

Dynamic Abbreviation of MMPI-I Chinese Version Based on the Machine Learning Predictions

Rui Chi^{1#}, Wentian Dong^{1#}, Chuan Shi¹, Yong He¹, Lin Cui³, Junyue Wang³, Jiaqi Li³, Xiaolu Zhou³, Bin Yu³, Jun Ji^{2,3*}

¹Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing, China

²College of Computer Science and Technology, Qingdao University, Qingdao, China

³Beijing Wanling Pangu Science and Technology Ltd., Beijing, China

Keywords: MMPI-I Chinese Version, Machine Learning, Abbreviated Scale, Psychological Examination

Abstract

MMPI was first introduced in the early 20th century and was extensively used by clinicians all over the world. Nevertheless, its application and popularization have been hindered in China since it is utterly cumbersome for patients to complete due to the time consumption of this assessment. In order to serve Chinese patients with accuracy and precision, MMPI-I form Chinese Version was introduced and translated by Dr. Song from Chinese Academy of Sciences in 1979 and then widely used in China. The goal of this study was to obtain an abbreviated version of MMPI-I Chinese Version without altering its functionalities and validities. In this study, the original MMPI-I was abbreviated by applying specific scales that could be predicted based on machine learning methods according to ten clinical scales. The sample contains a total amount of 10,679 subjects that were compiled by Peking University Sixth Hospital. All the subjects were fitted into 5,110 models specified by each gender, and all models were trained and validated with 10-fold cross-validation. By comparing the performances of each gender applied in different models, the least complicated and most credible abbreviated form of MMPI-I was obtained. By setting the thresholds of T-scores to 30 and 70 respectively, 17.5% and 24.1% items could be reduced with sensitivity and specificity greater than 75% for both males and females. In this case, the abbreviated MMPI-I can effectively improve the quality of mental assessments by notably decreasing the time consumption of the test.

Introduction

With the rapid development of society, the stress and pressure of human beings have leaped and are now considered a matter of continuing concern. Thus, personalities and mental status assessments are gradually assimilated into various aspects of our daily life including professional achievements, job searches, and most importantly, healthcare [1]. In the meantime, the applications of machine learning algorithms on psychological topics, such as bipolar disorder diagnosis [2], schizophrenia diagnosis [3], autism diagnosis [4] etc., have become a significant field of research. The MMPI was originally developed by the clinical psychologist S. R.

Hathaway and the neuropsychiatry scholar J. C. McKinley in 1937 at the University of Minnesota [5] to quantitatively diagnose mental health disorders and differentiate people with distinct personality traits. The original version by McKinley and Hathaway contains a total of 566 true/false questions, including 16 duplicated factors for the credibility of answers. Compared with SCL-90, which, the Chinese norms were revised in our recent studies [6], MMPI-I measures similar distinct dimensions of mental health but contains a unique evaluation of results validity [7]. As a result, MMPI-I has become the most popular personality assessment for its high reliability and strong validity. As it has been unanimously accepted internationally, it was also translated into simplified Chinese and revised with Chinese norms [8].

However, numerous aspects of this assessment have been considered challenging by people, even for those with magnificent patience. The development of an abbreviated version of MMPI-I Chinese Version has never been stopped for decades [9], and plenty of abbreviated form have been tested [10], whereas none of them retain the structure and function of MMPI-I entirely, specifically targeting the Chinese population. Even some of the most notable abbreviation versions such as MMPI-168 [11] and the Psychological Health Inventory (PHI) [12] did not quite obtain the needs for it to deploy effectively among Chinese patients.

In this study, in order to compile a shortened clinical diagnostic version of MMPI targeting the Chinese population and realize the dynamic abbreviation function for diagnosis practice, we researched the first 399 questions that were widely adopted for clinical diagnosis, including 10 clinical scales and 3 validity scales from the MMPI-I Chinese Version. Then we applied Machine Learning techniques [13] on these questions to construct 1,022 score prediction models by providing different combinations of items for each gender, based on the data analysis of 10,679 psychological test records from Peking University Sixth Hospital. In this study, the criteria for the item selection of the abbreviated version were to keep as few questions as possible on the premise of ensuring prediction performance.

Materials and Methods

Data Sample

A sample of 10,679 MMPI-I Chinese Version test data was selected and provided by the Peking University Sixth Hospital. The research data were collected mainly from 2020.6 to 2021.8 via Xinjiang app, adopted by Peking University Sixth Hospital for auxiliary clinical usage. The model consists of 4,601 male participants aging from 11.9 to 89.7, and 6,078 female participants with an age range between 11.8-year-old and 82.8-year-old (S2 Table 1). The data consisted no protected health information and all data were masked for research use.

All selected subjects had completed the entire assessment of the first 399 MMPI-I items for clinical diagnosis as required [14]. Distinct proper subsets of the 399 questions formed ten clinical scales and three validity scales. The clinical scales include "hypochondriasis" Hy, "depression" D, "hysteria" Hs, "psychopathic deviate" Pd, "masculinity-femininity" Mf, "paranoia" Pa, "psychasthenia" Pt, "schizophrenia" Sc, "hypomania" Ma, and "social introversion" Si, while "lie" L, "infrequency or fake bad" F, and "defensiveness" K are the three validity scales. With these collected answers to MMPI-I, clinicians can compute the scores for the ten clinical scales of each unit and determine whether the unit behaved positive or negative by converting original scores to T-scores in order to compare them to specific thresholds. All the extremely low or high T-scores are considered positive. The principle of this abbreviation was to accurately predict some of the clinical scales by items in all other scales, which are also questions that remained in the abbreviated version of MMPI-I. Therefore, all ten clinical scores of any test participant could be computed by requiring the participant to complete much fewer questions.

Abbreviation Procedure

In this study, the MMPI-I Chinese Version was abbreviated following the procedures illustrated in S1 Fig 1. First, we listed all possible combinations of clinical scales that could be selected as parameters of a prediction model. Each combination corresponds to a complementary set of scales that the chosen parameters could potentially predict. After that, we randomly split data into ten folds for cross-validation. For each fold of data and each set of parameter scales, linear discriminant analysis models for both males and females were constructed to predict the scores for the complementary set of scales by questions in the parameter scales. Next, we computed the mean sensitivities and the mean specificities with T-score-thresholds at 30 and 70 among all ten models for each scale combination. Finally, the list of combinations was sorted first by the descending number of scales successfully predicted. The sums of mean sensitivity and mean specificity over 170%, and then by the ascending number of the questions in the parameter scales, the unsuccessfully predicted and the validity scales. The abbreviated version of MMPI-I Chinese Version consists of the items on the top of the sorted list.

Before the final decision of these procedures, we first attempted to sort the importance by information entropies and predict unimportant scales in order by important scales. We had also compared all prediction models that estimate the score of one scale by all nine others. However, since all these attempts had not provided us with any valid predictions for assessment abbreviation, we eventually selected the method of constructing models for scores of scales, combining with age, gender, and items of those in the complement as parameters.

Performance Evaluation

According to the Chinese user manual of MMPI [8], clinicians usually select the T-score of 70 as the threshold to identify positive personalities for apparent sickness or characteristics. In this case, we used 70 as

the threshold to generate the confusion matrices, plot 10-fold ROC curves, and calculate the sensitivities and specificities for our models [15].

This study would be considered as a successful prediction if the sum of the mean sensitivity and the mean specificity of the 10-fold regression models exceeded 170%.

Results

In this study, in order to effectively find the most valid models, we first computed the sum of mean specificity and mean sensitivity for each predicted scale in each model. Among all potential models, one model for males and one model for females performed the best, with sums of sensitivity and specificity over 170% for four scales in the male model and five scales in the female model [16].

Among all qualified models for male, the model predicting scores for the scales D, Hy, Pa, and Sc with age and questions in Hs, Pd, Mf, Pt, Ma, and Si as parameters requires the least number of items no matter if we include items from the validity scales or not. In this circumstance, there were 329 items in total in the abbreviated form, where 252 questions were in clinical scales.

The performances and complexities of these best-performed models for females were measured. Including the validity questions, the model with the minor complexity was to predict scores of Hs, D, Pd, Pt, and Sc by age and items of Hy, Mf, Pa, Ma, and Si [17]. This abbreviation form of MMPI-I Chinese Version necessarily requires 303 questions in total, among which 230 items are contained in the five clinical scales.

Inasmuch as the subjects are based on clinical data, the reliability of the data is irrefutable. In this case, the prediction results of models from samples with a scale score less than 30 could not be counted as valid results. Therefore, we only left with prediction results of models from samples with a scale score of more than 70. Using the selected model for males, the ROC curves of the 10-fold datasets prediction performances of each scale are demonstrated in S1 Fig 2 [18], and the ROC curves using the best-performed female model are as

displayed in S1 Fig 3 [19]. The mean sensitivity and mean specificity of 10-fold for each gender of each scale was presented in S2 Table 2.

Conclusion

To summarize, the abbreviated results in this study could be a valid alternative to the original MMPI-I Chinese Version. It has come to a result that any male participants seeking medical treatment at the Peking University Sixth Hospital would necessarily complete 329 items from Hs, Pd, Mf, Pt, Ma, Si, L, F, and K scales [15]. On the other hand, female patients need to answer 303 questions from Hy, Mf, Pa, Ma, Si, L, F, and K scales.

Acknowledgements

Declarations

This study was supported by the Beijing Wanling Pangu Technology LLC., the Peking University Sixth Hospital, and funded by Qingdao University.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not Applicable

Availability of data and material

Ethics committee of Peking University Sixth Hospital states that data is designed for internal research usage only and is not disclosed to any third parties.

Authors' contributions

RC, WD and CS contributed to the design and conception of this study. XZ and YH gathered related materials and organized raw data. LC and JW conducted statistical analysis. JL prepared all figures and tables. JJ and BY conducted overview and revision of the manuscript. All authors read and approved the final manuscript.

Abbreviations

MMPI, The Minnesota Multiphasic Personality Inventory

MMPI-I, The Minnesota Multiphasic Personality Inventory form Chinese Version

References

1. Achenbach TM, Ivanova MY, Rescorla LA, Turner LV, Althoff RR.
Internalizing/externalizing problems: review and recommendations for clinical and research applications. *J Am Acad Child Adolesc Psychiatry* (2016) 55:647-56. doi: 10.1016/j.jaac.2016.05.012.
2. Ma, Y., Ji, J., Huang, Y., Gao, H., Li, Z., Dong, W., Zhou, S., Zhu, Y., Dang, W., Zhou, T., Yu, H., Yu, B., Long, Y., Liu, L., Sachs, G., & Yu, X. (2019). Implementing machine learning in bipolar diagnosis in China. *Translational psychiatry*, 9(1), 305. <https://doi.org/10.1038/s41398-019-0638-8>
3. Liu, J, He, Y., Wang, J., Quan, W., Tian, J., Feng, C., Yu, H., Nan, C., & Ji, J. (2019). Application of convolutional neural networks in the diagnosis of schizophrenia. *Chinese Journal of Behavioral Medicine and Brain Science*, 28(7): 622-626. DOI: 10.3760/cma.j.issn.1674-6554.2019.07.009.
4. Abbas H, Garberson F, Glover E, et al. Machine learning approach for early detection of autism by combining questionnaire and home video screening[J]. *Journal of the American Medical Informatics Association*, 2017, 25(8):1000-1007.
5. Schiele, B. C.; Baker, A. B.; Hathaway, S. R. (1943). "The Minnesota multiphasic personality inventory." *Journal-Lancet* (63): 292–297. ISSN 0096-0233.
6. Dang, W., Xu, Y., Ji, J., Wang, K., Zhao, S., Yu, B., Liu, J., Feng, C., Yu, H., Wang, W., Yu, X., Dong, W., & Ma, Y. (2021). Study of the SCL-90 Scale and Changes in the Chinese Norms [Original Research]. *Frontiers in Psychiatry*, 11(1505). <https://doi.org/10.3389/fpsy.2020.524395>
7. Xun Gong.(2019). Meta Analysis of Scl-90 Survey Results of Mental Health of Chinese College Students in Recent Years. *International Conference on Advancement of the Theory and Practices in Education* (in Finland).
8. Song W. (1989). "The Minnesota multiphasic personality inventory user manual". *Institute of Psychology, Chinese Academy of Sciences* (in Chinese). Newmark, C. S., Cook, L., Clarke, M., & Faschingbauer, T. R. (1973). Application of Faschingbauer's abbreviated MMPI to psychiatric

inpatients. *Journal of Consulting and Clinical Psychology*, 41(3), 416–421.

<https://doi.org/10.1037/h0035200>

9. Overall, J. E., Higgins, W., & de Schweinitz, A. (1976). Comparison of differential diagnostic discrimination for abbreviated and standard MMPI. *Journal of Clinical Psychology*, 32(2), 237–245.
[https://doi.org/10.1002/1097-4679\(197604\)32:2<237::AID-JCLP2270320208>3.0.CO;2-9](https://doi.org/10.1002/1097-4679(197604)32:2<237::AID-JCLP2270320208>3.0.CO;2-9)
10. Overall, J. E., & Gomez-Mont, F. (1974). "The MMPI-168 for Psychiatric Screening". *Educational and Psychological Measurement*, 34(2): 315–319. <https://doi.org/10.1177/001316447403400211>
11. Song, W., Mo W. (1992). "The Compilation of Psychological Health Inventory (PHI)". *Psychological Science* (2): 36-40. ISSN: 1671-6981 (in Chinese).
12. Achenbach TM, Ivanova MY, Rescorla LA, Turner LV, Althoff RR. Internalizing/externalizing problems: review and recommendations for clinical and research applications. *J Am Acad Child Adolesc Psychiatry* (2016) 55:647-56. doi: 10.1016/j.jaac.2016.05.012
13. Svanum Soren. A handbook on cross-national MMPI research: By J.N. Butcher and P. Paneheri (Minneapolis: University of Minnesota Press, 1976)[J]. Pergamon, 1977, 1(2): 1-2.
14. Quinlan R J. Induction of Decision Trees[J]. *Machine Learning*, 1986, 1(1):81-106.
15. Charles E. Metz, Basic principles of ROC analysis, *Seminars in Nuclear Medicine*, Volume 8, Issue 4, 1978, Pages 283-298, ISSN 0001-2998, [https://doi.org/10.1016/S0001-2998\(78\)80014-2](https://doi.org/10.1016/S0001-2998(78)80014-2).
16. Haslam, S. Alexander, and Craig McGarty. *Research methods and statistics in psychology*. Sage, 2014.
17. Gendreau, P., Grant, B. A., Leipziger, M., & Collins, S. (1979). Norms and recidivism rates for the MMPI and selected experimental scales on a Canadian delinquent sample. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 11(1), 21–31.
<https://doi.org/10.1037/h0081569>
18. Friedman J H. Greedy function approximation: a gradient boosting machine[J]. *The Annals of Statistics*, 2001, 29(5):1189-1232. Fisher, R.A. The use of multiple measurements in taxonomic problems[J]. *Annals of Eugenics*, 1936, 7(2):179-188

19. Altman D G, Bland J M. Statistics Notes: Comparing several groups using analysis of variance BMJ 1996; 312 :1472 doi:10.1136/bmj.312.7044.1472

Supporting Information

S1 Figure 1. The flow diagram as an illustration of the abbreviation procedure

S1 Figure 2. The ROC curves of the 10-fold datasets prediction performances of each scale using the selected model for male

S1 Figure 3. The ROC curves of the 10-fold datasets prediction performances of each scale using the selected model for female

S2 Table 1. Sample description. Abbreviations: Q1, first quartile; Q3, third quartile; IQR, inter-quartile range.

S2 Table 2. Mean sensitivity and mean specificity of the best model on the 10-fold subsample. Abbreviation: Hy, hypochondriasis; D, depression; Hs, hysteria; Pd, psychopathic deviate; Pt, psychasthenia; Sc, schizophrenia; Ma, hypomania; Si, social introversion.