

Effective Health Care Program

Future Research Needs Paper
Number 12

Future Research Needs for Treatment of Obstructive Sleep Apnea



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Future Research Needs Paper

Number 12

Future Research Needs for Treatment of Obstructive Sleep Apnea

**Identification of Future Research Needs From Comparative Effectiveness Review
No. 32**

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10055 I

Prepared by:

Tufts Evidence-based Practice Center, Tufts Medical Center, Boston, MA

Investigators:

Ethan M. Balk, M.D., M.P.H., *Co-Project Lead*
Mei Chung, Ph.D., M.P.H., *Co-Project Lead*
Jeffrey A. Chan, B.S., *Research Associate*
Denish Moorthy, M.B.B.S., M.S., *Research Associate*
Kamal Patel, M.P.H., M.B.A., *Research Associate*
Thomas W. Concannon, Ph.D., *Investigator*
Sara J. Ratichek, M.A., *Project Coordinator*
Lina Kong Win Chang, B.S., *Research Assistant*

**AHRQ Publication No. 12-EHC033-EF
February 2012**

This report is based on research conducted by the Tufts Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10055 I). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care researchers and funders of research make well-informed decisions in designing and funding research and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of scientific judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical research and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted, for which further reproduction is prohibited without the specific permission of copyright holders.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators has any affiliation or financial involvement that conflicts with the material presented in this report.

Suggested Citation: Balk EM, Chung M, Chan JA, Moorthy D, Patel K, Concannon TW, Ratichek SJ, Chang LKW. Future Research Needs for Treatment of Obstructive Sleep Apnea. Future Research Needs Paper No. 12. (Prepared by the Tufts Evidence-based Practice Center under Contract No. 290-2007-10055 I.) AHRQ Publication No. 12-EHC033-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2012.
www.effectivehealthcare.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

An important part of evidence reports is to not only synthesize the evidence, but also to identify the gaps in evidence that limited the ability to answer the systematic review questions. AHRQ supports EPCs to work with various stakeholders to identify and prioritize the future research that is needed by decisionmakers. This information is provided for researchers and funders of research in these Future Research Needs papers. These papers are made available for public comment and use and may be revised.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The evidence reports undergo public comment prior to their release as a final report.

We welcome comments on this Future Research Needs document. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang M.D., M.P.H.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Elisabeth U. Kato, M.D., M.R.P.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgements

Stakeholder Panel

Elise Berliner, Ph.D.
Agency for Healthcare Research and Quality
Rockville, MD

Carolyn M. D'Ambrosio, M.D., FCCP,
DABSM
Director, The Center for Sleep Medicine
Associate Professor of Medicine
Tufts University School of Medicine
Tufts Medical Center
Boston, MA

Gary D. Foster, Ph.D.
Laura Carnell Professor of Medicine and
Public Health
Director, Center for obesity Research and
Education
Temple University
Philadelphia, PA

Daniel J. Gottlieb, M.D., M.P.H.
Director, Sleep Disorders Center, VA
Boston Healthcare System
Associate Professor of Medicine, Boston
University School of Medicine
Boston, MA

Edward Grandi
Executive Director
American Sleep Apnea Association
Washington, DC

Christian Guilleminault M.D., D.Biol.
Sleep Medicine Division, Stanford
University
Stanford, CA

Joseph A. Hanak, M.D.
Chief, Adult Physical Medicine and
Rehabilitation
Tufts Medical Center
Boston, MA

Sheri Katz, D.D.S.
President of the American Academy of
Dental Sleep Medicine
Darien, IL

Meir Kryger M.D., FRCPC
Physician, Connecticut VA Healthcare
System
Professor, Yale University School of
Medicine
Hartford, CT

Tracy R. Nasca
Founder and Senior Vice President
Talk About Sleep, Inc.
Chaska, MN

Allan I. Pack, M.B.Ch.B., Ph.D.
The John Miclot Professor of Medicine
Chief, Division of Sleep Medicine
Director, Center for Sleep and Circadian
Neurobiology
Perelman School of Medicine at the
University of Pennsylvania
Philadelphia, PA

Sanjay R. Patel M.D., M.S.
Division of Sleep Medicine
Brigham and Women's Hospital
Boston, MA

Jeffrey R. Prinsell, D.M.D., M.D.
Oral & Maxillofacial Surgery
Marietta, GA

Patrick J Strollo, Jr., M.D., FCCP, FAASM
Medical Director, UPMC Sleep Medicine
Center
Professor of Medicine and Clinical and
Translational Science
Division of Pulmonary, Allergy, and Critical
Care Medicine
University of Pittsburgh School of Medicine
Pittsburgh, PA

Victor G. Villagra, M.D., FACP
Chief Medical Officer
National Rural Electric Cooperative
Association
Arlington, VA

Thomas Weakley
Director of Operations Owner-Operator
Independent Drivers Association (OOIDA)
Foundation, Inc.
Grain Valley, MO

Lichuan Ye, Ph.D., R.N.
Boston College William F. Connell School
of Nursing
Boston, MA

Robert W. Zavoski, M.D., M.P.H.
Medical Director
Medical Care Administration
Connecticut Department of Social Services
Hartford, CT

An anonymous Medical Director.

An anonymous Federal agency stakeholder.

Treatment for Obstructive Sleep Apnea: Future Research Needs

Structured Abstract

Background. Obstructive sleep apnea (OSA) is an important public health issue, with challenges for diagnosis and treatment. A recent Comparative Effectiveness Review (CER) found numerous areas with insufficient or low strength of evidence.

Purpose. With the assistance of a panel of representative stakeholders, to identify and prioritize future research needs topics for treatment of OSA.

Methods. Twenty-two panel members represented six stakeholder categories: patients and the public; providers; purchasers of health care; payers; policymakers, and principal investigators. Building on future research needs topics derived from the CER, stakeholders nominated additional topics for discussion. Nominated topics were discussed by stakeholders (excluding product makers) on a secure Web site discussion board. At the close of the discussion period, stakeholders nominated their top 10 future research needs topics based on the Agency for Healthcare Research and Quality Effective Health Care Program selection criteria. From these nominations, the highest priority future research needs were determined and were elaborated upon to include possible study designs to address the topics.

Future Research Needs Topics. The high-priority future needs topics included:

1. What is the impact of treatment of sleep-disordered breathing on major long-term clinical outcomes, including mortality, cardiovascular disease, and diabetes?
 - a. What are long-term outcomes of mandibular advancement device (MAD) treatment?
2. Cost-effectiveness analysis of a management strategy (diagnosis of symptomatic or high-risk patients through treatments of patients diagnosed with OSA), specifically for patients with mild to moderate disease severity.
 - a. Research on continuous positive airway pressure (CPAP) devices that are both economical as well as clinically effective.
3. Comparative trials of different sleep apnea treatments based on patient characteristics
 - a. Trials of CPAP stratified by disease severity.
 - b. Trials of non-CPAP treatments stratified by disease severity.
 - c. Comparison of alternative treatments for patients who do not tolerate CPAP.
4. Trials to improve compliance with CPAP, MAD, and other treatments, particularly evaluating cognitive therapy approaches.
5. What is the association between sleep apnea severity and long-term clinical outcomes?

Fourteen other future research needs topics were discussed.

Challenges. Stakeholder participation in the online discussion board was low. Discussions were begun by only five stakeholders, and only 41 percent of stakeholders participated in the online discussion. The median number of comments across topics was only two. Topic nomination was

done by 17 stakeholders (77 percent). Lessons learned from this Future Research Needs panel discussion can be applied to future panels.

Contents

Executive Summary	ES-1
Background	1
Context	1
Evidence Gaps	1
Methods	4
Identification of Evidence Gaps	4
Stakeholder Panel	4
Identification and Invitation of Individual Stakeholders	5
Introduction of Process to the Stakeholder Panel	5
Iterative Process To Identify Future Research Needs Topics	6
Use of Microsoft® SharePoint	6
Approach to Prioritization	9
Approach to Research Question Development and Considerations for Potential Research Designs	10
Results	11
Composition of Stakeholder Panel	11
Stakeholder Participation and Required Methods Modifications	11
Research Needs	13
High Priority Future Research Needs Topics	13
High Priority Future Research Needs Topic 1	13
High Priority Future Research Needs Topic 2	18
High Priority Future Research Needs Topic 3	20
High Priority Future Research Needs Topic 4	25
High Priority Future Research Needs Topic 5	27
Second-Tier Future Research Needs Topics	30
Second-Tier Future Research Needs Topic 6	30
Second-Tier Future Research Needs Topic 7	30
Second-Tier Future Research Needs Topic 8	31
Second-Tier Future Research Needs Topic 9	32
Other Future Research Needs Topics	32
Other Future Research Needs Topic 10	32
Other Future Research Needs Topic 11	32
Other Future Research Needs Topic 12	32
Other Future Research Needs Topic 13	32
Other Future Research Needs Topic 14	33
Other Future Research Needs Topic 15	33
Other Future Research Needs Topic 16	33
Other Future Research Needs Topic 17	33
Other Future Research Needs Topic 18	33
Other Future Research Needs Topic 19	33
Discussion	34
Challenges in Stakeholder Involvement	34
References	37
Acronyms	38

Tables

Table A. Evidence Gaps Affecting Conclusions for the Key Questions ES-2
Table B. List of Future Research Needs Topics ES-4
Table 1. Evidence Gaps Affecting Conclusions for the Key Questions 3
Table 2. Stakeholders Panel for Treatment of Sleep Apnea 11
Table 3. List of Future Research Needs Topics 14

Figures

Figure A. Analytic Framework for the Treatment of Obstructive Sleep Apnea in Adults
With Evidence Gaps ES-2
Figure 1. Analytic Framework for the Treatment of Obstructive Sleep Apnea in Adults
With Evidence Gaps 2
Figure 2. SharePoint Homepage: Treatment..... 7
Figure 3. Sample Discussion Thread: Treatment..... 8
Figure 4. Sample Size Calculations for an RCT of MAD Versus CPAP With Varying Estimates
of Risk Ratios and Control Rates for Combined Cardiovascular Outcomes 16
Figure 5. Sample Size Calculation for an RCT of Behavioral Interventions To Improve
Compliance With CPAP, MAD, or Other Treatments, Compared With Usual Care 27

Appendixes

Appendix A. AHRQ’s Effective Health Care Program Selection Criteria for New Research

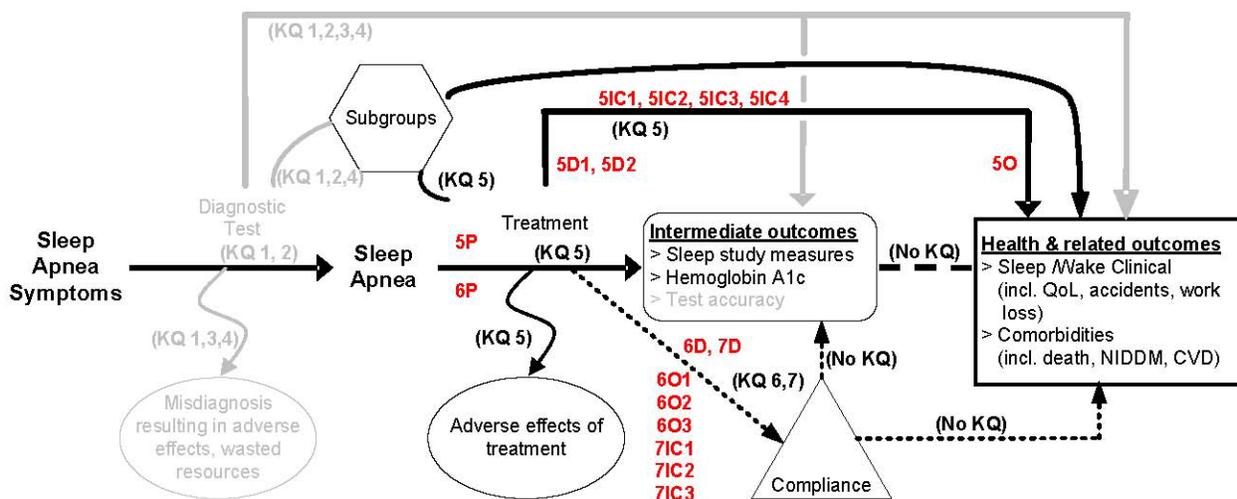
Executive Summary*

Background

The current Future Research Needs (FRN) project was launched shortly after completion of the comparative effectiveness review (CER) on obstructive sleep apnea (OSA). OSA is an important public health issue, due to the considerable mortality and morbidity associated with the condition. The commonly used methods for diagnosing and treating OSA are cumbersome, resource intensive, and often inconvenient for the patient. The Tufts Evidence-based Practice Center (EPC) conducted a CER on diagnostic tools, characteristics of OSA that are predictive of poor outcomes, and treatments for OSA. For the purpose of the FRN process, the original OSA CER was divided into two overarching sections: diagnosis and treatment. This document describes the FRN for treatment of OSA; an accompanying parallel report describes the FRN for diagnosis. The Background, Methods, and description of the challenges are nearly identical between the two reports.

Figure A is an analytic framework to visualize areas of the systematic review in which evidence gaps were identified. Table A summarizes the evidence gaps identified in our review of the treatment of OSA in adults (the CER's Key Questions 5–7). These gaps in the evidence review limited the ability to make conclusions on the questions asked; thus they formed the initial FRN topics.

Figure A. Analytic framework for the treatment of obstructive sleep apnea in adults with evidence gaps



CVD = cardiovascular disease; D = study design; IC = Intervention and comparator; KQ = Key Question; NIDDM = noninsulin dependent diabetes mellitus; O = outcome; P = population; QoL = quality of life
 The analytic framework above highlights the evidence gaps in red that were identified as affecting the conclusions for the respective key questions in the CER. The alphanumeric code for the gaps corresponds to the detailed gaps that are listed in Table 1. The first number of the code corresponds to the key question, the following letters represent the PICOD domains, and the last numerical corresponds to the number on the list for that particular key question and domain. Where there is only one gap identified, the last number is dropped. Grayed out portions of the analytic framework are diagnosis-related questions that are covered in the companion report.³

* Please refer to the main report for references.

Table A. Evidence gaps affecting conclusions for the Key Questions

Key Question	Category	Evidence Gap
Key Question 5. What is the comparative effect of different treatments for obstructive sleep apnea in adults?	Population	5P: MAD tested primarily in a narrow population.
	Intervention/ Comparator	5IC1: Limited data directly comparing different interventions, except autotitrating CPAP vs. fixed CPAP. 5IC2: Except for CPAP and MAD, limited data comparing treatments to control (no treatment). 5IC3: Surgery inadequately tested in unbiased studies compared with CPAP (or no surgery). 5IC4: No subgroup analyses available for any treatment, addressing either subquestion.
Key Question 5a. Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?	Outcomes	5O: Almost no data on clinical outcomes.
	Design	5D1: Incomplete reporting and inadequate analyses common. 5D2: Few high quality studies with little susceptibility to bias.
Key Question 5b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?	Population	6P: Limited data on patients with mild- to moderate-severity OSA.
	Outcomes	6O1: No studies evaluated predictors of compliance with MAD or other treatments. 6O2: Inadequate evidence related to nonsleep measure predictors of poor compliance with CPAP. 6O3: Variable definitions of compliance used.
	Design	6D: Studies commonly selectively reported analyses, in particular failing to report tested but nonsignificant associations.
Key Question 6. In obstructive sleep apnea patients prescribed nonsurgical treatments, what are the associations of pretreatment patient-level characteristics with treatment compliance?	Intervention/ Comparator	7IC1: Large heterogeneity of interventions limited conclusions. 7IC2: Trials rarely used standardized approaches to design treatments. The effects of specific interventions have not been confirmed. 7IC3: Trials addressed only compliance with CPAP.
	Design	7D: Few high-quality studies with little susceptibility to bias.

CPAP = continuous positive airway pressure device; MAD = mandibular advancement device; OSA = obstructive sleep apnea

Methods

Stakeholder Involvement

FRN topics nominated for prioritization were identified based on recommendations from the original CER and through an iterative process utilizing a diverse stakeholder panel. After a formal recruitment process, participating panelists were asked to review, discuss, and nominate for prioritization CER-derived FRN topics, as well as submit additional topics for FRN consideration. The panel consisted of patient advocates, providers, purchasers, payers, policy makers, researchers and product makers, representing the full range of stakeholders who may use research evidence in health care and public health decisionmaking.

Use of Microsoft® SharePoint

The EPC developed a secure, password-protected Microsoft® SharePoint Web site , primarily to host the online stakeholder discussions of the FRN topics. The SharePoint Web site

also functioned as a platform for Tufts EPC staff to post project announcements and reference documents.

Throughout the open discussion period, stakeholders could use SharePoint to submit new FRN topics; a direct email to EPC staff was also an option. The EPC reviewed and refined all stakeholder-submitted topics. EPC staff posted these topics, along with the CER-derived FRN topics, on the SharePoint Web site for stakeholder discussion.

Each FRN topic was posted as a separate “discussion board.” For their discussions, stakeholders were asked to consider four dimensions of need related to the proposed topic: (1) importance; (2) desirability of research/duplication; (3) feasibility, and (4) potential impact. Stakeholder participation was encouraged and monitored throughout the discussion period, and Tufts EPC staff monitored discussion boards daily to ensure appropriateness and relevance of all comments.

Approach to Prioritization

After the close of the online discussion period, stakeholders were asked to identify up to 10 FRN topics—considering the topic’s importance, desirability, feasibility, and potential impact—that were of highest priority and that met the Effective Health Care Program selection criteria. Nomination was conducted by email and individual phone calls. After nomination, the EPC grouped similar topics into overarching topics and categorized them into four groups: (1) high-priority topics; (2) second-tier priority topics; (3) topics of little interest to stakeholders, and (4) topics not meeting Effective Health Care Program appropriateness criteria.

Research Question Development and Considerations for Potential Research Design

For each “high-priority” FRN topic, EPC staff considered the range of study designs that would best address the topic. To determine candidate study designs, the feasibility of the study designs, and sample size calculations, we followed the structure laid out in the Future Research Needs document “Framework for Considering Study Designs for Future Research Needs.”⁴ For each topic, we described our assumptions about the most appropriate PICOD (Population, Intervention, Comparator, Outcomes, study Design) criteria, in particular describing the advantages and disadvantages of various potential research designs. We specifically considered the feasibility of the research questions focusing on potential sample size, time, and recruitment issues.

Results

Table B lists the final FRN topics.

Table B. List of future research needs topics

High Priority Future Research Needs Topics
1. What is the impact of treatment of sleep-disordered breathing on major long-term clinical outcomes, including mortality, cardiovascular disease, and diabetes? a. What are long-term outcomes of mandibular advancement devices (MAD) treatment?
2. Cost-effectiveness analysis of a management strategy (diagnosis of symptomatic or high-risk patients through treatments of patients diagnosed with OSA), specifically for patients with mild to moderate disease severity a. Research on CPAP devices that are both economical as well as clinically effective
3. Comparative studies of different sleep apnea treatments based on patient characteristics a. Analyses of CPAP stratified by disease severity b. Analyses of non-CPAP treatments stratified by disease severity c. Comparison of alternative treatments for patients who do not tolerate CPAP
4. Trials to improve compliance with CPAP, MAD, and other treatments, particularly evaluating cognitive therapy approaches
5. What is the association between sleep apnea severity and long-term clinical outcomes?
Second-Tier Future Research Needs Topics
6. What are the barriers to, and predictors of, compliance with different treatments?
7. Direct comparison of compliance rates with different interventions and incorporation of compliance into an overall comparison of effective treatment
8. Trials to evaluate weight-loss programs as an adjunctive treatment for sleep apnea a. What is the value of bariatric surgery for treatment of sleep apnea?
9. What is the value of including a sleep medicine specialist in the management of the patient with obstructive sleep apnea?
Other Future Research Needs Topics
10. What is consumer willingness to pay for treatment, to identify consumer preferences for strategies to treat sleep apnea?
11. What are the financial barriers to access to treatment?
12. Role of surgery for treatment of OSA a. Comparison of surgery versus CPAP b. Role of orthognathic surgery (corrective jaw surgery) c. Comparison of genio-tubercle advancement versus dental devices
13. Evaluation of postoperative CPAP for all patients with OSA or at high risk of OSA undergoing any surgery with sedation
14. Trials comparing CPAP versus pharmaceutical interventions
15. Trials comparing different CPAP masks
16. Trials comparing CPAP versus oropharyngeal exercises
17. Trials comparing different degrees of mandibular advancement
18. Studies of factors influencing therapist decisions concerning CPAP mask choice
19. Research into how to maintain patients in sleep apnea studies (where dropout rates are unacceptably high)

High-Priority Future Research Needs Topic 1

What is the impact of treatment of sleep-disordered breathing on major long-term clinical outcomes, including mortality, cardiovascular disease, and diabetes?

- a. What are long-term outcomes of mandibular advancement devices (MAD) treatment?

There was insufficient evidence in the CER on the effect of treatment of OSA on clinical outcomes, with only three studies out of 190 comparative studies reporting clinical outcomes. The clinical outcomes of interest are mortality, cardiovascular disease, and diabetes, including surrogate or intermediate markers of the latter two conditions. In assessment of comparative effectiveness, the two overarching comparisons of interest would be autotitrating (and other nonfixed) continuous positive airway pressure (CPAP) devices versus fixed CPAP devices and

CPAP devices versus the mandibular advancement device (MAD). To address the overarching FRN topic and its subcomponents, different study designs are proposed.

Randomized Controlled Trials

Any future trials of CPAP or MAD should study clinical outcomes. However, these could result in the use of large resources in terms of cost, time, and effort, and the decision to use these resources for conducting trials has to be balanced against the value of information that can be gained from studying clinical outcomes. Based on the event rate in the control arm and the assumed value of relative risk, the number of trial subjects that would be required would vary from less than 100 to around 2,000. Regardless of the relative difference in effects between groups, trials using mortality as an outcome would need a sample size in the thousands, while using surrogate outcomes would need smaller sample sizes. The trial duration and followup for mortality as an outcome would be measured in years, while trials focusing on surrogate outcomes for cardiovascular disease and diabetes can be relatively short term, for example 6 to 12 months. Loss to followup would be a major concern, along with compliance where many patients cross over from one treatment to another.

Post Hoc Analyses of Existing Observational Studies

A second-tier option for using existing data to evaluate the effects of different treatment on clinical outcomes would be to perform post hoc regressions or subgroup analyses within observational cohort studies like the Framingham Heart Study, the Wisconsin Sleep Study, or the Nurses Health Study. However, such post hoc analyses may be susceptible to type I error (falsely significant associations) due to multiple testing. These analyses of existing data can be done quickly with modest resources.

Observational Studies

Though less informative than randomized trials for the evaluation of comparative effectiveness, observational studies would be less resource intensive in assessing the impact of various treatments. Case-control studies of patients treated for OSA, with or without a clinical outcome (e.g., a myocardial infarction), could also address this question, though they would be subject to all the biases and other problems of any case-control study. Statistical adjustments like propensity score matching could be used to mimic random assignment. Retrospective data could be gathered quickly though prospective cohorts would take relatively more time and effort.

High-Priority Future Research Needs Topic 2

Cost-effectiveness analysis of a management strategy (diagnosis of symptomatic or high-risk patients through treatments of patients diagnosed with OSA), specifically for patients with mild to moderate disease severity

- a. Research on CPAP devices that are both economical as well as clinically effective

Cost-effectiveness analysis allows the comparison of different interventions on similar benefit, cost, and utility scales. Benefit and cost estimates can have both internal and external validity. The objective of this FRN is to establish better evidence about the costs and benefits of alternative treatment strategies for individual patients with mild to moderate disease severity.

This FRN is distinguished from FRN topic 3 (below) on three dimensions: (1) it assesses the expected value (not just benefit) of selected treatments for OSA; (2) it is designed to address these questions for patients with mild to moderate disease, an understudied group; and (3) it proposes to assess expected value from both the societal and patient perspectives. Out-of-pocket expenses for patients with mild to moderate disease were of considerable concern to patients in our stakeholder group. This FRN seeks to address that concern. The companion report presents a similar discussion on the costs and benefits of alternative diagnosis strategies.

Systematic Review

Cost-effectiveness analyses were not addressed by the CER. Conducting a systematic review may be the first step to ascertain the level of existing evidence.

Cost–Benefit Analysis

A quality-adjusted cost–benefit analysis is recommended, comparing the incremental costs and benefits of different treatments to each other and to a base case of no active treatment. Benefit, utility, and cost estimates may be derived from previous clinical trial data, where available, and otherwise from observational data. These estimates should include clinical, work-related, accident, and quality of life outcomes. Out-of-pocket patient costs should also be included in cost estimates. A societal perspective should be assumed in the main analysis and the patient perspective in a subanalysis. A variety of interventions should be considered, including but not limited to fixed and alternative CPAP, MAD, oropharyngeal and bariatric surgery (both as indicated), phased and combined interventions, and no treatment. Evidence for the effect of most of these treatments on patient-related outcomes is of low strength or insufficient; therefore, sensitivity analyses may be of great value, including ad hoc sensitivity estimates of how effective an intervention must be to be a cost-effective option. Because a cost-effectiveness analysis can draw from previously collected data, the cost, size and duration of such studies can be limited. However, as discussed by the stakeholders, it may be challenging to gather all the relevant data.

High-Priority Future Research Needs Topic 3

Comparative studies of different sleep apnea treatments based on patient characteristics

- a. Analyses of CPAP stratified by disease severity
- b. Analyses of non-CPAP treatments stratified by disease severity
- c. Comparison of alternative treatments for patients who do not tolerate CPAP

For all comparisons between treatments, the CER concluded that the strength of evidence is insufficient to determine which patients might benefit most from treatment or from which treatment. Trials of sleep apnea treatments do not compare their effectiveness in different patient subgroups, or groups stratified by obesity measures, sex, and other patient characteristics. There is a need for subgroup analyses to help clinicians base treatment decisions on baseline characteristics. These analyses can help maximize early effective, tolerable treatment, to better match treatments to specific patient types, and to minimize costly trial and error.

Patient-Level Meta-analysis

For several treatments (particularly CPAP and MAD), there are adequate trials to allow meta-analysis, but patient-level meta-analysis would be required to evaluate the differential effects of treatments in different populations. This approach allows unbiased regressions and subgroup analyses, as could be conducted in any individual trial, avoiding ecological fallacy and providing sufficient power to account for interactions. The analyses could be done quickly and with relatively few resources. However, acquiring the data is the usual main obstacle to conducting patient-level meta-analyses.

Post Hoc Analyses of Existing Trials

Reanalysis of existing trials could also provide unbiased regressions and subgroup analyses. These could also be conducted quickly and with limited resources. Acquiring the data from single studies can be relatively easy (or a moot point if investigators reanalyze their own data). Such post hoc analyses may be susceptible to both type I error (falsely significant associations) due to multiple testing or type II error (falsely nonsignificant associations) due to lack of statistical power.

Randomized Controlled Trials

In general, there is limited or no need for future randomized controlled trials (RCTs) of CPAP or MAD. But future RCTs could investigate whether an interaction term (between the main effect and the predictor variable—the subgroup factor) is statistically significant. For example, the analysis of interest would be whether the relative effectiveness of CPAP and MAD are different in different subpopulations. An a priori RCT would be more definitive than a post hoc analysis, but such a trial would need to be substantially larger than an equivalent trial evaluating only the main effect.

Observational Studies

Prospective or retrospective observational studies could evaluate the predictors of effective treatments using multivariable regression to determine which patient characteristics are associated with more successful treatment. Observational studies, particularly retrospective ones, could be conducted with fewer resources than RCTs, but would still need to be large to be powered for the multivariable analyses and would suffer from the inherent biases of observational data.

High-Priority Future Research Needs Topic 4

Trials to improve compliance with CPAP, MAD, and other treatments, particularly evaluating cognitive therapy approaches

The CER concluded that there was a low strength of evidence that some specific adjunct interventions may improve CPAP compliance compared with usual care. None of the included trials reported results on clinical outcomes. No general type of intervention was found to be more promising than others, and these trials generally had small sample sizes with less than 1 year of followup. Particular interventions discussed by the stakeholders included intensive patient education, social support programs engaging spouse or other family support, cognitive behavioral therapy, feedback from daily adherence monitoring, and peer support in a Web-based

fashion. The patient advocates on the stakeholder panel described a demand by patients for better training, education, and other means to improve compliance.

RCTs of Multicomponent Behavioral Interventions

A well-done RCT will produce the most convincing results, and if patient eligibility criteria and study setting are realistic, it should be fairly applicable to the majority of patients. Several study design considerations need to be addressed before a RCT can be properly designed and carried out, including using established metrics and objective (not self- or provider-reported) data to measure compliance, adopting behavioral change models and theories for designing interventions, and conducting program evaluation during trial period. It would be of greater interest to show long-term (several years) effects on compliance, since in most cases, treatment can be expected to be lifelong.

High-Priority Future Research Needs Topic 5

What is the association between sleep apnea severity and long-term clinical outcomes?

The CER reported strong evidence demonstrating that baseline AHI >30 events/hr is an independent predictor of all-cause mortality over several years of followup. Evidence for the association of baseline AHI and other long-term clinical outcomes, however, is lacking. To address this research gap, the CER recommended patient-level meta-analyses of existing data from available large cohorts of individuals who had sleep testing. Stakeholders on the FRN panel echoed the need for evidence on the association of baseline AHI and other long-term clinical outcomes—particularly to inform design of studies to assess long-term efficacy of treatment.

Patient-Level Meta-analysis

Standard meta-analysis of existing studies is not feasible due to the large degree of clinical heterogeneity in existing studies. Patient-level meta-analysis could overcome these problems by incorporating the differences into a regression analysis and performing a unitary analysis. Existing studies such as the Sleep Heart Health Study and the Wisconsin Sleep Cohort Study can be assessed to determine whether their data include a baseline measure of sleep apnea severity (e.g., AHI), and relevant demographic and clinical variables to allow for patient-level meta-analysis. This approach allows for unbiased regressions and subgroup analyses. Meta-regression could be used to evaluate associations between sleep apnea severity, and other patient characteristics, with long-term clinical outcomes. The analyses could be done quickly and with relatively few resources. However, acquiring the data is the usual main obstacle to conducting patient-level meta-analyses.

Post Hoc Analyses of Existing Studies

A second-tier option would be to perform post hoc regressions or subgroup analyses within already published cohort studies. These could also be conducted quickly and with limited resources. Acquiring the data from single studies can be relatively easy. Such post hoc analyses may be susceptible to both type I error (falsely significant associations) due to multiple testing or type II error (falsely nonsignificant associations) due to lack of statistical power.

Prospective Natural History Study

If existing studies are inappropriate for post hoc or meta-analyses, prospective natural history studies can be conducted to investigate potential associations between sleep apnea severity and long-term clinical outcomes. Studies would require assessment of baseline sleep apnea severity in a specified patient population and documentation of all relevant demographic and clinical variables over time. Moderate resources may be necessary to recruit and follow a large cohort over a long timeframe.

Discussion

We implemented a Web-based discussion board in preference over a series of teleconferences because of what we believed would be advantages of the online approach. Particular advantages include greater flexibility for stakeholders; full participation by all stakeholders in all discussions; a full record of all discussions; less time expenditure by stakeholders; and less resource expenditure by EPC staff. However, we encountered low participation rate of stakeholders during most stages of the project. In contrast, conducting an 1.5-hour teleconference for three consumer stakeholders produced nine topic nominations, highlighting the difference in impact between the two communication platforms. Overall, although SharePoint offered the convenience of asynchronous collaboration it seemed to lack appropriate incentives to engage stakeholders.

Several possible solutions exist for these problems; including reducing the number of stakeholders; reverting to conducting a series of teleconferences; using focus groups to allow a full, simultaneous discussion of each topic; combining teleconferences with online discussions; or numerous other similar approaches. Regardless of which type of discussion is held, there may be advantages to limiting the size of the stakeholder panel and to choosing stakeholders who show enthusiasm in joining the panel and participating in discussions.

In summary, the online Web site discussion of numerous topics by a large stakeholder panel was only moderately successful. Various approaches to improve participation and discussion are possible, including increased use of teleconferences, restricting the size and members of the stakeholder panel, and other approaches. More experience with different approaches is needed.

Background

Context

Comparative Effectiveness Reviews (CER)—systematic reviews of existing research on the effectiveness, comparative effectiveness, and comparative harms of different health care interventions—are intended to provide relevant evidence to inform real-world health care decisions for patients, providers, and policymakers. In addition to synthesizing the evidence, CERs also identify the gaps in evidence that limit the ability to answer the key research questions. As part of an effort beginning in 2010, the Agency for Healthcare Research and Quality supports its Evidence-based Practice Centers (EPCs) to work with various stakeholders to further develop and prioritize the future research needed by decisionmakers.² This process is new, and the methods to delineate future research needs (FRN) are not yet fully developed. The current report describes the first experience of the Tufts EPC in a stakeholder-driven process to identify and nominate for prioritization FRN topics (other than a pilot process with a limited involvement of a small number of stakeholders).

The FRN document is intended to inform and support researchers and those who fund research to ultimately enhance the body of comparative effectiveness evidence so that it is useful for decisionmakers. This document describes the process of developing a prioritized list of research needs with considerations of the advantages and disadvantages of various potential research designs to help researchers and funders develop future research proposals or solicitations, respectively. This process begins with identification of evidence gaps from the original CER, followed by the addition of other areas potentially requiring further research, nomination for prioritization of these evidence gaps by stakeholders, and development of potential study designs for the highest priority topics. Although researchers and funders of research are the end-users of the report, the resulting research is meant to improve health care decisions; therefore, the stakeholders for this process include patients, clinicians, research investigators, payers, and policymakers.

The current FRN project was launched shortly after completion of the CER on obstructive sleep apnea (OSA). OSA is an important public health issue, due to the considerable mortality and morbidity associated with the condition. The commonly used methods for diagnosing and treating OSA are cumbersome, resource intensive, and often inconvenient for the patient. The Tufts EPC conducted a CER on diagnostic tools, characteristics of OSA that are predictive of poor outcomes, and treatments for OSA.¹ For the purpose of the FRN process, the original OSA CER was divided into two overarching sections: diagnosis and treatment. This document describes the FRN for treatment of OSA; an accompanying parallel report describes the FRN for diagnosis.³ For the most part, the Background, Methods, and description of the challenges are nearly identical between the two reports.

Evidence Gaps

Figure 1 is an analytic framework to visualize areas of the systematic review in which evidence gaps were identified. Table 1 summarizes the evidence gaps identified in our review of the diagnosis of OSA in adults (the CER's Key Questions 5–7). These gaps in the evidence review limited the ability to make conclusions on the questions asked; thus they formed the initial FRN topics.

Table 1. Evidence gaps affecting conclusions for the Key Questions

Key Question	Category	Evidence Gap
Key Question 5. What is the comparative effect of different treatments for obstructive sleep apnea in adults?	Population	5P: MAD tested primarily in a narrow population.
	Intervention/ Comparator	5IC1: Limited data directly comparing different interventions, except autotitrating CPAP vs. fixed CPAP. 5IC2: Except for CPAP and MAD, limited data comparing treatments to control (no treatment). 5IC3: Surgery inadequately tested in unbiased studies compared with CPAP (or no surgery). 5IC4: No subgroup analyses available for any treatment, addressing either subquestion.
Key Question 5a. Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?	Outcomes	5O: Almost no data on clinical outcomes.
	Design	5D1: Incomplete reporting and inadequate analyses common. 5D2: Few high quality studies with little susceptibility to bias.
Key Question 5b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?	Population	6P: Limited data on patients with mild- to moderate-severity OSA.
	Outcomes	6O1: No studies evaluated predictors of compliance with MAD or other treatments. 6O2: Inadequate evidence related to nonsleep measure predictors of poor compliance with CPAP. 6O3: Variable definitions of compliance used.
	Design	6D: Studies commonly selectively reported analyses, in particular failing to report tested but nonsignificant associations.
Key Question 6. In obstructive sleep apnea patients prescribed nonsurgical treatments, what are the associations of pretreatment patient-level characteristics with treatment compliance?	Intervention/ Comparator	7IC1: Large heterogeneity of interventions limited conclusions. 7IC2: Trials rarely used standardized approaches to design treatments. The effects of specific interventions have not been confirmed. 7IC3: Trials addressed only compliance with CPAP.
	Design	7D: Few high-quality studies with little susceptibility to bias.

CPAP = continuous positive airway pressure device; MAD = mandibular advancement device; OSA = obstructive sleep apnea

Methods

The Methods section describes the a priori protocol methods as planned prior to convening the stakeholder panel. Much of the process did not go as planned, so several major modifications were made to the protocol during the process of developing the Future Research Needs (FRN) document. These modifications are described in the first part of the Results section. A description of lessons learned and suggestions for further modifications to the methods are in the Discussion section.

Identification of Evidence Gaps

We used an iterative process with a stakeholder panel to identify FRN topics for prioritization. From the Research Needs section of the original Comparative Effectiveness Review (CER), the Evidence-based Practice Center (EPC) generated an initial list of FRN topics and then solicited additional topics from the stakeholder panel.

Stakeholder Panel

We considered seven stakeholder categories to build a panel representing the full range of stakeholders who may use research evidence in health care and public health decisionmaking.

1. **Patients and the Public:** This group represents current and potential consumers of patient-centered health care and population-focused public health. This group also includes caregivers, family members and patient advocacy organizations, all of whom address the interests of consumers.
2. **Providers:** This group includes individuals (e.g., nurses, physicians, mental health counselors and other providers of care and support services) and organizations (e.g. hospitals, clinics, community health centers, community based organizations, pharmacies, EMS agencies, skilled nursing facilities, schools) that provide care to patients and populations.
3. **Purchasers:** This group includes employers, the self-insured, government, and other entities responsible for underwriting the costs of healthcare.
4. **Payers:** This group represents insurers, Medicare and Medicaid, individuals with deductibles, and others responsible for reimbursement for interventions and episodes of care.
5. **Policymakers:** This group includes organizations such as the White House, the U.S. Department of Health and Human Services, Congress, States, professional associations, and intermediary groups that collate and distribute information to policymakers.
6. **Principal investigators, researchers, and research funders.**
7. **Product makers:** This group represents drug and device manufacturers

These categories are not necessarily mutually exclusive. In concept, each individual represents at least one key type of health care and health care research decisionmaker. Any single person or entity may have several roles and may be responsible for different types of decisions. For example, some health care purchasers are also payers, and conversely, some payers also provide care. Patients and their advocates may be providers or employers with policymaking responsibilities, and so on. In addition, each of these seven stakeholder types may be focused on

applying CER at the patient level or at the population level. Patient-level decisions include questions pertaining to what treatment would be best for a given patient at a particular time. Population level decisions include questions pertaining to what services, resources, policies or other alternatives are best for groups of patients and entire communities that are connected by practice setting, geography, clinical domain or other cluster. To be patient centered, decisions made about groups of patients must recognize both the diversity of needs across populations and the heterogeneity of individuals within populations.

Identification and Invitation of Individual Stakeholders

We compiled a list of potential stakeholders from three sources: (1) those individuals who were or had been invited to be either a member of the sleep apnea report's Key Informant panel, Technical Expert Panel, or peer reviewer; (2) a previously compiled list of stakeholders assembled for a 2010 Stakeholder Forum on Comparative Effectiveness Research for the National Institutes of Health Clinical and Translational Science Awards program; and (3) a list of stakeholders compiled in 2010 for the Agency for Healthcare Research and Quality (AHRQ) Effective Health Care Program. We also solicited recommendations for stakeholders from selected government agencies, professional organizations, and other representative bodies. We selected people who would potentially fall within one or more of the stakeholder categories for either OSA diagnosis or treatment. With the assistance of our local domain expert, we initially selected the most promising individuals based on their perceived interest in the topic, their level of previous participation in discussions on the topic, and their fit into the stakeholder categories.

Individuals who met criteria to be a stakeholder were contacted directly by email with a brief description of our project, an invitation to be a stakeholder for one or both stakeholder panels (diagnosis and treatment), and the Executive Summary of our sleep apnea CER. Potential stakeholders were also telephoned, as necessary, to solicit their interest. Other individuals on our list were contacted specifically requesting suggestions for appropriate stakeholders. These people, in turn, were contacted by email. In addition, through Internet searches and focused searches in MEDLINE, we found other potential stakeholders or individuals who could suggest stakeholders. In particular, we contacted senior level administrators at various governmental, nongovernmental, and professional organizations, including but not limited to, the National Association of Community Health Centers, the Veterans Administration, Blue Cross Blue Shield, Centers for Medicare and Medicaid Services, Academy Health, the National Heart Lung Blood Institute, the Office of Minority Health, and the Centers for Disease Control and Prevention. Patients were solicited from our Key Informants list, personal contacts, and an Internet search of sleep apnea advocacy, support, or discussion groups. Selected potential stakeholders were invited to participate in both Future Research Needs Prioritization projects on obstructive sleep apnea (OSA) (diagnosis and treatment), which were run concurrently. All stakeholders completed a standard disclosure of interest form.

We also contacted manufacturers from the list of companies that were sent Scientific Information Packets for the original sleep apnea report. These companies were asked to provide potential FRN topics with rationales, but were not invited to be stakeholders.

Introduction of Process to the Stakeholder Panel

Along with an invitation letter, we distributed the executive summary and the Future Research section of the original CER¹ to the invited stakeholders. The original Key Questions, summary of evidence table, and the implications for future research sections in the executive

summary were highlighted. The purpose of the FRN project and expectations for the input from the stakeholders were outlined clearly in the invitation letter.

We scheduled teleconferences to allow general introductions and for the EPC to explain the purpose and process of the FRN topic development process, after compiling the stakeholder panel.

Iterative Process To Identify Future Research Needs Topics

The EPC reviewed the CER Future Research section and, from this, developed a series of FRN topics. We wrote a brief rationale statement for each. These formed the initial list of FRN topics.

After stakeholders submitted their disclosure of interest statements, we invited them to submit FRN topics to the EPC. For each FRN topic, stakeholders were asked to provide a brief rationale (maximum 250 words) considering the four dimensions of need as listed below under Approach to Prioritization. We reviewed submitted FRN topics, and planned to categorize each topic into one of three categories: “definitely relevant to Key Questions,” “not relevant to Key Questions,” or “unclear.” We planned to combine duplicate or similar FRN topics together into one topic. For FRN topics categorized as “unclear,” we asked the stakeholders to provide additional information or clarifications. All topics were distributed to stakeholders. Throughout the stakeholder panel discussion period, stakeholders were invited to submit new FRN topics.

Use of Microsoft® SharePoint

Two separate Microsoft SharePoint Web sites were created to: (1) host the FRN discussion; (2) submit additional topics for discussion; and (3) nominate topics for future research. SharePoint resources were provided through the Tufts Clinical and Translational Science Institute, of which the Tufts EPC is a member. One Web site was created for the treatment topics discussed in this document and a separate Web site for the parallel diagnosis discussion. SharePoint was chosen as it offered stakeholders the flexibility of an asynchronous online discussion forum. Additionally, the secure Web site provided Tufts staff the most control, in terms of site content, structure and functionality, when compared to other publicly available discussion platforms. The secure, password-protected Web site was housed behind the Tufts firewall, accessible to stakeholders via invitation only. The SharePoint Web site also served as a platform for Tufts EPC staff to post project announcements, a welcome video, and important reference documents including instructions to stakeholders on how to navigate and use the Web site, an FRN project overview, the Executive Summary and Key Questions of the full sleep apnea CER, lists of participating stakeholders, initial FRN discussion topics, common abbreviations and acronyms, announcements, and reference documents. The documents on the Web site were also emailed to all stakeholders. In addition to the EPC staff and the stakeholders, the AHRQ Task Order Officer was given access to the Web site.

The FRN topic discussion boards were the primary feature of the SharePoint Web site. Discussion boards were prominently placed front and center on the Web site’s homepage (Figure 2). Links to individual discussion boards were also strategically placed on the Web site’s main navigation toolbar for direct access from any page on the Web site.

Figure 2. SharePoint homepage: Treatment



Sleep Apnea Future Research Needs: Treatment

All Sites Advanced Search

View All Site Content

Home Page

Sleep Apnea Diagnosis Website

Contact Tufts Staff

Submit FRN Topic for Discussion

Sleep Apnea Treatment Topic Discussions

- 1. Association of apnea severity with long-term outcomes
- 2. Comparative trials of different sleep apnea treatments based on patient characteristics
- 3. Comparisons of alternative treatments for patients not tolerating fixed CPAP
- 4. Long-term trial of mandibular advancement devices
- 5. Studies to predict adherence to CPAP, surgery and mandibular advancement devices
- 6. Trials comparing CPAP with surgery
- 7. Trials comparing different degrees of mandibular advancement
- 8. Trials comparing oropharyngeal exercises to CPAP
- 9. Trials comparing pharmaceutical interventions to CPAP
- 10. Trials comparing suction tongue-retaining devices with nonsuction tongue-retaining devices

Sleep Apnea Treatment FRN Discussion Board (click on topic below to join discussion) ▾

Subject	Replies
11. Trials of CPAP stratified by severity of sleep apnea	7
13. Trials of non-CPAP treatments stratified by severity of sleep apnea	6
4. Long-term trial of mandibular advancement devices	6
7. Trials comparing different degrees of mandibular advancement	1
6. Trials comparing CPAP with surgery	3
9. Trials comparing pharmaceutical interventions to CPAP	2
8. Trials comparing oropharyngeal exercises to CPAP	5
14. Trials to evaluate weight loss programs as an additional interventions	4
10. Trials comparing suction tongue-retaining devices with nonsuction tongue-retaining devices	2
12. Trials of effective treatments to improve adherence with both CPAP and mandibular advancement devices	5
1. Association of apnea severity with long-term outcomes	4
3. Comparisons of alternative treatments for patients not tolerating fixed CPAP	11
5. Studies to predict adherence to CPAP, surgery and mandibular advancement devices	3
2. Comparative trials of different sleep apnea treatments based on patient characteristics	7
15. Comparisons of different CPAP masks	5
16. Studies of factors influencing patient & therapist decisions concerning mask (CPAP interface) choice for OSA therapy	4
17. Assess the value of post-operative use of CPAP (after any surgery) in those with OSA or likely OSA	4
18. Value of having a sleep medicine specialist involved in the treatment of OSA	4
19. Role of bariatric surgery in OSA patient	4
20. Role of orthognathic surgery (corrective jaw surgery) in OSA patients	1
21. Comparison of genio-tubercle advancement and dental devices for treatment of OSA with respect to long-term results	0
22. Cost-effectiveness of a management strategy (diagnosis [of symptomatic or high-risk patients] through treatment [of patients diagnosed with OSA]), specifically for patients with mild-to-moderate disease severity	2
23. Studies to improve adherence with CPAP and other treatments, particularly evaluating cognitive therapy approaches	3
24. Direct comparisons of compliance rates with different interventions, and thus comparisons of effective treatment	3
25. What are barriers to compliance?	2
26. Research on CPAP devices that are both economical as well as clinically effective	1
27. What are the financial barriers to access to treatment?	0

Site Actions ▾

Announcements ▾

4. Sleep Apnea report now available at AHRQ website

8/10/2011 1:10 PM

by tuftsepc

The complete report on the diagnosis and treatment of OSA is now available on the AHRQ website. It is available in various forms - the full report, the executive summary, the Clinicians Guide, and the Consumers Guide. Click the link titled "AHRQ..."

3. Presentation about FRN project added to Reference Documents

8/5/2011 4:37 PM

by tuftsepc

Thank you for your participation in the conference calls. For participants who were unable to join the two conference calls we had this week, we have uploaded the presentation made at the call, titled "Overview of Future Research needs project" in...

1. Welcome to the Sleep Apnea Treatment Website

8/5/2011 4:26 PM

by tuftsepc

Welcome to the Tufts Evidence-based Practice Center website for "Future Research Needs in the Diagnosis of Obstructive Sleep Apnea". For more instructions on navigating the website click here: [Instructions to Stakeholders](#)

2. List of participants added to Reference Documents

8/5/2011 4:25 PM

by tuftsepc

Please see the document titled "6. List of Participants" in the Reference Documents section.

Video: Welcome to the Future Research Needs Project ▾



04:02

Figure 3. Sample discussion thread: Treatment

Tufts CTSI Tufts Clinical and Translational Science Institute

Sleep Apnea Future Research Needs: Treatment

Advanced Search

Sleep Apnea Future Research Needs: Treatment

Sleep Apnea Future Research Needs: Treatment > Sleep Apnea Treatment FRN Topics > Comparative trials of different sleep apnea treatments based on patient characteristics

Sleep Apnea Treatment FRN Topics

View: **Threaded**

Threading

Created By

Replies

Edited: 7/19/2011 5:18 PM by [redacted] [View Properties](#) [Reply](#)

2. Comparative trials of different sleep apnea treatments based on patient characteristics

Comparative trials of different sleep apnea treatments based on patient characteristics

Rationale: Trials of sleep apnea treatments (CPAP, surgery, oral mandibular devices, and drugs) do not compare the effectiveness of these treatments in different patient sub-groups that are based on obesity, sex, and other patient characteristics. Trials of sleep apnea treatment need to be stratified by these patient characteristics to identify the patient groups which would benefit most from these interventions.

Instructions: Please comment on this topic for future research in treatment/diagnosis of sleep apnea. Please consider 1) the topic's importance; 2) whether the topic has already been sufficiently researched; 3) the feasibility of conducting research on this topic; and 4) its potential impact.

Posted: 8/6/2011 1:33 PM by [redacted] [View Properties](#) [Reply](#)

important area of research to better match treatments to specific patient types (obesity, morphometrics, severity)

[Show Quoted Messages](#)

Posted: 8/9/2011 7:43 AM by [redacted] [View Properties](#) [Reply](#)

Agree it would be important and useful to clarify which subgroups of patients as defined by obesity level, sex, age group, and OSA severity derive benefit from each treatment and then assess the comparative effectiveness of the various treatments.

[Show Quoted Messages](#)

Posted: 8/10/2011 1:38 PM by tuftsepc [View Properties](#) [Reply](#)

Which do people think might be a better approach (not that they're mutually exclusive)?

1. Evaluating large comparative studies (that used sufficiently broad eligibility criteria) to determine which patients (based on race, sex, BMI, etc.) most benefited from which treatment?
2. Comparisons of treatments specifically in narrow sub-populations of patients (eg, BMI>40, mild symptoms/relatively low AHI)?

[Show Quoted Messages](#)

Each FRN topic was uploaded as a separate “discussion board.” These were structured to allow stakeholders to comment on both the original FRN topic as well as other stakeholder comments (Figure 3). For their discussions, stakeholders were asked to consider four dimensions of need related to the proposed topic (see Appendix A and Approach to Prioritization section, below). Stakeholder participation was monitored and discussions were moderated daily by Tufts EPC staff to ensure appropriateness and relevance of all comments. The discussion boards were initially scheduled to be open for a 2-week period. During the open period, stakeholders were regularly encouraged (two to four times per week) to contribute to the discussion boards as well as submit additional topics for discussion. Stakeholders were encouraged to go through all topic areas, and provide comments and feedback on all topics on the Web site.

We planned to use the SharePoint Web site to nominate FRN topics using a process similar to topic submission described above. This strategy was revised due to limited stakeholder engagement with the SharePoint Web site (as described in the Results section).

At the conclusion of the project, all online discussions and email communications were archived for transparency purposes.

Approach to Prioritization

The stakeholders were asked to consider four dimensions of need related to the proposed topic. These four dimensions come from the AHRQ Effective Health Care Program selection criteria (Appendix A). These dimensions and the Effective Health Care Program guidance on them were described in detail in the leadup to topic submission, discussion and selection of FRN topics.

- Importance
- Desirability of Research/Duplication
- Feasibility
- Potential Impact

A fifth dimension, Appropriateness, was evaluated by EPC program staff after submission of initial FRN topics.

After the close of the online discussion, the stakeholders were asked to identify up to 10 FRN topics that were of highest priority and that met the Effective Health Care Program selection criteria. The original plan was to conduct this nomination step on the Web site, but as described in the Results section, nomination was conducted by email and individual phone calls. After nomination, the EPC grouped similar topics into overarching topics and edited the names of the title for clarity and consistency. Based on the stakeholder nomination, the EPC categorized the overarching topics into four groups:

- High-priority FRN topics: Clearly of interest to stakeholders (based on high levels of nominations). A consensus of stakeholders expressed that these topics are of high priority. We aimed for about five topics, but we used natural breaks in the rankings of the topics, rather than strictly defining the numbers of topics in this category. For these topics, the background, stakeholder discussion, and study design considerations are fully elaborated in the Results section.
- Second-tier FRN priority topics: Of interest to a substantial number of stakeholders, but lacking a consensus that these were high-priority topics. For these topics, only a summary of the stakeholder discussion is presented.

- Other topics: Of relatively little interest to stakeholders. These include the remaining topics (see the next category for exceptions) that were nominated by few if any stakeholders. Topics in this group are listed in the Results section without further discussion.
- Does not meet the Effective Health Care Program Appropriateness criteria. The EPC will move any topics that do not meet these minimum requirements for inclusion in the FRN report into this category for potential further discussion. These topics are not explicitly listed in the Results.

Approach to Research Question Development and Considerations for Potential Research Designs

For each high-priority FRN topic, we considered the range of study designs that would best address the topic. We did this taking into account the PICOD criteria (Population, Intervention, Comparator, Outcomes, study Design). For each topic, we described our assumptions about the most appropriate PICOD criteria, in particular describing the advantages and disadvantages of various potential research designs. We specifically considered the feasibility of the research questions focusing on potential sample size, time, and recruitment issues. For selected topics, we consulted with our local domain expert, who was also a stakeholder.

To determine candidate study designs, the feasibility of the study designs, and sample size calculations, we followed the structure laid out in the Future Research Needs document Framework for Considering Study Designs for Future Research Needs.⁴

Briefly, candidate study designs will differ across types of research needs. Effectiveness or efficacy of treatments can be most definitively addressed in randomized trials, and secondarily in well-conducted nonrandomized comparative observational studies. In contrast, eliciting patient preferences can be meaningfully performed with nonexperimental designs (e.g., in a survey or focus group). Furthermore, observational studies may be most appropriate to enhance generalizability and determine effectiveness, as opposed to efficacy alone. Each final FRN topic was assessed as to context of the research question and a determination as to whether evaluation of efficacy or effectiveness is of greater need. This informs the choice of study design.

When a simple randomized trial was deemed to be an appropriate study design to address a FRN topic, we performed sample size calculations using standard formulae for a two-sided chi-squared test at the 0.05 level of significance. We assumed an allocation ratio of 1:1, no loss to followup, no crossover between treatments, and no sequential monitoring. For studies with dichotomous outcomes, we determined a range of reasonable control rates (event rates in the comparator arm) and relative effects (risk ratios) and based on these calculated a range of scenarios and minimum sample sizes. For studies with continuous outcomes, we range of reasonable mean differences in effect size (e.g., hours of sleep) between arm and standard deviations of the differences and based on these calculated a range of scenarios and minimum sample sizes. For continuous outcomes we also a conservative range of equivalence (the range between the smallest mean difference that would be clinically significant and the largest difference that could be reasonably expected). Where possible, these assumed values were derived from the existing evidence. All power calculations were set at 90 percent power to detect a significant result. We report all our assumptions.

Results

Composition of Stakeholder Panel

Through our multipronged approach to compiling the stakeholder panel, we enlisted a total of 27 stakeholders, of whom 22 were on the treatment panel. Table 2 summarizes the number and types of stakeholders on the panel.

Table 2. Stakeholders panel for treatment of sleep apnea

Category	Subcategory	No. of Stakeholders
Patients and the public	Patient advocates or caregivers	4
	Current patients	
	Transportation sector employers	
Providers	Hospital administrator	9
	Clinicians – Primary care	
	Clinicians – Pulmonary care	
	Clinicians – Dental care	
	Clinicians – Psychiatry	
	Clinicians – Nursing	
	Clinicians – Sleep medicine	
	Clinicians – Oropharyngeal surgery	
Clinician – Obesity and Metabolism		
Purchasers	Public purchasers - e.g., Veterans Administration, Federal Employees Health Benefits Plan	1
Payers	Private insurers	3
	Medicaid (at the State level)	
Policymakers	Agency for Healthcare Research and Quality	3
	Federal agencies	
	Professional organizations, Guideline developers	
Principal investigators	Clinical research	2
	Health services/policy research	
TOTAL		22

Stakeholder Participation and Required Methods Modifications

As noted above, this was the first instance of the Tufts Evidence-based Practice Center (EPC) producing a Future Research Needs (FRN) document with a large number of stakeholders. As such, the process was to a large extent a test of the success of the specific protocol we envisaged. Here we describe the substantive changes that we had to make to the protocol methods.

Overall, the use of the Sharepoint Web site was only moderately successful. The majority of stakeholders were slow to sign on to the Web site, were slow to begin reviewing the ongoing discussions, participated only minimally in discussions, and did not offer new FRN topics. Furthermore, a small number of participants were not able to log in to the Sharepoint Web site because of technical difficulties.

There were 30 FRN topics discussed and prioritized; 14 of these were added by the EPC based on the Comparative Effectiveness Review (CER) Future Research section; 7 were added online (or via email to the EPC) by stakeholders; 9 were added by the EPC after a teleconference with the consumer advocates (see below).

The discussion period was started 12 days after we received signed disclosure of interest forms from almost all the stakeholders (with the exception of the patient stakeholder who we were delayed in recruiting). All participants were contacted 5 days before the online discussion board was made available. We informed them of the impending start of the online discussions as confirmed that their email addresses would be available to them during the course of the discussion. The original plan was for the stakeholder discussion to last 2 weeks; in reality, the discussion period was extended to 27 days to allow further stakeholder involvement. Six of 22 stakeholders signed in to the Web site within 2 days. The median time until stakeholders signed in was 4.5 days. Six stakeholders did not log in to the online discussion site at all. The last stakeholder signed in on day 23 of the discussion. Across topics, discussions were started by only five stakeholders; nine stakeholders participated in the online discussions (added any comments). Not including comments or questions added by the EPC, for the 30 topics discussed online, the median number of comments by stakeholders was 2, ranging from 0 (for 11 topics) to 7. After the discussion period ended, 17 of 22 stakeholders participated in topic nomination (which was conducted by email instead of on the Web site, as originally planned; see below).

To improve participation, after 1 week of discussion, we invited all stakeholders (from both panels) to participate in teleconferences where the EPC reviewed the materials that had been sent to them by email and that were available on the Web site, including the goals of the project, the stakeholders' responsibilities, the main criteria for selecting and discussing FRN topics, and a review of how to use the Sharepoint Web site. Two calls were scheduled, for which 13 of 27 stakeholders joined (from both diagnosis and treatment panels). To increase participation, we also sent numerous email reminders, offered to have stakeholders email us their comments that we would add to the discussion board, and also answered phone queries on technical issues related to the Web site. Furthermore, because several stakeholders could not easily access the Web site (primarily because of overseas travel), we compiled plain text versions of all the FRN topics and discussions, which we placed directly into the text of emails (not as attachments). This was emailed to all stakeholders.

Near the true end of the discussion period, we noted that the consumer stakeholders (the patient, the patient advocate, and the representative from the transportation industry) had not participated in the discussion. Upon reviewing the discussions to date, we thought that they might have been too scientifically technical for the lay stakeholders. We contacted them directly, had our suspicions confirmed, and organized a separate teleconference for the three of them (with the EPC). This 1.5-hour teleconference produced nine new topics. The EPC organized the discussion into FRN topics, summarized the separate discussions, sent the summary to the lay stakeholders, and then uploaded the topics and discussions to the Web site. A summary of the call was also emailed to all stakeholders. No stakeholders added further to these topic discussions.

After the discussion period was closed, the EPC decided that it was not worthwhile to attempt to use the Web site to have stakeholders nominate the topics. Instead this was done by email. The final list of topics and the text of the discussions were emailed to stakeholders. The order of the topics was randomized once, instead of separate randomization for each stakeholder, since the randomization process was too time consuming. (We chose to maintain links between

the list of topics and the topic discussions, rather than manually reorder each stakeholder list.) To prevent procrastination and further delay, we asked stakeholders to nominate the topics within 4 days, by Friday, with the expectation that responses would be in by Monday. Eight of 22 stakeholders sent in the topic nominations by Monday. It was likely that some of the delay in response was due to a weather-related phenomenon (Hurricane Irene) so we extended the timeline and sent out a reminder email after 1 week. Two of remaining 14 stakeholders sent in their nominations after receiving the email. Subsequently, followup phone calls were made to the remaining stakeholders to solicit their nominations. The final list of nominated topics was received the following Monday. Topics were nominated by 17 of 22 stakeholders. Stakeholders were asked to nominate up to 10 topics; they nominated between 4 and 10 topics each.

Research Needs

Based on the methods described above, we organized the FRN topics into three categories (see Methods section): High priority FRN topics; Second-tier priority FRN topics; and Other FRN topics. Topics that did not meet minimum requirements for inclusion are not presented. The FRN topics are as listed in Table 3.

High-Priority Future Research Needs Topics

High-Priority Future Research Needs Topic 1

What is the impact of treatment of sleep-disordered breathing on major long-term clinical outcomes, including mortality, cardiovascular disease, and diabetes?

- a. What are long-term outcomes of mandibular advancement devices (MAD) treatment?

Background

This FRN topic represents two topics that were nominated and discussed separately by the stakeholders. After nomination of topics by the stakeholder panel, the topics were combined into one overarching topic with different components.

One of the main research needs identified by the CER on treatment of sleep apnea was the use of clinical outcomes in comparative studies. Only 3 of 190 studies reported clinical outcomes. However, nonrandomized comparative studies of various treatment modalities were not reviewed for an assessment of their impact on outcomes. It was recommended that future research should focus on comparative studies with long-term followup and clinical outcomes. This was also an opinion echoed by the stakeholders on the FRN panel.

Table 3. List of future research needs topics

High Priority Future Research Needs Topics
1. What is the impact of treatment of sleep-disordered breathing on major long-term clinical outcomes, including mortality, cardiovascular disease, and diabetes? a. What are long-term outcomes of mandibular advancement devices (MAD) treatment?
2. Cost-effectiveness analysis of a management strategy (diagnosis of symptomatic or high-risk patients through treatments of patients diagnosed with OSA), specifically for patients with mild to moderate disease severity a. Research on CPAP devices that are both economical as well as clinically effective
3. Comparative studies of different sleep apnea treatments based on patient characteristics a. Analyses of CPAP stratified by disease severity b. Analyses of non-CPAP treatments stratified by disease severity c. Comparison of alternative treatments for patients who do not tolerate CPAP
4. Trials to improve compliance with CPAP, MAD, and other treatments, particularly evaluating cognitive therapy approaches
5. What is the association between sleep apnea severity and long-term clinical outcomes?
Second-Tier Future Research Needs Topics
6. What are the barriers to, and predictors of, compliance with different treatments?
7. Direct comparison of compliance rates with different interventions and incorporation of compliance into an overall comparison of effective treatment
8. Trials to evaluate weight-loss programs as an adjunctive treatment for sleep apnea a. What is the value of bariatric surgery for treatment of sleep apnea?
9. What is the value of including a sleep medicine specialist in the management of the patient with obstructive sleep apnea?
Other Future Research Needs Topic
10. What is consumer willingness to pay for treatment, to identify consumer preferences for strategies to treat sleep apnea?
11. What are the financial barriers to access to treatment?
12. Role of surgery for treatment of OSA a. Comparison of surgery versus CPAP b. Role of orthognathic surgery (corrective jaw surgery) c. Comparison of genio-tubercle advancement versus dental devices
13. Evaluation of postoperative CPAP for all patients with OSA or at high risk of OSA undergoing any surgery with sedation
14. Trials comparing CPAP versus pharmaceutical interventions
15. Trials comparing different CPAP masks
16. Trials comparing CPAP versus oropharyngeal exercises
17. Trials comparing different degrees of mandibular advancement
18. Studies of factors influencing therapist decisions concerning CPAP mask choice
19. Research into how to maintain patients in sleep apnea studies (where dropout rates are unacceptably high)

Stakeholder Discussion

The first nominated topic suggests that clinical outcomes should be the outcome of interest in comparisons of various treatments for sleep apnea, and the second topic focused on long-term outcomes for mandibular advancement devices (MAD). For assessing long-term clinical outcomes, the stakeholders' suggested study design was a database analysis from a prospective cohort study like the Framingham Heart Study or Nurses Health Study. For assessing clinical benefits from use of MAD, a randomized trial or comparative study was the suggested study design with MAD serving as the intervention arms and continuous positive airway pressure (CPAP) or other oral devices serving as the comparator arm.

Proposed Study Designs

To address the overarching FRN topic and its subcomponents, different study designs are reasonable and would address different aspects of this topic. In assessment of comparative effectiveness, the two overarching comparisons of interest would be autotitrating (and other

nonfixed) CPAP versus fixed CPAP and CPAP versus MAD. Sham or no treatment would not be a practical or ethical treatment option for a long-term study given the known effect of CPAP and MAD on sleep measures and symptoms. The clinical outcomes of interest are mortality, cardiovascular disease, and diabetes, including surrogate or intermediate markers of the latter two conditions.

Randomized Controlled Trials

Value of Study Design

Only one of 89 trials studied clinical outcomes. Any future trials of CPAP or MAD should study clinical outcomes. However, the resources required for the necessary long-term clinical outcomes would be great and loss to followup would be a major concern. Given the poor compliance of many patients with obstructive sleep apnea (OSA) treatments, it would also be likely that over the long term, many patients would cross from one treatment to another (and back again), complicating any analyses of these trials.

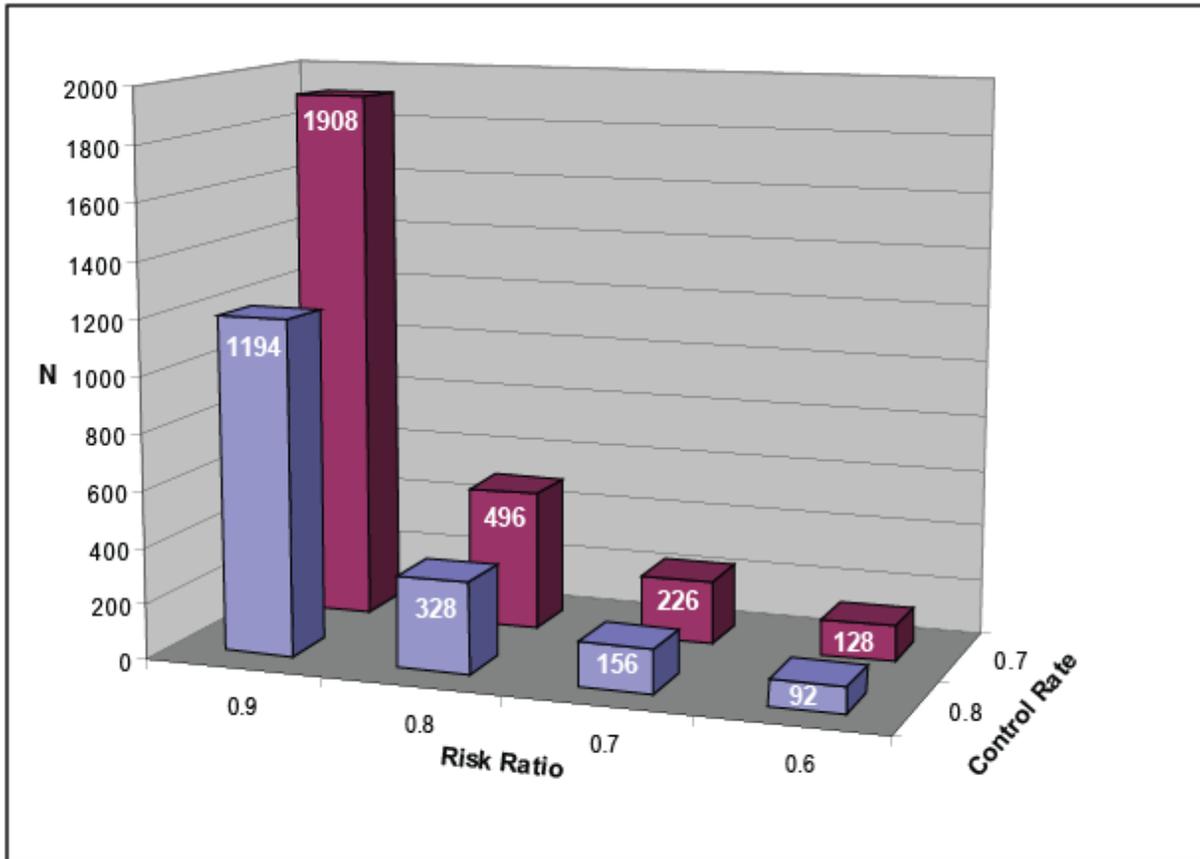
Resource Use, Size, and Duration

The primary analysis of interest in such a study would be the comparative effectiveness of various treatments on mortality, cardiovascular disease, and diabetes. These could result in the use of large resources in terms of cost, time and effort, and the decision to use these resources for conducting trials has to be balanced against the value of information that can be gained from studying clinical outcomes. While a study of mortality would require large sample sizes and long periods of followup, cardiovascular disease and diabetes could be followed up using both surrogate markers as well as hard clinical outcomes. For example, for the comparison between CPAP and MAD, the analysis of interest could be the effectiveness of CPAP and MAD on measures of glucose metabolism and insulin insensitivity, as markers of impaired glucose tolerance and incipient diabetes. The number of patients that would need to be enrolled in a trial would depend on the relative effect of the interventions and the event rate for the clinical outcome in the comparator arm. Regardless of the relative difference in effects between groups, trials using mortality as an outcome would need a sample size in the thousands, while using surrogate outcomes would need smaller sample sizes.

To evaluate sample size for trials, we used the data from trials between MAD and CPAP to conduct a power calculation using standard formulae for a two-sided equivalence test. Given the lack of any data to suggest an expected relative effect of treatment on a composite cardiovascular outcome, we assumed that it would be equal to the effect seen in trials for improvement in apnea-hypopnea index (AHI)—greater than 50 percent reduction in the AHI to less than 5 events/hr). We set the range of relative risk reduction that was seen in the report, which ranged from 0.6 to 0.9. From the CER, we also obtained the event rates of outcomes in the control arm (in this case, the control arm was CPAP) which was 0.7 to 0.8. The required total sample size (1:1 ratio in a two-arm randomized controlled trial [RCT]) for each value of relative risk and the control rate is presented in Figure 4.

The trial duration and followup for mortality as an outcome would be measured in years, while trials focusing on surrogate outcomes for cardiovascular disease and diabetes can be relatively short term, for example 6 to 12 months.

Figure 4. Sample size calculations for an RCT of MAD versus CPAP with varying estimates of risk ratios and control rates for combined cardiovascular outcomes



Ability To Recruit

Thousands of patients have already been recruited into OSA treatment trials. Patients are usually sufficiently symptomatic that they are likely to be interested in treatment. There are numerous treatment options, without clear guidance as to who should receive which treatment. However, these trials have used sleepiness and laboratory indicators like sleepiness scores and AHI as outcomes. When considering clinical outcomes, there is a potential for a large loss in followup. This would necessitate a much larger recruitment population, with an extended recruitment period, which may present as a resource constraint.

Ethical Issues

The primary ethical issue (beyond the standard ethical issues in conducting a randomized trial) pertains to whether there is equipoise. An argument can be made that another trial that evaluates only CPAP or MAD for the same sleep-related outcomes that the existing 89 trials have evaluated may be unethical since the value of the treatment is known and the trial would not change the conclusions from the body of evidence. However, if a new trial truly addresses a question that is inadequately addressed, such as how these treatments would affect objective clinical outcomes and to what degree, then there is adequate equipoise.

Post Hoc Analyses of Existing Observational Studies

Value of Study Design

A second-tier option for using existing data to evaluate the effects of different patient characteristics would be to perform post hoc regressions or subgroup analyses within observational cohort studies like the Framingham Heart Study, Wisconsin Sleep Study or Nurses Health Study. The CER did not include observational study designs when evaluating the comparative effects of different treatments. It is not clear whether the information on treatments for sleep apnea is captured in adequate detail in these studies. It is known that the outcomes of interest, that is, mortality, cardiovascular disease and diabetes, have been well documents in the study populations. However, if the required information is available, a retrospective analysis of impact of sleep apnea treatments would be invaluable in providing insight. However, such *post hoc* analyses may be susceptible to type I error (falsely significant associations) due to multiple testing Nevertheless, they would form a good basis to determine which treatments and outcomes could be further investigated a priori in future trials. This analysis could be performed either by the original study investigators or other researchers with access to the study data.

Resource Use, Size, and Duration

Post hoc analyses of existing data can be done quickly with modest resources.

Ethical Issues

There should be no ethical issues involved in reanalyzing existing study data, provided that consent has been obtained from study participants to use their anonymized data for future studies.

Observational Studies

Value of Study Design

Though less informative than randomized trials for the evaluation of comparative effectiveness, observational studies would be less resource intensive in assessing impact of various treatments. Case-control studies of patients treated for OSA, with or without a clinical outcome (e.g., a myocardial infarction), could also address this question, though they would be subject to all the biases and other problems of any case-control study. Efforts to mimic a random assignment in prospective observational studies, like propensity score matching, could play a part in a priori design of an observational study as well as in the post hoc analysis of the data from these studies.

Resource Use, Size, and Duration

Retrospective data could be gathered quickly and easily by chart review or similar approaches. Prospective cohorts would take somewhat more time and effort, but would still be less resource-intensive than a trial. Since outcomes generally change rapidly (within a few weeks or months) after the start of treatment, data could be collected fairly rapidly, whether retrospective or prospective. The only exception to this rule occurs when mortality is the outcomes, as the timeframe is then extended to years. A retrospective study is more practical when this is the case, even though it is susceptible to significant bias. The sample size will depend on the estimated relative effect of the treatment, and the event rate in the comparator arm

These studies are better suited for mortality outcomes as it is less resource intensive to follow up patients for a period of a few years.

Ability To Recruit

Patient recruitment should be straightforward and relatively simple, as there would be little added burden for them by entering a study, beyond a formalized assessment of symptoms, other outcomes of interest, and patient characteristics, all of which are reasonable parts of normal patient care.

Ethical Issues

There should be no substantive ethical issues involved in these observational studies.

High-Priority Future Research Needs Topic 2

Cost-effectiveness analysis of a management strategy (diagnosis of symptomatic or high-risk patients through treatments of patients diagnosed with OSA), specifically for patients with mild to moderate disease severity

- a. Research on CPAP devices that are both economical as well as clinically effective

Background

This FRN topic represents two topics that were nominated and discussed separately by the stakeholders. After nomination of topics by the stakeholder panel, the topics were combined into one overarching topic.

The topic as proposed covers both diagnosis and treatment strategies for a complete cost-effectiveness analysis. Since diagnosis and treatment FRN topics are being dealt with separately, here we focus primarily on the treatment strategy component of the cost-effectiveness analysis. A discussion of a cost-effectiveness analysis of treatment strategies can be found in the companion report.³ This FRN is distinguished from FRN 3 (below) on three dimensions: (1) it assesses the expected value (not just benefit) of selected treatments for OSA; (2) it is designed to address these questions for patients with mild to moderate disease, an understudied group; and (3) it proposes to assess expected value from both the societal and patient perspectives. Out-of-pocket expenses for patients with mild to moderate disease were of considerable concern to patients in our stakeholder group. This FRN seeks to address that concern.

It would likely be of value for future cost-effectiveness analyses to incorporate both phases of OSA management (diagnosis and treatment). Cost-effectiveness analysis allows the comparison of different interventions on similar benefit, cost and utility scales. If clinical trial data are available on the treatments and populations of interest, benefit and cost estimates have both internal and external validity.

Stakeholder Discussion

The objective of this research need is to establish better evidence about the costs and benefits of alternative management strategies for individual patients with mild to moderate disease severity. Ultimately, the goal is to develop evidence about care management from efficient testing high-risk individuals, accurate and low-cost diagnosis of OSA, through effective and low-cost treatment of patients. Issues that should be considered in a cost-effectiveness model include

the patient-related outcomes, including measures of functional status such as productivity (both absenteeism and presenteeism—working in spite of illness, with resulting poor work performance), quality of life, and work safety. The feasibility of such an analysis will depend on combining clinical trial results with health insurance data on health care utilization, employer data on absenteeism, and potentially auto insurance data on motor vehicle accidents in a large cohort of patients with OSA. Other issues that should be considered are the patients' costs for using OSA treatments. An example of such a cost, which would not be included in most analyses, was the cost to a long-haul truck driver who requires extra battery backup since a CPAP device can drain the truck's battery overnight.

Proposed Study Designs

Systematic Review

Cost-effectiveness analyses were not addressed by the CER. Conducting a systematic review may be the first step to ascertain the level of existing evidence.

Cost–Benefit Analysis

On the basis of the best evidence identified in the CER, a quality-adjusted cost-benefit analysis is recommended, comparing the incremental costs and benefits of different treatments to each other and to a base case of no active treatment. Benefit, utility, and cost estimates may be derived from previous clinical trial data, where available, and from observational data where trial data are not available. These estimates should include not only the standard clinical outcomes, but also work-related, accident, and quality of life outcomes. In the absence of cost estimates, charges may be derived from administrative data and adjusted cost-to charge ratios. Out-of-pocket patient costs should also be included.

Analytic Approach

Preference should be given to a Markov-chain, discrete events analysis. The proper outcome measure is quality-adjusted life-years (QALYs) gained. Future QALYs and costs should be discounted over the follow-up period. Adjustment should be made for major outcomes and adverse events, using rate and utility estimates. Probabilistic sensitivity analyses should be conducted on estimated outcome rates, utilities and cost estimates. A societal perspective should be assumed in the main analysis. The patient perspective should be assumed in a subanalysis, given high out-of-pocket costs for devices and other treatments, substantial concerns about adherence (due to inconvenience and discomfort of treatment), and the effects of OSA on work and lifestyle.

Interventions

Treatment strategies should be tested both with and without the involvement of a sleep medicine specialist, where possible. Phased treatment combinations should also be considered. A short list of treatment strategies to be compared could include

- No treatment
- Fixed CPAP
- Fixed CPAP followed by alternative devices if fixed CPAP fails
- Initial use of alternative CPAP devices (including autotitrating CPAP)
- CPAP with interventions to improve compliance

- CPAP followed by MAD, if necessary
- MAD alone
- MAD with interventions to improve compliance
- MAD followed by CPAP, if necessary
- Bariatric surgery, when indicated
- Surgery, when indicated

Numerous other treatments and combinations of treatments are also used, including strategies that base initial treatment choice on patient and disease characteristics.

Unfortunately, as the CER highlights, the evidence for the effect of most of these treatments on any patient-related outcomes is of low strength or insufficient. Thus, the sensitivity analyses may be of greatest value, by describing how effective an intervention would have to be to be a cost-effective option for patients with OSA.

Resource Use, Size and Duration

Because a cost-effectiveness analysis can draw from previously collected data, the cost, size and duration of such studies can be limited. However, as discussed by the stakeholders, it may be challenging to gather all the relevant data.

Ethical, Legal and Social Issues

No new data collection is proposed and therefore the direct risk to patients is minimal.

High-Priority Future Research Needs Topic 3

Comparative studies of different sleep apnea treatments based on patient characteristics

- Analyses of CPAP stratified by disease severity
- Analyses of non-CPAP treatments stratified by disease severity
- Comparison of alternative treatments for patients who do not tolerate CPAP

Background

This FRN topic represents four topics that were nominated and discussed separately by the stakeholders. After nomination of topics by the stakeholder panel, the topics were combined into one overarching topic with different components.

Key Question 5 in the CER asked about the comparative effect of different treatments. The two subquestions asked about whether the comparative effect varies by different presenting characteristics. Specifically:

- Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?
 - Characteristics: Age, sex, race, weight, bed partner, airway, other physical characteristics, and specific comorbidities

- Obstructive sleep apnea severity or characteristics: Baseline questionnaire (and similar tools) results, formal testing results (including hypoxemia levels), baseline quality of life, positional dependency
 - Other: specific symptoms
- b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?

For all comparisons between treatments, the CER concluded that the strength of evidence is insufficient to determine which patients might benefit most from treatment in general (i.e., compared to no treatment) or from which treatment (from direct comparisons of different treatments). Trials of sleep apnea treatments (CPAP, surgery, MAD, and drugs) do not compare the effectiveness of these treatments in different patient subgroups that are based on obesity, sex, and other patient characteristics.

Stakeholder Discussion

The overarching theme of the discussion in the CER and among the stakeholders was that there is a need for subgroup analyses to help clinicians base treatment decisions on baseline characteristics to maximize early effective, tolerable treatment, to better match treatments to specific patient types, and to minimize costly trial and error. Two major categories of studies could be performed to address this question: (1) comparisons of treatments specifically within narrow subpopulations of patients, or (2) subgroup analyses of large comparative studies. The discussants agreed that it would be a better use of research resources to conduct large RCTs that recruit a wide range of patients and then to perform prespecified subgroup analyses. Specific baseline patient characteristics (or subgroups) of interest among stakeholders included: race, sex, obesity level (body mass index), age group, disease severity, morphometrics (e.g., measurements of the jaw), and interactions within these categories (e.g., obese vs. nonobese in blacks and whites separately).

A different approach may be warranted for evaluation of effectiveness of CPAP (versus no treatment) for disease severity. It was noted that trials of CPAP for severe OSA have already been done and convincingly show improvement in short-term and sleep outcomes. Further trials are not necessary to show the value of CPAP (compared to no treatment) for this group of patients. In contrast, additional studies are needed to assess the short- and long-term value of CPAP in patients with mild to moderate OSA severity. This caveat, though, does not hold for comparisons of non-CPAP interventions to control, where further trials are necessary for all groups of patients. However, the CER found that MAD was effective at improving symptoms and sleep study measures, primarily in patients with moderate to severe OSA (as defined by AHI). Furthermore, trials comparing different active interventions should include a wide range of patients based on OSA severity to allow evaluation of the relative effectiveness of different treatments across the range of disease severity.

Also of interest, are future studies to delineate the most effective treatment in the subgroup of patients who have not tolerated CPAP. To date, the large majority of trials either explicitly or implicitly included only patients newly diagnosed with OSA. Thus, the current evidence may not be fully applicable to patients who have tried and failed CPAP. Related questions of interest include comparisons of different types of CPAP (e.g., of different alternatives to fixed CPAP), evaluation of different adjunctive treatments (e.g., humidification), and comparisons of different non-CPAP treatments. In particular, it was noted that oxygen via nasal prongs, as a standalone

treatment, is commonly prescribed by physicians for patients who do not tolerate CPAP, but this treatment has not been adequately tested. Of note, oxygen alone as a treatment was not evaluated in the CER.

Proposed Study Designs

To address the overarching FRN topic and its subcomponents, different study designs are reasonable and would address different aspects of this topic.

Patient-Level Meta-analysis

Value of Study Design

For comparisons of several of the treatments (particularly CPAP, autotitrating CPAP, and MAD), there were adequate trials to allow meta-analysis. However, meaningful conclusions about the relative effects of treatments on different subgroups of patients could not be ascertained from these meta-analyses because of the risk of ecological fallacy—erroneous conclusions based on ascribing the study mean estimate (e.g., mean AHI) to all participants in a study. But the large body of existing studies could be used to evaluate subgroup effects, without the risk of ecological fallacy, if patient-level meta-analyses could be conducted.

Briefly, typical meta-analysis combines the mean or overall estimates of effects from different studies. In contrast, patient-level meta-analysis uses the same patient-level data available to the original study investigators from multiple trials, and combines these. This approach allows unbiased regressions and subgroup analyses, as could be conducted in any individual trial. However, acquiring the data is the usual main obstacle to conducting patient-level meta-analyses.

The outcomes of interest (for this and all study designs discussed here) will be those outcomes that have been commonly obtained and which clinicians use to determine treatment effectiveness. These will primarily include AHI (both continuous and dichotomized [e.g., improvement by a certain amount or improvement to below a certain threshold]), other sleep measures, and symptoms including the Epworth Sleepiness Scale. Other outcomes of interest will be compliance and adverse events or discomfort.

Resource Use, Size, and Duration

If a researcher, funder, or other entity can convince the principal investigators of the largest, best CPAP and MAD trials to share their patient-level data for the purpose of meta-analysis, most of the patient characteristics could be quickly evaluated, with modest resources, regarding their association with effects.

Ethical Issues

Meta-analysis uses existing data. So long as the principal investigators agree to the use of the original data, there should be no ethical issues.

Post Hoc Analyses of Existing Trials

Value of Study Design

A second tier option for using existing data to evaluate the effects of different patient characteristics would be to perform post hoc regressions or subgroup analyses within already-published trials. This could be performed either by the original study investigators or other

researchers with access to the study data. To maximize accuracy and validity, such analyses should be done in multiple large, high-quality trials. However, such post hoc analyses may be susceptible to both type I error (falsely significant associations) due to multiple testing or type II error (falsely nonsignificant associations) due to lack of statistical power. Nevertheless, they would form a good basis to determine which patient characteristics could be further investigated a priori in future trials.

Resource Use, Size, and Duration

Post hoc analyses of existing data can be done quickly with modest resources.

Ethical Issues

There should be no ethical issues involved in reanalyzing existing study data.

Randomized Controlled Trials

Value of Study Design

The CER found 43 trials comparing CPAP to no or sham treatment, 21 trials comparing autotitrating CPAP with fixed CPAP, 10 trials comparing MAD to no or sham treatment, and 10 trials comparing CPAP to MAD. However, only 5 of these 84 trials were rated good quality. Regardless, it is unlikely that any new trial addressing these comparisons would substantially change the current conclusions about the effect and relative effect of CPAP, autotitrating CPAP, and MAD on measures of sleep and sleep quality. Thus the primary purpose of future studies should be to address this FRN topic, namely which group of patients would benefit most (or least) from treatment. Ideally, the a priori choice of patient characteristics would be based on existing patient-level meta-analyses or post hoc analyses of existing trials.

Of note, the EPC agrees with the stakeholders that it would be less desirable to conduct separate trials in narrow subgroups of patients. This would require indirect comparisons across studies where one has to assume that the differences in study findings are due to a single, explicit difference in population (e.g., obese vs. nonobese), even though it may be very likely that differences are due to multiple other, often subtle, differences between studies. Direct comparisons are always more reliable than indirect comparisons.

Resource Use, Size, and Duration

The primary analysis of interest in such a study would be whether the interaction term (between the main effect and the predictor variable) is statistically significant. Using as an example the comparison between CPAP and MAD, the goal would not be to demonstrate again that both CPAP and MAD are effective, but CPAP is superior. Instead, the analysis of interest would be whether the relative effectiveness of CPAP and MAD are different in different subpopulations of patients (e.g., baseline AHI <30 events/hr versus ≥ 30 events/hr) or analyses of the effect based on the predictor variable on a continuous scale (e.g., body mass index). The number of patients that would need to be enrolled in a trial to find a significant interaction would depend on the distribution of the predictor in the study sample and the strength of the interaction. Studies with a small minority of patients in the predictor category (e.g., if 10 percent were obese) would require a larger strength of interaction (e.g., a larger difference in the effect between the obese and nonobese) than studies that were more evenly balanced. Regardless of the actual distribution of study subjects and the relative difference in effects between groups, such trials would need to be substantially larger than an equivalent trial evaluating only the main effect.

Focusing on intermediate outcomes, such as AHI or measures of sleepiness, trials can be relatively short term, for example 2 to 3 months.

Ability to Recruit

Thousands of patients have already been recruited into OSA treatment trials. Patients are usually sufficiently symptomatic that they are likely to be interested in treatment. There are numerous treatment options, without clear guidance as to who should receive which treatment. For all these reasons, it is likely that there will not be major obstacles in recruitment to future trials.

Ethical Issues

The primary ethical issue (beyond the standard ethical issues in conducting a randomized trial) pertains to whether there is equipoise. An argument can be made that another trial that evaluates only CPAP or MAD for the same sleep-related outcomes that the existing 84 trials have evaluated may be unethical since the value of the treatment is known and the trial would not change the conclusions from the body of evidence. However, if a new trial truly addresses a question that is inadequately addressed, such as which patients would not benefit from treatment, or which treatment would work best in which patients, then there is adequate equipoise.

Observational Studies

Value of Study Design

Though less informative than randomized trials for the evaluation of comparative effectiveness, analyses of observational studies could be fruitful to evaluate predictors of effective treatment. Either prospective or retrospective cohort (single treatment group) data could be evaluated, using multivariable regression, to determine which patient characteristics (e.g., baseline body mass index) are associated with more successful treatment (e.g., clinically significant improvement in sleepiness). Case-control studies of patients treated for OSA, with or without a clinical outcome (e.g., a myocardial infarction), could also address this question, though they would be subject to all the biases and other problems of any case-control study. For any study design, the analyses that would be clinically useful would be predictors of outcomes based on baseline characteristics (e.g., relative risk of outcome in obese vs. nonobese), not *vice versa* (e.g., the average weight of successfully treated vs. untreated patients). The latter analyses are commonly reported, but do not provide useful insights for clinicians making treatment decisions.

Resource Use, Size, and Duration

Retrospective data could be gathered quickly and easily by chart review or similar approaches. Prospective cohorts would take somewhat more time and effort, but would still be less resource-intensive than a trial. Since symptoms generally change rapidly (within a few weeks) after the start of treatment, data could be collected fairly rapidly, whether retrospective or prospective. The sample size will depend on the number of regression covariates and the event rate (for dichotomous outcomes).

Ability To Recruit

Patient recruitment should be straightforward and relatively simple, as there would be little added burden for them by entering a study, beyond a formalized assessment of symptoms, other

outcomes of interest, and patient characteristics, all of which are reasonable parts of normal patient care.

Ethical Issues

There should be no substantive ethical issues involved in these observational studies.

High-Priority Future Research Needs Topic 4

Trials to improve compliance with CPAP, MAD, and other treatments, particularly evaluating cognitive therapy approaches

Background

Key Question 7 in the CER asked about the efficacy of interventions to improve compliance with device use. Specifically:

What is the effect of interventions to improve compliance with device (positive airway pressure, oral appliances, positional therapy) use on clinical and intermediate outcomes?

The CER concluded that there was a low strength of evidence that some specific adjunct interventions may improve CPAP compliance compared with usual care. None of the included trials reported results on clinical outcomes. The CER identified 18 trials that evaluated a wide variety of interventions to improve CPAP compliance, but did not find trials of MAD or other treatments of sleep apnea. The 18 trials of CPAP were mostly applicable to patients initiating CPAP with severe disease (AHI >30 events/hr who were obese (body mass index >30 kg/m²). They showed inconsistent effects across a wide variety of interventions. No general type of intervention was found to be more promising than others. However, compared with usual care, several interventions were shown to significantly increase hours of CPAP use per night in some studies. These included intensive support or literature (designed for patient education), cognitive behavioral therapy (given to patients and their partners), telemonitoring, and a habit-promoting audio-based intervention. However, these trials generally had small sample sizes (<40 patients in each arm) with less than 1 year of followup.

Stakeholder Discussion

Particular interventions discussed by the stakeholders included intensive patient education, social support programs engaging spouse or other family support, cognitive behavioral therapy, feedback from daily adherence monitoring, and peer support in a Web-based fashion. Stakeholders discussed that such interventions have been found to be effective in other domains such as weight control, drug addiction, and medication-taking behavior, but have not been adequately tested for OSA treatment. These types of interventions were discussed as being of greater importance than further modifications of CPAP and other devices.

Proposed Research Design

RCTs of multicomponent or multidimensional behavioral interventions aimed to improve long-term compliance with CPAP, MAD, and other treatments of sleep apnea.

Value of Study Design

A well-done RCT will produce the most convincing results, and if patient eligibility criteria and study setting are realistic, should be fairly applicable to the majority of patients. Several study design considerations need to be addressed before a RCT can be properly designed and carried out. These considerations are described as follows:

- Use established metrics to measure compliance. Commonly used metrics in the published literature include the continuous outcome hours of increased use of the device and the dichotomized outcome of at least 4 hours of use per night for more than 70 or 80 percent of the nights per week. However, it remains unclear how to interpret the clinical significance of these metrics. The relationships between improvements in compliance (using these measures) and clinical outcomes or quality of life have not been established.
- Use objective measures of compliance. Due to the nature of behavioral interventions, blinding of patients to the interventions cannot be achieved. Therefore, it is important to use objective measures of compliance that do not rely on self- or provider-reported data. An example of an objective measure is automatically recorded usage data from built-in memory chips in CPAP devices.
- Use behavioral change models and theories for designing interventions. Behavioral change models are multicomponent and multidimensional, involving patients, family members, peers, care providers, and changes to the patient's physical environment. Many interventions to promote health-related behaviors have utilized behavioral change models, and have been shown to be effective in changing patients' behaviors. For example, cognitive behavioral therapy (e.g., focusing on coping skills or environmental manipulation) was shown to be effective in improving compliance with CPAP treatment.⁵ All behavioral interventions should be piloted among the targeted patient population. Ideally, qualitative studies, such as focus groups, should be conducted to assess patients' preferences or feedback and barriers to the behavioral interventions as part of program evaluation.
- Conduct program evaluation during trial period. Program evaluation—a systematic method for collecting, analyzing and utilizing information regarding delivery of intervention program—is essential to identify barriers to either the implementation or effectiveness of the behavioral interventions.

Resource Use, Size, and Duration

The primary analysis of interest in such a RCT would be whether the compliance with CPAP, MAD, or other treatments is statistically significantly improved compared with usually care (no behavioral intervention). We conducted a power calculation using standard formulae for a two-sided equivalence test for mean difference in hours increase in use (as a continuous measure for compliance) between two independent samples. We set the range of equivalence (the range between the smallest mean difference that would be clinically significant and the largest difference that could be reasonably expected) from 0.5 hours to 8 hours of use. From the CER, we obtained the range of mean difference (from 1 hour to 3 hours of use) and the range of standard deviation of mean differences for the power calculation. The required total sample sizes (assuming a 1:1 ratio in a two-arm RCT) for pairs of assumptions using three mean differences and three standard deviation pair are presented in Figure 5. It should be noted that, most trials in the CER found a mean difference of less than 2 hours increase in CPAP use. Some trials in the CER found a significant improvement in compliance with CPAP use within 1 month, but other

trials found that it took much longer (6 to 9 months). Regardless, it would be of greater interest to show long-term (several years) effects on compliance, since in most cases, treatment can be expected to be life-long.

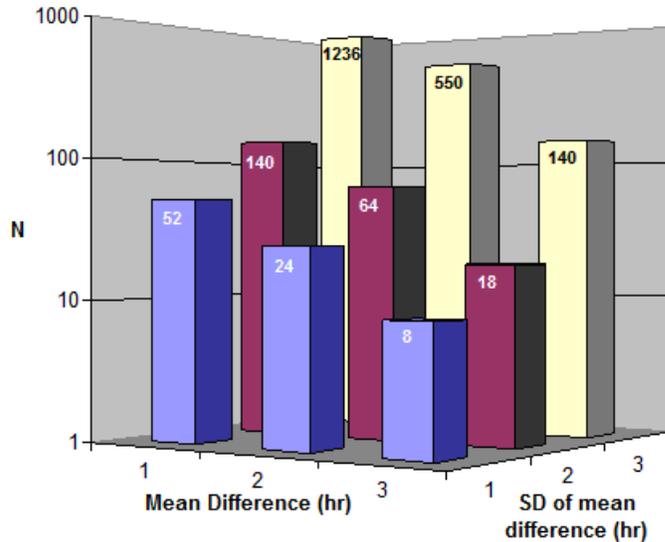
Ability To Recruit

Poor compliance with either CPAP or MAD is very common among patients with OSA. The patient advocates on the stakeholder panel described a demand by patients for better training, education, and other means to improve compliance. Therefore, it should be relatively easy to recruit patients with OSA who are being treated with devices into a trial.

Ethical Issues

There should be no substantive ethical issues involved beyond the standard ethical issues in conducting a randomized trial. The clinical value of interventions designed to improve compliance remains unclear.

Figure 5. Sample size calculation for a RCT of behavioral interventions to improve compliance with CPAP, MAD, or other treatments, compared with usual care



hr = hours (of use per night), N = number of trial participants (total), SD = standard deviation

High-Priority Future Research Needs Topic 5

What is the association between sleep apnea severity and long-term clinical outcomes?

Background

The CER reported strong evidence from four studies (three of which were of good quality) demonstrating that baseline AHI >30 events/hr is an independent predictor of all-cause mortality over several years of followup. Evidence for the association of baseline AHI and other long-term clinical outcomes, however, is lacking. To address this research gap, the CER recommended patient-level meta-analyses of existing data from available large cohorts of individuals who had

sleep testing, such as the Sleep Heart Health Study⁶ and the Wisconsin Sleep Cohort Study.⁷ However, other research study options could address the topic.

Stakeholder Discussion

Stakeholders on the FRN panel echoed the need for evidence on the association of baseline AHI and other long-term clinical outcomes—particularly to inform design of studies to assess long-term efficacy of treatment (e.g., early intervention to treat snoring in individuals with low baseline AHI). Stakeholders recommended prospective natural history studies designed to determine whether patient characteristics (e.g., age, sex, race/ethnicity), in addition to disease severity, play a role in determining long-term clinical outcomes.

Proposed Research Designs

To address the overarching FRN topic and its subcomponents, different study designs are reasonable and would address different aspects of this topic.

Patient-Level Meta-analyses

Value of Study Design

Standard meta-analysis of existing studies is not feasible, as discussed in the CER, because of the large degree of clinical heterogeneity in the existing studies, both in terms of their participant eligibility criteria and their analytic methods (e.g., adjusting for different variables). However, patient-level meta-analysis could overcome these problems, by incorporating the differences into a regression analysis and performing a unitary analysis. Existing studies of individuals receiving sleep testing—such as the Sleep Heart Health Study and the Wisconsin Sleep Cohort Study—can be assessed to determine whether their data include a baseline measure of sleep apnea severity (e.g., AHI), and relevant demographic and clinical variables to allow for patient-level meta-analysis. In general, this type of meta-analysis would combine original patient-level data from multiple studies, allowing for unbiased regressions and subgroup analyses. Meta-regression could be used to evaluate associations between sleep apnea severity, and other patient characteristics, with long-term clinical outcomes. Typically, acquiring data is the main obstacle to conducting patient-level meta-analyses. Clinical outcomes of interest would include incident clinical events, quality of life, and psychological or neurocognitive measures.

Resource Use, Size, and Duration

If appropriate data exist and are available, modest resources would be needed to evaluate associations between sleep apnea severity, in combination with other relevant patient characteristics, and long-term clinical outcomes.

Ethical Issues

Meta-analysis uses existing data. With permission to use original study data, there should be no ethical issues.

Post Hoc Analyses of Existing Studies

Value of Study Design

A second-tier option for using existing data to investigate potential associations between sleep apnea severity and long-term clinical outcomes would be to perform post hoc regressions

or subgroup analyses within already-published cohort studies. This could be performed either by the original study investigators or other researchers with access to the study data. To maximize accuracy and validity, such analyses should be done separately in multiple large, high-quality studies. The primary reasons this approach would be a second-tier option to patient-level meta-analysis are that they would be less well powered to show an association and they would be intrinsically less generalizable than an analysis of combined cohorts. However, they may be more feasible since they would not require collaboration of multiple sets of study investigators. Such post hoc analyses may be susceptible to both type I error (falsely significant associations) due to multiple testing or type II error (falsely nonsignificant associations) due to lack of statistical power. Nevertheless, they would form a good basis to determine which demographic and clinical variables could be further investigated a priori in future studies.

Resource Use, Size, and Duration

Post hoc analyses of existing data can be done quickly with modest resources.

Ethical Issues

There should be no ethical issues involved in reanalyzing existing study data.

Prospective Natural History Study

Value of Study Design

Prospective natural history studies can be conducted to investigate potential associations between sleep apnea severity and long-term clinical outcomes, particularly, if existing studies are inappropriate for post hoc or meta-analyses, or they lack data on relevant patient characteristics or outcomes of interest. Studies would require assessment of baseline sleep apnea severity in a specified patient population and documentation of all relevant demographic and clinical variables over time.

Resource Use, Size, and Duration

Moderate resources may be required to recruit and follow a large cohort over a long timeframe.

Ability To Recruit

Patient recruitment should be straightforward and relatively simple, as there would be little added burden for individuals entering a study beyond AHI measurements and documentation of relevant demographic and clinical variables—all of which are components of normal patient care. However, there is potential for large data loss due to long-term followup.

Ethical Issues

Few ethical issues are likely to occur as the study involves components of normal patient care.

Second-Tier Future Research Needs Topics

Second-Tier Future Research Needs Topic 6

What are the barriers to, and predictors of, compliance with different treatments?

Background

Key Question 6 of the CER asked about predictors of treatment compliance. Based on a total of five large cohort studies, there is a moderate strength of evidence that more severe OSA as measured by higher AHI is associated with greater compliance with CPAP use, there is a moderate strength of evidence that a higher measure of sleepiness is also associated with improved compliance, and the strength of evidence is insufficient regarding potential predictors of compliance with MAD.

Stakeholder Discussion

This FRN topic represents two topics that were nominated and discussed separately by the stakeholders. After nomination of topics by the stakeholder panel, the topics were combined into one overarching topic. The original topics were: Studies to predict adherence to CPAP, surgery and MAD; and What are the barriers to compliance?

The original rationale for the first topic was that high quality studies are necessary to determine which factors predict adherence with CPAP and MAD, to which predict successful outcomes after surgery. One stakeholder argued that the topic was moot since treatment must be started and CPAP is superior and thus the clear first choice. Another stakeholder commented that it is very difficult to predict adherence behavior but that it may understanding which factors could influence adherence may allow for designs of better interventions to promote adherence.

The discussion about the more general question of what are the barriers to compliance described several possible barriers, including cognitive impairment due to sleepiness (prior to treatment), a lack of followup care, the value of rapid training on the use of the equipment, the importance of reimbursement for patient education, and a need to demonstrate return-on-investment for interventions that improve compliance. It was stated that diabetes education is reimbursed, but sleep apnea reimbursement is not.

Second-Tier Future Research Needs Topic 7

Direct comparison of compliance rates with different interventions and incorporation of compliance into an overall comparison of effective treatment

Background

Key Question 7 of the CER asked about treatments to improve compliance. Based on 18 trials to improve CPAP compliance, there is a low strength of evidence that some specific adjunct interventions may improve CPAP compliance among overweight patients with more severe OSA who are initiating CPAP treatment. However, studies are heterogeneous and no general type of intervention (e.g., education) was more promising than others.

Stakeholder Discussion

The stakeholder discussion revolved around the need to improve compliance with CPAP to allow for more effective treatment among patients outside the research study setting. It was noted that truck drivers in particular have problems with compliance due to the difficulties of incorporating CPAP into their routine, including the need to sleep in the cab of their trucks without backup power generation. It was put forth that the key reason for poor compliance is lack of proper patient education and aftercare, and that physicians do not have adequate training in educating patients about sleep apnea management. It was stated that nonadherence has been stable at 50 to 60 percent for the past two decades. There was discussion about how cognitive behavioral therapy needs to be researched further as a tool to improve compliance. Trials are needed to compare cognitive behavioral therapy to usual care and different interventions to each other.

Second-Tier Future Research Needs Topic 8

Trials to evaluate weight-loss programs as an adjunctive treatment for sleep apnea

- a. What is the value of bariatric surgery for treatment of sleep apnea?

Background

Key Question 5 of the CER included analyses of weight-loss programs and bariatric surgery. Based on a three trials, there is a low strength of evidence to show that some intensive weight loss programs are effective treatment for OSA in obese patients. A single nonrandomized study provided insufficient evidence regarding bariatric surgery.

Stakeholder Discussion

This FRN topic represents two topics—on weight loss programs and bariatric surgery—that were nominated and discussed separately by the stakeholders. After nomination of topics by the stakeholder panel, the topics were combined into one overarching topic. The rationale regarding weight-loss programs noted that obesity is a risk factor for sleep apnea and that the incremental value of adding weight-loss programs to usual care should be examined in RCTs. It was noted that weight-loss alone has been shown to have a modest effect on AHI, so coupling weight loss with usual care could be synergistic, but that the current evidence is inadequate. Of interest were how practitioners can help patients to improve successful weight loss and which factors predict beneficial effect.

It was also noted that bariatric surgery has been shown to greatly improve AHI in cohorts not selected for sleep apnea, but that the effect specifically in OSA patients needs to be demonstrated. It was discussed that single-treatment cohort studies may be sufficient to demonstrate an effect and that patients would likely not be willing to be randomized to no surgery. The expected effect is on AHI and sleepiness is expected to be large enough to mitigate the need for a trial. One stakeholder discussed how it would be important to categorize patients based on the cause of the OSA, since obesity alone may not be the primary cause in many patients. One group would be those with the known physical anomalies including airway size, tongue, jaw, soft palate and nasal constriction and obesity. The other patient group would be those who are obese only and have no known tongue, jaw, soft palate and nasal constriction issues.

Second-Tier Future Research Needs Topic 9

What is the value of including a sleep medicine specialist in the management of the patient with obstructive sleep apnea?

Background

The CER identified no studies that addressed this topic.

Stakeholder Discussion

The discussion consisted primarily of the rationale. It was noted that patients with OSA are increasingly being treated by primary care physicians. However, most of the research has been done in highly specialized sleep centers with patients under the supervision of a sleep medicine specialist. Thus, it remains unknown how effective interventions may be without the sleep medicine specialist involved. Studies comparing treatment management with and without sleep medicine specialists are of interest (primary care physician alone, primary care physician and specialist together, and specialist alone). Outcomes of interest include adherence, quality of life, cost, and cardiovascular outcomes.

Other Future Research Needs Topics

Other Future Research Needs Topic 10

What is consumer willingness to pay for treatment, to identify consumer preferences for strategies to treat sleep apnea?

Other Future Research Needs Topic 11

What are the financial barriers to access to treatment?

Other Future Research Needs Topic 12

Role of surgery for treatment of OSA

- a. Comparison of surgery versus CPAP
- b. Role of orthognathic surgery (corrective jaw surgery)
- c. Comparison of genio-tubercle advancement versus dental devices

Other Future Research Needs Topic 13

Evaluation of postoperative CPAP for all patients with OSA or at high risk of OSA undergoing any surgery with sedation

Other Future Research Needs Topic 14

Trials comparing CPAP versus pharmaceutical interventions

Other Future Research Needs Topic 15

Trials comparing different CPAP masks

Other Future Research Needs Topic 16

Trials comparing CPAP versus oropharyngeal exercises

Other Future Research Needs Topic 17

Trials comparing different degrees of mandibular advancement

Other Future Research Needs Topic 18

Studies of factors influencing therapist decisions concerning CPAP mask choice

Other Future Research Needs Topic 19

Research into how to maintain patients in sleep apnea studies (where dropout rates are unacceptably high)

This topic was not considered to be high priority by any stakeholder.

Discussion

Challenges in Stakeholder Involvement

We implemented a Web-based discussion board in preference over a series of teleconferences because of what we believed would be advantages of the online approach. These included: greater flexibility for stakeholders to participate in the discussion at times convenient to them, including nonworking hours; a platform that would allow everyone to have a more equal voice, where discussion would not be led by the stakeholder who verbally dominated; full participation by all stakeholders in all discussions, not just those that occurred during calls they were able to attend; a full record of all discussions, without the need to summarize verbal discussions, which inevitably leads to omissions and other errors; less time expenditure by stakeholders who would not be asked to sit through numerous 60- to 90-minute phone calls; considerably less resource expenditure by Evidence-based Practice Center (EPC) staff, where each teleconference hour translates into 7 person-hours of resources; and less opportunity for “multitasking” during discussion (e.g., answering emails while participating in teleconferences).

However, we faced several challenges when implementing the SharePoint Web site discussion board. Perhaps the greatest challenge was the low participation rate of stakeholders during most stages of the project. When the discussion period was announced, few participants logged onto the Web site within the first week. Because of the low log-in rate, combined with the low number of page views for most stakeholders, the discussion period had to be extended. Despite the extension of the discussion period, low participation rates persisted. To increase participation, we sent several email reminders to stakeholders. After the discussion period was closed, participation rates remained low during the nomination period, despite simplification of the process by asking for nomination by email. We subsequently telephoned remaining stakeholders, which increased total participation. Overall, although SharePoint offered the convenience of asynchronous collaboration it seemed to lack appropriate incentives to engage stakeholders. Notably, the seemingly egalitarian Web-based workspace failed to connect patient advocates with the other stakeholder representatives. Assembling patient advocates independently via group teleconference proved more effective, but limited the integration of patient perspective into the SharePoint discussion forums.

Apparent barriers to using the Sharepoint Web site more fully included technical trouble logging in (e.g., while traveling); apparent reluctance by stakeholders to devote unscheduled time to sign in to the Web site, review the discussions and comment; hesitancy commenting in the online discussion by stakeholders without a technical background; and possibly a lack of engagement and interaction from communicating by text instead of verbally. The large number of topics may have been too great a burden for stakeholders to read at a single sitting. In addition, the timing of the project during the summer months may have limited participation because of vacation schedules. A large part of the problem of poor participation may have been simply that the Web-based approach did not force people to schedule a time to participate (as they would for a teleconference), thus there may have been a lack of perceived urgency. Furthermore, for almost all topics, the comments were so sparse that there was little sense of a discussion for stakeholders to participate in.

Several possible solutions exist for these problems. One potential approach that might improve discussion on an online discussion board would be to increase the depth of topic discussion by reducing the number of stakeholders. This could be accomplished by purposely

limiting the number of participants to those who show the ability and proven willingness to participate in discussions at early stages. For example, a run-in phase could allow inclusion of stakeholders, who could be given the opportunity to offer new topics and start discussions, but then further discussion would continue with only the interested stakeholders. In effect, this is what occurred, but an a priori plan to limit the number of participating stakeholders could help to make the process more explicit and possibly shorten the time period required for discussion. A potential downside of this approach is that it could be seen as more explicitly biased since certain stakeholders would be dropped from further participation. Also, the practical problem of how to cordially disinvite stakeholders from further participation would have to be overcome.

Alternatively, the EPC could revert to the more standard and basic approach of conducting a series of teleconferences, where all available stakeholders could join any or all calls. The teleconferences could have either a loose agenda where any new or old topic could be addressed, or a more structured agenda where each call would focus on a general category of Future Research Needs (FRN) topics. We used this approach near the end of the discussion period, when we noted that the consumer stakeholders (the patient, the patient advocate, and the representative from the transportation industry) had not participated in the discussion. We organized a separate teleconference for the three of them and a 1.5 hour teleconference produced nine new topics. The EPC would need to work to ensure that all participating stakeholders are given multiple opportunities to fully express their thoughts. The EPC would also have to accurately summarize the calls in an unbiased fashion, a nontrivial task. The summaries would have to be in a form that could be relatively easily used by the stakeholders.

A similar approach could be to use focus groups to allow a full, simultaneous discussions of each topic. Focus groups could be formed by either collecting stakeholders with similar backgrounds together to focus on specific topics or by randomizing stakeholders to focus groups in order to obtain a balance of viewpoints. These approaches could allow more in-depth discussion of focused topics. However, potential downsides include lessening the participation of less vocal or less topic-knowledgeable individuals and setting up separate subgroup discussions that not all stakeholders have equal access to. This approach may be somewhat simpler to organize and would require less time commitment by stakeholders than multiple teleconferences for all stakeholders.

An approach that combined both teleconferences with online discussions may be able to combine the advantages of the two discussion types, without many of the disadvantages of each. However, this approach is likely to be more time-consuming than the original timeline planned (though possibly less time-consuming than the actual discussion period for the current project). One variation could be an initial gathering of FRN topics and rationale text online, followed by teleconferences to further discuss each topic. The teleconferences would be summarized by the EPC staff and uploaded to the Web site. Further discussion would then be encouraged on line. Clearly, numerous other similar approaches are possible.

Regardless of which type of discussion were held, there may be advantages to limiting the size of the stakeholder panel that could improve discussion. We took the approach of being expansive to include as many voices as feasible. However, in future, it may be better to limit the panel to a single member of each stakeholder category. Attempts should be made to choose stakeholders who show enthusiasm in joining the panel and participating in discussions. However, unfortunately, across all the stakeholders we did not find a strong correlation between the level of enthusiasm upon invitation and the degree of participation in discussions.

In summary, the online Web site discussion of numerous topics by a large stakeholder panel was only moderately successful. Various approaches to improve participation and discussion are possible, including increased use of teleconferences, restricting the size and members of the stakeholder panel, and other approaches. More experience with different approaches is needed.

References

1. Balk EM, Moorthy D, Obadan NO, et al. Diagnosis and Treatment of Obstructive Sleep Apnea in Adults. Comparative Effectiveness Review No. 32. (Prepared by Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-1). AHRQ Publication No. 11-EHC052-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2011.
2. Agency for Healthcare Research and Quality. Future Research Needs Projects. 2010. www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayProduct&productID=521. Accessed September 6, 2011.
3. Balk EM, Chung M, Moorthy D, Chan JA, Patel K, Concannon TW, Ratichek SJ, Chang LKW. Diagnosis of Obstructive Sleep Apnea: Future Research Needs. Future Research Needs Paper No. 11. (Prepared by Tufts Evidence-based Practice Center under Contract No. 290-2007-10055 I.) AHRQ Publication No. 12-EHC031-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2012. www.effectivehealthcare.gov/reports/final.cfm.
4. Agency for Healthcare Research and Quality. Framework for Considering Study Designs for Future Research Needs (FRN). Draft Comparative Effectiveness Review. Rockville, MD: Agency for Healthcare Research and Quality; September 2011. www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=774.
5. Richards D, Bartlett DJ, Wong K, et al. Increased adherence to CPAP with a group cognitive behavioral treatment intervention: a randomized trial. *Sleep*. 2007 May 1;30(5):635-40.
6. Punjabi NM, Caffo BS, Goodwin JL, et al. Sleep-disordered breathing and mortality: a prospective cohort study. *PLoS Medicine/Public Library of Science*. 2009 Aug;6(8):e1000132.
7. Young T, Finn L, Peppard PE, et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. *Sleep*. 2008 Aug 1;31(8):1071-78.

Acronyms

AHI	Apnea-hypopnea index
AHRQ	Agency for Healthcare Research and Quality
CER	Comparative Effectiveness Review
CPAP	Continuous positive airway pressure (device)
EPC	Evidence-based Practice Center
FRN	Future research needs
MAD	mandibular advancement device
OSA	Obstructive sleep apnea
QALY	Quality-adjusted life-years
RCT	Randomized controlled trial

Appendix A. AHRQ's Effective Health Care Program Selection Criteria for New Research

1. Appropriateness	1a. Represents a health care drug, intervention, device, technology, or health care system/setting available (or soon to be available) in the United States
	1b. Relevant to 1013 enrollees (Medicare, Medicaid, S-CHIP, other federal health care programs)
	1c. Represents one of the priority conditions designated by the Department of Health and Human Services (DHHS)
2. Importance	2a. Represents a significant disease burden; large proportion or priority population
	2b. Is of high public interest; affects health care decisionmaking, outcomes, or costs for a large proportion of the US population or for a priority population in particular
	2c. Was nominated/strongly supported by one or more stakeholder groups
	2d. Represents important uncertainty for decisionmakers
	2e. Incorporates issues around both clinical benefits and potential clinical harms
	2f. Represents important variation in clinical care, or controversy in what constitutes appropriate clinical care
	2g. Represents high costs due to common use, to high unit costs, or to high associated costs to consumers, to patients, to health care systems, or to payers
3. Desirability of New Research/Duplication	3. Would not be redundant (i.e., the proposed new research is not sufficiently researched by AHRQ or others, considering both completed and in-process research)
4. Potential Impact	4a. Potential for significant health impact: - To improve health outcomes - To reduce significant variation in clinical practices known to be related to quality of care - To reduce unnecessary burden on those with health care problems
	4b. Potential for significant economic impact: - To reduce unnecessary or excessive costs
	4c. Potential for change: - The proposed topic exists within a clinical, consumer, or policymaking context that is amenable to evidence-based change - A product from the EHC program could be an appropriate vehicle
	4d. Potential risk from inaction: - Unintended harms from lack of prioritization of a nominated topic
	4e. Addresses inequities, vulnerable populations (including issues for patient subgroups)
	4f. Addresses a topic that has clear implications for resolving important dilemmas in health and health care decisions made by one or more stakeholder groups
	5. Capacity
	5b. Utilizes existing AHRQ resources or builds desired additional research capacity or decisional support for the EHC Program
	5c. Costs associated with the likely study design are reasonable considering limited program resources