# CASE STUDIES OF UTILIZATION OF THE MIND MONITORING SYSTEM (MIMOSYS) USING VOICE AND ITS FUTURE PROSPECTS

**Shuji Shinohara[1], Yasuhiro Omiya[2], Naoki Hagiwara[2], Mitsuteru Nakamura[1], Masakazu Higuchi[1], Takashi Kirita[2], Takeshi Takano[2], Shunji Mitsuyoshi[1], Shinichi Tokuno[1]**

[1]Graduate School of Medicine, The University of Tokyo, 113-0033, Japan, e-mail: {shinokan99, m-nakamura, higuchi, mitsuyoshi, tokuno}@m.u-tokyo.ac.jp
[2]Research and Product Development, PST Inc., 231-0023, Japan, e-mail: {omiya, hagiwara, kirita.takashi, takano}@medical-pst.com

**Abstract:** *We developed a method to measure the mental health condition of speakers based on the emotional components contained in the voice; we named the method Mind Monitoring System (MIMOSYS). Voice is input into MIMOSYS, and MIMOSYS outputs two vocal indices: vitality, a short-term index, and mental activity, which is calculated from the long-term tendency of vitality.*
*In this article, we first present an overview of MIMOSYS. We then present a development case of a smartphone app that utilizes MIMOSYS. Following this, we present case studies in which MIMOSYS vocal indices were used. The first study is on the relation of vitality and mental activity with the Beck Depression Inventory, a questionnaire index widely used for diagnosing depression. The second study is on the relation of vitality with Brain-Derived Neurotrophic Factor, blood indices, and the General Health Questionnaire 30 (GHQ30), a questionnaire index regarding neurosis. These studies demonstrated that the vocal indices of MIMOSYS exhibit identification tendencies similar to those of questionnaire indices and blood indices.*

**Keywords:** *Mind Monitoring System (MIMOSYS), Mental healthcare, telemonitoring, Beck Depression Inventory (BDI), Brain-Derived Neurotrophic Factor (BDNF), General Health Questionnaire (GHQ)*
PACS numbers: 89.65.Gh, 89.75.Fb, 05.45.Tp

## 1. INTRODUCTION

In modern society, economic losses caused by the mental health disorders that individuals experience have become an international issue, and there is a call for appropriate measures [1] [2]. In order to address this issue, a mental health screening method that can be used on a daily basis at a low cost is necessary. Currently, the main means of assessing mental health include medical interviews by professionals such as physicians and self-administered questionnaires such as the General Health Questionnaire (GHQ) [3] and Beck Depression Inventory (BDI) [4]. However, there is a limit on the number of medical interviews that can be conducted by professionals, and reporting bias is an issue with self-administered questionnaires. Here, reporting bias indicates the selective concealment or exposure by the subject, of specific information such as medical history or smoking history. Moreover, although assessment using biomarkers such as saliva and blood has been studied [5][6], the issues of costs and burden on the subjects remain.

On the other hand, pathophysiology analysis using vocal data is drawing attention with the recent spread of smartphones [7]. Its non-invasive nature, in addition to the fact that it can be conveniently conducted remotely, as it does not require a dedicated device, is an advantage of voice analysis using smartphones. From this point of view, the authors have been engaged in the development of a method that estimates stress conditions and depression based on vocal data [8] [9].

Specifically, as stress exerts an impact on humans' emotions [10], we developed a method to measure the mental health of a speaker based on the variations in emotional components extracted from the voice rather than directly analyzing stress conditions based on vocal data; we named it Mind Monitoring System (MIMOSYS) [9]. Voice is input into MIMOSYS, and MIMOSYS outputs two vocal indices of vitality – a short-term index, and mental activity, which is calculated from the long-term tendency of vitality.

In this article, we first present an overview of MIMOSYS. We then present a development case of a smartphone app that uses MIMOSYS. Following this, we present case studies in which MIMOSYS vocal indices were used. Finally, we describe the future prospects.

## 2. OVERVIEW OF THE MIND MONITORING SYSTEM (MIMOSYS)

In this chapter, we present an overview of the algorithm of MIMOSYS [9]. MIMOSYS is a system that measures the mental health condition of a person based on voice. For example, with regard to Major Depressive Episodes, the DSM-V lists characteristics such as loss of interest or pleasure and continued depressed mood, in which one experiences sadness and emptiness. Conversely, a higher proportion of the pleasure component than that of sadness in emotions can be considered as an indication of reasonable mental condition. From these perspectives, MIMOSYS estimates the mental health condition of a speaker based on the balance and variations in the emotional components in the speaker's voice.

MIMOSYS first calculates the degree of intensity of each of the four emotional components ("calmness," "anger," "joy," and "sorrow") in voiced speech on eleven-point scale of zero to ten using Sensibility Technology Ver.3.0 (AGI Inc., Tokyo, Japan), a vocal emotion recognition technology [11-13]. It also calculates the intensity of "excitement" on a ten-point scale of one to ten.

Next, based on the intensity of these five indices, vivacity and relaxation are calculated. It is noteworthy that vivacity is calculated from the joy and sorrow components, while relaxation is calculated from the calmness and excitement components. Finally, vitality is calculated from vivacity and relaxation.

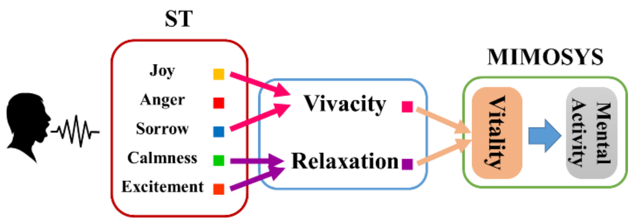



**Fig. 1** Data calculation flow of MIMOSYS. (Adapted from reference [14])

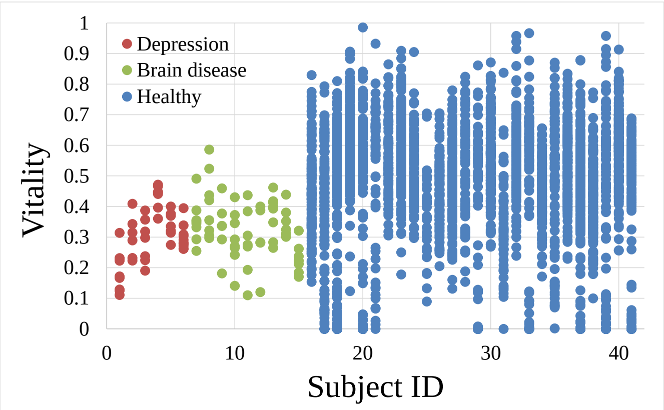



**Fig. 2** Example of vitality measurement. The horizontal axis represents subject ID, and the vertical axis represents vitality. The subjects are ordered from left to right as follows: patients with depression; brain disease patients with cerebral infarction, intracerebral hemorrhage, etc.; and healthy subjects. The data are arranged vertically because separate measurements of each subject were collected. (Adapted from reference [15])

In general, "vitality" can be defined in variously and can imply various concepts. However, here, vitality can be briefly defined as a scale on which patients experiencing depression or cerebral infarction score low, while healthy individuals score high. MIMOSYS calculates mental activity in addition to vitality as an index of mental health. These indices are output as real values in the interval [0.0, 1.0].

The main dissimilarity between vitality and mental activity is the duration of the measurement period. Vitality estimates the degree of mental health based on emotional components ("calmness," "anger," "joy," and "sorrow") as well as "excitement" contained in short-term vocal data such as a phone call.

On the other hand, mental activity is calculated based on data of vitality accumulated over a certain period of time, such as two weeks. Vitality varies depending on the circumstances at the time of measurement in a manner similar to variations in blood pressure while at rest immediately after exercise. We aim for a more accurate assessment of mental health by introducing mental activity in a manner similar to enhancing accuracy of determination of hypertension through long-term monitoring of blood pressure. Figure 1 illustrates the calculation flow.

Moreover, Figs. 2 and 3 illustrate measurement examples of the vitality and mental activity of 26 healthy subjects and 15 patients (six patients with depression and nine cerebral infarction patients). Here, the horizontal axis represents subject ID, and the vertical axis represents the respective index values (vitality and mental activity). The subjects are ordered from left to right as follows: patients with depression; brain disease patients with cerebral infarction, intracerebral hemorrhage, etc.; and healthy subjects. It is noteworthy that in Fig. 2, because separate measurements of each subject were collected, the data for vitality are arranged vertically.

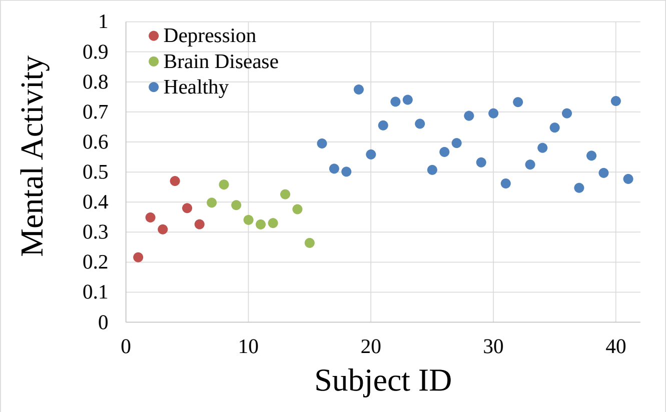



**Fig. 3** Example of mental activity measurement. The horizontal axis represents subject ID, and the vertical axis represents mental activity. The subjects are ordered from left to right as follows: patients with depression; brain disease patients with cerebral infarction, intracerebral hemorrhage, etc.; and healthy subjects. (Adapted from reference [15])

**Table 1** Identification performance of MIMOSYS with respect to healthy subjects and patients (depression and brain disease)

| Index | AUC | Sensitivity | Specificity |
|---|---|---|---|
| Vitality | 0.80 | 0.94 | 0.64 |
| Mental Activity | 0.99 | 1.0 | 0.92 |

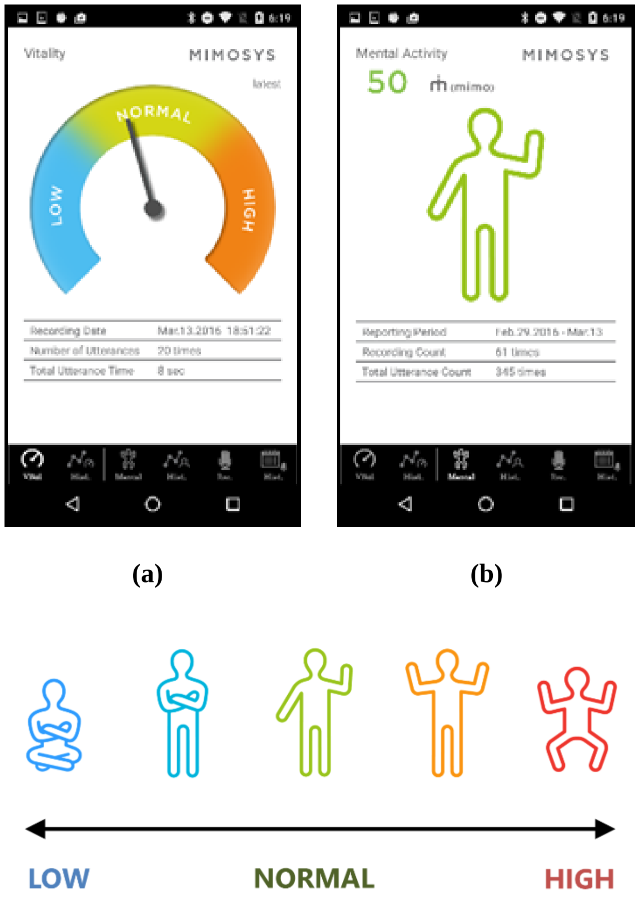
We present the Area under the Curve (AUC) as well as the sensitivity and specificity against the Receiver Operating Characteristic (ROC). Regarding vitality, AUC was 0.80, while sensitivity and specificity were 0.94 and 0.64, respectively. Meanwhile, regarding mental activity, AUC was 0.99, while sensitivity and specificity were 1.0 and 0.92, respectively. It has, thus, been demonstrated that mental activity, which is a long-term index, enhances identification performance. In particular, enhancement in specificity is significant.

## 3. IMPLEMENTATION OF MIMOSYS AS A SMARTPHONE APP

In this chapter, we present a development case of a smartphone app that uses MIMOSYS. We took note of telephone calls in which we routinely speak out loud and implemented MIMOSYS as a smartphone app [16].

This system consists of the following processes:

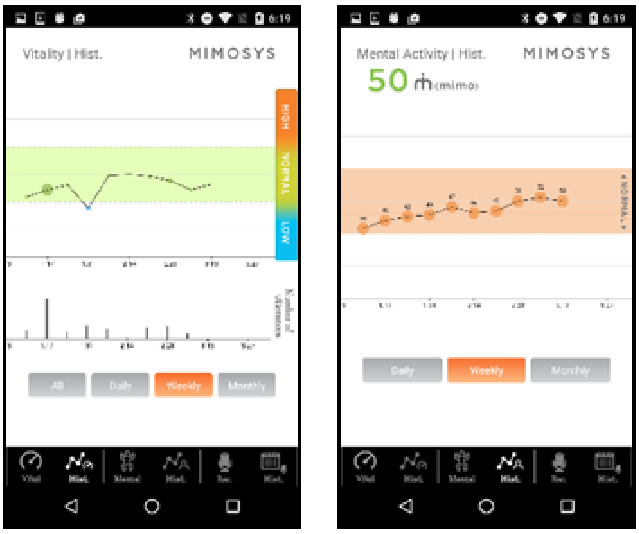
1. Audio recording
2. Analysis of health condition based on recorded voice
3. Accumulation of analysis results
4. Deletion of audio used for analysis
5. Presentation of accumulated analysis results to the user




(a) (b)

(c)

**Fig. 4** Examples of the smartphone app's display of analysis results: (a) vitality, (b) mental activity, and (c) human-type animation representing mental activity. (Adapted from reference [16])




(a) (b)

**Fig. 5** Examples of the smartphone app's display of the history of analysis results: (a)vitality and (b)mental activity. (Adapted from reference [16])

Figure 4(a) {vitality}, Fig. 4(b) {mental activity}, Fig. 5(a) {history of vitality}, and Fig. 5(b) {history of mental activity} are illustrated as examples of display of analysis results by this app.

In the smartphone app, both vitality and mental activity are converted from real values within [0.0, 1.0] to integer values within [0,100].

As illustrated in Figs. 4(b) and (c), mental activity is displayed on a five-point scale using human-type animation, according to the mental activity level. Moreover, as illustrated in Fig. 5, history of both vitality and mental activity are displayed on line charts.

Currently, a large-scale demonstration experiment is being conducted using this smartphone app [17].

## 4. BDI AND MIMOSYS INDICES (VITALITY AND MENTAL ACTIVITY)

In this chapter, we present a study [14] on the relation between the BDI, which is widely used for diagnosing depression, and MIMOSYS vocal indices (vitality and mental activity).

As part of this study, we collected audio data from 50 subjects (39 males and 11 females) over a period of approximately 2 months using the MIMOSYS smartphone app, which we described in the previous chapter. We also conducted the BDI test at the time when audio acquisition was started. However, because BDI scores are likely to vary in the medium- to long-term, at the time of analysis, we used audio data from the two weeks since the BDI test was performed. Of the data on 50 subjects, those on 48 subjects were valid for analysis.

First, based on the BDI scores, the subjects were divided – according to reference [18] – into two groups: the "low risk of disease" group with a score of 18 or less and the "high risk of disease" group with a score higher than 18. There were forty-three subjects in the low risk of disease group and five in the high risk of disease group.

Figure 6 illustrates the vitality of the two groups. The average values of vitality for the low risk of disease group and high risk of disease group were 0.37 (SD = 0.17, N = 1221) and 0.21 (SD = 0.094, N = 159), respectively. The results of the t-test demonstrated a significant difference between the two groups (t(307) = 16.89, p = 1.03E-46 ).

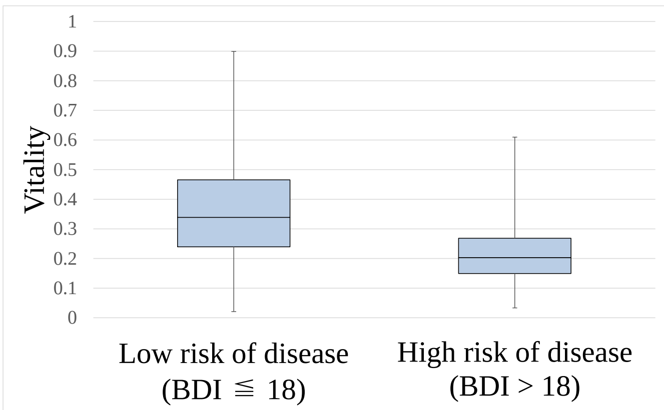



**Fig. 6** Box and whisker plots of vitality score by fixed phrase for the low risk of disease and high risk of disease groups. (Adapted from reference [14])

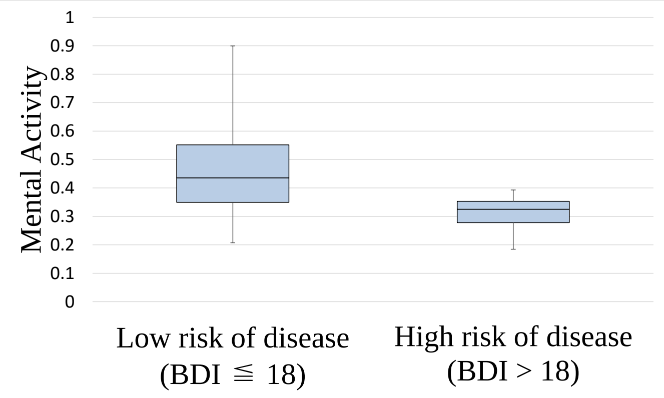



**Fig. 7** Box and whisker plots of mental activity score by fixed phrase for the low risk of disease and high risk of disease groups. (Adapted from reference [14])

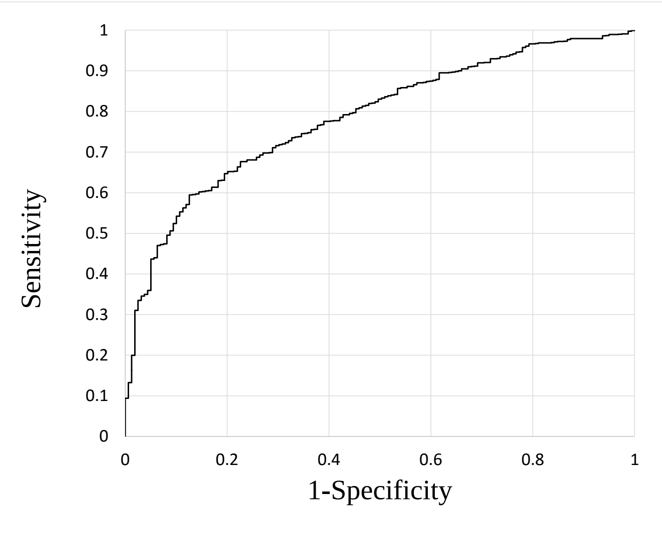



**Fig. 8** ROC curve used in discriminating between low and high risk of disease using vitality index. (Adapted from reference [14])

**Table** 2 Identification performance of MIMOSYS regarding low risk of disease group and high risk of disease group.

| Index | AUC | Sensitivity | Specificity |
|---|---|---|---|
| Vitality | 0.78 | 0.80 | 0.64 |
| Mental Activity | 0.81 | 1.0 | 0.61 |

Further, we illustrate the mental activity of the two groups in Fig. 7. The average values of mental activity for the low risk of disease group and high risk of disease group were 0.46 (SD = 0.15, N = 43) and 0.31 (SD = 0.08, N =5), respectively. The results of the t-test demonstrated a significant difference between the two groups ($t(8) = -3.6$, $p = 0.007$).

We present the AUC, sensitivity, and specificity of the ROC curve in Table 2 to determine the identification performance with regard to vitality and mental activity for the low risk of disease group and high risk of disease group. The AUC of vitality was 0.78, and the sensitivity and specificity were 0.80 and 0.64, respectively. Meanwhile, the AUC of mental activity was 0.81, and the sensitivity and specificity were 1.0 and 0.61, respectively. Figures 8 and 9 reveal the ROC curve of vitality and mental activity.

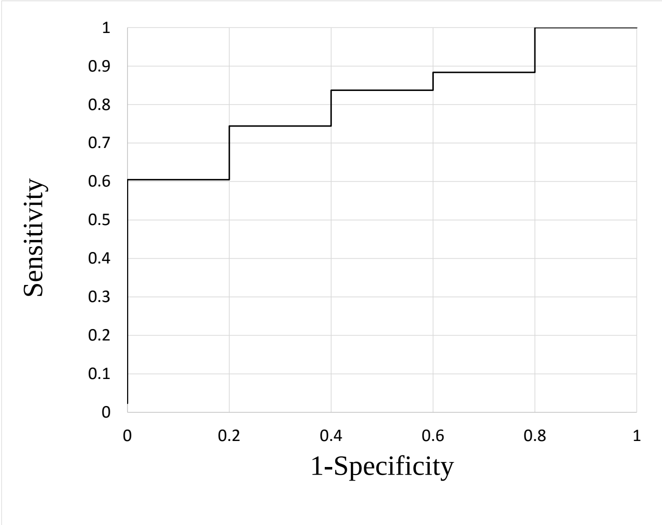



**Fig. 9** ROC curve used in discriminating between low and high risk of disease using mental activity index. (Adapted from reference [14])

Thus, the AUCs of vitality and mental activity were both approximately 0.8 indicating reasonable identification performance.

## 5. MIMOSYS, BLOOD INDICES, AND QUESTIONNAIRE INDICES

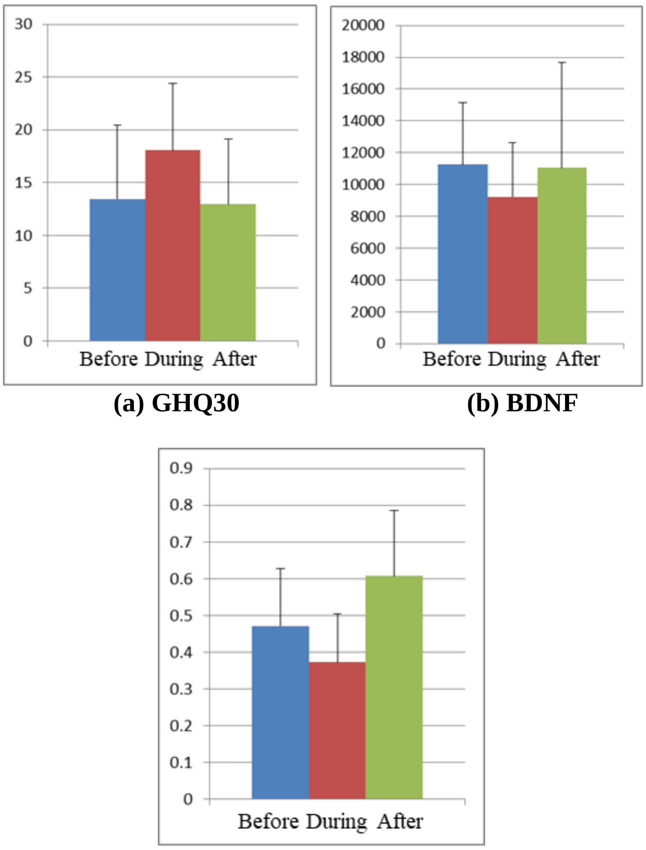
It is known that subjects with depression or those that are under stress exhibit decreased blood Brain Derived Neurotrophic Factor (BDNF) concentrations [19]. Similar to the BDI, the GHQ-30 is also a self-administered questionnaire that has been widely used for diagnosing depression.

In this chapter, we present studies [19] [20] on the relation of these two indices, that is, BDNF, a blood index, and the GHQ-30, a questionnaire index, with vitality, a vocal index.

In this study, the subjects were the members of the Self-Defense Force airborne brigade at the time of a ranger training program during which they were under extreme stress.
The training was carried out over a period of nine weeks.
Blood collection, audio recording and a self-administered questionnaire were performed three times, once each before, during (three weeks into the training program) and after (three to five days after the training program had ended) the training program.

Figure10 presents the values of (a) the self-administered questionnaire GHQ-30, (b) blood index BDNF, and (c) vocal index "vitality" before, during, and after the Self-Defense Force ranger training program. The data before, during, and after the training program are ordered from left to right in each figure. Here, it should be noted that the higher the GHQ-30 scores, the higher the stress or depression tendency. On the other hand, the lower the vitality and BDNF, the higher the stress or depression tendency. In effect, the GHQ-30 is an index with a magnitude relation that is in reverse to the other two.

The figures indicate a similar tendency in which each index exhibits the highest stress levels during training.




**(a) GHQ30** **(b) BDNF**

**(c) Vitality**

**Fig. 10** Comparison of the results of (a) the self-administered questionnaire GHQ-30, (b) blood index BDNF, and (c) vocal index "vitality" before, during, and after the Self-Defense Force ranger training program. (Adapted from the reference [20])

## 6. SUMMARY AND FUTURE PROSPECTS

In this article, we presented an overview and a few research and development cases of MIMOSYS. MIMOYS is likely to be capable of distinguishing between patients with depression and brain disease from healthy individuals based solely on voice, with high accuracy. Moreover, it is likely that it will serve as an alternative to other indices that measure the degree of stress and depression. For example, blood index BDNF or questionnaire indices such as the BDI and GHQ-30 screening solely based on voice is superior to blood tests in terms of costs and the burden on the subjects; moreover, it overcomes the issue of reporting bias that is present in self-administered questionnaires. Implementation of MIMOSYS as a smartphone app to analyze voice calls permits us to monitor daily mental health condition, which can result in early detection of depression, etc.

Currently, we are working on applying the MIMOSYS technology to field of the occupational medicine [21] and the development of automobiles [22]. With respect to application in the industrial hygiene field, we are developing a system that can result in the introduction of intervention at the appropriate time using self-administered questionnaires and a stress resilience program, as we monitor the stress levels of the employees of IT companies by MIMOSYS [21]. In addition, regarding application to the development of automobiles, in the past, the effects of driving a car have mostly been studied in light of negative aspects such as fatigue and drowsiness. Instead, we focus on positive aspects such as altering the mood to work on an application for developing a comfortable and safe automobile [22].

Furthermore, we are working on applying speech pathology analysis to depression and stress as well as to other diseases. Vitality and mental activity measured by MIMOSYS are imprecise indices to categorize healthy individuals and patients with depression or brain disease, and they are not suitable for detailed differentiation of diseases. Therefore, we are conducting research on its use for the differential diagnosis of a number of diseases and monitoring the course of diseases by directly extracting characteristic values unique to the diseases from voice without applying a vocal emotion recognition technology. In addition to mental disorders such as depression and bipolar disorder, we are currently conducting research on diseases and disorders that are likely to cause variation in the voice, such as neurological disorders (Parkinson's disease, etc.), dementia (including Alzheimer's disease), and dysarthria (vocal cord polyps, etc.).

Currently, feature values that are likely to enable us to distinguish patients with Parkinson's disease from healthy individuals [23, 24] and feature values that enable us to identify severity of depression [25], etc. have been recommended. In the future, we aim to develop a system that enables the differentiation of Parkinson's disease, dementia, and depression.

## 7. REFERENCES


[1] Kessler R. C., Akiskal H. S., Ames M., Birnbaum H., Greenberg P., Hirschfeld R. M. A., Jin R., Merikangas K. R., Simon G. E., Wang P. S. (2006). *Prevalence and effects of mood disorders on work performance in a nationally representative sample of U.S. workers*, Am. J. Psychiatry, 163, 1561-1568.

[2] World Health Organization (2004). *The Global Burden of Disease: 2004 update*, WHO Press, Geneva, Switzerland, 46-49.

[3] Goldberg D.P. (1978). *Manual of the General Health Questionnaire*, NFER Publishing, Windsor, England.

[4] Beck A. T., Ward C. H., Mendelson M., Mock J., Erbaugh J. (1961). *An Inventory for Measuring Depression*, Arch. Gen. Psychiatry, 4, 561-571.

[5] Izawa S., Sugaya N., Shirotsuki K., Yamada K. C., Ogawa N., Ouchi Y., Nagano Y., Suzuki K., Nomura S. (2008). *Salivary dehydroepiandrosterone secretion in response to acute psychosocial stress and its correlations with biological and psychological changes*, Biol. Psychol., 79(3), 294-298.

[6] Sekiyama A. (2007), *Interleukin-18 is involved in alteration of hipothalamic-pituitary-adrenal axis activity by stress*, SOBP Annu. Meeting., San Diego, USA.

[7] Shinohara S., Omiya Y., Nakamura M., Hagiwara N., Higuchi M., Mitsuyoshi S., Tokuno S. (2017). *Multilingual evaluation of voice disability index using pitch rate*, ASTESJ, 2(3), 765-772.

[8] Shinohara S., Mitsuyoshi S., Nakamura N., Omiya Y., Tsumatori G., Tokuno S. (2015). *Validity of a voice-based*

*evaluation method for effectiveness of behavioural therapy,* Pervasive Computing Paradigms for Mental Health, Springer International Publishing, 43-51.

[9] Shinohara S., Nakamura M., Omiya Y., Hagiwara N., Mitsuyoshi S., Tokuno S. *A mental health assessment method based on emotional level derived from voice,* in preparation.

[10] Lazarus R. S. (1993). *From psychological stress to the emotions: A history of changing outlooks,* Annu. Rev. Psychol., 44, Jan. 1-21.

[11] Mitsuyoshi S. (2003). *Emotion recognizing method, sensibility creating method, device, and software,* U.S. Patent 7340 393, Sep. 25, 2003.

[12] Mitsuyoshi S., Ren F., Tanaka Y., Kuroiwa S. (2006) *Non-verbal voice emotion analysis system,* Int. J. Innovative Comp. Info. and Control, 2, 819-830.

[13] Mitsuyoshi S., Tanaka Y., Ren F., Shibasaki K., Kato M., Murata T., Minami T., Yagura H. (2007). *Emotion voice analysis system connected to the human brain,* IEEE NLP-KE2007, Beijing, China, 479-484.

[14] Hagiwara N., Omiya Y., Shinohara S., Nakamura M., Yasunaga H., Mitsuyoshi S., Tokuno S. (2017). *Validity of mind monitoring system as a mental health indicator using voice*, ASTESJ, 2(3), 338-344.

[15] Tokuno S. (2016). *Onsei byoutai bunsekigaku*., Saibo, 48(14), 9-12. In Japanese

[16] Omiya Y., Hagiwara N., Shinohara S., Nakamura M., Mitsuyoshi S., Tokuno S. (2016). *Development of mind monitoring system using call voice,* Neuroscience 2016, San Diego, 2016.11.12-16.

[17] Tokuno S., Omiya Y., Shinohara S., Nakamura M., Hagiwara N., Mitsuyoshi S. (2016). *Psychological impact of Kumamoto earthquake by voice analysis using a smart phone application,* Neuroscience 2016, San Diego, 2016.11.12-16.

[18] Beck A. T., Steer R. A., Garbin M. G. J. (1988). *Psychometric properties of the Beck Depression Inventory Twenty-five years of evaluation,* Clin. Psych. Review, 8, 77–100.

[19] Suzuki G., Tokuno S., Nibuya M., Ishida T., Yamamoto T., Mukai Y., et al. (2014). *Decreased plasma brain-derived neurotrophic factor and vascular endothelial growth factor concentrations during military training,* PLoS ONE, 9(2), e89455.

[20] Tokuno S. (2015). *Stress evaluation by voice: a novel stress evaluation technology,* Annual Bilateral Behavioral Health Conference (Kanagawa), 2015.6.23

[21] Miyazaki K. (2016). *Kigyoban sutoresu regiriensu puroguramu to sono koka,* Kokoro to shakai, 47(3), 26-31. In Japanese

[22] Okazaki T. (2016) *Onsei byoutai bunseki gijutsu no jidousya eno ouyo,* Kokoro to shakai, 47(3), 32-37. In Japanese

[23] Shinohara S., Tokuno S. (2017). *Novel voice indicator for distinguishing parkinson's disease,* EMBC'2017, Jeju Island, Korea, 2017.7, 11-15.

[24] Omiya Y., Hagiwara N. (2017). *Algorithm to distinguish between articulatory disorder, depression and Parkinson's disease by voice,* EMBC'2017, Jeju Island, Korea, 2017.7, 11-15.

[25] Shinohara S., Omiya Y., Nakamura M., Higuchi M., Hagiwara N., Takano T., Toda H., Saito T., Tanichi M., Yoshino A., Mitsuyoshi S., Tokuno S. *Major depression index derived from the relationship between hurst exponent and zero crossing rate in voice,* Neuroscience 2017, accepted.