


# BMJ Open Physician Mental Workload Scale in China: Development and Psychometric Evaluation

Chuntao Lu,<sup>1,2</sup> Yinhuan Hu ,<sup>1</sup> Qiang Fu,<sup>3</sup> Samuel Governor,<sup>3</sup> Liuming Wang,<sup>4</sup> Chao Li,<sup>2</sup> Lu Deng,<sup>1</sup> Jinzhu Xie<sup>5</sup>

**To cite:** Lu C, Hu Y, Fu Q, *et al.* Physician Mental Workload Scale in China: Development and Psychometric Evaluation. *BMJ Open* 2019;**9**:e030137. doi:10.1136/bmjopen-2019-030137

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-030137>).

Received 28 February 2019  
Revised 06 September 2019  
Accepted 26 September 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

<sup>2</sup>Jingmen NO.2 People's Hospital, Jingmen, China

<sup>3</sup>Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University, Saint Louis, Missouri, USA

<sup>4</sup>Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

<sup>5</sup>The Third People's Hospital of Hubei Province, Wuhan, China

**Correspondence to**  
Prof. Yinhuan Hu;  
[hyh288@hotmail.com](mailto:hyh288@hotmail.com)

## 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, item-total score correlation coefficient, dimension-total score correlation coefficient and indices of confirmatory factor analysis.

**Results** Six dimensions (mental demands, physical demands, temporal demands, perceived risk, frustration level and performance) and 12 items were identified in the instrument. For reliability, Cronbach's  $\alpha$  for the whole scale was 0.81. For validity, the corrected item-content validity index of each item ranged from 0.85 to 1, item-total score correlation coefficients ranged from 0.31 to 0.75, and the correlation coefficients between the dimensions and total score ranged from 0.37 to 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.

## INTRODUCTION

Internationally, there has been a focus on the relationship between physicians' workload and their health.<sup>1</sup> Physicians' health is highly associated with their workload.<sup>2</sup> Excessive workload impacts physicians' health<sup>3,4</sup> and increases the risk of work-related musculoskeletal disorders.<sup>5,6</sup> High workload is related to adverse effects in the form of medical errors<sup>7</sup> and adverse incidents.<sup>8</sup> Physician workload can negatively contribute to patients' perceived quality of care<sup>9</sup> and affect patient

## 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.

satisfaction<sup>10</sup> and safety.<sup>11,12</sup> It is possible that these stressors have reached a point where they pose a severe problem for the entire healthcare system,<sup>13</sup> as physicians' unreasonable and overwhelming workload has adverse effects on physicians, patients and healthcare organisations.<sup>14</sup>

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

Different tools have been proposed to assess mental workload. Previous research established a brief instrument with six items to measure physician mental workload.<sup>22</sup> The NASA-Task Load Index (NASA-TLX) scale,

which is widely used in measuring mental workload,<sup>23</sup> has proven to be a sensitive, valid and reliable instrument<sup>24</sup> and can be used in human factor research.<sup>25</sup> Researcher has localised it as a 29-item questionnaire in Spain to measure workers' mental workload.<sup>26</sup> The existing body of research on NASA-TLX suggests that it can be used to measure nurse workload in healthcare settings.<sup>27–29</sup> 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.<sup>30</sup> 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.<sup>24</sup> 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.<sup>31</sup> Together, these tools provide essential insights into workload measurement in healthcare management, especially in nurse workload measurement. However, the workload of physicians is essentially different from the workload of the nurses and other 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.<sup>32</sup> Data from several studies suggest that most physicians work more than 10 hours a day<sup>33</sup> 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.<sup>32</sup> Physicians have been abused, injured and, in extreme cases, murdered by patients or their relatives in hospitals across China,<sup>34</sup> 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.<sup>35</sup>

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 physicians' workload. Six dimensions and 15 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 activity) and revised the descriptions of all items. Then, there were six dimensions (physical demands, mental demands, temporal demands, perceived risk, frustration level and performance) and 12 items, which consisted of a prescale ranging from 0 to 100.

In the presurvey 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, coefficient of variation, correlation analysis,<sup>36</sup> Cronbach's  $\alpha$ <sup>37</sup> and exploratory factor analysis (EFA).<sup>38</sup>

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 12 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 and ninety-six physicians voluntarily participated in the survey before March 2018; 11 invalid samples were deleted.

The detailed scale instructions indicated that our scale was anonymous, that participation was voluntary and that

**Table 1** Dimensions and items of physician mental workload scale

Dimensions and item	Endpoints (0–100)
<b>A: mental demands</b>	
A1 How demanding are the cognitive activities (eg, sensation, perception, remembering, thinking, calculating and attention) required during your medical work?	no workload <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
A2 How demanding are emotion and feeling (eg, empathy, sympathy, enthusiasm and negative emotion restraining) during your medical work?	no workload <span style="float: right;">heaviest workload</span>  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 <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
<b>B: physical demands</b>	
B1 How demanding are the physical activities required (eg, standing, stationary, controlling and repetitive action) in your medical work?	no workload <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
B2 How intensive is the physical activity during your medical work? (Is the work restful or laborious?)	no workload <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
<b>C: temporal demands</b>	
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 <span style="float: right;">heaviest workload</span>  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 <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
<b>D: perceived risk</b>	
D1 How risky do you perceive (eg, medical disputes) your medical work to be?	no workload <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
<b>E: frustration level</b>	
E1 How depressed or frustrated do you feel in your medical work?	no workload <span style="float: right;">heaviest workload</span>  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 <span style="float: right;">heaviest workload</span>  0 10 20 30 40 50 60 70 80 90 100
<b>F: performance</b>	
F1 How successful do you think you are in accomplishing the goals in your medical work?	heaviest workload <span style="float: right;">no workload</span>  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 <span style="float: right;">no workload</span>  0 10 20 30 40 50 60 70 80 90 100

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 12 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 (eg, mental demands vs physical demands, mental demands vs temporal demands and

so on). Respondents chose which of the two dimensions in each of the 15 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.

### Statistical analysis

Descriptive statistics are used to show the characteristics of the respondents, including gender, age and educational level (ie, PhD degree, master's degree and undergraduate), job title (ie, senior, middle and junior), work years, hospital level (ie, tertiary hospitals, secondary hospitals and 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.<sup>37</sup>

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.<sup>39</sup>

The test of construct validity was performed using the correlation coefficient method, EFA and confirmatory factor analysis (CFA). Item–total score correlation coefficient, dimension–total score correlation coefficient and dimension–dimension correlation coefficient were used. Items with item–total score correlation coefficient below 0.40 should be revised or removed from the scale. The correlation coefficients among dimensions should be lower than the dimension–total score correlation coefficients. 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.<sup>40</sup> 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)  $\leq 0.05$ ; 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)  $> 0.90$ .<sup>41</sup> Statistical analyses were performed with SPSS V.21 and AMOS V.17 (IBM Corp, Armonk, New York, USA).

### Patient and public involvement

Our participants were physicians working in hospitals. They took part in the presurvey 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 and 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](#).

### Reliability of physician mental workload scale

Each of the six components demonstrated at least satisfactory internal consistency (higher than 0.70), with Cronbach's  $\alpha$  in the range of 0.70–0.90. The Cronbach's  $\alpha$  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 to 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 score 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 score had a range of 0.37–0.72, showing that the dimensions and total scores had good convergent validity as well. Additionally, the correlation coefficients among the dimensions were lower than the correlation coefficients between the dimensions and the total score, which indicated that the scale had good discriminant validity ([table 4](#)).

### 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  $\chi^2$  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 three 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 two 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 two items related to strength requirements and the intensity of physical activity, with a factor loading in the range of 0.84–0.90.



**Table 2** Respondents' characteristics

Variable	Valid sample	Percentage (%)
<b>Gender</b>		
Male	200	51.9
Female	185	48.1
<b>Age (years)</b>		
<45	258	67.0
45–55	121	31.4
>55	6	1.6
<b>Educational level</b>		
PhD degree	36	9.4
Master's degree	48	12.5
Bachelor's degree	184	47.8
Below bachelor's college	117	30.3
<b>Job title</b>		
Senior title	83	21.6
Middle title	146	37.9
Junior title	156	40.5
<b>Work years in current institution (year)</b>		
1–5	122	31.7
6–10	91	23.6
11–15	58	15.1
≥16	114	29.6
<b>Hospital level</b>		
Tertiary hospital	130	33.8
Secondary hospital	124	32.2
First-tier hospital	131	34.0
<b>Work hours per week</b>		
≤40	73	19.0
41–60	152	39.5
>60	160	41.5
<b>Number of outpatients serviced per day</b>		
<20	135	35.1
20–50	171	44.4
>50	79	20.5
<b>Self-perceived health status</b>		
Poor	65	16.9
Fair	242	62.9
Good	78	20.2

Component 4, 'temporal demands', constituted two items that asked about the ratio of required time to available time and the frequency of multitasking, 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

**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.  
I-CVI, item-content validity index.

of 0.85–0.90. Component 6, 'perceived risk', included only one 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.

#### CFA 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 ( $\leq 0.05$ ). 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  $\alpha$  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,

**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

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.<sup>42</sup> Another study reported that subjective assessments of mental workload might not provide an accurate estimation of the performance dimension.<sup>26</sup> 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.<sup>43</sup> According to statistics, 96% of medical staff were abused or injured in 2012.<sup>44</sup> The physician–patient relationship is becoming increasingly fragile and has reached an unprecedented poor level in China.<sup>45</sup> 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.<sup>46</sup> 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.<sup>47</sup> Research has reported that physicians may feel stressed when poor scheduling leaves them pressed for time.<sup>48</sup> Mental workload encompasses the subjective experience of a given task load.<sup>49</sup> High task demands require considerable time and mental effort and represent a heavy workload for physicians.<sup>50</sup> The worse physicians' experience of their task load, the higher their mental workload is.<sup>51</sup>

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.

**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

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. Meanwhile, 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 6 dimensions and 12 items. The use of this scale can help us identify the main stressors in physician mental workload and implement targeted optimisation 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.

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

**Contributors** YH and QF designed the study; ChuL and JX performed formal analysis; YH obtained funding; JX and LD took part in the investigation; LW, ChuL and LD were involved in data cleaning; ChuL wrote the original draft; SG, QF and ChaL contributed to the interpretation of the results; and ChaL and ChuL performed critical revisions of the manuscript; all authors have read and approved the final manuscript.

**Funding** This work was supported by the National Natural Science Foundation of China (grant number 71774062).

**Disclaimer** The funder had no role in the study design, data collection, analysis, decision to publish the manuscript or manuscript preparation.

**Competing interests** None declared.

**Patient consent for publication** 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.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Yinhuan Hu <http://orcid.org/0000-0001-7717-5261>

## REFERENCES

- Györfy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians. *Hum Resour Health* 2016;14:12.
- Lu L, Gao M, Yang T, et al. The updated physician's Pledge and Chinese junior physicians. *Lancet* 2018;391.
- Karahan A, Kav S, Abbasoglu A, et al. Low back pain: prevalence and associated risk factors among hospital staff. *J Adv Nurs* 2009;65:516–24.
- Smith-Miller CA, Shaw-Kokot J, Curro B, et al. An integrative review: fatigue among nurses in acute care settings. *J Nurs Adm* 2014;44:487–94.
- Arvidsson I, Simonsen JG, Balogh I, et al. Discrepancies in pain presentation caused by adverse psychosocial conditions as compared to pain due to high physical workload? *Work* 2012;41:2472–5.
- Mänty M, Möller 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;71:543–8.
- Ridley C. Relating nursing workload to quality of care in child and adolescent mental health inpatient services. *Int J Health Care Qual Assur* 2007;20:429–40.
- 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;8:644–6.
- Michtalik HJ, Yeh H-C, Pronovost PJ, et al. Impact of attending physician workload on patient care: a survey of hospitalists. *JAMA Intern Med* 2013;173:375–7.
- Garcia-Lacalle J. Balancing physician workload and patient satisfaction: a dilemma for hospital managers. *Public Money & Management* 2010;30:183–8.
- Steyrer J, Schifflinger 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;38:306–16.
- Pedroja AT, Blegen MA, Abravanel R, et al. The relationship between hospital systems load and patient harm. *J Patient Saf* 2014;10:168–75.
- Williams ES, Rondeau KV, Xiao Q, et al. Heavy physician workloads: impact on physician attitudes and outcomes. *Health Serv Manage Res* 2007;20:261–9.
- Hart SG. Theory and measurement of human workload. *J Oper Res Soc* 1988;39:725–34.
- Cain B. *A review of the mental workload literature*. Toronto: Defence Research And Development, 2007.
- Lysaght RJ, Hill SG, Dick AO, et al. Operator workload: comprehensive review and evaluation of operator workload methodologies. Analytics Inc willow grove PA 1989.
- Theorell T, Hammarström A, Aronsson G, et al. A systematic review including meta-analysis of work environment and depressive symptoms. *BMC Public Health* 2015;15:738.
- 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.
- Olson K, Sinsky C, Rinne ST, et al. Cross-sectional survey of workplace stressors associated with physician burnout measured by the Mini-Z and the Maslach burnout inventory. *Stress Health* 2019;35:157–75.
- 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.
- European Commission. European PACT for health and mental health. Available: [http://ec.europa.eu/health/mental\\_health/policy\\_es](http://ec.europa.eu/health/mental_health/policy_es) [Accessed 24th Sep 2018].
- 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.
- Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv Psychol* 1998;52:139–83.
- 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.
- Hill SG, Iavecchia HP, Byers JC, et al. Comparison of four subjective workload rating scales. *Hum Factors* 1992;34:429–39.
- Rubio-Valdehita S, López-Núñez MI, López-Higes R, et al. Development of the CarMen-Q questionnaire for mental workload assessment. *Psicothema* 2017;29:570–6.
- Colligan L, Potts HWW, 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;84:469–76.
- 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;1:131–43.
- Malekpour F, Malekpour AR, Mohammadian Y, et al. Assessment of mental workload in nursing by using NASA-TLX. *J Urmia Nurs Midwifery Fac* 2014;11:892–9.



- 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. Available: <https://www.copsoq-network.org/assets/Uploads/annex1-Dimensions-and-items-in-the-COPSOQ-III-questionnaire-060718.pdf> [Accessed 14th Jun 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 Environ Res Public Health* 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 Plann Manage* 2017;32:270–84.
- 34 The Lancet. Ending violence against doctors in China. *Lancet* 1764;2012.
- 35 Chinese Hospital Association. Chinese patient safety goal 2014–2015. Available: <http://www.cha.org.cn/plus/view.php?aid=14626> [Accessed 24th Sep 2018].
- 36 Power M, Quinn K, Schmidt S, *et al*. Development of the WHOQOL-old module. *Qual Life Res* 2005;14:2197–214.
- 37 Kline RB. *Principles and practice of structural equation modeling*. 3rd edn. 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. *Nurs Res* 1986;35:382–5.
- 40 Dilorio CK. *Measurement in health behavior: methods for research and evaluation*. John Wiley & Sons, 2006.
- 41 Tabachnick B, Fidell LS. *Using multivariate statistics*. 6th edn. Boston: Pearson, 2013: 720–6.
- 42 DiDomenico A, Nussbaum MA. Effects of different physical workload parameters on mental workload and performance. *Int J Ind Ergon* 2011;41:255–60.
- 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;33:636–42.
- 44 Yueju L. Violence against doctors in China. *Lancet* 2014;384:745.
- 45 Sun T, Gao L, Li F, *et al*. Workplace violence, psychological stress, sleep quality and subjective health in Chinese doctors: a large cross-sectional study. *BMJ Open* 2017;7:e017182.
- 46 Wen J, Cheng Y, Hu X, *et al*. 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 medicine by Chinese physicians. Available: <http://www.cmda.net/u/cms/www/201807/06181247ffex.pdf> [Accessed 14th Jun 2019].
- 48 Horner RD, Szaflarski JP, Ying J, *et al*. Physician work intensity among medical specialties: emerging evidence on its magnitude and composition. *Med Care* 2011;49:1007–11.
- 49 Charles RL, Nixon J. Measuring mental workload using physiological measures: a systematic review. *Appl Ergon* 2019;74:221–32.
- 50 Mosaly PR, Mazur LM, Yu F, *et al*. Relating Task Demand, Mental Effort and Task Difficulty with Physicians' Performance during Interactions with Electronic Health Records (EHRs). *Int J Hum Comput Interact* 2018;34:467–75.
- 51 Colle HA, Reid GB. Context effects in subjective mental workload ratings. *Hum Factors* 1998;40:591–600.