Environ. Sci., Hokkaido University	11 (1)	1~10	June 1988

The Twelve-Item General Health Questionnaire Among Japanese Workers

Noboru Iwata[1], Yuko Okuyama[2], Yoshikazu Kawakami[2] and Kazuo Saito[1]

[1] Department of Hygiene and Preventive Medicine, Division of Social Environment, Graduate School of Environmental Science, Hokkaido University

[2] First Department of Internal Medicine, Hokkaido University School of Medicine

Abstract

The psychometric properties and the score distributions of 12 selected items from the Japanese version General Health Questionnaire (GHQ-12) were examined by using data obtained from 2,070 workers in a public institution. The internal consistency of the scale was at a satisfying level for males, but not for females. A principal factor analysis with varimax rotation of the GHQ-12 yielded two factors for both sexes: psychological distress and social dysfunction. No sex difference in the factor structure was found. The percentage distribution of the GHQ-12 scores was considerably skewed to the right. No sex difference was observed in either the mean score or the proportion of the subjects with a GHQ-12 score above the standard cut-off point. Correlational analysis of the GHQ-12 factors indicated that the factors derived from the GHQ-12 were related to various aspects of neurotic symptoms.

Key Words: General Health Questionnaire, factor analysis, self-rating psychiatric scale, Japanese workers

1. Introduction

The General Health Questionnaire (GHQ) was was principally designed as a test for the screening of non-psychotic psychiatric illness[1,2]. The GHQ is one of the more efficient inventories of its type[3] and is now the most widely used measurement for surveys of psychiatric epidemiology in the U. K. and other countries. It has been used in not only in-patient and out-patient clinics, but also in general medical care facilities, in the screening of community populations[4,5], in other areas of mental health studies[6,7], and in a number of countries, including Australia[8], Brazil[9], China[10], Finland[11], Jamaica[12], Spain[13], Taiwan[14], and Yugoslavia[15].

In a comparison between neurotic patients and normal controls, Nakagawa and Daibo[16] reported that the validity of the Japanese version GHQ was at the same level of that of the original version. Iwata and Saito[17] made clear that the Japanese version GHQ had a concurrent validity with the Todai Health Index (THI)[18] and a good level of internal consistent reliability among Japanese workers.

Among the five versions of the GHQ, the 12-item version (GHQ-12) is the

shortest. It takes only a short time to complete; thus the GHQ-12 seems to be easy to apply to both epidemiologic community surveys and occupational studies because of its brevity and lower frequency of missing data[6].

As with other psychometric tests, several factor-analytic studies on the GHQ have been carried out for the data of the U. K., the U. S., and the Australian populations. The authors demonstrated that the factor structure of the full-item version GHQ of Japanese workers was relatively similar to those of other studies based on the U. K. population[19].

However, studies on the psychometric properties, including factor-analytic study, of the GHQ-12 have been few. Thus, in the present study, the authors examine these issues based on the data of Japanese workers. Relationships of the factors derived from the GHQ-12 to several indices of the THI are also tested. These procedures allow us to discuss the nature of dimensions measured by the GHQ-12. This is the first analysis of the GHQ-12 of Japanese.

2. Materials and Methods

Two thousand one hundred and ninety workers in a public institution in Hokkaido were asked to complete the Japanese version GHQ translated by Nakagawa et al.[1], the THI, and other questions. The survey was carried out in September 1986. Complete data of the GHQ-12 were obtained from 2,142 subjects (97.8%; males: 1,986, females: 152, and unknown: 4). Among them, 2,070 subjects aged 20-59 years were analyzed in this study.

Table 1. Corrected item-whole correlations* and reliability coefficients of the Japanese version GHQ-12 by sex

the GHQ-12 irems	Males (n=1,927)	Females (n=143)
1 Able to concentrate	.42[c]	.35[c]
2 Lost much sleep over worry	.49[c]	.38[c]
3 Playing a useful part in things	.39[c]	.21[a]
4 Capable of making decisions	.40[c]	.21[a]
5 Constantly under strain	.60[c]	.47[c]
6 Couldn't overcome difficulties	.58[c]	.56[c]
7 Enjoy your normal activities	.41[c]	.27[b]
8 Able to face up to problems	.40[c]	.12
9 Feeling unhappy and depressed	.69[c]	.69[c]
10 Losing confidence in yourself	.65[c]	.66[c]
11 Thinking yourself worthless	.58[c]	.54[c]
12 Feeling reasonably happy	.22[c]	−.04
Mean corrected item-whole correlations:	.49	.37
Alpha coefficients:	.824	.736

a, b, c: significance levels of the correlation coefficient at $p<0.05$, 0.01, 0.001, respectively.
*: corrected item-whole correlation means the correlation coefficient between an item and the total score summing up the scores of the remaining eleven items.

The THI was developed as a general health questionnaire with the purpose of supplementing the Cornell Medical Index, which has been widely used in Japan[18,20]. It has 12 scales and three discriminant-function (DF) values to estimate the tendencies to neurosis (NEUR), psychosomatic disease (PSD), and schizophrenia (SCHIZO)[13,21].

Corrected item-whole correlations and Cronbach's alpha coefficients[22] of the GHQ-12 were calculated based on the 0-1-2-3 Likert score for each response. Corrected item-whole correlation means that the correlation coefficient between an item and the total score, the sum of the response scores for the remaining eleven items. A principal factor analysis with varimax rotation was carried out by sex. A squared multiple correlation was used as the initial communality estimate of an item. Factor score estimates were taken by the method of least squares that has been recognized to have the highest reliability and to be theoretically most clear among the several methods used for calculating the factor score[23]. The score of the GHQ-12 was calculated by sex-age group based on the 0-0-1-1 GHQ scoring.

3. Results and Discussion

(1) *Psychometric properties of the Japanese version GHQ-12*

Corrected item-whole correlations of the GHQ-12 are presented by sex in Table 1. The correlation coefficients ranged from 0.22 (item number 12) to 0.69 (9) for males (mean of the coefficients=0.49) and from −0.03 (12) to 0.69 (9) for females (mean=0.37). The values were also calculated by sex-age group divided into age categories by 10-year intervals. Among males, the values were about the same; mean value was 0.46 for both those aged 20-29 and 30-39 years and 0.52 for the other age-groups. Among females, the mean values ranged from 0.36 to 0.48. The correlations of the last item were smaller than those of the other items for most sex-age groups. Especially for females aged 20-29 and 40-49 years the values were negative.

The alpha coefficients of the scale are also presented in Table 1; the value was 0.82 for males and 0.73 for females. The values in the case when individual items were deleted were approximately equal to the value of the whole items; the range was from 0.79 to 0.84 for males and 0.66-0.78 for females. The values calculated by sex-age group also showed the similar levels: the value was 0.81 (range: 0.77-0.82) for males aged 20-29 and 30-39 years and 0.84 (0.81-0.86) for males aged 40-49 and 50-59 years; for females in the same age-groups the values were 0.72 (0.63-0.79), 0.85 (0.81-0.86), 0.72 (0.66-0.77), and 0.72 (0.65-0.74), respectively.

The internal consistency of the GHQ-12 was at the same level between males aged 20-29 years and 30-39 years and between those aged 40-49 years and 50-59 years. For females the consistency was higher for those aged 30-39 years than for the other age-groups and the latter groups showed about the same level.

A tendency for the internal consistency of the GHQ-12 to be higher for males than for females was observed. When the internal consistencies of the Japanese version of this sample were compared to those of the U. K. version of the U. K. employees, the levels for males in both countries were approximately the same,

Table 2. The loadings of the GHQ-12 items on the factors by sex: Varimax rotated factor pattern

Items	Factors					
	Males			Females		
	I	II	Communality	I	II	Communality
1	.31	.34	.21	.29	.36	.21
2	.57	.11	.33	.50	−.52	.25
3	.14	.59	.31	.06	.53	.29
4	.14	.62	.40	.08	.43	.19
5	.67	.17	.48	.58	.02	.34
6	.68	.13	.48	.71	−.03	.50
7	.19	.53	.32	.07	.59	.35
8	.17	.57	.35	−.00	.37	.14
9	.76	.22	.62	.77	.15	.62
10	.77	.15	.61	.78	.14	.62
11	.66	.17	.47	.70	.04	.50
12	.02	.45	.21	−.17	.43	.21
$\sum h^2$	3.035	1.814	4.848	2.910	1.310	4.220
(variance %)	(25.3%)	(15.1%)	(40.4%)	(24.3%)	(10.9%)	(35.2%)

Principal factor analysis was used.
Factor loadings greater than 0.350 are underlined.
$\sum h^2$: sum of squares of the factor loadings.
variance %: percentage of the total variance explained by the factor variance.

while those of Japanese females were lower than those of their U. K. counterparts[6]. As a whole, the internal consistency of the Japanese version GHQ-12 appears to be at a satisfying level for males, but for females its internal consistency is insufficient.

An initial (unrotated) principal factor solution of the GHQ-12 yielded two factors for each sex before the eigenvalue fell below 1.0. Thus, these factors were rotated. Table 2 indicates varimax rotated factor loadings of the GHQ-12 items by sex. Kaiser's measure of sampling adequacy (MSA)[24] values were of sufficient levels; the overall MSA value was 0.88 for males and 0.76 for females and no item had an MSA value below 0.5, which would be considered unacceptable for factor analysis[24] for both sexes. Although the values of females were lower than those of males, these values indicate the adequacy of the factor analysis performed in the present study. Conceptualization of the extracted factors was according to the items having greater loadings. Only loadings equal to or greater than 0.350 are regarded as significant. The factors derived from the GHQ-12 for males were as follows:

Factor 1: Psychological distress (Six items having significant loadings. This

factor explained 25.3% of total variance). There were high loadings on items concerned with "been losing confidence in yourself (item number 10)," "been feeling unhappy and depressed (9)," "felt you couldn't overcome your difficulties (6)," and "been thinking of yourself as a worthless person (11)." They were related to loss of confidnence. The items also included "felt constantly under strain (5)" and "lost much sleep over worry (2)." They are deemed to represent various aspects of psychological distress.

Factor 2: Social dysfunction (Five items, 15.1%). The items included "gelt capable of making decisions about things (4)," "felt that you are playing a useful part in things (3)," "been able to face up to your problems (8)," "been able to enjoy your normal day-to-day activities (7)," and "been feeling reasonably happy, all things considered (12)." They are related to social functioning.

For females the factor structure was the same as that of males, except for the allocation of the first item, which was "been able to concentrate on whatever you're doing."

No age difference of the factor structure was found either, except for the first item. The contents of the factors were not strongly effected by the item, since the loading of the first item was not always so much higher; it ranged from 0.19 to 0.43 for the first factor and from 0.30 to 0.41 for the second factor among four age-groups. For females, age difference of the factor structure was not examined, since the four age-groups consisted of too few subjects for factor analysis.

The authors compare the present factor structure with that of the Australian population reported by Worsley and Gribbin[25]. They derived three factors from the GHQ-12 in a sample Australian population; anhednia-sleep disturbance (allocated items were 1, 2, 5, 6, 9, 12), social functioning (1, 3, 4, 7, 8, 12), and loss of confidence (6, 9, 10, 11), from the first, second and third factor, respectively. A

Table 3. Comparison of factor structures derived from the GHQ-12

Worsley & Gribbin ('77)	The present study
Australia	Japan
both sexes householders	both sexes public officials
poor performance	Psychological distress
Social performance	Social dysfunction
Loss of confidence	
Total explained variance: 61.7%	35.2-40.2%

The column headings indicate the author, country, and nature of the sample for each study in order. Names of the derived factors appear below each column heading in order, beginning with the first factor. Total variance accounted for by each analysis is shown at the bottom of the table.

Both results were obtained by means of a principal factor analysis with varimax rotation.

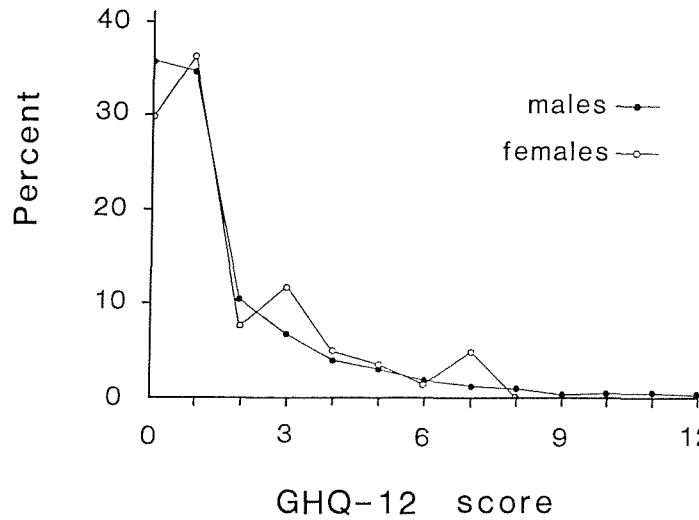

Figure 1. Percentage distribution of the GHQ-12 score in a sample of Japanese workers.

brief comparison of the present factor components with their study shows the following : (1) The items significantly loaded on the present first factor are allocated to both the first and third factors reported by Worsley and Gribbin. (2) The present second factor is equivalent to the second one reported by Worsley and Gribbin. Since fewer factors were derived from the GHQ-12 than those of Worsley and Gribbin, the contents of derived factors in the present study are more comprehensive than those of the Australian population.

(2) *Sex-age differences of the Japanese verison GHQ-12 score*

Table 4 shows the parameters of the score distribution and the proportion of subjects with a GHQ-12 score above the standard vut-off point of 2/3 by sex-age group. Sex and age distributions of the subjects are shown in the left column of Table 4. The Japanese verison GHQ-12 had a distribution considerably skewed to the right. Skewness ranged from 1.81 to 2.73 for males and from 1.02 to 1.65 for females. Score distributions of males were more skewed that those of females. These findings were the same as with the Japanese verison GHQ-60[26].

Females had a greater proportion of subjects with a GHQ-12 score above the cut-off point than males by χ^2 test, while no significant difference was observed for the mean score by the Mann-Whitney U-test. Differences were found for both mean score and the proportion among four age-groups of males by the Kruskal-Wallis H-test and χ^2 test. Among males, the older the subjects grew, the lower the scores became. Among females, this tendency was not found, though there were few subjects for each age-group. Females aged 30-39 years appeared to have more psychiatric symptoms than the other sex-age groups.

Table 4. Mean and standard deviation of the GHQ-12 and proportion of subjects with a score of 3 or more

Subjects		Mean	S. D.	Mode	Skewness	score>2 (%)
Whole sample	(n=2,070)	1.47	1.93	0	2.15	19.5%
Males total	(n=1,927)	1.46	1.94	0	2.21	18.9%
20 s	(n=430)	1.73	2.12	0	1.81	24.4%
30 s	(n=677)	1.46	1.81	1	2.01	19.8%
40 s	(n=288)	1.42	1.92	1	2.40	16.0%
50 s	(n=532)	1.26	1.93	0	2.73	15.0%
Females total	(n=143)	1.67	1.88	1	1.42	26.6%
20 s	(n=56)	1.79	2.06	1	1.49	26.8%
30 s	(n=19)	1.95	2.07	0	1.13	36.8%
40 s	(n=29)	1.55	2.05	0	1.65	20.7%
50 s	(n=39)	1.46	1.48	1	1.02	25.6%

a, b, c

S. D.: standard deviation.
score>2 (%): proportion of subjects with a score of 3 or more.
a: significance level of $p<0.0001$ by the Kruskal-Wallis H-test among four age-groups.
b: significance level of $p<0.001$ by the χ^2 test among four age-groups.
c: significance level of $p<0.05$ by the χ^2 test between sexes.

(3) *Relationships between the GHQ-12 factors and the THI*

As one hundred and fifty-seven males and 24 females with incomplete answers to the THI were omitted, this procedure was taken for the data of the remaining 1,770 males and 119 females. Table 5 presents Pearson's correlations of the GHQ-12 factors to the 12 scales and the two DF-values of the THI by sex. Correlations of these factors with the DF-value for SCHIZO were not calculated because of the higher probability of incorrect screening ratios[27].

For males, all of these correlations reached a high level of significance. For females, the second factor showed fewer significant correlations than that of males, partially because there were fewer subjects. As a whole, the correlation coefficients were higher for males than for females and the correlations of the first factor were always higher than those of the scond factor. The correlations of the "depressiveness" and "mental instability" scales and the DF-value for NEUR to the GHQ-12 factors were higher than those of the other THI scales and DF-values for both sexes.

According to partial correlations of the GHQ-12 factors to the THI scales and the DF-values, the "depressiveness" scale showed the highest correlations with the GHQ-12 factors among the THI scales (0.33 with the first factor and 0.20 with the second one for males, 0.11 and 0.32 respectively, for females), though its correlation with the first factor of females did not reach a significance level. The "mental instability" scale was significantly correlated with the first factor for both sexes; 0.13 for males and 0.33 for females. The DF-value for NEUR had significant correlations with the GHQ-12 factors, except for the second factor of females.

Table 5. Correlations between the GHQ and the THI

	GHQ factors			
	Males		Females	
	I	II	I	II
〈THI scale〉				
SUSY	.47[c]	.21[c]	.51[c]	.06
RESP	.28[c]	.13[c]	.36[c]	.07
EYSK	.38[c]	.13[c]	.31[b]	−.02
MOUT	.30[c]	.08[c]	.38[c]	.05
DIGE	.37[c]	.16[c]	.35[c]	.14
IMPU	.45[c]	.20[c]	.45[c]	.07
LISC	−.36[c]	−.24[c]	−.32[c]	−.19[a]
MENT	.58[c]	.23[c]	.64[c]	.21[a]
DEPR	.63[c]	.33[c]	.58[c]	.35[c]
AGGR	−.28[c]	−.12[c]	−.20[a]	−.05
NERV	.32[c]	.06[a]	.30[b]	.08
LIFE	.37[c]	.21[c]	.34[c]	.05
〈DF-value〉				
NEUR	.53[c]	.28[c]	.53[c]	.25[a]
PSD	.41[c]	.17[c]	.44[c]	.08

Males: n=1,770, Females: n=119.

a, b, c: significant correlation at the level of $p<0.05$, $p<0.01$, $p<0.001$ respectively.

Abbreviations. SUSY: many physical symptoms, RESP: respiratory organ complaints, EYSK: eye and skin complaints, MOUT: mouth and evacuation complaints, DIGE: digestive organ complaints, IMPU: impulsiveness, LISC: lie scale, MENT: mental instability, DEPR: depressiveness, AGGR: aggressiveness, NERV: nervousness, LIFE: irregularity of daily life, NEUR: discriminant-function value of neurosis, PSD: discriminant-function value of psychosomatic disease.

In summary, this study presents basic data about the psychometric properties and the score distribution of the Japanese verison GHQ-12 in a sample of Japanese workers. The internal consistency of the GHQ-12 was at a satisfying level. Factors derived from the GHQ-12 were labeled "psychological distress" and "social dysfunction" for both seves. Although a difference was found in the number of derived factors, the contents of the factors in this study were more comprehensive when compared to those of the Australian population. These results were consistent with those of the GHQ-60[19]. Pearson's correlation coefficients and partial correlations of the GHQ-12 factors to the THI scales and the DF-values of the THi indicated that the contents of the GHQ-12 were particularly correlated with other indices of depressive symptoms and neurotic illness, especially for the first factor labeled "psychological fistress." The GHQ-12 is considered to be a useful instrument which can measure neurotic symptoms in epidemiologic surveys.

The limitations of the present results are apparent, since there are few females

in this study. Thus, these findings are deemed to be applicable only to males. Further study based on more females is needed to examine the basic characteristics of the Japanese version GHQ-12.

Acknowledgements

The authors wish to acknowledge the club members of the "Hokkaido University Wandervögel Verein" for assistance in making the dataset analyzed in this study at the Hokkaido University Computing Center.

References

1) Goldberg, D. P. (1972): *The detection of psychiatric illness by questionnaire.* Maudsley Monograph No. 21. Oxford University Press: London. Translated In: Nakagawa, Y., ed. (1982). *The theory behind understanding psychiatric and neurotic symptoms using a questionnaire and its clinical applications,* Monograph of NIMH Japan, Part 1: Ichikawa. (in Japanese)
2) Goldberg, D. P. (1978): *Manual of the General Health Questionnaire.* NFER-Nelson Publishing Company: Windsor.
3) Goldberg, D. P. (1974): Psychiatric disorders, *Lancet* **23**, 1245-1247.
4) Gath, D., Osborn, M., Iles, S., Day, A., Bond, A. and Passingham, C. (1987): Psychiatric disorder and gynaecological symptoms in middle ages women: A community survey, *Brit. J. Psychiat.* **294**, 213-218.
5) Tarnopolsky, A., Hand, D. J., McLean, E. K., Roberts, H. and Wiggins, R. D. (1979): Validity and uses of a screening questionnaire (GHQ) in the community, *Brit. J. Psychiat.* **134**, 508-515.
6) Banks, M. H., Clegg, C. W., Jackson, P. R., Kemp, N. J., Stafford, E. M. and Wall, T. D. (1980): The use of General Health Questionnaire as an indicator of mental health in occupational studies, *J. Occup. Psychol.* **53**, 187-194.
7) Eagles, J. M., Craig, A., Rawlinson, F., Restall, D. B., Beattie, J. A. G. and Besson, J. A. O. (1987): The psychological well-being of supporters of the demented elderly, *Brit. J. Psychiat.* **150**, 293-298.
8) Tennant, C. (1977): The General Health Questionnaire: A valid index of psychological impairment in Australian populations, *Med. J. Australia* **2**, 392-394.
9) Mari, J. J. and Williams, P. (1984): Minor psychiatric discorder in primary care in Brazil: A pilot study, *Psychol. Med.* 14, 223-227.
10) Chan, D. W. and Chan, T. S. C. (1983): Reliability, validity and the structure of the General Health Questionnaire in a Chinese context, *Psychol. Med.* **13**, 363-371.
11) Mattila, M. J. and Salokangas, R. K. R. (1987): Mental health in the population approaching retirement age, *Acta Psychiat. Scand.* **75**, 195-201.
12) Harding, T. W. (1976): Validating a method of psychiatric case identification in Jamaica, *Bull. World Health Organ.* **54**, 225-231.
13) Munoz, P. E., Vazquez, J. L., Pastrana, E., Rodriguez, F. and Oneca, C. (1978): Study of the validity of Golderg's 60-item G. H. Q. in its Spanish version, *Soc. Psychiatry* **13**, 99-104.
14) Cheng, T. A. (1985): A pilot study of mental disorders in Taiwan, *Psychol. Med.* **15**, 195-24.
15) Radovanovic, Z. and Eric, L. (1983): Validity of the General Health Questionnaire in a Yugoslav student population, *Psychol. Med.* **13**, 205-207.
16) Nakagawa, Y. and Daibo, I. (1982): *Tests of the validity and reliability of the Japanese*

version General Health Questionnaire and its clinical applications. In: Nakagawa, Y., ed. *The theory behind understanding psychiatric and neurotic symptoms using a questionnaire and its clinical applications,* Monograph of NIMH Japan, Part 2: Ichikawa. (in Japanese)

17) Iwata, N. and Saito, K. (1987): Relationships of the Todai Health Index to the General Health Questionnaire and the Center for Epidemiologic Studies Depression Scale, *Jpn. J. Hyg.* **42**, 865-873. (with Japanese summary)

18) Aoki, S., Suzuki, S. and Yanai, H. (1974): A new trial of making a health and personality inventory, THPI, *Jap. J. Behaviormetr.* **2**, 41-53. (in Japanese with English summary)

19) Iwata, N., Okuyama, Y., Kawakami, Y. and Saito, K. (1988): Factor structure of the General Health Questionnaire in a sample of Japanese workers, *Ind. Health* **26**, 69-74.

20) Aoki, S. (1980): Study of the validity of the health questionnaire, THI: (1) The quantitative and qualitative differences of subjective complaints by sex and age, *Jpn. J. Hyg.* **34**, 751-765. (in Japanese with English summary)

21) Aoki, S. (1980): Study of the validity of the health questionnaire, THI: (2) The means of evaluation in the discriminative diagnosis of psychosomatic disease by THI, *Jpn. J. Hyg.* **34**, 766-776. (in Japanese with English summary)

22) Cronbach, L. J. (1951): Coefficient alpha and the internal structure of tests, *Phychometrika* **16**, 297-334.

23) Tsujioka, B. and Fujimura, K. (1975): Construction of personality questionnaire scales free from social desirability factors: Factor score method, *Jap. J. Educ. Psychol.* **23**, 206-212. (in Japanese with English summary)

24) Cenry, B. A. and Kaiser, H. F. (1977): A study of a measure of sampling adequacy for factor-analytic correlation matrices, *Multi. Behav. Res.* **12**, 43-47.

25) Worsley, A. and Gribbin, C. C. (1977): A factor analytic study of the twelve item General Health Questionnaire, *Australian and N. Z. J. Psychiatry* **11**, 269-272.

26) Iwata, N., Okuyama, Y., Kawakami, Y. and Saito, K. (1988): Psychiatric symptoms and related factors in a sample of Japanese workers, *Psychol. Med.* **18**. (in press)

27) Iwata, N., Okuyama, Y., Kawakami, Y. and Saito, K. (1987): An analysis of mental and somatic symptoms in public institution workers in Hokkaido: An epidemiological survey by means of the Todai Health Index, THI, *J. North. Occ. Health* **36**, 21-27. (in Japanese with English summary)

(Received 5 January 1988)