Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.jzus.zju.edu.cn; www.springerlink.com
E-mail: jzus@zju.edu.cn

Methodology and cohort profile for the Hangzhou Lumbar Spine Study: a study focusing on back health in a Chinese population*

Xiao-jian HU[1], Lun-hao CHEN[1], Michele C. BATTIÉ[2], Yue WANG[†‡1]

[1]Spine Lab, Department of Orthopedic Surgery, the First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China

[2]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton AB T6G 2G4, Canada

[†]E-mail: wangyuespine@zju.edu.cn

Received Oct. 5, 2017; Revision accepted Nov. 5, 2017; Crosschecked June 6, 2018

Abstract: Back pain is a worldwide health problem, adding a tremendous burden to modern societies. However, little information on back health is available in China, even though a quarter of the world's population is Chinese. To enhance knowledge in this area, we designed and initiated the Hangzhou Lumbar Spine Study, which is a cross-sectional study of a general sample of mainland Chinese with focusing on disc degeneration, Modic changes, endplate lesions, and back pain. The study consists of a structured questionnaire to measure back pain history and lifetime exposure to suspected risk factors, magnetic resonance imaging of the lumbar spine, bone mineral density study of the spine and hip, and DNA sample analysis. Here we briefly introduce the study methodology, report the test-retest reliability of the questionnaire, and describe the cohort profile to date. Since May 2014, 301 randomly selected subjects (male/female, 122/179; mean age, 51.0 years; range, 20–87 years) have been recruited. Tests-retests of the questionnaire, completed by 40 participants, revealed good reliability. To our knowledge, the Hangzhou Lumbar Spine Study is the first population-based epidemiological study conducted to characterize lumbar spinal phenotypes and back pain, their interaction, and their associations with lifetime environmental exposure, in mainland Chinese. Epidemiological information obtained from a reliable questionnaire, magnetic resonance (MR) imaging data, dual energy X-ray absorptiometry (DXA) measurements, and DNA analysis may serve as a valuable reference for future studies on back health, particularly for mainland Chinese.

Key words: Back pain; Population-based study; Disc degeneration; Modic change; Endplate lesion; Methodology
https://doi.org/10.1631/jzus.B1700484 **CLC number:** R681.5

1 Introduction

Back pain is a common health problem, with over two thirds of adults suffering from back pain sometime in their lives (Deyo and Weinstein, 2001). Back pain-related healthcare costs, work loss, and psychosocial problems result in a tremendous social burden in developed countries (Andersson, 1999). However, little information regarding back pain epidemiology is available in the developing country of China.

A basic challenge to devising effective approaches to lessen the burden of back pain on individuals and society is that a clear pathoanatomical cause cannot be identified in most back pain patients (Abraham and Killackey-Jones, 2002). Despite numerous studies of the etiology and pathogenesis of back pain, much remains unclear (Balagué et al., 2012). Of the various suspected underlying pathologies, disc degeneration has long been thought to be a major culprit (Battié and Videman, 2006). Yet most

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 81371995), the Key Platform Project of the Health Department of Zhejiang Province (No. 2015ZDA014), and the 2015 Qianjiang Talent Program of Zhejiang Province, China
ORCID: Yue WANG, https://orcid.org/0000-0002-6580-809X

studies have not revealed strong associations between disc degeneration and back pain (Jensen et al., 1994; Salminen et al., 1995). While disc degeneration is a genetically dominated condition studied mostly in populations of European descent living in developed countries (Battié and Videman, 2006), the epidemiology of disc degeneration in mainland Chinese remains largely unexplored.

Modic changes (MCs) or endplate signal changes represent another spinal phenotype increasingly suspected of being associated with back pain. Like disc degeneration, the prevalence of MCs varies considerably, from 19% to 59% in patients with lumbar degenerative disorders, but is considerably lower in asymptomatic individuals (Rahme and Moussa, 2008). While some studies suggest that MCs are age-related degenerative factors unrelated to back pain (Jarvik et al., 2005; Kuisma et al., 2007; Jarvinen et al., 2015), others report that Type I or Type II MCs, or both, are associated with back pain (Kjaer et al., 2005a, 2005b; Albert and Manniche, 2007; Kuisma et al., 2009). Further, there is evidence that MCs may be an independent risk factor for episodes of disabling back pain (Maatta et al., 2015). Such inconsistent findings confuse clinical interpretations of MCs and, thus, warrant further investigation. In addition, to date there has been no population-based epidemiological study investigating the etiology or pathogenesis of MCs.

Endplate lesions, a group of pathological findings newly recognized in cadaveric vertebrae, appear to be common spinal phenotypes, which may play an important role in both disc degeneration and back pain (Wang et al., 2012a, 2012b, 2013). Yet, the clinical significance of endplate lesions needs to be confirmed in vivo using standard radiological approaches, and so do their determinants. We postulated that modern high-resolution magnetic resonance (MR) imaging may be able to detect endplate lesions, and the presence of endplate lesions would relate to disc degeneration and pain.

Conflicting evidence on the clinical relevance of spinal phenotypes, such as disc degeneration, MCs, and endplate lesions, may be due to the tremendous variation in study samples and definitions and measures of both spinal phenotypes and back pain. For example, most studies on MCs used convenient samples (Kaapa et al., 2012), patients (Kuisma et al., 2009), or highly selected populations (Kuisma et al., 2007), resulting in limited comparability and generalizability of the findings. Among the few studies on MCs and back pain using general population samples, some used only T2-weighted MR images (MacGregor et al., 2004; Mok et al., 2016) or studied only men (Wang et al., 2012c) or women (MacGregor et al., 2004), others used samples of general populations of a specific age (Kjaer et al., 2005a, 2005b). A Japanese study investigated mainly elderly subjects (mean age, 66 years), for whom endplate pathologies were not the primary focus (Teraguchi et al., 2015). Moreover, most early studies used MR machines of low resolution (0.2 to 1.5 T), which limited the visualization and measurement of spinal phenotypes. In addition, few studies used structured questionnaires to measure lifetime environmental exposure, leaving the etiologies of most spinal phenotypes unstudied.

The Hangzhou Lumbar Spine Study, which is a cross-sectional study of a general population sample of mainland Chinese, is designed to enhance current knowledge of disc degeneration, MCs, endplate lesions, and back pain. Specific purposes of the study include: (1) to determine epidemiological characteristics of disc degeneration, endplate lesions, MCs, and back pain in Chinese; (2) to explore the associations between lifetime environmental exposure and the presence of endplate lesions and MCs; (3) to determine the relationships among endplate lesions, MCs, disc degeneration, and back pain; and (4) to identify genotypes associated with lumbar spinal phenotypes in Chinese. The aims of the present paper are to introduce the Hangzhou Lumbar Spine Study, present the study methodology, describe the study sample recruited to date, and conduct a test-retest reliability of the structured questionnaire.

2 Materials and methods

The Hangzhou Lumbar Spine Study is a cross-sectional study of adults randomly selected from a typical Chinese community in Hangzhou, the capital city of Zhejiang Province in eastern China. There are about six million registered residents in Hangzhou City. The study data acquisition comprises four parts: a structured questionnaire to measure back pain history and lifetime exposure to suspected risk factors, MR imaging of the lumbar spine, dual energy X-ray absorptiometry (DXA) to measure the bone mineral density (BMD) of the spine and hip, and DNA sample collection and analysis (Fig. 1). The study protocol was approved by the Ethical Committee at the First Affiliated Hospital of Zhejiang University (Hangzhou, China).

2.1 Subject recruitment

The study began recruiting subjects in May 2014. All subjects were recruited from the Caihe community, a typical Chinese community in Hangzhou. The Caihe community was founded in 1987, nine years after the reformation and opening-up of China. The first generation of Caihe residents was mainly labor workers from various factories, suburban farmers, office workers, and veterans. With rapid urban expansion and a real estate boom after 1998, many well-educated young couples moved into the community, further enriching the population. After 30 years of development, the community now is large and mature, with about 43000 registered residents in an area of 3.5 km^2. Unlike some new communities that consist mainly of young and well-educated immigrants, registered residents in Caihe cover a full range of ages and typical occupations, with diverse exposure to various suspected occupational risk factors for lumbar pathologies. Therefore, the Caihe community provides an ideal population pool from which to draw a representative sample of the general Chinese population for an epidemiological study.

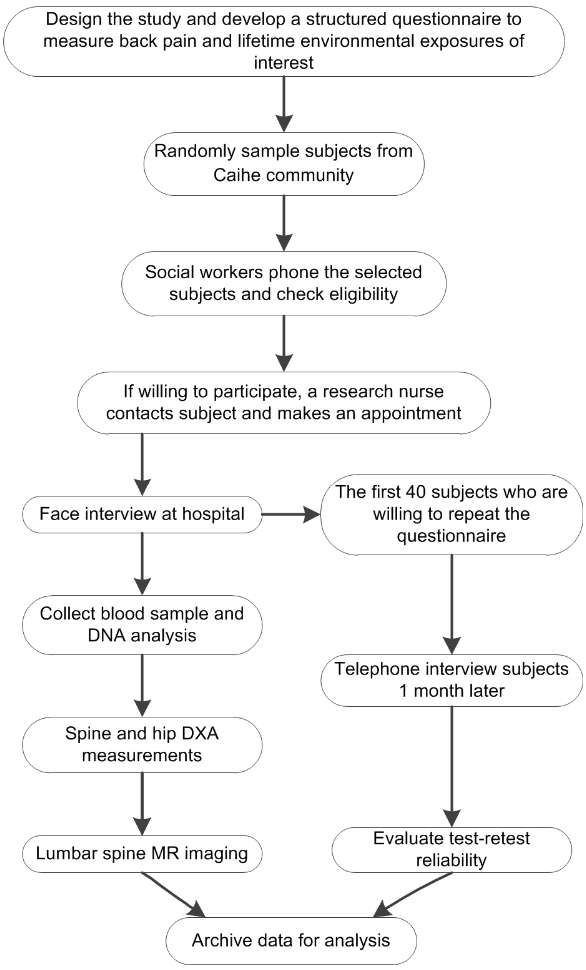

Fig. 1 Overview workflow for the Hangzhou Lumbar Spine Study

We used simple random sampling to recruit subjects from the community in an effort to decrease sampling bias. Subjects were randomly selected from the residency registry using the random function in Excel 2010 (Microsoft). A community worker employed by the study then contacted the selected subjects to introduce the study and ask if they would be willing to participate. If willing, study inclusion criteria were checked.

In addition to being a Han Chinese resident of Caihe and randomly selected from the registry, study subject inclusion criteria included being older than 18 years of age, having no contraindications for MR imaging or DXA (e.g. no lumbar spine instrumentation), and being willing to participate in the study as a volunteer. Subjects were randomly selected regardless of back pain or any other factors. Our goal was to recruit 800 to 1000 participants in five years.

If inclusion criteria were met, a single trained research nurse called the candidate participant to make an appointment to conduct the interview, followed by a 5-ml blood draw for a DNA sample. Then the subject underwent lumbar spine MR imaging and DXA. These studies were usually conducted on the same day. All volunteers provided written informed consent prior to data collection.

2.2 Structured questionnaire

A structured questionnaire was developed to measure lifetime environmental exposure of relevance and other suspected risk factors for lumbar spinal phenotypes, and to record each subject's history of back injury or trauma and back pain. The questions were adapted mainly from the Twin Spine Study (Battié et al., 1995), the Multinational Musculoskeletal

Inception Cohort Study Statement (Pincus et al., 2008), and the Delphi Study Recommendations for conducting prevalence studies of back pain (Dionne et al., 2008). The interview includes items on: (1) demographic characteristics; (2) general lifestyle and health behaviors; (3) occupational history and associated physical demands; and (4) back pain history. Although cognitive and psychosocial factors are important in studying back pain, we did not include related items in the questionnaire as the main focus of the study is the etiologies of lumbar spine MR findings and their associations with back pain. Furthermore, a concise and focused questionnaire can save time and cost for such a large scale epidemiological study. After the interview was developed and finalized by two of the authors (MCB and YW) in English, the questionnaire was translated into Chinese.

2.2.1 Demographics, general lifestyle, and health behaviors

Body weight and height were measured immediately prior to the interview. General demographic data were then collected, including education categorized according to the World Health Organization Health Survey (World Health Organization, 2002), and followed by questions on smoking, alcohol consumption, and vehicle use.

Evidence suggests that pack-years, the number of cigarettes smoked daily multiplied by years of smoking, can appropriately reflect the degree of smoking exposure (Pincus et al., 2008). To this end, two items were designed to identify whether subjects smoked regularly and the ages of onset and cessation of smoking, and a third item estimated the number of cigarettes smoked per day.

The Alcohol Use Disorders Identification Test, which demonstrates excellent sensitivity and validity for both alcohol use disorders and drinking problems (Bush, 1998), was simplified to measure alcohol exposure. The first item identified whether the subject drank regularly, the second indicated alcohol type (wine of low, mild, or high alcohol content), and the third determined the duration of alcohol consumption.

Overall exposure of vehicular vibration was evaluated by the hours spent per day driving or as a passenger in motorized vehicles, the type of motorized vehicles, driving distance, and the number of years using motored vehicles. As bicycles are commonly used for transportation in China, we also asked subjects the number of hours per day spent riding a bicycle.

Sports and exercise history was evaluated using the methods reported in the Twin Spine Study (Battié et al., 1995). To facilitate recall, subjects were queried about regular exercise or sports involvement during each school education period (primary school, middle school, and high school), and then during each decade after 20 years of age. For each period, exercise type, frequency, number of months exercised per year, and number of years of participation were recorded to estimate overall exposure to various sports and exercises. In addition, any exercise-related “back injury” was specifically recorded. Finally, any other physical activities which were regularly performed, but not covered under occupation and exercise, were also noted.

2.2.2 Occupational physical demands and history

Overall lifetime occupational demands were determined from the physical demands of the jobs subjects had for the majority of their working years (Battié et al., 1995). Occupational physical demands were classified as sedentary work (spending nearly all day sitting with little lifting), light physical demands (involving sitting and standing or walking activities, and light lifting generally less than 5 kg), medium physical demands (lifting and handling of weights generally less than 20 kg), heavy physical demands (frequent lifting of materials between 20–40 kg), and very heavy physical demands (frequent lifting of materials over 40 kg). The years of having each level of physical demands were recorded. Subjects were further asked how much of their work was done in awkward twisted or bent postures (none, occasionally, >1 h, or majority of the day). In addition, subjects were asked if they had physically hard work before 18 years of age.

Questions pertaining to occupational history and back injury history were mainly modified from the Twin Spine Study (Battié et al., 1995). The age at which the subject began his or her first full-time job was recorded. Each job the subject subsequently held was categorized with respect to material handling, and time spent sitting, standing, and walking, and on-the-job driving/riding per day and per week was recorded. The amount of weight most commonly lifted and the frequency of lifting were recorded to calculate a summary indicator of lifting per day. Working postures were categorized as working overhead or standing, slight bending/twisting, moderate bending/twisting, and extreme bending/twisting, and time spent in each position per day was recorded. Moreover, the "most physically demanding activity of the job" was documented. In addition, commute time (minutes to and from work) and transportation mode were recorded. Finally, subjects were asked whether they had suffered from back-related injuries, the symptom duration, and whether their choice of work or a change of jobs was ever influenced by their health.

2.2.3 Low back pain history

The questions pertaining to back pain came primarily from the recommendations of the Delphi Study by Dionne et al. (2008) and the Twin Spine Study (Battié et al., 2007). Low back pain was defined as pain or soreness in the area between the lower ribs and buttock crease, which was further illustrated using a diagram during the interview (Dionne et al., 2008).

Dionne's questionnaire was used to measure the presence, frequency, duration, and severity of current back pain and back pain over the past four weeks (Dionne et al., 2008). Pain related to fever or menstruation was excluded. Questions for back problems over the past 12 months were adapted from the Twin Spine Study (Battié et al., 1995). In addition, six items were designed to measure lifetime back pain and related history (Battié et al., 1995). The age of onset was recorded and space was provided to record detailed information as needed.

2.3 Blood sample collection and DNA analysis

Immediately following the interview, a 5-ml blood sample was collected and stored at −80 °C. DNA was extracted within one month by a laboratory technician using standard approaches and the extracted DNA samples were archived at −20 °C for future analysis.

2.4 DXA and BMD measurements

Lumbar spine and hip BMDs were measured with DXA (Lunar Prodigy, GE, USA) using standard protocols. Results were printed and archived together with interview data collection forms.

2.5 Lumbar spine MR imaging

The lumbar spine was imaged using a 3.0 T MR scanner (Philips Medical Systems, DA Best, the Netherlands). The protocol included T1W, T2W, and spectral presaturation with inversion recovery (SPIR) sequences of sagittal images, as well as axial T2W images of each lumbar intervertebral disc from L1/2 through L5/S1. Sagittal T2W images were acquired using a turbo spin-echo sequence with a repetition time (TR) of 3500 ms, echo time (TE) of 100 ms, echo train length of 17, and two acquisitions. The matrix size was 300×229, field of view (FOV) 24 cm×24 cm, with a slice thickness of 3 mm and intersection gap of 0.3 mm. Sagittal T1W images were acquired using a turbo spin-echo sequence with a TR of 442 ms, TE of 8 ms, echo train length of 4, and four acquisitions. The matrix size was 268×191, FOV 24 cm×24 cm, with a slice thickness of 3 mm and intersection gap of 0.3 mm. Fat suppression images were acquired using SPIR sequence with T1 suppression. This technique is based on a frequency-selective inversion pre-pulse tuned to fat resonance and with an inversion time adjusted to 220 ms for optimal fat suppression.

In summary, 15 T1W sagittal images, 15 T2W sagittal images, and 15 fat suppression sagittal images were obtained for each lumbar spine, and 3 axial T2W images for each lumbar disc using a 3.0 T MR scanner. Lumbar spine MR images were downloaded in DICOM format and archived for future study.

2.6 Test-retest reliability of the interview data

At the end of the interview, subjects were asked if they would be willing to be contacted again to repeat the questionnaire after one month, a time interval selected to prevent carry-over effects (Weir, 2005). Test-retest reliability response was evaluated using data from the first 40 subjects who agreed to repeat the interview. To ease the burden on the study volunteers, the second interview was conducted over the telephone by the same research nurse who conducted the face-to-face interview. The research nurse did not have access to the prior responses during the repeated interviews, and data analysis was performed by another investigator.

2.7 Statistical analysis

Statistical analyses were performed using STATA (Version 13.0, StataCorp LP, TX, USA).

Descriptive statistics was used to document the characteristics of the study sample and kappa statistics to examine the reliability of interview responses. For binary variables, test-retest agreement was measured using Cohen's kappa (κ). For ordinal variables, weighted κ was calculated. For continuous variables, intra-class correlation coefficients (ICCs) were calculated. It has been suggested that κ or ICC values of less than 0.20 indicate slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement, and 0.81 to 1.00 excellent agreement (Landis and Koch, 1977).

3 Results

To date, 1368 residents have been randomly selected from the resident registry. They included 661 (48.3%) men and 707 (51.7%) women (mean age, 48.2 years; standard deviation, 16.9 years; range, 18–104 years). Of these, 367 had invalid phone numbers in the registry, leaving 1001 to contact for possible participation (480 men, 521 women; mean age, 49.5 years; standard deviation, 14.9 years; range, 18–93 years). Among them, 605 subjects declined participation, 72 were college students who were living in other cities and unable to take part in, 8 were unable to travel to the data collection site (too old, diseased, or disabled), 9 were ineligible for MR imaging (5 pregnant, 2 with metal instrumentations, and 2 with a cardiovascular stent), and 6 completed the interview but failed to finish MR imaging (4 withdrew participation and 2 had contraceptive devices). As a result, 301 subjects (122 (40.5%) men and 179 (59.5%) women; mean age, 51.0 years; standard deviation, 15.2 years; range, 20–87 years) participated and completed all the proposed examinations (Table 1). The overall volunteer rate was 30.1%.

In terms of demographic characteristics and health behaviors, the 40 subjects who participated in the test-retest reliability study were similar to the larger study sample from which they came (Table 1). Among the 56 items tested from the interview, substantial to excellent agreement was observed in the responses to 47 items (83.9%), moderate agreement in 7 items (12.5%), and fair agreement or less in 2 items (3.6%).

More specifically, test-retest reliability coefficients ranged from 0.56 to 1.00 for demographic characteristics, general lifestyle, health behaviors, and sports and exercise history (Table 2). Other than hours spent riding a bicycle (κ=0.56), responses to most items on vehicle use had substantial to excellent agreement (κ or ICC=0.66–0.94). For exercise type and duration and related injury, agreement was generally substantial (κ or ICC=0.63–0.79; Table 2).

The reliability measurements for occupational demands and occupational history revealed moderate to excellent agreement (κ or ICC=0.58–1.00, Table 3). Agreement on items assessing the duration and physical demands of each specific occupation during the subject's lifetime is presented in Table 3. When more items were detailed to measure degrees of bending or twisting postures in a job, agreement was fair or substantial for most variables (ICC=0.33–0.74), with the lowest reliability coefficients found for time spent in various positions (Table 3). For recalled sudden work-related back pain or injury, test-retest agreement was substantial (κ=0.74).

Moderate to excellent test-retest agreement was found for different definitions of back pain and their related frequency (Table 4). Items for various definitions of low back pain experienced during the subject's lifetime revealed substantial to excellent agreement (κ=0.60–1.00; Table 4).

4 Discussion

Chinese make up about one quarter of the world's population. Yet, to date there has been no population-based study specifically designed to investigate the lumbar spine and back health of people in developing mainland China. The epidemiological characteristics of back pain and spinal phenotypes in mainland Chinese, therefore, remain largely unexplored, as do the risk factors of various spine pathologies. To complement knowledge in this area, we designed and initiated the Hangzhou Lumbar Spine Study. In this paper, we briefly introduced the methodology of the study, including the measurements of lifetime environmental exposure of interest and back pain history, and their test-retest reliability, and described the cohort profile to date.

Table 1 Demographic characteristics and health behaviors of study subjects

Variable	Current cohort in the Hangzhou Lumbar Spine Study (*n*=301)	Subjects who participated in the reliability study (*n*=40)	*P* value[a]
Age (year)	51.0±15.2	49.2±13.3	0.49[b]
Gender			0.50
Male	122 (40.5%)	14 (35.0%)	
Female	179 (59.5%)	26 (65.0%)	
Education			0.95
No formal school	5 (1.7%)	0	
Less than primary school	1 (0.3%)	0	
Primary school completed	13 (4.0%)	2 (5.0%)	
Secondary school completed	69 (22.9%)	9 (22.5%)	
High school or equivalent completed	71 (23.6%)	12 (30.0%)	
College to undergraduate degree	133 (44.2%)	16 (40.0%)	
Post graduate degree completed	9 (3.0%)	1 (2.5%)	
Occupation			0.84
Retired / Been unemployed / Unemployed student	126 (41.9%)	14 (35.0%)	
Legislator / Senior official / Manager	30 (10.0%)	3 (7.5%)	
Professional	53 (17.6%)	8 (20.0%)	
Technician / Associate professional	30 (10.0%)	4 (10.0%)	
Clerk	11 (3.7%)	2 (5.0%)	
Service/sales work	32 (10.6%)	6 (15.0%)	
Agricultural or fishery worker	1 (0.3%)	0	
Plant/machine operator	3 (1.0%)	0	
Craft or trades worker	10 (3.3%)	1 (20.0%)	
Elementary worker	0	0	
Armed forces	0	0	
Others	5 (1.7%)	2 (5.0%)	
Smoking			0.15
Yes	76 (22.0%)	6 (15.0%)	
No	225 (78.0%)	34 (85.0%)	
Alcohol consumption			0.59
Yes	134 (44.5%)	16 (40.0%)	
No	167 (55.5%)	24 (60.0%)	

[a] χ^2 test; [b] *t*-test. Values are expressed as number (percentage) or mean±standard deviation

China has undergone a continuous economic boom over the past 30 years. Underlying "made in China" were dramatic changes in occupational and environmental exposure and social structures. Surprisingly, however, little attention has been paid to the occupational health of modern Chinese. Back pain is within such a neglected area and epidemiological data on back health are currently absent for mainland Chinese. The Hangzhou Lumbar Spine Study aims to determine epidemiological characteristics of various lumbar spine phenotypes, their associations with each other and with back pain, and to elucidate their environmental and genetic determinants in mainland Chinese. Knowledge derived will contribute to the understanding of the etiology and pathogenesis of spinal phenotypes and back pain specifically in the developing country of China.

The current study is one of the few studies on the lumbar spine that have used a structured questionnaire to measure lifetime environmental and occupational exposure of interest. Most items used in the questionnaire were modified from the Twin Spine Study (Battié et al., 1995). As reported in the Twin Spine Study, overall the test-retest agreement was substantial for the majority of items (83.9%) proposed in the structured questionnaire. For lifetime occupational history, most items showed substantial agreement other than the degree of bending or twisting posture (ICC=0.36–0.79). It is challenging to estimate the extent and time of various postures during work, which are prone to recall bias (Halpern et al., 2001). Showing pictures or diagrams to demonstrate posture and load features, as carried out in the Twin Spine Study, may have been helpful to increase the recall

Table 2 Test-retest reliability for demographic characteristics, general lifestyle, health behaviors, and sports and exercise history in Caihe community

Item	κ or ICC	95% CI
Education		
What is the highest level of education you completed?	1.00	
Smoking		
Do you smoke or have you at some time smoked regularly?	0.81	0.71, 0.92
No / Yes, but <12 packs / Yes		
Smoking's pack year	1.00	
Alcohol consumption		
Do you drink or have you at some time drunk regularly?	0.96	0.89, 1.00
Never / Occasionally a little / Often (several times a week) / One time every day / >1 time every day		
Vehicle use		
How many hours per day do you spend, on average, driving or as a passenger in motorized vehicles?	0.94	0.90, 0.97
How many years have you spent about this much time in motorized vehicles?	0.80	0.66, 0.89
How many hours per day do you spend, on average, riding a bicycle?	0.56	0.31, 0.74
Do you drive?	0.94	0.83, 1.00
If yes, how many kilometers did you drive per year?	0.65	0.43, 0.80
If yes, how many years have you driven?	0.66	0.44, 0.80
Sports and exercise history		
Exercise type?	0.67	0.65, 0.73
None / Running / Walking / Ping pong / Badminton / Basketball / Football / Tennis / Yoga / Dance / Swimming / Gym / Others		
Ask if ever exercised in the time of primary, junior and high school and each decade after 20 years old respectively?	0.64	0.40, 0.88
Number of years exercised?	0.66	0.44, 0.80
Number of months in a year?	0.72	0.53, 0.84
Exercise frequency (number of days exercised in a week)?	0.63	0.40, 0.78
Did you have any related injuries when exercise?	0.79	0.39, 1.00

κ: Cohen's kappa; ICC: intra-class correlation coefficient; 95% CI: 95% confidence interval

reliability (Halpern et al., 2001), but we did not incorporate this method in the interview.

There is no universally accepted definition or measurement of back pain, and variation may result in limited comparability and generalizability of research findings and conclusions. We combined the questionnaires in the Twin Spine Study and the Delphi Study (Battié et al., 1995; Dionne et al., 2008) to measure various back pain phenotypes. Consistent with previous reports, the test-retest study revealed substantial to excellent agreement for back pain of various definitions (κ=0.60–0.86). We expect these comprehensive measures may help to address some specific research questions of back pain.

There are three particular strengths of the Hangzhou Lumbar Spine Study. First, unlike some other studies which recruited subjects by public advertising, subjects of the Hangzhou Lumbar Spine Study were randomly selected from a registry of a typical mature community. Given the culture and current socioeconomic status in China, public advertising is likely to attract a disproportionate number of volunteers with back pain problems. Random sampling could minimize selection bias and increase the representativeness of recruited subjects as a sample of the targeted population pool, and thus enhance the generalizability of the study results. Second, a structured questionnaire, adapted from some well-known studies, was specifically developed to measure lifetime exposure of suspected risk factors for back pain and spinal phenotypes. The questionnaire was completed in a face to face interview by a single trained research nurse, resulting in substantial to excellent agreement in retests for most items. Third, a specific protocol was used to image the lumbar spine with a 3.0 T MR scanner, resulting in 60 MR images for each lumbar spine. Such elaborate high-resolution MR images may have substantially increased the visibility of various

Table 3 Test-retest reliability for occupational demands and occupational history in Caihe community (n=40)

Item	κ or ICC	95% CI
Occupational demands		
Which of the following best describes the physical demands of the jobs you have had for the majority of your working years?	0.58	0.37, 0.64
Sedentary work / Light physical demands / Medium physical demands / Heavy physical demands / Very heavy physical demands		
In total, about how many years have you held jobs with the level of physical demands that you noted above?	0.74	0.56, 0.85
Of the jobs just noted, which of the following best describes how much of the work was done in awkward twisted or bent postures?	0.60	0.50, 0.62
Almost none / Occasionally worked in awkward twisted or bent positions / Spent over an hour during the work day in awkward twisted or bent positions / Spend the majority of the day in awkward twisted or bent positions		
During the majority of your working years, have the physical demands on your back (e.g. lifting, twisting, bending) been greatest during your employment or during home and leisure activities?	0.70	0.49, 0.92
Did you do physically hard work when you were young (under 18 years)?	0.72	0.37, 1.00
Occupational history		
During the last 12 months, what has been your main occupation?	0.61	0.50, 0.72
Are you currently?	0.91	0.80, 1.00
Employed full-time / Part time job / Unemployed / Retired / Student / Other		
If currently working, how do you evaluate your present working conditions?	0.72	0.64, 0.85
Very dissatisfied / Dissatisfied / Fair, neither satisfied nor dissatisfied / Satisfied / Very satisfied		
Age when you had your first full-time job	0.44	0.07, 0.70
Job category	0.59	0.43, 0.61
Working years	0.64	0.47, 0.76
Hours in sitting at work (not including in car)	0.80	0.70, 0.88
Hours in on-job driving/riding	0.95	0.92, 0.97
Vehicle type	0.86	0.85, 0.88
None / Car / Motorcycle / Truck / Bus / Tractor / Others		
Most common weight multiplied by frequency in lifting per day	0.49	0.28, 0.65
Heaviest weight, at least once a week	0.72	0.58, 0.82
Hours in overhead or stand work positions	0.36	0.13, 0.55
Hours in slight twisting and bending positions	0.33	0.10, 0.53
Hours in moderate bending/twisting	0.46	0.24, 0.63
Hours in deep bending and twisting	0.79	0.68, 0.86
Hours in standing/walking when work	0.59	0.41, 0.73
Work hours per week	0.67	0.52, 0.79
Commute time in going to work	0.67	0.51, 0.78
Mode of transportation	0.70	0.62, 0.82
Associated injuries when work	0.65	0.21, 1.00
Work-related sudden back pain/injuries	0.74	0.55, 0.94
Has your health ever influenced your choice, or a change of jobs?	1.00	

κ: Cohen's kappa; ICC: intra-class correlation coefficient; 95% CI: 95% confidence interval

spinal phenotypes and their measurement accuracy. In addition, the comprehensive back pain questionnaire, which is typically absent in most epidemiological studies of the spine, was adapted to measure the presence, degree, and frequency of various back pain phenotypes. Using detailed, comprehensive methodologies, the Hangzhou Lumbar Spine Study is expected to provide scientific evidence for many research questions related to the lumbar spine and back pain.

There are also some limitations of the Hangzhou Lumbar Spine Study. Although subjects were randomly recruited from a large community, the study does not represent all Chinese in the mainland. There

Table 4 Test-retest reliability for back pain history in Caihe community (*n*=40)

Item	*κ* or ICC	95% CI
Low back pain today		
Do you have low back pain (pain between the lower ribs and buttock crease) today? (yes/no)	0.86	0.56, 1.00
How would you rate your back symptoms today on a scale from 0 to 10, with 0 being no pain and 10 being the worst pain imaginable?	0.96	0.92, 0.98
Low back pain over the past 4 weeks		
Have you had pain in your lower back in the past 4 weeks?	0.82	0.62, 1.00
Have you had pain that goes down the leg?	1.00	
Low back pain over the past 12 months		
How often have you experienced low back pain? Daily / Not daily, but at least once a week / Not weekly, but at least once / A month / Several times a year / Two to three times a year / Once a year / None at all	0.60	0.42, 0.68
How would you rate your worst low back ache/pain over the past 12 months on a scale from 0 to 100, with 0 being no pain and 100 being the worst pain imaginable?	0.68	0.48, 0.82
Low back in lifetime		
Have you ever fractured your spine?	0.66	0.03, 1.00
Have you ever had sudden back pain lasting more than 1 d related to an injury/trauma or an unusually heavy load?	0.88	0.72, 1.00
Have you ever had sudden back pain lasting more than 1 d without an injury/trauma or an unusually heavy load?	0.60	0.32, 0.89
Have you ever had spine surgery?	1.00	
Were you ever unable to work (or do your normal activities) for at least a month due to back pain or back injury?	1.00	

κ: Cohen's kappa; ICC: intra-class correlation coefficient; 95% CI: 95% confidence interval

are 56 ethnicities in China; we recruited only Han Chinese subjects. Further, the overall participation rate was low, which may introduce sampling bias. Unfortunately, this is an issue with other studies as well due to low volunteer rates (Moore et al., 2016) or recruitment through public media (Mok et al., 2010). There are many reasons underlying the low participation rate in our study. About a quarter of randomly selected subjects could not be contacted through phone numbers in the registry, largely due to the transitory nature of cell phone numbers used. The most common reason for residents declining participation in the study was their unwillingness to spend the time required. Most subjects younger than 24 years old were currently studying in other cities and were unavailable to participate. Also, although a similar percentage of men and women were selected in random sampling, many more women participated in our study than men, which is a common phenomenon in subject recruitment (Todd et al., 1983). Finally, in the current study measurement of psychological factors is limited, which could be of interest for studying back pain.

5 Conclusions

To our knowledge, the Hangzhou Lumbar Spine Study is the first population-based epidemiological study conducted for understanding spinal phenotypes and back health in mainland Chinese. As the study continues and the sample size grows, the data set will be used to test many related hypotheses. Epidemiological information obtained from a reliable questionnaire, MR imaging data, DXA measurements, and DNA analysis may serve as a valuable reference for future studies on back health.

Acknowledgements

We are indebted to Ms. Ying YUAN for her tremendous work conducting the study. Without her valuable input as the study's research nurse, the study would not be possible. We thank the clinic staff at the Department of Orthopedic Surgery, Radiology, Clinical laboratory Center at the First Affiliated Hospital of Zhejiang University, and the community staff in Caihe, for their continuous assistance in conducting the study.

Compliance with ethics guidelines

Xiao-jian HU, Lun-hao CHEN, Michele C. BATTIÉ, and Yue WANG declare that they have no conflicts of interest.

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Informed consent was obtained from all subjects for being included in the study.

References

Abraham I, Killackey-Jones B, 2002. Lack of evidence-based research for idiopathic low back pain: the importance of a specific diagnosis. *Arch Intern Med*, 162 (13):1442-1444. https://doi.org/10.1001/archinte.162.13.1442

Albert HB, Manniche C, 2007. Modic changes following lumbar disc herniation. *Eur Spine J*, 16(7):977-982. https://doi.org/10.1007/s00586-007-0336-8

Andersson GBJ, 1999. Epidemiological features of chronic low-back pain. *Lancet*, 354(9178):581-585.

Balagué F, Mannion AF, Pellise F, et al., 2012. Non-specific low back pain. *Lancet*, 379(9814):482-491. https://doi.org/10.1016/S0140-6736(11)60610-7

Battié MC, Videman T, 2006. Lumbar disc degeneration: epidemiology and genetics. *J Bone Joint Surg Am*, 88(Suppl 2):3-9.

Battié MC, Videman T, Gibbons LE, et al., 1995. Determinants of lumbar disc degeneration: a study relating lifetime exposures and magnetic resonance imaging findings in identical twins. *Spine*, 20(24):2601-2612. https://doi.org/10.1097/00007632-199512150-00001

Battié MC, Videman T, Levalahti E, et al., 2007. Heritability of low back pain and the role of disc degeneration. *Pain*, 131(3):272-280. https://doi.org/10.1016/j.pain.2007.01.010

Bush K, 1998. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med*, 158(16):1789-1795. https://doi.org/10.1001/archinte.158.16.1789

Deyo RA, Weinstein JN, 2001. Low back pain. *New Engl J Med*, 344(5):363-370.

Dionne CE, Dunn KM, Croft PR, et al., 2008. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine*, 33(1):95-103. https://doi.org/10.1097/BRS.0b013e31815e7f94

Halpern M, Hiebert R, Nordin M, et al., 2001. The test-retest reliability of a new occupational risk factor questionnaire for outcome studies of low back pain. *Appl Ergon*, 32(1): 39-46. https://doi.org/10.1016/S0003-6870(00)00045-4

Jarvik JG, Hollingworth W, Heagerty PJ, et al., 2005. Three-year incidence of low back pain in an initially asymptomatic cohort: clinical and imaging risk factors. *Spine*, 30(13):1541-1548. https://doi.org/10.1097/01.brs.0000167536.60002.87

Jarvinen J, Karppinen J, Niinimaki J, et al., 2015. Association between changes in lumbar Modic changes and low back symptoms over a two-year period. *BMC Musculoskelet Disord*, 16:98. https://doi.org/10.1186/s12891-015-0540-3

Jensen MC, Brant-Zawadzki MN, Obuchowski N, et al., 1994. Magnetic resonance imaging of the lumbar spine in people without back pain. *New Engl J Med*, 331(2):69-73. https://doi.org/10.1056/NEJM199407143310201

Kaapa E, Luoma K, Pitkaniemi J, et al., 2012. Correlation of size and type of Modic types 1 and 2 lesions with clinical symptoms: a descriptive study in a subgroup of patients with chronic low back pain on the basis of a university hospital patient sample. *Spine*, 37(2):134-139. https://doi.org/10.1097/BRS.0b013e3182188a90

Kjaer P, Leboeuf-Yde C, Sorensen JS, et al., 2005a. An epidemiologic study of MRI and low back pain in 13-year-old children. *Spine*, 30(7):798-806. https://doi.org/10.1097/01.brs.0000157424.72598.ec

Kjaer P, Leboeuf-Yde C, Korsholm L, et al., 2005b. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine*, 30(10):1173-1180. https://doi.org/10.1097/01.brs.0000162396.97739.76

Kuisma M, Karppinen J, Niinimaki J, et al., 2007. Modic changes in endplates of lumbar vertebral bodies prevalence and association with low back and sciatic pain among middle-aged male workers. *Spine*, 32(10):1116-1122. https://doi.org/10.1097/01.brs.0000261561.12944.ff

Kuisma M, Karppinen J, Haapea M, et al., 2009. Modic changes in vertebral endplates: a comparison of MR imaging and multislice CT. *Skeletal Radiol*, 38(2):141-147. https://doi.org/10.1007/s00256-008-0590-9

Landis JR, Koch GG, 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159-174. https://doi.org/10.2307/2529310

Maatta JH, Wadge S, MacGregor A, et al., 2015. ISSLS prize winner: vertebral endplate (Modic) change is an independent risk factor for episodes of severe and disabling low back pain. *Spine*, 40(15):1187-1193. https://doi.org/10.1097/BRS.0000000000000937

MacGregor AJ, Andrew T, Sambrook PN, et al., 2004. Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. *Arthritis Care Res*, 51(2):160-167. https://doi.org/10.1002/art.20236

Mok FP, Samartzis D, Karppinen J, et al., 2010. ISSLS prize winner: prevalence, determinants, and association of Schmorl nodes of the lumbar spine with disc degeneration: a population-based study of 2449 individuals. *Spine*, 35(21):1944-1952. https://doi.org/10.1097/BRS.0b013e3181d534f3

Mok FP, Samartzis D, Karppinen J, et al., 2016. Modic changes of the lumbar spine: prevalence, risk factors, and association with disc degeneration and low back pain in a large-scale population-based cohort. *Spine J*, 16(1):32-41. https://doi.org/10.1016/j.spinee.2015.09.060

Moore C, Bolton T, Walker M, et al., 2016. Recruitment and representativeness of blood donors in the INTERVAL randomised trial assessing varying inter-donation intervals.

Trials, 17(1):458.
https://doi.org/10.1186/s13063-016-1579-7

Pincus T, Santos R, Breen A, et al., 2008. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum*, 59(1):14-24.
https://doi.org/10.1002/art.23251

Rahme R, Moussa R, 2008. The Modic vertebral endplate and marrow changes: pathologic significance and relation to low back pain and segmental instability of the lumbar spine. *Am J Neuroradiol*, 29(5):838-842.
https://doi.org/10.3174/ajnr.A0925

Salminen JJ, Erkintalo M, Laine M, et al., 1995. Low back pain in the young. A prospective three-year follow-up study of subjects with and without low back pain. *Spine*, 20(19): 2101-2108.
https://doi.org/10.1097/00007632-199510000-00006

Teraguchi M, Yoshimura N, Hashizume H, et al., 2015. The association of combination of disc degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study. *Spine J*, 15(4):622-628.
https://doi.org/10.1016/j.spinee.2014.11.012

Todd M, Davis KE, Cafferty TP, 1983. Who volunteers for adult development research?: research findings and practical steps to reach low volunteering groups. *Int J Aging Hum Dev*, 18(3):177-184.
https://doi.org/10.2190/AYK2-QRJ7-N8PU-M9B5

Wang Y, Videman T, Battié MC, 2012a. ISSLS prize winner: lumbar vertebral endplate lesions: associations with disc degeneration and back pain history. *Spine*, 37(17):1490-1496.
https://doi.org/10.1097/BRS.0b013e3182608ac4

Wang Y, Videman T, Battié MC, 2012b. Modic changes: prevalence, distribution patterns, and association with age in white men. *Spine J*, 12(5):411-416.
https://doi.org/10.1016/j.spinee.2012.03.026

Wang Y, Battié MC, Videman T, 2012c. A morphological study of lumbar vertebral endplates: radiographic, visual and digital measurements. *Eur Spine J*, 21(11):2316-2323.
https://doi.org/10.1007/s00586-012-2415-8

Wang Y, Videman T, Battié MC, 2013. Morphometrics and lesions of vertebral end plates are associated with lumbar disc degeneration: evidence from cadaveric spines. *J Bone Joint Surg Am*, 95(5):e26.
https://doi.org/10.2106/JBJS.L.00124

Weir JP, 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 19(1):231-240.
https://doi.org/10.1519/00124278-200502000-00038

World Health Organization, 2002. Reducing Risks, Promoting Healthy Life: the World Health Report 2002. http://www.who.int/whr/2002/en [Accessed on Sept. 10, 2015].

中文概要

题　目：基于中国一般人群腰痛的杭州腰椎研究的方法学及进展介绍

目　的：开展基于大样本人群的杭州腰椎疾患研究，研究国人一般人群腰痛、腰椎间盘退变、终板病损、Modic 改变的流行病学特征以及病理机制。

创新点：开展了中国大陆第一个基于大样本的有关腰痛、腰椎间盘退变、终板病损、Modic 改变的流行病学研究；开发了一份可信度高的包含腰痛病史以及相关风险因素的调查问卷；采用高分辨率 3.0 T 核磁共振来分辨腰椎病变特征。

方　法：从杭州采荷社区中随机抽取人群，完成调查问卷腰痛相关信息记录、腰椎核磁共振扫描、腰椎及股骨颈骨密度测量以及血液样本抽取。从参与人群重复抽取 40 人进行问卷可信度测量。调查问卷包含的信息包括，人群一般资料、吸烟史、腰痛史、职业史以及运动史，用来记录腰痛发生情况以及相关风险因素。腰椎核磁共振扫描用于观察椎间盘退变、Modic 改变以及终板病损等腰椎特征性病变。骨密度测量用来研究与腰椎特征性病变的联系。血液样本用于后续遗传因素的分析。

结　论：目前已完成 301 名社区人员调查问卷腰痛相关信息记录、腰椎核磁共振扫描、腰椎及股骨颈骨密度测量以及血液样本抽取；有关腰痛以及相关风险因素的调查问卷可信度高；杭州腰椎疾患研究可对国人腰痛、腰椎间盘退变、终板病损、Modic 改变的流行病学特征以及进一步的病理机制研究提供有价值的信息。

关键词：腰痛；大样本研究；椎间盘退变；Modic改变；终板缺损；方法学