Multilevel temperature-controlled radiofrequency for obstructive sleep apnea: Extended follow-up

DAVID L. STEWARD, MD, EDWARD M. WEAVER, MD, MPH, and B. TUCKER WOODSON, MD, Cincinnati, Ohio, Seattle, Washington, and Milwaukee, Wisconsin

OBJECTIVE: To determine long-term effectiveness of multilevel (tongue and palate) temperature-controlled radiofrequency tissue ablation (TCRFTA) for patients with obstructive sleep apnea syndrome (OSAS).

STUDY DESIGN AND SETTING: Prospective, 2-institution case series. Twenty-nine subjects with mild to moderate OSAS and who were at least 1 year from completion of multilevel TCRFTA were included, representing a subset of subjects who were enrolled in a previously published controlled trial. Exclusion criteria for this extended follow-up study included any additional treatment for OSAS after completion of TCRFTA.

RESULTS: Median follow-up was 23 months. Daytime sleepiness and OSAS-related quality of life were significantly improved at extended follow-up (both \( P < 0.001 \)). Median reaction time testing and apnea-hypopnea index (AHI) were also significantly improved at extended follow-up (both \( P = 0.03 \) and 0.01). Body mass index was unchanged (\( P = 0.94 \)).

CONCLUSIONS: Multilevel TCRFTA treatment of mild to moderate OSAS resulted in prolonged improvement in daytime somnolence, OSAS-related quality of life, psychomotor vigilance, and AHI in this group of subjects at extended follow-up. (Otolaryngol Head Neck Surg 2005;132:630-635.)

Obstructive sleep apnea syndrome (OSAS) is thought to affect between 2% and 4% of adults, resulting in significant morbidity and mortality. Morbidity results primarily from cardiovascular disease; quality of life deficits; and performance deficits caused by loss of alertness and daytime somnolence, increasing risk for motor vehicle accidents. However, successful treatment of OSAS with mechanical devices or surgery has been shown to ameliorate these risks of OSAS.

Nasal continuous positive airway pressure (CPAP), a mechanical device, is an efficacious treatment for severe OSAS. However, long-term effectiveness is limited by inadequate treatment adherence in many patients, especially those with milder OSAS. Other devices such as oral appliances are less efficacious than CPAP, and effectiveness is also dependent upon nightly adherence. Surgical management of OSAS has the advantage of not being dependent upon nightly use of a device but often produces incomplete correction or relapse of symptoms over time. Further, traditional operations are associated with significant perioperative morbidity.

Temperature-controlled radiofrequency tissue ablation (TCRFTA) of the tongue base and palate has been shown to improve OSAS in a randomized sham-placebo-controlled trial with minimal morbidity. Extended follow-up of patients undergoing tongue base TCRFTA for OSAS demonstrated persistent improvements in measures of general health status (short form-36) and daytime sleepiness. This study was designed to assess the long-term effectiveness of multilevel (tongue and palate) TCRFTA to improve quality of life, symptoms, reaction time, and physiology for patients with mild to moderate OSAS. Our hypothesis is that TCRFTA treatment provides long-term improvements in these outcomes, using validated measures.

MATERIALS AND METHODS

Study Design

This was a prospective 2-institutional case series with extended follow-up (minimum, 1 year) involving subjects treated with multilevel (tongue and palate) TCRFTA as part of a previously published randomized, CPAP and sham-placebo controlled trial. This study differs from previous publications involving these patients in that it studies long-term outcomes of TCRFTA.
treatment, whereas previous publications reported on only short-term outcomes.

**Subjects**

Twenty-nine subjects with mild to moderate OSAS (apnea-hypopnea index [AHI] between 5 and 40) and who were at least 1 year from completion of multilevel TCRFTA were included (minimum, 3 total treatment sessions, including at least 1 to tongue base and 1 to palate). This represents a subset of subjects who were enrolled in the original trial and included those who were originally randomized to TCRFTA or those who elected to cross over to active TCRFTA after completion of sham placebo or nasal CPAP therapy. Detailed inclusion criteria for the original study were published elsewhere. Exclusion criteria for this long-term follow-up study included any additional treatment for OSAS after completion of multilevel TCRFTA (nasal CPAP, surgery, and so on). All potentially eligible subjects were offered inclusion in this study. Twenty-nine of 46 potentially eligible subjects agreed to participate (63%). Of those 17 potentially eligible subjects who did not participate, 5 could not be located (moved or bad phone number), 4 refused, and 8 reported that they were too busy or missed scheduled appointments and didn’t follow-up despite multiple phone calls. Three patients were excluded because of additional treatment after completion of TCRFTA (2 CPAP, 1 surgery). Three additional patients undergoing crossover treatment were excluded because they underwent fewer than 3 total treatment sessions or were not multilevel (both tongue and palate). Crossover treatment was not part of the original protocol, and decisions regarding treatment after original protocol completion were individualized such that not all patients underwent crossover treatment with TCRFTA. This study was designed and performed with approval from local institutional review boards and in compliance with the Health Insurance Portability and Accountability Act. All subjects provided informed consent.

**Sleep Studies**

Sleep studies used the home Autoset PDS system (ResMed Corp, San Diego, CA), which was used at both screening baseline and long-term follow-up. Definitions of sleep study indices were published elsewhere.

**Intervention**

TCRFTA was performed with a radiofrequency generator (Gyrus-ENT, Memphis, TN) under local anesthesia in the medical office. Palate treatments consisted of 650 J delivered to midline and 325 J to each side, to create 3 nonoverlapping lesions per treatment session. Tongue-base treatments consisted of 600 to 1100 J, delivered to 1 to 3 sites (nonoverlapping).

**Outcomes**

The primary outcome measures were chosen to represent meaningful measurements of quality of life, daytime sleepiness, and reaction times (psychomotor vigilance). Surrogate outcome measures, including sleep respiratory parameters, were secondary outcomes. Lowest saturation did not improve with TCRFTA in the parent trial, so long-term results are not shown (no improvement). OSAS-specific quality of life was measured with the validated Functional Outcomes of Sleep Questionnaire (FOSQ) and with the validated Symptoms of Nocturnal Obstruction and Related Events (SNORE25) questionnaire. Daytime sleepiness was measured with the validated Epworth Sleepiness Scale (ESS). Reaction times were measured with the validated Psychomotor Vigilance Task (PVT-192; Ambulatory Monitoring Inc, Ardsley, NY) with total test time of 10 minutes and stimulus interval of 2 to 10 seconds. The slowest reaction time (SRT) was defined as the mean 10% slowest reaction times, analyzed as 1/SRT to minimize the contribution of very long lapses. Median and fastest (mean 10% fastest) reaction times were also measured. Acoustic pharyngometry measurements were obtained but are not reported because of poor reproducibility of data.

Long-term outcomes were measured at least 11 months after completion of TCRFTA treatments. Short-term outcomes data are available only in the patients who were originally assigned to TCRFTA treatment and were measured at least 8 weeks after the completion of treatment. Short-term outcomes were compared with long-term outcomes to assess the stability of the observed short-term improvements.

**Data Management and Statistical Methods**

Data were collected on case report forms at each site. Copies were mailed to the sponsor’s data coordinators, who entered the data and visually checked for accuracy. The principal investigator at each treatment site verified data accuracy. Data also were checked statistically, and inconsistencies were resolved with the raw data at each site.

Data are presented as the mean ± SD. Effect sizes were calculated as (posttreatment mean – pretreatment mean)/(pretreatment SD) as per Kazis et al. A positive sign denotes improvement with treatment; a negative sign denotes worsening. Changes in outcome measures were analyzed by 1-sided paired t test for normally distributed variables and by 1-sided sign test for nonnormally distributed variables. One-
sided tests were used because improvement with treatment was expected. Changes in morphologic data were analyzed by 2-sided tests, because changes were not expected. Normality was tested by Shapiro-Wilk W test, and significance was tested by Shapiro-Francia W test. A variable was considered nonnormal if it failed any 1 of these tests. Baseline and changes in outcome variables were reanalyzed after stratifying for significant covariates, and significant findings are presented. The data were analyzed with Stata 8/SE software (Stata Corp., College Station, TX). A P value of <0.05 was considered statistically significant.

RESULTS

Baseline characteristics of the study sample are summarized in Table 1. In general, the subjects were overweight, middle-aged males with moderate OSAS and excessive daytime sleepiness.

TCRFTA treatment is summarized in Table 2. Of the 29 patients who participated in long-term follow-up, 12 were originally randomized to TCRFTA treatment, 9 crossed over after nasal CPAP, and 8 crossed over after sham placebo. Relative to those originally randomized to TCRFTA treatment, crossover patients received fewer tongue treatment sessions (2.2 ± 0.8 sessions versus 4.8 ± 0.6 sessions, P < 0.001) and less total tongue energy (3500 ± 1800 J versus 10,400 ± 1300 J, P < 0.001). Median long-term follow-up after completion of TCRFTA treatment was 23 months (95% confidence interval, 19 – 26 months; range, 11 – 31 months). Median short-term follow-up in the original TCRFTA group was 9 weeks (95% confidence interval, 8 – 11 weeks; range, 8 – 26 weeks).

Results of differences between pretreatment baseline and long-term follow-up are summarized in Table 3. All outcomes were improved at long-term follow-up, with quality of life, daytime sleepiness, median reaction time, and AHI each being statistically significant. These improvements were not related to changes in body mass index or neck circumference (each unchanged).

The standardized magnitudes of the treatment effect for each outcome measure are presented graphically in Figure 1. By convention, a positive effect size denotes improvement in outcome measurement; such that treatment effects are considered negligible if less than 0.2, small if between 0.2 and 0.5, medium if between 0.5 and 0.8, and large if greater than 0.8. On average, OSAS-related quality of life, daytime sleepiness, and AHI were all significantly improved at extended follow-up, with large treatment effects. Reaction time testing was improved with small effect sizes.

We stratified on tongue energy dose to examine low-dose (n = 17, 3700 ± 1900 J) and high-dose (n = 11, 10,700 ± 600 J) subjects. The palate energy was comparable between these groups (2100 ± 900 J versus 2400 ± 300 J, respectively). The low-dose and high-dose groups both improved comparably on all subjective measures (effect sizes, 0.67-1.10, all P < 0.05). The high-dose group had consistently greater benefits on the objective outcome measures (effect sizes, 0.32-0.98), with AHI and apnea index significant (P < 0.05) and slowest and median reaction times trending to significance (P < 0.10). The low-dose group had

---

**Table 1. Baseline characteristics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Norms</th>
<th>N</th>
<th>Mean ± SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td>29</td>
<td>49 ± 9</td>
<td>29-64</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td></td>
<td>29</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index (kg/m²)</td>
<td>&lt; 25</td>
<td>29</td>
<td>29 ± 4</td>
<td>19-36</td>
</tr>
<tr>
<td>Neck circumference (cm)</td>
<td>≤ 42</td>
<td>28</td>
<td>41 ± 3</td>
<td>34-46</td>
</tr>
<tr>
<td>Functional Outcome of Sleep Q</td>
<td>&gt;17.8</td>
<td>29</td>
<td>15.6 ± 2.3</td>
<td>10.0-19.6</td>
</tr>
<tr>
<td>SNORE25</td>
<td></td>
<td>28</td>
<td>1.74 ± 0.62</td>
<td>0.52-2.80</td>
</tr>
<tr>
<td>Epworth Sleepiness Scale</td>
<td>&lt;10</td>
<td>29</td>
<td>12.9 ± 4</td>
<td>6-22</td>
</tr>
<tr>
<td>1/Slowest reaction time (1/msec)</td>
<td>≥ 2.88</td>
<td>29</td>
<td>2.78 ± 0.58</td>
<td>1.53-4.15</td>
</tr>
<tr>
<td>Median reaction time (msec)</td>
<td>≤ 231</td>
<td>29</td>
<td>228 ± 29</td>
<td>189-306</td>
</tr>
<tr>
<td>Fastest reaction time (msec)</td>
<td>≤ 191</td>
<td>29</td>
<td>183 ± 18</td>
<td>159-233</td>
</tr>
<tr>
<td>Apnea-Hypopnea Index (events/h)*</td>
<td>&lt; 5</td>
<td>22</td>
<td>19 ± 11</td>
<td>5-40</td>
</tr>
<tr>
<td>Apnea Index (events/h)*</td>
<td>&lt; 5</td>
<td>22</td>
<td>5 ± 6</td>
<td>0-22</td>
</tr>
</tbody>
</table>

SD, standard deviation; SNORE25, Symptoms of Nocturnal Obstruction and Related Events questionnaire.

*Based on home sleep study.

---

**Table 2. Radiofrequency treatment summary**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tongue</th>
<th>Palate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sessions</td>
<td>3.3 ± 1.5</td>
<td>2.0 ± 0.6</td>
</tr>
<tr>
<td>Total energy (Joules)</td>
<td>6500 ± 3800</td>
<td>2200 ± 800</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation.
mostly negligible to small benefits for objective outcomes (effect sizes, –0.02 to 0.20), except in the case of AHI (effect size, 0.75), and none were statistically significant (all P/h0.15).

Short-term outcome measurements were available only for subjects who were originally assigned TCRFTA treatment (n/h12) and only for quality of life, sleepiness, and reaction time outcomes. In this group, there were no significant changes from short-term (median 9 weeks) to long-term (median 26 months) follow-up for these outcome measures (all P/h0.40). In fact, the point estimate of each outcome improved slightly at long-term follow-up, relative to short-term follow-up (effect size range, 0.07 – 0.28).

DISCUSSION

The results of this study suggest that multilevel (tongue and palate) TCRFTA results in prolonged subjective and objective improvements across treatment outcomes of OSAS in this multi-institutional series of patients. The prolonged symptomatic improvement in this series is consistent with that reported for TCRFTA treatment of the tongue base in a smaller pilot series of subjects with OSAS.9 These data further support the long-term effectiveness of TCRFTA treatment in subjects with OSAS, specifically with respect to quality of life, symptoms, vigilance, and respiratory parameters.

There are several important potential limitations to this study. There is a potential selection bias among those choosing to participate in this follow-up study. Only 3 subjects were excluded because of additional treatment for OSAS after completion of TCRFTA, which suggests that this exclusion criterion introduced minimal selection bias. It was necessary to exclude these subjects to determine the long-term treatment effect of TCRFTA alone. It is possible that the research participants who declined crossover TCRFTA or who chose not to participate in this follow-up study had worse outcomes than did those who chose to participate. However, comparison of baseline (all subjects) and short-term outcomes (original TCRFTA subjects) indicated no significant differences between those who did and did not participate in long-term follow-up (analyses not shown).

The case series study design limits our ability to detect the true long-term treatment effect. As shown in

Table 3. Baseline versus long-term outcomes

<table>
<thead>
<tr>
<th>Outcome variables</th>
<th>N*</th>
<th>Change†</th>
<th>Effect size‡</th>
<th>P value§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional Outcome of Sleep Q</td>
<td>29</td>
<td>1.8 ± 1.9</td>
<td>0.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SNORE25</td>
<td>28</td>
<td>−0.69 ± 0.61</td>
<td>1.11</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Epworth Sleepiness Scale</td>
<td>29</td>
<td>−3.4 ± 3.9</td>
<td>0.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>1/Slowest reaction time (1/msec)</td>
<td>27</td>
<td>0.15 ± 0.51</td>
<td>0.28</td>
<td>0.07</td>
</tr>
<tr>
<td>Median reaction time (msec)</td>
<td>27</td>
<td>−6.1 ± 18.0</td>
<td>0.25</td>
<td>0.03</td>
</tr>
<tr>
<td>Fastest reaction time (msec)</td>
<td>27</td>
<td>−2.9 ± 12.2</td>
<td>0.16</td>
<td>0.22</td>
</tr>
<tr>
<td>Apnea-hypopnea index (events/h)</td>
<td>20</td>
<td>−10.4 ± 11.3</td>
<td>0.87</td>
<td>0.01</td>
</tr>
<tr>
<td>Apnea index (events/h)</td>
<td>18</td>
<td>−2.8 ± 6.2</td>
<td>0.41</td>
<td>0.11</td>
</tr>
<tr>
<td>Other Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck circumference (cm)</td>
<td>26</td>
<td>−0.2 ± 1.2</td>
<td>0.08</td>
<td>0.31</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>29</td>
<td>0.0 ± 1.4</td>
<td>0.00</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation. SNORE25, Symptoms of Nocturnal Obstruction and Related Events questionnaire.

*N = number of subjects with both baseline and long-term outcome data available for each variable.
†Change = long-term outcome – baseline outcome.
‡Effect size = (long-term outcome – baseline outcome)/(baseline standard deviation). Positive effect size indicates improvement; negative indicates worsening.
§P value based on one-sided paired t-test for comparison of means for normally distributed outcome variables or sign test for comparison of medians for non-normally distributed outcome variables (2-sided for other variables). P < 0.05 is significant (bold).

Fig 1. Long-term effect sizes after multilevel TCRFTA treatment. Effect size = (long-term outcome – baseline outcome)/(baseline SD). 1/SRT = slowest reaction time (reciprocal); RT = median reaction time; FRT = fastest reaction time; AI = apnea index. *Statistically significant (P < 0.05) effects.

Thus, no relapse was observed for these clinically important outcomes.

Otolaryngology–Head and Neck Surgery Volume 132 Number 4 STEWARD et al
the parent randomized trial and another blinded TCRFTA trial, there are placebo and blinding effects, which cannot be controlled in a case series. The methodological ideal would be to follow both TCRFTA and placebo patients long-term. However, with demonstrated treatment benefit of TCRFTA relative to placebo at 9 weeks, it is unethical to leave placebo patients untreated for the long term. A goal of this study was to test the hypothesis that the proven treatment effect at 9 weeks persists over an extended time. The results from this case series support this hypothesis.

A further limitation of this study is the use of home sleep studies rather than full in-laboratory studies. The home study provides sleep respiratory parameters that are limited, but sleep study variables were secondary outcomes from the outset. The improvement on respiratory parameters is consistent with the benefits observed for other outcome measures (reaction time, symptoms, and quality of life). The same home sleep study was used for both screening and baseline, which raises the risk of regression to the mean. However, comparison of the screening home sleep study with subsequent baseline polysomnography (used in the parent trial) revealed no consistent patterns of regression to the mean (analyses not shown). At long-term follow-up, we chose to measure respiratory outcome variables with the home sleep study to minimize patient burden and because of budgetary constraints.

The small sample leaves little statistical power to detect important differences in stratified analyses. Thus, we are limited in our ability to identify important treatment or covariate features that impact long-term outcome. The nonrandomized study design further compromises our ability to test the independent treatment effect of these covariates, because confounding variables may distort the observed effect of these covariates. We did carry out various stratified analyses to begin to examine the influence of other variables on outcome.

The subset analysis comparing short-term to long-term outcomes suggests that relapse did not occur. In fact, long-term outcomes were slightly better than short-term outcomes, despite stability of neck size and body mass index. The improvement at 26 months relative to 9 weeks was not statistically significant. Thus, these data do not indicate that outcomes continue to improve significantly more than 9 weeks after treatment is completed, but they do suggest that improvements at 9 weeks do not deteriorate by 26 months.

There was a bimodal distribution of total energy delivered to the tongue. The low-dose subgroup had improvements in long-term subjective outcomes that were comparable to those in the high-dose subgroup. However, on every long-term objective outcome measured (ie, reaction times and sleep respiratory parameters), the high-dose subgroup had long-term improvement superior to that of the low-dose group. This pattern suggests a dose–response effect of the total tongue energy: a positive effect with a low dose and a greater effect with a high dose. Previous research has demonstrated incremental improvement after additional tongue and palate TCRFTA after tongue treatment alone. The current study suggests that additional tongue energy alone improves outcome.

Subset analysis stratified on AHI (mild versus moderate to severe) suggested that both subgroups had significant improvements in the SNORE25 quality of life measurement (data not shown). Only the more severe OSAS group had large treatment effects and significant improvements in the other subjective outcomes and in the sleep respiratory outcomes. It is possible that more severe disease simply allows greater room for improvement. For example, the baseline apnea index was 0.9 events per hour in the mild subgroup and 10.1 events per hour in the moderate to severe subgroup. Thus, a change of −0.9 was the greatest possible improvement in the mild subgroup, whereas a change of −10.1 was possible in the moderate to severe subgroup. The original parent study demonstrated improved short-term outcomes in mild to moderate OSAS patients. The current study suggests that TCRFTA may also be useful in more severe OSAS patients, even though TCRFTA did not resolve OSAS in any of these subjects. Further prospective study will help delineate the role of TCRFTA in severe OSAS patients.

CONCLUSION

The results of this extended follow-up study suggest that multilevel TCRFTA provides significant prolonged improvements in OSAS-related quality of life, daytime sleepiness, psychomotor vigilance, and AHI. These data add to an enlarging body of evidence supporting the effectiveness of TCRFTA treatment for patients with OSAS.

REFERENCES


