Rhinology, 41, 2-6, 2003

Quantitative assessment of gustatory function in a clinical context using impregnated "taste strips"*

C. Mueller[1,2], S. Kallert[1], B. Renner[1], K. Stiassny[2], A.F.P. Temmel[2], T. Hummel[3], G. Kobal[1]

[1] Department of Pharmacology, University of Erlangen-Nuremberg, Krankenhausstr. 9, D-91054 Erlangen, Germany
[2] Department of Otorhinolaryngology, University of Vienna, AKH Wien, Währinger Gürtel 18-20, A-1090 Wien, Austria
[3] Smell and Taste Clinic, Department of Otorhinolaryngology, University of Dresden Medical School, Fetscherstr. 74, 01307 Dresden, Germany

SUMMARY

Assessment of gustatory sensitivity in a clinical setting is the prerequisite for correct diagnosis and adequate treatment of taste dysfunction. Despite of this, no taste test has been established for the routine clinical testing. The aim of the present study was to create a protocol which is easy to administer. The presently used technique is based on strips made from filter paper which were impregnated with different taste solutions (four concentrations each for sweet, sour, salty and bitter). These strips are placed on the tongue and subjects are asked to identify the taste quality. After establishing the concentration range of the taste solutions, the test was tried in 69 subjects. Each subject received eighteen taste strips (four concentrations of each taste quality plus two blanks) in a pseudo-randomized sequence. Results from this new procedure correlated significantly with the results of the well established extensive three-drop-technique (r_{69}=0.67). Repeated measures indicated good reproducibility of the results for the taste strips (r_{69}=0.68). These data suggest the usefulness of this new technique in routine clinical practice. Major advantages are long shelf-life, convenience of administration, short time needed for testing (approximately 8 min), and the possibility to test each side of the tongue separately.

Key words: taste strips, taste, gustation, quantitative, clinical

INTRODUCTION

Most patients suffering from disturbances of smell function only, report both smell and taste dysfunction prior to diagnosis (Deems et al., 1991). Because most people are not aware that flavor perception is largely mediated by retronasal stimulation of olfactory receptors during mastication and deglutition, sometimes hyposmia/anosmia is ascribed to losses in taste only. This confusion of gustatory and olfactory mediated sensations makes it often necessary to test both smell and taste function in order to reach an accurate diagnosis of chemosensory disturbances. However, in contrast to well established methods for the investigation of the first cranial nerve (Doty et al., 1984; Hummel et al., 1997a; Kobal et al., 2000) no quantitative test with different taste concentrations exists for the clinical assessment of gustatory function. Although the three-drop-method (Henkin et al., 1963) could be administered it has not been established in clinical routine due to certain disadvantages: the procedure is time-consuming, requires specially trained staff, the solutions need to be freshly prepared, and the taste solution is dissolved and diluted immediately after it is dropped on the surface of the tongue, which makes it difficult to detect localized losses of taste perception. In many clinical situations patients are only asked about their taste sensitivity or tested with a supra-threshold concentration of each taste quality. This lack of quantitative taste testing also results in a lack of the tracking of taste dysfunction over time. In other words, for the routine clinical testing a taste test would be needed which (1) has a long shelf-life, (2) is rigourously validated, and (3) is sufficiently easy to administer.

Thus, the aim of the present study was to create a protocol which would fulfil the criteria mentioned above. The presently used technique is based on strips made from filter paper which are soaked with different taste solutions (four concentrations each for sweet, sour, salty and bitter) and dried. This should

* Received for publication: April 15, 2002; accepted: July 17, 2002

result in a profile of taste perception in a time of not more than ten minutes. For validation the correlation with an established procedure (three-drop-method; Henkin et al., 1963) was used; tests were carried out at two different days in order to determine the test-retest-reliability. Finally normative data for the interpretation of the results of the taste test should be obtained.

MATERIAL AND METHODS

Human subjects

The study was carried out at the Departments of Pharmacology at the University of Erlangen-Nuremberg, Germany, and Otorhinolaryngology at the University of Vienna, Austria, between May 2001 and February 2002. It was conducted according to the guidelines of the Declaration of Helsinki on biomedical research involving human subjects. The investigation involved 69 participants (36 female, 33 male, mean age 29 years, range 15-75 years; 15-34 years: 57 subjects, 35-54: 10 subjects, 55-75: 2 subjects). One hour prior to testing subjects were asked not to eat or drink anything except water, not to smoke, and not to brush teeth. Subjects with diseases which might affect taste perception (Schiffman, 1983), such as endocrine disturbances (thyroid gland dysfunction, diabetes mellitus, Cushing`s syndrome), internal diseases (chronic renal failure, cirrhosis of the liver), middle ear affections, xerostomia or depression were not eligible for the study.

Gustatory function

For the assessment of gustatory function taste strips were used which are made of paper soaked in taste solutions and dried on a slowly rotating wheel. The length of a taste strip is 8 cm and an area of 2 cm^2 is impregnated with a taste stimulant. Experiments in highly trained observers from our laboratory (who were familiar with chemosensory perception) helped to establish those concentrations which were perceived as equi-intense for the three-drop-method and the taste strips. The lowest concentrations of each taste quality should be identified by half of the healthy subjects only; the highest concentration should be identified by approximately 100% of the subjects. The following concentrations were used for the taste strips: sweet: 0.4, 0.2, 0.1, 0.05 g/ml sucrose; sour: 0.3, 0.165, 0.09, 0.05 g/ml citric acid; salty: 0.25, 0.1, 0.04, 0.016 g/ml sodium chloride; bitter: 0.006, 0.0024, 0.0009, 0.0004 g/ml quinine-hydrochloride. For the three-drop-method the following concentrations were used: sweet: 0.4, 0.2, 0.1, 0.05 g/ml sucrose; sour: 0.075, 0.041, 0.0225, 0.0125 g/ml citric acid; salty: 0.25, 0.1, 0.04, 0.016 g/ml sodium chloride; bitter: 0.0015, 0.0006, 0.0002, 0.0001 g/ml quinine hydrochloride.

To obtain a measure of test-retest-reliability, the same tests (three-drop-method, and taste strips) were administered on two days separated by a mean interval of 12 days.

Three-drop-method

Using a 10µl pipette three drops of liquid were placed at the middle of the tongue at a distance of approximately 1.5 cm from the tip. As with administration of the taste strips (see below), subjects were then allowed to close the mouth. One drop contained a taste solution and the two others solvent only (distilled water); the sequence of administration was randomized across trials. Testing started with the lowest concentration. The subjects' task was to identify the drop which contained taste solution, and to indicate the taste quality. Then the mouth was rinsed with a sip of tap water. Using the mehod of ascending limits the threshold was noted as the concentration step which had been identified in 3 consecutive trials. Thus, scores for each taste qualities ranged between 0 and 4. The result for the entire test was the sum of the results for individual taste qualities (range 0 to 16).

Taste strips

As with the three-drop-method (see above) four concentrations were used for each taste quality resulting in a maximum total score of 16, and 4 for each taste quality. The taste strips (plus two blank strips without taste) were presented in increasing concentrations in a randomized order (see Table 1) and placed in approximately the same position where the drops were applied. Then subjects were asked to close the mouth and choose one of five possible answers on a form (sweet, sour, salty, bitter, no taste). Before assessment of each taste strip the mouth was rinsed with water.

Statistical analysis

SPSS 10.0 was employed for statistical evaluation. Correlational analyses were performed using Spearman statistics. The alpha level was set at 0.05. "Bland & Altman plots" (Bland and Altman, 1999) were used for graphical presentation of the repeatability of the two methods.

RESULTS

Mean values of the taste strips and the 3-drop-method derived from the 69 healthy subjects are presented in Table 2 (see also

1	2	1	2
Sweet 4	Sour 4	Salty 2	Bitter 2
Bitter 4	Salty 4	Sour 2	Sweet 2
Salty 4	Bitter 4	Sweet 2	Sour 2
Sour 4	Sweet 4	Sour 1	blank
Sour 3	Salty 3	blank	Salty 2
Sweet 3	blank	Salty 1	Sweet 1
blank	Bitter 3	Bitter 2	Sour 1
Bitter 3	Sweet 3	Bitter 1	Salty 1
Salty 3	Sour 3	Sweet 1	Bitter 1

Table 1. Sequences in which taste strips were applied. Each of the two sequences of eighteen taste strips (four concentrations of each taste quality plus two blanks) were applied in a pseudo-randomized order starting with the lowest concentration (1=highest concentration, 4=lowest concentration). Subjects had to choose between "sweet", "sour", "salty", "bitter", and "no taste". Before administration of each strip subjects took a sip of cool tap water.

Taste strips	mean	SD
Sweet	3.3	0.8
Sour	3.0	0.8
Salty	3.1	0.9
Bitter	3.0	1.1
Total score	12.4	2.3
3-drop-method	mean	SD
Sweet	3.5	0.8
Sour	3.5	0.8
Salty	3.5	1.0
Bitter	3.0	1.1
Total score	13.5	2.6

Table 2. Mean values and standard deviations (SD) of the test results separately for taste strips and the 3-drop-method. The maximum score was 4 for each taste quality and 16 for the total score (n=69).

Taste strips Percentile	Sweet	Sour	Salty	Bitter	Total score
5th	2.0	1.5	1.0	1.0	8.5
10th	2.0	2.0	2.0	1.0	9.0
50th	4.0	3.0	3.0	3.0	13.0
90th	4.0	4.0	4.0	4.0	15.0
95th	4.0	4.0	4.0	4.0	15.0
3-drop-method Percentile	Sweet	Sour	Salty	Bitter	Total score
5th	1.5	2.0	1.0	0.5	7.5
10th	2.0	2.0	2.0	2.0	9.0
50th	4.0	4.0	4.0	3.0	14.0
90th	4.0	4.0	4.0	4.0	16.0
95th	4.0	4.0	4.0	4.0	16.0

Table 3. Normative values derived from 69 healthy volunteers for taste strips and the 3-drop-method. For both tests the maximum score was 4 for each taste quality and 16 for the total score. Results below the 10th percentile yield hypogeusia (compare Doty et al., 1984; Kobal et al., 2000).

Figure 1). Normative values are listed in Table 3. In terms of the definition of hypogeusia the 10th percentile may be used to separate normogeusic from hypogeusic subjects (compare Doty et al., 1984; Kobal et al., 2000).

The percentage of correctly identified taste strips for the highest concentrations was 100% for sweet, 99% for sour, 96% for salty, and 99% for bitter (see Table 4). For the lowest concentrations it was 54% for sweet, 36% for sour, 51% for salty, and 52% for bitter.

Analyses revealed significant correlations between the results of the three-drop-method and the taste strip test. The coefficient of correlation of the total scores was r_{69}=0.67; for indvidual taste qualities it was r_{69}=0.35 (sweet), r_{69}=0.34 (sour), r_{69}=0.42 (salty), and r_{69}=0.54 (bitter) (for all correlations p<0.01).

For the taste strips the correlation coefficient for test and retest of the total scores was r_{69}=0.68; for individual taste qualities it was r_{69}=0.43 (sweet), r_{69}=0.40 (sour), r_{69}=0.34 (salty), and r_{69}=0.56 (bitter) (for all correlations p<0.01). For the 3-drop-

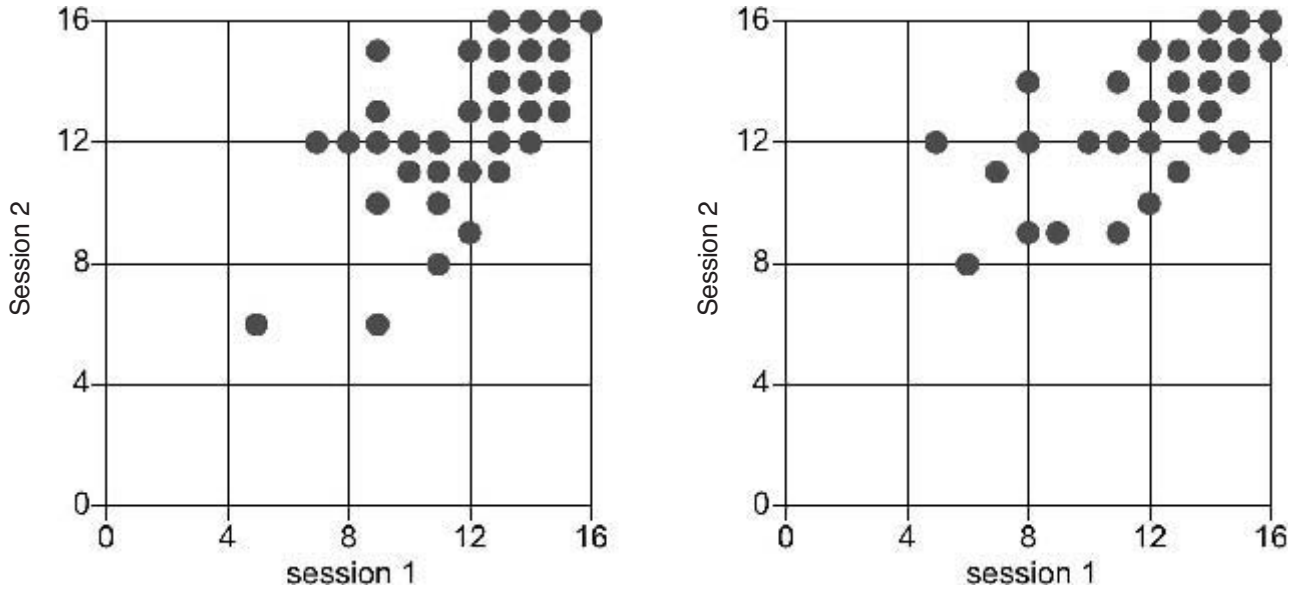

Figure 1. Results for taste tests obtained by means of taste strips (left) or the 3-drop-method ("taste drops"). Scores from session 1 are plotted against scores from session 2.

	correct	no taste	Sweet	Sour	Salty	Bitter
Sweet 1	100.0	0		0	0	0
Sweet 2	94.1	4.4		0	0	1.5
Sweet 3	85.4	11.6		0	1.5	1.5
Sweet 4	53.5	34.8		4.4	4.4	2.9
Sour 1	98.5	0	0		1.5	0
Sour 2	91.3	0	0		2.9	5.8
Sour 3	78.2	4.4	0		10.1	7.3
Sour 4	36.2	31.9	2.9		20.3	8.7
Salty 1	95.6	0	0	4.4		0
Salty 2	88.3	1.5	0	7.3		2.9
Salty 3	81.1	2.9	0	7.3		8.7
Salty 4	50.7	21.7	1.5	20.3		5.8
Bitter 1	98.5	1.5	0	0	0	
Bitter 2	85.4	10.1	1.5	1.5	1.5	
Bitter 3	66.6	24.6	1.5	4.4	2.9	
Bitter 4	52.1	40.6	0	4.4	2.9	
Blank	65.2		2.9	13.0	7.3	11.6

Table 4. Results listed for each of the sixteen concentrations of taste strips from 69 healthy subjects. Taste concentrations are numbered as follows: 1=highest, 4=lowest concentration. The first column ("correct") shows the percentage of correctly identified taste strips. Note that the lowest concentration of each taste quality (e.g. sour 4 or salty 4) is not identified by approximately half of the healthy subjects. The second column ("no taste") holds the percentage of taste strips which had been mistakenly identified as having no taste. The four columns "sweet, sour, salty, bitter" indicate the percentage of confusion with other taste qualities.

method the correlation coefficient for test and retest of the total scores was r_{69}=0.74; for individual taste qualities it was r_{69}=0.50 (sweet), r_{69}=0.36 (sour), r_{69}=0.37 (salty), and r_{69}=0.61 (bitter) (for all correlations p<0.01).

Investigation of repeatability using methods described by Bland and Altman (1999) revealed similar coefficients of repeatability for taste strips and the 3-drop-method (3.49 and 3.46, respectively). Neither an absolute systematic error nor a proportional error could be detected in the relation between measures from sessions 1 and 2. In addition, variation of the results did not depend on the magnitude of measurements (compare Figure 2).

To get an impression of the clinical usefulness we investigated 6 patients with lateralized loss of gustatory function following surgery of the middle ear (5 women, 1 man; age range 17-52 years). While all of these patients were able to identify suprathreshold tastants applied to the whole mouth, gustatory testing with the "taste strips" clearly differentiated between the healthy and the lesioned side (t-test: p=0.001). Differences in the number of correctly identified taste strips ranged from 5 to 11 (mean 6.8) between healthy and lesioned side.

DISCUSSION

The aim of the present study was to create a quantitative taste test which yields reliable data in a clinical surrounding in not more than ten minutes. Moreover the test should have a long shelf-life, be easily transportable, and allow to test each side of the tongue separately. An adequate tool to meet these requirements appeared to be taste strips which were impregnated with different concentrations of taste solutions. In order to enable a short time of examination, the taste strips were presented in a pseudo-randomized order. In this design each taste strip is presented only once, starting with the lowest concentration. In addition, the presently established test-retest-reliability of 0.68 compares well to other taste tests (Mattes, 1988; Doty, 1992; Ahne et al., 2000).

Nishimoto and collegues introduced a similar test in 1996 when they used salt-impregnated taste strips. However, due to the missing evaluation of sweet, sour and bitter taste this procedure seemed to be incomplete for clinical purposes. Paper-discs made from filter paper have also been used by Tomita and co-workers (1986) in order to define the size of the stimu-

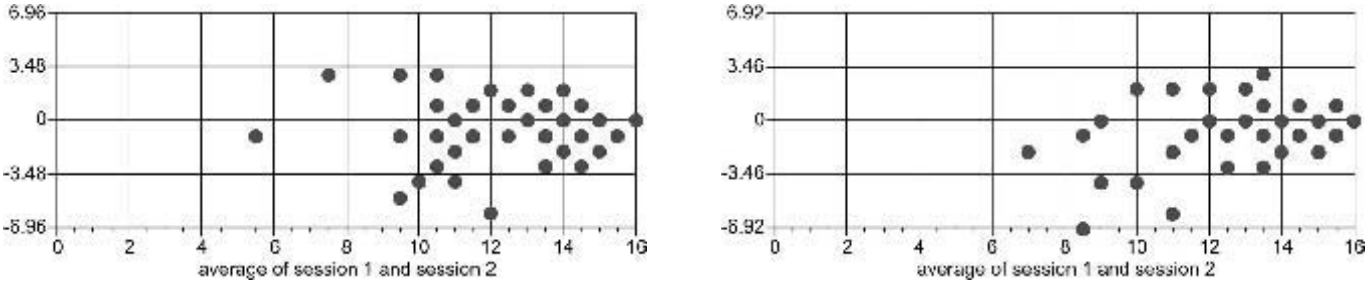

Figure 2. "Bland & Altman plot", separately for taste strips (left) and the 3-drop-method ("taste drops"; right). Differences between scores from session 1 and scores from session 2 are plotted against the average scores from the two sessions; 95% confidence intervals of differences between scores from session 1 and session 2 are indicated by horizontal lines (taste strips: 95 % confidence interval: +/- 3.48; "taste drops": 95 % confidence interval: +/- 3.46). A small number of outliers (n=4) is found for both tests with taste strips and taste drops. The plots indicate good reproducibility of the data; differences between results from sessions 1 and 2 are within the 95% confidence interval. Neither an absolute systematic error nor a proportianal error could be detected in the relation between measures from sessions 1 and 2. In addition, the variation of the results apparently does not depend on the magnitude of measurements.

lated surface of the tongue. In this protocol paper-discs are soaked with different taste solutions immediately before placement on the tongue. This elegant technique, however, does not remedy the problem that taste solutions have to be prepared freshly which is difficult in a clinical setting.
More recent work has led to the development of a whole-mouth gustatory test based on tasting tablets (Ahne et al., 2000). Here, subjects receive 28 tablets (6 concentrations for sweet, sour, salty, and bitter, respectively, plus four blanks) which requires 15-20 minutes. This test is easy to administer and has a very long shelf-life. However, it does not allow for regional testing which is of special importance after surgery of the middle ear with possible damage to the chorda tympani. Another interesting contribution was made by the development of a test which uses thin edible wavers (Hummel et al., 1997b). Thin wavers made from flour and water contain suprathreshold concentrations of tastants. While they may be used for both regional and whole-mouth testing they are not designed for the quantitative assessment of gustatory function.

Interestingly, we noticed a frequent confusion of salty and sour (see Table 4) which compares to previous work (Ahne et al., 2000). This may be due to the similarity of the two taste qualities since both of them produce a slight tingling/stinging on the tongue. It is well known that sour and salty taste is associated with a certain degree of irritation (Gilmore and Green, 1993). Furthermore, both tastants are frequently used in combination in foods with rich flavor, e.g. in dressings such as vinegar and salt. This well-known combination and the similarity of the two tastants with regard to irritation of the tongue could be a reason for the mutual confusion of these two taste qualities.

The results of the present study indicate the usefulness of the taste strips in a clinical context. Normative values (see Table 3) of the described test procedure are those of healthy, mainly young volunteers, who did not report any taste dysfunction, diseases relevant for taste perception, or medication. These data should be completed according to an age-related distribution in future publication. However, normal taste perception requires correct identification of the taste strip with the highest concentration.
Taken together, the described protocol for the quantitative evaluation of the human gustatory function exhibits several advantages, e.g. long shelf-life, convenience of administration, short time needed for testing, good reproducibility of the results, and the possibility to test each side of the tongue separately. Future work will focus on the establishment of normative data which will involve testing in patients with taste dysfunction.

ACKNOWLEDGEMENTS
This research was supported by an initiative of the working group on olfaction and gustation of the German Society for Otorhinolaryngology, Head and Neck Surgery, and by Theodor-Koerner-Fonds, Vienna, Austria (Dr. C. Mueller).

REFERENCES

1. Ahne G, Erras A, Hummel T, Kobal G (2000) Assessment of Gustatory Function by Means of Tasting Tablets. Laryngoscope 110: 1396-1401.
2. Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8: 135-160.
3. Deems DA, Doty RL, Gregg S, Moore-Gillon V, Shaman P, Mester AF, Kimmelman CP, Brightman VJ, Snow, Jr JB (1991) Smell and taste disorders, A Study of 750 Patients From the University of Pennsylvania Smell and Taste Center. Arch Otolaryngol Head Neck Surg 117: 519-528.
4. Doty RL, Shaman P, Dann M (1984) Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. Physiol Behav 32: 489-502.
5. Doty RL (1992) Diagnostic tests and assessment. J Head Trauma Rehabil 7: 47-65.
6. Gilmore MM, Green BG (1993) Sensory irritation and taste produced by NaCl and citric acid: effects of capsaicin desensitization. Chem Senses 18: 257-272.
7. Henkin RI, Gill, Jr JR, Bartter FC (1963) Studies on Taste Thresholds in Normal Man and in Patients with Adrenal Cortical Insufficiency: The Role of Adrenal Cortical Steroids and of Serum Sodium Concentration. J Clin Invest 42: 727-735.
8. Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G (1997a) 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. Chem Senses 22: 39-52.
9. Hummel T, Erras A, Kobal G (1997b) A test for the Screening of Taste Function. Rhinology 35: 146-148.
10. Kobal G, Klimek L, Wolfensberger M, Gudziol H, Temmel A, Owen CM, Seeber H, Pauli E, Hummel T (2000) Multicenter investigation of 1,036 subjects using a standardized method for the assessment of olfactory function combining tests of odor identification, odor discrimination, and olfactory thresholds. Eur Arch Otorhinolaryngol 257: 205-211.
11. Mattes RD (1988) Reliability of psychophysical measures of gustatory function. Percept Psychophys 43: 107-114
12. Nishimoto K, Hirota R, Egawa M, Furuta S (1996) Clinical Evaluation of Taste Dysfunction Using a Salt-Impregnated Taste Strip. ORL 58: 258-261.
13. Schiffman SS (1983) Taste and smell in disease (first of two parts). NEJM 308: 1275-1279.
14. Tomita H, Ikeda M, Okuda Y (1986) Basis and Practice of Clinical Taste Examinations. Auris Nasus, Larynx 13 (Suppl.1): S1-S15.

Christian Mueller, M.D.
Department of Otorhinolaryngology
University of Vienna, AKH Wien
Währinger Gürtel 18-20
A-1090 Wien
Austria

Fax: + 43-1-40400-3332
E-mail: chr.mue@gmx.at