THE RISE AND FALL(?) OF UPPP FOR SLEEP APNEA COPYRIGHT NOTICE

Washington University grants permission to use and reproduce the The Rise and Fall(?) of UPPP for Sleep Apnea as it appears in the PDF available here without modification or editing of any kind solely for end user use in investigating rhinosinusitis in clinical care or research (the "Purpose"). For the avoidance of doubt, the Purpose does not include the (i) sale, distribution or transfer of the The Rise and Fall(?) of UPPP for Sleep Apnea or copies thereof for any consideration or commercial value; (ii) the creation of any derivative works, including translations; and/or (iii) use of the The Rise and Fall(?) of UPPP for Sleep Apnea as a marketing tool for the sale of any drug. All copies of the The Rise and Fall(?) of UPPP for Sleep Apnea shall include the following notice: "All rights reserved. Copyright 2009. Washington University in St. Louis, Missouri." Please contact Jay Piccirillo (314-362-8641) for use of the The Rise and Fall(?) of UPPP for Sleep Apnea for any other intended purpose.

"All rights reserved. Copyright 2009. Washington University in St. Louis, Missouri."
The Rise and Fall (?) of UPPP for Sleep Apnea

Jay Piccirillo, MD, FACS
Director, Clinical Outcomes Research Office
Department of Otolaryngology-Head and Neck Surgery
Washington University
Obstructive Sleep Apnea

- First described in 1973(?)
- Before 1981, treatment options:
  - Tracheotomy
  - Tonsillectomy & adenoidectomy
  - Nasal surgery
- In 1981, new treatments options:
  - CPAP (effectively a reverse vacuum cleaner motor)
  - UPPP
**UPPP**

- Ikematsuo first described UPPP in Japan 1952
- Fujita et al first described UPPP in US 1981
- UPPP represented a major surgical advance in the treatment of snoring and OSAS
Fujita UPPP

- Sample size 12
- Mean Apnea Index
  - Pre UPPP 54
  - Post UPPP 28
  - Apnea Index < 20 8/12 (67%)

*Otolaryngol Head Neck Surg* 1981;89:923-934
After Fujita’s introduction much interest and excitement about procedure

UPPP predominate therapy for snoring and sleep apnea

In the 1980’s and early ’90’s - initial good response among carefully selected patients
Sher et al reported on patients who were carefully selected for UPPP and demonstrated by post operative polysomnography (PSG) that 72% of patients had marked improvements in their mean apnea index (AI). Eighty seven percent of these patients showed greater than 50% reduction in AI

*Laryngoscope.* 1985;95:1483-1487
UPPP – A PANACEA

By middle – late 1980’s, UPPP was thought to be the best treatment for sleep apnea and was applied widely. CPAP treatment was terribly cumbersome, and the limitations of UPPP were not routinely recognized.
Relapse Rates

- Mid ’90’s long-term results seemed less promising

- Significant number of patients who underwent UPPP demonstrated polysomnogram findings consistent with relapse of OSA
- **Larson**
  - Six months after surgery, 40% (20/50)
  - Twenty-one months after surgery, 61% (30/49)

- **Janson**
  - Six months after surgery 36% (9/25)
  - Four to eight years after surgery 52% (13/25)
  - None of the 7 patients with preoperative apnea-hypopnea index of more than 40 were responders

_Laryngoscope_. 1994 Nov;104(11 Pt 1):1362-8
Efficacy of Surgical Modifications of the Upper Airway in Adults with OSA

Literature Review

- MEDLINE search 1966 - 1993
- 37 UPPP articles identified; N=992
- 17 articles contained patient-level data on 345 patients

*Sleep* 1996;19(2):156-177
The Efficacy of Surgical Modifications of the Upper Airway in Adults With Obstructive Sleep Apnea Syndrome

An American Sleep Disorders Association Review

Aaron E. Sher¹, Kenneth B. Schechtman² and Jay F. Piccirillo³

Summary: This paper, which has been reviewed and approved by the Board of Directors of the American Sleep Disorders Association, provides the background for the Standards of Practice Committee's parameters for the practice of sleep medicine in North America. The intent of this paper is to provide an overview of the surgical treatment of obstructive sleep apnea syndrome, to provide the basis for the American Sleep Disorders Association's practice parameters on this subject and to share our findings of metaanalysis of previously published studies regarding uvulopalatopharyngoplasty. We searched MEDLINE from January 1966 through April 1993, with an update in February 1995, to provide a review of the application of surgical modifications of the upper airway to treat adults with obstructive sleep apnea syndrome. Operations to treat obstructive sleep apnea syndrome include nasal septal reconstruction; uvulopalatopharyngoplasty; uvulopalatopharyngoglossoplasty; laser midline glossectomy; lingual-plasty; inferior sagittal mandibular osteotomy and genioglossal advancement; with hyoid myotomy and suspension (the entire process is referred to as GAHM); maxillomandibular osteotomy and advancement, and tracheotomy. Papers included in metaanalysis provided preoperative and postoperative polysomnographic data on at least nine patients treated with uvulopalatopharyngoplasty for their obstructive sleep apnea. Analysis of the uvulopalatopharyngoplasty papers revealed that this procedure is, at best, effective in treating less than 50% of patients with obstructive sleep apnea syndrome. The site of pharyngeal narrowing or collapse, although identified by different and unvalidated methods, has a marked effect on the probability of success of uvulopalatopharyngoplasty. Patients who achieve a favorable response with uvulopalatopharyngoplasty tend to have less severe obstructive sleep apnea than those who do not. For patients who demonstrate retrolingual narrowing or collapse, other surgical modifications have been described, such as lingualplasty, GAHM, and maxillomandibular osteotomy and advancement. The studies to support the use of the surgical treatment of obstructive sleep apnea syndrome contain biases related to small sample size, limited follow-up and patient selection.
# UPPP Does Not Work?

Table 8—Response rates correlated to definition of response based on location of pharyngeal narrowing or collapse

<table>
<thead>
<tr>
<th>Variable used to measure response</th>
<th>Type I (n = 111)</th>
<th>Type II or III (n = 57)</th>
<th>Unknown location (n = 177)</th>
<th>p value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% Decrease in AI</td>
<td>65/78 (83.3)</td>
<td>4/21 (19.0)</td>
<td>60/97 (61.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI</td>
<td>47/70 (67.1)</td>
<td>9/38 (23.7)</td>
<td>56/97 (57.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in either AI or RDI</td>
<td>83/109 (76.1)</td>
<td>12/57 (21.1)</td>
<td>100/171 (58.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI and a postoperative RDI &lt; 20 or a 50% decrease in AI and a postoperative AI &lt; 10</td>
<td>57/109 (52.3)</td>
<td>3/57 (5.3)</td>
<td>77/171 (45.0)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

137/337 = 40.7%

Sleep 1996;19:156-177
Table 8—Response rates correlated to definition of response based on location of pharyngeal narrowing or collapse

<table>
<thead>
<tr>
<th>Variable used to measure response</th>
<th>Type I (n = 111)</th>
<th>Type II or III (n = 57)</th>
<th>Unknown location (n = 177)</th>
<th>p value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% Decrease in AI</td>
<td>65/78 (83.3)</td>
<td>4/21 (19.0)</td>
<td>60/97 (61.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI</td>
<td>47/70 (67.1)</td>
<td>9/38 (23.7)</td>
<td>56/97 (57.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in either AI or RDI</td>
<td>83/109 (76.1)</td>
<td>12/57 (21.1)</td>
<td>100/171 (58.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI and a postoperative RDI &lt; 20 or a 50% decrease in AI and a postoperative AI &lt; 10</td>
<td>57/109 (52.3)</td>
<td>3/57 (5.3)</td>
<td>77/171 (45.0)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

$^a$ Log-linear model, log-ratio of type I and II or III

137/337 = 40.7%

Sleep 1996;19:156-177
Table 8—Response rates correlated to definition of response based on location of pharyngeal narrowing or collapse

<table>
<thead>
<tr>
<th>Variable used to measure response</th>
<th>Type I (n = 111)</th>
<th>Type II or III (n = 57)</th>
<th>Unknown location (n = 177)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% Decrease in AI</td>
<td>65/78 (83.3)</td>
<td>4/21 (19.0)</td>
<td>60/97 (61.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI</td>
<td>47/70 (67.1)</td>
<td>9/38 (23.7)</td>
<td>56/97 (57.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in either AI or RDI</td>
<td>83/109 (76.1)</td>
<td>12/57 (21.1)</td>
<td>100/171 (58.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI and a postoperative RDI &lt; 20 or a 50% decrease in AI and a postoperative AI &lt; 10</td>
<td>57/109 (52.3)</td>
<td>3/57 (5.3)</td>
<td>77/171 (45.0)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

137/337 = 40.7%

Sleep 1996;19:156-177
Table 7—Baseline and percentage-change data based on location of pharyngeal narrowing or collapse, expressed as mean ± standard deviation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type I (n = 111)</th>
<th>Type II or III (n = 57)</th>
<th>Unknown location (n = 177)</th>
<th>p value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apnea index (AI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>38.9 ± 26 (n = 80)</td>
<td>59.9 ± 29 (n = 21)</td>
<td>46.4 ± 32 (n = 102)</td>
<td>0.004</td>
</tr>
<tr>
<td>Percentage change&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-74.6 ± 27 (n = 77)</td>
<td>-22.8 ± 29 (n = 20)</td>
<td>-53.6 ± 47 (n = 94)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Respiratory disturbance index (RDI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>56.6 ± 29 (n = 72)</td>
<td>64.5 ± 24 (n = 39)</td>
<td>47.8 ± 30 (n = 103)</td>
<td>0.096</td>
</tr>
<tr>
<td>Percentage change&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-32.7 ± 61 (n = 68)</td>
<td>-6.5 ± 47 (n = 37)</td>
<td>-32.1 ± 58 (n = 96)</td>
<td>0.002</td>
</tr>
<tr>
<td>Minimum O₂ saturation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>63.0 ± 15.8 (n = 55)</td>
<td>61.9 ± 20 (n = 20)</td>
<td>66.9 ± 20 (n = 68)</td>
<td>0.935</td>
</tr>
<tr>
<td>Percentage change</td>
<td>24.7 ± 45 (n = 55)</td>
<td>12.7 ± 45 (n = 12)</td>
<td>32.3 ± 79 (n = 67)</td>
<td>0.049</td>
</tr>
</tbody>
</table>
UPPPP Helps

Palate Obstruction

Adapted from Table 7
Which Is It??

- UPPP does not work: Mostly failures?
- UPPP works: improvement?
UPPP (and sleep surgery) a FAILURE

- This still is a popular view
  - “Success” rates = failure
  - Other sleep surgery besides UPPP less studied and largely ignored
    - Barbara Phillips in JCSM 2005

- Is it true that UPPP (and sleep surgery) is a failure?
Confusion About Efficacy of UPPP

- Mixed results that were interpreted as demonstrating failure – reflecting bias against surgery
- Judgment based on inadequate or inappropriate outcome measures
- Unrealistic claims of benefit of sleep surgery
- Study design and other methodological weaknesses
Clinically Important Endpoints

- Death
- Cardiovascular disease
- Motor vehicle accidents
- Quality of life
5-Year Survival

Adapted from Table 2

Chest 1988;94:9-14
UPPP Survival

5-Year Survival

- Untreated
- Treated

Adapted from Table 2

Chest 1988;94:9-14
**UPPP Survival**

5-Year Survival

<table>
<thead>
<tr>
<th>AI &gt; 20</th>
<th>AI &lt; 20</th>
<th>UPPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.87</td>
<td>0.96</td>
<td>0.85</td>
</tr>
</tbody>
</table>

**UPPP:**
- AI > 60
- UPPP NOT INDICATED

Adapted from Table 2

*Chest 1988;94:9-14*
UPPP Survival

AI > 5

Figure 1

UPPP NOT BAD!!

Chest 1994;105:155-9
Figure 2

UPPP Survival
AI >20

Chest 1994;105:155-9

UPPP NOT BAD!!
UPPP Survival

Survival of veterans with sleep apnea: Continuous positive airway pressure versus surgery

EDWARD M. WEAVER, MD, MPH, CHARLES MAYNARD, PHD, and BEVAN YUEH, MD, MPH, Seattle, Washington

OBJECTIVES: Continuous positive airway pressure (CPAP) improves sleep apnea survival. We tested whether CPAP is associated with better survival than uvulopalatopharyngoplasty (UPPP).

STUDY DESIGN AND METHODS: This retrospective cohort database study included all sleep apnea patients treated with CPAP or UPPP in Veteran Affairs facilities from October 1997 through September 2001. Treatment groups were compared with Cox regression, adjusting for age, gender, race, year treatment was initiated, and comorbidity. Sleep apnea severity and CPAP use data were not available.

RESULTS: By September 2002, 1339 (7.1%) of 18,754 CPAP patients and 71 (3.4%) of 2,072 UPPP patients were dead ($P < 0.001$). After adjustment, CPAP patients had 31% (95% confidence interval, 3% to 67%, $P = 0.03$) higher probability of being dead at any time, relative to UPPP patients.

CONCLUSIONS: UPPP confers a survival advantage over CPAP, after adjustment for age, gender, race, year of treatment, and comorbidity. However, we were unable to adjust for sleep apnea severity or CPAP use. Surgical treatment should be considered in sleep apnea patients who use CPAP inadequately. (Otolaryngol Head Neck Surg 2004;130:659-65.)
UPPP Survival

Figure 1

P<0.0001 by Log Rank
UPPP Mortality

Adjusted* Hazard Ratio of Death

P = 0.03

*Adjusted for age, sex, race, year, comorbidity.

Adapted from Table 3

Oto-HNS 2004;130:659-65
CPAP v UPPP

Outcome

Bad

Good

UPPP CPAP
**CPAP v UPPP**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Users</th>
<th>Non-Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bad</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Increased Incidence of Cardiovascular Disease in Middle-aged Men with Obstructive Sleep Apnea
A 7-Year Follow-up

Yüksel Peker, Jan Hedner, Jeanette Norum, Holger Kraiczi, and Jan Carlson
Sleep Laboratory, Department of Pulmonary Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden

The incidence of a cardiovascular disease (CVD) was explored in a consecutive sleep clinic cohort of 182 middle-aged men (mean age, 46.8 ± 9.3; range, 30–69 years in 1991) with or without obstructive sleep apnea (OSA). All subjects were free of hypertension or other CVD, pulmonary disease, diabetes mellitus, psychiatric disorder, alcohol dependency, as well as malignancy at baseline. Data were collected via the Swedish Hospital Discharge Register covering a 7-year period before December 31, 1998, as well as questionnaires. Effectiveness of OSA treatment initiated during the period as well as age, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP) at baseline, and smoking habits were controlled. The incidence of at least one CVD was observed in 22 of 60 (36.7%) cases with OSA (overnight oxygen desaturations of 30 or more) compared with in 8 of 122 (6.6%) subjects without OSA (p < 0.001). In a multiple logistic regression model, significant predictors of CVD incidence were OSA at baseline (odds ratio [OR] 4.9; 95% confidence interval [CI], 1.8–13.6) and age (OR 23.4; 95% CI, 2.7–197.5) after adjustment for BMI, SBP, and DBP at baseline. In the OSA group, CVD incidence was observed in 21 of 37 (56.8%) incompletely treated cases compared with in 1 of 15 (6.7%) efficiently treated subjects (p < 0.001). In a multiple regression analysis, efficient treatment was associated with a significant risk reduction for CVD incidence (OR 0.1; 95% CI, 0.0–0.7) after adjustment for age and SBP at baseline in the OSA subjects. We conclude that the risk of developing CVD is increased in middle-aged OSA subjects independently of age, BMI, SBP, DBP, and smoking. Furthermore, efficient treatment of OSA reduces the excess CVD risk and may be considered also in relatively mild OSA without regard to daytime sleepiness.
Figure 2: UPPP & Cardiovascular Disease

- CPAP (n=14): 64%
- UPPP (n=22): 50%

- 36%: 50%
Does Uvulopalatopharyngoplasty Inhibit Automobile Accidents?

Per-Olle Haraldsson, MD; Christer Carenfelt, MD; Michael Lysdahl, MD; Claes Tingvall, PhD

Patients with rhonchopathy, which includes obstructive sleep apnea syndrome (OSAS), who report sleepy spells at the wheel do poorly on simulated monotonous driving tests and have a twofold to threefold increase in traffic accidents. To assess whether drivers with rhonchopathy (heavy snoring, sleep disturbances, and daytime sleepiness) cause fewer automobile accidents after uvulopalatopharyngoplasty (UPPP), the car accident rate for the first 5 years after surgery was compared to the rate of the 5 years immediately before the operation. Data were collected by means of a self-report questionnaire. Fifty-six patients with rhonchopathy were compared to 142 controls without rhonchopathy who had been subjected to nasal surgery. The response rates were 96% and 94%, respectively. The reported habitual sleepiness while driving had disappeared in 87% (P<.001) of drivers who had the problem preoperatively. The accident risk reduction (corrected for mileage) in patients was almost four times greater than the reduction in controls (P<.001) after surgery. The relative rate of patients involved in any single-car accident fell by 77% (P<.05), and the relative rate of single-car accidents fell by 83% (P<.001). It is concluded that drivers with rhonchopathy have an increased risk for car accidents, especially single-car accidents, but that this risk returns to normal after UPPP.
# UPPP & Accidents

## Table V

<table>
<thead>
<tr>
<th>Accident Type</th>
<th>Relative Risk: OSA v Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-UPPP</td>
</tr>
<tr>
<td>Single-car</td>
<td>9.6</td>
</tr>
<tr>
<td>Multiple-car</td>
<td>1.9</td>
</tr>
<tr>
<td>All</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Corrected for driving exposure. All p<0.001
UPPP Works

- Reduced mortality
- Reduced cardiovascular disease
- Reduced motor vehicle crashes
- Improved symptoms
- Improved quality of life??
What About UPPP Failures?

- Why do they fail?
- Can we fix them?
### Table 8—Response rates correlated to definition of response based on location of pharyngeal narrowing or collapse

<table>
<thead>
<tr>
<th>Variable used to measure response</th>
<th>Type I (n = 111)</th>
<th>Type II or III (n = 57)</th>
<th>Unknown location (n = 177)</th>
<th>p value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% Decrease in AI</td>
<td>65/78 (83.3)</td>
<td>4/21 (19.0)</td>
<td>60/97 (61.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI</td>
<td>47/70 (67.1)</td>
<td>9/38 (23.7)</td>
<td>56/97 (57.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in either AI or RDI</td>
<td>83/109 (76.1)</td>
<td>12/57 (21.1)</td>
<td>100/171 (58.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>50% Decrease in RDI and a postoperative RDI &lt; 20 or a 50% decrease in AI and a postoperative AI &lt; 10</td>
<td>57/109 (52.3)</td>
<td>3/57 (5.3)</td>
<td>77/171 (45.0)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*UPPP fails when other anatomy obstructs*
Study Design and Other Methodological Weaknesses

- Inadequate sample size and little statistical power
- Failure to include confidence intervals
- Uncontrolled studies
- Inadequate follow-up
- Results with uncertain generalizability
- Quality of life
- Multiple endpoints
- Missing data and missing or inconsistent definitions
- Biased baseline data

Schechtman, Sher, Piccirillo Sleep 1995;18(8):659-666
Study Design and Other Methodological Weaknesses

- Goals of Methodological Review
  - Evaluate the quality of evidence
  - Stimulate higher quality research

- Note: most of these same weaknesses exist in the non-surgical sleep apnea literature as well.
Inadequate Sample Size and Little Statistical Power

- Sample size ranged from 9 to 90 (mean 27, median 22)

- Insufficient power to detect clinically significant differences

- The power of a statistical test is the probability that the test will find a significant difference when there really is a difference between groups
Inadequate Sample Size and Little Statistical Power

- Power is affected by sample size
- The sample size needed to achieve a specific level of power depends on:
  - The amount of variation in the population
  - The p-value for statistical significance
  - The degree of difference the investigator wishes to be able to detect
- Small sample size impairs a study’s ability to detect differences
Failure to Include Confidence Intervals

- Only 1 of the 37 (3%) UPPP papers included confidence intervals
- Confidence intervals bridge the gap between statistical significance and clinical significance
- The 95% confidence interval is the range of values in which the investigator can be 95% confident that the true mean of the population falls
Failure to Include Confidence Intervals

- Confidence intervals provide information about the range of the magnitude of difference.

- Allows one to determine if the difference could be clinically significant or not.
Uncontrolled Studies

- Only 4 of 37 (11%) studies used control groups
- None were randomized clinical trials
- Without a control group, there is no way to determine if outcomes are a result of therapeutic intervention or other factors
- The “improvements” in outcome measures seen in some of the uncontrolled studies may reflect day-to-day variability
Uncontrolled Studies

- Randomization ensures that other unforeseen variables that may affect outcome are balanced between the two groups.

- Randomized controlled trials are the gold standard of clinical research.

- However, randomization raises logistical and ethical challenges.
Inadequate Follow-up

- 7 papers (19%) provide no information on follow-up length
- 5 (14%) provided only minimum follow-up length
- 25 (67%) provided mean follow-up length
Inadequate Follow-up

- Mean follow-up length for the 25 papers was 5 months
- With short follow-up times, no way of knowing if the benefits of UPPP are maintained over time
Results with Uncertain Generalizability

- Difficult to assess the generalizability of results when data is from patients who are not typical of the population of interest

- Occurs in:
  - Prospective studies when outcome data are missing for nonrandom reasons related to the outcomes
  - Retrospective studies that require follow-up information as an entry criterion

- Bias can result when patients with missing follow-up are excluded from retrospective studies
Results with Uncertain Generalizability

- 12 (32%) papers were definitely retrospective
- 11 (30%) were definitely prospective
- 5 (14%) were probably retrospective
- 5 (14%) were probably prospective
- 4 (11%) were indeterminate
Results with Uncertain Generalizability

- Because of missing follow-up information
  - 15 (40%) papers were “definitely biased”
  - 12 (32%) were “definitely not biased”
  - 3 (8%) were probably biased
  - 3 (8%) were probably not biased
  - 4 (11%) were indeterminate

- In biased papers, results were based on an unrepresentative sample
Many articles reported on subjective outcome measure

None used validated subjective outcome measures

The results of objective tests often correlate poorly with symptoms of sleep apnea and the functional impairments associated with it

Quality of life measures give clinical relevance to results
Multiple Endpoints

- The mean number of endpoints was 7
- 8 (22%) papers had 10 or more endpoints
- Multiple endpoints increase the incidence of Type I or “False Positive” error
- Significant results in some of the studies with large numbers of endpoints may have occurred by chance
Missing Data and Missing or Inconsistent Definitions

- 12 (32%) of the UPPP papers did not provide mean follow-up data
- 14 (38%) did not provide sufficient information to determine whether the study was prospective or retrospective
- 12 (32%) papers did not report age of patients
Missing Data and Missing or Inconsistent Definitions

- 15 (40%) papers did not report gender
- 22 (59%) papers did not define OSA
- At least 5 different definitions among the papers that defined OSA
Biased Baseline Data

- When a minimum baseline value of an outcome measure is required as a criterion for enrollment in a study, it is important to perform a baseline assessment separate from the screening assessment.
- Results on screening assessments may reflect day-to-day variability.
- Setting minimum baseline values for enrollment biases the baseline data towards higher values.
Biased Baseline Data

- If the screening values are also used as baseline values, baseline values will be biased estimates of true values.

- Post-treatment values will tend to be lower than pre-treatment values even when there is no therapeutic effect.
Biased Baseline Data

- All 37 papers required a minimum baseline value of an outcome measure for enrollment

- None indicated use of separate screening and baseline assessments
UPPP Works!

- When tonsil and palate obstructs airway
- To improve clinically important outcomes
UPPP Does Not Work

- Consistently to cure all comers
- By itself when tongue obstructs
Conclusion

- Early promise of UPPP raised hopes

- Subsequent studies demonstrated limitations

- Misconceptions about benefit of UPPP due to
  - Bias against surgery
  - Inadequate or inappropriate outcome measures
  - Unrealistic claims of benefit of surgery
  - Methodological weaknesses of the published literature and possible misinterpretation
Summary

- There is a role for sleep surgery
- When patients are selected properly sleep surgery is beneficial
- Selective in our approach to sleep surgery