Usefulness of the Odor Stick Identification Test for Japanese Patients with Olfactory Dysfunction

Yoshiteru Hashimoto[1], Keijiro Fukazawa[1], Megumi Fujii[2], Sadamu Takayasu[1], Toshihiko Muto[3], Sachiko Saito[4], Yasuhiro Takashima[5] and Masafumi Sakagami[1]

[1]Department of Otorhinolaryngology, Hyogo College of Medicine, Nishinomiya City, Hyogo 663-8501, Japan, [2]Department of Otorhinolaryngology, Meiwa Hospital, 4-31 Agenaruo-cho, Nishinomiya-city, Hyogo 663-8186, Japan, [3]Department of Otorhinolaryngology, Hyogo Prefectural Kaibara Hospital, 1-5208 Kaibara, Kaibara-cho, Hikami-gunn, Hyogo 669-3395, Japan, [4]National Institute of Advanced Industrial Science and Technology, Neuroscience Research Institute, Central 6, 1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan and [5]Takasago International Corporation, Japan, Nissei Aroma Square 17F, 5-37-1 Kamata, Ota-ku, Tokyo 144-8721, Japan

Correspondence to be sent to: Yoshiteru Hashimoto, Department of Otolaryngology, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya City, Hyogo 663-8501, Japan. e-mail: y-h-1975@hyo-med.ac.jp

Abstract

The odor stick identification test (OSIT) is a new test of olfactory function recently developed for Japanese people. The purpose of the present study was to evaluate this test in relation to T&T olfactometry and the cross-cultural smell identification test (CC-SIT) by applying to 110 Japanese patients with olfactory disturbance. The averaged recognition thresholds for five odorants in T&T olfactometry, the number of correct answers in the CC-SIT and the rates of identification of 13 odorants in the OSIT were compared. The visual analogue scale (VAS) was also used to evaluate symptoms. The rate of identification of OSIT showed high and significant correlation coefficients with the averaged recognition thresholds of T&T olfactometry (-0.766, $P < 0.001$), with the number of correct answers in CC-SIT (0.754, $P < 0.001$) and with the VAS score (0.591, $P < 0.001$). In addition, on the identification performance measured by OSIT, we found significant differences between all pairs of four degrees of olfactory dysfunction except for one pair. Thus, we conclude that OSIT is useful for evaluating olfactory dysfunction in Japanese people.

Key words: cross-cultural smell identification test, odor stick identification test, olfactory disturbance, T&T olfactometry, visual analogue scale

Introduction

T&T olfactometry (Zusho, 1983) is the standard test used for measuring olfactory dysfunction in Japan. This test enables the measurement of olfactory detection and recognition thresholds and is the only method currently covered by Japanese health insurance (Miwa *et al.*, 2001; Fujii *et al.*, 2002a,b). Unfortunately, this test kit has not been well accepted for medical use because it requires a special large-scale deodorant device to avoid strong odor pollution and, moreover, measurement is a time-consuming process. Several other procedures have been presented and used worldwide to evaluate olfactory ability. The University of Pennsylvania smell identification test (UPSIT; Doty *et al.*, 1984) and the cross-cultural smell identification test (CC-SIT; Doty *et al.*, 1996) have been introduced as smell identification tests in the United States and the 'Sniffin' sticks' test has been used in Germany (Kobal *et al.*, 1996). However,

some odors included in these tests are unfamiliar to the Japanese and consequently the olfactory ability of Japanese subjects may not be measured correctly by these tests.

The odor stick identification test (OSIT) is a new type of smell identification test recently developed for the Japanese and can present 13 odors, all of which are familiar to the Japanese (Saito *et al.*, 2001, 2003). A significant correlation coefficient between the average recognition thresholds for T&T olfactometry and the rate of identification with OSIT has been reported in Japanese subjects with normal olfaction (Saito *et al.*, 2001). However, the clinical application of OSIT to patients with olfactory dysfunction has not yet been described. The purpose of the present study was to clarify the clinical utility of the OSIT for Japanese patients with olfactory dysfunction.

Materials and methods

This study was performed on 110 consecutive patients with olfactory dysfunction at the Department of Otolaryngology, Hyogo College of Medicine, from March 2001 to June 2002. The patients consisted of 44 men and 66 women with a mean age of 58.6 years. The patients were examined with the OSIT, T&T olfactometry and CC-SIT. Symptoms of olfactory disturbance were evaluated with the visual analogue scale (VAS; McCormack *et al.*, 1988). The causative diseases were chronic sinusitis in 38 cases, upper respiratory infection in 31 cases, nasal allergy in 10 cases, rhinitis with inflammation of the olfactory cleft in six cases, head trauma in four cases, acute sinusitis in one case, subarachnoid hemorrhage in one case and unknown etiology in 19 cases.

Odor Stick Identification Test (OSIT)

This test includes 13 odorants: perfume, rose, milk, Japanese orange, curry, roasted garlic, putrid smell, fermented beans/sweat socks, gas for cooker, menthol, Indian ink, wood and Japanese cypress (*hinoki*). These odors were chosen from odors familiar to the Japanese. The odorant for each odor was selected from pure chemicals, essential oils, or mixed odorants produced by the Takasago International Corporation. These odorants were packed in microcapsules, which were mixed in paste. Each paste was hardened in the form of a lipstick, which was called the 'odor stick' (Figure 1). Each odor stick was painted in a 2 cm circle on a thin paraffin paper by the examiner. The examiner folded this paper in half and rubbed it to grind the microcapsules and then passed it to a patient. The patient next opened and sniffed the paper. The patient was then given six alternatives: four odor names, including one correct name in principle, 'not detected' and 'unknown'; the odorant of fermented beans/sweat socks was an exception. First, we asked patients to choose a correct one out of the four odor names. We suggested that patients select 'unknown' or 'not detected', when it was difficult to select one from the four odor names.
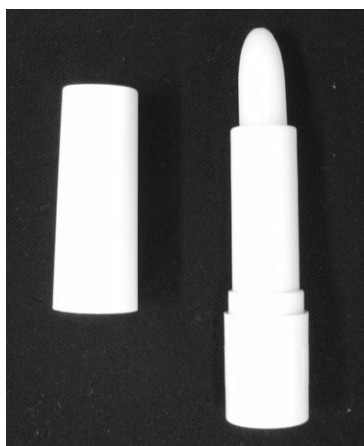

Figure 1 Example of the odor stick.

'Unknown' indicates that the presented odorant was detected but not recognized. This procedure was repeated for each odorant. The original method proposed using an answer sheet showing the difference between four alternatives of odor names and the two other ones, 'unknown' and 'not detected', by drawing a broken line between four alternatives of odor names and others or by putting others in parentheses. In the present study, we did not draw a broken line, but explained the difference verbally.

T&T olfactometry

T&T olfactometry is the standard Japanese olfactory test and is routinely performed clinically (Zusho, 1983). This test consists of five standard odors: (A) roses, (B) burning, (C) sweat, (D) fruits and (E) vegetable chips. The concentration is prepared at eight degrees, except (B), which is prepared at seven degrees (Zusho, 1983; Fujii *et al.*, 2002a,b). This test measures the detection and recognition thresholds. It is possible to evaluate subjects' olfactory ability as the average score of five odorants. In this study, the average recognition threshold was defined based on the result of T&T olfactometry by the mean recognition thresholds for five odorants (Takagi, 1987).

Cross-cultural smell identification test (CC-SIT)

The cross-cultural smell identification test (CC-SIT) consists of twelve odorants, based on items from the University of Pennsylvania smell identification test (UPSIT). The 12 items (odorants) were selected as multicultural odorants from the UPSIT. The 'scratch-and-sniff' technique employed was based on mechanical release of micro-encapsulated odorants. Smell identification was based on a multiple choice test from a list of four items (Doty *et al.*, 1996).

Visual analogue scale (VAS)

The visual analogue scale (VAS) has been used in psychological research since the early twentieth century (Hayes and Patterson, 1921; Freyd, 1923). In the present study, we used VAS to measure the symptoms of olfactory dysfunction. VAS consists of a 10 cm line, both ends of which have statements of the maximal and minimal extremes (McCormack *et al.*, 1988). The subjects were asked to indicate their feelings by marking the line at the appropriate point between the two extreme statements, defined as 'anosmia' and 'normal' in this study. The scales were scored by measuring the distance from the minimal endpoint to the mark, on a predetermined measurement interval. The most commonly chosen interval was in millimeters with a 10 cm line, producing a 100 point scale (McCormack *et al.*, 1988).

Procedure

OSIT was performed by either of two of the eight authors. T&T and VAS were performed by a well-trained assistant. The OSIT and CC-SIT did not require a deodorant device and did not cause odor pollution. On the other hand, a

deodorant device was required for T&T olfactometry. About 10 min were required to examine olfaction with OSIT, 12 min with the CC-SIT and 25 min with T&T olfactometry.

Data analysis

We calculated Pearson's correlation coefficients between the each pair among the OSIT identification rate, the average recognition threshold by T&T olfactometry, CC-SIT score and VAS score. We also evaluated OSIT identification rate and CC-SIT score according to the degree of olfactory disturbance measured by T&T olfactometry, with one-way analysis of variance (ANOVA) and the post-hoc Tukey–Kramer test. In all cases, P-values below 0.05 were regarded as statistically significant.

Results

Correlation among the results of OSIT, T&T olfactometry and CC-SIT

The correlations were –0.766 between the identification rate of OSIT and the average recognition threshold by T&T olfactometry ($P < 0.001$, $n = 110$; Figure 2A), –0.673 between CC-SIT score and the average recognition

threshold by T&T olfactometry ($P < 0.001$, $n = 90$; Figure 2B) and 0.754 between the identification rate of OSIT and CC-SIT score ($P < 0.001$, $n = 90$; Figure 2C). These findings suggested that the three tests were all significantly correlated.

Correlation between the results of VAS and each test

The correlations were 0.591 between the identification rate of OSIT and VAS score ($P < 0.001$, $n = 110$; Figure 3A), –0.479 between the average recognition threshold by T&T olfactometry and VAS score ($P < 0.001$, $n = 110$; Figure 3B) and 0.555 between CC-SIT score and VAS score ($P < 0.001$, $n = 90$; Figure 3C).

The relations of OSIT to the degree of olfactory dysfunction

The degree of olfactory acuity was determined, for each participant, based upon the average recognition thresholds of five odorants assessed by T&T olfactometry. The average thresholds of 1.1–2.5 were defined as the 'mild' degree of olfactory dysfunction, 2.6–4.0 as 'moderate', 4.1–5.5 as 'serious' and 5.6–5.8 as 'anosmia'. ANOVA revealed a significant difference of the identification performance among the degrees of olfactory dysfunction in both OSIT
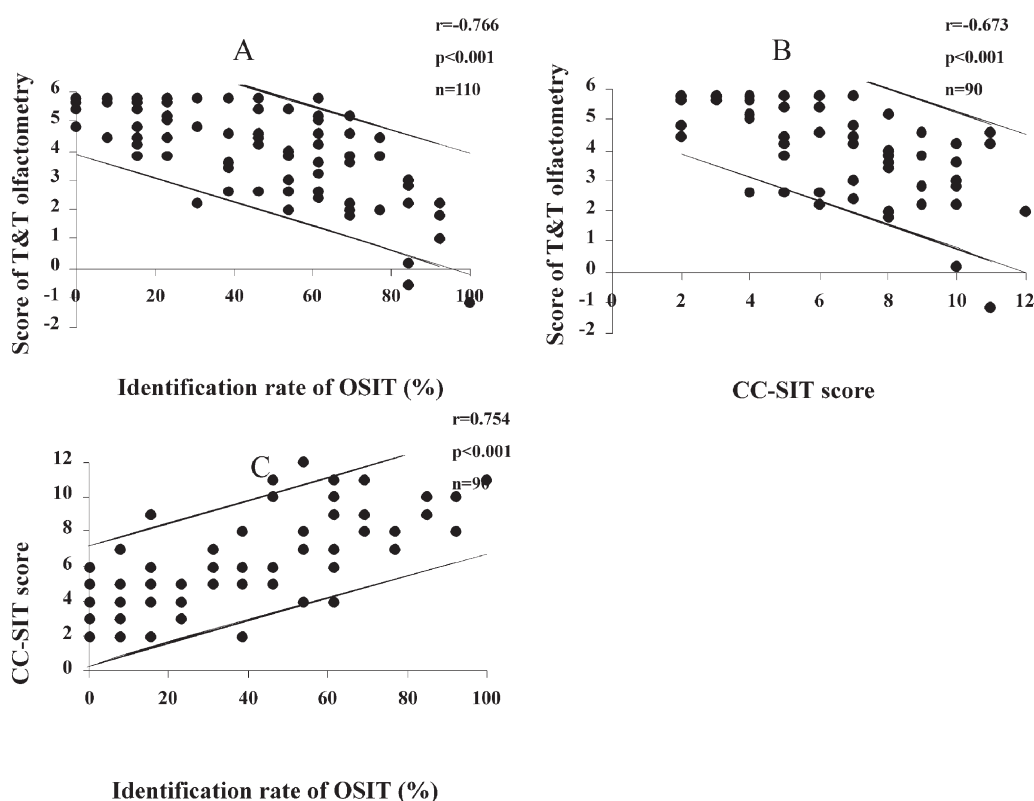
Figure 2   **(A)** Correlation between the results of OSIT and T&T olfactometry. The correlation coefficient between the results of OSIT and T&T olfactometry was –0.766 ($P < 0.001$, $n = 110$). The upper and lower lines represent 95% prediction intervals for the normal data. **(B)** Correlation between the results of CC-SIT and T&T olfactometry. The correlation coefficient between the results of CC-SIT and T&T olfactometry was –0.673 ($P < 0.001$, $n = 90$). The upper and lower lines represent 95% prediction intervals for the normal data. **(C)** Correlation between the results of OSIT and CC-SIT. The correlation coefficient between the results of OSIT and CC-SIT was 0.754 ($P < 0.001$, $n = 90$). The upper and lower lines represent 95% prediction intervals for the normal data.
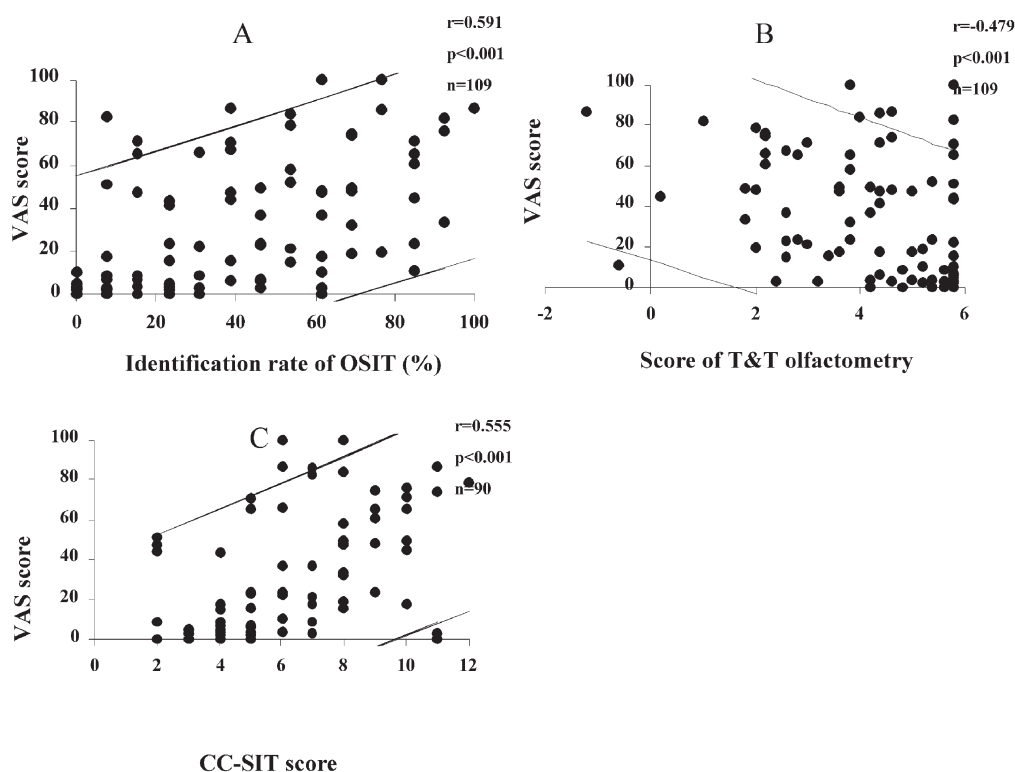
Figure 3 (A) Correlation between the results of OSIT and VAS. The results of OSIT and VAS were significantly correlated ($r = 0.591$, $P < 0.001$, $n = 110$). The upper and lower lines represent 95% prediction intervals for the normal data. **(B)** Correlation between the results of T&T olfactometry and VAS. The results of T&T olfactometry and VAS were significantly correlated ($r = -0.479$, $P < 0.001$, $n = 110$). Upper and lower lines represent 95% prediction intervals for normal data. **(C)** Correlation between the results of CC-SIT and VAS. The results of CC-SIT and VAS were significantly correlated ($r = 0.555$, $P <0.001$, $n = 90$). The upper and lower lines represent 95% prediction intervals for the normal data.

$[F(3, 102) = 105, P < 0.01]$ and CC-SIT $[F(3, 84) = 87, P < 0.01$; Figure 4A,B]. A Tukey–Kramer's post-hoc test revealed significant differences between all pairs of four degrees of olfactory dysfunction except for one pair ('mild' and 'moderate') in the identification performance of OSIT (Figure 4A); however, a Tukey–Kramer's post-hoc test could not reveal significant differences between 'mild' and 'moderate' and between 'moderate' and 'serious' in the CC-SIT score (Figure 4B).

The rate of correct identification of each odorant in OSIT

We calculated the rate of identification of each odorant in the 'mild' and 'moderate' groups of the degree of olfactory dysfunction in order to examine the validity of odorants. The mean rate of identification of 13 odorants was 60.5% and the rates of identification of two odorants, Japanese orange and gas for cooker, were <45.4% (mean rate – SD; Figure 5).

The rate of selected answers for each odorant in OSIT in the 'mild' and 'moderate' groups

The rates of selected answers for each odorant in OSIT in the 'mild' and 'moderate' groups are shown in Figure 6. The rates of 'unknown' responses for the two odorants (Japanese

orange and gas for cooker) were 23.3 ± 0% (average ± SEM) and the rates for the other odorants were 10.9 ± 1.89%. The two odorants, whose correct answer rates in OSIT were lower than 'the mean rate – SD' in the 'mild' and 'moderate' groups, showed higher rate of 'unknown' responses than the that for any one of the other odorants; however, we did not apply statistics because of the small sample size of one group.

Discussion

The only test of olfaction currently recognized by Japanese national health insurance is T&T olfactometry. However, this test has not been widely used in Japan for three main reasons. First, T&T olfactometry requires large-scale ventilation equipment to avoid strong odor pollution. Secondly, this test usually presents only five odors, which is insufficient to evaluate odor identification ability. Lastly, this test takes a relatively long time (~30 min). OSIT, on the other hand, does not require the use of a deodorant device, examines 13 kinds of odorants familiar to the Japanese in their daily life and only requires ~10 min to perform. Therefore, we consider OSIT to be a more efficient test for examining olfactory ability in routine daily clinical practice.
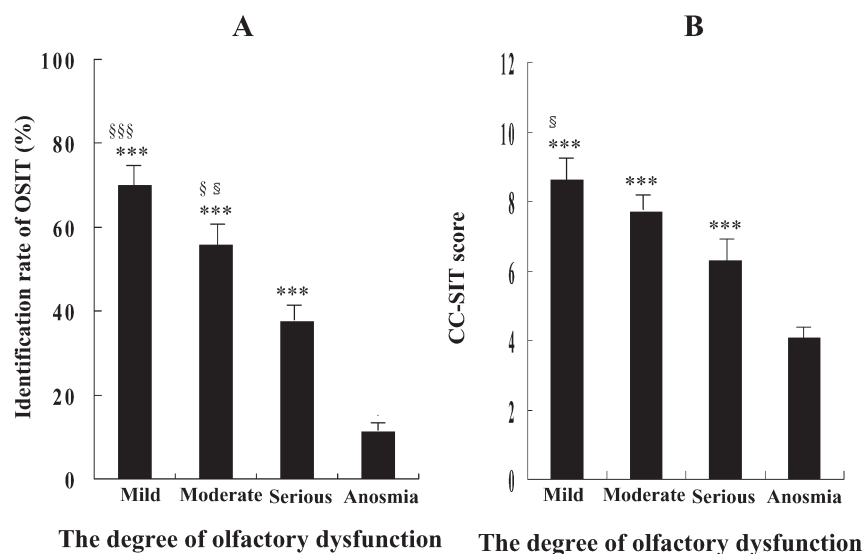
A

B

Figure 4 **(A)** The mean rate of identification (+SE) of OSIT for 'mild ($n = 10$)', 'moderate (n = 20)', 'serious ($n = 28$)' and 'anosmia ($n = 48$)' groups in the degree of olfactory dysfunction. ANOVA showed a significant main effect of the condition of olfactory dysfunction, a post-hoc test showed significant differences between all pairs of four degrees of olfactory dysfunction except for one pair ('mild' and 'moderate'). ***Significant difference with 'anosmia' by $P < 0.001$. Significant diference with 'serious' by §§§$P < 0.001$ and §§$P < 0.01$. **(B)** Averaged CC-SIT score (+SE) for 'mild ($n = 8$)', 'moderate (n = 17)', 'serious ($n = 23$)' and 'anosmia ($n = 40$)' groups in the degree of olfactory dysfunction. ANOVA showed a significant main effect of the condition of olfactory dysfunction and a post-hoc test showed a significant difference of identification between all pairs of four degrees of olfactory dysfunction except for two pairs ('mild' and 'moderate', 'moderate' and 'serious'). §Significant difference with 'serious' by $P < 0.05$.

OSIT is a smell identification test, while T&T olfactometry measures smell recognition threshold and detection threshold simultaneously. However, the high and significant correlation was observed between the identification rate of OSIT and the average recognition threshold by T&T olfactometry and between the rate of identification by OSIT and the VAS score. These results support the notion that OSIT well reflects subjective severity of olfactory disturbance.

There were significant differences among the each rate of identification by OSIT in the degrees of olfactory disturbance defined by T&T olfactometry without the one pair 'mild' and 'moderate' (Figure 4). A sample size of 'mild' was smaller than that of the other degrees, therefore, one needs to consider the significant difference between the 'mild' and 'moderate' with the larger sample size of 'mild'. On the other hand, there were no significant differences not only between the 'mild' and 'moderate', but also between the 'moderate' and 'serious' olfactory disturbances in the CC-SIT score (Figure 4). We attribute this discrepancy to differences in used odorants and alternatives to be answered. Familiar odorants for Japanese are used in OSIT. As the alternatives to be answered, CC-SIT is a forced multiple choice test from a list of four items, while OSIT includes the additional choices of 'not detected' and 'unknown'. Thus, OSIT can probably estimate the olfactory dysfunction of a patient more accurately than CC-SIT and can be used not only as a screening test for olfactory dysfunction, but also to measure the degree of olfactory ability.
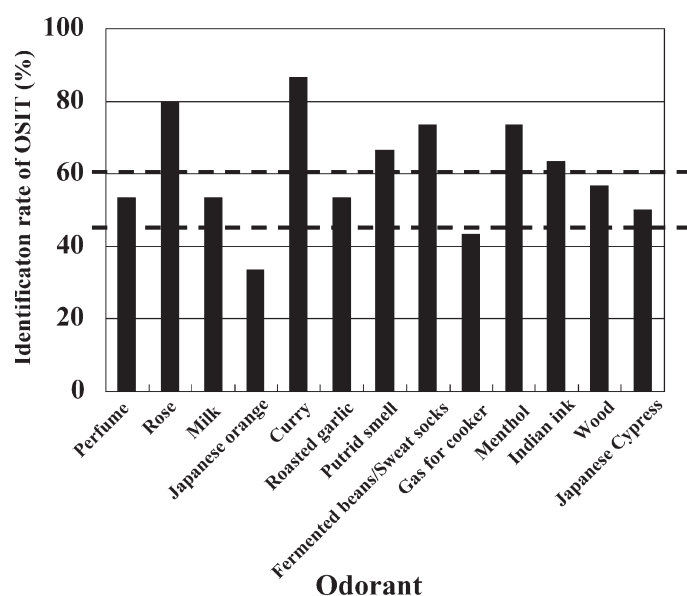

Figure 5 The mean rate of identification of each odorant by OSIT in the integrated group of 'mild' and 'moderate' ($n = 30$). The upper horizontal dashed line indicates the averaged mean rate (60.5%) and the lower line indicates the averaged mean rate minus standard deviation (SD), that is 45.4%. The rates of identification of two odorants ('Japanese orange' and 'gas for cooker') were <45.4% in the groups of 'mild' and moderate'.

The rates of identification in OSIT were approximately in accordance with the degree of olfactory acuity (Figure 4). However, two particular odorants, 'Japanese orange' and
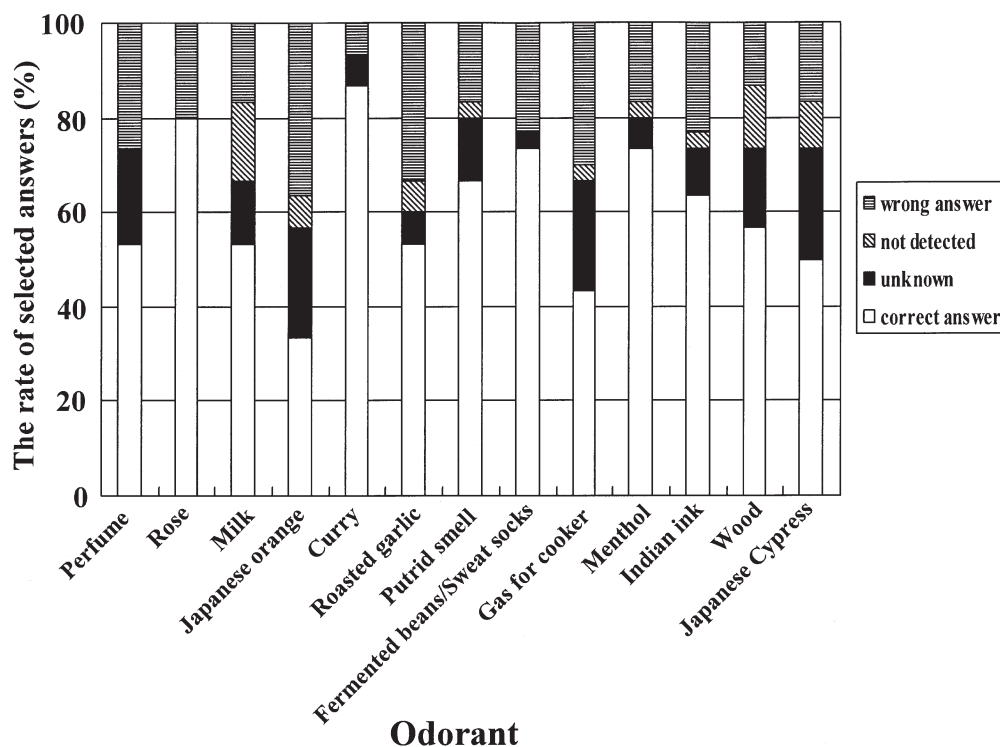
Figure 6 The rate of selected answers by patients for each odorant in OSIT for the integrated group of 'mild' and 'moderate' (*n* = 30). The rate of the 'unknown' answer for each odor was relatively high in two odorants: 'Japanese orange' and 'gas for cooker'.

'gas for cooker', the rates of identification were lower and the rates of 'unknown' response were higher than that of any of the other 11 odorants (Figures 5 and 6). This may suggest that patients tend to choose the 'unknown' response when it is difficult to select the correct name of odorant. The subjective intensity of each odorant in OSIT was fitted approximately from moderate to strong intensity and the percent correct of identification was >80% in normal subjects, but the identification rate of 'Japanese orange' was the lowest (Saito *et al.*, 1998, 2003). Although the 13 odorants in OSIT were familiar to the Japanese, we think it was still difficult for patients with olfactory dysfunction to identify these two odorants. In the case of 'gas for cooker', however, we need to consider also the tendency of its odor to fade away immediately when the paraffin paper is opened.

Conclusion

OSIT was an easy, fast and compact olfactory test suitable for clinical use in screening for olfactory disturbance. OSIT could measure the degree of olfactory disturbance similar to T&T olfactometry. Patients with olfactory dysfunction may find it difficult to identify two odorants, 'Japanese orange' and 'gas for cooker'. OSIT may be adapted for patients in other countries by choosing suitable, recognizable odors.

Acknowledgements

References

Doty, R.L., Shaman, P., Kimmelman, C.P. and Dann, M.S. (1984) *University of Pennsylvania smell identification test: a rapid quantitative olfactory function test for clinic*['clinic' OK?]. Laryngoscope, 94, 176–178.

Doty, R.L., Marcus, A. and Lee, W.W. (1996) *Development of the 12-items cross-cultural smell identification test (CC-SIT)*. Laryngoscope, 106, 353–356.

Freyd, M. (1923) *The graphic rating scale*. J. Educat. Psychol., 14, 83–102.

Fujii, M., Fukazawa, K., Takayasu, S. and Sakagami, M. (2002a) *Olfactory dysfunction in patients with head trauma*. Auris Nasus Larynx, 27, 35–40.

Fujii, M., Fukazawa, K., Hatta, C., Yasuno, H. and Sakagami, M. (2002b) *Olfactory acuity after total laryngectomy*. Chem. Senses, 27, 117–121.

Hayes, M.H. and **Patterson, D.G.** (1921) *Experimental development of the graphic rating method*. Psychol. Bull., 18, 98–99.

Kobal, G., Hummel, Th., Sekinger, B., Barz, S., Roscher, S. and Wolf, S. (1996) *'Sniffin' sticks': screening of olfactory performance*. Rhinology, 34, 222–226.

McCormack, H.M., Horne, D.J.L. and Sheather, S. (1988) *Clinical application of visual analogue scales: a critical review*. Psychol. Med., 18, 1007–1019.

Miwa, T., Furukawa, M., Tsukatani, T., Costanzo, R.M., DiNardo, L.J. and Reiter, E.R. (2001) *Impact of olfactory impairment on quality of life and disability.* Arch. Otolaryngol. Head Neck Surg., 29, 35–40.

Saito, S., Tsuchiya, N., Mise, M., Yoshida, Y., Kobayakawa, T., Ayabe, S., Yamaguchi, Y. and Takashima, Y. (1998) *A smell test for Japanese people: a study using a new stick-type tool for odor presentation.* Jpn. J. Taste Smell Res., 5, 323–326 [in Japanese].

Saito, S., Masuda, Y., Kobayakawa, T., Gotow, N., Mizoguchi, C. and Takashima, Y. (2001) *Relation between the Japanese standardized olfactory test and the odor stick identification test for Japanese in* *108 Japanese participants aged 20 to 81 years*. Jpn. J. Taste Smell Res., 8, 143–149 [in Japanese].

Saito, S., Ayabe-Kanamura, S., Naito, N., Gotow, N., Kobayakawa, T., Mise, M. and **Takashima, Y.** (2003) *Development of an odor identification test for Japanese people: verification of stick type and card type.* J. Jpn. Assoc. Odor Environ., 34, 1–6 [in Japanese].

Takagi, S.F. (1987) *A standardized olfactometer in Japan. A review over ten years*. Ann. N. Y. Acad. Sci., 510, 113–118.

Zusho, H. (1983) *Olfactometry in Japan.* Rhinology, 21, 281–285.