

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems of Factory Workers and Miners

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-057102
Article Type:	Original research
Date Submitted by the Author:	20-Sep-2021
Complete List of Authors:	Liu, Tao; Xinjiang Medical University, School of Public Health Lu, Yaoqin; Xinjiang Medical University, School of Public Health; Urumqi Center for Disease Control and Prevention Liu, Qi; Xinjiang Medical University, School of Public Health Yan, Huan; Xinjiang Medical University, Department of Nutrition and Food Hygiene; Xinjiang Autonomous Academy of Instrumental Analysis
Keywords:	MENTAL HEALTH, PREVENTIVE MEDICINE, PUBLIC HEALTH

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

# 1 Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems 2 of Factory Workers and Miners

3 Yaoqin Lu,<sup>1,2</sup> Qi Liu,<sup>1</sup> Huan Yan<sup>3,4</sup>, and Tao Liu<sup>1</sup>

4 <sup>1</sup>Department of Toxicology, School of Public Health, Xinjiang Medical University, Urumqi 830011,  
5 Xinjiang, China.

6 <sup>2</sup>Urumqi Center for Disease Control and Prevention, Urumqi 830026, Xinjiang, China.

7 <sup>3</sup>Department of Nutrition and Food Hygiene, College of Public Health, Xinjiang Medical University,  
8 Urumqi, Xinjiang, China.

9 <sup>4</sup>Xinjiang Engineering Technology Research Center for Green Processing of Nature Product Center,  
10 Xinjiang Autonomous Academy of Instrumental Analysis, Urumqi, Xinjiang, China.

11  
12 Correspondence should be addressed to Tao Liu; xjmult@163.com

## 13 Abstract

14 **Objective** A nomogram for predicting the risk of mental health problems was established in a population  
15 of factory workers and miners, in order to quickly calculate the probability of a worker suffering from  
16 mental health problems.

17 **Methods** A cross-sectional survey of 7,500 factory workers and miners in Urumqi was conducted by  
18 means of an electronic questionnaire using cluster sampling method. Participants were randomly  
19 assigned to the training group (70%) and the validation group (30%). Questionnaire-based survey was  
20 conducted to collect information. A least absolute shrinkage and selection operator (LASSO) regression  
21 model was used to screen the predictors related to the risk of mental health problems of the training  
22 group. Multivariate logistic regression analysis was applied to construct the prediction model. Calibration  
23 plots and receiver operating characteristic-derived area under the curve (AUC) were used for model  
24 validation. Decision curve analysis (DCA) was applied to calculate the net benefit of the screening model.

25 **Results** A total of 7,118 participants met the inclusion criteria and the data were randomly divided into  
26 a training group (n=4,955) and a validation group (n=2,163) in a ratio of 3:1. A total of 23 characteristics  
27 were included in this study and LASSO regression selected 12 characteristics such as education,  
28 professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working  
29 years, marital status, and work schedule as predictors for the construction of the nomogram. In the  
30 validation group the Brier score was 0.176, the calibration slope was 0.970 and the calibration curve of  
31 nomogram showed a good fit. The AUC of training group and verification group were 0.785 and 0.784  
32 respectively.

33 **Conclusion** The nomogram combining these 12 characteristics can be used to predict the risk of suffering  
34 mental health problems, providing a useful tool for quickly and accurately screening the risk of mental  
35 health problems.

36  
37 **Key words** Mental health; Predictor; Nomogram; Risk; Factory workers and miners

### 39 **Strengths and limitations of this study**

- 40 1. To our knowledge, this was a new model to develop and assess the likelihood of mental health
- 41 problems in a group of factory workers and miners.
- 42 2. This study provided a viable and easy-to-apply tool including factors that were closely related to
- 43 factory workers and miners for identifying workers at risk of mental health problems.
- 44 3. The results of this study showed good agreement and good discrimination between predictions and
- 45 observations.
- 46 4. We had considered many influential factors including demographics, job burnout, occupational stress
- 47 and occupational exposure factors, but we were still not certain whether all possible influences were
- 48 covered.
- 49 5. While the robustness of our nomogram was extensively validated internally in the same population,
- 50 external validation was lacking for other populations in other regions and countries.

51

### 52 **1. Introduction**

53

54 The World Health Organization (WHO) defines health as a state of complete physical, mental and social  
55 well-being and not merely the absence of disease or weakness <sup>[1]</sup>. Obviously, health is an organic unity  
56 of physical and mental well-being. People with good mental health are the precondition for the normal  
57 operation of our society. However, with the acceleration of people's pace of life, people are facing an  
58 increasing risk of poor health, which has become a global public health problem <sup>[2]</sup>. Mental health  
59 problems can not only take a toll on physical health such as increasing the risk of communicable and  
60 non-communicable diseases and even causing unintentional or intentional harm to others <sup>[3]</sup>, but can also  
61 have a negative impact on the economy. For example, mental health disorders represent a growing part  
62 of the global burden of disease <sup>[4]</sup>, with statistics showing that nearly one billion people worldwide  
63 currently suffer from a mental disorder, and mental illness is ranked as one of the leading causes of the  
64 global burden of disease <sup>[5]</sup>. Moreover, one study has estimated that due to the impact of mental illness,  
65 the global economy loses US \$1 trillion every year <sup>[6]</sup>.

66

67 As researchers around the world have delved into the field of mental health, factors such as gender,  
68 income levels, environment and education have been found to be associated with people's mental health  
69 problems <sup>[7-10]</sup>. Moreover, employment is also strongly associated with quality of life, higher self-esteem  
70 and fewer psychiatric symptoms <sup>[11]</sup>. In addition, in the context of the global challenges of climate change,  
71 an increasing number of scholars have been examining the epidemiological links between mental health  
72 and environmental factors. Some studies have suggested that mental health may be influenced by ambient  
73 temperature, and an association has been found between environmental pollutants, particularly fine  
74 particulate matter, and mental health problems <sup>[12]</sup>. A relevant study shows that with short-term exposure  
75 to ambient air pollution is associated with increased emergency room visits due to depression or suicide  
76 attempts <sup>[13]</sup>. Furthermore, other factors associated with mental health include sleep, diabetes, coronary  
77 artery disease and cardiovascular disease <sup>[14-15]</sup>. It is worth noting that job burnout and occupational stress  
78 are closely linked to mental health. Job burnout is an exhaustion state of physical and psychological that

1  
2  
3 79 often occurs in the work environment, and has a high correlation with depression. A large study of  
4 80 physicians found that of the 10.3% who met criteria for a major depressive episode, 50.7% were also  
5 81 affected by symptoms of burnout (OR 2.99) and indicated that worsening depression leads to a higher  
6 82 likelihood of burnout symptoms <sup>[16]</sup>. Occupational stress refers to a work environment where non-  
7 83 reciprocity of effort and reward may lead to strong negative emotions and distress. Related research has  
8 84 shown that the combination of high effort and low reward and over-commitment increases the risk of  
9 85 mental health problems such as depression <sup>[17]</sup>. Apparently, it is necessary to include the CMBI and ERI  
10 86 in this study to predict the risk of mental health problems among factory workers and miners. In addition,  
11 87 the CMBI and ERI questionnaires consist of 15 and 23 items respectively, which are a smaller number  
12 88 of items compared to the 90 items of the Symptom Check list-90 (SCL-90) questionnaire. However, there  
13 89 are few studies that include these influences in a more comprehensive way in the practice of detecting  
14 90 mental health. Therefore, more accurate identification of mental health problems in populations requires  
15 91 a questionnaire that include a wider range of factors affecting factory workers and miners' mental health  
16 92 problems.

17 93  
18 94 Factory workers and miners are a special group of workers with a relatively low overall level of education  
19 95 and are highly prone to suffering from mental health problems due to limited social support, excessive  
20 96 workload and irregular lifestyles, as well as occupational hazards such as noise and coal dust that they  
21 97 inevitably need to face in their working environment <sup>[18-19]</sup>. China has the world's largest group of factory  
22 98 workers and miners, about 6 million <sup>[20]</sup>, who are regularly involved in occupational hazards. Mental  
23 99 health problems which need to require a long process are known to be a syndrome caused by chronic  
24 100 stress. Factory workers and miners, represented by those engaged in coal mining, have a mental burden  
25 101 rating of 8.3, one of the highest mental burdens among 150 occupations <sup>[21]</sup>. This explains the high level  
26 102 of mental health problems among mine workers in previous studies, making the identification and  
27 103 treatment of mental health problems even more important. Therefore, it is essential to provide a viable  
28 104 and easy-to-apply tool for identifying workers at risk of mental health problems and thus for timely  
29 105 interventions.

30 106  
31 107 The Symptom Checklist-90 (SCL-90), which is widely used in psychiatric outpatient examinations, has  
32 108 a high degree of validity in evaluating various mental health surveys <sup>[22-23]</sup>. However, this questionnaire  
33 109 has 90 items, and in practice of our previous studies, it has been found to be complex and time-consuming  
34 110 to complete, requiring a high degree of patience and cooperation from the respondents. In addition, the  
35 111 questionnaire is slightly less targeted to the group of factory workers and miners, and lacks entries  
36 112 relating to the particular working environment of factory workers and miners. Nowadays, there is  
37 113 growing recognition that mental health plays an important role in achieving global development goals  
38 114 and the WHO has included mental health in the Sustainable Development Goals. However, there are  
39 115 currently no relevant studies that use objective indicators to form a nomogram for predicting mental  
40 116 health. Therefore, the aim of our study is to develop and validate an easy-to-use nomogram that combines  
41 117 objective information on the demographics, job burnout, occupational stress and occupational hazards to  
42 118 comprehensively and accurately predict the prevalence of mental health problems among factory workers

1  
2  
3 119 and miners.  
4  
5 120

## 6 121 **2. Materials and Methods**

### 7 122 8 9 123 **2.1. Participants**

10 124  
11 125 Participants in this cross-sectional survey were workers from factories and mining enterprises in the  
12 126 Urumqi region, who were recruited using a whole-group sampling method. A total of 7,500 participants  
13 127 in the Urumqi were surveyed from January to May 2019, covering all districts and counties in the Urumqi  
14 128 region, including Tianshan District, Shaibak District, Xinshi District, Shuimogou District, Toutunhe  
15 129 District, Dabancheng District, Middong District and Urumqi County.  
16  
17  
18  
19 130

20 131 The exclusion criteria were the following: (I) factory workers and miners in non-Urumqi area, (II)  
21 132 working history of factories and mining enterprises less than 1 year, (III) a confirmed diagnosis of a  
22 133 mental health problem and a history of treatment and use of psychotropic medication. Questionnaires  
23 134 with missing data were also excluded from the analysis based on discussion and agreement among the  
24 135 subject members. A total of 7,500 questionnaires were distributed and 7,315 questionnaires were returned,  
25 136 representing a return rate of 97.5%. After checking the validity and integrity of the questionnaires, 7,118  
26 137 questionnaires were confirmed as valid, with an effective rate of 97.3%. All participants understood the  
27 138 purpose of the study and voluntarily participated in the study.  
28  
29  
30  
31  
32 139

### 33 140 **2.2. Research Methods**

#### 34 141 35 142 **2.2.1. Assessment of mental health**

36 143  
37 144 The SCL-90, designed by Derogatis and his colleagues, was widely used in the mental health field [24],  
38 145 which contains 90 items across nine dimensions: somatization, obsessive-compulsive symptoms,  
39 146 interpersonal sensitivity, depression, anxiety, hostility, horror, bigotry and mental illness. The SCL-90  
40 147 has been used extensively in previous studies and has relatively high reliability and validity [25]. The  
41 148 questionnaire uses a Likert 5-point scale, with a score of 0 point indicating none and 4 points indicating  
42 149 severe. A total score above 160, a score above 2 on any item, or a positive item above 43 indicates the  
43 150 presence of a psychological abnormality. In this survey, Cronbach  $\alpha$  was 0.99, the half-reliability  
44 151 coefficient was 0.98, and the KMO was 0.994.  
45  
46  
47  
48  
49  
50  
51 152

#### 52 153 **2.2.2. Assessment of occupational stress**

53 154  
54 155 This survey evaluated occupational stress in factory workers and miners through the Effort–Reward  
55 156 Imbalance (ERI) model developed by Siegrist [26]. The ERI scale consists of three subscales: effort (E, 6  
56 157 items), reward (R, 11 items) and over commitment (6 items), for a total of 23 items. A Likert 5-level  
57 158 scoring method (1, "highly disagree" to 5, "highly agree") is used to grade the items in the questionnaire  
58  
59  
60

1  
2  
3 159 with the same weight for each item. The effort–return index  $ERI = E/R \times C$ , where C is the adjustment  
4 160 coefficient, and the value is 6/11. ERI values greater than 1, equal to 1, and less than 1 correspond to  
5 161 high pay–low return, pay–return balance, and low pay–high return, respectively. Moreover, the higher  
6 162 the ERI value, the greater the occupational stress [27]. In this survey, Cronbach  $\alpha$  was 0.94, the half-  
7 163 reliability coefficient was 0.93 and the KMO was 0.956.  
8  
9 164

### 10 165 **2.2.3. Assessment of job burnout**

11 166  
12 167 In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess  
13 168 job burnout, which has good reliability and validity [28]. CMBI is composed of 15 items in three  
14 169 dimensions: emotional exhaustion (5 items), depersonalization (5 items) and reduced personal  
15 170 accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete  
16 171 compliance and 7 points indicating complete non-compliance. According to the critical value (emotional  
17 172 exhaustion  $\geq 25$ , depersonalization  $\geq 11$ , personal achievement reduction  $\geq 16$ ), the levels of occupational  
18 173 burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to  
19 174 or above the critical value), moderate (any two aspects are equal to or higher than the critical values),  
20 175 and severe (three aspects are equal to or higher than the critical values) [29]. In this survey, Cronbach  $\alpha$   
21 176 was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.  
22  
23 177

### 24 178 **2.2.4. Candidate predictors**

25 179  
26 180 Trained investigators obtained information on demographics, job burnout, occupational stress, mental  
27 181 health and occupational exposure factors through on-site face-to-face collection of an electronic version  
28 182 of the questionnaire. Covariates included in this study: 1) demographic information: gender, ethnicity,  
29 183 education level, professional title, work schedule, marital status, monthly income, age, working years,  
30 184 labor contracts, working hours per day, and working hours per week; 2) occupational exposure factors:  
31 185 coal dust, silica dust, asbestos dust, benzene, lead, noise, and brucellosis; 3) questionnaires: ERI, CMBI;  
32 186 4) chronic diseases: diabetes, hypertension.  
33  
34 187

35 188 Sex was defined as male or female; ethnicity was defined as Han and other; education level was defined  
36 189 as junior high school and below, high school, junior college or bachelor's degree or above; labor contracts  
37 190 was defined as signed or unsigned; professional title was defined as no, primary, middle or senior; work  
38 191 schedule was defined as day shift, night shift, shift or day and night shifts; marital status was defined as  
39 192 unmarried, married, divorced or widowed; monthly income (yuan) was defined as <3000, 3000~, 4000~,  
40 193 5000~, 6000~, 7000~ or 8000~; age (years) was defined as <25, 25~, 30~, 35~, 40~ or 45~; working  
41 194 years was defined as ~5, 5~, 10~, 15~, 20~, 25~ or 30~; working hours per day (hours) was defined as  
42 195  $\leq 7$  or  $> 7$ ; working days per week (days) was defined as  $\leq 5$  or  $> 5$ ; exposure to coal dust, silica dust,  
43 196 asbestos dust, benzene, lead, noise, brucellosis were all defined as yes or no; ERI was defined as yes or  
44 197 no; CMBI was defined as none, mild, moderate and severe; hypertension and diabetes were both defined  
45 198 as yes or no.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



199

200 **2.3. Statistical analysis**

201

202 Categorical variables were described as counts and percentages, and chi square test or Fisher exact test  
 203 was used to compare categorical variables between groups. 70% of participants were randomly assigned  
 204 to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute  
 205 shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were  
 206 used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which  
 207 predictive models were constructed. A nomogram for predicting was generated according to the selected  
 208 characteristics. In addition, forest plot was drawn to visually depict the P-value, OR and 95% CI for the  
 209 selected validations. Statistically significant predictors were applied to develop a prediction model for  
 210 the risk of mental health problems among factory workers and miners by introducing all selected factors  
 211 and analyzing the statistical significance levels of them. We used calibration plots and receiver operating  
 212 characteristic (ROC) curves to show the calibration and discrimination of our final model. Brier scores  
 213 for overall performance, calibration slopes were used to assess the predictable accuracy of the model.  
 214 Decision curve analysis (DCA) was applied to calculate the net benefit of the nomogram. Statistical  
 215 analysis was performed using the open-source R software Version 3.6.1 (<http://www.r-project.org>). The  
 216 significance level ( $\alpha$ ) set at 0.05.

217

218 **3. Results**

219

220 **3.1. Participant characteristics**

221

222 A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training  
 223 group (n=4,955) and a validation group (n=2,163). Over half of all participants (65.31%) were male,  
 224 57.31% of the population was over 35 years of age and 78.32% of the subjects were married, showing  
 225 that factory workers and miners are generally older and most of them have spouses. The majority of them  
 226 had completed high school (83.94%), while a smaller percentage had completed undergraduate education  
 227 (22.98%), indicating that the group of factory workers and miners as a whole was not well educated. The  
 228 total number of workers (n, %) exposed to coal dust, silica dust, asbestos dust, benzene, lead, noise and  
 229 brucellosis in the factory and mining enterprises were 377 (5), 730 (10), 981 (14), 1,981 (28), 373 (5),  
 230 4,942 (69) and 121 (2) respectively, with the total number of workers exposed to noise amounting to  
 231 4,942, or 69% of the total population surveyed. The demographic, job burnout, occupational stress and  
 232 occupational exposure factors for the training and validation groups are shown in Table 1. The results  
 233 showed that there were no significant statistical differences between the two groups of characteristic  
 234 variables, except for coal dust and CMBI, indicating that the baseline levels were largely consistent  
 235 between the two groups.

Table 1 Characteristics of the study participants

Variables	Total (n = 7118)	train (n = 4955)	test (n = 2163)	<i>p</i>
-----------	------------------	------------------	-----------------	----------

1					
2					
3	Sex, n (%)				
4					
5	Male	4649 (65)	3216 (65)	1433 (66)	0.284
6	Female	2469 (35)	1739 (35)	730 (34)	
7					
8	Ethnicity, n (%)				
9	Han	5762 (81)	3982 (80)	1780 (82)	0.061
10	Other	1356 (19)	973 (20)	383 (18)	
11					
12	Education level, n (%)				
13	Junior high school and below	1143 (16)	804 (16)	339 (16)	0.765
14	High school	1406 (20)	988 (20)	418 (19)	
15	Junior college	2933 (41)	2038 (41)	895 (41)	
16	Bachelor's degree or above	1636 (23)	1125 (23)	511 (24)	
17					
18	Professional title, n (%)				
19	None	2854 (40)	1983 (40)	871 (40)	0.923
20	Primary	1644 (23)	1149 (23)	495 (23)	
21	Middle	1618 (23)	1133 (23)	485 (22)	
22	Senior	1002 (14)	690 (14)	312 (14)	
23					
24	Work schedule, n (%)				
25	Day shift	3986 (56)	2801 (57)	1185 (55)	0.585
26	Night shift	270 (4)	187 (4)	83 (4)	
27	Shift	2058 (29)	1412 (28)	646 (30)	
28	Day and night shifts	804 (11)	555 (11)	249 (12)	
29					
30	Marital status, n (%)				
31	Unmarried	1104 (16)	762 (15)	342 (16)	0.218
32	Married	5575 (78)	3906 (79)	1669 (77)	
33	Divorced	390 (5)	255 (5)	135 (6)	
34	Widowed	49 (1)	32 (1)	17 (1)	
35					
36	Monthly income (yuan), n (%)				
37	<3000	1799 (25)	1246 (25)	553 (26)	0.966
38	3000~	2418 (34)	1682 (34)	736 (34)	
39	4000~	1600 (22)	1125 (23)	475 (22)	
40	5000~	752 (11)	520 (10)	232 (11)	
41	6000~	288 (4)	201 (4)	87 (4)	
42	7000~	148 (2)	106 (2)	42 (2)	
43	8000~	113 (2)	75 (2)	38 (2)	
44					
45	Age (years), n (%)				
46	<25	431 (6)	297 (6)	134 (6)	0.173
47	25~	786 (11)	519 (10)	267 (12)	
48	30~	956 (13)	684 (14)	272 (13)	
49	35~	866 (12)	617 (12)	249 (12)	
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					

1						
2						
3		40~	849 (12)	588 (12)	261 (12)	
4		45~	3230 (45)	2250 (45)	980 (45)	
5						
6		Working years (years), n (%)				
7		<5	1170 (16)	794 (16)	376 (17)	0.248
8		5~	1065 (15)	736 (15)	329 (15)	
9		10~	997 (14)	721 (15)	276 (13)	
10		15~	389 (5)	273 (6)	116 (5)	
11		20~	763 (11)	538 (11)	225 (10)	
12		25~	1293 (18)	878 (18)	415 (19)	
13		30~	1441 (20)	1015 (20)	426 (20)	
14						
15		Labor contracts, n (%)				
16		Signed	6641 (93)	4624 (93)	2017 (93)	0.955
17		Unsigned	477 (7)	331 (7)	146 (7)	
18						
19		Working hours per day (hours), n (%)				
20		≤7	1161 (16)	814 (16)	347 (16)	0.712
21		>7	5957 (84)	4141 (84)	1816 (84)	
22						
23		Working days per week (days), n (%)				
24		≤5	4442 (62)	3107 (63)	1335 (62)	0.446
25		>5	2676 (38)	1848 (37)	828 (38)	
26						
27		Diabetes, n (%)				
28		Yes	429 (6)	298 (6)	131 (6)	0.988
29		No	6689 (94)	4657 (94)	2032 (94)	
30						
31		Hypertension, n (%)				
32		Yes	1330 (19)	929 (19)	401 (19)	0.861
33		No	5788 (81)	4026 (81)	1762 (81)	
34						
35		Coal dust, n (%)				
36		Yes	377 (5)	244 (5)	133 (6)	0.039
37		No	6741 (95)	4711 (95)	2030 (94)	
38						
39		Silica dust, n (%)				
40		Yes	730 (10)	523 (11)	207 (10)	0.223
41		No	6388 (90)	4432 (89)	1956 (90)	
42						
43		Asbestos dust, n (%)				
44		Yes	981 (14)	691 (14)	290 (13)	0.570
45		No	6137 (86)	4264 (86)	1873 (87)	
46						
47		Benzene, n (%)				
48		Yes	1981 (28)	1360 (27)	621 (29)	0.287
49		No	5137 (72)	3595 (73)	1542 (71)	
50						
51		Lead, n (%)				
52		Yes	373 (5)	246 (5)	127 (6)	0.128
53						
54						
55						
56						
57						
58						
59						
60						

Noise, n (%)	No	6745 (95)	4709 (95)	2036 (94)	
	Yes	4942 (69)	3420 (69)	1522 (70)	0.270
Brucellosis, n (%)	No	2176 (31)	1535 (31)	641 (30)	
	Yes	121 (2)	86 (2)	35 (2)	0.800
ERI, n (%)	No	6997 (98)	4869 (98)	2128 (98)	
	Yes	3147 (44)	2173 (44)	974 (45)	0.372
CMBI, n (%)	No	3971 (56)	2782 (56)	1189 (55)	
	No	959 (13)	674 (14)	285 (13)	0.033
	Mild	2667 (37)	1813 (37)	854 (39)	
	Moderate	2900 (41)	2031 (41)	869 (40)	
	Severe	592 (8)	437 (9)	155 (7)	

236

### 237 3.2. Feature selection

238

239 The lambda was smallest at 0.01801 as seen from the lasso results when there were 12 characteristics,  
 240 which were education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working  
 241 hours per day, working years, marital status, and work schedule based on the results of the questionnaires  
 242 on demographics, occupational stress, job burnout and occupational exposure factors (Figure 1).

243

### 244 3.3. Results of logistic regression model

245

246 The 12 features obtained from the LASSO regression were incorporated into a multivariate logistic  
 247 regression model and the regression results were shown in Table 2. It was clear from the results that  
 248 education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per  
 249 day, working years, marital status, and work schedule were independent determinants of risk for mental  
 250 health problems. In addition, there was no evidence of multicollinearity between the covariates included  
 251 in the model. The forest plot showed that the selected 12 features all contain items with  $P < 0.05$ , among  
 252 which the degree of severe of CMBI (OR, 19.84; 95% CI, 13.88-28.34;  $p < 0.001$ ) had the greatest impact  
 253 on the risk of mental health problems among factory workers and miners (Figure 2).

254

Table 2 Predictive factors of risk for mental health problems among factory workers and miners

Variable	$\beta$	S.E.	OR(CI95%)	Wald	<i>P</i>	VIF
Intercept	-2.33	0.25	0.10(0.06,0.16)	-9.357	0	-
Education level						
Junior school and below VS High school	0.34	0.13	1.41(1.10,1.81)	2.727	0.006**	2.28

1							
2							
3	Junior school and below VS Junior						
4	college	0.44	0.11	1.56(1.24,1.95)	3.850	< 0.001***	2.79
5							
6	Junior school and below VS Bachelor's						
7	degree or above	0.38	0.13	1.46(1.13,1.87)	2.953	0.003**	2.51
8							
9	Professional title						
10	None VS Primary	0.15	0.09	1.16(0.97,1.39)	1.582	0.114	1.35
11	None VS Middle	0.05	0.09	1.05(0.87,1.26)	0.519	0.604	1.34
12	None VS Senior	0.27	0.11	1.30(1.06,1.61)	2.458	0.014*	1.32
13							
14							
15	Work schedule						
16	Day and night shifts VS Day shift	-0.38	0.11	0.69(0.55,0.85)	-3.364	0.001**	2.70
17	Day and night shifts VS Night shif	0.01	0.20	1.01(0.68,1.49)	0.044	0.965	1.30
18	Day and night shifts VS Shift	0.01	0.12	1.01(0.81,1.27)	0.107	0.915	2.47
19							
20	Marital status						
21	Unmarried VS Married	0.16	0.13	1.18(0.91,1.52)	1.263	0.206	2.29
22	Unmarried VS Divorced	0.55	0.19	1.73(1.20,2.51)	2.918	0.004**	1.69
23	Unmarried VS Widowed	0.69	0.43	1.99(0.85,4.64)	1.586	0.113	1.09
24							
25	Age						
26	~25 VS 25~	-0.02	0.20	0.98(0.66,1.47)	-0.083	0.934	3.09
27	~25 VS 30~	-0.02	0.22	0.98(0.64,1.50)	-0.090	0.929	4.79
28	~25 VS 35~	0.56	0.23	1.76(1.13,2.74)	2.503	0.012*	5.01
29	~25 VS 40~	0.33	0.23	1.39(0.88,2.21)	1.419	0.156	4.97
30	~25 VS 45~	0.23	0.22	1.26(0.81,1.95)	1.018	0.308	10.93
31							
32	Working years						
33	~5 VS 5~	0.44	0.14	1.55(1.18,2.05)	3.114	0.002**	2.27
34	~5 VS 10~	0.06	0.15	1.06(0.78,1.43)	0.366	0.714	2.48
35	~5 VS 15~	0.06	0.20	1.06(0.72,1.56)	0.305	0.760	1.79
36	~5 VS 20~	0.29	0.18	1.33(0.95,1.88)	1.641	0.101	2.65
37	~5 VS 25~	0.48	0.17	1.61(1.15,2.25)	2.782	0.005**	3.99
38	~5 VS 30~	0.20	0.16	1.22(0.89,1.68)	1.239	0.216	3.90
39							
40	Working hours per day						
41	≤7 VS >7	-0.50	0.09	0.61(0.50,0.73)	-5.363	< 0.001***	1.15
42							
43	Diabetes						
44	No VS Yes	0.43	0.14	1.53(1.16,2.03)	2.974	0.003**	1.05
45							
46	Hypertension						
47	No VS Yes	0.52	0.09	1.69(1.42,2.00)	5.885	< 0.001***	1.11
48							
49	Asbestos dust						
50	No VS Yes	0.44	0.10	1.55(1.28,1.87)	4.474	< 0.001***	1.03
51							
52	ERI						
53							
54							
55							
56							
57							
58							
59							
60							

1								
2								
3								
4		No VS Yes	0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***	1.05
5	CMBI							
6		No VS Mild	0.26	0.12	1.30(1.03,1.64)	2.175	0.003**	2.73
7								
8		No VS Moderate	1.30	0.11	3.67(2.93,4.59)	11.361	< 0.001***	2.83
9								
10		No VS Severe	2.99	0.18	19.84(13.88,28.34)	16.41	< 0.001***	1.44
11								

255 Note:  $\beta$  is the regression coefficient. "\*\*\*\*" indicates  $P < 0.001$ , "\*\*\*" indicates  $P < 0.01$ , "\*\*" indicates  $P < 0.05$ .

### 256

### 257 3.4. Development of an individualized prediction model

258

259 Based on the results of the multivariate analysis, predictors such as education, professional title, age,  
 260 CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working years, marital status,  
 261 and work schedule were included in the nomogram. A model incorporating the above independent  
 262 predictors was developed and represented as a nomogram in Figure 3. Each variable in nomogram was  
 263 assigned a score, and the cumulative sum of each 'point' was the 'total score'. The "total score"  
 264 corresponded to the "predictable likelihood", which was the predicted probability of mental health  
 265 problems among factory workers and miners as suggested by our design of the nomogram.

266

267 As an example of the use of nomogram: a randomly selected sample from the training group, one with  
 268 no professional title, day shift, no diabetes or hypertension, Junior college, <5 of working years, >7 of  
 269 working hours per day, married, no exposed to asbestos dust, <25 years of age, no ERI, mild of CMBI,  
 270 with a calculated total score of 174 and a corresponding risk probability of 8.27% for mental health  
 271 problems.

### 272

### 273 3.5 The validation of calibration

274

275 Model validation was carried out in the validation group. The prediction accuracy of the model was  
 276 assessed by two aspects. (1) The Brier score for overall performance, which assessed the difference  
 277 between observed and predicted values, with values closer to 0 indicating better predictive ability. (2)  
 278 The calibration slope used for modal calibration, which assessed the agreement between the observed  
 279 and predicted values, with values closer to 1 indicating better performance. The accuracy measurements  
 280 for the bias correction were validated by the model with a Brier score of 0.176 and a calibration slope of  
 281 0.970, respectively (Figure 4). The prediction accuracy of the model was relatively high.

### 282

### 283 3.6 The validation of discrimination

284

285 ROC was plotted for the training and validation groups, and the AUC of training and the verification  
 286 groups were 0.785 and 0.784, respectively (Figure 5). The AUC of training and the verification groups  
 287 were both greater than 0.75, showing a good discrimination.

288

### 289 3.7 Decision Curve Analysis

290

291 As shown in the DCA of the risk of mental health problems nomogram in Figure 6, the model for  
292 predicting the risk of mental health problems for factory workers and miners in this study was more  
293 practically relevant if the threshold probability of patients was >10%.

294

### 295 4. Discussion

296

297 In this study, we developed and validated an easy-to-use nomogram model as a new method for  
298 predicting the risk of mental health problems among factory workers and miners. To the best of our  
299 knowledge, this is the first study to establish an objective indicators nomogram combination model based  
300 on mental health survey. Our study included common demographic, job burnout, occupational stress,  
301 chronic diseases and occupational exposure factors to distinguish whether the respondents suffer from  
302 mental health problems. This nomogram showed good accuracy and discrimination.

303

304 LASSO is suitable for analyzing a large number of clinical factors and avoiding over-fitting [30]. In our  
305 study, a total of 23 candidate variables were used to construct the nomogram, which were reduced to 12  
306 potential predictor variables by using the LASSO regression method. The nomogram could be a useful  
307 tool to better identify patients with mental health problems, as it not only covered comprehensive  
308 information, including demographic information, job burnout, occupational stress, chronic diseases and  
309 occupational exposure factors closely related to factory workers and miners, but also was simple to  
310 operate and easy to use. Therefore, the possibility of early intervention for patients with high-risk mental  
311 health problems will be increased by covering multiple information and easy to use nomogram modal,  
312 especially for factory workers and miners with poor working conditions, relatively low levels of  
313 education and low patience.

314

315 Mental health problems were very common in the group of factory workers and miners, and the  
316 prevalence of mental health of them was found to be 37.08% in our study. Notably, the CMBI showed  
317 the most significant score (score = 100) and the ERI also had a high score (score = 43) in mental health  
318 problem incidence risk nomogram, which indicated that both of them were relatively important factors  
319 for mental health problems among the group of factory workers and miners. Our finding was consistent  
320 with other studies that had shown that occupational stress was a significant predictor of anxiety and was  
321 negatively associated with mental health. In addition, there is a high correlation between burnout and  
322 depression [31].

323

324 In line with previous studies, working years was also an important influential factor in this study. Related  
325 study has shown that employment could improve patients' mental health, while unemployment could  
326 lead to a deterioration in mental health [32]. In China, workers' working years is an important aspect of  
327 employment, and researchers have studied this aspect and found that precarious employment is a source  
328 of stress for individuals and predisposes them to mental health problems [33]. In addition, environmental

1  
2  
3 329 factors were also one of the influential factors of mental health problems in our study. Relevant studies  
4 330 have found that exposure to air pollution is associated with increased suicide risk and depressive  
5 331 symptoms [34]. Hypertension and diabetes were the influential factors in this study. A study has shown  
6 332 that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher  
7  
8 333 than in the non-diabetic population [35]. Furthermore, there is a recognized association between  
9 334 hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear  
10  
11 335 [36].

12  
13 336

14 337 Factory workers and miners were inevitably exposed to occupational hazards such as benzene and  
15 338 asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers  
16 339 are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous  
17 340 enterprises, involving more than 30 different types of operations, of which factory workers and miners  
18 341 is the one [37]. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for  
19 342 mental health problems in this study. Our study found that the work schedules of factory workers and  
20 343 miners were vary and the phenomenon of night shifts was very common, which inevitably affected their  
21 344 normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health  
22 345 and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-  
23 346 regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could  
24 347 suffer from negative emotions such as anxiety and depression [38]. Professional title and education level  
25 348 were also important influences on mental health issues. In the workplace, generally speaking, the higher  
26 349 the professional title and education level, the higher the status of the worker in the company and the  
27 350 greater the role played in the position. The number of studies on socio-economic status and mental health  
28 351 had increased in recent years. Some of these studies have shown that major depression is higher in the  
29 352 low socio-economic status group [39]. It has also been suggested that education itself is the best indicator  
30 353 of socio-economic status [40]. Marital status was one of the influential factors for mental health problems.  
31 354 Many studies have found an association between mental health and gender, marital status, lifestyle and  
32 355 working conditions, and it has been shown that poor mental health in women is associated with divorce  
33 356 or widowhood [41]. In this study, working more than seven hours a day was a determinant factor on mental  
34 357 health problems, which was consistent with other studies that had shown that long working hours could  
35 358 have a negative impact on employees' mental health and that excessive workloads could increase workers'  
36 359 fatigue, which in turn could lead to anxiety and depression [42].

37  
38  
39 360

40 361 In China, there are many problems in identifying people with mental health problems due to uneven and  
41 362 imperfect levels of medical development across regions. Some studies have shown that in mainland  
42 363 China, general practitioners, surgeons and primary health care workers often have little or no mental  
43 364 health training, which prevents them from providing basic mental health services [43]. Non-mental health  
44 365 professionals in general hospitals learn about mental illness on their own, rather than learning about it  
45 366 during their formal education<sup>44</sup>. Therefore, this study designed a simple and comprehensive nomogram  
46 367 to address the issue of timely detection and effective interventions for people with mental health problems,  
47 368 so that people at risk of mental health problems could easily calculate their probability of suffering from  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 369 mental health problems without the help of medical staff. This study has several strengths. First, to our  
4 370 knowledge, this is the first model to develop and assess the likelihood of mental health problems in a  
5 371 group of factory workers and miners. Secondly, the nomogram in this study includes demographic  
6 372 information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors  
7 373 that are closely related to factory workers and miners, allowing for a more accurate assessment of the  
8 374 risk of morbidity among them, as well as providing a methodological reference for other related studies.  
9 375

#### 376 **Patient and public involvement**

377 Neither patients nor members of the public had any involvement in the design of this study.

378

379 **Acknowledgements** The authors are grateful to all participants and investigators.

380

381 **Contributions** Y.L., Q.L., and T.L. are responsible for conceptualization; Y.L. is responsible for  
382 methodology, software, formal analysis, resources, and visualization; Q.L. and T.L. are responsible for  
383 the original draft preparation; Q.L. and H.Y. are responsible for reviewing; Q.L. is responsible for editing;  
384 T.L. is responsible for supervision. Yaoqin Lu and Qi Liu contributed equally to this work.

385

386 **Funding** This study was supported by National Natural Science Foundation of China, grant number  
387 81760581 and Public Health and Preventive Medicine, the 13th Five-Year Plan Key Subject of Xinjiang  
388 Uygur Autonomous Region.

389

390 **Competing interests** None declared.

391

392 **Patient consent for publication** Not required.

393

394 **Ethics approval** The study was approved by the ethics committee of Urumqi Center for Disease Control  
395 and Prevention

396

397 **Data availability statement** The data used to support the findings of this study are available from the  
398 corresponding author upon request.

399

#### 400 **References**

- 401 [1] WHO Terminology Information System [online glossary] [http://www.who.int/health-systems-](http://www.who.int/health-systems-performance/docs/glossary.html)  
402 [performance/docs/glossary.html](http://www.who.int/health-systems-performance/docs/glossary.html).  
403 [2] Wang Y, Liu X, Qiu J, Wang H, Liu D, Zhao Z, Song M, Song Q, Wang X, Zhou Y, Wang W.  
404 Association between ideal cardiovascular health metrics and suboptimal health status in Chinese  
405 population. *Sci Rep* 2017;7:14975.  
406 [3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental  
407 health. *Lancet*. 2007;370:859–77.  
408 [4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive  
409 symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. *Health*

- 1  
2  
3 410 *Place* 2016;42:37–46.
- 4 411 [5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,  
5 412 Degenhardt L, Dicker D, *et al.* Global, regional, and national incidence, prevalence, and years lived  
6 413 with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A  
7 414 systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;386:743–800.
- 8 415 [6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health? *Int J*  
9 416 *Environ Res Public Health* 2020;17:7530.
- 10 417 [7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative  
11 418 income, gender, and location. *China Econ Rev* 2018;48:83–101.
- 12 419 [8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:  
13 420 Results from the PATH Through Life Survey. *Soc Sci Med.* 2009;69:229–237.
- 14 421 [9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and  
15 422 Subjective Well-being? *J Environ Econ Manag* 2017;85:81–94.
- 16 423 [10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. *World*  
17 424 *Psychiatry.* 2015;14:36–42.
- 18 425 [11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson  
19 426 S.J. IPS Supported employment: a review. *J Vocat Rehabil* 2014;40:1–13.
- 20 427 [12] Jia Z, *et al.* Exposure to ambient air particles increases the risk of mental disorder: findings from a  
21 428 natural experiment in Beijing. *Int J Environ Res Public Health* 2018;15:160.
- 22 429 [13] Szyszkwicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency  
23 430 department visits for suicide attempts in vancouver, Canada. *Environ Health Insights* 2010;4:79–86.
- 24 431 [14] Michael J. S. International classification of sleep disorders-third edition : highlights and  
25 432 modifications. *Chest* 2014;146:1387–1394.
- 26 433 [15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial  
27 434 infarction: gender differences. *Open Nurs J* 2018;12:205–214.
- 28 435 [16] Wurm W, Vogel K, Holl A, *et al.* Depression-burnout overlap in physicians. *PLoS One*  
29 436 2016;11:e0149913.
- 30 437 [17] Porru F, Robroek S J W, Bültmann U, *et al.* Mental health among university students: The  
31 438 associations of effort-reward imbalance and overcommitment with psychological distress. *J Affect*  
32 439 *Disord* 2021;282:953-961.
- 33 440 [18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,  
34 441 and burnout among physician assistant students: a multi-institutional study. *Health Professions Edu*  
35 442 2020;6:420–427.
- 36 443 [19] Hua D, Kong Y, Li W, Han Q, Zhang X, Zhu LX, Wan SW, Liu Z, Shen Q, Yang J, He HG, Zhu J.  
37 444 Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during  
38 445 the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. *Eclinical Med*  
39 446 2020;24:100424.
- 40 447 [20] Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP. The estimation of the number of  
41 448 underground coal miners and the annual dose to coal miners in China. *Health Phys* 2007;93:127–  
42 449 132.
- 43 450 [21] Yong X, Gao X, Zhang Z, *et al.* Associations of occupational stress with job burn-out, depression  
44 451 and hypertension in coal miners of xinjiang, china: A cross-sectional study. *BMJ open* 2020;10:  
45 452 e036087.
- 46 453 [22] P. Bech, J. Bille, S. B. Møller, L. C. Hellström, S. D. Østergaard. Psychometric validation of the

- 1  
2  
3 454 Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal  
4 455 sensitivity. *J Affect Disord* 2014;160:98–103.
- 6 456 [23] J. Zhang, X. Zhang. Chinese college students' SCL-90 scores and their relations to the college  
7 457 performance. *Asian J Psychiatr* 2013;6:134–140.
- 8 458 [24] Derogatis L, Lipman R.S, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary  
9 459 report. *Psychopharmacol Bull* 1973;9:13–28.
- 11 460 [25] Crespo-Maraver M, Doval E, Fernã N.J, Gimã Nez-Salinas J, Prat G, Bonet P. Caregiver's health:  
12 461 adaption and validation in a Spanish population of the Experience of Caregiving Inventory  
13 462 (ECI) *Gac. Sanit* 2018;33:348–355.
- 15 463 [26] Y. Li, X. Sun, H. Ge, J. Liu, L. Chen. The status of occupational stress and its influence the quality  
16 464 of life of coppernickel miners in Xinjiang, China. *Int J Environ Res Public Health* 2019;3:353–362.
- 17 465 [27] Siegrist J, Wege N, Pühlhofer F, Wahrendorf M. A short generic measure of work stress in the era  
18 466 of globalization: Effort-reward imbalance. *Int Arch Occup Environ Health* 2009;82:1005–1013.
- 20 467 [28] Zhang Z, Lu Y, Yong X, *et al.* Effects of occupational radiation exposure on job stress and job  
21 468 burnout of medical staff in xinjiang, China: A cross-sectional study. *Med Sci Monit* 2020;26:  
22 469 e927848.
- 24 470 [29] Freudenberger HJ. Staff burnout. *J Soc Issues* 1974;30:159–165.
- 25 471 [30] Hepp T, Schmid M, Gefeller O, Waldmann E, Mayr A. Approaches to regularized regression - a  
26 472 comparison between gradient boosting and the lasso. *Methods Inf Med* 2016;55:422–430.
- 28 473 [31] Occupational Stress and Employees Complete Mental Health: A Cross-Cultural Empirical Study
- 29 474 [32] Knapp M and Wong G. Economics and mental health: the current scenario. *World Psychiatry*  
30 475 2020;19:3–14.
- 32 476 [33] Benach, J, Vives, A, Amable, M, Vanroelen, C, Tarafa, G, & Muntaner, C. Precarious employment:  
33 477 understanding an emerging social determinant of health. *Annu Rev Public Health* 2014;35:229–53.
- 34 478 [34] Bakian A. V. *et al.* Acute air pollution exposure and risk of suicide completion. *Am J Epidemiol*  
35 479 2015;181:295–303.
- 37 480 [35] Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type  
38 481 1 diabetes: systematic literature review. *Diabet Med* 2006;23:445–448.
- 40 482 [36] Gilsanz P, Karter AJ, Beerl MS, Quesenberry CP, Whitmer RA. The bidirectional association  
41 483 between depression and severe hypoglycemic and hyperglycemic events in type 1  
42 484 diabetes. *Diabetes Care* 2018;41:446–452.
- 43 485 [37] Lu Y, Zhang Z, Yan H, *et al.* Effects of occupational hazards on job stress and mental health of  
44 486 factory workers and miners: A propensity score analysis. *BioMed Res Int* 2020;2020:1754897.
- 46 487 [38] Shi L, Liu Y, Jiang T, *et al.* Relationship between mental health, the clock gene, and sleep quality  
47 488 in surgical nurses: A cross-sectional study. *BioMed Res Int* 2020;2020:4795763.
- 49 489 [39] Sallis JF, Saelens BE, Frank LD *et al.* Neighborhood built environment and income: examining  
50 490 multiple health outcomes. *Soc Sci Med* 2009;68:1285–1293.
- 51 491 [40] Winkleby MA, Jatulis DE, Frank E *et al.* Socioeconomic status and health: how education, income,  
52 492 and occupation contribute to risk-factors for cardiovascular disease. *Am J Public*  
53 493 *Health* 1992;82:816–820.
- 55 494 [41] Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis P.N, Mavreas V. Prevalence  
56 495 and sociodemographic associations of common mental disorders in a nationally representative  
57 496 sample of the general population of Greece. *BMC Psychiatry* 2013;13:163.
- 59 497 [42] Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG *et al.* Long  
60

1  
2  
3 498 working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II  
4 499 study. *Psychol Med* 2011;41:2485–2494.

5 500 [43] Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, *et al.* Prevalence, treatment, and associated  
6 501 disability of mental disorders in four provinces in China during 2001–05: an epidemiological  
7 502 survey. *Lancet* 2009;373:2041–2053.

8 503 [44] Wu Q, Luo X, Chen S, *et al.* Mental health literacy survey of non-mental health professionals in six  
9 504 general hospitals in hunan province of China. *PloS one* 2017;12:e0180327.

10 505

11 506 **Figure legends**

12 507 **Fig.1. Feature selection using the LASSO binary logistic regression model.** (A) Feature selection for the LASSO  
13 508 binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against  
14 509 lambda by validating the optimal parameter lambda in the LASSO model. Dotted vertical lines were drawn based  
15 510 on 1 SE of the minimum criteria (the 1-SE criteria). (B) Feature selection was performed using the LASSO binary  
16 511 logistic regression model. A Coefficient profile was plotted based on the lambda series in Figure 1(A), and 12  
17 512 features with non-zero coefficients were selected by optimal lambda.

18 513

19 514 **Fig.2. The forest plot of the OR of the selected feature.**

20 515

21 516 **Fig.3. Developed mental health problems incidence risk nomogram.** The mental health problems incidence risk  
22 517 nomogram was developed in the array, with education, professional title, age, CMBI, ERI, asbestos dust,  
23 518 hypertension, diabetes, working hours per day, working years, marital status, and work schedule incorporated.

24 519

25 520 **Fig.4. Calibration curves of the mental health problems incidence risk nomogram prediction in validation**  
26 521 **group.** The x-axis represents the predicted risk of mental health problems. y-axis represents the actual diagnosed  
27 522 risk of mental health problems. The diagonal dashed line represents the perfect prediction of the ideal model. The  
28 523 solid lines represent the performance of the column plots, where closer to the diagonal dashed line indicates a better  
29 524 prediction.

30 525

31 526 **Fig.5. ROC curves for training and validation groups.** The y-axis represents the true positive rate of risk  
32 527 prediction. The x-axis represents the false positive rate of risk prediction. The ROC curves for the training and  
33 528 validation groups are shown in black and red.

34 529

35 530 **Fig.6. Decision curve analysis for mental health problems incidence risk nomogram.** The y-axis measures the  
36 531 net benefit. The solid red line represents nomogram of the risk of developing a mental health problem. The light blue  
37 532 dashed line represents the hypothesis that all participants were diagnosed with a mental health problem. The black  
38 533 dashed line represents the hypothesis that there is no risk of a mental health problem. The DCA showed that using  
39 534 this mental health problem incidence risk nomogram in the current study to predict mental health problem incidence  
40 535 risk increase in benefit than the intervention all patients or no intervention all patient if the threshold probability of  
41 536 a patients and a doctor is >10%.

42 537

43 538

44 539

45 540

46 541

47 542

48 543

49 544

50 545

51 546

52 547

53 548

54 549

55 550

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

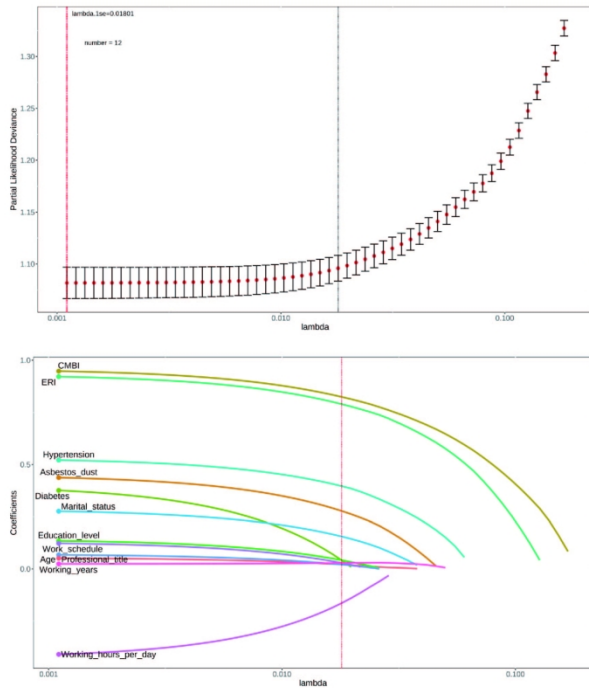
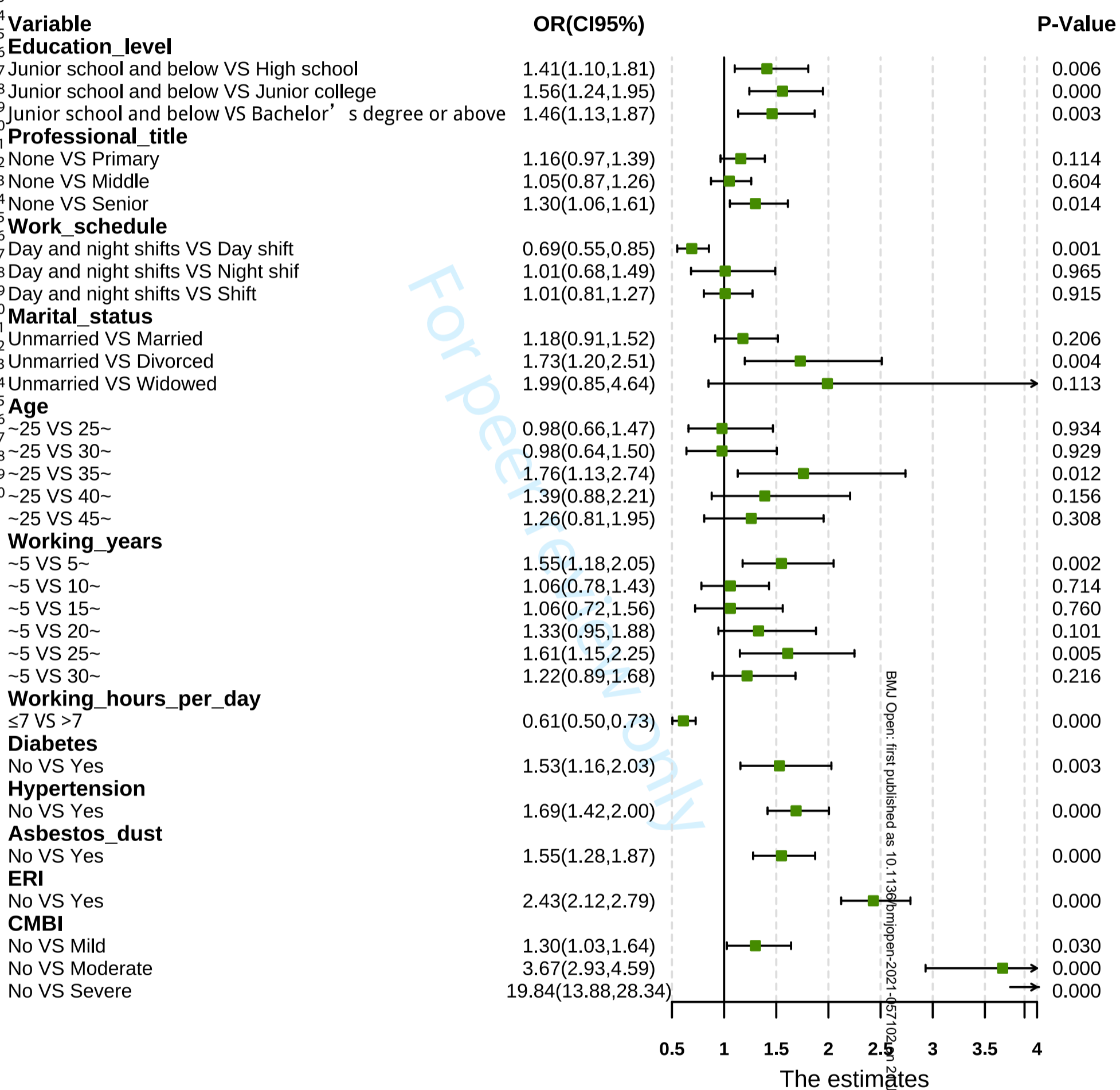


Fig.1. Feature selection using the LASSO binary logistic regression model.  
210x297mm (300 x 300 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



BMJ Open: first published as 10.1136/bmjopen-2021-025710 on 20 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 24, 2024 by guest. Protected by copyright.

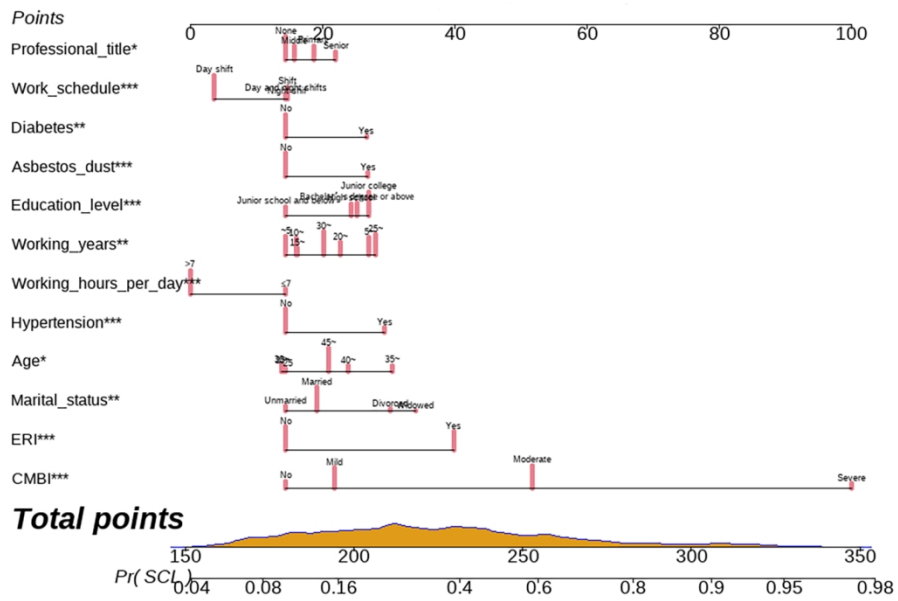
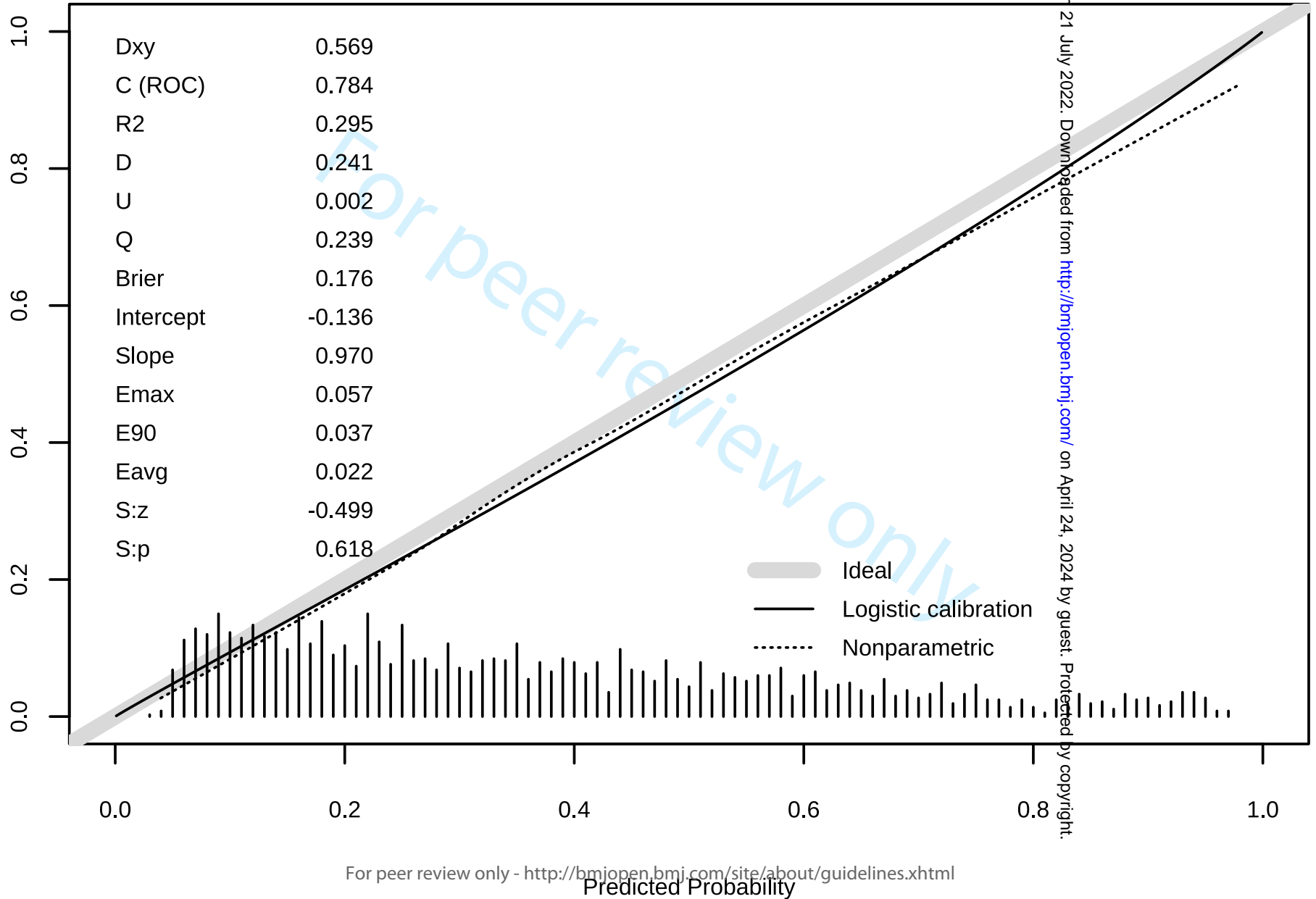


Fig.3. Developed mental health problems incidence risk nomogram.

352x230mm (118 x 118 DPI)

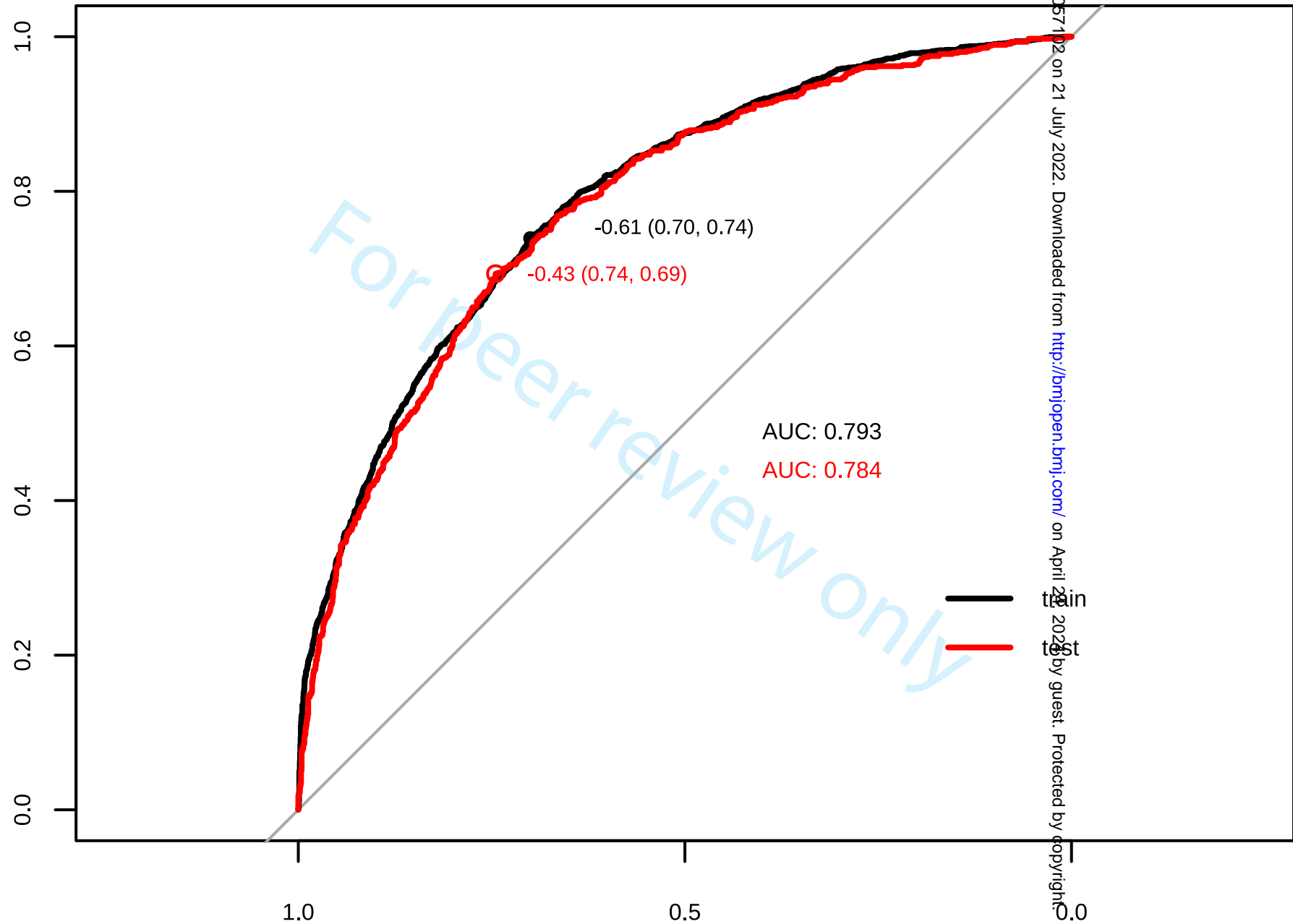


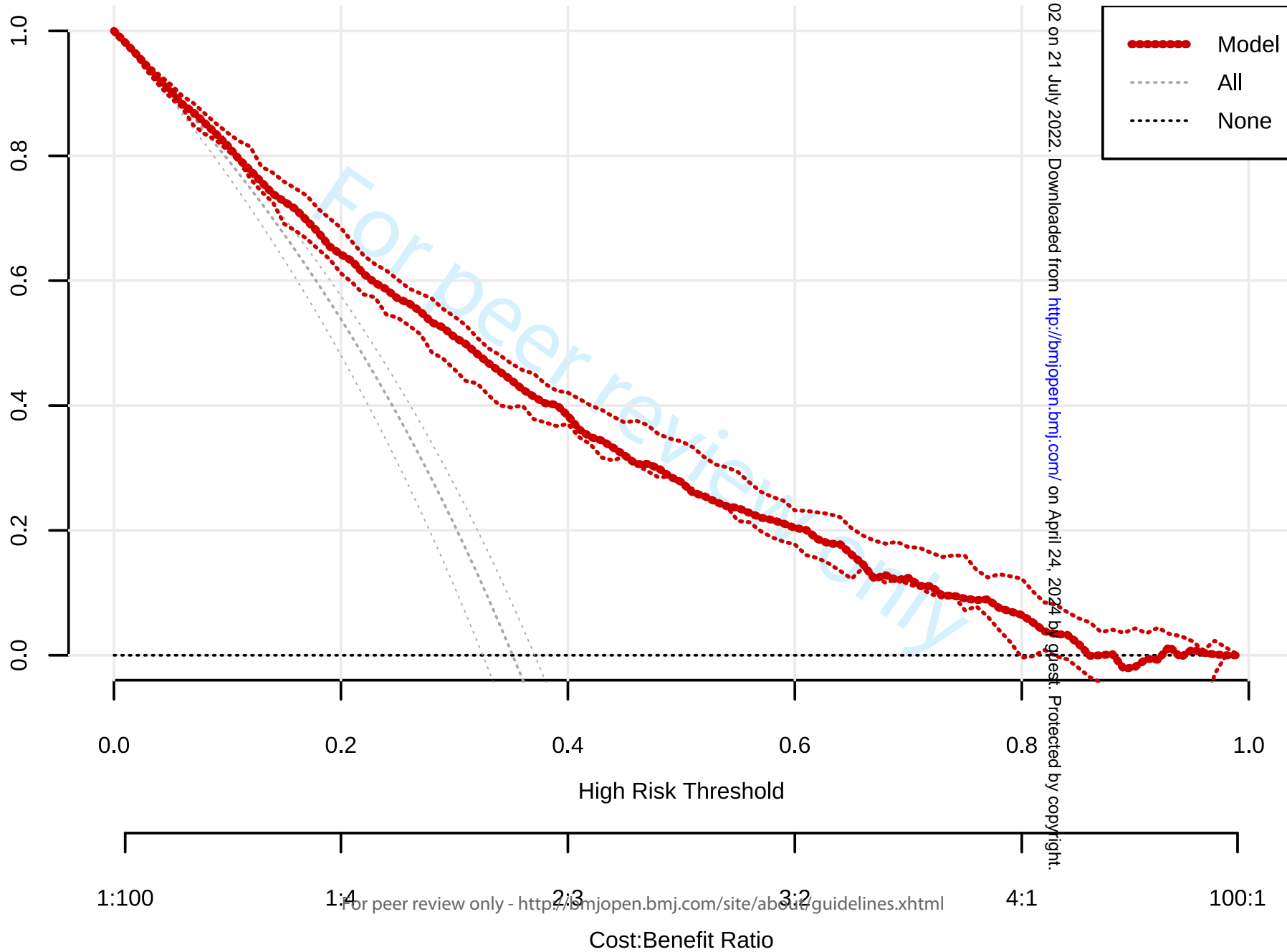
pen-2021-057102 or 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

For peer review only





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	1	Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems of Factory Workers and Miners
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1	A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training group (n=4,955) and a validation group (n=2,163) in a ratio of 3:1. A total of 23 characteristics were included in this study and LASSO regression selected 12 characteristics such as education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working years, marital status, and work schedule as predictors for the construction of the nomogram. In the validation group the Brier score was 0.176, the calibration slope was 0.970 and the calibration curve of nomogram showed a

For peer review only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

For peer review only

good fit, indicating good agreement between predictions and observations. The AUC of training group and verification group were 0.785 and 0.784 respectively, which showed good discrimination. The DCA suggested that the nomogram for predicting the risk of mental health problems among factory workers and miners was more practical when the risk threshold for mental health problems was 10% for intervention.

**Introduction**

Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	2	Factory workers and miners are a special group of workers with a relatively low overall level of education and are highly prone to suffering from mental health problems due to limited social support, excessive workload and irregular lifestyles, as well as occupational hazards such as noise and coal dust that they inevitably need to face in their working environment <sup>18-19</sup> . China
----------------------	---	--	---	--

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

For peer review only

has the world's largest group of factory workers and miners, about 6 million<sup>20</sup>, who are regularly involved in occupational hazards. Mental health problems which need to require a long process are known to be a syndrome caused by chronic stress. Factory workers and miners, represented by those engaged in coal mining, have a mental burden rating of 8.3, one of the highest mental burdens among 150 occupations<sup>21</sup>. This explains the high level of mental health problems among mine workers in previous studies, making the identification and treatment of mental health problems even more important. Therefore, it is essential to provide a viable and easy-to-apply tool for identifying workers at risk of mental health problems and thus for timely interventions.

Objectives 3 State specific objectives, including any prespecified hypotheses

3 Therefore, the aim of our study is to develop and validate an easy-to-use nomogram that combines objective information on the

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

					demographics, job burnout, occupational stress and occupational hazards to comprehensively and accurately predict the prevalence of mental health problems among factory workers and miners.
<b>Methods</b>					
Study design	4	Present key elements of study design early in the paper		3	The selection of participants. The quality of the questionnaires. The results of agreement and discrimination between predictions and observations in this nomogram.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection		4	Participants in this cross-sectional survey were workers from factories and mining enterprises in the Urumqi region, who were recruited using a whole-group sampling method. A total of 3,619 enterprises in the Urumqi were surveyed from January to May 2019, covering all districts and counties in the Urumqi region, including Tianshan District, Shaibak District, Xinshi District, Shuimogou District, Toutunhe District, Dabancheng District,

				Middong District and Urumqi County.
Participants	6	<p>(a) <i>Cohort study</i>—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up</p> <p><i>Case-control study</i>—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</p> <p><i>Cross-sectional study</i>—Give the eligibility criteria, and the sources and methods of selection of participants</p>	4	<p>The exclusion criteria were the following: (I) factory workers and miners in non-Urumqi area, (II) working history of factories and mining enterprises less than 1 year, (III) a confirmed diagnosis of a mental health problem and a history of treatment and use of psychotropic medication. Questionnaires with missing data were also excluded from the analysis based on discussion and agreement among the subject members.</p>
		<p>(b) <i>Cohort study</i>—For matched studies, give matching criteria and number of exposed and unexposed</p> <p><i>Case-control study</i>—For matched studies, give matching criteria and the number of controls per case</p>		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4-	<p>2.2.1. Assessment of mental health</p> <p>2.2.2. Assessment of occupational stress</p> <p>2.2.3. Assessment of job burnout</p> <p>2.2.4. Candidate predictors</p>
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	6	Categorical variables were described as counts and percentages, and chi square test or Fisher exact test was used to compare categorical variables

For peer review only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

		<p>between groups. 70% of participants were randomly assigned to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which predictive models were constructed. A nomogram for predicting was generated according to the selected characteristics.</p>
<p>Bias</p>	<p>9 Describe any efforts to address potential sources of bias</p>	<p>6 We used calibration plots and receiver operating characteristic (ROC) curves to show the calibration and discrimination of our final model. Brier scores for overall performance, calibration slopes were used to assess the predictable accuracy of the model. Decision curve analysis (DCA) was applied to calculate the net benefit of the nomogram. Statistical analysis was performed using the open-source R software Version 3.6.1 (<a href="http://www.r-">http://www.r-</a></p>



					project.org). The significance level ( $\alpha$ ) set at 0.05.	
Study size	10	Explain how the study size was arrived at				
Continued on next page						
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why			7	Sex was defined as male or female; ethnicity was defined as Han and other; education level was defined as junior high school and below, high school, junior college or bachelor's degree or above; labor contracts was defined as signed or unsigned; professional title was defined as no, primary, middle or senior; work schedule was defined as day shift, night shift, shift or day and night shifts; marital status was defined as unmarried, married, divorced or widowed; monthly income (yuan) was defined as <3000, 3000~, 4000~, 5000~, 6000~, 7000~ or 8000~; age (years) was defined as <25, 25~, 30~, 35~, 40~ or 45~; working years was defined as ~5, 5~, 10~, 15~, 20~, 25~ or 30~; working hours per day (hours) was defined as $\leq 7$ or $> 7$ ; working days per week (days) was defined as $\leq 5$ or $> 5$ ; exposure to coal dust, silica dust, asbestos dust, benzene, lead, noise, brucellosis were all defined as yes or no; ERI

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

				was defined as yes or no; CMBI was defined as none, mild, moderate and severe; hypertension and diabetes were both defined as yes or no.
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	6	Categorical variables were described as counts and percentages, and chi square test or Fisher exact test was used to compare categorical variables between groups. 70% of participants were randomly assigned to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which predictive models were constructed. A nomogram for predicting was generated according to the selected characteristics. In addition, forest plot was drawn to visually depict the P-value, OR and 95% CI for the selected validations.
		(b) Describe any methods used to examine subgroups and interactions		

For peer review only

1				
2		(c) Explain how missing data were addressed	4	A total of 7,500 questionnaires were distributed and 7,315 questionnaires were returned, representing a return rate of 97.5%. After checking the validity and integrity of the questionnaires, 7,118 questionnaires were confirmed as valid, with an effective rate of 97.3%. All participants understood the purpose of the study and voluntarily participated in the study.
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20		(d) Cohort study—If applicable, explain how loss to follow-up was addressed		
21		Case-control study—If applicable, explain how matching of cases and controls was addressed		
22		Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy		
23				
24		(e) Describe any sensitivity analyses		
25				
26	<b>Results</b>			
27	Participants	13* (a) 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	4	7500 participants volunteered for the survey Issued a total of 7500 questionnaires Collected a total of 7315 questionnaires 7118 valid and integrated questionnaires
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39		(b) Give reasons for non-participation at each stage	4	
40		(c) Consider use of a flow diagram	4	
41				
42				
43				
44				
45				
46				

1  
2 Descriptive data 14\* (a) Give characteristics of study participants (eg demographic, clinical, social) and information on  
3 exposures and potential confounders  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

6

A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training group (n=4,955) and a validation group (n=2,163). Over half of all participants (65.31%) were male, 57.31% of the population was over 35 years of age and 78.32% of the subjects were married, showing that factory workers and miners are generally older and most of them have spouses. The majority of them had completed high school (83.94%), while a smaller percentage had completed undergraduate education (22.98%), indicating that the group of factory workers and miners as a whole was not well educated. The total number of workers (n, %) exposed to coal dust, silica dust, asbestos dust, benzene, lead, noise and brucellosis in the factory and mining enterprises were 377 (5), 730 (10), 981 (14), 1,981 (28), 373 (5), 4,942 (69) and 121 (2) respectively, with the total number of workers

6/bmjopen-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

For peer review only

exposed to noise amounting to 4,942, or 69% of the total population surveyed. The demographic, job burnout, occupational stress and occupational exposure factors for the training and validation groups are shown in Table 1. The results showed that there were no significant statistical differences between the two groups of characteristic variables, except for coal dust and CMBI, indicating that the baseline levels were largely consistent between the two groups.

(b) Indicate number of participants with missing data for each variable of interest

(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)

Outcome data 15\* *Cohort study*—Report numbers of outcome events or summary measures over time

*Case-control study*—Report numbers in each exposure category, or summary measures of exposure

*Cross-sectional study*—Report numbers of outcome events or summary measures

A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training group (n=4,955) and a validation group (n=2,163). Over half of all participants (65.31%) were male, 57.31% of the population was over 35 years of age and 78.32% of the subjects were married, showing that factory workers and miners are

generally older and most of them have spouses. The majority of them had completed high school (83.94%), while a smaller percentage had completed undergraduate education (22.98%), indicating that the group of factory workers and miners as a whole was not well educated. The total number of workers (n, %) exposed to coal dust, silica dust, asbestos dust, benzene, lead, noise and brucellosis in the factory and mining enterprises were 377 (5), 730 (10), 981 (14), 1,981 (28), 373 (5), 4,942 (69) and 121 (2) respectively, with the total number of workers exposed to noise amounting to 4,942, or 69% of the total population surveyed. The demographic, job burnout, occupational stress and occupational exposure factors for the training and validation groups are shown in Table 1. The results showed that there were no significant statistical differences between the two groups of characteristic variables, except for coal dust and CMBI, indicating that the baseline levels were largely consistent between the two groups.

For peer review only

6/bmjopen-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Main results	16	<i>(a)</i> 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	6	Categorical variables were described as counts and percentages, and chi square test or Fisher exact test was used to compare categorical variables between groups. 70% of participants were randomly assigned to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which predictive models were constructed.
		<i>(b)</i> Report category boundaries when continuous variables were categorized	6	
		<i>(c)</i> If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		

Continued on next page

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses		
<b>Discussion</b>				
Key results	18	Summarise key results with reference to study objectives	13	Therefore, this study designed a simple and comprehensive nomogram to address the issue of timely detection and effective interventions for people with mental health problems, so that people at risk of mental health problems could easily calculate their probability of suffering from mental health problems without the help of medical staff.
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	13	This study also has several limitations. First, we have considered many influential factors including demographics, job burnout, occupational stress and occupational exposure factors, but we are still not certain whether all possible influences are covered. Secondly, while the robustness of our nomogram was extensively validated internally in the same population, external validation is lacking for other populations in other regions and countries. Nomogram need to be externally assessed in a wider population.

For peer review only

6/bmjopen-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13	The nomogram we proposed contains 12 characteristics related to demographics, job burnout, occupational stress and occupational hazard factors. The nomogram combining these 12 characteristics for the risk of mental health problems can be used to predict the risk of suffering mental health problems, providing a useful tool for quickly and accurately screening the risk of mental health problems among factory workers and miners.
Generalisability	21	Discuss the generalisability (external validity) of the study results		
<b>Other information</b>				
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	14	This work was supported by National Natural Science Foundation of China, grant number 81760581 and Public Health and Preventive Medicine, the 13th Five-Year Plan Key Subject of Xinjiang Uygur Autonomous Region.

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

# BMJ Open

## Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems of Factory Workers and Miners

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-057102.R1
Article Type:	Original research
Date Submitted by the Author:	17-Mar-2022
Complete List of Authors:	Lu, Yaoqin; Xinjiang Medical University, School of Public Health; Urumqi Center for Disease Control and Prevention Liu, Qi; Xinjiang Medical University, School of Public Health Yan, Huan; Xinjiang Medical University, Department of Nutrition and Food Hygiene; Xinjiang Autonomous Academy of Instrumental Analysis Liu, Tao; Xinjiang Medical University, School of Public Health
<b>Primary Subject Heading</b>:	Mental health
Secondary Subject Heading:	Public health
Keywords:	MENTAL HEALTH, PREVENTIVE MEDICINE, PUBLIC HEALTH

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

# 1 Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems 2 of Factory Workers and Miners

3 Yaoqin Lu,<sup>1,2</sup> Qi Liu,<sup>1</sup> Huan Yan<sup>3,4</sup>, and Tao Liu<sup>1</sup>

4 <sup>1</sup>Department of Toxicology, School of Public Health, Xinjiang Medical University, Urumqi 830011,  
5 Xinjiang, China.

6 <sup>2</sup>Urumqi Center for Disease Control and Prevention, Urumqi 830026, Xinjiang, China.

7 <sup>3</sup>Department of Nutrition and Food Hygiene, College of Public Health, Xinjiang Medical University,  
8 Urumqi, Xinjiang, China.

9 <sup>4</sup>Xinjiang Engineering Technology Research Center for Green Processing of Nature Product Center,  
10 Xinjiang Autonomous Academy of Instrumental Analysis, Urumqi, Xinjiang, China.

11  
12 Correspondence should be addressed to Tao Liu; xjmult@163.com

13 Yaoqin Lu and Qi Liu contributed equally to this work

## 14 Abstract

15 **Objective** A nomogram for predicting the risk of mental health problems was established in a population  
16 of factory workers and miners, in order to quickly calculate the probability of a worker suffering from  
17 mental health problems.

18 **Methods** A cross-sectional survey of 7,500 factory workers and miners in Urumqi was conducted by  
19 means of an electronic questionnaire using cluster sampling method. Participants were randomly  
20 assigned to the training group (70%) and the validation group (30%). Questionnaire-based survey was  
21 conducted to collect information. A least absolute shrinkage and selection operator (LASSO) regression  
22 model was used to screen the predictors related to the risk of mental health problems of the training  
23 group. Multivariate logistic regression analysis was applied to construct the prediction model. Calibration  
24 plots and receiver operating characteristic-derived area under the curve (AUC) were used for model  
25 validation. Decision curve analysis (DCA) was applied to calculate the net benefit of the screening model.

26 **Results** A total of 7,118 participants met the inclusion criteria and the data were randomly divided into  
27 a training group (n=4,955) and a validation group (n=2,163) in a ratio of 3:1. A total of 23 characteristics  
28 were included in this study and LASSO regression selected 12 characteristics such as education,  
29 professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working  
30 years, marital status, and work schedule as predictors for the construction of the nomogram. In the  
31 validation group the Brier score was 0.176, the calibration slope was 0.970 and the calibration curve of  
32 nomogram showed a good fit. The AUC of training group and verification group were 0.785 and 0.784  
33 respectively.

34 **Conclusion** The nomogram combining these 12 characteristics can be used to predict the risk of suffering  
35 mental health problems, providing a useful tool for quickly and accurately screening the risk of mental  
36 health problems.

38 **Key words** Mental health; Predictor; Nomogram; Risk; Factory workers and miners

39

40 **Strengths and limitations of this study**

- 41 1. This is the first study to develop an easy-to-use nomogram to predict the mental health risks of factory
- 42 workers and miners.
- 43 2. The AUC of training group and verification group were 0.785 and 0.784 respectively, showing
- 44 moderate discriminatory and calibration power.
- 45 3. This nomogram model's variables are more comprehensive, including demographics, burnout,
- 46 occupational stress and occupational hazards.
- 47 4. We had considered many influential factors, but we were still not certain whether all possible
- 48 influences were covered.
- 49 5. There is a lack of external validation in other populations in other regions and countries.

50

51 **1. Introduction**

52

53 The World Health Organization (WHO) defines health as a state of complete physical, mental and social  
54 well-being and not merely the absence of disease or weakness <sup>[1]</sup>. Obviously, health is an organic unity  
55 of physical and mental well-being. People with good mental health are the precondition for the normal  
56 operation of our society. However, with the acceleration of people's pace of life, people are facing an  
57 increasing risk of poor health, which has become a global public health problem <sup>[2]</sup>. Mental health  
58 problems can not only take a toll on physical health such as increasing the risk of communicable and  
59 non-communicable diseases and even causing unintentional or intentional harm to others <sup>[3]</sup>, but can also  
60 have a negative impact on the economy. For example, mental health disorders represent a growing part  
61 of the global burden of disease <sup>[4]</sup>, with statistics showing that nearly one billion people worldwide  
62 currently suffer from a mental disorder, and mental illness is ranked as one of the leading causes of the  
63 global burden of disease <sup>[5]</sup>. Moreover, one study has estimated that due to the impact of mental illness,  
64 the global economy loses US \$1 trillion every year <sup>[6]</sup>.

65

66 As researchers around the world have delved into the field of mental health, factors such as gender,  
67 income levels, environment and education have been found to be associated with people's mental health  
68 problems <sup>[7-10]</sup>. Moreover, employment is also strongly associated with quality of life, higher self-esteem  
69 and fewer psychiatric symptoms <sup>[11]</sup>. In addition, in the context of the global challenges of climate change,  
70 an increasing number of scholars have been examining the epidemiological links between mental health  
71 and environmental factors. Some studies have suggested that mental health may be influenced by ambient  
72 temperature, and an association has been found between environmental pollutants, particularly fine  
73 particulate matter, and mental health problems <sup>[12]</sup>. A relevant study shows that with short-term exposure  
74 to ambient air pollution is associated with increased emergency room visits due to depression or suicide  
75 attempts <sup>[13]</sup>. Furthermore, other factors associated with mental health include sleep, diabetes, coronary  
76 artery disease and cardiovascular disease <sup>[14-15]</sup>. It is worth noting that job burnout and occupational stress  
77 are closely linked to mental health. Job burnout is an exhaustion state of physical and psychological that

1  
2  
3 78 often occurs in the work environment, and has a high correlation with depression. A large study of  
4 79 physicians found that of the 10.3% who met criteria for a major depressive episode, 50.7% were also  
5 80 affected by symptoms of burnout (OR 2.99) and indicated that worsening depression leads to a higher  
6 81 likelihood of burnout symptoms <sup>[16]</sup>. Occupational stress refers to a work environment where non-  
7 82 reciprocity of effort and reward may lead to strong negative emotions and distress. Related research has  
8 83 shown that the combination of high effort and low reward and over-commitment increases the risk of  
9 84 mental health problems such as depression <sup>[17]</sup>. Apparently, it is necessary to include the CMBI and ERI  
10 85 in this study to predict the risk of mental health problems among factory workers and miners. However,  
11 86 there are few studies that include these influences in a more comprehensive way in the practice of  
12 87 detecting mental health. Therefore, more accurate identification of mental health problems in populations  
13 88 requires a questionnaire that include a wider range of factors affecting factory workers and miners'  
14 89 mental health problems.  
15 90

16 91 Factory workers and miners are a special group of workers with a relatively low overall level of education  
17 92 and are highly prone to suffering from mental health problems due to limited social support, excessive  
18 93 workload and irregular lifestyles, as well as occupational hazards such as noise and coal dust that they  
19 94 inevitably need to face in their working environment <sup>[18-19]</sup>. Through a review of the literature, our group  
20 95 found that coal dust, crystalline silica and noise pollution were common causes of health problems for  
21 96 workers in underground mines <sup>[20]</sup>. And, exposure to coal mine dust is a significant cause of  
22 97 pneumoconiosis in coal miners <sup>[21]</sup>. In addition, asbestos is one of the major occupational hazards in the  
23 98 daily work of workers in the construction and automotive industries <sup>[22]</sup>. China has the world's largest  
24 99 group of factory workers and miners, about 6 million <sup>[23]</sup>, who are regularly involved in occupational  
25 100 hazards. Mental health problems which need to require a long process are known to be a syndrome caused  
26 101 by chronic stress. Factory workers and miners, represented by those engaged in coal mining, have a  
27 102 mental burden rating of 8.3, one of the highest mental burdens among 150 occupations <sup>[24]</sup>. This explains  
28 103 the high level of mental health problems among mine workers in previous studies, making the  
29 104 identification and treatment of mental health problems even more important. Therefore, it is essential to  
30 105 provide a viable and easy-to-apply tool for identifying workers at risk of mental health problems and  
31 106 thus for timely interventions.  
32 107

33 108 There are many studies on mental health <sup>[25-26]</sup>; however, the results of previous studies lack consistency  
34 109 and mostly discuss factors influencing mental health, and most of them are single-center studies that  
35 110 focus on only certain aspects of mental health. Our study included common demographics, job burnout,  
36 111 occupational stress, chronic illness and occupational exposure factors to distinguish whether respondents  
37 112 suffered from mental health problems. In addition, there is a small body of literature that develops and  
38 113 validates a risk nomogram between depression and suicide to support timely intervention by clinicians.  
39 114 And the sample sizes of the two relevant studies were small, 474 and 273 depressed patients respectively  
40 115 <sup>[27-28]</sup>. Today, there is increasing recognition of the important role of mental health in achieving global  
41 116 development goals, and WHO has included mental health in the Sustainable Development Goals.  
42 117 However, there are no relevant studies that have used objective indicators for factory workers and miners

1  
2  
3 118 to form a nomogram to predict mental health. Therefore, to bridge this gap in the literature and provide  
4 119 additional information for the prevention of mental health problems, we conducted a multicenter  
5 120 investigation to develop and validate an easy-to-use nomogram that combines objective information on  
6 121 demographics, job burnout, occupational stress and occupational hazards to comprehensively and  
7 122 accurately predict the prevalence of mental health problems among factory workers and miners.  
8  
9

10 123

## 11 124 **2. Materials and Methods**

12 125

### 13 126 **2.1 Calculation of sample size**

14 127

15 128 The sample size formula for the present illness rate survey,  $n = \frac{z_{\alpha/2}^2 \times pq}{\delta^2}$ , p is the present-hazard rate,  
16 129  $q=1-p$ ,  $\delta$  is the tolerance error, generally taken as 0.1p,  $z_{\alpha/2}$  is the significance test statistic,  $z_{\alpha/2}=1.96$   
17 130 for  $\alpha=0.05$ , then the formula is calculated as,  $n = 400 \times \frac{q}{p}$ . A cross-sectional study in Xinjiang showed  
18 131 that 38.27% of factory workers and miners had mental health problems [29]. And a study revealed that  
19 132 633 out of 1675 coal miners (37.8%) suffered from mental disorders between August 2018 and June  
20 133 2019[30]. In this study, we assumed a 30% prevalence of mental health problem to obtain the maximum  
21 134 required sample size. which would calculate a sample size of 934, taking into account non-response and  
22 135 a 20% loss of questionnaires, which would require approximately 1168 people.  
23  
24

25 136

### 26 137 **2.2. Participants**

27 138

28 139 Participants in this cross-sectional survey were factory workers and mines in the Urumqi region, and the  
29 140 survey covered all districts and counties in the Urumqi region to avoid selection bias as far as possible.  
30 141 Specifically, this survey was conducted by means of whole-group random sampling from January to May  
31 142 2019, and a total of 202 enterprises were selected, including 21 in Tianshan District, 30 in Shaibak  
32 143 District, 24 in Xinshi District, 22 in Shuimogou District, 56 in Jingkai District, 37 in Midong District, 9  
33 144 enterprises in Dabancheng District and 3 enterprises in Urumqi County.  
34  
35

36 145

37 146 The inclusion criteria were as follows: (1) workers working in mining enterprises or factories in Urumqi;  
38 147 (2) workers with a history of working for more than one year; (3) Workers with no history of mental  
39 148 illness and no history of taking psychotropic drugs.  
40  
41

42 149

43 150 The exclusion criteria were the following: (1) factory workers and miners in non-Urumqi area; (2)  
44 151 working history of factories and mining enterprises less than 1 year; (3) a confirmed diagnosis of a mental  
45 152 health problem and a history of treatment and use of psychotropic medication; (4) Questionnaires with  
46 153 missing data were excluded.  
47  
48

49 154

50 155 An online electronic questionnaire was created using the Questionnaire Star platform to collect data. The  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 156 survey was conducted by trained surveyors who explained the purpose, meaning, content and  
4 157 requirements of the questionnaire to all participants and provided on-site instructions to ensure the return  
5 158 rate of the questionnaire. All participants understood the purpose of the study and were willing to  
6 159 participate in the study. A total of 7,500 questionnaires were distributed and 7,315 questionnaires were  
7 160 returned, representing a return rate of 97.5%. After checking the validity and integrity of the  
8 161 questionnaires, 7,118 questionnaires were confirmed as valid, with an effective rate of 97.3%. A total of  
9 162 7,118 participants met the inclusion criteria and the data were randomly divided into a training group  
10 163 (n=4,955) and a validation group (n=2,163) (Figure 1).  
11  
12  
13  
14  
15

## 16 165 **2.3. Research Methods**

### 17 166 18 167 **2.3.1. Assessment of mental health**

19 168  
20 169 The SCL-90, designed by Derogatis and his colleagues, was widely used in the mental health field <sup>[31]</sup>,  
21 170 which contains 90 items across nine dimensions: somatization, obsessive-compulsive symptoms,  
22 171 interpersonal sensitivity, depression, anxiety, hostility, horror, bigotry and mental illness. The SCL-90  
23 172 has been used extensively in previous studies and has relatively high reliability and validity <sup>[32]</sup>. The  
24 173 questionnaire uses a Likert 5-point scale, with a score of 0 point indicating none and 4 points indicating  
25 174 severe. A total score above 160, a score above 2 on any item, or a positive item above 43 indicates the  
26 175 presence of a psychological abnormality <sup>[33]</sup>. In this survey, Cronbach  $\alpha$  was 0.99, the half-reliability  
27 176 coefficient was 0.98, and the KMO was 0.994.  
28  
29  
30  
31  
32  
33  
34

### 35 178 **2.3.2. Assessment of occupational stress**

36 179  
37 180 This survey evaluated occupational stress in factory workers and miners through the Effort–Reward  
38 181 Imbalance (ERI) model developed by Siegrist <sup>[34]</sup>. The ERI scale consists of three subscales: effort (E, 6  
39 182 items), reward (R, 11 items) and over commitment (6 items), for a total of 23 items. A Likert 5-level  
40 183 scoring method (1, "highly disagree" to 5, "highly agree") is used to grade the items in the questionnaire  
41 184 with the same weight for each item. The effort–return index  $ERI = E/R \times C$ , where C is the adjustment  
42 185 coefficient, and the value is 6/11. ERI values greater than 1, equal to 1, and less than 1 correspond to  
43 186 high pay–low return, pay–return balance, and low pay–high return, respectively. Moreover, the higher  
44 187 the ERI value, the greater the occupational stress <sup>[35]</sup>. In this survey, Cronbach  $\alpha$  was 0.94, the half-  
45 188 reliability coefficient was 0.93 and the KMO was 0.956.  
46  
47  
48  
49  
50  
51

### 52 190 **2.3.3. Assessment of job burnout**

53 191  
54 192 In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess  
55 193 job burnout, which has good reliability and validity <sup>[36]</sup>. CMBI is composed of 15 items in three  
56 194 dimensions: emotional exhaustion (5 items), depersonalization (5 items) and reduced personal  
57 195 accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete  
58  
59  
60



1  
2  
3 196 compliance and 7 points indicating complete non-compliance. According to the critical value (emotional  
4 197 exhaustion  $\geq 25$ , depersonalization  $\geq 11$ , personal achievement reduction  $\geq 16$ ), the levels of occupational  
5 198 burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to  
6 199 or above the critical value), moderate (any two aspects are equal to or higher than the critical values),  
7 200 and severe (three aspects are equal to or higher than the critical values) [37]. In this survey, Cronbach  $\alpha$   
8 201 was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.  
9 202

#### 10 203 **2.3.4. Candidate predictors**

11 204  
12 205 Trained investigators obtained information on demographics, job burnout, occupational stress, mental  
13 206 health and occupational exposure factors through on-site face-to-face collection of an electronic version  
14 207 of the questionnaire. Covariates included in this study: 1) demographic information: gender, ethnicity,  
15 208 education level, professional title, work schedule, marital status, monthly income, age, working years,  
16 209 labor contracts, working hours per day, and working hours per week; 2) occupational exposure factors:  
17 210 coal dust, silica dust, asbestos dust, benzene, lead, noise, and brucellosis; 3) questionnaires: ERI, CMBI;  
18 211 4) chronic diseases: diabetes, hypertension. Information on four areas, including demographic  
19 212 information, questionnaires, occupational hazards and chronic diseases, were filled in by participants  
20 213 through their own responses on the questionnaire star.  
21 214

22 215 Sex was defined as male or female; ethnicity was defined as Han and other; education level was defined  
23 216 as junior high school and below, high school, junior college or bachelor's degree or above; labor contracts  
24 217 was defined as signed or unsigned; professional title was defined as no, primary, middle or senior; work  
25 218 schedule was defined as day shift, night shift, shift or day and night shifts; marital status was defined as  
26 219 unmarried, married, divorced or widowed; monthly income (yuan) was defined as <3000, 3000~, 4000~,  
27 220 5000~, 6000~, 7000~ or 8000~; age (years) was defined as <25, 25~, 30~, 35~, 40~ or 45~; working  
28 221 years was defined as ~5, 5~, 10~, 15~, 20~, 25~ or 30~; working hours per day (hours) was defined as  
29 222  $\leq 7$  or  $> 7$ ; working days per week (days) was defined as  $\leq 5$  or  $> 5$ ; exposure to coal dust, silica dust,  
30 223 asbestos dust, benzene, lead, noise, brucellosis were all defined as yes or no; ERI was defined as yes or  
31 224 no; CMBI was defined as none, mild, moderate and severe; hypertension and diabetes were both defined  
32 225 as yes or no.  
33 226

#### 34 227 **2.4. Statistical analysis**

35 228  
36 229 Categorical variables were described as counts and percentages, and chi square test or Fisher exact test  
37 230 was used to compare categorical variables between groups. 70% of participants were randomly assigned  
38 231 to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute  
39 232 shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were  
40 233 used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which  
41 234 predictive models were constructed. A nomogram for predicting was generated according to the selected  
42 235 characteristics. In addition, forest plot was drawn to visually depict the P-value, OR and 95% CI for the  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

236 selected validations. Statistically significant predictors were applied to develop a prediction model for  
 237 the risk of mental health problems among factory workers and miners by introducing all selected factors  
 238 and analyzing the statistical significance levels of them. We used calibration plots and receiver operating  
 239 characteristic (ROC) curves to show the calibration and discrimination of our final model. Brier scores  
 240 for overall performance, calibration slopes were used to assess the predictable accuracy of the model.  
 241 Decision curve analysis (DCA) was applied to calculate the net benefit of the nomogram. Statistical  
 242 analysis was performed using the open-source R software Version 3.6.1 (<http://www.r-project.org>). The  
 243 significance level ( $\alpha$ ) set at 0.05.

244

### 245 3. Results

246

#### 247 3.1. Participant characteristics

248

249 A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training  
 250 group (n=4,955) and a validation group (n=2,163). Over half of all participants (65.31%) were male,  
 251 57.31% of the population was over 35 years of age and 78.32% of the subjects were married, showing  
 252 that factory workers and miners are generally older and most of them have spouses. The majority of them  
 253 had completed high school (83.94%), while a smaller percentage had completed undergraduate education  
 254 (22.98%), indicating that the group of factory workers and miners as a whole was not well educated. The  
 255 total number of workers (n, %) exposed to coal dust, silica dust, asbestos dust, benzene, lead, noise and  
 256 brucellosis in the factory and mining enterprises were 377 (5.3), 730 (10.3), 981 (14), 1,981 (27.8), 373  
 257 (5.2), 4,942 (69.4) and 121 (1.7) respectively, with the total number of workers exposed to noise  
 258 amounting to 4,942, or 69% of the total population surveyed. The demographic, job burnout,  
 259 occupational stress and occupational exposure factors for the training and validation groups are shown  
 260 in Table 1. The results showed that there were no significant statistical differences between the two  
 261 groups of characteristic variables, except for coal dust and CMBI, indicating that the baseline levels were  
 262 largely consistent between the two groups.

Table 1 Characteristics of the study participants

Variables	Total (n = 7118)	train (n = 4955)	test (n = 2163)	<i>p</i>
Sex, n (%)				
Male	4649 (65.3)	3216 (64.9)	1433 (66.3)	0.284
Female	2469 (34.7)	1739 (35.1)	730 (33.7)	
Ethnicity, n (%)				
Han	5762 (80.9)	3982 (80.4)	1780 (82.3)	0.061
Other	1356 (19.1)	973 (19.6)	383 (17.7)	
Education level, n (%)				
Junior high school and below	1143 (16.1)	804 (16.2)	339 (15.7)	0.765
High school	1406 (19.8)	988 (19.9)	418 (19.3)	
Junior college	2933 (41.2)	2038 (41.1)	895 (41.4)	

1					
2					
3	Bachelor's degree or above	1636 (23.0)	1125 (22.7)	511 (23.6)	
4	Professional title, n (%)				
5					
6	None	2854 (40.1)	1983 (40.0)	871 (40.3)	0.923
7	Primary	1644 (23.1)	1149 (23.2)	495 (22.9)	
8	Middle	1618 (22.7)	1133 (22.9)	485 (22.4)	
9	Senior	1002 (14.1)	690 (13.9)	312 (14.4)	
10					
11	Work schedule, n (%)				
12					
13	Day shift	3986 (56.0)	2801 (56.5)	1185 (54.8)	0.585
14	Night shift	270 (3.8)	187 (3.8)	83 (3.8)	
15	Shift	2058 (28.9)	1412 (28.5)	646 (29.9)	
16	Day and night shifts	804 (11.3)	555 (11.2)	249 (11.5)	
17					
18	Marital status, n (%)				
19					
20	Unmarried	1104 (15.5)	762 (15.4)	342 (15.8)	0.218
21	Married	5575 (78.3)	3906 (78.8)	1669 (77.2)	
22	Divorced	390 (5.5)	255 (5.1)	135 (6.2)	
23	Widowed	49 (0.7)	32 (0.6)	17 (0.8)	
24					
25	Monthly income (yuan), n (%)				
26					
27	<3000	1799 (25.3)	1246 (25.1)	553 (25.6)	0.966
28	3000~	2418 (34.0)	1682 (33.9)	736 (34.0)	
29	4000~	1600 (22.5)	1125 (22.7)	475 (22.0)	
30	5000~	752 (10.6)	520 (10.5)	232 (10.7)	
31	6000~	288 (4.0)	201 (4.1)	87 (4.0)	
32	7000~	148 (2.1)	106 (2.1)	42 (1.9)	
33	8000~	113 (1.6)	75 (1.5)	38 (1.8)	
34					
35	Age (years), n (%)				
36					
37	<25	431 (6.1)	297 (6.0)	134 (6.2)	0.173
38	25~	786 (11.0)	519 (10.5)	267 (12.3)	
39	30~	956 (13.4)	684 (13.8)	272 (12.6)	
40	35~	866 (12.2)	617 (12.5)	249 (11.5)	
41	40~	849 (11.9)	588 (11.9)	261 (12.1)	
42	45~	3230 (45.4)	2250 (45.4)	980 (45.3)	
43					
44	Working years (years), n (%)				
45					
46	<5	1170 (16.4)	794 (16.0)	376 (17.4)	0.248
47	5~	1065 (15.0)	736 (14.9)	329 (15.2)	
48	10~	997 (14.0)	721 (14.6)	276 (12.8)	
49	15~	389 (5.5)	273 (5.5)	116 (5.4)	
50	20~	763 (10.7)	538 (10.9)	225 (10.4)	
51	25~	1293 (18.2)	878 (17.7)	415 (19.2)	
52	30~	1441 (20.2)	1015 (20.5)	426 (19.7)	
53					
54					
55					
56					
57					
58					
59					
60					

1					
2					
3	Labor contracts, n (%)				
4	Signed	6641 (93.3)	4624 (93.3)	2017 (93.3)	0.955
5	Unsigned	477 (6.7)	331 (6.7)	146 (6.7)	
6					
7	Working hours per day (hours), n (%)				
8	≤7	1161 (16.3)	814 (16.4)	347 (16.0)	0.712
9	>7	5957 (83.7)	4141 (83.6)	1816 (84.0)	
10					
11	Working days per week (days), n (%)				
12	≤5	4442 (62.4)	3107 (62.7)	1335 (61.7)	0.446
13	>5	2676 (37.6)	1848 (37.3)	828 (38.3)	
14					
15	Diabetes, n (%)				
16	Yes	429 (6.0)	298 (6.0)	131 (6.1)	0.988
17	No	6689 (94.0)	4657 (94.0)	2032 (93.9)	
18					
19	Hypertension, n (%)				
20	Yes	1330 (18.7)	929 (18.7)	401 (18.5)	0.861
21	No	5788 (81.3)	4026 (81.3)	1762 (81.5)	
22					
23	Coal dust, n (%)				
24	Yes	377 (5.3)	244 (4.9)	133 (6.1)	0.039
25	No	6741 (94.7)	4711 (95.1)	2030 (93.9)	
26					
27	Silica dust, n (%)				
28	Yes	730 (10.3)	523 (10.6)	207 (9.6)	0.223
29	No	6388 (89.7)	4432 (89.4)	1956 (90.4)	
30					
31	Asbestos dust, n (%)				
32	Yes	981 (13.8)	691 (13.9)	290 (13.4)	0.570
33	No	6137 (86.2)	4264 (86.1)	1873 (86.6)	
34					
35	Benzene, n (%)				
36	Yes	1981 (27.8)	1360 (27.4)	621 (28.7)	0.287
37	No	5137 (72.2)	3595 (72.6)	1542 (71.3)	
38					
39	Lead, n (%)				
40	Yes	373 (5.2)	246 (5.0)	127 (5.9)	0.128
41	No	6745 (94.8)	4709 (95.0)	2036 (94.1)	
42					
43	Noise, n (%)				
44	Yes	4942 (69.4)	3420 (69.0)	1522 (70.4)	0.270
45	No	2176 (30.6)	1535 (31.0)	641 (29.6)	
46					
47	Brucellosis, n (%)				
48	Yes	121 (1.7)	86 (1.7)	35 (1.6)	0.800
49	No	6997 (98.3)	4869 (98.3)	2128 (98.4)	
50					
51	ERI, n (%)				
52	Yes	3147 (44.2)	2173 (43.9)	974 (45.0)	0.372
53	No	3971 (55.8)	2782 (56.1)	1189 (55.0)	
54					
55					
56					
57					
58					
59					
60					

CMBI, n (%)					
No	959 (13.5)	674 (13.6)	285 (13.2)	0.033	
Mild	2667 (37.5)	1813 (36.6)	854 (39.5)		
Moderate	2900 (40.7)	2031 (41.0)	869 (40.2)		
Severe	592 (8.3)	437 (8.8)	155 (7.2)		

263

### 264 3.2. Feature selection

265

266 The lambda was smallest at 0.01801 as seen from the lasso results when there were 12 characteristics,  
 267 which were education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working  
 268 hours per day, working years, marital status, and work schedule based on the results of the questionnaires  
 269 on demographics, occupational stress, job burnout and occupational exposure factors (Figure 2).

270

### 271 3.3. Results of logistic regression model

272

273 The 12 features obtained from the LASSO regression were incorporated into a multivariate logistic  
 274 regression model and the regression results were shown in Table 2. It was clear from the results that  
 275 education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per  
 276 day, working years, marital status, and work schedule were independent determinants of risk for mental  
 277 health problems. In addition, there was no evidence of multicollinearity between the covariates included  
 278 in the model. The forest plot showed that the selected 12 features all contain items with  $P < 0.05$ , among  
 279 which the degree of severe of CMBI (OR, 19.84; 95% CI, 13.88-28.34;  $p < 0.001$ ) had the greatest impact  
 280 on the risk of mental health problems among factory workers and miners (Figure 3).

281

Table 2 Predictive factors of risk for mental health problems among factory workers and miners

Variable	$\beta$	S.E.	OR(CI95%)	Wald	P	VIF
Intercept	-2.33	0.25	0.10(0.06,0.16)	-9.357	0	-
Education level						
Junior school and below VS High school	0.34	0.13	1.41(1.10,1.81)	2.727	0.006**	2.28
Junior school and below VS Junior college	0.44	0.11	1.56(1.24,1.95)	3.850	< 0.001***	2.79
Junior school and below VS Bachelor's degree or above	0.38	0.13	1.46(1.13,1.87)	2.953	0.003**	2.51
Professional title						
None VS Primary	0.15	0.09	1.16(0.97,1.39)	1.582	0.114	1.35
None VS Middle	0.05	0.09	1.05(0.87,1.26)	0.519	0.604	1.34
None VS Senior	0.27	0.11	1.30(1.06,1.61)	2.458	0.014*	1.32
Work schedule						
Day and night shifts VS Day shift	-0.38	0.11	0.69(0.55,0.85)	-3.364	0.001**	2.70
Day and night shifts VS Night shift	0.01	0.20	1.01(0.68,1.49)	0.044	0.965	1.30

Day and night shifts VS Shift	0.01	0.12	1.01(0.81,1.27)	0.107	0.915	2.47
Marital status						
Unmarried VS Married	0.16	0.13	1.18(0.91,1.52)	1.263	0.206	2.29
Unmarried VS Divorced	0.55	0.19	1.73(1.20,2.51)	2.918	0.004**	1.69
Unmarried VS Widowed	0.69	0.43	1.99(0.85,4.64)	1.586	0.113	1.09
Age						
~25 VS 25~	-0.02	0.20	0.98(0.66,1.47)	-0.083	0.934	3.09
~25 VS 30~	-0.02	0.22	0.98(0.64,1.50)	-0.090	0.929	4.79
~25 VS 35~	0.56	0.23	1.76(1.13,2.74)	2.503	0.012*	5.01
~25 VS 40~	0.33	0.23	1.39(0.88,2.21)	1.419	0.156	4.97
~25 VS 45~	0.23	0.22	1.26(0.81,1.95)	1.018	0.308	10.93
Working years						
~5 VS 5~	0.44	0.14	1.55(1.18,2.05)	3.114	0.002**	2.27
~5 VS 10~	0.06	0.15	1.06(0.78,1.43)	0.366	0.714	2.48
~5 VS 15~	0.06	0.20	1.06(0.72,1.56)	0.305	0.760	1.79
~5 VS 20~	0.29	0.18	1.33(0.95,1.88)	1.641	0.101	2.65
~5 VS 25~	0.48	0.17	1.61(1.15,2.25)	2.782	0.005**	3.99
~5 VS 30~	0.20	0.16	1.22(0.89,1.68)	1.239	0.216	3.90
Working hours per day						
≤7 VS >7	-0.50	0.09	0.61(0.50,0.73)	-5.363	< 0.001***	1.15
Diabetes						
No VS Yes	0.43	0.14	1.53(1.16,2.03)	2.974	0.003**	1.05
Hypertension						
No VS Yes	0.52	0.09	1.69(1.42,2.00)	5.885	< 0.001***	1.11
Asbestos dust						
No VS Yes	0.44	0.10	1.55(1.28,1.87)	4.474	< 0.001***	1.03
ERI						
No VS Yes	0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***	1.05
CMBI						
No VS Mild	0.26	0.12	1.30(1.03,1.64)	2.175	0.003**	2.73
No VS Moderate	1.30	0.11	3.67(2.93,4.59)	11.361	< 0.001***	2.83
No VS Severe	2.99	0.18	19.84(13.88,28.34)	16.41	< 0.001***	1.44

282 Note:  $\beta$  is the regression coefficient. “\*\*\*” indicates  $P<0.001$ , “\*\*” indicates  $P<0.01$ , “\*” indicates  $P<0.05$ .

283

### 284 3.4. Development of an individualized prediction model

285

286 Based on the results of the multivariate analysis, predictors such as education, professional title, age,

1  
2  
3 287 CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working years, marital status,  
4 288 and work schedule were included in the nomogram. A model incorporating the above independent  
5 289 predictors was developed and represented as a nomogram in Figure 4. Each variable in nomogram was  
6 290 assigned a score, and the cumulative sum of each 'point' was the 'total score'. The "total score"  
7 291 corresponded to the "predictable likelihood", which was the predicted probability of mental health  
8 292 problems among factory workers and miners as suggested by our design of the nomogram.  
9 293

10 294 As an example of the use of nomogram: a randomly selected sample from the training group, one with  
11 295 no professional title, day shift, no diabetes or hypertension, Junior college, <5 of working years, >7 of  
12 296 working hours per day, married, no exposed to asbestos dust, <25 years of age, no ERI, mild of CMBI,  
13 297 with a calculated total score of 174 and a corresponding risk probability of 8.27% for mental health  
14 298 problems.  
15 299

### 16 300 **3.5 The validation of calibration**

17 301  
18 302 Model validation was carried out in the validation group. The prediction accuracy of the model was  
19 303 assessed by two aspects. (1) The Brier score for overall performance, which assessed the difference  
20 304 between observed and predicted values, with values closer to 0 indicating better predictive ability. (2)  
21 305 The calibration slope used for modal calibration, which assessed the agreement between the observed  
22 306 and predicted values, with values closer to 1 indicating better performance. The accuracy measurements  
23 307 for the bias correction were validated by the model with a Brier score of 0.176 and a calibration slope of  
24 308 0.970, respectively (Figure 5). The prediction accuracy of the model was relatively high.  
25 309

### 26 310 **3.6 The validation of discrimination**

27 311  
28 312 ROC was plotted for the training and validation groups, and the AUC of training and the verification  
29 313 groups were 0.785 and 0.784, respectively (Figure 6). The AUC of training and the verification groups  
30 314 were both greater than 0.75, showing a good discrimination.  
31 315

### 32 316 **3.7 Decision Curve Analysis**

33 317  
34 318 As shown in the DCA of the risk of mental health problems nomogram in Figure 7, the model for  
35 319 predicting the risk of mental health problems for factory workers and miners in this study was more  
36 320 practically relevant if the threshold probability of patients was >10%.  
37 321

## 38 322 **4. Discussion**

39 323  
40 324 To our knowledge, this is the first study to develop an easy-to-use nomogram to predict the mental health  
41 325 risks of factory workers and miners. The nomogram developed using the training set data contain 12  
42 326 items for education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working  
43 327

1  
2  
3 327 hours per day, working years, marital status, and work schedule. In addition, validation has shown that  
4 328 nomogram model has good accuracy and discriminatory power. Our novel nomogram can be used in any  
5 329 setting to provide a rapid assessment of mental health risks and to help identify patients with mental  
6 330 health risks, saving time compared to previous mental health investigations and improving on the lack  
7 331 of entries in previous investigations related to the specific working environment of factory workers and  
8 332 miners. The AUC of training group and verification group were 0.785 and 0.784 respectively, showing  
9 333 moderate discriminatory and calibration power.  
10 334

11 335 A review of the literature found that the vast majority of studies constructed nomograms to predict  
12 336 clinical disorders, with less literature used to predict psychological problems. In a study to predict the  
13 337 correlates of suicide attempts in a Chinese population with major depressive disorder, the C-index was  
14 338 0.715 and the C-index in the internal validation set was 0.703, and the calibration curve of the column  
15 339 line plot also showed good agreement between the predicted and observed risk of suicide attempts. The  
16 340 variables in the nomogram included socio-demographic information and clinical variables including age,  
17 341 duration, number of episodes, age at onset, number of hospitalizations, characteristics of anxiety and  
18 342 psychiatric symptoms, marital status, income, education level and employment status [27]. In another  
19 343 study that created a nomogram to predict the risk of psychosocial and behavioral problems in children  
20 344 and adolescents during the COVID-19 pandemic, the C index exceeded 0.800 and the calibration curve  
21 345 also showed good predictive accuracy. The variables covered three subject areas, namely demographic  
22 346 information, the psychosocial impact of the epidemic such as homework time and sedentary time, and  
23 347 the Child Behaviour Checklist score (CBCL) for the evaluation of psychological problems [38]. In this  
24 348 study, 7,118 participants were randomly divided into a training group (n=4,955) and a validation group  
25 349 (n=2,163) in a ratio of 3:1, involving a total of 23 features, and 12 features were selected by LASSO  
26 350 regression. The nomogram could be a useful tool to better identify patients with mental health problems,  
27 351 as it not only covered comprehensive information, including demographic information, job burnout,  
28 352 occupational stress, chronic diseases and occupational exposure factors closely related to factory workers  
29 353 and miners, but also was simple to operate and easy to use. In the validation group the Brier score was  
30 354 0.176, the calibration slope was 0.970 and the calibration curve of nomogram showed a good fit. The  
31 355 AUC of training group and verification group were 0.785 and 0.784 respectively. Compared to the two  
32 356 studies above, our nomogram showed good accuracy and discrimination, and more comprehensive  
33 357 coverage in this nomogram model. Therefore, the possibility of early intervention for patients with high-  
34 358 risk mental health problems will be increased by covering multiple information and easy to use  
35 359 nomogram modal, especially for factory workers and miners with poor working conditions, relatively  
36 360 low levels of education and low patience.

37 361  
38 362 Mental health problems were very common in the group of factory workers and miners, and the  
39 363 prevalence of mental health of them was found to be 37.08% in our study. Notably, the CMBI showed  
40 364 the most significant score (score = 100) and the ERI also had a high score (score = 43) in mental health  
41 365 problem incidence risk nomogram, which indicated that both of them were relatively important factors  
42 366 for mental health problems among the group of factory workers and miners. Our finding was consistent



1  
2  
3 367 with other studies that had shown that occupational stress was a significant predictor of anxiety and was  
4 368 negatively associated with mental health. In addition, there is a high correlation between burnout and  
5 369 depression [39].  
6  
7 370

8  
9 371 In line with previous studies, working years was also an important influential factor in this study. Related  
10 372 study has shown that employment could improve patients' mental health, while unemployment could  
11 373 lead to a deterioration in mental health [40]. In China, workers' working years is an important aspect of  
12 374 employment, and researchers have studied this aspect and found that precarious employment is a source  
13 375 of stress for individuals and predisposes them to mental health problems [41]. In addition, environmental  
14 376 factors were also one of the influential factors of mental health problems in our study. Relevant studies  
15 377 have found that exposure to air pollution is associated with increased suicide risk and depressive  
16 378 symptoms [42]. Hypertension and diabetes were the influential factors in this study. A study has shown  
17 379 that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher  
18 380 than in the non-diabetic population [43]. Furthermore, there is a recognized association between  
19 381 hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear  
20 382 [44].  
21  
22 383

23  
24  
25  
26 384 Factory workers and miners were inevitably exposed to occupational hazards such as benzene and  
27 385 asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers  
28 386 are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous  
29 387 enterprises, involving more than 30 different types of operations, of which factory workers and miners  
30 388 is the one [45]. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for  
31 389 mental health problems in this study. Our study found that the work schedules of factory workers and  
32 390 miners were vary and the phenomenon of night shifts was very common, which inevitably affected their  
33 391 normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health  
34 392 and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-  
35 393 regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could  
36 394 suffer from negative emotions such as anxiety and depression [46]. Professional title and education level  
37 395 were also important influences on mental health issues. In the workplace, generally speaking, the higher  
38 396 the professional title and education level, the higher the status of the worker in the company and the  
39 397 greater the role played in the position. The number of studies on socio-economic status and mental health  
40 398 had increased in recent years. Some of these studies have shown that major depression is higher in the  
41 399 low socio-economic status group [47]. It has also been suggested that education itself is the best indicator  
42 400 of socio-economic status [48]. Marital status was one of the influential factors for mental health problems.  
43 401 Many studies have found an association between mental health and gender, marital status, lifestyle and  
44 402 working conditions, and it has been shown that poor mental health in women is associated with divorce  
45 403 or widowhood [49]. In this study, working more than seven hours a day was a determinant factor on mental  
46 404 health problems, which was consistent with other studies that had shown that long working hours could  
47 405 have a negative impact on employees' mental health and that excessive workloads could increase workers'  
48 406 fatigue, which in turn could lead to anxiety and depression [50].  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 407

4 408 In China, there are many problems in identifying people with mental health problems due to uneven and  
5 409 imperfect levels of medical development across regions. Some studies have shown that in mainland  
6 410 China, general practitioners, surgeons and primary health care workers often have little or no mental  
7 411 health training, which prevents them from providing basic mental health services<sup>[51]</sup>. Non-mental health  
8 412 professionals in general hospitals learn about mental illness on their own, rather than learning about it  
9 413 during their formal education<sup>[52]</sup>. Therefore, this study designed a simple and comprehensive nomogram  
10 414 to address the issue of timely detection and effective interventions for people with mental health problems,  
11 415 so that people at risk of mental health problems could easily calculate their probability of suffering from  
12 416 mental health problems without the help of medical staff. This study has several strengths. First, to our  
13 417 knowledge, this is the first model to develop and assess the likelihood of mental health problems in a  
14 418 group of factory workers and miners. Secondly, the nomogram in this study includes demographic  
15 419 information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors  
16 420 that are closely related to factory workers and miners, allowing for a more accurate assessment of the  
17 421 risk of morbidity among them, as well as providing a methodological reference for other related studies.  
18 422

19 423

### 20 424 **5. Limitations**

21 425

22 426 This study also has several limitations. Firstly, we have considered many influential factors including  
23 427 demographics, job burnout, occupational stress and occupational exposure factors, but we are still not  
24 428 certain whether all possible influences are covered. Secondly, while the robustness of our nomogram was  
25 429 extensively validated internally in the same population, external validation is lacking for other  
26 430 populations in other regions and countries. Nomogram need to be externally assessed in a wider  
27 431 population.

28 432

### 29 433 **Patient and public involvement**

30 434

31 435 Neither patients nor members of the public had any involvement in the design of this study.

32 436

33 437 **Acknowledgements** The authors are grateful to all participants and investigators.

34 438

35 439 **Contributions** Y.L., Q.L., and T.L. are responsible for conceptualization; Y.L. is responsible for  
36 440 methodology, software, formal analysis, resources, and visualization; Q.L. and T.L. are responsible for  
37 441 the original draft preparation; Q.L. and H.Y. are responsible for reviewing; Q.L. is responsible for editing;  
38 442 T.L. is responsible for supervision. Yaoqin Lu and Qi Liu contributed equally to this work.

39 443

40 444 **Funding** This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous  
41 445 Region (grant number 2020D01A27), the Postgraduate Innovation Project of Xinjiang Uyghur  
42 446 Autonomous Region (grant number XJ2021G215), the Outstanding Young Scientist Training Program  
43 of Urumqi Science and Technology Talent Project (grant number N/A), the Public Health and Preventive  
44 Medicine, the 13th Five-Year Plan Key Subject of Xinjiang Uygur Autonomous Region (grant number

45 447

1  
2  
3 447 N/A), The funders were not involved in the conception, design, analysis or interpretation of this study.  
4  
5 448

6 449 **Competing interests** None declared.  
7  
8 450

9 451 **Patient consent for publication** Not applicable.  
10  
11 452

12 453 **Ethics approval** The study was approved by the ethics committee of Urumqi Center for Disease Control  
13 and Prevention (20181123)  
14 454  
15 455

16 456 **Data availability statement** Data are available on reasonable request. The data used in this study are  
17 available from the corresponding authors on reasonable request.  
18 457  
19 458

## 20 459 **References**

- 21 460 [1] WHO Terminology Information System [online glossary] [http://www.who.int/health-systems-](http://www.who.int/health-systems-performance/docs/glossary.html)  
22 461 [performance/docs/glossary.html](http://www.who.int/health-systems-performance/docs/glossary.html).  
23 462 [2] Wang Y, Liu X, Qiu J, Wang H, Liu D, Zhao Z, Song M, Song Q, Wang X, Zhou Y, Wang W.  
24 463 Association between ideal cardiovascular health metrics and suboptimal health status in Chinese  
25 464 population. *Sci Rep* 2017;7:14975.  
26 465 [3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental  
27 466 health. *Lancet*. 2007;370:859–77.  
28 467 [4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive  
29 468 symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. *Health*  
30 469 *Place* 2016;42:37–46.  
31 470 [5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,  
32 471 Degenhardt L, Dicker D, *et al*. Global, regional, and national incidence, prevalence, and years lived  
33 472 with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A  
34 473 systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;386:743–800.  
35 474 [6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health? *Int J*  
36 475 *Environ Res Public Health* 2020;17:7530.  
37 476 [7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative  
38 477 income, gender, and location. *China Econ Rev* 2018;48:83–101.  
39 478 [8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:  
40 479 Results from the PATH Through Life Survey. *Soc Sci Med*. 2009;69:229–237.  
41 480 [9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and  
42 481 Subjective Well-being? *J Environ Econ Manag* 2017;85:81–94.  
43 482 [10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. *World*  
44 483 *Psychiatry*. 2015;14:36–42.  
45 484 [11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson  
46 485 SJ. IPS Supported employment: a review. *J Vocat Rehabil* 2014;40:1–13.  
47 486 [12] Jia Z, *et al*. Exposure to ambient air particles increases the risk of mental disorder: findings from a  
48 487 natural experiment in Beijing. *Int J Environ Res Public Health* 2018;15:160.  
49 488 [13] Szyszkwicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency  
50 489 department visits for suicide attempts in vancouver, Canada. *Environ Health Insights* 2010;4:79–86.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 490 [14] Michael J. S. International classification of sleep disorders-third edition : highlights and  
4 491 modifications. *Chest* 2014;146:1387–1394.
- 6 492 [15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial  
7 493 infarction: gender differences. *Open Nurs J* 2018;12:205–214.
- 8 494 [16] Wurm W, Vogel K, Holl A, *et al.* Depression-burnout overlap in physicians. *PLoS One*  
9 495 2016;11:e0149913.
- 11 496 [17] Porru F, Robroek S J W, Bültmann U, *et al.* Mental health among university students: The  
12 497 associations of effort-reward imbalance and overcommitment with psychological distress. *J Affect*  
13 498 *Disord* 2021;282:953-961.
- 15 499 [18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,  
16 500 and burnout among physician assistant students: a multi-institutional study. *Health Professions Edu*  
17 501 2020;6:420–427.
- 19 502 [19] Hua D, Kong Y, Li W, Han Q, Zhang X, Zhu LX, Wan SW, Liu Z, Shen Q, Yang J, He HG, Zhu J.  
20 503 Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during  
21 504 the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. *EClinical Med*  
22 505 2020;24:100424.
- 24 506 [20] Armah E K, Adedeji J A, Boafu B B, *et al.* Underground gold miner exposure to noise, diesel  
25 507 particulate matter and crystalline silica dust. *J Health Pollut* 2021; 11(29):210301.
- 27 508 [21] Hall N B, Blackley D J, Halldin C N, *et al.* Current review of pneumoconiosis among US coal  
28 509 miners. *Curr Environ Health Rep*, 2019; 6(3):137-147.
- 29 510 [22] Wickramatillake B A, Fernando M A, Frank A L. Prevalence of asbestos-related disease among  
30 511 workers in sri lanka. *Ann Glob Health*, 2019, 85(1):108.
- 32 512 [23] Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP. The estimation of the number of  
33 513 underground coal miners and the annual dose to coal miners in China. *Health Phys* 2007;93:127–  
34 514 132.
- 36 515 [24] Yong X, Gao X, Zhang Z, *et al.* Associations of occupational stress with job burn-out, depression  
37 516 and hypertension in coal miners of Xinjiang, China: A cross-sectional study. *BMJ open* 2020;10:  
38 517 e036087.
- 40 518 [25] P. Bech, J. Bille, S. B. Møller, L. C. Hellström, S. D. Østergaard. Psychometric validation of the  
41 519 Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal  
42 520 sensitivity. *J Affect Disord* 2014;160:98–103.
- 44 521 [26] J. Zhang, X. Zhang. Chinese college students' SCL-90 scores and their relations to the college  
45 522 performance. *Asian J Psychiatr* 2013;6:134–140.
- 46 523 [27] Liang S, Zhang J, Zhao Q, *et al.* Incidence trends and risk prediction nomogram for suicidal attempts  
47 524 in patients with major depressive disorder. *Front Psychiatry*, 2021, 12: 644038.
- 49 525 [28] Kan S K, Chen N N, Zhang Y L. Predicting the risk of suicide attempt in a depressed population:  
50 526 Development and assessment of an efficient predictive nomogram. *Psychiatry Res*, 2022, 310:  
51 527 114436.
- 53 528 [29] Lu Y, Zhang Z, Yan H, *et al.* Effects of occupational hazards on job stress and mental health of  
54 529 factory workers and miners: A propensity score analysis. *BioMed Res Int*, 2020, 2020: 1754897.
- 56 530 [30] Li X, Jiang T, Sun X, *et al.* The relationship between occupational stress, musculoskeletal disorders  
57 531 and the mental health of coal miners: The interaction between bdnf gene, tph2 gene polymorphism  
58 532 and the environment. *J Psychiatr Res*, 2021, 135: 76-85.
- 59 533 [31] Derogatis L, Lipman R.S, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary

- 1  
2  
3 534 report. *Psychopharmacol Bull* 1973;9:13–28.
- 4 535 [32] Crespo-Maraver M, Doval E, Fernã N.J, Gimã Nez-Salinas J, Prat G, Bonet P. Caregiver’s health:  
5 536 adaption and validation in a Spanish population of the Experience of Caregiving Inventory  
6 537 (ECI) *Gac. Sanit* 2018;33:348–355.
- 7 538 [33] Dang W, Xu Y, Ji J, *et al.* Study of the scl-90 scale and changes in the chinese norms. *Front*  
8 539 *Psychiatry*, 2020, 11: 524395.
- 9 540 [34] Siegrist J. Adverse health effects of high-effort/low-reward conditions. *J Occup. Health Psychol.*  
10 541 1996;1:27–41.
- 11 542 [35] Siegrist J, Wege N, Pühlhofer F, Wahrendorf M. A short generic measure of work stress in the era  
12 543 of globalization: Effort-reward imbalance. *Int Arch Occup Environ Health* 2009;82:1005–1013.
- 13 544 [36] Zhang Z, Lu Y, Yong X, *et al.* Effects of occupational radiation exposure on job stress and job  
14 545 burnout of medical staff in xinjiang, China: A cross-sectional study. *Med Sci Monit* 2020;26:  
15 546 e927848.
- 16 547 [37] Freudenberger HJ. Staff burnout. *J Soc Issues* 1974;30:159–165.
- 17 548 [38] Wang L, Chen L, Jia F, *et al.* Risk factors and prediction nomogram model for psychosocial and  
18 549 behavioural problems among children and adolescents during the covid-19 pandemic: A national  
19 550 multicentre study: Risk factors of childhood psychosocial problems. *J Affect Disord*, 2021, 294:  
20 551 128-136.
- 21 552 [39] Occupational Stress and Employees Complete Mental Health: A Cross-Cultural Empirical Study
- 22 553 [40] Knapp M and Wong G. Economics and mental health: the current scenario. *World Psychiatry*  
23 554 2020;19:3–14.
- 24 555 [41] Benach, J, Vives, A, Amable, M, Vanroelen, C, Tarafa, G, & Muntaner, C. Precarious employment:  
25 556 understanding an emerging social determinant of health. *Annu Rev Public Health* 2014;35:229-53.
- 26 557 [42] Bakian A. V. *et al.* Acute air pollution exposure and risk of suicide completion. *Am J Epidemiol*  
27 558 2015;181:295–303.
- 28 559 [43] Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type  
29 560 1 diabetes: systematic literature review. *Diabet Med* 2006;23:445–448.
- 30 561 [44] Gilsanz P, Karter AJ, Beerl MS, Quesenberry CP, Whitmer RA. The bidirectional association  
31 562 between depression and severe hypoglycemic and hyperglycemic events in type 1  
32 563 diabetes. *Diabetes Care* 2018;41:446–452.
- 33 564 [45] Lu Y, Zhang Z, Yan H, *et al.* Effects of occupational hazards on job stress and mental health of  
34 565 factory workers and miners: A propensity score analysis. *BioMed Res Int* 2020;2020:1754897.
- 35 566 [46] Shi L, Liu Y, Jiang T, *et al.* Relationship between mental health, the clock gene, and sleep quality  
36 567 in surgical nurses: A cross-sectional study. *BioMed Res Int* 2020;2020:4795763.
- 37 568 [47] Sallis JF, Saelens BE, Frank LD *et al.* Neighborhood built environment and income: examining  
38 569 multiple health outcomes. *Soc Sci Med* 2009;68:1285–1293.
- 39 570 [48] Winkleby MA, Jatulis DE, Frank E *et al.* Socioeconomic status and health: how education, income,  
40 571 and occupation contribute to risk-factors for cardiovascular disease. *Am J Public*  
41 572 *Health* 1992;82:816–820.
- 42 573 [49] Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis P.N, Mavreas V. Prevalence  
43 574 and sociodemographic associations of common mental disorders in a nationally representative  
44 575 sample of the general population of Greece. *BMC Psychiatry* 2013;13:163.
- 45 576 [50] Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG *et al.* Long  
46 577 working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II

578 study. *Psychol Med* 2011;41:2485–2494.

579 [51] Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, *et al.* Prevalence, treatment, and associated  
580 disability of mental disorders in four provinces in China during 2001–05: an epidemiological  
581 survey. *Lancet* 2009;373:2041–2053.

582 [52] Wu Q, Luo X, Chen S, *et al.* Mental health literacy survey of non-mental health professionals in six  
583 general hospitals in hunan province of China. *PloS one* 2017;12:e0180327.

584

585 **Figure legends**

586 **Fig.1. Flow diagram of the participants involved in this study**

587

588 **Fig.2. Feature selection using the LASSO binary logistic regression model.** (A) Feature selection for the LASSO  
589 binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against  
590 lambda by validating the optimal parameter lambda in the LASSO model. Dotted vertical lines were drawn based  
591 on 1 SE of the minimum criteria (the 1-SE criteria). (B) Feature selection was performed using the LASSO binary  
592 logistic regression model. A Coefficient profile was plotted based on the lambda series in Figure 1(A), and 12  
593 features with non-zero coefficients were selected by optimal lambda.

594

595 **Fig.3. The forest plot of the OR of the selected feature.**

596

597 **Fig.4. Developed mental health problems incidence risk nomogram.** The mental health problems incidence risk  
598 nomogram was developed in the array, with education, professional title, age, CMBI, ERI, asbestos dust,  
599 hypertension, diabetes, working hours per day, working years, marital status, and work schedule incorporated.

600

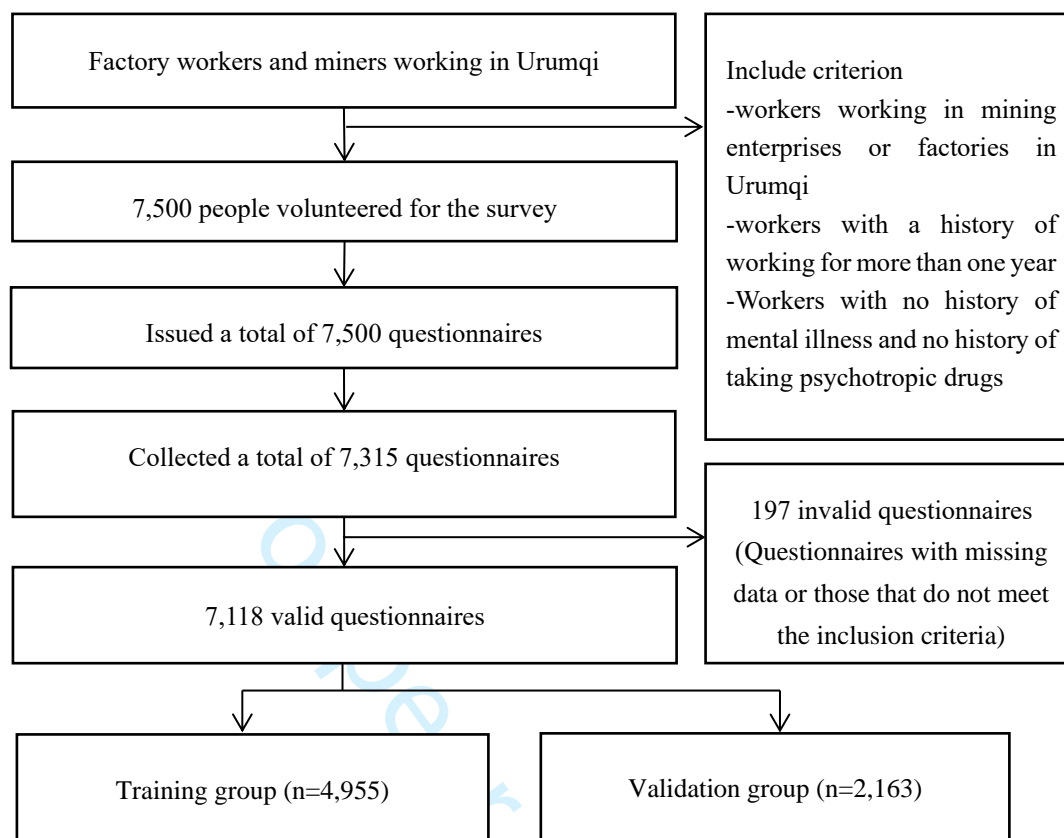
601 **Fig.5. Calibration curves of the mental health problems incidence risk nomogram prediction in validation**  
602 **group.** The x-axis represents the predicted risk of mental health problems. y-axis represents the actual diagnosed  
603 risk of mental health problems. The diagonal dashed line represents the perfect prediction of the ideal model. The  
604 solid lines represent the performance of the column plots, where closer to the diagonal dashed line indicates a better  
605 prediction.

606

607 **Fig.6. ROC curves for training and validation groups.** The y-axis represents the true positive rate of risk  
608 prediction. The x-axis represents the false positive rate of risk prediction. The ROC curves for the training and  
609 validation groups are shown in black and red.

610

611 **Fig.7. Decision curve analysis for mental health problems incidence risk nomogram.** The y-axis measures the  
612 net benefit. The solid red line represents nomogram of the risk of developing a mental health problem. The light blue  
613 dashed line represents the hypothesis that all participants were diagnosed with a mental health problem. The black  
614 dashed line represents the hypothesis that there is no risk of a mental health problem. The DCA showed that using  
615 this mental health problem incidence risk nomogram in the current study to predict mental health problem incidence  
616 risk increase in benefit than the intervention all patients or no intervention all patient if the threshold probability of  
617 a patients and a doctor is >10%.



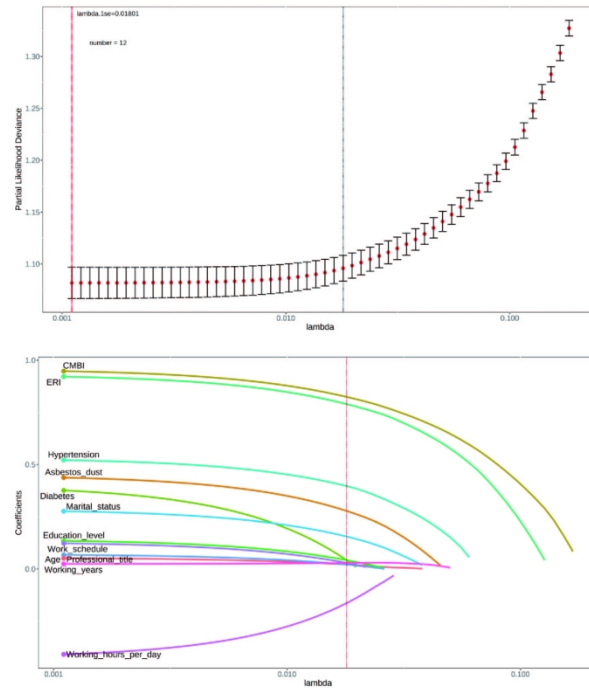
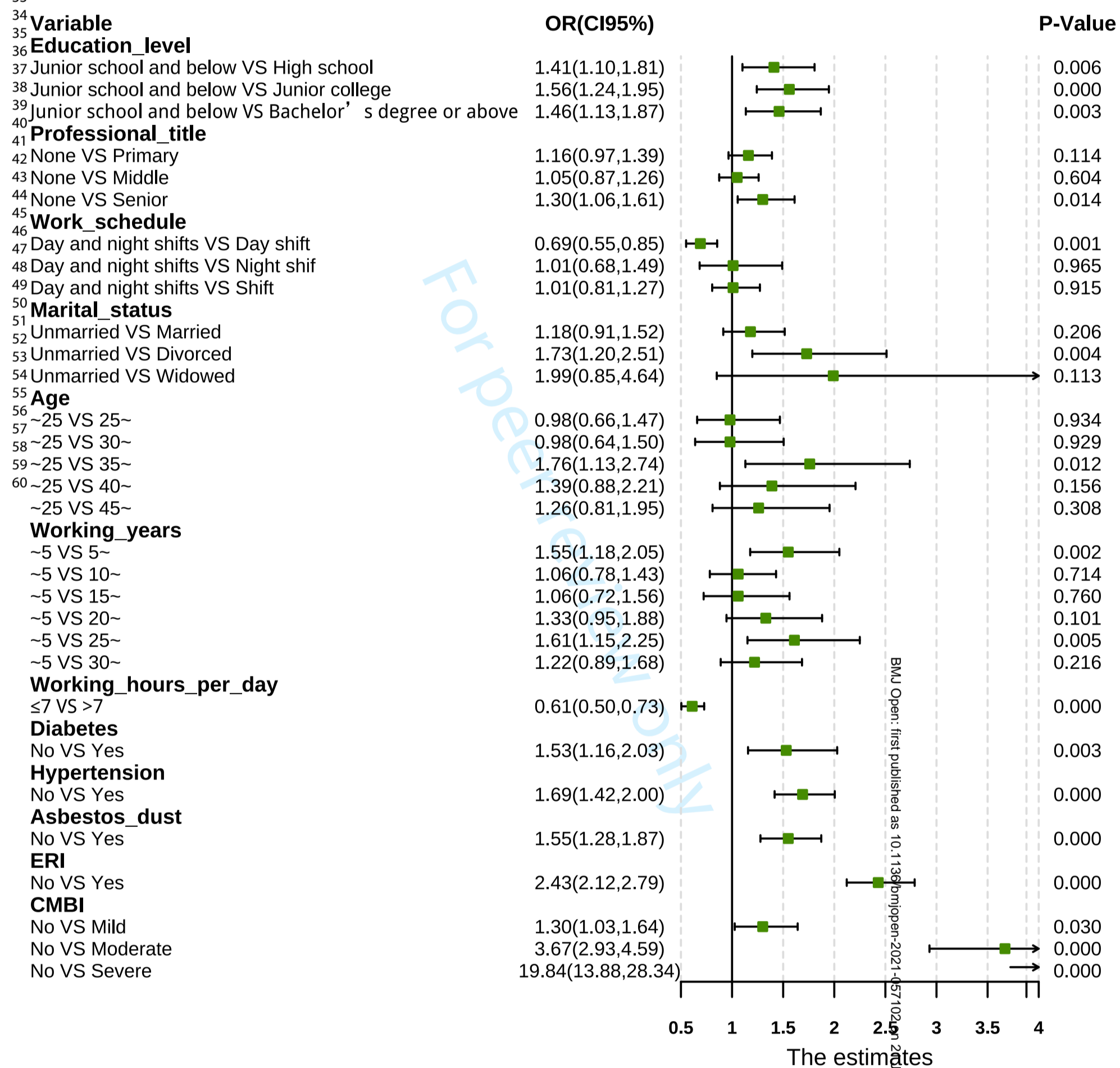


Fig.2. Feature selection using the LASSO binary logistic regression model. (A) Feature selection for the LASSO binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against lambda by validating the optimal parameter lambda in the LASSO model. Dotted vertical lines were drawn based on 1 SE of the minimum criteria (the 1-SE criteria). (B) Feature selection was performed using the LASSO binary logistic regression model. A Coefficient profile was plotted based on the lambda series in Figure 1(A), and 12 features with non-zero coefficients were selected by optimal lambda.

210x297mm (300 x 300 DPI)



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33



BMJ Open: first published as 10.1136/bmjopen-2021-025710 on 20 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

Points

CMBI

ERI

Marital status

Age

Hypertension

Working days per week

Working years

Education level

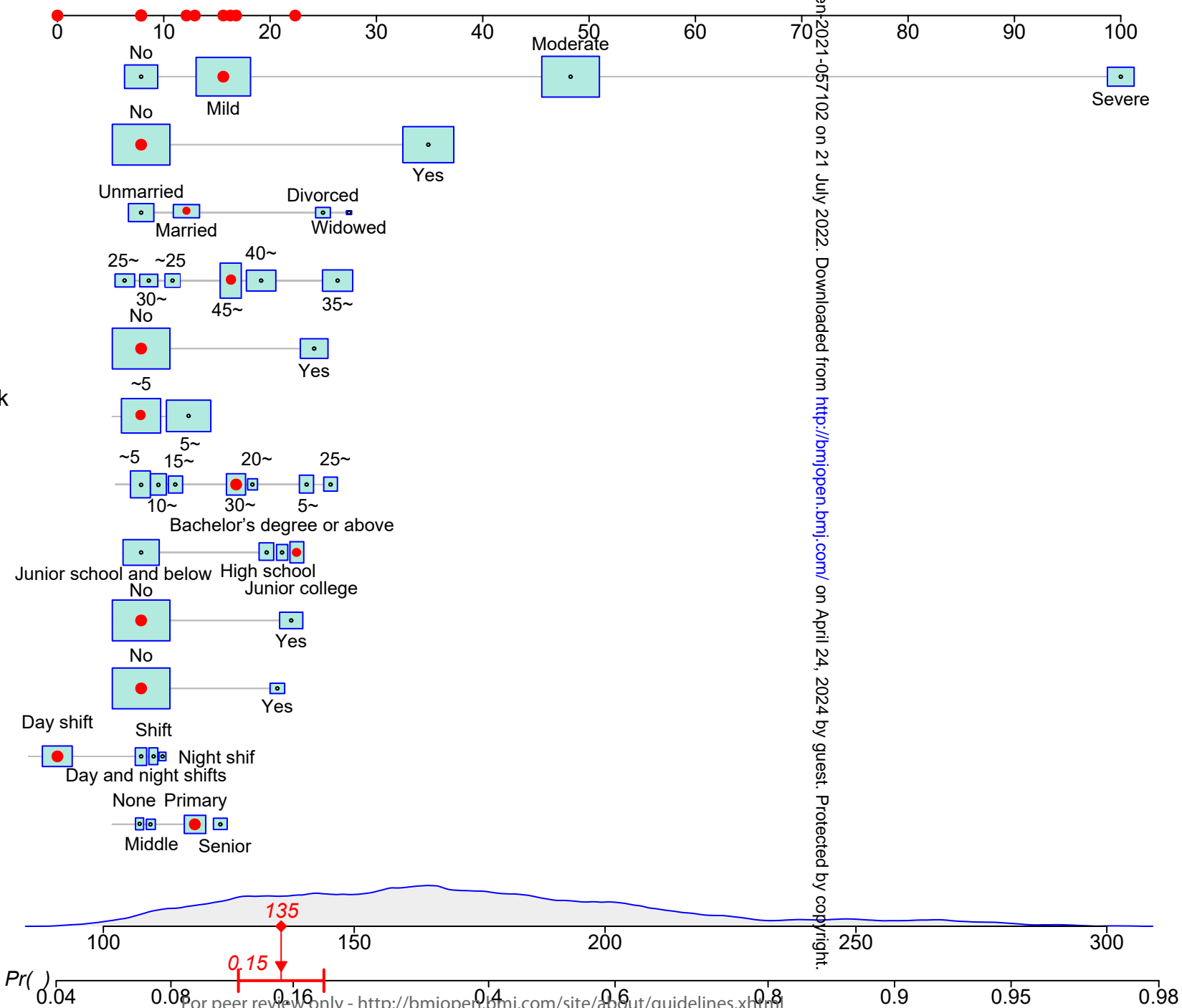
Asbestos dust

Diabetes

Work schedule

Professional title

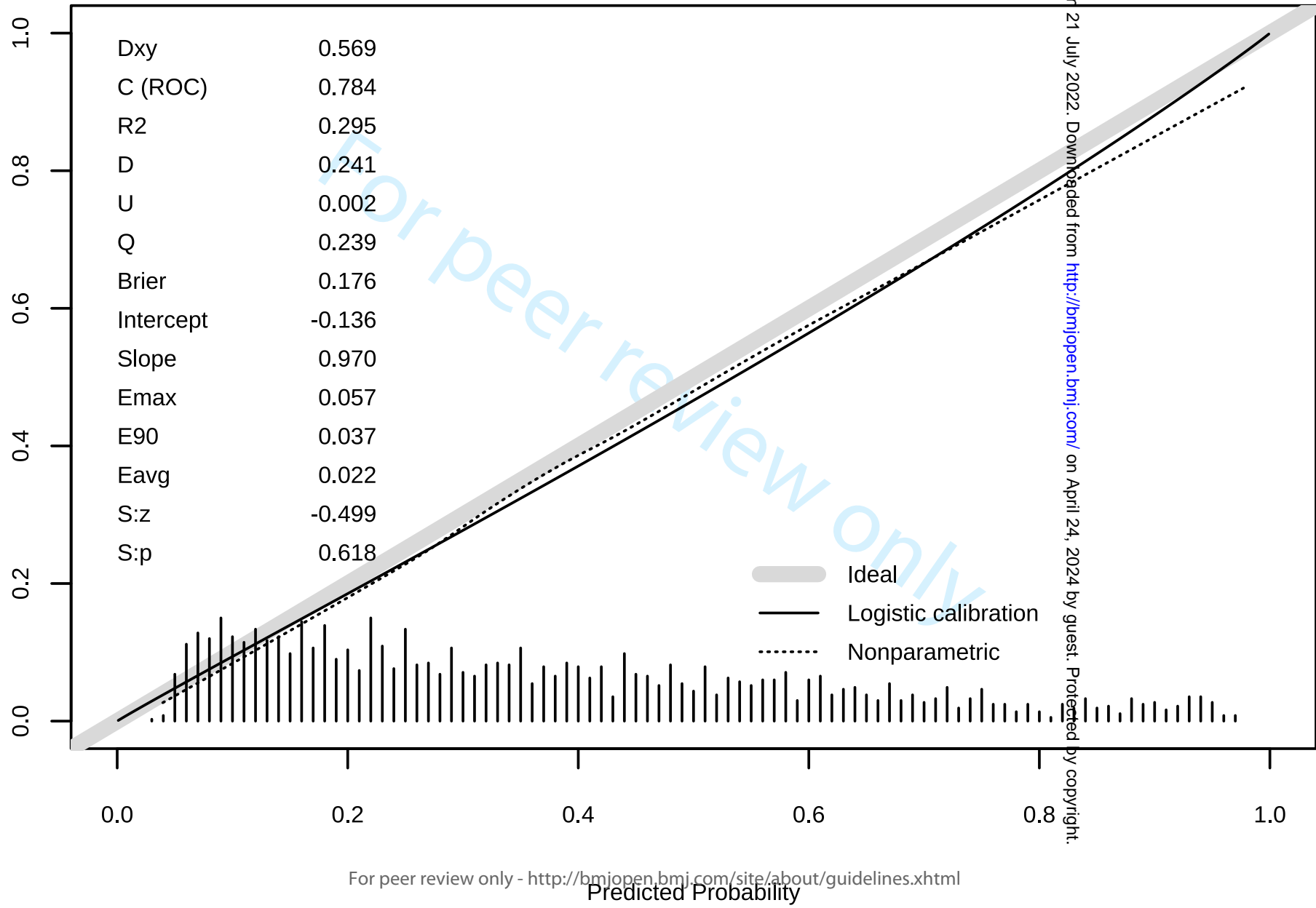
Total points

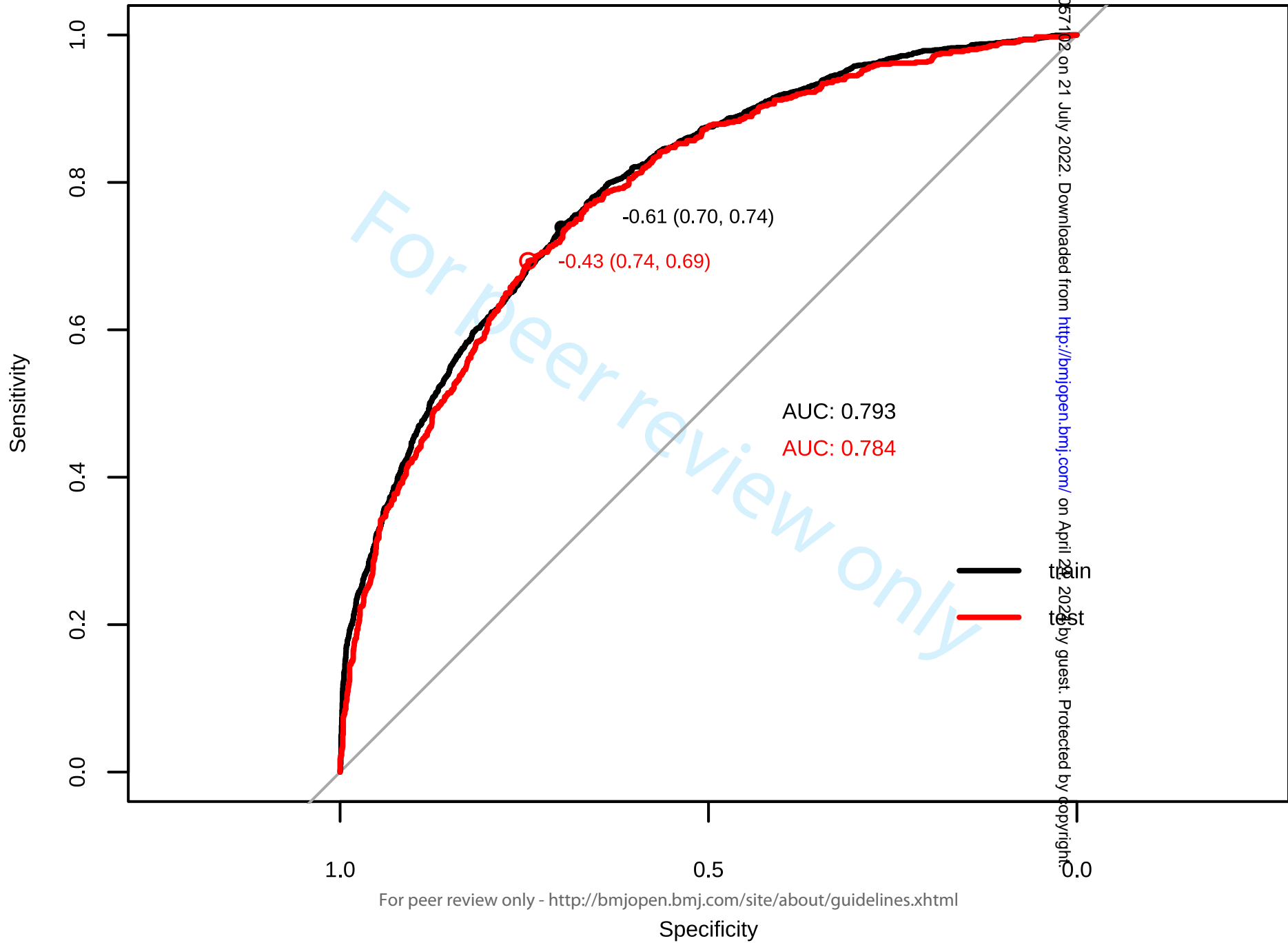


6/bmjopen-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

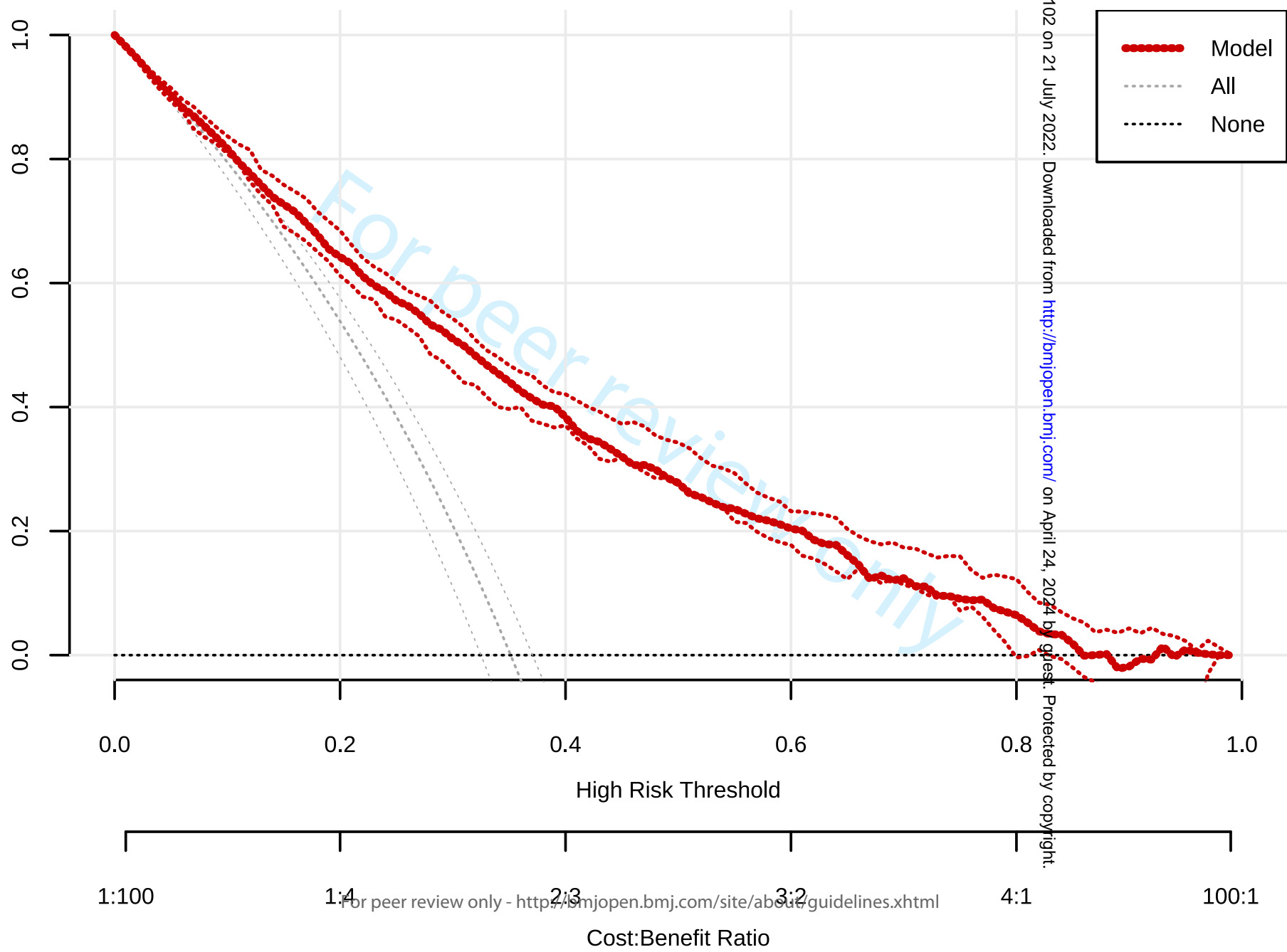
pen-2021-057102 or 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44





1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44



## STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
<b>Title and abstract</b>	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1	
<b>Introduction</b>				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	2	
Objectives	3	State specific objectives, including any prespecified hypotheses	3	
<b>Methods</b>				
Study design	4	Present key elements of study design early in the paper	4	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	4	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case		
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4	
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	4	
Bias	9	Describe any efforts to address potential sources of bias	4	
Study size	10	Explain how the study size was arrived at	4	

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	6
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	4
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed	4
		<i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) 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	7
		(b) Give reasons for non-participation at each stage	7
		(c) Consider use of a flow diagram	7
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	7
		(b) Indicate number of participants with missing data for each variable of interest	4
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	10
Main results	16	(a) 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	
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	12
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	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
<b>Other information</b>			
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	15

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).



# BMJ Open

## Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems of Factory Workers and Miners

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-057102.R2
Article Type:	Original research
Date Submitted by the Author:	24-Jun-2022
Complete List of Authors:	Lu, Yaoqin; Xinjiang Medical University, School of Public Health; Urumqi Center for Disease Control and Prevention Liu, Qi; Xinjiang Medical University Affiliated First Hospital, Postgraduate Education Management Section,; Xinjiang Medical University, School of Public Health Yan, Huan; Xinjiang Medical University, Department of Nutrition and Food Hygiene; Xinjiang Autonomous Academy of Instrumental Analysis Liu, Tao; Xinjiang Medical University, School of Public Health
<b>Primary Subject Heading</b>:	Mental health
Secondary Subject Heading:	Public health
Keywords:	MENTAL HEALTH, PREVENTIVE MEDICINE, PUBLIC HEALTH

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

# 1 Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems 2 of Factory Workers and Miners

3 Yaoqin Lu,<sup>1,2</sup> Qi Liu,<sup>1,5</sup> Huan Yan<sup>3,4</sup>, and Tao Liu<sup>1</sup>

4 <sup>1</sup>Department of Toxicology, School of Public Health, Xinjiang Medical University, Urumqi 830011,  
5 Xinjiang, China.

6 <sup>2</sup>Urumqi Center for Disease Control and Prevention, Urumqi 830026, Xinjiang, China.

7 <sup>3</sup>Department of Nutrition and Food Hygiene, College of Public Health, Xinjiang Medical University,  
8 Urumqi, Xinjiang, China.

9 <sup>4</sup>Xinjiang Engineering Technology Research Center for Green Processing of Nature Product Center,  
10 Xinjiang Autonomous Academy of Instrumental Analysis, Urumqi, Xinjiang, China.

11 <sup>5</sup>Graduate Education Management Section of Teaching and Scientific Research Department of the First  
12 Affiliated Hospital of Xinjiang Medical University, Urumqi 830017, Xinjiang, China.

13 Correspondence should be addressed to Tao Liu; xjmult@163.com

14 Yaoqin Lu and Qi Liu contributed equally to this work

## 15 Abstract

16 **Objective** A nomogram for predicting the risk of mental health problems was established in a population  
17 of factory workers and miners, in order to quickly calculate the probability of a worker suffering from  
18 mental health problems.

19 **Methods** A cross-sectional survey of 7,500 factory workers and miners in Urumqi was conducted by  
20 means of an electronic questionnaire using cluster sampling method. Participants were randomly  
21 assigned to the training group (70%) and the validation group (30%). Questionnaire-based survey was  
22 conducted to collect information. A least absolute shrinkage and selection operator (LASSO) regression  
23 model was used to screen the predictors related to the risk of mental health problems of the training  
24 group. Multivariate logistic regression analysis was applied to construct the prediction model. Calibration  
25 plots and receiver operating characteristic-derived area under the curve (AUC) were used for model  
26 validation. Decision curve analysis (DCA) was applied to calculate the net benefit of the screening model.

27 **Results** A total of 7,118 participants met the inclusion criteria and the data were randomly divided into  
28 a training group (n=4,955) and a validation group (n=2,163) in a ratio of 3:1. A total of 23 characteristics  
29 were included in this study and LASSO regression selected 12 characteristics such as education,  
30 professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working  
31 years, marital status, and work schedule as predictors for the construction of the nomogram. In the  
32 validation group the Brier score was 0.176, the calibration slope was 0.970 and the calibration curve of  
33 nomogram showed a good fit. The AUC of training group and verification group were 0.785 and 0.784  
34 respectively.

35 **Conclusion** The nomogram combining these 12 characteristics can be used to predict the risk of suffering  
36 mental health problems, providing a useful tool for quickly and accurately screening the risk of mental  
37 health problems.

39 **Key words** Mental health; Predictor; Nomogram; Risk; Factory workers and miners

40

41 **Strengths and limitations of this study**

42 1. This is the first study to develop an easy-to-use nomogram to predict the mental health risks of factory  
43 workers and miners.

44 2. The AUC of training group and verification group were 0.785 and 0.784 respectively, showing  
45 moderate discriminatory and calibration power.

46 3. This nomogram model's variables are more comprehensive, including demographics, burnout,  
47 occupational stress and occupational hazards.

48 4. We had considered many influential factors, but we were still not certain whether all possible  
49 influences were covered.

50 5. There is a lack of external validation in other populations in other regions and countries.

51

52 **1. Introduction**

53

54 The World Health Organization (WHO) defines health as a state of complete physical, mental and social  
55 well-being and not merely the absence of disease or weakness <sup>[1]</sup>. Obviously, health is an organic unity  
56 of physical and mental well-being. People with good mental health are the precondition for the normal  
57 operation of our society. However, with the acceleration of people's pace of life, people are facing an  
58 increasing risk of poor health, which has become a global public health problem <sup>[2]</sup>. Mental health  
59 problems can not only take a toll on physical health such as increasing the risk of communicable and  
60 non-communicable diseases and even causing unintentional or intentional harm to others <sup>[3]</sup>, but can also  
61 have a negative impact on the economy. For example, mental health disorders represent a growing part  
62 of the global burden of disease <sup>[4]</sup>, with statistics showing that nearly one billion people worldwide  
63 currently suffer from a mental disorder, and mental illness is ranked as one of the leading causes of the  
64 global burden of disease <sup>[5]</sup>. Moreover, one study has estimated that due to the impact of mental illness,  
65 the global economy loses US \$1 trillion every year <sup>[6]</sup>.

66

67 As researchers around the world have delved into the field of mental health, factors such as gender,  
68 income levels, environment and education have been found to be associated with people's mental health  
69 problems <sup>[7-10]</sup>. Moreover, employment is also strongly associated with quality of life, higher self-esteem  
70 and fewer psychiatric symptoms <sup>[11]</sup>. In addition, in the context of the global challenges of climate change,  
71 an increasing number of scholars have been examining the epidemiological links between mental health  
72 and environmental factors. Some studies have suggested that mental health may be influenced by ambient  
73 temperature, and an association has been found between environmental pollutants, particularly fine  
74 particulate matter, and mental health problems <sup>[12]</sup>. A relevant study shows that with short-term exposure  
75 to ambient air pollution is associated with increased emergency room visits due to depression or suicide  
76 attempts <sup>[13]</sup>. Furthermore, other factors associated with mental health include sleep, diabetes, coronary  
77 artery disease and cardiovascular disease <sup>[14-15]</sup>. It is worth noting that job burnout and occupational stress  
78 are closely linked to mental health. Job burnout is an exhaustion state of physical and psychological that

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

79 often occurs in the work environment, and has a high correlation with depression. A large study of  
80 physicians found that of the 10.3% who met criteria for a major depressive episode, 50.7% were also  
81 affected by symptoms of burnout (OR 2.99) and indicated that worsening depression leads to a higher  
82 likelihood of burnout symptoms <sup>[16]</sup>. Occupational stress refers to a work environment where non-  
83 reciprocity of effort and reward may lead to strong negative emotions and distress. Related research has  
84 shown that the combination of high effort and low reward and over-commitment increases the risk of  
85 mental health problems such as depression <sup>[17]</sup>. Apparently, it is necessary to include the CMBI and ERI  
86 in this study to predict the risk of mental health problems among factory workers and miners. However,  
87 there are few studies that include these influences in a more comprehensive way in the practice of  
88 detecting mental health. Therefore, more accurate identification of mental health problems in populations  
89 requires a questionnaire that include a wider range of factors affecting factory workers and miners'  
90 mental health problems.

91  
92 Factory workers and miners are a special group of workers with a relatively low overall level of education  
93 and are highly prone to suffering from mental health problems due to limited social support, excessive  
94 workload and irregular lifestyles, as well as occupational hazards such as noise and coal dust that they  
95 inevitably need to face in their working environment <sup>[18-19]</sup>. Through a review of the literature, our group  
96 found that coal dust, crystalline silica and noise pollution were common causes of health problems for  
97 workers in underground mines <sup>[20]</sup>. And, exposure to coal mine dust is a significant cause of  
98 pneumoconiosis in coal miners <sup>[21]</sup>. In addition, asbestos is one of the major occupational hazards in the  
99 daily work of workers in the construction and automotive industries <sup>[22]</sup>. China has the world's largest  
100 group of factory workers and miners, about 6 million <sup>[23]</sup>, who are regularly involved in occupational  
101 hazards. Mental health problems which need to require a long process are known to be a syndrome caused  
102 by chronic stress. Factory workers and miners, represented by those engaged in coal mining, have a  
103 mental burden rating of 8.3, one of the highest mental burdens among 150 occupations <sup>[24]</sup>. This explains  
104 the high level of mental health problems among mine workers in previous studies, making the  
105 identification and treatment of mental health problems even more important. Therefore, it is essential to  
106 provide a viable and easy-to-apply tool for identifying workers at risk of mental health problems and  
107 thus for timely interventions.

108  
109 There are many studies on mental health <sup>[25-26]</sup>; however, the results of previous studies lack consistency  
110 and mostly discuss factors influencing mental health, and most of them are single-center studies that  
111 focus on only certain aspects of mental health. Our study included common demographics, job burnout,  
112 occupational stress, chronic illness and occupational exposure factors to distinguish whether respondents  
113 suffered from mental health problems. In addition, there is a small body of literature that develops and  
114 validates a risk nomogram between depression and suicide to support timely intervention by clinicians.  
115 And the sample sizes of the two relevant studies were small, 474 and 273 depressed patients respectively  
116 <sup>[27-28]</sup>. Today, there is increasing recognition of the important role of mental health in achieving global  
117 development goals, and WHO has included mental health in the Sustainable Development Goals.  
118 However, there are no relevant studies that have used objective indicators for factory workers and miners

119 to form a nomogram to predict mental health. Therefore, to bridge this gap in the literature and provide  
120 additional information for the prevention of mental health problems, we conducted a multicenter  
121 investigation to develop and validate an easy-to-use nomogram that combines objective information on  
122 demographics, job burnout, occupational stress and occupational hazards to comprehensively and  
123 accurately predict the prevalence of mental health problems among factory workers and miners.

124

## 125 2. Materials and Methods

126

### 127 2.1 Calculation of sample size

128

129 The sample size formula for the present illness rate survey,  $n = \frac{z_{\alpha/2}^2 \times pq}{\delta^2}$ , p is the present-hazard rate,  
130  $q=1-p$ ,  $\delta$  is the tolerance error, generally taken as 0.1p,  $z_{\alpha/2}$  is the significance test statistic,  $z_{\alpha/2}=1.96$   
131 for  $\alpha=0.05$ , then the formula is calculated as,  $n = 400 \times \frac{q}{p}$ . A cross-sectional study in Xinjiang showed  
132 that 38.27% of factory workers and miners had mental health problems [29]. And a study revealed that  
133 633 out of 1675 coal miners (37.8%) suffered from mental disorders between August 2018 and June  
134 2019[30]. In this study, we assumed a 30% prevalence of mental health problem to obtain the maximum  
135 required sample size. which would calculate a sample size of 934, taking into account non-response and  
136 a 20% loss of questionnaires, which would require approximately 1168 people.

137

### 138 2.2. Participants

139

140 Participants in this cross-sectional survey were factory workers and mines in the Urumqi region, and the  
141 survey covered all districts and counties in the Urumqi region to avoid selection bias as far as possible.  
142 Specifically, this survey was conducted by means of whole-group random sampling from January to May  
143 2019, and a total of 202 enterprises were selected, including 21 in Tianshan District, 30 in Shaibak  
144 District, 24 in Xinshi District, 22 in Shuimogou District, 56 in Jingkai District, 37 in Midong District, 9  
145 enterprises in Dabancheng District and 3 enterprises in Urumqi County.

146

147 The inclusion criteria were as follows: (1) workers working in mining enterprises or factories in Urumqi;  
148 (2) workers with a history of working for more than one year; (3) Workers with no history of mental  
149 illness and no history of taking psychotropic drugs.

150

151 The exclusion criteria were the following: (1) factory workers and miners in non-Urumqi area; (2)  
152 working history of factories and mining enterprises less than 1 year; (3) a confirmed diagnosis of a mental  
153 health problem and a history of treatment and use of psychotropic medication; (4) Questionnaires with  
154 missing data were excluded.

155

156 An online electronic questionnaire was created using the Questionnaire Star platform to collect data. In  
157 the introductory section of the electronic questionnaire, we provide a paragraph stating that volunteers

1  
2  
3 158 can choose to continue answering the survey if they wish to participate and the relevant data will be used  
4 159 for scientific research, or refuse to answer if they do not wish to participate in the survey. In addition,  
5 160 this survey was conducted by trained surveyors who explained the purpose, meaning, content and  
6 161 requirements of the questionnaire to all participants and provided on-site instructions to ensure the return  
7 162 rate of the questionnaire. All participants understood the purpose of the study and were willing to  
8 163 participate in the study. A total of 7,500 questionnaires were distributed and 7,315 questionnaires were  
9 164 returned, representing a return rate of 97.5%. After checking the validity and integrity of the  
10 165 questionnaires, 7,118 questionnaires were confirmed as valid, with an effective rate of 97.3%. A total of  
11 166 7,118 participants met the inclusion criteria and the data were randomly divided into a training group  
12 167 (n=4,955) and a validation group (n=2,163) (Figure 1).  
13  
14  
15  
16  
17  
18

## 19 169 **2.3. Research Methods**

### 20 170 **2.3.1. Assessment of mental health**

21 171  
22 172 The SCL-90, designed by Derogatis and his colleagues, was widely used in the mental health field <sup>[31]</sup>,  
23 173 which contains 90 items across nine dimensions: somatization, obsessive-compulsive symptoms,  
24 174 interpersonal sensitivity, depression, anxiety, hostility, horror, bigotry and mental illness. The SCL-90  
25 175 has been used extensively in previous studies and has relatively high reliability and validity <sup>[32]</sup>. The  
26 176 questionnaire uses a Likert 5-point scale, with a score of 0 point indicating none and 4 points indicating  
27 177 severe. A total score above 160, a score above 2 on any item, or a positive item above 43 indicates the  
28 178 presence of a psychological abnormality <sup>[33]</sup>. In this survey, Cronbach  $\alpha$  was 0.99, the half-reliability  
29 179 coefficient was 0.98, and the KMO was 0.994.  
30  
31  
32  
33  
34  
35

### 36 181 **2.3.2. Assessment of occupational stress**

37 182  
38 183 This survey evaluated occupational stress in factory workers and miners through the Effort–Reward  
39 184 Imbalance (ERI) model developed by Siegrist <sup>[34]</sup>. The ERI scale consists of three subscales: effort (E, 6  
40 185 items), reward (R, 11 items) and over commitment (6 items), for a total of 23 items. A Likert 5-level  
41 186 scoring method (1, "highly disagree" to 5, "highly agree") is used to grade the items in the questionnaire  
42 187 with the same weight for each item. The effort–return index  $ERI = E/R \times C$ , where C is the adjustment  
43 188 coefficient, and the value is 6/11. ERI values greater than 1, equal to 1, and less than 1 correspond to  
44 189 high pay–low return, pay–return balance, and low pay–high return, respectively. Moreover, the higher  
45 190 the ERI value, the greater the occupational stress <sup>[35]</sup>. In this survey, Cronbach  $\alpha$  was 0.94, the half-  
46 191 reliability coefficient was 0.93 and the KMO was 0.956.  
47  
48  
49  
50  
51  
52

### 53 193 **2.3.3. Assessment of job burnout**

54 194  
55 195 In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess  
56 196 job burnout, which has good reliability and validity <sup>[36]</sup>. CMBI is composed of 15 items in three  
57 197 dimensions: emotional exhaustion (5 items), depersonalization (5 items) and reduced personal  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

198 accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete  
199 compliance and 7 points indicating complete non-compliance. According to the critical value (emotional  
200 exhaustion  $\geq 25$ , depersonalization  $\geq 11$ , personal achievement reduction  $\geq 16$ ), the levels of occupational  
201 burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to  
202 or above the critical value), moderate (any two aspects are equal to or higher than the critical values),  
203 and severe (three aspects are equal to or higher than the critical values) [37]. In this survey, Cronbach  $\alpha$   
204 was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.

#### 205 206 **2.3.4. Candidate predictors**

207  
208 Trained investigators obtained information on demographics, job burnout, occupational stress, mental  
209 health and occupational exposure factors through on-site face-to-face collection of an electronic version  
210 of the questionnaire. Covariates included in this study: 1) demographic information: gender, ethnicity,  
211 education level, professional title, work schedule, marital status, monthly income, age, working years,  
212 labor contracts, working hours per day, and working hours per week; 2) occupational exposure factors:  
213 coal dust, silica dust, asbestos dust, benzene, lead, noise, and brucellosis; 3) questionnaires: ERI, CMBI;  
214 4) chronic diseases: diabetes, hypertension. Information on four areas, including demographic  
215 information, questionnaires, occupational hazards and chronic diseases, were filled in by participants  
216 through their own responses on the questionnaire star.

217  
218 Sex was defined as male or female; ethnicity was defined as Han and other; education level was defined  
219 as junior high school and below, high school, junior college or bachelor's degree or above; labor contracts  
220 was defined as signed or unsigned; professional title was defined as no, primary, middle or senior; work  
221 schedule was defined as day shift, night shift, shift or day and night shifts; marital status was defined as  
222 unmarried, married, divorced or widowed; monthly income (yuan) was defined as <3000, 3000~, 4000~,  
223 5000~, 6000~, 7000~ or 8000~; age (years) was defined as <25, 25~, 30~, 35~, 40~ or 45~; working  
224 years was defined as ~5, 5~, 10~, 15~, 20~, 25~ or 30~; working hours per day (hours) was defined as  
225  $\leq 7$  or  $> 7$ ; working days per week (days) was defined as  $\leq 5$  or  $> 5$ ; exposure to coal dust, silica dust,  
226 asbestos dust, benzene, lead, noise, brucellosis were all defined as yes or no; ERI was defined as yes or  
227 no; CMBI was defined as none, mild, moderate and severe; hypertension and diabetes were both defined  
228 as yes or no.

#### 229 230 **2.4. Statistical analysis**

231  
232 Categorical variables were described as counts and percentages, and chi square test or Fisher exact test  
233 was used to compare categorical variables between groups. 70% of participants were randomly assigned  
234 to the training cohort and 30% to the validation cohort. Variables were screened using a least absolute  
235 shrinkage and selection operator (LASSO) regression and multivariate logistic regression models were  
236 used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which  
237 predictive models were constructed. A nomogram for predicting was generated according to the selected



1  
2  
3 238 characteristics. In addition, forest plot was drawn to visually depict the P-value, OR and 95% CI for the  
4 239 selected validations. Statistically significant predictors were applied to develop a prediction model for  
5 240 the risk of mental health problems among factory workers and miners by introducing all selected factors  
6 241 and analyzing the statistical significance levels of them. We used calibration plots and receiver operating  
7 242 characteristic (ROC) curves to show the calibration and discrimination of our final model. Brier scores  
8 243 for overall performance, calibration slopes were used to assess the predictable accuracy of the model.  
9 244 Decision curve analysis (DCA) was applied to calculate the net benefit of the nomogram. Statistical  
10 245 analysis was performed using the open-source R software Version 3.6.1 (<http://www.r-project.org>). The  
11 246 significance level ( $\alpha$ ) set at 0.05.  
12  
13  
14  
15  
16  
17

## 18 248 **2.5. Patient and public involvement**

19 249 Neither patients nor members of the public had any involvement in the design of this study.  
20  
21

## 22 251 **3. Results**

### 23 252 **3.1. Participant characteristics**

24 253  
25 254  
26 255 A total of 7,118 participants met the inclusion criteria and the data were randomly divided into a training  
27 256 group (n=4,955) and a validation group (n=2,163). Over half of all participants (65.31%) were male,  
28 257 57.31% of the population was over 35 years of age and 78.32% of the subjects were married, showing  
29 258 that factory workers and miners are generally older and most of them have spouses. The majority of them  
30 259 had completed high school (83.94%), while a smaller percentage had completed undergraduate education  
31 260 (22.98%), indicating that the group of factory workers and miners as a whole was not well educated. The  
32 261 total number of workers (n, %) exposed to coal dust, silica dust, asbestos dust, benzene, lead, noise and  
33 262 brucellosis in the factory and mining enterprises were 377 (5.3), 730 (10.3), 981 (14), 1,981 (27.8), 373  
34 263 (5.2), 4,942 (69.4) and 121 (1.7) respectively, with the total number of workers exposed to noise  
35 264 amounting to 4,942, or 69% of the total population surveyed. The demographic, job burnout,  
36 265 occupational stress and occupational exposure factors for the training and validation groups are shown  
37 266 in Table 1. The results showed that there were no significant statistical differences between the two  
38 267 groups of characteristic variables, except for coal dust and CMBI, indicating that the baseline levels were  
39 268 largely consistent between the two groups.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

278  
279  
280

Table 1 Characteristics of the study participants

Variables		Total (n = 7118)	train (n = 4955)	test (n = 2163)	<i>p</i>
Sex, n (%)	Male	4649 (65.3)	3216 (64.9)	1433 (66.3)	0.284
	Female	2469 (34.7)	1739 (35.1)	730 (33.7)	
Ethnicity, n (%)	Han	5762 (80.9)	3982 (80.4)	1780 (82.3)	0.061
	Other	1356 (19.1)	973 (19.6)	383 (17.7)	
Education level, n (%)	Junior high school and below	1143 (16.1)	804 (16.2)	339 (15.7)	0.765
	High school	1406 (19.8)	988 (19.9)	418 (19.3)	
	Junior college	2933 (41.2)	2038 (41.1)	895 (41.4)	
	Bachelor's degree or above	1636 (23.0)	1125 (22.7)	511 (23.6)	
Professional title, n (%)	None	2854 (40.1)	1983 (40.0)	871 (40.3)	0.923
	Primary	1644 (23.1)	1149 (23.2)	495 (22.9)	
	Middle	1618 (22.7)	1133 (22.9)	485 (22.4)	
	Senior	1002 (14.1)	690 (13.9)	312 (14.4)	
Work schedule, n (%)	Day shift	3986 (56.0)	2801 (56.5)	1185 (54.8)	0.585
	Night shift	270 (3.8)	187 (3.8)	83 (3.8)	
	Shift	2058 (28.9)	1412 (28.5)	646 (29.9)	
	Day and night shifts	804 (11.3)	555 (11.2)	249 (11.5)	
Marital status, n (%)	Unmarried	1104 (15.5)	762 (15.4)	342 (15.8)	0.218
	Married	5575 (78.3)	3906 (78.8)	1669 (77.2)	
	Divorced	390 (5.5)	255 (5.1)	135 (6.2)	
	Widowed	49 (0.7)	32 (0.6)	17 (0.8)	
Monthly income (yuan), n (%)	<3000	1799 (25.3)	1246 (25.1)	553 (25.6)	0.966
	3000~	2418 (34.0)	1682 (33.9)	736 (34.0)	
	4000~	1600 (22.5)	1125 (22.7)	475 (22.0)	
	5000~	752 (10.6)	520 (10.5)	232 (10.7)	
	6000~	288 (4.0)	201 (4.1)	87 (4.0)	
	7000~	148 (2.1)	106 (2.1)	42 (1.9)	
	8000~	113 (1.6)	75 (1.5)	38 (1.8)	
	>8000	107 (1.5)	78 (1.6)	29 (1.4)	
Age (years), n (%)	<25	431 (6.1)	297 (6.0)	134 (6.2)	0.173
	25~	786 (11.0)	519 (10.5)	267 (12.3)	
	30~	956 (13.4)	684 (13.8)	272 (12.6)	
	35~	866 (12.2)	617 (12.5)	249 (11.5)	
	40~	849 (11.9)	588 (11.9)	261 (12.1)	
	45~	3230 (45.4)	2250 (45.4)	980 (45.3)	
Working years (years), n (%)	<5	1170 (16.4)	794 (16.0)	376 (17.4)	0.248
	5~	1065 (15.0)	736 (14.9)	329 (15.2)	

	10~	997 (14.0)	721 (14.6)	276 (12.8)	
	15~	389 (5.5)	273 (5.5)	116 (5.4)	
	20~	763 (10.7)	538 (10.9)	225 (10.4)	
	25~	1293 (18.2)	878 (17.7)	415 (19.2)	
	30~	1441 (20.2)	1015 (20.5)	426 (19.7)	
Labor contracts, n (%)	Signed	6641 (93.3)	4624 (93.3)	2017 (93.3)	0.955
	Unsigned	477 (6.7)	331 (6.7)	146 (6.7)	
Working hours per day (hours), n (%)	≤7	1161 (16.3)	814 (16.4)	347 (16.0)	0.712
	>7	5957 (83.7)	4141 (83.6)	1816 (84.0)	
Working days per week (days), n (%)	≤5	4442 (62.4)	3107 (62.7)	1335 (61.7)	0.446
	>5	2676 (37.6)	1848 (37.3)	828 (38.3)	
Diabetes, n (%)	Yes	429 (6.0)	298 (6.0)	131 (6.1)	0.988
	No	6689 (94.0)	4657 (94.0)	2032 (93.9)	
Hypertension, n (%)	Yes	1330 (18.7)	929 (18.7)	401 (18.5)	0.861
	No	5788 (81.3)	4026 (81.3)	1762 (81.5)	
Coal dust, n (%)	Yes	377 (5.3)	244 (4.9)	133 (6.1)	0.039
	No	6741 (94.7)	4711 (95.1)	2030 (93.9)	
Silica dust, n (%)	Yes	730 (10.3)	523 (10.6)	207 (9.6)	0.223
	No	6388 (89.7)	4432 (89.4)	1956 (90.4)	
Asbestos dust, n (%)	Yes	981 (13.8)	691 (13.9)	290 (13.4)	0.570
	No	6137 (86.2)	4264 (86.1)	1873 (86.6)	
Benzene, n (%)	Yes	1981 (27.8)	1360 (27.4)	621 (28.7)	0.287
	No	5137 (72.2)	3595 (72.6)	1542 (71.3)	
Lead, n (%)	Yes	373 (5.2)	246 (5.0)	127 (5.9)	0.128
	No	6745 (94.8)	4709 (95.0)	2036 (94.1)	
Noise, n (%)	Yes	4942 (69.4)	3420 (69.0)	1522 (70.4)	0.270
	No	2176 (30.6)	1535 (31.0)	641 (29.6)	
Brucellosis, n (%)	Yes	121 (1.7)	86 (1.7)	35 (1.6)	0.800
	No	6997 (98.3)	4869 (98.3)	2128 (98.4)	
ERI, n (%)	Yes	3147 (44.2)	2173 (43.9)	974 (45.0)	0.372
	No	3971 (55.8)	2782 (56.1)	1189 (55.0)	
CMBI, n (%)	No	959 (13.5)	674 (13.6)	285 (13.2)	0.033
	Mild	2667 (37.5)	1813 (36.6)	854 (39.5)	
	Moderate	2900 (40.7)	2031 (41.0)	869 (40.2)	
	Severe	592 (8.3)	437 (8.8)	155 (7.2)	

281

282 **3.2. Feature selection**

283

284 The lambda was smallest at 0.01801 as seen from the lasso results when there were 12 characteristics,  
 285 which were education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working  
 286 hours per day, working years, marital status, and work schedule based on the results of the questionnaires  
 287 on demographics, occupational stress, job burnout and occupational exposure factors (Figure 2).

288

### 289 3.3. Results of logistic regression model

290

291 The 12 features obtained from the LASSO regression were incorporated into a multivariate logistic  
 292 regression model and the regression results were shown in Table 2. It was clear from the results that  
 293 education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per  
 294 day, working years, marital status, and work schedule were independent determinants of risk for mental  
 295 health problems. In addition, there was no evidence of multicollinearity between the covariates included  
 296 in the model. The forest plot showed that the selected 12 features all contain items with  $P < 0.05$ , among  
 297 which the degree of severe of CMBI (OR, 19.84; 95% CI, 13.88-28.34;  $P < 0.001$ ) had the greatest impact  
 298 on the risk of mental health problems among factory workers and miners (Figure 3).

299

Table 2 Predictive factors of risk for mental health problems among factory workers and miners

Variable	$\beta$	S.E.	OR(CI95%)	Wald	<i>P</i>	VIF
Intercept	-2.33	0.25	0.10(0.06,0.16)	-9.357	0	-
Education level						
Junior school and below VS High school	0.34	0.13	1.41(1.10,1.81)	2.727	0.006**	2.28
Junior school and below VS Junior college	0.44	0.11	1.56(1.24,1.95)	3.850	< 0.001***	2.79
Junior school and below VS Bachelor's degree or above	0.38	0.13	1.46(1.13,1.87)	2.953	0.003**	2.51
Professional title						
None VS Primary	0.15	0.09	1.16(0.97,1.39)	1.582	0.114	1.35
None VS Middle	0.05	0.09	1.05(0.87,1.26)	0.519	0.604	1.34
None VS Senior	0.27	0.11	1.30(1.06,1.61)	2.458	0.014*	1.32
Work schedule						
Day and night shifts VS Day shift	-0.38	0.11	0.69(0.55,0.85)	-3.364	0.001**	2.70
Day and night shifts VS Night shift	0.01	0.20	1.01(0.68,1.49)	0.044	0.965	1.30
Day and night shifts VS Shift	0.01	0.12	1.01(0.81,1.27)	0.107	0.915	2.47
Marital status						
Unmarried VS Married	0.16	0.13	1.18(0.91,1.52)	1.263	0.206	2.29
Unmarried VS Divorced	0.55	0.19	1.73(1.20,2.51)	2.918	0.004**	1.69
Unmarried VS Widowed	0.69	0.43	1.99(0.85,4.64)	1.586	0.113	1.09
Age						
~25 VS 25~	-0.02	0.20	0.98(0.66,1.47)	-0.083	0.934	3.09
~25 VS 30~	-0.02	0.22	0.98(0.64,1.50)	-0.090	0.929	4.79
~25 VS 35~	0.56	0.23	1.76(1.13,2.74)	2.503	0.012*	5.01

	~25 VS 40~	0.33	0.23	1.39(0.88,2.21)	1.419	0.156	4.97
	~25 VS 45~	0.23	0.22	1.26(0.81,1.95)	1.018	0.308	10.93
Working years							
	~5 VS 5~	0.44	0.14	1.55(1.18,2.05)	3.114	0.002**	2.27
	~5 VS 10~	0.06	0.15	1.06(0.78,1.43)	0.366	0.714	2.48
	~5 VS 15~	0.06	0.20	1.06(0.72,1.56)	0.305	0.760	1.79
	~5 VS 20~	0.29	0.18	1.33(0.95,1.88)	1.641	0.101	2.65
	~5 VS 25~	0.48	0.17	1.61(1.15,2.25)	2.782	0.005**	3.99
	~5 VS 30~	0.20	0.16	1.22(0.89,1.68)	1.239	0.216	3.90
Working hours per day							
	≤7 VS >7	-0.50	0.09	0.61(0.50,0.73)	-5.363	< 0.001***	1.15
Diabetes							
	No VS Yes	0.43	0.14	1.53(1.16,2.03)	2.974	0.003**	1.05
Hypertension							
	No VS Yes	0.52	0.09	1.69(1.42,2.00)	5.885	< 0.001***	1.11
Asbestos dust							
	No VS Yes	0.44	0.10	1.55(1.28,1.87)	4.474	< 0.001***	1.03
ERI							
	No VS Yes	0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***	1.05
CMBI							
	No VS Mild	0.26	0.12	1.30(1.03,1.64)	2.175	0.003**	2.73
	No VS Moderate	1.30	0.11	3.67(2.93,4.59)	11.361	< 0.001***	2.83
	No VS Severe	2.99	0.18	19.84(13.88,28.34)	16.41	< 0.001***	1.44

Note:  $\beta$  is the regression coefficient. “\*\*\*” indicates  $P < 0.001$ , “\*\*” indicates  $P < 0.01$ , “\*” indicates  $P < 0.05$ .

### 3.4. Development of an individualized prediction model

Based on the results of the multivariate analysis, predictors such as education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working years, marital status, and work schedule were included in the nomogram. A model incorporating the above independent predictors was developed and represented as a nomogram in Figure 4. Each variable in nomogram was assigned a score, and the cumulative sum of each 'point' was the 'total score'. The "total score" corresponded to the "predictable likelihood", which was the predicted probability of mental health problems among factory workers and miners as suggested by our design of the nomogram.

As an example of the use of nomogram: a randomly selected sample from the training group, one with no professional title, day shift, no diabetes or hypertension, Junior college, <5 of working years, >7 of working hours per day, married, no exposed to asbestos dust, <25 years of age, no ERI, mild of CMBI,

1  
2  
3 315 with a calculated total score of 174 and a corresponding risk probability of 8.27% for mental health  
4 316 problems.

5 317

### 6 318 **3.5 The validation of calibration**

7 319

8 320 Model validation was carried out in the validation group. The prediction accuracy of the model was  
9 321 assessed by two aspects. (1) The Brier score for overall performance, which assessed the difference  
10 322 between observed and predicted values, with values closer to 0 indicating better predictive ability. (2)  
11 323 The calibration slope used for modal calibration, which assessed the agreement between the observed  
12 324 and predicted values, with values closer to 1 indicating better performance. The accuracy measurements  
13 325 for the bias correction were validated by the model with a Brier score of 0.176 and a calibration slope of  
14 326 0.970, respectively (Figure 5). The prediction accuracy of the model was relatively high.

15 327

### 16 328 **3.6 The validation of discrimination**

17 329

18 330 ROC was plotted for the training and validation groups, and the AUC of training and the verification  
19 331 groups were 0.785 and 0.784, respectively (Figure 6). The AUC of training and the verification groups  
20 332 were both greater than 0.750, showing a good discrimination.

21 333

### 22 334 **3.7 Decision Curve Analysis**

23 335

24 336 As shown in the DCA of the risk of mental health problems nomogram in Figure 7, the model for  
25 337 predicting the risk of mental health problems for factory workers and miners in this study was more  
26 338 practically relevant if the threshold probability of patients was  $>10\%$ .

27 339

## 28 340 **4. Discussion**

29 341

30 342 To our knowledge, this is the first study to develop an easy-to-use nomogram to predict the mental health  
31 343 risks of factory workers and miners. The nomogram developed using the training set data contain 12  
32 344 items for education, professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working  
33 345 hours per day, working years, marital status, and work schedule. In addition, validation has shown that  
34 346 nomogram model has good accuracy and discriminatory power. Our novel nomogram can be used in any  
35 347 setting to provide a rapid assessment of mental health risks and to help identify patients with mental  
36 348 health risks, saving time compared to previous mental health investigations and improving on the lack  
37 349 of entries in previous investigations related to the specific working environment of factory workers and  
38 350 miners. The AUC of training group and verification group were 0.785 and 0.784 respectively, showing  
39 351 moderate discriminatory and calibration power.

40 352

41 353 A review of the literature found that the vast majority of studies constructed nomograms to predict  
42 354 clinical disorders, with less literature used to predict psychological problems. In a study to predict the

43 355

1  
2  
3 355 correlates of suicide attempts in a Chinese population with major depressive disorder, the C-index was  
4 356 0.715 and the C-index in the internal validation set was 0.703, and the calibration curve of the column  
5  
6 357 line plot also showed good agreement between the predicted and observed risk of suicide attempts. The  
7  
8 358 variables in the nomogram included socio-demographic information and clinical variables including age,  
9 359 duration, number of episodes, age at onset, number of hospitalizations, characteristics of anxiety and  
10 360 psychiatric symptoms, marital status, income, education level and employment status [27]. In another  
11 361 study that created a nomogram to predict the risk of psychosocial and behavioral problems in children  
12 362 and adolescents during the COVID-19 pandemic, the C index exceeded 0.800 and the calibration curve  
13 363 also showed good predictive accuracy. The variables covered three subject areas, namely demographic  
14 364 information, the psychosocial impact of the epidemic such as homework time and sedentary time, and  
15 365 the Child Behaviour Checklist score (CBCL) for the evaluation of psychological problems [38]. In this  
16 366 study, 7,118 participants were randomly divided into a training group (n=4,955) and a validation group  
17 367 (n=2,163) in a ratio of 3:1, involving a total of 23 features, and 12 features were selected by LASSO  
18 368 regression. The nomogram could be a useful tool to better identify patients with mental health problems,  
19 369 as it not only covered comprehensive information, including demographic information, job burnout,  
20 370 occupational stress, chronic diseases and occupational exposure factors closely related to factory workers  
21 371 and miners, but also was simple to operate and easy to use. In the validation group the Brier score was  
22 372 0.176, the calibration slope was 0.970 and the calibration curve of nomogram showed a good fit. The  
23 373 AUC of training group and verification group were 0.785 and 0.784 respectively. Compared to the two  
24 374 studies above, our nomogram showed good accuracy and discrimination, and more comprehensive  
25 375 coverage in this nomogram model. Therefore, the possibility of early intervention for patients with high-  
26 376 risk mental health problems will be increased by covering multiple information and easy to use  
27 377 nomogram modal, especially for factory workers and miners with poor working conditions, relatively  
28 378 low levels of education and low patience.

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39 380 Mental health problems were very common in the group of factory workers and miners, and the  
40 381 prevalence of mental health of them was found to be 37.08% in our study. Notably, the CMBI showed  
41 382 the most significant score (score = 100) and the ERI also had a high score (score = 43) in mental health  
42 383 problem incidence risk nomogram, which indicated that both of them were relatively important factors  
43 384 for mental health problems among the group of factory workers and miners. Our finding was consistent  
44 385 with other studies that had shown that occupational stress was a significant predictor of anxiety and was  
45 386 negatively associated with mental health. In addition, there is a high correlation between burnout and  
46 387 depression [39].

47  
48  
49  
50  
51  
52 389 In line with previous studies, working years was also an important influential factor in this study. Related  
53 390 study has shown that employment could improve patients' mental health, while unemployment could  
54 391 lead to a deterioration in mental health [40]. In China, workers' working years is an important aspect of  
55 392 employment, and researchers have studied this aspect and found that precarious employment is a source  
56 393 of stress for individuals and predisposes them to mental health problems [41]. In addition, environmental  
57 394 factors were also one of the influential factors of mental health problems in our study. Relevant studies

1  
2  
3 395 have found that exposure to air pollution is associated with increased suicide risk and depressive  
4 396 symptoms [42]. Hypertension and diabetes were the influential factors in this study. A study has shown  
5 397 that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher  
6 398 than in the non-diabetic population [43]. Furthermore, there is a recognized association between  
7 399 hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear  
8 400 [44].

9 401  
10 402 Factory workers and miners were inevitably exposed to occupational hazards such as benzene and  
11 403 asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers  
12 404 are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous  
13 405 enterprises, involving more than 30 different types of operations, of which factory workers and miners  
14 406 is the one [45]. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for  
15 407 mental health problems in this study. Our study found that the work schedules of factory workers and  
16 408 miners were vary and the phenomenon of night shifts was very common, which inevitably affected their  
17 409 normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health  
18 410 and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-  
19 411 regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could  
20 412 suffer from negative emotions such as anxiety and depression [46]. Professional title and education level  
21 413 were also important influences on mental health issues. In the workplace, generally speaking, the higher  
22 414 the professional title and education level, the higher the status of the worker in the company and the  
23 415 greater the role played in the position. The number of studies on socio-economic status and mental health  
24 416 had increased in recent years. Some of these studies have shown that major depression is higher in the  
25 417 low socio-economic status group [47]. It has also been suggested that education itself is the best indicator  
26 418 of socio-economic status [48]. Marital status was one of the influential factors for mental health problems.  
27 419 Many studies have found an association between mental health and gender, marital status, lifestyle and  
28 420 working conditions, and it has been shown that poor mental health in women is associated with divorce  
29 421 or widowhood [49]. In this study, working more than seven hours a day was a determinant factor on mental  
30 422 health problems, which was consistent with other studies that had shown that long working hours could  
31 423 have a negative impact on employees' mental health and that excessive workloads could increase workers'  
32 424 fatigue, which in turn could lead to anxiety and depression [50].

33 425  
34 426 In China, there are many problems in identifying people with mental health problems due to uneven and  
35 427 imperfect levels of medical development across regions. Some studies have shown that in mainland  
36 428 China, general practitioners, surgeons and primary health care workers often have little or no mental  
37 429 health training, which prevents them from providing basic mental health services [51]. Non-mental health  
38 430 professionals in general hospitals learn about mental illness on their own, rather than learning about it  
39 431 during their formal education [52]. Therefore, this study designed a simple and comprehensive nomogram  
40 432 to address the issue of timely detection and effective interventions for people with mental health problems,  
41 433 so that people at risk of mental health problems could easily calculate their probability of suffering from  
42 434 mental health problems without the help of medical staff. This study has several strengths. First, to our



1  
2  
3 435 knowledge, this is the first model to develop and assess the likelihood of mental health problems in a  
4 436 group of factory workers and miners. Secondly, the nomogram in this study includes demographic  
5 437 information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors  
6 438 that are closely related to factory workers and miners, allowing for a more accurate assessment of the  
7 439 risk of morbidity among them, as well as providing a methodological reference for other related studies.  
8  
9 440

## 11 441 **5. Limitations**

12 442  
13  
14 443 This study also has several limitations. Firstly, we have considered many influential factors including  
15 444 demographics, job burnout, occupational stress and occupational exposure factors, but we are still not  
16 445 certain whether all possible influences are covered. Secondly, while the robustness of our nomogram was  
17 446 extensively validated internally in the same population, external validation is lacking for other  
18 447 populations in other regions and countries. Nomogram need to be externally assessed in a wider  
19 448 population.  
20  
21 449

22 450 **Acknowledgements** The authors are grateful to all participants and investigators.

23 451  
24 452 **Contributions** Y.L., Q.L., and T.L. are responsible for conceptualization; Y.L. is responsible for  
25 453 methodology, software, formal analysis, resources, and visualization; Q.L. and T.L. are responsible for  
26 454 the original draft preparation; Q.L. and H.Y. are responsible for reviewing; Q.L. is responsible for editing;  
27 455 T.L. is responsible for supervision. Yaoqin Lu and Qi Liu contributed equally to this work.  
28  
29 456

30 457 **Funding** This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous  
31 458 Region (grant number 2020D01A27), the Postgraduate Innovation Project of Xinjiang Uyghur  
32 459 Autonomous Region (grant number XJ2021G215), the Outstanding Young Scientist Training Program  
33 460 of Urumqi Science and Technology Talent Project (grant number N/A), the Public Health and Preventive  
34 461 Medicine, the 13th Five-Year Plan Key Subject of Xinjiang Uygur Autonomous Region (grant number  
35 462 N/A), The funders were not involved in the conception, design, analysis or interpretation of this study.  
36  
37 463

38 464 **Competing interests** None declared.

39 465  
40 466 **Patient consent for publication** Not applicable.  
41  
42 467

43 468 **Ethics approval** The study was approved by the ethics committee of Urumqi Center for Disease Control  
44 469 and Prevention (20181123)  
45  
46 470

47 471 **Data availability statement** Data are available on reasonable request. The data used in this study are  
48 472 available from the corresponding authors on reasonable request.  
49  
50 473

## 51 474 **References**

52 475 [1] WHO Terminology Information System [online glossary] <http://www.who.int/health-systems->

- 1  
2  
3 476 performance/docs/glossary.html.
- 4 477 [2] Wang Y, Liu X, Qiu J, Wang H, Liu D, Zhao Z, Song M, Song Q, Wang X, Zhou Y, Wang W.  
5 478 Association between ideal cardiovascular health metrics and suboptimal health status in Chinese  
6 479 population. *Sci Rep* 2017;7:14975.
- 7 480 [3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental  
8 481 health. *Lancet*. 2007;370:859–77.
- 9 482 [4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive  
10 483 symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. *Health*  
11 484 *Place* 2016;42:37–46.
- 12 485 [5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,  
13 486 Degenhardt L, Dicker D, *et al*. Global, regional, and national incidence, prevalence, and years lived  
14 487 with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A  
15 488 systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;386:743–800.
- 16 489 [6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health? *Int J*  
17 490 *Environ Res Public Health* 2020;17:7530.
- 18 491 [7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative  
19 492 income, gender, and location. *China Econ Rev* 2018;48:83–101.
- 20 493 [8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:  
21 494 Results from the PATH Through Life Survey. *Soc Sci Med*. 2009;69:229–237.
- 22 495 [9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and  
23 496 Subjective Well-being? *J Environ Econ Manag* 2017;85:81–94.
- 24 497 [10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. *World*  
25 498 *Psychiatry*. 2015;14:36–42.
- 26 499 [11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson  
27 500 SJ. IPS Supported employment: a review. *J Vocat Rehabil* 2014;40:1–13.
- 28 501 [12] Jia Z, *et al*. Exposure to ambient air particles increases the risk of mental disorder: findings from a  
29 502 natural experiment in Beijing. *Int J Environ Res Public Health* 2018;15:160.
- 30 503 [13] Szyszkowicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency  
31 504 department visits for suicide attempts in vancouver, Canada. *Environ Health Insights* 2010;4:79–86.
- 32 505 [14] Michael J. S. International classification of sleep disorders-third edition : highlights and  
33 506 modifications. *Chest* 2014;146:1387–1394.
- 34 507 [15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial  
35 508 infarction: gender differences. *Open Nurs J* 2018;12:205–214.
- 36 509 [16] Wurm W, Vogel K, Holl A, *et al*. Depression-burnout overlap in physicians. *PLoS One*  
37 510 2016;11:e0149913.
- 38 511 [17] Porru F, Robroek S J W, Bültmann U, *et al*. Mental health among university students: The  
39 512 associations of effort-reward imbalance and overcommitment with psychological distress. *J Affect*  
40 513 *Disord* 2021;282:953-961.
- 41 514 [18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,  
42 515 and burnout among physician assistant students: a multi-institutional study. *Health Professions Edu*  
43 516 2020;6:420–427.
- 44 517 [19] Hua D, Kong Y, Li W, Han Q, Zhang X, Zhu LX, Wan SW, Liu Z, Shen Q, Yang J, He HG, Zhu J.  
45 518 Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during  
46 519 the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. *Eclinical Med*

- 520 2020;24:100424.
- 521 [20] Armah E K, Adedeji J A, Boafu B B, *et al.* Underground gold miner exposure to noise, diesel  
522 particulate matter and crystalline silica dust. *J Health Pollut* 2021; 11(29):210301.
- 523 [21] Hall N B, Blackley D J, Halldin C N, *et al.* Current review of pneumoconiosis among US coal  
524 miners. *Curr Environ Health Rep*, 2019; 6(3):137-147.
- 525 [22] Wickramatillake B A, Fernando M A, Frank A L. Prevalence of asbestos-related disease among  
526 workers in sri lanka. *Ann Glob Health*, 2019, 85(1):108.
- 527 [23] Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP. The estimation of the number of  
528 underground coal miners and the annual dose to coal miners in China. *Health Phys* 2007;93:127–  
529 132.
- 530 [24] Yong X, Gao X, Zhang Z, *et al.* Associations of occupational stress with job burn-out, depression  
531 and hypertension in coal miners of Xinjiang, China: A cross-sectional study. *BMJ open* 2020;10:  
532 e036087.
- 533 [25] P. Bech, J. Bille, S. B. Møller, L. C. Hellström, S. D. Østergaard. Psychometric validation of the  
534 Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal  
535 sensitivity. *J Affect Disord* 2014;160:98–103.
- 536 [26] J. Zhang, X. Zhang. Chinese college students' SCL-90 scores and their relations to the college  
537 performance. *Asian J Psychiatr* 2013;6:134–140.
- 538 [27] Liang S, Zhang J, Zhao Q, *et al.* Incidence trends and risk prediction nomogram for suicidal attempts  
539 in patients with major depressive disorder. *Front Psychiatry*, 2021, 12: 644038.
- 540 [28] Kan S K, Chen N N, Zhang Y L. Predicting the risk of suicide attempt in a depressed population:  
541 Development and assessment of an efficient predictive nomogram. *Psychiatry Res*, 2022, 310:  
542 114436.
- 543 [29] Lu Y, Zhang Z, Yan H, *et al.* Effects of occupational hazards on job stress and mental health of  
544 factory workers and miners: A propensity score analysis. *BioMed Res Int*, 2020, 2020: 1754897.
- 545 [30] Li X, Jiang T, Sun X, *et al.* The relationship between occupational stress, musculoskeletal disorders  
546 and the mental health of coal miners: The interaction between bdnf gene, tph2 gene polymorphism  
547 and the environment. *J Psychiatr Res*, 2021, 135: 76-85.
- 548 [31] Derogatis L, Lipman R.S, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary  
549 report. *Psychopharmacol Bull* 1973;9:13–28.
- 550 [32] Crespo-Maraver M, Doval E, Fernã N.J, Gimã Nez-Salinas J, Prat G, Bonet P. Caregiver's health:  
551 adaption and validation in a Spanish population of the Experience of Caregiving Inventory  
552 (ECI) Gac. *Sanit* 2018;33:348–355.
- 553 [33] Dang W, Xu Y, Ji J, *et al.* Study of the scl-90 scale and changes in the chinese norms. *Front*  
554 *Psychiatry*, 2020, 11: 524395.
- 555 [34] Siegrist J. Adverse health effects of high-effort/low-reward conditions. *J Occup. Health Psychol.*  
556 1996;1:27–41.
- 557 [35] Siegrist J, Wege N, Pühlhofer F, Wahrendorf M. A short generic measure of work stress in the era  
558 of globalization: Effort-reward imbalance. *Int Arch Occup Environ Health* 2009;82:1005–1013.
- 559 [36] Zhang Z, Lu Y, Yong X, *et al.* Effects of occupational radiation exposure on job stress and job  
560 burnout of medical staff in xinjiang, China: A cross-sectional study. *Med Sci Monit* 2020;26:  
561 e927848.
- 562 [37] Freudenberger HJ. Staff burnout. *J Soc Issues* 1974;30:159–165.
- 563 [38] Wang L, Chen L, Jia F, *et al.* Risk factors and prediction nomogram model for psychosocial and

- 564 behavioural problems among children and adolescents during the covid-19 pandemic: A national  
 565 multicentre study: Risk factors of childhood psychosocial problems. *J Affect Disord*, 2021, 294:  
 566 128-136.
- 567 [39] Occupational Stress and Employees Complete Mental Health: A Cross-Cultural Empirical Study
- 568 [40] Knapp M and Wong G. Economics and mental health: the current scenario. *World Psychiatry*  
 569 2020;19:3–14.
- 570 [41] Benach, J, Vives, A, Amable, M, Vanroelen, C, Tarafa, G, & Muntaner, C. Precarious employment:  
 571 understanding an emerging social determinant of health. *Annu Rev Public Health* 2014;35:229-53.
- 572 [42] Bakian A. V. *et al.* Acute air pollution exposure and risk of suicide completion. *Am J Epidemiol*  
 573 2015;181:295–303.
- 574 [43] Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type  
 575 1 diabetes: systematic literature review. *Diabet Med* 2006;23:445–448.
- 576 [44] Gilsanz P, Karter AJ, Beeri MS, Quesenberry CP, Whitmer RA. The bidirectional association  
 577 between depression and severe hypoglycemic and hyperglycemic events in type 1  
 578 diabetes. *Diabetes Care* 2018;41:446–452.
- 579 [45] Lu Y, Zhang Z, Yan H, *et al.* Effects of occupational hazards on job stress and mental health of  
 580 factory workers and miners: A propensity score analysis. *BioMed Res Int* 2020;2020:1754897.
- 581 [46] Shi L, Liu Y, Jiang T, *et al.* Relationship between mental health, the clock gene, and sleep quality  
 582 in surgical nurses: A cross-sectional study. *BioMed Res Int* 2020;2020:4795763.
- 583 [47] Sallis JF, Saelens BE, Frank LD *et al.* Neighborhood built environment and income: examining  
 584 multiple health outcomes. *Soc Sci Med* 2009;68:1285–1293.
- 585 [48] Winkleby MA, Jatulis DE, Frank E *et al.* Socioeconomic status and health: how education, income,  
 586 and occupation contribute to risk-factors for cardiovascular disease. *Am J Public*  
 587 *Health* 1992;82:816–820.
- 588 [49] Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis P.N, Mavreas V. Prevalence  
 589 and sociodemographic associations of common mental disorders in a nationally representative  
 590 sample of the general population of Greece. *BMC Psychiatry* 2013;13:163.
- 591 [50] Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG *et al.* Long  
 592 working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II  
 593 study. *Psychol Med* 2011;41:2485–2494.
- 594 [51] Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, *et al.* Prevalence, treatment, and associated  
 595 disability of mental disorders in four provinces in China during 2001–05: an epidemiological  
 596 survey. *Lancet* 2009;373:2041–2053.
- 597 [52] Wu Q, Luo X, Chen S, *et al.* Mental health literacy survey of non-mental health professionals in six  
 598 general hospitals in hunan province of China. *PloS one* 2017;12:e0180327.

## 600 Figure legends

601 **Fig.1. Flow diagram of the participants involved in this study**

602

603 **Fig.2. Feature selection using the LASSO binary logistic regression model.** (A) Feature selection for the LASSO  
 604 binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against  
 605 lambda by validating the optimal parameter lambda in the LASSO model. Dotted vertical lines were drawn based  
 606 on 1 SE of the minimum criteria (the 1-SE criteria). (B) Feature selection was performed using the LASSO binary  
 607 logistic regression model. A Coefficient profile was plotted based on the lambda series in Figure 1(A), and 12

1  
2  
3 608 features with non-zero coefficients were selected by optimal lambda.

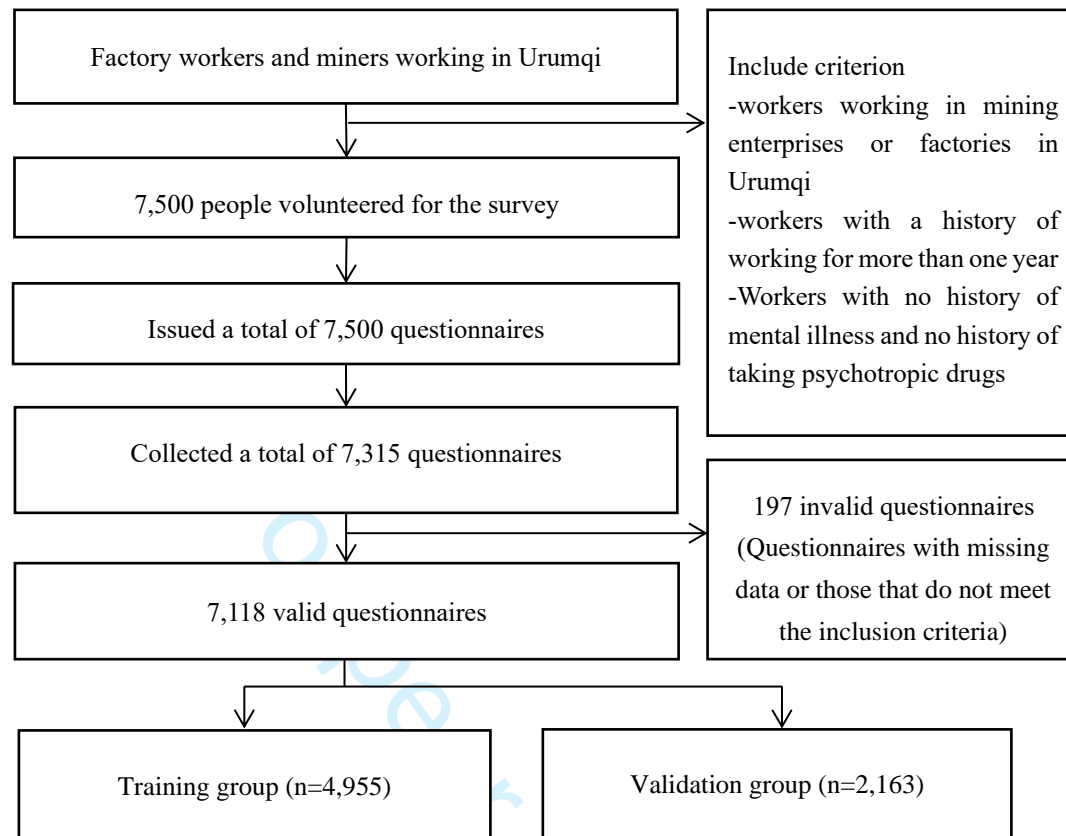
4 609  
5 610 **Fig.3. The forest plot of the OR of the selected feature.**

6 611  
7  
8 612 **Fig.4. Developed mental health problems incidence risk nomogram.** The mental health problems incidence risk  
9 613 nomogram was developed in the array, with education, professional title, age, CMBI, ERI, asbestos dust,  
10 614 hypertension, diabetes, working hours per day, working years, marital status, and work schedule incorporated.

11 615  
12 616 **Fig.5. Calibration curves of the mental health problems incidence risk nomogram prediction in validation**  
13 617 **group.** The x-axis represents the predicted risk of mental health problems. y-axis represents the actual diagnosed  
14 618 risk of mental health problems. The diagonal dashed line represents the perfect prediction of the ideal model. The  
15 619 solid lines represent the performance of the column plots, where closer to the diagonal dashed line indicates a better  
16 620 prediction.

17 621  
18 622 **Fig.6. ROC curves for training and validation groups.** The y-axis represents the true positive rate of risk  
19 623 prediction. The x-axis represents the false positive rate of risk prediction. The ROC curves for the training and  
20 624 validation groups are shown in black and red.

21 625  
22  
23 626 **Fig.7. Decision curve analysis for mental health problems incidence risk nomogram.** The y-axis measures the  
24 627 net benefit. The solid red line represents nomogram of the risk of developing a mental health problem. The light blue  
25 628 dashed line represents the hypothesis that all participants were diagnosed with a mental health problem. The black  
26 629 dashed line represents the hypothesis that there is no risk of a mental health problem. The DCA showed that using  
27 630 this mental health problem incidence risk nomogram in the current study to predict mental health problem incidence  
28 631 risk increase in benefit than the intervention all patients or no intervention all patient if the threshold probability of  
29 632 a patients and a doctor is >10%.



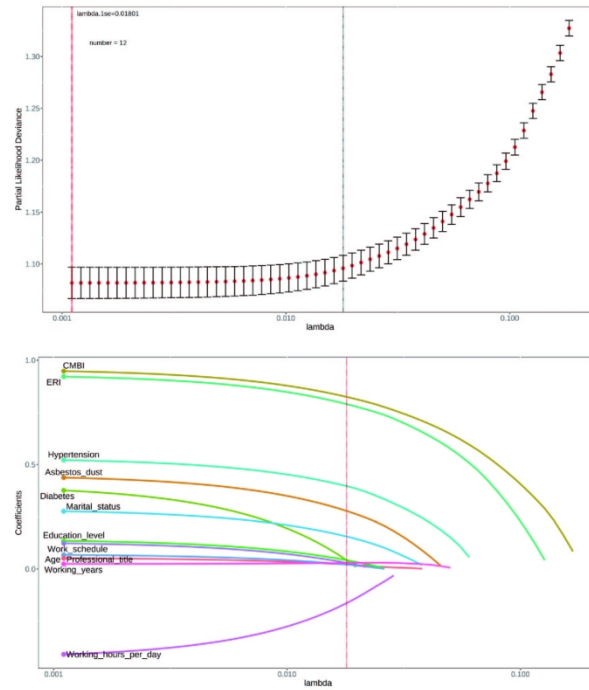
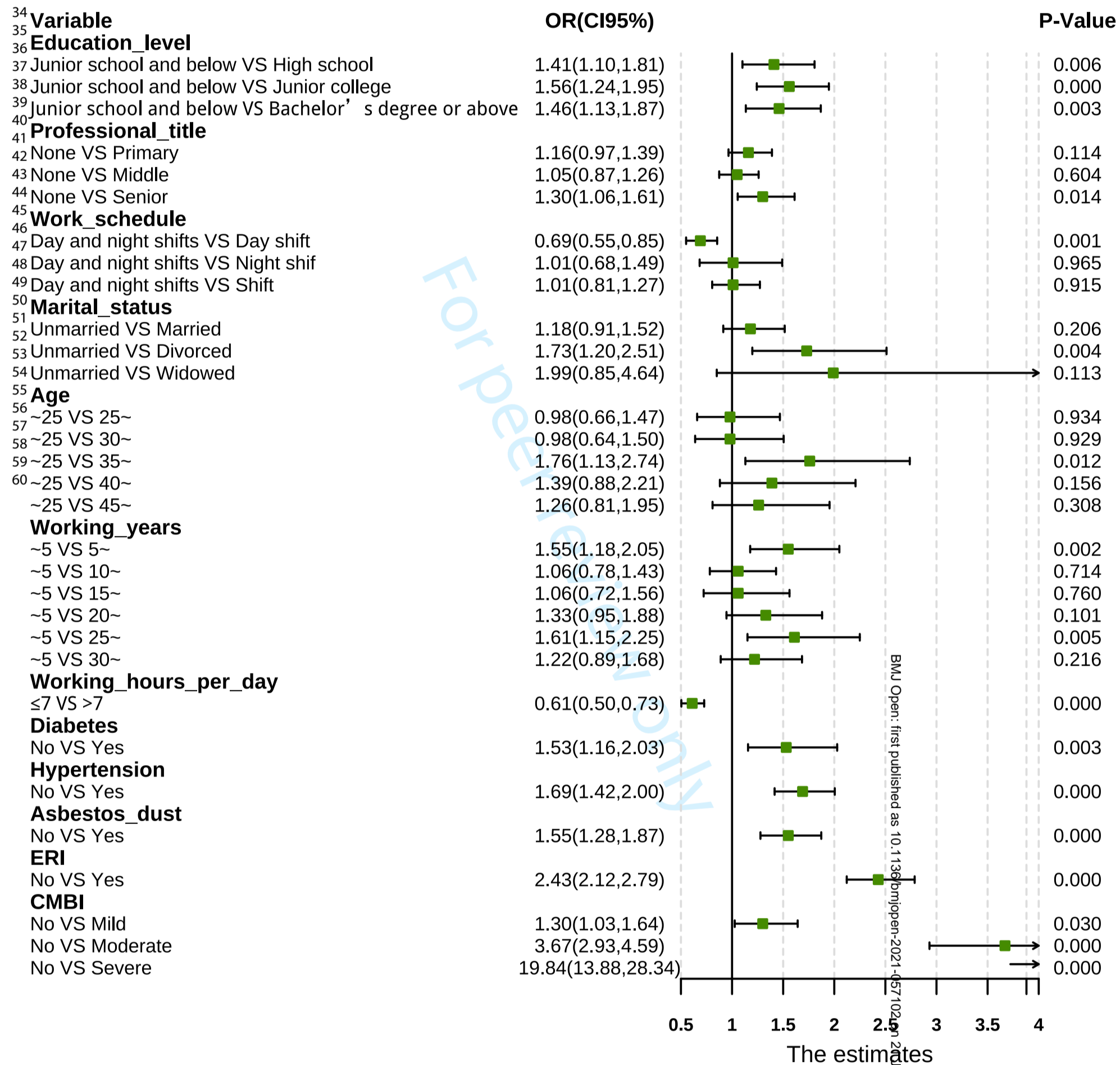


Fig.2. Feature selection using the LASSO binary logistic regression model. (A) Feature selection for the LASSO binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against lambda by validating the optimal parameter lambda in the LASSO model. Dotted vertical lines were drawn based on 1 SE of the minimum criteria (the 1-SE criteria). (B) Feature selection was performed using the LASSO binary logistic regression model. A Coefficient profile was plotted based on the lambda series in Figure 1(A), and 12 features with non-zero coefficients were selected by optimal lambda.

210x297mm (300 x 300 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33



BMJ Open: first published as 10.1136/bmjopen-2021-025710 on 20 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.



Points

CMBI

ERI

Marital status

Age

Hypertension

Working days per week

Working years

Education level

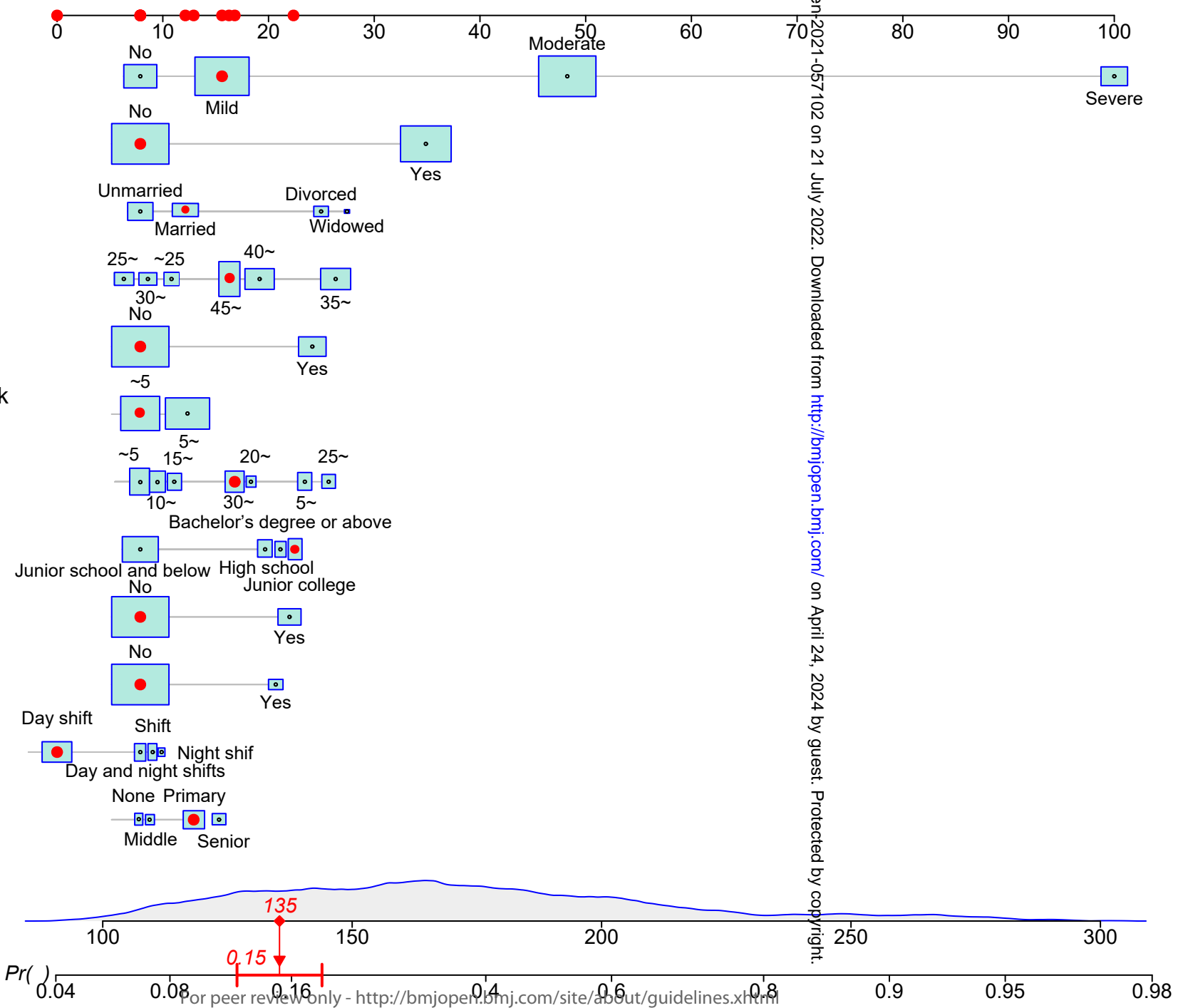
Asbestos dust

Diabetes

Work schedule

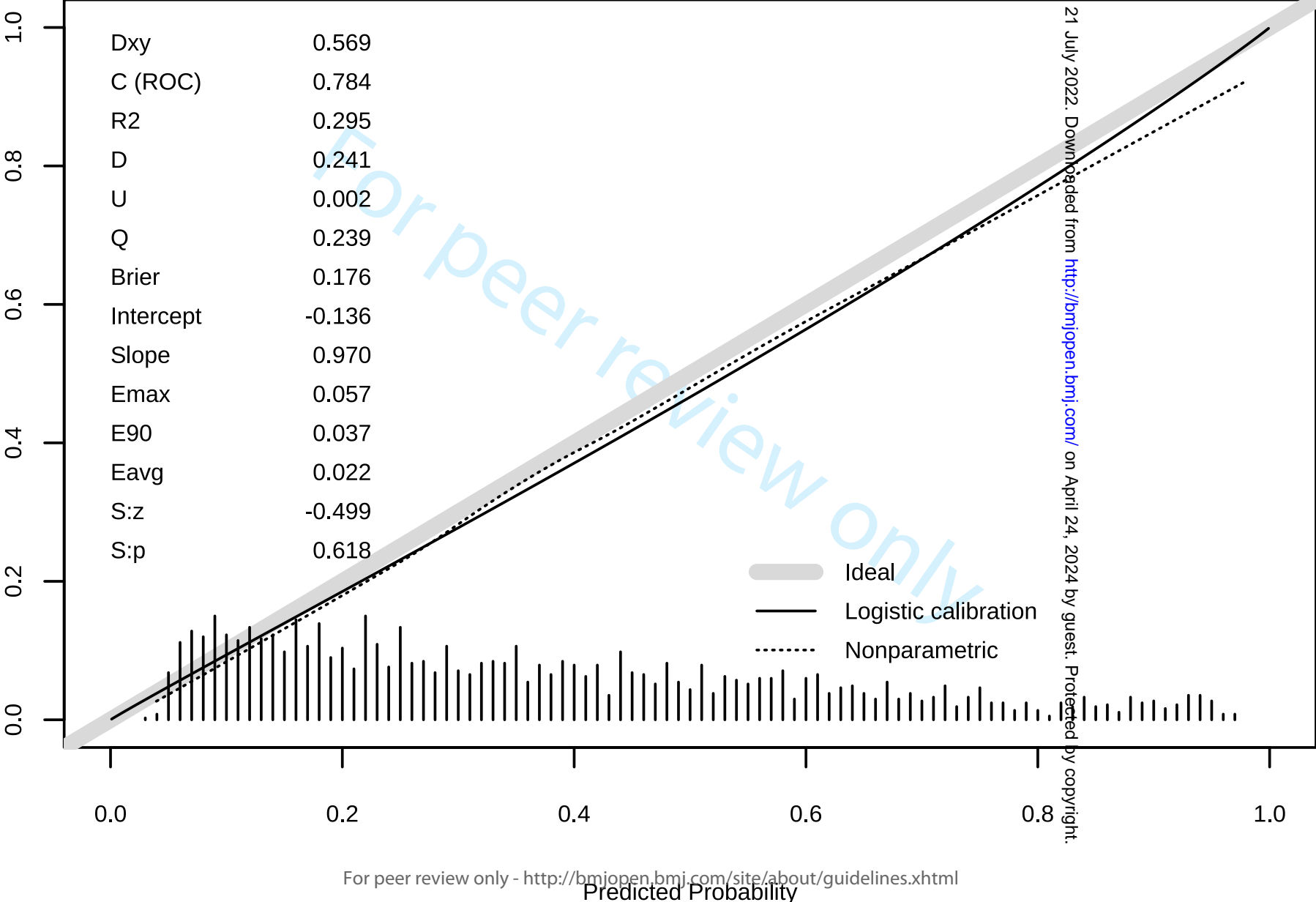
Professional title

**Total points**

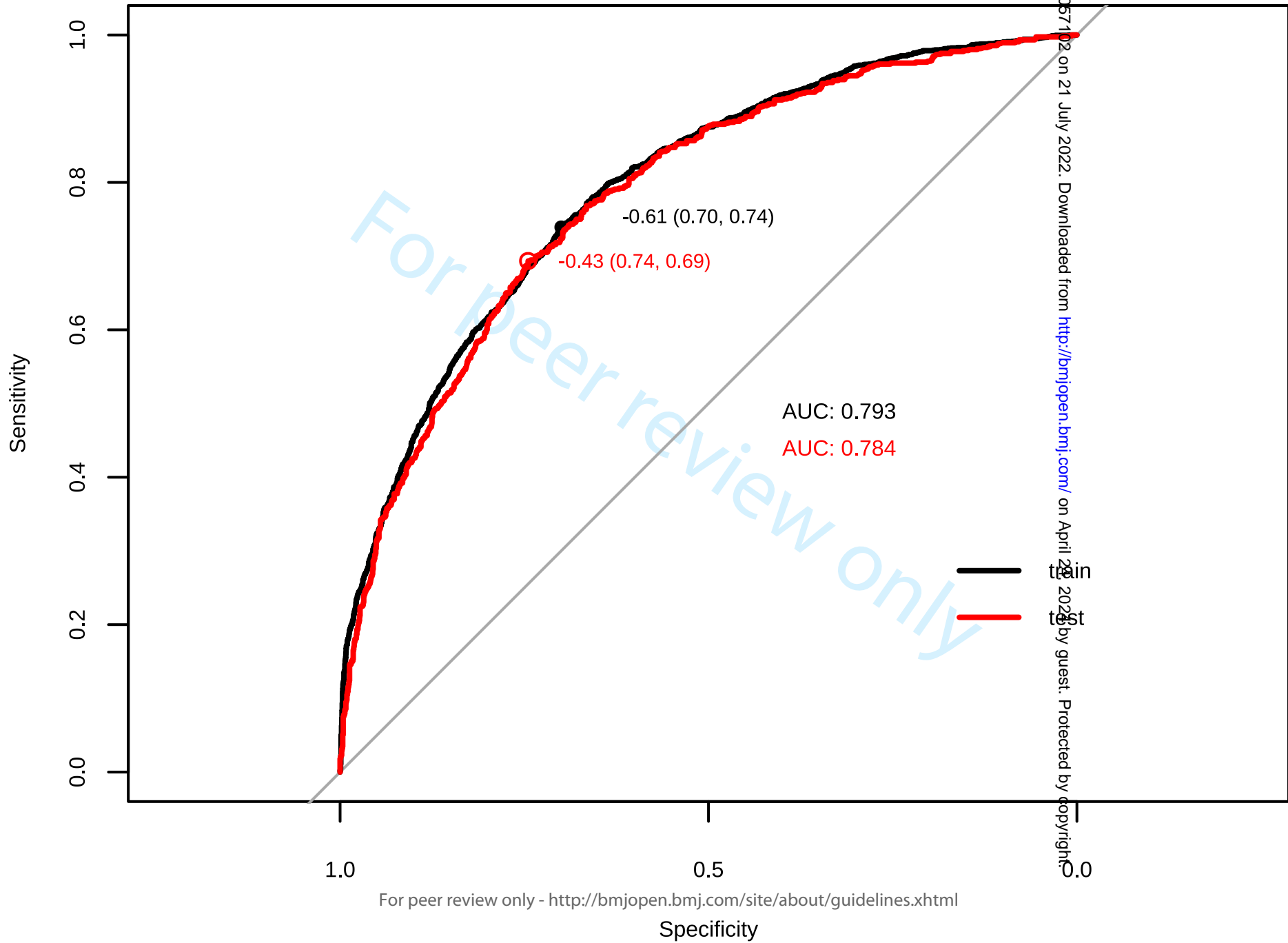


6/bmjopen-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

pen-2021-057102 or 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.

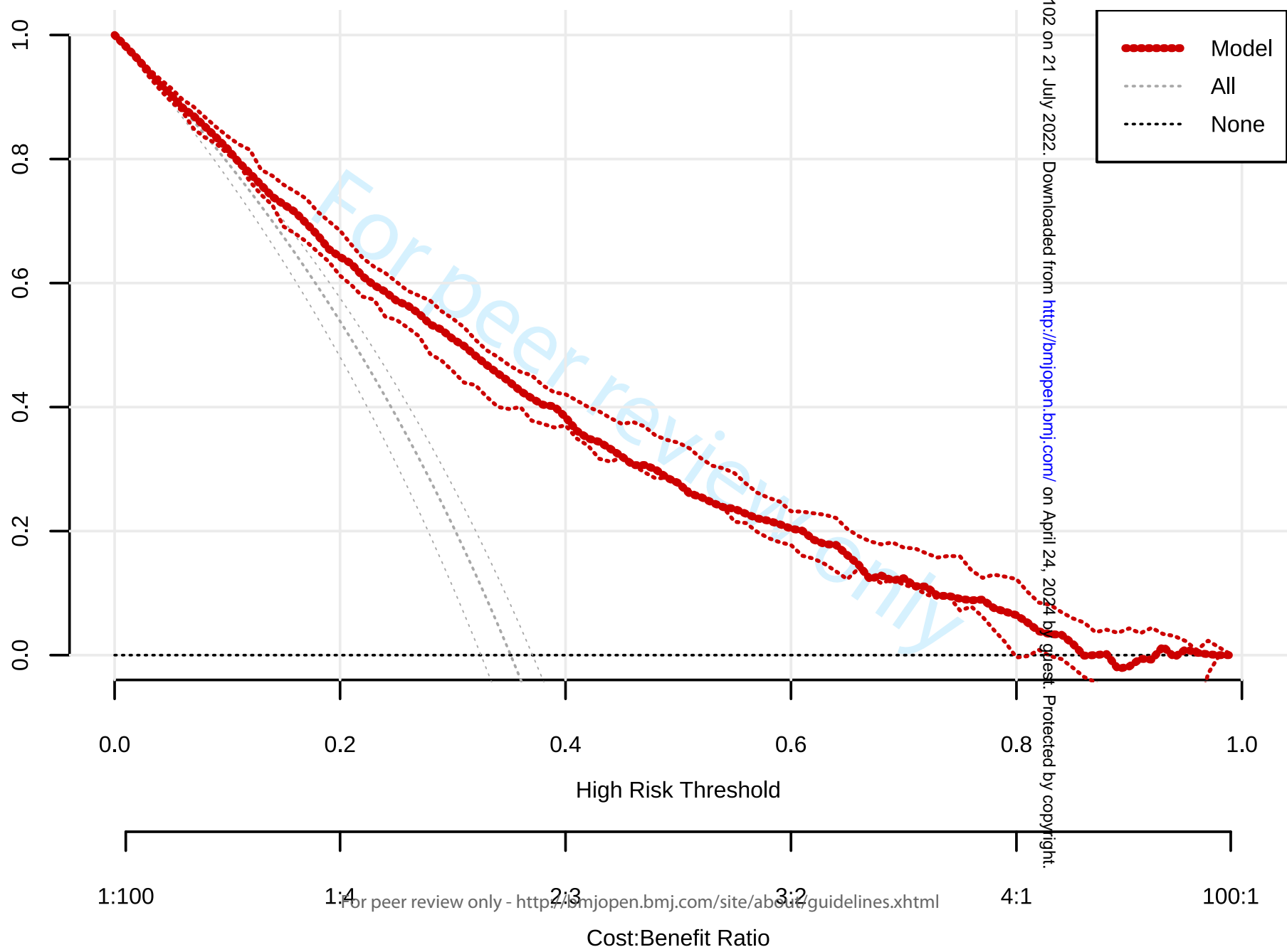


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44



per-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 28, 2024 by guest. Protected by copyright.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44



per-2021-057102 on 21 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 24, 2024 by guest. Protected by copyright.