BMJ Open

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030137
Article Type:	Research
Date Submitted by the Author:	28-Feb-2019
Complete List of Authors:	Lu, Chuntao; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Hu, Yinhuan; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Fu, Qiang; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Governor, Samuel ; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Wang, Liuming; Huazhong University of Science and Technology Tongji Medical College, Tongji Hospital Deng, Lu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Xie, Jinzhu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management
Keywords:	Physician, MENTAL HEALTH, Workload, Survey and Questionnaires, Hospitals, Public

Title page

Manuscript title:

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Corresponding author:

Yinhuan Hu

Address: No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Email: hyh288@hotmail.com

Telephone number: 13554285879

Co-author:

Chuntao Lu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Yinhuan Hu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Qiang Fu: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Samuel Governor: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Liuming Wang: Tongji Hospital, Tongji Medical College of Huazhong University of Science and Technology, No.1095 Jiefang Rd., Wuhan, Hubei 430030, P.R. China

Lu Deng ： School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Jinzhu Xie: The Third People's Hospital of Hubei Province, No.26 Zhongshan Rd., Wuhan, Hubei 430030, P.R. China.

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Abstract

Objective: The purpose of our study is to develop and perform reliability and validity assessments of mental workload scale for physicians in China.

Design: Three phases, involving 385 physicians from different-level of comprehensive public hospitals in China, were conducted in this research to develop this instrument. In the first phase, an initial item pool was developed through systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey. The third phase tested psychometric properties of the instrument, including reliability and validity.

Setting: Public hospitals in China

Participants: 385 physicians from different-level of comprehensive public hospitals in China took part in this survey in 2018.

Primary and secondary outcome measures: Cronbach's alpha was used to test the reliability of the scale. Content validity index, correlation coefficient analysis, exploratory factor analysis, and confirmatory factor analysis were conducted to test validity of the scale**.**

Results: Six dimensions (mental demand, physical demand, temporal demand, perceived risk, frustration level and performance) and twelve items were identified in the instrument. For reliability, Cronbach's α for the whole scale was 0.81. For validity, the corrected item-content validity index of each item ranged from 0.85-1 and correlation coefficients between dimensions and total scores had a range of 0.37-0.72.

The results of confirmatory factor analysis showed that the goodness of fit of the scale was reasonable. Thus, the scale had good reliability and validity.

Conclusion: The instrument showed acceptable psychometric properties and can be a useful instrument for diagnosing mental workload of physicians.

Keywords: Physician; Mental Health; Workload; Survey and Questionnaires; Hospitals, Public

Article Summary

Strengths and limitations of this study

This is the first study to develop a measurement about physician mental workload from subjective perspective in China.

The findings of this study have good validity and reliability.

There is potential reporting bias in the self-reported measurements of the workload among physicians.

There was no assessment of re-test reliability because of web-based survey method.

There existed a selection bias due to all respondents voluntarily rather than randomly took part in the survey.

Introduction

Internationally, there has been a focus on the workload that physicians have often to face and on the physical and mental health[1]. Physician health is highly associated with workload[2]. Excessive workload impacts physician's health[3,4]and increases the risk of work-related musculoskeletal disorders (MSDs)[5,6]. Exposure to workload has been shown to be related to adverse effects in medical errors[7] and adverse incidents[8]. Physician workload could be a negatively contributing factor to patients' perceived quality of care[9], and affects patient satisfaction[10] and safety[11,12]. It is possible that these stressors have reached a point where they pose a serious policy issue for the entire healthcare system[13]. Thus, unreasonable and overwhelming workload has adverse effects to physicians, patients and healthcare organizations[14].

Workload is thought to be multidimensional and multifaceted[15]. One aspect of workload includes the subjective psychological experiences of the human operator[16]. Mental workload has emerged as one of the most important occupational risk factors as well. Heavily mental workload can lead to serious health problems for workers (anxiety, burnout, cardiovascular diseases, digestive problems, etc.)[17], so as well to physicians, excessive mental workload can lead to inferior quality of care[18]. Currently, The European Pact for Mental Health and Welfare[19] has conducted mental workload assessments to promote physical and mental wellbeing.

Mental workload can be influenced by numerous factors that make a definitive

measurement difficult[20]. Different methods have been proposed to assess mental workload. Previous research has established a brief instrument with six items to measure physician mental workload[21]. The most widely used instrument to measure mental workload is NASA-Task Load Index (NASA-TLX) scale[22]which has proven to be a sensitive, valid, and reliable instrument[23] and can be used in human factors research[24]. The existing body of research on NASA-TLX suggests that it could be used to measure nurse workload[25-27]. In the same vein, the Subjective Workload Assessment Technique (SWAT) is a subjective rating technique with three dimensions of time load, mental effort load, and psychological stress load, which is used to assess mental workload as well[28]. It has been successfully applied in the mental workload assessment of several aircraft multitask conditions, such as assessing the mental workload of different systems of air defense[23]. Together these studies provide important insights into workload measurement in health care management. However, there is no specific instrument, to our knowledge, has been explored in physician mental workload in China.

Physicians, some of the major providers of health services, taking more and more responsibilities for patient care in Chinese health care management, have heavier workload, worse physical health, more mental strain and more intense relationships with patients[29]. Data from several studies suggest that most physicians' work more than 10 hours in a day[30] to manage outpatients and inpatients, on average, a physician in a tertiary hospital is responsible for 8.10 outpatients and 2.70 beds per day[29]. However, they even have been abused, injured, and in extreme cases,

murdered by patients or their relatives in hospitals across China[31], which resulted in extremely mental workload. Establishing the workload measurement system for medical personnel has been fitted into Chinese Patient Safety Goals by Chinese Hospital Association[32].

However, existing research about workload measurement instruments are concentrated on work time and objective workload in China, whereas the mental workload which is an indispensable problem has less relevant instrument in China. The purpose of this paper is to develop a scientific mental workload instrument, which can be used to measure or assess the actual mental workload of physicians in China.

Methods

Study design

The instrument was developed in three phases. In the first phase, an initial item pool was developed by integrating previous studies through systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey in 2017. The third phase was testing the psychometric properties of the instrument, including its reliability and validity, through a survey conducted in 2018 in comprehensive public hospitals in China.

Framework and items of generation and selection

Based on the framework of NASA-TLX and SWAT, we combined the status quo of

Chinese physicians' workload to determine the item pool. Six dimensions and fifteen items were sent to 20 experts for consultation. After a two-round consultation, it was suggested that we deleted four items, added a new item, and revised the description of all the items. Finally, the pre-scale consisted of six dimensions (physical demand, mental demand, temporal demand, effort, frustration level, and performance) and twelve items with an evaluation from 0 to 100.

In the pre-survey analysis, a sample of 80 physicians was surveyed with a web-based pre-scale during November and December 2017. Items were refined based on the following indexes or methods: critical ratio (CR), coefficient of variation (CV), correlation analysis[33], Cronbach's α[34], exploratory factor analysis (EFA)[35].

If an item was eliminated by any of the above methods, then the item was deleted or revised. Final scale consisted of six dimensions (mental demand, physical demand, temporal demand, perceived risk, frustration level, and performance) and twelve items (Table 1).

Data collection for testing validity and reliability of the scale

To test the validity and reliability of the developed scale, the samples size was 5-10 times the size of the items, were considered suitable[36]. Data was collected from the tertiary hospitals (valid sample size: 130), secondary hospitals (valid sample size: 124) and first-level hospitals (valid sample size: 131) from February 2018 to March 2018.

The scale included three parts. The first part of the scale was the principal twelve

items. The second part was used to collect the weights of each dimension. Every two dimensions formed a pair, and the respondents selected which of the two dimensions in a pair they considered to have contributed more to their workload. There were fifteen pairs the respondents needed to select from, and the weights of each dimension was equal to the number of times that dimension was selected divided by fifteen. The third part was the individual information about the physicians.

Items were scored from 0 to 100. Responses to the items are displayed in Table 1. The average scores of all items for a corresponding dimension were multiplied by the dimension weight to produce the dimension scores. The total scores were the sum of all dimension scores.

Table 1 Dimensions and items of physician mental workload scale

Dimensions and items	Endpoints (0-100)
Mental demand	
A1 How much cognitive activity (e.g. sensation, perception, remembering, thinking, calculating, attention etc.) was required during your medical work?	Little (lighter workload) /much (heavier workload)
A2 How much emotion and feeling was required (e.g. empathy, sympathy, enthusiasm, negative emotion restraining etc.) during your medical work?	Little (lighter workload) /much (heavier workload)
A3 How hard did you have to work to overcome difficulties in accomplishing your medical work?	Low (lighter workload) /high (heavier workload)
Physical demand	
B1 How much physical activity was required (e.g. standing, sedentary, controlling, repetitive action etc.) in your medical work?	Little (lighter workload) /much (heavier workload)
B2 How intensive or precise was the physical activity during your medical work? (Was the work restful or laborious? Was your muscle relaxed or tense?)	Low (lighter workload) /high (heavier workload)
Temporal demand	
C1 How much time pressure did you feel due to the ratio of required time to available time in your medical work (Was the pace slow and leisurely or rapid and frantic)?	Little (lighter workload) /much (heavier workload)
C2 How frequent did you have to complete multiple tasks at	Low (lighter workload) /high (heavier workload)

the same time (work overlap) in your medical work?	
Perceived risk	
D1 How much risk did you perceive (e.g. professional infection, medical dispute， uncertainty of medical treatment etc.) in your medical work?	Little (lighter workload) /much (heavier workload)
Frustration level	
E1 How depressed or frustrated did you feel in your medical work?	Low (lighter workload) /high (heavier workload)
E2 How anxious or irritated did you feel in your medical work?	Low (lighter workload) /high (heavier workload)
Performance	
F1 How successful do you think you were in accomplishing the goals in your medical work?	Low (heavier workload) /high (lighter workload)
F2 How satisfied were you with the outcome in your medical work?	Unsatisfactory(heavier workload) /satisfactory(lighte

Statistical analysis

Descriptive statistics were used to show the characteristics of the samples. For the reliability of the scale, Cronbach's alpha was used to assess the internal consistency of each instrument component. Values of 0.70 or greater for Cronbach's alpha were considered acceptable[34].

Content validity index (CVI) of each item was calculated to assess the accuracy of the scale using the scores of 1-4. Experts were invited to assess the items with a scale of 1 representing the item not relevant to corresponding dimension and 4 representing the item closely relevant to corresponding dimension. Corrected item-content validity index (I-CVI) and average scale-content validity index (S-CVI/Ave) were calculated. Corrected I-CVI of 0.78 or greater and S-CVI/Ave of 0.90 or greater were considered acceptable[37].

The test of construct validity was tested by correlation coefficient method,

Exploratory Factor Analysis and Confirmatory Factor Analysis. Items whose values for item-total correlation and dimension-total correlation below 0.40 were revised or removed from the scale. A Bartlett's test of sphericity scores lower than 0.05 and a Kaiser-Meyer–Olkin (KMO) measure of sampling adequacy higher than 0.70 and closer to 1 were considered appropriate for factor analysis[38]. Exploratory Factor Analysis and Confirmatory Factor Analysis was used to further explore and confirm the structure of the scale.

For the Confirmatory Factor Analysis, the criterion of model fit indices were listed as follows: $\chi 2/df < 3$; root mean square error of approximation (RMSEA) < 0.05; root of mean square residual (RMR)<0.05; goodness of fit (GFI) >0.90, comparative fit index (CFI)>0.90, Tucker-Lewis index (TLI)>0.90[39]. Statistical analyses were performed with SPSS V. 21 (IBM Corp., Armonk, NY, USA) and AMOS V.17 (IBM Corp., Armonk, NY, USA).

Participants and public involvement

We used a reliable and widely web-based survey (www.wjx.cn) in the form of quick-response code to survey physicians through sending the code to physician communication groups of each hospital. Participants were voluntarily to take part in this survey, before participating in the study, informed consent was provided to them.

Results

Sample characteristics

396 questionnaires were received and 11 questionnaires were excluded due to incomplete demographic information. The response rate was 97.2% (385/396). There were no issues of floor or ceiling effects as questions to every item were responded as required in the form of quick-response code. Characteristics of the participants are presented in Table 2.

Table 2 Respondents' characteristic of samples

Variable	Valid sample	Percentage（%）
Gender		
Male	200	51.95
Female	185	48.05
Age		
<45	258	67.01
45-55	121	31.43
>55	6	1.56
Marriage status		
Single	47	12.21
Married	329	85.45
Divorced/widowed	9	2.34
Educational level		
PhD degree	36	9.35
Master degree	48	12.47
Undergraduate	184	47.79
Junior college and below	117	30.39
Professional title		
Senior	83	21.56
Middle	146	37.92
Junior	156	40.52
Work years in his/her institutions (year)		
1-5	122	31.69
6-10	91	23.64
11-15	58	15.06

Variable	Valid sample	Percentage（%）
>16	114	29.61
Hospital level		
Tertiary hospitals	130	33.77
Secondary hospitals	124	32.21
First-level hospitals	131	34.02
Work hours per week		
<=40	73	18.96
41-60	152	39.48
>60	160	41.56
Number of daily service outpatients		
<20	135	35.06
20-50	171	44.42
>50	79	20.52
Self-rated health status		
Poor	65	16.89
Fair	242	62.86
Good	78	20.25

Reliability of physician mental workload scale

Each of the six components demonstrated at least satisfactory internal consistency higher than 0.70, with Cronbach's α in the range of 0.70-0.90. Cronbach's α for the whole scale was up to 0.81, greater than 0.80, indicating that the scale has a good reliability.

Validity of physician mental workload scale

The corrected I-CVI of each item ranged from 0.85-1 (Table 3), which is higher than 0.78. The CVI/Ave was 0.96, which is higher than 0.90. All of these showed a good content validity of the scale.

The correlation matrix between items and total scores was inspected to confirm the convergent validity, which was indicated by reasonable coefficients of 0.40 and

above, except for F1 and F2 (Table 3). Calculated correlation coefficients between dimensions and total scores had a range of 0.37-0.72, which was an additional index that showed that dimensions and total scores have a good convergent validity. Also, the correlation coefficients among the dimensions were lower than the correlation coefficients between the dimension-total scores, which indicated that the scale has a good discriminant validity (Table 4).

Table 3 Content validity and correlation coefficient of item-total scores of the scale

Items	Corrected I-CVI	Item-total correlations
Cognitive activity	1	0.57
Emotion and feeling	0.85	0.57
Overcoming difficulties	0.85	0.59
Physical activity	1	0.57
Intensity of physical activity	1	0.65
Time pressure	1	0.69
Multiple task	0.85	0.69
Risk	1	0.64
Depressed or frustrated	1	0.75
Anxious or irritated	1	0.75
Successful	1	0.33*
Satisfied	1	0.31*

*item-total scores were below than 0.4

Table 4 Correlation coefficient matrix between dimensions and total scores of the scale

Dimensions	Mental demand	Physical demand	Temporal demand	Perceived risk	Frustration Level	Performance	Total scores
Mental demand	1						
Physical demand	0.43	1					
Temporal demand	0.52	0.47	1				
Perceived risk	0.46	0.38	0.44	1			

Frustration level	0.40	0.36	0.54	0.51	1		
Performance	0.09	0.05	0.01	0.05	0.13	1	
Total scores	0.61	0.52	0.68	0.68	0.72	0.37	1

Exploratory factor analysis of physician mental workload scale

The KMO sample adequacy measurement was 0.81, which was higher than the recommended value of 0.70, and the Bartlett's test of sphericity with Chi square value 1950.70 and $p<0.000$. Thus, the data were suitable for factor analysis. Considering experts' suggestion, we select 6 principal components and the six-dimensional model explained 81.88% of total variance (Table 5).

The factor "mental demand" was developed from 3 items that asked for feeling or memory requirement, emotional requirement, and the effort input to overcome difficulties, with the factor loading in the range of 0.74-0.81. The factor "physical demand" consisted of 2 items that related to strength requirement and the intensity of work time with a factor loading in the range of 0.84-0.90. The factor "temporal demand" constituted 2 items that asked about the ratio of required time to available time and frequency of completing multiple tasks, with the factor loading in the range of 0.77-0.82.

The factor "perceived risk" included only 1 item that explained the perception of risk in conducting tasks (such as medical dispute and risk of being infectious) with a factor loading of 0.84. The factor "frustration level" consisted of 2 items that asked

about anxiety and level of depression or frustration, and the factor loading was in the range of 0.86-0.88. There were 2 items in the “performance demand” factor, which related to the sense of achievement and job satisfaction regarding work outcome, with the factor loading in the range of 0.85-0.90.

Table 5 Factor loadings for the rotated component matrix: varimax rotated components

Items	Components					
	1	2	3	4	5	6
A2	0.81					
A1	0.76					
A3	0.74					
E1		0.88				
E2		0.86				
B1			0.90			
B2			0.84			
C1				0.82		
C2				0.77		
F1					0.90	
F2					0.85	
D1						0.84

Confirmatory factor analysis of physician workload scale

The six-factor model obtained after Exploratory Factor Analysis was tested by Confirmatory Factor Analysis using maximum likelihood estimation method. The goodness-of-fit model was listed as follows: $\chi 2/df=1.84$, RMR= 0.04, and GFI=0.97, CFI=0.98, TLI=0.97, RMSEA=0.05. Referring to the criterion listed above, the model was a good fit for the data.

Discussion

The purpose of this study is to develop and explore the validity and reliability of mental workload scale for physicians. According to the results of the tests, the scale is reliable and valid, hence, it is considered as an effective instrument for assessing physician mental workload in Chinese comprehensive public hospitals. Results show a six-dimensional model which includes aspects related to mental demand, physical demand, temporal demand, perceived risk, frustration level and performance. Perceived risk and temporal demand are especially distinctive for Chinese physician mental workload.

Perceived risk, which is a different dimension and not included in framework of NASA-TLX and SWAT, is highly associated with physician mental workload in China. Medical practice is a special but a high-risk behavior because physicians do not only save and heal people, but also they put themselves at a risk of being infectious. Also in China, there tends to be an estranged relationship between physicians and patients which puts physicians at a dangerous risk of being assaulted in their line of work[40]. For instance, in recent years, many doctors have been assaulted, seriously injured and even murdered by patients or visitors in China. According to the statistics, 96% of medical staff have been abused or injured in 2012[41]. Moreover, the reports of these incidents by mass media have further exacerbated the conflict between doctors and patients. The physician-patient relationship is becoming more and more fragile and has reached an unprecedented poor level in China[42]. Without exaggeration, some physicians even wear helmet in the hospitals in Guangdong Province, which reflects

that physicians feel unsafe and have suffered from heavy psychological workloads during their work.

In China, the gap between healthcare demand and supply (thus doctor-patient ratio) in China has caused physicians in the secondary and tertiary hospital settings to become overworked[43]. Few physicians in tertiary hospitals can complete their work in regular 8 hours. Research has reported that physicians may feel stressed when poor scheduling leaves them pressed for time[44]. Chinese physicians play various roles during their work and they always need to work overtime, even though they conduct more than one task at the same time[43]. According to the White Paper on the Practice of Medicine by Chinese Physicians by the Chinese Medical Association in 2014, 32.7% of doctors had an average workweek over 60 hours[45]. High task demands require plenty of time, and evoke high mental effort and heavy workload for physicians[46]. Mental workload encompasses the subjective experience of a given task load[47], the higher the task load, the higher the mental workload[48].

Consistent with previous research on NASA-TLX[49], performance dimension shows a limited practical relevance since it is influenced by variations in physical load. In our study, the item-total scores of the two items in the dimension of performance were near to 0.4, and perhaps, would have been relevant in a reverse scoring. Other research reported that subjective assessments of mental workload may not provide accurate estimation of the performance dimension[17]. Considering this information, we retain the two items but revised their description.

Although we have attempted an accurate examination of the measurement

properties of the physician mental workload scale, there were still some limitations that merit discussions. One of the limitations was potential reporting bias in the self-reported measurements of the workload among physicians. Secondly, we could not conduct a re-test reliability because we used a web-based survey method. Thirdly, all respondents voluntarily decided to take part in the survey. This means physicians who were overburdened at the time of the survey may not have time to take part in the survey, resulting in a selection bias.

Notwithstanding these limitations, creating new scale items from a subjective perspective is of paramount importance in investigating Chinese physician's workload. Physician mental workload scale has acceptable preliminary psychometric properties with six dimensions and twelve items. The use of physician mental workload scale can help us to find the main stressors in physician mental workload and to implement targeted optimization strategies to mitigate these stressors in an effort to ameliorate the physical and mental health of physicians. This, consequently, will help us to improve the quality and efficiency of healthcare delivery in hospital settings.

Abbreviations

CVI: Content Validity Index;

RMR: Root of Mean Square Residual; GFI= Good of Fitness Index; CFI= Comparative Fit Index; TLI=Tucker-Lewis index; CR= Critical Ratio; CV=Coefficient of Variation;

NASA-TLX=NASA- Task Load Index; SWAT= Subjective Workload Assessment Technique;

EMRs= Electronic Medical Records

Funding Statement: This work was supported by the National Natural Science Foundation of China (grant number 71774062). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgement We would like to thank all the participants involved in this research for their time and contributions.

Author Contributors: YH and QF designed the study; CL and JX made formal analysis; YH obtained funding; CL, JX and LD took part in investigation; LW and LD were involved in data cleaning; CL wrote original draft; SG and QF contributed to the interpretation of the results and critical revision of the manuscript; All authors have read and approved the final manuscript.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Patient consent: Not required.

Data sharing: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Reference

1. Győrffy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians[J]. *Hum resour health*, 2016, 14: 12.
2. Lu L, Gao M, Yang T, et al. The updated Physician's Pledge and Chinese junior physicians. *The Lancet.* 2018;391:305.
3. Karahan A, Kav S, Abbasoglu A, et al. Low back pain: prevalence and associated risk factors among hospital staff. *J Adv Nurs.* 2009;65: 516-524.

4. Smith-Miller CA, Shaw-Kokot J, Curro B, et al. An integrative review: fatigue among nurses in acute care settings. *J Nurs Adm.* 2014;44:487-494.

5. Arvidsson I, Simonsen JG, Balogh I. Discrepancies in pain presentation caused by adverse psychosocial conditions as compared to pain due to high physical workload? *Work*. 2012;41:2472-2475.

6. Manty M, Moller A, Nilsson C, et al. Association of physical workload and leisure time physical activity with incident mobility limitations: a follow-up study. *Occup Environ Med*. 2014; 71:543-548.

7. Ridley C. Relating nursing workload to quality of care in child and adolescent mental health inpatient services. *Int J Health Care Qual Assur*.2007;20:429-440.

8. Michtalik HJ, Pronovost PJ, Marsteller JA, et al. Identifying potential predictors of a safe attending physician workload: a survey of hospitalists. *J Hosp Med*.2013; 8:644-646.

9. Michtalik HJ, Yeh HC, Pronovost PJ, et al. Impact of attending physician workload on patient care a survey of hospitalists. *JAMA Intern Med*.2013; 173:375-377.

10. Garcia-Lacalle J. Balancing physician workload and patient satisfaction: a dilemma for hospital managers. *Public Money & Management*.2010;30:183-188.

11. Steyrer J, Schiffinger M, Huber C,et al. Attitude is everything? The impact of workload, safety climate, and safety tools on medical errors: a study of intensive care units. *Health Care Manage R*.2013;38:306-316.

12. Pedroja AT, Blegen MA, Abravanel R. The relationship between hospital systems load and patient harm. *J Patient Saf*. 2014;10:168-175.

13. Williams E S, Rondeau K V, Xiao Q, et al. Heavy physician workloads: impact on physician attitudes and outcomes. *Health Care Manage R*, 2007, 20: 261-269.

14. Hart SG. Theory and Measurement of Human Workload. *J Oper Res Soc*.1988;*39*:725-734.

15. Cain B. A review of the mental workload literature. Toronto: Defence Research And Development, 2007.

16. Lysaght RJ, Hill SG, Dick AO, et al. Operator workload: Comprehensive review and evaluation of operator workload methodologies. Analytics inc willow grove pa,

1989.

17. Rubio-Valdehita S, Lopez-Nunez MI, Lopez-Higes R, et al. Development of the CarMen-Q Questionnaire for mental workload assessment. *Psicothema*. 2017; 29:570-576.

18. Weigl M, Müller A, Holland S, et al. Work conditions, mental workload and patient care quality: a multisource study in the emergency department. *BMJ Qual Saf.* 2016; 25: 499-508.

19. European Commission. European Pact for health and mental health. http://ec.europa.eu/health/mental_health/policy_es. (Accessed on 24th Sep 2018)

20. Jex HR. Measuring mental workload: Problems, progress, and promises. *Adv Psychol*, 1988, 52: 5-39.

21. Bertram DA, Opila DA, Brown JL, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*.1992;30: 95-104.

22. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol*. 1998;*52*:139-183.

23. Rubio S, Díaz E, Martín J, et al. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Appl Psychol.* 2004; 53: 61-86.

24. Hill SG, Iavecchia HP, Byers JC. Comparison of four subjective workload rating scales. *Hum Factors*.1992;*34*:429.

25. Colligan L, Potts HW, Finn CT, et al. Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *Int J Med Inform.*2015; 84:469-476.

26. Hoonakker P, Carayon P, Gurses A, et al. Measuring Workload Of ICU Nurses with a Questionnaire Survey: The NASA Task Load Index (TLX). *IIE Trans Healthc Syst Eng*. 2011;1:131-143.

27. Malekpour F, Malekpour A R, Mohammadian Y, et al. Assessment of mental workload in nursing by using NASA-TLX. *The Journal of Urmia Nursing and*

Midwifery Faculty.2014; 11: 892-899.

28. Reid GB, Nygren TE. The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Adv Psychol*.1988; *52*.
29. Zhou H, Han X, Zhang J, et al. Job satisfaction and associated factors among medical staff in tertiary public hospitals: Results from a national cross-sectional survey in China. *Int J Occup Env Heal*. 2018; 15: 1528.
30. Sun J, Ma J, Hu G, et al. Welfare, wellness, and job satisfaction of Chinese physicians: A national survey of public tertiary hospitals in China. *Int J Health Plan M*. 2017; 32: 270-284.
31. The Lancet, Ending violence against doctors in China. *The Lancet*. 2012; 379:1764.
32. Chinese Hospital Association. Chinese patient safety goal 2014-2015. http://www.cha.org.cn/plus/view.php?aid=14626 (Accessed on 24th Sep 2018)
33. Power M, Quinn K, Schmidt S, et al. Development of the WHOQOL-old module. *Qual Life Res*. 2005; 14 (10):2197-2214.
34. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York: Guilford publications, 2015:69-70.
35. Ford JK, MacCullum RC, Tait M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Pers Psychol*. 1986; 39:291-314.
36. Nunnally J C. Psychometric theory. New York: McGraw-Hill, 1967.
37. Lynn MR. Determination and quantification of content validity. *Nursing Res*. 1986; 35:382-385.
38. DiIorio C K. Measurement in health behavior: methods for research and evaluation. John Wiley & Sons, 2006.
39. Tabachnick B, Fidell LS. Using Multivariate Statistics. 6th ed., Pearson, Boston, 2013: 720-726.
40. Tang L, Guan M. Rise of Health Consumerism in China and Its Effects on Physicians' Professional Identity and the Physician-Patient Relationship and Communication. *Health commun*.2018; 33:636-642.
41. Yueju L. Violence against doctors in China. *The Lancet*. 2014; 384:745.
42. Sun T, Gao L, Shi Y. Workplace violence, psychological stress, sleep quality and

subjective health in chinese doctors: a large cross-sectional study. *BMJ OPEN*. 2017;1-8.

43. Wen J, Cheng Y, Hu X. Workload, burnout, and medical mistakes among physicians in China: A cross-sectional study. *Biosci trends*. 2016; 10:27-33.

44. Horner RD, Szaflarsk JP, Ying J. Physician Work Intensity Among Medical Specialties: Emerging Evidence on its Magnitude and Composition. *Med care*. 2011; 49:1007-1011.

45. Chinese Medical Doctor Association. White Paper on the Practice of Medi cine by Chinese Physicians. http://www.cmda.net/xiehuixiangmu/falvshiwubu/tongzhigonggao/2015-05-28/14587.html. (Accessed on 24th Sep 2018)

46. Mosaly PR, Mazur LM, Yu F. Relating Task Demand, Mental Effort and Task Difficulty with Physicians' Performance during Interactions with Electronic Health Records (EHRs). *Int J Hum-Comput Int.* 2017; 34:467-475.

47. Charles RL, Nixon J. Measuring mental workload using physiological measures: A systematic review[J]. Appl ergon. 2019; 74: 221-232.

48. Colle HA, Reid GB. Context effects in subjective mental workload ratings. Hum factors. 1998; 40: 591-600.

49. DiDomenico A, Nussbaum MA. Effects of different physical workload parameters on mental workload and performance. *Int J Ind Ergonom* 2011. 41;255-260.

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectional reporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Study design	#4	Present key elements of study design early in the paper	5
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	9

	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	n/a
Bias	#9	Describe any efforts to address potential sources of bias	17
Study size	#10	Explain how the study size was arrived at	9
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	n/a
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	8-9
	#12b	Describe any methods used to examine subgroups and interactions	n/a
	#12c	Explain how missing data were addressed	9
	#12d	If applicable, describe analytical methods taking account of sampling strategy	8-9
	#12e	Describe any sensitivity analyses	n/a
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	5-6
	#13b	Give reasons for non-participation at each stage	n/a
	#13c	Consider use of a flow diagram	n/a
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	9

	#14b	Indicate number of participants with missing data for each variable of interest	9
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	11-14
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
	#16b	Report category boundaries when continuous variables were categorized	n/a
	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Key results	#18	Summarise key results with reference to study objectives	14-16
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	16-17
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	17
Generalisability	#21	Discuss the generalisability (external validity) of the study results	17
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

BMJ Open

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030137.R1
Article Type:	Original research
Date Submitted by the Author:	04-Jul-2019
Complete List of Authors:	Lu, Chuntao; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Hu, Yinhuan; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Fu, Qiang; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Governor, Samuel ; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Wang, Liuming; Huazhong University of Science and Technology Tongji Medical College, Tongji Hospital Li, Chao; Jingmen No 2 People's Hospital Deng, Lu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Xie, Jinzhu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Mental health
Keywords:	Physician, MENTAL HEALTH, Workload, Survey and Questionnaires, Hospitals, Public

Title page

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Chuntao Lu Yinhuan Hu* Qiang Fu Samuel Governor Liuming Wang Chao Li Lu Deng Jinzhu Xie

Corresponding author:

*Yinhuan Hu

Address: No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Email: hyh288@hotmail.com

Telephone number: 13554285879

Co-author:

Chuntao Lu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Yinhuan Hu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Qiang Fu: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Samuel Governor: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Liuming Wang: Tongji Hospital, Tongji Medical College of Huazhong University of

Science and Technology, No.1095 Jiefang Rd., Wuhan, Hubei 430030, P.R. China

Chao Li: Jingmen No 2 People's Hospital. No.39 Xiangshan Rd., Jingmen, Hubei 448126, P.R. China

Lu Deng: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd.,Wuhan, Hubei 430030, P.R. China

Jinzhu Xie: The Third People's Hospital of Hubei Province, No.26 Zhongshan Rd., Wuhan, Hubei 430030, P.R. China.

Abstract

Objective: The purpose of our study is to develop and perform reliability and validity assessments of mental workload scale for physicians in China.

Design: Three phases, involving 396 physicians from different-level of comprehensive public hospitals in China, were conducted to develop the instrument. In the first phase, an initial item pool was developed through systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey. The third phase tested reliability and validity of the instrument.

Setting: Public hospitals in China.

Participants: A total of 396 physicians from different tiers of comprehensive public hospitals in China participated in this study in 2018.

Primary and secondary outcome measures: Cronbach's alpha, content validity index, correlation coefficient between items and dimensions, and indices of confirmatory factor analysis**.**

Results: Six dimensions (mental demand, physical demand, temporal demand, perceived risk, frustration level, and performance) and twelve items were identified in the instrument. For reliability, Cronbach's α for the whole scale was 0.81. For validity, the corrected item-content validity index of each item ranged from 0.85 to 1 and correlation coefficients between dimensions and total scores had a range of 0.37-0.72. The results of confirmatory factor analysis showed that the goodness of indices of the scale was reasonably well.

Conclusion: The instrument showed good reliability and validity, and it is useful for

diagnosing mental workload of physicians.

Keywords: Physician; Mental Health; Workload; Survey and Questionnaires; Hospitals, Public

Article Summary

Strengths and limitations of this study

This is the first study to develop a measurement about physician mental workload from a subjective perspective in China.

Qualitative and quantitative methods were involved in item selection.

There is potential a reporting bias in the self-reported measurements of the workload among physicians.

There existed a selection bias due to all respondents voluntarily rather than randomly took part in the survey.

Among six dimensions, perceived risk only included one item, which may result in measurement error.

Introduction

Internationally, there has been a focus on the relationship between physicians' workload and their health. [1] Physicians' health is highly associated with workload. [2] Excessive workload impacts physician's health [3-4] and increases the risk of work-related musculoskeletal disorders (MSDs). [5-6] Exposure to workload has been shown to be related to adverse effects in medical errors [7] and adverse incidents. [8] Physician workload could be a negatively contributing factor to patients' perceived quality of care, [9] and affects patient satisfaction [10] and safety. [11-12] It is possible that these stressors have reached a point where they pose a serious problem for the entire healthcare system. [13] Thus, the unreasonable and overwhelming workload has adverse effects to physicians, patients and healthcare organizations. [14]

Workload is thought to be multidimensional and multifaceted. [15] One aspect of workload includes the subjective psychological experiences of the human operator. [16] Mental workload as a kind of workload has emerged as one of the most important occupational risk factors, which results in burnout or anxiety. [17] Lack of control over workload was expected to correlate most highly with burnout. [18,19]Heavy mental workload can lead to serious health problems (cardiovascular diseases, digestive problems, etc.) for physicians as well. [17] Meanwhile, the excessive mental workload can also lead to an inferior quality of care service. [20] Currently, The European Pact for Mental Health and Welfare is conducting mental workload assessments to promote physical and mental wellbeing. [21]

Different tools have been proposed to assess mental workload. Previous research

has established a brief instrument with six items to measure physician mental workload. [22] The NASA-Task Load Index (NASA-TLX) scale which was widely used in measuring mental workload [23] has proven to be a sensitive, valid, and reliable instrument [24] and can be used in human factors research. [25] Researcher has localized it as a 29-item questionnaire in Spain to measure workers' mental workload. [26] The existing body of research on NASA-TLX suggested that it could be used to measure nurse workload in health care settings. [27-29] In the same vein, the Subjective Workload Assessment Technique (SWAT) is a subjective rating technique with three dimensions of time load, mental effort load, and psychological stress load, which is used to assess mental workload as well. [30] It has been successfully applied in the mental workload assessment of several aircraft multitask conditions, such as assessing the mental workload of different systems of air defense. [24] Copenhagen Psychosocial Questionnaire was a wide-spread tool used in the industrial or in the services branch in Europe, which included the main dimensions of the most influential psychosocial theories at work. [31] Together these tools provide important insights into workload measurement in health care management, especially in nurse workload measurement. However, there is still different workload between physicians and nurses essentially, meanwhile, these measurement was designed for other workers, so we do not ensure these tool can be directly used in physician mental workload measurement. Specifically, mental workload measurement needs to be developed for physicians.

With the increasing of patient demands for health, physicians tend to have a heavier workload, worse physical health, more mental strain and more intense

relationships with patients in China. [32] Data from several studies suggest that most physicians' work more than 10 hours in a day [33] to manage outpatients and inpatients. On average, a physician in a tertiary hospital is responsible for 8.10 outpatients and 2.70 beds per day. [32] However, they even have been abused, injured, and in extreme cases, murdered by patients or their relatives in hospitals across China, [34] which resulted in extremely mental workload. Establishing the workload measurement system for medical personnel has been incorporated into Chinese Patient Safety Goals by Chinese Hospital Association. [35]

Existing researches about workload measurement instruments are concentrated on objective workload in China, for example, the measurement of work time. Whereas physicians' mental workload is an indispensable problem in China, there are few instruments exploring physician mental workload in China. The purpose of this paper is to develop a scientific mental workload instrument, which can be used to measure or assess the actual mental workload of physicians in China.

Methods

Study design

The instrument was developed in three phases. In the first phase, an initial item pool was developed by integrating previous studies through a systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey in 2017. The third phase was testing the psychometric properties of the

instrument, including its reliability and validity, through a survey conducted in 2018 in comprehensive public hospitals in China.

Framework and items of generation and selection

Based on the framework of NASA-TLX and SWAT, we combined the current situation of Chinese physicians' workload to determine the item pool. Six dimensions and fifteen items were sent to 20 experts (including physicians, hospital managers, researchers, and human resource managers) for consultation. According to two-round expert consultation, we deleted four items, added a new item (the intensity of physical activities), and revised the description of all the items. Then, there were six dimensions (physical demand, mental demand, temporal demand, effort, frustration level, and performance) and twelve items, which consisted of the pre-scale with a range from 0 to 100.

In the pre-survey analysis, we selected 3 hospitals (1 tertiary hospital, 1 secondary hospital, and 1 first-tier hospital) through conveniently sampling. A sample of 80 physicians was surveyed with a web-based scale during November and December 2017. Finally, 74 samples were validated to conduct item selection. Items were refined based on the following indexes or methods: critical ratio (CR), coefficient of variation (CV), correlation analysis, [36] Cronbach's α, [37] and exploratory factor analysis (EFA). [38]

If an item was eliminated by any of the above methods, then the item was deleted or revised. The final scale consisted of six dimensions (mental demand,

physical demand, temporal demand, perceived risk, frustration level, and performance) and twelve items (Table 1).

Data collection for testing the validity and reliability of the scale

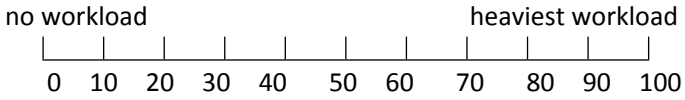
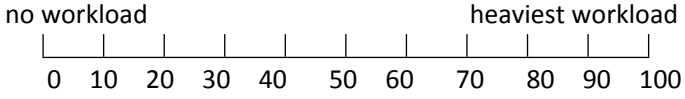
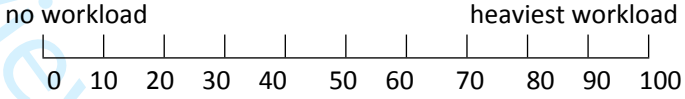
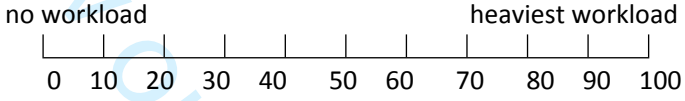
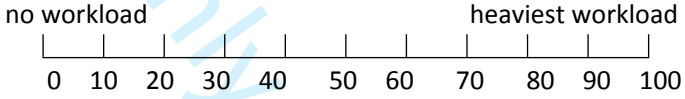
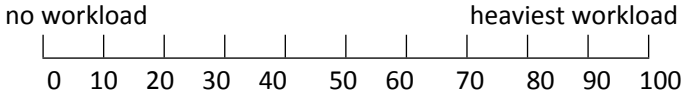
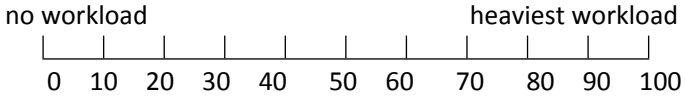
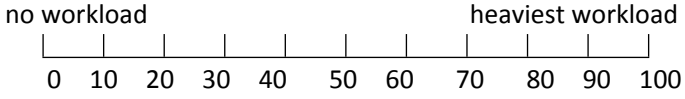
To test the validity and reliability of the developed scale, we planned to survey 400 respondents (physicians are working in hospitals) from different tier hospitals (2 tertiary hospitals, 2 secondary hospitals, and 2 first-tier hospitals). These hospitals were randomly selected from Hubei province, China. We used a website--wenjuanxing (www.wjx.cn) to survey physicians. The human resource at each participating hospital sent the access code to the physicians. 396 physicians voluntarily participated in the survey before March 2018, finally, 11 invalid samples were deleted.

There was a detailed description in the guidance of the scale, which showed that our scale was anonymous and all physicians were voluntary to answer this question. Meanwhile, we described that the aim of our survey was to develop a physician mental workload scale, so the results would not be used for other purposes. The physician mental workload scale included three parts. The first part of the scale was the principal twelve items. The second part was a table which included 15 pairs in our scale, which was used to collect the weights of each dimension. Every two dimensions formed a pair, and the respondents selected which of the two dimensions in a pair they considered to have contributed more to their workload. There were fifteen pairs the respondents needed to select from, and the weight of each dimension was equal to the number of times that dimension was selected divided by fifteen. The third part

was the individual information about the physicians.

The response endpoints of items are displayed in Table 1. Items were scored as follows: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, the average scores of all items for a corresponding dimension were multiplied by the dimension weight to produce the dimension scores, and then the total scores were the sum of all dimension scores.

Table 1 Dimensions and items of physician mental workload scale

Dimensions and item	Endpoints (0-100)
A Mental demand	
A1 How many cognitive activities (e.g. sensation, perception, remembering, thinking, calculating, attention etc.) were required during your medical work?	
A2 How much emotion and feeling were input (e.g. empathy, sympathy, enthusiasm, negative emotion restraining etc.) during your medical work?	
A3 How hard did you have to work to overcome difficulties in accomplishing your medical work?	
B Physical demand	
B1 How many physical activities were required (e.g. standing, sedentary, controlling, repetitive action etc.) in your medical work?	
B2 How intensive was the physical activity during your medical work? (Was the work restful or laborious?)	
C Temporal demand	
C1 How much time pressure did you feel in your medical work? (Daily medical work required time was more/less than available time?)	
C2 How frequently did you have to complete multiple tasks at the same time (work overlap) in your medical work?	
D Perceived risk	
D1 How much risk did you perceive (e.g. medical dispute etc.) in your medical work?	
E Frustration level	

E1 How depressed or frustrated did you feel in your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
E2 How anxious or irritated did you feel in your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
F Performance	
F1 How successful do you think you were in accomplishing the goals in your medical work?	heaviest workload ... no workload 0 10 20 30 40 50 60 70 80 90 100
F2 How satisfied were you with the outcome in your medical work?	heaviest workload ... no workload 0 10 20 30 40 50 60 70 80 90 100

Statistical analysis

Descriptive statistics were used to show the characteristics of the respondents, including gender, age, educational level (i.e. Ph.D. degree, Master degree, Undergraduate), job title (i.e. senior, middle, junior), work years, hospital level (i.e. tertiary hospitals, secondary hospitals, first-tier hospital), work hours per week, number of outpatients serviced per day, and self-perceived health status.

For the reliability of the scale, Cronbach's alpha was used to assess the internal consistency of each instrument component. Values of 0.70 or greater for Cronbach's alpha were considered acceptable. [37]

Content validity index (CVI) of each item was calculated to assess the accuracy of the scale using the scores of 1-4. Experts were invited to assess the items with a scale of 1 representing the item not relevant to corresponding dimension and 4 representing the item closely relevant to the corresponding dimension. Corrected item-content validity index (I-CVI) and average scale-content validity index (S-CVI/Ave)

were calculated. Corrected I-CVI of 0.78 or greater and S-CVI/Ave of 0.90 or greater were considered acceptable. [39]

The test of construct validity was tested by correlation coefficient method, exploratory factor analysis, and confirmatory factor analysis. Items whose values for item-total correlation and dimension-total correlation below 0.40 were revised or removed from the scale. A Bartlett's test of sphericity scores lower than 0.05 and a Kaiser-Meyer–Olkin (KMO) measure of sampling adequacy higher than 0.70 and closer to 1 were considered appropriate for factor analysis. [40] Exploratory factor analysis and confirmatory factor analysis was used to further explore and confirm the structure of the scale. For the exploratory factor analysis, we used a varimax rotated method to examine whether the structure matched the framework. For the confirmatory factor analysis, the criterion of model fit indices were listed as follows: $\chi 2/df<3$; root mean square error of approximation (RMSEA) < 0.05; root of mean square residual (RMR)<0.05; goodness of fit (GFI) >0.90, comparative fit index (CFI)>0.90, Tucker-Lewis index (TLI)>0.90. [41]Statistical analyses were performed with SPSS V. 21 (IBM Corp., Armonk, NY, USA) and AMOS V.17 (IBM Corp., Armonk, NY, USA).

Patient and public involvement

This project involves physicians in the second and third phase. All participants were voluntary and no incentives were provided for the participation. Meanwhile, Participants were not directly involved in the design and recruitment of this study. The results were not to be provided back to participants.

Results

Sample characteristics

396 responses (online survey) were received and 11 questionnaires were excluded due to incomplete demographic information. There were no issues of floor or ceiling effects as questions to every item were responded as required in the form of a web-based survey. The characteristics of the participants are presented in Table 2.

Table 2 Respondents' characteristic of samples

Variable	Valid sample	Percentage（%）
Gender		
Male	200	51.9
Female	185	48.1
Age		
<45	258	67.0
45-55	121	31.4
>55	6	1.6
Educational level		
Ph.D. degree	36	9.4
Master degree	48	12.5
Undergraduate	184	47.8
Junior college	117	30.3
Job title		
Senior	83	21.6
Middle	146	37.9
Junior	156	40.5
Work years in his/her institutions (year)		
1-5	122	31.7
6-10	91	23.6
11-15	58	15.1
⩾16	114	29.6
Hospital level		

Variable	Valid sample	Percentage（%）
Tertiary hospitals	130	33.8
Secondary hospitals	124	32.2
First-level hospitals	131	34.0
Work hours per week		
<=40	73	19.0
41-60	152	39.5
>60	160	41.5
Number of outpatients serviced per day		
<20	135	35.1
20-50	171	44.4
>50	79	20.5
Self-perceived health status		
Poor	65	16.9
Fair	242	62.9
Good	78	20.2

Reliability of physician mental workload scale

Each of the six components demonstrated at least satisfactory internal consistency higher than 0.70, with Cronbach's α in the range of 0.70-0.90. Cronbach's α for the whole scale was up to 0.81, which indicated that the scale has good reliability.

Validity of physician mental workload scale

The corrected I-CVI of each item ranged from 0.85-1 (Table 3), which is higher than 0.78. The S-CVI/Ave was 0.96, which is higher than 0.90. All of these showed good content validity of the scale.

The correlation matrix between items and total scores was inspected to confirm the convergent validity, which was indicated by reasonable coefficients of 0.40 and above, except for F1 and F2 (Table 3). Calculated correlation coefficients between dimensions and total scores had a range of 0.37-0.72, which was an additional index

that showed that dimensions and total scores have a good convergent validity. Also, the correlation coefficients among the dimensions were lower than the correlation coefficients between the dimension-total scores, which indicated that the scale has a good discriminant validity (Table 4).

Table 3 Content validity and correlation coefficient of item-total scores of the scale

Items	Corrected I-CVI	Item-total correlations
A1 Cognitive activity	1	0.57
A2 Emotion and feeling	0.85	0.57
A3 Overcoming difficulties	0.85	0.59
B1 Physical activity	1	0.57
B2 Intensity of physical activity	1	0.65
C1 Time pressure	1	0.69
C2 Multiple task	0.85	0.69
D1 Risk concern	1	0.64
E1 Depressed or frustrated	1	0.75
E2 Anxious or irritated	1	0.75
F1 Successful	1	0.33*
F2 Satisfied	1	0.31*

*item-total scores were below than 0.4

Table 4 Correlation coefficient matrix between dimensions and total scores of the scale

Dimensions	A	B	C	D	E	F	Total scores
A Mental demand	1						
B Physical demand	0.43	1					
C Temporal demand	0.52	0.47	1				
D Perceived risk	0.46	0.38	0.44	1			
E Frustration level	0.40	0.36	0.54	0.51	1		
F Performance	0.09	0.05	0.01	0.05	0.13	1	
Total scores	0.61	0.52	0.68	0.68	0.72	0.37	1

Exploratory factor analysis of physician mental workload scale

The KMO sample adequacy measurement was 0.81, which was higher than the recommended value of 0.70, and Bartlett's test of sphericity with Chi-square value 1950.70 and p<0.000. Thus, the data were suitable for factor analysis. Considering experts' suggestion, we selected 6 principal components in the exploratory factor analysis and results showed that the six-dimensional model explained 81.88% of the total variance (Table 5).

The component 1"mental demand" was developed from 3 items that asked for feeling or memory requirement, emotional requirement, and the effort input to overcome difficulties, with the factor loading in the range of 0.74-0.81. The component 2"frustration level" consisted of 2 items that asked about anxiety and level of depression or frustration, and the factor loading was in the range of 0.86-0.88. The component 3 "physical demand" consisted of 2 items that related to strength requirement and the intensity of work time with a factor loading in the range of 0.84-0.90.

The component 4 "temporal demand" constituted 2 items that asked about the ratio of required time to available time and frequency of completing multiple tasks, with the factor loading in the range of 0.77-0.82. There were 2 items in the "performance demand" component 5, which related to the sense of achievement and job satisfaction regarding work outcome, with the factor loading in the range of 0.85-0.90. The component 6 "perceived risk" included only 1 item that explained the perception of risk in conducting tasks (such as medical dispute and risk of being

infectious) with a factor loading of 0.84.

Table 5 Factor loadings for the rotated component matrix: varimax rotated components

Items	Components					
	1	2	3	4	5	6
A2 Emotion and feeling	0.81					
A1 Cognitive activity	0.76					
A3 Overcoming difficulties	0.74					
E1 Depressed or frustrated		0.88				
E2 Anxious or irritated		0.86				
B1 Physical activity			0.90			
B2 Intensity of physical activity			0.84			
C1 Time pressure				0.82		
C2 Multiple task				0.77		
F1 Successful					0.90	
F2 Satisfied					0.85	
D1 Risk concern						0.84

Confirmatory factor analysis of physician workload scale

The six-factor model obtained after exploratory factor analysis was tested by confirmatory factor analysis using the maximum likelihood estimation method. The goodness-of-fit model was listed as follows: $\chi 2/df=1.84$ (<3), RMR= 0.04 (<0.05), and GFI=0.97 (>0.9), CFI=0.98 (>0.9), TLI=0.97 (>0.9), RMSEA=0.05 ($\leqslant$0.05). Referring to the criterion listed above, the model was a good fit for the data.

Discussion

The purpose of this study was to develop a mental workload scale for physicians

and explore the validity and reliability of the scale. According to the results of the tests, the scale is reliable and valid, hence, it is considered as an effective instrument for assessing physician mental workload in Chinese comprehensive public hospitals. Results show a six-dimensional model which includes aspects related to mental demand, physical demand, temporal demand, perceived risk, frustration level, and performance. Compared with other relevant scales, this scale only includes 12 items, which have strength in the aspects of length, so it could be completed in a short time. As for its contents, the dimensions of perceived risk and temporal demand are especially distinctive for Chinese physician mental workload.

The whole Cronbach's α of the 12 items was beyond 0.7, which indicated that it had good reliability. For the content validity, the corrected I-CVI was higher than 0.78 and the S-CVI/Ave was more than 0.9, which showed that it had a good content validity. For the construct validity, except for F1 and F2, the correlation coefficient between item and total scores was more than 0.4, which showed that the construct validity was at a good level. The item F1 and F2 were to explain the aspect of the performance. In our study, the item-total scores of the two items in the dimension of performance were near to 0.4, and perhaps, would have been relevant in a reverse scoring. Consistent with previous research on NASA-TLX, performance dimension shows a limited practical relevance since it is influenced by variations in physical load. [42] Other research reported that subjective assessments of mental workload may not provide an accurate estimation of the performance dimension. [26] Considering this information, we retained the two items but revised their description.

The specific dimension perceived risk, which is a different dimension and not included in the framework of NASA-TLX and SWAT, is highly associated with physician mental workload in China. There tends to be an estranged relationship between physicians and patients which puts physicians at a dangerous risk of being assaulted, by patients or visitors in their line of work. [43] According to the statistics, 96% of medical staff have been abused or injured in 2012. [44] The physician-patient relationship is becoming more and more fragile and has reached an unprecedented poor level in China. [45] The tense relationship resulted in heavy psychological workloads during their work.

Another dimension of temporal demand is also fully specific. The gap between healthcare demand and supply (thus doctor-patient ratio) in China has caused physicians in the secondary and tertiary hospital settings to become overworked. [46] They always need to work overtime, even though they conduct more than one task at the same time. According to the report by the Chinese Medical Association in 2018, physicians in tertiary hospitals had an average workweek of 51.05 hours, which was more than legal 40 hours per week. [47] Research has reported that physicians may feel stressed when poor scheduling leaves them pressed for time. [48] Mental workload encompasses the subjective experience of a given task load. [49] High task demands require plenty of time and evoke high mental effort and heavy workload for physicians. [50] Then, the worse the experience of the task load, the higher the mental workload. [51]

Although we have attempted an accurate examination of the measurement properties of the physician mental workload scale by using qualitative and

quantitative methods, there were still some limitations that merit discussions. Firstly, among six dimensions, perceived risk only included one item, which may result in measurement error. Secondly, there was a potential reporting bias in the self-reported measurements of the workload among physicians. Thirdly, all respondents voluntarily decided to take part in the survey. This means physicians who were overburdened at the time of the survey may not have time to take part in the survey, resulting in a selection bias. Thus, all of these above need continued research to improve this scale, meanwhile, burnout was relevant to the mental workload, which was a direction for further exploration as well.

Conclusion

Creating new items from a subjective perspective is of paramount importance in investigating the Chinese physician's workload. Physician mental workload scale has acceptable preliminary psychometric properties with six dimensions and twelve items. The use of a physician mental workload scale can help us to find the main stressors in physician mental workload and to implement targeted optimization strategies to mitigate these stressors in an effort to ameliorate the physical and mental health of physicians. This, consequently, will help us to improve the quality and efficiency of healthcare delivery in hospital settings.

Abbreviations

CVI: Content Validity Index;

RMR: Root of Mean Square Residual; GFI= Good of Fitness Index; CFI= Comparative Fit

Index; TLI=Tucker-Lewis index; CR= Critical Ratio; CV=Coefficient of Variation; NASA-TLX=NASA- Task Load Index; SWAT= Subjective Workload Assessment Technique;

EMRs= Electronic Medical Records

Funding Statement: This work was supported by the National Natural Science Foundation of China (grant number 71774062). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgment We would like to thank all the participants involved in this research for their time and contributions.

Author Contributors: Yinhuan Hu and Qiang Fu designed the study; Chuntao Lu and Jinzhu Xie made formal analysis; Yinhuan Hu obtained funding; Jinzhu Xie and Lu Deng took part in investigation; Liuming Wang, Chuntao Lu and Lu Deng were involved in data cleaning; Chuntao Lu wrote original draft; Samuel Governor, Qiang Fu and Chao Li contributed to the interpretation of the results; Chao Li and Chuntao Lu made critical revision of the manuscript; All authors have read and approved the final manuscript.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Patient consent: Not required.

Ethics approval: The Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (IORG No. IORG0003571) gave the final approval for the study.

Data sharing: The datasets used and/or analyzed during the current study are

available from the corresponding author on reasonable request.

Reference

1. Győrffy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians[J]. *Hum resour health*, 2016, 14: 12.
2. Lu L, Gao M, Yang T, et al. The updated Physician's Pledge and Chinese junior physicians. *The Lancet.* 2018;391:305.
3. Karahan A, Kav S, Abbasoglu A, et al. Low back pain: prevalence and associated risk factors among hospital staff. *J Adv Nurs.* 2009;65: 516-524.
4. Smith-Miller CA, Shaw-Kokot J, Curro B, et al. An integrative review: fatigue among nurses in acute care settings. *J Nurs Adm.* 2014;44:487-494.
5. Arvidsson I, Simonsen JG, Balogh I. Discrepancies in pain presentation caused by adverse psychosocial conditions as compared to pain due to high physical workload? *Work*. 2012;41:2472-2475.
6. Manty M, Moller A, Nilsson C, et al. Association of physical workload and leisure time physical activity with incident mobility limitations: a follow-up study. *Occup Environ Med*. 2014; 71:543-548.
7. Ridley C. Relating nursing workload to quality of care in child and adolescent mental health inpatient services. *Int J Health Care Qual Assur*.2007;20:429-440.
8. Michtalik HJ, Pronovost PJ, Marsteller JA, et al. Identifying potential predictors of a safe attending physician workload: a survey of hospitalists. *J Hosp Med*.2013; 8:644-646.
9. Michtalik HJ, Yeh HC, Pronovost PJ, et al. Impact of attending physician workload on patient care a survey of hospitalists. *JAMA Intern Med*.2013; 173:375-377.
10. Garcia-Lacalle J. Balancing physician workload and patient satisfaction: a dilemma for hospital managers. *Public Money & Management*.2010;30:183-188.
11. Steyrer J, Schiffinger M, Huber C,et al. Attitude is everything? The impact of

workload, safety climate, and safety tools on medical errors: a study of intensive care units. *Health Care Manage R*.2013;38:306-316.

12. Pedroja AT, Blegen MA, Abravanel R. The relationship between hospital systems load and patient harm. *J Patient Saf*. 2014;10:168-175.
13. Williams E S, Rondeau K V, Xiao Q, et al. Heavy physician workloads: impact on physician attitudes and outcomes. *Health Care Manage R*, 2007, 20: 261-269.
14. Hart SG. Theory and Measurement of Human Workload. *J Oper Res Soc*.1988;*39*:725-734.
15. Cain B. A review of the mental workload literature. Toronto: Defence Research And Development, 2007.
16. Lysaght RJ, Hill SG, Dick AO, et al. Operator workload: Comprehensive review and evaluation of operator workload methodologies. Analytics inc willow grove pa, 1989.
17 Theorell T, Hammarström A, Aronsson G, et al. A systematic review including meta-analysis of work environment and depressive symptoms[J]. *BMC public health*, 2015, 15(1): 738.
18 Noseworthy J, Madara J, Cosgrove D, et al. Physician burnout is a public health crisis: a message to our fellow health care CEOs[J]. *Health Affairs Blog*, 2017.
19 Olson K, Sinsky C, Rinne S T, et al. Cross‐sectional survey of workplace stressors associated with physician burnout measured by the Mini‐Z and the Maslach Burnout Inventory[J]. *Stress and Health*, 2019, 35(2): 157-175.
20. Weigl M, Müller A, Holland S, et al. Work conditions, mental workload and patient care quality: a multisource study in the emergency department. *BMJ Qual Saf.* 2016; 25: 499-508.
21. European Commission. European Pact for health and mental health. http://ec.europa.eu/health/mental_health/policy_es. (Accessed on 24th Sep 2018)
22. Bertram DA, Opila DA, Brown JL, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*.1992;30: 95-104.

23. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol*. 1998;*52*:139-183.

24. Rubio S, Díaz E, Martín J, et al. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Appl Psychol.* 2004; 53: 61-86.

25. Hill SG, Iavecchia HP, Byers JC. Comparison of four subjective workload rating scales. *Hum Factors*.1992;*34*:429.

26. Rubio-Valdehita S, Lopez-Nunez MI, Lopez-Higes R, et al. Development of the CarMen-Q Questionnaire for mental workload assessment. *Psicothema*. 2017; 29:570-576.

27. Colligan L, Potts HW, Finn CT, et al. Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *Int J Med Inform.*2015; 84:469-476.

28. Hoonakker P, Carayon P, Gurses A, et al. Measuring Workload Of ICU Nurses with a Questionnaire Survey: The NASA Task Load Index (TLX). *IIE Trans Healthc Syst Eng*. 2011;1:131-143.

29. Malekpour F, Malekpour A R, Mohammadian Y, et al. Assessment of mental workload in nursing by using NASA-TLX. *The Journal of Urmia Nursing and Midwifery Faculty*.2014; 11: 892-899.

30. Reid GB, Nygren TE. The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Adv Psychol*.1988; *52*.

31. The COPSOQ III questionnaire. https://www.copsoq-network.org/assets/Uploads/annex1-Dimensions-and-items-in-the-COPSOQ-III-questionnaire-060718.pdf (Accessed on 14th June 2019)

32. Zhou H, Han X, Zhang J, et al. Job satisfaction and associated factors among medical staff in tertiary public hospitals: Results from a national cross-sectional survey in China. *Int J Occup Env Heal*. 2018; 15: 1528.

33. Sun J, Ma J, Hu G, et al. Welfare, wellness, and job satisfaction of Chinese physicians: A national survey of public tertiary hospitals in China. *Int J Health Plan M*. 2017; 32: 270-284.

34. The Lancet, Ending violence against doctors in China. *The Lancet*. 2012; 379:1764.

35. Chinese Hospital Association. Chinese patient safety goal 2014-2015. http://www.cha.org.cn/plus/view.php?aid=14626 (Accessed on 24th Sep 2018)

36. Power M, Quinn K, Schmidt S, et al. Development of the WHOQOL-old module. *Qual Life Res*. 2005; 14 (10):2197-2214.

37. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York: Guilford publications, 2015:69-70.

38. Ford JK, MacCullum RC, Tait M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Pers Psychol*. 1986; 39:291-314.

39.Lynn MR. Determination and quantification of content validity. *Nursing Res*. 1986; 35:382-385.

40. DiIorio C K. Measurement in health behavior: methods for research and evaluation. John Wiley & Sons, 2006.

41. Tabachnick B, Fidell LS. Using Multivariate Statistics. 6th ed., Pearson, Boston, 2013: 720-726.

42. DiDomenico A, Nussbaum MA. Effects of different physical workload parameters on mental workload and performance. *Int J Ind Ergonom* 2011. 41;255-260.

43. Tang L, Guan M. Rise of Health Consumerism in China and Its Effects on Physicians' Professional Identity and the Physician-Patient Relationship and Communication. *Health commun*.2018; 33:636-642.

44. Yueju L. Violence against doctors in China. *The Lancet*. 2014; 384:745.

45. Sun T, Gao L, Shi Y. Workplace violence, psychological stress, sleep quality and subjective health in chinese doctors: a large cross-sectional study. *BMJ OPEN*. 2017;1-8.

46. Wen J, Cheng Y, Hu X. Workload, burnout, and medical mistakes among physicians in China: A cross-sectional study. *Biosci trends*. 2016; 10:27-33.

47. Chinese Medical Doctor Association. White Paper on the Practice of Medi cine by Chinese Physicians. http://www.cmda.net/u/cms/www/201807/06181247ffex.pdf. (Accessed on 14th June 2019)

48. Horner RD, Szaflarsk JP, Ying J. Physician Work Intensity Among Medical Specialties:

Emerging Evidence on its Magnitude and Composition. *Med care*. 2011; 49:1007-1011.

49. Charles RL, Nixon J. Measuring mental workload using physiological measures: A systematic review[J]. Appl ergon. 2019; 74: 221-232.
50. Mosaly PR, Mazur LM, Yu F. Relating Task Demand, Mental Effort and Task Difficulty with Physicians' Performance during Interactions with Electronic Health Records (EHRs). *Int J Hum-Comput Int.* 2017; 34:467-475.
51. Colle HA, Reid GB. Context effects in subjective mental workload ratings. Hum factors. 1998; 40: 591-600.

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Methods			

Study design	#4	Present key elements of study design early in the paper	5
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	7
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-10
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	n/a
Bias	#9	Describe any efforts to address potential sources of bias	17
Study size	#10	Explain how the study size was arrived at	7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	n/a
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	9-10
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	n/a
Statistical methods	#12c	Explain how missing data were addressed	7
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	9-10
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed	6-7

		eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	
Participants	#13b	Give reasons for non-participation at each stage	n/a
Participants	#13c	Consider use of a flow diagram	n/a
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	11
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	7
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	12-15
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	15-16
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	17-18
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	17

Generalisability	#21	Discuss the generalisability (external validity) of the study results	18
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	19

BMJ Open

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030137.R2
Article Type:	Original research
Date Submitted by the Author:	07-Aug-2019
Complete List of Authors:	Lu, Chuntao; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management; Jingchu University of Technology Affiliated Central Hospital Hu, Yinhuan; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Fu, Qiang; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Governor, Samuel ; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Wang, Liuming; Huazhong University of Science and Technology Tongji Medical College, Tongji Hospital Li, Chao; Jingchu University of Technology Affiliated Central Hospital Deng, Lu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Xie, Jinzhu; The Third People's Hospital of Hubei Province
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Mental health
Keywords:	Physician, MENTAL HEALTH, Workload, Survey and Questionnaires, Hospitals, Public

Title page

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Chuntao Lu Yinhuan Hu* Qiang Fu Samuel Governor Liuming Wang Chao Li Lu Deng Jinzhu Xie

Corresponding author:

*Yinhuan Hu

Address: No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Email: hyh288@hotmail.com

Telephone number: 13554285879

Co-author:

Chuntao Lu: 1.School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China.

2. Jingchu University of Technology Affiliated Central Hospital, No.39 Xiangshan Rd., Jingmen, Hubei 448126, P.R. China

Yinhuan Hu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Qiang Fu: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Samuel Governor: Department of Epidemiology and Biostatistics, College for Public

Health and Social Justice, Saint Louis University, Missouri, MO, United States

Liuming Wang: Tongji Hospital, Tongji Medical College of Huazhong University of Science and Technology, No.1095 Jiefang Rd., Wuhan, Hubei 430030, P.R. China

Chao Li: Jingchu University of Technology Affiliated Central Hospital. No.39 Xiangshan Rd., Jingmen, Hubei 448126, P.R. China

Lu Deng： School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Jinzhu Xie: The Third People's Hospital of Hubei Province, No.26 Zhongshan Rd., Wuhan, Hubei 430030, P.R. China.

Abstract

Objective: The purpose of our study is to develop and perform reliability and validity assessments of mental workload scale for physicians in China.

Design: Three phases, involving 396 physicians from different-level of comprehensive public hospitals in China, were conducted to develop the instrument. In the first phase, an initial item pool was developed through a systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey. The third phase tested the reliability and validity of the instrument.

Setting: Public hospitals in China.

Participants: A total of 396 physicians from different tiers of comprehensive public hospitals in China participated in this study in 2018.

Primary and secondary outcome measures: Cronbach's alpha, content validity index, the correlation coefficient between items and dimensions, and indices of confirmatory factor analysis**.**

Results: Six dimensions (mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance), and twelve items were identified in the instrument. For reliability, Cronbach's α for the whole scale was 0.81. For validity, the corrected item-content validity index of each item ranged from 0.85 to 1, and correlation coefficients between dimensions and total scores had a range of 0.37-0.72. The results of the confirmatory factor analysis showed that the goodness of indices of the scale was reasonably well.

Conclusion: The instrument showed good reliability and validity, and it is useful for

diagnosing the mental workload of physicians.

Keywords: Physician; Mental Health; Workload; Survey and Questionnaires; Hospitals, Public

Article Summary

Strengths and limitations of this study

This is the first study to develop a measurement about physician mental workload from a subjective perspective in China.

Qualitative and quantitative methods were involved in item selection.

There is a potential reporting bias in the self-reported measurements of the workload among physicians.

There existed a selection bias due to all respondents voluntarily rather than randomly took part in the survey.

Among six dimensions, perceived risk only included one item, which may result in measurement error.

Introduction

Internationally, there has been a focus on the relationship between physicians' workload and their health. [1] Physicians' health is highly associated with the workload. [2] Excessive workload impacts physician's health [3-4] and increases the risk of work-related musculoskeletal disorders (MSDs). [5-6] Exposure to workload is related to adverse effects in medical errors [7] and adverse incidents. [8] Physician workload could be a negatively contributing factor to patients' perceived quality of care, [9] and affects patient satisfaction [10] and safety. [11-12] It is possible that these stressors have reached a point where they pose a severe problem for the entire healthcare system. [13] Thus, the unreasonable and overwhelming workload has adverse effects on physicians, patients, and healthcare organizations. [14]

The workload is thought to be multidimensional and multifaceted. [15] One aspect of workload includes the subjective psychological experiences of the human operator. [16] Mental workload as a kind of workload has emerged as one of the most critical occupational risk factors, which results in burnout or anxiety. [17] Lack of control over workload was expected to correlate most highly with burnout. [18,19]Heavy mental workload can lead to serious health problems (cardiovascular diseases, digestive problems, etc.) for physicians as well. [17] Meanwhile, the excessive mental workload can also lead to an inferior quality of care service. [20] Currently, The European Pact for Mental Health and Welfare is conducting mental workload assessments to promote physical and mental wellbeing. [21]

Different tools have been proposed to assess mental workload. Previous research

has established a brief instrument with six items to measure physician mental workload. [22] The NASA-Task Load Index (NASA-TLX) scale, which was widely used in measuring mental workload [23] has proven to be a sensitive, valid, and reliable instrument [24] and can be used in human factors research. [25] The researcher has localized it as a 29-item questionnaire in Spain to measure workers' mental workload. [26] The existing body of research on NASA-TLX suggested that it could be used to measure nurse workload in health care settings. [27-29] In the same vein, the Subjective Workload Assessment Technique (SWAT) is a subjective rating technique with three dimensions of time load, mental effort load, and psychological stress load, which is used to assess mental workload as well. [30] It has been successfully applied in the mental workload assessment of several aircraft multitask conditions, such as assessing the mental workload of different systems of air defense. [24] Copenhagen Psychosocial Questionnaire was a wide-spread tool used in the industrial or the branch of the service in Europe, which included the main dimensions of the most influential psychosocial theories at work. [31] Together these tools provide essential insights into workload measurement in health care management, especially in nurse workload measurement. However, there is still different workload between physicians and nurses essentially, meanwhile, these measurements were designed for other workers, so we do not ensure these tool can be directly used in physician mental workload measurement. Specifically, mental workload measurement needs to be developed for physicians.

With the increasing of patient demands for health, physicians tend to have a

heavier workload, worse physical health, more mental strain, and more intense relationships with patients in China. [32] Data from several studies suggest that most physicians work more than 10 hours a day [33] to manage outpatients and inpatients. On average, a physician in a tertiary hospital is responsible for 8.10 outpatients and 2.70 beds per day. [32] However, they even have been abused, injured, and in extreme cases, murdered by patients or their relatives in hospitals across China, [34] which resulted in extremely mental workload. Establishing the workload measurement system for medical personnel has been incorporated into Chinese Patient Safety Goals by the Chinese Hospital Association. [35]

Existing researches about workload measurement instruments are concentrated on objective workload in China, for example, the measurement of work time. Whereas physicians' mental workload is an indispensable problem, there are few instruments exploring physician mental workload in China. The purpose of this paper is to develop a scientific mental workload instrument, which can be used to measure or assess the actual mental workload of physicians.

Methods

Study design

The instrument was developed in three phases. In the first phase, an initial item pool was developed by integrating previous studies through a systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and

a pilot survey in 2017. The third phase was testing the psychometric properties of the instrument, including its reliability and validity, through a study conducted in 2018 in comprehensive public hospitals in China.

Framework and items of generation and selection

Based on the framework of NASA-TLX and SWAT, we combined the current situation of Chinese physicians' workload to determine the item pool. Six dimensions and fifteen items were sent to 20 experts (including physicians, hospital managers, researchers, and human resource managers) for consultation. According to two-round expert consultation, we deleted four items, added a new item (the intensity of physical activities), and revised the description of all the items. Then, there were six dimensions (physical demands, mental demands, temporal demands, effort, frustration level, and performance) and twelve items, which consisted of the pre-scale with a range from 0 to 100.

In the pre-survey analysis, we selected three hospitals (one tertiary hospital, one secondary hospital, and one first-tier hospital) through conveniently sampling. A sample of 80 physicians was surveyed with a web-based scale during November and December 2017. Finally, 74 samples were validated to conduct item selection. Items were refined based on the following indexes or methods: critical ratio (CR), coefficient of variation (CV), correlation analysis, [36] Cronbach's α, [37] and exploratory factor analysis (EFA). [38]

If an item was eliminated by any of the above methods, then the item was

deleted or revised. The final scale consisted of six dimensions (mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance), and twelve items (Table 1).

Data collection for testing the validity and reliability of the scale

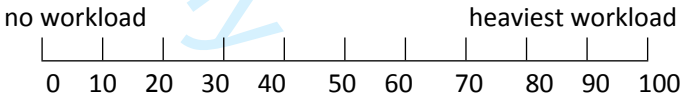
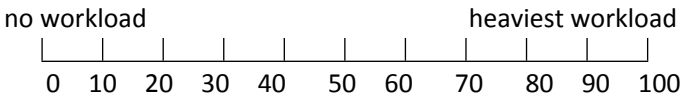
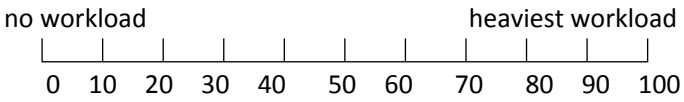
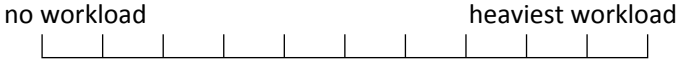
To check the validity and reliability of the developed scale, we planned to survey 400 respondents (physicians are working in hospitals) from different tier hospitals (two tertiary hospitals, two secondary hospitals, and two first-tier hospitals). These hospitals were randomly selected from Hubei province, China. We used a website--wenjuanxing (www.wjx.cn), which is a wildly used platform to conduct surveys in China, to develop an electronic questionnaire to survey physicians. The electronic questionnaire is a kind of access code or website, and respondents can scan the access code or click the website by their phones to complete the questionnaire. We sent the access code or website to the human resources at each participating hospital. Then, human resources sent the access code to the physicians' online communication group at each hospital. Three hundred ninety-six physicians voluntarily participated in the survey before March 2018; finally, 11 invalid samples were deleted.

There was a detailed description of the guidance of the scale, which showed that our scale was anonymous, and all physicians were voluntary to answer this question. Meanwhile, we described that our survey aimed to develop a physician mental workload scale, so the results would not be used for other purposes. The physician mental workload scale included three parts. The first part of the scale included twelve

items, and respondents need to score them one by one. The second part was a table which included 15 pairs in our scale, which was used to collect the weights of each dimension. Every two dimensions formed a pair (e.g., Mental demands vs. Physical demands, Mental demands vs. Temporal demands, etc.). Thus, there were fifteen pairs the respondents needed to select from. The respondents need to choose a dimension which was contributed more to their workload in each pair. Then, the weight of each dimension was equal to the number of times that dimension was selected divided by fifteen. The third part of the sclale was the individual information about the physicians.

The response endpoints of items are displayed in Table 1. Items were scored as follows: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, the average scores of all items for a corresponding dimension were multiplied by the dimension weight to produce the dimension scores, and then the total scores were the sum of all dimension scores.

Table 1 Dimensions and items of physician mental workload scale

Dimensions and item	Endpoints (0-100)
A Mental demands	
A1 How demanding of cognitive activities (e.g., sensation, perception, remembering, thinking, calculating, attention, etc.) were required during your medical work?	
A2 How demanding of emotion and feeling were input (e.g., empathy, sympathy, enthusiasm, negative emotion restraining, etc.) during your medical work?	
A3 How hard did you have to work to overcome difficulties in accomplishing your medical work?	
B Physical demands	
B1 How demanding of physical activities were	

required (e.g., standing, stationary, controlling, repetitive action, etc.) in your medical work?	0 10 20 30 40 50 60 70 80 90 100
B2 How intensive was the physical activity during your medical work? (Was the work restful or laborious?)	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
C Temporal demands	
C1 How pressured did you feel about work time in your medical work? (Daily medical work required time was more/less than available time?)	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
C2 How frequently did you have to complete multiple tasks at the same time (work overlap) in your medical work?	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
D Perceived risk	
D1 How risky did you perceive (e.g., medical dispute, etc.) in your medical work?	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
E Frustration level	
E1 How depressed or frustrated did you feel in your medical work?	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
E2 How anxious or irritated did you feel in your medical work?	no workload heaviest workload 0 10 20 30 40 50 60 70 80 90 100
F Performance	
F1 How successfully do you think you were in accomplishing the goals in your medical work?	heaviest workload no workload 0 10 20 30 40 50 60 70 80 90 100
F2 How satisfied were you with the outcome in your medical work?	heaviest workload no workload 0 10 20 30 40 50 60 70 80 90 100

Statistical analysis

Descriptive statistics were used to show the characteristics of the respondents, including gender, age, educational level (i.e. Ph.D. degree, Master degree, Undergraduate), job title (i.e. senior, middle, junior), work years, hospital-level (i.e. tertiary hospitals, secondary hospitals, first-tier hospital), work hours per week, number of outpatients serviced per day, and self-perceived health status.

For the reliability of the scale, Cronbach's alpha was used to assess the internal consistency of each instrument component. Values of 0.70 or higher for Cronbach's alpha were considered acceptable. [37]

Content validity index (CVI) of each item was calculated to assess the accuracy of the scale using the scores of 1-4. Experts were invited to evaluate the items with a scale of 1 representing the item not relevant to corresponding dimension and 4 representing the item closely related to the corresponding dimension. Corrected item-content validity index (I-CVI) and average scale-content validity index (S-CVI/Ave) were calculated. Corrected I-CVI of 0.78 or higher and S-CVI/Ave of 0.90 or greater were considered acceptable. [39]

The test of construct validity was tested by correlation coefficient method, exploratory factor analysis, and confirmatory factor analysis. Items whose values for item-total correlation and dimension-total correlation below 0.40 were revised or removed from the scale. A Bartlett's test of sphericity scores lower than 0.05 and a Kaiser-Meyer–Olkin (KMO) measure of sampling adequacy higher than 0.70 and closer to 1 were considered appropriate for factor analysis. [40] Exploratory factor analysis and confirmatory factor analysis were used to explore and confirm the structure of the scale. For the exploratory factor analysis, we used a varimax rotated method to examine whether the structure matched the framework. For the confirmatory factor analysis, the criterion of model fit indices were listed as follows: $\chi 2/df<3$; root mean square error of approximation (RMSEA) < 0.05; root of mean square residual (RMR)<0.05; goodness of fit (GFI) >0.90, comparative fit index (CFI)>0.90, Tucker-

Lewis index (TLI)>0.90. [41]Statistical analyses were performed with SPSS V. 21 (IBM Corp., Armonk, NY, USA) and AMOS V.17 (IBM Corp., Armonk, NY, USA).

Patient and public involvement

Our participants were physicians working in hospitals. They took part in the pre-survey and formal survey to complete our scale. All participants were voluntary, and no incentives were provided for their participation. Meanwhile, Participants were not directly involved in the design and recruitment of this study. The results were not provided back to participants.

Results

Sample characteristics

Three hundred ninety-six responses (online survey) were received, and 11 questionnaires were excluded due to incomplete demographic information. There were no issues about floor or ceiling effects as questions to every item were responded as required in the form of a web-based survey. The characteristics of the participants are presented in Table 2.

Table 2 Respondents' characteristic of samples

Variable	Valid sample	Percentage（%）
Gender		
Male	200	51.9
Female	185	48.1
Age		
<45	258	67.0

Variable	Valid sample	Percentage（%）
45-55	121	31.4
>55	6	1.6
Educational level		
Ph.D. degree	36	9.4
Master degree	48	12.5
Bachelor degree	184	47.8
Below bachelor college	117	30.3
Job title		
Senior title	83	21.6
Middle title	146	37.9
Junior title	156	40.5
Work years in his/her institutions (year)		
1-5	122	31.7
6-10	91	23.6
11-15	58	15.1
>=16	114	29.6
Hospital level		
Tertiary hospitals	130	33.8
Secondary hospitals	124	32.2
First-tier hospitals	131	34.0
Work hours per week		
<=40	73	19.0
41-60	152	39.5
>60	160	41.5
Number of outpatients serviced per day		
<20	135	35.1
20-50	171	44.4
>50	79	20.5
Self-perceived health status		
Poor	65	16.9
Fair	242	62.9
Good	78	20.2

Reliability of physician mental workload scale

Each of the six components demonstrated at least satisfactory internal consistency higher than 0.70, with Cronbach's α in the range of 0.70-0.90. Cronbach's α for the whole scale was up to 0.81, which indicated that the scale has excellent

reliability.

The validity of physician mental workload scale

The corrected I-CVI of each item ranged from 0.85-1 (Table 3), which is higher than 0.78. The S-CVI/Ave was 0.96, which is higher than 0.90. All of these showed good content validity of the scale.

The correlation matrix between items and total scores were inspected to confirm the convergent validity, which was indicated by reasonable coefficients of 0.40 and above, except for F1 and F2 (Table 3). Calculated correlation coefficients between dimensions and total scores had a range of 0.37-0.72, which was an additional index that showed dimensions and total scores have good convergent validity. Also, the correlation coefficients among the dimensions were lower than the correlation coefficients between the dimension-total scores, which indicated that the scale has a good discriminant validity (Table 4).

Table 3 Content validity and correlation coefficient of item-total scores of the scale

Items	Corrected I-CVI	Item-total correlations
A1 Cognitive activity	1	0.57
A2 Emotion and feeling	0.85	0.57
A3 Overcoming difficulties	0.85	0.59
B1 Physical activity	1	0.57
B2 Intensity of physical activity	1	0.65
C1 Time pressure	1	0.69
C2 Multiple tasks	0.85	0.69
D1 Risk concern	1	0.64
E1 Depressed or frustrated	1	0.75
E2 Anxious or irritated	1	0.75
F1 Successful	1	0.33*
F2 Satisfied	1	0.31*

*item-total scores were below than 0.4

Table 4 Correlation coefficient matrix between dimensions and total scores of the scale

Dimensions	A	B	C	D	E	F	Total scores
A Mental demands	1						
B Physical demands	0.43	1					
C Temporal demands	0.52	0.47	1				
D Perceived risk	0.46	0.38	0.44	1			
E Frustration level	0.40	0.36	0.54	0.51	1		
F Performance	0.09	0.05	0.01	0.05	0.13	1	
Total scores	0.61	0.52	0.68	0.68	0.72	0.37	1

Exploratory factor analysis of physician mental workload scale

The KMO sample adequacy measurement was 0.81, which was higher than the recommended value of 0.70, and Bartlett's test of sphericity with Chi-square value 1950.70 and $p<0.000$. Thus, the data was suitable for factor analysis. Considering experts' suggestion, we selected six principal components in the exploratory factor analysis, and results showed that the six-dimensional model explained 81.88% of the total variance (Table 5).

The component 1"mental demands" was developed from 3 items that asked for feeling or memory requirement, emotional requirement, and the effort input to overcome difficulties, with the factor loading in the range of 0.74-0.81. The component 2"frustration level" consisted of 2 items that asked about anxiety and level of depression or frustration, and the factor loading was in the range of 0.86-0.88. The component 3 "physical demands" consisted of 2 items that related to strength requirement and the intensity of work time with a factor loading in the range of 0.84-

0.90.

The component 4 “temporal demands” constituted two items that asked about the ratio of required time to available time and frequency of completing multiple tasks, with the factor loading in the range of 0.77-0.82. There were two items in the “performance” component 5, which related to the sense of achievement and job satisfaction regarding work outcome, with the factor loading in the range of 0.85-0.90. The component 6 “perceived risk” included only 1 item that explained the perception of risk in conducting tasks (such as medical dispute and risk of being infectious) with a factor loading of 0.84.

Table 5 Factor loadings for the rotated component matrix: varimax rotated components

Items	Components					
	1	2	3	4	5	6
A2 Emotion and feeling	0.81					
A1 Cognitive activity	0.76					
A3 Overcoming difficulties	0.74					
E1 Depressed or frustrated		0.88				
E2 Anxious or irritated		0.86				
B1 Physical activity			0.90			
B2 Intensity of physical activity			0.84			
C1 Time pressure				0.82		
C2 Multiple tasks				0.77		
F1 Successful					0.90	
F2 Satisfied					0.85	
D1 Risk concern						0.84

Confirmatory factor analysis of physician workload scale

The six-factor model obtained after exploratory factor analysis was tested by

confirmatory factor analysis using the maximum likelihood estimation method. The goodness-of-fit model was listed as follows: χ2/df=1.84 (<3), RMR= 0.04 (<0.05), and GFI=0.97 (>0.9), CFI=0.98 (>0.9), TLI=0.97 (>0.9), RMSEA=0.05 (⩽0.05). Referring to the criterion listed above, the model was a good fit for the data.

Discussion

The purpose of this study was to develop a mental workload scale for physicians and explore the validity and reliability of the scale. According to the results of the tests, the scale is reliable and valid; hence, it is considered as an effective instrument for assessing physician mental workload in Chinese comprehensive public hospitals. Results show a six-dimensional model which includes aspects related to mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance. Compared with other relevant scales, this scale only includes 12 items, which have strength in the aspects of length so that it could be completed in a short time. As for its contents, the dimensions of perceived risk and temporal demands are uniquely distinctive for physician mental workload in China.

The whole Cronbach's α of the 12 items was beyond 0.7, which indicated that it had excellent reliability. For the content validity, the corrected I-CVI was higher than 0.78, and the S-CVI/Ave was more than 0.9, which showed that it had good content validity. For the construct validity, except for F1 and F2, the correlation coefficient between item and total scores was more than 0.4, which showed that the construct

validity was at a good level. The item-total scores of the two items in the dimension of performance were near to 0.4, and perhaps, would have been relevant in a reverse scoring. Consistent with previous research on NASA-TLX, performance dimension shows a limited practical relevance since variations influence it in physical load. [42] Other study reported that subjective assessments of mental workload might not provide an accurate estimation of the performance dimension. [26] Considering this information, we retained the two items but revised their description.

The specific dimension perceived risk, which is a different dimension and not included in the framework of NASA-TLX and SWAT, is highly associated with physician mental workload in China. There tends to be an estranged relationship between physicians and patients, which puts physicians at a dangerous risk of being assaulted, by patients or visitors in their line of work. [43] According to the statistics, 96% of medical staff have been abused or injured in 2012. [44] The physician-patient relationship is becoming more and more fragile and has reached an unprecedented poor level in China. [45] The tense relationship resulted in heavy psychological workloads during their work.

Another dimension of temporal demands is also entirely specific. The gap between healthcare demand and supply (thus doctor-patient ratio) in China has caused physicians in the secondary and tertiary hospital settings to become overworked. [46] They always need to work overtime, even though they conduct more than one task at the same time. According to the report by the Chinese Medical Association in 2018, physicians in tertiary hospitals had an average workweek of 51.05

hours, which was more than legal 40 hours per week. [47] Research has reported that physicians may feel stressed when poor scheduling leaves them pressed for time. [48] Mental workload encompasses the subjective experience of a given task load. [49] High task demands require plenty of time and evoke high mental effort and heavy workload for physicians. [50] Then, the worse the experience of the task load, the higher the mental workload. [51]

Although we have attempted an accurate examination of the measurement properties of the physician mental workload scale by using qualitative and quantitative methods, there were still some limitations that merit discussions. Firstly, among six dimensions, perceived risk only included one item, which may result in measurement error. Secondly, there was a potential reporting bias in the self-reported measurements of the workload among physicians. Thirdly, all respondents voluntarily decided to take part in the survey. Physicians who were overburdened at the time of the study may not have time to take part in the investigation, which resulted in a selection bias. Thus, all of these above need continued research to improve this scale. Meanwhile, burnout was relevant to the mental workload, which was a direction for further exploration as well.

Conclusion

Creating new items from a subjective perspective is of paramount importance in investigating the Chinese physician's workload. Physician mental workload scale has acceptable preliminary psychometric properties with six dimensions and twelve items.

The use of a physician mental workload scale can help us to find the main stressors in physician mental workload, and to implement targeted optimization strategies to mitigate these stressors so that the physical and mental health of physicians can be enhanced. This, consequently, will help us to improve the quality and efficiency of healthcare delivery in hospital settings.

Abbreviations

CVI: Content Validity Index;

RMR: Root of Mean Square Residual; GFI= Good of Fitness Index; CFI= Comparative Fit Index; TLI=Tucker-Lewis index; CR= Critical Ratio; CV=Coefficient of Variation;

NASA-TLX=NASA- Task Load Index; SWAT= Subjective Workload Assessment Technique;

EMRs= Electronic Medical Records

Funding Statement: This work was supported by the National Natural Science Foundation of China (grant number 71774062). The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Acknowledgment We would like to thank all the participants involved in this research for their time and contributions.

Author Contributors: Yinhuan Hu and Qiang Fu designed the study; Chuntao Lu and Jinzhu Xie made formal analysis; Yinhuan Hu obtained funding; Jinzhu Xie and Lu Deng took part in the investigation; Liuming Wang, Chuntao Lu and Lu Deng were involved in data cleaning; Chuntao Lu wrote the original draft; Samuel Governor, Qiang Fu and Chao Li contributed to the interpretation of the results; Chao Li and Chuntao Lu made

critical revision of the manuscript; All authors have read and approved the final manuscript.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that we have no competing interests.

Patient consent: Not required.

Ethics approval: The Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (IORG No. IORG0003571) gave the final approval for the study.

Data sharing: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Reference

1. Győrffy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians[J]. *Hum resour health*, 2016, 14: 12.
2. Lu L, Gao M, Yang T, et al. The updated Physician's Pledge and Chinese junior physicians. *The Lancet.* 2018;391:305.
3. Karahan A, Kav S, Abbasoglu A, et al. Low back pain: prevalence and associated risk factors among hospital staff. *J Adv Nurs.* 2009;65: 516-524.
4. Smith-Miller CA, Shaw-Kokot J, Curro B, et al. An integrative review: fatigue among nurses in acute care settings. *J Nurs Adm.* 2014;44:487-494.
5. Arvidsson I, Simonsen JG, Balogh I. Discrepancies in pain presentation caused by adverse psychosocial conditions as compared to pain due to high physical workload? *Work*. 2012;41:2472-2475.
6. Manty M, Moller A, Nilsson C, et al. Association of physical workload and leisure

time physical activity with incident mobility limitations: a follow-up study. *Occup Environ Med*. 2014; 71:543-548.

7. Ridley C. Relating nursing workload to quality of care in child and adolescent mental health inpatient services. *Int J Health Care Qual Assur*.2007;20:429-440.

8. Michtalik HJ, Pronovost PJ, Marsteller JA, et al. Identifying potential predictors of a safe attending physician workload: a survey of hospitalists. *J Hosp Med*.2013; 8:644-646.

9. Michtalik HJ, Yeh HC, Pronovost PJ, et al. Impact of attending physician workload on patient care a survey of hospitalists. *JAMA Intern Med*.2013; 173:375-377.

10. Garcia-Lacalle J. Balancing physician workload and patient satisfaction: a dilemma for hospital managers. *Public Money & Management*.2010;30:183-188.

11. Steyrer J, Schiffinger M, Huber C, et al. Attitude is everything? The impact of workload, safety climate, and safety tools on medical errors: a study of intensive care units. *Health Care Manage R*.2013;38:306-316.

12. Pedroja AT, Blegen MA, Abravanel R. The relationship between hospital systems load and patient harm. *J Patient Saf*. 2014;10:168-175.

13. Williams E S, Rondeau K V, Xiao Q, et al. Heavy physician workloads: impact on physician attitudes and outcomes. *Health Care Manage R*, 2007, 20: 261-269.

14. Hart SG. Theory and Measurement of Human Workload. *J Oper Res Soc*.1988;*39*:725-734.

15. Cain B. A review of the mental workload literature. Toronto: Defence Research And Development, 2007.

16. Lysaght RJ, Hill SG, Dick AO, et al. Operator workload: Comprehensive review and evaluation of operator workload methodologies. Analytics inc willow grove pa, 1989.

17 Theorell T, Hammarström A, Aronsson G, et al. A systematic review including meta-analysis of work environment and depressive symptoms[J]. *BMC public health*, 2015, 15(1): 738.

18 Noseworthy J, Madara J, Cosgrove D, et al. Physician burnout is a public health crisis: a message to our fellow health care CEOs[J]. *Health Affairs Blog*, 2017.

19 Olson K, Sinsky C, Rinne S T, et al. Cross‐sectional survey of workplace stressors associated with physician burnout measured by the Mini‐Z and the Maslach Burnout Inventory[J]. *Stress and Health*, 2019, 35(2): 157-175.

20. Weigl M, Müller A, Holland S, et al. Work conditions, mental workload and patient care quality: a multisource study in the emergency department. *BMJ Qual Saf.* 2016; 25: 499-508.

21. European Commission. European Pact for health and mental health. http://ec.europa.eu/health/mental_health/policy_es. (Accessed on 24th Sep 2018)

22. Bertram DA, Opila DA, Brown JL, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*.1992;30: 95-104.

23. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol*. 1998;*52*:139-183.

24. Rubio S, Díaz E, Martín J, et al. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Appl Psychol.* 2004; 53: 61-86.

25. Hill SG, Iavecchia HP, Byers JC. Comparison of four subjective workload rating scales. *Hum Factors*.1992;*34*:429.

26. Rubio-Valdehita S, Lopez-Nunez MI, Lopez-Higes R, et al. Development of the CarMen-Q Questionnaire for mental workload assessment. *Psicothema*. 2017; 29:570-576.

27. Colligan L, Potts HW, Finn CT, et al. Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *Int J Med Inform.*2015; 84:469-476.

28. Hoonakker P, Carayon P, Gurses A, et al. Measuring Workload Of ICU Nurses with a Questionnaire Survey: The NASA Task Load Index (TLX). *IIE Trans Healthc Syst Eng*. 2011;1:131-143.

29. Malekpour F, Malekpour A R, Mohammadian Y, et al. Assessment of mental workload in nursing by using NASA-TLX. *The Journal of Urmia Nursing and*

Midwifery Faculty.2014; 11: 892-899.

30. Reid GB, Nygren TE. The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Adv Psychol*.1988; *52*.
31. The COPSOQ III questionnaire. https://www.copsoq-network.org/assets/Uploads/annex1-Dimensions-and-items-in-the-COPSOQ-III-questionnaire-060718.pdf (Accessed on 14th June 2019)
32. Zhou H, Han X, Zhang J, et al. Job satisfaction and associated factors among medical staff in tertiary public hospitals: Results from a national cross-sectional survey in China. *Int J Occup Env Heal*. 2018; 15: 1528.
33. Sun J, Ma J, Hu G, et al. Welfare, wellness, and job satisfaction of Chinese physicians: A national survey of public tertiary hospitals in China. *Int J Health Plan M*. 2017; 32: 270-284.
34. The Lancet, Ending violence against doctors in China. *The Lancet*. 2012; 379:1764.
35. Chinese Hospital Association. Chinese patient safety goal 2014-2015. http://www.cha.org.cn/plus/view.php?aid=14626 (Accessed on 24th Sep 2018)
36. Power M, Quinn K, Schmidt S, et al. Development of the WHOQOL-old module. *Qual Life Res*. 2005; 14 (10):2197-2214.
37. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York: Guilford publications, 2015:69-70.
38. Ford JK, MacCullum RC, Tait M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Pers Psychol*. 1986; 39:291-314.
39. Lynn MR. Determination and quantification of content validity. *Nursing Res*. 1986; 35:382-385.
40. DiIorio C K. Measurement in health behavior: methods for research and evaluation. John Wiley & Sons, 2006.
41. Tabachnick B, Fidell LS. Using Multivariate Statistics. 6th ed., Pearson, Boston, 2013: 720-726.
42. DiDomenico A, Nussbaum MA. Effects of different physical workload parameters on mental workload and performance. *Int J Ind Ergonom* 2011. 41;255-260.
43. Tang L, Guan M. Rise of Health Consumerism in China and Its Effects on Physicians'

Professional Identity and the Physician-Patient Relationship and Communication. *Health commun*.2018; 33:636-642.

44. Yueju L. Violence against doctors in China. *The Lancet*. 2014; 384:745.
45. Sun T, Gao L, Shi Y. Workplace violence, psychological stress, sleep quality and subjective health in Chinese doctors: a large cross-sectional study. *BMJ OPEN*. 2017;1-8.
46. Wen J, Cheng Y, Hu X. Workload, burnout, and medical mistakes among physicians in China: A cross-sectional study. *Biosci trends*. 2016; 10:27-33.
47. Chinese Medical Doctor Association. White Paper on the Practice of Medi cine by Chinese Physicians. http://www.cmda.net/u/cms/www/201807/06181247ffex.pdf. (Accessed on 14th June 2019)
48. Horner RD, Szaflarsk JP, Ying J. Physician Work Intensity Among Medical Specialties: Emerging Evidence on its Magnitude and Composition. *Med Care*. 2011; 49:1007-1011.
49. Charles RL, Nixon J. Measuring mental workload using physiological measures: A systematic review[J]. Appl ergon. 2019; 74: 221-232.
50. Mosaly PR, Mazur LM, Yu F. Relating Task Demand, Mental Effort and Task Difficulty with Physicians' Performance during Interactions with Electronic Health Records (EHRs). *Int J Hum-Comput Int.* 2017; 34:467-475.
51. Colle HA, Reid GB. Context effects in subjective mental workload ratings. Hum factors. 1998; 40: 591-600.

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Methods			

Study design	#4	Present key elements of study design early in the paper	5
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	7
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-10
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	n/a
Bias	#9	Describe any efforts to address potential sources of bias	17
Study size	#10	Explain how the study size was arrived at	7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	n/a
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	9-10
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	n/a
Statistical methods	#12c	Explain how missing data were addressed	7
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	9-10
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed	6-7

		eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	
Participants	#13b	Give reasons for non-participation at each stage	n/a
Participants	#13c	Consider use of a flow diagram	n/a
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	11
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	7
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	12-15
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	15-16
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	17-18
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	17

Generalisability	#21	Discuss the generalisability (external validity) of the study results	18
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	19

BMJ Open

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-030137.R3
Article Type:	Original research
Date Submitted by the Author:	06-Sep-2019
Complete List of Authors:	Lu, Chuntao; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management; Jingchu University of Technology Affiliated Central Hospital Hu, Yinhuan; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Fu, Qiang; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Governor, Samuel ; Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology and Biostatistics Wang, Liuming; Huazhong University of Science and Technology Tongji Medical College, Tongji Hospital Li, Chao; Jingchu University of Technology Affiliated Central Hospital Deng, Lu; Huazhong University of Science and Technology Tongji Medical College, School of Medicine and Health Management Xie, Jinzhu; The Third People's Hospital of Hubei Province
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Mental health
Keywords:	Physician, MENTAL HEALTH, Workload, Survey and Questionnaires, Hospitals, Public

Title page

Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Chuntao Lu Yinhuan Hu* Qiang Fu Samuel Governor Liuming Wang Chao Li Lu Deng Jinzhu Xie

Corresponding author:

*Yinhuan Hu

Address: No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Email: hyh288@hotmail.com

Telephone number: 13554285879

Co-author:

Chuntao Lu: 1.School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China.

2. Jingchu University of Technology Affiliated Central Hospital, No.39 Xiangshan Rd., Jingmen, Hubei 448126, P.R. China

Yinhuan Hu: School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Qiang Fu: Department of Epidemiology and Biostatistics, College for Public Health

and Social Justice, Saint Louis University, Missouri, MO, United States

Samuel Governor: Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Missouri, MO, United States

Liuming Wang: Tongji Hospital, Tongji Medical College of Huazhong University of Science and Technology, No.1095 Jiefang Rd., Wuhan, Hubei 430030, P.R. China

Chao Li: Jingchu University of Technology Affiliated Central Hospital. No.39 Xiangshan Rd., Jingmen, Hubei 448126, P.R. China

Lu Deng： School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, No.13 Hangkong Rd., Wuhan, Hubei 430030, P.R. China

Jinzhu Xie: The Third People's Hospital of Hubei Province, No.26 Zhongshan Rd., Wuhan, Hubei 430030, P.R. China.

Abstract

Objective: The purpose of our study is to develop a mental workload scale for physicians in China and assess the scale's reliability and validity.

Design: The instrument was developed over three phases involving 396 physicians from different tiers of comprehensive public hospitals in China. In the first phase, an initial item pool was developed through a systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey. The third phase tested the reliability and validity of the instrument.

Setting: Public hospitals in China.

Participants: A total of 396 physicians from different tiers of comprehensive public hospitals in China participated in this study in 2018.

Primary and secondary outcome measures: Cronbach's alpha, content validity index, the correlation coefficient between the items and dimensions, and indices of confirmatory factor analysis**.**

Results: Six dimensions (mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance) and twelve items were identified in the instrument. For reliability, Cronbach's α for the whole scale was 0.81. For validity, the corrected item-content validity index of each item ranged from 0.85 to 1, and the correlation coefficients between the dimensions and total scores ranged from 0.37-0.72. The results of the confirmatory factor analysis showed that the goodness-of-fit indices of the scale were satisfactory.

Conclusion: The instrument showed good reliability and validity, and it is useful for

diagnosing the mental workload of physicians.

Keywords: Physician; Mental Health; Workload; Survey and Questionnaires; Hospitals, Public

Article summary

Strengths and limitations of this study

This is the first study to develop a measurement of physician mental workload from a subjective perspective in China.

Qualitative and quantitative methods were involved in item selection.

There was a potential reporting bias in the self-reported measurements of physician workload.

There was a selection bias due to all respondents voluntarily rather than randomly participating in the survey.

Among the six dimensions, perceived risk included only one item, which may have resulted in measurement error.

Introduction

Internationally, there has been a focus on the relationship between physicians' workload and their health. [1] Physicians' health is highly associated with their workload. [2] Excessive workload impacts physicians' health [3-4] and increases the risk of work-related musculoskeletal disorders (MSDs). [5-6] High workload is related to adverse effects in the form of medical errors [7] and adverse incidents. [8] Physician workload can negatively contribute to patients' perceived quality of care [9] and affect patient satisfaction [10] and safety. [11-12] It is possible that these stressors have reached a point where they pose a severe problem for the entire healthcare system, [13] as physicians' unreasonable and overwhelming workload has adverse effects on physicians, patients, and healthcare organizations. [14]

Workload is thought to be multidimensional and multifaceted. [15] One aspect of workload includes the subjective psychological experiences of the human operator. [16] Mental workload has emerged as one of the most critical occupational risk factors that results in burnout or anxiety. [17] A lack of control over workload is expected to correlate closely with burnout. [18,19] Heavy mental workload can lead to serious health problems (cardiovascular diseases, digestive problems, etc.) for physicians[17] and an inferior quality of care service. [20] Currently, The European Pact for Mental Health and Welfare is conducting mental workload assessments to promote physical and mental wellbeing. [21]

Different tools have been proposed to assess mental workload. Previous research established a brief instrument with six items to measure physician mental

workload. [22] The NASA-Task Load Index (NASA-TLX) scale, which is widely used in measuring mental workload [23], has proven to be a sensitive, valid, and reliable instrument [24] and can be used in human factor research. [25] Researcher has localized it as a 29-item questionnaire in Spain to measure workers' mental workload. [26] The existing body of research on NASA-TLX suggests that it can be used to measure nurse workload in health care settings. [27-29] In the same vein, the Subjective Workload Assessment Technique (SWAT) is a subjective rating technique with three dimensions—time load, mental effort load, and psychological stress load—and is used to assess mental workload. [30] It has been successfully applied in assessing the mental workload of several aircraft multitasking conditions, such as in assessing the mental workload required by different systems of air defence. [24] The Copenhagen Psychosocial Questionnaire is a widespread tool used in the industrial and service branches in Europe, and its main dimensions include the most influential psychosocial theories at work. [31] Together, these tools provide essential insights into workload measurement in health care management, especially in nurse workload measurement. However, the workload of physicians and nurses is essentially different from the workload of the workers that previous measurements were designed to assess. Thus, it remains unclear whether these tools can be directly used in measuring physician mental workload, and a mental workload measurement must be developed for physicians.

With increasing patient health demands, physicians tend to have a heavier workload, worse physical health, more mental strain, and more intense relationships

with patients in China. [32] Data from several studies suggest that most physicians work more than 10 hours a day [33] to manage outpatients and inpatients. On average, a physician in a tertiary hospital is responsible for 8.10 outpatients and 2.70 beds per day. [32] Physicians have been abused, injured, and, in extreme cases, murdered by patients or their relatives in hospitals across China, [34] which results in psychological stress. Establishing a workload measurement system for medical personnel has been incorporated into the Chinese Patient Safety Goals by the Chinese Hospital Association. [35]

Existing studies on workload measurement instruments are concentrated on assessing objective workload in China, for example, measuring work time. While physicians' mental workload is a critical problem, there are few instruments exploring this problem in China. The purpose of this paper is to develop a scientific mental workload instrument that can be used to assess the mental workload of physicians.

Methods

Study design

The instrument was developed in three phases. In the first phase, an initial item pool was developed by integrating previous studies through a systematic literature review. The second phase consisted of two rounds of Delphi expert consultations and a pilot survey in 2017. The third phase involved testing the psychometric

properties of the instrument, including its reliability and validity, through a study conducted in 2018 in comprehensive public hospitals in China.

Framework and item generation and selection

We combined the dimensions of the NASA-TLX and SWAT frameworks to determine the item pool so that it would measure the current situation of Chinese physicans' workload. Six dimensions and fifteen items were sent to 20 experts (including physicians, hospital managers, researchers, and human resource managers) for consultation. In accordance with the findings from two rounds of expert consultation, we deleted four items, added a new item (the intensity of physical activities), and revised the descriptions of all items. Then, there were six dimensions (physical demands, mental demands, temporal demands, effort, frustration level, and performance) and twelve items, which consisted of a pre-scale ranging from 0 to 100.

In the pre-survey analysis, we selected three hospitals (one tertiary hospital, one secondary hospital, and one first-tier hospital) through convenience sampling. A sample of 80 physicians was surveyed with a web-based scale during November and December 2017. Finally, a valid sample of 74 physicians was used for item selection. Items were refined based on the following indexes or methods: critical ratio (CR), coefficient of variation (CV), correlation analysis, [36] Cronbach's α, [37] and exploratory factor analysis (EFA). [38]

If an item was eliminated by any of the above methods, then the item was

deleted or revised. The final scale consisted of six dimensions (mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance) and twelve items (Table 1).

Data collection for testing the validity and reliability of the scale

To check the validity and reliability of the developed scale, we planned to survey 400 respondents (physicians working in hospitals) from different tiers of hospitals (two tertiary hospitals, two secondary hospitals, and two first-tier hospitals). These hospitals were randomly selected from Hubei province, China. We used wenjuanxing (www.wjx.cn), a widely used website for conducting surveys in China, to develop an electronic questionnaire with which to survey physicians. Respondents could scan the access code or click on the website using their phones to access and complete the electronic questionnaire. We sent the access code and website to the human resource managers at each participating hospital, who then sent the access code to the physicians' online communication group at each hospital. Three hundred ninety-six physicians voluntarily participated in the survey before March 2018; eleven invalid samples were deleted.

The detailed scale instructions indicated that our scale was anonymous, that participation was voluntary, and that our survey aimed to develop a physician mental workload scale, so the results would not be used for other purposes. The physician mental workload scale included three parts. The first part of the scale included twelve items that respondents scored one by one. The second part was a

table that included 15 pairs of dimensions and was used to collect the weights of each dimension. Every two dimensions formed a pair (e.g., Mental demands vs. Physical demands, Mental demands vs. Temporal demands, etc.). Respondents chose which of the two dimensions in each of the fifteen pairs contributed more to their workload. Then, the weight of each dimension was equal to the number of times that dimension was selected divided by 15. The third part of the scale was designed to collect physicians' individual characteristics.

The response endpoints of the items are displayed in Table 1. Items were scored as follows: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100. The average scores of all items for a corresponding dimension were multiplied by the dimension weight to produce the dimension scores, and then, the total scores were calculated as the sum of all dimension scores.

Table 1 Dimensions and items of physician mental workload scale

Dimensions and item	Endpoints (0-100)
A Mental demands	
A1 How demanding are the cognitive activities (e.g., sensation, perception, remembering, thinking, calculating, attention) required during your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
A2 How demanding are emotion and feeling (e.g., empathy, sympathy, enthusiasm, negative emotion restraining) during your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
A3 How hard do you have to work to overcome difficulties in accomplishing your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
B Physical demands	
B1 How demanding are the physical activities required (e.g., standing, stationary, controlling, repetitive action) in your medical work?	no workload ... heaviest workload 0 10 20 30 40 50 60 70 80 90 100
B2 How intensive is the physical activity during	no workload ... heaviest workload

your medical work? (Is the work restful or laborious?)	0 10 20 30 40 50 60 70 80 90 100
C Temporal demands	
C1 How much pressure do you feel related to work time in your medical work? (Daily medical work required time is more/less than available time.)	no workload — heaviest workload 0 10 20 30 40 50 60 70 80 90 100
C2 How frequently do you have to complete multiple tasks at the same time (work overlap) in your medical work?	no workload — heaviest workload 0 10 20 30 40 50 60 70 80 90 100
D Perceived risk	
D1 How risky do you perceive (e.g., medical disputes) your medical work to be?	no workload — heaviest workload 0 10 20 30 40 50 60 70 80 90 100
E Frustration level	
E1 How depressed or frustrated do you feel in your medical work?	no workload — heaviest workload 0 10 20 30 40 50 60 70 80 90 100
E2 How anxious or irritated do you feel in your medical work?	no workload — heaviest workload 0 10 20 30 40 50 60 70 80 90 100
F Performance	
F1 How successful do you think you are in accomplishing the goals in your medical work?	heaviest workload — no workload 0 10 20 30 40 50 60 70 80 90 100
F2 How satisfied are you with the outcomes of your medical work?	heaviest workload — no workload 0 10 20 30 40 50 60 70 80 90 100

Statistical analysis

Descriptive statistics are used to show the characteristics of the respondents, including gender, age, and educational level (i.e., Ph.D. degree, master's degree, undergraduate), job title (i.e., senior, middle, junior), work years, hospital level (i.e., tertiary hospitals, secondary hospitals, first-tier hospital), work hours per week, number of outpatients serviced per day, and self-perceived health status.

For the reliability of the scale, Cronbach's alpha was used to assess the internal

consistency of each instrument component. Values of 0.70 or higher for Cronbach's alpha were considered acceptable. [37]

The content validity index (CVI) of each item was calculated to assess the accuracy of the scale using scores of 1-4. Experts were invited to evaluate the items, with a score of 1 representing an item not relevant to the corresponding dimension and a score of 4 representing an item closely related to the corresponding dimension. The corrected item-content validity index (I-CVI) and average scale-content validity index (S-CVI/Ave) were calculated. A corrected I-CVI of 0.78 or higher and an S-CVI/Ave of 0.90 or greater were considered acceptable. [39]

The test of construct validity was performed using the correlation coefficient method, EFA, and confirmatory factor analysis (CFA). Items with item-total correlation and dimension-total correlation values below 0.40 were revised or removed from the scale. Bartlett's test of sphericity scores lower than 0.05 and a Kaiser-Meyer–Olkin (KMO) score of sampling adequacy higher than 0.70 and close to 1 were considered appropriate for factor analysis. [40] EFA and CFA were used to explore and confirm the structure of the scale. For the EFA, we used the varimax rotation method to examine whether the structure matched the framework. For the CFA, the criteria for the model fit indices were as follows: $\chi 2/df<3$; root mean square error of approximation (RMSEA) < 0.05; root mean square residual (RMR) <0.05; goodness-of-fit index (GFI) >0.90, comparative fit index (CFI) >0.90, and Tucker-Lewis index (TLI) >0.90. [41] Statistical analyses were performed with SPSS V. 21 (IBM Corp., Armonk, NY, USA) and AMOS V.17 (IBM Corp., Armonk, NY, USA).

Patient and public involvement

Our participants were physicians working in hospitals. They took part in the pre-survey and formal survey to complete our scale. Participation was voluntary, and no incentives were provided for participation. Participants were not directly involved in the design or recruitment of this study. The results were not provided to participants.

Results

Sample characteristics

Three hundred ninety-six responses (online survey) were received, and 11 were excluded due to incomplete demographic information. There were no issues related to floor or ceiling effects as the questions for every item were responded to in the form of a web-based survey. The characteristics of the participants are presented in Table 2.

Table 2 Respondents' characteristics

Variable	Valid sample	Percentage（%）
Gender		
Male	200	51.9
Female	185	48.1
Age		
<45	258	67.0
45-55	121	31.4
>55	6	1.6
Educational level		
Ph.D. degree	36	9.4
Master's degree	48	12.5

Variable	Valid sample	Percentage（%）
Bachelor's degree	184	47.8
Below bachelor's college	117	30.3
Job title		
Senior title	83	21.6
Middle title	146	37.9
Junior title	156	40.5
Work years in current institution (year)		
1-5	122	31.7
6-10	91	23.6
11-15	58	15.1
>=16	114	29.6
Hospital level		
Tertiary hospital	130	33.8
Secondary hospital	124	32.2
First-tier hospital	131	34.0
Work hours per week		
<=40	73	19.0
41-60	152	39.5
>60	160	41.5
Number of outpatients serviced per day		
<20	135	35.1
20-50	171	44.4
>50	79	20.5
Self-perceived health status		
Poor	65	16.9
Fair	242	62.9
Good	78	20.2

Reliability of physician mental workload scale

Each of the six components demonstrated at least satisfactory internal consistency (higher than 0.70), with Cronbach's α in the range of 0.70-0.90. The Cronbach's α for the whole scale reached as high as 0.81, which indicated that the scale had excellent reliability.

Validity of physician mental workload scale

The corrected I-CVI of each item ranged from 0.85-1 (Table 3), which was higher than 0.78. The S-CVI/Ave was 0.96, which was higher than 0.90. All of these values supported the good content validity of the scale.

The correlation matrix between items and total scores was inspected to confirm the convergent validity, which was indicated by reasonable coefficients of 0.40 and above, except for F1 and F2 (Table 3). The calculated correlation coefficients between dimensions and the total scores had a range of 0.37-0.72, showing that the dimensions and total scores had good convergent validity. Additionally, the correlation coefficients among the dimensions were lower than the correlation coefficients between the dimension-total scores, which indicated that the scale had good discriminant validity (Table 4).

Table 3 Content validity and correlation coefficient of item-total scores of the scale

Items	Corrected I-CVI	Item-total correlations
A1 Cognitive activity	1	0.57
A2 Emotion and feeling	0.85	0.57
A3 Overcoming difficulties	0.85	0.59
B1 Physical activity	1	0.57
B2 Intensity of physical activity	1	0.65
C1 Time pressure	1	0.69
C2 Multiple tasks	0.85	0.69
D1 Risk concern	1	0.64
E1 Depressed or frustrated	1	0.75
E2 Anxious or irritated	1	0.75
F1 Successful	1	0.33*
F2 Satisfied	1	0.31*

*Item-total scores were below 0.4.

Table 4 Correlation coefficient matrix between dimensions and total scores of the

scale

Dimensions	A	B	C	D	E	F	Total scores
A Mental demands	1						
B Physical demands	0.43	1					
C Temporal demands	0.52	0.47	1				
D Perceived risk	0.46	0.38	0.44	1			
E Frustration level	0.40	0.36	0.54	0.51	1		
F Performance	0.09	0.05	0.01	0.05	0.13	1	
Total scores	0.61	0.52	0.68	0.68	0.72	0.37	1

Exploratory factor analysis of physician mental workload scale

The KMO sample adequacy measurement was 0.81, which was higher than the recommended value of 0.70. Bartlett's test of sphericity value with the chi-square values was 1950.70 ($p<0.000$). Thus, the data were suitable for factor analysis. Considering the experts' suggestions, we selected six principal components in the EFA, and the results showed that the six-dimensional model explained 81.88% of the total variance (Table 5).

Component 1, "mental demands", was developed from 3 items that asked about feeling or memory requirements, emotional requirements, and the effort required to overcome difficulties, with a factor loading in the range of 0.74-0.81. Component 2, "frustration level", consisted of 2 items that asked about anxiety and levels of depression or frustration, and the factor loading was in the range of 0.86-0.88. Component 3, "physical demands", consisted of 2 items related to strength requirements and the intensity of work time, with a factor loading in the range of 0.84-0.90.

Component 4, “temporal demands”, constituted two items that asked about the ratio of required time to available time and the frequency of multi-tasking, with a factor loading in the range of 0.77-0.82. There were two items in “performance” component 5, which related to the sense of achievement and job satisfaction regarding work outcomes, with the factor loading in the range of 0.85-0.90. Component 6, “perceived risk”, included only 1 item that explained the perception of risk in conducting tasks (such as medical disputes and risk of being infectious), with a factor loading of 0.84.

Table 5 Factor loadings for the rotated component matrix: varimax rotated components

Items	Components					
	1	2	3	4	5	6
A2 Emotion and feeling	0.81					
A1 Cognitive activity	0.76					
A3 Overcoming difficulties	0.74					
E1 Depressed or frustrated		0.88				
E2 Anxious or irritated		0.86				
B1 Physical activity			0.90			
B2 Intensity of physical activity			0.84			
C1 Time pressure				0.82		
C2 Multiple tasks				0.77		
F1 Successful					0.90	
F2 Satisfied					0.85	
D1 Risk concern						0.84

Confirmatory factor analysis of physician workload scale

The six-factor model obtained after EFA was tested by CFA using the maximum likelihood estimation method. The goodness-of-fit model was as follows: $\chi 2/df=1.84$

(<3), RMR=0.04 (<0.05), GFI=0.97 (>0.9), CFI=0.98 (>0.9), TLI=0.97 (>0.9), and RMSEA=0.05 (≤0.05). Based on these criteria, the model was a good fit for the data.

Discussion

The purpose of this study was to develop a mental workload scale for physicians and explore its validity and reliability. The test results show that the scale is reliable and valid; hence, it is considered an effective instrument for assessing physician mental workload in Chinese comprehensive public hospitals. The results show a six-dimensional model that includes aspects related to mental demands, physical demands, temporal demands, perceived risk, frustration level, and performance. In contrast to other relevant scales, this scale includes only 12 items; thus, its length is a strength because it can be completed in a short time. As for the scale's contents, the dimensions of perceived risk and temporal demands are uniquely distinctive for physician mental workload in China.

The Cronbach's α of the whole scale was higher than 0.7, which indicated that the scale had excellent reliability. Additionally, the corrected I-CVI was higher than 0.78, and the S-CVI/Ave was more than 0.9, which showed that it had good content validity. For the construct validity, except for F1 and F2, the correlation coefficient between the item and total scores was more than 0.4, which showed that the construct validity was good. The item-total scores of the two items in the dimension of performance were near 0.4 and perhaps would have been relevant with reverse

scoring. Consistent with previous research on NASA-TLX, the performance dimension shows limited practical relevance since variations influence it in terms of physical load. [42] Another study reported that subjective assessments of mental workload might not provide an accurate estimation of the performance dimension. [26] Considering this information, we retained the two items but revised their description.

The specific dimension perceived risk, which is not included in the NASA-TLX or SWAT frameworks, is highly associated with physician mental workload in China. There tends to be an estranged relationship between physicians and patients, which puts physicians at a dangerous risk of being assaulted by patients or visitors. [43] According to statistics, 96% of medical staff were abused or injured in 2012. [44] The physician-patient relationship is becoming increasingly fragile and has reached an unprecedented poor level in China. [45] This tense relationship results in heavy psychological workload during physicians' work.

Another dimension, temporal demands, is also highly specific. The gap between healthcare demand and supply (and thus the doctor-patient ratio) in China has caused physicians in secondary and tertiary hospital settings to become overworked. [46] They frequently need to work overtime and perform more than one task at the same time. According to a report by the Chinese Medical Association in 2018, physicians in tertiary hospitals had an average workweek of 51.05 hours, which was more than the legal 40 hours per week. [47] Research has reported that physicians may feel stressed when poor scheduling leaves them pressed for time. [48] Mental

workload encompasses the subjective experience of a given task load. [49] High task demands require considerable time and mental effort and represent a heavy workload for physicians. [50] The worse physicians' experience of their task load, the higher their mental workload is. [51]

Although we have attempted an accurate examination of the measurement properties of the physician mental workload scale by using qualitative and quantitative methods, there are still some limitations that merit discussion. First, among the six dimensions, perceived risk included only one item, which may have resulted in measurement error. Second, there was a potential reporting bias in the self-reported measurements of workload among physicians. Third, all respondents voluntarily decided to take part in the survey. Physicians who were overburdened at the time of the study may not have had time to take part in the investigation, which could have resulted in selection bias. Thus, these findings reveal the need for continued research to improve this scale. Burnout is also relevant to mental workload and is another direction for further exploration.

Conclusion

Creating new items from a subjective perspective is of paramount importance in investigating Chinese physicians' workload. The physician mental workload scale has acceptable preliminary psychometric properties, with six dimensions and twelve items. The use of this scale can help us identify the main stressors in physician mental workload and implement targeted optimization strategies to mitigate these

stressors in order to enhance the physical and mental health of physicians. Doing so will consequently improve the quality and efficiency of healthcare delivery in hospital settings.

Abbreviations

CVI=Content Validity Index;

EFA= Confirmatory Factor Analysis; CFA= Exploratory Factor Analysis;

RMR=Root Mean Square Residual; GFI=Good-of-Fit Index; CFI=Comparative Fit Index; TLI=Tucker-Lewis Index; CR=Critical Ratio; CV=Coefficient of Variation;

NASA-TLX=NASA Task Load Index; SWAT=Subjective Workload Assessment Technique;

Funding Statement: This work was supported by the National Natural Science Foundation of China (grant number 71774062). The funder had no role in the study design, data collection, analysis, decision to publish the manuscript, or manuscript preparation.

Acknowledgment We would like to thank all of the participants involved in this research for their time and contributions.

Author Contributors: Yinhuan Hu and Qiang Fu designed the study; Chuntao Lu and Jinzhu Xie performed formal analysis; Yinhuan Hu obtained funding; Jinzhu Xie and Lu Deng took part in the investigation; Liuming Wang, Chuntao Lu and Lu Deng were involved in data cleaning; Chuntao Lu wrote the original draft; Samuel Governor, Qiang Fu and Chao Li contributed to the interpretation of the results; and Chao Li and Chuntao Lu performed critical revisions of the manuscript; All authors have read

and approved the final manuscript.

Competing interests: We have read and understood BMJ's policy on the declaration of interests and declare that we have no competing interests.

Patient consent: Not required.

Ethics approval: The Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (IORG No. IORG0003571), gave the final approval for the study.

Data sharing: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Reference

1. Győrffy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians[J]. *Hum resour health*, 2016, 14: 12.
2. Lu L, Gao M, Yang T, et al. The updated Physician's Pledge and Chinese junior physicians. *The Lancet.* 2018;391:305.
3. Karahan A, Kav S, Abbasoglu A, et al. Low back pain: prevalence and associated risk factors among hospital staff. *J Adv Nurs.* 2009;65: 516-524.
4. Smith-Miller CA, Shaw-Kokot J, Curro B, et al. An integrative review: fatigue among nurses in acute care settings. *J Nurs Adm.* 2014;44:487-494.
5. Arvidsson I, Simonsen JG, Balogh I. Discrepancies in pain presentation caused by adverse psychosocial conditions as compared to pain due to high physical workload? *Work*. 2012;41:2472-2475.
6. Manty M, Moller A, Nilsson C, et al. Association of physical workload and leisure time physical activity with incident mobility limitations: a follow-up study. *Occup Environ Med*. 2014; 71:543-548.

7. Ridley C. Relating nursing workload to quality of care in child and adolescent mental health inpatient services. *Int J Health Care Qual Assur*.2007;20:429-440.

8. Michtalik HJ, Pronovost PJ, Marsteller JA, et al. Identifying potential predictors of a safe attending physician workload: a survey of hospitalists. *J Hosp Med*.2013; 8:644-646.

9. Michtalik HJ, Yeh HC, Pronovost PJ, et al. Impact of attending physician workload on patient care a survey of hospitalists. *JAMA Intern Med*.2013; 173:375-377.

10. Garcia-Lacalle J. Balancing physician workload and patient satisfaction: a dilemma for hospital managers. *Public Money & Management*.2010;30:183-188.

11. Steyrer J, Schiffinger M, Huber C, et al. Attitude is everything? The impact of workload, safety climate, and safety tools on medical errors: a study of intensive care units. *Health Care Manage R*.2013;38:306-316.

12. Pedroja AT, Blegen MA, Abravanel R. The relationship between hospital systems load and patient harm. *J Patient Saf*. 2014;10:168-175.

13. Williams E S, Rondeau K V, Xiao Q, et al. Heavy physician workloads: impact on physician attitudes and outcomes. *Health Care Manage R*, 2007, 20: 261-269.

14. Hart SG. Theory and Measurement of Human Workload. *J Oper Res Soc*.1988;*39*:725-734.

15. Cain B. A review of the mental workload literature. Toronto: Defence Research And Development, 2007.

16. Lysaght RJ, Hill SG, Dick AO, et al. Operator workload: Comprehensive review and evaluation of operator workload methodologies. Analytics inc willow grove pa, 1989.

17 Theorell T, Hammarström A, Aronsson G, et al. A systematic review including meta-analysis of work environment and depressive symptoms[J]. *BMC public health*, 2015, 15(1): 738.

18 Noseworthy J, Madara J, Cosgrove D, et al. Physician burnout is a public health crisis: a message to our fellow health care CEOs[J]. *Health Affairs Blog*, 2017.

19 Olson K, Sinsky C, Rinne S T, et al. Cross-sectional survey of workplace stressors

associated with physician burnout measured by the Mini-Z and the Maslach Burnout Inventory[J]. *Stress and Health*, 2019, 35(2): 157-175.

20. Weigl M, Müller A, Holland S, et al. Work conditions, mental workload and patient care quality: a multisource study in the emergency department. *BMJ Qual Saf.* 2016; 25: 499-508.
21. European Commission. European Pact for health and mental health. http://ec.europa.eu/health/mental_health/policy_es. (Accessed on 24th Sep 2018)
22. Bertram DA, Opila DA, Brown JL, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*.1992;30: 95-104.
23. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv Psychol*. 1998;*52*:139-183.
24. Rubio S, Díaz E, Martín J, et al. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Appl Psychol.* 2004; 53: 61-86.
25. Hill SG, Iavecchia HP, Byers JC. Comparison of four subjective workload rating scales. *Hum Factors*.1992;*34*:429.
26. Rubio-Valdehita S, Lopez-Nunez MI, Lopez-Higes R, et al. Development of the CarMen-Q Questionnaire for mental workload assessment. *Psicothema*. 2017; 29:570-576.
27. Colligan L, Potts HW, Finn CT, et al. Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *Int J Med Inform.*2015; 84:469-476.
28. Hoonakker P, Carayon P, Gurses A, et al. Measuring Workload Of ICU Nurses with a Questionnaire Survey: The NASA Task Load Index (TLX). *IIE Trans Healthc Syst Eng*. 2011;1:131-143.
29. Malekpour F, Malekpour A R, Mohammadian Y, et al. Assessment of mental workload in nursing by using NASA-TLX. *The Journal of Urmia Nursing and Midwifery Faculty*.2014; 11: 892-899.

30. Reid GB, Nygren TE. The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Adv Psychol*.1988; *52*.

31. The COPSOQ III questionnaire. https://www.copsoq-network.org/assets/Uploads/annex1-Dimensions-and-items-in-the-COPSOQ-III-questionnaire-060718.pdf (Accessed on 14th June 2019)

32. Zhou H, Han X, Zhang J, et al. Job satisfaction and associated factors among medical staff in tertiary public hospitals: Results from a national cross-sectional survey in China. *Int J Occup Env Heal*. 2018; 15: 1528.

33. Sun J, Ma J, Hu G, et al. Welfare, wellness, and job satisfaction of Chinese physicians: A national survey of public tertiary hospitals in China. *Int J Health Plan M*. 2017; 32: 270-284.

34. The Lancet, Ending violence against doctors in China. *The Lancet*. 2012; 379:1764.

35. Chinese Hospital Association. Chinese patient safety goal 2014-2015. http://www.cha.org.cn/plus/view.php?aid=14626 (Accessed on 24th Sep 2018)

36. Power M, Quinn K, Schmidt S, et al. Development of the WHOQOL-old module. *Qual Life Res*. 2005; 14 (10):2197-2214.

37. Kline RB. Principles and practice of structural equation modeling. 3rd ed. New York: Guilford publications, 2015:69-70.

38. Ford JK, MacCullum RC, Tait M. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Pers Psychol*. 1986; 39:291-314.

39.Lynn MR. Determination and quantification of content validity. *Nursing Res*. 1986; 35:382-385.

40. DiIorio C K. Measurement in health behavior: methods for research and evaluation. John Wiley & Sons, 2006.

41. Tabachnick B, Fidell LS. Using Multivariate Statistics. 6th ed., Pearson, Boston, 2013: 720-726.

42. DiDomenico A, Nussbaum MA. Effects of different physical workload parameters on mental workload and performance. *Int J Ind Ergonom* 2011. 41;255-260.

43. Tang L, Guan M. Rise of Health Consumerism in China and Its Effects on

Physicians' Professional Identity and the Physician-Patient Relationship and Communication. *Health commun*.2018; 33:636-642.

44. Yueju L. Violence against doctors in China. *The Lancet*. 2014; 384:745.
45. Sun T, Gao L, Shi Y. Workplace violence, psychological stress, sleep quality and subjective health in Chinese doctors: a large cross-sectional study. *BMJ OPEN*. 2017;1-8.
46. Wen J, Cheng Y, Hu X. Workload, burnout, and medical mistakes among physicians in China: A cross-sectional study. *Biosci trends*. 2016; 10:27-33.
47. Chinese Medical Doctor Association. White Paper on the Practice of Medi cine by Chinese Physicians. http://www.cmda.net/u/cms/www/201807/06181247ffex.pdf. (Accessed on 14th June 2019)
48. Horner RD, Szaflarsk JP, Ying J. Physician Work Intensity Among Medical Specialties: Emerging Evidence on its Magnitude and Composition. *Med Care*. 2011; 49:1007-1011.
49. Charles RL, Nixon J. Measuring mental workload using physiological measures: A systematic review[J]. Appl ergon. 2019; 74: 221-232.
50. Mosaly PR, Mazur LM, Yu F. Relating Task Demand, Mental Effort and Task Difficulty with Physicians' Performance during Interactions with Electronic Health Records (EHRs). *Int J Hum-Comput Int.* 2017; 34:467-475.
51. Colle HA, Reid GB. Context effects in subjective mental workload ratings. Hum factors. 1998; 40: 591-600.

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Methods			

Study design	#4	Present key elements of study design early in the paper	5
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	7
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-10
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	n/a
Bias	#9	Describe any efforts to address potential sources of bias	17
Study size	#10	Explain how the study size was arrived at	7
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	n/a
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	9-10
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	n/a
Statistical methods	#12c	Explain how missing data were addressed	7
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	9-10
Statistical methods	#12e	Describe any sensitivity analyses	n/a
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed	6-7

		eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	
Participants	#13b	Give reasons for non-participation at each stage	n/a
Participants	#13c	Consider use of a flow diagram	n/a
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	11
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	7
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	12-15
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
Main results	#16b	Report category boundaries when continuous variables were categorized	n/a
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	#18	Summarise key results with reference to study objectives	15-16
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	17-18
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	17

Generalisability	#21	Discuss the generalisability (external validity) of the study results	18
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	19