Correlation between self-administered psychological test and emotion measured by voice analysis

Mitsuteru NAKAMURA[1,*], Shuji SHINOHARA[1], Yasuhiro OMIYA[2], Shunji MITSUYOSHI[1], Hiroaki TAKAGI[3], Akira USHIWATARI[4], and Shinichi TOKUNO[1]

[1]Graduate School of Medicine, The University of Tokyo, Industry & Trade Center Building 905, 2 Yamashita-cho, Naka-ku, Yokohama, Kanagawa, Japan

[2]PST inc., Industry & Trade Center Building 905, 2 Yamashita-cho, Naka-ku, Yokohama, Kanagawa, Japan

[3]Trypro inc., Asahi mansion 201, 6-6-8 Daita, Setagaya-ku, Tokyo, Japan

[4]Cii Software Development Co.,Ltd., 1-6-10 Kawagishi, Toda, Saitama, Japan

Keywords: Mental health, Screening, Emotion, Self-administered psychological test, Voice emotion analysis

Abstract. It is well known that people's way of expressing emotion alters when they have a mental disorder. We investigated the relationship between scores on a self-administered psychological test, namely the General Health Questionnaire 30 Items Version (GHQ30), and emotion as recognized by voice. We asked participants to undertake the GHQ30 test twice with an approximately 80 day period in between, and we recorded their voices for one week before and after each test in order to obtain their emotion indices. Regarding the results, the emotion index "anger" showed a correlation with the GHQ30 scores, and the emotion index "excitement" showed a correlation with the GHQ30 subscale "suicidal ideation and depression." A comparison of the amount of change between the first and second administration of the test, on the other hand, revealed a correlation between the amount of change in the sorrow index and the amount of change in GHQ30 scores. It also revealed a correlation between the amount of change in the joy index and the amount of change in the GHQ30 subscales "general disorder trends," "physical conditions," and "social activity disorders." These results demonstrated that mental health state is reflected on emotion indices as recognized by voice analysis, suggesting that emotion as recognized by voice analysis may be useful as a monitor of mental health state.

Introduction

The losses associated with mental disorder are unignorably enormous and global in scale [1], and appropriate measures must be taken to address the problem. A necessary method for dealing with the problem is to provide mental health screening that is low-cost and can be used on a daily basis. At present, the main ways to evaluate mental health are consultation with a physician or other expert, and self-administered questionnaires such as the General Health Questionnaire 30 Items Version (GHQ30)[2], Beck Depression Inventory [3] and others. However, each of these has their respective drawbacks; there is a limit to the number of opportunities to have a consultation with an expert, and self-administered questionnaires have the problem of reporting bias (conscious or unconscious underestimation by the respondent). The use of biomarkers in mental health evaluations has also been studied, but there remain issues concerning their cost and the burden placed upon participants.

There is another possible method for inferring mental health. Given that mental health has an impact on emotions [4], conversely, monitoring changes in emotion may be used to infer mental health state.

There are a number of machine systems for monitoring emotion, including facial expression [5] and voice [6] [7] [8] [9] analysis. These systems are influenced by individual differences, but Sensibility Technology (ST) [6] [7] [8] [9] eliminates this problem by analyzing emotion in voice non-verbally. The emotions of individuals in situations considered to be highly stressful were monitored in a preceding study [10], but this study did not evaluate stress and mental health factors other than variance in the participants' situations, and so the results were based only on an inference of the objective stress inflicted on the participants. Another experiment [11], which was on a much larger scale, made use of voice recordings, GHQ30-based psychological testing, and professional interviews with some of the participants. However, due to the large scale of this experiment, only one voice recording was made for each participant, and the study failed to identify a significant correlation between emotion as recognized by voice analysis and GHQ30 results or professional diagnosis. In the present study, mental health state was evaluated twice with GHQ30 with an interval period of approximately 80 days, and the emotion indices as recognized by ST were also obtained based on sequential voice recordings. Therefore, in our comparison of the data, we were able to give consideration to changes over time.

Methods

With the support of the company concerned, we conducted this study on approximately 50 truck drivers. We recorded their voices immediately before work and immediately after work, and requested them to administer the GHQ30 test. The number of participants from whom we obtained both voice data and GHQ30 data was 26. Of these, 20 were male and 6 were female. The GHQ30 test was conducted twice. The second test was conducted approximately 80 days after the first. We distributed the GHQ30 questionnaires among the participants and collected them as and when they were completed. Accordingly, the response date varies between the individual participants. With regard to the voice data, we recorded the participants' voices sequentially and analyzed the voice data pertaining to around one week before and after the response dates of the two GHQ30 tests, respectively.

The voice data we recorded was work-related content. The participants' utterances were made during dialogues with a manager. To record the voices, we used a handheld microphone (PM-240, TOA, Kobe, Japan) connected via a USB audio interface (X2u, SHURE, Niles, IL, USA) to a laptop PC installed with an application designed for this experiment. Regarding the audio format, we used a linear pulse code modulation with a sampling frequency of 11025Hz and 16-bit quantization.

Regarding the audio data, we used ST to analyze each utterance based on units lasting from one breathe to the next. The principle we used to demarcate utterances was whether the audio volume continuously exceeds a set threshold or falls short of it. The indices we analyzed using ST were the five emotions of excitement, calm, anger, joy, and sorrow. We first worked out the averages of these emotion indices with respect to each sequential audio unit, which comprised a recording immediately before and another immediately after work, and then worked out the averages of the audio data results during two weeks subject to analysis . We then compared the average values for the emotion indices obtained with the GHQ30 scores conducted during that period. The GHQ30 includes six subscales: general disorder trends, physical conditions, sleep disorders, social activity disorders, anxiety and dysthymia, suicidal ideation and depression. We calculated the scores for each of these subscales and compared the data. We used R 3.2.0 [12] to calculate the correlation coefficient and the significance level.

Result

We compared each participant's GHQ30 scores obtained with their emotion indices as determined by voice analysis, and then investigated correlations among the whole sample of participants. Table 1 shows the correlation coefficients between the emotion indices obtained from the voice analysis and the GHQ30 scores; these coefficients do not reflect the timing of the recording and GHQ30 responses.

The total scores of GHQ30 showed a weak correlation with the anger index. Some of the GHQ30 subscales showed a correlation with the anger and excitement indices. There was a particularly strong correlation between the anger index and the sleep disorders subscale.

Table 1. Correlation between the GHQ30 scores and the emotion indices. Asterisks on a correlation coefficient show significance level. * shows the coefficient is significant at 0.1 level, ** show the coefficient is significant at 0.05 level, and *** show the coefficient is significant at 0.01 level. Sample size is 52 (twice for each participant).

Subscale of GHQ30	Emotion indices				
	Excitement	Calm	Anger	Joy	Sorrow
General disorder trends	0.020	-0.002	0.095	-0.030	0.010
Physical conditions	0.097	-0.105	0.105	0.045	0.103
Sleep disorders	-0.006	0.045	0.362***	-0.059	-0.186
Social activity disorders	-0.136	0.202	0.062	-0.183	-0.160
Anxiety and dysthymia	0.088	-0.058	0.260*	0.052	-0.099
Suicidal ideation and depression	0.306**	-0.216	0.125	0.190	0.081
Total	0.094	-0.041	0.240*	0.014	-0.058

Since the same participants undertook a second test around 80 days later, we obtained the amount of change in the emotion indices and GHQ30 scores pertaining to these participants, respectively, and calculated the correlation coefficients between the amount of change in emotional indices and the amount of change in GHQ30 scores (Table 2). The amount of change in total scores of GHQ30 was correlated with the amount of change in sorrow index (Figure 1). In addition, while it fell short of the significance level of 0.1, there was a weak correlation between the amount of change in anger index and the amount of change in joy index. We also observed correlations between the amount of change in GHQ30 subscales and the amount of change in sorrow and joy indices, the correlation between the amount of change in anger index and the amount of change in sleep disorders subscale being particularly strong.

Table 2. Correlation between the amount of change in GHQ30 scores and the amount of change in emotion indices. Asterisks on a correlation coefficient show significance level. * shows the coefficient is significant at 0.1 level, ** show the coefficient is significant at 0.05 level, and *** show the coefficient is significant at 0.01 level. Sample size is 26.

Subscale of GHQ30	Emotion indices				
	Excitement	Calm	Anger	Joy	Sorrow
General disorder trends	-0.148	-0.004	-0.000	-0.360*	0.586***
Physical conditions	-0.213	0.037	0.140	-0.356*	0.417**
Sleep disorders	0.094	-0.182	0.433**	0.091	0.001
Social activity disorders	-0.218	0.249	-0.197	-0.378*	0.190
Anxiety and dysthymia	0.084	-0.141	0.251	0.091	0.012
Suicidal ideation and depression	0.015	-0.002	0.173	-0.161	0.168
Total	-0.117	-0.023	0.249	-0.317	0.421**

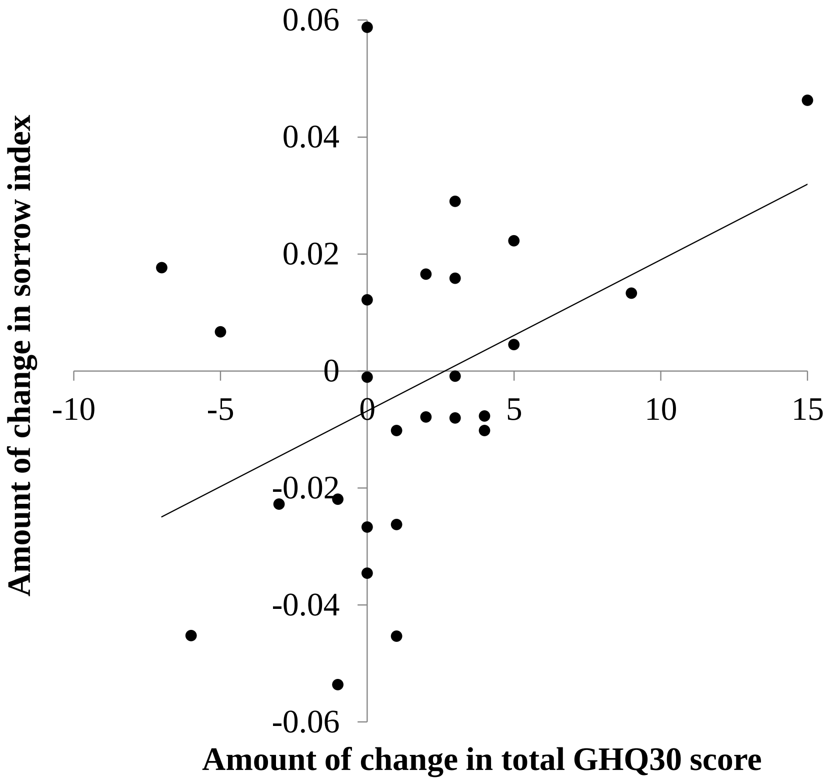

Figure 1. Scatter plot of the amount of change in GHQ30 scores and the amount of change in sorrow index. The correlation coefficient between these two variables is 0.421. The line in the plot shows the regression line.

Discussion

Regarding the comparison between the GHQ30 scores and the emotion indices, the emotions joy and sorrow did not necessarily show strong correlations when their values per se were compared with the GHQ30 scores, but each did show some degree of correlation when change over time for individual participants was measured. It is assumed that the influence of mental health state upon emotion varies between individuals. However, by observing changes over time for individual participants, respectively, we were able to offset individual differences and identify common trends among these individuals.

In the case of excitement and anger on the other hand, the correlation between these indices and the GHQ30 scores became scarcely observable after we included the amount of change between the two tests. Measuring amount of change narrows the range of available values, increasing the influence of factors outside the correlations. These emotion indices may have been affected more significantly by the decreased data range rather than the negation of individual differences resulted from amount of change measured.

The GHQ30 subscale sleep disorders showed a strong correlation with the anger index and scores per se and also a strong correlation in terms of amount of change. This finding suggests that the anger emotion index may be an indicator of not just general mental health state, but also of sleep disorders.

Conclusion

With the aim of devising a method of mental health screening that is low-cost and can be implemented on a daily basis, we examined the relationship between emotion as recognized by voice analysis and mental health state based on self-administered questionnaires. We found a correlation between the total GHQ30 score and the anger index. We also found correlation between the GHQ30 subscales and anger and excitement indices. When we compared the amount of change in scores between the two administrations of the GHQ30 test taken around 80 days apart from each other with the amount of change in emotion indices before and after the time of these administrations, we found that the amount of change in total GHQ30 score was correlated with that of the sorrow index. We also observed a correlation in terms of amount of change between some of the GHQ30 subscales and sorrow index as well as joy index.

These results suggest that emotions as recognized from voice analysis could be used to infer the mental health of the speaker. Inasmuch as the emotion index anger directly reflects mental health state, examining changes in sorrow index for a specific individual may provide a means to ascertain the changes in the mental health state of that individual.

Because we used voice analysis as a means to recognize emotion, we were able to gather the data at extremely low cost. By making use of telephone conversation settings, it may become possible to monitor the mental state of the user without them being overly conscious about it, and so we believe that such technology will enable screening on a daily basis.

References

[1] R. C. Kessler, S. Aguilar-Gaxiola, J. Alonso, S. Chatterji, S. Lee, J. Ormel, T. B. Üstün, P. S. Wang, The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys, Epidemiol. Psichiatr. Soc. 18 (2009) 23-33.

[2] D. P. Goldberg, Manual of the General Health Questionnaire, NFER Publishing, Windsor, England, 1978.

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

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

[5] Y. Tian, T. Kanade, J. F. Cohn, "Facial Expression Recognition," in Handbook of Face Recognition, Springer-Verlag London, London, United Kingdom, 2011.

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

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

[8] S. Mitsuyoshi, U.S. Patent 7,340,393. (2003)

[9] S. Mitsuyoshi, Development of Verbal Analysis Pathophysiology, Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal 5 (2015) 11-16.

[10] S. Tokuno, G. Tsumatori, S. Shono, E. Takei, T. Yamamoto, G. Suzuki, S. Mitsuyoshi, and M. Shimura, Usage of emotion recognition in military health care, DSR 2011, Singapore, Singapore, (2011) 1-5.

[11] S. Tokuno, Stress Evaluation by Voice: From Prevention to Treatment in Mental Health Care, Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal 5 (2015) 30-35.

[12] R. Ihaka and R. Gentleman, R: a language for data analysis and graphics, Journal of Computational and Graphical Statistics 5 (1996) 299-314.