Received: 17 October 2021 | Revised: 16 December 2021 | Accepted: 17 January 2022
DOI: 10.1002/lio2.747

ORIGINAL RESEARCH

Laryngoscope
Investigative Otolaryngology

Comparing the diagnostic accuracy of audiometric Weber test and tuning fork Weber test in patients with conductive hearing loss

Siti Nazira Abdullah MD[1] | Mohd Normani Zakaria PhD[2] | Rosdan Salim MMed ORL-HNS[1] | Mohd Khairi Md Daud MMed ORL-HNS[1] | Nik Adilah Nik Othman MMed ORL-HNS[1]

[1]Department of Otorhinolaryngology-Head and Neck Surgery, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[2]Audiology Programme, School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

Correspondence
Nik Adilah Nik Othman, Department of Otorhinolaryngology-Head and Neck Surgery, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.
Email: adilahkk@usm.my

Abstract

Objectives: Weber test is typically conducted using tuning forks, but an audiometer can also be used for a similar purpose. Compared to the tuning fork Weber (TFW) test, performing the audiometric Weber (AW) test offers many advantages. Nevertheless, AW and TFW tests' performance compared to pure-tone audiometry (PTA) has yet to be studied. The present study aimed to determine the accuracy and agreement between the AW and TFW tests compared to PTA.

Methods: In this observational cross-sectional study, 74 participants (aged 12-67 years) with unilateral conductive hearing loss (CHL) or bilateral asymmetrical CHL were enrolled. The TFW test was performed according to the established protocol at 256 and 512 Hz. For the AW test, the bone vibrator was placed in the middle of the forehead, where 250 and 500 Hz frequencies were tested. TF and AW test results were then compared with the expected lateralization from the respective PTA results.

Results: At 256 Hz (or 250 Hz), the overall accuracy values of TFW and AW tests were 81.1% and 86.5%, respectively. At 512 Hz (or 500 Hz), the overall accuracy results of TFW and AW tests were 85.1% and 82.4%, respectively. In addition, the kappa statistics revealed substantial agreements between the two tests and PTA ($k = .63\text{-}.72$).

Conclusion: Both AW and TFW tests are reasonably accurate in assessing patients with CHL. It is recommended for audiologists to perform the simple AW test to verify incomplete or questionable audiograms that are commonly encountered in clinical practice.

Level of evidence: Level 3b.

KEYWORDS
accuracy, audiometric weber test, conductive hearing loss, masking dilemma, tuning fork weber test

1 | INTRODUCTION

Pure-tone audiometry (PTA) has been regarded as the gold standard test for detection of hearing thresholds.[1–3] With the use of a sophisticated audiometer, the severity of hearing loss and type of hearing loss across speech frequencies for each ear are conveniently documented by PTA.[3] Prior to the wide application of PTA in clinical settings, the hearing status was assessed using the tuning fork tests.[1,4] In fact, tuning fork tests such as Weber and Rinne tests are still commonly used by otorhinolaryngologists as they are inexpensive, simple to administer, and reasonably sensitive in detecting hearing loss.[1] Of note, the sensitivity and specificity of the tuning fork Weber (TFW) test can be as high as 78% and 99%, respectively.[5,6] When it is combined with the Rinne test, the overall diagnostic accuracy improved.[7]

Despite its diagnostic usefulness, PTA has several limitations. In particular, masking problems and "false" air-bone gaps (ABGs) may occur, which would affect its diagnostic accuracy.[1–3] In cases of bilateral large ABGs in PTA, overmasking is likely to occur, and masked bone conduction (BC) thresholds cannot be obtained (resulting in incomplete PTA results).[3] With the widely used B-71 bone vibrator, better than expected BC thresholds would be produced (leading to false ABGs). Specifically, the bone vibrator would produce vibrotactile (VT) sensation at low frequencies (as low as 25 dB).[3] In this regard, due to lower BC thresholds, inappropriate ABGs are produced, which can be misdiagnosed as conductive hearing loss (CHL).[8] The false ABGs in PTA could also be contributed by harmonic distortions produced by the respective bone vibrator (which may appear as low as 20 dB at low frequencies).[9] The presence of harmonic distortions may influence the BC thresholds and create false ABGs.

The application of the tuning fork Weber (TFW) test would be useful in verifying questionable audiograms. For example, if the true (masked) BC thresholds could not be obtained due to the masking problems (in cases of bilateral large ABGs), a centralized perception in the TFW test would support the diagnosis of bilateral CHL. On the other hand, a lateralized perception (to one ear) would indicate the presence of sensorineural hearing loss (SNHL) (in the other ear). The presence of false ABGs could also be confirmed with the use of the TFW test, particularly in cases of unilateral hearing loss. In this regard, the combination of the PTA and Weber tests would be useful in solving "complex" clinical cases.

It is worth mentioning that the Weber test can also be carried out using the readily available bone vibrator of the audiometer. Known as the audiometric Weber (AW) test, it works in the same way as the TFW test (with the bone transducer placed in the middle of the forehead). However, in comparison to the TFW test, the AW test has several advantages, including multiple frequencies can be tested, the sound presentation can be controlled at intended intensity levels, and it offers consistent force on the forehead surface area.[10–12]

In view of the advantages of the AW test, it is not known whether the AW test would outperform the TFW test when assessing patients with CHL. As such, the diagnostic performance of TFW and AW tests (in comparison to PTA) has not been systematically studied. Essentially, the aim of the present study was to determine the accuracy and the agreement of TFW and AW tests in comparison to PTA.

2. MATERIALS AND METHODS

2.1. Ethical considerations

Prior to the data collection, a written consent form was obtained from each participant. The study procedure was approved by the institutional ethics committee, in accordance with the 1975 Declaration of Helsinki and its later amendments.

2.2. Study design and participants

The present observational cross-sectional study followed the STARD reporting guideline. Seventy-four eligible participants were recruited among patients of the Otorhinolaryngology (ORL) Clinic of a University Hospital. The inclusion criteria were those with the age of at least 12 years and those with unilateral CHL (with ABG of at least 15 dB at any one frequency) or bilateral asymmetrical CHL (with the difference between right and left AC thresholds of at least 15 dB at two adjacent frequencies).[13] In addition, patients with mixed hearing loss, wound and skin diseases affecting the head area, as well as those who were unable to give appropriate behavioral responses, were excluded from the study.

2.3. Study procedure

The instruments used were 256 and 512 Hz tuning forks made of steel and kept in room air.[14–16] A two-channel audiometer (GSI 61, Grason-Stadler, Inc., USA) with the respective transducers was used for the PTA and AW tests. It was regularly calibrated based on the American National Standards Institute (ANSI S3.6, 1996). Additionally, a calibrated 226-Hz tympanometer (AT235H, Interacoustics, Denmark) was also used to verify the diagnosis of CHL.

The PTA testing was conducted by experienced audiologists in a dedicated sound-proof room. TDH-39 headphones and a B-71 bone vibrator were used to measure the AC and BC thresholds, respectively. The respective thresholds at specific frequencies were recorded (in dB HL) according to the established protocol.[3] Appropriate masking procedures (for both AC and BC tests) were performed as applicable.

The TFW test was performed in a quiet room within the ORL clinic (with the background noise level of less than 60 dB SPL). Figure 1 shows an example of the TFW test procedure. The tuning fork (256 or 512 Hz) was held by the tester at its stem, and either prong was struck against the tester's elbow at its upper one-third end. The stem of the vibrating tuning fork was then placed firmly in the midline of the forehead with the same pressure. Subsequently, the participant was asked whether the tone was heard in the left ear, right

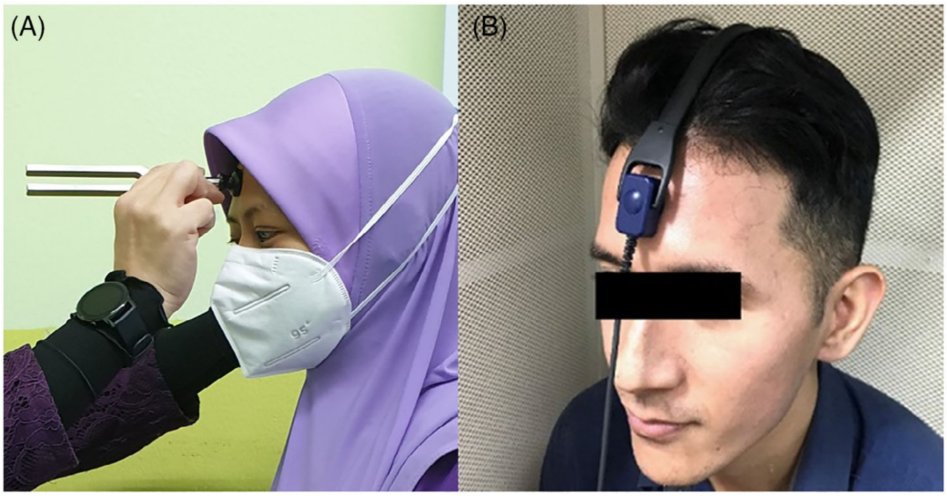

FIGURE 1 The procedure for tuning fork Weber (TFW) test (left panel) and audiometric Weber (AW) test (right panel) employed in the present study

ear, or middle of the head ("central" perception). The TFW test results for both frequencies (256 and 512 Hz) were obtained from all participants.

An example of the AW test procedure is shown in Figure 1. As depicted, the B-71 bone oscillator was placed in the midline of the forehead. A pure tone (at 250 or 500 Hz) was presented at 20 dB SL (i.e., 20 dB above the hearing threshold of that particular frequency) using the audiometer.[12] Similar to the TFW test, each participant was required to report whether the tone was perceived in the left ear, right ear, or middle of the head. The AW test results (for 250 and 500 Hz) were recorded from all respective participants. Of note, the PTA and Weber tests were conducted on the same day.

2.4. Statistical analysis

Descriptive statistics were used to analyze the demographic information. The AC thresholds, BC thresholds, and ABGs of the PTA were described in terms of their mean, *SD* and range. The data for the TFW and AW tests consisted of categorical responses (i.e., "left," "right," and "central"). Using the formula proposed by Rubinstein and Klein,[4] the predicted lateralization response was obtained based on the PTA results. In this regard, to determine the accuracy of TFW and AW tests in comparison to PTA, the sound lateralization results (from the TFW and AW tests) were compared with the predicted lateralization responses. For each participant, the TFW and AW test results were considered "correct" if they were similar to the predicted lateralization results and vice versa. The accuracy of these tests was calculated manually by the researcher and was considered good if the value is at least 80%.[17] To compare the accuracy of these two tests at the respective frequencies, McNemar's test was used. The *p* value of less than 0.05 was considered statistically significant. Likewise, weighted Kappa agreement analysis was used to assess the agreement between the lateralization responses (from the TFW and AW tests) and the expected lateralization results (based on the aforementioned formula). Kappa (k) value of ≥.81 indicates almost perfect agreement, .61 to .80 implies substantial agreement, .41 to .60 indicates moderate agreement, .21 to .40 demonstrates fair agreement, .01 to .20 implies slight agreement, and ≤0 indicates poor agreement.[18] All the data were analyzed using the SPSS software version 24 (SPSS, Inc., Chicago, IL). Additionally, it was also of interest to know the performance of TFW and AW tests in patients with different ABGs. As such, the participants were then categorized into either smaller or larger ABG groups based on the median value at specific frequencies (calculated from ABGs of the worse ear). The accuracy and agreement results were documented accordingly.

3. RESULTS

3.1. Demographic data

The age of the participants ranged from 12 to 67 years, with a mean of 31.3 years. Table 1 reveals the demographic information of the participants. As shown, of 74 participants, 42 (56.8%) were males, and most were Malays (90.5%). Table 2 shows the mean, *SD*, and range for the AC thresholds, BC thresholds, and ABG of the participants for the worse ear. As indicated, the BC thresholds were all within the normal limits. Significant ABGs were noted at all frequencies, and the highest value was observed at 250 Hz (up to 75 dB).

3.2. TFW and AW tests results

Table 3 shows the accuracy and agreement results of TFW and AW tests (in comparison to the predicted lateralization results based on the PTA). Overall, at 256 Hz (or 250 Hz) frequency, the accuracy results of the TFW and AW tests were 81.1% and 86.5%, respectively. Nevertheless, McNemar's test found the accuracy of the two tests to be not statistically different ($p = 0.302$). In terms of the agreement with the PTA, the kappa values were 0.63 and 0.72 for the TFW and AW tests, respectively. On the other hand, at 512 Hz (or 500 Hz)

TABLE 1 Demographic information of participants involved in the present study ($n = 74$)

Demographic data	Variables	*n*
Gender	Males	42
	Females	32
Ethnicity	Malay	67
	Chinese	4
	Others	2
	Indian	1
Background of education	Primary school	8
	Secondary school	45
	Tertiary education	21
Occupations	Students	35
	Housemakers	16
	Professionals	9
	Self-employed	4
	Retired	4
	Others	4
	Unemployed	2
Type of hearing loss/ laterality	Unilateral CHL	59
	Asymmetrical CHL	15
Diagnosis	Mucosal chronic otitis media	30
	Middle ear effusion	16
	Cholesteatoma	7
	Acute suppurative otitis media	4
	Perforated tympanic membrane	4
	Otomycosis	3
	Otitis externa	2
	Atelectatic tympanic membrane	2
	Hemotympanum	1
	Goldenhar syndrome	1
	Others	4

Abbreviation: CHL, conductive hearing loss.

frequency, the accuracy values of the TFW and AW tests were 85.1% and 82.4%, respectively. As revealed by McNemar's test, the performance of the two tests was comparable ($p = 0.804$). The kappa values were 0.72 and 0.68 for the TFW and AW tests, respectively.

The accuracy and agreement outcomes were also determined separately in participants with smaller or larger ABGs. Based on the ABGs of the PTA, the median values were 45 and 32.5 dB for 250 and 500 Hz frequencies, respectively. Based on these median values, smaller and larger ABG groups were established. It is worth noting that a comparable number of participants was achieved after the establishment of the groups. For 250 Hz frequency, the smaller and larger ABG groups had 35 and 39 participants, respectively. For this frequency, the mean ABGs were 26.4 ± 10.0 and 55.5 ± 9.7 dB for the smaller and larger ABG groups, respectively. While for 500 Hz frequency, each group consisted of 37 participants (with mean ABGs of 21.6 ± 6.5 dB and 48.6 ± 10.8 dB for the smaller and larger ABG groups, respectively).

For the smaller ABG group, at 256 Hz (or 250 Hz) frequency, the accuracy results of the TFW and AW tests were 77.7% and 88.5%, respectively (Table 3). As shown by McNemar's test, the accuracy of the AW test was found to be significantly better ($p = .031$). In line with this, the kappa values were .44 and .74 for the TFW and AW tests, respectively. On the contrary, at 512 Hz (or 500 Hz) frequency, the accuracy values of TFW and AW tests were 84.2% and 78.3%, respectively. Nevertheless, the performance of these tests was found to be comparable ($p = .727$ by McNemar's test). In terms of the agreement, the kappa values were .68 and .54 for the TFW and AW tests, respectively.

As shown in Table 3, for the larger ABG group, the accuracy values of 81.5% and 84.6% were noted for the TFW and AW tests, respectively, at 256 Hz (or 250 Hz). However, McNemar's test revealed an insignificant statistical result when the accuracy of the two tests was compared ($p = 1.000$). The kappa values of .55 and .64 were found for the TFW and AW tests, respectively. Likewise, at 512 Hz (or 500 Hz) frequency, the accuracy of the TFW and AW tests were 86.1% and 89.1%, respectively. McNemar's test found the accuracy of the two tests to be not statistically different ($p = 1.000$). The kappa agreement values of .69 and .71 were found for the TFW and AW tests, respectively.

4. DISCUSSION

The present study was performed to unveil the diagnostic accuracy of the AW and TFW tests in comparison to the expected results from PTA. To our knowledge, this is the first study to report the accuracy of the TFW and AW tests in participants with CHL. In many studies concerning the Weber test, the term "sensitivity" has been frequently used to indicate the ability of the test to show correct lateralization in patients with CHL or SNHL.[1,6,7] Nevertheless, in the present study, the term accuracy was preferred for a similar purpose. It is worth mentioning that a more general definition for sensitivity is the ability of a test to correctly identify people who have a specific medical condition.[19] The term sensitivity is perhaps sensible to indicate the ability of the Weber test to correctly detect the abnormal ear among those with unilateral CHL. In contrast, for participants with bilateral asymmetrical CHL (in which both ears are affected), the term sensitivity may not be entirely appropriate to indicate the ability of the test to "find" the worse ear. Accuracy is a more general term and can be defined as the closeness of measured value to the true value.[20] As such, accuracy is the preferred term to demonstrate the performance of the TFW and AW tests in the present study.

Based on the results obtained in the present study, several patterns were observed. First, regardless of frequency, the overall accuracy results of the TFW and AW tests were found to be good (>80%). These findings were supported by the respective kappa agreement results indicating substantial agreement between the tests. These

TABLE 2 Mean, *SD*, and range for AC thresholds, BC thresholds, and ABG of the patients for the worse ear

	AC (dB HL)		BC (dB HL)		ABG (dB)	
Frequency (Hz)	Mean (*SD*)	Range	Mean (*SD*)	Range	Mean (*SD*)	Range
250	48.9 (15.8)	20–80	4.6 (9.8)	−10 to 20	44.4 (6.0)	5–75
500	43.8 (15.6)	20–80	10 (9.0)	−10 to 20	33.8 (6.7)	5–70
1000	41.6 (16.6)	10–80	7.8 (8.4)	−10 to 20	33.8 (8.2)	5–70
2000	38.6 (12.9)	10–65	16.0 (4.9)	0–20	22.6 (8.0)	0–50
4000	43.1 (16.7)	10–80	12.9 (7.9)	−5 to 20	30.2 (8.8)	0–65

Abbreviations: ABG, air-bone gap; AC, air conduction; BC, bone conduction.

TABLE 3 Accuracy and Kappa agreement results for TFW test and AW test for specific ABG categories

	TFW test			AW test		
ABG category	Frequency (Hz)	Accuracy (%)	Kappa (k)	Frequency (Hz)	Accuracy (%)	Kappa (k)
Overall	256	81.1	0.63	250	86.5	0.72
	512	85.1	0.72	500	82.4	0.68
Smaller ABG	256	77.7	0.44	250	88.5	0.74
	512	84.2	0.68	500	78.3	0.54
Larger ABG	256	81.5	0.55	250	84.6	0.64
	512	86.1	0.69	500	89.1	0.71

Abbreviations: ABG, air-bone gap; AW, audiometric Weber; TFW, tuning fork weber.

results implied that both Weber tests were comparatively accurate in identifying patients with unilateral CHL or bilateral asymmetrical CHL. Second, even though the accuracy was found to be comparable between the two tests, a better agreement result was found for the AW test ($k = .72$) relative to the TFW test ($k = .63$) at 250 Hz (or 256 Hz) frequency. On the contrary, at 512 Hz (or 500 Hz), the TFW test produced a higher agreement value ($k = .72$), than the AW test ($k = .68$). It is uncertain as to why different agreement results were produced by the two Weber tests, as well as the "dependency" on frequency. These findings were rather unexpected. Due to many advantages, it was initially thought that better study results would be seen in the AW test when assessing patients with CHL at all frequencies. However, at 512 Hz (or 500 Hz) frequency, the TFW test produced better outcomes than the AW test.

In a study by Thompson,[12] AW and Rinne tests (at 1000 Hz) were carried out on 185 patients (aged 12-75 years) with various hearing loss conditions. As reported, in patients with unilateral CHL, the error for the AW test was 23.5% (indicating <80% accuracy). In fact, the highest error (i.e., 60%) was noted when assessing those with bilateral CHL. The author then concluded that the AW test was of little diagnostic value when assessing patients with hearing loss.[12] When compared with the present study's findings, the lower accuracy results found in the study by Thompson might be due to methodological differences between the studies. Apart from differences in the number of samples and audiometric characteristics, the present study used 250 and 500 Hz frequencies in the AW test, while only 1000 Hz frequency was tested in the study by Thompson.

Based on the existing literature, the accuracy (or sensitivity) outcomes of the TFW test reported in past studies were lower than 80%. Generally, higher accuracy results were reported with the use of a tuning fork at 512 Hz.[1] The highest sensitivity of the TFW test (78%) was reported by Shuman et al.[6] when assessing patients with unilateral idiopathic sudden SNHL. Of 250 patients, 198 (78%) reported correct lateralization results. The authors then stated that the TFW test (at 512 Hz) was a reliable predictor of the PTA results.[6] Surprisingly, the accuracy results of the TFW test were lower in those with unilateral CHL.[1] In a study by Stankiewicz and Mowry,[5] patients (264 ears) with various hearing loss conditions were tested with the TFW test (in conjunction with Rinne and Bing tests). The performances of 256, 512, and 1024 Hz tuning forks were compared as intended. It was found that in those with unilateral CHL, the accuracy of the TFW test was 54% for 512 Hz frequency. The accuracy results for 256 Hz and 1024 Hz frequencies were 43% and 46%, respectively.[5] In the present study (in which the majority of participants had unilateral CHL), the overall accuracy of the TFW test was 81.1% at 256 Hz and 85.1% at 512 Hz. Again, methodological differences (e.g., sample size, audiometric characteristics, etc.) between the studies may have contributed to the different findings.

As revealed, comparatively better accuracy results were found in participants with larger ABGs than those with smaller ABGs. These findings were sensible and in line with those of previous studies.[1,21] Even though the expected lateralization could be perceived when the difference between ears is as small as 2.5 to 4 dB,[1] better lateralization performance would be expected if the interaural differences are larger (as characterized by larger ABGs). It is worth noting that for the smaller ABG group, the accuracy of the AW test (i.e., 88.5%) was found to be significantly higher than that of the TFW test (i.e., 77.7%) ($p = 0.031$ by McNemar's test) at 256 Hz (or 250 Hz) frequency. This suggests that the AW test can provide more accurate results when testing patients with small ABGs at this frequency. The superior performance of the AW test (relative to the TFW test) might be contributed by the fact that the sound can be presented at intended intensity levels constantly (without loudness decay). Nevertheless, further research is warranted to study the performance of the two Weber tests when testing CHL patients with much smaller ABGs (i.e., <15 dB).

As an effort to produce reliable results, the present study recruited participants who were at least 12 years of age. In view of this, less reliable results had been reported when conducting the Weber test among younger children (≤11 years).[21,22] The tuning fork test results are also influenced by the material of the tuning fork and techniques used.[1] In relation to the present study, a recommended material of the tuning fork was used (i.e., steel).[14] Furthermore, the TFW and AW tests were performed by one dedicated tester (who was well trained in conducting tuning fork tests), and a consistent standard technique was applied.

In the present study, around 10% to 20% of the participants did not show correct sound lateralization. Note that they had either unilateral CHL or bilateral asymmetrical CHL (with at least a 15-dB difference between ears). Since the sound lateralization would be perceived even when there is a small difference between the ears (2.5-4 dB),[1] all the participants should be able to correctly lateralise the perceived sound (in both the TFW and AW tests). In this regard, there might be some "hidden" SNHL in the worse ear (that could not be detected by PTA), contributing to the incorrect sound lateralization.[23]

The present study had several limitations. Only low frequencies hearing was tested in both Weber tests. In this regard, it is of interest to know the accuracy of other frequencies (e.g., ≥1000 Hz) in the AW test. Unlike the TFW test, utilizing higher frequencies in the AW test is possible as there is no decay issue. The use of high-frequency stimuli might be beneficial in assessing patients with high-frequency CHL (due to increased middle ear mass conditions) or high-frequency SNHL, which is subjected to further research. Furthermore, the AW test was performed using the conventional B-71 bone transducer. In view of this, it is interesting to unveil the performance of the AW test in patients with hearing loss when tested with the newly designed B-81 bone vibrator that has better acoustical characteristics.[9]

5. CONCLUSION

Even though some differences were noted (in terms of frequency), the overall accuracy of both the AW and TFW tests is reliable when assessing patients with unilateral CHL or bilateral asymmetrical CHL. The accuracy results are supported by substantial agreements between the two Weber tests and the respective PTA results. The AW test can be a good alternative for clinicians who are not well trained in performing tuning fork tests or if the tuning forks are not available. As such, it is recommended for audiologists to conduct the simple AW test in conjunction with other audiological tests to verify the PTA results as masking problems and questionable audiograms are commonly encountered in clinical practice.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Siti Nazira Abdullah https://orcid.org/0000-0002-4066-2098
Mohd Normani Zakaria https://orcid.org/0000-0002-3694-3460
Rosdan Salim https://orcid.org/0000-0002-8724-9014
Mohd Khairi Md Daud https://orcid.org/0000-0002-5056-6446
Nik Adilah Nik Othman https://orcid.org/0000-0002-9800-0898

REFERENCES

1. Kelly EA, Li B, Adams ME. Diagnostic accuracy of tuning fork tests for hearing loss: a systematic review. *Otolaryngol Head Neck Surg*. 2018; 159(2):220-230.
2. Wan Mohamad WN, Romli M, Awang MA, Lih AC, Abdullah R, Zakaria MN. The presence of unusual bone conduction thresholds in pure tone audiometry. *Indian J Otol*. 2020;26(1):54-57.
3. Katz J, Chasin M, English K, Hood L, Tillery KL. *Handbook of Clinical Audiology*. Lippincott Williams & Wilkins; 2015.
4. Rubinstein M, Klein L. The Weber test: its significance in assessing the true value of bone conduction. *Acta Otolaryngol*. 1957;48(3): 266-275.
5. Stankiewicz JA, Mowry HJ. Clinical accuracy of tuning fork tests. *Laryngoscope*. 1979;89(12):1956-1963.
6. Shuman AG, Li X, Halpin CF, Rauch SD, Telian SA. Tuning fork testing in sudden sensorineural hearing loss. *JAMA Intern Med*. 2013;173(8): 706-707.
7. Miltenburg DM. The validity of tuning fork tests in diagnosing hearing loss. *J Otolaryngol*. 1994;23(4):254-259.
8. Boothroyd A, Cawkwell S. Vibrotactile thresholds in pure tone audiometry. *Acta Otolaryngol*. 1970;69(6):381-387.
9. Eichenauer A, Dillon H, Clinch B, Loi T. Effect of bone-conduction harmonic distortions on hearing thresholds. *J Acoust Soc Am*. 2014; 136(2):EL96-EL102.
10. Markle DM, Fowler EP Jr, Moulonguet H. The audiometer Weber test as a means of determining the need for the type of, masking. *Ann Otol Rhinol Laryngol*. 1952;61(3):888-900.
11. Sonnenschein R. Fundamental principles of functional hearing tests: with recent developments in tuning forks and sounding rods. *Arch Otolaryngol*. 1933;18(5):599-613.
12. Thompson AK. Audiometric Weber and Rinne tests as compared to pure-tone thresholds. *J S Afr Speech Hear Assoc*. 1974;21(1):63-70.
13. Durakovic N, Valente M, Goebel JA, Wick CC. What defines asymmetric sensorineural hearing loss? *Laryngoscope*. 2019;129(5):1023-1024.
14. MacKechnie CA, Greenberg JJ, Gerkin RC, et al. Rinne revisited: steel versus aluminum tuning forks. *Otolaryngol Head Neck Surg*. 2013; 149(6):907-913.
15. Bunch C, Fletcher H, Fowler E, Guild S, MacFarlan D, Pohlman A. Report of the committee on methods of testing the hearing by bone conduction. *Ann Otol Rhinol Laryngol*. 1936;45(3):800-821.
16. Woodruff EC. A study of the effects of temperature upon a tuning fork. *Phys Rev (Series I)*. 1903;16:325-355.
17. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203-211.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.

19. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Front Public Health*. 2017;5:307.
20. Binder EM, Krska R. *Romer Labs Guide to Mycotoxins*. Anytime Publishing Services; 2012.
21. Capper JW, Slack RW, Maw AR. Tuning fork tests in children (an evaluation of their usefulness). *J Laryngol Otol*. 1987;101(8):780-783.
22. Behn A, Westerberg BD, Zhang H, Riding KH, Ludemann JP, Kozak FK. Accuracy of the Weber and Rinne tuning fork tests in evaluation of children with otitis media with effusion. *J Otolaryngol*. 2007;36(4):197-202.
23. Blakley BW, Siddique S. A qualitative explanation of the Weber test. *Otolaryngol Head Neck Surg*. 1999;120(1):1-4.

How to cite this article: Abdullah SN, Zakaria MN, Salim R, Md Daud MK, Nik Othman NA. Comparing the diagnostic accuracy of audiometric Weber test and tuning fork Weber test in patients with conductive hearing loss. *Laryngoscope Investigative Otolaryngology*. 2022;7(2):523-529. doi:10.1002/lio2.747