BMJ Open

Pediatric Adenotonsillectomy Trial for Snoring (PATS): protocol for a randomised controlled trial to evaluate the effect of adenotonsillectomy in treating mild obstructive sleep-disordered breathing

Rui Wang [iD] ,[1,2] Jessie P. Bakker,[3] Ronald D. Chervin,[4] Susan L. Garetz,[5] Fauziya Hassan,[6] Stacey L. Ishman,[7,8] Ron B. Mitchell,[9] Michael G. Morrical,[10] Syed K. Naqvi,[9] Jerilynn Radcliffe,[11,12] Emily I. Riggan,[13] Carol L. Rosen,[14] Kristie Ross,[14] Michael Rueschman,[10] Ignacio E. Tapia,[11,15] H. Gerry Taylor,[14,16] David A. Zopf,[5] Susan Redline[3]

For numbered affiliations see end of article.

Correspondence to
Dr Rui Wang;
rwang@hsph.harvard.edu

ABSTRACT

Introduction Mild obstructive sleep-disordered breathing (oSDB), characterised by habitual snoring without frequent apnoeas and hypopnoeas on polysomnography, is prevalent in children and commonly treated with adenotonsillectomy (AT). However, the absence of high-level evidence addressing the role of AT in improving health and behavioural outcomes has contributed to significant geographical variations in care and potential for surgery to be both overused and underused.

Methods and analysis The Pediatric Adenotonsillectomy Trial for Snoring (PATS) is a single-blinded, multicentre randomised controlled trial designed to evaluate the effect of AT in treating mild oSDB. Four hundred sixty eligible children, aged 3.0–12.9 years old, will be randomised to either early adenotonsillectomy or to watchful waiting with supportive care (WWSC) with a 1:1 ratio. The study's coprimary endpoints are (1) change from baseline in executive behaviour relating to self-regulation and organisation skills as measured by the Behavioural Rating Inventory of Executive Function (BRIEF) Global Composite Score (GEC); and (2) change from baseline in vigilance as measured on the Go-No-Go (GNG) signal detection parameter (d-prime). A mixed effects model will be used to compare changes in the BRIEF GEC score and GNG score at 6 and 12 months from baseline between the AT arm and the WWSC arm.

Ethics and dissemination The study protocol was approved by the institutional review board (IRB) at Children's Hospital of Philadelphia (CHOP) on 3 October 2014 (14–0 11 214). The approval of CHOP as the central IRB of record was granted on 29 February 2016. The results will be published in peer-reviewed journals and presented at academic conferences. The data collected from the PATS study will be deposited in a repository (National Sleep Research Resource, sleepdata.org) after completion of the study to maximise use by the scientific community.

Strengths and limitations of this study

► Evaluation of the benefit and adverse effects of surgical intervention versus watchful waiting, including assessment of associated healthcare utilisation (HCU).
► Adoption of coprimary endpoints that include parent-reported and objectively collected performance-based neurocognitive measures.
► Collection of a large variety of data from multiple sources (child, caregiver, teacher and neighbourhood geocode) and across multiple domains (neuro-behaviour, polysomnography, actigraphy, symptoms, quality of life, anthropometry, blood pressure, HCU, tobacco exposure and immunoglobulin titres).
► Supported by an Informatics and Data Management Core that develops and integrates cutting-edge, open-source web development tools and dynamic research data.
► Double blinding was not feasible for a surgical trial in children.

Trial registration number NCT02562040; Pre-results.

INTRODUCTION

Obstructive sleep-disordered breathing (oSDB) is common in the paediatric population and is associated with significant morbidity.[1] Adenotonsillectomy (AT), the second most common surgery performed under general anaesthesia in children (more than 289 000 times per year in the USA),[2] is generally considered the first line treatment for oSDB in otherwise healthy children aged 2–18 years with adenotonsillar hypertrophy.[3]

This procedure is often performed on children with symptoms of oSBD without polysomnographic evidence of frequent apnoeas or hypopnoeas.[4] The single randomised controlled study examining outcomes of paediatric AT for obstructive sleep apnoea (OSA) syndrome (Childhood Adenotonsillectomy Trial, CHAT) included only children with polysomnographically-documented OSA.[5] The CHAT study found that AT compared with watchful waiting resulted in improved behaviour, quality of life, OSA symptoms and polysomnographic parameters, but did not lead to significant improvement in objective measures of attention or executive function. Of note, almost half of the children not undergoing AT had polysomnographic resolution of the OSA over a 7-month period.[5]

While screening children for inclusion in the CHAT study, it became apparent that almost half of symptomatic children considered to be surgical candidates for AT had what are often considered less severe forms of oSDB, including snoring, flow limitation or mild OSA (obstructive Apnoea–Hypopnoea Index (oAHI) score of <3).[6] These entities could be grouped together and classified as mild sleep-disordered breathing (SDB) (mild oSDB). Evidence to date has shown little correlation between severity of oSBD and neurocognitive morbidity.[6–8] However, several studies have demonstrated that mild oSDB is associated with more severe neurobehavioural impairment that is more easily reversed with appropriate intervention.[7 8] Rigorously controlled data are not available on the benefits of AT for mild oSDB, or for treating younger children, who may be most sensitive to the effects of sleep problems due to developmental plasticity. Lack of data has led to huge geographical variability in the USA with regard to the management of mild oSDB, with the rate of AT per 10 000 children varying from 28.9 in the West to 125.1 in the South.[9] Unnecessary surgery may expose children to risk, and the healthcare system to considerable costs. Conversely, withholding effective treatment from children could result in substantial short-term and long-term health burdens to the children, their family and society. Effective and timely treatment could also potentially reduce healthcare costs associated with symptoms and comorbidities that are exacerbated by the presence of mild oSDB. Understanding the role of treatment for mild oSDB is of especial importance, given the increased prevalence of SDB among vulnerable groups of children, such as racial minorities.[10] Filling these gaps in knowledge is critical to inform clinical guidelines, decision-making and appropriate use of interventions in populations most likely to benefit.

The goal of the Pediatric Adenotonsillectomy Trial for Snoring (PATS, 'The impact of treatment of mild SDB in children's health') is to provide high-quality evidence regarding the effects of surgical intervention versus watchful waiting (observation) on a group of healthy children with nocturnal obstructive symptoms whose polysomnograms demonstrate mild oSDB. This study was specifically designed to evaluate the effectiveness of AT as well as associated healthcare utilisation (HCU) in children with mild oSDB. In this article, we present the PATS protocol V.19 (19 February 2019) and describe the unique challenges in designing a randomised controlled trial (RCT) of a surgical intervention in young children, including selecting appropriate outcomes, determining approaches for collection of HCU data across geographically diverse US sites and optimising data collection in studies of young children.

METHODS AND ANALYSIS
Study overview
PATS is a multicentre, randomised, single-blinded 12-month intervention study that compares the impact of AT on measures of behaviour, quality of life, sleep-related symptoms, polysomnographic findings and HCU in children with mild oSDB (figure 1). Children with symptoms of mild oSDB are recruited from each site's otolaryngology, sleep, pulmonary and/or general paediatric clinics. At baseline, participants undergo neurobehavioural testing and polysomnography (PSG) and assessment of patient-reported outcomes (sleepiness, quality of life and sleep quality), anthropometry and blood pressure. All measures are repeated at 6 and 12 months, except that the PSG is only repeated at 12 months. In addition to visits at baseline and 6 and 12 months, participants receive monthly telephone calls to maximise retention and to collect interim data on symptoms and HCU. The study started enrolment in June 2016. As of 26 August 2019, 344 children have been randomised. The reporting of the PATS protocol follows the Standard Protocol Items: Recommendations for Intervention Trials reporting guidelines,[11] which were developed in 2013 to establish the minimum content of a clinical trial protocol.

Study aims and endpoints
The primary objectives were to determine the effect of early adenotonsillectomy (eAT) versus watchful waiting with supportive care (WWSC) on a coprimary outcome: executive function assessed by a parent behaviour rating (Behaviour Rating Inventory of Executive Function Global Executive Composite, second edition or preschool version; Behavioural Rating Inventory of Executive Function (BRIEF)2/P Global Composite Score (GEC)[12 13] and children's sustained attention or vigilance as assessed by signal detection parameter (d-prime) for performance on the Continuous Performance Test from the Go-No-Go (GNG) task.[14]

The secondary objectives were to determine the effect of eAT versus WWSC on oSDB symptoms and quality of life. We also will track and compare group changes in HCU occurring within each site's medical system and externally, as well as filled prescriptions. Exploratory analyses propose assessment of changes in anthropometry and blood pressure, and identification of factors that moderate the response to AT.
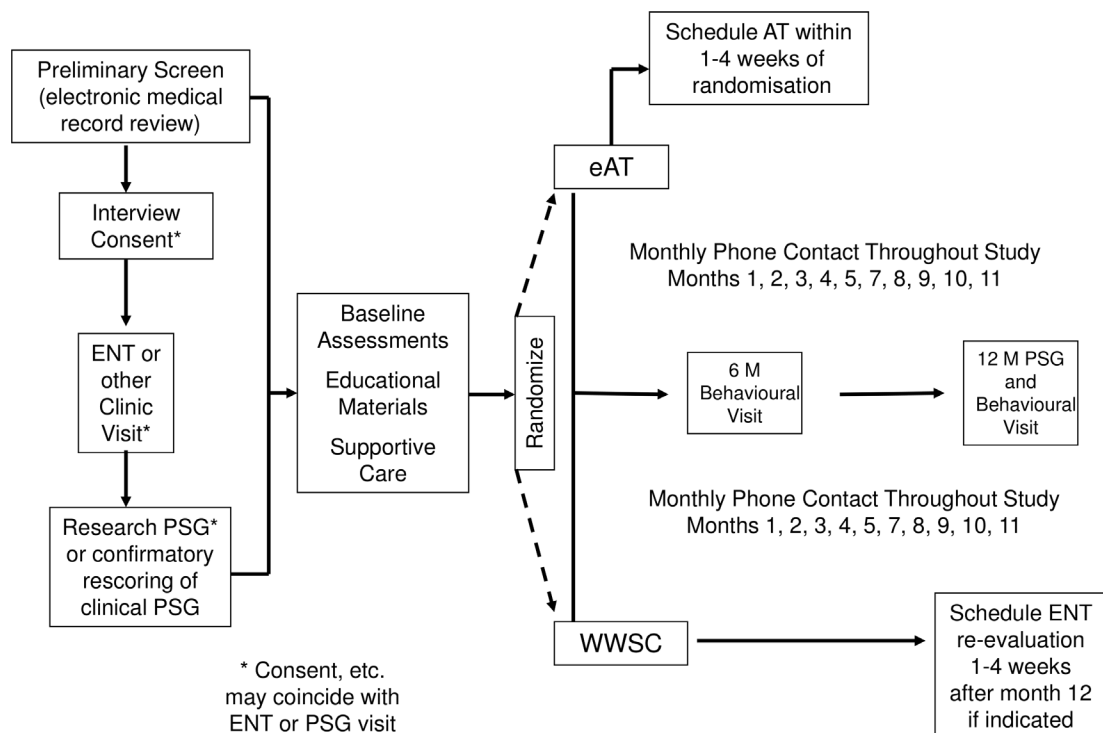
Figure 1 An overview of the study design. AT, adenotonsillectomy; eAT, early adenotonsillectomy; ENT, Otolaryngology (ear, nose and throat); PSG, polysomnography; WWSC, watchful waiting with supportive care.

Secondary outcomes include measurements from a range of domains: objective performance testing, behavioural scales, quality of life, physical examination and HCU. Potential effect modifiers to be evaluated include demographics, sleep duration and efficiency, asthma/atopy measures, secondhand smoke exposure and measures of family functioning. A detailed list of these outcomes and assessment procedures is provided in table 1.

Study organisation

The study is supported by a Data Coordination Center (DCC)/Sleep Reading Center (Brigham and Women's Hospital, Boston, Massachusetts, USA), charged with development of the study's statistical design and monitoring plans, construction and management of the study database and study materials, and generation of statistical reports to investigators and the PATS Data and Safety Monitoring Board (DSMB). The Sleep Reading Center is charged with centralised PSG scoring and generation of standardised PSG variables. A surgical quality assurance core is based at the University of Michigan, Ann Arbour, Michigan, USA. A neuropsychology core is provided by psychologists at two sites (Children's Hospital of Philadelphia (CHOP), Pennsylvania, USA, and Nationwide Children's Hospital Research Institute, Columbus, Ohio, USA).

The study is also supported by a Clinical Coordinating Center (CCC), CHOP, Pennsylvania, charged with overseeing the activities at the clinical sites and regulatory approvals, and providing clinical expertise. Clinical sites are each headed by a sleep medicine physician or an

otolaryngologist and, together with their local research team (study coordinators and sleep laboratory staff), are responsible for recruitment and follow-up of participants. Initially, five clinical sites (CHOP, Pennsylvania, USA; Cincinnati Children's Medical Center, Ohio, USA; Rainbow Babies & Children's Hospital at University Hospitals, Cleveland Medical Centre, Ohio, USA; University of Michigan Health System, Ann Arbour, Michigan, USA; University of Texas Southwestern Medical Centre, Dallas, Texas, USA) were identified to participate in the study. In July 2018, two new sites (Children's Hospital, Boston, Massachusetts, USA, and Children's Hospital of the King's Daughters, Norfolk, Virginia, USA) were added to improve subject accrual. In June 2019, Boston Children's Hospital was closed to accrual on DSMB's recommendation due to its slow accrual, resulting in six recruitment sites.

Study governance is through a steering committee with representation from each participating site, key quality control cores and National Health Lung Blood Institute (NHLBI) programme staff. An executive committee, consisting of the study chair, the DCC directors, CCC director and the NHLBI project officer, who regularly meets by telephone to address emerging issues. Subcommittees are organised to address the multiple quality control and monitoring needs of the study: surgical quality control, neuropsychology quality control, polysomnography quality, recruitment and operations, and publications and presentations. An independent DSMB, with expertise in paediatric ethics, surgery, sleep apnoea, clinical trials and biostatistics, appointed by and

Table 1 Primary and secondary endpoints, prespecified candidate moderators

Primary endpoints	BRIEF2/P Global Executive Composite Score[12 13]
	GNG sustained attention d-prime parameter[14]
Secondary endpoints	
Objective performance testing	GNG inhibitory control d-prime[14]
	Fine motor coordination: NIH-Toolbox 9-Hole Pegboard Dexterity Test
Behavioural scale	Executive function: BRIEF 2/P meta-cognition and emotional regulation summary scores and subscales for parent and teacher reports
	Behaviour: Child Behaviour Checklist[23] summary scale and subscores, parent and teacher ratings
	Attention: Conners 3 Short Form (caregiver and teacher versions) Global Index T score and subscales[24]
Sleep-disordered breathing symptoms	PSQ-SRBD Scale total score[25]
	Sleepiness: Epworth Sleepiness Scale modified for children summary score and PSQ-SRBD sleepiness scale[26]
	Snoring: the patch snoring sensor
Quality of life	Generic: Paediatric Quality of Life Inventory total score and subscores[27]
	Disease specific: Obstructive Sleep Apnea-18 (OSA-18) total score
Physical exam	Measurements of weight; height; body mass index; waist, hip and neck circumferences
	Systolic, diastolic and mean blood pressure levels
Healthcare use	Medications, healthcare visits (scheduled and unscheduled), ascertained from caregiver reports, EMR surveillance, billing and pharmacy records and hospitalisations
Potential effect modifiers	
Demographics	Race, SES (parent education, family income, financial stress rating scale and geocode data on neighbourhood characteristics)
Sleep duration and efficiency	Objective assessment by 7-day wrist actigraphy
Asthma/atopy	IgE, International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire, review of EMR and parent interview (using NHLBI asthma definitions based on a history of asthma and use of asthma medications)

Continued

Table 1 Continued

Secondhand smoke exposure	Urinary cotinine
Family functioning cluster	Family functioning (Family Assessment Device, short form); parenting style (Parenting Style Questionnaire); parent perception of stress (Parenting Stress Index Fourth Edition, short form); medical literacy (Rapid Estimate of Adult Literacy in Medicine, Revised) and discrimination (Experiences of Discrimination)

BRIEF, Behavioural Rating Inventory of Executive Function; EMR, electronic medical record; GNG, Go-No-Go; NIH, National Institutes of Health; PSQ-SRBD, Paediatric Sleep Questionnaire: Sleep-Related Breathing Disorder.

reporting directly to the NHLBI, reviews quarterly reports and meets semiannually to assess the emerging data and make recommendations. A board-certified paediatric sleep medicine physician is continuously available as an independent medical monitor (MM).

Sample population and enrolment

This study recruits children with symptoms of mild oSDB and their caregivers. The inclusion criteria are (1) ages 3.0–12.9 years at the time of screening; (2) diagnosis of mild oSDB defined as (a) parent report of habitual snoring that occurs most of the night on at least three nights per week and has been present for at least 3 months (on average occurring more than three nights per week or more than one-half the sleep time) and (b) obstructive apnoea index of <1/hour and oAHI of <3/hour and no oxyhaemoglobin desaturation of <90% in conjunction with obstructive events, confirmed on nocturnal, laboratory-based PSG; (3) tonsillar hypertrophy of ≥2 based on a standardised scale of 0–4; (4) determination to be a candidate for AT by ear, nose and throat (ENT) evaluation; and (5) primary indication for AT is nocturnal obstructive symptoms. As in all RCTs, equipoise about randomisation in PATS is required on the part of participants, their families and their clinicians (ENT surgeons).

The exclusion criteria are (1) previous tonsillectomy; (2) recurrent tonsillitis that merits prompt AT per the American Academy of Otolaryngology Head and Neck Surgery Clinical Practice Guidelines[3]; (3) severe obesity (body mass index (BMI) z-score of ≥3); (4) severe chronic health conditions that might hamper participation or confound key variables under study, including but not limited to (a) severe cardiopulmonary disorders; (b) bleeding disorders; (c) sickle cell disease; (d) epilepsy requiring medication; (e) other severe chronic health problems such as diabetes or narcolepsy; (f) mental retardation or being assigned to a self-contained classroom for all academic subjects; (g) known genetic, craniofacial, neurological or psychiatric conditions likely to affect the airway, cognition or behaviour; and (h) psychiatric

or behavioural disorders requiring or likely to require initiation of new medication, therapy or other specific treatment during the 12-month trial period (other than attention-deficit hyperactivity disorder (ADHD)); children with ADHD are included, but those with autism spectrum disorder or those with global development impairment are excluded; (5) current use of psychotropic medication (other than medications for ADHD), hypnotics, antihypertensives or growth hormone; chronic corticosteroids are allowable, although children with a burst of oral corticosteroid therapy for asthma are deferred until corticosteroids are no longer prescribed and it has been 30 days since the last dose; medication use during the study is captured on a monthly basis via phone or in-person interviews using a structured case report form completed by research coordinators; in addition, prescriptions that are filled within the local medical system are captured by accessing HCU data on a quarterly basis; (6) history of severe developmental disability or Adaptive Behavioural Assessment System score of <60; (7) parent/guardian unable to accompany the child on the night of the PSG; (8) family planning to move out of the area within the year; (9) family does not speak English or Spanish well enough to complete the behavioural and performance measures; and (10) child being in foster care.

Study interventions

Depending on the randomised treatment assignment, participants are assigned to either WWSC or eAT. Within 4 weeks of randomisation, participants randomised to the eAT arm undergo surgery under general anaesthesia as part of routine clinical care. Surgery is performed by board-certified otolaryngologists with or without the assistance of resident physicians in an accredited otolaryngology training programme. Prior to the surgical procedure, tonsillar size is graded using a standardised scale of 0–4.[15] Extent of adenoid tissue is graded as mild (0%–33%), moderate (34%–66%) or severe (67%–100%) obstruction of the posterior choanae intraoperatively in subjects undergoing AT. Complete bilateral tonsillectomy and removal of obstructing adenoid tissue are performed by cold dissection, monopolar electrocautery or any other recognised surgical technique.

Regardless of the treatment assignment, all participants undergo sleep and healthy lifestyle education. Standardised materials recommended by the National Institutes of Health (NIH) and paediatric professional sleep societies are used to reinforce optimal sleep health, and educational play is encouraged by providing take-home materials addressing sleep health. Other supportive care is provided at initial evaluation and as needed throughout the course of the trial. For example, participants identified as having suboptimal asthma or nasal allergy control will be referred to their primary care physician for management and further treatment of these problems.

After 12 months, children who did not undergo AT who have a 12-month PSG showing concerns for oSDB or whose parent reports ongoing symptoms/concerns are referred back to ENT for further clinical management (such as AT, if still indicated) as per standard clinical care.

Blinding

As in the Childhood Adenotonsillectomy Trial (CHAT),[16] the use of a surgical intervention prevents blinding of the child, parent and certain staff members because performing sham surgery in children raises ethnical and feasibility concerns. PATS adopts a similar approach where the principal investigators (PIs) at each site (other than sites at which the PI is a surgeon), and study coordinators who directly collect primary outcomes are blinded to study treatment. In addition, all DCC and CCC staff, except for those responsible for statistical analyses, data management and AE adjudication and communication, are blinded. The responsibilities of blinded and unblinded staff at each site have been clearly delineated, and a structured format for communication was established to minimise the impact of the unblinding on study outcomes and study progress.

Neurobehavioural testing

To ensure reliable collection of neurobehavioural test data, much attention was directed at developing a rigorous protocol for training research assistants to properly administer the tests. Initial training was provided by in-person review and demonstration of procedures. Examiners later reviewed administration procedures, practiced the assessments with other team members and made video recordings of an assessment conducted with a child volunteer. To ensure fidelity of test administration, the videos were reviewed by one of the two psychologists in the neurobehavioral core, with feedback provided and additional assessments required if procedures did not meet specific competency criteria. The challenge in testing young children, some of whom had limited attention spans and difficulty in following through on test instructions, was addressed by selecting engaging tests that were 'hands-on' and could be easily understood by children as young as 3 years. Testing procedures included defined opportunities for children to practice, with repetition of instructions. Recognising that despite these procedures, there would still be some variation in engaging children, protocols were developed to allow the examiners to document behaviours that may have contributed to test performance, such as inattention or off-task behaviours (eg, 'Child pushes button repeatedly without reference to the screen').

Informatics and Data Management Core (IDMC)

PATS is supported by an innovative IDMC that develops and integrates cutting-edge, open-source web development tools and dynamic research data, providing robust and highly interactive tools for multicenter studies, clinical trials and data repositories. These tools are developed

and updated on a weekly basis using a continuous deployment methodology based on the agile software development framework. The core provides thorough documentation of the software and the deployment architecture in the form of online version-controlled documentation and web-based video tutorials. Electronic data entry is supported by the software program Slice (https://sliceable.org), which excels at dynamic in-application reporting and provides researchers, individual clinical sites, the DSMB and the sponsor, a live snapshot of the current state of the database. Slice also provides robust project management tools, including the ability to easily create and track participant/study schedules. Data interoperability is handled by the Spout JavaScript Objective Notation (https://www.json.org/) data dictionary framework to modularise data definitions into small, maintainable versioned data element descriptors. Finally, the IDMC promotes data liberation, enabling researchers to export all data they have entered at any point in a useable format that can be imported into a new system of their choice.

HCU and electronic medical record (EMR) surveillance

The study addressed challenges in collecting consistent and complete HCU from multiple, diverse medical centres that use different EMRs and from families who may use healthcare services outside of PATS clinical sites. To comprehensively identify episodes of HCU, a surveillance approach was developed that includes the following: (1) a semistructured interview undertaken on a monthly basis with caregivers when information is gathered regarding any HCU 'billed and filled' (ie, any healthcare encounter and any filled prescription); (2) the local EMR is queried approximately quarterly in order to ensure that no internal HCU (encounters or prescriptions within the local medical system) was missed during caregiver interviews; and (3) attempts are made to receive medical reports based on any caregiver reports of external HCU (encounters or prescriptions outside of the local medical system that are not visible in the local EMR). Prior to study start, metadata for common paediatric HCU events were identified to develop a standardised HCU data dictionary which was supplied to each participating site. A medical record analyst from each site was asked to develop an electronic query designed to pull appropriate data at planned intervals. HCU data are entered into a cumulative electronic log by the unblinded coordinator from each site, encompassing hospitalisations (reason, location and number of days), unscheduled and scheduled outpatient visits, and filled medication prescriptions. Given the marked variability in EMR and resources across sites, procedures for undertaking the quarterly EMR queries vary: some sites have an analyst working directly with project staff to request a batch of data, whereas other sites train coordinators to query their EMR using script developed by an analyst. The quarterly EMR queries each cover a period of 4 months, such that there is always an overlap across queries, as there can be delays in data

being populated in the EMR. Completed logs (deidentified apart from the inclusion of dates) are transferred to the DCC via an encrypted data transfer method; source data are maintained on a secure server at each site.

Statistical considerations

A total of 460 children are randomised to one of the two treatment arms in a 1:1 ratio. Factors identified to possibly influence treatment response include child's age (reflecting developmental differences in neurobehaviour and potential sensitivity to oSDB), weight status (a comorbidity that may portend less effective surgical responses) and race (based on prior data indicating suboptimal surgical response of African–American compared with white children).[6] Therefore, randomisation is stratified by the following factors within site: age (<5 years vs >5 years), overweight status (BMI>85th percentile) and race (African–American vs other). Stratification provides greater assurance that the comparison groups will be similar with respect to these variables. However, given the overall sample size of 460 and the relatively large number of strata (eight strata within each of the seven sites), the expected total number of subjects within each stratum is too small (about 8) to use standard randomisation approaches such as permuted blocks. To ensure that treatment arms are balanced with respect to these factors, as well as for the number of subjects in each group, we use a dynamic randomisation method, Pocock and Simon's minimisation method.[17] Specifically, for each eligible participant, based on the value of his/her stratification factors, the participant will have a 30% chance to be allocated randomly to one of the two treatment arms and a 70% chance to be allocated to the arm that minimises the differences in number of participants across two treatment arms within each stratum deterministically. We have implemented this randomisation algorithm in our data management system (Slice).

In CHAT, greater improvements in the BRIEF score were observed in the eAT versus WWSC arm, but we could not rule out the possibility that these improvements were influenced by parental expectations. Therefore, in PATS, we elected a coprimary outcome that included one objective, performance-based neurocognitive measure (the GNG d-prime score for sustained attention) and one composite behavioural outcome (the BRIEF2/P GEC score). To maintain a study-wise significance level of 5% for analysis of coprimary endpoints, we use a sequentially rejective method, the Holm's method, which has been shown to be uniformly more powerful than the Bonferroni procedure. In the case of two tests using an overall alpha of 0.05, the comparison with the largest difference will be tested at the 0.025 level. If it is rejected, the comparison with the second difference will be tested at the 0.05 level. For the BRIEF score change from baseline to 12 months, we used prior CHAT data and assumed a relatively large 3.7-point difference in change scores between the two arms, an 11.5-point SD at baseline and a correlation between the baseline and the follow-up

measurements of 0.73. For the GNG change score, we assumed a smaller 0.33 difference between the two arms, a baseline SD of 0.77, and a correlation between baseline and follow-up measurements of 0.48.[14] Based on the experience in the CHAT study, we assume a drop-out rate of 15% at 6 months and an additional 5% attrition at 12 months. Using these estimates and methods described in Hedeker *et al*[18] for sample size estimation for longitudinal designs with attrition, we estimated a total sample size of 460 participants, resulting in 390 and 368 evaluable subjects at 6 and 12 months, respectively. Our sample size with the assumed attrition rate has 98% power to detect a difference of 3.7 points in the BRIEF 2/P GEC change score and 98% power to detect a difference of 0.33 points in the GNG change score between treatment groups at a significance level of 2.5% and 5.0%, respectively. In designing this study, we chose the sample size so that the study will have ample power for testing the primary and key secondary hypotheses and adequate power to detect moderate to large moderation effects.

Primary analyses will follow the 'intention-to-treat' principle and use a mixed effects model relating outcomes and treatment group indicators. Time (0, 6 and 12 months) will be modelled as a categorical variable to allow separate comparisons of intervention effect at 6 and 12 months. Missing data will be handled through multiple imputation[19] or inverse probability weighting.[20] Continuous secondary outcomes will be analysed in the similar fashion as the primary outcome. For endpoints related to HCU, we will consider models that account for potential data dispersion and possible preponderance of zeros (eg, zero-inflated negative binomial models). Statistical tests of treatment by covariate interaction will be performed to assess whether treatment effect varies by age, baseline weight, atopy/asthma status, secondhand smoke, socioeconomic status (SES), family functioning or race.

Safety and data monitoring

The study is monitored routinely for issues of data quality, study conduct (including recruitment and follow-up rates), data quality and adverse events (AEs). Of particular concern are attrition and cross-over rates which, if excessive, could jeopardise the integrity of the study. A special category of event, denoted as 'treatment failure' was used in the CHAT study and is also used in PATS. Treatment failures are identified using prespecified thresholds for defining changes in behaviour or health likely attributable to persistent mild oSDB, adjudicated by an independent MM. AE surveillance, adjudication and reporting follows the requirements of NHLBI and the central reliant institutional review board (IRB) at CHOP, as well as any site-specific IRB requirements. Quarterly reports addressing these issues of study conduct, data quality, AEs and treatment failures are provided to the steering committee, the DSMB and NHLBI. Given that the patient population consists of children who are otherwise healthy, with mild oSDB, and that the intervention is considered a standard clinical intervention, we do not anticipate that the interim analysis will yield efficacy data compelling enough to require early termination. Therefore, we will monitor the BRIEF2/P GEC score and GNG score, the coprimary outcomes, in planned interim analyses of efficacy. We plan to perform one interim analysis after half of the study population has completed their 12-month evaluations. Based on our recruitment projections, most of the accrual will be complete at this time, and therefore early stopping may not be relevant. To create a formal framework for assessment of interim results, the Haybittle-Peto boundary will be used[21]; that is, interim results for comparisons of the BRIEF2/P score and GNG score between treatment groups will be considered sufficient to consider early termination only if at least one of the between-group differences are statistically significant using a family-wide significance level of 0.001. The Haybittle-Peto stopping rule allows the final analysis to be evaluated at a 5% level of significance.[21 22]

PATIENT AND PUBLIC INVOLVEMENT STATEMENT

The design of this study is informed by the experience of participants and their families in the CHAT study, where study staff heard of interest in understanding the impact of snoring on quality of life and cognition. The teachers of the participating children provided standardised neurobehavioural assessments during the course of the study. The conduct of the study is overseen by the DSMB that includes representatives from the National Institutes of Health. We plan to use data from this study to disseminate information directly to patients through educational modules, blogs and an online forum available in a sleep apnoea patient portal (MyApnea.Org) that has enrolled over 17 000 patients and their family members to learn more about sleep apnoea and ongoing sleep apnoea research.

ETHICS AND DISSEMINATION

The potential consequences of deferring surgery and treating oSDB conservatively are unclear and provide the rationale for this randomised controlled trial. There is great physician and geographical differences regarding whether or not oSDB is treated surgically. In many centres, children with a normal PSG do not undergo AT, and in other centres, children who snore do not undergo PSG (precluding distinction of OSA from oSDB). All options, including refusal to be in the study in order to obtain more immediate treatment, as well as potential risks of surgery, will be discussed with the participants and their families. At the end of the trial, participants will have a final PSG, and children with persistent symptoms of SDB or new abnormalities on PSG will be referred for clinical management.

The study protocol, IRB# 14–0 11 214, was approved by the IRB at CHOP on 3 October 2014. Following NIH policies, it was decided that the CHOP IRB would be the study's single central IRB. Participating sites provided

reliance agreements allowing the CHOP IRB to act as the IRB of record for their institutions. The relying institutions remain responsible for ensuring compliance with the CHOP IRB's determinations and with the Terms of its Office of Human Research Protections, approved Federal Wide Assurance. The approval of CHOP as the central IRB of record was granted on 29 February 2016. Each clinical centre is responsible for ensuring that informed consent is obtained from each participant according to the guidelines of its IRB. Informed consent (signed and dated by the participant's parent/guardian) must be obtained prior to initiation of any study related activity.

Proposed protocol changes are presented to the steering committee to allow all members to benefit from the scientific debate generated in these discussions. Proposed changes can be implemented only after the steering committee reaches a majority vote and the NHBLI project officer approves of the proposed changes. Once a proposed change has been approved, the CCC and DCC will coordinate all activities required to implement the change via the issuance of a protocol amendment document and revised protocol. Substantive changes to the protocol require approval from the DSMB before implementation.

To maintain patient confidentiality, participants are identified to the DCC only by patient identification numbers and no personal information will be transmitted to the DCC. Furthermore, data for reports and publications will be provided in aggregate or blinded form without the identification of individual patients. At the clinical sites and participating centres, all data will be (1) kept in confidential locked files, (2) identified by participant identification number only, and (3) kept separately from identifying information used for participant tracking and follow-up contacts.

The results will be published in peer-reviewed journals and presented at academic conferences, as well as directly to patients through a web portal (MyApnea.org). The data collected from the PATS study will be deposited in a repository (National Sleep Research Resource; sleepdata.org) after completion of the study to maximise use by the scientific community.

DISCUSSION

Mild oSDB is of great clinical and public health relevance given its high prevalence and potential impacts to health and well-being of children, their families and the healthcare system. A paucity of data from randomised clinical trials has led to fundamental questions regarding the role of AT in children with mild oSDB, contributing to large geographical variations in care and potential for surgery to be both overused and underused. PATS was designed to resolve uncertainties on management approaches for paediatric mild oSDB by addressing several critical issues:(1) assess outcomes of importance to children and their families, in particular, the patient-reported outcomes of behaviour, quality of life and sleep disturbances; (2)

examine differences in treatment responses among children who are at increased risk of mild oSDB, such as preschool children, minorities and children with asthma or obesity; (3) evaluate HCU as an understudied outcome in this condition and (4) assess moderating influences of secondhand smoke, insufficient or irregular sleep, SES and family functioning. Meeting the study goals requires collection of a large variety of data from multiple sources (child, caregiver, teacher and neighbourhood geocode) and across multiple domains (neurobehaviour, PSG, actigraphy, sleep-related symptoms, quality of life, anthropometry, blood pressure, HCU, cotinine and immunoglobulin titres).

Several challenges present in the design of PATS: (1) The targeted study population include children aged 3.0–12.9 years at the time of screening. In accord with the rapid development of children in this age range, age-specific forms are available for both the BRIEF and GNG tests. Therefore, the tests given at each visit are age-dependent. During the 12 month follow-up period, children may move from one age group to another age group, resulting in different age-specific tests used at baseline and at the 12-month follow-up. However, the test scores are normalised to each age category, making them comparable across different groups. Furthermore, age has been chosen as a stratification factor to ensure balance across treatment arms within each age category. Effect modification by age will also be examined as a prespecified subgroup analysis. (2) Recall bias may present, especially when reporting behaviour problems; parents may differ in their vigilance in monitoring their children's behaviour problems or AEs in general and willingness to discuss these issues with the study coordinator, and such differences may not be balanced by treatment arm. When analysing safety data, sensitivity analyses may be needed to quantify the potential effect of such bias. (3) Double blinding is not possible in a study of surgical treatment in children. Parents and children cannot be feasibly blinded to surgery. The use of a caregiver-reported outcome is of concern in this setting as responses may reflect treatment expectations. We attempted to address this concern by including an objective test (GNG) as a coprimary outcome, as well as collecting comparable behavioural data from the child's teachers, who may be unaware of treatment. To minimise bias due to unblinded staff, we established structured communication protocols between blinded and unblinded personnel at each site. Nonetheless, unblinding may occur especially considering the study's frequent contact points between parents and study personnel (three visits and monthly phone calls). Every effort is made to prevent unblinding and any unblinding episodes are documented to facilitate the interpretation of study findings. (4) As in any clinical trial, cross-over and loss to follow-up will be inevitable despite attempts at best practice. While cross-over does not threaten the validity of the intent-to-treat primary analysis, it may dilute the treatment effect and reduce the study power. The rate of cross-over is closely monitored and its effect on study power

will be assessed. (5) HCU data are from diverse academic healthcare centres in the USA, where costs are difficult to directly assess due to the discrepancies between costs and charges. Therefore, our analyses will quantify key HCU events (eg, hospitalisations, clinic visits and medications), which will provide a proxy for costs.

In summary, PATS will provide evidence on whether children with mild oSDB benefit from surgery by randomising children to the two most common managements: AT or observation. The findings will have key implications for disease management, including the need for preoperative PSGs to distinguish oSDB from obstructive sleep apnoea, the potential to reduce practice and geographic variability in the management of oSDB and the understanding of response to surgery in African–American children and in lower socioeconomic status families, in order to optimise their management and reduce health disparities. Moreover, the design of PATS provides a model for conducting a surgical trial in children across a large age range studied with both caregiver reported and objectively measured outcomes while also assessing a wide range of other outcomes, such as HCU and potential effect modification by several host and environmental factors. Salient statistical considerations include plans for analysis of a coprimary outcome without excessive loss of power, use of a dynamic randomisation method to address multiple strata of interest in the context of a modest sample size, analysis of complementary caregiver and teacher reports, and interim safety analyses that minimally impact study power.

Author affiliations
[1] Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, Massachusetts, USA
[2] Department of Biostatistics, Harvard University T. H. Chan School of Public Health, Boston, Massachusetts, USA
[3] Division of Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
[4] Sleep Disorders Center and Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA
[5] Sleep Disorders Center and Department of Otolaryngology-Head and Neck Surgery, University of Michigan, Ann Arbor, Michigan, USA
[6] Sleep Disorders Center and Division of Pediatric Pulmonology, University of Michigan, Ann Arbor, Michigan, USA
[7] Divisions of Otolaryngology-Head and Neck Surgery and Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA
[8] Department of Otolaryngology-Head and Neck Surgery, College of Medicine, University of Cincinnati, Cincinnati, Ohio, USA
[9] Department of Otolaryngology, Head and Neck Surgery, UT Southwestern and Children's Medical Center Dallas, Dallas, Texas, USA
[10] Division of Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, Massachusetts, USA
[11] Department of Pediatrics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[12] Center for Human Phenomic Science, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[13] Department of Otolaryngology, Eastern Virginia Medical School, Norfolk, Virginia, USA
[14] Department of Pediatrics, University Hospitals Rainbow Babies and Children's Hospital, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA
[15] Division of Pulmonary Medicine, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[16] Department of Pediatrics, Abigail Wexner Research Institute at Nationwide Children's Hospital and The Ohio State University, Columbus, Ohio, USA

ORCID iD
Rui Wang http://orcid.org/0000-0001-5007-193X

REFERENCES
1 O'Brien LM, Gozal D. Behavioural and neurocognitive implications of snoring and obstructive sleep apnoea in children: facts and theory. *Paediatr Respir Rev* 2002;3:3–9.
2 Hall MJ, Schwartzman A, Zhang J, *et al*. Ambulatory surgery data from hospitals and ambulatory surgery centers: United States, 2010. *Natl Health Stat Report* 2017;102:1–15.
3 Mitchell RB, Archer SM, Ishman SL, *et al*. Clinical Practice Guideline: Tonsillectomy in Children (Update)-Executive Summary. *Otolaryngol Head Neck Surg* 2019;160:187–205.
4 Mitchell RB, Pereira KD, Friedman NR. Sleep-Disordered breathing in children: survey of current practice. *Laryngoscope* 2006;116:956–8.
5 Chervin RD, Ellenberg SS, Hou X, *et al*. Prognosis for spontaneous resolution of OSA in children. *Chest* 2015;148:1204–13.
6 Marcus CL, Moore RH, Rosen CL, *et al*. A randomized trial of adenotonsillectomy for childhood sleep apnea. *N Engl J Med* 2013;368:2366–76.
7 Chervin RD, Ruzicka DL, Giordani BJ, *et al*. Sleep-Disordered breathing, behavior, and cognition in children before and after adenotonsillectomy. *Pediatrics* 2006;117:e769–78.
8 Emancipator JL, Storfer-Isser A, Taylor HG, *et al*. Variation of cognition and achievement with sleep-disordered breathing

in full-term and preterm children. *Arch Pediatr Adolesc Med* 2006;160:203–10.

9 Boss EF, Marsteller JA, Simon AE. Outpatient tonsillectomy in children: demographic and geographic variation in the United States, 2006. *J Pediatr* 2012;160:814–9.

10 Rosen CL, Larkin EK, Kirchner HL, *et al*. Prevalence and risk factors for sleep-disordered breathing in 8- to 11-year-old children: association with race and prematurity. *J Pediatr* 2003;142:383–9.

11 Chan A-W, Tetzlaff JM, Altman DG, *et al*. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.

12 Gioia GA IP, Guy SC, Kenworthy L. In: Lutz FL, ed. *BRIEF2: behavior rating inventory of executive function*. 2nd edn, 2015.

13 Isquith PK, Crawford JS, Espy KA, *et al*. Assessment of executive function in preschool-aged children. *Ment Retard Dev Disabil Res Rev* 2005;11:209–15.

14 Orchinik LJ, Taylor HG, Espy KA, *et al*. Cognitive outcomes for extremely preterm/extremely low birth weight children in kindergarten. *J Int Neuropsychol Soc* 2011;17:1067–79.

15 Friedman M, Hwang MS. Brodsky and Friedman scales and clinical tonsil size grading in children. *JAMA Otolaryngol Head Neck Surg* 2015;141:1–8.

16 Redline S, Amin R, Beebe D, *et al*. The childhood adenotonsillectomy trial (CHAT): rationale, design, and challenges of a randomized controlled trial evaluating a standard surgical procedure in a pediatric population. *Sleep* 2011;34:1509–17.

17 Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 1975;31:12.

18 Hedeker D, Gibbons RD, Waternaux C. Sample size estimation for longitudinal designs with attrition: comparing time-related contrasts between two groups. *Journal of Educational and Behavioral Statistics* 1999;24:70–93.

19 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–99.

20 Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc* 1995;90:106–21.

21 Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. *Br J Radiol* 1971;44:793–7.

22 Peto R, Pike MC, Armitage P, *et al*. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;34:585–612.

23 Achenbach TM RL. *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University ofVermont Research Center for Children, Youth, & Families, 2001.

24 C C. *Conners (Manual)*. 3rd edn. San Antonio: Pearson, 2009.

25 Chervin RD, Hedger K, Dillon JE, *et al*. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep Med* 2000;1:21–32.

26 Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness scale. *Sleep* 1991;14:540–5.

27 Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the pediatric quality of life inventory version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001;39:800–12.