Assessment of olfaction in multiple sclerosis: evidence of dysfunction by olfactory evoked response and identification tests

C H Hawkes, B C Shephard, G Kobal

Abstract
Objective—To resolve whether the olfactory pathways are affected in multiple sclerosis.
Methods—Olfaction was assessed by: (1) The University of Pennsylvania smell identification test (UPSIT, which uses microencapsulated odours that are released when scratched with a pencil) in 72 patients with multiple sclerosis and 96 controls, (2) olfactory evoked potentials (OEP) to 20 ppm H₂S by volume, and 50% CO₂ in air for 45 patients with multiple sclerosis and 47 controls. The abnormality rate in patients with multiple sclerosis for both tests (1) and (2) was compared with that for visual evoked potentials measured using a standard chequerboard technique.
Results—By comparison with controls, patients exhibited significantly low scores on the smell identification test with 15% of patients scoring outside the 95% confidence intervals for controls. The UPSIT was occasionally abnormal when the visual evoked potential (VEP) was normal. In general UPSIT scores correlated well with the H₂S-evoked response in controls and patients. For H₂S, there was a statistically significant increase of latency and decrease of amplitude for patients compared with controls. Increased H₂S latency and reduced UPSIT score correlated with greater disability on conventional rating scales. Overall, H₂S responses were abnormal in about one quarter of patients with multiple sclerosis. The sensitivity of UPSIT and OEP was similar although disorder on one test did not necessarily indicate abnormality in the other. The visual evoked potential was found to be a more sensitive indicator of disease than OEP or UPSIT.
Conclusion—These findings confirm the existence of olfactory dysfunction in multiple sclerosis and validate a new evoked potential technique.

Keywords: olfactory evoked potentials; smell identification; multiple sclerosis

Many have suggested that the olfactory pathway is spared in multiple sclerosis. It was claimed that the first cranial nerve is preserved because of the anatomical nature of myelin basic protein. Olfactory disorder was present in only 1% of patients with multiple sclerosis according to one observer. Dotty et al found microsma in seven of 31 patients with multiple sclerosis based on UPSIT tests. Another worker used serial binary dilutions of amyl acetate and nitrobenzene to determine the olfactory acuity of 40 patients with multiple sclerosis. The mean olfactory threshold of patients with multiple sclerosis was compared with that of 24 age and sex matched controls and no difference was found. Odours used by this author had some trigeminal component that would tend to reduce test specificity. In a further study using the UPSIT together with a match to sample test, patients with multiple sclerosis scored as high as the controls in both procedures. However, the average age of the patients with multiple sclerosis was 47 years, whereas that of the controls was 63 years, and there was a preponderance of women in the multiple sclerosis sample which might have masked some abnormalities in multiple sclerosis. Some have failed to find any plaques in the olfactory tracts of brains from patients with multiple sclerosis but there is pathological evidence that the olfactory tract can be affected. In a series of 22 patients with multiple sclerosis using various odours, it was found that anosmia or hyposmia was demonstrable in 10 patients and descriptive impairment in another. Eight of the 10 patients with hyposmia or anosmia were only detectable by use of odorants which have minimal or no apparent intranasal trigeminal activity. A similar percentage abnormality was found in another study of odour identification.

Olfaction has been one of the most difficult human sensory modalities to evaluate by passively delivered stimuli, chiefly because early techniques predisposed to simultaneous trigeminal stimulation that inevitably corrupted the stimulus. It is now possible to activate the olfactory nerve selectively by use of an olfactometer, described below, which generates a specific olfactory evoked potential (OEP) from the brain. Although some data are available for healthy subjects, there are few published studies for OEP in neurological disease states and none for multiple sclerosis. The
disease was selected (a) to compare results between a standardised psychophysical measure with those obtained from OEP in a clinical condition in which evoked potentials in other modalities are known to be altered; (b) to compare the relative sensitivity of these measures to one another and to visual evoked potential measures; and (c) to help clarify whether multiple sclerosis is accompanied by meaningful loss of olfactory function.

Methods

After local ethics committee approval and informed consent of patients and controls we undertook the procedures described below.

UNIVERSITY OF PENNSYLVANIA SMELL IDENTIFICATION TEST (UPSIT)

We studied 72 patients with multiple sclerosis (43 women; 29 men mean age 43.9 years). The nose was checked for patency by simply asking them to blow through one nostril at a time. A questionnaire was administered to all subjects to probe for previous history of nasal disease, head injury, use or misuse of drugs, and other conditions (for example, endocrine or hepatic disease) which are occasionally associated with hyposmia. The period of disease varied from 1 to 33 years with a mean duration of 11 years. Most (61) fell into the clinically definite category with four probable and seven possible cases. The University of Pennsylvania smell identification test (UPSIT) was used. This test employs strips of microencapsulated odours that are released when scratched with a pencil. For each of 40 strips, the odour must be identified from four given alternatives. Our normative data comprised 96 healthy people, 39 men, 57 women, aged 17–78 years. They were mainly derived from healthy members of hospital staff and British Telecom. Normal values for over 1500 American subjects are available, but it is essential to establish a United Kingdom control group because some test odours might be unfamiliar to United Kingdom residents. Smokers and non-smokers were included in control and patient groups.

OLFACTORY EVOKED POTENTIALS (OEPs)

The term olfactory evoked potential (OEP) will be used in this article to refer to both H₂S and CO₂ responses. CO₂ has no odour and is a stimulant of trigeminal nerve endings in the nose.

The olfactory stimulator we used is similar in construction and design to that described by Kobal except that our solenoid valves have a slower response time. Technical details of the olfactometer are given elsewhere. In brief, the device was set up to deliver either H₂S (20 ppm by volume) or CO₂ (50% by volume) without altering the chemical or thermal conditions in the nose. This is achieved by mixing a pulse of the stimulant in a constantly flowing air stream with controlled temperature and humidity. The flow rate of the airstream is 140 ml/s and the stimulus duration of both gases is 200 ms. Subjects breathe through the mouth to avoid variation in stimulus concentration due to nasal breathing. This is monitored by a thermistor taped close to the nose and if nasal breathing is discovered the trace is rejected. The cerebral response is recorded at CZ and PZ using the standard 10/20 system referenced to A1. The electro-oculogram is simultaneously averaged and any trial contaminated by eye blinks or other muscle artefacts—for example, from head or transmitted body tremor—is discarded. Typically a 2048 ms epoch is used of which 600 ms is presignal. Filters are set at 1–50 Hz and amplification at 6.25 µV for H₂S or 12.5 µV for CO₂. Eleven artefact free signals are averaged on a Nihon-Kohden Neuropak-4 machine, which uses 256 data points per channel. The averaged trace is smoothed by one pass of a nine point filter (equivalent to a low pass filter of 27.5 Hz). Latency measurements are unaffected by this filter setting but there is slight reduction of amplitude. Latencies are measured from the start of the trigger pulse. The stimulus reaches the nostril approximately 10 ms afterwards; because this interval is short in comparison with the initial N1 response, it was ignored in latency calculations. A simple visual tracking task is performed to maintain a constant state of alertness. Headphones delivering white noise are worn to mask out auditory clues from switching devices.

Figure 1 shows a sample of normal CO₂ and H₂S averaged responses derived from four healthy subjects. The P1 and N2 responses are often undetectable and N1 is sometimes indistinct even in healthy people. Latencies are measured to the first negative peak (N1) and second positive trough (P2). If the N1 response is not clearly visible its position is estimated from the point of downward (positive) deviation from baseline of the P2 response. Stimuli are presented unilaterally; the most patent nostril is usually selected for testing but if results are obtained from both nostrils then the mean latency and amplitude is used unless specified otherwise. Amplitude is measured from the peak of N1 to the trough of P2. In house normative values for visual evoked potentials (VEPs) were taken from a series of 70 healthy controls. Our technique of VEP recording uses a standard 16° radius chequerboard pattern back projected on to a screen viewed at a distance of 1 m. The square size is 50 seconds and luminance 137 and 14 candelas/m².

CONTROL GROUP

The control group was derived from the 96 in the UPSIT group who had been screened as described above. From them we obtained 47 subjects who had successful tracings to both H₂S and CO₂ (50%). Of these there were 18 males and 29 females, mean age 45.6 years. We attempted to test all those who volunteered for the UPSIT but there were several dropouts because of the more invasive nature of the OEP and organisational difficulties.

PATIENTS WITH MULTIPLE SCLEROSIS

We initially examined 45, also derived from the 72 in the UPSIT group defined above. They were obtained from inpatients or outpatients...
ona consecutive basis. Again all were asked to volunteer but several refused or were too disabled to be tested. In particular those with pronounced cerebellar tremor were unsuited to the procedure. There were 19 men, 26 women, mean age 42.6 years. All had clinically definite forms of the disease according to the McAlpine criteria (1972). The duration of disease ranged from 1–25 years with a mean of 13 years. Patients were classified according to the Kurtzke functional scale (KFS) and expanded disability status scale (17). Twenty two patients were on regular medication or had received a course of methyl prednisolone within the preceding six months. Typically they were receiving an antispastic agent—such as dantrolene, baclofen (liaoresal) or diazepam (Valium). Three were on antidepressants.

STATISTICAL METHODS
The overall difference between the groups, in terms of UPSIT score, N1, and P2 latency from start of the stimulus, and N1-P2 amplitude was assessed using regression analysis. This allows for the effect of age on the outcome and eliminates the need to match on age. For UPSIT scores the proportion of correct answers (p) was first transformed using the logit transformation (log(p/(1-p)) as p was skewed towards 1, especially for the controls. The effect of this is to increase the scatter (variability) of UPSIT scores close to 40 (the maximum score). This gives the variability of UPSIT score around the regression line a more normal distribution—one of the requirements for a valid regression analysis. An additional effect of this transformation is that the rapid decrease of UPSIT score with age, which appears as a possibly quadratic curve on the standard (linear) scale of measurement, will appear as a linear effect on this scale of measurement.

An evoked potential was considered absent if it was not possible to discern a clear response from the background noise in an artefact free recording. It was only possible to execute at most two averaged runs per gas in view of the prolonged nature of the tests in disabled patients. Hence the mean of two runs was taken.

Relations between UPSIT scores and OEP data and between KFS, EDSS, and OEP data were examined using Spearman’s rank correlation coefficient.

Results
UNIVERSITY OF PENNSYLVANIA SMELL IDENTIFICATION TEST (UPSIT)
Seventy two patients with multiple sclerosis and 96 controls were evaluated. Regression analysis after logit transformation (fig 2) showed no curved (quadratic) component on this scale of measurement (P=0.07) and a significant age related influence in both groups (P<0.0001). The age effect was not significantly different between patient and control groups (P=0.51). The regression line for patients was significantly lower than that for controls (P=0.0001), confirming that patients had impaired olfaction at all ages.

Out of the 72 patients with multiple sclerosis for whom UPSIT scores were obtained 11 (15%) scored outside the 95% prediction limit for controls. Sixty six of the 72 who had UPSIT were also evaluated by visual evoked potential. Fifty four (82%) of these were abnormal (outside 95% limits); 11 (17%) were normal; one equivocal. Of the 54 with abnormal VEP only seven (13%) had abnormal UPSIT. Of the 12 with a normal visual evoked potential, three (5%) had UPSIT scores outside the 95% limit for controls and the remaining nine were normal.
The control group consisted of 47 subjects with clear responses to both gases. There were two apparently healthy people not included who had no visible response to H$_2$S probably due to blink and gas intolerance artefact. Of the 45 patients examined, a distinct OEP to H$_2$S was obtained in 37 (16 men, 21 women). In the remaining eight the responses could not be obtained either because of artefacts or absent response. Figure 3 shows a sample of averaged CO$_2$ and H$_2$S responses of varying latency from four patients with multiple sclerosis. The response to 50% CO$_2$ was measured in 34 patients. Of these 27 were clear; seven were ill defined because of unavoidable recording artefacts chiefly because of intolerance to 50% CO$_2$. All had responses to lower concentrations of this gas. Only 26 had acceptable tracings to both gases, and these were used for comparison with OEP in the control group.

The N1 and P2 distributions were normal but the amplitude distribution was skewed. It was decided to use log transformation on all three variables as this not only normalised the amplitude distribution but stabilised the variances between the two groups for N1 and P2. For each transformed variable, regression analysis was carried out adjusting for age to see if any significance could be attributed to differences due to diagnosis—that is, whether multiple sclerosis or control. For the 26 patients and 47 controls, analysis showed that for H$_2$S a significant difference existed between diagnostic groups for N1 and P2 latency and N1-P2 amplitude (table 1). None of the CO$_2$ variables was significant. For N1 and P2 latency, regression lines on age were constructed for both gases. For CO$_2$ there were two of 26 patients with delayed N1 latency and four of 26 delayed to P2 representing five separate patients (8%-15% outside 95% confidence limits). For H$_2$S there were six of 26 delayed to N1 and three of 26 to P2 representing six separate patients (11%-23% abnormal). Of the 45 patients in whom OEP were attempted 41 had visual evoked potentials (VEP) undertaken within the preceding six months. Of these 34 (75%) were delayed (more than 2 SD from mean) and seven were normal. In all instances of intact CO$_2$ or H$_2$S response the VEP was also normal. There were no instances of abnormal CO$_2$ or H$_2$S responses in the presence of normal VEP but there were many examples of the converse. In four instances in which an individual CO$_2$ response was abnormal the H$_2$S latency was intact; in four subjects the H$_2$S response was delayed and the corresponding CO$_2$ was normal; in five occasions both CO$_2$ and H$_2$S latencies were abnormal in the same patient.

The UPSIT scores for patients and controls were compared with OEP for patients and controls with responses to both gases. There was a weak correlation between latency and amplitude measurement for H$_2$S (N1 latency $r = -0.322$, $P<0.006$; N1-P2 amplitude $r = 0.357$, $P<0.002$) and with CO$_2$ P2 latency ($r = -0.254$, $P<0.033$). In five instances in which CO$_2$ responses were delayed the UPSIT score was abnormal in only two and another six with delayed H$_2$S responses all had normal UPSIT score.
Disease severity was graded on the Kurtzke functional scale (KFS) and expanded disability status scale (EDSS). Correlations were then calculated between these two measurements and both H$_2$S and UPSIT data. There was a significant positive correlation of increasing disability on EDSS with N1 and P2 latency (N1: $r = 0.44$, $P < 0.01$; P2: $r = 0.42$, $P < 0.02$) but not with amplitude, indicating that more disabled patients have a longer latency response. The EDSS correlated reasonably well with UPSIT score ($r = 0.42$, $P < 0.0003$). KFS did not correlate with the H$_2$S or UPSIT measurements.

Discussion

The present findings suggest that some patients with multiple sclerosis have impaired olfaction—that is, about 25% on the basis of OEP and 15% on the basis of UPSIT. Our value for identification tests is slightly lower than the 23% incidence found by Doty but much less than the 35% described by Wender and Szmeja and 45% documented by Pinching. This variation may well relate to patient selection in a hospital environment. As stated, many of our patients had received a course of steroids which could have improved their olfactory tests either at the level of the nose or central olfactory pathways. It is likely that impaired smell function may occur more often during relapse as reduced olfactory ability is sometimes noted by patients during their attacks and may improve with remission or administration of steroids. As shown in figure 2, UPSIT scores for multiple sclerosis are generally low and fall off with age as do those of the controls. Although the patients with multiple sclerosis are clearly a different population only 15% had scores outside the 95% limit of prediction for controls. The absence of selective impairment to any of the 40 odours suggests a generalised olfactory disorder in a minority of patients—a situation quite unlike that found in Parkinson’s disease, in which the defect is severe and certain odours seem to be identified with more difficulty than others. The presence of low UPSIT score in a few of those patients with normal visual evoked potentials emphasises that a different modality is being measured. However, the VEP was abnormal far more often than UPSIT.

The OEP measurements show abnormalities affecting H$_2$S latency and to a lesser extent, amplitude. This phenomenon is similar to that found in pattern evoked visual potentials in multiple sclerosis and polyneuropathy of demyelinating type. Only about one quarter of patients with multiple sclerosis had delayed H$_2$S responses. This may be an underestimate as very disabled patients were less able to be tested, especially those with head tremor. This contrasts with the 75% abnormality on VEP in our patients. Disordered CO$_2$ potentials were less consistent depending on the method of analysis and did not invariably imply abnormal H$_2$S response. Conversely an impaired H$_2$S response was sometimes accompanied by normal CO$_2$ values, implying that the two tests

<table>
<thead>
<tr>
<th>Table 1 Comparison of regression line data from only controls and patients with responses to both gases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H$_2$S</strong></td>
</tr>
<tr>
<td>N1</td>
</tr>
<tr>
<td>P2</td>
</tr>
<tr>
<td>Amp</td>
</tr>
<tr>
<td><strong>CO$_2$</strong></td>
</tr>
<tr>
<td>N1</td>
</tr>
<tr>
<td>P2</td>
</tr>
<tr>
<td>Amp</td>
</tr>
</tbody>
</table>

Note that differences are expressed in logs. All H$_2$S results are significant but none of the CO$_2$ data are. N1 latency in ms; P2 latency in ms; Amp = N1-P2 amplitude; n=26 patients.
were evaluating different aspects of neuronal function.

In our initial OEP analysis with smaller numbers there was a higher percentage of patients with prolongation of H₂S latency, probably because more severely disabled patients were studied. If the OEP analysis is restricted to just those with responses to both gases then H₂S OEP was more sensitive than CO₂ in detecting abnormality. If all patients with multiple sclerosis are compared with controls then there are significant latency differences to CO₂ as well. Clearly a satisfactory trigeminal response can be obtained and abnormalities of this would not be surprising in view of the frequency of facial numbness in multiple sclerosis. Patient and control CO₂ values deteriorated with age. Aging alone is associated with worsening H₂S and CO₂ responses in both groups and for H₂S the explanation is the well documented deterioration of smell sense with age; CO₂ responses behave in similar manner.

The concentration of H₂S used in our study was higher than that used by Kobal and Hummel—that is, 20 ppm compared with 2 ppm. We elected for a higher concentration to facilitate detection of the evoked response. None of our six anosmic patients studied informally could smell H₂S at this concentration nor was an evoked response discernible to this gas, but all responded to CO₂. A lower concentration of H₂S (for example, 2 ppm) may suffice. Despite differences in gas concentration and solenoid valves comparison of our data with unpublished control data from Erlangen shows similar latencies. The amplitude measurements are slightly smaller for Ipswich patients, which may relate to the use of faster reacting valves in Erlangen.

It should be emphasised that there are more technical difficulties in obtaining an olfactory evoked response in comparison with other procedures such as visual, auditory, and somatosensory potentials. Apart from the need for frequent checking of apparatus it is rarely practical even in healthy people to average more than 20 signals because of the necessarily prolonged interstimulus interval required (60 s). There were two apparently healthy subjects from the UPSIT control group with no response to H₂S. We are unsure of the reason—perhaps intolerance to the smell of H₂S and the fact that they were both elderly subjects (62 and 78 years old). The averaged response could be replicated only once in patients with multiple sclerosis because of the frequent tendency to fatigue and poor concentration. Many patients were taking drugs which had the potential to interfere with the evoked responses. Steroid use has been discussed.

Both antidepressants and antispastic agents taken by our patients could impair olfactory function either by provoking drowsiness or by a direct effect on the olfactory pathways, and we have not been able to allow for this source of possible error.

The clear advantage of OEP is its objectivity in comparison with psychophysical tests; hence a demented patient could be evaluated provided he was cooperative. Insufficient is known about the effect of circadian and diurnal variation on these evoked potentials. It is only feasible to study a few odours (just one in each case) and theoretically more subtle defects to a particular odour group might be missed. The UPSIT evaluates the entire olfactory process and is a useful screening test. The existence of patients with normal UPSIT and abnormal H₂S responses confirms that different aspects are measured. Whether the OEP/UPSIT will have value as a diagnostic procedure in patients with multiple sclerosis is open to debate. The presence of abnormal UPSIT scores and normal visual evoked potentials in some people implies that olfactory assessment may be of some diagnostic value; however, no one had an abnormal H₂S or CO₂ evoked response in association with a normal VEP.

Compared with VEP relatively few patients with multiple sclerosis have abnormality on either UPSIT or OEP, presumably because the olfactory pathways are less often the site of demyelination than the optic nerve. It is possible that undetected olfactory abnormalities were present beyond the limits of our technique but the similarity of the OEP/UPSIT data and comparability of our findings with others suggests that the procedure is reasonable. More abnormalities might be exposed in the OEP by use of other olfactory stimulants such as vanilla or phenyl ethyl alcohol.

Olfactory evoked response studies have been introduced into clinical work only recently and our study is one of the first applied to a neurological disease. It confirms that olfaction is impaired in multiple sclerosis and validates the earlier pathological findings of Peters and McDonald.8

We thank the East Anglian Regional Health Authority for supporting this project (grant No 42) and Miss A Petruckevitch, Institute of Neurology, London for statistical support. We are also grateful to Dr Geoff Barrett and Dr Brian Youill, Institute of Neurology, for helpful comments on this article.

---


Assessment of olfaction in multiple sclerosis: evidence of dysfunction by olfactory evoked response and identification tests
C H Hawkes, B C Shephard and G Kobal

J Neurol Neurosurg Psychiatry 1997 63: 145-151
doi: 10.1136/jnnp.63.2.145

Updated information and services can be found at:
http://jnnp.bmj.com/content/63/2/145

These include:

References
This article cites 14 articles, 5 of which you can access for free at:
http://jnnp.bmj.com/content/63/2/145#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Immunology (including allergy) (1943)
Multiple sclerosis (934)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/