Using primary and routinely collected data to determine prevalence and patterns of multimorbidity in rural China: a representative cross-sectional study of 6474 Chinese adults

Xinyi Zhang,[a,s] Tingzhuo Liu,[a,s] Zhifang Li,[b] Jiajuan Yang,[c] Huinan Hou,[d] Tianyou Hao,[e] Pei Zhang,[c] Chi Hu,[c] Mingjia Bao,[f] Pengpeng Ye,[g] Shangzhi Xiong,[h,i] Wei Tian,[a] Guangcan Yan,[a] Jing Zhang,[a] Yue Wang,[a] Wei Jiang,[g] Anqi Ge,[a] Yonghui Pan,[j] Devarsetty Praveen,[h,k,l] David Peiris,[h] Xiaoqi Feng,[m] Ding Ding,[n] Lijing L. Yan,[i] Xiaolin Xu,[o] Hanbin Zhang,[p,q] Yongchen Wang,[r,***] Wenjing Tian,[a,**] and Maoyi Tian[a,r,*]

[a]School of Public Health, Harbin Medical University, Harbin, China
[b]School of Public Health, Changzhi Medical College, Changzhi, China
[c]Yichang Centre for Disease Control and Prevention, Yichang, China
[d]Jiamusi Centre for Disease Control and Prevention, Jiamusi, China
[e]Heping Hospital Affiliated to Changzhi Medical College, Changzhi, China
[f]Heilongjiang Provincial Centre for Disease Control and Prevention, Harbin, China
[g]National Centre for Non-communicable Disease Control and Prevention, Chinese Centre for Diseases Control and Prevention, Beijing, China
[h]The George Institute for Global Health, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia
[i]Global Health Research Centre, Duke Kunshan University, Kunshan, China
[j]Division of Neurology, The First Affiliated Hospital of Harbin Medical University, Harbin, China
[k]The George Institute for Global Health, India, Hyderabad, India
[l]Prasanna School of Public Health, Manipal Academy of Higher Education, Manipal, India
[m]School of Population Health, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia
[n]Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia
[o]School of Public Health, Zhejiang University, Hangzhou, China
[p]European Centre for Environment and Human Health, University of Exeter, United Kingdom
[q]Environmental Research Group, MRC Centre for Environment and Health, Faculty of Medicine, Imperial College London, UK
[r]Division of General Practice, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

The Lancet Regional Health - Western Pacific 2025;54: 101272

Published Online xxx

https://doi.org/10.1016/j.lanwpc.2024.101272

Summary

Background In China, rising chronic diseases has coincided with the increasing burden of multimorbidity, particularly for vulnerable populations. Limited primary data are available to understand the prevalence and patterns of multimorbidity, especially in resource-limited rural areas. This study aims to conduct robust evaluations of the prevalence and patterns of multimorbidity among rural adults in China, and to compare the differences in prevalence and patterns when using primary data alone versus in combination with routinely collected data.

Methods This cross-sectional study was conducted in three provinces in China, with two counties per province and 40 villages per county, resulted in a total of 240 villages. Participants were randomly selected and stratified by sex and age in each village. Multimorbidity, defined as the coexistence of two or more diseases in same individual, was assessed through data collection involving primary data (face-to-face questionnaire, physical examination and fasting blood sample collection) and routinely collected data (health insurance claims, hospital electronic records and infectious disease surveillance system). Multimorbidity prevalence and patterns were compared based on 1) primary data alone and 2) primary data complemented by routinely collected data.

Findings A total of 6474 individuals participated in this study (50.9% women, mean age 57.1). Combining routinely collected data with the primary data increased the prevalence of all single disease conditions. Multimorbidity prevalence rose from 35.7% with primary data alone to 44.4% with the addition of routinely collected data. The top three dyad multimorbidity patterns (hypertension with heart disease, stroke, or chronic digestive diseases) remained consistent between the two ascertainment methods, while triad pattern rankings had a substantial shift. According to

*Corresponding author. School of Public Health, Harbin Medical University, 157 Baojian Road, Harbin, China.
**Corresponding author. School of Public Health, Harbin Medical University, 157 Baojian Road, Harbin, China.
***Corresponding author. Division of General Practice, The Second Affiliated Hospital of Harbin Medical University, Harbin Medical University, 246 Xuefu Road, Harbin, China.
E-mail addresses: maoyi.tian@hrbmu.edu.cn (M. Tian), tianwenjing@ems.hrbmu.edu.cn (W. Tian), yongchenwang@163.com (Y. Wang).
[s]Joint first author.

blood pressure measurements, over 40% of participants had elevated blood pressure and may have undiagnosed hypertension. Over 20% may have undiagnosed mental health disorders base on the questionnaires, and nearly 10% with undiagnosed chronic kidney disease as indicated by blood testing.

Interpretation The utilisation of primary data combined with routinely collected data provided a robust estimation of multimorbidity burden in three rural regions in China. Yet, the prevalence may still have been underestimated due to inaccuracies in self-reported data and underdiagnosis of diseases. Future research should incorporate routinely collected data for more robust epidemiological evidence of multimorbidity.

Funding Harbin Medical University Leading Talent Grant (31021220002) and National Natural Science Foundation of China (72074065 and 72474063).

Keywords: Multimorbidity; Routinely collected data; China

Research in context

Evidence before this study

We conducted a literature search to identify studies evaluating the prevalence and patterns of multimorbidity in China. We searched PubMed using the terms "multimorbidity" "multiple chronic conditions" and "China" covering the period from database inception to 1 September 2024, with no language restrictions. In China, research on multimorbidity has been expanding, particularly regarding its prevalence, patterns, and risk factors. We identified several studies exploring multimorbidity in various Chinese populations. Two systematic reviews revealed similar multimorbidity patterns but had different prevalence rates. These variations were primarily due to differing definitions of multimorbidity and the reliance on secondary data analysis, with limited multimorbidity studies using primary data collection. Additionally, the frequent use of self-reported disease diagnosis raises concerns about report bias. To date, no studies in China have robustly combined primary data and routinely collected data to estimate the prevalence and patterns of multimorbidity, particularly among rural populations.

Added value of this study

This study estimates the prevalence and patterns of multimorbidity in rural China by combining self-reported data with routinely collected data. It highlights the potential of routinely collected data to enhance understanding of the status of multimorbidity, by identifying common disease patterns such as hypertension combined with other conditions, as well as non-traditional multimorbidity patterns such as chronic digestive system diseases combined with other conditions. The physical examination data and pathology tests used in the study also revealed that many participants still had unscreened or undiagnosed conditions, leading to a major gap in the detection and identification of multimorbidity cases.

Implications of all the available evidence

The integration of primary data with routinely collected data enhances the estimation of multimorbidity prevalence and patterns. Findings of this study align with the existing literature, indicating that combining multiple data sources can mitigate the report bias commonly present in multimorbidity studies, thereby improving the accuracy of prevalence estimates. Future research should focus on improving data collection practices and ensuring the integration of high-quality data to provide a more accurate assessment of multimorbidity prevalence and to further enhance the effectiveness of public health interventions.

Introduction

The global aging population has led to an increased prevalence of chronic diseases, resulting in a rise in individuals with multiple co-existing chronic conditions, often referred to as "multimorbidity".[1] Multimorbidity is commonly defined as the co-occurrence of two or more chronic conditions within an individual, without any implicit ordering or necessary causal relationship between the conditions.[2,3] In China, one in four adults has multimorbidity based on limited estimates, creating a major public health concern with a projected heavy burden on the healthcare systems.[4–6] People with multimorbidity are more prone to adverse health outcomes, including loss of physical function, disability, and premature death.[7,8] But the existing clinical practices and primary healthcare services focus on single diseases, leading to inappropriate care and potentially harmful treatment due to interactions among the coexisting diseases.[9,10] Robust evidence on the prevalence of multimorbidity and the patterns of the disease combinations is essential for improving the treatment and management of people with multimorbidity.[11,12]

In China, an increasing number of studies have estimated the prevalence and patterns of multimorbidity.[4] The current multimorbidity studies in China are mainly based on secondary data analysis from existing cohort studies, with a lack of primary data collection focused on multimorbidity.[13] Moreover, the studies with primary data collection mainly depend on self-reported data which may affect the estimation of the disease prevalence due to report bias.[13,14] Patients with mental health disorders also become a growing public health challenge globally, but few studies examined the multimorbidity with inclusion of mental health disorders in China.[15] In addition, studies addressing multimorbidity in rural areas with constrained resources, are needed, whose results may reduce health inequity in understanding, treating, and managing multimorbidity.[16,17]

Routinely collected data refer to those that have been collected from various data sources, which are not originally intended for addressing specific research questions, such as administrative records, health insurance claims data, hospital electronic medical records, public health datasets and disease surveillance systems.[18–20] Over the past decade, routinely collected data were widely used in research involving observational studies and large-scale randomised controlled trial in China.[21–23] Such data are increasingly viewed as valuable resources for clinical research, epidemiological studies, and health system research, offering a cost-effective alternative to resource-intensive primary data collection.[24,25] In this study, we aim to comprehensively assess the prevalence and patterns of multimorbidity including non-communicable diseases, chronic infectious diseases, and mental health conditions among rural Chinese adults using the combination of primary data and routinely collected data.

Methods

This study was a large-scale cross-sectional survey conducted in rural areas of three Chinese provinces. This study was coordinated by Harbin Medical University in collaboration with Changzhi Medical College, Jiamusi Centre for Disease Control and Prevention, and Yichang Centre for Disease Control and Prevention, and received approval from the ethics committee at Harbin Medical University (HMUIRB2022005PRE). Prior to the enrolment, all participants completed written informed consent. The study was registered in the Chinese Clinical Trial Registry (ChiCTR2300069860).

Settings

This study was conducted in 6 counties of 3 provinces (Jiaoqu county and Huachuan county from Heilongjiang province, Shangdang county and Huguan county from Shanxi province, and Yidu county and Zhijiang county from Hubei province. Three study provinces were selected based on their economic development levels and geographical locations (Supplementary Figure S1) to ensure representative rural healthcare system. Within each province, the rural healthcare systems are relatively homogeneous. Counties and villages were therefore selected primarily based on their proximity to the research team (prior collaboration and willingness to participate) to maximise the feasibility of implementing the study. In each county, 40 villages were selected with a total of 240 villages participating in this study. The villages in China serve as a fundamental organisational unit for its rural population, while in urban areas, this unit refers to a community. Each village in China has its boundaries, designated head and a cluster of population. But the size of the population or the land areas of each village can vary.

Sample size

We estimated that a sample size of 6000 participants from 240 villages would provide the study with a 90% power (a two-sided alpha = 0.05) to detect a multimorbidity prevalence of at least 30.9% in four different groups stratified by sex (male and female) and age (30–59, 60 and above), accounting for an estimated intraclass correlation of 0.02, equivalent to a design effect of 1.5 and a 10% non-response rate.[13]

Participants and sampling

Individuals aged 30 years and above and permanently live in the selected villages were eligible to participate into the study. Participants who were pregnant, had difficulties communicating with the study team, required intensive medical attention, or unable to provide informed consent were excluded from the study. Recruitment of the participants were done by making a random selection from a list of age eligible (aged 30 and above) village members stratified by sex (female and male) and age (aged 30–59 and 60 and above) to ensure proportional representation. Village doctors, the village level primary health care providers, were instructed to invite participants sequentially from the random list to the survey. For any participant who was not in the village on the day of the survey, the next participant on the random list was invited (Supplementary Figure S2).

Definition of multimorbidity

In this study, we defined multimorbidity as the co-existing of two or more chronic diseases or health conditions.[26–28] We included 21 diseases to assess multimorbidity, comprising 19 non-communicable diseases (including hypertension, chronic digestive system diseases, heart disease, arthritis, stroke, chronic lung disease, chronic back pain, oral health disorders, diabetes, eye disease, chronic kidney disease, thyroid disease, ear disease, cancer, osteoporosis, anxiety, depression, epilepsy and dementia) and two communicable diseases (Tuberculosis and hepatitis B).

Primary data collection

Data collection was conducted between April 14th and August 29th 2023. Primary data collection consisted of a face-to-face questionnaire and clinical measurements (physical examination and fasting blood sample collection). First, the face-to-face questionnaire was adapted from a modified version of the Multimorbidity Assessment Questionnaire for Primary Care,[29] which consists of a list of questions to collect participants' self-reported history of 20 diseases (hepatitis B was not included in the questionnaire). In addition, a nine-item Patient Health Questionnaire (PHQ-9) and a seven-item Generalised Anxiety Disorder Scale (GAD-7) were used to screen and measure severity of depression and anxiety respectively.[30–32] The Chinese versions of these questionnaires have demonstrated good reliability and validity in assessing depression and anxiety in various populations.[33–36] Population at high risk of depression was defined as an assessment score ≥5 from PHQ-9, while population at high risk of anxiety was defined as an assessment score ≥5 from GAD-7. Second, the physical examination included the measurement of height, weight, waist circumference of the participants, using uniformed scales and tapes, and twice blood pressure (with a minimum minimal interval of 5 min between the measurements), using an electronic blood pressure monitor (Omron J710), were measured. Standardised training and operating procedures were implemented to ensure consistent physical examination across all sites. All equipment were centrally procured by the study team and calibrated before field data collection was commenced in each study province. Body mass index (BMI) was calculated as weight (kilogram) divided by the square of height (meter). The blood pressure was calculated as the average of two measurements. Participants with an increased risk of high blood pressure were defined as those with a systolic blood pressure (SBP) ≥130 mmHg and/or a diastolic blood pressure (DBP) ≥80 mmHg.[37,38] Lastly, the fasting blood samples were collected after an overnight fast and centrifuged within 2 h of collection. Blood serum sample tests included fasting glucose and creatinine. The serum creatinine was used to identify participants at risk of chronic kidney disease (CKD), while fasting blood glucose was tested to screen participants at risk of diabetes.[39] All blood samples were sent to a central laboratory located in Heping Hospital Affiliated to Changzhi Medical College in Changzhi, Shanxi Province for testing to ensure consistency in both the testing process and standards. Participants with an increased risk of diabetes were defined as having a fasting glucose level ≥7.0 mmol/L, while participants with an increased risk of CKD were defined as having an estimated glomerular filtration rate (eGFR) <90 mL/min/1.73 m^2 (CKD stage two or above).[40] Overall, for conditions such as anxiety and depression, hypertension, diabetes, and CKD, if a participant had measurements that exceeded the threshold for these conditions, but had not yet received a respective diagnosis, the participant was considered to be at risk.

Routinely collected data collection

The routinely collected data were extracted from three data sources, including health insurance claims from the health insurance bureau, hospital electronic health records from local medical institutions, and infectious disease surveillance system from the local Centers for Disease Control and Prevention. The International Classification of Diseases, Tenth Revision (ICD-10) coding system is used in all routinely collected data, after a respective standard diagnosis for all disease conditions. Despite fragmented health information systems,[41] participants' unique national IDs were used to link with the data sources. All disease information extracted from routinely collected data was used to complement the missing reports due to report bias in the face-to-face questionnaire data collection, except for hepatitis B, which was only assessed through routinely collected data. In addition, all extracted disease conditions from the routinely collected data were mapped to the 21 disease conditions collected in the face-to-face questionnaire (Supplementary Table S1). Supplementary Figure S3 illustrated the data collection and data outcomes for primary data and routinely collected data.

Statistical analysis

Descriptive analysis was conducted to present the participants' characteristics, anthropometric and blood serum testing results, anxiety and depression assessment results and the prevalence and patterns of multimorbidity prevalence, in all participants and by provinces. Continuous variables were presented as mean and standard deviation (SD), while categorical variables were presented as percentages. The prevalence was calculated based on primary data only and the combination of primary data and routinely collected data. Dyad multimorbidity patterns referred to the co-existence of two disease conditions, whereas triad patterns were the co-existence of three disease conditions. All dyad patterns and the top ten triad multimorbidity patterns were presented, and their prevalence were calculated. We further compared the differences in prevalence and patterns between those using primary data and the combination of primary data and routinely collected data. The heatmap for the dyad multimorbidity patterns were produced using Excel (Microsoft Excel 2024). Statistical analyses were conducted in Stata 15.0 software (Stata Statistical Software: College Station, TX: STATA Corp LP). All analyses were based on the 21 disease conditions, listed on the face-to-face questionnaire, to provide a consistent and standardised approach to assess multimorbidity.

Role of the funding source

The funder played no role in the study design, data collection, data analysis, data interpretation, the writing of the report, or the decision to submit the paper for publication.

Results

Characteristics of study participants

A total of 6474 participants were recruited from 240 villages across three provinces (2208 from Shanxi, 2095 from Hubei and 2171 from Heilongjiang) (Table 1). The mean age of all participants was 57.1 years, with 38.9% (2516/6474) aged 60 or above and 50.9% (3295/6474) were females. 13.3% (860/6474) participants had an educational level of senior high school or above, and 89.5% (5793/6474) were married. Only one-third of the participants had an annual household income more than 35,000 Chinese yuan (approximately 5000 USD). Almost all participants (96.4%) were covered by urban and rural resident basic medical insurance. A total of 6346 participants (98%) provided blood samples for testing, the missing samples were primarily due to participants not fasting at the time of blood collection. The average blood pressure was 131.9 mm Hg for systolic and 84.1 mm Hg for diastolic, with 50.3% (3258/6474)had a systolic blood pressure over 130 mm Hg and 63.5% (4112/6474) with a diastolic blood pressure over 80 mm Hg. The mean fasting blood glucose level was 5.8 mmol/L, with 10.4% (659/6474) of participants exhibiting levels above 7 mmol/L. The mean BMI was 24.8 kg/m^2 and the mean estimated glomerular filtration rate (eGFR) was 118.5 mL/min/1.73 m^2, with about 10% of the participants below 90 mL/min/1.73 m^2. More than 19% of participants were considered to have mild or more severe depression, and about 15% experienced more than mild anxiety. The overall prevalence of multimorbidity was 35.7% (2314/6474) based on primary self-reported data, which increased to 44.4% (2876/6474) in combination with routinely collected data (Table 1). The prevalence of multimorbidity by gender, age and other subgroups is provided in Supplementary Figure S4.

The prevalence of single conditions

Based on primary self-reported data, the five most prevalent conditions were hypertension (34%), chronic digestive system diseases (14.2%), heart disease (14.1%), stroke (10.4%), and arthritis (10.0%). Diabetes, oral health disorders, chronic lung diseases, chronic back pain, and eye diseases, each affected 5%–10% of the study population. In combination with routinely collected data, the most significant absolute increases in prevalence were observed in chronic lung diseases (6.6% increase), chronic digestive system diseases (6.5% increase), stroke (5.4% increase), chronic back pain (5.2% increase), and oral health disorders (3.9% increase). Additionally, hypertension, eye diseases, and chronic kidney diseases had an increased prevalence of between 1% and 2% (Fig. 1).

Moreover, over 75% of the participants were identified as having an increased risk for hypertension (SBP ≥130 mmHg and/or DBP ≥80 mmHg), yet only 36.0% (2330/6474) of these individuals were diagnosed, either self-reported or through the routinely collected data. Additionally, 19.7% (1277/6474) and 16.1% (1044/6474) of the participants were at a risk for depression and anxiety (severity level of mild or above) based on PHQ-9 and GAD-7, respectively, although diagnosed depression and anxiety were below 1%. Blood testing results further found that 14.7% (952/6474) of the participants were at risk for diabetes (fasting glucose level ≥7.0 mmol/L), and 13.8% (891/6474) were at risk for CKD (eGFR <90 mL/min/1.73 m^2). Notably, 9.2% (597/6474) of those with eGFR <90 mL/min/1.73 m^2 had not received a clinical diagnosis, and 4.6% (299/6474) of those at risk for diabetes mellitus remain undiagnosed (Fig. 2).

Dyad multimorbidity patterns

The prevalence of dyad multimorbidity patterns using primary self-reported data alone or in combination with routinely collected data was shown in Fig. 3. With self-reported data alone, the five most prevalent dyad patterns were hypertension and heart disease (8.1%), hypertension and stroke (6.5%), hypertension and chronic digestive system diseases (5.5%), hypertension and diabetes (5.4%), and hypertension and arthritis (4.8%). In combination with routinely collected data, the top five dyad patterns remained unchanged with slight change in the ranking. The prevalence of top five patterns increased, hypertension and heart disease increased from 8.1% to 9.9%, hypertension and chronic digestive system diseases increased from 5.5 to 8.5%, and hypertension and stroke increased from 6.5% to 8.0%. Hypertension and arthritis (from 4.8% to 6.6%) tied for the fourth place, followed by hypertension with diabetes slightly increase from 5.4% to 6.1%. The prevalence of chronic digestive system diseases combined with chronic lung disease was the same as that of hypertension combined with chronic lung disease, with the former increasing from 1.9% to 5.9% and the latter rising from 2.9% to 5.9%. Moreover, significant increases were observed in several dyad patterns involving chronic digestive system diseases, including combinations with heart disease (from 3.6% to 5.3%), arthritis (from 3.1% to 5.3%), and chronic back pain (from 1.8% to 5.3%).

Triad multimorbidity patterns

Using primary self-reported data alone, hypertension was included in all leading triad multimorbidity patterns. In combination with routinely collected data, the prevalence of hypertension and heart disease and stroke

	All (n = 6474)	Shanxi (n = 2208)	Hubei (n = 2095)	Heilongjiang (n = 2171)
Female, n (%)	3295 (50.9)	1152 (52.2)	1064 (50.8)	1079 (49.7)
Age (y), mean (SD)	57.1 (10.5)	57.7 (10.5)	57.0 (10.4)	56.6 (10.5)
Age range, n (%)				
30–59	3958 (61.1)	1283 (58.1)	1288 (61.5)	1387 (63.9)
60+	2516 (38.9)	925 (41.9)	807 (38.5)	784 (36.1)
Ethnic minority, n (%)	290 (4.5)	76 (3.5)	151 (7.2)	63 (2.9)
Level of education, n (%)				
Primary school and lower	2659 (41.1)	1014 (45.9)	614 (29.3)	1031 (47.5)
Junior high school	2955 (45.6)	966 (43.8)	1068 (51.0)	921 (42.4)
Senior high school and above	860 (13.3)	228 (10.3)	413 (19.7)	219 (10.1)
Marital status, n (%)				
Never married	159 (2.5)	60 (2.7)	58 (2.8)	41 (1.9)
Married	5793 (89.5)	1965 (89.0)	1922 (91.7)	1906 (87.8)
Divorced	116 (1.8)	19 (0.9)	25 (1.2)	72 (3.3)
Widowed	406 (6.3)	164 (7.4)	90 (4.3)	152 (7.0)
Annual income, Chinese Yuan, n (%)				
<9999	1568 (24.2)	622 (28.2)	320 (15.3)	626 (28.8)
10,000–19,999	1422 (22.0)	475 (21.5)	340 (16.2)	607 (28.0)
20,000–34,999	1446 (22.3)	514 (23.3)	502 (24.0)	430 (19.8)
≥35,000	2038 (31.5)	597 (27.0)	933 (44.5)	508 (23.4)
Medical insurance, n (%)				
URBMI	6244 (96.4)	2147 (97.2)	1996 (95.3)	2101 (96.8)
UEBMI	158 (2.4)	37 (1.7)	88 (4.2)	33 (1.5)
Commercial Health Insurance	292 (4.5)	4 (0.2)	109 (5.2)	179 (8.2)
No medical insurance	74 (1.1)	23 (1.0)	10 (0.5)	41 (1.9)
Systolic blood pressure (mm Hg), mean (SD)	131.9 (19.8)	134.5 (20.0)	128.2 (18.2)	132.8 (20.4)
<130 mm Hg, n (%)	3216 (49.7)	982 (44.5)	1177 (56.2)	1057 (48.7)
≥130 mm Hg, n (%)	3258 (50.3)	1226 (55.5)	918 (43.8)	1114 (51.3)
Diastolic blood pressure (mm Hg), mean (SD)	84.1 (11.7)	83.8 (11.8)	81.6 (10.8)	86.9 (11.7)
<80 mm Hg, n (%)	2362 (36.5)	834 (37.8)	930 (44.4)	598 (27.5)
≥80 mm Hg, n (%)	4112 (63.5)	1374 (62.2)	1165 (55.6)	1573 (72.5)
Body mass index (kg/m^2), mean (SD)	24.8 (3.7)	25.3 (3.6)	23.9 (3.6)	25.0 (3.8)
Fasting blood glucose (mmol/L)[a], mean (SD)	5.8 (1.8)	5.6 (1.7)	5.6 (1.6)	6.1 (2.0)
<7 mmol/L, n (%)	5687 (89.6)	1910 (90.5)	1918 (91.8)	1859 (86.6)
≥7 mmol/L, n (%)	659 (10.4)	201 (9.5)	171 (8.2)	287 (13.4)
eGFR (mL/min/1.73 m^2)[a], mean (SD)	118.5 (31.3)	117.6 (34.8)	113.7 (28.3)	124.2 (29.4)
≥90 mL/min/1.73 m^2, n (%)	5672 (89.4)	1952 (92.5)	1768 (84.6)	1952 (91.0)
60–89 mL/min/1.73 m^2, n (%)	606 (9.5)	146 (6.9)	281 (13.5)	179 (8.3)
<60 mL/min/1.73 m^2, n (%)	68 (1.1)	13 (0.6)	40 (1.9)	15 (0.7)
PHQ-9, n (%)				
None depression	5211 (80.5)	1746 (79.1)	1915 (91.4)	1550 (71.4)
Mild depression	970 (15.0)	367 (16.6)	157 (7.5)	446 (20.5)
Moderate depression	201 (3.1)	77 (3.5)	17 (0.8)	107 (4.9)
Severe depression	92 (1.4)	18 (0.8)	6 (0.3)	68 (3.1)
GAD-7, n (%)				
None anxiety	5458 (84.3)	1798 (81.4)	1974 (94.2)	1686 (77.7)
Mild anxiety	870 (13.4)	355 (16.1)	116 (5.5)	399 (18.4)
Moderate anxiety	107 (1.7)	50 (2.3)	3 (0.1)	54 (2.5)
Severe anxiety	39 (0.6)	5 (0.2)	2 (0.1)	32 (1.5)
Prevalence of multimorbidity, priamry data only, n (%)				
≥2 disease conditions	2314 (35.7)	773 (35.0)	681 (32.5)	860 (39.6)
≥3 disease conditions	1192 (18.4)	357 (16.2)	327 (15.6)	508 (23.4)
≥4 disease conditions	606 (9.4)	159 (7.2)	152 (7.3)	295 (13.6)

(Table 1 continues on next page)

	All (n = 6474)	Shanxi (n = 2208)	Hubei (n = 2095)	Heilongjiang (n = 2171)
(Continued from previous page)				
Prevalence of multimorbidity, priamry data with routinely collected data, n (%)				
≥2 disease conditions	2876 (44.4)	818 (37.0)	1146 (54.7)	912 (42.0)
≥3 disease conditions	1696 (26.2)	393 (17.8)	753 (35.9)	550 (25.3)
≥4 disease conditions	1015 (15.7)	175 (7.9)	511 (24.4)	329 (15.2)

URBMI, Urban and Rural Resident Basic Medical Insurance; UEBMI, Urban Employee Basic Medical Insurance. [a]The total number of blood samples for testing was 6346 (2111 from Shanxi, 2089 from Hubei and 2146 from Heilongjiang).

Table 1: **Characteristics and anthropometric measurement results of all study participants and by province.**

(from 2.41% to 3.38%), and hypertension and chronic digestive system diseases and heart disease (from 2.09% to 3.17%) remained the leading two triad patterns, whereas the prevalence of four patterns increased more than two folds, including chronic digestive system diseases and chronic lung disease and chronic back pain (from 0.34% to 2.55%), chronic digestive system diseases and arthritis and chronic back pain (from 0.73% to 2.12%), hypertension and chronic digestive system diseases and chronic lung disease (from 0.96% to 2.97%) and hypertension and chronic digestive system diseases and chronic back pain (from 0.99% to 2.64%) (Fig. 4).

Discussion

This cross-sectional study investigated the prevalence and patterns of multimorbidity in rural China, using primary data and routinely collected data. Our findings revealed large disparities in multimorbidity prevalence and patterns by applying various data sources, highlighting potential gaps in the current understanding multimorbidity, namely, the neglection of mental health disorders in prior multimorbidity assessment, underdiagnosis of chronic kidney disease, hypertension and diabetes, and the underestimation in disease prevalence solely based on self-reported data. The study addressed the need and opportunities to integrate routinely collected data in future assessment of multimorbidity.

The prevalence of multimorbidity in our study, using primary self-reported data alone, was comparable with a recent systematic review, which found a prevalence of 36.1% among the Chinese population (IQR 19.6, 48.8).[4] But in combination with routinely collected data, the prevalence increased substantially. This might be explained by the following four reasons. Firstly, the review included younger adults (aged 18–30), the lower burden of multimorbidity in this age group may attenuated the overall prevalence. Secondly, the prevalence of multimorbidity depends on its definition. For example, the number or the type of disease conditions included in the assessment can influence the measurement of multimorbidity.[42] Thirdly, the approaches to collect data, such as the anthropometric measurement, pathology test, blood pressure measurement, screening questionnaire of

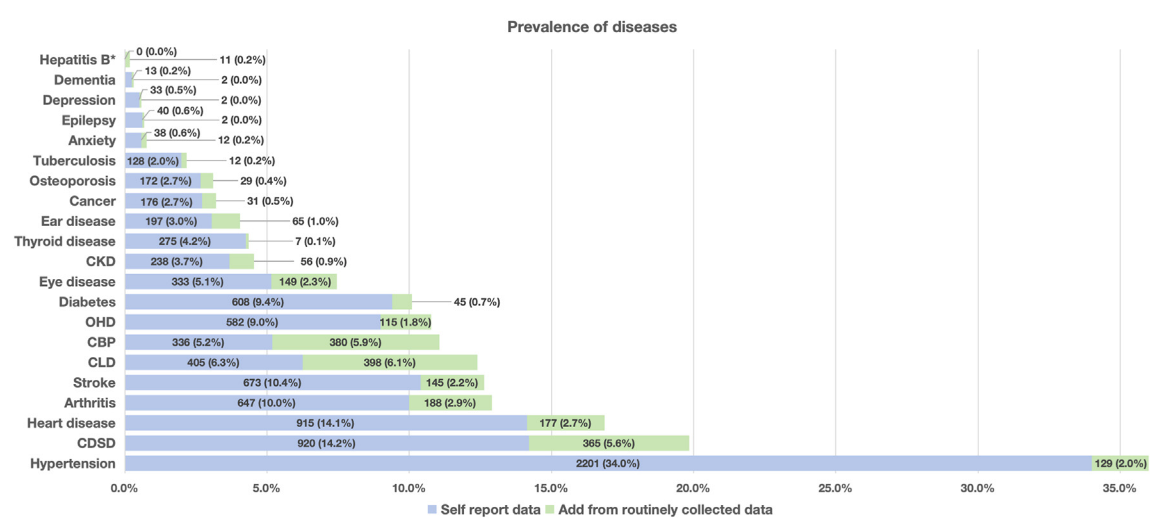

Fig. 1: Prevalence of single disease conditions based on primary self-reported data and routinely collected data. CDSD, Chronic digestive system diseases; OHD, Oral health disorders; CLD, Chronic lung diseases; CBP, Chronic back pain; CKD, Chronic kidney diseases; *: Primary self-reported data not available.

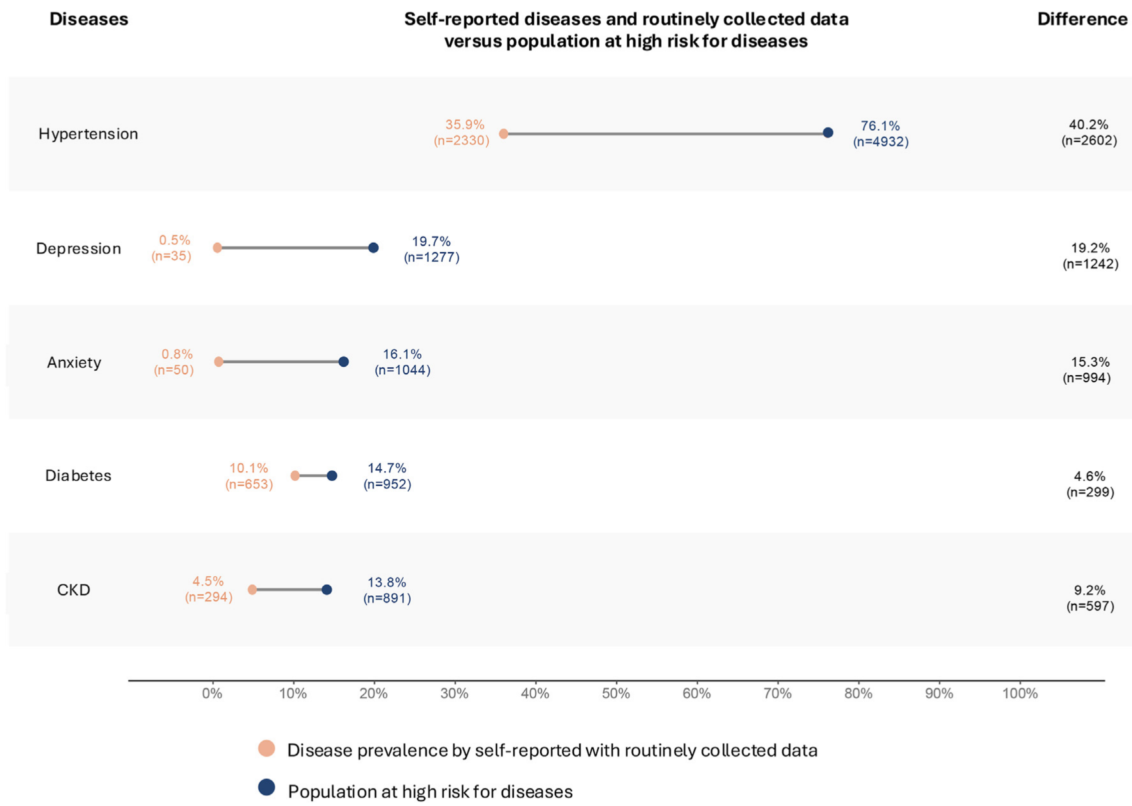

Fig. 2: Differences in the prevalence of selected disease conditions between identified using participants' self-reported data in combination with routinely collected data and identified using clinical measurements and assessment tool.

disease conditions, can further increase the prevalence of multimorbidity due to underdiagnosis.[43] Lastly, the use of multiple data sources provides a less biased assessment of disease conditions, and enhancing the robustness of multimorbidity assessment. For example, a study in Wales involving 2.3 million participants, showed that relying on a single data source may underestimate the prevalence of multimorbidity.[44] Unlike previous findings,[13,43,45] in our study, hypertension and chronic digestive system diseases became an emerging combination in dyad or triad multimorbidity patterns, after supplemented with the routinely collected data. This might be due to increased high prevalence of hypertension and chronic digestive system diseases in combination with routinely collected data. Furthermore, our study identified disparities in multimorbidity prevalence across the three provinces, which may be attributed to variations in socioeconomic status. Evidence from a prior study in China demonstrated that while the prevalence of multimorbidity increased in higher socioeconomic groups from 2011 to 2015, multimorbidity was more prevalent in economically disadvantaged regions.[46]

Most existing research on multimorbidity heavily relies on self-reported data. In this study, we found that the prevalence of multimorbidity significantly increased after supplementing with routinely collected data. Previous studies in China and Australia employing this approach also observed increases in disease prevalence.[47,48] In addition, we found that the supplementation with routinely collected data had a more pronounced impact on the prevalence among multimorbidity patients with three or more co-existing diseases. This suggests the difficulty in accurately reporting health status increases as the number of co-existing conditions rises, resulting in a higher report bias. Moreover, patients may fail to report conditions they have become accustomed to or those that are asymptomatic. Instead, patients tend to focus on conditions requiring active or ongoing treatment.[49,50] We also found that the shift in multimorbidity patterns was mainly caused by the change in the prevalence of a few disease conditions before and after incorporating the routinely collected data. High prevalent disease conditions are always significant contributors among common multimorbidity patterns.[27,28] As reflected in this study, chronic digestive system diseases, stroke, oral health disorders, chronic lung disease, and other major chronic diseases or conditions were largely underreported in the face-to-face questionnaire. This is likely due to a lack of awareness about these diseases or conditions, leading

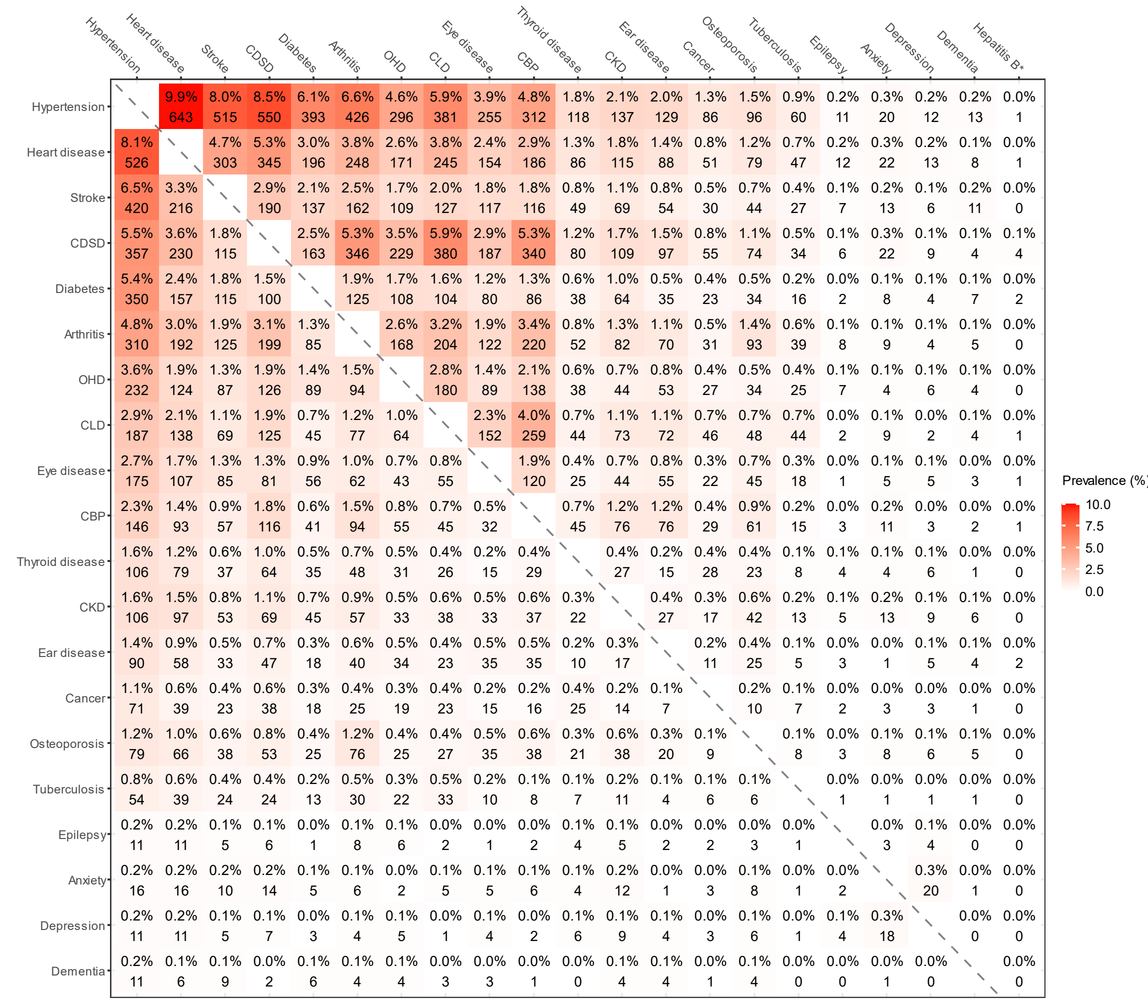

Fig. 3: Dyad multimorbidity patterns and prevalence with primary self-reported data alone and in combination with routinely collected data. CDSD, Chronic digestive system diseases; OHD, Oral health disorders; CLD, Chronic lung disease; CBP, Chronic back pain; CKD, Chronic kidney disease. Bottom left panel data resource: primary self-reported data alone; Top right panel data resource: primary self-reported data in combination with routinely collected data. Number in each square indicates the prevalence of that corresponded dyad multimorbidity pattern.

individuals not to recognise associated symptoms or to perceive their conditions as not severe enough to report.[51–53] Additionally, the cognitive and psychosocial conditions of the study participants may lead to variability in the interpretation of questions asked by an interviewer, despite a standardised operation protocol was adapted.[50] Also, given the nature of the study design, with a focus on rural population, the low health literacy level among this population may lead to underreporting of the diseases.[54]

Although many diseases showed substantial changes in prevalence following supplementing with routinely collected data, the prevalence of hypertension and diabetes remained relatively stable compared to the other prevalent diseases. This may be due to high awareness of patients with hypertension and diabetes, reducing the likelihood of recall bias or underreporting.[55] The better awareness in hypertension and diabetes may be attributed to the successful implementation of China's National Essential Public Health Service Package since 2009, which provides screening, management, and referral for patients with hypertension and diabetes by the primary healthcare providers.[56] Nevertheless, this study still found that up to 40.2% and 4.6% of individuals at risk for hypertension or diabetes may remain unscreened or undiagnosed. In particular, a significant number of individuals had measured SBP ≥130 mmHg and/or DBP ≥80 mmHg, This may be attributed to the use of the latest hypertension diagnosis standards, while there remains a gap between the guideline and the criteria currently employed in clinical practice in China, where the clinical diagnosis of

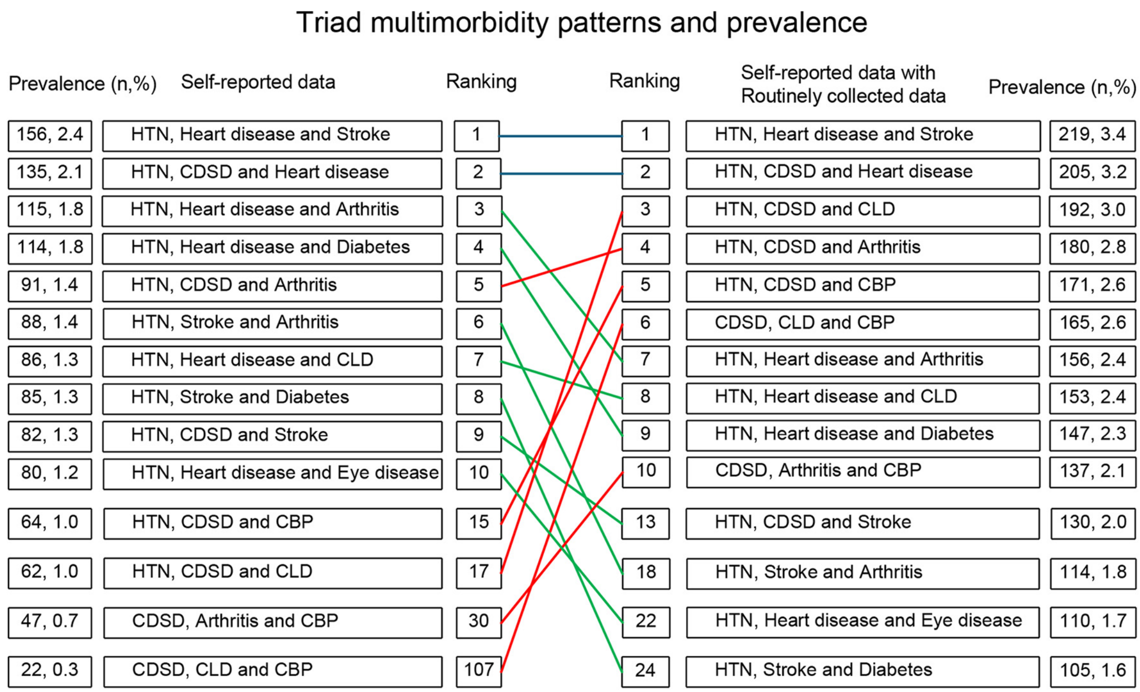

Fig. 4: The leading 10 triad multimorbidity patterns and prevalence with primary self-reported data alone and in combination with routinely collected data. HTN, Hypertension; CDSD, Chronic digestive system diseases; OHD, Oral health disorders; CLD, Chronic lung disease; CBP, Chronic back pain; CKD, Chronic kidney disease.

hypertension still follows the standard of SBP ≥140 mmHg and/or DBP ≥90 mmHg.[37,57] In addition to hypertension, this study also found that many participants with diabetes, mental health disorders, and chronic kidney disease might remain unscreened or undiagnosed. Although routinely collected data can support to address underreporting caused by report bias in the face-to-face questionnaire, many diseases may still go undetected and undiagnosed due to insufficient public health awareness, inadequate screening, or suboptimal disease management.[58,59] Our study also revealed that the self-reported prevalence of anxiety and depression was low, while the number of participants screened to have mild or above anxiety and depression was high and similar to that in a nationally representative survey of more than 30,000 individuals in China.[60] This highlights an urgent need to further strengthen mental health screening in resource-limited rural areas.[61]

Strengths and limitations

The strengths of this study include the random selection of participants stratified by age and sex from 240 villages across three Chinese provinces to maximize sample representativeness in the study areas. Additionally, the use of routinely collected data, together with physical examinations, disease screening questionnaires and blood tests complements the primary self-reported data and reduces potential report bias and thus provides a more robust assessment of multimorbidity prevalence and patterns. This study also has a few limitations. Firstly, due to the outmigration of younger male individuals from rural areas to work outside the villages, despite stratification by age and gender, the average age of the 30–59 age group lies closer to its upper limit. Thus, the prevalence and patterns of multimorbidity among the younger male group may not be accurately represented. Secondly, in this study, multimorbidity was defined as the co-existence of two or more chronic diseases, including chronic non-communicable diseases, chronic infectious diseases and mental health disorders. However, due to the protection of personal privacy, chronic infectious diseases, such as HIV and hepatitis C were not collected, which may underestimate the impact of chronic infectious diseases on multimorbidity. Thirdly, self-medication with Traditional Chinese Medicine may affect participants' awareness of their actual health conditions and their intent to seek for formal diagnosis and treatment in medical facilities. This would subsequently result in omissions in self-reported diseases, leading to underreporting or underestimation of multimorbidity prevalence. Lastly, the routinely collected data may be inaccurate due to factors such as the misuse or unauthorised use of health insurance, which can lead to mis-recording in health claims. In addition, routinely collected data generally still rely on manual data entry at the initial stage, making it inevitably prone to human errors. For instance, the data entry for hospital electronic medical record systems may experience inaccuracies in disease coding or diagnostic records.

Recommendations

The current health system and clinical practices primarily focus on treating single diseases, which is increasingly inadequate and sometimes even harmful for patients with multimorbidity. It is essential to understand the current disease landscape to assess multimorbidity. We recommend considering the complement of routinely collected data with primary data in future studies, to better understand multimorbidity epidemiology and guide future development and implementation of interventions. However, the effectiveness of this approach depends on the completeness, accessibility, quality and accuracy of the data.[62–64] Future research should account for potential errors during diagnostic and administrative processes, conducting validation work and adhering to best practices when using routinely collected data.[65] Additionally, screening and increasing awareness for hypertension, diabetes, mental health disorders, and chronic kidney disease should be prioritised at the primary healthcare level, especially in rural areas, to reduce the potential health inequities. We also recommend specialists develop generalist skills, while general practitioners gain specialist knowledge to treat and manage multimorbidity. Given the increasing burden of mental health disorders, it is important to integrate physical and mental health services, as they remain fragmented in China. Both clinical practice and research should focus on the needs of patients experiencing both physical and mental health conditions.

Conclusion

The prevalence of multimorbidity is currently underestimated in studies solely using self-reported data. Combining primary data with routinely collected data is feasible, but there is a need to enhance the integration of routinely collected data to further improve its accessibility and quality. Using multiple data sources may offer a replicable approach for assessing multimorbidity in diverse settings and populations, with a fast growing of information technology worldwide. Addressing the gaps in data collection and integration is critical for advancing our understanding of multimorbidity, leading to more accurate and robust assessment of the burden of multimorbidity, prioritising healthcare needs and improve the outcomes of patients with multimorbidity, not only in China, but beyond.

Contributors

MT, XZ and TL led the conception of this study. ZL, JY, HH, TH, PZ, MB, PY, SX, YW, WJ, AG, YP, XF, LY, HZ, YW and WT provided suggestions and technical supports for the development of the protocol. XZ completed the draft of this manuscript, and MT, SX, YP and WT made substantial edits to the draft. ZL, JY, HH, TH, PZ and MB provided critical local coordinations and implementations. XZ and TL verified the data and had access to the raw data. MT had the final responsibility for the decision to submit the manuscript for publication. All co-authors provided critical suggestions for the writing and approved the submission of this paper.

Data sharing statement

The study data and related documents are available upon reasonable request from the corresponding author, Professor Maoyi Tian, via email at maoyi.tian@hrbmu.edu.cn.

Editor note

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

All authors declare no conflict of interests for this study.

Acknowledgements

The study is supported by Harbin Medical University Leading Talent Grant (31021220002), received by Maoyi Tian and National Natural Science Foundation of China (72074065 and 72474063), received by Yongchen Wang. We would like to thank all the participants and interviewers for their valuable time and contribution to this study.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanwpc.2024.101272.

References

1 van den Akker M, Buntinx F, Metsemakers JF, Roos S, Knottnerus JA. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol*. 1998;51(5):367–375.
2 Tugwell P, Knottnerus JA. Multimorbidity and Comorbidity are now separate MESH headings. *J Clin Epidemiol*. 2019;105:vi–viii.
3 Nicholson K, Almirall J, Fortin M. The measurement of multimorbidity. *Health Psychol*. 2019;38(9):783–790.
4 Hu Y, Wang Z, He H, Pan L, Tu J, Shan G. Prevalence and patterns of multimorbidity in China during 2002-2022: a systematic review and meta-analysis. *Ageing Res Rev*. 2024;93:102165.
5 Tran PB, Kazibwe J, Nikolaidis GF, Linnosmaa I, Rijken M, van Olmen J. Costs of multimorbidity: a systematic review and meta-analyses. *BMC Med*. 2022;20(1):234.
6 Zhao Y, Atun R, Anindya K, et al. Medical costs and out-of-pocket expenditures associated with multimorbidity in China: quantile regression analysis. *BMJ Glob Health*. 2021;6(2):e004042.
7 Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med*. 2007;22 Suppl 3(Suppl 3):391–395.
8 Mujica-Mota RE, Roberts M, Abel G, et al. Common patterns of morbidity and multi-morbidity and their impact on health-related quality of life: evidence from a national survey. *Qual Life Res*. 2015;24(4):909–918.
9 Ermogenous C, Green C, Jackson T, Ferguson M, Lord JM. Treating age-related multimorbidity: the drug discovery challenge. *Drug Discov Today*. 2020;25(8):1403–1415.
10 Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA*. 2005;294(6):716–724.
11 Moffat K, Mercer SW. Challenges of managing people with multimorbidity in today's healthcare systems. *BMC Fam Pract*. 2015;16:129.
12 Wallace E, Salisbury C, Guthrie B, Lewis C, Fahey T, Smith SM. Managing patients with multimorbidity in primary care. *BMJ*. 2015;350:h176.
13 Zhang X, Padhi A, Wei T, et al. Community prevalence and dyad disease pattern of multimorbidity in China and India: a systematic review. *BMJ Glob Health*. 2022;7(9):e008880.
14 Xie D, Wang J. Comparison of self-reports and biomedical measurements on hypertension and diabetes among older adults in China. *BMC Public Health*. 2020;20(1):1664.

15 Cheng C, Yang CY, Inder K, Wai-Chi Chan S. Urban-rural differences in mental health among Chinese patients with multiple chronic conditions. *Int J Ment Health Nurs*. 2020;29(2):224–234.
16 Zhang C, Xiao S, Shi L, et al. Urban-rural differences in patterns and associated factors of multimorbidity among older adults in China: a cross-sectional study based on apriori algorithm and multinomial logistic regression. *Front Public Health*. 2021;9:707062.
17 Zhou Y, Guo Y, Liu Y. Health, income and poverty: evidence from China's rural household survey. *Int J Equity Health*. 2020;19(1):36.
18 Spasoff RA. *Epidemiologic methods for health policy*. Oxford University Press; 1999.
19 Mc Cord KA, Al-Shahi Salman R, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials*. 2018;19(1):29.
20 Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885.
21 Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol*. 2011;40(6):1652–1666.
22 Neal B, Wu Y, Feng X, et al. Effect of salt substitution on cardiovascular events and death. *N Engl J Med*. 2021;385(12):1067–1077.
23 Xiong S, Jiang W, Wang Y, et al. Using routinely collected data to determine care cascades of hypertension and type-2 diabetes management in China: a cross-sectional study. *Lancet Reg Health West Pac*. 2024;45:101019.
24 de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract*. 2006;23(2):253–263.
25 Nicholls SG, Langan SM, Sørensen HT, Petersen I, Benchimol EI. The RECORD reporting guidelines: meeting the methodological and ethical demands of transparency in research using routinely-collected health data. *Clin Epidemiol*. 2016;8:389–392.
26 Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet*. 2012;380(9836):37–43.
27 Fan J, Sun Z, Yu C, et al. Multimorbidity patterns and association with mortality in 0.5 million Chinese adults. *Chin Med J (Engl)*. 2022;135(6):648–657.
28 Violan C, Foguet-Boreu Q, Flores-Mateo G, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS One*. 2014;9(7):e102149.
29 Pati S, Hussain MA, Swain S, et al. Development and validation of a questionnaire to assess multimorbidity in primary care: an Indian experience. *BioMed Res Int*. 2016;2016:6582487.
30 Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–1097.
31 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–613.
32 Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med*. 2007;146(5):317–325.
33 Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2014;36(5):539–544.
34 Shih YC, Chou CC, Lu YJ, Yu HY. Reliability and validity of the traditional Chinese version of the GAD-7 in Taiwanese patients with epilepsy. *J Formos Med Assoc*. 2022;121(11):2324–2330.
35 Gong Y, Zhou H, Zhang Y, et al. Validation of the 7-item Generalized Anxiety Disorder scale (GAD-7) as a screening tool for anxiety among pregnant Chinese women. *J Affect Disord*. 2021;282:98–103.
36 Zhang C, Wang T, Zeng P, et al. Reliability, validity, and measurement invariance of the general anxiety disorder scale among Chinese medical university students. *Front Psychiatry*. 2021;12:648755.
37 National Center for Cardiovascular Diseases. [Chinese clinical practice guidelines of hypertension]. *Chin J Cardiol*. 2022;50(11):1050–1095.
38 Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of cardiology/American heart association task force on clinical practice guidelines. *Hypertension*. 2018;71(6):1269–1324.
39 Boss K, Stolpe S, Müller A, et al. Effect of serum creatinine difference between the Jaffe and the enzymatic method on kidney disease detection and staging. *Clin Kidney J*. 2023;16(11):2147–2155.
40 Levey AS, Coresh J, Bolton K, et al. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis*. 2002;39(2 Suppl 1):S1–S266.
41 Xiong S, Jiang W, Meng R, et al. Factors associated with the uptake of national essential public health service package for hypertension and type-2 diabetes management in China's primary health care system: a mixed-methods study. *Lancet Reg Health West Pac*. 2023;31:100664.
42 Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med*. 2012;10(2):142–151.
43 Sun ZJ, Fan JN, Yu CQ, et al. [Prevalence, patterns and long-term changes of multimorbidity in adults from 10 regions of China]. *Zhonghua Liuxingbingxue Zazhi*. 2021;42(5):755–762.
44 MacRae C, Morales D, Mercer SW, et al. Impact of data source choice on multimorbidity measurement: a comparison study of 2.3 million individuals in the Welsh National Health Service. *BMC Med*. 2023;21(1):309.
45 Yao SS, Cao GY, Han L, et al. Prevalence and patterns of multimorbidity in a nationally representative sample of older Chinese: results from the China health and retirement longitudinal study. *J Gerontol A Biol Sci Med Sci*. 2020;75(10):1974–1980.
46 Zhao Y, Atun R, Oldenburg B, et al. Physical multimorbidity, health service use, and catastrophic health expenditure by socioeconomic groups in China: an analysis of population-based panel data. *Lancet Glob Health*. 2020;8(6):e840–e849.
47 Liu H, Zhao Y, Qiao L, et al. Consistency between self-reported disease diagnosis and clinical assessment and under-reporting for chronic conditions: data from a community-based study in Xi'an, China. *Front Public Health*. 2024;12:1296939.
48 Navin Cristina TJ, Stewart Williams JA, Parkinson L, Sibbritt DW, Byles JE. Identification of diabetes, heart disease, hypertension and stroke in mid- and older-aged women: comparing self-report and administrative hospital data records. *Geriatr Gerontol Int*. 2016;16(1):95–102.
49 Topp J, Andrees V, Heesen C, Augustin M, Blome C. Recall of health-related quality of life: how does memory affect the SF-6D in patients with psoriasis or multiple sclerosis? A prospective observational study in Germany. *BMJ Open*. 2019;9(11):e032859.
50 Khare SR, Vedel I. Recall bias and reduction measures: an example in primary health care service utilization. *Fam Pract*. 2019;36(5):672–676.
51 Cong S, Yao JY, Fan J, et al. [Analysis on awareness of chronic obstructive pulmonary disease (COPD) status and related knowledge in patients with COPD in China, 2014-2015]. *Zhonghua Liuxingbingxue Zazhi*. 2020;41(7):1034–1040.
52 Gikas A, Triantafillidis JK. The role of primary care physicians in early diagnosis and treatment of chronic gastrointestinal diseases. *Int J Gen Med*. 2014;7:159–173.
53 Frost M, Wraae K, Gudex C, et al. Chronic diseases in elderly men: underreporting and underdiagnosis. *Age Ageing*. 2012;41(2):177–183.
54 Ran M, Peng L, Liu Q, Pender M, He F, Wang H. The association between quality of life(QOL) and health literacy among junior middle school students: a cross-sectional study. *BMC Public Health*. 2018;18(1):1183.
55 Zhang M, Shi Y, Zhou B, et al. Prevalence, awareness, treatment, and control of hypertension in China, 2004-18: findings from six rounds of a national survey. *BMJ*. 2023;380:e071952.
56 Xiong S, Cai C, Jiang W, et al. Primary health care system responses to non-communicable disease prevention and control: a scoping review of national policies in Mainland China since the 2009 health reform. *Lancet Reg Health West Pac*. 2022;31:100390.
57 Yin R, Yin L, Li L, et al. Hypertension in China: burdens, guidelines and policy responses: a state-of-the-art review. *J Hum Hypertens*. 2022;36(2):126–134.
58 Zhao Y, Crimmins EM, Hu P, et al. Prevalence, diagnosis, and management of diabetes mellitus among older Chinese: results

from the China Health and Retirement Longitudinal Study. *Int J Public Health*. 2016;61(3):347–356.
59 Lu J, Lu Y, Wang X, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 1·7 million adults in a population-based screening study (China PEACE Million Persons Project). *Lancet*. 2017;390(10112):2549–2558.
60 Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry*. 2019;6(3):211–224.
61 Que J, Lu L, Shi L. Development and challenges of mental health in China. *Gen Psychiatr*. 2019;32(1):e100053.
62 Bernardi FA, Alves D, Crepaldi N, Yamada DB, Lima VC, Rijo R. Data quality in health research: integrative literature review. *J Med Internet Res*. 2023;25:e41446.
63 Syed R, Eden R, Makasi T, et al. Digital health data quality issues: systematic review. *J Med Internet Res*. 2023;25:e42615.
64 Liu Y, Xiao S, Yin X, et al. Nation-wide routinely collected health datasets in China: a scoping review. *Public Health Rev*. 2022;43:1605025.
65 Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: the importance of high-quality diagnostic coding to research. *CMAJ*. 2017;189(33):E1054–e1055.