheung and Rensvold, 2002; Horn and McArdle, 1992; Vandenberg and Lance, 2000). While the configural invariance is satisfied, the next step is to set the factor loadings to be equal across the two age groups. This is generally referred to as metric invariance. Metric invariance is built on configural invariance by requiring that in addition to the equality of the structural form, the factor loadings are also equivalent across the two age groups. The next level of constraint that adds to the invariance is the intercepts. It was also referred to as scalar equivalence (Mullen, 1995), indicating the existence of strong factorial invariance (Meredith, 1993). While the mean is further constrained across the groups, it is referred to as strict factorial invariance, indicating systematic group differences in means matrices are due to group differences in common factor score distributions (Yoon and Millsap, 2007). The factorial invariance testing was carried out using function cfa from the package lavaan (Rosseel, 2012) with the specification of group.equal to "loadings," "intercepts," and "means" for metric, scalar, and strict invariance, respectively.

3. Results

3.1. Descriptive statistics

Descriptive statistics were generated using the function dfSummary from the package summarytools (Comtois, 2021). Table 1 shows the mean and standard deviation of the 23 MMSE items for the study and the two age groups of young-old and old-old.

3.2. EFA

3.2.1. Factorability

The KMO measure of sampling adequacy was 0.9, indicating the adequacy of undertaking factor analysis. Similarly, the Bartlett test of sphericity also indicated the sufficiency numbers of significant correlations that it was unlikely the population correlation matrix was an identity matrix ($\chi^2 = 69{,}555$; $P < 0.001$). The individual MSA that ranged from 0.77 to 0.97 also gave the same conclusion.

3.2.2. Determining number of factors

The various methods of determining the number of factors gave diverse outcomes that ranged from 1 to 7. However, the majority of the fit indicators showed either a six-factor or a seven-factor solution. The heterogeneous correlation correlogram with the correlation coefficients shown in Figure A1 and without correlation coefficients shown in Figure A2 indicated either a six-factor or a seven-factor solution. The scree plot and Horn's parallel analysis showed a six-factor solution (Figure A3). Out of the six fit indices from Table A1, RMSEA, BIC, and SABIC suggested seven factors. While the Velicer's minimum average partial procedure criteria (map) showed a five-factor model, the two vss1 and vss2 indicated a one-factor and a two-factor solution, respectively. The Kaiser's greater than 1 rule indicated a six-factor solution with a moderately high 61.57% cumulative percentage explained (Table A2). In summary, both the six-factor and seven-factor models

Table 1. MMSE items and older group descriptive statistics

MMSE item	Mean	Standard deviation
Orientation		
Morning, noon, afternoon, and night	0.99	0.12
Month	0.95	0.22
Mid-autumn festival (month and date)	0.94	0.25
Four seasons	0.97	0.18
Home – region/village	0.98	0.15
Short/delayed recall/registration		
Registration of desk	0.96	0.20
Registration of apple	0.95	0.22
Registration of dress	0.94	0.24
Delayed recall		
Recall of desk – repetition	0.84	0.37
Recall of apple – repetition	0.84	0.37
Recall of dress – repetition	0.80	0.40
Calculation		
Calculation – 20-3	0.95	0.21
Calculation – 20-3-3	0.89	0.32
Calculation – 20-3-3-3	0.88	0.32
Calculation – 20-3-3-3-3	0.85	0.35
Calculation – 20-3-3-3-3-3	0.86	0.35
Language		
Verbal repetition a sentence	0.95	0.21
Name pen	0.99	0.08
Name watch	0.99	0.08
Comprehend instruction		
Instruction 1 – Pick paper with right hand	0.97	0.16
Instruction 2 – Fold the paper	0.97	0.16
Instruction 3 – Place paper on the floor	0.98	0.13
Visuospatial/drawing		
Draw intersecting pentagons	0.64	0.48
Age		
All (65 and above)	80.5	10.11
Young-old (65–79)	72.2	4.27
Old-old (80 and above)	88.9	6.85

were two more likely exploratory factor models, suggested by the majority of the fit indicators.

3.3. CFA

Based on the theoretical cognitive function classification of MMSE together with the results of EFA, the base reference factor model was set to the seven-factor model while the six-factor model was specified as the supporting model. The main difference between the six-factor and seven-factor was mainly due to the further breakdown of the six MMSE items into two constructs of language and instruction. This led to stating five CFA models with the seven-factor oblique model as the hypothesized model and the rest as competing CFA models. The first proposed competing CFA was a one-factor model. This model indicated all the 23 MMSE items were formed into a single MMSE construct. The six-factor oblique CFA was the next proposed competing CFA model. While the fit of

the seven-factor was better than the six-factor solution, the orthogonal seven-factor was included to examine whether an unrelated of the seven-factor constructs of MMSE was better than the seven-factor oblique model. While the one-factor CFA model was that all the 23 items were grouped under one overall MMSE construct, a second-order seven-factor CFA also indicated an overall MMSE construct however loaded under the 7 MSSE constructs instead of the 23 MMSE items.

Table 2 shows the results of the proposed seven-factor oblique CFA model and the four competing CFA models. The fit indices CFI and TLI of the hypothesized seven-factor oblique model (1.0; 1.0) were higher than the four competing models. Similarly, the RMSEA and SRMR of the hypothesized model also indicated a better fit than the four competing models with the lowest values (0.009; 0.033). These results indicated the proposed seven-factor oblique model fitted better than the four competing models. Table 3 displays the latent correlations of the seven factors together with their reliability indicator, Zumbo's alpha. Since all the MMSE items were binary coded, the ordinal Zumbo's alpha was reported (Zumbo, Gadermann, and Zeisser, 2007). The factor correlations for all the seven domains of MMSE were all positive. These positive coefficients indicated an older adult that possessed a high MMSE construct in one cognitive function domain tended to also possess high in another domain. For instance, the short recall was moderately correlated with language with a positive correlation coefficient of 0.64, indicating that an older adult that has a high recall was also high in language capability. These results supported the proposed theoretical CFA that these seven MMSE domains were distinct yet associated. The Zumbo's alpha coefficients for all the six MMSE constructs were all high in reliability with values all higher than 0.89.

Figure 1 shows the graphical representation of the seven-factor oblique CFA model. These seven factors were orientation, short recall, delayed recall, calculation, language, comprehend instruction, and visuospatial. The standardized factor loading coefficients were indicated on top of the arrow that ran from the construct to the MMSE items. All the factor loadings were high in value except for visuospatial. These high loadings indicated the high association of the MMSE items to the respective latent factor MMSE construct. The error residuals were printed after the items, on top of the arrow that ran from error terms (E1 to E23) to the MMSE items. These error residuals were low in value also indicating that the MMSE items were low in measurement errors when they were loaded into the appropriate MMSE construct. The double arrows that ran between the seven constructs indicated it represents an oblique CFA model.

3.4. Factorial invariance

Table 4 summarizes the results of factorial invariance of the two age groups. The CFI for all the four invariance conditions was all at a high value of 0.999. Similar results

Table 2. Fit indices of hypothesized cfa and competing models.

Model	χ^2	df	CFI	TLI	RMSEA	SRMR
Hypothesized seven-factor oblique	341*	209	1.000	1.000	0.009	0.034
Competing seven-factor second order	605*	223	1.000	0.999	0.014	0.044
Competing six-factor oblique	609*	215	1.000	0.999	0.015	0.050
Competing one-factor	9196*	230	0.976	0.973	0.069	0.123
Competing seven-factor orthogonal	77470*	230	0.789	0.768	0.201	0.467

χ^2, Chi-square statistics; df, degrees of freedom; CFI, comparative fit index, TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square. *$P < 0.01$.

Table 3. Factor correlations and Zumbo's alpha for the seven domains of MMSE.

MMSE construct	1.	2.	3.	4.	5.	6.	7.	Zumbo's alpha
1. Orientation								0.93
2. Short recall	0.53							0.96
3. Delay recall	0.48	0.57						0.94
4. Calculation	0.57	0.58	0.50					0.98
5. Language	0.56	0.64	0.54	0.63				0.91
6. Instruction	0.32	0.40	0.34	0.40	0.51			0.89
7. Visuospatial	0.38	0.37	0.35	0.50	0.39	0.29		-

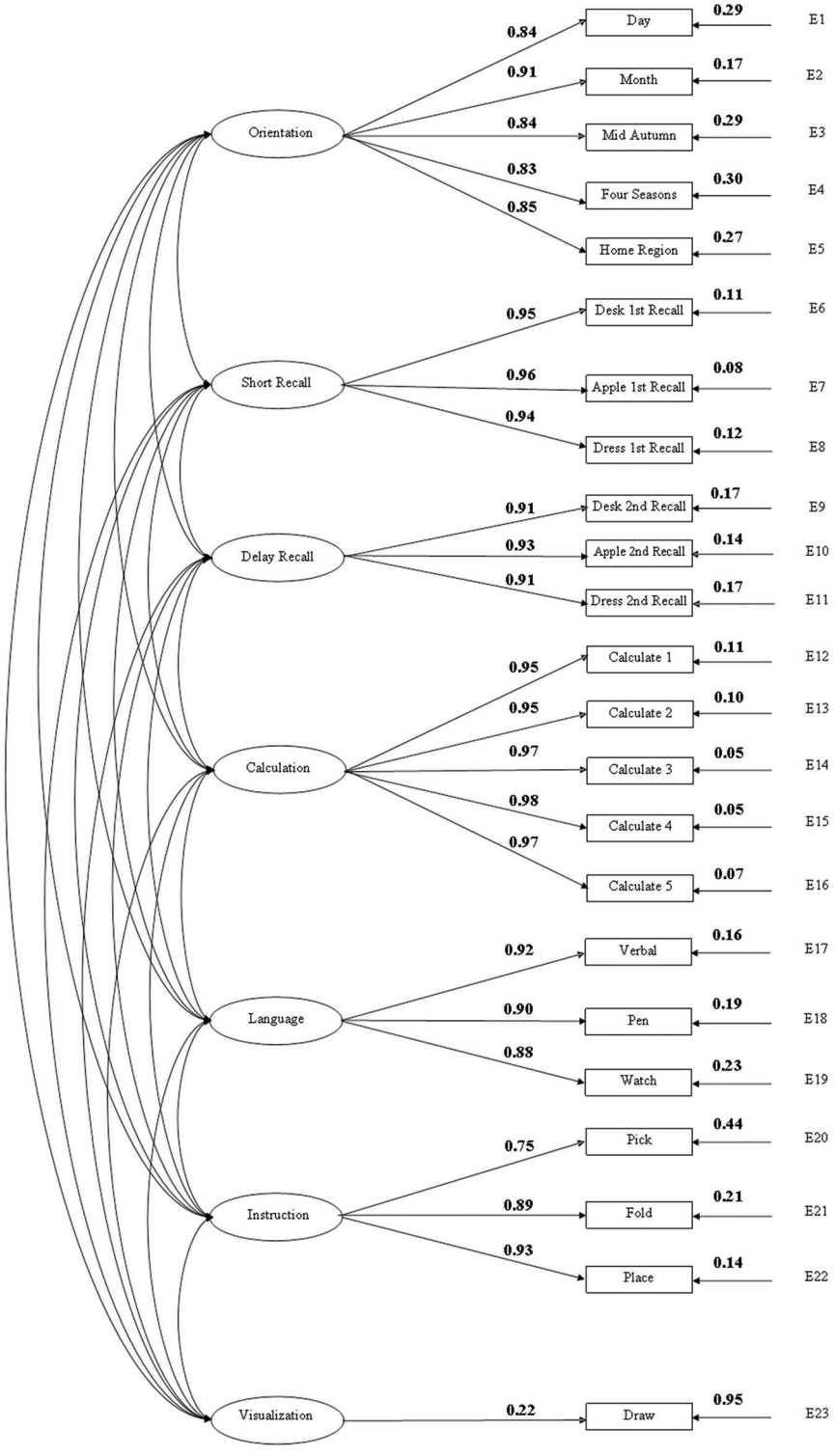


Figure 1. Mini-Mental State Examination seven-factor oblique confirmatory factor model.

were also found for the TLI. Both RMSEA and SRMR were low at a satisfactory level. These results indicated that the level of invariance for the MMSE was extremely good.

Table 4. Factorial invariance – young-old and old-old.

Factorial invariance	CFI	TLI	RMSEA	SRMR
Configural	0.999	0.999	0.009	0.057
Metric	0.999	0.998	0.015	0.066
Scalar	0.999	0.999	0.010	0.060
Strict	0.999	0.999	0.012	0.059

CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation; SRMR, standardized root mean squared residual.

4. Conclusion and discussion

This study explored and validated the factor structure of MMSE through EFA, CFA, and the factorial invariance test. The validation results indicated a seven-factor oblique CFA best fitted the MMSE inventory, specifying seven cognitive functions inherently within it: Orientation, short recall, delayed recall, calculation, language, comprehend instruction, and visuospatial. These results were similar to MMSE literature on their cognitive functions but differed in the number of dimensions. The factorial invariance confirmed the hypothesized CFA was at a high invariance level showing almost no measurement differences between the young-old and old-old. The reliability results also indicated that these 23 MMSE indicators that formed the seven factors were of high reliability.

One main finding of the present study is the statistical results of the CFA indicated the alignment between the 23 MMSE items with the theoretical expected seven cognitive functions. For instance, the three items of registration of desk, apple, and dress were fittingly grouped under the short recall cognitive function, and the five subtraction calculation items properly formed the calculation cognitive function. This finding did not appear in the MMSE literature that generally reported a low dimension (e.g., Fillenbaum, Heyman, Wilkinson, *et al*., 1987). One plausible reason is that the MMSE literature was not predominantly to establish the dimensions of the MMSE inventory with the main purpose being to relate the MMSE items to the respective cognitive functions they belonged to (e.g., Fillenbaum, Heyman, Wilkinson, *et al*., 1987; Tinklenberg, Brooks, Tanke, *et al*., 1990) but reluctantly used the EFA to confirm a low dimension to form a summated MMSE score. Another probable reason is that more updated appropriate EFA procedures to determine the number of dimensions used in the present paper were not adopted even for the more recent papers (e.g., Baek, Kim, Park, *et al*., 2016).

The present paper also revealed the characteristics of the MMSE inventory possess a duality factorial structure that could be viewed as a seven dimension of cognitive functions, and also as a general higher-order all-inclusive cognitive function construct within which the seven dimensions of cognitive functions were grouped under it. This duality property was reflected in the results of CFA. The seven-factor CFA model turned out as the best model in which the MMSE items were grouped under their respective cognitive functions according to the theoretical grouping showed not only the alignment with theory expectation but also indicating the first property of this inventory that these seven cognitive functions were associated but were separate constructs. The second best fit CFA, the second-order CFA with a slightly lower fit, indicated the possibility to view this inventory as a second-order cognitive function construct. The moderate to high factor correlations of the seven constructs within the seven oblique CFA further showed evidence that these seven constructs were positively associated but were different in their cognitive functionality. Similarly, the low fit of the seven-factor orthogonal CFA also indicated the unlikeliness that the seven MMSE constructs were unrelated. The practical implication of this duality property is that MMSE can be viewed as an overall indicator or as separate seven distinct but related cognitive functions.

Another crucial conclusion and inference from the CFA result is that it indicates the routine way of generating an MMSE score, whether it is viewed as separate seven constructs or an overall higher-order construct, the summated score to generate an overall MMSE score or subscale MMSE scores by summing the items is not an appropriate procedure. The earlier studies on the validation of MMSE often use an all-inclusive MMSE summated score to represent an index in measuring the level of cognitive function by summing the MMSE items according to the number of items correctly answered (e.g., Park, Kwon, Jung, *et al*., 2012 used domain MMSE summated scores). The limitations of the summated score were well noted in the measurement literature. The main limitation is that summated score does not take care of measurement errors. When the summated score is further subdivided into a few categorical levels using a cutoff arbitrary decision to distinguish the risk level of cognitive impairment, additional measurement errors are introduced and converting a continuous score to a categorical variable

produces a loss of information. However, this procedure was commonly found in studies that used MMSE (Deb and Braganza, 1999; Osterweil, Mulford, Syndulko, *et al*., 1994; Yi and Vaupeo, 2002). More seriously, the deriving of an optimal cutoff score (e.g., Wong and Fong, 2009) has nothing to do with the validity of the MMSE inventory and its dimensionality, and the relationship to the cognitive functions but a vaguely derived score is generated that contains measurement errors.

The results of the factorial invariance indicated the MMSE inventory possessed a high degree of invariance between the two age groups which were seldom found in the factorial invariance literature. The configural invariance indicated that both the young-old and old-old have the same factorial structure. This gave the conclusion that the structural form between MMSE items and their cognitive functions is the same for the two age groups. The metric invariance further provided evidence that the factor loadings were also invariant across the two age groups. That is the weights that indicated the association between the MMSE items and the cognitive function construct were also maintained, reflected by the equality of their respective factor loadings. Scale invariance further qualified the latent MMSE construct was also with the same degree of measurement errors across the two age groups. The last invariance, the strong factorial invariance, is a prerequisite to testing for the equality of latent means. In the presence of this invariance, the comparison of latent means becomes unambiguous (Cheung and Rensvold, 2002), indicating that the MMSE could be used with high confidence for latent model analysis, where systematic group differences in means matrices are due to group differences in common factor score distributions (Yoon and Millsap, 2007). In summary, the results of the factorial invariance assured the use of the MMSE as a seven latent dimensions construct and suggested that moving away from the commonly used summated score approach which contained measurement errors to latent modeling is a better direction for further analysis using latent models.

For further analyses and follow-up after validation, the recommendation is to set CFA as the base to establish the measurement component of MMSE. When the structural component is considered, a latent approach is recommended to use the various latent models for further analyses. For instance, using the structural equation model to relate MMSE items and the cognitive function constructs to a group of covariates or examining the effect of cognitive function constructs on a medical condition. The latent modeling followed after CFA has several advantages. First, the number of MMSE items to include in an MMSE inventory becomes a researcher's choice that the option to choose a subset of MMSE items that are more related to the study does not restrict to a complete list of MMSE items. Second, the creation of the various cognitive functions allows for addressing the effects separately from the overall cognitive function or an individual cognitive subscale. For instance, examining the effect of the overall cognitive function on a medical condition and concurrently exploring the effect of the calculation subscale on the same medical condition. Third, removing the measurement errors is automatically inbuilt into the latent model.

In summary, this paper recommends validation of MMSE using EFA, CFA, and factorial invariance test and showed the results of the validation that MMSE is more appropriate than the MMSE literature that only concentrated on EFA. While this systematic approach was commonly used and already established in the measurement, psychology, and education literature, it is recommended for future use of MMSE. These procedures avoid all the limitations discussed in the paper. It takes into account measurement errors, relates the MMSE items more appropriately to the theoretical cognitive function setting, creates competing CFA models that are appropriately set up before testing, and tests for invariances.

Funding

This study was not supported by any grant.

Conflicts of interest

No conflicts of interest were reported by all others.

Authors' contributions

TKT conducted the analysis and drafted the introduction, methods, results, and discussions of the manuscript. QF reviewed, amended, and provided recommendations on the organization of the manuscript.

Ethical approval

The human data used in our study are a publicly available survey dataset that can be downloaded from the webpage: https://cpha.duke.edu/research/chinese-longitudinal-healthy-longevity-survey-clhls.

Availability of supporting data

The CLHLS dataset is in open access on the webpage: https://cpha.duke.edu/research/chinese-longitudinal-healthy-longevity-survey-clhls.

References

Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, *et al*. (2021). Mini-mental state examination (MMSE) for the early detection of dementia in people with mild cognitive

impairment (MCI). *Cochrane Database of Systematic Reviews*, 7(7):CD010783.

https://doi.org/10.1002/14651858.CD010783.pub3

Baek MJ, Kim K, Park YH, *et al.* (2016). The validity and reliability of the MMSE-2 for detecting mild cognitive impairment and Alzheimer's disease in a Korean population. *PLoS One*, 11(9):e0163792.

https://doi.org/10.1371/journal.pone.0163792

Bartlett MS. (1951). The effect of standardization on a Chi-Square approximation in factor analysis. *Biometrika*, 38:337-344.

https://doi.org/10.1093/biomet/38.3-4.337

Bentler PM. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107:238-246.

https://doi.org/10.1037/0033-2909.107.2.238

Bentler PM. (1995). *EQS 5* [*Computer Program*]. Encino, CA: Multivariate Software Inc.

Bollen K and Lennox R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110:305-314.

https://doi.org/10.1037/0033-2909.110.2.305

Bollen KA. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Brown TA. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

Browne MW and Cudeck R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, 21:230-258.

https://doi.org/10.1177/0049124192021002005

Burke SL, Grudzien A, Burgess A, *et al.* (2021). The utility of cognitive screeners in the detection of dementia spectrum disorders in Spanish-speaking populations. *Journal of Geriatric Psychiatry and Neurology*, 34(2):102-118.

https://doi.org/10.1177/0891988720915513

Cattell RB. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1:245-276.

https://doi.org/10.1207/s15327906mbr0102_10

Cheung GW and Rensvold RB. (2002). Evaluating goodness of fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2):233-255.

https://doi.org/10.1207/s15328007sem0902_5

CLHLS. (2020). *Chinese Longitudinal Healthy Longevity Survey, Data Collection and Quality Control Procedures*. Available from: https://sites.duke.edu/centerforaging/programs/chinese-longitudinal-healthy-longevity-survey-clhls/design-method-study-population/design-method-data-collection-quality-control-procedures. [Last accessed on 2022 Jan 26].

https://doi.org/10.1007/978-1-4020-6752-5_2

Comtois D. (2021). *summarytools: Tools to Quickly and Neatly Summarize Data. R package version 1.0.0*. Available from: https://CRAN.R-project.org/package=summarytools.[Last accessed on 2022 Jan 26].

Davey RJ and Jamieson S. (2004). The validity of using the mini-mental state examination in NICE dementia guidelines. *Journal of Neurology Neurosurgery and Psychiatry*, 75:343-344.

Deb S and Braganza J. (1999). Comparison of rating scales for the diagnosis of dementia in adults with Down's syndrome. *Journal of Intellectual Disability Research*, 43(5):400-407.

https://doi.org/10.1046/j.1365-2788.1999.043005400.x

Dinno A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of simulated data. *Multivariate Behavioral Research*, 44(3):362-388.

https://doi.org/10.1080/00273170902938969

Fillenbaum GG, Heyman A, Wilkinson WE, *et al.* (1987). Comparison of two screening tests in Alzheimer's disease. The correlation and reliability of the mini-mental state examination and the modified blessed test. *Archives of Neurology*, 44:924-927.

https://doi.org/10.1001/archneur.1987.00520210026014

Finch WH. (2020). Using fit statistic differences to determine the optimal number of factors to retain in an exploratirt factor analysis. *Educational and Psychological Measurement*, 80(2):217-241.

https://doi.org/10.1177/0013164419865769

Folstein MF, Folstein SE, and McHugh PR. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12:189-198.

https://doi.org/10.1016/0022-3956(75)90026-6

Fox, J. (2019). *polycor: Polychoric and Polyserial Correlations. R package version 0.7-10*. Available from: https://CRAN.R-project.org/package=polycor. [Last accessed on 2022 Jan 26].

Garrodp LE, Abad FJ, and Ponsoda V. (2013). A new look at Horn's parallel analysis with ordinal variables. *American Psychological Association*, 18(4):454-474.

https://doi.org/10.1037/a0030005

Glorfeld LW. (1995). An Improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3):377-393.

https://doi.org/10.1177/0013164495055003002

Gu D. (2005). *Assessment of Data Quality of the Chinese Longitudinal Healthy Longevity Survey in 2002*. The CLHLS Technical Report No.2005-1.

Gu D, Feng Q, Chen H, *et al.* (2021). Chinese longitudinal healthy longevity survey. In: D Gu and ME Dupre (eds.). Encyclopedia of Gerontology and Population Aging.

Switzerland AG, Cham: Springer Nature. p.957-970.

Hayton JC, Allen DG, and Scarpello V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7(2):191-205.

https://doi.org/10.1177/1094428104263675

Horn JL and McArdle JJ. (1992). A practical and theoretical guide to MI in aging research. *Experimental Aging Research*, 18:117-144.

https://doi.org/10.1080/03610739208253916

Horn JL. (1965). A rationale and a test for the number of factors in factor analysis. *Psychometrika*, 30:179-185.

https://doi.org/10.1007/bf02289447

Hsieh SW, Wu DW, Wang CW, *et al.* (2021). Poor cognitive function is associated with obstructive lung diseases in Taiwanese adults. *International Journal of Environmental Research and* Public Health, 18:2344.

https://doi.org/10.3390/ijerph18052344

Hu L and Bentler PM. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6:1-55.

https://doi.org/10.1080/10705519909540118

Jackson DJ, Gillaspy JA, and Purc-Steogebsib R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1):6-23.

https://doi.org/10.1037/a0014694

Jöreskog KG. (1993). Testing structural equation models. In: KA Bollen and JS Long (eds.). *Testing Structural Equation Models*. Newbury Park, CA: Sage, p.294-316.

Kaiser HF and Rice J. (1974). Little jiffy, mark Iv. *Educational and Psychological Measurement*, 34(1):111-117.

https://doi.org/10.1177/001316447403400115

Kaiser HF. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20:114-151. http://dx.doi.org/10.1177/001316446002000116

Kaiser HF. (1970). A second generation little jiffy. *Psychometrika*, 35(4):401-415.

https://doi.org/10.1007/BF02291817

Lane FC, Anderson B, Ponce HF, *et al.* (2012). Factorial invariance of LibQUAL as a measure of library service quality over time. *Library and Information Science Research*, 34:22-30.

https://doi.org/10.1016/j.lisr.2011.07.004

Langa KM, Llewellyn DJ, Lang IA, *et al.* (2009). Cognitive health among older adults in the United States and in England. *BMC Geriatrics*, 9:23. https://doi.org10.1186/1471-2318-9-23

Lezak MD, Howieson DB, and Loring DW. (2004). *Neuropsychological Assessment*. 4th ed. New York: Oxford University Press.

Liau AK, Chow D, Tan TK, *et al.* (2010). Development and validation of the Personal strengths inventory using exploratory and confirmatory factor analyses. *Journal of Psychoeducational Assessment*, 29:14-26.

https://doi.org/10.1177/0734282910365648

Lim M, Li Z, Xie H, *et al.* (2019). An Asian study on clinical and psychological factors associated with personal recovery in people with psychosis. *BMC Psychiatry*, 19:256.

https://doi.org/10.1186/s12888-019-2238-9

Meredith W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58:525-543.

https://doi.org/10.1007/bf02294825

Mitchell AJ. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*, 43(4):411-431.

https://doi.org/10.1016/j.jpsychires.2008.04.014

Mullen M. (1995). Diagnosing measurement equivalence in cross-national research. *Journal of International Business Studies*, 3:573-596.

https://doi.org/10.1057/palgrave.jibs.8490187

Osterweil D, Mulford P, Syndulko K, *et al.* (1994). Cognitive function in old and very old residents of a residential facility: Relationship to age, education, and dementia. *Journal of the American Geriatric Society*, 42(7):766-773.

https://doi.org/10.1111/j.1532-5415.1994.tb06539.x

Park MH, Kwon DY, Jung JM, *et al.* (2012). Mini-mental status examination as predictors of mortality in the elderly. *Acta Psychiatrica Scandinavica*, 127:298-304.

https://doi.org/10.1111/j.1600-0447.2012.01918.x

Pett MA, Lackey NR, and Sullivan JJ. (2003). *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. Thousand Oaks, CA: Sage.

https://doi.org/10.4135/9781412984898

Pomplun M and Omar MH. (2001). The factorial invariance of a test of reading comprehension across groups of limited English proficient students. *Applied Measurement in Education*, 14:261-283.

https://doi.org/10.1207/s15324818ame1403_4

Putnick DL and Bornstein MH. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Development Review*, 41:71-90.

https://doi.org/10.1016/j.dr.2016.06.004

Reilly S, Challis D, Burns A, *et al.* (2004). The use of assessment scales in old age psychiatry services in England and Northern Ireland. *Aging and Mental Health*, 8(3):249-55.

https://doi.org/10.1080/13607860410001669787

Revelle W and Rocklin T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14:403-414.

https://doi.org/10.1207/s15327906mbr1404_2

Revelle W. (2021). *psych: Procedures for Personality and Psychological Research*. Evanston, Illinois, USA: Northwestern University. Available from: https://CRAN.R-project.org/package=psych Version = 2.1.9. [Last accessed on 2022 Jan 26].

Revuelta J, Franco-Martinez A, and Ximénez C. (2021). Nominal factor analysis of situational judgment tests: Evaluation of latent dimensionality and factorial invariance. *Educational and Psychological Measurement*, 81(6):1054-1088.

https://doi.org/10.1177/0013164421994321

Rosseel Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2):1-36.

https://doi.org/10.18637/jss.v048.i02

Schürer S, van Ophuysen, S, and Behrmann, L. (2021). A new instrument for assessing cohesion in primary and grammar school classes factorial structure and measurement invariance of the GruKo4. *Journal of Psychoeducational Assessment*, 39(3):271-285.

https://doi.org/10.1177/0734282920966825

Schwartz G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6:461-464.

Sclove LS. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52:333-343.

https://doi.org/10.1007/bf02294360

Shigemori K, Ohgi S, Okuyama E, *et al*. (2010). The factorial structure of the mini-mental state examination (MMSE) in Japanese dementia patients. *BMC Geriatrics*, 10:36.

https://doi.org/10.1186/1471-2318-10-36

Shulman KI, Herrmann N, Brodaty H, *et al*. (2006). IPA survey of brief cognitive screening instruments. *International Psychogeriatrics*, 18:281-294.

https://doi.org/10.1017/s1041610205002693

Tinklenberg J, Brooks JO, Tanke ED, *et al*. (1990). Factor analysis and preliminary validation of the mini-mental state examination from a longitudinal perspective. *International Psychogeriatrics*, 2(2):123-134.

https://doi.org/10.1017/s1041610290000382

Tucker LR and Lewis C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1):1-10.

https://doi.org/10.1007/BF02291170

van der Meulen M, Dirven L, Bakunina K, *et al*. (2021). MMSE is an independent prognostic factor for survival in primary central nervous system lymphoma. *Journal of Neurooncology*, 152:357-362.

https://doi.org/10.1007/s11060-021-03708-8

Vandenberg RJ and Lance CE. (2000). A review and synthesis of the MI literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3:4-69.

https://doi.org/10.1177/109442810031002

Velicer WF and Jackson DN. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25:1-28.

https://doi.org/10.1207/s15327906mbr2501_1

Velicer WF. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41:321-327.

https://doi.org/10.1007/bf02293557

Wong SS and Fong NK. (2009). Reliability and validity of the telephone version of the cantonese mini-mental state examination (T-SMMSE) when used with elderly patients with and without dementia in Hong Kong. *International Psychogeriatrics*, 1(2):345-353.

https://doi.org/10.1017/s1041610209008588

Yi Z and Vaupel JW. (2002). Functional capacity and self-evaluation of health and life of oldest old in China. *Journal of Social Issue*, 58:733-748.

https://doi.org/10.1111/1540-4560.00287

Yoon M and Millsap RE. (2007). Detecting violations of factorial invariance using data-based specification searches: A monte carlo study. *Structural Equation Modeling*, 14(3):435-463.

https://doi.org/10.1080/10705510701301677

Zumbo BD, Gadermann AM, and Zeisser C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, 6(1):21-29.

https://doi.org/10.22237/jmasm/1177992180

Appendix

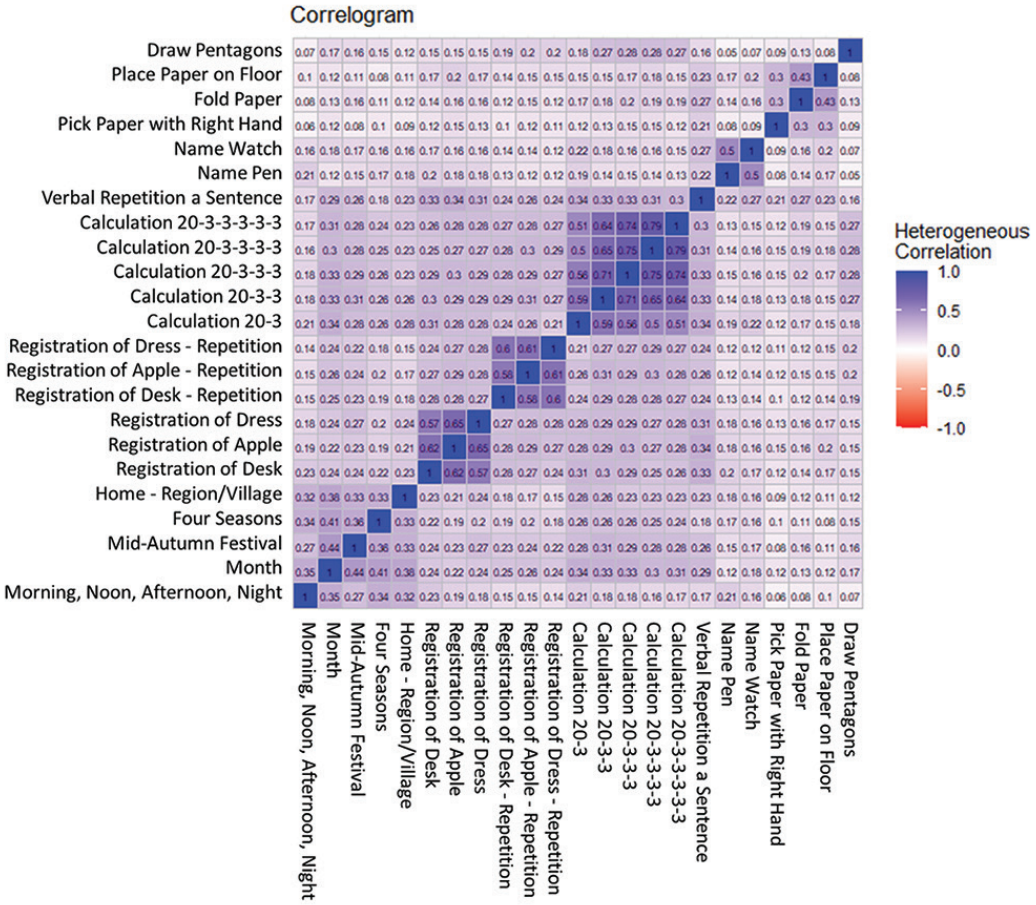


Figure A1. Correlogram with correlation coefficient printed.

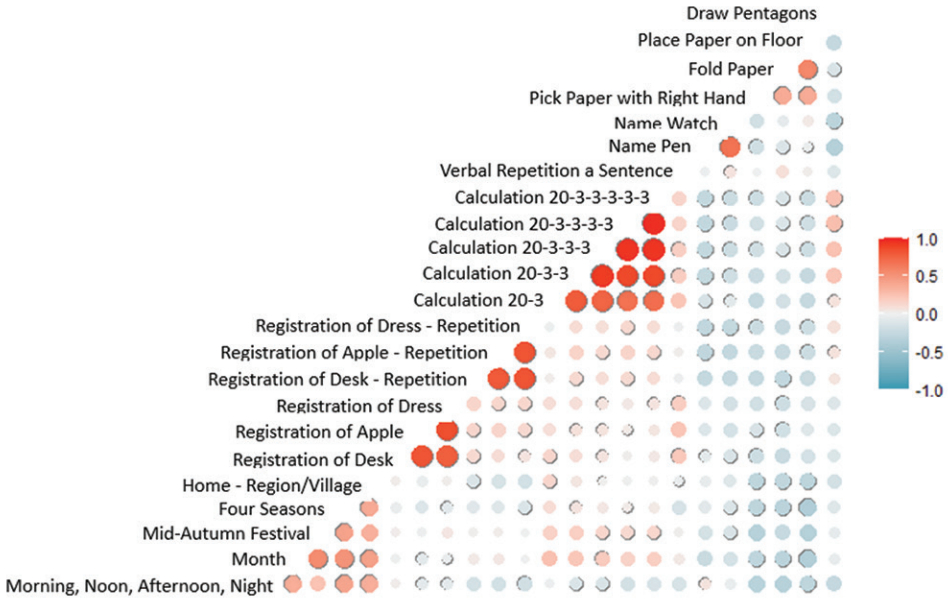


Figure A2. Correlogram without correlation coefficient.

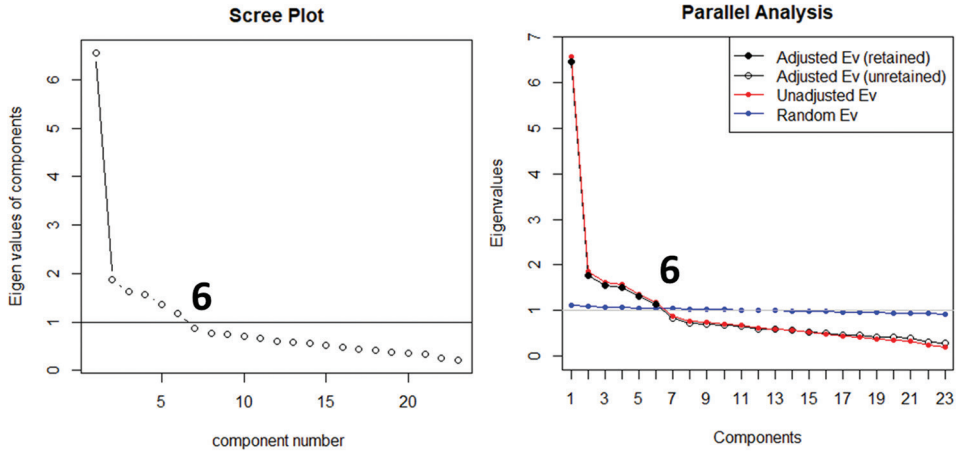


Figure A3. Scree plot and Horn's parallel analysis.

Table A1. Indicators to the determination of the number of factors.

No. of factors	RMSEA	BIC	SABIC	map	vss1	vss2
1	0.121	26100	26831	0.022	0.71	0
2	0.098	14752	15413	0.017	0.55	0.76
3	0.085	9703	10297	0.018	0.52	0.74
4	0.067	4938	5469	0.019	0.5	0.68
5	0.051	2055	2526	0.016	0.51	0.70
6	0.036	351	764	0.018	0.53	0.70
7	0.022	−444	−85	0.024	0.53	0.67

RMSEA, root mean square error of approximation; BIC, Bayesian information criterion; SABIC, sample-adjusted BIC; map, Velicer's MAP criteria; vss1, very simple structure 1; vss2, very simple structure 2.

Table A2. Number of factors, eigenvalue, % and cumulative % variance explained.

No. of factors	Eigenvalue	% of variance	Cumulative % of variance
1	6.56	28.52	28.52
2	1.87	8.12	36.63
3	1.62	7.04	43.67
4	1.57	6.82	50.49
5	1.36	5.93	56.42
6	1.19	5.15	61.57
7	0.87	3.77	65.34
8	0.76	3.30	68.64
9	0.74	3.20	71.84
10	0.70	3.03	74.87
11	0.67	2.91	77.78
12	0.61	2.63	80.42
13	0.58	2.53	82.95
14	0.56	2.45	85.40
15	0.52	2.25	87.65
16	0.48	2.07	89.72
17	0.43	1.88	91.60
18	0.42	1.80	93.40
19	0.37	1.63	95.03
20	0.36	1.55	96.58
21	0.34	1.46	98.04
22	0.24	1.05	99.10
23	0.21	0.90	100.00

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Self-reported hearing loss, hearing aid use, and cognitive function among U.S. older adults

Jessica S. West[1,2]*, Sherri L. Smith[1,2,3], and Matthew E. Dupre[1,2,4]

[1]Center for the Study of Aging and Human Development, Duke University, Durham, North Carolina, USA

[2]Department of Population Health Sciences, Duke University, Durham, North Carolina, USA

[3]Department of Head and Neck Surgery and Communication Sciences, Duke University Durham, North Carolina, USA

[4]Department of Sociology, Duke University, Durham, North Carolina, USA

Abstract

There has been increasing attention to the role of hearing loss as a potentially modifiable risk factor for Alzheimer's disease and related dementias. However, more nationally-representative studies are needed to understand the co-occurring changes in hearing loss and cognitive function in older adults over time, and how hearing aid use might influence this association. The purpose of this report is to examine how age-related changes in hearing loss and hearing aid use are associated with trajectories of cognitive function in a nationally-representative sample of U.S. older adults. We used 11 waves of longitudinal data from the Health and Retirement Study (HRS) from 1998 to 2018 to examine changes in self-reported hearing loss, hearing aid use, and cognitive function in adults 65 and older by race and ethnicity. Results from mixed models showed that greater levels of hearing loss were associated with lower levels of cognitive function at age 65 in non-Hispanic White, non-Hispanic Black, and Hispanic older adults. We also found that the associations diminished across age in White and Black individuals; but remained persistent in Hispanic individuals. The use of hearing aids was not associated with cognitive function in Black older adults but appeared protective for White and Hispanic older adults. Overall, the findings from this report suggest that the timely identification of hearing loss and subsequent acquisition of hearing aids may be important considerations for reducing declines in cognitive function that manifests differently in U.S. population subgroups.

Keywords: Hearing loss; Cognitive decline; Dementia; Racial/ethnic disparities; Longitudinal trajectories

***Corresponding author:**
Jessica S. West,
(jessie.west@duke.edu)

Citation: West JS, Smith SL, Dupre ME. (2022). Self-reported hearing loss, hearing aid use, and cognitive function among U.S. older adults. International Journal of Population Studies, 8(1):16-24. https://doi.org/10.36922/ijps.v8i1.1308

Received: March 30, 2022

Accepted: May 27, 2022

Published Online: June 15, 2022

Copyright: © 2022 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations

1. Introduction

Alzheimer's disease and related dementias (ADRDs) currently affect more than 55 million people worldwide and are projected to impact nearly 80 million people by 2030 (Gauthier, Rosa-Neto, Morais, *et al.*, 2021). The prevention of ADRD is a global public health priority and identifying modifiable risk factors for cognitive decline will contribute to the development of effective interventions (Gauthier, Rosa-Neto, Morais, *et al.*, 2021). Hearing loss has received increasing attention as a potentially modifiable

risk factor for cognitive decline (Livingston, Sommerlad, Orgeta, *et al.*, 2017). In the United States, hearing loss is one of the most common health problems in later life (Whitson, Cronin-Golomb, Cruickshanks, *et al.*, 2018), impacting more than 25% of individuals aged 65–74 and upward of 50% of individuals aged 75 and older (CDC, 2017).

Although there is accumulating evidence that hearing loss is linked to cognitive decline (Whitson, Cronin-Golomb, Cruickshanks, *et al.*, 2018), the findings have been disparate, and the explanations for the associations have varied in the literature (see Wayne and Johnsrude 2015 for a comprehensive review). In particular, most of the research on hearing and cognition has examined cross-sectional measures of hearing loss – either at baseline (Alattar, Bergstrom, Laughlin, *et al.*, 2019; Deal, Betz, Yaffe, *et al.*, 2017; Ge, McConnell, Wu, *et al.*, 2021; Golub, Brickman, Ciarleglio, *et al.*, 2020; Golub, Luchsinger, Manly, *et al.*, 2017; Lin, Yaffe, Xia, *et al.*, 2013) or at the end of the follow-up period (Deal, Sharrett, Albert, *et al.*, 2015). Consequently, these studies do not account for the co-occurring changes in hearing status and cognition that can occur over time. Indeed, research using longitudinal measures of hearing and cognitive function has shown that hearing loss is associated with lower baseline performance on cognitive tests, as well as accelerated declines in cognition compared to those with no hearing loss (Maharani, Dawes, Nazroo, *et al.*, 2018b; 2019). However, more research is needed from a life course perspective to better understand the short- and long-term impact of hearing loss to potentially identify individuals who may be at greater risk of cognitive decline over time.

Studies have shown that the prevalence of hearing loss in the United States is highest in White adults, followed by Hispanic and Black adults, respectively (Agrawal, Platz, Niparko, *et al.*, 2008). Studies have also shown that Black and Hispanic older adults have higher risks for ADRD compared with White older adults (Babulal, Quiroz, Albensi, *et al.*, 2019). However, most studies that have examined racial/ethnic differences in the impact of hearing loss on cognition have simply controlled for race and/or ethnicity and have not considered possible differences among these groups (Brenowitz, Kaup, Lin, *et al.*, 2019; Curhan, Willett, Grodstein, *et al.*, 2019; Lin, 2011; Lin, Metter, O'Brien, *et al.*, 2011; Wallhagen, Strawbridge, and Shema, 2008). It is also unclear whether and to what extent the associations change with increasing age among these population groups. Thus, there remains limited evidence of racial/ethnic differences in the longitudinal association between hearing loss and cognitive decline at the national level.

The purpose of this report is to examine how age-related changes in hearing loss and hearing-aid use are associated with trajectories of cognitive function in U.S. older adults. Using nationally-representative longitudinal data of adults aged 65 and older from 1998 to 2018, we characterize how levels of self-reported hearing and hearing aid use are related to changes in cognitive function for 20 years in non-Hispanic White, non-Hispanic Black, and Hispanic older adults. We also account for a wide array of sociodemographic, behavioral, and health-related factors that may contribute to the associations.

2. Data and methods

2.1. Data

We analyzed 11 waves of publicly-available RAND Health and Retirement Study (HRS) data from 1998 to 2018 (RAND Center for the Study of Aging, 2021). Sponsored by the National Institute on Aging (grant number U01AG009740), the HRS is a nationally-representative prospective study of U.S. adults over the age of 50 that has collected biennial data over the past 30 years. Specific details of the multistage sampling design, data collection techniques, and response rates have been documented extensively elsewhere (Sonnega, Faul, Ofstedal, *et al.*, 2014). The current analysis is limited to 38,231 participants who were eligible to participate in 1998-2018. We limited our analysis to adults aged 65 and older who were administered the measures for cognitive functioning in the HRS (described below) and aged 85 and younger to minimize the potential influence of selective survival at advanced ages. We further limited our analysis to participants who identified themselves as Hispanic, non-Hispanic Black or African American (hereafter referred to as Black), or non-Hispanic White (n = 21,076). Approximately 3% of the sample had missing data on at least one measure of hearing or cognitive function and were omitted. Full details of the inclusion/exclusion criteria for the analytic sample are provided in a flow chart (Figure 1). The final sample included 20,545 individuals who provided a total of 90,990 observations for analysis.

2.2. Measures

Our primary dependent variable was cognitive function. Cognitive function was ascertained in HRS participants using an adapted version of the Telephone Interview for Cognitive Status (TICS) which was modeled after the Mini-Mental State Examination (Ofstedal, Fisher, and Herzog, 2005). Beginning in 1998, all age-eligible participants received the full set of cognitive performance tests – which included six tasks that measured (i) speed of mental processing, (ii) memory, (iii) working memory, (iv) orientation, (v) knowledge, and (vi) language (Ofstedal, Fisher, and Herzog, 2005). Correct responses from each

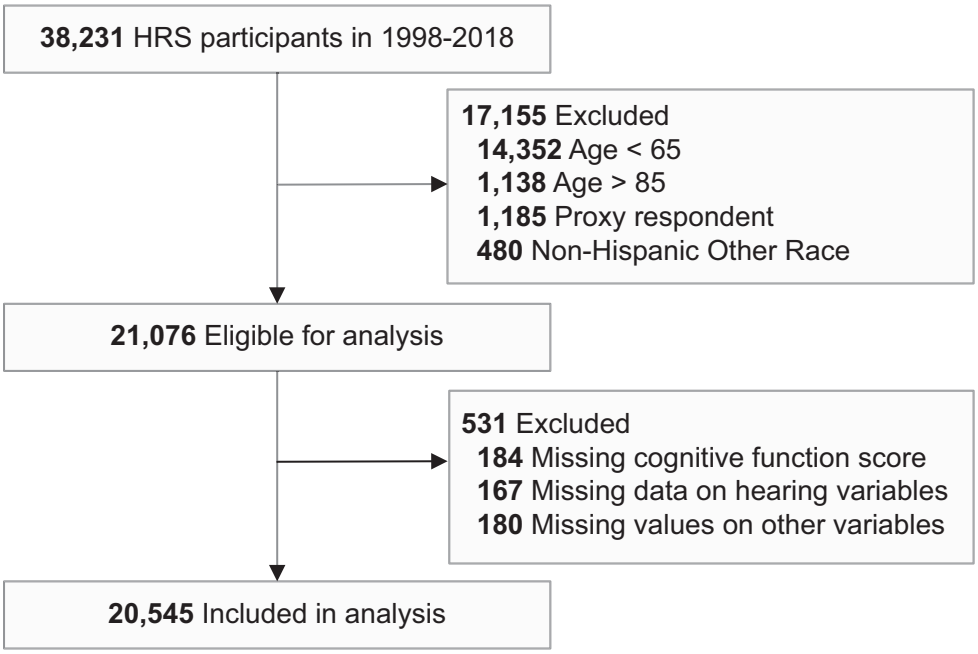


Figure 1. Study participants from the Health and Retirement Study, 1998 – 2018.
HRS: Health and Retirement Study. Note: In the HRS, the non-Hispanic Other race category includes participants who self-identified as American Indian, Alaska Native, Asian, Native Hawaiian, or Pacific Islander. The HRS combines these individuals into an "other" category to protect participant confidentiality due to their small sample sizes.

measure were combined to form a total cognitive score that ranges from 0 to 35 (Ofstedal, Fisher, and Herzog, 2005). Self-reported hearing was measured using a five-point scale (excellent, very good, good, fair, or poor) to assess participants' overall level of hearing (while wearing a hearing aid, if applicable). The use of a hearing aid was also included and dichotomized (0 = no and 1 = yes). Race and ethnicity were assessed using two self-reported measures: 1) do you consider yourself Hispanic or Latino (0=no, 1=yes), and 2) what race do you consider yourself to be: White, Black or African American, American Indian, Alaska Native, Asian, Native Hawaiian, Pacific Islander, or something else? Due to small sample sizes, participants who self-identified as American Indian, Alaska Native, Asian, Native Hawaiian, or Pacific Islander were combined into an "other" category by the HRS to protect participant confidentiality (and were omitted from analysis). We thus use the measures for race and ethnicity to classify individuals as Hispanic, non-Hispanic Black, or non-Hispanic White.

Additional covariates included sex (0 = female and 1 = male), marital status (married [reference], divorced, widowed, or never married), household size (continuous), and whether the participant lives in the South (reference) versus other geographic regions. Several socioeconomic variables were also included: wealth (logged continuous variable), which was measured using total household assets minus debt (Bugliari, Nancy, Chris, *et al*., 2016) and years of education (continuous). Health behaviors included smoking status (never [reference], past, or current smoker) and number of alcoholic drinks per day (0 – 4 [reference] or 5 or more) (McKee, Stransky, and Reichard, 2018). Health status included number of limitations in activities of daily living (ADL, continuous) and indicators for ever being diagnosed by a doctor for high blood pressure, diabetes, cancer, lung disease, heart disease, or stroke. We included an indicator for participants who died during the study period (0 = alive and 1 = died) and a continuous variable for the number of waves present to account for attrition as previously shown (Brown, O'Rand, and Adkins, 2012; Yang and Land, 2013). With the exception of sex, race/ethnicity, and the indicators for attrition, all measures were included as time-varying variables in the models described below.

2.3. Analytic strategy

Overall distributions of study variables were calculated separately for each racial/ethnic group. A series of linear mixed models were used to estimate trajectories of cognitive function from age 65 to 85 (Rabe-Hesketh and Skrondal, 2012). The mixed models (i.e., multilevel mixed-effects linear regression models) used a hierarchical framework to incorporate the individuals' repeated observations (level 1) nested within individuals (level 2)

to estimate (1) the within-individual change in cognitive function with age and (2) the between-person differences in baseline cognitive function (intercept) and the change in cognitive function (slope) with increasing age. Preliminary tests of model fit indicated that a quadratic function best parameterized the age-related patterns of cognitive decline among older adults – a finding that is consistent with prior research (Cloutier, Chertkow, Kergoat, *et al*., 2015; Maharani, Dawes, Nazroo, *et al*., 2019). Preliminary analyses also indicated significant racial/ethnic differences in hearing and/or cognitive function; therefore, the mixed models were estimated separately in each racial/ethnic group. Bayesian Information Criteria (BIC) and Akaike Information Criteria (AIC) statistics indicated that a quadratic function remained the best model fit for all three groups. Differences by sex were also assessed and no significant interactions were found.

All mixed models included adjustments for sex, marital status, household size, geographic region, education, employment, wealth, smoking status, alcohol consumption, ADLs, disease diagnoses, and indicators for mortality/attrition. We also tested for interactions among age, hearing loss, and hearing aid use to account for possible changes in the associations over time. The final models retain the significant associations that provided the best model fit (assessed by BIC and AIC statistics) for each racial/ethnic group. Estimates were then plotted to facilitate interpretation.

3. Results

Overall sample distributions for the study period (1998 – 2018) are shown by race/ethnicity in Table 1. Overall, White older adults exhibited significantly higher cognitive function (22.7, SD = 4.7) compared with Black (18.9, SD = 5.5) and Hispanic (19.1, SD = 5.2) older adults. A larger proportion of Hispanic older adults reported hearing loss (fair/poor hearing, 32.4%) than White (22.3%) and Black (20.1%) older adults. In terms of hearing aid use, a larger proportion of White participants (13.7%) reported wearing hearing aids compared with both Hispanic (8.3%) and Black (5.5%) participants. More White participants died during the study period (39.3%) compared to Black (35.2%) and Hispanic (27.7%) participants.

Table 2 presents the results from the mixed models showing the age-related associations among hearing level, hearing aid use, and cognitive function in White, Black, and Hispanic older adults in the study. To facilitate interpretation of the findings, the results from Table 2 are plotted in Figure 2 for four major categories of hearing and hearing aid use: (i) excellent hearing, unaided; (ii) excellent hearing, aided; (iii) poor hearing, unaided; and (iv) poor hearing, aided. For White older adults, worse self-rated hearing ($P < 0.001$) and wearing a hearing aid ($P < 0.001$) were associated with lower cognitive scores at age 65 compared to those who reported better hearing or wearing a hearing aid, respectively. White participants with excellent (unaided) hearing had an average cognitive score of 24.8 at age 65, compared to 23.9 for excellent (aided), 23.8 for poor (unaided), and 23.0 for poor (aided) among White participants. Results also showed that the differences in cognitive function associated with self-rated hearing and wearing a hearing aid diminished at later ages.

For Black older adults, worse self-rated hearing ($P < 0.001$) was associated with lower cognitive scores at age 65 compared to those who reported better hearing. The use of hearing aids was not significantly associated with cognitive function in Black participants ($P = 0.52$). In Figure 2, Black older adults with excellent (unaided) hearing had an average cognitive score of 20.9, compared to 20.8 for those with excellent (aided) hearing, 19.6 for poor (unaided), and 19.5 for poor (aided). Consistent with the finding for White older adults, the association between self-rated hearing and cognition declined among Black older adults at later ages.

For Hispanic older adults, self-rated hearing ($P = 0.008$) and hearing-aid use ($P < 0.001$) were associated with lower cognitive scores at age 65. Hispanic participants with excellent (unaided) hearing had an average cognitive score of 20.7, while those with excellent (aided) hearing had scores of 18.9, those with unaided (unaided) had scores of 20.2 and those with poor (aided) had scores of 18.4. Results also showed that the differences in cognitive function associated with wearing a hearing aid (but not self-rated hearing) diminished at later ages.

4. Discussion

In this brief report, we examined 20 years of longitudinal data to characterize how age-related changes in hearing loss and hearing aid use were associated with trajectories of cognitive decline in a large U.S. national sample. Overall, the results suggested that levels of self-reported hearing and hearing aid use had complex associations with cognitive function that varied across age and across major racial/ethnic groups in the U.S.

Three major findings were observed. First, White older adults had overall better cognitive function at all ages compared with Black and Hispanic older adults. This finding is consistent with previous research showing that Black and Hispanic populations are at higher risk for developing ADRD compared to White adults (Babulal, Quiroz, Albensi, *et al*., 2019). Second, older adults who reported worse hearing had lower levels of cognitive

Table 1. Characteristics of study participants in the Health and Retirement Study, 1998 – 2018 (*n*=20,545)

	Non-Hispanic White	Non-Hispanic Black	Hispanic
Number of participants	15,319	3,235	1,991
Number of observations	70,202	12,883	7,905
Age, mean (SD), y	73.7 (5.7)	72.6 (5.6)	72.5 (5.5)
Sex (male)	42.3	37.1	40.7
Self-rated hearing, mean (SD)	1.7 (1.1)	1.7 (1.0)	2.0 (1.1)
Self-rated hearing			
Excellent	14.5	12.9	12.0
Very good	27.2	26.1	17.3
Good	36.0	40.9	38.4
Fair	16.9	15.8	26.7
Poor	5.4	4.3	5.7
Hearing aid use (yes)	13.7	5.5	8.3
Cognitive score, mean (SD)	22.7 (4.7)	18.9 (5.5)	19.1 (5.2)
Marital status			
Married/partnered	64.0	43.2	59.8
Divorced	8.4	18.6	13.5
Widowed	24.3	32.7	23.4
Never married	2.3	5.5	3.3
Household size, mean (SD)	1.9 (0.8)	2.2 (1.3)	2.6 (1.5)
Live in South	37.6	57.7	45.6
Wealth, median, US$	248,000	50,000	50,000
Education, mean (SD), y	12.9 (2.6)	11.4 (3.3)	8.6 (4.6)
Employment status			
Employed	13.9	13.2	11.6
Unemployed	13.4	16.7	26.9
Retired	72.7	70.1	61.5
Smoking status			
Never smoker	41.7	41.7	47.6
Past smoker	48.8	46.1	42.9
Current smoker	9.6	12.2	9.5
Heavy drinking	0.9	0.9	2.5
ADL, mean (SD)	0.3 (0.8)	0.5 (1.0)	0.5 (1.1)
Ever diagnosed with:			
High blood pressure	58.5	77.6	64.5
Diabetes	19.6	34.0	36.4
Cancer	19.1	14.8	11.1
Lung disease	11.6	8.9	6.1
Heart disease	30.4	25.8	20.0
Stroke	9.6	11.2	7.4
Study follow-up			
Years of follow-up, mean (SD)	12.2 (5.2)	11.3 (5.4)	11.2 (5.3)
Died during study period	39.3	35.2	27.7

Values reported as percentages or means (standard deviation [SD]); and are reported for the pooled data.

Table 2. Mixed model estimates of age-related changes in cognitive function associated with self-reported hearing loss and hearing aid use in U.S. older adults, HRS 1998 – 2018

	Non-Hispanic White (*n*=15,319)	Non-Hispanic Black (*n*=3,235)	Hispanic (*n*=1,991)
Age	−0.11 (0.01)***	−0.14 (0.02)***	−0.10 (0.03)***
Age^2	−0.01 (0.00)***	−0.01 (0.00)***	−0.01 (0.00)***
Hearing loss	−0.24 (0.03)***	−0.32 (0.06)***	−0.13 (0.05)**
Age*Hearing loss	0.01 (0.00)***	0.02 (0.01)*	
Hearing aid	−0.84 (0.14)***	−0.14 (0.22)	−1.82 (0.50)***
Age*Hearing aid	0.08 (0.03)**		0.29 (0.10)**
Age^2*Hearing aid	−0.00 (0.00)*		−0.01 (0.00)*
Constant	17.46 (0.18)***	13.93 (0.40)***	16.82 (0.36)***
Random-effects parameters			
Variance (age)	0.05 (0.00)	0.04 (0.00)	0.05 (0.01)
Variance (intercept)	7.22 (0.18)	11.34 (0.52)	10.84 (0.62)
Covariance (age, intercept)	−0.16 (0.02)	−0.20 (0.04)	−0.24 (0.05)
Residual variance	7.57 (0.05)	9.03 (0.14)	8.18 (0.16)
BIC	371,899.30	71,148.27	43,088.90
AIC	371,569.60	70,894.50	42,844.77

HRS, Health and Retirement Study; BIC, Bayesian information criterion; AIC, Akaike information criterion. All models included sex, marital status, household size, geographic region, education, employment, wealth, smoking status, alcohol consumption, ADLs, disease diagnoses, and indicators for attrition (mortality during the study and number of waves present). *P<0.05, **P<0.01, ***P<0.001.

function than those who reported better hearing, regardless of race or ethnicity. Much of the previous research examining longitudinal changes in hearing loss and cognitive status has assessed hearing ability at only one point in time (Alattar, Bergstrom, Laughlin, *et al.*, 2019; Deal, Betz, Yaffe, *et al.*, 2017; Deal, Sharrett, Albert, *et al.*, 2015; Ge, McConnell, Wu, *et al.*, 2021; Lin, Yaffe, Xia, *et al.*, 2013), thus limiting our understanding of age-related changes in hearing status and cognition. In addition, studies that have included multiple measures of both hearing and cognitive function (Maharani, Dawes, Nazroo, *et al.*, 2018b; 2019) did not account for differences related to race and/or ethnicity. Our analysis extends this research by showing that the association between hearing and cognition is patterned differently by race/ethnicity.

Although hearing loss has a negative association with cognitive function at baseline (age 65) for all three groups, the impact of hearing loss on cognitive decline wanes with age for both White and Black adults. This pattern is generally consistent with the age-as-leveler hypothesis (House, Lepkowski, Kinney, *et al.*, 1994), which suggests that stressors (e.g., hearing loss) may be more impactful for health (e.g., cognition) at earlier ages, but dimmish at older ages. For Hispanic older adults, however, we found that the association between level of hearing and cognitive function persisted across age. This pattern is consistent with the notion of persistent inequality, which suggests that the magnitude of the health differential remains stable across age (Haas and Rohlfsen, 2010).

Finally, the findings for hearing aid use differed by racial/ethnic group. The results suggested that wearing hearing aids may be protective for White older adults in slowing cognitive decline among those in the early old age. This general finding is consistent with previous research showing that individuals experience a more gradual decrease in episodic memory decline after beginning to wear hearing aids compared to before (Maharani, Dawes, Nazroo, *et al.*, 2018a). Hearing aid use appears particularly beneficial for Hispanics – those who report either excellent or poor aided hearing experience a small improvement in cognitive functioning from ages 65 to 75. In contrast, our results suggest that self-reported hearing aid use is not particularly impactful for cognitive functioning among Black adults. We suspect this finding may be due, in part, to the relatively low prevalence of hearing aid use among Black adults (5.5%), which has been well documented. The research has shown that the average delay in adopting hearing aids following hearing aid candidacy is nearly 10 years, ranging from 8.6 years for White adults to 15.2 years for non-White adults (Simpson, Matthews, Cassarly, *et al.*, 2019). Thus, the low/delayed adoption of hearing aids among Black adults may result in these individuals missing the cognitive benefits associated with hearing aids.

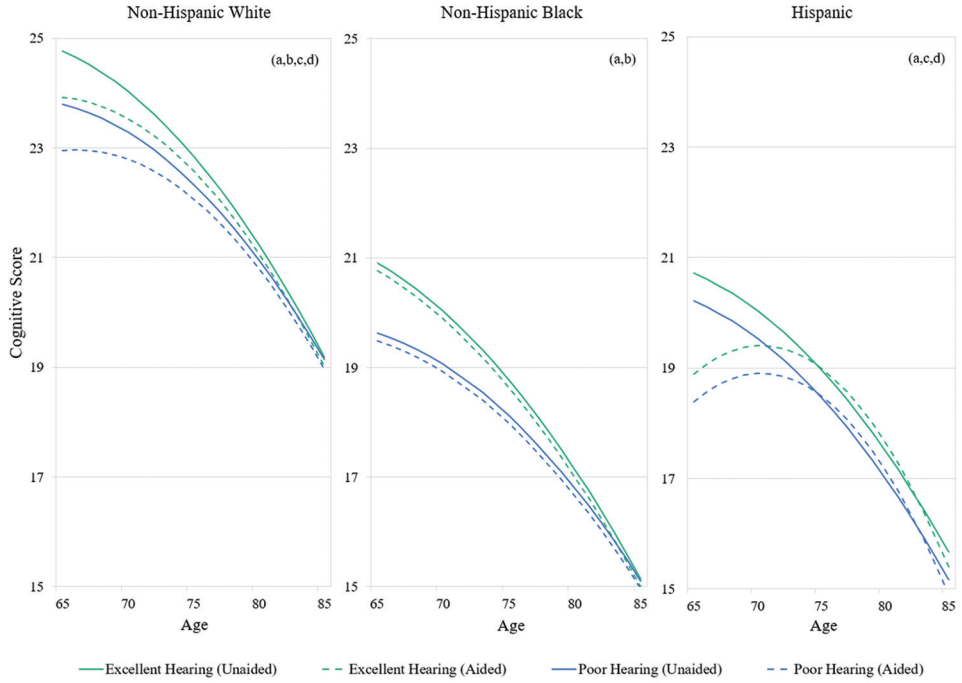


Figure 2. Plots of age-related changes in cognitive function associated with self-reported hearing loss and hearing aid use in U.S. older adults, HRS 1998 – 2018.
HRS: Health and Retirement Study. Results were calculated from estimates reported, as shown in Table 2. (a) Statistically significant intercept difference ($P < 0.05$) related to hearing loss. (b) Statistically significant slope difference ($P < 0.05$) related to hearing loss. (c) Statistically significant intercept difference ($P < 0.05$) related to hearing aid use. (d) Statistically significant slope difference ($P < 0.05$) related to hearing aid use.

Several limitations of this analysis should be noted. First, hearing function and hearing aid use were self-reported measures. Although pure-tone audiometry is the gold standard for clinically assessing hearing sensitivity (West, Smith, and Dupre, 2020), studies also suggest that it may not be able to measure the experience of hearing disability in real world environments (Demeester, Topsakal, Hendrickx, *et al.*, 2012), especially regarding an individual's reported listening comprehension in group conversations (Gatehouse and Noble, 2004) or in noisy environments (Kramer, Kapteyn, Festen, *et al.*, 1996). Furthermore, audiometric data are rarely assessed on repeated occasions, particularly among the general population at a national level. A second limitation of this study is the temporality of the measures. Although the current data are longitudinal, measures of hearing and cognitive function were both assessed at the same time points (i.e., interviews) over the study period. Therefore, the results of this study should be considered associational, and we remain guarded in forming causal interpretations of the findings. Finally, although we extend prior research on hearing and cognitive function by examining major population subgroups (non-Hispanic White, non-Hispanic Black, and Hispanic older adults), we lacked sufficient data on other U.S. racial/ethnic groups. There is some evidence to suggest that Asian Americans with hearing loss are less likely to receive a hearing test compared with White, Black, and Hispanic individuals; and substantially less likely to use hearing aids compared to White adults (Choi, Kari, Friedman, *et al.*, 2018). Additional research is needed to further understand these racial/ethnic differences in hearing, access to hearing healthcare, and their possible implications for changes in cognitive function at older ages.

5. Conclusion

In summary, the current findings extend previous research on the link between hearing loss and cognition by providing national-level evidence of age-related changes in hearing status and cognition, and how these associations vary by major racial/ethnic groups in the United States. Timely identification of hearing loss and subsequent acquisition of hearing aids are important considerations for reducing the burden of ADRD.

Funding

J.S. West is supported by the Duke Aging Center Postdoctoral Research Training Grant (NIA T32 AG000029).

Conflicts of interest

The authors have no conflicts of interest to declare.

Authors' contributions

Conceived and designed the study and analyzed the data: J.S. West, M.E. Dupre. Wrote the manuscript: J.S. West, M.E. Dupre, S.L. Smith.

Ethical approval

This study did not involve research on human subjects.

Availability of supporting data

The data used in this study are publicly available.

References

Agrawal Y, Platz EA, and Niparko JK. (2008). Prevalence of hearing loss and differences by demographic characteristics among US adults: Data from the national health and nutrition examination survey, 1999-2004. *Archives of Internal Medicine*, 168(14):1522-1530.

https://doi.org/10.1001/archinte.168.14.1522

Alattar AA, Bergstrom J, Laughlin GA, *et al.* (2019). Hearing impairment and cognitive decline in older, community-dwelling adults. *The Journals of Gerontology Series A Biological Sciences and Medical Sciences*, 75(3):567-573.

https://doi.org/10.1093/gerona/glz035

Babulal GM, Quiroz YT, Albensi BC, *et al.* (2019). Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: Update and areas of immediate need. *Alzheimers and Dementia*, 15(2):292-312.

Brenowitz WD, Kaup AR, Lin FR, *et al.* (2019). Multiple sensory impairment is associated with increased risk of dementia among black and white older adults. *The Journals of Gerontology Series A, Biological Sciences and Medical Sciences*, 74(6):890-896.

https://doi.org/10.1093/gerona/gly264

Brown TH, O'Rand AM, and Adkins DE. (2012). Race-ethnicity and health trajectories: Tests of three hypotheses across multiple groups and health outcomes. *Journal of Health and Social Behavior*, 53(3):359-377.

https://doi.org/10.1177/0022146512455333

Bugliari D, Nancy C, Chris C, *et al.* (2016). *RAND HRS Data Documentation, Version P*. Labor and Population Program: RAND Center for the Study of Aging.

CDC. (2017). Table A-6a. Age-Adjusted Percentages (with Standard Errors) of Hearing Trouble, Vision Trouble, and Absence of Teeth among Adults Aged 18 and Over, by Selected Characteristics: United States, 2017. In: *Summary Health Statistics: National Health Interview Survey, 2017*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

https://doi.org/10.3886/icpsr09228

Choi JS, Kari E, Friedman RA, *et al.* (2018). Prevalence of hearing loss and hearing care use among asian americans: A nationally representative sample. *Otology and Neurotology*, 39(2):158-167.

https://doi.org/10.1097/mao.0000000000001668

Cloutier S, Chertkow H, Kergoat MJ, *et al.* (2015). Patterns of cognitive decline prior to dementia in persons with mild cognitive impairment. *Journal of Alzheimer's Disease*, 47(4):901-913.

https://doi.org/10.3233/jad-142910

Curhan SG, Willett WC, Grodstein F, *et al.* (2019). Longitudinal study of hearing loss and subjective cognitive function decline in men. *Alzheimers and Dementia*, 15(4):525-533.

https://doi.org/10.1016/j.jalz.2018.11.004

Deal JA, Betz J, Yaffe K, *et al.* (2017). Hearing impairment and incident dementia and cognitive decline in older adults: The health ABC study. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 72(5):703-709.

https://doi.org/10.1093/gerona/glw069

Deal JA, Sharrett AR, Albert MS, *et al.* (2015). Hearing impairment and cognitive decline: A pilot study conducted within the atherosclerosis risk in communities neurocognitive study. *American Journal of Epidemiology*, 181(9):680-690.

https://doi.org/10.1093/aje/kwu333

Demeester K, Topsakal V, Hendrickx JJ, *et al.* (2012). Hearing disability measured by the speech, spatial, and qualities of hearing scale in clinically normal-hearing and hearing-impaired middle-aged persons, and disability screening by means of a reduced SSQ (the SSQ5). *Ear and Hearing*, 33(5):615-616.

https://doi.org/10.1097/aud.0b013e31824e0ba7

Gatehouse S and Noble W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology*, 43(2):85-99.

https://doi.org/10.1080/14992020400050014

Gauthier S, Rosa-Neto P, Morais JA, *et al.* (2021). *World Alzheimer Report 2021: Journey Through the Diagnosis of Dementia*. London, England: Alzheimer's Disease International.

Ge S, McConnell ES, Wu B, *et al.* (2021). Longitudinal association between hearing loss, vision loss, dual sensory loss, and cognitive decline. *Journal of the American Geriatrics Society*, 69(3):644-650.

https://doi.org/10.1111/jgs.16933

Golub JS, Brickman AM, Ciarleglio AJ, *et al.* (2020). Audiometric age-related hearing loss and cognition in the hispanic community health study. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 75(3):552-560.

https://doi.org/10.1093/gerona/glz119

Golub JS, Luchsinger JA, Manly JJ, *et al.* (2017). Observed hearing loss and incident dementia in a multiethnic cohort. *Journal of the American Geriatrics Society* 65(8):1691-1697.

https://doi.org/10.1111/jgs.14848

Haas SA and Rohlfsen L. (2010). Life course determinants of racial and ethnic disparities in functional health trajectories. *Social Science and Medicine*, 70(2):240-250.

https://doi.org/10.1016/j.socscimed.2009.10.003

House JS, Lepkowski JM, Kinney AM, *et al.* (1994). The social stratification of aging and health. *Journal of Health and Social Behavior*, 35(3):213-234.

https://doi.org/10.2307/2137277

Kramer SE, Kapteyn TS, Festen JM, *et al.* (1996). The relationships between self-reported hearing disability and measures of auditory disability. *Audiology*, 35(5):277-287.

https://doi.org/10.3109/00206099609071948

Lin FR, Metter EJ, O'Brien RJ, *et al.* (2011). Hearing loss and incident dementia. *Archives of Neurology*, 68(2):214-220.

Lin FR, Yaffe K, Xia J, *et al.* (2013). Hearing loss and cognitive decline in older adults. *JAMA Internal Medicine*, 173(4):293-299.

https://doi.org/10.1001/jamainternmed.2013.1868

Lin FR. (2011). Hearing loss and cognition among older adults in the United States. *The Journals of Gerontology: Series A Biological Sciences and Medical Sciences*, 66(10):1131-1136.

https://doi.org/10.1093/gerona/glr115

Livingston G, Sommerlad A, Orgeta V, *et al.* (2017). Dementia prevention, intervention, and care. *The Lancet*, 390(10113):2673-2734.

Maharani A, Dawes P, Nazroo J, *et al.* (2018a). Longitudinal relationship between hearing aid use and cognitive function in older americans. *Journal of the American Geriatrics Society*, 66(6):1130-1136.

https://doi.org/10.1111/jgs.15363

Maharani A, Dawes P, Nazroo J, *et al.* (2018b). Visual and hearing impairments are associated with cognitive decline in older people. *Age and Ageing*, 47(4):575-581.

https://doi.org/10.1093/ageing/afy061

Maharani A, Dawes P, Nazroo J, *et al.* (2019). Associations between self-reported sensory impairment and risk of cognitive decline and impairment in the health and retirement study cohort. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 75(6):1230-1242.

https://doi.org/10.1093/geronb/gbz043

McKee MM, Stransky ML, and Reichard A. (2018). Hearing loss and associated medical conditions among individuals 65 years and older. *Disability and Health Journal*, 11(1):122-125.

https://doi.org/10.1016/j.dhjo.2017.05.007

Ofstedal MB, Fisher GG, and Herzog AR. (2005). *Documentation of Cognitive Functioning Measures in the Health and Retirement Study*. Ann Arbor, MI: Survey Research Center, University of Michigan.

https://doi.org/10.7826/isr-um.06.585031.001.05.0010.2005

Rabe-Hesketh S and Skrondal A. (2012). *Multilevel and Longitudinal Modeling Using Stata, Volume I: Continuous Responses*. 3rd ed. College Station, Texas: StataCorpLP.

https://doi.org/10.1177/0962280214527532

RAND Center for the Study of Aging. (2021). *RAND HRS Longitudinal File 2018* (*V1*). Available from: https://www.rand.org/well-being/social-and-behavioral-policy/centers/aging/dataprod/hrs-data.html [Last accessed on 2022 May 25].

Simpson AN, Matthews LJ, Cassarly C, *et al.* (2019). Time from hearing aid candidacy to hearing aid adoption: A longitudinal cohort study. *Ear and Hearing*, 40(3):468-476.

https://doi.org/10.1097/aud.0000000000000641

Sonnega A, Faul JD, Ofstedal MB, *et al.* (2014). Cohort profile: The health and retirement study (HRS). *International Journal of Epidemiology*, 43(2):576-585.

https://doi.org/10.1093/ije/dyu067

Wallhagen MI, Strawbridge WJ, and Shema SJ. (2008). The relationship between hearing impairment and cognitive function: A 5-year longitudinal study. *Research in Gerontological Nursing*, 1(2):80-86.

https://doi.org/10.3928/19404921-20080401-08

Wayne RV and Johnsrude IS. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Research Reviews*, 23(Pt B):154-166.

https://doi.org/10.1016/j.arr.2015.06.002

West JS, Smith SL, and Dupre ME. (2020). Hearing loss. In: D Gu and M Dupre (eds.). *Encyclopedia of Gerontology and Population Aging*. Cham: Springer International Publishing. p.1-11.

https://doi.org/10.1007/978-3-319-69892-2_1135-1

Whitson HE, Cronin-Golomb A, Cruickshanks KJ, *et al.* (2018). American geriatrics society and national institute on aging bench-to-bedside conference: Sensory impairment and cognitive decline in older adults. *Journal of the American Geriatrics Society*, 66(11):2052-2058.

https://doi.org/10.1111/jgs.15506

Yang Y and Land KC. (2013). *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications*. Boca Raton, FL: CRC Press.

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Association between school dropouts, early marriages, childbearing, and mental health in early adulthood of women: Evidence from a cohort study in Bihar, India

Ravita Yadav[1], Preeti Dhillon[2], Archana Kujur[3], and Sangram Kishor Patel[4]*

[1]International Institute for Population Sciences, Mumbai, India

[2]Department of Survey Research and Data Analytics, International Institute for Population Sciences, Mumbai, Maharashtra, India

[3]Department of Geography, Central University of Karnataka, Aland Road, Kalaburagi, Karnataka, India

[4]Poverty, Gender and Youth Program, Population Council, New Delhi, India

Abstract

School dropouts, early marriages, and low age at childbearing are issues still prevalent in Indian states like Bihar, which may be responsible for poor mental health among young adults. The present study examined the effect of life-course events such as school dropouts, early marriages, and early childbearing on mental health status at later ages (23–28 years). Using data from the Understanding the Lives of Adolescents and Young Adults consisting of a 2360 adolescent (ages 15 – 19) girl cohort interviewed in 2007 and re-interviewed at ages 23 – 28 in 2016 from the state of Bihar, India, we applied ordinal logistic regression models in analyzing factors associated with mental health status. Women who never attended school, or dropped out from school, and who got married before age 19 showed a poorer mental health status in their young adulthood (22 – 28 years) compared to their respective counterparts who attended a school and who got married at age 19 or older. As compared to women who had a child before age 19, those who did not have any child, or who had children after 20 years of age were more likely to have poor mental health. Working women, high self-efficacy of women, and women who have decision-making power showed better mental health outcomes as compared to their respective counterparts. To enhance psychological well-being of young women, the study recommends continue education and delaying marriage as the programmatic keys with attention to improving young women's autonomy and gender role attitudes and reducing societal pressure for bearing first child soon after marriage.

Keywords: Mental health; Depression; School dropout; Early marriage; Early childbearing; Adolescents; Women empowerment; India

***Corresponding author:**
Sangram Kishor Patel
(sangramkishor@gmail.com)

Citation: Yadav R, Dhillon P, Kujur A, Patel SK. (2022). Association between school dropouts, early marriages, childbearing, and mental health in early adulthood of women: Evidence from a cohort study in Bihar, India. *International Journal of Population Studies*, 8(1):25-37. https://doi.org/10.36922/ijps.v8i1.1280

Received: January 17, 2022

Accepted: June 14, 2022

Published Online: June 27, 2022

Copyright: © 2022 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

According to the World Health Organization, globally, 14% of adolescents (10 – 19) suffer from mental health conditions (WHO, 2021) and nearly 9.8 million young Indians

aged between 13 and 17 years need active interventions (Gururaj, Varghese, Benegal, *et al.*, 2016). The burden of mental health problems in India is 2443 disability-adjusted life years (DALYs) per 10,000 population and the age-adjusted suicide rate is 21.1 per 10,000 population (Prakash, 2021). According to the National Mental Health Survey 2015 – 2016, conducted by the National Institute of Mental Health and Neurosciences, 9.8 million teenagers in the age-group 13 – 17 years suffer depression and other mental health issues (Roy, Bharati, and Chakravarty, 2019). Adolescence is a time of change and transition: Changes to hormones and to the body, changes in the social environment, and changes to the brain and the mind, all occur in this transition phase of life (Blakemore, 2019). Post-traumatic disorder, exposure to violence, and a combination of genetic and environmental influences at an early age lead to higher rates of adult mental health problems, including personality disorders, major depression, substance use, and suicidal behavior (Chauhan, Srivastava, Kumar, *et al.*, 2021; Kumar, Srivastava, Mishra, *et al.*, 2020; Patel, Saggurti, Pachauri, *et al.*, 2015). Many mental health problems emerge in mid-to-late-adolescence. According to a large-scale meta-analysis, worldwide, the onset of the first mental problem occurs before age 14 in one-third of individuals; and by age 18 half of them get affected, and before age 25, it is seen in nearly 63% population (Solmi, Radua, Olivola, *et al.*, 2021).

School is a place where a child spends most of their time and it is a known fact that during this time, cognitive, social, and emotional skill development takes place. According to the chief-executive officer of child rights and you, the dropout rate of girls rises significantly after secondary education. One in every five girls enrolled drops out after class eighth (Deb, Chatterjee, and Walsh, 2020). Stress to perform well in academics, household factors, early marriages and pregnancies, cultural factors, substance use, low self-esteem, and feeling of isolation from the school environment are some of the core reasons for early dropouts (Dangal, 2006; Kishore and Shaji, 2012; Shahidul and Karim, 2015). Young women cannot enjoy their day-to-day life normally because they have high discipline levels and more remarkable persistence to achieve their goals and find success (Hjorth, Bilgrav, Frandsen, *et al.*, 2016). A study in India shows that for women, merely participating in secondary education may lead to life-long autonomy and participation in decision-making (Marphatia, Reid, and Yajnik, 2019). Among the member countries of Organization for Economic Co-operation and Development, around 20% of young people end their education before reaching the upper secondary level, which means one in five students have a higher risk of facing unemployment, morbidity, mortality, and poverty compared with their cohorts who complete their education. The State of Bihar of India, which had a literacy rate of 61.8% (overall), 71.2% for men, and 51.5% for women, according to the 2011 census, was one among the four states which performed poorly in school dropouts. The other three states are Jharkhand, Uttar Pradesh, and Arunachal Pradesh.

Another vital life event is a marriage that plays a key role in the transition to adulthood that affects the lives of young people. Marriage is widely recognized as a benefit for mental health due to good social support. However, the apparent benefit may vary across the lifetime of an individual. The timing of an event may be more consequential than its occurrence (Elder, 1995); this may be especially true for marriage where it is culturally agreed that there is an appropriate time to do it. Early or child marriage is a strong social custom particularly for girls in South Asian countries, including India where the mean age of marriage for girls married below age 18 was 16.5 years in 2011 (ORG and CCI, 2015). However, there has been an improvement in recent years in child marriage where the percentage of women married before age 18 has reduced from 27% in 2015–2016 to 23% in 2019–2021 (IIPS and ICF, 2020). However, still with early marriage, the transition from childhood into adulthood is lost for many young women, particularly in states like Bihar. Although early marriage affects both boys and girls, this transition in India, where mostly women leave their parental house, comes with the change in status and position within the household, and a onset of early childbearing. Therefore, early school dropouts, early marriage, early childbearing affect psychological well-being; however, these associations over the early life course have not been examined in an Indian setting. Family characteristics such as early divorce, parental education, and occupation also predict early dropouts and also predict childhood psychiatric disorders (Gilman, Kawachi, Fitzmaurice, *et al.*, 2002; Le Strat, Dubertret, and Le Foll, 2011; Lorant, Deliège, Eaton, *et al.*, 2003; Muntaner, Eaton, Diala, *et al.*, 1998; Muntaner, Eaton, Miech, *et al.*, 2004).

Mental health remains a neglected issue in India. The data constraint and limited understanding of the causal effects of mental illness and life cycle events such as school dropouts, marriage, and childbearing do not allow for interventions to tackle this problem. This study attempts to understand the various life-course events such as educational attainment or dropout, marriage, and childbearing at early ages on reported mental health status of women in early adulthood.

2. Data and Methods

2.1. Data sources

We used data from the Understanding the Lives of Adolescents and Young Adults (UDAYA), a state-level longitudinal survey for Uttar Pradesh and Bihar, India, conducted by the Population Council. UDAYA was designed to provide estimates for the state as a whole as well as for the urban and rural areas of the state instead of district or sub-district levels. This paper is based on data drawn from a survey of Bihar participants in Youth in India: Situation and Needs Study (IIPS and Population Council, 2010) conducted in six states (Andhra Pradesh, Bihar, Jharkhand, Maharashtra, Rajasthan, and Tamil Nadu) in 2006–2007 (Youth stand adolescents who aged 15 – 19 at that time were followed up after almost 8–9 years later in 2015–2016 by Population Council). The follow-up study was done to assess factors that determine accumulation or losses of assets and adolescents' quality of transitions to adulthood. The corresponding data set provides longitudinal information for adolescents who were 15–19 years old in 2007 and were young adults aged 23–28 years in 2015–2016. Three categories of adolescents were tracked – unmarried adolescent girls and boys aged 15–19 in 2007, and married girls aged 15–19 in 2007 (Santhya, Acharya, Pandey, *et al.*, 2017). In total, 601 married girls, 1759 unmarried girls, and 563 unmarried boys were followed up successfully in 2015–2016. To fulfill the targeted objectives, the study has extracted 2360 sample sizes (only girls), out of total female sample of 3188 at baseline. The newly generated datasets included all girls who belonged to the 15–19 age-group in the baseline (2007) and the same girl cohort who were in the age group 23 – 28 years, in the follow-up (2015 – 2016).

2.2. Measures

2.2.1. Outcome variables

Mental health status: The outcome variable was constructed from the measure of psychological well-being that was obtained from the 12-item General Health Questionnaire (GHQ-12), a popular and widely used and validated scale to measure the mental health status of the population (Montazeri, Harirchi, Shariati, *et al.*, 2003). The questionnaire consists of 12 items, each one assessing the severity of a mental problem over the past 1 month using a dichotomous (0–1) scoring style. The positive items were scaled as 0 for "Yes" and 1 for "No," and the negative ones as 1 for "Yes" and 0 for "No." The 12-items are: (a) Being able to concentrate; (b) lost sleep over worry; (c) feeling that you are playing a useful role; (d) feeling that you are capable of making a decision; (e) feeling constantly under strains; (f) feeling that you can't overcome difficulties; (g) being able to enjoy your normal day-to-day activities; (h) being able to face up to your problems; (i) feeling unhappy and distressed; (j) losing confidence in yourself; (k) thinking of yourself as a worthless person; and (l) feeling reasonably happy, all things considered. The score was used to generate a total score ranging from 0 to 12. Using the generated scores, a mental health status index was constructed, which was divided into three categories (Goldberg and Williams, 1988). The first category with score "0" is considered "normal mental health," the second category with scores "1-6" is considered "moderate mental health," and the third category "7-12" is considered "poor mental health." The internal consistency of the items was assessed using Cronbach's alpha, and alpha equal to or > 0.75 was considered satisfactory (Taber, 2018). Therefore, these mental health questions were suitable for constructing mental health index.

2.2.2. Explanatory variable

2.2.2.1. Educational attainment

The variable defined as current status of schooling was computed at both baseline and follow-up surveys. The education variable considers the status of female schooling at the time of the survey and is classified into three categories. The first category is women who never attended school (never attended). The second category is dropouts-women who dropped out of school before completing 12 standards (dropout before completing 12 class), and the third category consists of women who completed or were continuing their schooling at the time of the survey (completed or continuing).

2.2.2.2. Marriage and childbearing-related variables

Marital status (never married, currently married), the age at first marriage (below 19 years, 19 years or above), age at first birth (below 20 years, 20 years and above, no child, unmarried), parity, and age of last child (no child, 2 years and below, above 2 years, and unmarried) were measured at follow-up data. Age at first birth was not readily available in the data; therefore, this variable was constructed by subtracting the age of the first child from the age of the woman and then categorized as – women who gave birth to their first child before age 20; and after 20 years. Parity is defined as the total number of times a woman gives birth to a child, alive, or dead. In the case of parity, unmarried women were categorized with the women having zero parity to find out the mental health status of women having zero parity as compared to women having parity one or two or three plus.

The age of the last child was computed using the birth order and age of every child. This variable was useful to

compare the mental condition of women with no child, with children < 2 years of age and children older than 2 years of age. This variable is used to compare the impact of raising younger children and children older than 2 years on the women's mental health since younger children require more attention and care from their mothers.

2.2.3. Control variables from baseline

Background characteristics variables of the females taken from the baseline survey are as follows; Religion (Hindu and Non-Hindu), caste (scheduled caste/scheduled tribes, other backward class [OBC], and General), family type (Nuclear family and Non-nuclear family), parents' education (educated and not educated) parents' working status (No and Yes), and parents' life (One of the parents or both died and both are alive). Data provided a household's wealth index variable and constructed from the household's conditions and amenities (IIPS and Population Council, 2010).

2.2.4. Control variables from follow-up

Variables at the follow-up survey are also the key characteristics of the respondents that could be closely linked with their mental health. The following variables are included respondent's age (23–24, 25–26, and 26–27), place of residence (rural and urban), household wealth index (Poor, Middle Rich), working status (Yes vs. No), self-efficacy (High vs. low), and decision-making power (Yes vs. No). The self-efficacy is a variable that is determined by combining two variables, expressing an opinion to older adults in the family, confronting if something went wrong, that is, (1) whether the respondent is able to express an opinion (1-never, 2-sometimes, and 3-often) to older persons in the family and (2) whether she confronts (1-stay quiet, 2-sometimes confront, and 3-always confront) if someone says or does something wrong to her. We constructed self-efficacy score by summing the responses to these questions and then dichotomized it as high versus low. This question was asked by both unmarried and married women. We considered four items relevant to decisions about matters related to their own lives, all are measured at follow-up – 'decision in spending money, about making major household purchases, about health care for herself and whether she should work or stay at home. The score was assigned as follows: (1) Others only, (2) jointly with others, (3) alone, and an additive score for decision-making was constructed by summing the three responses from the above-mentioned items. The higher the score, the higher the decision-making in adolescence.

Husband's education (no education, primary, secondary, higher secondary and above, DK) and spousal age gap were also modeled. Spousal age gap (at the follow-up survey) is calculated by subtracting the wife's age from the husband's age and was categorized in four groups "wife older than husband or wife 2 years younger than husband)," "wife three to 6 years younger to husband," "wife more than 6 years younger than husband," and "do not know" (DK) (husband's age was not reported). The categorization of spousal age gap is done based on the distribution of data. Further, being younger than husband (more than 6 years) is also a proxy indicator of low level of empowerment of young girls. Aforesaid variables – such as education, self-efficacy, decision-making power, working status, and spousal age gap are proxy indicators of empowerment of young women.

2.3. Statistical Analysis

We used bivariate analyses (cross-tabulations and Chi-square tests) and ordered logistic regression (OLR) models to examine school dropouts, early marriages, and early childbearing on mental health status at follow-up. OLR is being used when the dependent variable is ordinal (i.e., the variable has a meaningful order with more than two categories or levels). Here, the dependent variable (Mental Health status) has three ordinal categories in nature, that is, normal, moderate, and poor mental health status. The poor mental health status is coded as the highest rank, whereas the normal is coded as the lowest rank. Three multivariate models were fitted to understand the effects of sociodemographic variables measured in adolescence (Molde-1), marriage and childbearing in adolescence (Model-2), and variables measured in young adulthood (Model-3) on mental health outcomes in young adulthood at follow-up. All analysis was carried out using STATA 15.0.

3. Results

The prevalence of women's mental health status by GHQ-12 at baseline and follow-up surveys were shown in Table 1. Cronbach alpha values suggested a higher consistency in reporting depressive symptoms in both rounds. The pattern of reported symptoms changed considerably over the period. During adolescence, the statement "Not felt capable of making decisions" was reported by almost 21% of adolescent girls, although such a proportion reduced to merely 13% at follow-up. Similarly, the percentage of girls who felt cannot overcome difficulties reduced from 15% to 12% over the cohort. On the other hand, to all other statements reporting of depressive symptoms increased over the period. For instance, girls who felt under strain increased from 6% in adolescence to 19% in follow-up, and those who were unhappy and depressed increased from 5% to 18%. Consistently, almost 10% of girls reported that they were not able to face up problems at both rounds of the survey. Overall, mental health status of adolescents worsened from adolescence (15–19) to young adulthood

(23–28), while worsening was more visible in unmarried girls (Figure 1).

Findings of the frequency distributions and the associations between variables measured at adolescence and the mental health outcomes at young adulthood are presented in Table 2. The results revealed that school attendance at adolescence showed a positive association with later mental health status, which was indicated by the likelihood of poor mental health. As compared to girls who never attended school, those who dropped out from school in adolescence had lower adjusted odds ratio (AOR) (0.80, 95% CI: 0.64, 0.99) and those who continued their schooling had lower AOR (0.73, 95%CI: 0.59, 0.91).

The effect of early marriage and childbearing on mental health status in early adulthood (at age 23–28 years in follow-up) after controlling for the effect of socioeconomic variables is shown in Table 3 which contains results from bivariate and multivariate OLR. Age at the first marriage equal to or greater than 19 was associated with 29% (AOR: 0.71, 95% CI: 0.56, 0.90) lower odds as compared to ages at the first marriage younger than 19. In comparison with girls who had the first childbirth before age 20 years, those who experienced childbearing after age 20 had nearly 24% higher risk (AOR: 1.24, 95%CI 1.00, 1.54) and those who did not have any child had 64% higher risk (AOR: 1.64, 95%CI: 1.15, 2.35) of having poor mental health status at age 23–28 years. In addition, it reduced with the increase in number of children a women have.

Table 4 presents the results from the analysis to assess the effect of current schooling status and other women

Table 1. Response to statements of GHQ-12 by adolescent girl's cohort at baseline and follow-up surveys in Bihar, India.

GHQ-12	Adolescence (Baseline)	Young adulthood (Follow-up)
	N=2360	N=2360
Not able to concentrate	1.5	8.3
Lost sleep over worry	4.2	13.8
Not felt playing a useful role	4.1	7.3
Not felt capable of making decisions	21.1	13.2
Felt constantly under strain	5.6	18.8
Felt cannot overcome difficulties	14.5	11.5
Not able to enjoy a normal day to day activities	1.9	10.5
Not able to face up problems	10.4	9.8
Unhappy and depressed	4.5	17.6
Losing self-confidence	4.3	9.1
Thinking of self as worthless	5.0	9.1
Reasonably not happy	2.5	13.7
Cronbach alpha	0.64	0.85

Marital status was taken from the baseline survey.

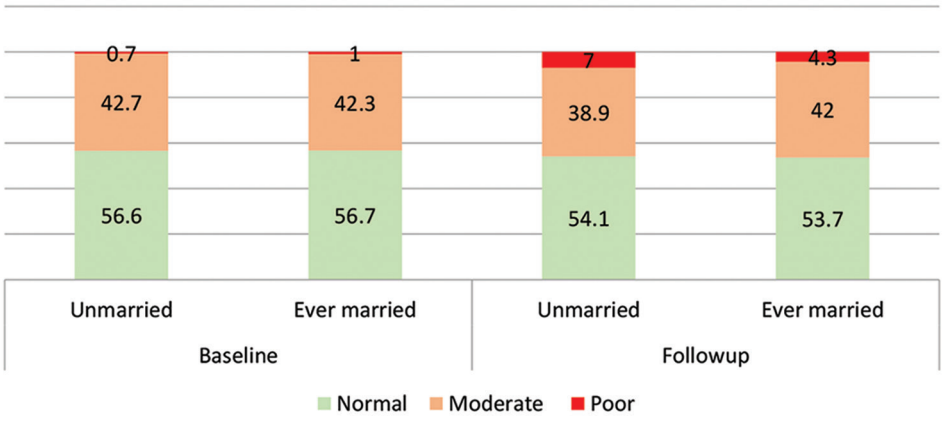


Figure 1. Changes in the mental health status of unmarried and ever married girls from adolescence to early adulthood (from baseline to follow-up).

Table 2. Effects of factors measured in adolescence on mental health status in later young adulthood (aged 23–28 years) of women in Bihar

Background characteristics measured at baseline (at age 15–19 years)	Mental health status at follow-up (%)				N	Results from ordered logistic regression[#]
	Normal	Moderate	Poor	Level of significance[##]		AOR (95% CI)
School attendance						
Never attended®	55.4	38.7	5.9	***	576	1.00
Dropped out	49.9	43.3	6.8		988	0.80 (0.64,0.99)**
Completed or continuing	58.1	36.2	5.8		796	0.73 (0.59,0.91)**
Religion						
Hindu®	52.5	41.1	6.4	*	1,971	1.00
Muslim	61.2	33.3	5.6		389	0.75 (0.59,0.92)***
Caste						
SC/ST®	50.5	42.0	7.5	**	488	1.00
OBC	54.6	39.4	6.1		1,549	0.92 (0.71,1.16)
General	56.5	38.3	5.2		323	0.97 (0.61,1.28)
Family type						
Nuclear®	55.2	38.3	6.5		1,074	1.00
Non-nuclear Family	53.0	41.0	6.1		1,286	1.00 (0.88,1.23)
Mother's education						
Not educated®	52.7	40.8	6.5	***	1,954	1.00
Educated	60.1	34.7	5.3		406	0.76 (0.58,0.95)**
Father's education						
Not educated®	53.0	40.9	6.1		1,213	1.00
Educated	55.0	38.6	6.5		1,147	1.11 (0.94,1.37)
Mother's working status						
No	54.4	39.6	6.0		1,129	-
Yes	53.6	39.9	6.5		1,231	
Father's working status						
No®	46.8	44.7	8.6	***	271	1.00
Yes	54.9	39.1	6.0		2,089	0.65 (0.53,1.10)**
Parent's survival status						
One of the parents or both died®	51.5	42.3	6.2		259	1.00
Both alive	54.3	39.4	6.3		2,101	1.16 (0.93,2.17)
Total	54.0	39.7	6.3		2,360	

[#]Based on multivariate ordinal logistic regression models. AOR, adjusted odds ratios are based on controlling all variables in the table; ®: reference category; CI: confidence interval; [##]based on Chi-square test; ***$P < 0.01$; ** $P < 0.05$; * $P < 0.10$.

empowerment-related variables measured at young adulthood after controlling socioeconomic variables measured in adolescence. In comparison to the girls who never attended school, dropped out girls from school had 23% (AOR: 0.77, 95% CI: 0.62, 0.94), and who were continuing schooling till follow-up had 45% (AOR: 0.55, 95% CI: 0.39, 0.79) lower risk of being in poor mental health status.

The socioeconomic control variables also show the effect on mental health outcomes in early adulthood of women (Table 2). The young adult women from Muslim religion had a 28% (AOR = 0.72, 95% CI: 0.56, 0.92) lower risk of having poor mental health than the Hindus. A higher risk of being in poor mental health status was observed in early adulthood of women from SCs/STs and OBCs as compared to general category. The education and occupation of parents also affect the mental health status of children. The risk of being in poor mental health status was 26% (AOR = 0.74, 95% CI: 0.58, 0.95) lower in young adulthood

Table 3. Effects of marriage and childbearing in adolescence on mental health status in later young adulthood (aged 23–28 years) of women in Bihar, India

Background characteristics measured at follow up	Mental health status (%)				N	Results from ordered logistic regression[#]
	Good	Moderate	Poor	Level of significance[##]		AOR (95% CI)
Age						
23–24®	53.0	40.0	7.0	***	1,311	1.00
25–26	56.0	38.3	5.7		727	0.98 (0.82,1.18)
27–28	53.3	42.2	4.4		322	0.82 (0.63,1.04)
Marital Status						
Never married®	54.0	39.2	6.8	***	155	1.00
Ever married	54.0	39.8	6.2		2,204	0.82 (0.44,1.52)
Age at first marriage						
<19 years®	52.4	41.5	6.1	***	1,603	1.00
≥19 years	58.2	35.2	6.6		602	0.71 (0.56,0.90)***
Unmarried	54.0	39.2	6.8		155	1.00
Husband's education						
No education®	55.2	38.2	6.6	***	547	1.00
Primary	56.8	37.7	5.5		597	1.10 (0.85,1.42)
Secondary and above	54.5	40.1	5.4		938	1.33 (1.03,1.72)**
DK	30.7	55	14.2		122	2.61 (1.75,3.91)***
Unmarried	54.0	39.2	6.8		155	1.00
Spousal age gap						
Below 2 years®	52.4	41.3	6.3	***	593	1.00
3–6 years	60.2	35.9	3.9		937	0.79 (0.64,0.99)**
7+years	56.2	37.4	6.4		307	1.04 (0.78,1.39)
DK	38.8	49.2	12		368	1.62 (1.22,2.15)***
Unmarried	54.0	39.2	6.8		155	1.00
Age at first birth						
<20 years®	55.9	39.9	4.3		910	1.00
≥20 years	55.2	38.6	6.2		963	1.24 (1.00,1.54)**
No child	49.1	40.3	10.6		333	1.64 (1.15,2.35)***
Unmarried	54.0	39.2	6.8		155	1.00
Parity						
0	49.1	40.3	10.6	**	333	
1	52.7	39.6	7.7		409	-
2	56.6	38.2	5.2		721	-
3+	54.2	40.9	4.9		897	-
Age of last child						
No child®	49.1	40.3	10.6		333	1.00
≤2 years	54.9	39.7	5.4		1,222	0.83 (0.52,1.30)
>2 years	52.2	40.8	6.9		650	0.91 (0.57,1.46)
Unmarried	54.0	39.2	6.8		155	0.64 (0.38,1.10)
Total	54.0	39.8	6.3		2,360	N = 2360

[#]Based on multivariate ordinal logistic regression models by controlling background characteristics such as school attendance, religion, caste, father's working status from baseline; Parity was not taken in model as it was correlated with age, age at first birth. ®: reference category; CI: confidence interval; DK: don't know; [##]based on Chi-square test; ***$P < 0.01$; **$P < 0.05$; *$P < 0.10$.

Table 4. Factors at follow-up associated with mental health among adolescent girls (aged 23–28 years) in Bihar, India, 2016

Background characteristics measured at follow up (at age 23–28 years)	Mental health status (%)				N	Results from ordered logistic regression[#]
	Good	Moderate	Poor	Level of significance[##]	2,360	AOR (95% CI)
School attendance						
Never attended®	50.1	43.2	6.7	***	1,170	1.00
Dropout before completing 12 class	56.5	37.5	6.0		1,023	0.77 (0.62,0.94)**
Completed or continuing	60.1	34.6	5.3		167	0.55 (0.39,0.79)***
Age						
23–24®	53.0	40.0	7.0		1,311	1.00
25–26	56.0	38.3	5.7		727	0.96 (0.80,1.15)
27–28	53.3	42.2	4.4		322	0.83 (0.65,1.07)
Place of residence						
Urban®	56.2	39.6	4.2		333	1.00
Rural	53.6	39.8	6.6		2,027	0.95 (0.78,1.15)
Wealth status						
Poor Quin®	54.0	40.4	5.6		780	1.00
Middle Quin	54.6	38.7	6.7		1,061	1.06 (0.85,1.30)
Rich Quin	52.5	41.1	6.4		519	1.10 (0.84,1.43)
Spousal education						
No education®	55.2	38.2	6.6	***	547	1.00
Primary	56.8	37.7	5.5		597	1.06 (0.82,1.36)
Secondary and above	54.5	40.1	5.4		938	1.18 (0.91,1.53)
DK	30.7	55	14.2		122	3.05 (2.05,4.53)***
Unmarried	54.0	39.2	6.8		155	1.44 (0.99,2.08)*
Working status						
Yes®	53.1	39.5	7.4	**	954	1.00
No	54.6	39.9	5.5		1,406	0.92 (0.77,1.11)
Self-efficacy						
High®	56.3	38.3	5.5	***	2,010	1.00
Low	40.6	48.5	10.9		350	1.79 (1.42,2.26)***
Decision making						
Others®	51.2	41.6	7.2	*	1,045	1.00
Alone	56.2	38.3	5.6		1,315	0.86 (0.73,1.02)
Total	54.0	39.8	6.3		2,360	

[#]Based on multivariate ordinal logistic regression models by controlling background characteristics such religion, caste, father's working status and parents survival status from baseline; ®: reference category; CI: confidence interval; DK: don't know; [##]based on Chi-square test; ***$P < 0.01$; **$P < 0.05$; *$P < 0.10$.

of women whose mothers were educated as compared to women whose mothers were uneducated (Table 2). Husbands' education was positively related to the good mental health status of women. Young women who were more than 6 years younger to husband had poor mental health as compared to those who had almost 3–6 years of spousal age gap (Table 3). However, self-efficacy of young women is an important factor that improves their mental health status. As compared to high self-efficacy girls, others had a high risk (AOR: 1.79, 95% CI: 1.42, 2.26) to face poor mental health in early adulthood (Table 4).

4. Discussion

Using data from the UDAYA study consisting of a 2360 adolescent (ages 15 – 19) girl cohort interviewed in 2007 and reinterviewed at ages 23–28 in 2016 from the state of Bihar, India, this study examined the effect of life-course events such as school dropouts, early marriages, and

early childbearing on mental health status at later ages (23–28 years). Our findings on the effect of educational attainment in adolescence (baseline at age 15–19) and in early adulthood (follow-up at age 23–28) on mental health status in early adulthood (23–28 years) are consistent with prior research that found significant association between dropout rate and mental health problems in China (Wang, Yang, He, *et al.*, 2015). School attendance in adolescence and also in early adulthood affects mental health status in early adulthood. From these findings, it is visible that continued schooling in adolescence and getting higher education until young adulthood is positively associated with good mental health status in early adulthood. Issues such as poverty, low academic achievement, lack of motivation, high rates of early marriage and pregnancy, unsupportive family, and domestic violence have also been found to be the main causes of school dropouts (Jensen and Thornton, 2003; Velez and Saenz, 2001), and these factors may also have negative consequences on mental health even for longer period as evident in the present study. Dropping out of school affects a person's life in different ways. Inability to complete high school education can lead to negative outcomes such as unemployment, underemployment, and poverty which results in poor mental health in early adulthood as also seen in this study.

Life-course events such as marriage and childbearing have a huge impact on the life of an individual especially psychologically both positively and negatively. The present findings show that women who are ever married have better mental health as compared to their unmarried counterparts at their young adulthood. Studies have shown that marriage enhances psychological well-being (Strohschein, McDonough, Monette, *et al.*, 2005; Waite, 1995) and improves social support, and it also connects people to other individuals, social groups, and social institutions which are themselves a source of benefits (Stolzenberg, Blair-Loy, and Waite, 1995). Another study by Haarasilta *et al.* (2004) found that unmarried or not-cohabiting young adults showed higher odds of having major depressive episodes. The present results did not find any significant effect of marital status on mental health status in young adulthood. The proportion of women who experienced separation, divorce, or widowhood was lower in Bihar, India; hence, mental health was status that was not analyzed separately for these groups. Mental illness among the widows is generally associated with the financial condition whereas the separated or divorced women can have stress or strain due to financial difficulties as well as childcare commitments and diminished social support (Wade and Pevalin, 2004).

There are many studies in developing countries showing a high incidence of early marriage. In addition, parents and other family members arranged the marriage of teen girls makes them vulnerable to mental health problems as they do not have the freedom to speak or share in the broader aspect which results in a psychotic disorder such as immobility, lack of confidence, and nervousness (Ahmed, Khan, Alia, *et al.*, 2013). It is found in this study that women who marry in their teenage have poor mental health as they are hit by early termination of education, due to low level of education, they may not be able to find work in the formal sector, face social and physical vulnerability, early childbearing which affect the life of a female in various dimensions (Jacubowski, 2008), that is, directly responsible for the poor mental health status of women at her later ages (Roest, 2016). People who get married at older ages have a higher level of maturity and emotional behavioral skills, use more effective strategies to cope and conflicts and problems, and, therefore, are expected to have greater marital satisfaction means better mental health status (Ghosh, Lahiri, and Datta, 2017; Hajihasani and Sim, 2019). Therefore, the present finding on the negative effect of late marriage on their poor mental health status in early adulthood is well supported by prior studies. This finding is also supporting the recent ongoing discussion on increase of legal age at marriage from 18 to 21 years in India. In this direction, Government of India has introduced the Prohibition of Child Marriage (Amendment), Act, 2021 to increase the legal age at marriage from 18 to 21 years (Government of India, 2021).

In the present research, it is found that young married women who have children after the age of 20 or who did not have any child had a poorer mental health status. The pressure to have children among young married women who do not have children and the pressure to cope up with the rearing and caring of a child are both situations in India that can have a lot of mental pressure on women (Marphatia, Ambale, and Reid, 2017). Fertility pressures from the family particularly from in-laws' are common especially among adolescent girls and in the states such as like Uttar Pradesh and Bihar (Dixit, Bhan, Benmarhnia, *et al.*, 2021). Women who had child after age 20 may have younger children at the current age (22 – 29) when mental health status being assessed, and, therefore, having young kids have the feeling of responsibility and caregiving due to which they have some kind of mental pressure (Hank, 2010; Aitken, Hewitt, Keogh, *et al.*, 2016).

A higher proportion of young women from SCs/STs experiencing poor mental health was found in the present study, which is similar to the finding in most of the studies (Das, Do, Friedman, *et al.*, 2009; Gaur, Vohra, Subash, *et al.*, 2003; Kohrt, 2009; Mohindra, Haddad, and Narayana, 2006). Findings on the effect

of a husband's education suggest that highly educated husband's wives are more likely to have poor mental health. This may be due to male dominance through education (Pai, Godboldo-Brooks and Edington, 2010). However, on the contrary, wives 3–5 years younger than their husbands have better mental health than their peers who are either older than or 2 years younger than their husbands. However, a significant proportion of women did not know about their husbands' age and education. It is noteworthy that women who do not know their husband's age and their husband's education are more vulnerable to poor mental health. It is expected in a rural Indian society where marriages are performed at an early age, mostly arranged by the family with no consent of girls, where gender roles are also very critical (Patel, Santhya and Haberland, 2021). In such cases, women are less likely to know about their husband's characteristics. Further, the present findings that women's empowerment (work status, self-efficacy, decision making) improves mental health outcomes suggest a need to focus on girl's empowerment at younger ages. In this way, events such as drop-out from school, early marriage, and early childbirth may be delayed, which will improve their mental health outcomes in young adulthood.

The present study is based on two state samples and, therefore, results may not be generalized for other states of India. Variables, including outcome variables of mental health status and husband characteristics, are based on the reporting from individuals and, therefore, may have some response biasedness. Nevertheless, the study findings are crucial and have programmatic relevance.

5. Conclusions

The present study identified the effect of life course events such as school drop-out, marriage, childbearing, and parental characteristics on the mental health outcomes at young adulthood. We also assessed the association of empowerment of young women with their mental health outcomes. Young women who never attended school or dropped out from school have a poorer mental health status in their young adulthood (22–28 years). Young women who completed their higher education showed a better mental health status. Young women having educated mothers and fathers with the job were less likely to have poor mental health. Women whose marriage was dissolved due to the death of their spouse, separation, or divorce were most likely to have poor mental health. Delaying age at marriage positively affects mental health in early adulthood as girls who got married after age 19 showed better mental health outcomes. On the contrary, compared to women who had their first child before age 19, women who did not have a child, or who had their first child after 20 years of age, were more likely to have poorer mental health. This suggests that the societal pressure of bearing a child just after marriage is a norm that might affect the mental health of women who could not bear it soon after marriage. Further, young mothers who gave birth after age 20 years are likely to report more mental health issues in that new childcare is found to be more associated with postpartum depression than those who already had older children. The current findings indicate that not bearing a child, bearing a child late, and postpartum depression affect mental health negatively and require a programmatic attention on young motherhood.

Another emerging finding of the study is the association of empowerment-related variables with mental health. Young women who did not know their husband's age or education had poor mental health outcomes. Surprisingly, women with highly educated husbands are more likely to have poor mental health. Working women, women who have better self-efficacy and have a decision-making power show better mental health outcomes. Therefore, the study recommends continued attention to improving young women's autonomy and gender role attitudes as it directly affects their psychological well-being.

Acknowledgments

This research uses data from the study on "Understanding the Lives of Adolescents and Young Adults (UDAYA) in Bihar and Uttar Pradesh," collected by the Population Council. Therefore, the authors are thankful to the Population Council for providing the data.

Funding

We declare that we did not receive any funding for this work.

Conflict of Interest

The authors declare that they do not have any competing interest.

Author Contribution

Conception and design of the study: RY and PD; analysis and/or interpretation of data: RY; drafting the manuscript: RY and AK; revising the manuscript critically for important intellectual content: PD and SKP; reading and approving the manuscript: all authors.

Availability of Supporting Data

The use is secondary in nature and is freely available to everyone on request at https://www.popcouncil.org/. UDAYA follow-up data and survey tools can be obtained from the Harvard Dataverse.

Ethics Statement

Not applicable.

Consent to Publish

Not applicable.

References

Ahmed S, Khan S, Alia M, *et al.* (2013). Psychological impact evaluation of early marriages. *International Journal of Endorsing Health Science Research*, 1(2):84-86.

https://doi.org/10.29052/IJEHSR.v1.i2.2013.84-86

Aitken Z, Hewitt B, Keogh L, *et al.* (2016). Young maternal age at first birth and mental health later in life: Does the association vary by birth cohort? *Social Science and Medicine*, 157:9-17.

https://doi.org/10.1016/j.socscimed.2016.03.037

Blakemore SJ. (2019). Adolescence and mental health. *Lancet*, 393(10185):2030-2031.

https://doi.org/10.1016/S0140-6736(19)31013-X

Chauhan S, Srivastava S, Kumar P, *et al.* (2021). Interaction of substance use with physical activity and its effect on depressive symptoms among adolescents. *Journal of Substance Use*, 26:524-530.

https://doi.org/10.1080/14659891.2020.1851411

Dangal G. (2006). Teenage pregnancy: Complexities and challenges. *JNMA Journal of the Nepal Medical Association*, 45(162):262-272.

https://doi.org/10.31729/jnma.513

Das J, Do QT, Friedman J, *et al.* (2009). Mental health patterns and consequences: Results from survey data in five developing countries. *The World Bank Economic Review*, 23(1):31-55.

https://doi.org/10.1093/wber/lhn010

Deb S, Chatterjee P, and Walsh K. (2010). Anxiety among high school students in India: Comparisons across gender, school type, social strata and perceptions of quality time with parents. *Australian Journal of Educational and Developmental Psychology*, 10:18-31.

Dixit A, Bhan N, Benmarhnia T, *et al.* (2021). The association between early in marriage fertility pressure from in-laws' and family planning behaviors, among married adolescent girls in Bihar and Uttar Pradesh, India. *Reproductive Health*, 18:60.

https://doi.org/10.1186/s12978-021-01116-9

Elder GH Jr. (1995). The life course paradigm: Social change and individual development. In: *Examining Lives in Context: Perspectives on the Ecology of Human Development*. Washington, DC: American Psychological Association.

https://doi.org/10.1037/10176-003

Gaur DR, Vohra AK, Subash S, *et al.* (2003). Prevalence of psychiatric morbidity among 6 to 14 years old children. *Indian Journal of Community Medicine*, 28(3):133.

Ghosh S, Lahiri S, and Datta N. (2017). Understanding happiness and psychological wellbeing among young married women in rural India. *Journal of Comparative Family Studies*, 48(1):113-131. Available from:

https://doi.org/10.3138/jcfs.48.1.113

Gilman SE, Kawachi I, Fitzmaurice GM, *et al.* (2002). Socioeconomic status in childhood and the lifetime risk of major depression. *International Journal of Epidemiology*, 31(2):359-367.

https://doi.org/10.1093/ije/31.2.359

Goldberg D and Williams P. (1988). *A User's Guide to the General Health Questionnaire*. Windsor, UK: NFER-Nelson.

Government of India. (2021). Bill No. 163 of 2021. *The Prohibition of Child Marriage* (*Amendment*) *Bill, 2021*. New Delhi: As Introduced in Lok Sabha, Government of India.

Gururaj G, Varghese M, Benegal V, *et al.* (2016). *National Mental Health Survey of India, 2015-16: Summary*. Bengaluru: National Institute of Mental Health and Neurosciences, p.1-48.

Haarasilta LM, Marttunen MJ, Kaprio JA, *et al.* (2004). Correlates of depression in a representative nationwide sample of adolescents (15-19 years) and young adults (20-24 years). *The European Journal of Public Health*, 14(3):280-285.

Hajihasani M and Sim T. (2019). Marital satisfaction among girls with early marriage in Iran: Emotional intelligence and religious orientation. *International Journal of Adolescence and Youth*, 24(3):297-306.

https://doi.org/10.1080/02673843.2018.1528167

Hank K. (2010). Childbearing history, later-life health, and mortality in Germany. *Population Studies*, 64(3):275-291.

Hjorth CF, Bilgrav L, Frandsen LS, *et al.* (2016). Mental health and school dropout across educational levels and genders: A 4.8-year follow-up study. *BMC Public Health*, 16(1):1-12.

https://doi.org/10.1186/s12889-016-3622-8

IIPS and ICF. (2020). *National Family Health Survey* (*NFHS-5*), *2019-20: Factsheet India*. Mumbai: International Institute for Population Sciences (IIPS) and ICF. International Institute for Population Sciences and ICF.

IIPS and Population Council. (2010). *Youth in India: Situation and Needs 2006-07*. Mumbai: International Institute for Population Sciences and Population Council.

Jacubowski N. (2008). Marriage is not a safe place: Heterosexual marriage and HIV-related vulnerability in Indonesia. *Culture, Health and Sexuality*, 10(1):87-97.

https://doi.org/10.1080/13691050601058573.

Jensen R and Thornton R. (2003). Early female marriage in the developing world. *Gender and Development*, 11(2):9-19.

https://doi.org/10.1080/741954311

Kishore AN and Shaji KS. (2012). School dropouts: Examining the space of reasons. *Indian Journal of Psychological Medicine*, 34(4):318-323.

https://doi.org/10.4103/0253-7176.108201

Kohrt BA. (2009). Vulnerable social groups in post conflict settings: A mixed methods policy analysis and epidemiology study of caste and psychological morbidity in Nepal. *Intervention*, 7(3):239-264.

https://doi.org/10.1097/WTF.0b013e3283346426

Kumar P, Srivastava S, Mishra P, *et al.* (2020). Does depressive symptoms, physical inactivity and substance use catalyze the suicidal tendency among adolescents? Evidence from a cross-sectional study. *Children and Youth Services Review*, 119:105661.

https://doi.org/10.1016/j.childyouth.2020.105661

Le Strat Y, Dubertret C, and Le Foll B. (2011). Child marriage in the United States and its association with mental health in women. *Pediatrics*, 128(3):524-530.

https://doi.org/10.1542/peds.2011-0961

Lorant V, Deliège D, Eaton W, *et al.* (2003). Socioeconomic inequalities in depression: A meta-analysis. *American Journal of Epidemiology*, 157(2):98-112.

https://doi.org/10.1093/aje/kwf182

Marphatia AA, Ambale GS, and Reid AM (2017) Women's Marriage age matters for public health: A review of the broader health and social implications in South Asia. *Frontiers in Public Health*, 5:269.

https://doi.org/10.3389/fpubh.2017.00269

Marphatia AA, Reid AM, and Yajnik CS. (2019). Developmental origins of secondary school dropout in Rural India and its differential consequences by sex: A biosocial life-course analysis. *International Journal of Educational Development*, 66:8-23.

https://doi.org/10.1016/j.ijedudev.2018.12.001

Mohindra KS, Haddad S, and Narayana D. (2006). Women's health in a rural community in Kerala, India: Do caste and socioeconomic position matter? *Journal of Epidemiology and Community Health*, 60(12):1020-1026.

https://doi.org/10.1136/jech.2006.047647

Montazeri A, Harirchi AM, Shariati M, *et al.* (2003). The 12-item general health questionnaire (GHQ-12): Translation and validation study of the Iranian version. *Health and Quality of Life Outcomes*, 1(1):1-4.

https://doi.org/10.1186/1477-7525-1-66

Muntaner C, Eaton WW, Diala C, *et al.* (1998). Social class, assets, organizational control and the prevalence of common groups of psychiatric disorders. *Social Science and Medicine*, 47(12):2043-2053.

https://10.1016/s0277-9536(98)00309-8

Muntaner C, Eaton WW, Miech R, *et al.* (2004). Socioeconomic position and major mental disorders. *Epidemiologic Reviews*, 26(1):53-62.

https://doi.org/10.1093/epirev/mxh001

Office of the Registrar General and Census Commissioner of India. (2015). *Adolescent and Youth Population*. New Delhi: Office of the Registrar General and Census Commissioner, Government of India. Available from: http://www.censusindia.gov.in 2011census PCA_AY 2011 [Last accessed on 2022 May 25].

Pai CW, Godboldo-Brooks A, and Edington DW. (2010). Spousal concordance for overall health risk status and preventive service compliance. *Annals of Epidemiology*, 20(7):539-546.

Patel SK, Santhya KG, and Haberland N. (2021). What shapes gender attitudes among adolescent girls and boys? Evidence from the UDAYA longitudinal study in India. *PLoS One*, 16(3):e0248766.

https://doi.org/10.1371/journal.pone.0248766

Patel SK, Saggurti N, Pachauri S, *et al.* (2015). Correlates of mental depression among female sex workers in Southern India. *Asia Pacific Journal of Public Health*, 27(8):809-819.

https://doi.org/10.1177/1010539515601480

Prakash IJ. (2021). Ageing India: Psychological concerns and responses. In: *Gerontological Concerns and Responses in India*. Singapore: Springer, p.25-42.

https://doi.org/10.1007/978-981-16-4764-2_4

Roest J. (2016). Child marriage and early child-bearing in India: Risk factors and policy implications. *Young Lives Policy Paper*, 10:12-34. https://www.younglives.org.uk [Last accessed on 2022 Apr 23].

Roy P, Bharati P, and Chakravarty S. (2019). Adolescent obesity and mental health status in Asian countries: A systematic review. In: *Data Scie SDGs: Challenges, Opportunities and Realities*. Berlin: Springer Nature.

Santhya KG, Acharya R, Pandey N, *et al.* (2017). *Understanding the Lives of Adolescents and Young Adults* (*UDAYA*) *in Bihar, India*. New Delhi: Population Council. https://www.efaidnbmnnnibpcajpcglclefindmkaj; https://www.popcouncil.org/uploads/pdfs/2017pgy_udaya-biharreport.pdf [Last accessed on 2022 Apr 23].

Shahidul SM and Karim AH. (2015). Factors contributing to school dropout among the girls: A review of literature. *European Journal of Research and Reflection in Educational Sciences*, 3(2):25-35.

Solmi M, Radua J, Olivola M, *et al.* (2021). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, 27:281-295.

https://doi.org/10.1038/s41380-021-01161-7

Stolzenberg RM, Blair-Loy M, and Waite LJ. (1995). Religious participation in early adulthood: Age and family life cycle effects on church membership. *American Sociological Review*, 60(1):84-103.

https://doi.org/10.2307/2096347

Strohschein L, McDonough P, Monette G, *et al.* (2005). Marital transitions and mental health: are there gender differences in the short-term effects of marital status change? *Social Science and Medicine*, 61(11):2293-2303.

https://doi.org/10.1016/j.socscimed.2005.07.020

Taber KS. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48:1273-1296.

https://doi.org/10.1007/s11165-016-9602-2

Velez W and Saenz R. (2001). Toward a comprehensive model of the school leaving process among Latinos. *School Psychology Quarterly*, 16(4):445.

https://doi.org/10.1521/scpq.16.4.445.19897

Wade TJ and Pevalin DJ. (2004). Marital transitions and mental health. *Journal of Health and Social Behavior*, 45(2):155-170.

http://doi.org/10.1177/002214650404500203

Waite LJ. (1995). Does marriage matter? *Demography*, 32(4):483-507.

https://doi.org/10.2307/2061670

Wang H, Yang C, He F, *et al.* (2015). Mental health and dropout behavior: A cross-sectional study of junior high students in Northwest Rural China. *International Journal of Educational Development*, 41:1-12.

https://doi.org/10.1016/j.ijedudev.2014.12.005

WHO. (2021). *Adolescent Mental Health*. Geneva: World Health Organization. https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [Last accessed on 2022 May 25].

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

The travails of gerontology education in Malta: Challenges and possibilities

Marvin Formosa*

Department of Gerontology and Dementia Studies, Faculty for Social Wellbeing, University of Malta, Msida, Malta

Abstract

As in recent decades, Malta has experienced an increase in both the number of available university programs in ageing studies and graduate students, the aim of this article is to evaluate the country's efforts in ensuring a trained workforce in gerontology, geriatrics, and dementia education. While Malta punches above its weight as far as gerontology education is concerned, one also notes a number of shortcomings. The country is still devoid of a clear space for professional gerontologists to put in practice all their knowledge, and unfortunately both public and private employers are still highly unaware of the skills that professional gerontologists can bring toward the improvement of the quality of life and well-being of older persons living either in the community or long-term care. Moreover, curricula remain hindered by two key limitations. Primarily, there is a disproportionate Western bias in the choice of theories and practices in all realms of ageing studies. Second, that no full-time faculty member at the Department of Gerontology and Dementia Studies is a geriatrician, and that such faculty members all service the University on a visiting basis. In this respect, this chapter recommends three key and urgent strategies for gerontology education in Malta. These include establishing gerontology as a discipline in its own right is long overdue, founding gerontology as a *bona fide* profession, and accrediting gerontology.

Keywords: Geriatrics; Dementia studies; Ageing; Higher education; University

***Corresponding author:**
Marvin Formosa,
(marvin.formosa@um.edu.mt)

Citation: Formosa M. (2022). The travails of gerontology education in Malta: Challenges and possibilities. International Journal of Population Studies, 8(1):38-48.
https://doi.org/10.36922/ijps.v8i1.1311

Received: April 4, 2022

Accepted: June 13, 2022

Published Online: June 27, 2022

Copyright: © 2022 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Some 15 years ago, Askham *et al.* (2007) anticipated that despite the increasing number and percentage of persons aged 65-plus, many of whom require community or residential care services, the demand for gerontology education will fall and that there is a strong possibility for many gerontology departments to close. True to their word, financial restraints have affected the global number of gerontology and geriatric programs, to the extent that in many countries, such programs are closing or have become stagnant (Silverstein and Fitzgerald, 2017). However, this is far from the case in Malta. Malta is a European Union (EU) Member State and at end of 2020, Malta's total population reached 516,100 persons (National Statistics Office, 2021). While 18.9% (97.418 persons) were aged 65+, the 80+, and 90+ cohorts reached 4.3% (22,183 persons) and 0.6% (3207 persons), respectively. Such demographic trends are due, to a large extent, to the country's rising levels of life expectancy. While at the beginning of the 20th century life

expectancy in Malta was around 43/46 years for men/women, life expectancy for those born in 2020 had reached 80.3 and 84.5 years for men and women, respectively (National Statistics Office, 2021). The need for old-age care and/or geriatric care in Malta, and hence, the development of gerontology education programs, will rise considerably considering that demographic projections indicate that Malta will be one of the fastest ageing countries in the EU. During the 2013–2060 period, the 65-plus/80-plus cohort will reach 28.5%/10.5% of the total population in 2060, from 17.5%/3.8% in 2013 (+11.0%/+6.7%) – thus, implying that the number of older persons requiring communicate services to age-in-place or residing in residential long-term care will increase in the coming three decades (European Commission, 2015).

During the 2010s, Malta witnessed an exponential increase in applications from students to read for university programs in gerontology, geriatrics, and dementia care. Moreover, contrary to the situation in the United Kingdom where "employers are also decreasingly likely to allow employees the time off to study or to contribute toward the fees" (Askham, Gilhooly, Parkette, *et al*., 2007, p. 46), the opposite has occurred in Malta as both the public and private sectors have released, and even sponsored, many of their personnel to enroll at the University's Department of Gerontology and Dementia Studies. Comparative data on the global number of students in gerontology education are unavailable to-date. However, Silverstein and Fitzgerald (2017) pointed out that "financial constraints and a variety of other factors have affected the number of gerontology and geriatric programs available around the world" and while "in some countries, the field is flourishing in terms of educational programs and jobs for graduates… in other programs are closing, never existed or have become stagnant" (p. 1). In June 2022, the Department of Gerontology and Dementia Studies was offering four distinct programs in gerontology education – namely, the Higher Diploma in Gerontology and Geriatrics, Master of Gerontology and Geriatrics, Master of Arts in Ageing and Dementia Studies, and Doctorate. Through such academic programs, the department aims at developing and deepening interdisciplinary scientific teaching, education, and research in gerontology, geriatrics, and dementia, and is part of the Faculty for Social Well-being. Indeed, the department

> …puts emphasis on didactic teaching on ageing welfare policy, dementia studies, social and biological theories of ageing, quality of life in old age, and researching ageing and later life. Spearheaded by the academic output of its academic staff, the department lies at the forefront of scholarly activity and publications, as well as European and international collaborations related to ageing, older adults, and later life. A major focus of the department is the role of the interdisciplinary team practice in geriatric and dementia care.
>
> University of Malta, n.d.

This article traces the development of gerontology education in Malta in the years 1990 – 2020. This period witnessed an increase in the number of educational programs in ageing studies, and subsequently, a proliferation of students graduating in gerontology, geriatrics, and dementia studies. While the subsequent section delineates global developments in the field of gerontology education, the third part presents the emergence and consolidation of graduate programs in gerontology, geriatrics, and dementia studies. The fourth section embeds such developments in a critical commentary by highlighting the field's successes and limitations, and the final concluding section puts forward a range of the future recommendations for policy action.

2. Gerontology education in malta

University programs in gerontology are most common in the United Kingdom, United States, Canada, Nordic countries, and Australia. However, gerontology programs can also be found in countries as diverse as Germany, China, Israel, Singapore, Turkey, Guatemala, Uruguay, Venezuela, Taiwan, South Africa, Mexico, and of course Malta. The range of focuses, functions, and specialties of gerontology education leads to a key query: *What are the mutual competencies in these gerontology educational programmes*? The construction of a framework of competencies for gerontology education is most advanced in the United States, where the Association for Gerontology in Higher Education (AGHE) put forward three categories of competencies:

> Category I competencies represent the essential orientation to the field of gerontology, are foundational and expected to be broadly represented in Associate, Bachelors, Master's degree, and gerontology certificate programs. Category II competencies are "interactional" competencies that capture the processes of knowing and doing across the fields of gerontology and are also expected to be broadly represented in the above types of educational programs. Category III competencies are meant to capture the most relevant skills for contexts of employment in the variety of sectors and areas that gerontologists may work, including education.
>
> AGHE, 2014, p. 10

The Institute of Gerontology was set up in 1987 as an academic institution within the University of Malta to develop interdisciplinary teaching and research in gerontology and geriatrics. The main objective of the institute

was to offer training programs to respond to the urgent demand for trained personnel in the field of ageing welfare. While the government had launched a range of community care services (such as senior centers, home help services, the Telecare service, and domiciliary health services), institutional long-term care services, and rehabilitation and geriatric hospitals, the social and health-care professionals responsible for such services and institutions had sparse knowledge on the principles and practicalities of gerontological and geriatric care, despite many holding a graduate professional background (Formosa, 2017). To mitigate against a shortcoming, the October 1990–June 2010 period saw the Institute of Gerontology offering a Postgraduate Diploma in Gerontology and Geriatrics, and Master of Gerontology and Geriatrics, for professionals working in the field of ageing. Admission requirements included either a degree or professional qualification in social and behavioral sciences/medicine, social work, and allied health professions obtained from a university or from any other authority recognized by the University of Malta; or any other professional qualification deemed acceptable by the University of Malta, if it is accompanied by adequate work experience (at least 5 years) in the care of older persons. However, regulations also included the proviso that a candidate without the above qualifications may be admitted to the program as a mature student if he/she has an exceptional degree of experience in the field of ageing or the care of older persons and has the necessary background to profit from the program. The program of the postgraduate diploma, together with information on its number of European Credit Transfer and Accumulation System (ECTS), which is a tool of the European Higher Education Area for making studies and courses more transparent, was as follows:

Compulsory study-Units. Eight study units (four ECTS) requiring 18 and 50 h of lecturing and study hours, respectively. The eight core modules, evenly balanced in both areas of gerontology and geriatrics, covered the following areas: population aging; sociology of aging; health promotion; biology and physiology of aging; psychology of aging; research and evaluation; clinical aspects of old age; and medical and social rehabilitation.

Elective study-units. Choosing five study units (four ECTS) requiring 18 and 50 h of lecturing and study hours, respectively, from the following elective study units: social policies and strategies; medical problems; program and services: international comparison; psychogeriatrics; geriatric supportive services; social welfare with older persons; epidemiology; statistics and computing; income security for older persons; and nutrition.

Practice placements. Participants were required to have an internship of 240 h divided into four practicums of 60 h each. Practical placements of both observational and hands on type were aimed at enabling the participants to apply theoretical knowledge to practical solutions (eight ECTS).

Dissertation. Dissertation on an approved area of study chosen in consultation with a supervisor. The dissertation (15,000 words) required a sustained effort in defining the chosen area of investigation, in researching the issue at hand, in integrating the empirical data collected into a wider context of the subject, and in drawing conclusions and recommendations (30 ECTS).

Final examinations. Three final comprehensive examinations (30 ECTS)

Source: modified from van Rijsselt, Parkatti, and Troisi, 2007, p. 89 – 90

The key goal of the field placement, deemed as a learning experience, was to assist the students make an operative transition from the classroom to a practical situation. The rationale was that the theoretical knowledge obtained in the classroom will be synthesized during the field placement experience in an actual and unpredictable "real world" site. The filed placement would thus provide the students with an opportunity to develop and rehearse skills related to professional goals and objectives of gerontology and foster the required proficiencies in the social and health care of older persons. Admission requirements for the Master of Gerontology and Geriatrics required an average mark of 70% or higher in the Postgraduate Diploma in Gerontology and Geriatrics (van Rijsselt, Parkatti, and Troisi, 2007). The master degree program was entirely research based. Following a successful proposal, students were required to author a 50,000-word dissertation that researched a topic of either a gerontological or geriatric concern.

In 2005, following Malta's accession in the EU 1 year beforehand the Institute of Gerontology changed its name to the European Centre for Gerontology, and in October 2010 streamlined its graduate programs in gerontology education in line with the Bologna Agreement (Formosa, 2015). As a result, the Postgraduate Diploma and Master of Gerontology and Geriatrics were merged into one program of study – namely, the Master of Gerontology and Geriatrics. The newly formed model, which is still retained at the time of writing, comprised of the following study program:

Study-Units. Eleven study units (five ECTS) requiring 28 and 70 h of lecturing and study hours, respectively: Sociology of ageing; quality of life; biological issues in old age; health promotion; research and evaluation; social policies, programs and services in the field of

ageing; clinical conditions and health-care services in geriatrics medicine; psychological and psychogeriatric issues; social interactions in later life; pharmacological issues in later life; and gerontology: Multidisciplinary and interdisciplinary approaches.

Field placements. Participants are required to have an internship of 240 h divided into four placements of 60 h each (5 ECTS).

Dissertation. Dissertation on an approved area of study chosen in consultation with a supervisor (20,000 words) (30 ECTS).

University of Malta, n.d.

Admission requirements were modified to include either a bachelor degree obtained with at least second class honors, or a professional qualification considered by the board to be comparable to a degree, in a discipline pertaining to one of the caring professions; or a bachelor degree obtained with at least second class honors; or a professional qualification considered by the board to be comparable to a degree deemed by the board as relevant and adequate for the applicant to follow the course with profit; or a bachelor degree obtained with third class honors if applicants are also in possession of other qualifications, including relevant experience, following their first cycle degree; or in exceptional cases, a professional qualification in one of the caring professions, together with at least 5 years' experience, which together are deemed by the Admissions Board, on the advice of the Faculty Admissions Committee, to be comparable to the level of a first degree. Regulations allow students who do not wish to continue with the research component, as well those who do not achieve at least a 65% average mark in their examination results, to be awarded the postgraduate diploma.

The year 2010 also witnessed the launch of a Doctoral (PhD) program in gerontology which was juxtaposed midway the North American route for gaining a doctoral degree in gerontology and the British path for the award of a Doctor of Philosophy. While the PhD program comprised a 3-year research initiative, where students have to write a 100,000-word dissertation and originality is a key requirement, as generally found in the British path, it also included intensive and interdisciplinary doctoral training in the biological sciences, health and medical issues, psychological theories and mechanisms, public policy issues, social theories, and most importantly, methodology and research issues. To-date, titles of completed dissertations included *Implementing person-centered dementia care in a rehabilitation hospital through an appreciative inquiry approach* (Scerri, 2018) and *Moral reasoning in end-of-life decision-making for persons in end-stage dementia* (Dimech, 2019). Ongoing research emphases for doctorate study include masculinity and caregiving, gay and lesbian older persons in long-term care, Montessori activities for persons with dementia, and active ageing in long-term care.

In June 2014, the Institute of Gerontology migrated to the Faculty for Social Well-being, and another name change took place, this time to Department for Gerontology. The department continued offering the Master of Gerontology and Geriatrics, but also launched two new graduate programs: Higher Diploma in Gerontology and Geriatrics (October 2015) and the Master of Arts in Ageing and Dementia Studies (February 2016). The Higher Diploma in Gerontology and Geriatrics targets students who have completed their higher education and are looking to expand their knowledge of gerontology and geriatrics through a recognized university program. Admission requirements included the general entry requirements for entrance to university. At the time of writing, the higher diploma comprised of 19 study-units and three field placements:

Study-Units. Nineteen study units of four ECTS each requiring 18 and 50 h of lecturing and study hours, respectively. Study-units included: Key principles in social gerontology; physiological and medical issues in old age; community services for older people; biological aspects; research methods; economic and social aspects of ageing; health-care professionals in old age; theoretical issues in ageing policy; mental health issues in later life; researching ageing and later life; food and nutrition in later life; familial networks and informal care; pensions and their sustainability in Malta; social rehabilitation in later life; introduction to abuse and neglect; income security, social protection and poverty prevention; diversity and discrimination in later life; long-term services for population ageing; educational gerontology; multi-disciplinary health services for older people; and recognizing and preventing elder abuse in long-term settings. Another study-unit, legal issues in later life, comprised a two ECTS (10 h of lectures).

Field placements. Participants are required to have an internship of 180 h divided into three placements of 60 h each (12 ECTS).

University of Malta, n.d.

On completion of the higher diploma, graduates are expected to be able to provide higher levels of social and health-care services to older persons which, in turn, will function to improve the quality of service to older adults. Most especially, candidates will be able to meet the needs of prospective employers by being knowledgeable in embracing the roles of financial planning for older persons, engaging in case work, and assisting in the organization

and administration of social and related services in community and residential care. The Master of Arts in Ageing and Dementia Studies reflected the urgent need for a trained workforce in dementia care. The objective of the degree was to reinforce and mature the understanding, skills, competencies, and attitudes of students working between the interface of ageing and dementia. The Master of Ageing and Dementia Studies was planned and developed to provide students with an in-depth, research-based knowledge of dementia, including theory, innovative and best practices, and policy issues, as well as a grounding in academic and research skills. Thus, promising to act as a catalyst for candidates' professional development with meeting the needs of both present and incoming cohorts of older persons with dementia. The program of study was as follows:

> *Study-Units*. Eleven study units of five ECTS requiring 28 and 70 h of lecturing and study hours, respectively. Study-units included: Social policies, programs and services in the field of ageing; clinical conditions and health-care services in geriatrics medicine; ageing: Psychological and psychogeriatric issues; social interactions in later life; pharmacological issues in later life; and gerontology: Multidisciplinary and interdisciplinary approaches.
>
> *Practice placements*. Participants are required to have an internship of 240 h divided into four practicums of 60 h each. (5 ECTS).
>
> *Dissertation*. Dissertation on an approved area of study chosen in consultation with a supervisor (25,000 words). (30 ECTS).
>
> University of Malta, n.d.

Admission criteria included the same qualifications require for entry in the Master of Gerontology and Geriatrics *except* the proviso where a professional qualification in one of the caring professions, together with at least 5 years' experience, is deemed to be comparable to the level of a first degree which was dropped. The Master Degree of Ageing and Dementia Studies makes a valuable contribution to the development of dementia care knowledge, research, and practice is designed in a way to promote collaboration between disciplines of nursing, medicine, occupational therapy, social work, and other professionals working with persons living with dementia and their families/significant others. Table 1 provides a summary of admissions requirements and study programs for each degree.

A comparative analysis of Malta's gerontology education to what is available in other countries finds that programs tend to focus substantially on demographic trends and policy frameworks of the geographical context they are located it, and that Malta's programs tick all the boxes as far as classroom lectures, practice placements and research dissertations are concerned. While other programs tend to miss on either practice placements or research dissertations as is the case in Mexico, or even both as is the case in Scotland, some programs also focus more deeply on issues that are of a major interest in the region such as gerontology accreditation in North America, assistive technology in Asia, and informal care in Europe.

In March 2019, the department experienced another name change, from "Department of Gerontology" to the "Department of Gerontology and Dementia Studies" so as to reflect the additional focus of dementia studies in its vision and objectives. Figure 1 illustrates a timeline highlighting the major changes/additions, name changes, degree offerings, and historical moments in Malta's travails in gerontology education. Indeed, while in 2016, the government tasked the department to plan, coordinate, and implement a 14-h training program to all nurses in the public sector and working with older persons, one resident academic collaborates with the university to organize biannual training programs on dementia care for informal caregivers.

3. Possibilities and challenges of gerontology education in malta

The Department of Gerontology and Dementia Studies survived and even thrived in a period where many gerontological department and graduate programs closed down. A case in point is the European Master's Program in Gerontology which was disbanded by VU-University Amsterdam in 2010 due to the fact that getting the program accredited was a costly exercise, especially in the context of the then worldwide financial crisis (Aartsen, 2011). Following Hertz *et al.*'s (2007) typology, the focus of the Department of Gerontology and Dementia Studies is mainly in the education of undergraduate and graduate students, and continuing education for both aspiring and working professionals, in gerontology, geriatrics, and dementia care. While its function is to ensure that all tiers in Malta's workforce are trained in these areas, its unique tripartite specialties include policy making, gerontology services, and dementia care. The Maltese situation is thus more similar to the United States context than the European one considering the range of more specialized degrees and certificate programs in ageing studies.

Returning back to Askham *et al.* (2007) predictions, it suffices to state that while the department included only six students in 2007, this figure reached 51 in 2020. Alumni statistics show that the period October 1990–April 2022 hosted as much as 224 Maltese and 132 international

Table 1. Gerontology education programs offered by the University of Malta (1990 – current)

Name	Admission Requirements	Study program	
		Taught part	Research part
Postgraduate Diploma in Gerontology and Geriatrics (1990 – 2010)	Either an undergraduate degree or professional qualification in social and behavioral sciences/medicine, social work and allied health professions; or any other professional qualification deemed acceptable by the University of Malta, if it is accompanied by adequate work experience in the care of older persons	13 study units (four ECTS each), a placement (eight ECTS), and final examinations (30 ECTS)	Long essay (20,000 words) (30 ECTS)
Master of Gerontology and Geriatrics (1990 – 2010)	Postgraduate Diploma in Gerontology and Geriatrics issued by the University of Malta	None	Dissertation (50,000 words) (60 ECTS).
Postgraduate Diploma in Gerontology and Geriatrics (2010 – current)	Either an undergraduate degree or professional qualification in social and behavioral sciences/medicine, social work and allied health professions; or any other professional qualification deemed acceptable by the University of Malta, if it is accompanied by adequate work experience in the care of older persons	Eleven study units and one placement (five ECTS each)	None
Master of Gerontology and Geriatrics (2010 – current)		Eleven study units and one placement (five ECTS each)	Dissertation (20,000 words) (30 ECTS).
Higher Diploma in Gerontology and Geriatrics (2015 – current)	General entry requirements for entrance to university or being aged 23 years or over on the condition of a successful interview carried out by the members of the board of studies	19/1 study units (4/2 ECTS) and three placements (five ECTS each)	None
Master of Arts in Ageing and Dementia Studies (2016 – current)	Any undergraduate degree	10 study units and two placements (five ECTS each)	Dissertation (25,000 words) (30 ECTS).
Doctor of Philosophy (Gerontology) (2010 – current)	Any master degree	None	Dissertation (100,000 words) (30 ECTS).

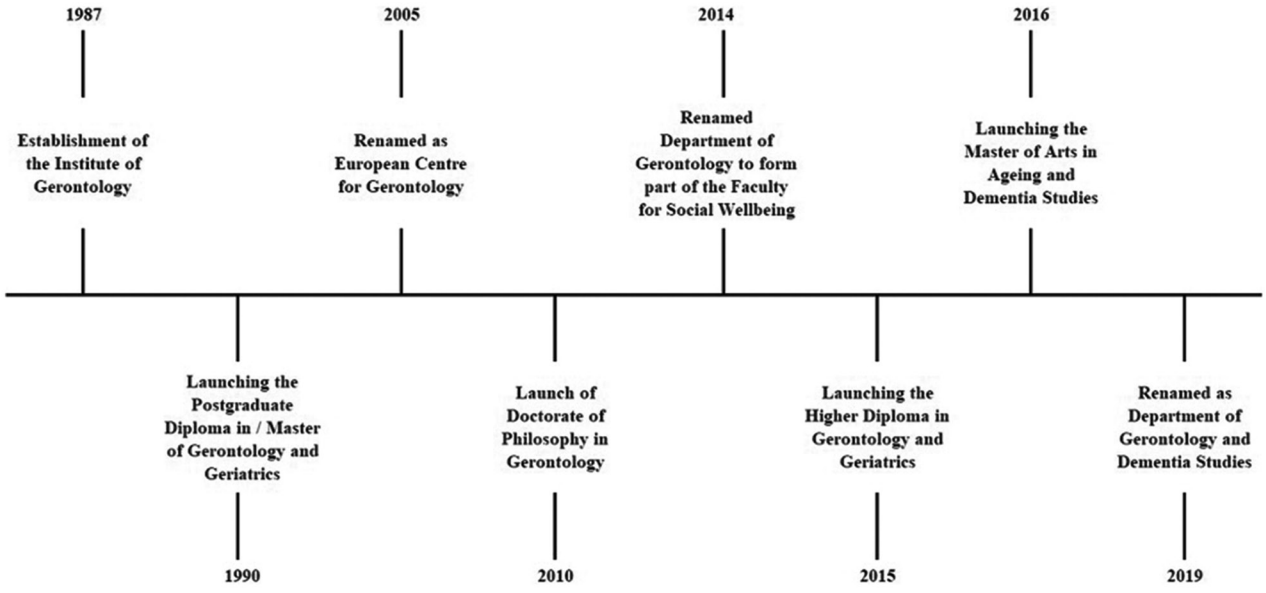


Figure 1. Time of developments in gerontology education, University of Malta, Malta (1987 – current).

students (356 in total). Such fruitful developments were not the result of coincidence or good fortune but followed detailed policy planning to ensure that the country has a sufficient and trained workforce to meet the social and health-care needs of the unprecedented increase of older persons in Malta. All the department's degree programs

are of 90 ECTS and were designed to be as flexible and modular as possible, so that students can cease their studies at the end of any semester and then rejoin in the future years. The academic content of the degree programs was planned and formulated to follow a "progressive academic learning concept" (Russell, 1990), with their special focus starting with basic gerontological knowledge, such as theories, concepts and research, and introducing academic and social debates on relevant gerontological issues, and then going on to the accumulation of academic knowledge, competencies, and skills. To authenticate the quality of its programs and to ensure that academic standards are maintained, especially in consideration that the University of Malta is the only university in the country, the department engages the services of a number of international external examiners for a maximum period of 3 years. Moreover, the department collaborates closely with a number of foreign universities, as well as the World Health Organization and United Nations' organizations, in particular with the International Institute on Ageing, United Nations - Malta. While some degree programs are available full-time, such as the Higher Diploma and Master of Gerontology and Geriatrics, the Masters of Arts in Ageing and Dementia Studies and the Doctorate are available on a part-time basis. Hence, the department has the potential to meet the academic needs and interests of both regular students interested in graduating in ageing studies as well as working professionals. The presence of the multi-national student body also serves to provide added value and attraction to the degree programs. As reported in the United States context, it is

> …useful to engage international students in class discussions to learn about diverse aging process in other cultural settings. This exchange also occurs at the student level where international students share aging experiences in their cultures with [local] students, thus building a comparative perspective and a more globalized knowledge base for the field.
>
> Mwangi, Yamashita, Ewen, *et al.*, 2012, p. 213

The appeal of the Department of Gerontology and Dementia Studies is also partly the result of effective lobbying and marketing tactics on behalf of its resident academics. Stratagems included a televised 14-session series on the impact of ageing and dementia on individuals and caregivers, the publication of information booklets on dementia and how to care for a relative with dementia, producing a street-theatre production on dementia across a number of villages and towns across Malta, and coordinating extensive training in dementia management and care for all nurses working with people with dementia in public care homes and geriatric institutions. Such efforts culminated in successfully lobbying the Junior Ministry for Active Ageing to issue six scholarships for personnel in the public service to read for a Masters of Arts in Ageing and Dementia Studies. At the same time, the department was intrinsically involved in the planning and writing of key policies and program initiatives documents on active ageing, dementia care, and care homes for older persons. On one hand, the *National Strategic Policy for Active Ageing* (Parliamentary Secretariat for Rights of Persons with Disability and Active Ageing, 2013), *National Dementia Strategy* (ibid., 2015a), and the *National Minimum Standards for Care Homes for Older Persons* (ibid., 2015b) were all spearheaded by resident academics within the Department (Scerri, 2015; Formosa, 2018). The launch of these three policy directions was instrumental in leading a number of workers in the field of ageing, especially those working in care homes and geriatric institutions, to enroll in a graduate program in either gerontology and geriatrics or dementia care. It is noteworthy that the strategic objectives for the section on workforce development included the following items related to gerontology education:

a. In collaboration with the University of Malta and the Malta College of Arts, Science, and Technology (MCAST) develops study units on the medical, social, psychological, and economic aspects of dementia for students undergoing health and social care training programs. An interdisciplinary team approach will be fostered and emphasized using appropriate teaching methodologies.
b. Provide patient-centered dementia care training to all health-care professionals caring for individuals with dementia. This will include continuous professional development programs for skills updating.
c. Provide a training program for caregivers. This may form part of a yearly program organized by professionals in the dementia field. Various aspects of dementia care including patient-caregiver relationship, stress management, and communication will be included in the study.
d. Support accredited information technology platforms that facilitate online dementia training for health-care professionals and caregivers.
e. Provide training opportunities in the optimal delivery of palliative care to health and social care professionals working with individuals with dementia.
f. Ensure that curricula at all levels of training are coherent to develop a well-trained multidisciplinary workforce.
g. Ensure that all dementia training programs delivered outside the University of Malta and the Malta College of Arts, Science, and Technology (MCAST) (and other recognized institutions) are fully accredited and classified within the Malta Qualification Framework.

> Parliamentary Secretariat for Rights of Persons with Disability and Active Ageing, 2015a, p. 71

Indeed, it is valuable that public entities, non-governmental organizations and private companies engaged in ageing welfare trust that the provision of such policies and standards cannot be rightfully attained in the absence of a learned and trained workforce. On the other hand, faculty members were instrumental in advising the government in the conceptualization, organization, and long-term implementation of program initiatives such as the Dementia Helpline, the University of the Fourth Age, and Dementia Intervention Team (Formosa, 2019).

However, this is not the same as saying that there is no space for improvement or any shortcomings in gerontology education in Malta. The fact that the University of Malta offers specific degrees in ageing studies implies that its social and health-care workforce – being in either social work, nursing, or oral health to mention some – is not well versed in gerontological and geriatric issues. Due to its microstate status, Malta focuses much more on on-the-job training for people working with seniors and starts people sooner in positions that, comparably, in other high-income countries would require a specialization in gerontology either at Bachelor or Master levels. In fact, most students who enroll to read for such studies are already working in the field of ageing, and hence, would have acquired their responsibility before any specialty training. Although the curriculum of the department's programs was set by local and international experts in ageing studies and also includes external evaluations, it remains hindered by two key limitations. One inadequacy is the disproportionate American and Anglo biases in the choice of theories and practices in all realms of ageing studies. There is little, if any, debate on the manner that established paradigms – such as active ageing, successful ageing and productive ageing – are located in cross-cultural spaces. While there are studies that highlight the role played by culture in ageing positively, such research is not listed in the reading lists of gerontology study-units, with hardly any mention of how geriatric and dementia care is hinged on the value system of the inquirer of cultural constructs or their social construction of reality" (Bowling, 1993, p. 449). Since all paradigms in ageing studies are value-laden and culturally constructed concepts, it is unfortunate that students are not required to problematize the imbued cultural hegemony.

Indeed, gerontology education requires an impetus to "elicit understandings of how [paradigms] are conceptualized by people in different cultural contexts," requiring "a framework that is capable of identifying and explaining the relationships between cultural values and understandings of the two concepts by people in diverse groups," and shedding "light on the intracultural variations, including both differences and similarities, among people in the same culture" (Tam, 2014, p. 889). Unfortunately, despite a sharp focus on the political economy of ageing, feminist gerontology, and gay and lesbian ageing, a cross-cultural component is not a strong point in the department's curricula. A second failing refers to the fact that no full-time faculty member at the Department of Gerontology and Dementia Studies is a geriatrician, and that such faculty members all work on a visiting basis (Formosa and Galea, 2020). Despite the fact that the Maltese population is ageing rapidly, there has been a slow development of geriatrics in academia, so that the teaching of geriatrics, especially at an undergraduate level, seems to have emerged as simply an afterthought. In the context of a looming geriatrician shortage, the "geriatrician" of the primary care workforce to prepare for an ageing population is especially warranted (Friedman, Gillespie, Medina-Walpole, *et al*., 2013). Such a state of affairs leads to two urgent necessities:

> (1) The need to train a growing number of health-care professionals in the care of older patients in the community and in the different types of elder care facilities. A greater number of coordinated programs are needed based on previous successful experiences… (2) The need to attract young fellows with academic potential to enter the discipline of geriatrics. Preclinical exposure to elders living within the community or in institutions appears to enhance the motivation of young medical students to become geriatricians…
>
> Michel and Cha, 2015, p. 1010

The lack of fulltime geriatricians as faculty members also implies that clinical training experiences a lack of facilities and infrastructure, in that lecturing and clinical placement hours have to be adopted to the geriatricians' availability rather than the other way around, and it is not uncommon that students arrive to the patient's bedside in big groups. This not only serves to hinge negatively on satisfaction of the older patients at the hospital but is ineffective as far as the students' educational experience is concerned. In this respect, the department's track record in collaborating with other medical societies – such as rehabilitation, oncogeriatrics, and swallowing disorders to mention some – is not a favorable one. Moreover, the absence of full-time academic posts in geriatrics has led to a lack of departmental discussion as what should be the basic knowledge, attitudes, and skills that students in gerontology education should have developed and possessed throughout studies. In 2000, a committee of the American Geriatrics Society (2000) highlighted that the strategies for teaching geriatrics may change according to the resources available at each institution, material, and

human. It also proposed that students have contact with healthy as well as frail older patients to avoid developing an inaccurate stereotype of older persons as being considered synonymous for ill persons. Unfortunately, such recommendations are still to be implemented in the department's geriatric curriculum.

4. Conclusions

Given the global trends in population ageing, one can never overemphasize the need for a workforce of professionals trained in gerontology, geriatrics, and dementia care. Although each geographical region and continent has certainly its own special quandaries, "each country [require] people with adequate knowledge and skills to care for its older population and to create opportunities for older adults to prosper in later life and continue to make valuable contributions to their communities" (Silverstein and Fitzgerald, 2017, p. 1). Hence, it is not surprising that over the past decade, it has become more common to witness calls for occupational positions, issued by both welfare service agencies and universities, noting a preference for a qualification in gerontology or a related field to ageing studies. This does not, however, mean that gerontology has become a well-established career path, and a Masters or Doctorate in Gerontology still does not guarantee its holder a smooth entry in either ageing services, gerontological care and rehabilitation, or academia. The key challenge facing holders of academic qualifications in gerontology is competition from more traditional disciplines, since those completing more traditional programs tend to be more identified with their main disciplines, which have many more members, and larger organizations to serve the membership. At the same time, the development of broad-spectrum master degrees and doctorates, such as in public health, also put gerontology education on the back foot since such programs are more attractive to potential candidates as they promise potential employment in a wide range of fields ranging from pediatrics to geriatrics. Overcoming such a state of affairs, requires three strategies – namely, establishing gerontology as a discipline in its own right, founding gerontology as a *bona fide* profession, and accrediting gerontology to mitigate against the fact that gerontology graduates are at a serious competitive disadvantage compared with graduates from programs that hold a licensure agreement – strategies that are elaborated upon in the final parts of this article.

The field's most pressing objective is to establish gerontology as a discipline in its own right, something that is surely long overdue. The field has long been in a rapid stage of development and the discussion on gerontology's evolution and future cannot be postponed any longer. Admittedly, the tension between traditional disciplinary perspectives and the multidimensional issues in ageing on one hand, and structural barriers that impeded interdisciplinary knowledge development and translation on the other, have hindered gerontology's development as an integrative discipline. However, in recent years, one has witnessed the development of unique social, psychological, and biological theories (Bengston and Settersten, 2016); specialist research methodology focusing on longitudinal life course data through distinctive strategies such as event history analysis and hierarchical linear modelling (Weil, 2017); and formal organizations that promote professional socialization and information dissemination (Formosa, 2021). For instance, the International Association of Gerontology and Geriatrics now includes as much as 84 members amid increasing degree programs and specialized publications from 72 different countries with a combined membership of more than 50,000 professionals and students (International Association of Gerontology and Geriatrics, n.d.).

A second key issue is related to the journey of gerontology education concerns the varied efforts by scholars to establish gerontology as a *bona fide* profession. There is no doubt that specialists with qualifications in gerontology education have the appropriate training and skills to become employed professionals in the field. After all, "the knowledge and competencies gained through gerontology education and training is unique, specialized, and distinguishable from that of other disciplines and professions [so] they should be brought together under a professional framework" (Pianosi and Payne, 2014, p. 835). It is only professionals that can affect the quality of care for older people because only they hold advanced proficiencies and academic preparation. Moreover, it is professional bodies that positively affect the public's perceptions and expectations about standards of care, particularly in contrast to other social and health-care providers who have not had enough training in gerontology. This absence of a professional status for gerontology is a key reason whereby students steer away from pursuing studies in ageing since "agencies that provide services for the aged do not require an academic background in aging studies for employment" (Maiden, Horowitz, and Howe, 2005, p. 5). One can mitigate against this quandary if gerontology academics associate more effectively with prospective employers and professional organizations, to clarify the skills and competencies that graduates in gerontology, geriatrics, and dementia care possess, so that gerontology degree recipients can expect to receive consideration in hiring decisions.

One key obstacle in professionalizing gerontology is that most graduates in gerontology education remain intimately and steadfastly connected with their primary professional roles and occupations. As it was argued,

> …many of those working in the social and health services and health-care field are subject to more developed, registered professions such as social work and nursing. These practitioners have developed and firmly implemented competencies, knowledge, defined occupational roles, and professional practices to serve more generic or generalized segments of the population's needs. Some of these professions have (recently) developed subspecialties in an effort to incorporate aging specific concepts and, to some degree, gerontological knowledge and competencies (e.g., social work and nursing).
>
> Pianosi and Payne, 2014, p. 835

The final challenge for gerontology education is that of accrediting gerontology so as to mitigate against the just mentioned reason that gerontology graduates are at a serious competitive disadvantage compared with graduates from programs that receive licensure, such as social work, nurses, and counselors. Accrediting gerontology is a first step for achieving certification. To push the professionalization of gerontologists forward, a formalized accreditation system needs to be created. This system would inform the public as to the roles and competencies that professional gerontologists possess, while also acting as the liaison between potential clients, academic education or training of gerontologists, and the practical and professional roles, they work within and the public. Indeed, not only

> …professionalizing gerontology has the potential to not only legitimize the value of gerontological knowledge, education, and competencies but [also] to… standardize, monitor, evaluate, and potentially regulate those calling themselves gerontologists. This would undoubtedly result in an enhanced quality of life for older individuals and an aging society. Providing educated, knowledgeable, and monitored professionals to the field of aging would be gerontology's main goal.
>
> Pianosi and Payne, 2014, p. 835

Indeed, only by reacting and resolving the above-mentioned quandaries can professional gerontologists provide ageing societies with highly educated and accountable individuals that would provide high quality services to older adults in a wide variety of settings and roles.

Funding

This research received no specific grant from any funding agency, commercial entity, or not-for-profit organization.

Conflict of interest

The author has no conflicts of interest to declare.

Ethics statement

Not applicable as this study involves theoretical work and commentary.

Availability of supporting data

No data from secondary sources were utilized for the writing of this paper.

References

Aartsen M. (2011). European master's program in gerontology (EuMaG): Goals, curriculum, and students. *Gerontology and Geriatrics Education*, 32(3):225-232.

https://doi.org/10.1080/02701960.2011.598975

American Geriatrics Society. (2000). Core competencies for the care of older patients: Recommendations of the American Geriatrics Society. *Academic Medicine*, 75(3):252-255.

Askham J, Gilhooly M, Parkette T, *et al*. (2007). Speculations on the future of taught Masters courses in gerontology. *Gerontology and Geriatrics Education*, 27(3):27-47.

https://doi.org/10.1300/J021v27n03_03

Association for Gerontology in Higher Education. (2014). *Gerontology Competencies for Undergraduate and Graduate Education*. Available from: https://www.aghe.org/resources/gerontologycompetencies-for-undergraduate-and-graduate-education [Last accessed on 2017 Jul 01].

Bengston V and Setterston AR Jr. (eds.) (2016). *Handbook of Theories of Aging*. 3rd ed. New York: Springer.

Bowling A. (1993). The concepts of successful and active aging. *Family Practice*, 10(4):449-53.

https://doi.org/10.1093/fampra/10.4.449

Dimech J. (2019). *Moral Reasoning in End-of-life Decision-making for Persons in End-stage Dementia*. Unpublished Dissertation in Fulfilment for the Doctor of Philosophy, Department of Gerontology and Dementia Studies, University of Malta.

European Commission. (2015). *The 2015 Ageing Report: Economic and Budgetary Projections for the EU 28 Member States (2013-2060)*. Luxembourg: European Commission.

Formosa M and Galea R. (2020). Critical educational gerontology at a senior centre in Malta: Possibilities and limitations for critical consciousness. *Educational Gerontology*, 46(2):59-71.

https://doi.org/10.1080/03601277.2020.1711587

Formosa M. (2015). *Ageing Policy in Malta: Issues, Policies and Future Trends*. Malta: Book Distributors Limited.

Formosa M. (2017). Responding to the active ageing index: Innovations in active ageing policies in Malta. *Journal of*

Population Ageing, 10(1):87-99.

https://doi.org/10.1007/s12062-016-9163-1

Formosa M. (2018). National policies for healthy ageing in Malta: Achievements and limitations. *Healthy Aging Research*, 7(1):1-6.

http://dx.doi.org/10.1097/HXR.0000000000000017

Formosa M. (2019). Active ageing in the fourth age: The experiences and perspectives of older persons in long-term care. *Geopolitical Social Security and Freedom Journal*, 2(1):78-92.

https://doi.org/10.2478/gssfj-2019-0008

Formosa M. (2021). Building evidence for the impact of older adult learning on active ageing: A quantitative study. *Studies in Adult Education and Learning*, 27(2):53-74.

https://doi.org/10.4312/as/9934

Friedman SM, Gillespie SM, Medina-Walpole AM, *et al*. (2013). "Geriatricizing" hospitalists: Identifying educational opportunities. *Gerontology and Geriatrics Education*, 34(4):409-420.

Hertz JE, Douglass C, Johnson A, *et al*. (2007). An emerging typology of academic interdisciplinary gerontology centers in the United States. *Gerontology and Geriatrics Education*, 28(1):19-35.

https://doi.org/10.1300/J021v28n01_02

International Association of Gerontology and Geriatrics. (n.d.). *International Association of Gerontology and Geriatrics*. Available from: https://www.iagg-er.net [Last accessed on 2020 Feb 01].

Maiden JR, Horowitz B and Howe LJ. (2005). Workforce training and education gaps in gerontology and geriatrics: What we found in New York State. *Gerontology and Geriatrics Education,* 31(4):328-348.

Michel JP and Cha HB. (2015). Filling the educational geriatric gap around the world. *The Journal of Post-Acute and Long-Term Care Medicine*, 16(12):1010-1013.

https://doi.org/10.1016/j.jamda.2015.08.020

Mwangi, MS, Yamashita T, Ewen HH, *et al*. (2012). Globalization of gerontology education: Current practices and perceptions for graduate gerontology education in the United States. *Gerontology and Geriatrics Education*, 33(2):198-217.

National Statistics Office. (2021). *World Population Day: 11 July 2021*. Available from: https://nso.gov.mt/en/News_Releases/Documents/2021/07/News2021_122.pdf [Last accessed on 2021 Apr 04].

Parliamentary Secretariat for Rights of Persons with Disability and Active Ageing. (2013). *National Strategic Policy for Active Ageing: Malta 2014-2020*. Available from: https://activeageing.gov.mt/wp-content/uploads/2021/05/ActiveAgeingPolicyEN.pdf [Last accessed on 2021 Apr 04].

Parliamentary Secretariat for Rights of Persons with Disability and Active Ageing. (2015a). *National Dementia Strategy: Malta 2015-2023*. Available from: https://www.alzheimer-europe.org/sites/default/files/2021-10/Malta%20National%20Dementia%20Strategy%202015-2023.pdf [Last accessed on 2021 Apr 04].

Parliamentary Secretariat for Rights of Persons with Disability and Active Ageing. (2015b). *National Minimum Standards for Care Homes for Older People*. Available from: https://activeageing.gov.mt/wp-content/uploads/2021/02/NationalMinimumStandardsforCareHomeEN.pdf [Last accessed on 2021 Apr 04].

Pianosi B and Payne H. (2014). The professionalization of gerontology: A Canadian perspective. *Educational Gerontology*, 40(11):834-846.

https://doi.org/10.1080/03601277.2014.904146

Russell C. (1990). A multidisciplinary course in gerontology: Sociological issues and teaching strategies. *Educational Gerontology*, 16(2):151-163.

https://doi.org/10.1080/0380127900160203

Scerri A. (2018). *Implementing Person-centred Dementia Care in a Rehabilitation Hospital through an Appreciative Inquiry Approach*. Unpublished Dissertation in Fulfilment for the Doctor of Philosophy, Department of Gerontology and Dementia Studies, University of Malta.

Scerri C. (2015). Dementia in Malta: Trends, policy and professional development. In: M Formosa and C Scerri (eds.). *Population ageing in Malta: Multidisciplinary Perspectives*. Malta: Malta University Press, p.161-191.

Silverstein N and Fitzgerald K. (2017). Educating a new generation of professionals in aging worldwide. *Gerontology and Geriatrics Education*, 38(1):1-4.

https://doi.org/10.1080/02701960.2016.1276015

Tam M. (2014) Understanding and theorizing the role of culture in the conceptualizations of successful aging and lifelong learning. *Educational Gerontology*, 40(12):881-893.

https://doi.org/10.1080/03601277.2014.907072

University of Malta. (n.d.). *Department of Gerontology and Dementia Studies*. Available from: https://www.um.edu.mt/socialwellbeing/gerontology [Last accessed on 2020 Feb 01].

van Rijsselt RJ, Parkatti T, and Troisi J. (2007). European initiatives in postgraduate education in gerontology. *Gerontology and Geriatrics Education*, 27(3):79-97.

https://doi.org/10.1300/J021v27n03_06

Weil J. (2017*). Research Design in Aging and Social Gerontology: Quantitative, Qualitative, and Mixed Methods*. London: Routledge.

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Association between cognitive functioning and active life engagement: A time-use study of older adults in rural China

Huijun Liu[1], Yaolin Pei[2]*, and Bei Wu[2]

[1]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, Shaanxi Province, China

[2]Rory Meyers College of Nursing, New York University, New York City, New York, USA

Abstract

This study aimed to examine the pattern of active life engagement and the association between cognitive functioning and active life engagement among older adults in rural China. Two waves of panel data with the previous day's activities in a time-use survey were collected among older adults age 60 and older in rural China. Logistic and OLS regressions were used to examine the impacts of cognitive functioning on participation and intensity in six types of activities. The overall active life engagement level of older adults in rural China was relatively low. Cognitive functioning and its decline significantly associated with the active life engagement. Older adults with higher cognitive functioning were more likely to engage in household work, recreational activity, and socially connected activity, and the decline in cognitive functioning was also significantly associated with the lower likelihood of engaging in household work, recreational activity, physical activity, and lower intensity of socially connected activity. Participation in diverse life activities is an important component of successful aging. The findings of this study suggest the need for increasing awareness of the influence of cognition on daily activities. Future interventions need to consider cognitive health to maximize active life engagement in Chinese rural older adults.

Keywords: Cognitive functioning; Decline; Time-use; Active life engagement; Successful aging

***Corresponding author:**
Yaolin Pei
(yp22@nyu.edu)

Citation: Liu H, Pei Y, Wu B. (2022). Association between cognitive functioning and active life engagement: A time-use study of older adults in rural China. *International Journal of Population Studies*, 8(1):49-58.
https://doi.org/10.36922/ijps.v8i1.1301

Received: March 17, 2022

Accepted: June 13, 2022

Published Online: June 27, 2022

Copyright: © 2022 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Active life engagement which has been defined as having two major components "remaining involved in activities that are meaningful and purposeful" and "maintaining close relationships" (Rowe and Kahn, 1998) – is a useful umbrella under which to categorize productive activities, socially connected activity, recreational activity, and physical activity. Although opportunities for active life engagement may decline with age, continuing participation in diverse life activities is an important component of successful aging and can promote the health and well-being of older adults (Dong, Li, and Simon, 2014; Rowe and Kahn, 1997; Rowe and Kahn, 1998). However, the ability to perform daily activities such as cognitive functioning may determine the kind of activities that older adults can participate. Identifying which types of activities are mostly determined by the cognitive functioning is critical for interventions of promoting successful aging.

It is evident that active life engagement of older adults is constrained by their external environment and individual resources, such as facilities, age, and socioeconomic status, as well as their own health conditions (Mejía, Ryan, Gonzalez, *et al*., 2017). At the same time, although not inevitable, cognitive impairment become more common in late life. Therefore, in addition to improve active life engagement before the onset of functional decline, it is also important to promote active life engagement in the wake of impairments in cognitive functioning.

Maintaining active life engagement is a significant challenge in rural China, due to a lack of resources at individual and community level. This issue is exacerbated in the face of rapid increase of aging population in rural areas (Chen and Liu, 2009; Peng, 2011). The previous studies on this issue have mainly focuses on Western countries with cross-sectional designs, and little is known about the relationship between cognitive impairment and active life engagement among older adults in rural China. Furthermore, research on the influence of cognitive impairment on active life engagement is only limited among older adults with some chronic diseases (Ben Ari, Johansson, Ytterberg, *et al*., 2014; Desrosiers, Demers, Robichaud, *et al*., 2008; Lenze, Munin, Dew, *et al*., 2004).

Diary-based methods that capture individuals' time-use for a single day allows researchers to detail the participation of daily activities that are comprised active life engagement and address the limitations of the previous studies that used aggregated measures of time-use (Freedman, Cornman, Carr, *et al*., 2019; McKenna, Liddle, Brown, *et al*., 2009). Using this approach, each participant is asked to recall what he/she was doing on the previous day, specific activities are clearly outlined and participation in daily activities can be distinguished. In addition, time-use diaries are less likely to be influenced by biases related to social desirability and are more accurate and precise than aggregated measures (e.g., general survey questions) in understanding the daily life of participants (Brenner, 2011; Sabbath, Matz-Costa, Rowe, *et al*., 2016).

Using a two-wave representative sample of older adults in rural China, the purpose of this study is to examine the pattern of active life engagement among rural Chinese older adults and investigate the impact of level and declines of cognitive functioning on different kinds of activities.

1.1. The rural China setting

Rural older Chinese population represent a unique opportunity for studying the relationship between cognitive functioning and active life engagement. The majority of rural Chinese older adults may have to make a living by participating in the paid work beyond age 65 and until their health deteriorates, due to the out-migration of their adult children, and a lack of comprehensive social security system (Zhang, 2010; Zhou, 2012). In addition, due to significant rural-urban disparities in socioeconomic development, rural villages are less likely to have cultural, wellness and recreational facilities (e.g., senior activity centers and chess room) than communities in urban areas (Liu, 2016). Moreover, rural Chinese older adults are less likely to participate in volunteering activities compared to urban China, because most volunteer programs in China are organized by government agencies and mainly targeted to retired professionals (Luo, Pan, and Zhang, 2019).

1.2. Cognitive functioning and active life engagement

The International Classification of Functioning, Disability and Health Model (ICF model) developed by the World Health Organization, provides a relevant conceptual framework for studying the association between functional ability and active life engagement in older adults (the World Health Organization, 2001). The phenomenon of participation in activities not only is related to impairments but also links to the environment individuals reside (Arnadottir, Gunnarsdottir, Stenlund, *et al*., 2011). According to this model, active life engagement is constrained by cognitive functioning, as well as contextual factors that include personal (e.g., age and gender) and external environmental factors (e.g., region). The ICF Model has been widely used by rehabilitation researchers to identify risk factors (e.g., vision, physical functioning, and cognitive functioning) for active life engagement and evaluate rehabilitation success after their therapeutic interventions among older adults with chronic diseases (Alma, Van der Mei, Melis-Dankers, *et al*., 2011; Anaby, Miller, Eng, *et al*., 2009; Ben Ari, Johansson, Ytterberg, *et al*., 2014; Cimarolli, Boerner, Reinhardt, *et al*., 2017; Desrosiers, Demers, Robichaud, *et al*., 2008; Lenze, Munin, Dew, *et al*., 2004). It is noted that the ICF model could be applied to general populations of older adults, not only those with impairments or chronic disease (Levasseur, Desrosiers, and Tribble, 2007).

Cognitive impairment may limit everyday activities in older adults. For those rural Chinese older adults, withdrawal from economy activities and social activities is generally not voluntary but due to the decline of physical and cognitive functioning (Arnadottir, Gunnarsdottir, Stenlund, *et al*., 2011; Rosso, Taylor, Tabb, *et al*., 2013). However, there is limited research on the association between cognitive functioning and active life engagement among rural older adults in China.

Impairments in cognitive functioning may decrease rural older adults' ability to participate in economic and physical activities that need considerable motor skills and function in decision making. Some evidence shows that health status is associated with paid work participation, but is not associated with other activities (Sabbath, Matz-Costa, Rowe, *et al*., 2016). Severe levels of limitations in cognitive functioning may also influence participation in activities that need some cognitive reserves, such as household work, caregiving activities, and socially connected activity. For example, individuals with possible and probable dementia were less likely to engage in social connecting activities (Parisi, Roberts, Szanton, *et al*., 2017).

Among older adults who have hip fracture, cognitive impairment is associated with less exercises (Lenze, Munin, Dew, *et al*., 2004). In a study that included adults with multiple sclerosis, cognitive impairment was associated with restrictions in participations in domestic, leisure, and mobility activities (Ben Ari, Johansson, Ytterberg, *et al*., 2014). While these studies were conducted in outpatient clinics and hospitals, limited studies have focused on community-dwelling older adults. More recently, in a systematic review and meta-analysis, Stolwyk et al., (2021) and his colleagues found that the decline of cognitive functioning is associated with reduced active life engagement (Stolwyk, Mihaljcic, Wong, *et al*., 2021). Therefore, we expect that low level cognitive functioning and its decline are associated with low levels of active life engagement among older adults in rural China.

2. Data and Methods

2.1. Data collection

Data were drawn from the longitudinal study on "Well-being of Elderly Survey in Anhui Province (WESAP)," a panel survey conducted every 3 years between 2001 and 2018 in rural townships of Chaohu, Anhui province. A total of seven waves of surveys were completed by trained interviewers. Using a stratified multistage sampling design, 1800 older adults aged 60 and older from 72 randomly selected villages within six rural townships were interviewed in the first wave in 2001 and 1715 provided valid information. The number of original respondents who participated in the follow-up surveys in 2003, 2006, 2009, and 2012 was 1391, 1067, 807, and 605, respectively. To replace the deceased and lost to follow-up elders, the 2009 and 2015 surveys also randomly recruited 420 (aged 60–68) and 480 (aged 60–65) refreshment cohorts from the same villages using the same sampling design.

This study focused on respondents who participated in the sixth (2015) and the seventh (2018) waves of the survey. The sixth wave served as the baseline in the current analysis, because it was the first time that the time-use diary data were specifically collected in the WESAP questionnaire survey. Pooling data from Wave 6 and 7 allowed us to examine, whether the baseline cognitive functioning and its decline were associated with time-use patterns 3 years later. Our sample consists of 1243 older adults age 60 and older who were interviewed in 2015. After removing 128 respondents who died before the 2018 survey, 95 who were lost to follow-up in 2018, and five cases with missing values in daily time-use and functioning, a total of 1015 older adults were included in this study. Compared with the excluded samples, the included ones were younger, more likely to be married and living with others, with higher income and more social support.

2.2. Measures

2.2.1. Outcome: Time-use on daily activities

The pattern of time-use which reflects active life engagement of older adults was assessed by a comprehensive 24-h recall measure. Following the general time-use data collection guidelines, this instrument was developed for participants' time used during the day before the date of survey completion. Investigators applied this time-use instrument to calculate the amount of time spent on each activity from a list of 17 items, such as housekeeping, caregiving, working at home, working away from home, leisure time, and sleep. To explore how the functional health of older adults specifically influence different domains of their active life engagement in a rural China's social setting, we grouped activity items into the following five categories: economic work (i.e., paid work and earning money from farming, business, manufacturing), household work (i.e., the activities for household maintenance such as cooking, cleaning, doing laundry, and caregiving activities to grandchildren, parents, or spouse), recreational activities (i.e., watching television/movie, reading books/newspapers, listening to the radio, and surfing the internet), physical activities (i.e., sports, walking, or dancing), and socially connected activities (i.e., making phone calls, chatting, and playing chess). Sedentary time for rest (e.g., napping or doze off) was also included as a reversed measure of active life engagement. These groupings distinguished the two dimensions of daily activities (social vs. solitary and active vs. sedentary) based on prior studies of time-use among older adults and the active engagement literature (Lee, Chi, and Palinkas, 2019; Lennartsson and Silverstein, 2001; Simone and Haas, 2013). Following the two-dimension measure of daily activities, the more time spent on both social and active activities (e.g., economic work), social activities (e.g., social connected activities), or active activities (e.g., physical activities), the less time spent on both solitary and sedentary activities (e.g., sedentary time), and the more active life engagement is.

In our study, the diary measure of time-use may not add up to 24 h, because some categories of time-use including volunteering that rarely happened in rural China and personal care activities such as sleeping, bathing, or eating at home were not considered in this study.

The amount of time spent on the activities of each category was calculated. For each activity, we measured with both participation and intensity (or daily duration of participation). Participation was assessed by a binary indicator of whether an individual participated in a given activity. For the subset who participated, the number of hours spent on that activity was used to assess intensity. We, then, grouped activities into the six broad categories of the interest for the current study and recalculated these measures.

2.2.2. Key predictors: Cognitive functioning

Cognitive functioning was considered as the key predictors of active life engagement in this study, which was measured both at the baseline level and its decline in follow-up.

The level of cognitive functioning was measured with the Chinese version of the Mini-Mental State Examination (MMSE). The MMSE mainly focuses on four dimensions of cognitive functioning: orientation, calculation, language or comprehension, and recall (Folstein, Folstein, and McHugh, 1975). In the Chinese version, we, further, modified some items according to the rural Chinese cultural and socioeconomic conditions, to improve the validity and the reliability of the scale implemented in the rural Chinese population. For example, the participants were asked to write a sentence, which is impossible for illiterate older adults to complete. We had to drop this item. The modified Chinese version of MMSE reached a high reliability in this study (cronbach's $\alpha = 0.85$). The combined score of MMSE ranges from 0 to 24, with higher scores indicating better cognitive functioning. The decline of cognitive functioning was coded as 1 if the score of cognitive function in 2018 is lower than that in 2015, otherwise was coded as 0.

2.2.3. Confounders: Sociodemographic variables and physical health

The sociodemographic variables of age, gender, marital status, education, and household income, which had been identified as important factors of health and time-use of older people, were controlled in the data analysis. We included age as a continuous variable and gender as a dichotomous variable (male = 1). Marital status was coded as married/living with partner as 1 and 0 as otherwise. Educational attainment was constructed as a dichotomous variable: "illiterate" or "literate" (primary school or above). Household income was assessed by the total amount of earnings of the individual and his/her spouse in the previous 12 months, including pensions, part-time income, and earnings from self-employed activities. It was transformed using in form in regression models.

We also included living arrangement, social support, presence of chronic disease(s), and Activities of Daily Life (ADLs) as potential confounders. Living arrangement was coded as a dichotomous variable (living alone = 1 and others = 0). Social support was measured by three indicators (continuous variables). The first one asked respondent "how many relatives and friends are contacted at least once within 1 month." The second one asked respondent "how many relatives and friends that you feel safe to discuss personal issues are there." The third one asked respondent "how many relatives and friends that you can look for help are there." Three dichotomous variables were used to indicate the presence of each of the following chronic diseases: diabetes, hypertension, and cardiovascular disease (heart disease or stroke). ADLs were assessed through a series of questions about activities of daily living and instrumental activities of daily living (ranges from 15 to 45), with a higher score indicating the older person exhibits better capability of performing these activities.

2.3. Analytical methodology

This study adopted both univariate and multivariate analyses of our two-wave panel data. Descriptive statistics were presented to summarize sample characteristics, cognitive functioning, and the time-use pattern. We also conducted two sets of analyses to examine the impacts of cognitive functioning level in 2015 and its decline on the likelihood and the intensity of participating different activities in 2018. According to the two measures of each activity, we first run logistic regressions to model the impacts of cognitive functioning level in 2015 and its decline on probability of participating in these activities or prevalence of participation in 2018. Then, we applied OLS regressions among participants to model the intensity (or daily duration) of participating in each category of activities, which given that our outcome was a count of the number of hours spent. All models were adjusted for potential confounders including sociodemographic characteristics, living arrangement, social support, presence of chronic disease, ADLs in 2015, and respective activities in 2015. In addition, normal test showed that the distributions of hours spent in six activities were skewed. Ln format was used to transform the dependent variables in OLS regression. All analyses were conducted using Stata 15 software.

3. Results

The characteristics of the sample in 2015 and functional health of sampled older adults are shown in Table 1. The

average age of the sampled older men and women was 70 years old (range between 60 and 98), with over 70% married and 51% male. <40% respondents were literate. On average, the annual income of sampled household was about 6640 Yuan (Chinese Currency, equivalent to $980) and 20% of the respondents lived alone. In terms of the presence of chronic disease, about 9% of the respondents had diabetes, almost 40% had hypertension, and over 20% had cardiovascular disease. Table 1 also presents the cognitive functioning (level in 2015 and decline in follow-up) of sampled older men and women. On average, almost half of the respondents reported a decline in their cognitive functioning.

Table 2 shows the pattern of active life engagement in 2018 among older adults in rural China. In 2018, about 56% of respondents reported that they were engaged in economic work and generally worked approximately 4 h; household work was performed by 75% of respondents; over 72% were engaged in recreational activity but only 37% performed physical activity; less than half (46%) participated in socially connected activity; and 88% of respondents spent 5 h on average for sedentary time.

Table 3 presents results of logistic regression for engagement of six types of activities after controlling engagement of six types of activities at baseline, respectively, socialdemographic characteristics, and physical health in 2015. Estimates suggest that older adults with higher level of cognitive functioning were more likely to engaging in household work, recreational activity, and social connected activity. The association between cognitive functioning and engagement of physical activity was also significant at marginal level (10%). Maintaining their cognitive functioning without decline was also significantly associated with higher possibility of engaging in household work, recreational activity, and physical activity. However, decline of cognitive functioning in follow-up also increased the possibility of sedentary time at marginal level.

Table 1. Sample characteristics in 2015 and cognitive functioning

Variables	Mean (SE)/%	Coding/ranging
Gender	50.84	0=women; 1=men
Age	70.05 (7.77)	60–98
Marital status	72.65	0=single; 1=married
Education attainment	36.67	0=illiterate; 1=literate
Annual income	6639.54 (9243.69)	0–100000
Living arrangement	20.00	0=living with others; 1=living alone
Social support 1	3.19 (0.07)	0–10
Social support 2	2.88 (0.07)	0–9
Social support 3	3.43 (0.08)	0–10
Having diabetes	8.63	0=no; 1=onset
Having hypertension	38.73	0=no; 1=onset
Having cardiovascular disease	21.18	0=no; 1=onset
ADLs	41.77 (5.46)	15–45
Cognitive functioning in 2015	15.74 (4.83)	0–24
Cognitive functioning decline in 2018	47.59	0=no; 1=yes

Table 2. Time spent on daily activities among older adults in rural China in 2018

Activities	Percentage	Hours/mean (SD)
Economic work	56.29	4.17 (4.81)
Household work	75.25	4.06 (4.42)
Recreational activities	72.89	3.18 (3.49)
Physical activities	37.43	1.04 (2.24)
Socially connected activity	45.97	2.35 (3.31)
Sedentary time	88.31	5.11 (3.82)

Table 3. Estimates of logistic regression for engagement of six types of activities in 2018 (n=1015)

Variables	Economic work	Household work	Recreational activity	Physical activity	Sedentary time	Social connection
Engagement of activities in 2015	2.03***	2.17***	2.21***	2.05***	1.36	1.99***
Gender (Women for reference)	1.42*	0.18***	1.36+	1.01	1.15	0.84
Age	0.93***	0.97*	1.00	1.00	0.99	1.02*
Marital status (Single for reference)	1.15	0.68	0.75	0.82	0.80	1.07
Education (Illiterate for reference)	0.92	1.00	1.21	0.95	0.97	0.91
Annual income	1.13*	0.99	1.08	0.95	0.96	0.88*
Living alone	0.75	1.24	0.73	0.94	0.70	1.49+
Social support 1	0.95	0.98	0.97	1.06	0.95	0.99
Social support 2	1.04	0.95	0.97	1.02	1.08	0.97
Social support 3	0.98	1.01	1.06	0.96	0.89*	0.97
Having diabetes	1.46	1.59	1.49	1.22	0.94	1.54+
Having hypertension	0.86	1.03	0.95	1.33*	0.95	1.16
Having cardiovascular disease	1.11	0.93	1.04	1.26	0.89	0.88
ADLs	1.09***	1.06***	1.02	1.04*	0.99	1.04*
Cognitive function	1.02	1.09***	1.08***	1.03+	0.96	1.06**
Cognitive decline	0.81	0.66*	0.56***	0.73*	1.44+	1.00
Constant	0.74	1.45	0.15	0.07*	88.73*	0.03**
LR Chi2	271.04	182.40	116.00	57.67	18.20	78.14
Pseudo R2	0.1949	0.1606	0.0978	0.0430	0.0248	0.0558

Significance levels: +$P < 0.1$, *$P < 0.05$, *$P < 0.01$, ***$P < 0.001$.

Only among the subsets who participated in each of the six types of activities in 2018 (duration hours >0), we run the OLS regression models to examine the impacts of cognitive functioning level in 2015 and its decline on intensity of participating six types of activities, respectively. The estimates of OLS regression after controlling participating intensity of six types of activities at baseline (including duration hours = 0), respectively, socialdemographic characteristics, and physical health are as shown in Table 4. The level of cognitive functioning was not significantly correlated to participating intensity of any of the six types of activity besides that the association between cognitive functioning and intensity of social connected activity was significant at 10% level. Cognitive functioning decline in follow-up only significantly reduced the frequency of participating social connected activity.

4. Discussion

Active life engagement has been identified as a critical indicator of successful aging. The study examined the patterns of time-use in older adults of rural China using the two waves of time diary data from rural China. On an absolute level, participants had low levels of active life engagement with more than 70% of rural older adults participating in sedentary recreational activities, 88% spending 5 h on average for sedentary time. Only 38% and 46% engaged in physical activity and socially connected activity, respectively. The impacts of cognitive functioning on probability and intensity of active life engagement varied by activities' type. After controlling baseline participation or intensity of activities, respectively, socialdemographic characteristics and physical health, cognitive functioning, and its decline significantly associated with house work, recreational activity, physical activity, and socially connected activity. These findings highlight the importance of promoting cognitive functioning to facilitate the active life engagement.

Consistent with some previous studies (Chou, Chow, and Chi, 2004; Dong, Li, and Simon, 2014; Gauthier and Smeeding, 2003), we found that the majority of older adults (over 70%) in rural China averagely spend more than 3 h in sedentary recreational activities (such as watching TV and reading or listening to radio) or just in sedentary time for rest (such as napping or doze off). Only about one third and 45% of respondents engaged in physical activity and socially connected activity, respectively. However, more than half of rural Chinese older adults still averagely spent 4 h on economic work. The previous studies suggested

Table 4. Estimates of OLS regression for intensity of six types of activities in 2018

Variables	Economic work	Household work	Recreational activity	Physical activity	Sedentary time	Social connection
Intensity of activities in 2015	0.16***	0.07+	0.17***	0.08	0.12**	0.07+
Gender (Women for reference)	0.05	−0.46***	0.13*	0.07	−0.01	−0.08
Age	−0.03***	0.01+	0.01+	0.00	0.01**	0.01**
Marital status (Single for reference)	0.11	0.14+	0.08	−0.03	0.12+	−0.00
Education (Illiterate for reference)	0.08	−0.06	0.07	−0.19*	−0.00	0.06
Annual income	0.01	0.02	−0.04*	−0.05	−0.01	−0.04*
Living alone	−0.10	−0.04	0.11	−0.07	0.11+	0.13
Social support 1	−0.01	0.00	−0.01	0.00	−0.03**	−0.02
Social support 2	0.02	0.01	−0.01	−0.00	0.01	−0.00
Social support 3	−0.01	−0.01	0.01	−0.01	−0.01	−0.00
Having diabetes	0.10	−0.06	0.03	−0.02	0.07	−0.10
Having hypertension	−0.06	0.07	0.04	0.03	−0.02	0.11+
Having cardiovascular disease	−0.05	0.07	0.06	−0.02	0.04	−0.01
ADLs	0.01+	0.01	0.00	−0.01	−0.02***	0.01
Cognitive function	−0.01	0.00	−0.00	0.01	−0.01	0.01+
Cognitive decline	0.02	−0.03	−0.06	−0.01	0.04	−0.14*
Constant	2.97***	0.54	0.97*	1.27*	1.94***	0.72
F	11.84	7.78	3.59	0.46	8.11	2.30
R square	0.2544	0.1428	0.0736	0.0422	0.1286	0.0756
N	572	764	740	380	896	468

Significance levels: $+P < 0.1$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

that the majority of rural Chinese older adults may need to participate in economic activities until their health deteriorates due to the lack of financial security (Zhang, 2010; Zhou, 2012). In a systematic review, estimated sedentary time which include sitting time and sedentary recreational activities, is half of that measured objectively (Harvey, Chastin, and Skelton, 2015). The self-reported dairy data in this study may underestimate the actual time in sedentary time. Overall, the participants exhibited low levels of active life engagement.

The present study found that lower level of cognitive functioning and its decline was associated with lower likelihood of engaging in many activities, such as household work, recreational activity, physical activity, and socially connected activity, and associated with a lower intensity of socially connected activity among Chinese older adults. Our study extends the previous findings on the association between cognitive functioning and active life engagement from those with impairments or chronic disease to general community-dwelling older adults. In addition, in line with previous evidence that limitations in physical functioning predicted lower participation across domains of leisure activities (Janke, Davey, and Kleiber, 2006) and a decline in functional capacity was associated with lower rates or ceasing of participation in some leisure activities and social connecting (Janke, Davey, and Kleiber, 2006; Lefrancois, Leclerc, and Poulin, 1997; Strain, Grabusic, Searle, *et al*., 2002). Our study provide evidence that cognitive functioning is also a significant factor that affects individual's time-use pattern or life engaging behaviors in rural China.

Acting as a basic function or ability for maintaining personal autonomy and independence, cognitive functioning is a precondition for older adults to actively participate in most categories of physical, productive, or social-engaged activities. The significant effects of cognitive functioning on household work, recreational activity, and socially connected activity suggest that a threshold in cognitive ability is required for some kinds of activities, which are usually less physical demanding but need more skills in comprehension, decision-making, and organizing. More importantly, though our findings are consistent with previous studies that cognitive functioning is positively associated with frequency of reading newspapers, books, and magazine (Chou, Chow, and Chi, 2004), we found that both cognitive functioning and cognitive decline were not

associated with more likelihood of spending time idling or doing nothing (sedentary time). This implies that the higher percentage and intensity among older adults in rural China who spent in sedentary time are not caused by their limited cognitive functioning, but may due to a lack of resources in individuals and the region of rural China. According to the IFM model, in addition to functioning ability, active life engagement is also constrained by external resources (the World Health Organization, 2001). It is highly possible that poor resource in those rural Chinese older adults make them have few choices, but to participate in sedentary time. Another possibility is that older adults in rural China does not have positive life style due to their low levels of education.

5. Strengthens and Limitations

An important strength of this study is that a 24-h recall time diary was used to collect participants' time-use on all types of activities, which was considered to be more precise and accurate than the questions on time spent collected from general social surveys or stylized retrospective studies (Brenner, 2011; Sabbath, Matz-Costa, Rowe, *et al*., 2016; Ver Ploeg, Altonji, Bradburn, *et al*., 2000). In addition, the data used in this study come from a seven-wave longitudinal survey which has overall high participation level (over 95% responded to the survey) and low rate of lost follow-up (<8%). Our cohort study with the latest two-wave data allowed us to examine the association between changes of functional health and life engagement. Furthermore, because the baseline functional ability was measured before the collection of the study outcomes, this survey data enables us to examine whether cognitive functioning could predict older adults' active life engagement in later time.

However, we need to acknowledge several limitations in this study. First, functional health status is very likely to be both a precursor and a product of healthy time-use, creating a cycle of advantage or disadvantage (Aartsen, Smits, Van Tilburg, *et al*., 2002; Schooler and Mulatu, 2001). The present study only examined the impacts of functional status at baseline and its decline on life engagements in follow-up by controlling the baseline time-use. We will examine the potential reciprocal relationship between functional health and life engagement in the future. Future research should also explore the causal relationship between cognitive decline trajectories and time-use pattern. A second limitation stems from the fact that our data come from a well-defined area of central China, which is thought to largely reflect the social and cultural conditions of poor rural areas. The study findings may not be generalizable to older adults in other rural areas of China. A third limitation is that the six categories of activities are grouped based on the social context of rural China and some activities such as volunteer work (which was considered an important indicator of social engagement) were not considered due to very low participation rate. Therefore, we were not able to examine the association between functional health and volunteering. In addition, one inevitable limitation of longitudinal research is attrition and the self-selection of participants over time. The individuals who dropped out of the study were more vulnerable than those included in our analyses. The excluded sample may have underestimated the effect of cognitive functioning on the life engagement of older parents. Finally, the R square in some of our regression models was very low which may be attributed to the miss of some key variables. Those key variables for engagement of physical activity and sedentary time should be explored in future study.

6. Conclusions

This study shows the overall low levels of active life engagement among older adults in rural China. Cognitive functioning and its decline impact the pattern of active life engagement among older people in rural China, which suggests that the active life engagement of older adults was limited by their declining cognitive functioning. However, their associations with different types of activities were less consistent. The findings of this study suggest the need for increasing awareness of the influence of cognition on daily activities. The early detection of cognitive impairment presents a good opportunity for interventions to maintain older adults' cognitive functioning in rural China. Given that active life engagement is critical to improve healthy life expectancy and the quality of life in older adults, it is important to increase their positive engagement. For example, we did not find the effect of cognitive functioning on sedentary time. Future programs in rural China need to enrich the daily activities based on their cognitive functioning among older adults.

Funding

This work was supported by the Major Project of National Social Science Foundation of China (21ZDA103).

Conflicts of Interest

The Author(s) declare(s) that there is no conflicts of interest.

Authors' Contributions

All authors participated in (a) the study conception and design, or the analysis and interpretation of data, (b) the drafting of the article or its critical revision for important intellectual content, and/or (c) approval of the version to

be published.

Availability of Supporting Data

The data used in this study is not available for the public yet.

Ethics Statement

Ethical approval was obtained from the Ethical Review Committee of Xi'an Jiaotong University.

References

Aartsen MJ, Smits CH, Van Tilburg T, *et al.* (2002). Activity in older adults: Cause or consequence of cognitive functioning? A longitudinal study on everyday activities and cognitive performance in older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(2):153-162.

https://doi.org/10.1093/geronb/57.2.P153

Alma MA, Van der Mei SF, Melis-Dankers BJ, *et al.* (2011). Participation of the elderly after vision loss. *Disability and Rehabilitation*, 33(1):63-72.

https://doi.org/10.3109/09638288.2010.488711

Anaby D, Miller WC, Eng JJ, *et al.* (2009). Can personal and environmental factors explain participation of older adults? *Disability and Rehabilitation*, 31(15):1275-1282.

https://doi.org/10.1080/09638280802572940

Arnadottir SA, Gunnarsdottir ED, Stenlund H, *et al.* (2011). Participation frequency and perceived participation restrictions at older age: Applying the international classification of functioning, disability and health (ICF) framework. *Disability and Rehabilitation*, 33(22-23):2208-2216.

https://doi.org/10.3109/09638288.2011.563818

Ben Ari E, Johansson S, Ytterberg C, *et al.* (2014). How are cognitive impairment, fatigue and signs of depression related to participation in daily life among persons with multiple sclerosis? *Disability and Rehabilitation*, 36(23):2012-2018.

https://doi.org/10.3109/09638288.2014.887797

Brenner PS. (2011). Investigating the biasing effect of identity in self-reports of socially desirable behavior. *Sociological Focus*, 44(1):55-75.

https://doi.org/10.1080/00380237.2011.10571388

Chen F and Liu G. (2009). Population aging in China. In: P Uhlenberg (ed.). *International Handbook of Population Aging*. New York: Springer, p.157-172.

Chou K, Chow NW, and Chi I. (2004). Leisure participation amongst Hong Kong Chinese older adults. *Ageing and Society*, 24(4):617-629.

https://doi.org/10.1017/S0144686X04001813

Cimarolli VR, Boerner K, Reinhardt JP, *et al.* (2017). A population study of correlates of social participation in older adults with age-related vision loss. *Clinical Rehabilitation*, 31(1):115-125.

https://doi.org/10.1177/0269215515624479

Desrosiers J, Demers L, Robichaud L, *et al.* (2008). Short-term changes in and predictors of participation of older adults after stroke following acute care or rehabilitation. *Neurorehabilitation and Neural Repair*, 22(3):288-297.

https://doi.org/10.1177/1545968307307116

Dong X, Li Y, and Simon MA. (2014). Social engagement among US Chinese older adults findings from the PINE Study. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 69(Suppl 2):S82-S89.

https://doi.org/10.1093/gerona/glu152

Folstein MF, Folstein SE, and McHugh PR. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189-198.

Freedman VA, Cornman JC, Carr D, *et al.* (2019). Time Use and experienced wellbeing of older caregivers: A sequence analysis. *The Gerontologist*, 59(5):e441-e450.

https://doi.org/10.1093/geront/gny175

Gauthier AH and Smeeding TM. (2003). Time use at older ages: Cross-national differences. *Research on Aging*, 25(3):247-274.

https://doi.org/10.1177/0164027503025003003

Harvey JA, Chastin SF, and Skelton DA. (2015). How sedentary are older people? A systematic review of the amount of sedentary behavior. *Journal of Aging and Physical Activity*, 23(3):471-487.

Janke M, Davey A, and Kleiber D. (2006). Modeling change in older adults' leisure activities. *Leisure Sciences*, 28(3):285-303.

https://doi.org/10.1080/01490400600598145

Lee Y, Chi I, and Palinkas LA. (2019). Retirement, leisure activity engagement, and cognition among older adults in the United States. *Journal of Aging and Health*, 31(7):1212-1234.

https://doi.org/10.1177/0898264318767030

Lefrancois R, Leclerc G, and Poulin N. (1997). Predictors of activity involvement among older adults. *Activities, Adaptation and Aging*, 22(4):15-29.

https://doi.org/10.1300/J016v22n04_03

Lennartsson C and Silverstein M. (2001). Does engagement with life enhance survival of elderly people in Sweden? The role of social and leisure activities. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 56(6):S335-S342.

https://doi.org/10.1093/geronb/56.6.S335

Lenze EJ, Munin MC, Dew MA, *et al.* (2004). Adverse effects of depression and cognitive impairment on rehabilitation

participation and recovery from hip fracture. *International Journal of Geriatric Psychiatry*, 19(5):472-478.

https://doi.org/10.1002/gps.1116

Levasseur M, Desrosiers J, and Tribble D. (2007). Comparing the disability creation process and international classification of functioning, disability and health models. *Canadian Journal of Occupational Therapy*, 74(Suppl 5):233-242.

https://doi.org/10.1177/000841740707405S02

Liu X. (2016). How does social activities affect the rural elderly life satisfaction? *Population and Economics*, 2(215):40-47.

Luo Y, Pan X, and Zhang Z. (2019). Productive activities and cognitive decline among older adults in China: Evidence from the China health and retirement longitudinal study. *Social Science and Medicine*, 229:96-105.

https://doi.org/10.1016/j.socscimed.2018.09.052

Mejía ST, Ryan LH, Gonzalez R, *et al*. (2017). Successful aging as the intersection of individual resources, age, environment, and experiences of well-being in daily activities. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 72(2):279-289.

https://doi.org/10.1093/geronb/gbw148

McKenna K, Liddle J, Brown A, *et al*. (2009). Comparison of time use, role participation and life satisfaction of older people after stroke with a sample without stroke. *Australian Occupational Therapy Journal*, 56(3):177-188.

https://doi.org/10.1111/j.1440-1630.2007.00728.x

Parisi JM, Roberts L, Szanton SL, *et al*. (2017). Valued activities among individuals with and without cognitive impairments: Findings from the national health and aging trends study. *The Gerontologist*, 57(2):309-318.

Peng X. (2011). China's demographic history and future challenges. *Science*, 333(6042):581-587.

https://doi.org/10.1126/science.1209396

Rosso AL, Taylor JA, Tabb LP, *et al*. (2013). Mobility, disability, and social engagement in older adults. *Journal of Aging and Health*, 25(4):617-637.

https://doi.org/doi.org/10.1177/0898264313482489

Rowe JW and Kahn RL. (1997). Successful aging. *The Gerontologist*, 37(4):433-440.

https://doi.org/10.1093/geront/37.4.433

Rowe JW and Kahn RL. (1998). *Successful Aging*. New York: Pantheon Books.

Sabbath EL, Matz-Costa C, Rowe JW, *et al*. (2016). Social predictors of active life engagement: A time-use study of young-old French adults. *Research on Aging*, 38(8):864-893.

https://doi.org/10.1177/0164027515609408

Schooler C and Mulatu MS. (2001). The reciprocal effects of leisure time activities and intellectual functioning in older people: A longitudinal analysis. *Psychology and Aging*, 16(3):466-482.

https://doi.org/10.1037/0882-7974.16.3.466

Simone PM and Haas AL. (2013). Frailty, leisure activity and functional status in older adults: Relationship with subjective well being. *Clinical Gerontologist*, 36(4):275-293.

https://doi.org/10.1080/07317115.2013.788114

Stolwyk RJ, Mihaljcic T, Wong DK, *et al*. (2021). Poststroke cognitive impairment negatively impacts activity and participation outcomes: A systematic review and meta-analysis. *Stroke*, 52(2):748-760.

Strain LA, Grabusic CC, Searle MS, *et al*. (2002). Continuing and ceasing leisure activities in later life: A longitudinal study. *The Gerontologist*, 42(2):217-223.

https://doi.org/10.1093/geront/42.2.217

Ver Ploeg M, Altonji J, Bradburn N, *et al*. (2000). *Time-use Measurement and Research: Report of a Workshop*. Washington, DC: National Academies Press.

World Health Organization. (2001). *International Classification of Functioning, Disability and Health: ICF*. Geneva: World Health Organization.

Zhang W. (2010). A study on the effects on labor participation of the Chinese elderly. *Population and Economics*, 1(178):85-92.

Zhou C. (2012). The analysis of elderly adriculture labor supply and its determinants in developed rural area. *Population and Economics*, 5(194):95-101.

ACCSCIENCE PUBLISHING

International Journal of Population Studies

REVIEW ARTICLE

Intergenerational relationships and caregiving burden among family caregivers of older adults: A systematic review

Xiachun Ke, and Yuanyuan Fu*

School of Social Development and Public Policy, Beijing Normal University, Beijing, China

Abstract

Although intergenerational relationships play a significant role in maintaining the well-being of both caregivers and care recipients, few studies have examined how intergenerational relationships influence caregiving burden, especially for adult-child caregivers. A systematic search of electronic databases, including Scopus, Web of Science, and PsycINFO, was conducted across quantitative studies published in English in 2012 – 2022 to provide a summary of (1) the operationalization of caregiver burden and intergenerational relationships used in this field and (2) the effect of intergenerational relationships on caregiver burden. Intergenerational relationships are rarely defined, and they are operationalized in multiple and diverging ways. The eight included papers showed varying results but generally indicated that intergenerational relationships exert a great influence on caregiver burden. Further examination of the impact of intergenerational relationships on caregiver burden is an indispensable prerequisite for interventions that could positively influence the health outcome of adult-child caregivers.

Keywords: Adult-child caregivers; Caregiver burden; Caregiving burden; China; Intergenerational relationship; Older adults; Systematic review

***Corresponding author:**
Yuanyuan Fu
(fuyuanyuan00@163.com)

Citation: Ke X, Fu Y. (2022). Intergenerational relationships and caregiving burden among family caregivers of older adults: A systematic review. *International Journal of Population Studies*, 8(1):59-69.
https://doi.org/10.36922/ijps.v8i1.1320

Received: April 22, 2022

Accepted: July 25, 2022

Published: August 12, 2022

Copyright: © 2022 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations

1. Introduction

Demographic aging, as a result of the decline in fertility rates and the increase in life expectancies, has brought serious economic and health challenges globally. The United Nations has reported that the proportion of the world's population over 60 years or old is projected to increase from 14% in 2022 to 22% in 2050 (United Nations, 2022), which will undoubtedly create a huge care demand worldwide. However, in many countries, institutional care and community-based care have not been well developed, and family caregivers still bear the primary responsibility for providing long-term care for older adults (Hoffmann and Rodrigues, 2010).

Family members, as untrained informal caregivers, are faced with practical challenges related to learning new skills and adopting new responsibilities, including providing practical help, personal care, psychological support, transport, and coordination of treatment (Stenberg, Ruland, and Miaskowski, 2010), which may strain caregivers' physical, financial, and psychosocial resources (Pearlin, Mullan, Semple, *et al.*, 1990). Caregiver burden has been the most commonly reported negative outcome in caregiving

(Aung, Aung, Moolphate, *et al.*, 2021; Tramonti, Barsanti, Ghicopulos, *et al.*, 2013; Ugargol and Bailey, 2018), which may lead to emotional distress (Clyburn, Stones, Hadjistavropoulos, *et al.*, 2000), poor health, and decreased quality of life (Iecovich, 2008).

Researchers have recently shown great interest in how the relationship between caregivers and care recipients influences family caregivers' burden. However, most studies have focused on spousal caregivers (Liu and Lou, 2019; Swinkels, Broese van Groenou, de Boer, *et al.*, 2019; Wagner and Brandt, 2018), whereas fewer studies have focused on adult-child caregivers and the impacts of intergenerational relationships on caregiver burden.

Despite the growing interest in this area, review studies are sparse. Solomon *et al.* (2015) reviewed how the relationship quality affects the well-being of mother-daughter care dyads. When the relationship quality is positive, mother-daughter dyads enjoy rewards and mutuality. When the relationship quality is ambivalent or negative, burden, conflicts, and blames conspire, creating a destructive cycle. A recent systematic review and meta-analysis paper focused on the relationship between filial piety and caregiving burden among adult children (Pan, Chen, and Yang, 2022). Filial piety is a key Chinese culture value that determines children's attitudes and behaviors toward their parents (Dong, Zhang, and Simon, 2014), and older Chinese tend to have a strong perception of filial piety, which prescribes that adult children should take the responsibility of caring for them at home (Tang, Wu, Yeung, *et al.*, 2009). Filial piety greatly influences people's behaviors as well as intergenerational relationships (Kim, Kim, and Hurh, 1991). Pan *et al.* (2022) concluded that stronger filial piety may lessen caregiver burden after analyzing 12 studies. Existing literature indicated that intergenerational relationships are related to caregiver burden, unfortunately, no review studies are available to provide a systematic and comprehensive review of the empirical studies about the effect of intergenerational relationships on caregiver burden. Therefore, the purpose of this systematic review is to clarify (1) the operationalization of caregiver burden and intergenerational relationships used in this field and (2) the effect of intergenerational relationships on caregiving burden among adult children. A clearer understanding of this relationship may help validate theories and guide future studies in this field.

2. Methods

2.1. Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page, McKenzie, Bossuyt, *et al.*, 2021) serves as the framework for this systematic review. The PRISMA statement published in 2009 is a reporting guideline designed to help systematic reviewers transparently report why the review was done, what the authors did, and what they found (Moher, Liberati, Tetzlaff, *et al.*, 2009). Over the past decade, advances in systematic review methodology and terminology have necessitated an update to the guideline. The PRISMA 2020 statement replaces the 2009 statement and includes new reporting guidance that reflects advances in methods to identify, select, appraise, and synthesize studies.

A series of literature searches were conducted in June 2022 using PsycINFO, Web of Science, and Scopus to investigate papers that met our inclusion criteria. Our search terms were ["caregiver" OR "caregiving" OR "carer" OR "support provider" OR "caring"] AND ["elder" OR "elderly" OR "old age" OR "older" OR "aged" OR "geriatric" OR "aging" OR "seniors"] AND ["burden"] AND ["filial" OR "intergenerational" OR "trans-generational" OR "multi-generational" OR "inter-generational" OR "child-parent" OR "parent-child"] in either [TITLE] or [ABSTRACT]. The database search limiters involved being written in English and published between 2012 and 2022.

2.2. Inclusion and exclusion criteria

Articles were required to meet the following inclusion criteria: (a) Adult-child caregivers of older adults aged 60 and over were the study population; (b) intergenerational relationships within the adult-child caregiver-care recipient dyads and caregiver burden comprised the focus of the study; and (c) original quantitative studies were included, but we removed qualitative studies, randomized controlled trials, case–control studies, systematic reviews, and meta-analyses. That is because the purpose of this systematic review is to clarify (1) the operationalization of caregiver burden and intergenerational relationships used in this field and (2) the effect of intergenerational relationships on caregiving burden among adult children. For this purpose, compared to other research designs, quantitative studies can help us get a more specific understanding of the operationalization and specific relationship.

The results of the systematic search are shown in the PRISMA flow diagram (Moher, Liberati, Tetzlaff, *et al.*, 2009; Page, McKenzie, Bossuyt, *et al.*, 2021) (Figure 1). We identified a total of 724 records in the initial search, and three additional articles were added through citations and references from retrieved publications and researchers' knowledge. After removing 151 duplicates and 513 irrelevant records through the title and abstract screening, 63 articles were retained for full-text sorting. Fifty-five full-text articles were excluded for the following reasons: (a) Two

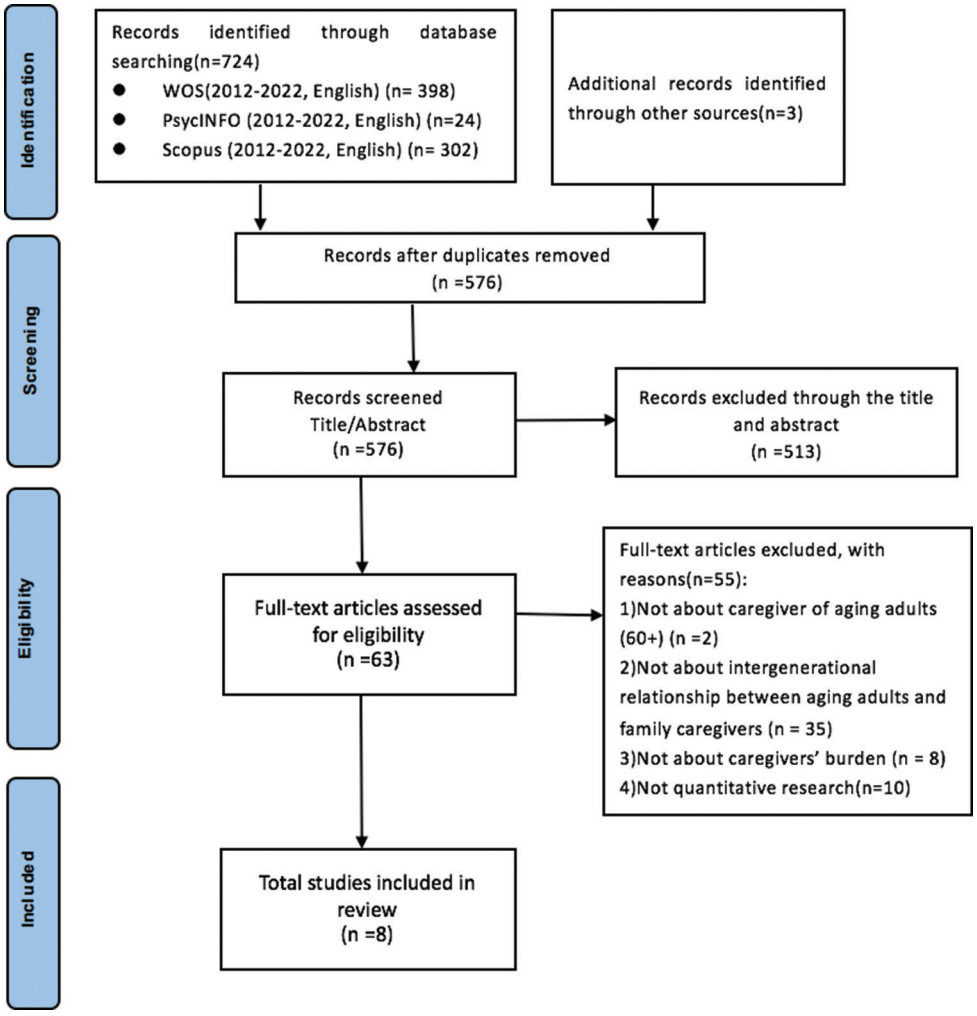


Figure 1. Literature search and selection process.

studies were not about caregivers of 60+ older adults; (b) 35 studies did not focus on intergenerational relationships; (c) eight studies did not focus on caregiver burden, and (d) 10 studies were not quantitative. The two authors performed the entire selection process independently to ensure consistent and complete screenings of all eligible studies. Eight studies were finally included in the structured full paper review for data extraction and synthesis. The detailed data extraction results are shown in Table 1.

3. Results

3.1. Study characteristics

We utilized a pre-developed data extraction template to synthesize study characteristics from each study, and a summary of the eight included studies is found in Table 1.

All eight studies used a cross-sectional design. Concerning the characteristics of the care recipients, only Enright *et al.* (2020) targeted older adults with dementia, whereas other studies did not specifically restrict the health status of the care recipients. Regarding the characteristics of caregivers, five studies included only adult children as caregivers (Aires, Mocellin, Fengler, *et al.*, 2017; Lin, Chen, and Li, 2012; Liu and Bern-Klug, 2016; Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022; Wu, Liu, Cao, *et al.*, 2021). The kinship between caregivers and care recipients in the other three studies was more diverse. Lopez-Anuarbe and Kohli (2019) focused on male caregivers, with sons accounting for approximately 50%, and del-Pino-Casado *et al.* (2014) focused on primary home caregivers of older relatives, with children accounting for 60.1%. In both studies, the researchers conducted separate analyses of the children subgroup and thus were included in this review. Enright *et al.* (2020) adopted a small sample size; the participant sample included 58 informal caregivers of persons with dementia, including 20 adult children. This study was also included in this review because of the relatively large proportion of children in the sample, which

Table 1. Detailed data extraction result.

Authors; year	Study design; location; year	Sample size; kinship	Study aim	Basic theory	Intergenerational relationship-related variable (V); measurement (M)	Caregiver burden measurement	Summary of main results
(Aires, Mocellin, Fengler, *et al.*, 2017)	Cross-sectional; Brazil; 2015	• 100 caregiver children of older adults • daughters (*n*=74) others (*n*=26)	Analyze the association between filial responsibility and the overload of the children when caring for their older parents	NA	• V=care behaviors • M=instrumental support (ADLS/IADLS), emotional support, financial support, keeping company, and visits	Caregiver Burden Inventory (CBI; Novak and Guest, 1989)	Children who were more involved with the ADLs and provided financial support showed higher levels of overload
(del-Pino-Casado, Millán-Cobo, Palomino-Moral, *et al.*,2014)	Cross-sectional; Spain; 2010	• 208 primary home caregivers of older relatives (65 years of age and over) • offspring (*n*=125) spouse (*n*=64) others (*n*=19)	Analyze the effect of cultural factors on the subjective burden of primary home caregivers of older relatives	NA	• V=reciprocity • M= "The care recipient is very grateful, and this gratifies and compensates me." 5-point Likert-type scale (strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree)	Caregiver Strain Index (CSI; Robinson, 1983)	The subjective burden was negatively associated with reciprocity in the offspring group
(Enright, O'Connell, Branger, *et al.*, 2020)	Cross-sectional; Canada; 2009	• 58 informal caregivers of persons with dementia • sons (*n*=3) daughters (*n*=17) spouse (*n*=35) others (*n*=3)	Examine the associations between informal caregivers' perception of identity change in their care partner, the quality of the caregiver/care recipient relationship, and caregiver burden	NA	• V=relationship quality • M=Burns Relationship Satisfaction Scale (BRSS; Burns, Sayers, and Moras, 1994; Heyman, Sayers, and Sayers, 1994)	Zarit Burden Interview (ZBI; Zarit, Orr and Zarit, 1985)	Current caregiver/care recipient relationship quality significantly moderates the association between perceived identity change and caregiver burden
(Lin, Chen, and Li, 2012)	Cross-sectional; Taiwan; NA	• 502 adult children who were married and had parents who were over 65 years old • sons (*n*=198) daughters (*n*=304)	Figure out the relationship among depression, parent-children relationships, and caregiver burden	Conservation of resources theory	• V=parent-child relationship satisfaction • M= "I get along well with my parent" and "I am satisfied with the relationship with my parent." 5-point Likert-type scale (strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree)	Burden Assessment Scale (BAS; Reinhard and Horwitz, 1995)	Caregiver's burden significantly influences caregiver's levels of depression; a better parent-child relationship would weaken the relationship between burden and depression
(Liu and Bern-Klug, 2016)	Cross-sectional; China; 2005(CLHLS)	• 895 dyads of adult-child caregivers and their oldest parent • sons (*n*=392) daughters (*n*=503)	Examine one dimension of caregiver burden "worry about performance" (WaP) and investigate whether variables inspired by Pearlin's stress process model help to explain the variation of WaP	Pearlin's Stress Process Model	• V=caregiver's perception of emotional closeness with the parent • M= "I have a close relationship with my parent." (1= completely agree, 2= agree, 3= disagree, and 4= completely disagree)	Worry about performance: two items from the ZBI, that is, "Do you feel you should be doing more for your parent?" and "Do you feel you should do a better job in caring for your parent?"	Emotionally close relationship with the parent will significantly increase WaP

(Cont'd...)

https://doi.org/10.36922/ijps.v8i1.1320

Table 1. ***(Continued).***

Authors; year	Study design; location; year	Sample size; kinship	Study aim	Basic theory	Intergenerational relationship-related variable (V); measurement (M)	Caregiver burden measurement	Summary of main results
(Lopez-Anuarbe and Kohli, 2019)	Cross-sectional; USA; 2011 and 2015 National Study of Caregiving (NSOC) database	• 2007 caregivers in 2011 and 2204 in 2015 • sons (54% in 2011 and 53% in 2015) husbands/spouses/ partners (26% in 2011 and 28% in 2015) others (21% in 2011 and 20% in 2015)	Figure out the impact of caregiver characteristics, tasks, and resources on caregiving burden of male caregivers as spouses, sons, and other caregivers	The stress appraisal model	• V=felt appreciated by the care recipient • M= "Does the elderly care recipient appreciate you?" (binary: Yes or no)	Is helping the recipient emotionally/ financially/physically difficult for you? (binary: Yes or no)	For sons, feeling appreciated by the care recipient was significantly linked to lower emotional burden in 2015, but not significant in 2011 and other kinds of burden
(Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022)	Cross-sectional; Brazil; NA	• 138 caregivers who were children of elderly people who needed help due to physical, cognitive, or both types of dependency • sons (10.9%) daughters (89.1%)	Verify (a) if a positive interaction and the presence of conflicts in dyads with an adult-child caregiver and an older adult care recipient are related to caregivers' mental health problems (specifically burden and depression) and (b) if the presence of burden and depression in a child caregiver can predict the absence of positive interaction and the presence of conflicts in the dyad	NA	• V=dyadic relationship • M=Brazilian version of the Dyadic Relationship Scale (Sebern and Whitlatch,2007) 11 items divided into two independent subscales: "Positive interaction" (6 items) and "Conflict" (5 items)	ZBI	• The positive interaction between the dyad and the presence of conflicts were, respectively, negatively and positively correlated with burden • The presence of burden was predictive of a worse interaction between the dyad and the presence of conflicts
(Wu, Liu, Cao, *et al.*, 2021)	Cross-sectional; China; 2005 (CLHLS)	• 168 dyads of adult-child caregivers and their parents • sons (n=108) daughters (n=60)	Describe the characteristics of oldest-old Chinese with disability and their adult-child caregivers and the extent to which these characteristics are associated with caregiver burden	NA	• V1=the relationship with the care recipient • V2=willingness to provide care • M1=not very close versus very close • M2=willing to do versus without patience/need respite care/unwilling to do	ZBI	• No significant association was found between close relationships and burden • No significant association was found between willingness to provide care and burden

NA: Not available, specified, or reported

https://doi.org/10.36922/ijps.v8i1.1320

matched our research focus. In addition, regarding sample site and size, three studies were based on data from China (Lin, Chen, and Li, 2012; Liu and Bern-Klug, 2016; Wu, Liu, Cao, *et al.*, 2021), and some studies used a national database (Liu and Bern-Klug, 2016; Lopez-Anuarbe and Kohli, 2019; Wu, Liu, Cao, *et al.*, 2021).

3.2. Relevant theoretical framework

The majority of studies in our review did not adopt a well-established conceptual framework. However, the Pearlin's Stress Process Model (Pearlin, Mullan, Semple, *et al.*, 1990) was used in Liu and Bern-Klug's study (2016). In the Pearlin's Stress Process Model, secondary stressors refer to stressful experiences triggered by primary stressors. This model indicated that the psychosocial resources of caregivers, like the closeness with the care recipients, may influence the secondary stressors.

In addition, the study by Lopez-Anuarb and Kohli (2019) was based on the caregiving stress appraisal model (Yates, Tennstedt, and Chang, 1999), which draws on both the Stress Process Model developed by Pearlin *et al.* (Pearlin, Mullan, Semple, *et al.*, 1990) and the Appraisal Model proposed by Lawton *et al.* (Lawton, Kleban, Moss, *et al.*, 1989; Lawton, Moss, Kleban, *et al.*, 1991). The Appraisal Model relates how caregivers appraise stressful situations with their reactions to them. Therefore, if the caregiver feels that his/her personal and family relationship with the recipient is positive and in line with his/her caregiver duties, family roles, and available resources, his/her burden will be easier to bear and may be lower.

Another study (Lin, Chen, and Li, 2012) was based on the conservation of resources (COR) theory (Hobfoll, 1989). Based on COR, the threat of resource loss becomes a stressor for the individual, which will give rise to psychological distress. In contrast, an individual who has sufficient resources is better at meeting challenges and preventing negative influences from stress. Therefore, positive parent-child relationships will improve caregivers' psychological resources that help mitigate the stress of caring, while negative relationships will consume caregivers' resources and lead to stress.

3.3. Tools for caregiver burden and intergenerational relationships

This systematic review tries to unravel the relationship between intergenerational relationships and caregiver burden, so we further explored how researchers have operationalized these two concepts. Six studies used caregiver burden as an outcome variable in their studies (Aires, Mocellin, Fengler, *et al.*, 2017; del-Pino-Casado, Millán-Cobo, Palomino-Moral, *et al.*, 2014; Enright, O'Connell, Branger, *et al.*, 2020; Liu and Bern-Klug, 2016; Lopez-Anuarbe and Kohli, 2019; Wu, Liu, Cao, *et al.*, 2021). One study used burden as a predictor variable for caregiver depression (Lin, Chen, and Li, 2012), whereas another study viewed burden as both an outcome and a predictor factor in the dyadic relationship (Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022). Regarding measurement tools, four studies used or partially used the Zarit Burden Interview (ZBI) to measure caregiver burden (Enright, O'Connell, Branger, *et al.*, 2020; Liu and Bern-Klug, 2016; Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022; Wu, Liu, Cao, *et al.*, 2021). This scale was developed by Zarit in the 1980s to evaluate the caregiver burden of dementia patients and is composed of 22 items, with a total score range of 0 – 88 (higher scores indicate heavier caregiver burden) (Zarit, Orr and Zarit, 1985). Other scales included the Burden Assessment Scale (BAS), the Caregiver Burden Inventory (CBI), and the Caregiver Strain Index (CSI). Only one study used a binary question rather than a scale to confirm whether caregivers had emotional/financial/physical burdens (Lopez-Anuarbe and Kohli, 2019). In conclusion, in most of these studies, the researchers tended to assess caregiver burden from a comprehensive perspective, including subjective burden and objective burden.

On the other hand, the tools measuring intergenerational relationships were more diverse. For the following analysis, measurements have been divided into three categories based on the solidarity perspective and conflict perspective: Structural-associational solidarity, affectual solidarity, and intergenerational conflict.

3.3.1. Structural-associational solidarity

Queluz *et al.* (2022) used the Dyadic Relationship Scale, which includes a subscale of positive interaction, including items such as "I learned good things about him/her" and "communication between us has improved." The study by Aires *et al.* (2017) also explored the role of care behaviors in affecting the overload of caregivers, such as instrumental support (ADLS/IADLS), emotional support, financial support, keeping company, and visits.

3.3.2. Affectual solidarity

Two studies (Liu and Bern-Klug, 2016; Wu, Liu, Cao, *et al.*, 2021) used a single question in the CLHLS questionnaire to evaluate caregivers' perception of emotional closeness between adult children and older parents. In addition, Enright *et al.* (2020) used the Burns Relationship Satisfaction Scale (BRSS), which consists of 13 items that assess satisfaction in various areas of the relationship (i.e., handling finances and degree of affection and caring). The total scores are the sum of the items and range from 0 to

78, with higher scores indicating greater satisfaction. In addition, in the study by Lin *et al.* (2012), adult children's satisfaction with their intergenerational relationship was determined by assessing how they felt regarding two items: "I get along well with my parent" and "I am satisfied with the relationship with my parent."

3.3.3. Intergenerational conflict

In this review, we found that researchers measured the extent of intergenerational conflict from both the adult-child and parental perspectives. del-Pino-Casado *et al.* (2014) and Lopez-Anuarbe and Kohli (2019) both asked caregivers how they feel appreciation from their parents. In the study by Queluz *et al.* (2022), conflict was assessed by items such as "I felt depressed when I had problems with my relationship" and "I felt angry toward him/her." Wu *et al.* (2021) also asked about caregivers' willingness to provide care and divided the answer into willing to do versus without patience/need respite care/unwilling to do.

In summary, although intergenerational relationships are multidimensional and complex, by combining these eight studies, we found that tools for intergenerational relationships were based on two major dimensions, solidarity and conflict. In addition, comprehensive scales were used for measurement in some studies (Enright, O'Connell, Branger, *et al.*, 2020; Queluz, de Santis, de Fatima Kirchner, *et al.*,2022).

3.4. The role of intergenerational relationships in affecting caregiver burden

We summarized the mechanisms of the role of intergenerational relationships in affecting caregiver burden through these eight studies and found that intergenerational relationships can be both stressors of burden and important psychosocial resources for caregivers.

3.4.1. Intergenerational relationships as a stressor of burden

First, with regard to structural-associational solidarity, Aires *et al.* (2017) found that involvements with ADLs and the provision of financial support were positively correlated with burden. Queluz *et al.* (2022) found that a positive interaction between the dyads was negatively correlated with burden. Second, concerning intergenerational conflict, the subjective burden was negatively associated with reciprocity and appreciation from older parents. When adult children receive more positive feedback but not negative emotions from their parents, they feel less burden (del-Pino-Casado, Millán-Cobo, Palomino-Moral, *et al.*, 2014; Lopez-Anuarbe and Kohli, 2019; Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022). Third, we found some contradictory conclusions on how affectual solidarity influences caregiver burden. Liu and Bern-Klug (2016) found that an emotionally close relationship with the parent increased worries about performance, which means that adult children feel more worries and emotional burden when they have more affectual closeness with their parents. However, based on the same database as in Liu and Bern-Klug's study, Wu *et al.* (2021) found that relationship closeness with the care recipient and willingness to provide care was not associated with burden. Last but not least, it deserves more attention that although most of the studies try to determine how intergenerational relationships affect caregiver burden, we found that some studies try to find a bidirectional relationship between intergenerational relationships and caregiver burden. For example, Queluz *et al.* (2022) found that not only were the positive interaction between the dyads and the presence of conflicts negatively and positively correlated with burden, respectively, but also that burden was predictive of a worse interaction between the dyad and of the presence of conflicts.

3.4.2. Intergenerational relationships as psychosocial resources for caregivers

In addition, researchers have also explored the impact of intergenerational relationships as a psychosocial resource for caregivers when faced with challenges and difficulties in caregiving. Lin *et al.* (2012) found that caregiver burden significantly improved caregiver depression and that intergenerational relationships moderated the effect of burden on depression. The relationship between caregiving burden and his or her level of depression was weaker when participants had a better parent-child relationships. Enright *et al.* (2020) explored the relationship among perceived identity change, intergenerational relationship quality, and caregiver burden using a mediation model. They found that relationship quality mediated the association between perceived identity change and caregiver burden. Caregiver-reported changes in the identity predicted significantly led to a decrease in relationship quality, which predicted increases in caregiver burden levels.

In summary, the mechanisms by which intergenerational relationships impact adult children's caregiving burden remain unclear and need further exploration through more comprehensively designed empirical research.

4. Discussion

To the best of our knowledge, this systematic review is the first attempt to compare the effects of intergenerational relationships on caregiver burden across studies focusing on family caregivers of older adults. Some studies have pointed out a significantly positive relationship between intergenerational relationships and caregiver burden. However, some studies reported a negative association

between intergenerational relationships and caregiver burden. In addition, some studies have found mediating and moderating roles of intergenerational relationships in the stress process. Several aspects relevant to the findings need to be discussed to facilitate interpretation.

4.1. Measurement of intergenerational relationships

In terms of measurement, the majority of the existing papers adopt well-established scales to measure caregiver burden. However, when measuring intergenerational relationships, only a few studies adopt scales (Enright, O'Connell, Branger, *et al.*, 2020; Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022). Scales to measure caregiver burden included the ZBI, the BAS, the CBI, and the CSI. Intergenerational relationships are a multidimensional concept relative to concepts such as attachment, emotional exchange, emotional support, intimacy, love, mutuality, reciprocity, and so on (Solomon, Hansen, Baggs, *et al.*, 2015). However, the majority of these eight studies did not adopt well-established scales to measure intergenerational relationships. Based on previous findings, we found that current measurements of intergenerational relationships can be divided into the following three main categories: Structural-associational solidarity, affectual solidarity, and intergenerational conflict. Therefore, we believe that in subsequent studies, reliable and valid scales should be used to measure intergenerational relationships to gain a more comprehensive understanding. Several scales have been used to measure relationships, such as the Burns Relationship Satisfaction Scale (Burns, Sayers, and Moras, 1994; Heyman, Sayers, and Sayers, 1994) and the Dyadic Relationship Scale (Sebern and Whitlatch, 2007). However, it is worth noting that not only adult children can be primary caregivers of older adults but also spouses, other relatives, or even friends. Therefore, we suggest that a reliable and valid relationship scale specifically for child-parent care dyads is necessary. For example, Bai (2018) developed and validated an Intergenerational Relationship Quality Scale for Aging Chinese Parents based on solidarity, conflict, and ambivalence models, which enables researchers and service practitioners to accurately measure the relationship quality between older people and their adult children.

4.2. Impact of intergenerational relationships on caregiver burden

Some studies pointed out a significantly positive relationship between intergenerational relationships and caregiver burden. For example, when adult children give parents more support or feel closer to them, they report more burden (Aires, Mocellin, Fengler, *et al.*, 2017; Liu and Bern-Klug, 2016). However, some studies have pointed out the significantly negative effect of intergenerational relationships on caregiver burden. For example, when adult children gain more positive feedback and more positive interactions from their parents, they report less burden (del-Pino-Casado, Millán-Cobo, Palomino-Moral, *et al.*, 2014; Lopez-Anuarbe and Kohli, 2019; Queluz, de Santis, de Fatima Kirchner, *et al.*, 2022). In addition, a not-significant association between intergenerational relationships and caregiver burden was found in one study (Wu, Liu, Cao, *et al.*, 2021). Some researchers have pointed out the importance of intergenerational relationships as a psychological resource for caregivers. A good intergenerational relationship can mitigate the effects of depression on caregiver burden (Lin, Chen, and Li, 2012) and can be a protective factor when facing complex caregiving difficulties (Enright, O'Connell, Branger, *et al.*, 2020). There are two probable reasons why the conclusions vary among studies. On the one hand, as stated above, due to the lack of an accurate operational definition of intergenerational relationships, it cannot be determined whether these inconsistencies in results arise from methodological issues or reflect a weak overall effect of intergenerational relationships on caregiver burden. On the other hand, various cultural contexts can result in different conclusions. These eight studies were based on data from Asia, Europe, North America, and South America. People's understanding and social norms about the parent-child relationships may differ greatly from culture to culture. For example, compared to Western culture, filial piety serves as a core familial norm in the Chinese social value system, which claims that children should provide their older parents with a broad range of support and care (Ikels, 2004). Thus, the impact of intergenerational relationships on caregiver burden in different cultural contexts may be different and deserves more exploration in future studies.

4.3. Pearlin's stress process model

The majority of studies in this review were not based on a well-established conceptual framework. However, the Pearlin's Stress Process Model (Pearlin, 1999; Pearlin, Lieberman, Menaghan, *et al.*, 1981; Pearlin, Mullan, Semple, *et al.*, 1990) is one of the most commonly used frameworks to explain the variation in caregiver health outcomes, including caregiver burden in long-term care studies. In addition, Pearlin *et al.* (1990) pointed out two important concepts relative to intergenerational relationships in this model, that is, family conflict and social support.

First, the Pearlin's Stress Process Model in 1981 reshaped the sociological study of stress and mental health by emphasizing the role of chronic stressors and resources on psychological distress. Pearlin *et al.* (1981) pointed out that disruptive life events might lead to some chronic stressors, such as family conflict. Family conflict might lead

to mental health outcomes, such as burden. In addition, Pearlin *et al.* (1981) hypothesized that social support might inhibit or buffer the effects of stressors on mental health. That is, social support may be either a mediator or moderator in the relationship between stressors and burden. Second, following his seminal publication in 1981, Pearlin's Stress Process Model in 1990 was applied in the caregiving context. In this updated model, Pearlin *et al.* (1990) pointed out that family conflict might be a kind of secondary stressor and that social support is an important mediator for individuals, without explaining this mediating mechanism clearly. However, Pearlin's Stress Process Model in 1999 named social support as "moderating resources" rather than "mediating resources," which is quite different from the version in 1990. Pearlin (1999) pointed out that social support may reduce or contain the intensity of a stressor, inhibit the emergence of secondary stressors, and cushion the effect of the stressors on outcomes. In summary, Pearlin *et al.* held the view that family conflict is a stressor for caregiver burden. However, it's difficult to find a clear way with social support in influencing caregiver burden in his models. Does social support mediate the relationship between social context, stressors, and stress outcomes? Or does social support moderate the effect of stressors on stress outcomes? In our review of empirical studies, we also found the same puzzle that the role of intergenerational relationships on caregiver burden was inconclusive.

4.4. Limitations

The findings of this systematic review should be interpreted with caution due to several limitations. First, considering that demographics are highly related to intergenerational relationships and could drastically change over decades, the searches of the articles were limited to three databases (Web of Science, Scopus, and PsycINFO) over a specific period (2012 – 2022) to obtain the most recent studies. Therefore, future studies can extend the period or databases to obtain more relevant literature. Second, to accomplish our research aim, this review included only quantitative studies, and all the studies had cross-sectional designs. We think that it would be interesting for future studies to include different study designs, such as qualitative studies, randomized controlled trials, case–control studies, and so on, which will help facilitate a better understanding of the complex relationship between intergenerational relationships and caregiver burden.

5. Conclusions

Overall, this systematic review has contributed to the understanding of the relationship between intergenerational relationships and caregiver burden over the past decade. First, measurements of intergenerational relationships have not been clarified, and we suggested that reliable and valid scales should be used to measure intergenerational relationships to gain a more comprehensive understanding. Second, the relationship between intergenerational relationships and caregiver burden shows complexity and great variation; intergenerational relationships can directly alleviate or exacerbate caregiver burden, or they can be important psychological resources. Finally, we suggested that future studies could pay more attention to theories and use some theoretical frameworks to help explain the variation in caregiver health outcomes, such as the Pearlin's Stress Process Model. In conclusion, this systematic review shows that further examining the impact of intergenerational relationships on caregiver burden is an indispensable prerequisite for interventions that could positively influence the health outcomes of caregivers. For this purpose, more studies are needed to deepen the understanding of the different aspects of intergenerational relationships and caregiver burden.

Acknowledgments

None.

Funding

This study was supported by Beijing Education Sciences Fourteenth 5-year Planning Young Scholar Project (Ref. No. BACA21111).

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to research, authorship, and/or publication of this article.

Authors' Contributions

Ke X. and Fu Y. worked together to make substantial contributions to the conceptualization and design of the study, analysis of the data, drafting the manuscript, and revising the manuscript.

Ethics Statement

This study did not involve human subjects. The human data in this study are secondary data from the previous studies in public databases.

Availability of Supporting Data

All data are secondary data from publicly available data sources.

References

Aires M, Mocellin D, Fengler FL, *et al.* (2017). Association between filial responsibility when caring for parents and

the caregivers overload. *Revista Brasileira de Enfermagem*, 70(4):767-774.

https://doi.org/10.1590/0034-7167-2017-0133

Aung T, Aung MN, Moolphate S, *et al.* (2021). Caregiver burden and associated factors for the respite care needs among the family caregivers of community dwelling senior citizens in Chiang Mai, Northern Thailand. *International Journal of Environmental Research and Public Health*, 18(11):5873.

https://doi.org/10.3390/ijerph18115873

Bai X. (2018). Development and validation of a multidimensional intergenerational relationship quality scale for aging Chinese parents. *The Gerontologist*, 58(6):e338-e348.

https://doi.org/10.1093/geront/gnx156

Burns DD, Sayers SL, and Moras K. (1994). Intimate relationships and depression: is there a causal connection? *Journal of Consulting and Clinical Psychology*, 62(5):1033-1043.

https://doi.org/10.1037//0022-006x.62.5.1033

Clyburn LD, Stones MJ, Hadjistavropoulos T, *et al.* (2000). Predicting caregiver burden and depression in Alzheimer's disease. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences*, 55(1):S2-S13.

https://doi.org/10.1093/geronb/55.1.s2

del-Pino-Casado R, Millán-Cobo MD, Palomino-Moral PA, *et al.* (2014). Cultural correlates of burden in primary caregivers of older relatives: A cross-sectional study. *Journal of Nursing Scholarship*, 46(3):176-186.

https://doi.org/10.1111/jnu.12070

Dong X, Zhang M, and Simon MA. (2014). The expectation and perceived receipt of filial piety among Chinese older adults in the Greater Chicago area. *Journal of Aging and Health*, 26(7):1225-1247.

https://doi.org/10.1177/0898264314541697

Enright J, O'Connell M, Branger C, *et al.* (2020). Identity, relationship quality, and subjective burden in caregivers of persons with dementia. *Dementia*, 19(6):1855-1871.

https://doi.org/10.1177/1471301218808607

Heyman RE, Sayers SL, and Bellack AS. (1994). Global marital satisfaction versus marital adjustment: An empirical comparison of three measures. *Journal of Family Psychology*, 8(4):432-446.

https://doi.org/10.1037/0893-3200.8.4.432

Hobfoll SE. (1989). Conservation of resources. A new attempt at conceptualizing stress. *The American Psychologist*, 44(3):513-524.

https://doi.org/10.1037//0003-066x.44.3.513

Hoffmann F and Rodrigues R. (2010). Informal Carers: Who takes care of them? *European Centre for Social Welfare Policy and Research*. Available from https://www.euro.centre.org/downloads/detail/1256 [Last accessed on 2022 Jul 24].

Iecovich E. (2008). Caregiving burden, community services, and quality of life of primary caregivers of frail elderly persons. *Journal of Applied Gerontology*, 27(3):309-330.

https://doi.org/10.1177/0733464808315289

Ikels C. (2004). *Filial Piety: Practice and Discourse in Contemporary East Asia*. Redwood City, California: Stanford University Press.

Kim KC, Kim S, and Hurh WM. (1991). Filial piety and intergenerational relationship in Korean immigrant families. *The International Journal of Aging and Human Development*, 33(3): 233-245.

https://doi.org/10.2190/Y91P-UNGR-X5E1-175K

Lawton MP, Kleban MH, Moss M, *et al.* (1989). Measuring caregiving appraisal. *Journal of Gerontology*, 44(3):P61-P71.

https://doi.org/10.1093/geronj/44.3.p61

Lawton MP, Moss M, Kleban MH, *et al.* (1991). A two-factor model of caregiving appraisal and psychological well-being. *Journal of Gerontology*, 46(4):P181-P189.

https://doi.org/10.1093/geronj/46.4.p181

Lin WF, Chen LH, and Li TS. (2012). Adult children's caregiver burden and depression: The moderating roles of parent-child relationship satisfaction and feedback from others. *Journal of Happiness Studies*, 14(2):673-687.

http://doi.org/10.1007/s10902-012-9348-0

Liu H and Lou VW. (2019). Transitioning into spousal caregiving: Contribution of caregiving intensity and caregivers' multiple chronic conditions to functional health. *Age Ageing*, 48(1):108-114.

https://doi.org/10.1093/ageing/afy098

Liu J and Bern-Klug M. (2016). "I should be doing more for my parent:" Chinese adult children's worry about performance in providing care for their oldest-old parents. *International Psychogeriatrics*, 28(2):303-315.

https://doi.org/10.1017/S1041610215001726

Lopez-Anuarbe M and Kohli P. (2019). Understanding male caregivers' emotional, financial, and physical burden in the United States. *Healthcare*, 7(2):72.

https://doi.org/10.3390/healthcare7020072

Moher D, Liberati A, Tetzlaff J, *et al.* (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4):264-269.

https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Novak M and Guest C. (1989). Application of a multidimensional caregiver burden inventory. *The Gerontologist*, 29(6):798-803.

https://doi.org/10.1093/geront/29.6.798

Page MJ, McKenzie JE, Bossuyt PM, *et al.* (2021). The PRISMA

2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1):1-11.

https://doi.org/10.1186/s13643-021-01626-4

Pan Y, Chen R, and Yang D. (2022). The relationship between filial piety and caregiver burden among adult children: A systematic review and meta-analysis. *Geriatric Nursing*, 43:113-123.

https://doi.org/10.1016/j.gerinurse.2021.10.024

Pearlin LI, Lieberman MA, Menaghan EG, *et al.* (1981). The stress process. *Journal of Health and Social Behavior*, 22(4):337-356.

https://doi.org/10.2307/2136676

Pearlin LI, Mullan JT, Semple SJ, *et al.* (1990). Caregiving and the stress process: An overview of concepts and their measures. *The Gerontologist*, 30(5):583-594.

https://doi.org/10.1093/geront/30.5.583

Pearlin LI. (1999). The stress process revisited. In: CS Aneshensel and JC Phelan (eds.). *Handbook of the Sociology of Mental Health*. Boston, MA: Springer, p. 395-415.

https://doi.org/10.1007/0-387-36223-1_19

Queluz FNF, de Santis L, de Fatima Kirchner L, *et al.* (2022). The influence of the quality of filial relationship on the mental health of children caregivers. *Trends in Psychology*, 30(3):455-467.

https://doi.org/10.1007/s43076-022-00140-3

Reinhard SC and Horwitz AV. (1995). Caregiver burden: Differentiating the content and consequences of family caregiving. *Journal of Marriage and the Family*, 57(3):741-750.

https://doi.org/10.2307/353928

Robinson BC. (1983). Validation of a caregiver strain index. *Journal of Gerontology*, 38(3):344-348.

https://doi.org/10.1093/geronj/38.3.344

Sebern MD and Whitlatch CJ. (2007). Dyadic relationship scale: A measure of the impact of the provision and receipt of family care. *The Gerontologist*, 47(6):741-751.

https://doi.org/10.1093/geront/47.6.741

Solomon DN, Hansen L, Baggs JG, *et al.* (2015). Relationship quality in non-cognitively impaired mother-daughter care dyads: A systematic review. *Journal of Family Nursing*, 21(4):551-578.

https://doi.org/10.1177/1074840715601252

Stenberg U, Ruland CM, and Miaskowski C. (2010). Review of the literature on the effects of caring for a patient with cancer. *Psychooncology*, 19(10):1013-1025.

https://doi.org/10.1002/pon.1670

Swinkels JC, van Groenou MIB, de Boer A, *et al.* (2019). Male and female partner-caregivers' burden: Does it get worse over time? *Gerontologist*, 59(6):1103-1111.

https://doi.org/10.1093/geront/gny132

Tang CS, Wu AMS, Yeung D, *et al.* (2009). Attitudes and intention toward old age home placement: A study of young adult, middle-aged, and older Chinese. *Ageing International*, 34(4):237-251.

https://doi.org/10.1007/s12126-009-9047-2

Tramonti F, Barsanti I, Ghicopulos I, *et al.* (2013). Family functioning and caregiver burden in Alzheimer's disease. *Family Science*, 4(1):161-168.

https://doi.org/10.1080/19424620.2013.871741

Ugargol AP and Bailey A. (2018). Family caregiving for older adults: Gendered roles and caregiver burden in emigrant households of Kerala, India. *Asian Population Studies*, 14(2):194-210.

https://doi.org/10.1080/17441730.2017.1412593

United Nations. (2022). World Population Prospects 2022. New York: The United Nations. Available from: https://www.un.org/development/desa/pd/content/publications [Last accessed on 2022 Jul 24].

Wagner M and Brandt M. (2018). Long-term care provision and the well-being of spousal caregivers: An analysis of 138 European regions. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences*, 73(4):e24-e34.

https://doi.org/10.1093/geronb/gbx133

Wu Y, Liu Y, Cao X, *et al.* (2021). Burden of caregivers who care for oldest-old parents with disability: A cross-sectional study. *Geriatric Nursing*, 42(4):792-798.

https://doi.org/10.1016/j.gerinurse.2021.04.013

Yates ME, Tennstedt S, and Chang BH. (1999). Contributors to and mediators of psychological well-being for informal caregivers. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences*, 54(1):P12-P22.

https://doi.org/10.1093/geronb/54b.1.p12

Zarit S, Orr NK, and Zarit JM. (1985). *The Hidden Victims of Alzheimer's Disease: Families under Stress*. New York: New York University Press.

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Disaggregating the longitudinal association between urbanization and body weight in Chinese adults over 1991 – 2015

Hongwei Xu*

Department of Sociology, CUNY-Queens College, New York, United States

Abstract

Urbanization is widely viewed as a major contextual force behind the rising prevalence of overweight and obese people in developing countries. Research in China often conflates between-community difference and within-community change - two separate processes of urbanization that are related to body weight gain. Capitalizing on longitudinal and multilevel data from the 1991 to 2015 China Health and Nutrition Survey, the present study disaggregated the association between change in a community-level urbanicity index and change in individual-level body weight status over time in Chinese adults aged 18–65 years. A positive longitudinal relationship was confirmed between urbanicity and body weight in men, but varied in women by the choice of anthropometric measure. However, for both men and women, such an overall association was largely driven by preexisting between-community differences in the level of urbanization rather than an intrinsic within-community urbanization process. This pattern is robust against two different disaggregation methods. These findings together confirm the inadequate simplicity of the conventional model of community effects on health and nutrition.

Keywords: Community; Overweight; Obesity

***Corresponding author:**
Hongwei Xu,
(hongwei.xu@qc.cuny.edu)

Citation: Xu, H. (2022). Disaggregating the longitudinal association between urbanization and body weight in Chinese adults over 1991–2015. *International Journal of Population Studies*, 8(1):70-82. https://doi.org/10.36922/ijps.v8i1.334

Received: March 9, 2022

Accepted: August 16, 2022

Published Online: September 8, 2022

Copyright: © 2022 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Being overweight or obese has raised a public health concern for the Chinese population (Ng *et al*., 2014). According to disease surveillance data collected by the Chinese Center for Disease Control and Prevention, the overweight rate among Chinese adults aged 18 years and older nearly doubled from 16.4% in 1992 to 30.1% in 2012, and the obese rate more than tripled from 3.6% to 11.9% during the same period (Chinese CDC, 2015; Wang, 2005). Reasons for population-level weight gain in China are multifaceted and the subject of considerable debate among scholars. Nonetheless, prior research suggests that rising incomes, higher-fat diets, reduced physical activity, and cultural ideals regarding desirable weight all play a role. Because higher levels of urbanization are likely to include a shift from occupations requiring strenuous physical activities to those with more sedentary activities, an increase in automotive use for job commuting and daily activities, more affordable food markets for meat and cooking oil, and easier access to Western fast-food restaurants, many weight-gain-related changes in China are thought

to stem from urbanization, one of the most dramatically changing features of communities in contemporary China (Monda *et al*., 2007; Monda *et al*., 2008; Xu *et al*., 2013).

In China, urbanization can be driven by flows of migrants from rural villages to cities for better life opportunities, resulting in rapid increases in population size and density in existing urban areas. Meanwhile, rural villages can experience *in situ* urbanization fueled by the development of township and village enterprises and the inflow of foreign investment. *In-situ* village urbanization involves changes in the economic structure, wherein the labor force shifts from agricultural activities to activities in manufacturing and service sectors. *In situ* village urbanization also includes changes in the built environment, wherein previously rural areas of farmland and farmhouses are converted into urban areas with modern roads designed for automobiles, factories, and residential and commercial centers (Zhu, 2000; Zhu, 2002).

In situ urbanization involves within-community longitudinal changes in demographics, socioeconomic conditions, and built environment, while urbanization through rural-to-urban migration can be seen as between-community cross-sectional differences in these characteristics. Therefore, *in situ* urbanization is the context often related to the conventional wisdom about the effects of changing community characteristics on individual-level weight gain. That is, increasing *in situ* urbanization in China is thought to have a strong impact on altering individual physical activity patterns in the urbanized areas, which, in turn, is thought to drive the increasing prevalence of overweight and obesity in China (Monda *et al*., 2007; Monda *et al*., 2008). As part of the urbanization process, local food environment is also likely to shift, usually from a high-fiber, low-fat, and low-energy diet to a low-fiber, high-fat, and high-energy diet, contributing to the so-called "obesogenic environments" – contexts that promote obesity by encouraging both physical inactivity and excessive energy intake (Swinburn *et al*., 1999; Mehta & Chang, 2008).

Investigation into the *in situ* urbanization process and its implications for within-community weight gain requires longitudinal data at both the individual and community levels. The China Health and Nutrition Survey (CHNS) has been a valuable data source in this regard. However, as reviewed below, previous analyses of the CHNS data focused on an overall association between community-level urbanization and individual-level body weight status without making either a conceptual or an analytical distinction between within-community urbanization processes and pre-existing between-community differences. Thus, an observed overall association can be theoretically misleading because it may be largely driven by between-community differences, including substantial pre-existing gaps in urbanization levels between rural and urban areas and an uneven pace of urbanization across the entire country (Xie & Hannum, 1996; Yeh *et al*., 2011).

This study seeks to improve our understanding about the overall association between urbanization and weight gain in Chinese adults by disaggregating it into two components: (1) the between-community component that captures differences in level of urbanization at baseline and rate of urbanization over time across communities; and (2) the within-community component that reflects the *in situ* urbanization process and corresponds to the conventional theory of neighborhood effects on health (Roux, 2004; Entwisle, 2007). Drawing on longitudinal and multilevel data from the CHNS, the present study prospectively examines body weight changes in Chinese adults from 1991 to 2015. The analysis of weight change relies on both general adiposity and abdominal adiposity with physically measured anthropometric data, minimizing simplistic erroneous inference that can result from sole reliance on body weight measures (Xu *et al*., 2013). Two disaggregation methods are compared in assessing the relative strengths of between- and within-community components in explaining the longitudinal association between urbanization and weight gain.

2. Prior research and limitations

An ongoing longitudinal panel study first conducted in 1989, the CHNS now includes more than 7000 households across 15 provinces and municipal cities in contemporary China that vary substantially in geography, economic development, public resources, and health indicators. In addition, detailed community-level data are collected from local officials. The long survey period makes the CHNS data extremely valuable for studying the relation between urbanization and body weight status, not only because it usually takes some time for community characteristics to evolve as a result of human activities but also because the study period of the CHNS (1990s and 2000s) is when China experienced unprecedented social changes and urbanization. Therefore, it is not surprising that previous research on community-level urbanization and weight gain in China heavily relies on analyses of the CHNS data.

That being said, a few studies using the CHNS data examined only cross-sectional associations between urbanization and body weight status. For example, Van de Poel *et al*. (2009) performed separate cross-sectional analyses of respondents aged 16 years and older from the 1991 and 2004 waves of CHNS. They found that, compared with residents of communities in the bottom tercile of an urbanicity index, the risk of being overweight (body mass

index [BMI] ≥25) was significantly higher for residents of communities in the top tercile. Thompson *et al.* (2015), who analyzed the adult sample from the 2009 wave of the CHNS, found that living in a community in the top tercile of an urbanicity index was associated with a greater likelihood of having a high waist-to-height ratio (WHtR) (>0.5) than living in a community in the bottom tercile for men but not for women. The cross-sectional nature of these studies means that the findings are based on between-community differences only.

The most common approach for longitudinal analyses of the urbanization-weight association is to estimate a hierarchical linear model (for continuous measures of body weight) or logistic model (for dichotomous measures of body weight) that adjusts for the multilevel data structure in CHNS, nesting measurement occasions (level 1) within individuals (level 2), who, in turn, are nested within communities (level 3). Applying this approach to the adult sample (ages 18–59) in 1991–2009 CHNS, Jaacks *et al.* (2013) found that a two standard deviation increase in community urbanization was related to a 0.23 unit (kg/m^2) increase in BMI and a 3% increase in the odds of being overweight (BMI ≥25). The same approach has also been applied to examining the diet and physical activity mediators through which urbanization affects body weight status. For example, using 1991–1997 CHNS data, Monda *et al.* (2007) found that a 1-unit change in urbanization score was related to a 7% and a 6% increase in the odds of light or moderate (versus heavy) occupational activity for men and women ages 18–55, respectively. Using the 1991–2011 CHNS data, Ng *et al.* (2014) measured physical activity for adults aged 18 – 60 at work and at home using the metabolic equivalent of task hours per week. They found that higher urbanization scores were associated with lower occupational physical activity for both men and women, and lower domestic physical activity for women, but higher domestic physical activity for men.

However, alternative model specifications have led to inconclusive findings. Focusing on a subset of women who had their BMI measured in both the 1991 and 2004 waves of the CHNS, and who were not overweight or obese in 1991, Jones-Smith and Popkin (2010) estimated a logistic model with robust standard errors and found significantly greater odds of being overweight (BMI ≥25) for women in communities with greater increases in urbanization scores over the study period than for women in communities within the lowest quintile of urbanization in 1991 and that experienced no increase over time. In contrast, applying a difference-in-differences estimator with fixed-effects to the 1991–2004 CHNS data, Van de Poel *et al.* (2012) found no significant association between urbanization (defined as movement from below to above the median of an urbanicity index) and obesity (BMI >30) among adults aged 18 years and older. Gordon-Larsen *et al.* (2014) identified latent class trajectories of adult BMI in the 1991–2011 CHNS data and found that baseline urbanicity was not associated with upward BMI trajectories for men or women. After dividing communities into terciles of urbanicity scores, they found a positive association between the 10-year change in community urbanicity scores and residents' rate of being overweight (BMI ≥25) in the least urbanized communities at the baseline, but a negative association in the most urbanized communities at the baseline.

Despite using these different estimation strategies, the previous research has paid little attention to explicitly modeling the within-community process of urbanization separately from between-community differences in the baseline level or the rate of urbanization. The mixed findings in the literature may be partially explained by different estimation strategies that make inferences about different components of the urbanization process that are related to body weight status in different ways. For example, the hierarchical modeling approach draws inference from pooling together within- and between-community variability over time (Monda *et al.*, 2007; Jaacks *et al.*, 2013; Ng *et al.*, 2014), whereas the difference-in-differences approach and its variant adopted by other researchers (Jones-Smith & Popkin, 2010; Van de Poel *et al.*, 2012; Gordon-Larsen *et al.*, 2014) adjusts for between-community difference in the baseline level of urbanization but leveraged on between-community difference in the rate of urbanization.

The disaggregation analysis in this study will shed new light on the methodological issues (different modeling strategies that may or may not correctly specify different processes of urbanization) that lead to inconsistent findings obtained from the same data source in the literature. More importantly, it will help clarify the mismatch between the conceptual model of within-community process (i.e., *in situ* urbanization) and the statistical model that conflates within- and between-community processes. The disaggregation of between- and within-community components permits a precise and unambiguous test of the longitudinal association between *in situ* urbanization and weight gain.

3. Data and methods

3.1. Sample

Subjects for this study were adults ages 18–65 in the CHNS. Although the CHNS data are not nationally representative, households were randomly selected from a diverse set of nine provinces in northeast, central, and south China. Together, these nine provinces are home to more than

40% of China's population, or 549 million people. Survey communities were drawn through a stratified, multistage random sampling process from cities, suburbs, towns, and villages designated by China's National Bureau of Statistics. In each community, 20 households were randomly selected and all household members were interviewed. The response rate at the individual level is 88%. In addition to individual-level data, the CHNS collects background characteristics of the survey communities. Details on the design and sampling of the CHNS are available elsewhere (Popkin *et al*., 2010).

This study draws on data from nine waves of the survey (1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015) that span more than 20 years. Three municipal cities were added to the sampling frame in 2011 and three more provinces were added in 2015. Respondents from these new sites were excluded, because they might obscure the long-time trends in the other provinces since 1991. A total number of 69,582 (93.5%) out of 74,419 age-eligible person-year observations had valid values of body weight and height between 1991 and 2015 (69,175), or valid values of waist circumference (WC) between 1993 (see the next section) and 2015 (60,760), or both (60,353). Among them, 1818 (2.6%) observations were excluded due to missing data on covariates. The overall analytical sample consisted of 67,764 person-year observations (35,255 females and 32,509 males) contributed by 21,029 individual adults (10,983 female and 10,046 male) from 241 communities. To maximize the statistical power, the final sample size was allowed to vary depending on the number of valid values of each outcome variable.

3.2. Measures

The dependent variable, body weight status, is captured in two ways: (1) BMI, calculated from body weight (in kilograms) and height (in centimeters), which taps general obesity and (2) WC, waist-to-hip ratio (WHpR), and WHtR, all of which tap abdominal obesity. All the anthropometric measures were taken by experienced healthcare workers. While widely used as an indicator for measuring whole-body obesity, BMI is not the best measure for abdominal fat accumulation. In several populations, measures of abdominal obesity, such as WC, WHpR, and WHtR, were found to be superior to BMI for predicting cardiovascular disease risk and for obesity screening (Yusuf *et al*., 2004; Li *et al*., 2006; Knowles *et al*., 2011).

We also examine dichotomous outcomes of body weight status. Overweight status is defined as a BMI ≥24 kg/m^2, using the Chinese Center for Disease Control and Prevention guidelines (2003). Abdominal obesity is defined as a WC ≥90 cm for men and WC ≥85 cm for women (JCDCG, 2007); a WHpR ≥0.9 for men and WHpR ≥0.85 for women (the World Health Organization, 2008); or a WHtR >0.5 (Browning *et al*., 2010). As a robustness check, overweight is also defined as a BMI ≥ 25 kg/m^2, using the World Health Organization guidelines (1998) and abdominal obesity is WC ≥90 cm for men and WC ≥80 cm for women according to the International Diabetes Federation guidelines (IDF, 2006). The CHNS did not collect data on WC until 1993, and thus, all measures of body weight status that involve WC are from solely the 1993–2015 period.

The primary independent variable of interest is a community-level urbanicity index designed specifically for the CHNS data (Jones-Smith & Popkin, 2010). Capitalizing on the rich community-level data in the CHNS, the urbanicity index captures 12 dimensions of urbanization for each community in each wave, including population density, economic activity, traditional markets, modern markets, transportation infrastructure, environmental sanitation, communications, housing conditions, average education level, socioeconomic diversity, health infrastructure, and social services. Each of these dimensions is measured by one or multiple variables and assigned ten possible points, resulting in a maximum value of 120 points summed across the 12 dimensions, with higher values indicating greater urbanization. Detailed information on this index is available elsewhere (Monda *et al*., 2007; Jones-Smith & Popkin, 2010).

To account for biological differences in body weight status and weight gain trajectory, the full sample is divided into female and male subsamples. Other demographic control variables include age (in years), age-squared, and marital status. Socioeconomic backgrounds are captured by respondents' educational attainment (no formal education, primary school, middle school, high school, or college and above) and annual household per capita income (adjusted for inflation and divided into quartiles). Dummy variables indexing provinces and survey waves are included to adjust for spatial and temporal fixed effects, respectively.

3.3. Methods

To assess the robustness of the findings, two methods that vary in their strengths and limitations were employed to disaggregate between- and within-community components of the urbanization-body weight associations. The first, standard method is to rescale the time-varying urbanicity index by community-specific mean centering. Borrowing the notations from Curran and Bauer (2011), let z_{jt} be the urbanicity index score for community j at time t, the mean-centered score, denoted by $\dot{z}_{jt}$, can be calculated as:

$$\dot{z}_{jt} = z_{jt} - \overline{z}_j \quad (1)$$

where $\bar{z}_j$ is the observed sample mean urbanicity index score for community j throughout its entire observation period. Estimates of disaggregated within- and between-community differences can be obtained by regressing the outcome variable on $\dot{z}_{jt}$ and $\bar{z}_j$. Let y_{ijt} be a continuous body weight measure for individual i living in community j at time t, and X_{ijt} be a set of individual- and household-level control variables. A three-level random effects model is fitted to incorporating the hierarchical data structure as the following:

$$y_{ijt} = \beta_0 + \beta X_{ijt} + \gamma_1 \bar{z}_j + \gamma_2 \dot{z}_{jt} + u_j + u_i + \varepsilon_{ijt} \quad (2)$$

where β_0 is an intercept, u_i and u_j represent person- and community-level random intercepts, respectively, and ε_{ijt} denotes residues. The coefficient γ_1 captures the relation between average levels of urbanicity index and average levels of body weight measures pooling over communities. In contrast, the coefficient γ_2 captures the mean relation between a given community's time-specific deviation in urbanization (from its overall level of urbanization) and the time-specific body weight status among the residents living in that community.

The validity of the standard method relies on the assumption that the community-level urbanicity index is unrelated to time (Curran & Bauer, 2011). In other words, it assumes potential growth in the body weight outcome (i.e., y_{ijt}), but no systematic change in urbanization itself, aside from random variations, with the passage of time. This assumption is unlikely to hold because not only has China been one of the fastest urbanizing countries in the world since the 1980s (United Nations, 2015) but also the sampled communities in CHNS have gained considerable growth in their urbanicity index scores between 1991 and 2015 (see the results section).

The second method, referred to as growth curve disaggregation, allows the community-level urbanicity index to change as a function of time and allows the rates of change to vary randomly over communities. These assumptions represent a more realistic scenario, in which the initial level of urbanicity and the pace of urbanization over time differ across communities. A growth curve model for the relationship between the time-varying community-level urbanicity index and time can be specified as the following:

$$\begin{aligned} z_{jt} &= \alpha_0 + \alpha_1 T_{jt} + \delta_{0j} + \delta_{1j} T_{jt} + r_{jt} \\ &= (\alpha_0 + \delta_{0j}) + (\alpha_1 + \delta_{1j}) T_{jt} + r_{jt} \end{aligned} \quad (3)$$

where T_{jt} is the measure of time at time t for community j, which permits the possibility that all the sampled communities are not surveyed at all of the same points in time (i.e., unbalanced longitudinal data with respect to time); α_0 is the mean intercept, indicating the average urbanicity index across all communities and all time points; α_1 is the mean slope, indicating the average rate of change in urbanization over time; δ_{0j} represents the deviation of community j's intercept from the grand mean α_0; δ_{1j} represents the deviation of community j's slope from the mean slope α_1; and the residual term r_{jt} captures the within-community variability of urbanization around the community-specific mean. Between-community differences at any given point in time are determined by both between-community variability in the intercept (δ_{0j}) and between-community variability in the slope (δ_{1j}). Instead of centering at the observed community-specific mean as in Equation (1), each community- and time-specific z_{jt} can be deviated from its model's implied value (Curran & Bauer, 2011):

$$\hat{r}_{jt} = z_{jt} - \left(\hat{\alpha}_0 + \hat{\delta}_{0j}\right) - \left(\hat{\alpha}_1 + \hat{\delta}_{1j}\right) T_{jt} \quad (4)$$

where $\hat{\alpha}_0$, $\hat{\delta}_{0j}$, $\hat{\alpha}_1$, $\hat{\delta}_{1j}$, and $\hat{r}_{jt}$ are the estimates of the coefficients defined in Equation (3). Between-community differences averaged over time can be represented by:

$$\hat{\delta}_{0j} = (\bar{z}_j - \hat{\alpha}_0) - \left(\hat{\alpha}_1 + \hat{\delta}_{1j}\right) \bar{T}_j \quad (5)$$

where $\bar{T}_j$ is the mean value of time for community j and allows unbalanced time structure. The values of $\hat{\delta}_{0j}$ and $\hat{r}_{jt}$ can then be substituted into Equation (2) for $\bar{z}_j$ and $\dot{z}_{jt}$, respectively.

Three multilevel random-intercept models, in which repeated measures (level-1) were nested within individuals (level-2) which, in turn, were nested within communities (level-3), were fitted to each outcome variable. Model 1 assessed the overall association between body weight status and the urbanicity index. Model 2 applied the standard method to rescaling urbanicity index scores and disaggregating the overall association into between- and within-community differences. Model 3 adopted the growth curve disaggregation method instead. Multilevel linear models were estimated for continuous measures of body weight status, while multilevel logistic models were estimated for binary measures of body weight status.

4. Results

4.1. Descriptive statistics of body weight status

The gender-stratified secular trends of the continuous and dichotomous measures of body weight status presented in Figures 1 and 2, indicate that both Chinese men and women have grown heavier over two decades. For example, Figure 1 shows that the average BMI increased

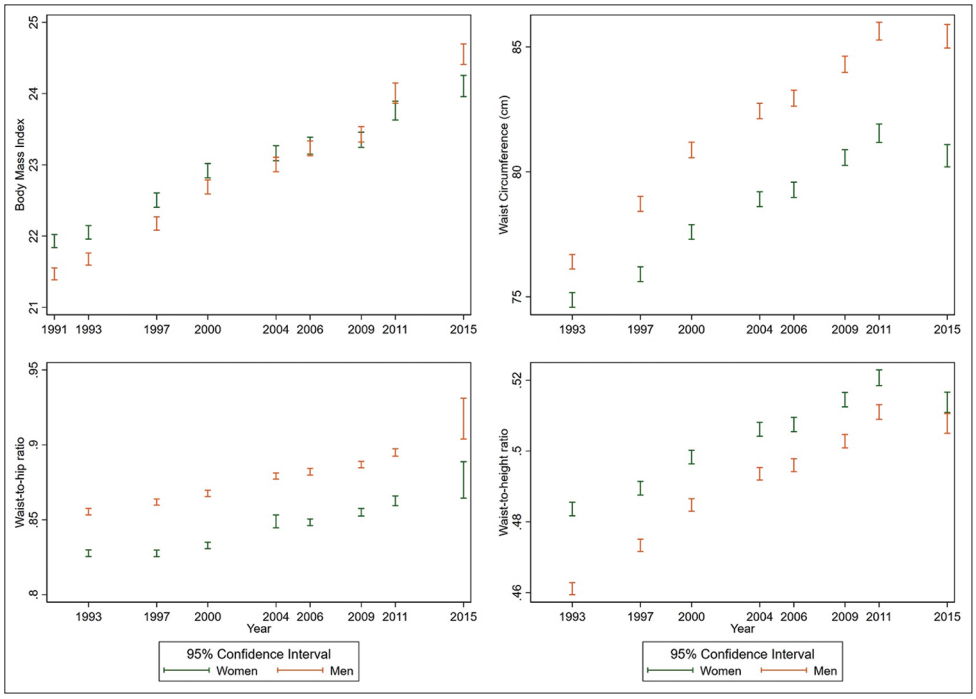


Figure 1. Trends of continuous measures of body weight status among Chinese men and women (18–65 years older).

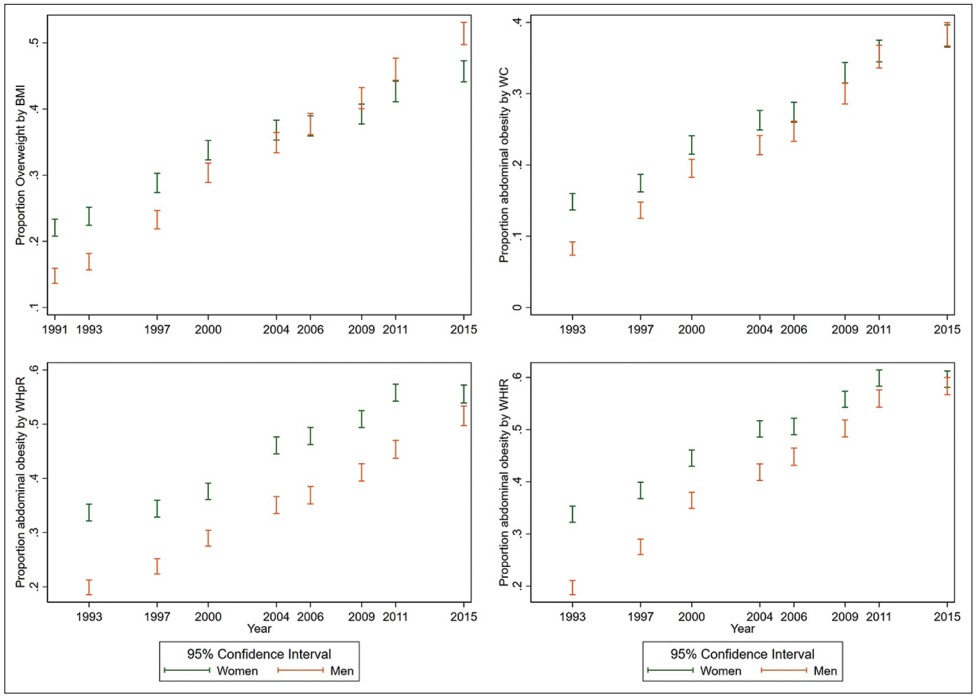


Figure 2. Trends of prevalence of overweight and abdominal obesity among Chinese men and women (18-65 years older). BMI, body mass index; WC, waist circumference; WHpR, waist-to-hip ratio; WHtR, waist-to-height ratio.

from 21.9 in 1991 to 24.1 in 2015 for women and from 21.5 to 24.6 during the same period for men (Table S1 in Supplementary File). For both men and women, the 95% confidence intervals of the sample average BMI in 2015 do not overlap with those in 1991, suggesting statistically significant secular increases. Figure 2 shows that using a cutoff of BMI ≥24 kg/m^2, the sample percentage of overweight women nearly doubled from 22.1% in 1991 to

45.7% in 2015 (Table S2 in Supplementary File). During the same period, the sample percentage of overweight men more than doubled from 14.8% to 51.4%. Weight changes are also seen in the measures related to abdominal obesity. For example, the average WC increased by about 8% from 74.9 cm in 1993 to 80.6 cm in 2015 for women and by about 12% from 76.4 cm to 85.4 cm for men during the same period. Using WC ≥90 cm for men and WC ≥85 cm for women as cutoffs, the sample percentage of abdominal obesity grew by more than 150% from 14.8% in 1993 to 38.1% in 2015 for women and more than tripled from 8.3% to 38.3% for men during the same period.

The secular trend of gender gap in weight gain varies across different measures of body weight status. Among the continuous measures depicted in Figure 1, men had on average significantly lower BMI and WHtR than women in the early 1990s, but the gender gaps narrowed over time. In fact, the gender gap in average BMI lost its statistical significance by 2000 and men's average BMI gradually surpassed that of women thereafter. Men had on average significantly higher WC and WHpR than women in the early 1990s, with the gender gap growing for WC but remaining relatively stable for WHpR in the study period. In contrast, men had generally lower rates of abdominal obesity according to WC and WHpR cutoffs between 1993 and 2015, but the gender gap in WC-based rate of abdominal obesity converged by 2006 (Figure 2). The gendered patterns across measures of weight growth between the continuous and dichotomous measures of WC and WHpR may be attributed to gender-specific cutoffs of abdominal obesity.

4.2. Descriptive statistics of independent variables

Figure 3 shows the trends of average urbanicity index values and the associated 95% confidence intervals, stratified by four community types used in the CHNS sampling stage. Rapid urbanization took place in the sampled communities, with the average urbanicity index score growing by almost 50% from 46.4 in 1991 to 71.5 in 2015 (Table S3 in Supplementary File). The pace of urbanization was fastest in village communities, where the average urbanicity index score rose by nearly 70% from 34.8 to 59 over two decades. Nevertheless, Figure 3 indicates that the average level of urbanization in village communities remained significantly lower than in any other type of community throughout the period of 1991-2015. In addition, the average level of urbanization in suburban communities was significantly lower than that in city communities throughout the same period, and significantly lower than that in town communities during most of the period. These findings suggest that despite a growing urbanicity within each type of community, the pace of urbanization was uneven across different types of communities – sustaining the between-community gaps in the level of urbanization over two decades, which might be attributable to public policies and investments in favor of urban development during China's reform era (Xie & Hannum, 1996). Therefore, the observed growth in population-level body weight described above could have been driven by both within- and between-community differences in urbanization.

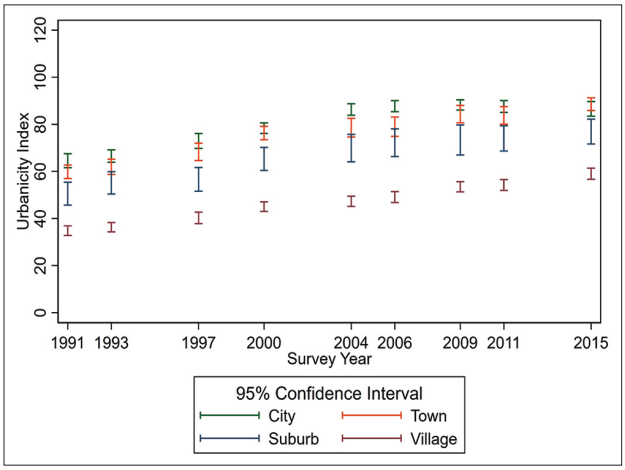


Figure 3. Trend of urbanicity index at the community level, stratified by community types.

Table 1 shows that the average age of the respondents was around 42 years; more than 80% were married; and most had completed elementary or middle school education, with men on average having a higher level of education than women. The person-year observations in the analytical sample were by and large evenly distributed across household income quartiles, provinces, and survey waves.

4.3. Regression results for continuous outcomes

Table 2 reports multilevel regression disaggregation results for the longitudinal associations between urbanization and continuous measures of body weight status. To preserve the space, only the main coefficient estimates of interest were presented. Full regression estimates from selected models can be found in Table S4 in Supplementary File and full results from other models can be requested from the author. In the female subsample, the urbanicity index score was significantly associated with BMI and WC (Model 1), but unrelated to WHpR or WHtR. Both associations were attributed to between-community difference but not within-community difference in urbanicity index, regardless of which disaggregation method was employed (Models 2 and 3). After disaggregation, WHpR was not associated with either between- or within-community difference in urbanicity index; neither was WHtR.

Table 1. Descriptive statistics of the control variables.

Control variables	Women		Men	
Age in years (mean [SD])	42.5	(12.6)	42.1	(13.0)
Marital status (%)				
Never married	9.8		15.1	
Married	84.9		82.0	
Divorced/widowed	5.3		2.9	
Educational attainment (%)				
No school	16.3		4.1	
Elementary school	29.6		25.8	
Middle school	31.9		40.3	
High school	17.6		22.9	
College or above	4.6		7.0	
Per capita household income (%)				
1st quartile	25.2		24.5	
2nd quartile	25.0		24.9	
3rd quartile	25.0		25.0	
4th quartile	24.8		25.6	
Province (%)				
Liaoning	9.7		9.3	
Heilongjiang	8.7		8.8	
Jiangsu	12.0		11.6	
Shandong	11.0		10.7	
Henan	11.6		11.4	
Hubei	11.4		11.6	
Hunan	11.4		11.3	
Guangxi	12.4		13.3	
Guizhou	11.8		12.0	
Wave (%)				
1991	11.0		10.7	
1993	10.3		10.1	
1997	12.0		13.0	
2000	11.0		11.0	
2004	11.4		11.4	
2006	11.2		11.0	
2009	11.3		11.3	
2011	10.7		10.4	
2015	11.1		11.2	
N of person-year observations	32,573		30,880	

The patterns were different for men. Urbanicity index was significantly and positively related to all the four continuous measures of body weight status. For WC, WHpR, and WHtR, the overall associations were completely attributable to between-community difference in urbanization, regardless of the choice of disaggregation method. For BMI, the overall association was driven by both between- and within-community differences, although the latter played a relatively minor role and was only marginally significant at $\alpha = 0.1$ level. For example, according to the standard disaggregation, BMI would increase by 0.024 unit for every one-unit increase in between-community difference in urbanicity index, but only by 0.005 unit for every one-unit increase in within-community difference. The results were almost identical according to the growth curve disaggregation.

4.4. Regression results for overweight and abdominal obesity

Table 3 reports multilevel regression disaggregation results for the longitudinal associations of urbanization with the overweight and abdominal obesity measures. In the female subsample, urbanicity index was positively associated with both overweight- and WC-based abdominal obesity measures, but unrelated to WHpR or WHtR (Model 1). The standard disaggregation showed the association between urbanicity index score and the overweight measure being driven by both between- and within-community differences (Model 2), whereas the growth curve disaggregation suggested that within-community difference did not played any significant role (Model 3). In contrast, the two disaggregation methods consistently showed that the association between urbanicity index score and the WC-based abdominal obesity measure was entirely driven by between-community difference in urbanization.

In the male subsample, urbanicity index score was again significantly and positively related to all the four measures of overweight and abdominal obesity. For abdominal obesity, the two disaggregation methods found that the association was attributed to between- but not within-community difference, regardless of which measure was used. For the overweight measure, between-community difference again played a much stronger role than within-community difference (i.e., log odds = 0.048 vs. 0.011 according to the standard disaggregation, and 0.051 versus 0.01 according to the growth curve disaggregation).

4.5. Sensitivity analysis

Two sensitivity analyses were conducted. First, Table S5 in Supplementary File shows that similar results were obtained when alternative cutoff points were used to classify overweight (BMI ≥25 vs. 24 kg/m^2) and abdominal obesity (WC ≥80 vs. 85 cm in women). For both men and women, urbanicity index was positively related to the risks of being overweight and having abdominal obesity, and these associations were mainly driven by between- rather than within-community difference. Second, Table S6 in Supplementary File reports coefficient estimates from growth curve models of the continuous

Table 2. Three-level random-intercept linear models of longitudinal associations between urbanicity and body weight status in Chinese adults (18–65 years).

Key predictors	BMI	WC	WHpR	WHtR
Female sample	(*n*=35,065)	(*n*=30,709)	(*n*=30,203)	(*n*=30,661)
Model 1: Urbanicity index	0.005*	0.016*	−0.008	0.003
	(0.002)	(0.007)	(0.007)	(0.005)
Model 2: Standard disaggregation				
Between-community component	0.011***	0.025**	−0.008	0.000
	(0.003)	(0.009)	(0.008)	(0.006)
Within-community component	0.003	0.011	−0.008	0.004
	(0.003)	(0.012)	(0.017)	(0.008)
Model 3: Growth curve disaggregation				
Between-community component	0.013***	0.031***	−0.009	0.004
	(0.003)	(0.009)	(0.008)	(0.006)
Within-community component	0.002	0.010	−0.001	0.004
	(0.003)	(0.012)	(0.015)	(0.008)
Male sample	(*n*=32,309)	(*n*=28,324)	(*n*=27,622)	(*n*=28,280)
Model 1: Urbanicity index	0.011***	0.037***	0.020***	0.016
	(0.002)	(0.008)	(0.005)	(0.005)
Model 2: Standard disaggregation				
Between-community component	0.024***	0.091***	0.026***	0.037***
	(0.003)	(0.010)	(0.006)	(0.006)
Within-community component	0.005†	0.012	−0.002	0.004
	(0.003)	(0.012)	(0.011)	(0.007)
Model 3: Growth curve disaggregation				
Between-community component	0.025***	0.102***	0.029***	0.042***
	(0.003)	(0.009)	(0.006)	(0.005)
Within-community component	0.005†	0.007	−0.011	0.002
	(0.003)	(0.012)	(0.012)	(0.007)

BMI, body mass index (kg/m^2); WC, waist circumference (cm); WHpR, waist-to-hip ratio (multiplied by 100); WHtR, waist-to-height ratio (multiplied by 100). Robust standard errors are in parentheses. All the models adjusted for age, marital status, education, household income, provincial fixed effects, and time fixed effects. †P<0.1; *P<0.05; **P<0.01; ***P<0.001.

measures of body weight status. These growth curve models include not only random intercepts that capture intra-person correlation (of repeated measurements) and intra-community correlation (of clustered individuals) but also random coefficients for age that capture heterogeneous age effect. Despite slight changes in certain coefficient estimates and standard errors, the results are qualitatively the same as those reported in Table 2. Unfortunately, similar specification of growth curve models for dichotomous measures of body weight status failed to converge.

5. Discussion

Urbanization is widely viewed as a major contextual force behind the rising prevalence of obesity in developing countries (Hoffman, 2001; Ng *et al.*, 2014). China, a developing country in the midst of rapid urbanization and nutrition transition, is an ideal setting to assess the impact of urbanization on excess weight gain. Higher levels of urbanization are likely to include a shift from occupations requiring strenuous physical activities to those with more sedentary activities, an increase in automotive use for job commuting and daily activities, more affordable food markets for meat and cooking oil, and easier access to Western fast-food restaurants – all factors increasing body weight status. However, the previous research has reported inconsistent findings on the relationship between urbanization at the community level and body weight status at the individual level.

More important, the previous research usually does not make a conceptual distinction between different forms

Table 3. Three-level random-intercept logistic models of longitudinal associations of urbanicity with overweight and abdominal obesity in Chinese adults (18–65 years).

Key predictors	Overweight	Abdominal obesity based on		
		WC	WHpR	WHtR
Female sample	(*n*=35,065)	(*n*=30,709)	(*n*=30,203)	(*n*=30,661)
Model 1: Urbanicity index	0.014***	0.008**	−0.002	0.002
	(0.003)	(0.003)	(0.002)	(0.003)
Model 2: Standard disaggregation				
Between-community component	0.021***	0.012***	−0.004	0.002
	(0.004)	(0.003)	(0.002)	(0.003)
Within-community component	0.010*	0.004	−0.001	0.002
	(0.004)	(0.005)	(0.003)	(0.004)
Model 3: Growth curve disaggregation				
Between-community component	0.025***	0.014***	−0.004	0.004
	(0.004)	(0.003)	(0.002)	(0.003)
Within-community component	0.007	0.002	−0.001	0.001
	(0.005)	(0.005)	(0.004)	(0.004)
Male sample	(*n*=32,309)	(*n*=28,324)	(*n*=27,622)	(*n*=28,280)
Model 1: Urbanicity index	0.029***	0.021***	0.007***	0.015
	(0.003)	(0.003)	(0.002)	(0.003)
Model 2: Standard disaggregation				
Between-community component	0.048***	0.035***	0.012***	0.023***
	(0.003)	(0.003)	(0.002)	(0.003)
Within-community component	0.011*	0.005	0.002	0.007*
	(0.005)	(0.005)	(0.004)	(0.004)
Model 3: Growth curve disaggregation				
Between-community component	***	0.039***	0.013***	0.026***
	(0.004)	(0.003)	(0.002)	(0.003)
Within-community component	0.010†	0.002	0.000	0.004
	(0.005)	(0.005)	(0.004)	(0.004)

Overweight if body mass index≥24 kg/m^2. Abdominal obesity if waist circumference (WC) ≥90 cm in men or≥85 cm in women; waist-to-hip ratio (WHpR) ≥0.9 in men or≥0.85 in women; or waist-to-height ratio (WHtR) >0.5. Robust standard errors are in parentheses. All the models adjusted for age, marital status, education, household income, provincial fixed effects, and time fixed effects. †P<0.1; *P<0.05; **P<0.01; ***P<0.001.

of urbanization (i.e., *in situ* urbanization of community environment versus rural-to-urban migration) and thus empirically conflates preexisting between-community difference and intrinsic within-community change in relation to body weight gain. From the perspective of place effects on health, it is the within-community urban development that has a direct bearing on the conventional hypothesis about the relationship between urbanization and body weight changes. In contrast, between-community difference may encompass gaps in communities' baseline levels and rates of urbanization in relation to weight gain.

With prospective, longitudinal, and multilevel data, this study documented considerable weight gain among Chinese adults with respect to their average body size and fat distribution, as well as remarkable within-community urbanization over two decades. After taking into account individual- and household-level demographic and socioeconomic factors, regression estimates confirmed a positive longitudinal association between community-level urbanicity and individual-level body weight status, with noteworthy gender differences. For Chinese men, the positive weight gain-urbanization association holds irrespective of body weight measure (continuous or dichotomous, general overweight or abdominal obesity). For Chinese women, the statistical significance is sensitive to the choice of body weight measure.

Through disaggregation analysis, the overall longitudinal association between community-level urbanicity and

individuals' body weight was found to be largely driven by between-community difference. This finding holds for both men and women and is robust against the choice of disaggregation method or body weight measure. In other words, contrary to the conventional theory, little evidence supports that an increase in the level of urbanization within a community was related to weight gain among its residents. The observed positive association between urbanization and body weight status was heavily inferred from disparities between individuals living in different communities that differed in their baseline and trajectory of urbanization.

The lack of association between within-community urbanization and weight gain might reflect insufficient statistical power. For example, as evident in Figure 3 and in Supplementary File Table S3, the cross-sectional between-community variation in the urbanicity index score at any time point was considerably larger than the longitudinal within-community variation over any two consecutive waves of the CHNS. As a result, the CHNS data provided greater statistical power for researchers to detect potential between-community effect than within-community effect. In essence, this is similar to the drawback of fixed-effects models being less efficient than random-effects models of longitudinal data.

To be comparable with previous research, the present study did not disentangle the multiple dimensions of urbanization. The urbanicity index provides a single composite measure of 12 distinct aspects of urbanization and, hence, does not allow researchers to identify the heterogeneity in the process of urbanization that may have diverse implications for people's weight gain. For example, when a village develops into a town, not only would its transportation infrastructure and food environment change substantially but also more importantly, new urban social and economic structures would be established in place of the old rural ones. Agricultural work would be replaced by manufacture or service industry, and a physically active lifestyle might be replaced by a sedentary lifestyle, making villagers more susceptible to changes in body weight status than, say, city dwellers whose lifestyles would change less drastically as their urban community experiences such gradual changes as the opening of a new supermarket or subway station. Therefore, future research should develop and test theories that identify specific aspects of urbanization related to weight gain. Furthermore, the present study does not examine intermediate diet and physical activity outcomes that lie in the pathway to weight gain. However, the CHNS has made some changes in questionnaire design related to self-reported measures of diet and physical activities, making longitudinal analysis difficult. Future research using the CHNS data may specify a shorter study period during which consistent measures are available to decompose the associations between urbanization and these intermediate outcomes.

Despite these limitations, findings from this study highlight complex patterns of body weight changes in relation to urbanization as the Chinese society transitions from an old era of poverty and under-nutrition to a new one of affluence and over-nutrition. The findings also challenge the adequacy of the simple conventional model of community effects on health (Averett & Korenman, 1999; Entwisle, 2007). Understanding risks for obesity in adults depends not only on whether they live in urbanized communities but also on how obesogenic environments evolve as a result of human activities. In terms of policy implications, the findings suggest that limited public health resources to address the rising prevalence of obesity should not be evenly distributed but targeted at city and town communities where the level of urbanization continues to be high and migrants from rural areas are about to adopt new lifestyles.

Acknowledgments

None.

Funding

This study was supported by the Lieberthal Rogel Center for Chinese Studies at the University of Michigan under a faculty research grant and City University of New York – Queens College under a Cycle 50 PSC-CUNY Research Award (TRADA-50-116). The CHNS data collection has been funded by the National Institute for Health (NIH), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) for R01 HD30880, National Institute on Aging (NIA) for R01 AG065357, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) for R01DK104371 and R01HL108427, and the NIH Fogarty grant D43 TW009077 since 1989.

Conflict of interest

No conflicts of interest were reported by the author.

Author contributions

This is a sole-authored study.

Ethics approval and consent to participate

This study does not involve human subjects as defined by CUNY Human Research Protection Program (HRPP) because it only involves analysis of publicly available, secondary survey data. Therefore, this study does not require CUNY HRPP or IRB review.

Consent for publication

Not applicable.

Availability of data

The data used in this study are drawn from the CHNS and publicly available at the CHNS website: https://www.cpc.unc.edu/projects/china.

References

Averett, S. & Korenman, S. (1999). Black-white differences in social and economic consequences of obesity. *International Journal of Obesity*, 23(2), 166-173.

Browning, L. M., Hsieh, S. D., & Ashwell, M. (2010). A systematic review of waist-to-height ratio as a screening tool for the prediction of cardiovascular disease and diabetes: 0·5 Could be a suitable global boundary value. *Nutrition Research Reviews*, 23(2), 247-269.

https://doi.org/10.1017/S0954422410000144

Chinese Center for Disease Control and Prevention. (2003). *Chinese Guidelines on Prevention and Treatment of Overweight and Obesity in Adults*. Beijing, Chinese: Chinese Center for Disease Control and Prevention.

Chinese Center for Disease Control and Prevention. (2015). *Report of Nutrition and Chronic Diseases in Chinese Population-2015*. Beijing: People Health Publishing House.

Curran, P. J. & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62(1), 583-619.

https://doi.org/10.1146/annurev.psych.093008.100356.

Roux, A. V. (2004). Estimating neighborhood health effects: The challenges of causal inference in a complex world. *Social Science and Medicine*, 58(10), 1953-1960.

https://doi.org/10.1016/S0277-9536(03)00414-3

Entwisle, B. (2007). Putting people into place. *Demography*, 44(4), 687-703.

https://doi.org/10.1353/dem.2007.0045

Gordon-Larsen, P., Wang, H., & Popkin, B. M. (2014). Overweight dynamics in Chinese children and adults. *Obesity Reviews*, 15(Suppl 1), 37-48.

https://doi.org/10.1111/obr.12121

Hoffman, D. J. (2001). Obesity in developing countries: Causes and implications. *Food Nutrition and Agriculture*, 28, 35-44.

International Diabetes Federation. (2006). *IDF Consensus Worldwide Definition of the Metabolic Syndrome*. Brussels: International Diabetes Federation.

Jaacks, L. M., Gordon-Larsen, P., Mayer-Davis, E. J., Adair, L. S., & Popkin, B. (2013). Age, period and cohort effects on adult body mass index and overweight from 1991 to 2009 in China: The China health and nutrition survey. *International Journal of Epidemiology*, 42(3), 828-837.

https://doi.org/10.1093/ije/dyt052

Joint Committee for Developing Chinese Guidelines on Prevention and Treatment of Dyslipidemia in Adults (2007). Chinese guidelines on prevention and treatment of dyslipidemia in adults. *Chinese Journal of Cardiology*, 35(5), 390-419.

Jones-Smith, J. C. & Popkin, B. M. (2010). Understanding community context and adult health changes in China: Development of an urbanicity scale. *Social Science and Medicine*, 71(8), 1436-1446.

https://doi.org/10.1016/j.socscimed.2010.07.027

Knowles, K. M., Paiva, L. L., Sanche, S. E., Revilla, L., Lopez, T., Yasuda, M. B., *et al.* (2011). Waist circumference, body mass index, and other measures of adiposity in predicting cardiovascular disease risk factors among peruvian adults. *International Journal of Hypertension*, 2011, 931402.

https://doi.org/10.4061/2011/931402

Li, C., Ford, E. S., Mokdad, A. H., & Cook, S. (2006). Recent trends in waist circumference and waist-height ratio among US children and adolescents. *Pediatrics*, 118(5), e1390-e1398.

https://doi.org/10.1542/peds.2006-1062

Mehta, N. K. & Chang, V. W. (2008). Weight status and restaurant availability: A multilevel analysis. *American Journal of Preventive Medicine*, 34(2), 127-133.

https://doi.org/10.1016/j.amepre.2007.09.031

Monda, K. L., Adair, L. S., Zhai, F., & Popkin, B. M. (2008). Longitudinal relationships between occupational and domestic physical activity patterns and body weight in China. *European Journal of Clinical Nutrition*, 62(11), 1318-1325.

https://doi.org/10.1038/sj.ejcn.1602849

Monda, K. L., Gordon-Larsen, P., Stevens, J., & Popkin, B. M. (2007). China's transition: The effect of rapid urbanization on adult occupational physical activity. *Social Science and Medicine*, 64(4), 858-870.

https://doi.org/10.1016/j.socscimed.2006.10.019

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., *et al.* (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: A systematic analysis for the global burden of disease study 2013. *The Lancet*, 384(9945), 766-781.

https://doi.org/10.1016/S0140-6736(14)60460-8

Ng, S. W., Howard, A. G., Wang, H. G., Su, C., & Zhang, B. (2014). The physical activity transition among adults in China: 1991-2011. *Obesity Reviews*, 15(1), 27-36.

https://doi.org/10.1111/obr.12127

Popkin, B. M., Du, S., Zhai, F., & Zhang, B. (2010). Cohort profile: The China health and nutrition survey-monitoring and understanding socio-economic and health change in China, 1989-2011. *International Journal of Epidemiology*, 39(6), 1435-1440.

https://doi.org/10.1093/ije/dyp322

Swinburn, B., Egger, G., & Raza, F. (1999). Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine*, 29(6), 563-570.

https://doi.org/10.1006/pmed.1999.0585

Thompson, A. L., Adair, L., Zhang, B., & Popkin, B. (2015). Environmental, dietary, and behavioral factors distinguish Chinese adults with high waist-to-height ratio with and without inflammation. *The Journal of Nutrition*, 145(6), 1335-1344.

https://doi.org/10.3945/jn.114.206102

United Nations, Population Division. (2015). *World Urbanization Prospects: The 2014 Revision*, (*ST/ESA/SER.A/366*). New York: United Nations Population Division.

Van de Poel, E., O'Donnell, O., & Van Doorslaer, E. (2012). Is there a health penalty of China's rapid urbanization? *Health Economics*, 21(4), 367-385.

https://doi.org/10.1002/hec.1717

Van de Poel, E., O'Donnell, O., & Van Doorslaer, E. (2009). Urbanization and the spread of diseases of affluence in China. *Economics and Human Biology*, 7(2), 200-216.

https://doi.org/10.1016/j.ehb.2009.05.004

Wang, L. (2005). *Summary Report of China's 2002 National Nutrition and Health Survey*. Beijing: People's Medical Publishing House.

World Health Organization. (1998). *Obesity: Preventing and Managing the Global Epidemic*. Report of a WHO Consultation (WHO Technical Report Series 894). Geneva: World Health Organization WHO.

World Health Organization. (2008). *Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation*. Geneva: World Health Organization.

Xie, Y. & Hannum, E. (1996). Regional variation in earning inequality in reform-era urban China. *The American Journal of Sociology*, 101(4), 950-992.

Xu, H., Short, S. E., & Liu, T. (2013). Dynamic relations between fast-food restaurant and body weight status: A longitudinal and multilevel analysis of Chinese adults. *Journal of Epidemiology and Community Health*, 67(3), 271-279.

https://doi.org/10.1136/jech-2012-201157

Yeh, A. G., Xu, J., & Liu, K. (2011). *China's Post-Reform Urbanization: Retrospect, Policies and Trends*. New York: United Nations Population Fund (UNFPA) and the International Institute for Environment and Development (IIED).

Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F., *et al.* (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *Lancet*, 364(9438), 937-952.

https://doi.org/10.1016/S0140-6736(04)17018-9

Zhu, Y. (2000). *In situ* urbanization in rural China: Case studies from Fujian Province. *Development and Change*, 31(2), 413-434.

https://doi.org/10.1111/1467-7660.00160

Zhu, Y. (2002). Beyond large-city-centred urbanisation: *In situ* transformationof rural areas in Fujian Province. *Asia Pacific Viewpoint*, 43(1), 9-22.

https://doi.org/10.1111/1467-8373.00155

ACCSCIENCE PUBLISHING

International Journal of Population Studies

RESEARCH ARTICLE

Fertility by parity in China in the context of changing fertility policy

Yuanyuan Duan[1], and Wei Chen[2]*

[1]College of Humanities and Development Studies, China Agricultural University, Beijing, China,100193

[2]Center for Population and Development Studies, Renmin University of China, Beijing, China, 100872

Abstract

The study aims to investigate the dynamics of fertility by parity of Chinese women over the past seven decades under the context of changing fertility policy. Using data from population censuses, population sample surveys, and fertility surveys in China, the study estimates China's fertility by parity from 1949 to 2020 by adopting multiple fertility measures, including parity-specific total fertility rate, parity progression ratio, parity-progression-ratio-based total fertility rate, and cumulated cohort fertility rate, as well as the decomposition method. The study further evaluates the unique features of China's configuration of parity-specific fertility through an international comparative analysis of some Western countries based on data from the Human Fertility Database. It shows that in China, both the rigid fertility policy of restricting the number and timing/spacing of children implemented since the early 1980s and the recent relaxation of fertility policy of gradually easing the number and timing/spacing of children have had a significant impact on fertility patterns and levels, especially for parity two. However, the effect of fertility policy relaxation in a low-fertility context has been less sustainable than the earlier rigid fertility policy that contributed to the rapid decline in fertility for second and higher orders of parity. Under the joint influence of the Confucian fertility culture, rapid socioeconomic growth, and the internalization of long-standing strict fertility policies, China has formed a unique pattern of parity-specific fertility profile compared to those of some developed societies, with a universal progression to the first birth, a very low but policy sensitive progression to the second birth, and an extremely low progression to the third birth.

Keywords: Total fertility rate; Parity progression ratio; Fertility by parity; Fertility policy; China

***Corresponding author:**
Wei, Chen. (weichen@ruc.edu.cn)

Citation: Duan, Y., & Chen, W. (2022). Fertility by parity in China in the context of changing fertility policy. *International Journal of Population Studies*, 8(1):83-101. https://doi.org/10.36922/ijps.v8i1.348

Received: July 4, 2022

Accepted: September 19, 2022

Published Online: October 21, 2022

Copyright: © 2022 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Parity analysis is essential for a deeper understanding of fertility dynamics. The reduction of high-order births has been a main determinant of the initial decline toward low fertility, while the change of pace of postponement at birth order one and two has emerged as a crucial driver of recent fertility change among low-fertility settings (Bongaarts & Sobotka, 2012; Kohler, *et al.*, 2002). A rapid shift to delayed childbearing and a low probability of progression of a higher order after the first birth were regarded as the important features of lowest-low fertility in the late 1990s in European countries (Billari & Kohler, 2004). In

line with this, the increases of the first and second births resulted by a decline in the pace of fertility postponement have contributed substantially to the increases in the overall period total fertility rate (TFR) between 1998 and 2008 in European countries (Bongaarts & Sobotka, 2011).

Underlying the dynamics of period fertility, there has been a well-documented long-term decline in cohort fertility by parity across developed countries. In these countries, the reduction in the progression ratios to the third and higher orders was the main driver of fertility decline among women born between 1940 and 1955; the falling first or second birth played the major role in changes in fertility among women born between 1950 and 1970 (Zeman, *et al.*, 2018). Nevertheless, parity distribution of fertility in different countries may vary considerably even when they have similar levels of cohort fertility (Zeman, *et al.*, 2018).

As the most populous country, fertility in China is of global significance. Fertility in China has undergone significant changes since the foundation of People's Republic of China in 1949, which has been well-documented to be closely associated with its fertility policies. The extraordinary fertility transition in China since the 1970s is widely considered to be the product of a strict birth control policy interacting with socioeconomic development. However, there has been much confusion on the extent to which that fertility policy has been effective to lower the fertility because of the complexity associated with data quality and measurement issues since the 1990s. The recent adjustments of fertility policy after 2015 designed to lessen restrictions on childbearing have received wide attention, yet there have been doubts casted on how much the new policy could boost fertility. The control of parity and timing/spacing has been the key content of China's fertility policy and varied with the evolution of fertility policy, making the fertility parity analysis in China especially pertinent. The analysis of period/cohort level and patterns of fertility since 1949 in China from the parity perspective would help to illustrate the effects of the changing fertility policy on fertility during the period.

1.1. Evolution of fertility policy and its control of parity and timing/spacing

The evolution of China's fertility policy on parity and timing/spacing could be divided into four phases. In Phase 1, from 1949 to the early 1970s, there was neither control of number of children nor timing/spacing of births. Although China piloted its first official family planning program in the early 1950s, the primary focus of the policy was to promote the ideas of birth control and small families without explicit regulation on control of number of children and timing/spacing of births (Jiang & Liu, 2016; Lu & Zhai, 2009; Zhang & Cao, 2006). The policy was interrupted firstly by the Great Famine during the period from 1959 to 1961 and then by the early years of Cultural Revolution in the late 1960s.

In Phase 2, during the 1970s, there was control of the timing/spacing of birth orders one and two as well as the restrictions on third and higher birth orders. In the early 1970s, China resumed its nationwide fertility policy with the campaign slogan "later, longer, fewer" (wan xi shao), referring to later marriage, longer birth interval, and fewer number of children. By a later marriage, the minimum age of later marriage was set at 23 years old for women and 25 years old for men in rural areas and 25 for women and 28 for men in urban areas. By a longer birth interval, the minimum interval between the first birth and second birth was set at 3 years in 1978. By fewer children, the maximum number of children was initially three children for rural couples and two for urban couples, but two children for all couples since 1977 (Bongaarts & Greenhalgh, 1985).

In Phase 3, from 1980 to the early 2010s, the policy-allowed number of children was one child (or two or three in certain cases) and the stipulations of timing/spacing of births were also applied. The rigid one-child policy, under which every couple was allowed to have only one child, except very special circumstances, was implemented nationally in 1980. In the early stages of the implementation of the one-child policy, there were only a few circumstances in which a second child was allowed, mostly for couples in rural areas. The proportion of couples who were allowed to have a second child of the total number of couples was no more than 10% (Lu & Zhai, 2009). This rigid one-child policy was revised first in 1984 and then in 1986 by allowing more couples to have a second child but limiting third and higher-order births as well as unauthorized second births due to the strong resistance from residents in rural areas with widespread traditional values of large family and strong son preferences (Greenhalgh, 1986). From 1984 on, some rural couples were allowed to have a second child under certain conditions. The policy was further modified in 1988 to a "1.5-child policy": Couples in major rural areas were allowed to have a second child after a certain interval after their first child if that first child was a girl. This adjusted version of the one-child policy lasted from the mid-to-late 1980s until the end of 2013 (Zhao, 2015). If all couples follow fully the fertility policy, there would be 63% of them ending up with only one child in the late 1990s (Gu, *et al.*, 2007). However, the actual proportion of women having only one child estimated from the 2017 China Fertility Survey data was much lower. The cohort cumulative fertility rate was 1.6 – 1.7 for women aged 35–44 in 2017 who entered childbearing age in the late 1990s and early 2000s (Chen & Duan, 2019), with only 41% of them ending up with one child. Exemptions to the one-

child rule often came with a timing/spacing requirement, stipulating a minimum of age at second births or/and 4 or 6 years between the first and second birth. While from the beginning of the 21st century, restrictions on the timing/spacing have been gradually lifted. By the end of 2013, 18 provinces (autonomous regions and municipalities) had abolished controls on the interval between births (Zhang, *et al*., 2016).

In Phase 4, from 2014 to the present, both the control of parity two and the birth interval restrictions were gradually removed. At the end of 2013, the Chinese government implemented selective two-child policy, allowing couples where either wife or husband is a singleton to have the second child (Basten & Jiang, 2014). In 2015, the fertility policy was further adjusted to universal two-child policy, allowing all couples to have two children (Zeng & Hesketh, 2016). The Chinese government is now encouraging the birth of children with its new "three-child" policy accompanied by supportive measures, launched in 2021. The 2021 amendment provides general principles; the details of how these new family planning measures are to be implemented are left, as with the previous ones, to the discretion of respective provincial governments (Attané, 2022).

1.2. Debate on effects of fertility policy in China

Debates about the effects of fertility policy on China's fertility have come with the evaluation of fertility policy and have been ongoing for many years. As one of the most controversial policies in history, debate has raged over the positive and negative effects of the one-child policy. The importance of the one-child policy for fertility transition is controversial (Zhang & Cao, 2006). China's family planning policy has been claimed to have been effective, based on the evidence of the various estimates at different times of hundreds of millions of averted births and the decline in the fertility rate (Goodkind, 1992, 2017; Yang, *et al*., 2000; Zhao, 1991). However, these assertations have been contested with claims that the "later, longer, fewer" policy of 1973 played a critical role in driving down the fertility rate, and that the role of the one-child policy and its descendants from 1980 on was much less significant (Gietel-Basten, *et al*., 2019), and others argue in particular that the importance of the family planning program on fertility transition should not be overstated, especially from the 1990s (Zhao & Chen, 2011).

Since the policy relaxation with the selective two-child policy in 2013 and the universal two-child policy in 2015, a debate has been under way on the need to review the effectiveness of the relaxation of fertility policy. Under the two-child policy, the focus of academic and policy debate has been on whether the adjustments to fertility policy have significantly affected fertility, and how much China's fertility will increase under the new policy. After the implementation of selective two-child policy in 2013, some researchers argued that the policy was ineffective based on the fact that far fewer couples than expected have applied for permission to have a second child. Other researchers argued that the effect of the policy has been largely in line with the expectation based on the substantial increase in births after the implementation of selective two-child policy. They suggested that the reason for the low proportion of eligible couples who submit the application to have a second child was that the decision to have a second child takes time and the application did not happen in a short time (Zhai, *et al*., 2021). The implementation of universal two-child policy led to a rapid shift in scholarly attention to the new policy. The increases in the newborn and fertility rate in 2016, particularly for the second births, suggest that there has been a positive effect of the universal two-child policy in the short term (Shi, *et al*., 2018; Song, 2017; Yuan & Gao, 2017). While they may be based on different population projections resulting from the use of different data sources and assumptions, studies suggest that the universal two-child policy will affect fertility rates and thus slow population aging, but with only a moderate and temporary effect (Wang & Ge, 2016; Zeng & Hesketh, 2016; Zhang, *et al*., 2019). However, after the accumulative effect runs out, the number of births and fertility has fallen sharply since 2018. The downward trend in number of births and fertility was not effectively reversed even by the implementation of the three-child policy in 2021. Recent studies have consistently found that a large proportion of couples do not want to have a second child even they are allowed to (Feng, 2018; Jin, *et al*., 2016). The latest survey data from the National Bureau of Statistics show that China's fertility intention is only 1.8, which is significantly lower than the replacement level and lower than that of developed low-fertility countries (Yu, *et al*., 2021).

1.3. Current study

Fertility trends in China, and in particular the trends by different birth orders, are closely related to the fertility policy. In this paper, we, therefore, investigate the interrelation between period/cohort level as well as patterns of fertility since 1949 and fertility policy from the parity perspective.

Despite a large body of studies assessing the levels of and trends in overall fertility, there has been limited research on fertility by parity, especially after the 1980s. Based on reliable data sources and different fertility measures, the research on China's fertility levels in the 1970s and 1980s has provided valuable information for our understanding of fertility and the policy effects in these two decades. However, different beliefs about the credibility/quality of population survey data since the 1990s have led to a wide

divergence in scholars' understanding of fertility levels and trends in China. Estimations of TFR directly obtained from census or sample survey data are lower than those obtained based on adjusted census data or other data sources. Some scholars suggested that the underreporting of births is greater for the second and higher orders because births do not conform to the policy restriction, such as births with insufficient intervals or even over-born, are more likely to be underreported (Wang, 2003; Zhang & Su, 1995). Accordingly, the existing literature on fertility by parity has focused on a specific period, particularly the early years.

The period TFR has been the most common measure of the fertility level due to its simplicity and wide availability (Bongaarts & Feeney, 1998; Ma, *et al*., 1986a; McDonald & Kippen, 2007). However, the period TFR is hard to reveal the accurate effects of fertility policy. The period TFR is not able to distinguish a change in timing (or tempo) of cohort fertility from a change in the level (or quantum) of cohort fertility (Schoen, 2004). The period TFR is regarded as biased or distorted by tempo effects. That is, estimates of TFR are depressed during years when women delay childbearing and inflated in years when childbearing is accelerated (Bongaarts & Feeney, 1998). More specifically, under the interference of various period factors, the estimates of parity-specific TFR sometimes exceed one child per woman (e.g., Whelpton, 1945, 1954; Yao, 1995). For instance, the TFR for the first births was >1 in many Western countries in the 1940s and in China in the early 1980s, implying that women on average have more than one first birth, which is not interpretable (Ma, *et al*., 1986a; Rallu & Toulemon, 1994).

This study aims to evaluate China's fertility by parity since 1949 using multiple sources of data and by adopting multiple fertility measures, including period TFR by parity, period parity progression ratios (PPPRs), period parity-progression-ratio-based total fertility rate (PPTFR) by parity, cohort parity progression ratio (CPPR), and cohort cumulative fertility rate (CCFR) by parity as well as its decomposition. This enables us not only to assess the fertility dynamics by parity over the past 70 years but also to identify the possible effects of the changing fertility policy during the period. Performing this analysis is especially interesting given the recent adjustment in the family planning policy. The results are expected to provide important insights into fertility by parity in China by further comparing the parity distribution in China with other low- fertility settings.

2. Data and methods

2.1. Data sources

Multiple sources of data are used in this research. The fertility data of China mainly come from national population censuses performed in 1982, 1990, 2000, 2010, and 2020; the 1% national intercensal population sample surveys conducted in 1987, 1995, 2005, and 2015; national one- per-thousand annual sample surveys of population change from 1982 to 2020; and national retrospective fertility sample surveys conducted in 1982, 1988, 1992, 1997, 2001, and 2017. These population censuses and population sample surveys collected detailed information on age-parity-specific fertility rates, age at first marriage, and age when children were born for Chinese women. These data allow us to establish the historical fertility levels and patterns by parity. The availability of the Human Fertility Database allows us to compare fertility by parity in China with that in other low-fertility societies.

Considered the widely established fact of varying degrees of birth under-reporting in censuses and surveys since the early 1990s, this paper made an effort to adjust the estimates of period fertility rate from 2000 to 2020 based on the number of births published by China's National Bureau of Statistic (NBS). First, it estimated the age-specific number of women aged 15–49 from 2000 to 2020 based on the 2010 census. Second, with the assumption that the age-specific fertility schedules are accurate in the census/population sample survey data, it decomposed the births published by NBS into the different ages of women of childbearing age. It is worth noting that the NBS updated the number of births from 2011 to 2020 based on the 2020 census. Third, it obtained the adjusted estimates of age- specific fertility rate based on census/sample survey data. The adjusted estimates based on the 2017 China Fertility Survey followed the same process. Due to the possible significant overestimate of NBS-published births in the 1990s, we are not able to adjust the estimates of period fertility rate in the 1990s applying the same strategy (Chen, 2016; Zhao & Chen, 2011).

2.2. Methods

The main aim of the study is to examine the aggregate level of fertility by parity in China. Besides the most commonly available measure, the conventional TFR by parity, the study includes multiple measures of fertility: PPPR, PPTFR, CPPR, and CCFR. PPTFR provides more stable and consistent estimates than conventional TFR because they control not only for age but also for parity. CCFR provides a straightforward measure of the fertility of real groups of women and has been championed as the most appropriate measure to analyze fertility (Ryder, 1986).

The definitions and calculations of TFR and CCFR are well known, while those of PPTFR are more complex. The definitions of PPTFR can be understood by the way in which their calculation improves compared with TFR, whose calculation controls for women's age only. The construction of PPTFR further takes parity into

consideration, showing the total fertility rate of a population that is calculated based on age-parity-specific fertility rates (Rallu and Toulemon, 1994). The calculation of PPTFR requires the use of parity progression ratios and life tables. Supplementary information 1 presents a full description of the calculation process and formula for PPTFR. Detailed calculation methods can also be found in relevant literature (Feeney, 1983; Feeney & Yu, 1987; Ma, *et al.*, 1986b; Rallu & Toulemon, 1994).

Guided by the previous study on decomposition of cohort fertility in low-fertility countries (Zeman, *et al.*, 2018), this study also explores how parity progression ratios varied across cohorts and the contribution of changing parity progression ratios to the overall change in completed cohort fertility in China. The specific decomposition process is presented in supplementary information 2.

3. Results

3.1. Period TFR by parity and ages at the first and second births

Figure 1 plots the TFR and parity-specific TFR for Chinese women from 1949 to 2020 based on all datasets we used for this research. China has experienced a dramatic fertility transition between the 1950s and 1990s. Fertility in China fell to replacement level in the early 1990s and has remained low for almost three decades since then. The fertility transition in China can be roughly divided into four stages (Chen, 2003; Lu & Zhai, 2009): Pre-transitional fertility during the 1950s and 1960s; rapid fertility decline or first fertility transition in the 1970s; buffered fertility in the 1980s; and continued low-fertility or second fertility transition from the 1990s.

Fertility by parity shows different characteristics in different stages of China's fertility transition. In the first stage, TFR stabilized at a high level of about six in the 1960s due to the socioeconomic development and people's living being restored from a protracted war since the founding of the People's Republic of China. TFR in China fell sharply during the Great Famine from 1959 to 1961, with the lowest level of fertility of 3.29 children per woman during this period. The TFR recovered to 6.02 in 1962 and 7.46 in 1963 soon after the Great Famine and then fluctuated at around six children per woman in the late 1960s. In the 1950s and 1960s, the fertility for parities one and two fluctuated at around one, while fertility for parity three and above fluctuated at around four, except the period of Great Famine. The changes of fertility and fertility by parity were mainly resulted from the structure of childbearing age of women. The traditional fertility attitudes and low level of social development were the main forces to maintain the high level of fertility at the first stage.

In the second phase, the TFR in China has declined continuously and rapidly in the 1970s, plummeting from 5.71 children per woman in 1970 to 2.73 in 1979 due to the nationwide "later, longer, fewer" fertility policy launched in the early 1970s. By parity, the decline of fertility of third birth and above was substantial, falling steeply from 3.90 in 1970 to 1.18 in 1980. While the decline of fertility for parity one and parity two was modest during the same period, with the first-birth fertility rate declining from 0.96 to 0.67 in 1975, then increasing to 0.86 in 1979, and with the second-birth fertility rate declining from 0.90 in 1970 to 0.61 in 1977, then slightly rising to 0.68 in 1979. Due to the fertility policy during this period, there was a significant monotonic decline of fertility for parity three and above and a corresponding delay in the births for parities one and two. Accordingly, the decline in the second-birth fertility rate was related to the delay of the age at second birth, rather than a real decline of the level of second-birth fertility.

In the third phase, the TFR hovered between 2.20 and 2.86 in the 1980s, with two significant "rebounds" in 1982 and 1986, respectively. The second-birth fertility reversed from a downward trend in the 1970s to an upward trend in the 1980s. Affected by the one-child policy, the fertility for the second birth fell to its lowest level ever since the founding of the People's Republic of China, with a level of 0.59 in 1980, but it then fluctuated upward, reaching a peak of 0.86 in 1987 during this phase. Over the same period, the fertility rate for the first birth continued to be > 1 throughout the 1980s, with a peak level of 1.3 in 1982. The fertility rate for the third child and above showed a fluctuating downward trend, with a level of around 0.81 in 1981 and 1982, and then declining to 0.57 in 1987.

The fluctuant and even the rebound TFR was related to the reverse of the postponement trend in the ages at marriage and childbearing of young cohorts as well as the "compensation" of delayed childbearing of the old cohorts in the 1970s. This superimposed effect of childbearing for early and older cohorts offsets the decline in cohort fertility to some extent. The fertility rate for parity one continued to be > 1, which is a reflection of this superimposed effect. There were two main reasons for the reverse of postponement of women's marriage and childbearing in the 1980s. First, the new marriage law launched in 1980 set the legal age of marriage for women at 20 years old, which was 3 years lower than that regulated by the fertility policy in the 1970s, leading to an advance of marriage for women (Zha & Ji, 1984). Second, the household contract responsibility system in the context of economic reform, which was widely implemented in rural areas during this period, contributed to the tendency of rural residents to get married and to give births early. This was due to the

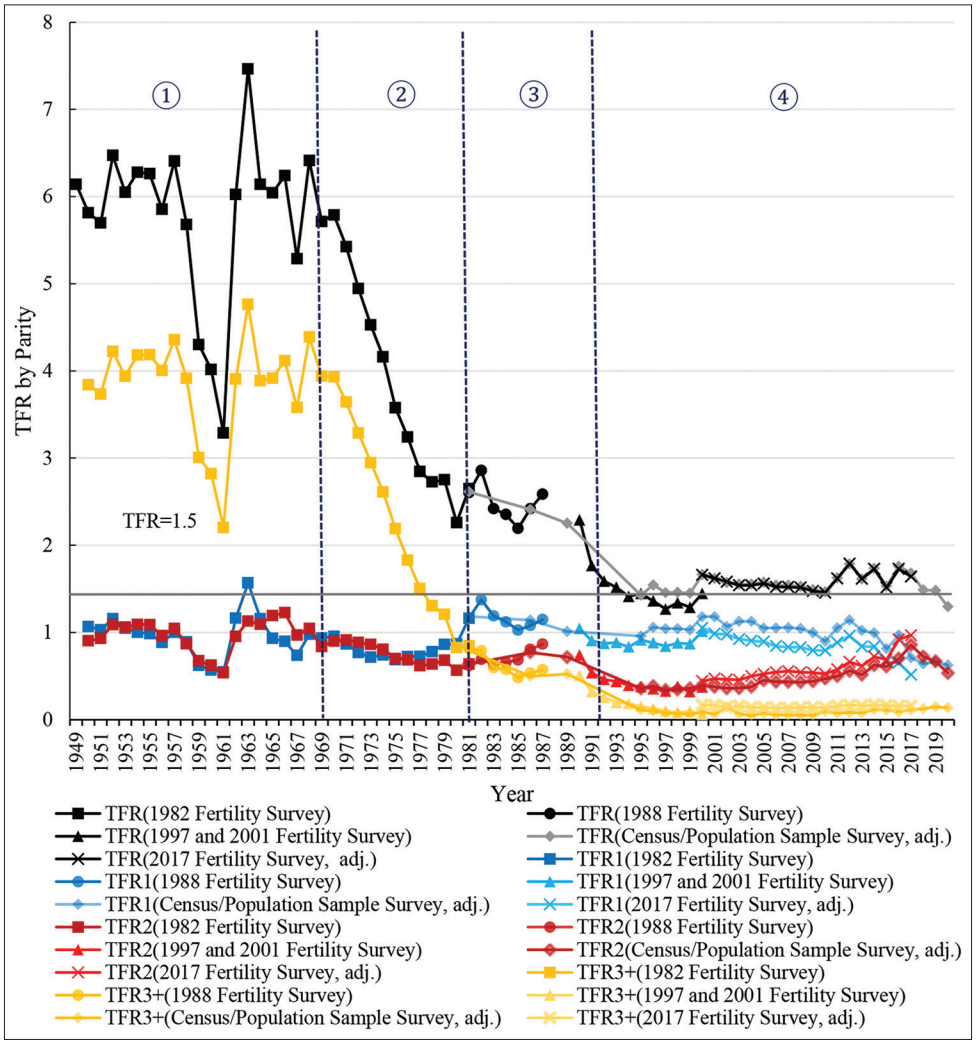


Figure 1. TFR and parity-specific TFR in China, 1949–2020

Sources: The 1982 Fertility Survey: Calculations from 1950 to 1962 made by Yao (1995) and NPFPC and CPDRC (2013), calculations from 1963 to 1981 made by Ma *et al.* (1986b) from the National One-per-Thousand Fertility Survey in 1982; The 1988 Fertility Survey: Calculations made by Yang *et al.* (1991) from the National Two-per-Thousand Fertility Survey in 1988; The 1997 and 2001 Fertility Surveys: Calculations made by Ding (2003) from the National Population and Reproductive Health Survey in 1997 and 2001; Censuses/Population Sample Surveys: Estimations from national population censuses performed in 1982, 1990, 2000, 2010 and 2020, the national one per cent intercensal sample surveys conducted in 1987, 1995, 2005 and 2015, and national one per thousand annual sample surveys of population change from 1982 to 2020 ; 2017 Fertility Survey: Authors' own estimations from the National Fertility Survey in 2017 and the population census conducted in 2020. It should be noted that we adjusted estimations of the Census/Population Sample Survey and the 2017 Fertility Survey from 2000 to 2020 based on the births published by China's National Bureau of Statistics (NBSC, 2022).

fact that family conditions depended to a large extent on the number of family workforce under the system. Both early marriage and childbearing were conducive to increase the number of workforce in the family (Su, 1992). As shown in Figure 2, the average age of women at the first and second births declined markedly in the early 1980s, from the age of 24.82 and 27.51 in 1975 to the age of 23.36 and 26.28 in 1981, respectively. The advance of age of women at first marriage led to a higher fertility rate in the 20–22 age group. The "compensatory" effect is mainly due to the fact that delayed marriage and childbearing in the 1970s creating a "deficit" that led to an increase in fertility in the corresponding age group in the 1980s (Gu, 1991; Zha & Ji, 1984). Other factors contributing to fertility fluctuations in the 1980s include the adjustments and improvements of fertility policies in the mid to late1980s as well as the entry of peak birth cohort of 1963 into reproductive ages (Wu, 2010).

It should also be noted that the fluctuations and increases in the TFR in the 1980s were mainly caused by fertility changes for the parities one and two, while the fertility rate for parity three and above declined. Compared with the first-birth fertility, the second-birth fertility was relatively flat. The reasons include: (1) Most urban women and a small proportion of rural women obeyed to the one-child policy rule and did not progress to their second births and (2) the

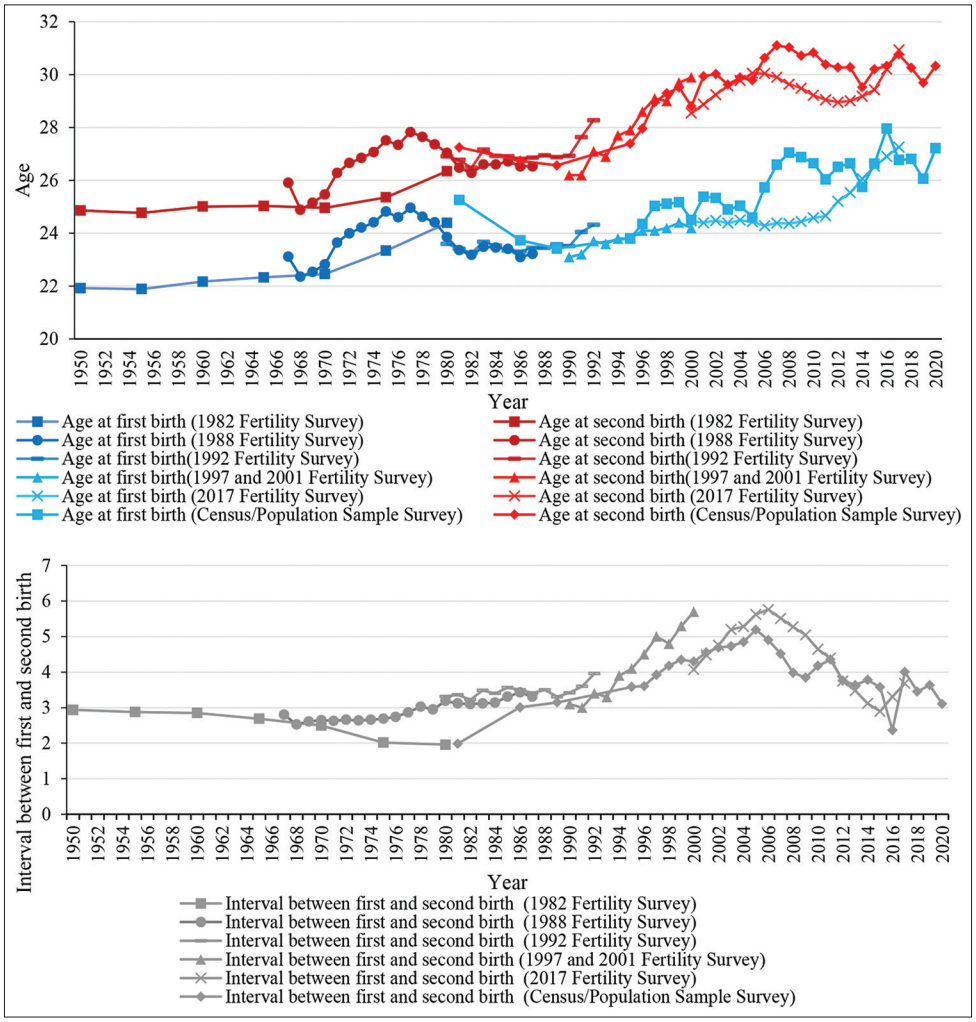


Figure 2. Mean ages at the first birth, second birth, and interval between the first and second birth of Chinese women, 1950 – 2020

Sources: The 1982 Fertility Survey: Calculations made by Liu and Zou (2011) from the National One-per-Thousand Fertility Survey in 1982; The 1988 Fertility Survey: Calculations made by Feeney and Feng (1993) from the National Two-per-Thousand Fertility Survey in 1982; The 1992 Fertility Survey: Tabulations from Statistics of the 1992 Fertility Sampling Survey in China, edited by Jiang (1995); The 1997 and 2001 Fertility Surveys: Calculations made by Ding (2003) from the National Population and Reproductive Health Surveys in 1997 and 2001; Censuses/Population Sample Surveys: Estimations from national population censuses performed in 1982, 1990, 2000, 2010 and 2020, and the national one per cent intercensal sample surveys conducted in 1987, 1995, 2005 and 2015, and the national one per thousand annual sample surveys of population change from 1982 to 2020; The 2017 Fertility Survey: Authors' own estimations from the National Fertility Survey in 2017.

interval between the first and second births of women who were allowed to have second births increased, especially between 1985 and 1987. The increase in the interval between the birth of the first child and the second child further delayed women's age at the second child, which led to a reduction in the period fertility rate of the second birth (Ding, 2003).

In the fourth stage, the second-birth fertility rate for Chinese women showed a rapid decline, falling to a level of around 0.33 by the end of the 1990s. During the same period, the first-birth fertility rate remained stable at 0.84 – 0.91. The trend in the parity-specific TFR for the third child and above also showed a rapid decline in the early 1990s, but remained stable in the late 1990s. Assuming that underreporting of births was more serious for the second child and higher orders, the actual second-birth fertility rate throughout the 1990s was higher than the level shown in Figure 1.

Despite the great differences among estimates of the TFR since 2000 based on different data sources, the rough trends of fertility are relatively consistent. The estimated first-birth fertility rate based on census/sample survey data was higher than that calculated based on the 2017 Fertility Survey data, while the estimates of the fertility rates for the second birth, third birth and above based on census/sample survey data were lower than those estimated based on the 2017 Fertility Survey data. An overall trend from estimates based on different data sources shows that the level of second-birth fertility began to fluctuate upwards after 2000, reaching its peak in 2016 before falling down.

This contrasts with the fluctuating decline in the first-birth fertility and the relative stability of fertility rate for the third birth and above over the same period.

To be specific, the changes in the second-child fertility after 2000 have been influenced by two factors: First, the reduction in the lifetime cohort fertility that has had a negative effect on the second-birth fertility; second, the gradual entry of the first only-child cohort into reproductive age since 2004 as well as the gradual relaxation of fertility policies, including the successive removal of the restriction on interval between first and second birth since 2002, the implementation of selective two-child policy in 2013 and the universal two-child policy in 2015. Due to the relaxation of fertility policies, the proportion of women in reproductive age who are allowed to have a second child has increased, which, in turn, has had a positive effect on the second-birth fertility. It is evident that the positive effect of the latter factor was greater than the negative effect of the former factor. With the removal of the two-child restriction in 2015, the second-birth fertility fell after a timely rise in 2016 and 2017, suggesting that the policy did have a significant but unsustainable "accumulative" effect.

There has been a significant delay in women's age at first and second births and an increase in the interval between the first and second births since the 1990s. This has been resulted from the further intensification of fertility policy in the 1990s as well as the increased opportunities for women to have higher education during this period. Women's age at first birth is closely associated with women's age at first marriage, while their age at second births is associated with both the age at first birth and the interval between the first and second births. The interval between the first and second births maintained around 3 years from the 1950s to the 1970s in the first and second stages of demographic transition, which is close to natural intervals. In the third stage in the 1980s under the one-child policy and its corresponding regulations on birth interval, the interval between the first and second births increased to around 4 years. From the mid-1990s, the interval between the first and second births increased and became higher than 4 years, reaching a peak in 2005–2006, and then declined as birth spacing controls were phased out. The interval remained at around 3 years recently.

3.2. Period fertility schedule by parity

Trends in overall fertility came from changing age-specific fertility. Figure 3 shows age-specific fertility rates by parity in 1970, 1980, 1989, 2000, 2010, and 2016. The fertility schedule of Chinese women saw a transition from a "natural fertility schedule" to a "controlled fertility schedule" with the fertility transition. During the 1950s and 1960s, the age-specific fertility schedule was an unrestricted "natural fertility schedule" (Feng, 1985), characterizing by "early, dense, and numerous" childbearing with women having children throughout the reproductive span (Song & Li, 1991). Following the implementation of "later, longer, fewer" policy in the 1970s and one-child policy in 1980, the age pattern of Chinese women shifted to a "controlled fertility schedule" (Feng, 1985). Since the mid-1990s, the age-specific fertility schedule has been characterized by a steady delay in the age at births and an increasing proportion of second births (Song & Tang, 2017). There was a significant change in the fertility schedule in 2016, with fertility rates of second births being higher than that of first births, which is clearly related to the implementation of universal two-child policy.

The schedule of age-specific fertility shows different characteristics for different parities. There was a pronounced change in age-specific fertility schedule for parity one before and after the start of the 21st century, shifting from a concentrated schedule to a flat one. The shift of fertility schedule for parity three and above occurred even earlier, between the 1970s and 1980s, with substantial declines at nearly all ages. The fertility schedule for parity three and above stabilized ever since the 1990s. Compared with the fertility schedules for parity one and for parity three and above, the changes of fertility schedule for parity two were more complex. In 1970, the peak of age-specific fertility rates for parity two of Chinese women occurred at age 23 with a fertility rate of 0.13, indicating an early childbearing and a higher level for parity two. The 1980s saw a delay in childbearing for parity two, with the peak fertility rate shifting to occur at age 25 with a fertility rate of 0.08. The schedule of age-specific fertility rate for parity two fluctuated in 1989, with the peak age advancing to 24 years. The age-specific fertility rates for parity two in 2000 and 2010 shows a substantial fall at nearly all ages compared to those for 1989. After the implementation of the universal two-child policy, there was a significant increase in age-specific fertility rate for parity two in 2016, especially for women aged 35 years old and above.

It is worth noting that there are differences in the schedules of estimates in 2000, 2010, and 2016 using different data sources. In general, estimates of age-specific fertility scheduled for parity two, parity three and above based on 2017 Fertility Survey data were higher than those obtained from census/sample survey data, while the opposite was true for age-specific fertility rates for parity one.

3.3. PPPRs

Figure 4 shows the PPPRs by parity from 1955 to 2016. The progression ratio to second births in China remained stable at around 0.98 from the 1950s to the late 1960s and

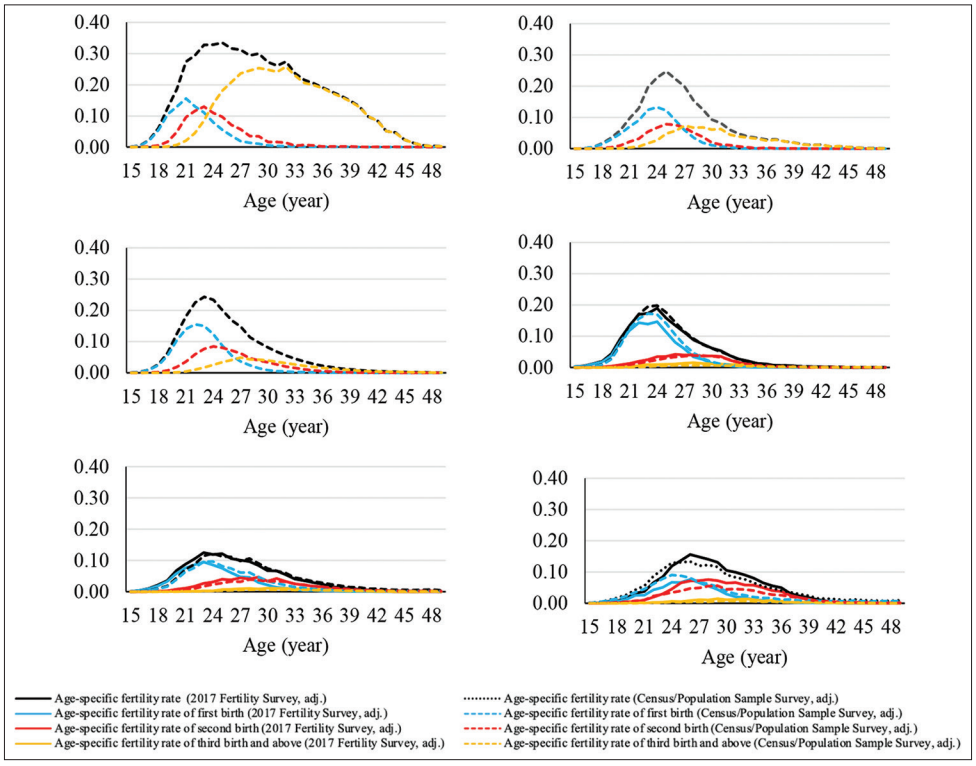


Figure 3. Schedules of total and parity-specific fertility of Chinese women in selected years

Sources: 1970: Tabulates from Fertility Data of China, edited by Yao (1995); 1980: Tabulates from Analysis of the National One-Per-Thousand-Population Sample Survey, edited by Population & Economics Editorial Board (1983); 1989: Tabulations from censuses conducted in 1990; 2000, 2010, 2016: Tabulates from censuses conducted in 2000, 2010, and national one per thousand annual sample surveys of population change in 2016, as well as authors' own estimation from the 2017 Fertility Survey.

declined slightly to 0.96 in 1979. That is, almost all Chinese women who had already had one child would progress to a second child during the 1950s to 1970s, indicating a widespread birth of two children. Since the introduction of the one-child policy in 1980, the progression ratio to second births of Chinese women declined appreciably to 0.80 in 1984 and rebounded slightly from 1985 to 1987 when the fertility policy in rural areas was adjusted and continued to decline afterward. The decline of the progression ratio to second births during this period undoubtedly reflects the impact of the fertility policy. The low level of social development and traditional fertility attitudes might have the opposite effect on the progression ratio to second births during this period, that is, inhibiting this decline. During the 1990s, there was a monotonous and rapid decline in the progression ratio to second births. The rapid decline during this period was partly influenced by the quality of the birth data. Given that the fertility policy was stable, it suggests that socioeconomic development factors played important roles to the decline of progression ratio to second births. The progression ratio to second births remained at around 0.45 in the 2000s, then rose to 0.53 in 2005 as the first generation of only-child entered their reproductive age, and then fluctuated between 0.53 and 0.61 from 2005 and 2015. With the implementation of the selective two-child policy in 2013 and the universal two-child policy in 2015, the progression ratio to second births began to fluctuate upward. In particular, after the universal two-child policy, the progression ratio to second births rose sharply from 0.54 in 2015 to 0.71 in 2016, reaching a peak of 0.74 in 2017.

The progression ratio to first births has remained stable at above 0.95 in the second half of the 20th century, except for the period of the Great Famine. Even though the progression ratio to first births has been slowly declining since the beginning of the 21st century, it has remained at 0.90 recently. The decline in the progression ratios to the third and fourth births was earlier and faster than the decline in the progression ratio to the second births, which occurred in the mid-1960s. It should be noted that the progression ratios to third and fourth births declined rapidly from the 1970s onwards, plummeting from 0.89 and 0.87 in 1970, to 0.36 and 0.30 in 1990, respectively, as a result of the "later, longer, fewer" policy that began in the early 1970s and one-child policy in the 1980s. After the 1990s, the progression ratios to the third and fourth births stabilized at around 0.20.

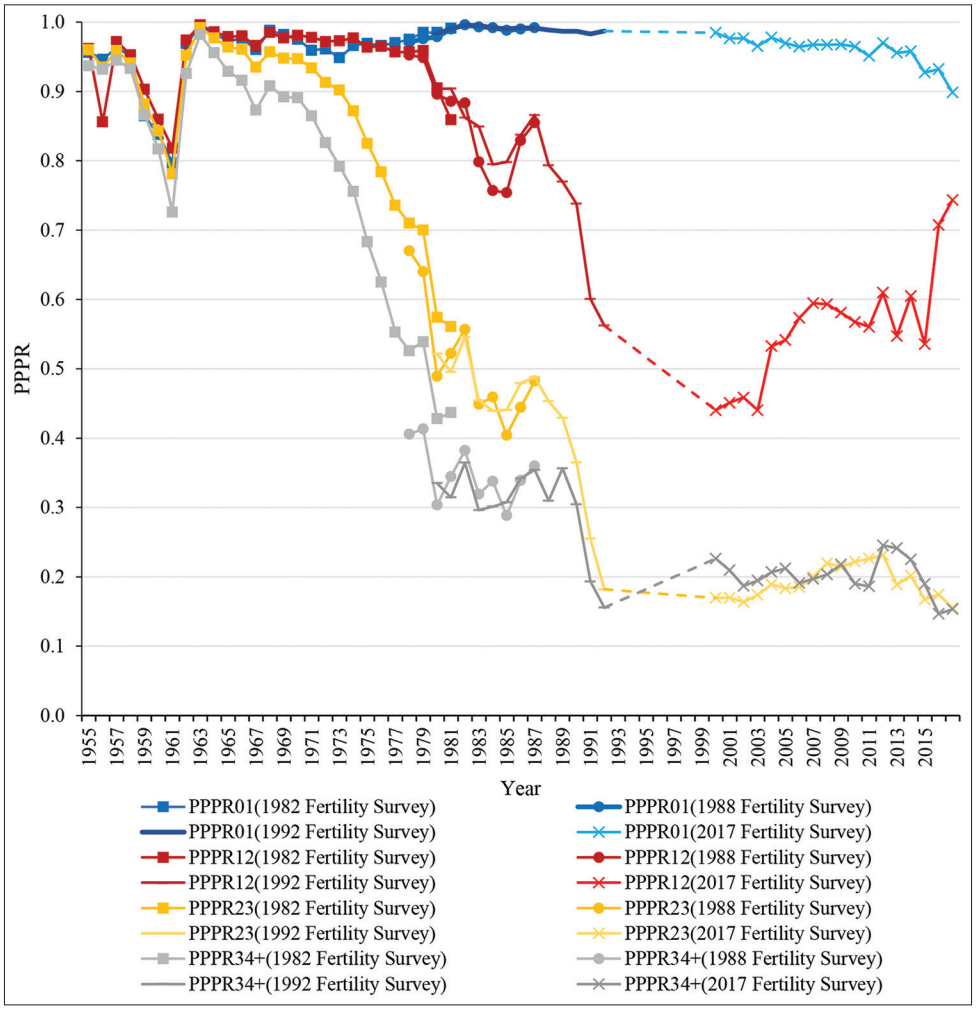


Figure 4. Period parity progression ratios of Chinese women, 1955–2017

Sources: The 1982 Fertility Survey: Calculations made by Feeney and Yu (1987) from the National One-per-Thousand Fertility Survey in 1982; The 1988 Fertility Survey: Calculations made by Yang *et al.* (1991) from the National Two-per-Thousand Fertility Survey in 1988; The 1992 Fertility Survey: Tabulations from Statistics of the 1992 Fertility Sampling Survey in China, edited by Jiang (1995); The 2017 Fertility Survey: Authors' own estimations from the 2017 Fertility Survey.

3.4. Period PPTFR by parity

In Figure 5, we compare the conventional TFR and PPTFR. There are remarkable similarity and difference between the two series. In terms of similarity, the two series show generally similar levels of fertility and move together over most of the period. Regarding the difference, the fertility trend indicated by PPTFR is flatter than that by conventional TFR. Supplementary information 3 presents the values of conventional TFR and PPTFR from 1949 to 2020.

By stages, the PPTFR is lower than the conventional TFR in the 1950s and early 1960s in the first stage, especially after the recovery from the Great Famine in 1962, but higher than the conventional TFR since the mid-1970s in the second stage. The conventional TFR tends to exaggerate the decline in fertility during this period, compared with those based on parity progressions (Feeney & Yu, 1987). In the third stage, the PPTFR is flatter than the conventional TFR again, showing a gap between the two series when the conventional TFR rebounded in 1982, 1986, and 1987, respectively. The conventional TFR suggests a reverse of fertility decline in the early 1980s, with the TFR rising substantially from 2.25 children per woman in 1980 to 2.65 in 1981. The PPTFR, however, appears to show continued decline of fertility, falling from 2.70 in 1980 to 2.65 in 1981. In the fourth stage, the conventional TFR showed a decreasing trend since the beginning of 21st century and a recovery with fluctuation from 2010 onwards.

Figure 6 shows the PPTFR by parity from 1955 to 2016. PPTFR for parity one is equivalent to PPPR to the first births. The PPTFRs for parity two and for parity three and above are flatter than the conventional TFRs for parity two and for

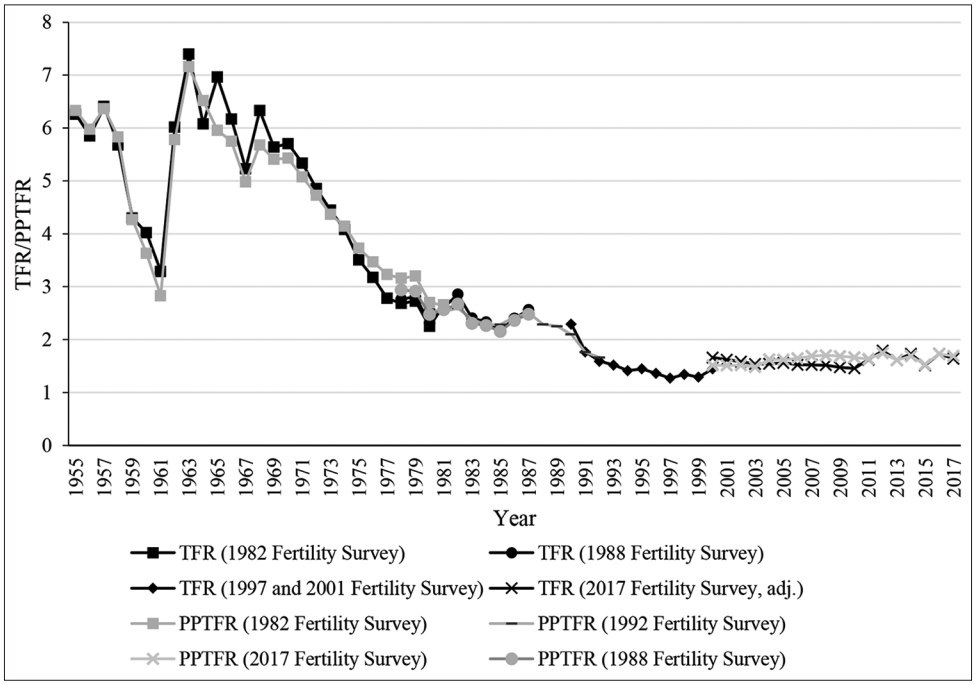


Figure 5. TFR and PPTFR in China, 1955–2017

Sources for TFR estimations: The 1982 Fertility Survey: Calculations from 1950 to 1962 made by Yao (1995) and NPFPC and CPDRC (2013), calculations from 1963 to 1981 made by Ma et al. (1986b) from the National One-per-Thousand Fertility Survey in 1982; The 1988 Fertility Survey: Calculations made by Yang et al. (1991) from the National Two-per-Thousand Fertility Survey in 1988; The 1997 and 2001 Fertility Surveys: Calculations made by Ding (2003) from the National Population and Reproductive Health Surveys in 1997 and 2001; Censuses/Population Sample Surveys: Estimations from national population censuses performed in 1982, 1990, 2000, 2010 and 2020, the national one per cent intercensal sample surveys conducted in 1987, 1995, 2005 and 2015, and national one per thousand annual sample surveys of population change from 1982 to 2020 ;The 2017 Fertility Survey: Authors' own estimations from the National Fertility Survey in 2017 and the population census conducted in 2020. It should be noted that we adjusted estimations of the 2017 Fertility Survey from 2000 to 2020 based on the births published by China's National Bureau of Statistics (NBSC, 2022).

parity three and above, correspondingly. The effects of fertility policy on fertility are evidently shown in the significant decline of PPTFR for parity three and above in the 1970s, the decline of PPTFR for parity two in the 1980s and 1990s, and the fluctuation upward of PPTFR for parity two in the 2010s.

3.5. Changes in CCFR and CPPR

There was a broadly monotonic downward trend in cohort cumulative fertility in China between women born in 1926 and 1977, except for a slight "rebound" in women born in 1961 and 1962. The average cohort cumulative fertility declined from about five children per woman in the late 1920s and early 1930s birth cohorts to around 1.7 children per woman in the 1970s birth cohorts (Table 1). The cohort cumulative fertility for parity one has remained stable at a level of 0.98 and above children per woman among all birth cohorts. A decline in the cohort cumulative fertility for parity two occurred first in the 1940s birth cohorts, declining from 0.96 children per woman in the early 1940s birth cohorts to 0.90 children per woman in the early 1950s birth cohorts, then declining rapidly to 0.66 children per woman in the 1960s birth cohorts, and finally declining to 0.57 children per woman in the early 1970s birth cohorts. The decline in cohort cumulative fertility for parity three and above emerged in earlier birth cohorts and to a greater extent than that for parity two. To be specific, the cohort cumulative fertility for parity three and above falling from 3.19 in 1930 birth cohort to 1.04 in 1950 birth cohort, further to 0.40 in 1960 birth cohort, and to < 0.20 in 1970 birth cohort.

As, shown in Figure 7, progression ratios to first births remained relatively stable and close to one among different birth cohorts of women. Accordingly, the dynamics of progression ratios from the first to second births were similar with cohort fertility for second births. Progression ratios to second births in women born during 1926 – 1951 remained above 0.9, showing a downward trend from the 1945 birth cohort. There was a large decline of progression ratios to second births between the 1951 and 1957 cohorts, followed by a "rebound" during 1960 and 1961 birth cohorts with a level around 0.7, fluctuated and declined thereafter, with the progression ratio to second births stabilizing at around 0.57 for the 1970–1977 birth cohorts.

Table 2 displays the contributions of changes in parity progression ratios to the first, second, third and later births in fertility changes of 1930, 1950, 1970 and 1977 births cohorts. The fertility decline in women born between 1930

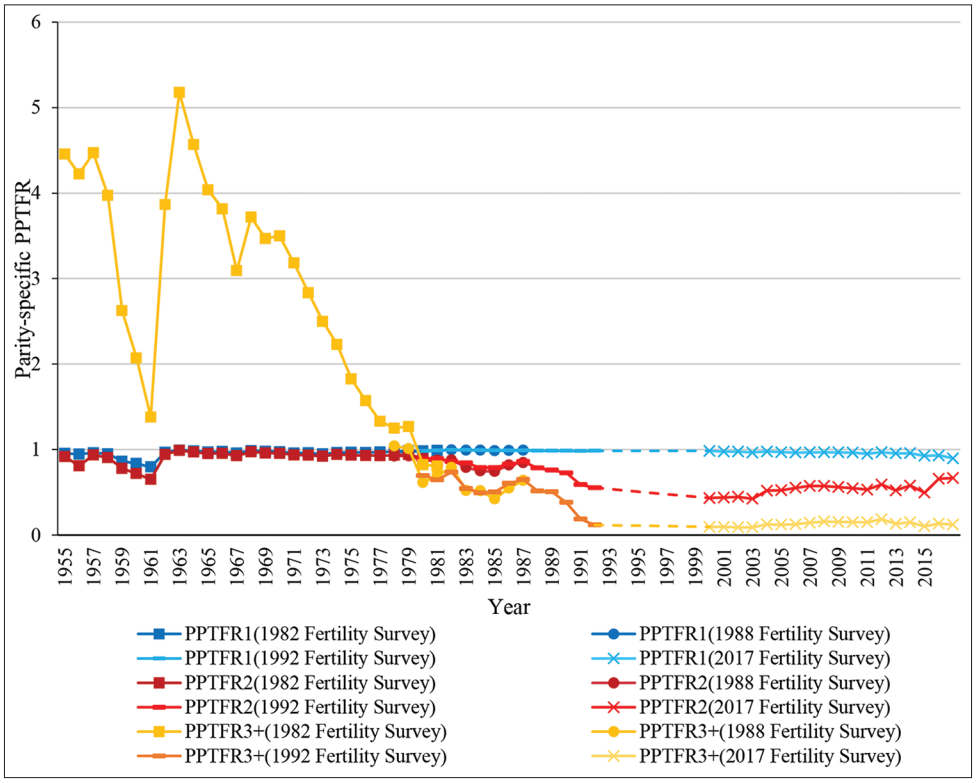


Figure 6. Parity-specific PPTFR in China, 1955–2017

Sources: The 1982 Fertility Survey: Calculations made by Feeney and Yu (1987) from the National One-per-Thousand Fertility Survey in 1982; The 1988 Fertility Survey: Calculations made by Yang et al. (1991) from the National Two-per-Thousand Fertility Survey in 1988; The 1992 Fertility Survey: Tabulations from Statistics of the 1992 Fertility Sampling Survey in China, edited by Jiang (1995); The 2017 Fertility Survey: Authors' own estimations from the 2017 Fertility Survey.

and 1950, from 5.10 to 2.93, was primarily driven by the decreasing transition to higher-order births. Specifically, the progression ratio to first births in 1950 cohort was slightly higher than that in 1930 cohort, contributing 0.09 to the fertility decline between 1930 to 1950 cohort. The progression ratios to the second births and third birth contributed −0.22 and −2.04 to this fertility decline, respectively. The fertility decline in the cohorts born between 1950 and 1970, from 2.93 to 1.75, was mostly due to falling progression ratio to second and higher orders. The contributions of progression ratios to the second and third births were −0.70 and −0.48, respectively.

3.6. The unique Chinese pattern of configurations of parity-specific fertility

Figure 8 shows TFR and parity-specific TFR in China, Japan, the U.S., Canada, and Czechia, 1949 – 2020, which confirms what we have discussed above. It is evident that fertility policy in China plays an important role in shaping the dynamics of fertility, especially for parity two and higher orders.

Figure 9 further shows that the progression ratio to first births for Chinese women was the highest among those countries except for the period of Great Famine from 1959 to 1961, while the parity progression ratio to second births was the lowest and parity progression ratio to third births was at a low level since the 1990s. It is worth noting that Chinese women's second-child fertility has been higher than those of Japan and Czechia since 2000 and was the highest during 2012 – 2017. Given that the parity progression ratio to second births was the lowest during this whole period except 2016 and 2017, the recent high level of second-child fertility of Chinese women is largely due to China's high parity progression ratio to first births.

In terms of cohort fertility, compared with those in developed low-fertility societies, the fertility levels of the 1940s and 1950s birth cohorts in China were higher, whereas the fertility levels of the 1970s birth cohorts were at the middle level. The CCFR for women in the 1970 cohort in China is 1.75, which is higher than those in East Europe (1.64), German-speaking countries (1.64), South Europe (1.57), and East Asia (including only Korea and Singapore, 1.68), and lower than those in Central Europe (1.86), West Europe (1.93), North Europe (2.00), and English-speaking non-European countries (2.05).

Table 1. Cohort cumulative fertility rates of Chinese women, born in 1926–1977

Cohort	CCFR	CCFR1	CCFR2	CCFR3+	Cohort	CCFR	CCFR1	CCFR2	CCFR3+
1926	4.84	0.96	0.90	2.98	1949	3.08	0.99	0.92	1.17
1927	4.97	0.96	0.91	3.09	1950	2.93	0.99	0.90	1.04
1928	5.05	0.97	0.92	3.16	1951-1957	-	-	-	-
1929	5.09	0.97	0.93	3.19	1958	2.11	0.99	0.67	0.44
1930	5.10	0.97	0.93	3.19	1959	2.07	0.99	0.66	0.42
1931	5.14	0.98	0.94	3.23	1960	2.05	0.99	0.66	0.40
1932	5.12	0.98	0.95	3.20	1961	2.11	0.99	0.68	0.43
1933	5.09	0.98	0.95	3.16	1962	2.10	0.99	0.70	0.41
1934	5.03	0.98	0.95	3.09	1963	2.01	0.99	0.67	0.35
1935	4.93	0.98	0.96	2.99	1964	1.99	0.99	0.67	0.33
1936	4.84	0.98	0.96	2.90	1965	1.95	0.99	0.66	0.30
1937	4.75	0.99	0.96	2.80	1966	1.93	0.99	0.66	0.28
1938	4.62	0.99	0.96	2.68	1967	1.89	0.99	0.64	0.27
1939	4.50	0.99	0.96	2.55	1968	1.83	0.99	0.62	0.23
1940	4.37	0.99	0.96	2.42	1969	1.78	0.99	0.59	0.20
1941	4.25	0.99	0.96	2.30	1970	1.75	0.99	0.57	0.19
1942	4.15	0.99	0.96	2.20	1971	1.73	0.99	0.56	0.18
1943	4.03	0.99	0.96	2.08	1972	1.73	0.99	0.56	0.18
1944	3.91	0.99	0.96	1.96	1973	1.73	0.99	0.57	0.18
1945	3.74	0.99	0.96	1.79	1974	1.73	0.98	0.57	0.17
1946	3.55	0.99	0.96	1.61	1975	1.72	0.98	0.56	0.17
1947	3.39	0.99	0.95	1.45	1976	1.72	0.98	0.56	0.18
1948	3.23	0.99	0.94	1.31	1977	1.70	0.98	0.56	0.16

Sources: Authors' own estimations from the 1990 census (1926 – 1950 cohorts) and the 2017 Fertility Survey (1958 – 1977 cohorts).

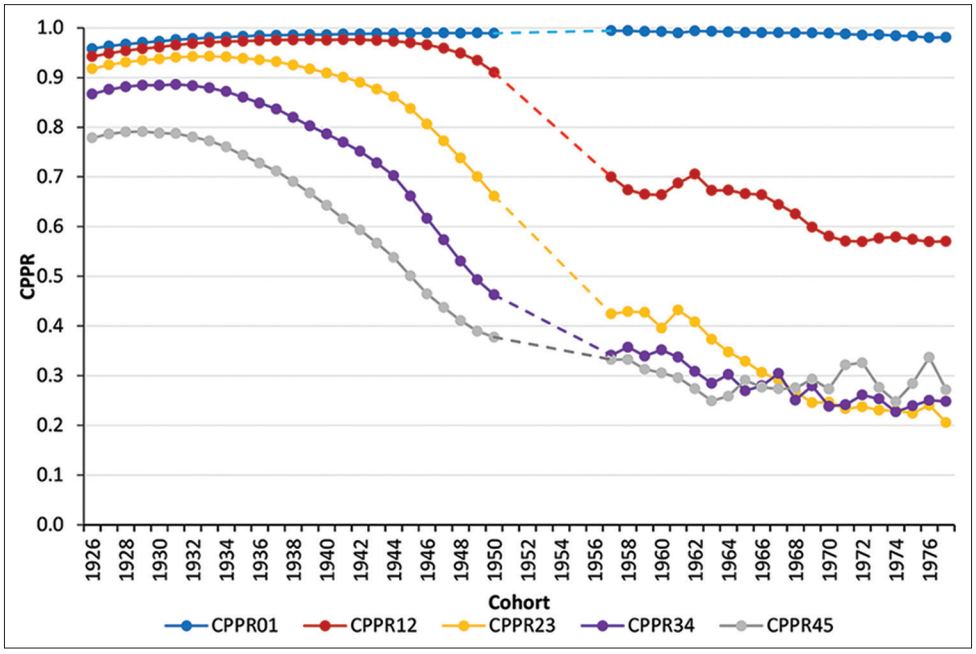


Figure 7. Cohort parity progression ratios of Chinese women, 1926–1977

Sources: Authors' own estimations from the 1990 census (1926 – 1950 cohorts) and the 2017 Fertility Survey (1958 – 1977 cohorts).

 https://doi.org/10.36922/ijps.v8i1.348

Table 2. Decomposition of changes in cohort cumulative fertility across cohorts

Cohort	dPPR01	dPPR12	dPPR23	Changes in CCFR
1930–1950	0.09	−0.22	−2.04	−2.17
1950–1970	0.00	−0.70	−0.48	−1.18
1970–1977	−0.01	−0.01	−0.03	−0.05

Sources: Authors' own estimations from the 1990 census (1926 – 1950 cohorts) and the 2017 Fertility Survey (1958 – 1977 cohorts).

The configurations of parity-specific fertility can vary greatly, even when summing up to similar cohort fertility levels. Zeman *et al.* (2018) have categorized the different configurations of parity progression ratios corresponding to a completed cohort fertility of 1.6 children per woman into "benchmark," "high childlessness," "one-child pattern," "stopping at two," and "polarized pattern" (Table 3). Compared to these patterns, the combination of configuration of parity-specific fertility in China is markedly different. The results of the 2017 Fertility Survey show that the fertility of women born in 1982 – 1983 is around 1.62 children per woman, with the progression ratios to the first, second, and third births at levels of around 0.95, 0.57, and 0.20, respectively. The "universal" nature of childbearing is evident in China with a proportion of childlessness < 5%. The progression ratio to first birth is higher than that in any other low-fertility developed country. Yet, progression ratios to second and third births are both low. That is, compared to low-fertility developed countries, the progression ratio to first births in China has a boosting effect on women's cohort fertility, while the progression ratio to the second and third births has the opposite effect. China builds on a unique mode to those based on developed low-fertility societies.

4. Concluding remarks

The fertility level and pattern in China have been shaped by fertility culture, socioeconomic development, as well as fertility policy. The traditional Confucian culture of marriage and fertility as well as low level of socioeconomic development in the 1950s and 1960s were the main reasons that fertility level and pattern for the second births were characterized by "early and universal". The average age at second births was around 25 years old and progression ratios from first to second birth remained stable at over 0.98. The "later, longer, fewer" policy has negligible impact on progression ratios from first to second births. Although there was a delay in the average age of women at second births, the progression ratios from first to second births remain almost unchanged. With the implementation and adjustment of strict restrictions on the parity and timing/spacing of second children in the 1980s, the second-birth fertility rate fluctuated. In contrast to a rebounded second-birth TFR, the progression ratio from first to second births declined. With the gradual removal of the interval between births and the number of births in the past decade or so, the second-birth fertility rate has seen an increase, especially after the implementation of the universal two-child policy in 2015, indicating a significant policy effect. However, the decline in the second-birth fertility rate after 2017 shows that socioeconomic development and changes in the childbearing values have become the main drivers of contemporary second-birth fertility in China, not the fertility policy.

A review of changes in the levels and patterns of second-birth fertility in China since 1949 shows that traditional fertility attitudes and low socioeconomic levels were responsible for high and widespread levels of second birth in the early years, while changes in fertility attitudes toward modernity and rapid economic development have dominated the recent low levels of second-birth fertility. Both the policy of controlling the number and timing/spacing of second child, which began in the 1980s, and the more recent policy of gradually easing the number and timing/spacing of second children have had an impact on the levels and patterns of second-birth fertility. However, the impact of policy relaxation in a low-fertility context has been less sustainable than the earlier control policies that effectively contributed to the decline in fertility for birth order two.

The high second-birth fertility in the recent decade in China, both period and cohort, was more likely to be associated with the universal progression ratio from 0 to 1, while the progression ratio from first to second births was still low. We have good reasons to believe there will be a continued delay of first births and even an increase in lifetime infertility rates which would lead to a decline in the progression ratio to first births, while the short-term "accumulating" effect on second births due to recent policy adjustments would quickly disappear. Until supportive measures are in place and take effect, a decline in second-birth fertility rate will be inevitable.

The configuration of parity-specific fertility in China differs markedly from that in the west. The "universal" childbearing is evident in China, where the progression ratio to first birth is higher than in any low-fertility developed country. In contrast, progression ratios from first to second and from second to third in China are both low since the 1990s. This is related to the internalization of long-standing strict fertility policies and rapid socioeconomic development. As the younger birth cohorts are more likely to be influenced by individualism and their fertility may become more diverse, the progression ratio to first birth is likely to decrease, and it is too early to be seen to what extent the progression ratios from first to second

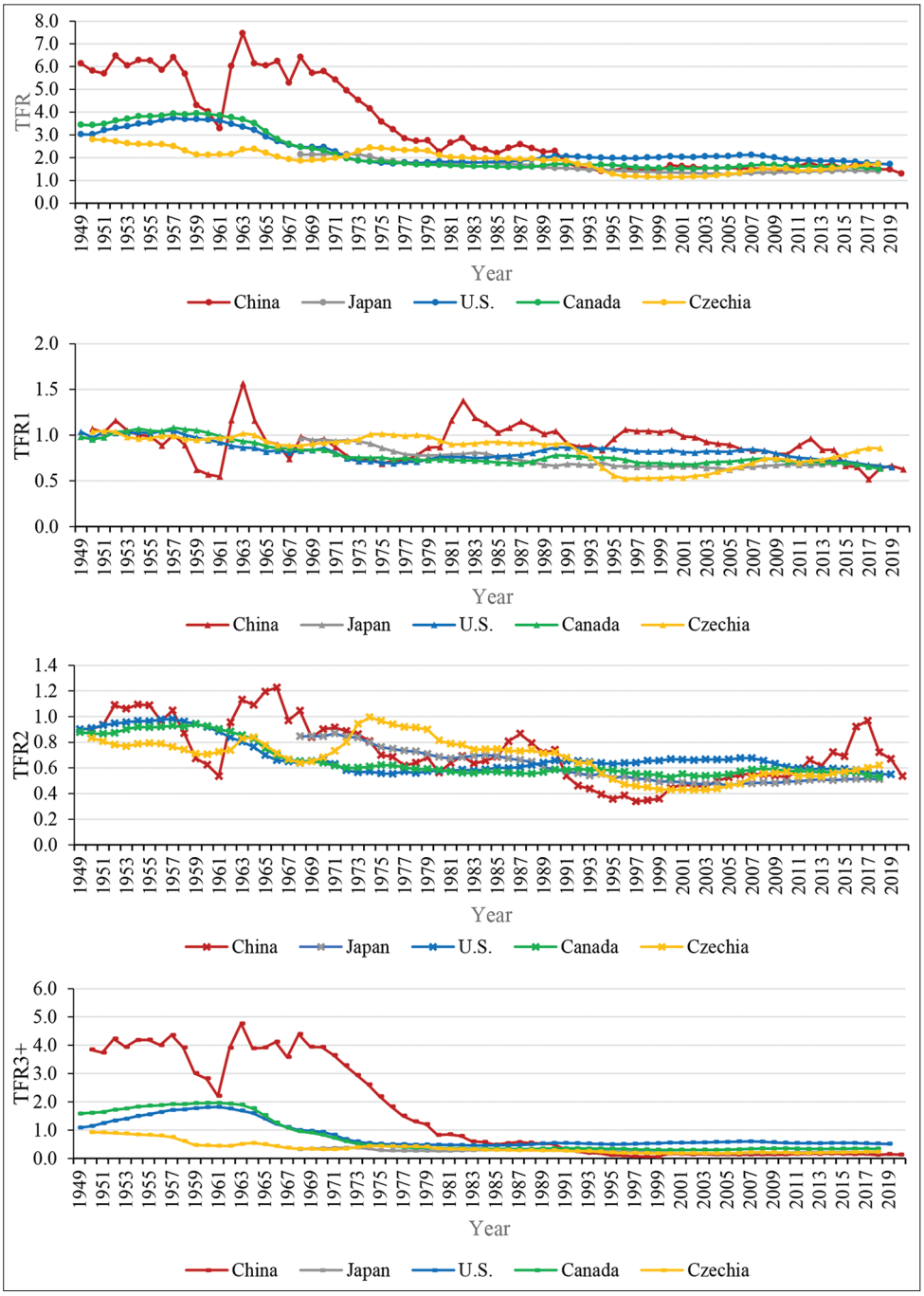


Figure 8. TFR and parity-specific TFR in China, Japan, the U.S., Canada, Czechia, 1949–2020

Sources: China: The 1982 Fertility Survey (1949–1981), the 1988 Fertility Survey (1982 – 1987), the 2001 Fertility Surveys (1990–1994), Censuses/ Population Sample Surveys (1989, 1995–1999, 2018–2020), the 2017 Fertility Survey (2000–2017), and a linear interpolation (1988). Japan, the U.S., Canada, and Czechia: Human Fertility Database.

births and from second to third births would follow the Western patterns, that is, to increase.

This study was limited by the lack of reliable sources of data and was, therefore, unable to estimate the reliable fertility level since the 1990s. Although efforts had been made to adjust estimates of period fertility rate using NBS-published births from 2000 to 2020, the study directly used the face values of data sources in the 1990s. Due to the births underreports in the data sources, the fertility might be under-estimated, especially for fertility of higher orders.

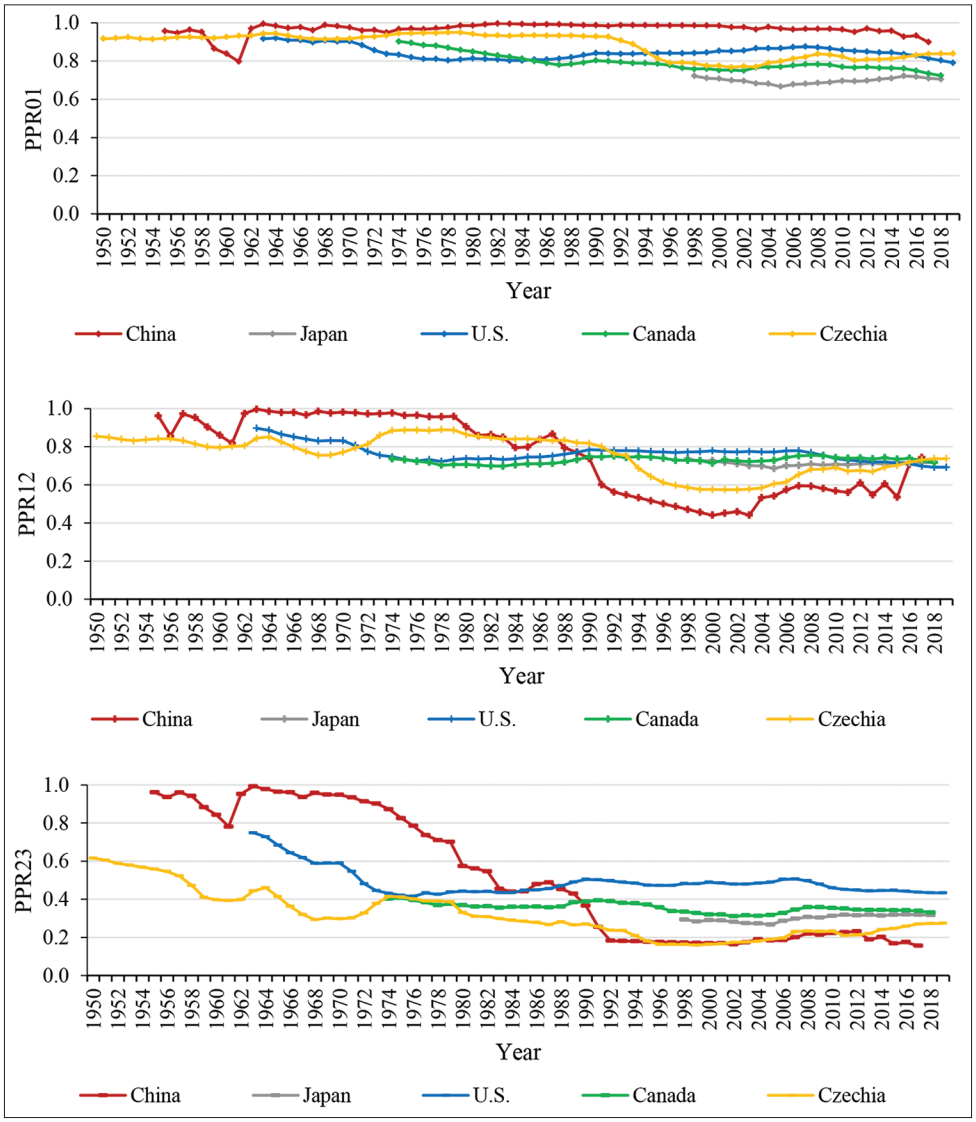


Figure 9. Parity-specific PPPR in China, Japan, the U.S., Canada, Czechia, 1950–2020

Sources: China: The 1982 Fertility Survey (1949–1981), the 1988 Fertility Survey (1982 – 1987), the 2001 Fertility Surveys (1990–1994), Censuses/ Population Sample Surveys (1989, 1995–1999, 2018–2020), the 2017 Fertility Survey (2000–2017), and a linear interpolation (1988). Japan, the U.S., Canada, and Czechia: Human Fertility Database.

Table 3. Different patterns of configurations of parity progression ratios corresponding to a completed cohort fertility of 1.6 children per woman

Patterns of PPR configurations	PPR01	PPR12	PPR23
Benchmark	0.80	0.72	0.30
High childlessness	0.70	0.85	0.37
One-child pattern	0.90	0.55	0.32
Stopping at two	0.85	0.75	0.15
Polarized pattern	0.75	0.68	0.45
Chinese pattern	0.95	0.57	0.20

Sources: The Chinese pattern: authors' own estimations from the 2017 Fertility Survey. All other patterns are estimations made by Zeman *et al.* (2018).

Despite the data limit, this study mainly aims to illustrate the dynamics of fertility by parity rather than to estimate the exact fertility values.

Acknowledgments

None.

Funding

This study was supported by a research project "China's Population Development Strategy in the New Era" from China National Social Sciences Foundation (Grant No. 22AZD083).

Conflict of interest

The authors have no conflicts of interest to declare.

Author contributions

Conceptualization: Yuanyuan Duan, Wei Chen
Formal analysis: Yuanyuan Duan
Writing – original draft: Yuanyuan Duan
Writing – review & editing: Wei Chen

Ethics approval and consent to participate

Not applicable as this study involves the analysis of secondary data.

Consent for publication

Not applicable.

Availability of data

Some of the data used in this paper are available to the public: (i) National population censuses: http://www.stats.gov.cn/tjsj/pcsj/; (ii) national 1‰ annual sample surveys of population change: http://www.stats.gov.cn/tjsj/ndsj/; and (iii) Human Fertility Database: https://www.humanfertility.org/cgi-bin/main.php

References

Attané, I. (2022). China's new three-child policy: What effects can we expect? *Population Societies*, 596(1):1-4.

Basten, S., & Jiang, Q. (2014). China's family planning policies: Recent reforms and future prospects. *Studies in Family Planning*, 45(4):493-509.

https://doi.org/10.1111/j.1728-4465.2014.00003.x

Billari, F.C., & Kohler, H.P. (2004). Patterns of low and lowest-low fertility in Europe. *Population Studies (Camb)*, 58(2):161-176.

https://doi.org/10.1080/0032472042000213695

Bongaarts, J., & Feeney, G. (1998). On the quantum and tempo of fertility. *Population and Development Review*, 24(2):271-291.

https://doi.org/10.2307/2807974

Bongaarts, J., & Greenhalgh, S. (1985). An alternative to the one-child policy in China. *Population and Development Review*, 11(4):585-617.

https://doi.org/10.2307/1973456

Bongaarts, J., & Sobotka, T. (2011). *Demographic Explanation for the Recent Rise in European Fertility: Analysis Based on the Tempo and Parity-Adjusted Total Fertility*. Vienna: Vienna Institute of Demography.

Bongaarts, J., & Sobotka, T. (2012). A demographic explanation for the recent rise in European fertility. *Population and Development Review*, 38(1):83-120.

https://doi.org/10.1111/j.1728-4457.2012.00473.x

Chen, W. (2003). *Induced Abortion in China: Trends, Patterns and Determinants.* (Doctor of Philosophy). Australia: Australian National University.

Chen, W. (2016). China's demographic estimation using the generalized stable population model. *Population Journal*, 38:5-13 [Article in Chinese].

Chen, W., & Duan, Y. (2019). Recent levels and trends of fertility in China. *Population Research*, 43(1):3-17 [Article in Chinese].

Ding, J. (2003). Analysis about the impact of changes in fertility pattern on fertility level for China between 1991-2000. *Population Research*, 2:55-60 [Article in Chinese].

Feeney, G. (1983). Population dynamics based on birth intervals and parity progression. *Population Studies* (*Camb*), 37(1):75-89.

https://doi.org/10.2307/2174381

Feeney, G., & Feng, W. (1993). Parity progression and birth intervals in China: The influence of policy in hastening fertility decline. *Population and Development Review*, 19(1):61-101.

https://doi.org/10.2307/2938385

Feeney, G., & Yu, J. (1987). Period parity progression measures of fertility in China. *Population Studies*, 41(1):77-102.

Feng, F. (1985). An exploration of the shift in fertility schedules in China. *Population and Economics*, 6:4-9 [Article in Chinese].

Feng, X. (2018). The willingness to have a second child: Distance from assumption to reality. *Journal of Xinjiang Normal University (Philosophy and Social Sciences)*, 39(1):115-123 [Article in Chinese].

Gietel-Basten, S., Han, X., & Cheng, Y. (2019). Assessing the impact of the "one-child policy" in China: A synthetic control approach. *PLoS One*, 14(11):e0220170.

Goodkind, D. (1992). *Estimates of averted Chinese births, 1971-1990*: *Comparisons of Fertility Decline, Family Planning Policy, and Development in Six Confucian Societies*: Australia: Research School of Social Sciences, Australian National University.

Goodkind, D. (2017). The astonishing population averted by China's birth restrictions: Estimates, nightmares, and reprogrammed ambitions. *Demography*, 54(4):1375-1400.

https://doi.org/10.1007/s13524-017-0595-x

Greenhalgh, S. (1986). Shifts in China's population policy, 1984-86: Views from the central, provincial, and local levels. *Population and Development Review*, 12(3):491-515.

https://doi.org/10.2307/1973220

Gu, B. (1991). Fertility trends in rural China in the 1980s from the two-per-thousand fertility survey. *Chinese Journal of*

Population Science, 3:13-17 [Article in Chinese].

Gu, B., Wang, F., Guo, Z., & Zhang, E. (2007). China's local and national fertility policies at the end of the twentieth century. *Population and Development Review*, 33(1):129-148.

Jiang, Q., & Liu, Y. (2016). Low fertility and concurrent birth control policy in China. *The History of the Family*, 21(4):551-577.

https://doi.org/10.1080/1081602X.2016.1213179

Jiang, Z. (1995). *Statistics of the* 1992 *Fertility Sampling Survey in China*. Beijing, China: China Population Publishing House [Article in Chinese].

Jin, Y., Song, J., & Chen, W. (2016). Women's fertility preference and intention in urban China: An empirical study on the nationwide two-child policy. *Population Research*, 40(6):22-37 [Article in Chinese].

Kohler, H.P., Billari, F.C., & Ortega, J.A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28(4):641-680.

Liu, S., & Zou, M. (2011). Birth interval between first and second child and its policy implications. *Population Research*, 35(2):83-93 [Article in Chinese].

Lu, Y., & Zhai, Z. (2009). *Sixty Years of New China Population*. Beijing: China Population Press [Article in Chinese].

Ma, Y., Wang, Y., & Yang, S. (1986a). The porpoal of progression model of population development and the establishment of index of parity progression-based total fertility rate. *Populaton and Economics*, 3:23-32 [Article in Chinese].

Ma, Y., Wang, Y., & Yang, S. (1986b). The proposal of progression model of population development and the establishment of index of parity progression-based total fertility rate. *Population and Economics*, 2:24-32 [Article in Chinese].

McDonald, P., & Kippen, R. (2007). *The intrinsic total fertility rate*: A New Approach to the Measurement of Fertility. New York: Population Association of America.

National Bureau of Statistics of China. (2022). *China Statistical Abstract* (*2021*). Beijing: China Statistics Press [Article in Chinese].

National Population and Family Planning Commission, China Population and Development Research Center. (2013). *Handbook on population and family planning data (2012)*. China, Beijing: China Population Publishing House [Article in Chinese].

Population and Economics Editorial Board. (1983). *An analysis of a national one-per-thousand-population sample survey in birth rate*. Changchun, China: Changchun Xinhua Printing Factory [Article in Chinese].

Rallu, J.L., & Toulemon, L. (1994). Period fertility measures: The construction of different indices and their application to France, 1946-89. *Population an English Selection*, 6:59-93.

Ryder, N.B. (1986). Observations on the history of cohort fertility in the United States. *Population and Development Review*, 12(4):617-643.

Schoen, R. (2004). Timing effects and the interpretation of period fertility. *Demography*, 41(4):801-819.

Shi, R., Chen, N., & Zheng, Q. (2018). Evaluation on the effect of childbearing policy adjustments in China. *Chinese Journal of Population Science*, (4):114-125+128 [Article in Chinese].

Song, J. (2017). Turning point: Where will China's fertility rate go?-based on European experience and enlightenment. *Exploration and Free Views*, 4:70-75 [Article in Chinese].

Song, J., & Tang, S. (2017). Characteristics and changes in women's fertility schedules in China since 1995. *Chinese Journal of Population Science*, 4:15-27+126 [Article in Chinese].

Song, T., & Li, C. (1991). Schedule of women's fertility in contemporary China. *Population Research*, 3:17-22 [Article in Chinese].

Su, R. (1992). Analysis of fertility trends among Chinese women in the 1990s. *Population and Economics*, 1:10-16.

Wang, J. (2003). Changes in fertility patterns and estimates of fertility levels in China, 1990-2000. *Chinese Journal of Population Science*, 4:36-42 [Article in Chinese].

Wang, J., & Ge, Y. (2016). Population trends in China under the universal two-child policy. *Population Research*, 40(6):3-21 [Article in Chinese].

Whelpton, P.K. (1945). Effect of increased birth rate on future population. *American Journal of Public Health and the Nations Health*, 35(4):326-333.

https://doi.org/10.2105/ajph.35.4.326

Whelpton, P.K. (1954). *Cohort Fertility: Native White Women in the United States*. Princeton: Princeton University Press.

Yang, K., Chen, S., & Wei, J. (2000). *The costs and benefits of China's family planning*. Beijing, China: People's Press [Article in Chinese].

Yang, S., Gu, B., & Xiao, Z. (1991). Trends of fertility in China. *Chinese Journal of Population Science*, 5:20-27 [Article in Chinese].

Yao, X. (1995). *Fertility data of China*. Beijing, China: China Population Publishing House [Article in Chinese].

Yu, J., Zhou, Y., & Xie, Y. (2021). Macro-level social factors and ideal number of children in China. *Population Research*, 45(6):45-61 [Article in Chinese].

Yuan, X., & Gao, Y. (2017). The adjustment of family policy and the change of new-born population. *Population and Family Planning*, 6:20-23 [Article in Chinese].

Zeman, K., Beaujouan, É., Brzozowska, Z., & Sobotka, T. (2018). Cohort fertility decline in low fertility countries: Decomposition using parity progression ratios. *Demographic Research*, 38:651-690.

Zeng, Y., & Hesketh, T. (2016). The effects of China's universal two-child policy. *The Lancet*, 388(10054):1930-1938.

Zha, R., & Ji, Y. (1984). Analysis of Chinese women's fertility. *Population Research*, 6:11-20 [Article in chinese].

Zhai, Z., Chen, J., & Li, L. (2015). New developments and future trends of births in China. *Population Research*, 39(2):48-56 [Article in Chinese].

Zhang, C., Liu, H., & Wang, X. (2016). Trends in the second birth interval in China since 1970. *Population Research*, 40(1):69-86 [Article in Chinese].

Zhang, E., & Su, R. (1995). Analysis of fertility trends in China in the 1990s. *Chinese Journal of Population Science*, 24(4):51-58 [Article in Chinese].

Zhang, W., & Cao, X. (2006). Family planning during Economic Reform era. In: Zhao, Z., & Guo, F., (eds.). *Transition and Challenge: China's Population at the Beginning of the 21st Century*. Oxford: Oxford University Press.

Zhang, X., Guo, F., & Zhai, Z. (2019). China's demographic future under the new two-child policy. *Population Research and Policy Review*, 38(4):537-563.

https://doi.org/10.1007/s11113-019-09519-0

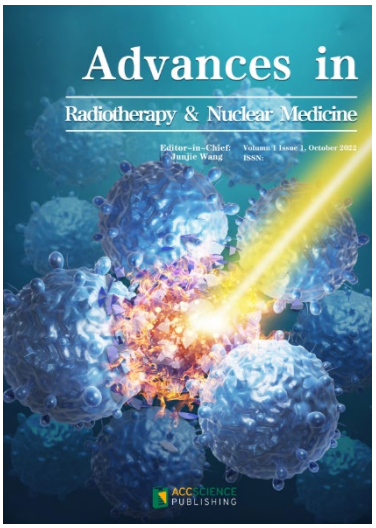
Zhao, X. (1991). How many births have been averted in China in the past 20 years? *In presentation at the International Seminar on Fertility and Contraception of China, Beijing, China*,8(5):10-11 [Article in Chinese].

Zhao, Z. (2015). Closing a sociodemographic chapter of Chinese history. *Population and Development Review*, 41(4):681-686.

https://doi.org/10.1111/j.1728-4457.2015.00090.x

Zhao, Z., & Chen, W. (2011). China's far below replacement fertility and its long-term impact: Comments on the preliminary results of the 2010 census. *Demographic Research*, 25:819-836.

Our Journals



Advances in Radiotherapy & Nuclear Medicine (ARNM) is a peer-reviewed and open-access journal that aims to publish and disseminate novel research in the breadth of neurology and neuroscience.

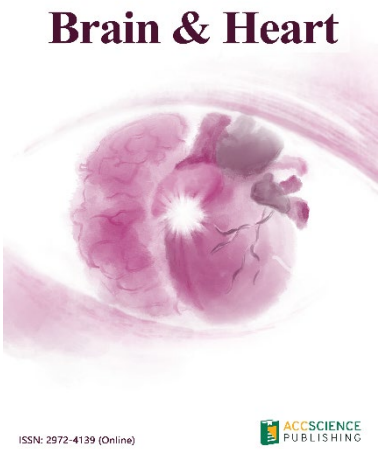
ARNM covers subject areas, including but not limited to the following:

- Conventional Radiotherapy (CR)
- Stereotactic Body Radiation Therapy (SBRT)
- Brachytherapy (BT)
- Boron Neutron Capture Therapy (BNCT)
- Particle Therapy (proton and heavy ions) (PT)
- Targeted and Immunotherapy (TI)
- Combined Modality Therapy (Heat therapy, electric field therapy, nursing, technology) (CMT)
- Radiation Biology (RB)
- Radiation Physics (RP)
- Innovative Radiation Technology (IRT)
- Positron Emission Tomography (PET)
- Radiopharmaceuticals and Radio-tracer (RR)
- Molecular Imaging and Radionuclide Therapy (MI & RT)
- Single-photon Emission Computed Tomography (SPETCT)

Brain & Heart focuses on neurocardiology, a neurology and cardiology-based interdisciplinary subject that studies the circulatory mechanism of the human body, as well as the mechanisms of the interplay between the cardiovascular system and the nervous system. The journal's scope includes:

Clinical and basic research on diseases related to the circulatory and nervous systems, such as: orthostatic dizziness, orthostatic hypotension, autonomic dysfunction, and the relationship between the autonomic nervous system and the circulatory function in cerebral degeneration;

Heart-brain research on patients with syncope, autonomic dysfunction, cryptogenic stroke, and stroke with atrial fibrillation; research on the relationship between structural heart diseases and nervous system diseases, the correlation between cardiac electrophysiology and abnormal organizational structures and the pathogenesis of stroke, as well as new ways of diagnosis, treatment and prevention of unexplained stroke.



Start a new journal

Write to us via email if you are interested to start a new journal with AccScience Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@accscience.com

https://accscience.com/journal/IJPS

Contact
www.accscience.com
9 Raffles Place, Republic Plaza 1 #06-00 Singapore 048619
Email: editorial@accscience.com
Phone: +65 8182 1586