

Challenges in Clinical Trials of SDB Devices

Edward M. Weaver, MD, MPH
University of Washington
VA Puget Sound

Disclosures

- No industry, commercial, or financial conflicts of interest
- Presenting my opinions

My Opinions

Informed by:

- My clinical and trials experience
- International Surgical Sleep Society (ISSS)
Surgery Trials Group consensus (10/2014 & 5/2017)
- American Academy of Sleep Medicine (AASM)
Sleep Apnea Quality Measures Task Force paper
– JCSM 2015; 11(3): 357-383
- AASM Sleep Surgery Clinical Practice Guideline
Task Force consensus (11/2017)
- Other published studies

Outline

- Panel 3 questions overview
 - Questions
 - Considerations
 - Sources to inform me
 - My proposals

Panel 3:

Clinical Trial Design Considerations for Therapeutic SDB Devices

Questions

- Control group?
- Minimum duration of the study (trial)?
- Objective primary outcomes?
- Clinically meaningful differences?
- Patient-reported outcomes?

Q1: What is the most appropriate control group?

- Considerations:
 - Evidence quality
 - Relevance
 - Feasibility
 - Ideally both randomized and observational trials

Q1: What is the most appropriate control group?

- Control proposal: ~~No treatment~~
 - Ideally placebo when possible
 - Delay to active treatment
 - Continued attempts at failed medical therapy

Q1: What is the most appropriate control group?

- Comparative effectiveness (treatment control) challenges:
 - Patient must be appropriate for both therapies
 - Likely small differences requires large sample
 - Equivalence study requires large sample
 - No single gold standard to compare against
 - Goal to identify therapies, not identify best

Q1: What is the most appropriate control group?

- **Baseline:**
 - Case series
 - Appropriate for truly new therapy
 - Trial hard to justify

Q2: What is the minimum duration of the study?

- Considerations:
 - Treatment acclimation / recovery
 - New steady state
 - CPAP, Oral Appliance, Surgery device, Wt loss device, other
 - Acclimation, recovery, weight stability, sleep stability
 - Longer is better but less feasible
 - Future information systems might make longer-term outcomes feasible
 - Registries, electronic health records, others

Q2: What is the minimum duration of the study?

- Duration proposal: 3 months steady state
 - Practical
 - Likely represents medium-term outcomes

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Considerations:
 - Primary endpoint might best be subjective
 - Clinical disease burden
 - Sleep apnea testing
 - Efficacy measures (not effectiveness measures)
 - Must account for treatment adherence
 - Does not fully reflect clinical disease burden

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Considerations:
 - Other objective tests (eg, MSLT, MWT, PVT, ...)
 - Availability, cost, relevance
 - Health outcomes
 - Impractical duration
 - Surrogate outcomes (eg, biomarkers)
 - Beware of not reflecting clinical outcomes

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Sources:
 - AASM Sleep Apnea Quality Measures TF
 - ISSS Surgery Trials Group
 - AASM Sleep Surgery Clinical Practice Guideline Task Force

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Objective parameter proposal: Sleep apnea testing

Standardized Outcome Measures

International Surgical Sleep Society
Surgery Trials Group
10/23/2014
5/4/2017

Charge

- Come up with a list of outcome measures that the surgery trials group (and others) will be willing and invested in collecting routinely on all patients.
 - Subjective
 - Objective
 - Separate from phenotyping variables

Approach

- Brainstorm list
 - Each team member:
 - What already collect routinely?
 - Why?
 - What else considering collecting?
 - These are measures that have been already considered and vetted on some level for this purpose

Approach

- Evaluate brainstorm list
 - Ease, availability
 - Value as outcome measure
 - Validity, reliability
 - Familiarity
 - Not based primarily on literature

Approach

- Proposed short list
 - Pros, cons, discussion points
 - Open for disc, additions, subtractions
 - Refine list with whole group
 - Goal: Final list we all can buy into
 - Note: Anyone can collect whatever else they want—we just want a minimum set that all will collect

Brainstormed List

- Questionnaires
- Exam
- Physiologic Measures
- Sleep Test
- Complications
- Indexes

Brainstormed Sleep Testing

- Many parameters
- Resp: AHI, AI, RDI
- O2: ODI, LSAT, MSAT, %TST<90%
- Sleep arch: SE, N3%, REM%
- Test type
- Others

Proposed Sleep Testing

- Apnea-hypopnea index
 - Pros: Expected, usually reported, contributes to SASI
 - Cons: Variable defn, unreliable, poor surrogate, home tests don't capture arousal hypopneas

Proposed Sleep Testing

- Apnea index
 - Pros: Standard defn, more obj than AHI, more reliable than AHI, more clinically important
 - Cons: Sometimes not reported

The severity of individual obstruction events is related to increased mortality rate in severe obstructive sleep apnea

ANU MURAJA-MURRO^{1,2}, ANTTI KULKAS³, MIKKO HILTUNEN¹,
SALLA KUPARI¹, TAINA HUKKANEN¹, PEKKA TIIHONEN¹,
ESA MERVAALA^{1,2} and JUHA TÖYRÄS^{1,4,5}

Based on the multivariate logistic regression analysis, obstruction severity was the only parameter which was related statistically significantly to mortality in the severe OSA category.

Proposed Sleep Testing

- Oxygen desaturation index
 - Pros: More objective (scored by computer), reliable, proxy for intermittent hypoxia (pathophys of morbidity)
 - Cons: Variable defn (3% v 4%), not always reported

Sleep-disordered Breathing and Cardiovascular Disease

An Outcome-based Definition of Hypopneas

Naresh M. Punjabi¹, Anne B. Newman², Terry B. Young³, Helaine E. Resnick⁴, and Mark H. Sanders⁵

riates. The adjusted prevalent odds ratios for quartiles of the hypopnea index using a 4% desaturation criterion were as follows: 1.00 (<1.10 events/h), 1.10 (1.01–3.20 events/h), 1.33 (3.21–7.69 events/h), and 1.41 (>7.69 events/h). Hypopnea measures based on less than 4% oxyhemoglobin desaturation or presence of arousals showed no association with cardiovascular disease.

Level 2

Punjabi, AJRCCM 2008;177:1150-5

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Objective parameter proposal: Sleep apnea testing
 - Apnea-hypopnea index (expected)
 - Apnea index (important, consistent, reliable)
 - Oxygen desaturation index 4% (clinical import)
 - In-lab vs home (Panel 2)

Q3: What objective parameter(s) should be used for the primary effectiveness endpoints?

- Objective parameter proposal:
Treatment adherence
 - Devices that depend on adherence for outcome
 - Use adherence to convert sleep testing efficacy to effectiveness measures
 - Mean nightly index = Index w/ device * % use device
 - Sleep 2011; 34(1): 105-110.

Q4: What are the clinically meaningful differences for the primary endpoints?

- Considerations:
 - Not formally defined for sleep testing
 - Clinical judgment
 - Anchor methods or effect sizes
 - Depends on severity of sleep apnea
 - Define *minimal* clinically important difference (MCID)

Q4: What are the clinically meaningful differences for the primary endpoints?

- Sources:
 - AASM Sleep Surgery Clinical Practice Guideline Task Force consensus (11/2/2017)
 - NHMRC Sleep Surgery RCT

Q4: What are the clinically meaningful differences for the primary endpoints?

- MCID proposal:
 - Sleep test parameters *change*: 10 events/hr
 - Balance between AASM Sleep Surgery Clinical Practice Guideline Task Force (11/2/2017) & NHMRC Sleep Surgery RCT
 - Meaningful shift of severity category with morbidity implications

Q4: What are the clinically meaningful differences for the primary endpoints?

- MCID proposal:
 - Sleep test parameters *change*: 10 events/hr
 - Corrected for adherence
 - AHI 20 → 0 in the lab requires use 50% of sleep time
 - AHI 40 → 0 in the lab requires use 25% of sleep time

Q4: What are the clinically meaningful differences for the primary endpoints?

- MCID proposal:
 - Alternatively, treat adherence as its own outcome variable
 - Treatment adherence: 4 hrs/night x 70% nights
 - Approximately 50% of recommended sleep time
 - Current standard for CPAP

Q4: What are the clinically meaningful differences for the primary endpoints?

- MCID proposal:
 - Treatment adherence *change*: 0.5 hours/night
 - Treatment adherence *change*: 10% nights used
 - Treatment acceptance *change*: 10% of people
 - Am Acad Sleep Medicine Surgery Clinical Practice Guideline Task Force (11/2/2017)

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- Considerations:
 - PRO important to patients
 - PRO are effectiveness measures
 - Practical
 - Validity, reliability, responsiveness, burden
 - Prone to bias (recall bias, placebo effect, etc.)

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- Considerations:
 - Symptoms
 - Health status
 - Functional status
 - Quality of life
 - Adverse effects

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- Considerations:
 - Sleep testing parameters do NOT reflect PRO
 - Device trials should include PRO

Polysomnography indexes are discordant with quality of life, symptoms, and reaction times in sleep apnea patients

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

CONCLUSIONS: PSG indexes are not consistently associated with sleepiness, quality of life, or reaction time, both at baseline and as outcome measures in patients with mild-moderate OSAS. PSG indexes may not quantify some important aspects of OSAS disease burden or treatment outcome. Clinically important outcomes should be measured directly.

Level 1

Weaver, Oto-HNS 2005; 132: 255-62

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- Sources:
 - Published studies
 - AASM Sleep Apnea Quality Measures TF
 - ISSS Surgery Trials Group
 - AASM Sleep Surgery Clinical Practice Guideline Task Force

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- PRO proposal: Include
 - Patient-reported benefit (preferably validated)
 - Adverse Effects:
 - Serious
 - Device specific

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- PRO suggestions: Symptoms
 - Snoring bother (MCID 25%)
 - Epworth Sleepiness Scale (MCID 2/24)

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- PRO suggestions: Functional status
 - Functional Outcomes of Sleep Questionnaire (FOSQ) (MCID 1 on scale 5-20)

Q5: What patient-reported outcomes (PRO) are appropriate in the evaluation of SDB devices?

- PRO suggestions: Quality of life
 - Symptoms of Nocturnal Obstruction & Related Events (SNORE-25) (MCID 0.5 on scale 0-5)
 - Global OSA-related Quality of Life Change (MCID 3 on 15-point scale)

Panel 3:

Clinical Trial Design Considerations for Therapeutic SDB Devices

Questions

- Control group?
- Minimum duration of the study?
- Objective primary outcomes?
- Clinically meaningful differences?
- Patient-reported outcomes?

Outline

- Panel 3 questions overview
 - Questions
 - Considerations
 - Sources to inform me
 - My proposals