Efficient Prediction of Vitamin B Deficiencies via Machine-learning Using Routine Blood Test Results in Patients With Intense Psychiatric Episode

Hidetaka Tamune1)2)3)5)*, Jumpei Ukita3)4)5), Yu Hamamoto1)2), Hiroko Tanaka1)2), Kenji Narushima1), Naoki Yamamoto1)

1) Department of Neuropsychiatry, Tokyo Metropolitan Tama Medical Center, Tokyo, Japan
2) Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
3) Mental Health Research Course, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
4) Department of Physiology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
5) H. Tamune and JU contributed equally to this work

* Correspondence:
Dr. Hidetaka Tamune, tamune-tky@umin.ac.jp

Abstract: 339 words (within 350 words)
Main text: 2660 words + 4 Tables + 5 Figures + 27 references
Keywords: Machine Learning; Random Forest Classifier; Vitamin B Deficiency; Folic Acid; Early Diagnosis; Decision support techniques or decision making.

Vitamin deficiency predicted by machine-learning

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background: Vitamin B deficiency is common worldwide and may lead to psychiatric symptoms; however, vitamin B deficiency epidemiology in patients with intense psychiatric episode has rarely been examined. Moreover, vitamin deficiency testing is costly and time-consuming, which has hampered effectively ruling out vitamin deficiency-induced intense psychiatric symptoms. In this study, we aimed to clarify the epidemiology of these deficiencies and efficiently predict them using machine-learning models from patient characteristics and routine blood test results that can be obtained within one hour.

Methods: We reviewed 497 consecutive patients deemed to be at imminent risk of seriously harming themselves or others over 2 years in a single psychiatric tertiary-care center. Machine-learning models (k-nearest neighbors, logistic regression, support vector machine, and random forest) were trained to predict each deficiency from age, sex, and 29 routine blood test results gathered in the period from September 2015 to December 2016. The models were validated using a dataset collected from January 2017 through August 2017.

Results: We found that 112 (22.5%), 80 (16.1%), and 72 (14.5%) patients had vitamin B1, vitamin B12, and folate (vitamin B9) deficiency, respectively. Further, the machine-learning models were well generalized to predict deficiency in the future unseen data, especially using random forest; areas under the receiver operating characteristic curves for the validation dataset (i.e. the dataset not used for training the models) were 0.716, 0.599, and 0.796, respectively. The Gini importance of these vitamins provided further evidence of a relationship between these vitamins and the complete blood count, while also indicating a hitherto rarely considered, potential association between these vitamins and alkaline phosphatase (ALP) or thyroid stimulating hormone (TSH).
Discussion: This study demonstrates that machine-learning can efficiently predict some vitamin deficiencies in patients with active psychiatric symptoms, based on the largest cohort to date with intense psychiatric episode. The prediction method may expedite risk stratification and clinical decision-making regarding whether replacement therapy should be prescribed. Further research includes validating its external generalizability in other clinical situations and clarify whether interventions based on this method could improve patient care and cost-effectiveness.
1. Introduction

Vitamin B deficiency is common worldwide and may lead to psychiatric symptoms\textsuperscript{1–4}. For example, meta-analyses have shown that patients with schizophrenia or first-episode psychosis have lower folate (vitamin B\textsubscript{9}) levels than their healthy counterparts\textsuperscript{4,5}. Moreover, vitamin therapy can effectively alleviate symptoms in a subgroup of patients with schizophrenia\textsuperscript{3,6–8}. However, the epidemiology of vitamin B deficiency in patients with active mental symptoms requiring immediate hospitalization has rarely been examined.

In a psychiatric emergency, psychiatrists should promptly distinguish treatable patients with altered mental status due to a physical disease from patients with an authentic mental disorder (international statistical classification of diseases and related health problems-10, ICD-10 code: F2-9). However, vitamin deficiency testing is very costly (around 60 dollars for each measurement of vitamin B\textsubscript{1} (vitB\textsubscript{1}), vitamin B\textsubscript{12} (vitB\textsubscript{12}), or folate in the U.S.; 15–25 dollars for each test in Japan) and usually requires at least two days. Therefore, an efficient, cost-effective method of predicting vitamin B deficiency is needed.

Although several studies have applied machine-learning to the prediction of diagnosis or treatment outcomes\textsuperscript{9–11}, no study using machine-learning has focused on vitamin B deficiencies. We herein explore whether vitB\textsubscript{1}, vitB\textsubscript{12}, and folate deficiencies can be predicted using a machine-learning classifier from patient characteristics and routine blood test results obtained within one hour based on a large cohort of patients requiring urgent psychiatric hospitalization.
2. Methods

2.1. Medical chart review

We reviewed consecutive patients admitted to the Department of Neuropsychiatry at Tokyo Metropolitan Tama Medical Center, one of the biggest psychiatric tertiary-care centers in Japan, between September 2015 and August 2017 under the urgent involuntary hospitalization law, which requires the immediate psychiatric hospitalization of patients at imminent risk of seriously harming themselves or others. The necessity of hospitalization was judged by designated mental health specialists. There were no exclusion criteria. The patient characteristics, ICD-10 codes, and laboratory data were gathered retrospectively.

Since the reference ranges for vitB_{1}, vitB_{12}, and folate are 70–180 nmol/L (30–77 ng/mL), 180–914 ng/L, and > 4.0 μg/L, respectively^{12}, a deficiency of the nutrients was defined as < 30 ng/mL, < 180 ng/L, and < 4.0 μg/L, respectively, unless otherwise stated. The odds ratios of each deficiency in each ICD-10 code were calculated assuming binomial distribution.

2.2. Classifiers and statistics

We compared four types of standard machine-learning classifiers; k-nearest neighbors, logistic regression, support vector machine, and random forest. Each type of classifier was trained to predict the deficiency of each substance from age, sex, and 29 routine blood variables (described in the Result section with values). For developing the models, any missing values were replaced using the mean. The classifiers were trained using the dataset populated in the period from September 2015 to December 2016 (the “Training set”). First, except for logistic regression, we optimized the hyperparameters of the classifier by selecting the best combination of hyperparameters that maximized the ”5-fold cross validation” accuracy, among many combinations within appropriate ranges. The cross-validation accuracy was computed as follows; in one session, the
classifiers were trained using 80% of the training set and evaluated on the withheld 20% of the training set. This session was performed five times so that every data would be withheld once. The accuracies were finally averaged across sessions to yield the cross-validation accuracy. By incorporating this process, the classifiers were generalized to unseen data (Graphical method is shown in Figure 1).

Using the optimized hyperparameters, the classifiers were then validated using data collected from January 2017 through August 2017 (the “Validation set”). We report the classification performance on the validation set in the results section unless otherwise stated.

We quantified the sensitivity, specificity, and accuracy (defined as the average of the sensitivity and the specificity on the optimal operating point) using receiver operating characteristic curves (ROCs). We also quantified the 95% confidence interval of the area under the ROCs (AUCs) and accuracy using 1000-times bootstrapping.

When investigating the Gini importance and the partial dependency, we retrained the classifiers using all datasets. All data analyses were performed using Python (2.7.10) with the Scikit-learn package (0.19.0) and R (3.4.2) with the edarf package (1.1.1) and pROC package (1.15.3).

2.3. Robustness verification

We verified the robustness of the prediction performances by three independent approaches. First, we compared the following two prediction performances; random forest classifiers trained and validated using the dataset from the F2 population, and random forest classifiers trained and validated using the dataset from the non-F2 population.

Second, we compared the prediction performances of several random forest classifiers trained and validated using the dataset where different cut-off values were used to define the vitamin deficiency. We chose other two cut-off values for each
vitamin based on previous reports\textsuperscript{14–16}, as well as pre-defined cut-off values (see also 2.1.).

Third, we trained and validated other random forest classifiers where the dataset was split in a different way. Here, the training set consisted of data between the 31\textsuperscript{st} of January, 2016 and August 2017 and the validation set consisted of data between September 2015 and the 31\textsuperscript{st} of January, 2016 so that the sample sizes of the training and validation sets were equal to those in the original split.

2.4. Subsampling analysis

We also examined the relationship between the dataset size and the generalization performance\textsuperscript{17}. In this analysis, we trained the random forest classifiers using X\% of the training set (X = 30, 35, 40, …, 95, and 100), and validated them using the validation set. The hyperparameters were identical to those used in the previous section. To remove sampling bias, this procedure was repeated 100 times for each value of X, where the training dataset was sampled randomly for each repetition. This results in obtaining 100 AUC scores for each X and for each vitamin. We plotted the AUC scores (averaged across the 100 repetition) versus X for each vitamin. Then the curve was fit with the following saturating function using Levenberg-Marquardt algorithm implemented as “curve_fit” function in the Scipy package (0.19.0).

\[
Y = \frac{a \cdot X}{X + b} + 0.5
\]

where Y is the AUC score, and a and b are the parameters to fit. Note that

\[
Y \to a + 0.5 \text{ as } X \to \infty, \text{ and } Y = 0.5 \text{ as } X = 0.
\]

2.5. Ethical considerations

Informed consent was obtained from participants using an opt-out form on the website. The study protocol was approved by the Research Ethics Committee, Tokyo
Metropolitan Tama Medical Center (Approval number: 28-8). The study complied with the Declaration of Helsinki and the STROBE statement.
3. Results

3.1. Eligible patients

During the 2-year study period, 497 consecutive patients (496 were Asian) were enrolled. The mean age (standard deviation, SD) was 42.3 (±15.4) years, and 228 patients (45.9%) were women. F2 (Schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders) was diagnosed in over 60% of the patients. The ICD-10 codes of the patients and the number of deficiencies at several cut-off values for vitB₁, vitB₁₂, and folate are shown in Table 1. According to the predefined cut-off values¹², 112 (22.5%), 80 (16.1%), and 72 (14.5%) patients exhibited a deficiency of vitB₁ (<30 ng/mL), vitB₁₂ (<180 ng/L), and folate (<4.0 μg/L), respectively. Vitamin B deficiencies in sub-groups are shown in Table 2. A summary of the full dataset is shown in Table 3. Detailed information (sub-datasets) is shown in Supplementary Table 1, 2, and 3 online. Histograms of vitB₁, vitB₁₂, and folate values are shown in Figure 2 A-C.

3.2. Prediction via machine-learning using routine blood test results

Machine-learning classifiers were trained to predict the deficiency of each substance from patient characteristics and routine blood test results. The classifiers were trained using the dataset gathered in the period from September 2015 to December 2016 (the “Training set”, n = 373), which was then validated from January 2017 through August 2017 (the “Validation set”, n = 124). By splitting the whole dataset in this way, the ratio of the training and validation sample size was 3:1, a commonly used ratio in machine-learning analyses.

AUCs for the validation set for each classifier are summarized in Table 4. Although the performance of the classifiers was similar except for the k-nearest neighbors, random forest yielded the highest AUC on average. Therefore, we focused on random forest in the following analysis.
The AUCs of the random forest classifiers were 0.716, 0.599, and 0.796, for 

vitB₁, vitB₁₂, and folate, respectively (Figure 2 D-F and Table 4). With some operative 

points on the ROC, the sensitivity, specificity, and accuracy for the validation set were 

calculated (Table 4. See also Supplementary Table 4 for training set and 

Supplementary Table 5 for different operating points). The 95% confidence interval 

(CI) of the AUC and accuracy was quantified using 1000-times bootstrapping. For 

random forest classifiers, the 95% CI of each value did not include 0.5, except for the 

AUC of vitB₁₂. 

Figure 3 shows the Gini importance (a–c) and partial dependency plots (d–f) 

for the eight most important variables for each substance. The results provided further 

evidence of a relationship between the vitamin B levels and complete blood count while 

also indicating the hitherto rarely considered, potential association between these 

vitamins and alkaline phosphatase (ALP) or thyroid stimulating hormone (TSH). 

3.3. Robustness verification 

We verified the robustness of the results by three independent means. First, we 

asked if the prediction performance was influenced by the ICD-10 categories. When the 

prediction performances were compared between the random forest classifiers trained 

using the dataset from the F2 population and the classifiers trained using the dataset 

from the other population, the AUC was not statistically different (DeLong’s test), 

except in the case of vitB₁ (see Supplementary Table 6). 

Second, we used different cut-off values to define the deficiency¹⁴–¹⁶. Although 

the AUC for the validation set, shown in Supplementary Table 7, tended to be higher 

when strict cut-off values were used, the obtained AUCs were not statistically 

significant (p > 0.05, DeLong’s test with Bonferroni correction). 

Third, we investigated if the prediction performance was influenced by the way 

the dataset was split into the training and validation set. Here, we trained and evaluated
random forest classifiers using a dataset split in a reversed way (see Methods for
details); The AUCs for the validation set were 0.771, 0.621, and 0.745 for vitB₁, vitB₁₂,
and folate, respectively; none were statistically different from the AUC trained using the
original setting (DeLong’s test), further demonstrating the robustness of the
performance.

3.4. Subsampling analysis

To estimate the number needed to saturate the performance, we examined the
relationship between the generalizability and the sample size. We randomly sampled
X% of the training set, trained random forest classifiers using the dataset, and evaluated
the generalization performance by AUCs using the validation set (X = 30, 35, 40, …, 95,
and 100; see Methods for details). As shown in Figure 4, the relationships between
AUC and the training size for vitB₁ and vitB₁₂ were almost saturated, whereas that for
folate is not saturated. To quantitatively understand this, we fitted each curve using a
saturating function formulated in equation (1) (see Methods for details). The fitted
parameters of equation (1) were as follows; for vitB₁, a = 0.186 and b = 0.074; for
vitB₁₂, a = 0.099 and b = 0.156; and for folate, a = 0.291 and b = 0.123. By using these
parameter values and extrapolating the curve, we then computed how many additional
samples are necessary to reach almost maximum performances. To reach 99% of the
maximum performance (i.e. \( Y = (a + 0.5) \times 0.99 \) in equation (1)), the training dataset to
be collected was 92.5%, 143%, and 341% of the training size in this study for vitB₁,
vitB₁₂, and folate, respectively. These quantitative analyses revealed that collecting
further similar datasets up to 1,000 patients (e.g. four years × hospitals with similar
scale as Tokyo Metropolitan Tama Medical Center) may increase and reproduce the
generalizability for folate, while the effect of collecting further dataset is expected to be
small for vitB₁ and vitB₁₂.
4. Discussion

4.1. Relevance of the present study

Based on the largest cohort to date of patients at imminent risk of seriously harming themselves or others, this study indicated that deficiency of certain vitamins can be predicted in an efficient manner via machine-learning using routine blood test results. The 29 routine blood variables are available at almost all hospitals/clinics and are necessary to rule out other comorbid physical problems. Given the large number of patients with vitamin B deficiencies, empirical therapy might be acceptable; however, risk stratification is preferred for personalized medicine and shared decision-making. The prediction method presented here may expedite clinical decision-making as to whether vitamins should be prescribed to a patient (Graphical abstract is shown in Figure 5).

Remarkably, the AUC of folate deficiency was 0.796. The robustness of folate prediction was also suggested by various independent methods and statistics. Folate has a potential to maintain neuronal integrity and is one of the homocysteine-reducing B-vitamins. Homocysteine may be linked to the etiology of schizophrenia, and vitamin B supplements have been reported to reduce psychiatric symptoms significantly in patients with schizophrenia. A recent meta-review has pointed out that the bioactivity of the supplement should be considered (e.g. methylfolate, which successfully crosses the blood-brain barrier, has been reported effective, whereas the effect of other forms of folate is equivocal). As our study does not present longitudinal clinical courses, an intervention effect of folate supplementation to the cohort based on our method remains to be clarified.

4.2. Biological mechanism prediction

To connect with biological knowledge, we compared four models with high interpretability in this study. Using the random forest classifiers, as shown in Figure 3,
we identified several items related to complete blood count as top hits. Notably, our classifier was blind to any biological knowledge, including the well-established association between anemia and vitamin B deficiency, including folate\textsuperscript{20}. The results provide further evidence of a relationship between vitamin B levels and the complete blood count and support the use of machine-learning to investigate novel, underlying biological mechanisms\textsuperscript{21}.

ALP and its metabolites indicate the vitamin B\textsubscript{6} status\textsuperscript{22}; low vitB\textsubscript{12} is potentially associated with low ALP\textsuperscript{23}. More generally, ALP may have a close and complicated relationship with the overall vitamin B group. Autoimmune disorders, especially thyroid disease, are commonly associated with pernicious anaemia\textsuperscript{24}, but there has been no established hypothesis regarding the causal relationships between thyroid disease and vitamin B deficiencies. The potential association between the levels of these vitamins and ALP or TSH awaits further study, both via investigations of populations and basic research\textsuperscript{25}.

### 4.3. Limitations

This study is subject to several limitations. First, the findings of this single-center retrospective study may have limited external generalizability, though internal generalizability was considered to the maximum extent. Second, the patients’ basic characteristics and long-term prognosis were not fully investigated due to administrative restrictions. Though there is similar involuntary treatment/admission in psychiatry worldwide, there is a gap between legislation and practice\textsuperscript{26}. Therefore, the extent to which this method can expedite clinical decision-making is unclear. Further, we did not investigate the relationship between serological values and the need for intervention. The lack of data for vitamin B deficiency in the Japanese general population hampered the comparison between the experimental cohort and their counterparts who lacked psychiatric symptoms. Establishing appropriate reference
values and an assessment method requires further investigation. Finally, we did not
assess the predictive value of other nutritional impairments, including vitamin B₆ and
homocysteine deficiency, which were previously shown to have a close link with
psychiatric symptoms³,⁵; however, our study provides fundamental data on nutritional
impairment based on the largest cohort of patients with intense psychiatric episode ever
assembled for this purpose and presents a potential framework for predicting nutritional
impairment using machine-learning.

4.4. Conclusion

The present report is, to the best of our knowledge, the first to demonstrate that
machine-learning can efficiently predict nutritional impairment. This study also
provided the possible application of machine-learning to investigate novel, underlying
biological mechanisms. Further research is needed to validate the external
generalizability of the findings in other clinical situations and clarify whether
interventions based on this method can improve patient care and cost-effectiveness.
5. Contribution to the Field Statement

Vitamin B deficiency is common worldwide and may lead to psychiatric symptoms; however, vitamin B deficiency epidemiology in patients with intense psychiatric symptoms has rarely been examined. Moreover, vitamin deficiency testing is costly and time-consuming. Based on the largest cohort to date of patients at imminent risk of seriously harming themselves or others, this study demonstrated that the deficiency of certain vitamins can be predicted in an efficient manner via machine-learning models from patient characteristics and routine blood test results obtained within one hour.

In detail, among the 497 patients investigated (over 60% was diagnosed with schizophrenia or related psychotic disorders), 22.5%, 16.1%, and 14.5% patients had a deficiency of vitamin B$_1$, B$_{12}$, and folate, respectively, by direct measurement. Further, the machine-learning models well generalized to predict the deficiency in unseen datasets; areas under the receiver operating characteristic curves (AUCs) for the validation dataset were 0.716, 0.599, and 0.796, respectively. The prediction method presented in this study may expedite risk stratification and clinical decision-making regarding whether replacement therapy should be prescribed. The results also provided further evidence for well-known relationships between these vitamins and the complete blood count and supported the application of machine-learning to investigate novel, underlying biological mechanisms.
6. Acknowledgements

We thank Mr. James Robert Valera for his assistance in editing this manuscript and all the staff for their care of the patients and their contributions to this study. This manuscript has been previously published as a preprint27 in medRxiv (https://doi.org/10.1101/19004317).

7. Author Contributions Statement

H. Tamune has full access to all data and takes responsibility for the integrity of the data. H. Tamune, JU, KN, and NY conceived the study. H. Tamune, YH, and H. Tanaka collected the data. JU performed the statistical analyses. H. Tamune and JU drafted the first version of the manuscript. All authors critically revised the manuscript for intellectual content and approved the final version.

8. Data Availability Statements

The source code will be available on https://github.com/ukky17/vitaminPrediction after acceptance. The datasets utilized in the current study are available from the corresponding author upon reasonable request.

9. Conflict of Interest Statement

The authors declare no conflict of interest, except for a scholarship grant awarded to JU from Takeda Science Foundation and Masayoshi Son Foundation.
References


Supporting Material List

- Supplementary Table 1 (related to Table 1). Divided patient distribution data (n = 497)
- Supplementary Table 2 (related to Table 2). Divided data of vitamin B deficiencies in sub-groups
- Supplementary Table 3 (related to Table 3). Divided dataset of age, sex, and 29 parameters
- Supplementary Table 4 (related to Table 4). Summary of sensitivity, specificity, and accuracy for the training set
- Supplementary Table 5 (related to Table 4). Sensitivities and specificities at other operating points
- Supplementary Table 6. Subgroup analyses
- Supplementary Table 7. AUC with different cut-off values
Legends

Figure 1: Graphical illustration of method of machine-learning

Figure 2: Histogram and ROC curves of each vitamin B value

(A-C) The histograms for vitamin B₁, vitamin B₁₂, and folate (vitamin B₉). Their medians (1ˢᵗ–3ʳᵈ quartile) are 35 (30–42) ng/mL, 285 (206–431) ng/L, and 7.2 (4.9–10.8) μg/L, respectively.

(D-F) ROC curves for vitamin B₁, vitamin B₁₂, and folate. Operating points used in Table 4 and Supplementary Table 5 are depicted in blue.

Abbreviations: Vit B₁, vitamin B₁; Vit B₁₂, vitamin B₁₂.

Figure 3: Gini importance and partial dependence plots of vitamin B deficiencies

The Gini importance (A-C) and partial dependency plots of the probability of deficiency (D-F) are shown for the eight most important variables for vitamin B₁, vitamin B₁₂, and folate (vitamin B₉). Combined with these, this machine-learning classifier without hypothesis also provided further evidence of a relationship between vitamin B levels and the complete blood count while also indicating a potential association between these vitamins and alkaline phosphatase (ALP) or thyroid-stimulating hormone (TSH).

Abbreviations: Vit B₁, vitamin B₁; Vit B₁₂, vitamin B₁₂; Hb, hemoglobin; Hct, hematocrit; WBC, white blood cell count; CK, creatine kinase; RDW.CV, red blood cell distribution width-coefficient variation; Plt, platelet; ALT, alanine transaminase; Lym, lymphocyte fraction; Cre, creatinine; Neu, neutrocyte fraction; γGTP, γ-glutamyltransferase; MCV, mean corpuscular volume; glu, plasma glucose.

Figure 4: Subsampling analysis

The AUC performances as a function of the dataset size is shown for each vitamin (mean ± SEM across 100 repetitions; see Methods for details).
Figure 5: Graphical abstract
Table 1. Patient distribution data (n = 497)

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>42.3 (15.4) years</td>
<td>Woman 228 (45.9%)</td>
<td>Asian 496</td>
</tr>
<tr>
<td></td>
<td>Man 269 (54.1%)</td>
<td>Others 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ICD-10 code</th>
<th>VitB1 [ng/mL]</th>
<th>VitB12 [ng/L]</th>
<th>Folate [μg/L]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0</td>
<td>&lt;20</td>
<td>&lt;150</td>
<td>&lt;3.0</td>
</tr>
<tr>
<td>F1</td>
<td>&lt;28</td>
<td>&lt;180*</td>
<td>&lt;4.0*</td>
</tr>
<tr>
<td>F2</td>
<td>&lt;30*</td>
<td>&lt;200</td>
<td>&lt;5.0</td>
</tr>
<tr>
<td>F3</td>
<td>&lt;40*</td>
<td>&lt;3.0</td>
<td>&lt;5.0</td>
</tr>
<tr>
<td>F4</td>
<td>&lt;50*</td>
<td>&lt;16.3</td>
<td>&lt;22.5</td>
</tr>
<tr>
<td>F5</td>
<td>&lt;60*</td>
<td>&lt;7.4</td>
<td>&lt;16.1</td>
</tr>
<tr>
<td>F6</td>
<td>&lt;70*</td>
<td>&lt;4.8</td>
<td>&lt;21.5</td>
</tr>
<tr>
<td>F7</td>
<td>&lt;80*</td>
<td>&lt;0.2</td>
<td>&lt;27.0</td>
</tr>
<tr>
<td>F8</td>
<td>&lt;90*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Age is shown as mean (SD). Asterisks show the predefined cut-off values for vitamin B1, vitamin B12, and folate (vitamin B9) based on a reference\textsuperscript{12}; different cut-off values based on previous reports\textsuperscript{14–16} are also presented for further investigation.

ICD-10 codes. F0, Mental disorders due to known physiological conditions; F1, Mental and behavioral disorders due to psychoactive substance use; F2, Schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders; F3, Mood disorders; F4, Anxiety, dissociative, stress-related, somatoform, and other non-psychotic mental disorders; F5, Behavioral syndromes associated with physiological disturbances and physical factors; F6, Disorders of adult personality and behavior; F7, Intellectual disabilities; F8, Pervasive and specific developmental disorders; F9, Behavioral and emotional disorders with onset usually occurring in childhood and adolescence.
Table 2. Vitamin B deficiencies in sub-groups

<table>
<thead>
<tr>
<th></th>
<th>F0</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F6</th>
<th>F7</th>
<th>F8</th>
<th>F9</th>
</tr>
</thead>
<tbody>
<tr>
<td>vitB&lt;sub&gt;1&lt;/sub&gt; &lt; 30 [ng/mL]</td>
<td>9</td>
<td>4</td>
<td>70</td>
<td>11</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>(32%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitB&lt;sub&gt;12&lt;/sub&gt; &lt; 180 [ng/L]</td>
<td>5</td>
<td>4</td>
<td>53</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>(18%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folate &lt; 4.0 [μg/L]</td>
<td>5</td>
<td>7</td>
<td>38</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>(18%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F0</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F6</th>
<th>F7</th>
<th>F8</th>
<th>F9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitB&lt;sub&gt;1&lt;/sub&gt; &lt; 30 [ng/mL]</td>
<td>1.68</td>
<td>0.80</td>
<td>1.12</td>
<td>0.78</td>
<td>0.79</td>
<td>1.10</td>
<td>1.15</td>
<td>0.48</td>
<td>0</td>
</tr>
<tr>
<td>[0.74-3.83]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitB&lt;sub&gt;12&lt;/sub&gt; &lt; 180 [ng/L]</td>
<td>1.14</td>
<td>1.24</td>
<td>1.35</td>
<td>0.69</td>
<td>1.21</td>
<td>0.18</td>
<td>1.32</td>
<td>0.73</td>
<td>0</td>
</tr>
<tr>
<td>[0.42-3.10]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folate &lt; 4.0 [μg/L]</td>
<td>1.30</td>
<td>3.16</td>
<td>0.70</td>
<td>0.65</td>
<td>2.81</td>
<td>0.59</td>
<td>1.32</td>
<td>1.04</td>
<td>0</td>
</tr>
<tr>
<td>[0.48-3.55]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Square brackets indicate the 95% confidence interval.

Abbreviations; see Table 1.
Table 3. Summary of full dataset of age, sex, and 29 parameters for machine-learning

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Units</th>
<th>Mean</th>
<th>SD</th>
<th>UN</th>
<th>mg/dL</th>
<th>Cre</th>
<th>mg/dL</th>
<th>0.7</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>years</td>
<td>42.3</td>
<td>15.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Woman n = 228</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T.bil</td>
<td>mg/dL</td>
<td>0.7</td>
<td>0.4</td>
</tr>
<tr>
<td>WBC</td>
<td>×10³/µL</td>
<td>8.2</td>
<td>2.8</td>
<td></td>
<td></td>
<td>Na</td>
<td>mmol/L</td>
<td>139</td>
<td>3</td>
</tr>
<tr>
<td>Hb</td>
<td>g/dL</td>
<td>13.7</td>
<td>1.7</td>
<td></td>
<td></td>
<td>Cl</td>
<td>mmol/L</td>
<td>105</td>
<td>4</td>
</tr>
<tr>
<td>Hct</td>
<td>%</td>
<td>40.3</td>
<td>4.5</td>
<td></td>
<td></td>
<td>K</td>
<td>mmol/L</td>
<td>3.7</td>
<td>0.4</td>
</tr>
<tr>
<td>MCV</td>
<td>fL</td>
<td>89</td>
<td>6.6</td>
<td></td>
<td></td>
<td>cor.Ca</td>
<td>mg/dL</td>
<td>9.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Plt</td>
<td>×10⁴/µL</td>
<td>24.9</td>
<td>6.3</td>
<td></td>
<td></td>
<td>CK</td>
<td>IU/L</td>
<td>514</td>
<td>1230</td>
</tr>
<tr>
<td>RDW.CV</td>
<td>%</td>
<td>13.5</td>
<td>1.3</td>
<td></td>
<td></td>
<td>AST</td>
<td>IU/L</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>Neu</td>
<td>%</td>
<td>70</td>
<td>11</td>
<td></td>
<td></td>
<td>ALT</td>
<td>IU/L</td>
<td>27</td>
<td>24</td>
</tr>
<tr>
<td>Lym</td>
<td>%</td>
<td>23</td>
<td>10</td>
<td></td>
<td></td>
<td>LDH</td>
<td>IU/L</td>
<td>239</td>
<td>91</td>
</tr>
<tr>
<td>Mono</td>
<td>%</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td>ALP</td>
<td>IU/L</td>
<td>224</td>
<td>81</td>
</tr>
<tr>
<td>Eo</td>
<td>%</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>γGTP</td>
<td>IU/L</td>
<td>37</td>
<td>63</td>
</tr>
<tr>
<td>Baso</td>
<td>%</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>Glu</td>
<td>mg/dL</td>
<td>112</td>
<td>40</td>
</tr>
<tr>
<td>TP</td>
<td>g/dL</td>
<td>7.2</td>
<td>0.6</td>
<td></td>
<td></td>
<td>CRP</td>
<td>mg/dL</td>
<td>0.4</td>
<td>0.9</td>
</tr>
<tr>
<td>Alb</td>
<td>g/dL</td>
<td>4.4</td>
<td>0.4</td>
<td></td>
<td></td>
<td>TSH</td>
<td>µIU/mL</td>
<td>1.7</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Two patients lacked age data (no photo ID was available), and one patient lacked biochemistry data (inappropriate sample processing). For machine-learning, the missing values were replaced using the mean.

Abbreviations: WBC, white blood cell count; Hb, hemoglobin; Hct, hematocrit; MCV, mean corpuscular volume; RDW.CV, red blood cell distribution width-coefficient variation; Plt, platelet; Neu, neutrocyte fraction; Lym, lymphocyte fraction; Mono, monocyte fraction; Eo, eosinocyte fraction; Baso, basocyte fraction; TP, total protein; Alb, albumin; UN, urea nitrogen; Cre, creatinine; T.bil, total bilirubin; Na, sodium; Cl, chlorine; TSH, thyroid-stimulating hormone.
chloride; K, potassium; cor.Ca, corrected calcium; CK, creatine kinase; AST, aspartate transaminase; ALT, alanine transaminase; LDH, lactate dehydrogenase; ALP, alkaline phosphatase; γGTP, γ-glutamyltransferase; Glu, plasma glucose; CRP, C-reactive protein; TSH, thyroid-stimulating hormone.
Table 4. Summary of AUC, sensitivity, specificity, and accuracy for the validation set

<table>
<thead>
<tr>
<th>Classifier</th>
<th>vitB₁</th>
<th>vitB₁₂</th>
<th>Folate</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-nearest neighbors</td>
<td>0.596</td>
<td>0.542</td>
<td>0.514</td>
<td>0.551</td>
</tr>
<tr>
<td></td>
<td>[0.483–0.702]</td>
<td>[0.394–0.705]</td>
<td>[0.383–0.651]</td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>0.715</td>
<td>0.602</td>
<td>0.754</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>[0.602–0.815]</td>
<td>[0.454–0.745]</td>
<td>[0.610–0.877]</td>
<td></td>
</tr>
<tr>
<td>Support vector machine</td>
<td>0.715</td>
<td>0.620</td>
<td>0.699</td>
<td>0.678</td>
</tr>
<tr>
<td></td>
<td>[0.613–0.814]</td>
<td>[0.472–0.763]</td>
<td>[0.536–0.842]</td>
<td></td>
</tr>
<tr>
<td>Random forest</td>
<td>0.716</td>
<td>0.599</td>
<td>0.796</td>
<td>0.704</td>
</tr>
<tr>
<td></td>
<td>[0.610–0.825]</td>
<td>[0.426–0.755]</td>
<td>[0.656–0.911]</td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity, specificity, and accuracy for random forest

<table>
<thead>
<tr>
<th></th>
<th>vitB₁</th>
<th>vitB₁₂</th>
<th>Folate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.594</td>
<td>0.316</td>
<td>0.667</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.783</td>
<td>0.943</td>
<td>0.917</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.688</td>
<td>0.629</td>
<td>0.792</td>
</tr>
<tr>
<td></td>
<td>[0.597–0.787]</td>
<td>[0.523–0.746]</td>
<td>[0.665–0.909]</td>
</tr>
</tbody>
</table>

Generalization performance of the classifiers was evaluated using AUC of the validation set for each type of classifiers. For random forest classifiers, sensitivity, specificity, and accuracy of the classification at the optimal operating points that maximized accuracy on the receiver operating characteristic curve of the validation set are also shown (see also Figure 2 D-F). Accuracy was defined as the average of the sensitivity and specificity. Square brackets indicate the 95% confidence interval. For further information, see Figure 2 and Supplementary Table 5.
Abbreviations: AUC, area under the receiver operating characteristic curve.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5

Admission

~Few hours

~Several days

497 consecutive patients with intense psychiatric episode

Outsourcing

Efficient prediction

Age, sex, 29 routine blood test results

Random forest classifier

AUC 0.796 (Folate)

Definitive serological values of vitB$_1$, vitB$_{12}$, folate

- Risk stratification
- Decision-making
- Potential biological mechanism suggestion