New Modification of Smell Identification Test for the Detection of Malingers: A Pilot Experimental Study

Reza Erfanian1, MD; Safa Taherkhani2, PhD; Hakima Abdullah1, MD; Saeed Sohrabpour1, MD; Hamed Emami1, MD; Mehdi Hoorang3, PharmD; Behrooz Amirzargar1, MD

1Otorhinolaryngology Research Center, Tehran University of Medical Sciences, Tehran, Iran; 2Department of Biomedical Engineering (Center of Excellence), Amirkabir University of Technology, Tehran, Iran; 3Health Policy Research Center, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran

Correspondence: Behrooz Amirzargar, MD; Valiasr Hospital, West end of Keshavarz Blvd., Postal code: 14197-33141, Tehran, Iran
Tel/Fax: +98 21 66581628
Email: amirzargarb@yahoo.com
Received: 24 November 2020
Revised: 30 January 2021
Accepted: 13 March 2021

Abstract

Background: A major problem with the University of Pennsylvania Smell Identification Test (UPSIT) is its poor sensitivity for malingering detection in a group of people familiar with the test mechanism. This study aimed to evaluate the modification of UPSIT to detect anosmia malingering.

Methods: This was a pilot experimental study conducted in 2019 in Tehran. The participants were 60 healthy subjects classified into two groups of 30 people. The first group was requested to deliberately feign a negative result on the Iranian version of UPSIT, Iran Smell Identification Test (ISIT) (malingering group). The second group consisted of participants, who did not scratch the odorant part of ISIT during the tests (anosmia group). ISIT was modified in two steps. At each step, one incorrect option was deleted from the available choices. The number of each group’s answers, altered away from the correct choice, was then calculated and compared.

Results: The coached malingering group participants were able to feign anosmia in the original ISIT exam. In the modified ISIT, the number of answers changed from correct to wrong during the second stage (from three available choices to two choices) was significantly higher in the anosmia group (P<0.001). In the ROC analysis, the area under the curve was 0.92 (P<0.001). The cutoff of 4.5 for this test showed 93% sensitivity, 82% specificity, and 90% PPV and NPV.

Conclusion: The ISIT is not capable of detecting malingering in the coached participants, yet by deleting the choices step-by-step, the sensitivity and specificity of the test increased.


Keywords ● Smell ● Malingering ● Anosmia

Introduction

Malingering is defined as the intentional creation of false or exaggerated physical or psychological symptoms stimulated through external motivations, such as eschewing work, gaining economic benefit, escaping criminal prosecution, or abusing drugs. The common strategy for detecting malingering is to utilize the techniques disclosing the behavior inconsistent with the alleged loss.

Traumatic brain injury is a major public health problem associated with high rates of death along with physical and
sensory impairment, such as olfactory dysfunction, which is a common consequence of craniofacial (CF) trauma.3 According to studies on olfactory dysfunction, 20% to 30% of patients with head trauma experience anosmia (post-traumatic complete loss of smell).4 Different factors have been reported to be associated with the olfactory disorder, including ruptures or stretches in the filaments of olfactory nerves passing through the ethmoid cribiform plate, lesions disrupting the olfactory bulbs, the primary or secondary olfactory cortex, brain connection areas, and the damage of the nasal mucosa following trauma.5-7 A claimed sensory loss may be the basis of the pursuit of disability paybacks. In such cases, although a thorough history can reveal discrepancies, they may not be adequate to demonstrate malingering.

There are several tests to measure olfactory dysfunction. The University of Pennsylvania Smell Identification Test (UPSIT) is one of the most commonly used tests for olfactory dysfunction and is considered as a gold standard of smell test owing to its reliability (r=0.94).8, 9 UPSIT was first employed in North America in 1984. It is a multiple-choice test, and there are four options per question. The subjects should choose one option, even if they do not sense any smell. The test takes about 10-15 minutes, containing microencapsulated odorants in a scratch and sniff format for each question. It comprises 40 questions, out of which the anosmic patients generally score around 10 out of 40 correct, as they choose their answers by chance. Malingering (or faking anosmia) should be considered as a possibility in patients scoring five or less.10

In Iran, ISIT, which comprises 40 items is utilized to confirm anosmia in forensic medicine, which proved to be reliable with a Cronbach’s alpha of 0.78. Since anosmia caused by an accident may incur full compensatory costs,11, 12 A person, who is malingering deliberately chooses incorrect options throughout the test, while an anosmic person usually answers approximately one-fourth of the questions correctly. A drawback of UPSIT is that if subjects are familiar with the test mechanism, they can therefore feign anosmia (malingering). Hence, a malingering knowing the test mechanism can effectively produce similar results to anosmia patients.13 This study aimed to evaluate the UPSIT modification to prevent anosmia malingering.

**Materials and Methods**

This was a pilot experimental study conducted in 2019 in Tehran. The subjects were 60 healthy adults with appropriate olfactory function enrolled in this study from January 2019 to March 2019. All the participants were otorhinolaryngology residents at Tehran University of Medical Sciences. Written informed consent was obtained from all of the participants, and the study was approved by the Ethics Committee of Tehran University of Medical Sciences considering the Declaration of Helsinki (Ethical approval code: IR.TUMS.AMIRALAM.REC.1398-10-22-AE). All the participants were examined for normal olfactory function according to ISIT prior to enrolling in the current study. The participants were excluded from the study, if they had a history of olfactory dysfunction, including anosmia, parosmia and phantosmia, severe septal deviation, nasal obstruction and polyposis, sinusitis, allergic rhinitis, history of nasal surgery, neuroleptic disease, and head trauma.

The participants were randomly classified into two groups of 30 based on block randomization. The participants of the first group were asked to feign ISIT (malingering group). The participants of the second group (anosmia group) were normal subjects answering the questions without scratching the odorant part of the exam booklet. The exam booklets applied in this study were bought from Saba Tajhiz Sabalan Medical Engineering Company in Tehran.17 ISIT exam was taken by both groups, and the participants were classified according to their ISIT score (0-5: malingering, 6-18: anosmia, 19-33: microsomia, 34-40: normosmia).

Subsequently, the test mechanism was explained to the malingering group, and ISIT was retaken by both groups. Participants were again classified based on their ISIT scores.

Afterward, the ISIT format was modified across three steps, and based on the modified version, the two groups were examined. In the modified test, for each question, two incorrect answers were randomly selected for stepwise elimination. These options were removed to limit the subject’s choices.

In step 1, the participants were asked to choose an answer for each question, even if they could not identify the smell. Afterward, in step 2, a single option was deleted and the participants were again requested to select one among the three remaining options; meanwhile, an incorrect or correct answer was recorded. Subsequently, in step 3, a further option was deleted and the participants were asked to select an option between the remaining two. The number of correct answers given at each step, the number of correct answers that were changed from correct to wrong, and the number of options changed from wrong to correct were calculated for each individual.
The calculated parameters for each participant included:

- **T1**: The number of correct answers in the first step comprising four different options.
- **T2**: The number of correct answers in the second step comprising three options following the elimination of an incorrect option.
- **T3**: The number of correct answers in the third step following the elimination of two incorrect options. The participants were asked to choose an option from the two remaining ones.
- **TF1**: The number of questions in which the subjects changed their correct option to an incorrect one once reduced from four to three.
- **TF2**: The number of questions in which the subjects changed their correct option to an incorrect one once reduced from three to two.
- **FT1**: The number of questions in which the subjects changed their incorrect option to the correct one once reduced from four to three.
- **FT2**: The number of questions in which the subjects changed their incorrect option to the correct one once reduced from three to two.
- **MT1**: Maximum correct answers in a row in the four-option step.
- **MT2**: Maximum correct answers in a row in the three-option step.
- **MT3**: Maximum correct answers in a row in the two-option step.
- **MF1**: Maximum incorrect answers in a row in the four-option step.
- **MF2**: Maximum incorrect answers in a row in the three-option step.
- **MF3**: Maximum incorrect answers in a row in the two-option step.

**Statistical Analysis**

Qualitative data was reported as percentage and frequency. The Chi square test was applied for the comparison of qualitative data. The Mann-Whitney U test and *t* test were utilized to compare the quantitative differences between the two groups. Box plots were employed to illustrate the distribution of the median, the 25th percentile, and the 75th percentile. ROC curve was used to determine cut-off, sensitivity, and specificity. *P* values of less than 0.05 were considered to be significant. All the data were analyzed via IBM SPSS Statistics software version 21 (IBM Corp., Armonk, NY, USA).

**Results**

The baseline characteristics of the participants are shown in Table 1. There was no statistically significant difference between the malingering and anosmia groups. All *P* values were more than 0.05.

Primarily, out of the 30 subjects in the malingering group, 12 subjects (40%) were successfully able to feign anosmia according to Iranian forensic medicine guidelines. The mean score and standard deviation was 5.4±4.3. Once being informed about the test format and how to answer its questions, all the 30 participants of the malingering group were successfully able to feign anosmia, according to Iranian forensic medicine guidelines. The mean score and standard deviation was 12.2±4.4, which was a significant increase from the baseline (*P*<0.001).

The number of correct answers in the first part of the modified test comprising four options was then examined. The median, 25th percentile, and 75th percentile for the anosmia group were 11, 9, and 14, respectively. For the malingering group, the median, 25th percentile, and 75th percentile were 13, 11, and 15, respectively (Figure 1A). In the second step of the modified test comprising three options, the median, 25th percentile, and 75th percentile for the anosmia group were 16, 13, and 17, respectively, while for the malingering group the same values were 17, 13, and 18, respectively (Figure 1B).

Once the correct answers were checked in the third step of the modified test, comprising two options, the medians of the 25th percentile and 75th percentile for the anosmia group were respectively 19, 17, and 23 and 20, 18, and 24 for the malingering group (Figure 1C).

The examination of the number of questions, in which subjects changed their correct choices to incorrect ones from the four-option step to the three-option step revealed that the median, 25th percentile, and 75th percentile for the anosmia group were respectively 7, 6, and 10 along with 5, 2, and 7 for the malingering group (Figure 2A). In terms of the number of questions, in which subjects changed correct choices to incorrect

### Table 1: Baseline characteristics of the participants in the anosmia and malingering groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Malingering Group</th>
<th>Anosmia Group</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age*</td>
<td>28.65±3.91</td>
<td>28.92±3.64</td>
<td>0.857</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12 (40%)</td>
<td>14 (46.67%)</td>
<td>0.169</td>
</tr>
<tr>
<td>Female</td>
<td>18 (60%)</td>
<td>16 (53.33%)</td>
<td></td>
</tr>
<tr>
<td>Primary score of ISIT*</td>
<td>38.13±1.74</td>
<td>37.82±2.14</td>
<td>0.795</td>
</tr>
</tbody>
</table>

*Data are presented as “mean±SD”. Age difference was analyzed via *t* test. Sex difference was analyzed with Chi square test. *P* value of <0.05 was considered to be statistically significant. ISIT: Iran Smell Identification Test
Modification of smell identification test for malingerers

ones from the three-option step to the two-option step the results indicated that for the anosmia group the median, 25th percentile, and 75th percentile were respectively 8, 5, and 9 and 3, 2, and 4 for the malingering group (figure 2B).

The statistical analysis revealed that the best parameter for identifying subjects with anosmia was the number of responses changed from correct to incorrect ones in the third step. These numbers were lower in the malingering group, since this group’s subjects knew that they had recognized the correct smell, and therefore they did not wish to change their correct option to an incorrect one (to increase the number of correct answers by decreasing the options). However, the anosmic subjects were more likely to change their choice, since they were unable to recognize, whether they were correct or incorrect leading to a naturally high number of wrong answers. In the

Figure 1: The box plots represent the number of correct answers from the first to the third step in the two groups of anosmia and malingering. T1=The number of correct answers in the first step presenting four options to the individual. T2=The number of correct answers in the second step presenting three options for the individual following the elimination of an incorrect option. T3=The number of correct answers in the third step with two eliminated incorrect options. The participant was asked to choose an option from the two remaining options.

Figure 2: The box plots show the answers that subjects changed from correct to incorrect. TF1=The number of questions, in which the subjects changed their correct option to an incorrect one once reduced from four to three. TF2=The number of questions, in which the subjects changed their correct option to an incorrect one once reduced from three to two.

Figure 3: The ROC curves demonstrate test sensitivity and specificity. TF2=The number of questions, in which the subjects changed their correct option to an incorrect one once reduced from three to two. FT1=The number of questions, in which the subjects changed their incorrect option to the correct one once reduced from four to three. FT2=The number of questions, in which the subjects changed their incorrect option to the correct one once reduced from three to two.
ROC analysis, the area under the curve (AUC) was 0.92 (P<0.001). The cut-off of 4.5 for this test showed 93% sensitivity and 82% specificity (figure 3). The Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were both 90%. The comparisons of mean scores of the parameters in the anosmia and malingering groups are represented in table 2.

**Discussion**

Our study demonstrated that the Iranian version of UPSIT, also known as ISIT, is not capable of detecting malingering in the coached participants. Herein, modified ISIT was employed to prevent anosmia malingering.

In the present study, ISIT was moderated across the three steps. At each step, an incorrect option was deleted, and the final answers were recorded. The number of answers changed from correct to incorrect options was calculated. AUC was obtained as 0.92. The cut-off of 4.5 for this test revealed a 93% sensitivity and 82% specificity. Without any knowledge about the mechanism behind the ISIT test, only 12 subjects were successfully able to feign anosmia to the national standards of forensic medicine. However, following the given training on the mechanism of the test, all the healthy subjects were able to feign anosmia successfully according to the forensic standards. The results implied that a step-by-step deletion of the available options (modification) increased the sensitivity and specificity of the test. The current work is different from many others, owing to the nature of the modification. The present study did not develop a new test, instead applied a new method to raise the sensitivity and specificity of the commonly used tool, UPSIT.

There are other indirect ways to confirm anosmia. In a study by Roberts and others, a clear decrease in olfactory activity in the anterior cortical regions of the patients with anosmia following trauma was reported. In these patients, a decline in the metabolic activity of olfactory areas was demonstrated in neuro-image studies. Bonanni and colleagues examined 25 anosmic patients following head trauma under an olfactory stimulation electroencephalogram and observed 17 cases, where olfactory reactions had stopped. Neuroradiological studies have shown evidence of a decrease in the volume of the olfactory bulb and the inferior frontal cortex in adults with olfactory dysfunction after traumatic brain injury. Meanwhile, none of these indirect smell tests are fully validated for the detection of anosmia owing to various mechanisms being potentially responsible for
traumatic and non-traumatic anosmia. An alternative test for olfaction is the Sniff Magnitude Test (SMT). This test quantifies the olfactory function with the measurement of the exploratory sniffing behavior in response to odor stimuli. The main output of SMT is the “sniff magnitude ratio”, defined as the mean sniff magnitude formed by the unpleasant odor stimuli divided by the mean sniff magnitude to nonodorized air equal to one in the anosmia. However, the simplicity of the test and the suboptimal olfactory reflex make it highly vulnerable to malingering.

In Iran, UPSIT has been standardized to measure the olfactory function. ISIT, the standardized 40-item smell identification test, is a modified version of UPSIT to assess the olfactory function in Iranian patients. Pouraghaei and others performed a study to compare the efficacy of SPECT and ISIT to detect real anosmia and malingering in forensic medicine. The sensitivity values of ISIT to detect malingering subjects and patients with anosmia were 66.6% and 87.5%, respectively. The specificity of ISIT tool to detect malingering subjects and anosmic patients was 90% and 75%, respectively. The sensitivity and specificity of the ISIT to detect hyposmia cases were obtained as 100%. Therefore, it was concluded that the use of ISIT for ruling out malingering cases was useful and efficient.

Various studies have tried to improve the detection of olfactory dysfunction malingerers. Mehdizadeh and colleagues developed a novel test to differentiate malingering subjects from the patients with anosmia, determining five substances (coffee, lemon, rosewater, thyme, and garlic) to be qualified odors within a 20-item odor discrimination test. Therein, subjects were forced to select the bottle with different odors from two other bottles. The test consisted of 20 items (60 bottles). It has been revealed that this test comprised 90% sensitivity, 55.71% specificity, 67.02% PPV, and 84.78% NPV. In another study, the scratch density for releasing the odorant from the microencapsulated odorant strip of UPSIT was utilized to differentiate malingerers. The malingerers had less scratch density in releasing the odorant.

The other psychophysical method to detect malingering is analyzing the response sequences of the examinee facing the different levels of smell stimulants or no stimulants. Linschoten and Harvey indicated that the correct classification of patients with anosmia and malingerers increased to 100% through the use of response-sequence analysis for discrimination and concluded that a maximum-likelihood adaptive staircase procedure, accompanied by response-sequence analysis could be considered as a powerful method to detect malingerers in evaluating the olfactory function. They utilized different butyl alcohol concentrations in each step and chose the next concentration based on the previous response. The examinee was forced to choose between two choices (the smell is present or not). The basis of this test was that the malingerers cannot produce truly random sequences of response. However, in other multiple choice psychophysical tests malingerers can choose the response prior to stimulus presentation and make a response sequence equal to anosmic patients.

A potential limitation of this study may relate to individual differences in working memory, which may influence responsiveness; however, these limitations are beyond the control of the researchers.

Conclusion

According to results, the olfactory forensic test applied in Iran, which is based on the Pennsylvania test, has a very high error rate for malingering subjects familiar with the test mechanism. This can lead to high financial costs and legal injustices for both society and the government. However, it was found that, by use of step-by-step deleting options, the sensitivity and specificity of the test increased. Thus, this method of malingering evaluation may have broader applications in multiple-choice psychometric tests in forensics medicine.

Acknowledgment

The research was fully sponsored by the Otorhinolaryngology Research Center of Tehran University of Medical Sciences.

Conflict of Interest: None declared.

References

PMID: 29446651.


22 Doty RL, Genow A, Hummel T. Scratch density differentiates microsmic from normosmic