

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 (<u>http://bmjopen.bmj.com</u>).

If you have any questions on BMJ Open's open peer review process please email <u>info.bmjopen@bmj.com</u>

# **BMJ Open**

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

Journal:	BMJ Open
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





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 <u>licence</u>.

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 <u>Creative Commons</u> 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.

relievon

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

1 2

#### 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.

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

36

3	
4	
5	
6	
6 7 8	
, Q	
0	
9	
10	
11	
12 13	
13	
14	
15	
16 17 18	
17	
18	
19	
20	
20 21	
22	
22	
24	
25	
26	
27	
24 25 26 27 28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
37 38	
39	
40	
41	
42	
43	
44	
45	
46	
40	
48 49	
49 50	
51	
52	
53	
54	
55	
56	
57	
58	
59	

#### 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 health41 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 to43 factory workers and miners for identifying workers at risk of mental health problems.

3. The results of this study showed good agreement and good discrimination between predictions andobservations.

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.

5. While the robustness of our nomogram was extensively validated internally in the same population,external validation was lacking for other populations in other regions and countries.

#### 52 1. Introduction

51

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 non-communicable diseases and even causing unintentional or intentional harm to others <sup>[3]</sup>, but can also 60 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>.

60

66

67 As researchers around the world have delved into the field of mental health, factors such as gender, income levels, environment and education have been found to be associated with people's mental health 68 69 problems [7-10]. 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 [13]. Furthermore, other factors associated with mental health include sleep, diabetes, coronary artery disease and cardiovascular disease [14-15]. It is worth noting that job burnout and occupational stress 77 78 are closely linked to mental health. Job burnout is an exhaustion state of physical and psychological that

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

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 [18-19]. China 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.

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

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 13	126	Urumqi region, who were recruited using a whole-group sampling method. A total of 7,500 participants
14	120	
15 16		in the Urumqi were surveyed from January to May 2019, covering all districts and counties in the Urumqi
10	128	region, including Tianshan District, Shaibak District, Xinshi District, Shuimogou District, Toutunhe
18	129	District, Dabancheng District, Middong District and Urumqi County.
19 20	130	
20 21	131	The exclusion criteria were the following: (I) factory workers and miners in non-Urumqi area, (II)
22	132	working history of factories and mining enterprises less than 1 year, (III) a confirmed diagnosis of a
23 24	133	mental health problem and a history of treatment and use of psychotropic medication. Questionnaires
24 25	134	with missing data were also excluded from the analysis based on discussion and agreement among the
26	135	subject members. A total of 7,500 questionnaires were distributed and 7,315 questionnaires were returned,
27 28	136	representing a return rate of 97.5%. After checking the validity and integrity of the questionnaires, 7,118
28 29	137	questionnaires were confirmed as valid, with an effective rate of 97.3%. All participants understood the
30	138	purpose of the study and voluntarily participated in the study.
31 32	139	
33	140	<ul> <li>2.2. Research Methods</li> <li>2.2.1. Assessment of mental health</li> </ul>
34	141	2.2. Research Methous
35 36		
37	142	2.2.1. Assessment of mental health
38	143	
39 40	144	The SCL-90, designed by Derogatis and his colleagues, was widely used in the mental health field <sup>[24]</sup> ,
41	145	which contains 90 items across nine dimensions: somatization, obsessive-compulsive symptoms,
42	146	interpersonal sensitivity, depression, anxiety, hostility, horror, bigotry and mental illness. The SCL-90
43 44	147	has been used extensively in previous studies and has relatively high reliability and validity [25]. The
45	148	questionnaire uses a Likert 5-point scale, with a score of 0 point indicating none and 4 points indicating
46 47	149	severe. A total score above 160, a score above 2 on any item, or a positive item above 43 indicates the
47 48	150	presence of a psychological abnormality. In this survey, Cronbach $\alpha$ was 0.99, the half-reliability
49	151	coefficient was 0.98, and the KMO was 0.994.
50 51	152	
52	153	2.2.2. Assessment of occupational stress
53	154	1
54 55	155	This survey evaluated occupational stress in factory workers and miners through the Effort-Reward
55 56	156	Imbalance (ERI) model developed by Siegrist <sup>[26]</sup> . The ERI scale consists of three subscales: effort (E, 6
57		
58 59	157	items), reward (R, 11 items) and over commitment (6 items), for a total of 23 items. A Likert 5-level
60	158	scoring method (1, "highly disagree" to 5, "highly agree") is used to grade the items in the questionnaire

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

#### 165 2.2.3. Assessment of job burnout

In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess job burnout, which has good reliability and validity <sup>[28]</sup>. CMBI is composed of 15 items in three dimensions: emotional exhaustion (5 items), depersonalization (5 items) and reduced personal accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete compliance and 7 points indicating complete non-compliance. According to the critical value (emotional exhaustion  $\geq 25$ , dependent dependence of a constraint dependence of a co burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to or above the critical value), moderate (any two aspects are equal to or higher than the critical values). and severe (three aspects are equal to or higher than the critical values) <sup>[29]</sup>. In this survey, Cronbach  $\alpha$ was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.

#### 178 2.2.4. Candidate predictors

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

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 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.

1

1	
2	
3	
4	
5	
6	
7	
, 8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
21 22	
23	
24	
25	
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	
54 55	
56	
57	
58	
59	
60	

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 (α) 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	57,210/ of the nonvelation was over 25 years of any and 79,220/ of the subjects were married showing

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

57					
58	Variables	Total ( $n = 7118$ )	train (n = 4955)	test ( $n = 2163$ )	р
59					
60					

Page	8	of	38
------	---	----	----

Sex, n (%)				
Male	4649 (65)	3216 (65)	1433 (66)	0.2
Female	2469 (35)	1739 (35)	730 (34)	
Ethnicity, n (%)				
Han	5762 (81)	3982 (80)	1780 (82)	0.0
Other	1356 (19)	973 (20)	383 (18)	
Education level, n (%)				
Junior high school and below	1143 (16)	804 (16)	339 (16)	0.7
High school	1406 (20)	988 (20)	418 (19)	
Junior college	2933 (41)	2038 (41)	895 (41)	
Bachelor's degree or above	1636 (23)	1125 (23)	511 (24)	
Professional title, n (%)				
None	2854 (40)	1983 (40)	871 (40)	0.9
Primary	1644 (23)	1149 (23)	495 (23)	
Middle	1618 (23)	1133 (23)	485 (22)	
Senior	1002 (14)	690 (14)	312 (14)	
Work schedule, n (%)				
Day shift	3986 (56)	2801 (57)	1185 (55)	0.5
Night shift	270 (4)	187 (4)	83 (4)	
Shift	2058 (29)	1412 (28)	646 (30)	
Day and night shifts	804 (11)	555 (11)	249 (12)	
Marital status, n (%)				
Unmarried	1104 (16)	762 (15)	342 (16)	0.2
Married	5575 (78)	3906 (79)	1669 (77)	
Divorced	390 (5)	255 (5)	135 (6)	
Widowed	49 (1)	32 (1)	17(1)	
Monthly income (yuan), n (%)				
<3000	1799 (25)	1246 (25)	553 (26)	0.9
3000~	2418 (34)	1682 (34)	736 (34)	
4000~	1600 (22)	1125 (23)	475 (22)	
5000~	752 (11)	520 (10)	232 (11)	
6000~	288 (4)	201 (4)	87 (4)	
7000~	148 (2)	106 (2)	42 (2)	
8000~	113 (2)	75 (2)	38 (2)	
Age (years), n (%)				
<25	431 (6)	297 (6)	134 (6)	0.1
25~	786 (11)	519 (10)	267 (12)	
30~	956 (13)	684 (14)	272 (13)	
35~	866 (12)	617 (12)	249 (12)	

2					
3 4	40~	849 (12)	588 (12)	261 (12)	
4 5	45~	3230 (45)	2250 (45)	980 (45)	
6	Working years (years), n (%)				
7 8	<5	1170 (16)	794 (16)	376 (17)	0.248
9	5~	1065 (15)	736 (15)	329 (15)	
0  1	10~	997 (14)	721 (15)	276 (13)	
2	15~	389 (5)	273 (6)	116 (5)	
3	20~	763 (11)	538 (11)	225 (10)	
4 5	25~	1293 (18)	878 (18)	415 (19)	
б	30~	1441 (20)	1015 (20)	426 (20)	
7 8	Labor contracts, n (%)				
o 9	Signed	6641 (93)	4624 (93)	2017 (93)	0.955
)	Unsigned	477 (7)	331 (7)	146 (7)	
1 2	Working hours per day (hours), n (%)				
3	≤7	1161 (16)	814 (16)	347 (16)	0.712
4 5	>7	5957 (84)	4141 (84)	1816 (84)	
6	Working days per week (days), n (%)				
7	≤5	4442 (62)	3107 (63)	1335 (62)	0.446
8 9	>5	2676 (38)	1848 (37)	828 (38)	
0	Diabetes, n (%)	2010 (30)		020 (00)	
1 2	Yes	429 (6)	298 (6)	131 (6)	0.988
3	No	6689 (94)	4657 (94)	2032 (94)	0.900
4	Hypertension, n (%)	(+()	()+()	2032 (94)	
5 6	Yes	1330 (19)	929 (19)	401 (19)	0.861
7	No	5788 (81)	4026 (81)	1762 (81)	0.001
3 9	Coal dust, n (%)	5788 (81)	4020 (81)	1702 (81)	
0	Yes	377 (5)	244 (5)	133 (6)	0.039
1	No			2030 (94)	0.039
2 3		6741 (95)	4711 (95)	2030 (94)	
4	Silica dust, n (%)	720 (10)	522 (11)	207(10)	0.222
5 6	Yes	730 (10)	523 (11)	207 (10)	0.223
7	No	6388 (90)	4432 (89)	1956 (90)	
8 9	Asbestos dust, n (%)				
9 0	Yes	981 (14)	691 (14)	290 (13)	0.570
1	No	6137 (86)	4264 (86)	1873 (87)	
2 3	Benzene, n (%)				
4	Yes	1981 (28)	1360 (27)	621 (29)	0.287
5	No	5137 (72)	3595 (73)	1542 (71)	
6 7	Lead, n (%)				
58	Yes	373 (5)	246 (5)	127 (6)	0.128
59 50					
50		8			

	No	6745 (95	) 4709 (9:	5)	2036 (94)			
Noise, n (	(%)							
	Yes	4942 (69	) 3420 (69	9)	1522 (70)	0.27		
	No	2176 (31	) 1535 (3	1)	641 (30)			
Brucellos	sis, n (%)							
	Yes	121 (2)	86 (2)		35 (2)	0.80		
	No	6997 (98	) 4869 (98	8)	2128 (98)			
ERI, n (%	(o)							
	Yes	3147 (44	) 2173 (44	4)	974 (45)	0.37		
	No	3971 (56	) 2782 (50	6)	1189 (55)			
CMBI, n	(%)							
	No	959 (13)	674 (14	•)	285 (13)	0.03		
	Mild	2667 (37	) 1813 (3'	7)	854 (39)			
	Moderate	2900 (41	) 2031 (4	1)	869 (40)			
	Severe	592 (8)	437 (9)	)	155 (7)			
236								
237	<b>3.2.</b> Feature selection							
238								
239	The lambda was smallest at 0.01	801 as seen f	rom the lasso results whe	en there were	12 characteristics,			
240	which were education, profession	al title, age, C	CMBI, ERI, asbestos dust,	, hypertension	, diabetes, working			
241	hours per day, working years, mar	rital status, an	d work schedule based or	n the results of	the questionnaires			
242	on demographics, occupational st	ress, job burr	nout and occupational exp	posure factors	(Figure 1).			
243 244	3.3. Results of logistic regression	n model						
245								
246	The 12 features obtained from t	he LASSO r	egression were incorpor-	ated into a m	ultivariate logistic			
247	regression model and the regress	sion results w	vere shown in Table 2. If	t was clear fro	om the results that			
248	education, professional title, age,	CMBI, ERI,	asbestos dust, hypertens	ion, diabetes,	working hours per			
249	day, working years, marital status	s, and work so	chedule were independen	t determinants	s of risk for mental			
250	health problems. In addition, there	e was no evic	lence of multicollinearity	between the c	ovariates included			
251	in the model. The forest plot show	ved that the so	elected 12 features all cor	ntain items wit	th $P < 0.05$ , among			
252	which the degree of severe of CM	BI (OR, 19.84	4; 95% CI, 13.88-28.34; p	o < 0.001) had	the greatest impact			
253	on the risk of mental health problems among factory workers and miners (Figure 2).							
254	254 Table 2 Predictive factors of risk for mental health problems among factory workers and miners							
	Variable	β S.I	E. OR(CI95%)	Wald	Р	VIF		
	Intercept	-2.33 0.2	0.10(0.06,0.16)	-9.357	0	-		
Educat	ion level							
Junior	school and below VS High school	0.34 0.1	3 1.41(1.10,1.81)	2.727	0.006**	2.28		
			0					
			9					

#### **BMJ** Open

Junior school and below VS Junior					-0.001***	
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 shif	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			(			
to onling hours per aug					- 0 001444	
$\leq$ 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 VC Voc	0.52	0.00	1 (0(1 42 2 00)	E 00E	< 0.001***	1 1 1
No VS Yes	0.52	0.09	1.69(1.42,2.00)	5.885		1.11
Asbestos dust						
No VS Yes	0.44	0.10	1.55(1.28,1.87)	4.474	<b>&lt;</b> 0.001***	1.03
ERI						
			10			

CMBI	No VS Yes	0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***
CMDI	No VS Mild	0.26	0.12	1.30(1.03,1.64)	2.175	0.003**
	No VS Moderate	1.30	0.11	3.67(2.93,4.59)	11.361	<b>&lt;</b> 0.001***
	No VS Severe	2.99	0.18	19.84(13.88,28.34)	16.41	<b>&lt;</b> 0.001***
255 256	Note: $\beta$ is the regression coeffic	ient. "***" inc	licates P<	:0.001, "**" indicates P<0	0.01, "*" indio	cates P<0.05.
257	3.4. Development of an ind	ividualized p	oredictio	on model		
258						
259	Based on the results of the	multivariate	analysis,	predictors such as edu	cation, pro	fessional title, age
260	CMBI, ERI, asbestos dust, h	ypertension,	diabetes,	working hours per day	, working y	ears, marital status
261	and work schedule were in	cluded in the	e nomog	gram. A model incorpo	orating the	above independen
262	predictors was developed an	d represented	l as a no	mogram in Figure 3. E	ach variable	in nomogram wa
263	assigned a score, and the	cumulative s	sum of	each 'point' was the 't	otal score'.	The "total score"
264	corresponded to the "predic	table likelih	ood", w	hich was the predicted	l probability	of mental health
265	problems among factory wor	kers and mir	iers as su	ggested by our design	of the nomo	gram.
266						
267	As an example of the use of	nomogram:	a randor	nly selected sample fro	m the traini	ng group, one witl
268	no professional title, day shi	ft, no diabete	es or hyp	ertension, Junior colleg	ge, <5 of we	orking years, >7 o
269	working hours per day, marr	ied, no expos	sed to as	bestos dust, <25 years	of age, no E	RI, mild of CMBI
270	with a calculated total score	e of 174 and	a corres	ponding risk probabili	ty of 8.27%	for mental healt
271	problems.					
272						
272 273	3.5 The validation of calibr	ation				
	3.5 The validation of calibr	ation				
273	<b>3.5 The validation of calibr</b> Model validation was carrie		validatio	on group. The prediction	on accuracy	of the model wa
273 274		d out in the		• • •		
273 274 275	Model validation was carrie	d out in the ) The Brier s	score for	overall performance,	which asses	ssed the difference
273 274 275 276	Model validation was carrie assessed by two aspects. (1	d out in the ) The Brier s cted values, y	score for with value	e overall performance, ues closer to 0 indication	which asses	edictive ability. (2
273 274 275 276 277	Model validation was carrie assessed by two aspects. (1 between observed and predi	d out in the ) The Brier s cted values, y for modal cal	score for with value ibration,	voverall performance, ues closer to 0 indication which assessed the ag	which asses ng better pro- reement bet	ssed the difference edictive ability. (2 ween the observed
273 274 275 276 277 278	Model validation was carrie assessed by two aspects. (1 between observed and predi The calibration slope used f	d out in the ) The Brier s cted values, for modal cal alues closer to	score for with valu ibration, o 1 indic	e overall performance, ues closer to 0 indicating which assessed the ag ating better performance	which asses ng better pro- reement bet e. The accur	ssed the difference edictive ability. (2 ween the observed racy measurement
273 274 275 276 277 278 279	Model validation was carrie assessed by two aspects. (1 between observed and predi The calibration slope used f and predicted values, with va	ed out in the ) The Brier s cted values, v for modal cal alues closer to validated by t	score for with valu ibration, to 1 indic he mode	voverall performance, ues closer to 0 indication which assessed the ag ating better performance of with a Brier score of 0	which asses ng better pro- reement bet e. The accur 0.176 and a	edictive ability. (2 ween the observed racy measurement calibration slope o
273 274 275 276 277 278 279 280	Model validation was carrie assessed by two aspects. (1 between observed and predi The calibration slope used f and predicted values, with va for the bias correction were v	ed out in the ) The Brier s cted values, v for modal cal alues closer to validated by t	score for with valu ibration, to 1 indic he mode	voverall performance, ues closer to 0 indication which assessed the ag ating better performance of with a Brier score of 0	which asses ng better pro- reement bet e. The accur 0.176 and a	edictive ability. (2 ween the observed racy measurements calibration slope o
273 274 275 276 277 278 279 280 281	Model validation was carrie assessed by two aspects. (1 between observed and predi The calibration slope used f and predicted values, with va for the bias correction were v	d out in the ) The Brier s cted values, y for modal cal alues closer to validated by t ). The predic	score for with valu ibration, to 1 indic he mode	voverall performance, ues closer to 0 indication which assessed the ag ating better performance of with a Brier score of 0	which asses ng better pro- reement bet e. The accur 0.176 and a	edictive ability. (2 ween the observed racy measurement calibration slope o
273 274 275 276 277 278 279 280 281 282	Model validation was carried assessed by two aspects. (1) between observed and predic The calibration slope used for and predicted values, with va- for the bias correction were va- 0.970, respectively (Figure 4)	d out in the ) The Brier s cted values, y for modal cal alues closer to validated by t ). The predic	score for with valu ibration, to 1 indic he mode	voverall performance, ues closer to 0 indication which assessed the ag ating better performance of with a Brier score of 0	which asses ng better pro- reement bet e. The accur 0.176 and a	edictive ability. (2 ween the observed racy measurement calibration slope o
273 274 275 276 277 278 279 280 281 282 283	Model validation was carried assessed by two aspects. (1) between observed and predic The calibration slope used for and predicted values, with va- for the bias correction were va- 0.970, respectively (Figure 4)	d out in the ) The Brier s cted values, v or modal cal alues closer to validated by t .). The predic <b>mination</b>	score for with valu ibration, o 1 indic he mode tion accu	e overall performance, ues closer to 0 indicatin which assessed the ag ating better performance with a Brier score of 0 uracy of the model was	which asses ng better pro- reement bet e. The accur 0.176 and a relatively h	ssed the difference edictive ability. (2 ween the observed racy measurement calibration slope o igh.
273 274 275 276 277 278 279 280 281 282 283 283 284	Model validation was carrie assessed by two aspects. (1) between observed and predi The calibration slope used f and predicted values, with va for the bias correction were v 0.970, respectively (Figure 4) <b>3.6 The validation of discri</b>	d out in the ) The Brier s cted values, v or modal cal alues closer to validated by t ). The predic <b>mination</b> ining and va	score for with valu ibration, o 1 indic he mode tion accu lidation	c overall performance, ues closer to 0 indicatin which assessed the ag ating better performance of with a Brier score of 0 uracy of the model was groups, and the AUC of	which asses ng better pro- reement bet e. The accur 0.176 and a relatively h	ssed the difference edictive ability. (2 ween the observed racy measurement calibration slope o igh.
273 274 275 276 277 278 279 280 281 282 283 283 284 285	Model validation was carrie assessed by two aspects. (1) between observed and predi The calibration slope used f and predicted values, with va for the bias correction were v 0.970, respectively (Figure 4) <b>3.6 The validation of discri</b> ROC was plotted for the tra	d out in the ) The Brier s cted values, for modal cal alues closer to validated by t ). The predic <b>mination</b> ining and va 4, respectivel	score for with valu ibration, o 1 indic he mode tion accu lidation y (Figur	r overall performance, ues closer to 0 indicatin which assessed the ag ating better performance of with a Brier score of 0 uracy of the model was groups, and the AUC of train	which asses ng better pro- reement bet e. The accur 0.176 and a relatively h	ssed the difference edictive ability. (2 ween the observed racy measurement calibration slope o igh.
273 274 275 276 277 278 279 280 281 282 283 284 283 284 285 286	Model validation was carried assessed by two aspects. (1) between observed and predic The calibration slope used f and predicted values, with va- for the bias correction were va- 0.970, respectively (Figure 4) <b>3.6 The validation of discri</b> ROC was plotted for the trans- groups were 0.785 and 0.784	d out in the ) The Brier s cted values, for modal cal alues closer to validated by t ). The predic <b>mination</b> ining and va 4, respectivel	score for with valu ibration, o 1 indic he mode tion accu lidation y (Figur	r overall performance, ues closer to 0 indicatin which assessed the ag ating better performance of with a Brier score of 0 uracy of the model was groups, and the AUC of train	which asses ng better pro- reement bet e. The accur 0.176 and a relatively h	ssed the difference edictive ability. (2 ween the observe racy measurement calibration slope c igh.

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

295 4. Discussion

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

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

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

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

factors were also one of the influential factors of mental health problems in our study. Relevant studies have found that exposure to air pollution is associated with increased suicide risk and depressive symptoms <sup>[34]</sup>. Hypertension and diabetes were the influential factors in this study. A study has shown that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher than in the non-diabetic population <sup>[35]</sup>. Furthermore, there is a recognized association between hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear <sup>[36]</sup>.

Factory workers and miners were inevitably exposed to occupational hazards such as benzene and asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous enterprises, involving more than 30 different types of operations, of which factory workers and miners is the one <sup>[37]</sup>. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for mental health problems in this study. Our study found that the work schedules of factory workers and miners were vary and the phenomenon of night shifts was very common, which inevitably affected their normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could suffer from negative emotions such as anxiety and depression [38]. Professional title and education level were also important influences on mental health issues. In the workplace, generally speaking, the higher the professional title and education level, the higher the status of the worker in the company and the greater the role played in the position. The number of studies on socio-economic status and mental health had increased in recent years. Some of these studies have shown that major depression is higher in the low socio-economic status group <sup>[39]</sup>. It has also been suggested that education itself is the best indicator of socio-economic status <sup>[40]</sup>. Marital status was one of the influential factors for mental health problems. Many studies have found an association between mental health and gender, marital status, lifestyle and working conditions, and it has been shown that poor mental health in women is associated with divorce or widowhood [41]. In this study, working more than seven hours a day was a determinant factor on mental health problems, which was consistent with other studies that had shown that long working hours could have a negative impact on employees' mental health and that excessive workloads could increase workers' fatigue, which in turn could lead to anxiety and depression [42].

In China, there are many problems in identifying people with mental health problems due to uneven and imperfect levels of medical development across regions. Some studies have shown that in mainland China, general practitioners, surgeons and primary health care workers often have little or no mental health training, which prevents them from providing basic mental health services [43]. Non-mental health professionals in general hospitals learn about mental illness on their own, rather than learning about it during their formal education<sup>44</sup>. 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

Page 15 of 38

#### BMJ Open

1		
2 3	369	mental health problems without the help of medical staff. This study has several strengths. First, to our
4		
5	370	knowledge, this is the first model to develop and assess the likelihood of mental health problems in a
6 7	371	group of factory workers and miners. Secondly, the nomogram in this study includes demographic
8	372	information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors
9	373	that are closely related to factory workers and miners, allowing for a more accurate assessment of the
10 11	374	risk of morbidity among them, as well as providing a methodological reference for other related studies.
12	375	
13	376	Patient and public involvement
14 15	377	Neither patients nor members of the public had any involvement in the design of this study.
15	378	
17	379	Acknowledgements The authors are grateful to all participants and investigators.
18 19	380	Acknowledgements the additions are grateriar to an participants and investigators.
20		
21	381	Contributions Y.L., Q.L., and T.L. are responsible for conceptualization; Y.L. is responsible for
22	382	methodology, software, formal analysis, resources, and visualization; Q.L. and T.L. are responsible for
23 24	383	the original draft preparation; Q.L. and H.Y. are responsible for reviewing; Q.L. is responsible for editing;
25	384	T.L. is responsible for supervision. Yaoqin Lu and Qi Liu contributed equally to this work.
26	385	
27 28	386	Funding This study was supported by National Natural Science Foundation of China, grant number
29	387	81760581 and Public Health and Preventive Medicine, the 13th Five-Year Plan Key Subject of Xinjiang
30	388	Uygur Autonomous Region.
31 32	389	
33	390	Uygur Autonomous Region. Competing interests None declared. Patient consent for publication Not required. Ethics approval The study was approved by the ethics committee of Urumqi Center for Disease Control
34	391	competing increases (volie declared.
35 36	392	Patient consent for publication Not required.
37	393	
38	394	Ethics approval The study was approved by the ethics committee of Urumqi Center for Disease Control
39	395	and Prevention
40 41	396	
42	397	Data availability statement The data used to support the findings of this study are available from the
43	398	corresponding author upon request.
44 45	399	
46	400	References
47		
48 49	401	[1] WHO Terminology Information System [online glossary] http://www.who.int/health-systems-
50	402	performance/docs/glossary.html.
51	403	[2] Wang Y, Liu X, Qiu J, Wang H, Liu D, Zhao Z, Song M, Song Q, Wang X, Zhou Y, Wang W.
52 53	404	Association between ideal cardiovascular health metrics and suboptimal health status in Chinese
54	405	population. <i>Sci Rep</i> 2017;7:14975.
55	406	[3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental
56 57	407	health. <i>Lancet</i> . 2007;370:859–77.
57 58	408	[4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive
59	409	symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. Health
60		

1 2		
2 3	440	
4	410	<i>Place</i> 2016;42:37–46.
5	411	[5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,
6	412	Degenhardt L, Dicker D, et al. Global, regional, and national incidence, prevalence, and years lived
7 8	413	with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A
9	414	systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743-800.
10	415	[6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health?. Int J
11	416	Environ Res Public Health 2020;17:7530.
12	417	[7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative
13 14	418	income, gender, and location. China Econ Rev 2018;48:83–101.
15	419	[8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:
16	420	Results from the PATH Through Life Survey. <i>Soc Sci Med.</i> 2009;69:229–237.
17	421	[9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and
18		
19 20	422	Subjective Well-being? J Environ Econ Manag 2017;85:81–94.
21	423	[10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. <i>World</i>
22	424	<i>Psychiatry</i> . 2015;14:36–42.
23	425	[11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson
24 25	426	SJ. IPS Supported employment: a review. J Vocat Rehabil 2014;40:1–13.
25 26	427	[12] Jia Z, et al. Exposure to ambient air particles increases the risk of mental disorder: findings from a
27	428	natural experiment in Beijing. Int J Environ Res Public Health 2018;15:160.
28	429	[13] Szyszkowicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency
29	430	department visits for suicide attempts in vancouver, Canada. Environ Health Insights 2010;4:79-86.
30 31	431	[14] Michael J. S. International classification of sleep disorders-third edition : highlights and
32	432	modifications. <i>Chest</i> 2014;146:1387–1394.
33	433	[15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial
34	434	infarction: gender differences. Open Nurs J 2018;12:205–214.
35 36	435	[16] Wurm W, Vogel K, Holl A, <i>et al.</i> Depression-burnout overlap in physicians. <i>PLoS One</i>
37	436	2016;11:e0149913.
38	430	[17] Porru F, Robroek S J W, Bültmann U, <i>et al.</i> Mental health among university students: The
39		
40	438	associations of effort-reward imbalance and overcommitment with psychological distress. J Affect
41 42	439	Disord 2021;282:953-961.
43	440	[18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,
44	441	and burnout among physician assistant students: a multi-institutional study. Health Professions Edu
45	442	2020;6:420–427.
46 47	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.
47	444	Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during
49	445	the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. EClinical Med
50	446	2020;24:100424.
51 52	447	[20] Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP. The estimation of the number of
52 53	448	underground coal miners and the annual dose to coal miners in China. Health Phys 2007;93:127-
55 54	449	132.
55	450	[21] Yong X, Gao X, Zhang Z, <i>et al.</i> Associations of occupational stress with job burn-out, depression
56	451	and hypertension in coal miners of xinjiang, china: A cross-sectional study. <i>BMJ open</i> 2020;10:
57 59	451	e036087.
58 59		
60	453	[22] P. Bech, J. Bille, S. B. Møller, L. C. Hellström, S. D. Østergaard. Psychometric validation of the
		15

Page 17 of 38

1

#### BMJ Open

2		
3	454	Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal
4 5	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 9	458	[24] Derogatis L, Lipman R.S, Covi L. SCL-90: an outpatient psychiatric rating scale-preliminary
10	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 14	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 18	465	[27] Siegrist J, Wege N, Pühlhofer F, Wahrendorf M. A short generic measure of work stress in the era
19	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 22	468	burnout of medical staff in xinjiang, China: A cross-sectional study. Med Sci Monit 2020;26:
22 23	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 27	472	comparison between gradient boosting and the lasso. <i>Methods Inf Med</i> 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. <i>World Psychiatry</i>
30 21	475	2020;19:3–14.
31 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. <i>Annu Rev Public Health</i> 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 36	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. <i>Diabet Med</i> 2006;23:445–448.
39 40	482	[36] Gilsanz P, Karter AJ, Beeri 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 44	485	[37] Lu Y, Zhang Z, Yan H, et al. Effects of occupational hazards on job stress and mental health of
45	486	factory workers and miners: A propensity score analysis. <i>BioMed Res Int</i> 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. <i>BioMed Res Int</i> 2020;2020:4795763.
48 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 <i>et al.</i> Socioeconomic status and health: how education, income,
52 53	492	and occupation contribute to risk-factors for cardiovascular disease. $Am J Public$
53 54	493	Health 1992;82:816–820.
55	494	[41] Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis P.N, Mavreas V. Prevalence
56 57	495	and sociodemographic associations of common mental disorders in a nationally representative
57 58	496	sample of the general population of Greece. <i>BMC Psychiatry</i> 2013;13:163.
59	497	[42] Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG <i>et al.</i> Long
60		
		16

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

- 500 [43] Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, *et al.* Prevalence, treatment, and associated
  501 disability of mental disorders in four provinces in China during 2001–05: an epidemiological
  502 survey. *Lancet* 2009;373:2041–2053.
  - 503 [44] Wu Q, Luo X, Chen S, *et al.* Mental health literacy survey of non-mental health professionals in six
    504 general hospitals in hunan province of China. *PloS one* 2017;12:e0180327.

#### 506 Figure legends

Fig.1. 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 weas plotted based on the lambda series in Figure 1(A), and 12 features with non-zero coefficients were selected by optimal lambda.

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

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

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

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

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

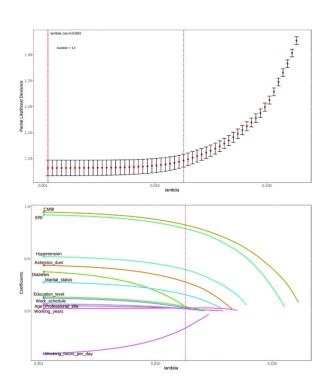


Fig.1. Feature selection using the LASSO binary logistic regression model.

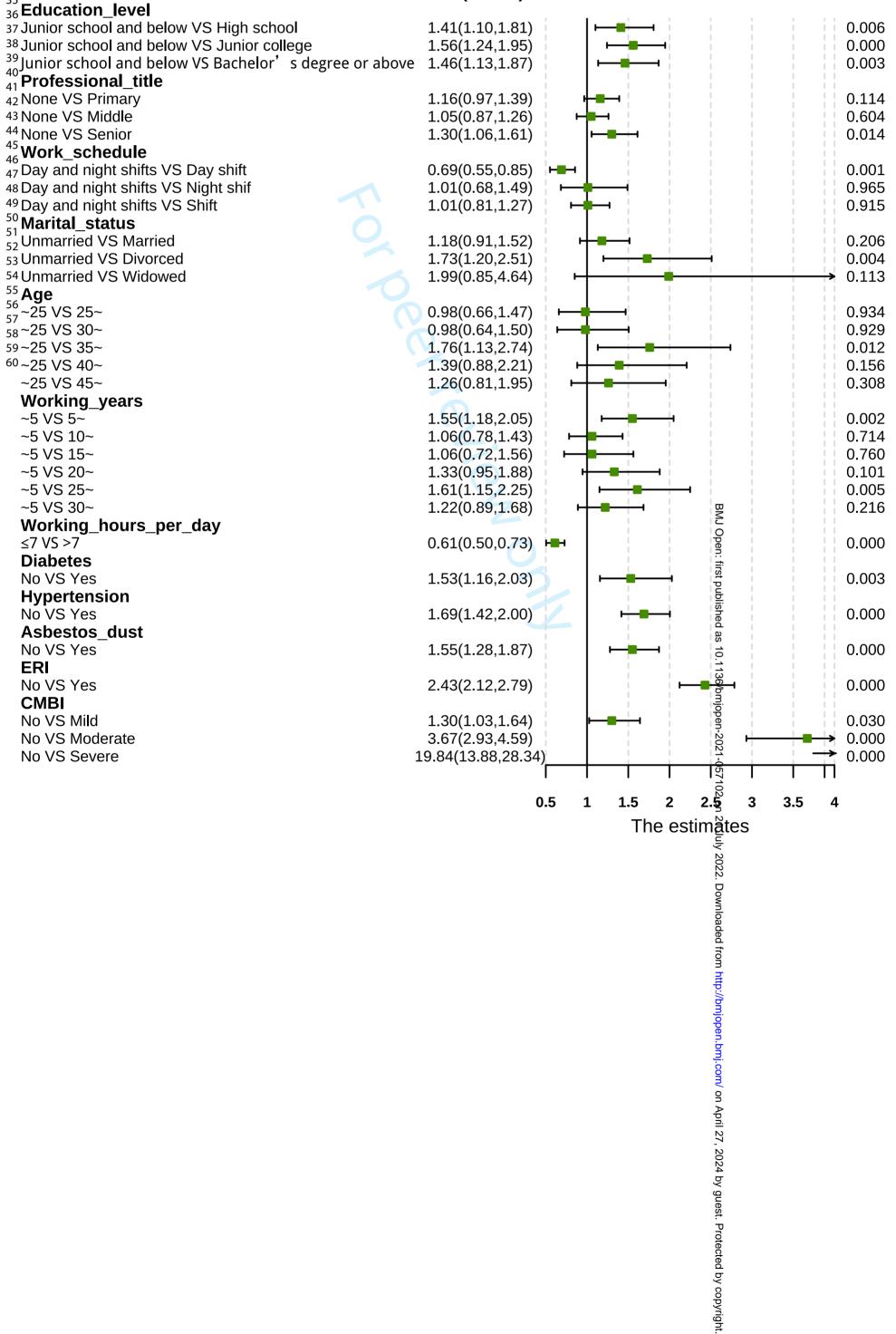
210x297mm (300 x 300 DPI)

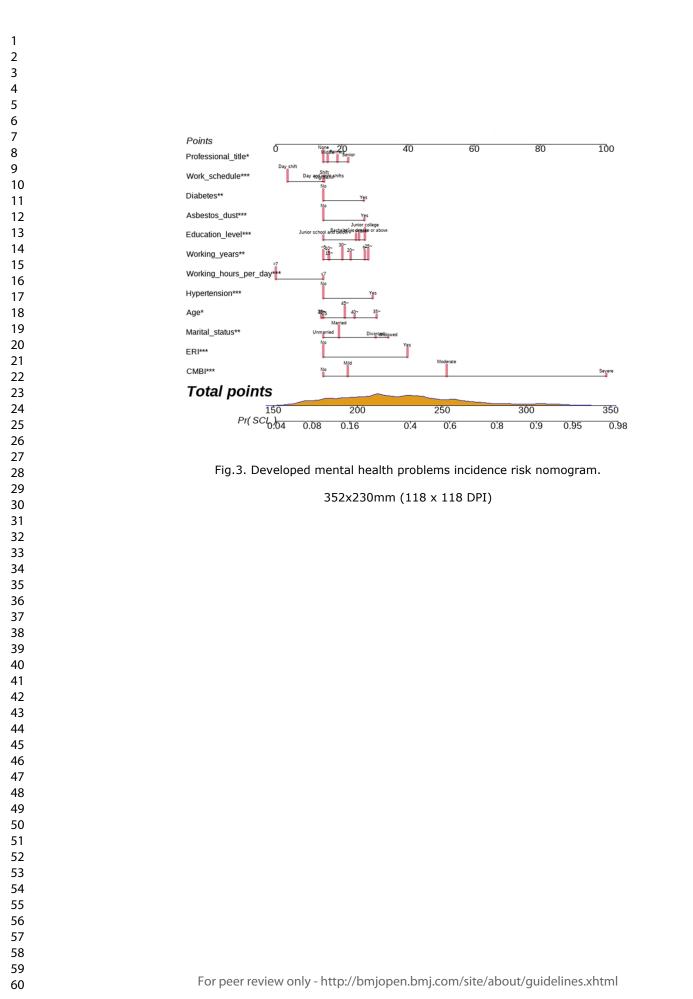
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	

# <sup>34</sup><sub>35</sub>Variable

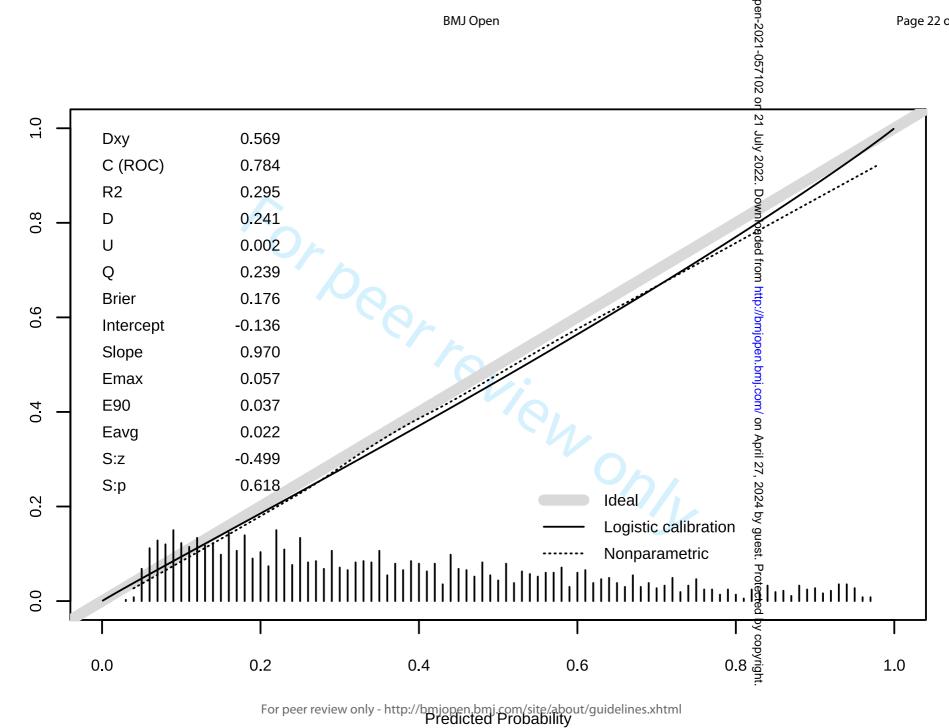
### OR(CI95%)

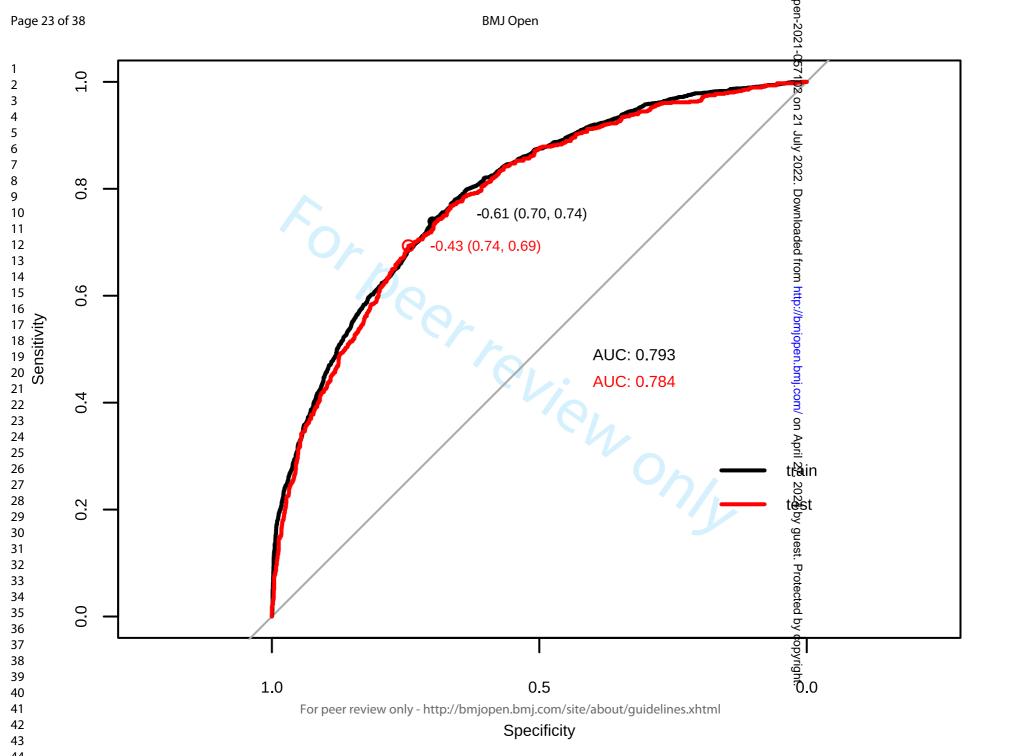
## P-Value

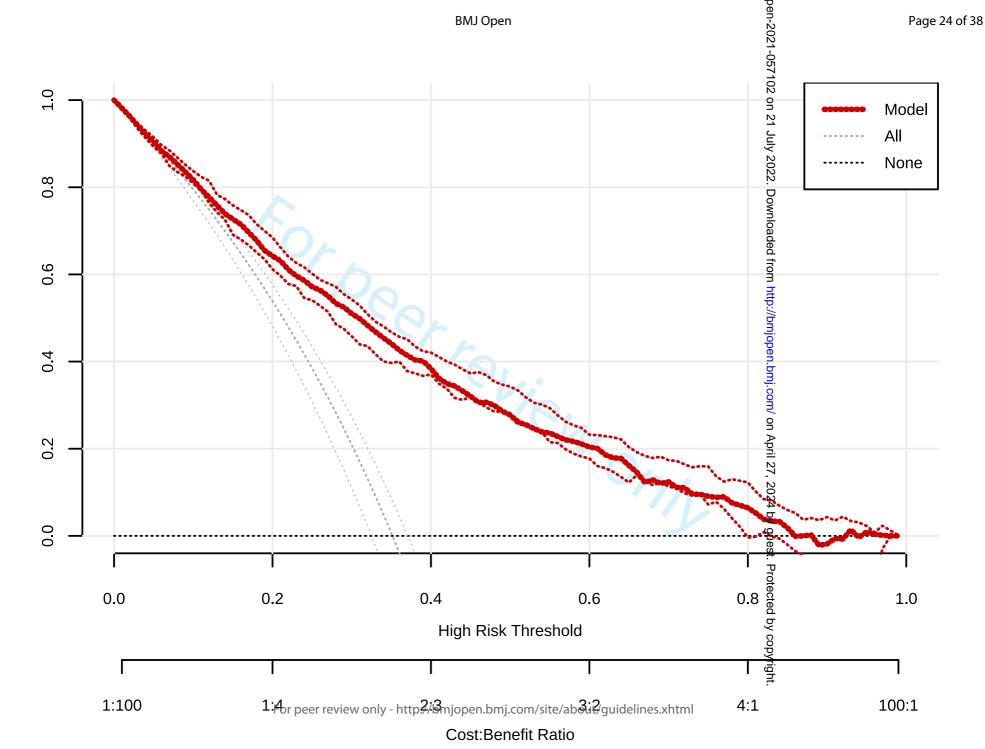




A A







A A

	tem	6/bmjopen-2021-057,02 o	Relevant text from manuscri
Title and abstract	No.         Recommendation           1         (a) Indicate the study's design with a commonly used term in the title or the abstract	<sup>1</sup> Page <sup>D</sup> No.	Development and Validation of
The and abstract	( <i>a</i> ) indicate the study's design with a commonly used term in the title of the abstract	L July	Nomogram for Predicting the
		21 July 2022.	Risk of Mental Health Problem
		2. D	of Factory Workers and Miner
	(b) Provide in the abstract an informative and balanced summary of what was done and what was		A total of 7,118 participants m
	(b) Provide in the abstract an informative and balanced summary of what was done and what was found	loade	the inclusion criteria and the da
		ed fro	were randomly divided into a
	(b) Provide in the abstract an informative and balanced summary of what was done and what was found	nloaded from http://bmjopen.bmj.com/ on April 27, 2024 by gues	training group (n=4,955) and a
		ttp://	validation group $(n=2,163)$ in
		bmjc	ratio of 3:1. A total of 23
		open	characteristics were included i
		.bmj	this study and LASSO regression
		com	selected 12 characteristics such
		on	education, professional title, a
		April	CMBI, ERI, asbestos dust,
		27,	hypertension, diabetes, workin
		202	hours per day, working years,
		4 by	marital status, and work sched
		gues	as predictors for the constructi
		÷	of the nomogram. In the
		rotec	validation group the Brier scor
		Protected by copyright.	was 0.176, the calibration slop
		ру сс	was 0.970 and the calibration
		ру	curve of nomogram showed a

	6/bmjopen-2021-057102 on 21 July 2022. Downloaded from http://bmjoper	good fit, indicating good
	-057	agreement between predictions
	102	and observations. The AUC of
	on 2	training group and verification
	1 Jul	group were 0.785 and 0.784
	y 20	respectively, which showed goo
	22. [	discrimination. The DCA
	Dow	suggested that the nomogram for
	nloac	predicting the risk of mental
	ded f	health problems among factory
	rom	workers and miners was more
	http:	practical when the risk threshol
	)//bm	for mental health problems was
	jopei	10% for intervention.
troduction	.bmj	
ackground/rationale 2 Explain the scientific background and rationale for the investigation being reported	2 6	Factory workers and miners are
	√ on	special group of workers with a
	Apri	relatively low overall level of
	.com/ on April 27, 2024 by gues	education and are highly prone
	202	suffering from mental health
	14 by	problems due to limited social
	gue	support, excessive workload an
	.+	irregular lifestyles, as well as
	rote	occupational hazards such as
	Protected by copyright.	noise and coal dust that they
	by c	inevitably need to face in their
	оруг	working environment <sup>18-19</sup> . China

Objectives

	21-(	has the world's largest group of
	)571	factory workers and miners, about
	02 0	6 million <sup>20</sup> , who are regularly
	in 21	involved in occupational hazards.
	July	Mental health problems which
	y 202	need to require a long process are
	22. E	known to be a syndrome caused
	Dowr	by chronic stress. Factory workers
	nloac	and miners, represented by those
	ded f	engaged in coal mining, have a
	rom	mental burden rating of 8.3, one
	http:	of the highest mental burdens
	//bm	among 150 occupations <sup>21</sup> . This
	jope	explains the high level of mental
	n.bn	health problems among mine
		workers in previous studies,
	m/ o	making the identification and
	n Ap	treatment of mental health
	oril 2	problems even more important.
	7, 20	Therefore, it is essential to
	024 t	provide a viable and easy-to-
	ру дг	apply tool for identifying workers
	lest.	at risk of mental health problems
	021-057102 on 21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 27, 2024 by guest. Protected by copyright	and thus for timely interventions.
tate specific objectives, including any prespecified hypotheses	3 ecte	Therefore, the aim of our study is
	d by	to develop and validate an easy-
	сор	to-use nomogram that combines
	Ŋrię	objective information on the

6/bmjopen-20

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		BMJ Open	/bmjopen-:	Pag
			6/bmjopen-2021-057102 on 21 July 2022.	demographics, job burnout, occupational stress and occupational hazards to comprehensively and accurately predict the prevalence of mental health problems among factory workers and miners.
Methods				
Study design	4	Present key elements of study design early in the paper	wnloaded from http://	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	Ddwnloaded from http://bmjopen.bmj.com/ on April 27, 2024 by guest. Protected by cdpyright.	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 3,61 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,

of 38	BMJ Open	⁄bmjopen	
		-2021-0571	Middong District and Urumqi County.
Participants	<ul> <li>6 (a) Cohort study—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</li> <li>(b) Cohort study—For matched studies, give matching criteria and number of exposed and the sources and methods of exposed and the sources and methods.</li> </ul>	6/bmjopen-2021-057102 on 21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April-27, 2024 by gues	The exclusion criteria were the following: (I) factory workers an 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 dat were also excluded from the analysis based on discussion and agreement among the subject members.
	unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	om/ on A	
Variables	<ul> <li>Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable</li> </ul>	÷	<ul> <li>2.2.1. Assessment of mental health</li> <li>2.2.2. Assessment of occupation stress</li> <li>2.2.3. Assessment of job burnou</li> <li>2.2.4. Candidate predictors</li> </ul>
Data sources/ measurement	<ul> <li>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</li> </ul>	Protected by copyrigh	Categorical variables were described as counts and percentages, and chi square test Fisher exact test was used to compare categorical variables

3 4

ВМЈ Орен	n mjopen-2
Bias 9 Describe any efforts to address potential sources of bias	n       between groups. 70% of participants were randomly assigned to the training cohort 30% to the validation cohort. Variables were screened using 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, fi which predictive models were constructed. A nomogram for predicting was generated according to the selected characteristics.         6       We used calibration plots and receiver operating characterist (ROC) curves to show the calibration and discrimination our final model. Brier scores fi overall performance, calibration slopes were used to assess the predictable accuracy of the mo Decision curve analysis (DCA was applied to calculate the ne benefit of the nomogram. Statistical analysis was perfor using the open-source R softw Version 3.6.1 (http://www.r-

Page	31	of	38
------	----	----	----

 6/bmjope

	10 Explain how the study size was arrived at	project.org). The significance level (α) set at 0.05.
Study size	10   Explain how the study size was arrived at   N	
Continued on next page		
Quantitative	11 Explain how quantitative variables were handled in the analyses. If applicable, describe which $5 = \frac{1}{2}$	Sex was defined as male or
variables	groupings were chosen and why	female; ethnicity was defined as
		Han and other; education level was
		defined as junior high school and
		below, high school, junior college
	oade oade	or bachelor's degree or above; labo
		contracts was defined as signed or
	PO S	unsigned; professional title was
	ter and ter an	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
	groupings were chosen and why	<3000, 3000~, 4000~, 5000~,
		6000~, 7000~ or 8000~; age (years
	17, 2	was defined as <25, 25~, 30~, 35~
	0024	$40\sim$ or $45\sim$ ; working years was
	в в в в в в в в в в в в в в в в в в в	defined as ~5, 5~, 10~, 15~, 20~,
	gue de la companya d La companya de la comp	25~ or 30~; working hours per day
	11 Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why 5 UV 202 Downloaded from http://brigoen.brig.com/ on April 27, 2024 by guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Guest Protected by Com/ on April 27, 2024 by Gu	(hours) was defined as $\leq 7$ or $> 7$ ;
	Prot	working days per week (days) was
		defined as $\leq 5$ or $>5$ ; exposure to
	ŭ 	coal dust, silica dust, asbestos dust
	× 8	benzene, lead, noise, brucellosis
	рут	were all defined as yes or no; ERI
	g ht.	
	·	

3 4

	BMJ Open	5/bmjopen-20	Page
		021-057102 on 21 J	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         Image: Control for confounding those used to control for confounding the conformation of the conformatio of the c	2 (a) Describe all statistical methods, including those used to control for confounding	4 by guest. Protected	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	by c	
		by copyright.	

	() Factoria tara attace attace d	6/bmjopen-2021-057102 on 21 July 2022.	
		<u></u>	A total of 7,500 questionnaires
	(c) Explain how missing data were addressed	4 1-05	
		710	were distributed and 7,315
		2 on	questionnaires were returned,
		21	representing a return rate of 97.5%
		July	After checking the validity and
		202	integrity of the questionnaires,
			7,118 questionnaires were
		own	confirmed as valid, with an
		load	effective rate of 97.3%. All
		led f	participants understood the purpos
		rom	of the study and voluntarily
		Downloaded from http:/	participated in the study.
	(d) Cohort study—If applicable, explain how loss to follow-up was addressed	/bm	
	Case-control study—If applicable, explain how matching of cases and controls was addressed	ope	
	<i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	bmjopen.bmj.c	
	( <u>e</u> ) Describe any sensitivity analyses	öm/	
Results		on A	
Participants 13	* (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined	4 II	7500 participants volunteered for
	for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	27, 2	the survey
		27, 2024 by gues	Issued a total of 7500
		by	questionnaires
		gue	Collected a total of 7315
		. <del></del>	questionnaires
		rote	7118 valid and integrated questionnaires
		Protected	questionnanes
	(b) Give reasons for non-participation at each stage	4 4 4	
	(c) Consider use of a flow diagram	4 <u>y</u>	
		ight.	
	9 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	-	

Page 34 of 38

	BMJ Open       6         *       (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders       6	Pag
Descriptive data 1	* (a) Give characteristics of study participants (eg demographic, clinical, social) and information on 6	A total of 7,118 participants met the
Descriptive data 1	exposures and potential confounders	inclusion criteria and the data were
	102	randomly divided into a training
		(n=1.955) and a validation
		group ( $n=2,163$ ). Over half of all
		participants (65.31%) were male,
	22.1	57.31% of the population was over
	Dow	35 years of age and 78.32% of the
	21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 27, 2024 by gues	subjects were married, showing the
		factory workers and miners are
	No in internet in the second sec	generally older and most of them
		have spouses. The majority of the
		had completed high school
		(83.94%), while a smaller
		percentage had completed
		undergraduate education (22.98%
		indicating that the group of factor
	n Ap	workers and miners as a whole wa
		not well educated. The total numb
	,7,20	of workers (n, %) exposed to coal
		dust, silica dust, asbestos dust,
	y y gu	benzene, lead, noise and brucellos
		in the factory and mining
	Protected by copyright	enterprises were 377 (5), 730 (10)
	ecte	981 (14), 1,981 (28), 373 (5), 4,94
	d by	(69) and 121 (2) respectively, with
		the total number of workers

je 35 of 38	BMJ Open	
	open-2	
		exposed to noise amounting to
	0571	4,942, or 69% of the total
		population surveyed. The
	on 2,	demographic, job burnout,
		occupational stress and
	20: 20:	occupational exposure factors for
		the training and validation groups
	Dowr	are shown in Table 1. The results
		showed that there were no
		significant statistical differences
	P P P P P P P P P P P P P P P P P P P	between the two groups of
		characteristic variables, except for
		coal dust and CMBI, indicating that
		the baseline levels were largely
	2022. Downloaded from http://bmjopen.bm	consistent between the two groups.
	(b) Indicate number of participants with missing data for each variable of interest	
	(c) <i>Cohort study</i> —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 $\square$	
	15*       Cohort study—Report numbers of outcome events or summary measures over time       Image: 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       6         Cross-sectional study—Report numbers of outcome events or summary measures       6         Update       0         Update	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
	די זי	participants (65.31%) were male,
	otteecte	57.31% of the population was over
	Protected by copyright.	35 years of age and 78.32% of the
		subjects were married, showing that
		factory workers and miners are

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 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.

6/bmjopen-2021-057102 on 21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 27, 2024 by guest. Protected by copyright

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

.v on April 27. :

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 neluded	2021-057102 on 21 July	Categorical variables were described as counts and percentages, and chi square test or Fisher exact test was used to
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 eriod	6/bmjopen-2021-057102 on 21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 27, 2024 by guest. Protect	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 logistr regression models were used to estimate risk ratios and corresponding 95% confidence intervals (CI) of risk factors, from which predictive models were constructed.
	by copyright.	
		5. Protected by copyright. For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

		BMJ Open	6/bmjopen-2021	Pa
Other analyses	17	Report other analyses done-eg analyses of subgroups and interactions, and sensitivity analyses	021-0	
Discussion			9571	
Discussion Key results Limitations	18	Summarise key results with reference to study objectives Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	-057102 on 21 July 2022. Downloaded from http://bmjopen.bmj.com/ on April 27, 2024 by guest. Protected by copyright	Therefore, this study designed a simple and comprehensive nomogram to address the issue of timely detection and effective interventions for people with ment health problems, so that people at risk of mental health problems 
			/ copyri	Nomogram need to be externally assessed in a wider population.

Page 39 of 38		BMJ Open	6/bmjopen-	
1 2 Interpretation 3 4 5 6 7 8 9 10 11 12 13 14 15	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	6/bmJopen-2021-05/102 on 21 July 2022. Downloaded from http://	
16 17 18 19 20 Generalisability	21	Discuss the generalisability (external validity) of the study results	- B	
21 Other informatio	on		open	
23 24 25 26 27 28 29 30	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	141.com/ on April 27, 2024	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.
31 32 33 *Give information s 34	sepa	rately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups	in cohoest.	
<ul> <li>35 Note: An Explanati</li> <li>36 checklist is best use</li> <li>37 http://www.annals.c</li> <li>39</li> <li>40</li> <li>41</li> </ul>	ed in	nd Elaboration article discusses each checklist item and gives methodological background and published conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmed and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at wy	icine.org	/, Annals of Internal Medicine at e-statement.org.
42 43 44 45 46		<b>15</b> For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xh	-	

# **BMJ Open**

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

Journal:	BMJ Open
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





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 <u>licence</u>.

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 <u>Creative Commons</u> 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.

R. O.

	Development and Validation of a Nomogram for Predicting the Risk of Mental Health Problems
:	of Factory Workers and Miners
:	Yaoqin Lu, <sup>1,2</sup> Qi Liu, <sup>1</sup> Huan Yan <sup>3,4</sup> , and Tao Liu <sup>1</sup>
	<sup>1</sup> Department of Toxicology, School of Public Health, Xinjiang Medical University, Urumqi 830011
	Xinjiang, China.
	<sup>2</sup> Urumqi Center for Disease Control and Prevention, Urumqi 830026, Xinjiang, China.
	<sup>3</sup> Department of Nutrition and Food Hygiene, College of Public Health, Xinjiang Medical University
	Urumqi, Xinjiang, China. <sup>4</sup> Xinjiang Engineering Technology Research Center for Green Processing of Nature Product Center
1	Xinjiang Autonomous Academy of Instrumental Analysis, Urumqi, Xinjiang, China.
1	This is a second s
1:	Correspondence should be addressed to Tao Liu; xjmult@163.com
1	Yaoqin Lu and Qi Liu contributed equally to this work
1	Abstract
1	Objective A nomogram for predicting the risk of mental health problems was established in a population
1	of factory workers and miners, in order to quickly calculate the probability of a worker suffering from
1	mental health problems.
1	Methods A cross-sectional survey of 7,500 factory workers and miners in Urumqi was conducted by
1	means of an electronic questionnaire using cluster sampling method. Participants were randomly
2	assigned to the training group (70%) and the validation group (30%). Questionnaire-based survey was
2	conducted to collect information. A least absolute shrinkage and selection operator (LASSO) regression
2	model was used to screen the predictors related to the risk of mental health problems of the training
2	group. Multivariate logistic regression analysis was applied to construct the prediction model. Calibration
2	plots and receiver operating characteristic-derived area under the curve (AUC) were used for mode
2	validation. Decision curve analysis (DCA) was applied to calculate the net benefit of the screening model
2	Results A total of 7,118 participants met the inclusion criteria and the data were randomly divided into
2	a training group (n=4,955) and a validation group (n=2,163) in a ratio of 3:1. A total of 23 characteristics
2	were included in this study and LASSO regression selected 12 characteristics such as education
2	professional title, age, CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working
3	years, marital status, and work schedule as predictors for the construction of the nomogram. In the
3	validation group the Brier score was 0.176, the calibration slope was 0.970 and the calibration curve or
3	nomogram showed a good fit. The AUC of training group and verification group were 0.785 and 0.784
3	respectively.
34	<b>Conclusion</b> The nomogram combining these 12 characteristics can be used to predict the risk of suffering
3	mental health problems, providing a useful tool for quickly and accurately screening the risk of menta
3	health problems.
3	

#### **BMJ** Open

2		
3	38	Key wo
4 5	39	
6	40	Strengt
7 8	41	1. This
o 9	42	workers
10	43	2. The
11 12	44	modera
13	45	3. This
14 15	46	occupat
16	47	4. We
17	48	influenc
18 19	49	5. There
20	50	
21 22	51	1. Intro
23	52	
24 25	53	The Wo
25 26	54	well-be
27	55	of phys
28 29	56	
30	57	operatio
31 32		increasi
32 33	58	problem
34	59	non-cor
35 36	60	have a 1
37	61	of the g
38	62	currentl
39 40	63	global t
41	64	the glob
42 43	65	
44	66	As rese
45	67	income
46 47	68	problem
48	69	and few
49 50	70	an incre
51	71	and env
52 53	72	tempera
53 54	73	particul
55	74	to ambi
56 57	75	attempt
58	76	artery d
59 60	77	are clos

ords Mental health; Predictor; Nomogram; Risk; Factory workers and miners ths and limitations of this study

is the first study to develop an easy-to-use nomogram to predict the mental health risks of factory s and miners.

AUC of training group and verification group were 0.785 and 0.784 respectively, showing te discriminatory and calibration power.

nomogram model's variables are more comprehensive, including demographics, burnout, tional stress and occupational hazards.

had considered many influential factors, but we were still not certain whether all possible ces were covered.

e is a lack of external validation in other populations in other regions and countries.

#### oduction

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

archers around the world have delved into the field of mental health, factors such as gender, levels, environment and education have been found to be associated with people's mental health ns <sup>[7-10]</sup>. Moreover, employment is also strongly associated with quality of life, higher self-esteem rer psychiatric symptoms <sup>[11]</sup>. In addition, in the context of the global challenges of climate change, easing number of scholars have been examining the epidemiological links between mental health ironmental factors. Some studies have suggested that mental health may be influenced by ambient ature, and an association has been found between environmental pollutants, particularly fine ate matter, and mental health problems <sup>[12]</sup>. A relevant study shows that with short-term exposure ent air pollution is associated with increased emergency room visits due to depression or suicide ts <sup>[13]</sup>. Furthermore, other factors associated with mental health include sleep, diabetes, coronary isease and cardiovascular disease [14-15]. It is worth noting that job burnout and occupational stress ely linked to mental health. Job burnout is an exhaustion state of physical and psychological that

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

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 [18-19]. Through a review of the literature, our group found that coal dust, crystalline silica and noise pollution were common causes of health problems for workers in underground mines <sup>[20]</sup>. And, exposure to coal mine dust is a significant cause of pneumoconiosis in coal miners<sup>[21]</sup>. In addition, asbestos is one of the major occupational hazards in the daily work of workers in the construction and automotive industries <sup>[22]</sup>. China has the world's largest group of factory workers and miners, about 6 million <sup>[23]</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>[24]</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.

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

Page 5 of 29

#### **BMJ** Open

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

#### 124 2. Materials and Methods

#### 126 2.1 Calculation of sample size

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, 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 for  $\alpha = 0.05$ , then the formula is calculated as,  $n = 400 \times \frac{q}{p}$ . A cross-sectional study in Xinjiang showed that 38.27% of factory workers and miners had mental health problems <sup>[29]</sup> And a study revealed that 633 out of 1675 coal miners (37.8%) suffered from mental disorders between August 2018 and June  $2019^{[30]}$ . In this study, we assumed a 30% prevalence of mental health problem to obtain the maximum required sample size, which would calculate a sample size of 934, taking into account non-response and a 20% loss of questionnaires, which would require approximately 1168 people.

#### 137 2.2. Participants

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

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

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

155 An online electronic questionnaire was created using the Questionnaire Star platform to collect data. The

survey was conducted by trained surveyors who explained the purpose, meaning, content and requirements of the questionnaire to all participants and provided on-site instructions to ensure the return rate of the questionnaire. All participants understood the purpose of the study and were willing to participate in the study. 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%. 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) (Figure 1).

#### 165 2.3. Research Methods

#### 167 2.3.1. Assessment of mental health

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

#### 178 2.3.2. Assessment of occupational stress

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

#### 190 2.3.3. Assessment of job burnout

192 In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess 193 job burnout, which has good reliability and validity <sup>[36]</sup>. CMBI is composed of 15 items in three 194 dimensions: emotional exhaustion (5 items), depersonalization (5 items) and reduced personal 195 accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete Page 7 of 29

#### **BMJ** Open

compliance and 7 points indicating complete non-compliance. According to the critical value (emotional exhaustion  $\geq 25$ , dependent dependence of the experiment reduction  $\geq 16$ , the levels of occupational burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to or above the critical value), moderate (any two aspects are equal to or higher than the critical values), and severe (three aspects are equal to or higher than the critical values) [<sup>37]</sup>. In this survey, Cronbach  $\alpha$ was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.

#### 2.3.4. Candidate predictors

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

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 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.

#### 2.4. Statistical analysis

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. Statistically significant predictors were applied to develop a prediction model for the risk of mental health problems among factory workers and miners by introducing all selected factors and analyzing the statistical significance levels of them. 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 (http://www.r-project.org). The significance level ( $\alpha$ ) set at 0.05.

**3. Results** 

#### 247 3.1. Participant characteristics

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 (5.3), 730 (10.3), 981 (14), 1,981 (27.8), 373 (5.2), 4,942 (69.4) and 121 (1,7) 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.

5 m - 5	
Table 1 Characteristics of the study participants	

Variables	Total (n = 7118)	train (n = 4955)	test (n = 2163)	р	
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)	0.06	
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)		

Page 9 of 29

1

# BMJ Open

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

2					
3	Labor contracts, n (%)				
4 5	Signed	6641 (93.3)	4624 (93.3)	2017 (93.3)	0.955
6	Unsigned	477 (6.7)	331 (6.7)	146 (6.7)	
7 8	Working hours per day (hours), n (%)				
9	≤7	1161 (16.3)	814 (16.4)	347 (16.0)	0.712
10	>7	5957 (83.7)	4141 (83.6)	1816 (84.0)	
11 12	Working days per week (days), n (%)				
13	≤5	4442 (62.4)	3107 (62.7)	1335 (61.7)	0.446
14 15	>5	2676 (37.6)	1848 (37.3)	828 (38.3)	
16	Diabetes, n (%)				
17 18	Yes	429 (6.0)	298 (6.0)	131 (6.1)	0.988
19	No	6689 (94.0)	4657 (94.0)	2032 (93.9)	
20	Hypertension, n (%)				
21 22	Yes	1330 (18.7)	929 (18.7)	401 (18.5)	0.861
23	No	5788 (81.3)	4026 (81.3)	1762 (81.5)	
24 25	Coal dust, n (%)				
26	Yes	377 (5.3)	244 (4.9)	133 (6.1)	0.039
27 28	No	6741 (94.7)	4711 (95.1)	2030 (93.9)	
28 29	Silica dust, n (%)			,	
30	Yes	730 (10.3)	523 (10.6)	207 (9.6)	0.223
31 32	No	6388 (89.7)	4432 (89.4)	1956 (90.4)	
33	Asbestos dust, n (%)			,	
34 35	Yes	981 (13.8)	691 (13.9)	290 (13.4)	0.570
36	No	6137 (86.2)	4264 (86.1)	1873 (86.6)	0.070
37	Benzene, n (%)	0107 (00.2)	1201 (00.1)	1075 (00.0)	
38 39	Yes	1981 (27.8)	1360 (27.4)	621 (28.7)	0.287
40	No	5137 (72.2)	3595 (72.6)	1542 (71.3)	0.207
41 42	Lead, n (%)	5157 (72.2)	3333 (12.0)	1542 (71.5)	
43	Yes	373 (5.2)	246 (5.0)	127 (5.9)	0.128
44	No	6745 (94.8)	4709 (95.0)	2036 (94.1)	0.120
45 46	Noise, n (%)	0743 (94.8)	4709 (95.0)	2030 (94.1)	
47	Yes	4942 (69.4)	3420 (69.0)	1522 (70.4)	0.270
48 49		· · · · ·		. ,	0.270
50	No Demociliaria en (94)	2176 (30.6)	1535 (31.0)	641 (29.6)	
51 52	Brucellosis, n (%)	101 (17)	9((1,7))	25 (1 ()	0.900
52 53	Yes	121 (1.7)	86 (1.7)	35 (1.6)	0.800
54	No	6997 (98.3)	4869 (98.3)	2128 (98.4)	
55 56	ERI, n (%)		2172 (42.0)	074 (45.0)	0.050
57	Yes	3147 (44.2)	2173 (43.9)	974 (45.0)	0.372
58 50	No	3971 (55.8)	2782 (56.1)	1189 (55.0)	
59 60					

BMJ Open

VIF

-

2.28

2.79

2.51

1.35

1.34

1.32

2.70

1.30

0.033

1							
2 3	CMBI, n (	<sup>(%)</sup>					
4 5		No	959	(13.5)	674 (13.6	)	285 (13.2)
6		Mild	2667	(37.5)	1813 (36.6	5)	854 (39.5)
7 8		Moderate	2900	(40.7)	2031 (41.0	))	869 (40.2)
9		Severe	592	(8.3)	437 (8.8)		155 (7.2)
10 11	263						
12	264	3.2. Feature selection					
13 14	265						
15	266	The lambda was smallest at 0.01	801 as s	een from	the lasso results when	n there were	12 characteristics,
16 17	267	which were education, profession	al title, a	ge, CMB	I, ERI, asbestos dust, l	nypertension	, diabetes, working
17 18	268	hours per day, working years, man	rital statu	is, and w	ork schedule based on	the results of	f the questionnaires
19	269	on demographics, occupational st	tress, job	burnout	and occupational expo	osure factors	(Figure 2).
20 21	270						
22	271	3.3. Results of logistic regressio	n model				
23 24	272			~ ~			
25	273	The 12 features obtained from t		-	1		-
26 27	274	regression model and the regress					
27	275	education, professional title, age,					- 1
29	276	day, working years, marital status			-		
30 31	277	health problems. In addition, ther					
32	278	in the model. The forest plot show					
33 34	279	which the degree of severe of CM				·	the greatest impact
35	280	on the risk of mental health probl	ems amo	ong facto	ry workers and miners	(Figure 3).	
36 37	281	Table 2 Predictive factors of	risk for r	nental he	alth problems among	factory work	ers and miners
38		Variable	β	S.E.	OR(CI95%)	Wald	P
39 40		Intercept	-2.33	0.25	0.10(0.06,0.16)	-9.357	0
41	Educati	on level					
42 43	Junior	school and below VS High school	0.34	0.13	1.41(1.10,1.81)	2.727	0.006**
44	Juni	or school and below VS Junior					
45 46		college	0.44				
47	Junior	conege	0.11	0.11	1.56(1.24,1.95)	3.850	<b>&lt;</b> 0.001***
48 49		school and below VS Bachelor's	0.11	0.11	1.56(1.24,1.95)	3.850	< 0.001***
49		-	0.38	0.11	1.56(1.24,1.95) 1.46(1.13,1.87)	3.850 2.953	< 0.001*** 0.003**
50	Profess	school and below VS Bachelor's					
50 51	Profess	school and below VS Bachelor's degree or above					
50 51 52	Profess	school and below VS Bachelor's degree or above ional title	0.38	0.13	1.46(1.13,1.87)	2.953	0.003**
50 51 52 53 54	Profess	school and below VS Bachelor's degree or above ional title None VS Primary	0.38 0.15	0.13 0.09	1.46(1.13,1.87) 1.16(0.97,1.39)	2.953 1.582	0.003** 0.114
50 51 52 53 54 55	Profess Work so	school and below VS Bachelor's degree or above ional title None VS Primary None VS Middle None VS Senior	0.38 0.15 0.05	0.13 0.09 0.09	1.46(1.13,1.87) 1.16(0.97,1.39) 1.05(0.87,1.26)	2.953 1.582 0.519	0.003** 0.114 0.604
50 51 52 53 54	Work s	school and below VS Bachelor's degree or above ional title None VS Primary None VS Middle None VS Senior	0.38 0.15 0.05	0.13 0.09 0.09	1.46(1.13,1.87) 1.16(0.97,1.39) 1.05(0.87,1.26)	2.953 1.582 0.519	0.003** 0.114 0.604

Day and night shifts	VS Shift	0.01	0.12	1.01(0.81,1.27)	0.107	0.915	2.4
Marital status							
Unmarried VS M	arried	0.16	0.13	1.18(0.91,1.52)	1.263	0.206	2.2
Unmarried VS Div	vorced	0.55	0.19	1.73(1.20,2.51)	2.918	0.004**	1.6
Unmarried VS Wi	dowed	0.69	0.43	1.99(0.85,4.64)	1.586	0.113	1.0
Age							
~25 VS 25~	-	0.02	0.20	0.98(0.66,1.47)	-0.083	0.934	3.0
~25 VS 30~	-	0.02	0.22	0.98(0.64,1.50)	-0.090	0.929	4.7
~25 VS 35~		0.56	0.23	1.76(1.13,2.74)	2.503	0.012*	5.0
~25 VS 40~		0.33	0.23	1.39(0.88,2.21)	1.419	0.156	4.9
~25 VS 45~		0.23	0.22	1.26(0.81,1.95)	1.018	0.308	10.
Working years							
~5 VS 5~		0.44	0.14	1.55(1.18,2.05)	3.114	0.002**	2.2
~5 VS 10~		0.06	0.15	1.06(0.78,1.43)	0.366	0.714	2.4
~5 VS 15~		0.06	0.20	1.06(0.72,1.56)	0.305	0.760	1.7
~5 VS 20~		0.29	0.18	1.33(0.95,1.88)	1.641	0.101	2.6
~5 VS 25~		0.48	0.17	1.61(1.15,2.25)	2.782	0.005**	3.9
~5 VS 30~		0.20	0.16	1.22(0.89,1.68)	1.239	0.216	3.9
Working hours per day							
≤7 VS >7	-	0.50	0.09	0.61(0.50,0.73)	-5.363	< 0.001***	1.1
Diabetes							
No VS Yes		0.43	0.14	1.53(1.16,2.03)	2.974	0.003**	1.0
Hypertension							
No VS Yes	1	0.52	0.09	1.69(1.42,2.00)	5.885	< 0.001***	1.1
Asbestos dust							
						< 0.001***	
No VS Yes	1	0.44	0.10	1.55(1.28,1.87)	4.474	<b>4</b> 0.001	1.0
ERI							
No VS Yes		0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***	1.0
CMBI							
No VS Mild		0.26	0.12	1.30(1.03,1.64)	2.175	0.003**	2.7
No VS Modera	ate	1.30	0.11	3.67(2.93,4.59)	11.361	< 0.001***	2.8
No VS Sever	e	2.99	0.18	19.84(13.88,28.34)	16.41	< 0.001***	1.4
$\begin{array}{c} 282 \\ 283 \end{array}$ Note: $\beta$ is the regre	ssion coefficient. "*	**" indi	icates P<	0.001, "**" indicates <i>P</i> <	0.01, "*" indi	cates <i>P</i> <0.05.	

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

#### **BMJ** Open

287 CMBI, ERI, asbestos dust, hypertension, diabetes, working hours per day, working years, marital status, 288 and work schedule were included in the nomogram. A model incorporating the above independent 289 predictors was developed and represented as a nomogram in Figure 4. Each variable in nomogram was 290 assigned a score, and the cumulative sum of each 'point' was the 'total score'. The "total score" 291 corresponded to the "predictable likelihood", which was the predicted probability of mental health 292 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, with a calculated total score of 174 and a corresponding risk probability of 8.27% for mental health problems.

#### 300 3.5 The validation of calibration

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

#### **3.6 The validation of discrimination**

ROC was plotted for the training and validation groups, and the AUC of training and the verification
groups were 0.785 and 0.784, respectively (Figure 6). The AUC of training and the verification groups
were both greater than 0.75, showing a good discrimination.

#### **3.7 Decision Curve Analysis**

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

#### 322 4. Discussion

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

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

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

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

#### **BMJ** Open

with other studies that had shown that occupational stress was a significant predictor of anxiety and was
negatively associated with mental health. In addition, there is a high correlation between burnout and
depression <sup>[39]</sup>.

In line with previous studies, working years was also an important influential factor in this study. Related study has shown that employment could improve patients' mental health, while unemployment could lead to a deterioration in mental health [40]. In China, workers' working years is an important aspect of employment, and researchers have studied this aspect and found that precarious employment is a source of stress for individuals and predisposes them to mental health problems [41]. In addition, environmental factors were also one of the influential factors of mental health problems in our study. Relevant studies have found that exposure to air pollution is associated with increased suicide risk and depressive symptoms <sup>[42]</sup>. Hypertension and diabetes were the influential factors in this study. A study has shown that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher than in the non-diabetic population <sup>[43]</sup>. Furthermore, there is a recognized association between hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear [44] 

Factory workers and miners were inevitably exposed to occupational hazards such as benzene and asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous enterprises, involving more than 30 different types of operations, of which factory workers and miners is the one <sup>[45]</sup>. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for mental health problems in this study. Our study found that the work schedules of factory workers and miners were vary and the phenomenon of night shifts was very common, which inevitably affected their normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could suffer from negative emotions such as anxiety and depression [46]. Professional title and education level were also important influences on mental health issues. In the workplace, generally speaking, the higher the professional title and education level, the higher the status of the worker in the company and the greater the role played in the position. The number of studies on socio-economic status and mental health had increased in recent years. Some of these studies have shown that major depression is higher in the low socio-economic status group <sup>[47]</sup>. It has also been suggested that education itself is the best indicator of socio-economic status <sup>[48]</sup>. Marital status was one of the influential factors for mental health problems. Many studies have found an association between mental health and gender, marital status, lifestyle and working conditions, and it has been shown that poor mental health in women is associated with divorce or widowhood [49]. In this study, working more than seven hours a day was a determinant factor on mental health problems, which was consistent with other studies that had shown that long working hours could have a negative impact on employees' mental health and that excessive workloads could increase workers' fatigue, which in turn could lead to anxiety and depression <sup>[50]</sup>.

In China, there are many problems in identifying people with mental health problems due to uneven and imperfect levels of medical development across regions. Some studies have shown that in mainland China, general practitioners, surgeons and primary health care workers often have little or no mental health training, which prevents them from providing basic mental health services [51]. Non-mental health professionals in general hospitals learn about mental illness on their own, rather than learning about it during their formal education<sup>[52]</sup>. 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. This study has several strengths. First, to our knowledge, this is the first model to develop and assess the likelihood of mental health problems in a group of factory workers and miners. Secondly, the nomogram in this study includes demographic information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors that are closely related to factory workers and miners, allowing for a more accurate assessment of the risk of morbidity among them, as well as providing a methodological reference for other related studies. 

#### 423 5. Limitations

This study also has several limitations. Firstly, 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.

#### 432 Patient and public involvement

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

435 Acknowledgements The authors are grateful to all participants and investigators.

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

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

# BMJ Open

2		
	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	
o 9 '	451	Patient consent for publication Not applicable.
10	452	
11	453	Ethics approval The study was approved by the ethics committee of Urumqi Center for Disease Control
12 13 '	454	and Prevention (20181123)
	455	
	456	Data availability statement Data are available on reasonable request. The data used in this study are
16 17	457	available from the corresponding authors on reasonable request.
	458	
19	459	References
20 21	460	[1] WHO Terminology Information System [online glossary] http://www.who.int/health-systems-
~~	461	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
	464	population. <i>Sci Rep</i> 2017;7:14975.
27	465	[3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental
28	466	health. Lancet. 2007;370:859–77.
	467	[4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive
31	468	symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. <i>Health</i>
32 33 <sup>4</sup>	469	<i>Place</i> 2016;42:37–46.
	470	[5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,
	471	Degenhardt L, Dicker D, et al. Global, regional, and national incidence, prevalence, and years lived
36 37 <sup>'</sup>	472	with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A
	473	systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743-800.
	474	[6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health?.Int J
40 41	475	Environ Res Public Health 2020;17:7530.
42 '	476	[7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative
	477	income, gender, and location. China Econ Rev 2018;48:83-101.
44 45	478	[8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:
46 '	479	Results from the PATH Through Life Survey. Soc Sci Med. 2009;69:229-237.
	480	[9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and
48 , 49	481	Subjective Well-being? J Environ Econ Manag 2017;85:81–94.
49 50	482	[10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. <i>World</i>
51 4	483	Psychiatry. 2015;14:36–42.
52 <i>.</i> 53	484	[11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson
54	485	SJ. IPS Supported employment: a review. J Vocat Rehabil 2014;40:1–13.
55	486	[12] Jia Z, et al. Exposure to ambient air particles increases the risk of mental disorder: findings from a
		notional and anima at in Dailing Let I Funition Day Dublic II. alth 2019.15.160
	487	natural experiment in Beijing. Int J Environ Res Public Health 2018;15:160.
58	488	[13] Szyszkowicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency
58		

1

1		
2 3		
4	490	[14] Michael J. S. International classification of sleep disorders-third edition : highlights and
5	491	modifications. Chest 2014;146:1387–1394.
6 7	492 493	[15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial infarction: gender differences. <i>Open Nurs J</i> 2018;12:205–214.
8	494	[16] Wurm W, Vogel K, Holl A, <i>et al.</i> Depression-burnout overlap in physicians. <i>PLoS One</i>
9	495	2016;11:e0149913.
10 11		
12	496	[17] Porru F, Robroek S J W, Bültmann U, <i>et al.</i> Mental health among university students: The
13	497	associations of effort-reward imbalance and overcommitment with psychological distress. J Affect
14	498	Disord 2021;282:953-961.
15 16	499	[18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,
16 17	500	and burnout among physician assistant students: a multi-institutional study. Health Professions Edu
18	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 22	504	the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. EClinical Med
23	505	2020;24:100424.
24	506	[20] Armah E K, Adedeji J A, Boafo B B, et al. Underground gold miner exposure to noise, diesel
25	507	particulate matter and crystalline silica dust. <i>J Health Pollut</i> 2021; 11(29):210301.
26 27	508	[21] Hall N B, Blackley D J, Halldin C N, <i>et al.</i> Current review of pneumoconiosis among US coal
27 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	510	
31		workers in sri lanka. Ann Glob Health, 2019, 85(1):108.
32 33	512	[23] Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP. The estimation of the number of
34	513	underground coal miners and the annual dose to coal miners in China. <i>Health Phys</i> 2007;93:127-
35	514	132.
36	515	[24] Yong X, Gao X, Zhang Z, et al. Associations of occupational stress with job burn-out, depression
37 38	516	and hypertension in coal miners of Xinjiang, China: A cross-sectional study. BMJ open 2020;10:
38 39	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.
43 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.
48 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. <i>Psychiatry Res</i> , 2022, 310:
51	527	114436.
52	528	[29] Lu Y, Zhang Z, Yan H, <i>et al.</i> Effects of occupational hazards on job stress and mental health of
53 54	520 529	
55		factory workers and miners: A propensity score analysis. <i>BioMed Res Int</i> , 2020, 2020: 1754897.
56	530	[30] Li X, Jiang T, Sun X, <i>et al.</i> 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 59	532	and the environment. J Psychiatr Res, 2021, 135: 76-85.
59 60	533	[31] Derogatis L, Lipman R.S, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary
		17

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2		
3	534	report. Psychopharmacol Bull 1973;9:13-28.
4 5	535	[32] Crespo-Maraver M, Doval E, Fernã N.J, Gimã Nez-Salinas J, Prat G, Bonet P. Caregiver's health:
6	536	adaption and validation in a Spanish population of the Experience of Caregiving Inventory
7	537	(ECI) Gac. Sanit 2018;33:348-355.
8 9	538	[33] Dang W, Xu Y, Ji J, et al. Study of the scl-90 scale and changes in the chinese norms. Front
9 10	539	<i>Psychiatry</i> , 2020, 11: 524395.
11	540	[34] Siegrist J. Adverse health effects of high-effort/low-reward conditions. J Occup. Health Psychol.
12	541	1996;1:27–41.
13 14	542	[35] Siegrist J, Wege N, Pühlhofer F, Wahrendorf M. A short generic measure of work stress in the era
15	543	of globalization: Effort-reward imbalance. Int Arch Occup Environ Health 2009;82:1005–1013.
16	544	[36] Zhang Z, Lu Y, Yong X, <i>et al.</i> Effects of occupational radiation exposure on job stress and job
17	545	burnout of medical staff in xinjiang, China: A cross-sectional study. Med Sci Monit 2020;26:
18 19	546	e927848.
20	547	[37] Freudenberger HJ. Staff burnout. J Soc Issues 1974;30:159–165.
21	548	[38] Wang L, Chen L, Jia F, <i>et al.</i> Risk factors and prediction nomogram model for psychosocial and
22 23	549	behavioural problems among children and adolescents during the covid-19 pandemic: A national
24	550	multicentre study: Risk factors of childhood psychosocial problems. J Affect Disord, 2021, 294:
25	551	128-136.
26 27	552	[39] Occupational Stress and Employees Complete Mental Health: A Cross-Cultural Empirical Study
27	553	[40] Knapp M and Wong G. Economics and mental health: the current scenario. <i>World Psychiatry</i>
29	554	2020;19:3–14.
30	555	[41] Benach, J, Vives, A, Amable, M, Vanroelen, C, Tarafa, G, & Muntaner, C. Precarious employment:
31 32	556	understanding an emerging social determinant of health. <i>Annu Rev Public Health</i> 2014;35:229-53.
33	557	[42] Bakian A. V. <i>et al.</i> Acute air pollution exposure and risk of suicide completion. <i>Am J Epidemiol</i>
34	558	2015;181:295–303.
35 36	559	[43] Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type
37	560	1 diabetes: systematic literature review. <i>Diabet Med</i> 2006;23:445–448.
38	561	[44] Gilsanz P, Karter AJ, Beeri MS, Quesenberry CP, Whitmer RA. The bidirectional association
39 40	562	between depression and severe hypoglycemic and hyperglycemic events in type 1
41	563	diabetes. Diabetes Care 2018;41:446–452.
42	564	[45] Lu Y, Zhang Z, Yan H, et al. Effects of occupational hazards on job stress and mental health of
43 44	565	factory workers and miners: A propensity score analysis. <i>BioMed Res Int</i> 2020;2020:1754897.
44	566	[46] Shi L, Liu Y, Jiang T, <i>et al.</i> Relationship between mental health, the clock gene, and sleep quality
46	567	in surgical nurses: A cross-sectional study. <i>BioMed Res Int</i> 2020;2020:4795763.
47	568	[47] Sallis JF, Saelens BE, Frank LD et al. Neighborhood built environment and income: examining
48 49	569	multiple health outcomes. Soc Sci Med 2009;68:1285–1293.
50	570	[48] Winkleby MA, Jatulis DE, Frank E <i>et al.</i> Socioeconomic status and health: how education, income,
51	571	and occupation contribute to risk-factors for cardiovascular disease. Am J Public
52 53	572	Health 1992;82:816–820.
54	573	[49] Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis P.N, Mavreas V. Prevalence
55	574	and sociodemographic associations of common mental disorders in a nationally representative
56 57	575	sample of the general population of Greece. <i>BMC Psychiatry</i> 2013;13:163.
58	576	[50] Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG et al. Long
59	577	working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II
60		18

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.

585 Figure legends

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

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 weas plotted based on the lambda series in Figure 1(A), and 12
features with non-zero coefficients were selected by optimal lambda.

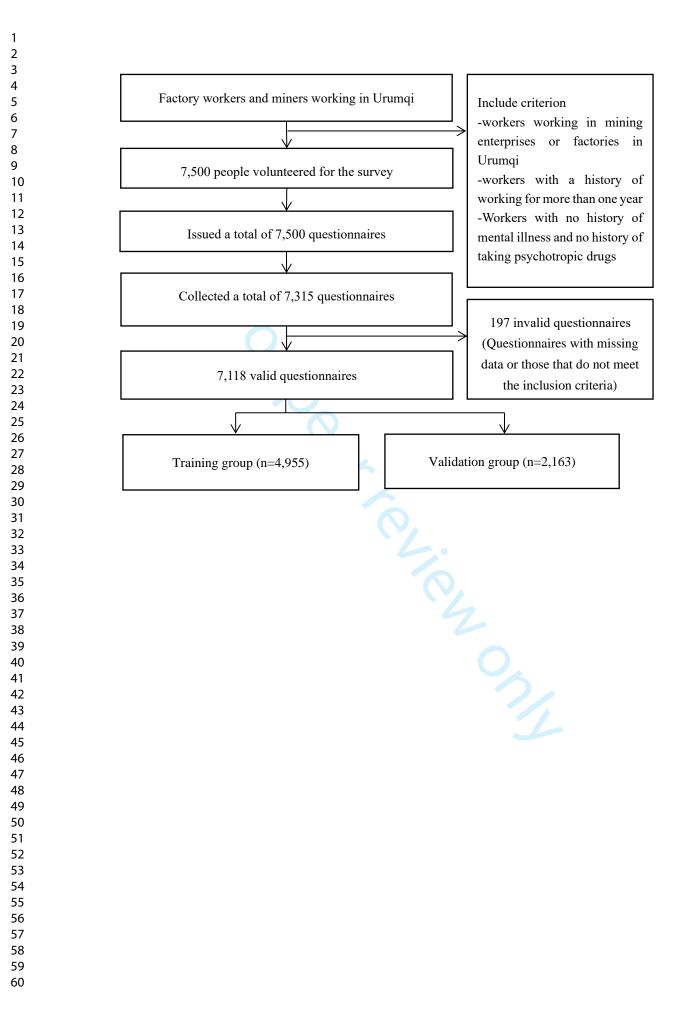
595 Fig.3. The forest plot of the OR of the selected feature.

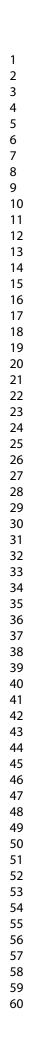
Fig.4. Developed mental health problems incidence risk nomogram. The mental health problems incidence risk
nomogram was developed in the array, with education, professional title, age, CMBI, ERI, asbestos dust,
hypertension, diabetes, working hours per day, working years, marital status, and work schedule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

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

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

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





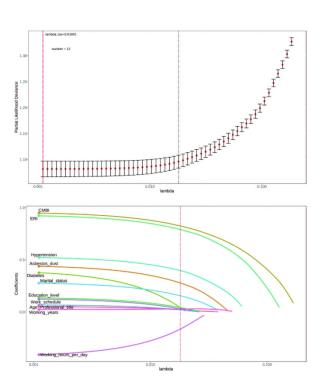
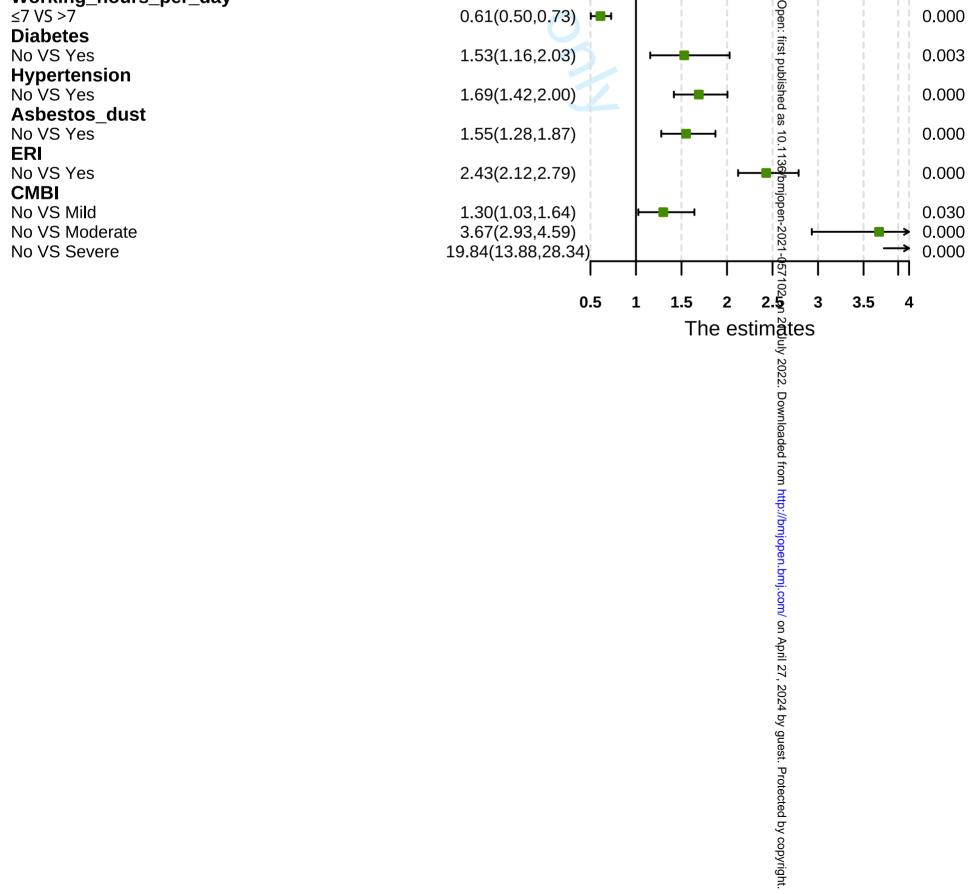
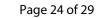


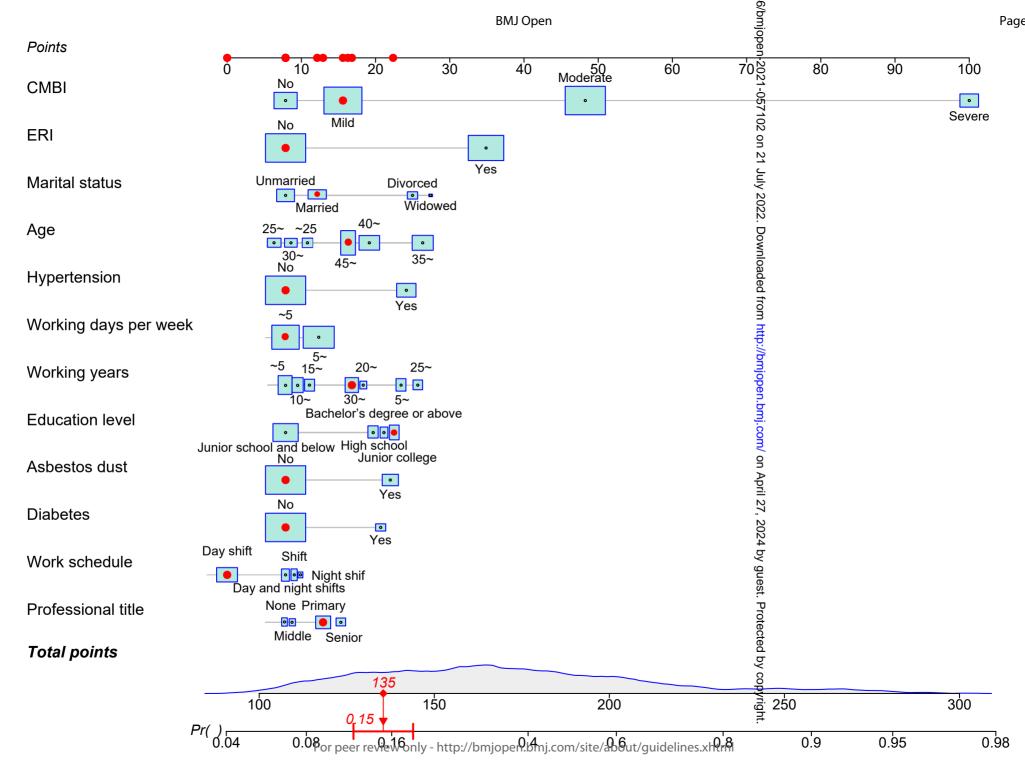
Fig.2. Feature selection using the LASSO binary logistic regression model. (A) Feature selection for the LASSO binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against lambda by validating the optimal parameter lambda in the LASSO model. Dotted 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 weas 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)

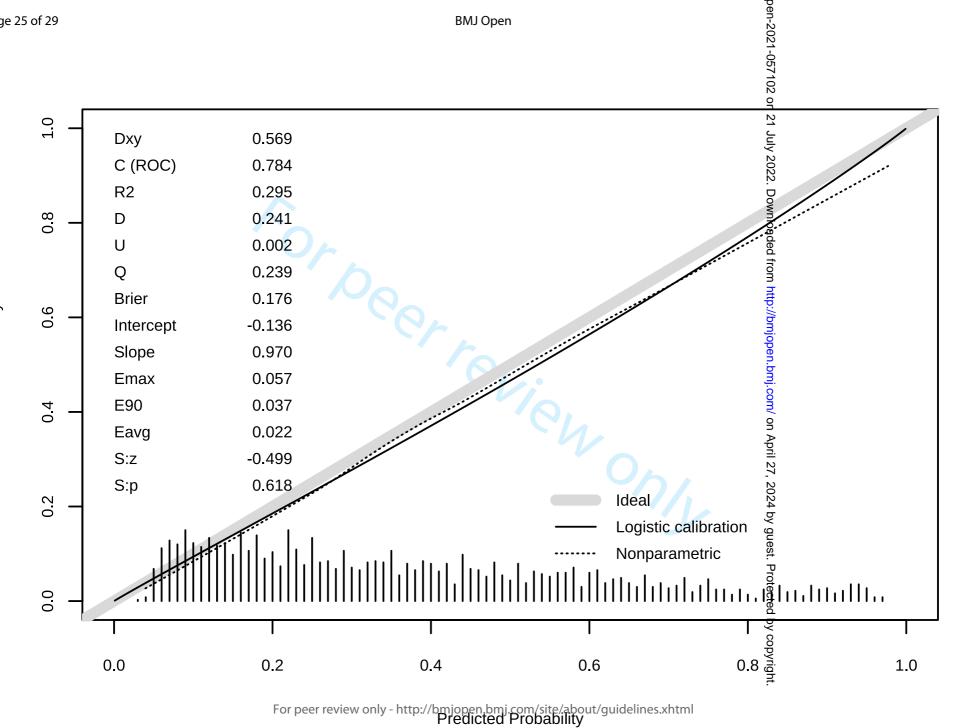
20			
21			
22			
23			
24 25			
26			
27			
28			
29			
30 31			
32			
33			
<sup>34</sup> <sub>25</sub> Variable	OR(Cl95%)		<b>P-Value</b>
<sup>35</sup> <sub>36</sub> Education_level			
<sup>36</sup> Junior school and below VS High school	1.41(1.10,1.81)		0.006
<sup>38</sup> Junior school and below VS Junior college	1.56(1.24, 1.95)		0.000
<sup>39</sup> Junior school and below VS Sumor conege	1.46(1.13,1.87)		0.000
<ul> <li><sup>39</sup>Junior school and below VS Bachelor's degree or above</li> <li><sup>40</sup>A<sub>1</sub><b>Professional_title</b></li> </ul>	1.40(1.13,1.07)		0.003
$_{41}$ FORSSIONAL III $_{42}$ None VS Primary	1.16(0.97,1.39)		0.114
<sup>42</sup> None VS Middle			0.604
<sup>43</sup> None VS Senior	1.05(0.87,1.26)		0.004
<sup>45</sup> Work, ashadula	1.30(1.06,1.61)		0.014
<sup>45</sup> Work_schedule			0.001
47 Day and night shifts VS Day shift	0.69(0.55,0.85)		0.001
48 Day and night shifts VS Night shif	1.01(0.68,1.49)		0.965
<sup>49</sup> Day and night shifts VS Shift	1.01(0.81,1.27)		0.915
<sup>50</sup> Marital_status	4 40(0 04 4 50)		0.000
52 Unmarried VS Married	1.18(0.91,1.52)		0.206
53 Unmarried VS Divorced	1.73(1.20,2.51)		0.004
54 Unmarried VS Widowed	1.99(0.85,4.64)		→ 0.113
<sup>55</sup> Age			0.004
<sup>56</sup> <sub>57</sub> ~25 VS 25~	0.98(0.66,1.47)		0.934
58~25 VS 30~	0.98(0.64,1.50)		0.929
59~25 VS 35~	1.76(1.13,2.74)		0.012
<sup>60</sup> ~25 VS 40~	1.39(0.88,2.21)		0.156
~25 VS 45~	1.26(0.81,1.95)		0.308
Working_years			
~5 VS 5~	1.55(1.18,2.05)	¦ I⊷⊨≡→+ ¦ ¦ ¦	0.002
~5 VS 10~	1.06(0.78,1.43)		0.714
~5 VS 15~	1.06(0.72,1.56)		0.760
~5 VS 20~	1.33(0 <mark>.95</mark> ,1.88)		0.101
~5 VS 25~	1.61(1.15,2.25)		0.005
~5 VS 30~	1.22(0.89,1.68)		0.216
Working_hours_per_day			
≤7 VS >7	0.61(0.50,0.73)	Here Ppen:	0.000
Diabetes			

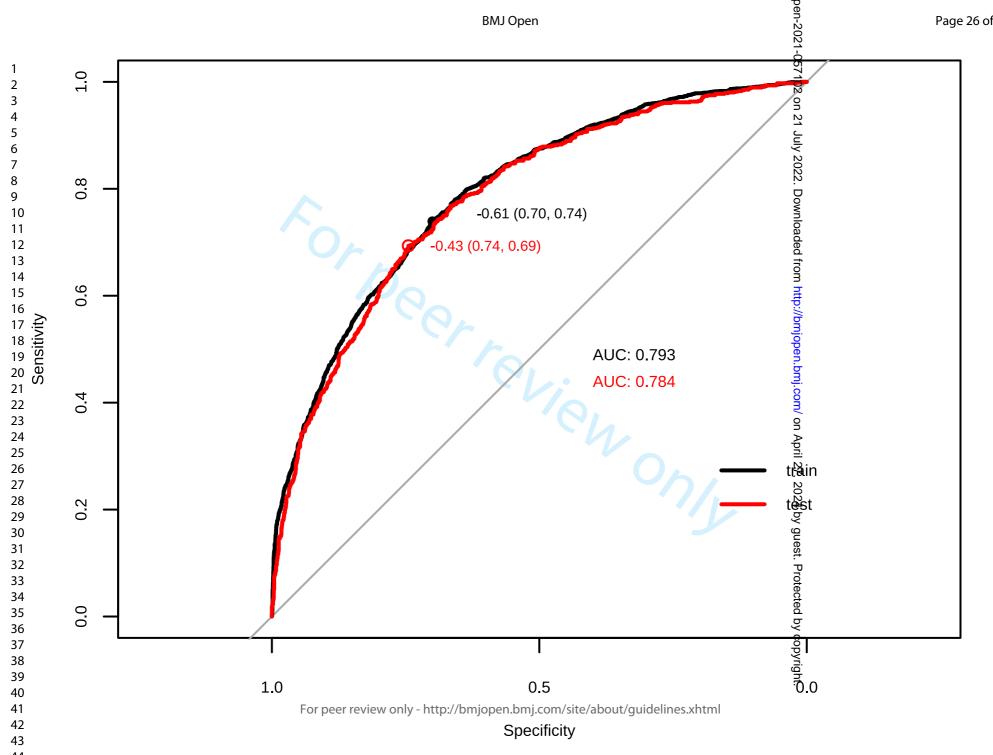


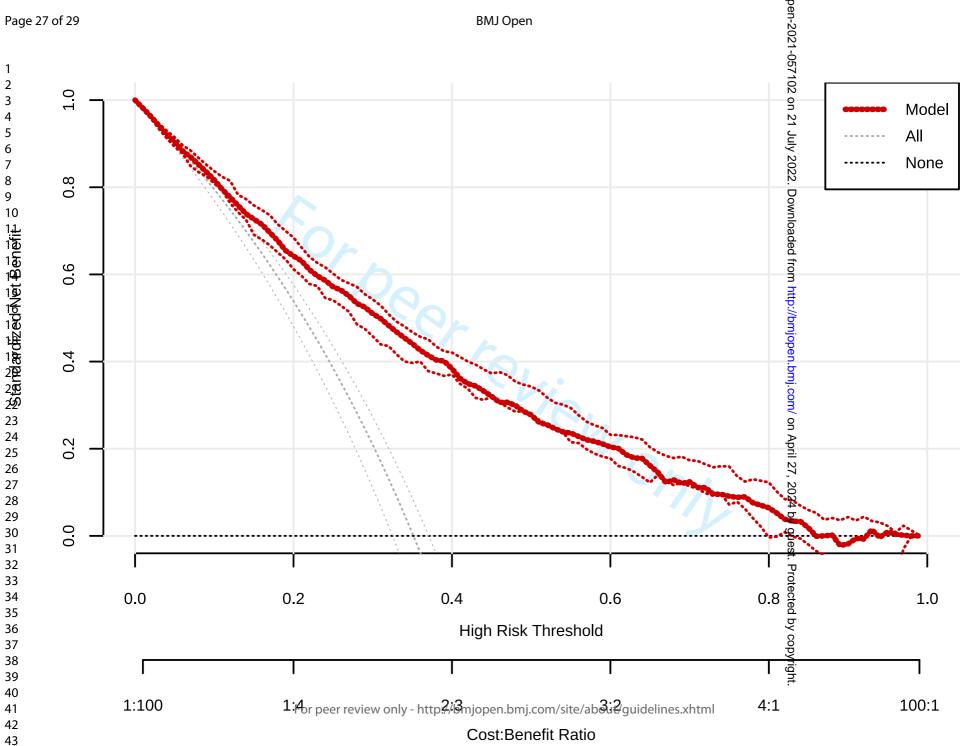












A A

STROBE Statement-checklist of items that should be included in reports of observation	onal studies
---	--------------

3

STROBE Statement	—che	cklist of items that should be included in reports of observational studies	6/bmjopen-2021-057	
	Item No.	Recommendation	<sup>N</sup> Page	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		
		( <i>b</i> ) Provide in the abstract an informative and balanced summary of what was done and what was found	21 July 2022	
Introduction			Do	
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	<u>vnlo</u> 2	
Objectives	3	State specific objectives, including any prespecified hypotheses	ad 3	
Methods			ed fro	
Study design	4	Present key elements of study design early in the paper	3 4	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4 ttp://bm	
Participants	6	<ul> <li>(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up</li> <li>Case-control study—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</li> <li>Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants</li> </ul>	2     3       2     3       4     4       4     4       4     4       5     4       4     4       4     4       5     4       4     4       4     4       4     4       4     4       5     4       4     4       4     4       4     4       5     4       4     4       4     4       4     4       4     4       4     4       5     5       6     6       6     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7       7     7	
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	il 27, 2024 by (	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	guest. 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	Protected by 4	
Bias	9	Describe any efforts to address potential sources of bias		
Study size	10	Explain how the study size was arrived at	copyright.	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

#### BMJ Open

f 29		BMJ Open	6/bmjopen-2021-057
			97 20 02
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	$\rightarrow$
Statistical	12	(a) Describe all statistical methods, including those used to control for confounding	8 6
methods		(b) Describe any methods used to examine subgroups and interactions	0 N
		(c) Explain how missing data were addressed	21 24 2022 22.
		(d) Cohort study—If applicable, explain how loss to follow-up was addressed	× 4
		Case-control study—If applicable, explain how matching of cases and controls was addressed	022
		Cross-sectional study—If applicable, describe analytical methods taking account of sampling	
		strategy	NW NI
		(e) Describe any sensitivity analyses	Download ad ed
Results		6	ed f
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined	from 7 http://www.analysia.com
		for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	http://
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on	<u>50</u> 7 7 5.5 5.5 7
1		exposures and potential confounders	, and the second s
		(b) Indicate number of participants with missing data for each variable of interest	<u> </u>
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	on
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time	April
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision	024
		(eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were	by g
		included	, 2010 24 by guest
		(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	otec
		period	ted
Continued on next page			Protected by copyright
			opy .
			rrigt

		BMJ Open	6/bmjop	Page 30
			ben-20	
Other analyses	17	Report other analyses done-eg analyses of subgroups and interactions, and sensitivity analyses	21-0	
Discussion			571	
Key results	18	Summarise key results with reference to study objectives	02 o	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss	n 21	15
		both direction and magnitude of any potential bias	l Ju	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of	ly 202	13
		analyses, results from similar studies, and other relevant evidence	022.	
Generalisability	21	Discuss the generalisability (external validity) of the study results	Do	13
Other informati	ion		vnlo	
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the	ade	15
		original study on which the present article is based	d fre	
			Ĕ	

\*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.

a the STROBE Initiative is available at www.strope-on April 27, 2024 by guest. Protected by copyright.

# **BMJ Open**

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

Journal:	BMJ Open
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





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 <u>licence</u>.

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 <u>Creative Commons</u> 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.

review only

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	<b>Results</b> 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
20 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
31	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	<b>Conclusion</b> 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.

1 2		
3	39	Key words Mental health; Predictor; Nomogram; Risk; Factory workers and miners
4 5	40	
6	41	Strengths and limitations of this study
7	42	1. This is the first study to develop an easy-to-use nomogram to predict the mental health risks of factory
8 9	43	workers and miners.
10	44	2. The AUC of training group and verification group were 0.785 and 0.784 respectively, showing
11 12	45	moderate discriminatory and calibration power.
13	46	3. This nomogram model's variables are more comprehensive, including demographics, burnout,
14 15	47	occupational stress and occupational hazards.
16	48	4. We had considered many influential factors, but we were still not certain whether all possible
17	49	influences were covered.
18 19	50	5. There is a lack of external validation in other populations in other regions and countries.
20	51	O,
21 22	52	1. Introduction
23	53	
24 25	54	The World Health Organization (WHO) defines health as a state of complete physical, mental and social
26	55	well-being and not merely the absence of disease or weakness <sup>[1]</sup> . Obviously, health is an organic unity
27 28	56	of physical and mental well-being. People with good mental health are the precondition for the normal
29	57	operation of our society. However, with the acceleration of people's pace of life, people are facing an
30	58	increasing risk of poor health, which has become a global public health problem <sup>[2]</sup> . Mental health
31 32	59	problems can not only take a toll on physical health such as increasing the risk of communicable and
33	60	non-communicable diseases and even causing unintentional or intentional harm to others <sup>[3]</sup> , but can also
34 35	61	have a negative impact on the economy. For example, mental health disorders represent a growing part
36	62	of the global burden of disease <sup>[4]</sup> , with statistics showing that nearly one billion people worldwide
37 38	63	currently suffer from a mental disorder, and mental illness is ranked as one of the leading causes of the
39	64	global burden of disease <sup>[5]</sup> . Moreover, one study has estimated that due to the impact of mental illness,
40	65	the global economy loses US \$1 trillion every year <sup>[6]</sup> .
41 42	66	
43	67	As researchers around the world have delved into the field of mental health, factors such as gender,
44 45	68	income levels, environment and education have been found to be associated with people's mental health
46	69	problems <sup>[7-10]</sup> . Moreover, employment is also strongly associated with quality of life, higher self-esteem
47 48	70	and fewer psychiatric symptoms <sup>[11]</sup> . In addition, in the context of the global challenges of climate change,
49	71	an increasing number of scholars have been examining the epidemiological links between mental health
50 51	72	and environmental factors. Some studies have suggested that mental health may be influenced by ambient
52	73	temperature, and an association has been found between environmental pollutants, particularly fine
53	74	particulate matter, and mental health problems <sup>[12]</sup> . A relevant study shows that with short-term exposure
54 55	75	to ambient air pollution is associated with increased emergency room visits due to depression or suicide
56	76	attempts <sup>[13]</sup> . Furthermore, other factors associated with mental health include sleep, diabetes, coronary
57 58	77	artery disease and cardiovascular disease <sup>[14-15]</sup> . It is worth noting that job burnout and occupational stress
59	78	are closely linked to mental health. Job burnout is an exhaustion state of physical and psychological that
60		

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

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 [18-19]. Through a review of the literature, our group found that coal dust, crystalline silica and noise pollution were common causes of health problems for workers in underground mines <sup>[20]</sup>. And, exposure to coal mine dust is a significant cause of pneumoconiosis in coal miners<sup>[21]</sup>. In addition, asbestos is one of the major occupational hazards in the daily work of workers in the construction and automotive industries <sup>[22]</sup>. China has the world's largest group of factory workers and miners, about 6 million <sup>[23]</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>[24]</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.

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

# **BMJ** Open

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

# 125 2. Materials and Methods

# **2.1 Calculation of sample size**

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

# 138 2.2. Participants

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

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.

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.

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

can choose to continue answering the survey if they wish to participate and the relevant data will be used for scientific research, or refuse to answer if they do not wish to participate in the survey. In addition, this survey was conducted by trained surveyors who explained the purpose, meaning, content and requirements of the questionnaire to all participants and provided on-site instructions to ensure the return rate of the questionnaire. All participants understood the purpose of the study and were willing to participate in the study. 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%. 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) (Figure 1).

- 169 2.3. Research Methods
- **2.3.1.** Assessment of mental health

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

- **2.3.2.** Assessment of occupational stress

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

# **2.3.3.** Assessment of job burnout

195 In this survey, the Chinese Maslach Burnout Inventory (CMBI) revised by Li et al. was used to assess 196 job burnout, which has good reliability and validity <sup>[36]</sup>. CMBI is composed of 15 items in three 197 dimensions: emotional exhaustion (5 items), depensionalization (5 items) and reduced personal Page 7 of 26

# **BMJ** Open

accomplishment (5 items). The score for each item ranges from 1 to 7, with 1 point indicating complete compliance and 7 points indicating complete non-compliance. According to the critical value (emotional exhaustion  $\geq 25$ , dependent dependence of a constraint dependence of a co burnout are divided into none (all aspects are below the critical value), mild (any one aspect is equal to or above the critical value), moderate (any two aspects are equal to or higher than the critical values), and severe (three aspects are equal to or higher than the critical values) [<sup>37]</sup>. In this survey, Cronbach  $\alpha$ was 0.89, the half-reliability coefficient was 0.86 and the KMO was 0.919.

#### 2.3.4. Candidate predictors

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

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 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.

### 2.4. Statistical analysis

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. Statistically significant predictors were applied to develop a prediction model for the risk of mental health problems among factory workers and miners by introducing all selected factors and analyzing the statistical significance levels of them. 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 (http://www.r-project.org). The significance level ( $\alpha$ ) set at 0.05.

# **2.5. Patient and public involvement**

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

**3. Results** 

## 3.1. Participant characteristics

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 (5.3), 730 (10.3), 981 (14), 1,981 (27.8), 373 (5.2), 4,942 (69.4) and 121 (1,7) 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.

	279					
	280					
		Table 1 Characteristics	of the study particip	oants		
	Varia	bles	Total $(n = 7118)$	train (n = 4955)	test (n = 2163)	р
	Sex, n (%)	Male	4649 (65.3)	3216 (64.9)	1433 (66.3)	0.28
		Female	2469 (34.7)	1739 (35.1)	730 (33.7)	
	Ethnicity, n (%)	Han	5762 (80.9)	3982 (80.4)	1780 (82.3)	0.06
		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)	0.76
		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)	0.92
		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)	0.58
	work seneduce, if (70)	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)	0.21
		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)	
Ν	Ionthly income (yuan), n (%)	<3000	1799 (25.3)	1246 (25.1)	553 (25.6)	0.96
		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)	
	Age (years), n (%)	<25	431 (6.1)	297 (6.0)	134 (6.2)	0.17
		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)	
١	Working years (years), n (%)	<5	1170 (16.4)	794 (16.0)	376 (17.4)	0.24
	// //	5~	1065 (15.0)	736 (14.9)	329 (15.2)	

	10	0.07(14.0)	721 (14 ()	27((12.8))	
	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
(110413), 11 (70)	>7	5957 (83.7)	4141 (83.6)	1816 (84.0)	
Working days per week	≤5	4442 (62.4)	3107 (62.7)	1335 (61.7)	0.446
(days), n (%)		4442 (02.4)	5107 (02.7)	1333 (01.7)	0.440
	>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)	0.86
	No	5788 (81.3)	4026 (81.3)	1762 (81.5)	
Coal dust, n (%)	Yes	377 (5.3)	244 (4.9)	133 (6.1)	0.03
	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)	0.28
	No	5137 (72.2)	3595 (72.6)	1542 (71.3)	
Lead, n (%)	Yes	373 (5.2)	246 (5.0)	127 (5.9)	0.128
2000, 1 (70)	No	6745 (94.8)	4709 (95.0)	2036 (94.1)	0.120
Noise, n (%)	Yes	4942 (69.4)	3420 (69.0)	1522 (70.4)	0.270
10150, 11 (70)	No	2176 (30.6)	1535 (31.0)	641 (29.6)	0.27
Brucellosis, n (%)	Yes	121 (1.7)	86 (1.7)	35 (1.6)	0.800
Dideenosis, II (70)		~ /	· · ·		0.800
	No	6997 (98.3)	4869 (98.3)	2128 (98.4)	0.27
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)	

# **BMJ** Open

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

# **3.3. Results of logistic regression model**

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

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

Variable	β	S.E.	OR(CI95%)	Wald	Р	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 shif	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*	5.0
			10			

	~25 VS 40~	0.33	0.23	1.39(0.88,2.21)	1.419	0.156	4
	~25 VS 45~	0.23	0.22	1.26(0.81,1.95)	1.018	0.308	1(
Working	g years						
	~5 VS 5~	0.44	0.14	1.55(1.18,2.05)	3.114	0.002**	2
	~5 VS 10~	0.06	0.15	1.06(0.78,1.43)	0.366	0.714	2
	~5 VS 15~	0.06	0.20	1.06(0.72,1.56)	0.305	0.760	1
	~5 VS 20~	0.29	0.18	1.33(0.95,1.88)	1.641	0.101	2
	~5 VS 25~	0.48	0.17	1.61(1.15,2.25)	2.782	0.005**	3
	~5 VS 30~	0.20	0.16	1.22(0.89,1.68)	1.239	0.216	3
Working	g hours per day						
	≤7 VS >7	-0.50	0.09	0.61(0.50,0.73)	-5.363	< 0.001***	1
Diabetes	3						
	No VS Yes	0.43	0.14	1.53(1.16,2.03)	2.974	0.003**	1
Hyperter	nsion						
	No VS Yes	0.52	0.09	1.69(1.42,2.00)	5.885	<b>&lt;</b> 0.001***	1
Asbestos	s dust						
	No VS Yes	0.44	0.10	1.55(1.28,1.87)	4.474	<b>&lt;</b> 0.001***	1
ERI							
	No VS Yes	0.89	0.07	2.43(2.12,2.79)	12.786	< 0.001***	1
CMBI							
	No VS Mild	0.26	0.12	1.30(1.03,1.64)	2.175	0.003**	2
	No VS Moderate	1.30	0.11	3.67(2.93,4.59)	11.361	< 0.001***	2
	No VS Severe	2.99	0.18	19.84(13.88,28.34)	16.41	< 0.001***	1

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,

# **BMJ** Open

with a calculated total score of 174 and a corresponding risk probability of 8.27% for mental health problems. 3.5 The validation of calibration Model validation was carried out in the validation group. The prediction accuracy of the model was assessed by two aspects. (1) The Brier score for overall performance, which assessed the difference between observed and predicted values, with values closer to 0 indicating better predictive ability. (2) The calibration slope used for modal calibration, which assessed the agreement between the observed and predicted values, with values closer to 1 indicating better performance. The accuracy measurements for the bias correction were validated by the model with a Brier score of 0.176 and a calibration slope of 0.970, respectively (Figure 5). The prediction accuracy of the model was relatively high. 3.6 The validation of discrimination ROC was plotted for the training and validation groups, and the AUC of training and the verification groups were 0.785 and 0.784, respectively (Figure 6). The AUC of training and the verification groups were both greater than 0.750, showing a good discrimination. **3.7 Decision Curve Analysis** As shown in the DCA of the risk of mental health problems nomogram in Figure 7, the model for predicting the risk of mental health problems for factory workers and miners in this study was more practically relevant if the threshold probability of patients was >10%. 4. Discussion 

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

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

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

Mental health problems were very common in the group of factory workers and miners, and the prevalence of mental health of them was found to be 37.08% in our study. Notably, the CMBI showed the most significant score (score = 100) and the ERI also had a high score (score = 43) in mental health problem incidence risk nomogram, which indicated that both of them were relatively important factors for mental health problems among the group of factory workers and miners. Our finding was consistent with other studies that had shown that occupational stress was a significant predictor of anxiety and was negatively associated with mental health. In addition, there is a high correlation between burnout and depression [39].

In line with previous studies, working years was also an important influential factor in this study. Related study has shown that employment could improve patients' mental health, while unemployment could lead to a deterioration in mental health <sup>[40]</sup>. In China, workers' working years is an important aspect of employment, and researchers have studied this aspect and found that precarious employment is a source of stress for individuals and predisposes them to mental health problems <sup>[41]</sup>. In addition, environmental factors were also one of the influential factors of mental health problems in our study. Relevant studies

# **BMJ** Open

have found that exposure to air pollution is associated with increased suicide risk and depressive symptoms <sup>[42]</sup>. Hypertension and diabetes were the influential factors in this study. A study has shown that the prevalence of depression in adults with type 1 diabetes (T1D) is approximately three times higher than in the non-diabetic population <sup>[43]</sup>. Furthermore, there is a recognized association between hyperglycemia and depression, but the underlying biological mechanisms of this association are unclear <sup>[44]</sup>.

Factory workers and miners were inevitably exposed to occupational hazards such as benzene and asbestos dust in their working environment. According to statistics, a total of nearly 2 million workers are exposed to various occupational hazards and over 16 million people worked in toxic and hazardous enterprises, involving more than 30 different types of operations, of which factory workers and miners is the one [45]. Similarly, the occupational hazard asbestos dust was selected as a predictor of risk for mental health problems in this study. Our study found that the work schedules of factory workers and miners were vary and the phenomenon of night shifts was very common, which inevitably affected their normal sleep. Some studies have shown that sleep problem is a risk factor for a variety of mental health and chronic diseases. Lack of sleep or poor sleep quality could lead to abnormalities in the body's self-regulatory functions and disturbances in the circadian rhythm of the biological clock, which in turn could suffer from negative emotions such as anxiety and depression [46]. Professional title and education level were also important influences on mental health issues. In the workplace, generally speaking, the higher the professional title and education level, the higher the status of the worker in the company and the greater the role played in the position. The number of studies on socio-economic status and mental health had increased in recent years. Some of these studies have shown that major depression is higher in the low socio-economic status group <sup>[47]</sup>. It has also been suggested that education itself is the best indicator of socio-economic status <sup>[48]</sup>. Marital status was one of the influential factors for mental health problems. Many studies have found an association between mental health and gender, marital status, lifestyle and working conditions, and it has been shown that poor mental health in women is associated with divorce or widowhood [49]. In this study, working more than seven hours a day was a determinant factor on mental health problems, which was consistent with other studies that had shown that long working hours could have a negative impact on employees' mental health and that excessive workloads could increase workers' fatigue, which in turn could lead to anxiety and depression <sup>[50]</sup>.

In China, there are many problems in identifying people with mental health problems due to uneven and imperfect levels of medical development across regions. Some studies have shown that in mainland China, general practitioners, surgeons and primary health care workers often have little or no mental health training, which prevents them from providing basic mental health services [51]. Non-mental health professionals in general hospitals learn about mental illness on their own, rather than learning about it during their formal education<sup>[52]</sup>. 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. This study has several strengths. First, to our

knowledge, this is the first model to develop and assess the likelihood of mental health problems in a
group of factory workers and miners. Secondly, the nomogram in this study includes demographic
information, job burnout, occupational stress, chronic illnesses, and also occupational exposure factors
that are closely related to factory workers and miners, allowing for a more accurate assessment of the
risk of morbidity among them, as well as providing a methodological reference for other related studies.

# 5. Limitations

This study also has several limitations. Firstly, 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.

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

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

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

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

468 Ethics approval The study was approved by the ethics committee of Urumqi Center for Disease Control469 and Prevention (20181123)

471 Data availability statement Data are available on reasonable request. The data used in this study are
 472 available from the corresponding authors on reasonable request.

58 474 **References** 

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

1		
2 3		
3 4	476	performance/docs/glossary.html.
5	477	[2] Wang Y, Liu X, Qiu J, Wang H, Liu D, Zhao Z, Song M, Song Q, Wang X, Zhou Y, Wang W.
6	478	Association between ideal cardiovascular health metrics and suboptimal health status in Chinese
7	479	population. Sci Rep 2017;7:14975.
8 9	480	[3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A. No health without mental
9 10	481	health. Lancet. 2007;370:859–77.
11	482	[4] Adjaye-Gbewonyo K, Avendano M, Subramanian S.V, Kawachi I. Income inequality and depressive
12	483	symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. <i>Health</i>
13 14	484	<i>Place</i> 2016;42:37–46.
14	485	[5] Vos T, Barber R.M, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A,
16	486	Degenhardt L, Dicker D, <i>et al.</i> Global, regional, and national incidence, prevalence, and years lived
17	480 487	with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A
18		
19 20	488	systematic analysis for the Global Burden of Disease Study 2013. <i>Lancet</i> 2015;386:743–800.
20	489	[6] Huang Z, Li T, Xu M. Are there heterogeneous impacts of national income on mental health? <i>Int J</i>
22	490	Environ Res Public Health 2020;17:7530.
23	491	[7] Asadullah M.N, Xiao S, Yeoh E. Subjective well-being in China, 2005–2010: The role of relative
24 25	492	income, gender, and location. China Econ Rev 2018;48:83-101.
25 26	493	[8] Butterworth P, Rodgers B, Windsor T.D. Financial hardship, socio-economic position and depression:
27	494	Results from the PATH Through Life Survey. Soc Sci Med. 2009;69:229-237.
28	495	[9] Zhang X, Zhang X, Chen X. Happiness in the Air: How Does a Dirty Sky Affect Mental Health and
29	496	Subjective Well-being? J Environ Econ Manag 2017;85:81–94.
30 31	497	[10] Wahlbeck K. Public mental health: The time is ripe for translation of evidence into practice. <i>World</i>
32	498	<i>Psychiatry</i> . 2015;14:36–42.
33	499	[11] Luciano AE, Drake RE, Bond GR, Becker DR, Carpenter-Song E, Lord S, Swarbrick P and Swanson
34	500	SJ. IPS Supported employment: a review. <i>J Vocat Rehabil</i> 2014;40:1–13.
35 36	501	[12] Jia Z, <i>et al.</i> Exposure to ambient air particles increases the risk of mental disorder: findings from a
37	502	natural experiment in Beijing. Int J Environ Res Public Health 2018;15:160.
38	502	[13] Szyszkowicz M, Willey J. B, Grafstein E, Rowe B. H, Colman I. Air pollution and emergency
39		
40	504	department visits for suicide attempts in vancouver, Canada. <i>Environ Health Insights</i> 2010;4:79–86.
41 42	505	[14] Michael J. S. International classification of sleep disorders-third edition : highlights and
43	506	modifications. <i>Chest</i> 2014;146:1387–1394.
44	507	[15] AbuRuz ME, Al-Dweik G. Depressive symptoms and complications early after acute myocardial
45	508	infarction: gender differences. Open Nurs J 2018;12:205–214.
46 47	509	[16] Wurm W, Vogel K, Holl A, et al. Depression-burnout overlap in physicians. PLoS One
48	510	2016;11:e0149913.
49	511	[17] Porru F, Robroek S J W, Bültmann U, et al. Mental health among university students: The
50	512	associations of effort-reward imbalance and overcommitment with psychological distress. J Affect
51 52	513	Disord 2021;282:953-961.
52 53	514	[18] Johnson AK, Blackstone SR, Skelly A, Simmons W. The relationship between depression, anxiety,
54	515	and burnout among physician assistant students: a multi-institutional study. <i>Health Professions Edu</i>
55	516	2020;6:420–427.
56	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.
57 58	518	Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during
59	519	the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. <i>EClinical Med</i>
60	517	the COVID-17 outbreak in wunan, Clinia, a large-searce cross-sectional study. EClinical mea

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

Page 19 of 26

128-136.

2020;19:3-14.

2015;181:295-303.

diabetes. Diabetes Care 2018;41:446-452.

# **BMJ** Open

behavioural problems among children and adolescents during the covid-19 pandemic: A national

multicentre study: Risk factors of childhood psychosocial problems. J Affect Disord, 2021, 294:

[39] Occupational Stress and Employees Complete Mental Health: A Cross-Cultural Empirical Study

[40] Knapp M and Wong G. Economics and mental health: the current scenario. World Psychiatry

[41] Benach, J, Vives, A, Amable, M, Vanroelen, C, Tarafa, G, & Muntaner, C. Precarious employment:

[42] Bakian A. V. et al. Acute air pollution exposure and risk of suicide completion. Am J Epidemiol

[43] Barnard KD, Skinner TC, Peveler R. The prevalence of co-morbid depression in adults with Type

[44] Gilsanz P, Karter AJ, Beeri MS, Quesenberry CP, Whitmer RA. The bidirectional association

[45] Lu Y, Zhang Z, Yan H, et al. Effects of occupational hazards on job stress and mental health of

factory workers and miners: A propensity score analysis. BioMed Res Int 2020;2020:1754897.

between depression and severe hypoglycemic and hyperglycemic events in type 1

1 diabetes: systematic literature review. Diabet Med 2006;23:445-448.

understanding an emerging social determinant of health. Annu Rev Public Health 2014;35:229-53.

1 2	
3	564
4 5	565
6	566
7	567
8 9	568
, 10	569
11	570
12 13	571
13 14	572
15	573
16 17	574
17	575
19	576
20 21	577
21	578
23	579
24 25	580
25 26	581
27	582
28	583
29 30	584
31	585
32	586
33 34	587
35	588
36	589
37 38	590
39	591
40	592
41 42	593
42 43	594
44	595
45	596
46 47	597
48	598
49	599
50 51	600
52	601
53	602
54 55	603
56	604
57	605
58 59	606
177	(07

60

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. 599 600 **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

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 weas plotted based on the lambda series in Figure 1(A), and 12

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

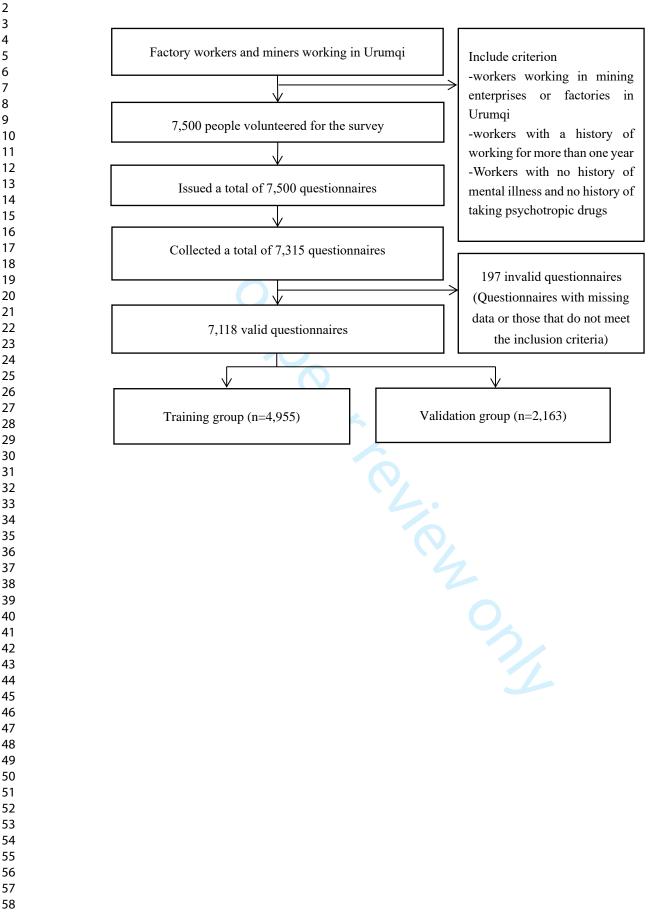
# Fig.3. The forest plot of the OR of the selected feature.

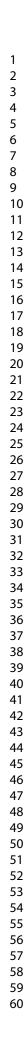
Fig.4. Developed mental health problems incidence risk nomogram. The mental health problems incidence risk
nomogram was developed in the array, with education, professional title, age, CMBI, ERI, asbestos dust,
hypertension, diabetes, working hours per day, working years, marital status, and work schedule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

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

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

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





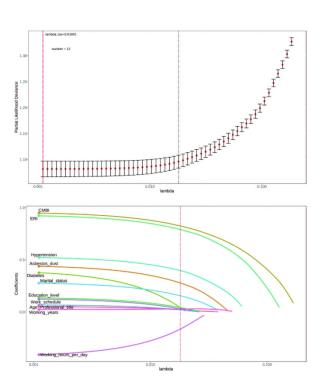
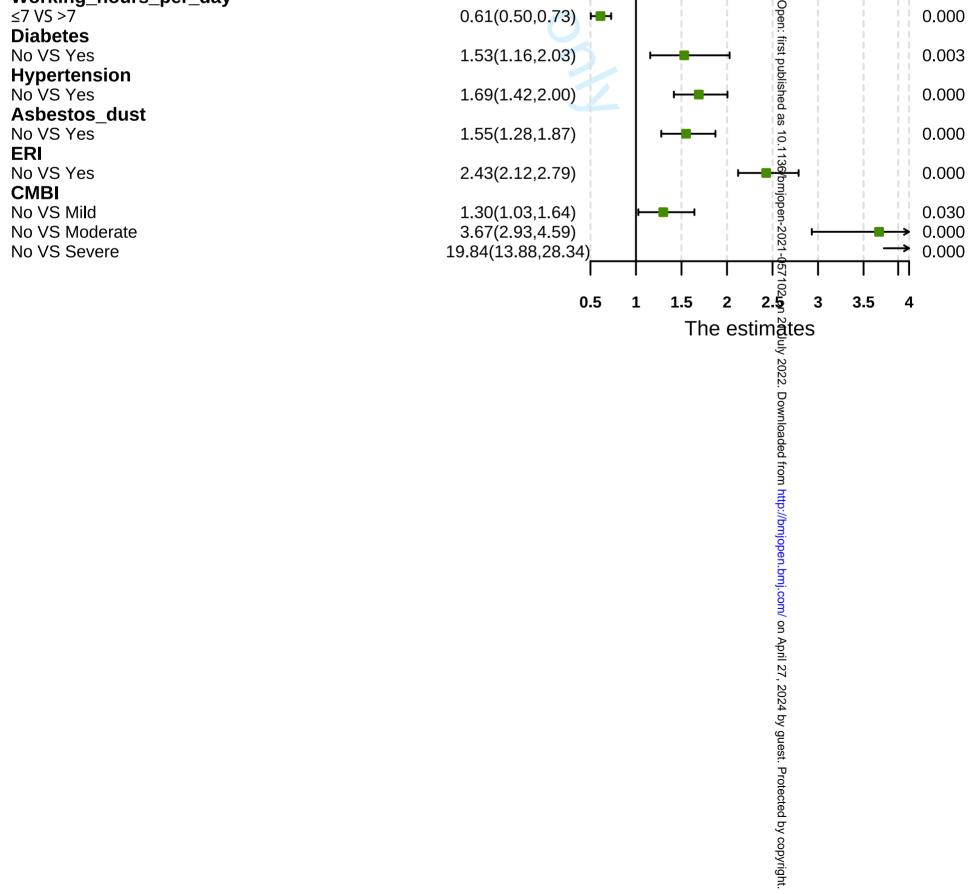
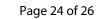


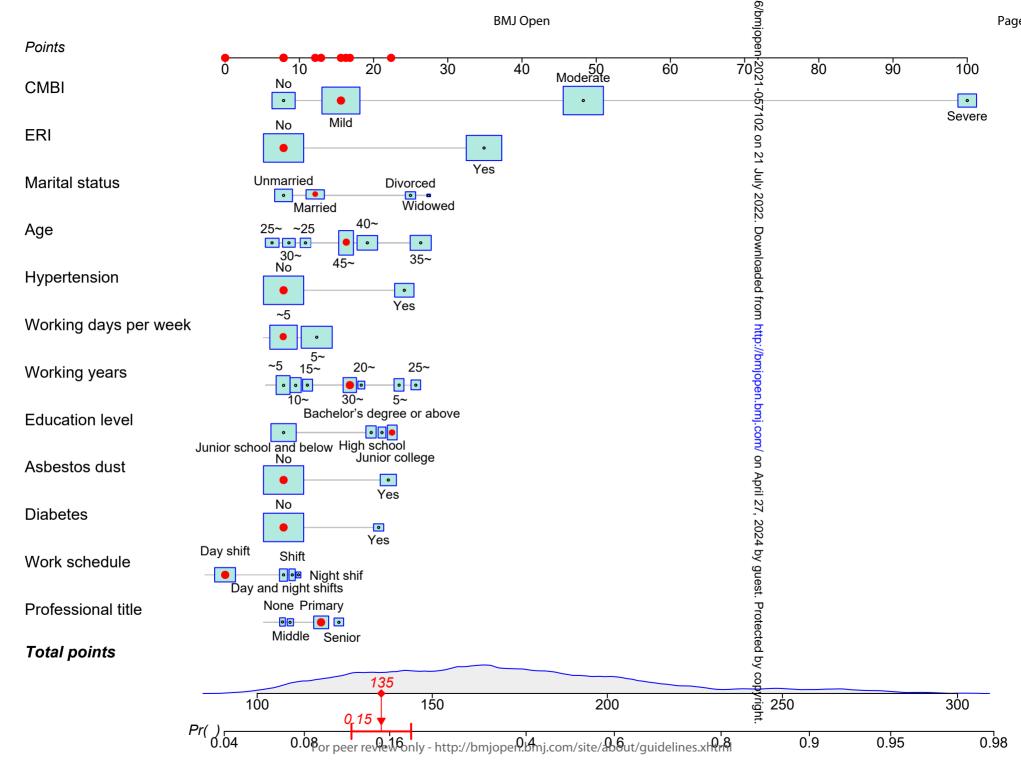
Fig.2. Feature selection using the LASSO binary logistic regression model. (A) Feature selection for the LASSO binary logistic regression model. The partial likelihood deviation (binomial deviation) curve was plotted against lambda by validating the optimal parameter lambda in the LASSO model. Dotted 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 weas 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)

12 13 14			
15			
16			
17 18			
19			
20			
21			
22			
23 24			
25			
26			
27			
28			
29 30			
31			
32			
33			
<sup>34</sup> <sub>25</sub> Variable	OR(Cl95%)		<b>P-Value</b>
<sup>33</sup> <sub>36</sub> Education_level			
<sup>37</sup> Junior school and below VS High school	1.41(1.10,1.81)		0.006
<sup>38</sup> Junior school and below VS Junior college	1.56(1.24,1.95)		0.000
<sup>39</sup> Junior school and below VS Bachelor's degree or above	1.46(1.13,1.87)		0.003
<sup>40</sup> <sub>41</sub> Professional_title			
42 None VS Primary	1.16(0.97,1.39)		0.114
<sup>43</sup> None VS Middle	1.05(0.87,1.26)		0.604
44 None VS Senior	1.30(1.06,1.61)		0.014
<sup>45</sup> <sub>46</sub> Work_schedule	1100(1100,1101)		0.011
$_{47}^{46}$ Day and night shifts VS Day shift	0.69(0.55,0.85)		0.001
48 Day and night shifts VS Night shif	1.01(0.68, 1.49)		0.965
<sup>49</sup> Day and night shifts VS Shift	1.01(0.81, 1.27)		0.915
<sup>50</sup> Marital_status	1.01(0.01,1.27)		0.010
<sup>51</sup> <sub>52</sub> Unmarried VS Married	1.18(0.91,1.52)		0.206
53 Unmarried VS Divorced	1.73(1.20,2.51)		0.004
<sup>54</sup> Unmarried VS Widowed	1.99(0.85,4.64)		→ 0.113
<sup>55</sup> <sub>56</sub> Age	1.00(0.00,4.04)		0.110
<sup>56</sup> <sub>57</sub> ~25 VS 25~	0.98(0.66,1.47)		0.934
57 25 VS 25 58~25 VS 30~	0.98(0.64,1.50)		0.929
59~25 VS 35~	1.76(1.13,2.74)		0.012
<sup>60</sup> ~25 VS 40~	1.39(0.88,2.21)		0.156
~25 VS 45~	1.26(0.81,1.95)		0.308
Working_years	1.20(0.01,1.95)		0.308
	1.55(1.18,2.05)		0.002
~5 VS 10~			0.002
~5 VS 10~ ~5 VS 15~	1.06(0.78,1.43) 1.06(0.72,1.56)		0.714
~5 VS 15~ ~5 VS 20~	. ,		0.780
~5 VS 20~ ~5 VS 25~	1.33(0.95,1.88) 1.61(1.15,2.25)		0.101
~5 VS 25~ ~5 VS 30~	1.61(1.15,2.25)		
	1.22(0.89,1.68)		0.216
<b>Working_hours_per_day</b> ≤7 VS >7	0 61 (0 50 0 72)		0.000
≤/ vs >/ Diabetes	0.61(0.50,0.73)		0.000
	1 52(1 16 2 02)	s l	0 002

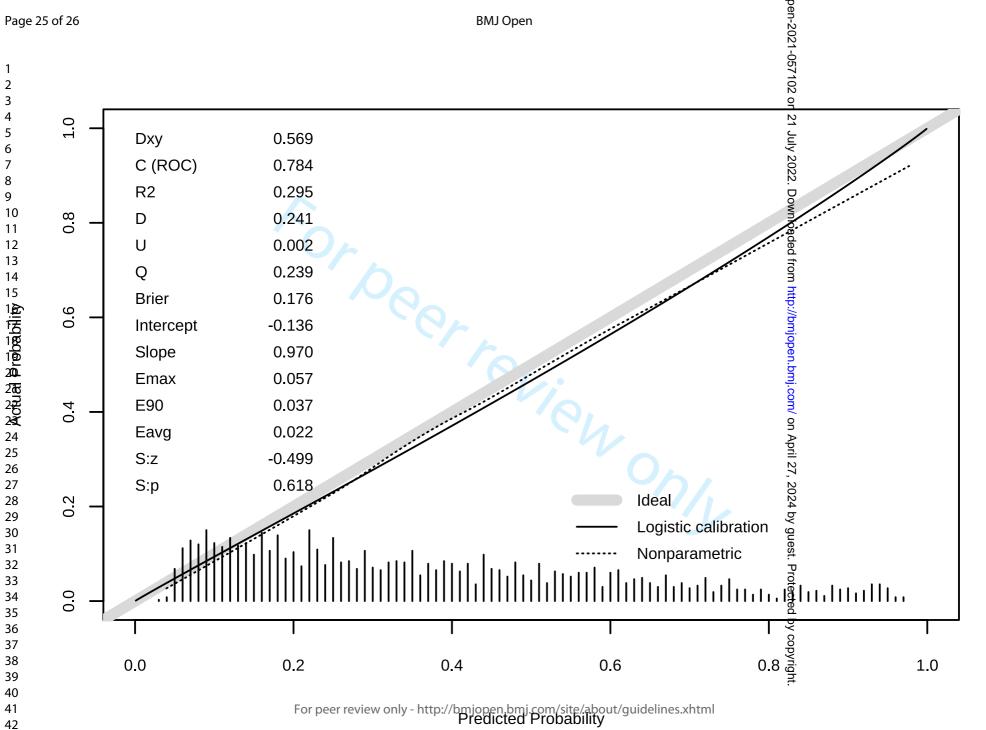


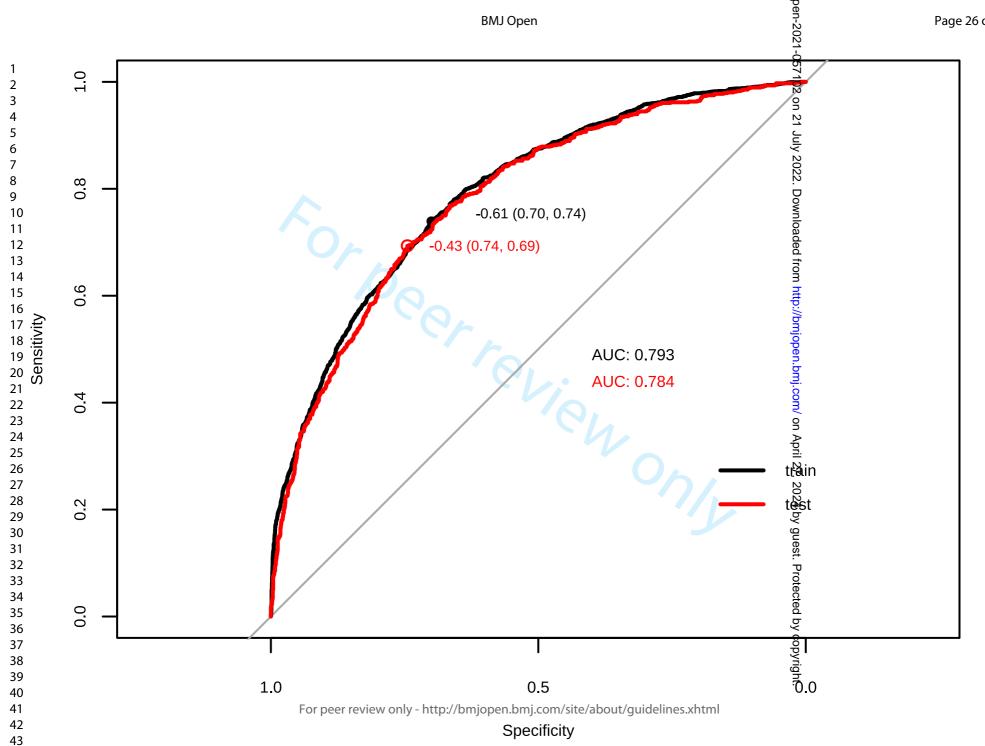




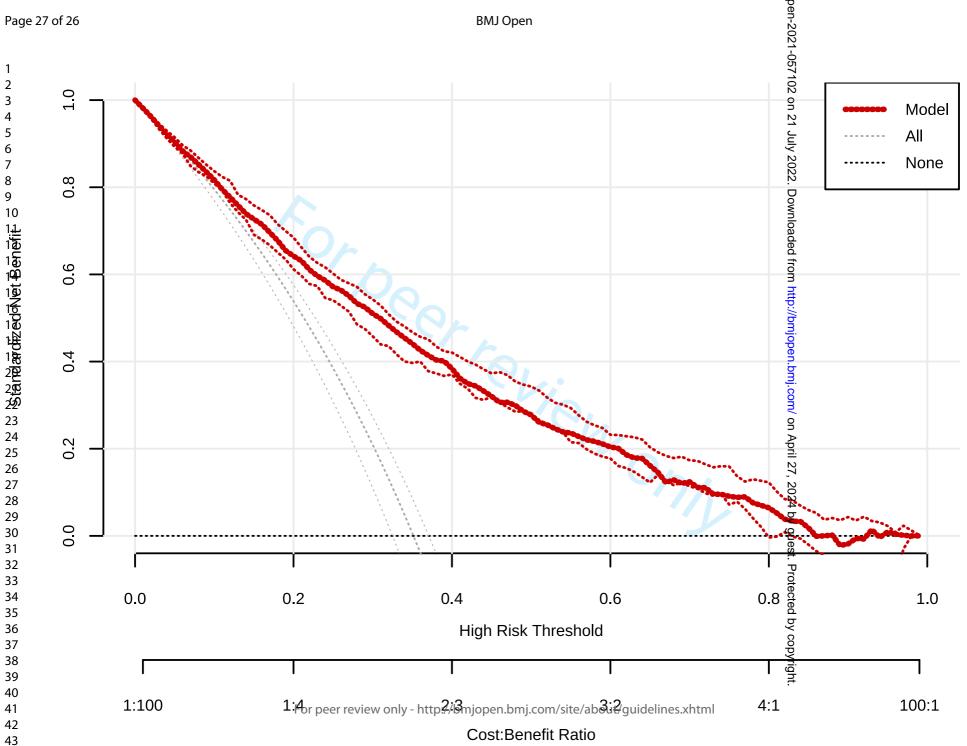


A A





. .



A A