

Using machine learning approach to predict suicide risk from a longitudinal cohort data in Japan.

Po-Han Chou^{1,2,9*}, Shao-Cheng Wang^{3,4,5,9}, Chi-Shin Wu^{6,7,*}, Masaya Ito⁸, Masaru Horikoshi⁸

1. Department of Psychiatry, China Medical University Hsinchu Hospital, China Medical University, Hsinchu, Taiwan.
2. Department of Psychiatry, China Medical University Hospital, China Medical University, Taichung, Taiwan.
3. Department of Psychiatry, Taoyuan General Hospital, Ministry of Health and Welfare, Taoyuan, Taiwan,
4. Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States.
5. Department of Medical Laboratory Science and Biotechnology, Chung Hwa University of Medical Technology, Tainan, Taiwan.
6. National Center for Geriatrics and Welfare Research, National Health Research Institutes, Miaoli, Taiwan.
7. Department of Psychiatry, National Taiwan University Hospital, Yunlin branch, Yunlin, Taiwan.
8. National Center for Cognitive-Behavior Therapy and Research, National Center of Neurology and Psychiatry
9. Equal contribution as first authors

Background

Several prognostic models of suicide risk have been published; however, few have been implemented in Japan using longitudinal cohort data.

Aims & Objectives

The aim of this study was to identify suicide risk factors for suicidal ideation in the Japanese population and to develop a machine-learning model to predict suicide risk in Japan.

Methods

Data was obtained from Wave1 Time 1 (November 2016) and Time 2 (March 2017) of the National Survey for Stress and Health in Japan, were incorporated into a suicide risk prediction machine-learning model, trained using 65 items related to trauma and stress. The study included 3090 and 2163 survey respondents >18 years old at Time 1 and Time 2, respectively. The mean (SD) age was 44.9 (10.9) years at Time 1 and 46.0 (10.7) years at Time 2. We analyzed the participants with increased suicide risk at Time 2 survey. Model performance, including the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity, were also analyzed.

Results

The model showed a good performance (AUC = 0.830, 95% confidence interval = 0.795–0.866). Overall, the model achieved an accuracy of 78.8%, sensitivity of 75.4%, specificity of 80.4%, positive predictive value of 63.4%, and negative predictive value of 87.9%. The most important risk factor for suicide risk was the participants' Suicidal Ideation Attributes Scale score, followed by the Sheehan Disability Scale score, Patient Health Questionnaire-9 scores, Cross-Cutting Symptom Measure (CCSM-suicidal ideation domain, Dissociation Experience Scale score, history of self-harm, Generalized Anxiety Disorder-7 score, Post-Traumatic Stress Disorder check list-5 score, CCSM-dissociation domain, and Impact of Event Scale-Revised scores at Time 1.

Discussion & Conclusion

This prognostic study suggests the ability to identify patients at a high risk of suicide using an online survey method. In addition to confirming several well-known risk factors of suicide, new risk measures related to trauma and trauma-related experiences were also identified, which may help guide future clinical assessments and early intervention approaches.