

CON: Thoughtful Steps Informed by More Comparative Effectiveness Research Is Needed in Home Testing

Sairam Parthasarathy, M.D., F.A.A.S.M.

Department of Medicine of the University of Arizona, Tucson, AZ; Arizona Respiratory Center of the University of Arizona, Tucson, AZ

Sir William Osler famously stated, “One special advantage of the skeptical attitude of mind is that a man is never vexed to find that after all he has been in the wrong.”¹ This statement may serve as a disclaimer when we discuss the pitfalls of portable (or out-of-center) testing for sleep apnea, but may also serve to heighten our sense of caution as we go down this path of disease management strategy. Admittedly, newer strategies are needed to reduce healthcare costs, but we cannot compromise on quality of care because, as providers, we—and not other stakeholders (third party payors, industry, or durable medical equipment companies)—are ultimately responsible for patients’ outcomes.² Stakeholder consensus can change the direction of health care policymaking and deviate far from accomplishing the originally intended objectives. Such policymaking is effected by “herding” practitioners and changing practice patterns through rewards and penalties.³ It may be superficial, yet true, to characterize this process as an “informed and real-world laboratory experiment” and thereby making this a less than perfect science. Unfortunately, the assessment of the impact of such new policies on both their intended effects and unintended consequences tends to lag by many years, and sometimes the full extent of the impact may never be ascertained. During the early period of policy implementation, practitioners are left to fend for themselves while ensuring adequate and appropriate care for their patients. This con debate is a mere outline of some of the concerns surrounding portable monitoring that would inform the practitioner and help with his or her clinical decision making. In order to do so, we need to address certain definition of terms that will be used, critically evaluate the scientific data, address areas of uncertainty, and propose future directions.

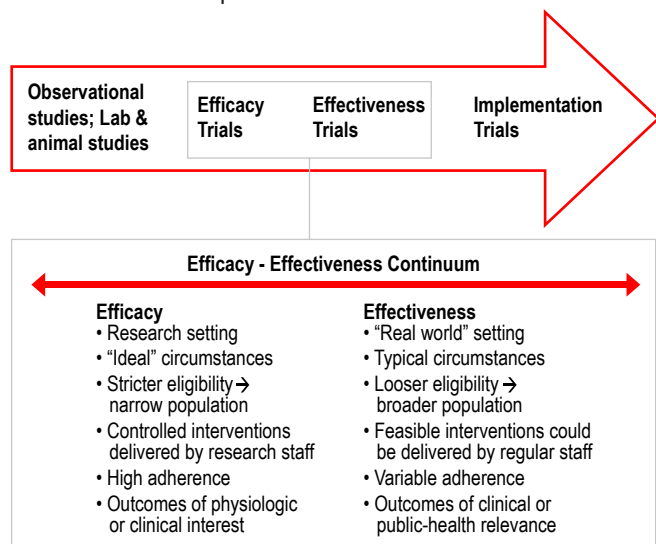
Efficacy of a diagnostic-treatment approach for a medical condition would be defined as the ability of an intervention to produce the desired beneficial effect in expert hands and under ideal circumstances. *Effectiveness* would be the ability of an intervention to produce the desired beneficial effect in the usual or “real-world” setting. Efficacy of an intervention is informed by randomized clinical trials (RCTs) that are designed to show whether an intervention produces the desired clinical outcome under optimal conditions, which includes a narrow population, delivered by research staff, with high treatment adherence. Effectiveness studies involve real-world clinicians and settings with loose eligibility criteria and variable treatment adherence (see **Figure 1**). The natural progression of clinical research proceeds from RCTs that inform us of the efficacy of an interven-

tion before such an intervention can be tested (or translated) into clinical effectiveness studies in the translation continuum (**Figure 1**). Translating practices that are found to be efficacious into routine practice settings to produce effective results helps develop evidence-based practice. While there have been well-designed RCTs that have assessed portable monitoring in combination with treatment, there are currently no clinical effectiveness studies.⁴⁻⁸ Moreover, we know that testing alone does not help improve patient outcomes. Patient outcomes are only influenced by testing when such testing is combined with an effective intervention.⁹ The effectiveness of the portable monitoring strategy would therefore be dependent on the implementation and performance of the positive airway pressure auto-titration device to determine treatment pressure outside the sleep laboratory.

Technology Assessment

Technology assessment is the cornerstone when we attempt to change practices in the out-of-center management of sleep apnea. While the technology of controlled trials may yield positive results, in the real world, the implementation of such technology may suffer from significant performance variations. Recently, Collop and colleagues performed the arduous task of technology assessment of portable devices against well-reasoned expectations for test characteristics. The benchmarks included a likelihood ratio for a positive test > 5 and sensitivity of 0.825 for detecting obstructive sleep apnea (defined as an apnea-hypopnea index ≥ 5 per hour of sleep), with the assumptions that the pre-test probability in the population was 0.85 and at least 66% of the population be diagnosed accurately as positive (**Figure 2**).¹⁰ In their methodical assessment of the devices, only 9 of 27 devices exceeded the desired LR+ve (likelihood ratio of a positive test) value > 5 ; and only 7 of these devices performed adequately on a consistent basis if a sensitivity of 0.825 was also taken into account.¹⁰ This is rather concerning when one realizes that such testing was conducted under controlled conditions by experienced research staff and that their real-world performance outside of the research setting is unknown. Moreover, if the assumptions were to change, namely, if the pre-test probability were to drop to 0.5, then the LR+ve would need to be greater than 20. In such a scenario, only 4 of the 21 devices surveyed would meet the desired test characteristics. At a minimum, such analyses prudently suggest that the provider needs to be very familiar with the device

Figure 1—Continuum of translation of research from observational to implementation studies

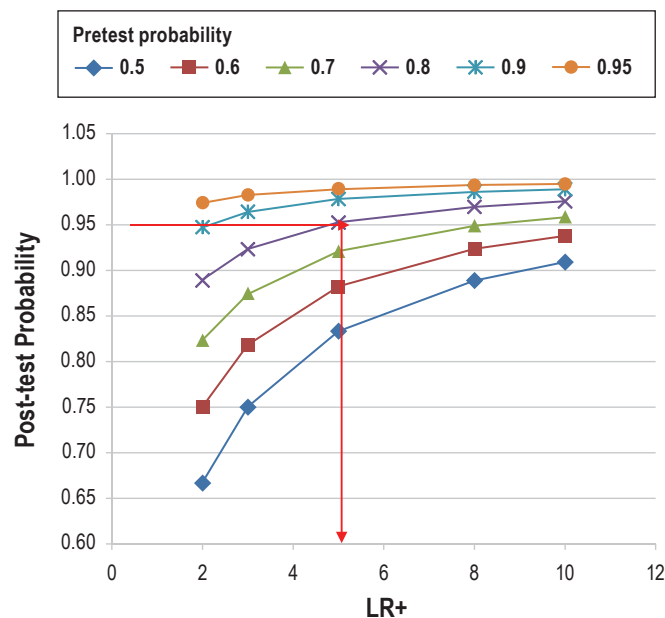


In the inset the characteristics that distinguish efficacy and effectiveness trials are shown.

and their own practice setting. While patients referred to sleep consultation may have high pre-test characteristics, currently evolving practice strategies of screening certain special populations—such as patients in a heart failure clinic or hypertension clinic—may be associated with reductions in pre-test probability when compared to a sleep clinic population with patients referred for suspected sleep apnea. Suffice to say, we have a lot to learn about device performance, and this is compounded by the constant introduction of newer devices and modifications to existing technologies.¹¹

An interesting aspect pertaining to the debate surrounding portable monitoring is that patient outcomes are dependent on the implementation and performance of the positive airway pressure auto-titration device used to determine treatment pressure outside the sleep laboratory. While such technology may work in expert hands during a clinical trial, widespread implementation requires (a) improved understanding of the devices by the providers, and (b) mechanisms for compensating the time and effort of such care delivery. With regards to the former, there is evidence that knowledge of autoPAP devices may not be uniform with regards to known contraindications in certain special populations (heart failure), limitations (nasal obstruction), and variability in device performance between manufacturers.¹²⁻¹⁴ Devices of various manufacturers can perform differently in response to the degree of air leak (due to poor mask fit or mouth opening), and this, in turn, can deleteriously impact the performance of such autoPAP devices and may lead to reduced adherence to PAP therapy.^{12,14,15} Mask fitting is usually performed in the sleep laboratory by an attendant sleep technician who can assess the performance of the mask interface during various body positions and sleep stages in a given patient. However, such an opportunity to try various masks during the course of the night is lost when one fits masks in a clinic setting and the nighttime performance and seal of such devices are unknown.

Figure 2—Relationship between the likelihood ratio of a positive test (x axis), pre-test probability (various colored lines), and post-test probability (y axis)



The red horizontal and vertical lines indicate that for a pre-test probability of 0.8 (purple line), and a desired post-test probability of 0.95, the portable monitoring device must have a LR+ ≥ 5 . Conversely, if the pre-test probability were 0.5 (blue line), and the desired post-test probability remained 0.95, then the LR+ needs to be > 20 (cannot be plotted in this graph). Reproduced with permission from Collop et al.¹⁰

Many autoPAP devices do provide estimations of degree of leak (as time spent in large leak or 90th percentile leak levels), but such information is available only upon seeing the patient back after, for instance, a 2-week trial period; and during that time a potential opportunity for early correction and adherence promotion may be lost. Additionally, there are financial limitations with regard to the number of masks that can be attempted for a given patient in an outpatient setting, which may make the cost of multiple mask exchange simply prohibitive. This is compounded by ongoing efforts to reduce mask exchange benefits by third party payors as a cost-saving strategy.

There are other issues that influence the generalizability of efficacy studies. Patients in clinical trials are carefully selected as opposed to the real-world population (**Table 1** provides categories of patients who were excluded). Moreover, the currently available efficacy studies are unblinded and are performed in an expert and resource-rich environment, which may not be immediately transferrable to the real-world; and even if this were feasible, may not be as effective. Additionally, data loss from monitoring devices and autoPAP devices are more easily corrected from a logistical standpoint with research staff manpower than that available in day-to-day practice.

A Call for Comparative-Effectiveness Research

The main thrust of this con article is to state that the area of portable monitoring and home-based treatment of sleep apnea direly needs comparative effectiveness research (CER) with cost-efficacy and tangible patient outcomes measurements be-

fore widespread institution into real-world practice. In the absence of such CER, currently, there are few modeling studies that are not favorable toward portable monitoring and home-based treatment initiation, and that requires careful attention.¹⁶⁻¹⁸

Pietzche and colleagues recently published their Markov model-based study which used a hypothetical model applied to 50-year-old men with a 50% chance of having moderate-severe sleep apnea.¹⁷ They performed 10-year and lifetime incremental cost-effectiveness ratio projections for laboratory based studies and portable monitoring.¹⁷ They factored rates and cost of cardiovascular events and motor vehicle accidents, health-related quality of life, CPAP adherence, and failure of portable monitoring or therapy. They found that full-night polysomnography in conjunction with CPAP therapy was the most cost-efficient strategy when compared to split-night or portable monitoring studies. In another study, Ayas and colleagues undertook a theoretical decision model to assess the pre-test probability above which it would be appropriate to use portable studies to “rule-in” disease in symptomatic patients with suspected sleep apnea.¹⁶ They factored variables such as failure to improve with CPAP therapy, CPAP intolerance, technical failure rate of portable monitoring, and sensitivity and specificity of portable monitoring devices for diagnosing sleep apnea. They performed sensitivity analysis by varying CPAP adherence and healthcare costs, and concluded that a pre-test probability above which portable sleep study was economical was greater than 0.68. Similarly, Masa and colleagues performed a study to determine the agreement between laboratory-based sleep study and portable monitoring for clinical decision-making in a sample of 366 participants.¹⁸ Therapeutic decisions (CPAP, no CPAP, or impossible decision) were based upon portable monitoring study OR laboratory sleep study, AND auxiliary clinical set of variables. After study subjects underwent diagnostic procedures (sleep study or portable monitoring) in a random manner, the same clinical decision-making procedure was repeated. The sensitivity and specificity of the therapeutic decision-making was modest at 73% and 77%, respectively, with an agreement level of 76% (sum of true positives and true negatives). They concluded that portable-monitoring based therapeutic decision making was *adequate* when AHI was high, but deficient in large population of patients with mild to moderate AHI. These studies would suggest that both the pre-test probability and the severity of sleep apnea can heavily influence the outcomes of a diagnostic strategy.

Portable monitoring affords an *adequate* level of management while potentially incurring higher lifetime costs! This—from a health policy standpoint—is quite concerning. If the afore-mentioned assumptions and modeling are correct, we are trading a superior-quality diagnostic test (in-laboratory sleep study) for an inferior quality test (portable monitoring) for *perceived* short-term gains in cost, when in fact, in the long-term, such a strategy may be more expensive. This is puzzling because in other fields of medicine, diagnostic tests have been constantly traded in for superior tests and not the reverse, as there are expenses and harms incurred by false-negative and false-positive testing—such as in pulmonary embolism.¹⁹ What is perhaps unique to management of sleep apnea is that the CPAP therapy is widely accepted as essentially harmless, and therefore we can live with false-positive testing by portable monitors and conceivably risk treating individuals with

Table 1—Exclusion criteria form efficacy studies of portable monitoring

- Neuromuscular disease
- Moderate to severe respiratory disease (breathlessness OR SpO₂ < 90%)
- Recent and unstable cardiovascular disease
- Individuals requiring bilevel PAP or O₂
- Clinically unstable chronic medical condition
- Nasal congestion and blockage
- Heart failure (moderate to severe)
- Morbid obesity
- O₂ use prior to OCT test; Hypercapnia
- Use of narcotics
- Restless legs
- Alpha-blockers

false-positive results. In contrast, in the management of pulmonary embolism, therapies such as systemic anticoagulation, vena caval filters, or thrombolytics have significant attendant risks; therefore, the diagnostic tools have evolved from less accurate (ventilation-perfusion lung scans) to more accurate (CT pulmonary angiography) tests. The driving factors for such change have traditionally been to improve quality and save lives, although cost-effectiveness studies ensued to help validate the shifting management paradigms. The question that can be legitimately posed at this juncture is, “Can clinical decision-making mandated by third-party payors requiring portable monitoring in all cases (even in mild sleep apnea) be considered as the wrong approach?” While many payors acknowledge the guideline-based exceptions to this approach for patients with significant comorbidities (morbid obesity, heart failure, and severe respiratory disease) or other suspected sleep disorders, the approach to a patient with suspected sleep apnea without such comorbidities or conditions needs to be rethought.

There are other intangibles that are not factored here. There are program development costs, such as the need to purchase new portable monitoring units, autoPAP devices, maintenance of such equipment, compatibility issues with existing systems, changing respiratory scoring rules and guidelines, and mechanisms for clinic-based therapy initiation and mask refitting. Such costs are often not factored into cost-effectiveness analyses. Moreover, there are other advantages to lab-based sleep studies that may be missed by portable monitoring—coexisting sleep disorders (periodic leg movement disorder, REM behavior sleep disorder, etc.) and other conditions (EEG abnormalities such as alpha-delta sleep or inter-ictal foci). Also, in patients with mild positional sleep apnea, an in-laboratory study may allow the technician to intervene during the course of the sleep study and request the patient to lie supine so as to increase the chance for performing a conclusive test and reduce the number of repeat tests. Collectively, these intangibles can add up to having an impact on the management of our patients.

One recognizes that patients may prefer portable monitoring in the home and that this dovetails nicely with the current mantra of the medical home concept. However, until home-based por-

table monitoring technology can advance to the point of providing similar if not superior testing to the laboratory setting, a sea change in that direction may be unwise. Only time is the arbiter of all truths. Until then, it is wise to consider this con portion of the debate as a minority report that demands more comparative effectiveness research. It would be apropos to end by quoting Sir Osler, who stated, "...we doctors [providers] have always been simple trusting folk! To have the placid faith of the simple believer, instead of the fighting faith of the aggressive doubter, has ever been our besetting sin in the matter of treatment."¹

CITATION

Parthasarathy S. Con: Thoughtful steps informed by more comparative effectiveness research is needed in home testing. *J Clin Sleep Med* 2013;9(1):9-12.

REFERENCES

- Osler W. An Address on the treatment of disease: Being the address in medicine before the Ontario Medical Association, Toronto, June 3rd, 1909. *Br Med J* 1909;2:185-9.
- Crossing the Quality Chasm: A New Health System for the 21st Century. National Academies Press, 2001.
- <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html>. Last accessed October 20, 2012.
- Antic NA, Buchan C, Esterman A, et al. A randomized controlled trial of nurse-led care for symptomatic moderate-severe obstructive sleep apnea. *Am J Respir Crit Care Med* 2009;179:501-8.
- Kuna ST, Gurubhagavatula I, Maislin G, et al. Noninferiority of functional outcome in ambulatory management of obstructive sleep apnea. *Am J Respir Crit Care Med* 2011;183:1238-44.
- Rosen CL, Auckley D, Benca R, et al. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. *Sleep* 2012;35:757-67.
- Berry RB, Hill G, Thompson L, et al. Portable monitoring and autotitration versus polysomnography for the diagnosis and treatment of sleep apnea. *Sleep* 2008;31:1423-31.
- Masa JF, Jimenez A, Duran J, et al. Alternative methods of titrating continuous positive airway pressure: a large multicenter study. *Am J Respir Crit Care Med* 2004;170:1218-24.
- Hulley SB, Cummings SR, Browner WS, et al. *Designing clinical research: an epidemiologic approach*, 2nd ed. Lippincott Williams & Wilkins, 2001.
- Collop NA, Tracy SL, Kapur V, et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. *J Clin Sleep Med* 2011;7:531-48.
- Garber AM. Modernizing device regulation. *N Engl J Med* 2010;362:1161-3.
- Coller D, Stanley D, Parthasarathy S. Effect of air leak on the performance of auto-PAP devices: a bench study. *Sleep Breath* 2005;9:167-75.
- Parthasarathy S, Habib M, Quan SF. How are automatic positive airway pressure and related devices prescribed by sleep physicians? A web-based survey. *J Clin Sleep Med* 2005;1:27-34.
- Farre R, Montserrat JM, Rigau J, et al. Response of automatic continuous positive airway pressure devices to different sleep breathing patterns: a bench study. *Am J Respir Crit Care Med* 2002;166:469-73.
- Valentin A, Subramanian S, Quan SF, et al. Air leak is associated with poor adherence to autoPAP therapy. *Sleep* 2011;34:801-6.
- Ayas NT, Fox J, Epstein L, et al. Initial use of portable monitoring versus polysomnography to confirm obstructive sleep apnea in symptomatic patients: an economic decision model. *Sleep Med* 2010;11:320-4.
- Pietzsch JB, Garner A, Cipriano LE, et al. An integrated health-economic analysis of diagnostic and therapeutic strategies in the treatment of moderate-to-severe obstructive sleep apnea. *Sleep* 2011;34:695-709.
- Masa JF, Corral J, Pereira R, et al. Therapeutic decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a large multicentric study. *Am J Respir Crit Care Med* 2011;184:964-71.
- Patel S, Kazerooni EA. Helical CT for the evaluation of acute pulmonary embolism. *AJR Am J Roentgenol* 2005;185:135-49.

ACKNOWLEDGMENTS

Dr. Parthasarathy received funding from the NIH/NHLBI (R01 HL 095748) during the time of preparation of this manuscript.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication November, 2012

Accepted for publication November, 2012

Address correspondence to: Sairam Parthasarathy, M.D., Associate Professor of Medicine, University of Arizona, 1501, North Campbell Ave., AHSC, Rm 2342D, Tucson, AZ 85724; Tel: (520) 626-6109; Fax: (520) 626-6970; E-mail: sparthamd@yahoo.com

DISCLOSURE STATEMENT

Dr. Parthasarathy's institution has received grant support from Philips-Respironics Corp.