BMJ Open

Comparison of predicted psychological distress among workers between artificial intelligence and psychiatrists: a cross-sectional study in Tsukuba Science City, Japan

Shotaro Doki [ID] ,[1] Shinichiro Sasahara,[1] Daisuke Hori,[1] Yuichi Oi,[1] Tsukasa Takahashi,[2] Nagisa Shiraki,[2] Yu Ikeda,[2] Tomohiko Ikeda,[2] Yo Arai,[2] Kei Muroi,[2] Ichiyo Matsuzaki[1,3]

Check for updates

[1]Faculty of medicine, University of Tsukuba, Tsukuba, Japan
[2]Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan
[3]International Institute for Integrative Sleep Medicine, University of Tsukuba, Tsukuba, Japan

Correspondence to
Dr Shinichiro Sasahara;
s-sshara@md.tsukuba.ac.jp

ABSTRACT

Objectives Psychological distress is a worldwide problem and a serious problem that needs to be addressed in the field of occupational health. This study aimed to use artificial intelligence (AI) to predict psychological distress among workers using sociodemographic, lifestyle and sleep factors, not subjective information such as mood and emotion, and to examine the performance of the AI models through a comparison with psychiatrists.

Design Cross-sectional study.

Setting We conducted a survey on psychological distress and living conditions among workers. An AI model for predicting psychological distress was created and then the results were compared in terms of accuracy with predictions made by psychiatrists.

Participants An AI model of the neural network and six psychiatrists.

Primary outcome The accuracies of the AI model and psychiatrists for predicting psychological distress.

Methods In total, data from 7251 workers were analysed to predict moderate and severe psychological distress. An AI model of the neural network was created and accuracy, sensitivity and specificity were calculated. Six psychiatrists used the same data as the AI model to predict psychological distress and conduct a comparison with the AI model.

Results The accuracies of the AI model and psychiatrists for predicting moderate psychological distress were 65.2% and 64.4%, respectively, showing no significant difference. The accuracies of the AI model and psychiatrists for predicting severe psychological distress were 89.9% and 85.5%, respectively, indicating that the AI model had significantly higher accuracy.

Conclusions A machine learning model was successfully developed to screen workers with depressed mood. The explanatory variables used for the predictions did not directly ask about mood. Therefore, this newly developed model appears to be able to predict psychological distress among workers easily, regardless of their subjective views.

Strengths and limitations of this study

► An artificial intelligence (AI) model for predicting psychological distress among workers was created using sociodemographic, lifestyle and sleep factors.

► The explanatory variables used for the predictions did not directly ask about mood; therefore, the developed model could predict psychological distress, regardless of their subjective views.

► The results of AI model were compared in terms of accuracy with predictions made by psychiatrists.

► It was difficult to generalise our results to wider populations because of the large number of researchers involved in this study.

► The model used in this study cannot indicate which parameters/features are most important for predicting the outcome of interest.

INTRODUCTION

An estimated 3.5 million people around the world are thought to suffer from depression. In terms of disability-adjusted life years, depression was ranked third in 2004 and is predicted to be ranked first in 2030.[1] If further measures are not taken to prevent depression, up to US$16 trillion in economic losses between 2010 and 2030 are expected.[2] In Japan, economic losses due to depression in the labour market were reported to total 2 trillion JPY (10 billion GBP) in 2005.[3] Therefore, preventive measures for depression are needed for both individual health and the economy.

In recent years, new technologies using artificial intelligence (AI) have been applied to the field of clinical medicine. As a result, AI has achieved higher accuracy than clinicians for the pathological diagnosis of lymph node metastasis in breast cancer.[4]

In the dermatological clinical setting, using dermoscopic images, AI has achieved higher accuracy than most dermatologists for the diagnosis of melanoma.[5] Another AI system that distinguishes malignant tumours from skin lesions has been shown to outperform board-certified dermatologists.[6]

AI has been applied to not only physical diseases but also mental illnesses. Applying AI in psychiatry has also been shown to be useful for diagnosis and treatment.[7] Using machine learning, body part movements have been applied to diagnose depressive disorder with 97.2% accuracy,[8] and, using a three-dimensional camera (Kinect; Microsoft, Redmond, Washington, USA), head movements, facial expressions and voice have been applied to diagnose moderate depressive disorder with 89.7% accuracy.[9] AI analysis of gait data has also been used to identify mood disorders.[10] Using a wearable device, a neural network and random forest model was found to detect stress levels with 92% accuracy and major depressive disorder with 87% accuracy.[11 12] Using neural language processing, the tendency towards depression was predicted by posts to the internet with 0.67 precision.[13] Electroencephalogram signals have also been reported to be capable of detecting major depressive disorder with high accuracy using a three-dimensional convolutional neural network.[14] However, from a neuroimaging point of view, the development of technology to detect mood disorders using AI is still in the exploratory stage.[15] In current neural network models, not only biological but also sociodemographic data are used for the prediction of depression.[16]

To our knowledge, no studies have been conducted to screen depression/psychological distress in a large population using machine learning. For earlier interventions to treat depression/psychological distress, a more rapid and effective screening tool is needed.

AI has been used to monitor safety and health among workers in the field of occupational health[17] and to assess sick-listed employees returning to work based on individual, occupational, support and psychological factors.[18]

Given this background, the present study aimed to use AI to predict psychological distress among workers using sociodemographic, lifestyle and sleep factors, not subjective information such as mood and emotion, and to examine the predictive accuracy of AI through a comparison with psychiatrists.

METHODS
Study design
In this cross-sectional study, workers took part in an online survey about psychological distress and living conditions. This study was conducted at Tsukuba Science City, Japan, from February through March 2017. An AI model for predicting psychological distress was created and the results were then compared in terms of accuracy with predictions made by psychiatrists.

Participants
The study participants were mainly researchers and office workers. Tsukuba Science City is located northeast of Tokyo in Ibaraki Prefecture. It has a population of about 230 000 and features two universities and a number of public and private research institutes. As of 2017, the Tsukuba Science City Network consisted of 89 organisations around the city, including research institutions, universities, educational foundations, local governments and private companies. A total of 53 organisations, mainly comprising research and educational institutions, with a total of 19 481 workers, agreed to participate.

Items
The Kessler Screening Scale for Psychological Distress (K6) was used to assess psychological distress. Previous studies have shown that the K6 has two cut-off points: a score of ≥5 points indicates moderate psychological distress and a score of ≥13 points indicates severe psychological distress.[19] According to Kessler *et al*, when the cut-off is set at 12/13 points, the sensitivity and specificity are 0.36 and 0.96, respectively, with an area under the curve value of 0.86 for detecting serious mental illness, as diagnosed by the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV).[20]

The explanatory variables used in this study were age, sex, marital status, employment status, type of organisation, type of job, education level, living with family members, household income, exercise habit, smoking status and sleep status. Sleep status was measured using the eight-item Athens Insomnia Scale (AIS) sleep questionnaire. Since this study aimed to detect psychological distress from sociodemographic, lifestyle and sleep factors, not by asking directly about mood, two of the AIS items asking about subjective mental status—'well-being during the day' and 'functioning capacity during the day'—were excluded from the analysis.

Neural network protocol
Two neural network models were created to determine the two depressive states: moderate psychological distress (a K6 Score of ≥5 points) and severe psychological distress (a K6 Score of ≥13 points). The model was repeated 10 times (models 1–10). Because the nature of psychological distress has been reported to differ between men and women,[21] we created separate models for each sex.

In total, 100 cases (the hold-out test set) were randomly selected for validation through a comparison between the trained AI models and psychiatrists. Machine learning for the model creation was performed using 80% of the total data as training data and 20% as validation data. A neural network was used to perform the training. The loss function was set to SparseCategoricalCrossentropy, which calculates the cross-entropy loss between the predictions and labels, the optimisation function was set to Adam, an optimisation tool combining RMSProp and momentum, and the activation function was set to the rectified linear unit activation function. The number of layers was set

to 3, the number of nodes in the middle layer was set to 100 and the number of epochs was set to 20. These conditions were determined with reference to previous studies.[4 10] Python V.3.7 (https://www.python.org/downloads/), Keras V.2.2.4 (https://keras.io/) and TensorFlow V.1.13.1 (https://www.tensorflow.org/) were used for the analysis. A schematic of the neural network model used in this study is shown in online supplemental material appendix A.

Psychiatrists

The number of psychiatrists to be used in the comparison was determined with reference to a previous systematic review comparing AI and clinicians that reported median numbers of five clinicians and four experts.[22] Six psychiatrists (A–F) participated in the study, five of whom were experts (two government-designated psychiatrists and three occupational health consultants). Two had more than 10 years of experience, three had more than 5 years of experience and one had less than 5 years of experience. The psychiatrists used the same information as that was used to validate the AI to predict moderate and severe psychological distress. The psychiatrists selected one of dichotomous variable of moderate/severe psychological distress or healthy for each individual participant. The psychiatrists independently evaluated the data from the 100 cases hold-out test set to calculate the accuracy, sensitivity and specificity.

Statistical analysis

The accuracy, sensitivity and specificity for each of the 10 AI models and psychiatrists were calculated using the hold-out test set. The results from the AI and psychiatrists were compared using analysis of variance (ANOVA).

Patient and public involvement statement

Neither patients nor the public were involved in this study.

RESULTS

Of the 19 481 potential participants, responses were received from 7255 (response rate: 37.2%). Overall, 4 respondents with missing data for age were excluded, leaving 7251 respondents (4574 (63.1%) males, 2677 (36.9%) females; mean age 44.3 years) for inclusion in the analysis. The participants' characteristics are shown in table 1. The majority (41.1%) was in teaching and research occupations, 47.0% exercised at least once a week and 10.8% were smokers. The median annual household income was 6–8 million JPY. The mean (SD) K6 Score was 5.6 (5.1), and the numbers of respondents classified as having moderate psychological distress (K6 Score ≥5) and severe psychological distress (K6 Score ≥13) were 3560 (49.1%) and 721 (9.9%), respectively. The mean K6 Score was significantly higher among women than among men ($p < 0.001$).

Machine learning was used to create 10 trained models with 100 validated cases each (table 2 and 3). For

Table 1 Characteristics of the participants in this study (N=7251)

	N (%)	Mean (SD)
Sex		
Male	4574 (63.1)	
Female	2677 (36.9)	
Age, years		44.3 (10.6)
Marital status		
Unmarried	1838 (25.4)	
Married	5134 (70.8)	
Divorced	226 (3.1)	
Bereaved	53 (0.7)	
Employment status		
Full time	5702 (78.6)	
Part-time	1378 (19.0)	
Temporary staff	171 (2.4)	
Type of organisation		
Public official	810 (11.2)	
Incorporated administrative agency	3065 (42.3)	
Private company	1306 (18.0)	
Others	2070 (28.5)	
Type of job		
Researcher/academic	2980 (41.1)	
Clerk/administration	2539 (35.0)	
Technician/engineer	1614 (22.2)	
Other	118 (1.6)	
Position		
Manager	1529 (21.1)	
Non-managerial	5722 (78.9)	
Living with family members		
Yes	5549 (76.5)	
No	1702 (23.5)	
Exercise habit		
Three times or more per week	1259 (17.4)	
One to two times per week	2146 (29.6)	
Less than once per week	3846 (53.0)	
Smoking status		
Smoker	782 (10.8)	
Non-smoker	6469 (89.2)	
Household income, JPY		
<2 million	199 (2.7)	
2–4 million	957 (13.2)	
4–6 million	1221 (16.8)	
6–8 million	1392 (19.2)	
8–10 million	1398 (19.3)	
10–12 million	1007 (13.9)	

Continued

Table 1 Continued

	N (%)	Mean (SD)
12–14 million	500 (6.9)	
>14 million	577 (8.0)	
K6 Score		5.6 (5.1)
Male		5.3 (5.0)
Female		6.0 (5.1)
<5 points	3691 (50.9)	
≥5 points	3560 (49.1)	
<13 points	6530 (90.1)	
≥13 points	721 (9.9)	

K6, Kessler Screening Scale for Psychological Distress.

moderate psychological distress, the accuracy, sensitivity, specificity and positive and negative predictive values were 65.2%, 64.6%, 65.9%, 65.9% and 65.1%, respectively. For severe psychological distress, the accuracy, sensitivity, specificity and positive and negative target rates were 89.9%, 17.5%, 96.2%, 29.7% and 93.1%, respectively. The results for the six psychiatrists are shown in table 2 and table 3. For moderate psychological distress, the accuracy, sensitivity and specificity were 64.4%, 51.4% and 78.0%, respectively. Among the government-designated psychiatrists, the accuracy, mean sensitivity and specificity were 62.0%, 46.4% and 78.2%, respectively. For severe psychological distress, the accuracy, sensitivity and specificity were 85.8%, 25.0% and 91.1%, respectively. Among the government-designated psychiatrists, the accuracy, mean sensitivity and specificity were 85.7%, 20.8% and 91.3%, respectively.

The results of the ANOVA conducted to compare the accuracy of the trained AI model with that of the psychiatrists revealed no significant difference for moderate psychological distress (p=0.263). On the other hand, the trained AI model showed significantly higher accuracy than did the psychiatrists for severe psychological distress (p=0.001) (figures 1 and 2).

DISCUSSION
In this study, we successfully created a machine learning model to screen workers with depressed mood. The accuracy was 65.2% for moderate psychological distress and 89.9% for severe psychological distress, which was higher than that for the psychiatrists. The explanatory variables used for the predictions did not directly ask about mood. Therefore, the developed model could be easily applied to the workers, regardless of their subjective views.

This model appears to be appropriate as a screening tool for psychological distress because it consists of questions that do not directly ask about mood and emotion. Many depression/psychological distress questionnaires ask about subjective mood, but due to stigma, many depressed patients remain undiagnosed or untreated.[23] Disclosing mental health status in the workplace has become an extraordinarily complex issue.[24] Although predicting psychological distress using only assessment items that do not directly ask about mood and emotion can be difficult, it can help prevent the manipulation of questionnaire responses due to stigma towards mental healthcare and inappropriate responses to questions asking about subjective mood. The model developed in this study appears useful as a screening tool for psychological distress because it avoids this issue.

However, due to its low sensitivity and specificity, our model still needs to be improved for practical use. The sensitivity and specificity of the K6 for the prediction of severe psychological distress have been reported to be 0.36 and 0.96, respectively,[20] and those of the Depression Anxiety Stress Scales, which detect depression in workers,

Table 2 Comparison of predictive performance between our artificial intelligence (AI) model and psychiatrists for moderate psychological distress

	AI (K6≥5)				Psychiatrists (K6≥5)		
Model no.	Accuracy (%)	Sensitivity	Specificity	Subject	Accuracy (%)	Sensitivity	Specificity
1	63	0.529	0.735	A	62	0.373	0.878
2	63	0.608	0.653	B	61	0.412	0.816
3	66	0.627	0.694	C	63	0.608	0.653
4	72	0.725	0.714	D	67	0.431	0.918
5	64	0.569	0.714	E	69	0.745	0.633
6	65	0.725	0.571	F	61	0.608	0.612
7	63	0.659	0.607				
8	65	0.765	0.531				
9	63	0.608	0.653				
10	68	0.647	0.714				
Mean (SD)	65.2 (2.750)	0.646 (0.071)	0.659 (0.066)		64.4 (3.078)	0.514 (0.133)	0.780 (0.123)

K6, Kessler Screening Scale for Psychological Distress.

Table 3 Comparison of predictive performance between our artificial intelligence (AI) model and psychiatrists for severe psychological distress

| Model no. | AI (K6≥13) | | | Psychiatrists (K6≥13) | | | |
	Accuracy (%)	Sensitivity	Specificity	Subject	Accuracy (%)	Sensitivity	Specificity
1	91	0.250	0.967	A	86	0.125	0.924
2	90	0.125	0.967	B	83	0.250	0.880
3	90	0.125	0.967	C	88	0.250	0.935
4	88	0.125	0.946	D	88	0.000	0.957
5	92	0.250	0.978	E	84	0.625	0.859
6	92	0.250	0.978	F	86	0.375	0.902
7	90	0.250	0.957				
8	88	0.125	0.946				
9	89	0.125	0.957				
10	89	0.125	0.957				
Mean (SD)	89.9 (1.375)	0.175 (0.061)	0.962 (0.011)		85.8 (1.863)	0.250 (0.197)	0.911 (0.033)

K6, Kessler Screening Scale for Psychological Distress.

have been reported to be 0.91 and 0.46, respectively.[25] The sensitivity and specificity of the model in this study were 0.18 and 0.96, respectively, and although the specificity was high, due to its low sensitivity, the accuracy of the model still needs to be improved before it can be applied as a screening tool. To improve the accuracy of our model, the amount of training data needs to be increased, ensemble learning has to be performed in conjunction with other classifiers (such as decision trees and support vector machines) and facial expression and speech analysis need to be combined.

Because of its high specificity for the prediction of severe psychological distress and because it performed better than did the psychiatrists, our trained AI model

Figure 1 Comparison between our artificial intelligence (AI) model and psychiatrists of the percentage of correct answers for moderate psychological distress (K6≥5). The results of analysis of variance revealed no significant difference between the two groups (p=0.263). Whiskers extend to the maximum value and minimum value. Data points exceeding a distance of 1.5 times the IQR below the first quartile or above the third quartile are shown as outliers. K6, Kessler Screening Scale for Psychological Distress.
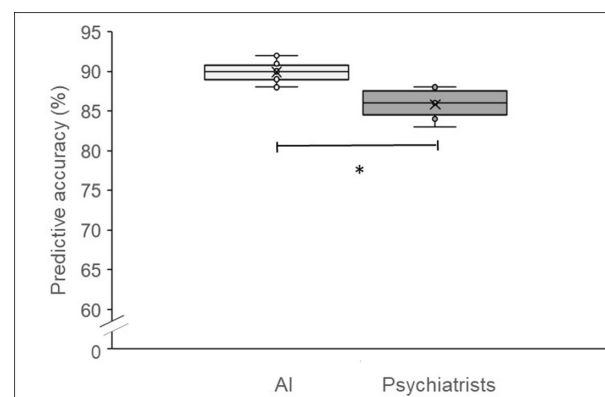
Figure 2 Comparison between our artificial intelligence (AI) model and psychiatrists of the percentage of correct answers for severe psychological distress (K6≥13). The results of analysis of variance revealed that the AI model achieved a significantly higher proportion of correct answers (p=0.001). Whiskers extend to the maximum value and minimum value. K6, Kessler Screening Scale for Psychological Distress.

shows potential for clinical applications. At this stage, it is difficult to use AI models as screening tools, but their high specificity suggests that they can aid clinicians in prediction. It is difficult for occupational physicians to predict psychological distress because they are not specialised in psychiatry.[26] The use of tools such as guidelines by general practitioners has been shown to improve the diagnosis and management of workers mental health.[27] It could be also useful for occupational physicians to interview and check for psychological distress in staff who receive a positive prediction from the AI model, and this could be useful in improving mental health in the workplace.

As a future prospect, we believe that the AI model developed in this study could be used to monitor the risk of psychological distress among workers in the field of occupational health. If smart devices could be used to obtain

information on heart rate, activity levels, sleep patterns and working hours automatically, the AI model would be able to provide a prediction without requiring workers to enter data manually. Our AI model could also be effective when combined with the use of a smartphone, as an association has been reported between depression and the frequency of smartphone use.[28] In addition, the use of smartphone apps has been shown to help prevent depression in workers.[29] In the future, we would like to obtain time-series data and create a more accurate model using a recurrent neural network.

A weak point of this study is that it is difficult to diagnose depression accurately based on assessment items that do not ask directly about mood and emotion. The diagnostic criteria for depression according to the eleventh revision of the International Classification of Diseases and the DSM-5 are based on subjective items, such as depressed mood and loss of pleasure and interest. Based on the results of this study, our AI model could be expected to be used as an adjunct to diagnosis rather than for a definitive diagnosis and in the future, as a screening tool for large populations.

This study did have some limitations. First, we used an online questionnaire to predict psychological distress as opposed to interviews with psychiatrists. Interviews with psychiatrists are necessary to make a more accurate diagnosis, but the cost of conducting such interviews would be extremely high because of the large amount of data. As the K6 was used to predict psychological distress, we could not exclude other illnesses that present with depression, such as bipolar disorder and schizophrenia. In addition, it was difficult to generalise our results to wider populations because of the large number of researchers involved in this study, as researchers tend to burn out and feel more qualitative and quantitative pressure compared with other types of occupations.[30 31] Finally, the AI black-box problem remains.[32] The model used in this study cannot indicate which parameters/features are the most important for predicting the outcome of interest.

ORCID iD
Shotaro Doki http://orcid.org/0000-0003-4026-9058

REFERENCES

1 World Health Organization. The global burden of disease : 2004 update, 2008. Available: https://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf?ua [Accessed 8 Oct 2020].

2 The Lancet commission. Lancet commission: inaction on mental health crisis will cost world $16 trillion by 2030, 2018. Available: https://globalmentalhealthcommission.org/press_release/lancet-commission-inaction-on-mental-health-crisis-will-cost-world-16-trillion-by-2030/ [Accessed 8 Oct 2020].

3 Sado M, Yamauchi K, Kawakami N, *et al*. Cost of depression among adults in Japan in 2005. *Psychiatry Clin Neurosci* 2011;65:442–50.

4 Ehteshami Bejnordi B, Veta M, Johannes van Diest P, *et al*. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 2017;318:2199–210.

5 Haenssle HA, Fink C, Schneiderbauer R, *et al*. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol* 2018;29:1836–42.

6 Fujisawa Y, Otomo Y, Ogata Y, *et al*. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *Br J Dermatol* 2019;180:373–81.

7 Cuzzolin F, Morelli A, Cîrstea B, *et al*. Knowing me, knowing you: theory of mind in AI. *Psychol Med* 2020;50:1057–61.

8 Joshi J, Dhall A, Goecke R. *Relative body parts movement for automatic depression analysis. humaine association conference on affective computing and intelligent interaction*. Geneva, 2013: 492–7.

9 Scherer S, Stratou G, Morency LP. Audiovisual behavior descriptors for depression assessment. in proceedings of the 15th ACM on International Conference on multimodal interaction 2013:135–40.

10 Deligianni F, Guo Y, Yang G-Z. From emotions to mood disorders: a survey on gait analysis methodology. *IEEE J Biomed Health Inform* 2019;23:2302–16.

11 Patlar Akbulut F, Ikitimur B, Akan A. Wearable sensor-based evaluation of psychosocial stress in patients with metabolic syndrome. *Artif Intell Med* 2020;104:101824.

12 Cho C-H, Lee T, Kim M-G, *et al*. Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: prospective observational cohort study. *J Med Internet Res* 2019;21:e11029.

13 Tung C, Lu W. Analyzing depression tendency of web posts using an event-driven depression tendency warning model. *Artif Intell Med* 2016;66:53–62.

14 Khan DM, Yahya N, Kamel N, *et al*. Automated diagnosis of major depressive disorder using brain effective connectivity and 3D convolutional neural network. *IEEE Access* 2021;9:8835–46.

15 Gao S, Calhoun VD, Sui J. Machine learning in major depression: from classification to treatment outcome prediction. *CNS Neurosci Ther* 2018;24:1037–52.

16 Sau A, Bhakta I. Artificial neural network (ANN) model to predict depression among geriatric population at a slum in Kolkata, India. *J Clin Diagn Res* 2017;11:VC01–4.

17 Uchida M, Noshita K, Koyama H. [Application of deep learning for the occupational health of VDT workers]. *Sangyo Eiseigaku Zasshi* 2019;61:256–60.

18 Na K-S, Kim E. A machine Learning-Based predictive model of return to work after sick leave. *J Occup Environ Med* 2019;61:e191–9.

19 Sakurai K, Nishi A, Kondo K, *et al*. Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan. *Psychiatry Clin Neurosci* 2011;65:434–41.

20 Kessler RC, Chiu WT, Demler O, *et al*. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National comorbidity survey replication. *Arch Gen Psychiatry* 2005;62:617–27.

21 Duijts SFA, Kant IJ, Landeweerd JA, *et al*. Prediction of sickness absence: development of a screening instrument. *Occup Environ Med* 2006;63:564–9.

22 Nagendran M, Chen Y, Lovejoy CA, *et al*. Artificial intelligence versus clinicians: systematic review of design, reporting Standards, and claims of deep learning studies. *BMJ* 2020;368:m689.

23 Smith K. Mental health: a world of depression. *Nature* 2014;515:180–1.

24 Stratton E, Choi I, Calvo R, *et al*. Web-Based decision aid tool for disclosure of a mental health condition in the workplace: a randomised controlled trial. *Occup Environ Med* 2019;76:595–602.

25 Nieuwenhuijsen K, de Boer AGEM, Verbeek JHAM, *et al*. The depression anxiety stress scales (DASS): detecting anxiety disorder and depression in employees absent from work because of mental health problems. *Occup Environ Med* 2003;60 Suppl 1:77i–82.

26 Seelig MD, Katon W. Gaps in depression care: why primary care physicians should hone their depression screening, diagnosis, and management skills. *J Occup Environ Med* 2008;50:451–8.

27 Mazza D, Chakraborty SP, Brijnath B, *et al*. Diagnosing and managing work-related mental health conditions in general practice: new Australian clinical practice guidelines. *Med J Aust* 2019;211:76–81.

28 Demirci K, Akgönül M, Akpinar A. Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. *J Behav Addict* 2015;4:85–92.

29 Deady M, Glozier N, Calvo R, *et al*. Preventing depression using a smartphone APP: a randomized controlled trial. *Psychol Med* 2020:1–10.

30 Padilla MA, Thompson JN. Burning out faculty at doctoral research universities. *Stress Health* 2016;32:551–8.

31 Kageyama T, Matsuzaki I, Morita N, *et al*. Mental health of scientific researchers I. characteristics of job stress among scientific researchers working at a research Park in Japan. *Int Arch Occup Environ Health* 2001;74:199–205.

32 Castelvecchi D. Can we open the black box of AI? *Nature* 2016;538:20–3.