Journal of Exercise Science & Fitness 19 (2021) 216–222

Contents lists available at ScienceDirect

Journal of Exercise Science & Fitness

journal homepage: www.elsevier.com/locate/jesf

Temporal trends in step test performance for Chinese adults between 2000 and 2014

Yang Liu [a, b], Mark S. Tremblay [c], Grant R. Tomkinson [d, e, *]

[a] School of Physical Education and Sport Training, Shanghai University of Sport, Shanghai, China
[b] Shanghai Research Center for Physical Fitness and Health of Children and Adolescents, Shanghai University of Sport, Shanghai, China
[c] Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada
[d] Department of Education, Health and Behavior Studies, University of North Dakota, Grand Forks, ND, USA
[e] Alliance for Research in Exercise, Nutrition and Activity (ARENA), Allied Health and Human Performance, University of South Australia, Adelaide, SA, Australia

A R T I C L E I N F O

Article history:
Received 27 March 2021
Received in revised form
6 July 2021
Accepted 6 July 2021
Available online 12 July 2021

Keywords:
China
Cardiorespiratory fitness
Exercise testing

A B S T R A C T

Backgroud/Objective: Cardiorespiratory endurance is an excellent marker of functional endurance and health among adults. The aim of this study was to estimate temporal trends in step test performance for Chinese adults between 2000 and 2014.
Methods: Apparently healthy adults aged 20–59 years were included. Nationally representative step test data (n = 603,977) from 2000, 2005, 2010, and 2014 were reported descriptively by the China Physical Fitness Surveillance Center. Temporal trends in means were estimated at the sex-age level for all adults and separate location/occupation groups using sample-weighted linear regression, with trends in distributional characteristics described visually and estimated as the ratio of coefficients of variation (CVs).
Results: Collectively, there was a negligible improvement in mean step test performance of 0.12 standardized effect sizes (95% confidence interval (95%CI): 0.11–0.13). Negligible to small improvements were observed for all age, sex, location, and occupation groups. Variability declined substantially over time (ratio of CVs (95%CI): 0.86 (0.86–0.86)), with negligible to large improvements in those below the 10th percentile, and negligible to moderate declines in those above the 90th percentile.
Conclusion: There have been negligible to large improvements in step test performance for low to average performing Chinese adults since 2000, which may be meaningful to public health because low endurance is an important risk factor for all-cause mortality.

Introduction

Cardiorespiratory endurance (CRE) is the ability to perform large-muscle, whole-body, moderate to vigorous intensity physical activity for long periods.[1] Low CRE is significantly associated with all-cause and cardiovascular mortality (independent of age, sex, smoking, alcohol, self-reported physical activity levels, socioeconomic status, comorbidities and other covariates),[2–5] cancer (e.g., respiratory, thoracic, central nervous system, hematological),[5] stroke,[6] type 2 diabetes,[7] hypertension,[8] poorer quality of life,[9] cognitive declines,[10] and functional impairments.[11] The American Heart Association[12] recommends CRE as a clinical vital sign because it substantially improves the net reclassification of patients in clinical settings when added to traditional risk factors.[13] This health-related evidence supports the promotion of aerobic exercise in national[14–16] and global[17] physical activity guidelines for adults. For these reasons, temporal trends in CRE may provide important insights into corresponding trends in population health.

CRE can be measured using a variety of maximal or submaximal field- or laboratory-based run/walk, cycling, or step tests. Although indirect calorimetry using expired gas analysis is the criterion measure of maximal oxygen uptake ($\dot{V}O_{2max}$) during progressive aerobic exercise to volitional fatigue, such testing is time-

* Corresponding author. Department of Education, Health and Behavior Studies, University of North Dakota, 2751 2nd Avenue North, Stop 8235, Grand Forks, ND, 58202, USA.
E-mail address: grant.tomkinson@und.edu (G.R. Tomkinson).

https://doi.org/10.1016/j.jesf.2021.07.001
1728-869X/

consuming, expensive, ecologically invalid, and requires participants to exercise to maximal effort, which increases the risk of adverse events for some individuals.[18,19] Submaximal exercise testing on the other hand is simple, inexpensive, can easily be administered to large groups of individuals simultaneously, and because exercise is performed at a lower intensity, the risk of adverse events is lower compared to maximal testing.[18–20] Step tests, which measure heart rate during recovery from a submaximal stepping task, are a feasible, ecologically valid, and scalable measure of functional endurance (i.e., the ability to perform submaximal ambulatory aerobic exercise using a familiar activity of daily living) for clinical screening and population surveillance.[18,21,22] In adults, step testing has moderate to very high criterion validity for estimating gas-analyzed $\dot{V}O_{2max}$ (in ml/kg/min),[18] a small standard error of estimate (<7% of mean measured $\dot{V}O_{2max}$),[18] nearly perfect test-retest reliability,[18,22,23] and is generally safe and well tolerated.[18,21,22]

We recently published a large meta-analysis of temporal trends in functional endurance for adults, which showed a negligible (per decade) international decline of 0.13 standardized (Cohen's) effect sizes (ES), or 1.6%, across 2,525,827 adults (aged 18–59 years) representing eight high- and upper-middle-income countries over the period 1967–2016.[24] As part of our meta-analysis, we found a small (per decade) decline in mean distance running performance (800-m (women) and 1000-m (men)) of 0.34 ES, or 3.8%, from 1985 to 2014 in a nationally representative sample of 499,229 Chinese adults aged 18–22 years. To the best of our knowledge, no study has reported on recent trends in functional endurance for a nationally representative sample of Chinese adults not aged between 18 and 22 years. Furthermore, it is not known whether there have been temporal trends in the distributional characteristics of adult functional endurance (i.e., trends in distributional variability and/or asymmetry). Such knowledge would help to identify whether trends in mean functional endurance were uniform (i.e., similar in those with low, medium, or high endurance) or non-uniform (i.e., dissimilar in those with low, medium, or high endurance) across the distribution. This issue is potentially important in understanding both how the trends in functional endurance came about and how best to address them (e.g., if declining).

Examination of national temporal trends in adult functional endurance will not only improve insights into population health for such countries, but also will facilitate between-country comparisons, improve knowledge of global and regional trends, provide insight for healthy public polices and interventions, and potentially predict future disease burden. Over the past two decades, the China Physical Fitness Surveillance Center has conducted national fitness surveillance of 20-to 59-year-olds in 2000, 2005, 2010, and 2014, and published summary statistics for functional endurance as step test performance in their "Report on National Physical Fitness Surveillance".[25–28] Using these nationally representative summary fitness data, the primary aim of this study was to estimate temporal trends (in means and distributional characteristics) in step test performance for Chinese adults (aged 20–59 years) between 2000 and 2014.

Methods

Participants and sampling procedures

Using complex, stratified, multistage probability cluster sampling, nationally representative samples of civilian, non-institutionalized Chinese adults (aged 20–59 years) were measured for functional endurance using submaximal bench stepping in 2000, 2005, 2010, and 2014. The sampling procedure across the four national surveys has been described in detail elsewhere.[29] Briefly, the sampling procedure comprised six stages: (1) covered all 31 provinces, autonomous regions, and municipalities across Mainland China; (2) the random selection of three cities from each province based on economic status, operationalized as city-level gross domestic product (GDP); (3) the random selection of three districts from urban areas, and three counties from rural areas, from selected cities; (4) the random selection of three streets from selected districts and three towns from selected counties; (5) the random selection of two residential communities from selected streets and two villages from selected towns; and (6) the systematic sampling (to ensure equal numbers) of eligible participants from selected residential communities and villages who had been living at their current residence for the past three years. All four national fitness surveys were conducted between April and October, with fitness test results statistically processed and reported as summary statistics (sample sizes, means, standard deviations (SDs), and percentiles).[25–28] Because only summary statistics were available and used in our analysis, this study was exempted from review by the university institutional review board at the authors' respective universities and from subject informed consent.

Prior to sample selection, all potential participants were asked whether they had any known genetic, chronic or acute diseases, and/or disabilities, which could affect their ability to participate in exercise. Unfortunately, no data on medical exclusions were available. Collectively, functional endurance data were available for 603,977 apparently healthy adults (i.e., free from known disease and/or disability; men, 50.1% $n = 302{,}386$; women, 49.9% $n = 301{,}591$) aged 20–59 years over the period 2000–2014. The mean sex-age-location-year-specific sample sizes were 3101 (range: 2726–3309) and 6336 (range: 5773–6684) for rural and urban adults, respectively. Among urban adults, the mean sex-age-occupation-year-specific sample sizes were 3143 (range: 2860–3303) and 3192 (range: 2879–3424) for manual and non-manual laborers, respectively.

Testing procedures

Functional endurance was estimated by trained technicians using a standardized single-stage, fixed-height, and fixed-rate step test protocol.[25–28] Participants were asked to follow the pre-testing guidelines, which included no strenuous exercise or heavy labor within 12 h of testing, dressing in sportswear (e.g., shorts and a short-sleeved shirt), and warming-up prior to exercise testing to avoid injury.[29] After 5 min of quiet rest, participants stepped up and down a 30-cm (men) or 25-cm (women) step for 3 min at the rate of 30 repetitions per minute, which was set by a metronome. When the signal started, each participant placed one foot on the step, stepped up, placed the other foot on the platform, straightened both legs and stood upright, and then stepped down, bringing down the same foot that was first placed up. This sequence constituted 1 repetition, with participants repeating the sequence for as long as possible, up to 3 min, with the test stopped by technicians (and the test duration subsequently recorded) if participants could not maintain the required stepping rate for 3 consecutive repetitions. After the completion of the test, participants were asked to sit quietly on a chair while heart rate (bpm) was measured by a fingertip pulse oximeter for three 30-s intervals at 1-, 2-, and 3-min post-exercise. Step test performance was calculated using the following step test index formula[30] (Eq. (1)):

$$STI = \frac{100t}{HR_1 + HR_2 + HR_3} \quad (1)$$

where STI = step test index, t = time (seconds) of the step test,

HR_1 = heart rate (bpm) at 1.0–1.5 min post-test, HR_2 = heart rate (bpm) at 2.0–2.5 min post-test and HR_3 = heart rate (bpm) at 3.0–3.5 min post-test.

Participants self-reported their age and sex. Location (rural and urban) was defined by the sampling area. Occupation (manual labor (e.g., construction worker, factory employee, waiter) and non-manual labor (e.g., teacher, doctor, office worker)) was classified (for urban adults only) by survey staff according to the sampling unit (e.g., if the sampling unit was a construction company, then all builders were classified as manual labor and all administration staff were classified as non-manual labor).

Statistical analyses

Temporal trends in mean step test performance were analyzed from summary statistics at the sex-age group level for all Chinese adults (aged 20–59 years), and for separate location (rural and urban) and occupation (manual labor and non-manual labor) groups, using sample-weighted linear regression models relating the year of testing to mean step test performance.[31] Linear models were used because they naturally summarized the overall temporal trends. Because STI data are interval, we expressed the overall linear trends as standardized ES (i.e., the regression coefficient divided by the pooled SD), which facilitated comparisons between groups and with published studies. To interpret the magnitude of rates of change, ES of 0.2, 0.5, and 0.8 were used as thresholds for small, moderate, and large, respectively, with ES < 0.2 considered to be negligible.[32] Positive rates of change indicated temporal improvements and negative rates of change indicated temporal declines.

Using procedures described elsewhere,[31] we calculated population-weighted mean changes (for 20-to 29-year-olds, 30-to 39-year-olds, 40-to 49-year-olds and 50-to 59-year-olds, plus young adults (20-to 39-year-olds), middle-aged adults (40-to 59-year-olds), and all adults), sex (men and women), location, and occupation groups by pooling the rates of change across all relevant groups using post-stratification population weighting.[33] Population weights were obtained from sex-age-specific population estimates for China in 2000 using United Nations data.[34]

Temporal trends in distributional characteristics were examined visually and as trends in the coefficient of variation (CV, the ratio of the SD to the mean). Trends in CVs were analyzed as the ratio of CVs, by dividing the 2014 CVs by the 2000 CVs using the procedure described elsewhere.[35] Ratios >1.1 indicated substantial increases in variability (i.e., the magnitude of variability in relation to the mean increased over time), ratios <0.9 indicated substantial declines in variability (i.e., the magnitude of variability in relation to the mean decreased over time), and ratios between 0.9 and 1.1 indicated negligible trends in variability (i.e., the magnitude of variability in relation to the mean did not change substantially over time).[36] Finally, distributional trends were examined visually using LOWESS (LOcally WEighted Scatter-plot Smoother) curves (tension = 66),[37] by plotting the standardized ES changes in step test performances between 2000 and 2014 for a range of percentiles (from the 3rd to the 97th) against the percentiles themselves.

Results

Collectively, there was a negligible improvement in step test performance of 0.12 ES (95%CI: 0.11–0.13) between 2000 and 2014, with negligible improvements observed for all age (except 50-to 59-year-olds who experienced a small improvement), sex, location, and occupation groups (Table 1). There were also negligible age- (ES [95% CI]: young adults, 0.07 [0.06–0.08]; middle-aged adults, 0.20 [0.19–0.21]), sex- (ES [95% CI]: men, 0.08 [0.07–0.09]; women, 0.17 [0.16–0.18]), location- (ES [95% CI]: rural, 0.05 [0.04–0.06]; urban, 0.16 [0.15–0.17]), and occupation-related (ES [95% CI]: manual labor, 0.14 [0.13–0.15]; non-manual labor, 0.18 [0.17–0.19]) temporal differences.

Trends in mean step test performance were not uniform across the population, as indicated by trends in distributional variability and asymmetry. Table 1 indicates that variability in step test performance declined substantially over the entire period (ratio of CVs [95%CI]: 0.86 [0.86 to 0.86]). Variability declined substantially in both sex (ratio of CVs [95%CI]: men, 0.88 [0.87–0.89]; women, 0.83 [0.82–0.84]) and location (ratio of CVs [95%CI]: rural, 0.84 [0.83–0.85]; urban, 0.88 [0.87–0.89]) groups; in middle-aged adults (ratio of CVs [95%CI]: 0.82 [0.81–0.83]) but not young adults (aged 20–29 years, ratio of CVs [95%CI]: 0.90 [0.89–0.91]); and in manual urban laborers (ratio of CVs [95%CI]: 0.85 [0.84–0.86]) but not non-manual urban laborers (ratio of CVs [95% CI]: 0.91 [0.90–0.92]). There were also trends in distributional asymmetry (Fig. 1). While negligible improvements were almost always observed in the average performers (i.e., at the 50th percentile), small to large (equivalent to 0.36–1.04 ES or 14–35 percentile points) improvements were observed in very low performers (3rd percentile), and negligible to small (equivalent to 0.11–0.48 ES or 4–18 percentile points) improvements were observed in low performers (10th percentile). In contrast, negligible changes (equivalent to −0.19 to 0.13 ES or −7 to 5 percentile points) and negligible to moderate declines (equivalent to −0.52 to −0.07 ES or −2 to −20 percentile points) were observed in high performers (90th and 97th percentiles, respectively).

Discussion

This study estimated temporal trends in functional endurance operationalized as step test performance for 603,977 Chinese adults aged 20–59 years between 2000 and 2014. The principal findings were that (a) there was a negligible population-level improvement across the 14-year period for all age, sex, location, and occupation groups (except 50-to 59-year-olds), equivalent to 0.12 ES; and (b) variability in step test performance declined substantially over time, as a result both ends of the distribution moving towards the middle. Our finding of improved step test performance, at least among the low and average performers, is suggestive of corresponding improvements in functional endurance of those in greatest need. Improved functional endurance may be important to public health, especially for those with low endurance (e.g., those at the low end of the distribution and older middle-aged adults), because low endurance is an important risk factor for all-cause mortality, chronic diseases and functional impairments,[2–11] and because China's population is aging.[38] Moreover, increased national efforts that promote health and fitness are required to further improve population-level functional endurance.

Explanation of main findings

We have previously argued that temporal trends in functional endurance are probably influenced by trends in a network of physiological, physical, behavioral, social, and/or environmental factors.[24,39] Because adult step test performance has moderate to very high criterion validity,[18] trends in step test performance are suggestive of corresponding trends in underlying $\dot{V}O_{2max}$ (and/or relative oxygen transport capability) and other aerobic factors (i.e., mechanical efficiency and fractional utilization of oxygen). Trends in psychosocial factors (e.g., motivation and/or tolerance of bench stepping aerobic exercise) may also be involved.[24,39]

Concurrent trends in body size/composition and physical

Table 1
Temporal trends in means and variability for step test performance among 20-to 59-year-old Chinese adults between 2000 and 2014.

Group	n	mean ± SD	Standardized ES change in means (95%CI)	Ratio of CVs (95%CI)
Age				
20–29 years	153,482	56.7 ± 8.8	0.07 (0.06–0.08)	0.90 (0.89–0.91)
30–39 years	153,107	57.5 ± 8.9	0.08 (0.07–0.09)	0.88 (0.87–0.89)
40–49 years	150,718	58.9 ± 9.7	0.16 (0.15–0.17)	0.83 (0.82–0.84)
50–59 years	146,670	58.8 ± 10.4	0.26 (0.25–0.27)	0.81 (0.80–0.82)
Sex				
Men	302,386	57.4 ± 9.4	0.08 (0.07–0.09)	0.88 (0.87–0.89)
Women	301,591	59.1 ± 9.9	0.17 (0.16–0.18)	0.83 (0.82–0.84)
Location				
Rural	198,493	59.3 ± 10.2	0.05 (0.04–0.06)	0.84 (0.83–0.85)
Urban	405,484	57.7 ± 9.3	0.16 (0.15–0.17)	0.88 (0.87–0.89)
Occupation				
Manual	201,167	57.9 ± 9.4	0.14 (0.13–0.15)	0.85 (0.84–0.86)
Non-manual	204,317	57.5 ± 9.2	0.18 (0.17–0.19)	0.91 (0.90–0.92)
All				
20–59 years	603,977	58.3 ± 9.7	0.12 (0.11–0.13)	0.86 (0.86–0.86)

Notes: Positive changes in means indicated temporal improvements in means and negative changes indicated temporal declines in means; standardized changes (ES) in means of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large, respectively, with ES < 0.2 considered to be negligible; ratio of CVs >1.1 indicated substantial temporal increases in variability and ratios <0.9 indicated substantial temporal declines. Occupational trends were only available for adults from urban areas. Mean ± SD step test index values are also shown.
Abbreviations: SD = standard deviation; CV = coefficient of variation; n = sample size; 95%CI = 95% confidence interval; ES = effect size.

activity levels are probably also involved.[24,39] Temporal data from China's Reports on National Physical Fitness Surveillance[25–28] indicate increases in both height and body mass (across all age, sex, location, and occupation strata) between 2000 and 2014. Increased height probably reflects increased leg and stride lengths,[40,41] resulting in a reduced energy requirement of bench stepping.[42] Increased body mass probably also reflects increases in both fat mass and fat-free muscle mass.[29,43] Increased fat mass will negatively affect step test performance by increasing the energy requirement of stepping without contributing to it. Conversely, increased fat-free muscle mass will positively affect step test performance by improving the metabolic potential of the exercising muscles (e.g., increased glycogen storage capacity, increased number of mitochondria) and possibly mechanical efficiency (e.g., reduced antagonist co-activation needed for stability).[44] Using data from China's Reports on National Physical Fitness Surveillance, Tian et al.[29] reported an increase in the prevalence of 20-to 59-year-old Chinese adults meeting the minimum leisure-time physical activity recommendation (i.e., 150 min of moderate-intensity exercise or 75 min of vigorous intensity exercise per week, or an equivalent combination) from 17.2% in 2000 to 22.8% in 2014. This temporal increase in leisure-time physical activity may have been influenced by *The National Fitness Program*, which was implemented under Chinese Sport Law in 1995 by the State Council of the People's Republic of China. This program encourages Chinese people to engage in regular physical activity to improve their physical fitness and health. However, it is unclear whether there have been concurrent trends in occupational, household, or transport physical activity levels, which significantly contribute to total physical activity levels. Assuming the temporal connection between leisure-time physical activity levels and step test performance is causal, then the national promotion of physical activity[16] might be a suitable population approach to improving the functional endurance of Chinese adults.

Our finding of a substantial decline in the variability of step test performances indicated that the improvement in mean step test performance was not uniform across the population. Our visual analysis indicated that this decline was because both ends of the distribution had moved towards the middle. In the absence of criterion-referenced health-related cut-points for functional endurance in Chinese adults, and in light of research showing that mortality and morbidity are disproportionately higher in the lowest quintile for functional endurance,[3,4] our finding of improved step test performance for adults below the 20th percentile suggests that Chinese adults in 2014 were collectively at reduced risk of negative health outcomes compared to adults in 2000, which potentially indicates a reduction in chronic disease burden.

Comparisons with other studies

Although studies examining temporal trends in mean adult functional endurance are scarce, specific to China, we have previously reported a small (per decade) decline in long-distance running performance in a nationally representative sample of young Chinese adults (aged 18–22 years) between 1985 and 2014.[24] While it is challenging to describe the temporal differences in functional endurance for Chinese adults between the present study and our previous work,[24] it is possible that differences in exercise modalities (bench stepping vs. running), effort criteria (submaximal vs. maximal), pacing strategies (externally fixed-rate stepping vs. self-selected running pace), medical screening criteria, age groups, and the span of testing years were involved.

Furthermore, studies examining temporal trends in the distributional characteristics of adult functional endurance, as opposed to mean values, are even rarer. To our knowledge, only one other study has examined trends in the distributional characteristics of adult functional endurance over a similar time period and using a similar statistical approach, which allows for direct comparison. Using nationally representative data on 103,505 older Japanese adults (aged 65–79 years) between 1998 and 2017, we recently found a substantial decline in the variability (i.e., ratio of CVs <0.9) of functional endurance (operationalized as 6-min walking distance (6MWD)), which indicated that the small (per decade) improvement in mean 6MWD was not uniform across the distribution.[35] Unlike in our present study, because we were unable to examine concurrent trends in distributional asymmetry, we could not determine whether the low performers, high performers, or both, had changed over time in older Japanese adults. We therefore recommend that future studies examining temporal trends in mean functional endurance also examine trends in distributional characteristics.

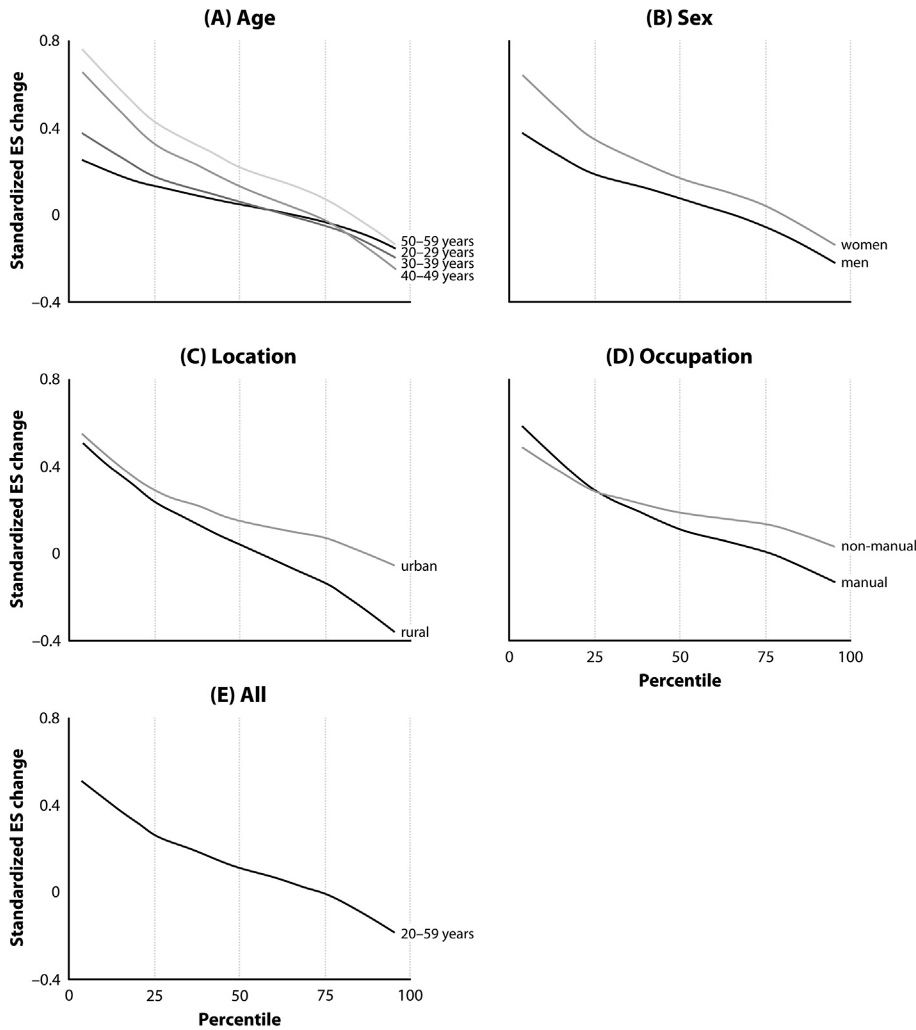

Fig. 1. Distributional trends in step test performance for 20-to 59-year-old Chinese adults between 2000 and 2014.
Notes: Distributional trends are shown for different (A) age, (B) sex, (C) location and (D) occupation groups, as well as (E) all Chinese adults aged 20–59 years; trends between 2000 and 2014 are shown as standardized ES, with positive values indicating improvements and negative values indicating declines; the solid lines are the LOWESS (LOcally WEighted Scatter-plot Smoother) curves (tension = 66), which are used to represent the trends at various percentiles (range: 3rd to 97th), with sloping lines indicating asymmetric trends (i.e., lines that sloped upwards from the bottom left to the top right indicated relatively larger improvements in the high performers, and lines that sloped downwards from the top left to the bottom right indicated relatively larger declines in the low performers) and flat (horizontal) lines indicated symmetric trends (i.e., uniform trends across all percentiles or step test performance levels).

Strengths and limitations

This was the first study to examine temporal trends in step test performance for Chinese adults. We used nationally representative fitness data from repeated cross-sectional samples, which were collected by trained technicians approximately every 5 years, at the same time of year, using standardized testing protocols and a consistent sampling strategy. The sample-weighted regression and post-stratified population weighting procedure adjusted our trends for sampling bias by incorporating the underlying population demographics, and our stratified analysis controlled for potential confounding (e.g., age, sex, location, occupation).

This study was however limited because we estimated temporal trends using only summary statistics (because raw data were not available), which did not impact our trends, but did mean we were unable to statistically account for the effects of trends in underlying mechanistic factors (e.g., body size, physical activity levels) or assess the influence of trends in step test performance on health outcomes. We were unable to examine period and cohort effects because we estimated trends using only summary statistics based on repeated cross-sectional samples rather than mixed longitudinal samples. We also could not determine whether our trends were affected by aggregation bias related to differences in age distributions within age bands. Furthermore, we could not extend our temporal analysis beyond 2014 because step testing performance was not measured in the subsequent 2019 national survey.

The overall response rate in each survey was similar (range: 76–84%).[29] Although pre-screening was performed prior to sample selection with only apparently healthy adults included, no additional pre-exercise screening was performed. Nonetheless, the published summary statistics for step test performance could have been biased if less fit/less healthy and physically impaired adults either opted out or were excluded. While this may have reduced the generalizability of our findings to the entire 20-to 59-year-old Chinese population, it is challenging to determine whether our trends were biased without evidence of temporal trends in the number of adults who either opted out or were medically excluded. Finally, while single-stage, fixed-height, and fixed-rate step tests are conducive to group testing, they may not present the same relative cardiorespiratory challenge to all participants which could affect the validity of the test, and may elicit vigorous exercise intensities leading to adverse events for some individuals, thus eliminating their benefit as a submaximal exercise test.[42]

Conclusions

Between 2000 and 2014 there was a negligible population-level improvement in mean step test performance for 20-to 59-year-old Chinese adults, which indicated improved functional endurance. We found negligible to small improvements in mean step test performance for all age, sex, location, and occupation groups, as well as trends in distributional characteristics, which indicated that both ends of the distribution had moved towards the middle. Improved functional endurance, particularly among those with low functional endurance, may be particularly meaningful because those with low endurance (i.e., below the 20th percentile) have the highest mortality risk. Moreover, despite tracking positively, further efforts are required to see continued improvement in adult functional endurance. China's historical national surveillance of functional endurance provides insight into trends in population health and fitness, assists in the evaluation of healthy public policies and interventions, and highlights the potential opportunity for other countries to engage in a cost-effective public health surveillance strategy.

Authors' contributions

YL had full access to the data, takes responsibility for the integrity of the data, contributed to the interpretation of results and critically reviewed the manuscript for important intellectual content; MST contributed to the interpretation of results and critically reviewed the manuscript for important intellectual content; and GRT developed the research question, designed the study, had full access to the data, takes responsibility for the integrity of the data, conducted the statistical analysis, and drafted the manuscript. All authors have read and approved the final version of the manuscript, agree to be accountable for all aspects of the work, and agree with the order of presentation of the authors.

Declaration of interest

The author(s) have no conflicts of interest relevant to this article.

Funding

This work was supported by Shanghai Key Laboratory of Human Performance (Shanghai University of Sport, No.11DZ2261100).

Data statement

The datasets analyzed in this review are available from the corresponding author on reasonable request.

Acknowledgments

We would like to thank Ms. Danqing Zhang and Ms. Youzhi Ke from Shanghai University of Sport for extracting the summary statistics from hard copy Reports on National Physical Fitness Surveillance.

References

1. Saltin B. Oxygen transport by the circulatory system during exercise in man. In: Keul J, ed. *Limiting Factors of Physical Performance*. Stuttgart: Thieme; 1973: 235–252.
2. Blair SN. Physical inactivity: the biggest public health problem of the 21st century. *Br J Sports Med*. 2009;43(1):1–2.
3. Blair SN, Kohl 3rd HW, Paffenbarger Jr RS, et al. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *J Am Med Assoc*. 1989;262(17):2395–2401. https://doi.org/10.1001/jama.262.17.2395.
4. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *J Am Med Assoc*. 2009;301(19):2024–2035. https://doi.org/10.1001/jama.2009.681.
5. Jensen MT, Holtermann A, Bay H, et al. Cardiorespiratory fitness and death from cancer: a 42-year follow-up from the Copenhagen Male Study. *Br J Sports Med*. 2017;51(18):1364–1369. https://doi.org/10.1136/bjsports-2016-096860.
6. Lee CD, Blair SN. Cardiorespiratory fitness and stroke mortality in men. *Med Sci Sports Exerc*. 2002;34(4):592–595. https://doi.org/10.1097/00005768-200204000-00005.
7. Crump C, Sundquist J, Winkleby MA, et al. Physical fitness among Swedish military conscripts and long-term risk for type 2 diabetes mellitus: a cohort study. *Ann Intern Med*. 2016;164(9):577–584. https://doi.org/10.7326/M15-2002.
8. Juraschek SP, Blaha MJ, Whelton SP, et al. Physical fitness and hypertension in a population at risk for cardiovascular disease: the Henry Ford ExercIse Testing (FIT) Project. *J Am Heart Assoc*. 2014;3(6), e001268. https://doi.org/10.1161/JAHA.114.001268.
9. Sloan RA, Sawada SS, Martin CK, et al. Associations between cardiorespiratory fitness and health-related quality of life. *Health Qual Life Outcome*. 2009;7:47. https://doi.org/10.1186/1477-7525-7-47.
10. Zhu N, Jacobs Jr DR, Schreiner PJ, et al. Cardiorespiratory fitness and cognitive function in middle age: the CARDIA study. *Neurology*. 2014;82(15):1339–1346. https://doi.org/10.1212/WNL.0000000000000310.
11. Patterson SL, Forrester LW, Rodgers MM, et al. Determinants of walking function after stroke: differences by deficit severity. *Arch Phys Med Rehabil*. 2007;88(1):115–119. https://doi.org/10.1016/j.apmr.2006.10.025.
12. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. 2016;134(24): e653–e699. https://doi.org/10.1161/CIR.0000000000000461.
13. Gupta S, Rohatgi A, Ayers CR, et al. Cardiorespiratory fitness and classification of risk of cardiovascular disease mortality. *Circulation*. 2011;123(13): 1377–1383. https://doi.org/10.1161/CIRCULATIONAHA.110.003236.
14. Australian Government Department of Health. Australia's physical activity & sedentary behaviour guidelines for adults (18–64 years). Available from: https://www1.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines#npa1864 . Accessed March 26, 2021.
15. Ross R, Chaput J-P, Giangregorio L, et al. Canadian 24-hour movement guidelines for adults aged 18–64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Appl Physiol Nutr Metabol*. 2020;45(10 Suppl. 2):S165–S179. https://doi.org/10.1139/apnm-2020-0467.
16. Chinese Center for Disease Control and Prevention. *Physical Activity Guidelines for Chinese Adults (Trial Version)*. Beijing: People's Medical Publishing House; 2011.
17. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med*. 2020;54(24): 1451–1462. https://doi.org/10.1136/bjsports-2020-102955.
18. Bennett H, Parfitt G, Davison K, et al. Validity of submaximal step tests to estimate maximal oxygen uptake in healthy adults. *Sports Med*. 2016;46(5): 737–750. https://doi.org/10.1007/s40279-015-0445-1.
19. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. *Phys Ther*. 2000;80(8):782–807. https://doi.org/10.1093/ptj/80.8.782.
20. Sartor F, Vernillo G, de Morree HM, et al. Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. *Sports Med*. 2013;43(9):865–873. https://doi.org/10.1007/s40279-013-0068-3.
21. Craig CL, Shields M, Leblanc AG, et al. Trends in aerobic fitness among Canadians, 1981 to 2007–2009. *Appl Physiol Nutr Metabol*. 2012;37(3):511–519. https://doi.org/10.1139/h2012-023.
22. Petrella RJ, Koval JJ, Cunningham DA, et al. A self-paced step test to predict aerobic fitness in older adults in the primary care clinic. *J Am Geriatr Soc*. 2001;49(5):632–638. https://doi.org/10.1046/j.1532-5415.2001.49124.x.
23. Yuan J, Fu Y, Zhang R, et al. The reliability and sensitivity of indices related to cardiovascular fitness evaluation. *Kinesiology*. 2008;40(2):138–145.
24. Lamoureux NR, Fitzgerald JS, Norton KI, et al. Temporal trends in the cardiorespiratory fitness of 2,525,827 adults between 1967 and 2016: a systematic review. *Sports Med*. 2019;49(1):41–55. https://doi.org/10.1007/s40279-018-1017-y.
25. Department of Mass Sport in Sport Commission of China. *Report on National Physical Fitness Surveillance*. Beijing: Beijing Sport University Press; 2000.
26. General Administration of Sport of China. *Report on National Physical Fitness Surveillance*. Beijing: Renmin Sport Press; 2005.
27. General Administration of Sport of China. *Report on National Physical Fitness Surveillance*. Beijing: Renmin Sport Press; 2010.
28. General Administration of Sport of China. *Report on National Physical Fitness Surveillance*. Beijing: Renmin Sport Press; 2014.

29. Tian Y, Jiang C, Wang M, et al. BMI, leisure-time physical activity, and physical fitness in adults in China: results from a series of national surveys, 2000–14. *Lancet Diabetes Endocrinol*. 2016;4(6):487–497. https://doi.org/10.1016/S2213-8587(16)00081-4.
30. General Administration Sport of China. *The Fitness Test Manual for Chinese Adults*. Beijing: People's Sport Publication; 2003.
31. Tomkinson GR, Kaster T, Dooley FL, et al. Temporal trends in the standing broad jump performance of 10,940,801 children and adolescents between 1960 and 2017. *Sports Med*. 2021;51(3):531–548. https://doi.org/10.1007/s40279-020-01394-6.
32. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. second ed. Mahwah, NJ: Lawrence Erlbaum; 1988.
33. Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of Populations: Methods and Applications*. Hoboken, NJ: John Wiley & Sons; 2013:143–188.
34. United Nations. *World Population Prospects 2019: Data Booklet*. ST/ESA/SER.A/424). New York: United Nations;; 2019.
35. Tomkinson GR, Kidokoro T, Dufner TJ, et al. Temporal trends in 6-minute walking distance for older Japanese adults between 1998 and 2017. *J Sport Health Sci*. 2020:S2095–S2546. https://doi.org/10.1016/j.jshs.2020.06.007. Jun 16.
36. Drinkwater EJ, Hopkins WG, McKenna MJ, et al. Modelling age and secular differences in fitness between basketball players. *J Sports Sci*. 2007;25(8):869–878. https://doi.org/10.1080/02640410600907870.
37. Cleveland W, Devlin S. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc*. 1988;83(403):596–610. https://doi.org/10.2307/2289282.
38. United Nations. *World Population Ageing 2019: Highlights*. ST/ESA/SER.A/430). New York: United Nations; 2019.
39. Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br J Sports Med*. 2019;53:478–486. https://doi.org/10.1136/bjsports-2017-097982.
40. Suwanachaiy S, Kulaputana O, Chaiwanichsiri D. Walk performance in Thai men and women: physical activity dependence. *Asian Biomed*. 2010;4:87–93. https://doi.org/10.2478/abm-2010-0010.
41. Woo J, Ho SC, Lau J, et al. Age-associated gait changes in the elderly: pathological or physiological? *Neuroepidemiology*. 1995;14(2):65–71. https://doi.org/10.1159/000109780.
42. Hansen D, Jacobs N, Bex S, et al. Are fixed-rate step tests medically safe for assessing physical fitness? *Eur J Appl Physiol*. 2011;111(2):2593–2599. https://doi.org/10.1007/s00421-011-1886-3.
43. Li H, Sui X, Huang S, et al. Secular change in cardiorespiratory fitness and body composition of women: the Aerobics Center Longitudinal Study. *Mayo Clin Proc*. 2015;90:43–52. https://doi.org/10.1016/j.mayocp.2014.08.020.
44. Mian OS, Thom JM, Ardigò LP, et al. Metabolic cost, mechanical work, and efficiency during walking in young and older men. *Acta Physiol*. 2006;186:127–139. https://doi.org/10.1111/j.1748-1716.2006.01522.x.