

The Face of Sleepiness: Improvement in Appearance after Treatment of Sleep Apnea

Ronald D. Chervin, M.D., M.S., F.A.A.S.M.¹; Deborah L. Ruzicka, R.N., Ph.D.¹; Arshia Vahabzadeh, B.S., CFPH¹; Margaret C. Burns¹; Joseph W. Burns, Ph.D.²; Steven R. Buchman, M.D.³

¹Sleep Disorders Center and Department of Neurology, University of Michigan, Ann Arbor, MI; ²Michigan Tech Research Institute, Michigan Technological University, Ann Arbor, MI; ³Craniofacial Anomalies Program and Division of Plastic Surgery, Department of Surgery, University of Michigan, Ann Arbor, MI

Study Objectives: Anecdote but no formal evidence suggests that facial appearance improves after hypersomnolent patients with obstructive sleep apnea are treated. We investigated whether masked volunteer raters can identify post- rather than pre-treatment images as looking more alert, and whether impressions are predicted by any objective changes on highly precise 3-dimensional digital photogrammetry.

Methods: Participants included 20 adults with obstructive sleep apnea on polysomnography and excessive sleepiness on Epworth Sleepiness Scales. Photogrammetry was performed before and after ≥ 2 months of adherent use of positive airway pressure. Twenty-two raters then assessed pre- and post-treatment facial images, paired side-by-side in random order.

Results: Subjects included 14 men and 6 women, with mean age 45 ± 11 (SD) years and mean baseline apnea/hypopnea index of 26 ± 21 . The 22 raters twice as often identified post-treatment rather than pre-treatment images to look more alert ($p = 0.0053$), more youthful ($p = 0.026$), more attractive

($p = 0.0068$), and more likely to reflect the treated state ($p = 0.015$). Photogrammetry documented post-treatment decreases in forehead surface volume and decreased infraorbital and cheek redness, but no narrowing of the interpalpebral fissure. Decreased deep NREM sleep at baseline, and pre- to post-treatment decrements in facial redness showed promise as predictors of improved subjective ratings for alertness.

Conclusions: Patients with obstructive sleep apnea are perceived to appear more alert, more youthful, and more attractive after adherent use of positive airway pressure. Objective changes in facial surface volume and color were identified. Post-treatment decrements in redness may inform subjective impressions of improved alertness.

Keywords: Obstructive sleep apnea, sleepiness, photography, face

Citation: Chervin RD; Ruzicka DL; Vahabzadeh A; Burns MC; Burns JW; Buchman SR. The face of sleepiness: improvement in appearance after treatment of sleep apnea. *J Clin Sleep Med* 2013;9(9):845-852.

The main symptom of nocturnal sleep disorders is often excessive daytime sleepiness. Although sleepiness can be a major contributor to decreased quality of life^{1,2} and mortality,³ patients frequently grow habituated to their sleepiness, ignoring it as well as any underlying disorder. One scarcely explored potential consequence that could motivate patients to address sleep disorders is a sleepy facial appearance. For example, anecdotal, cultural, and internet lore suggests that sleepiness causes dark, “baggy” or puffy circles under the eyes, and plastic surgeons regularly treat this condition without exploring possible causes.⁴ A recent study did show that on average, persons asked to rate facial photographs taken with and without immediately preceding experimental sleep deprivation could distinguish the former as looking more tired, less healthy, and less attractive.⁵ However, no peer-reviewed literature, to our knowledge, has explored whether sleep disorders may cause dark circles under the eyes, ptosis, or other specific changes in facial appearance. No published literature has examined the extent to which treatment for sleep disorders might diminish such facial changes, cause a subjective improvement in appearance, and perhaps motivate patients to treat their sleep disorders.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Clinicians have noted anecdotally that excessively sleepy patients with obstructive sleep apnea sometimes appear more alert and more youthful after they have been adherent to treatment with continuous positive airway pressure, but neither subjective nor objective changes in physical appearance have been studied in a formal manner previously. This study combined visual impressions of medical and lay raters with novel, precise, 3-dimensional digital photographic methods to determine whether improvement in alertness with treatment for sleep apnea can be demonstrated, and whether objective facial changes may explain the subjective ratings.

Study Impact: Both medical and lay raters identified post- vs. pre-treatment facial images as looking more alert, more youthful, and more attractive, and 3-dimensional digital photogrammetry documented objective volume and color changes, some of which showed promise as predictors of raters’ subjective ratings. Knowledge that facial appearance is likely to improve with use of positive airway pressure may offer clinicians new strategies to enhance adherence.

To address these questions in a preliminary manner, we used advanced photographic technology, specifically three-dimensional digital photogrammetry, which has been developed mainly to assess subtle changes in surface morphology after

plastic surgery.⁶ This technique has a high degree of precision and accuracy, though methodologic precautions must be followed to maximize image quality.⁷ We recruited subjectively sleepy adult patients who were diagnosed with obstructive sleep apnea (OSA).⁸ This was an ideal initial disease model because subjects could be tested before and after alleviation of severe sleepiness by home use of positive airway pressure. We tested the hypotheses that treatment of OSA leads to changes in (1) infraorbital, forehead, or cheek surface volumes, (2) skin color over the same regions, (3) size of the interpalpebral fissure, and (4) subjective appearance (level of alertness) as judged by independent human raters. In the absence of relevant peer-reviewed literature, the objective outcomes were selected based on common cultural assumptions and desire to include areas that do not show change in addition to those that do.

METHODS

Subjects

Subjects were recruited from the sleep disorders clinics at the University of Michigan Sleep Disorders Center. Subjects signed written informed consent prior to all study procedures, and those whose pictures are shown in this report kindly provided consent for this as well. This study was approved by the University of Michigan Medical School Institutional Review Board and registered with ClinicalTrials.gov (NCT01060436). Inclusion criteria were: (1) age 18-65 years, because facial changes and approaches to OSA treatment may differ for children and older adults; (2) presence of subjective excessive daytime sleepiness, as determined by an Epworth Sleepiness Scale⁹ score ≥ 10 ; (3) new diagnosis of OSA, as determined by the referring clinician and documented by a recent nocturnal polysomnogram showing ≥ 5 apneic events (apneas, hypopneas, or respiratory event-related arousals)/h of sleep, using American Academy of Sleep Medicine 2007 standards for scoring sleep and events¹⁰; and (4) plan in place to treat OSA by laboratory titration and then use of continuous positive airway pressure (CPAP) or bilevel positive airway pressure (bilevel PAP) at home. Exclusion criteria included comorbid sleep, medical, or psychiatric disorders or medications that could complicate treatment outcomes and data analysis by amelioration or exacerbation of sleepiness or OSA. Subjects with a prior history of facial plastic surgery or Botox injections, or who anticipated such treatments before completion of their final imaging session, were also excluded. Participants were compensated with a \$50 Amazon.com gift certificate after each of 2 imaging sessions.

Protocol

The protocol commenced with facial imaging and completion of Epworth scales on or before the night that a subject came to the sleep laboratory for a clinically indicated CPAP or bilevel PAP titration study. Two to four months after the patient had been using CPAP or BiPAP at home, adherence and good response were verified either in person or by phone. Successful treatment was judged on the basis of reported PAP use on a nightly or near-nightly basis and confirmation by electronic download of information from the PAP unit; reported absence of snoring, nasal congestion, discomfort, mask displacement

during the night, mouth leak, or mask leak; and reported improvement in daytime sleepiness (though Epworth scores were not required to improve). When necessary, efforts were made to address barriers to adherence. Adherent subjects who were treated successfully with CPAP for 2 months or longer were invited to complete the protocol by returning for their second facial imaging appointment, at which time they also completed another Epworth scale and were screened for any new medical or health developments, aside from treatment of OSA, that could have had impact on their level of daytime sleepiness.

Polysomnography

Sleep studies were performed in the Michael S. Aldrich Sleep Disorders Laboratory, a component of the University of Michigan Sleep Disorders Center. The Center is accredited by the American Academy of Sleep Medicine and all procedures followed AASM 2007 Scoring Manual recommendations.¹⁰ Hypopneas were scored by the alternative rule 4b.

Photogrammetry

Facial images were captured using a 2-pod, 8-camera 3dMD Face photographic system sold commercially for clinical and research use by 3dMD (Atlanta, GA, www.3dMD.com). This stereophotogrammetry system has been well validated for precision, error, and repeatability,¹¹ and more recently reviewed.¹² The instrument was calibrated at each use. Cameras were mounted in a fixed position, the same chair was used each time, the same camera-to-subject distance was maintained across sessions, and floor markings and set wall distances were used and checked before each imaging session to optimize consistency of pictures across sessions and participants. Indoor lighting and temperature remained consistent across sessions. Pre- and post-PAP images for any given subject were acquired at approximately the same time of day (within about 2 h). Prior to image acquisition, subjects were given towels, washcloths, and soap and were asked to remove all make-up. They sat in an adjustable armchair at a fixed distance from a tripod that supported 8 cameras, 4 in each pod. Subjects were instructed to sit erect and assume a neutral, relaxed facial pose while looking at a static target anchored at eye level above the camera system. Facial image measurements, volumetric, and area calculations were performed with 3dMD Vultus software. Images used for analysis were high quality with good resolution. Registration error between pre- and post-PAP images was generally about one-third the manufacturer-recommended limit, confirming that subjects had not moved during image acquisition and that positioning was essentially identical at the 2 sessions. Templates were devised to identify facial areas of interest, and modified slightly as necessary to fit individual features. The same templates were used for any given subject's pre- and post-PAP images.

Quantitative analyses focused on volume of specific facial areas, color, and size of the interpalpebral fissure. Volume calculations for 2 regions of interest—just beneath the lower eyelids and on the forehead (**Figure 1**)—were computed as differences between pre- and post-CPAP images (post minus pre). The manufacturer claims valid resolution of surface changes, either extrusion or intrusion, as small as 0.5 mm. For a target area, the software calculates post-minus-pre-treatment increases in volume for those pixels where volume has increased, and

post minus pre-treatment decreases in volume for those pixels where volume has decreased. To standardize measurements across individuals and allow aggregate analyses, we subtracted the decrease from the increase to calculate the entire regional change (cm^3), and divided this result by the surface area (cm^2) of the subject's template to produce the final metric, change in cm^3 per cm^2 of facial surface.

For color analysis, 2 facial regions were chosen, just below the lower eyelids and at the cheeks (**Figure 1**). Color captured in the digital images is not analyzed by 3dMD software. To compute the color changes, a JPEG image was exported from the 3dMD software. A computer code was written in MatLab (MathWorks, Natick, MA) to extract the RGB (red, green, blue) data from the designated region of the image. As color and intensity information are correlated in the RGB color space, the RGB data were converted to the orthogonal YCbCr color space. The YCbCr color space represents luminance (Y) as a weighted sum of the RGB data, and chrominance (blueness Cb and redness Cr) by subtracting the luminance from the B and R channels. As the directly computed luminance and chrominance separate the intensity and color information, the decomposition is a common choice for skin detection in automated image analysis where illumination-independent measures of color are desired.¹³ Color changes in this report were derived from the change in Cb and Cr.

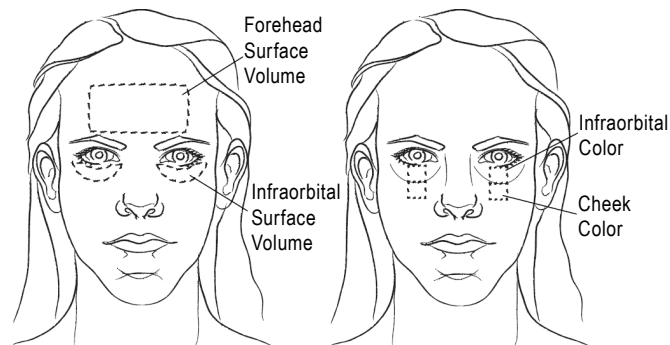
The size of the interpalpebral fissure—essentially a continuous measure to reflect any tendency toward ptosis—was assessed by measuring the maximal vertical interpalpebral distance, and dividing this by the intercanthal distance for the same eye. This ratio was computed and used in all analyses to provide automatic adjustment for any subtle change, between pre- and post-PAP images, in the distance between the camera and the subject.

Frontal images were rated by 22 volunteers masked to pre- vs. post-CPAP status. Twenty pairs of images, one pair for each subject, were presented to each rater using a Microsoft PowerPoint program. A given pair included a pre- and post-CPAP image, in randomized order, of the same subject. Each rater was asked to determine, for each pair of photographs, which face appeared most alert, most youthful, and most attractive, and to determine which was the post-CPAP image. The rater was free to choose one image as reflecting one attribute while choosing another image as reflecting another attribute. Each rater was given a \$5 gift certificate for every full group of 5 paired images for which the designations of pre- and post-PAP images were correct. Thus, any rater could earn between \$0 and \$20 dollars, provided to ensure similar motivation between raters to select carefully. Raters included 10 adult members of the local lay community, 4 men and 6 women, with two < 30 years old, three 30-39, two 40-49, three 50-59, and none 60-69. Raters also included 12 sleep or plastic surgery faculty, residents, medical students, or medical staff, with 7 men and 5 women in this group, three < 30 years old, five 30-39, one 40-49, two 50-59, and one 60-69.

Analysis

Subjective ratings for alertness were considered the primary outcome, whereas other subjective ratings (for youthfulness, attractiveness, and the “treated” condition) as well

Figure 1—Regions of the face over which volume (left) and color (right) were assessed

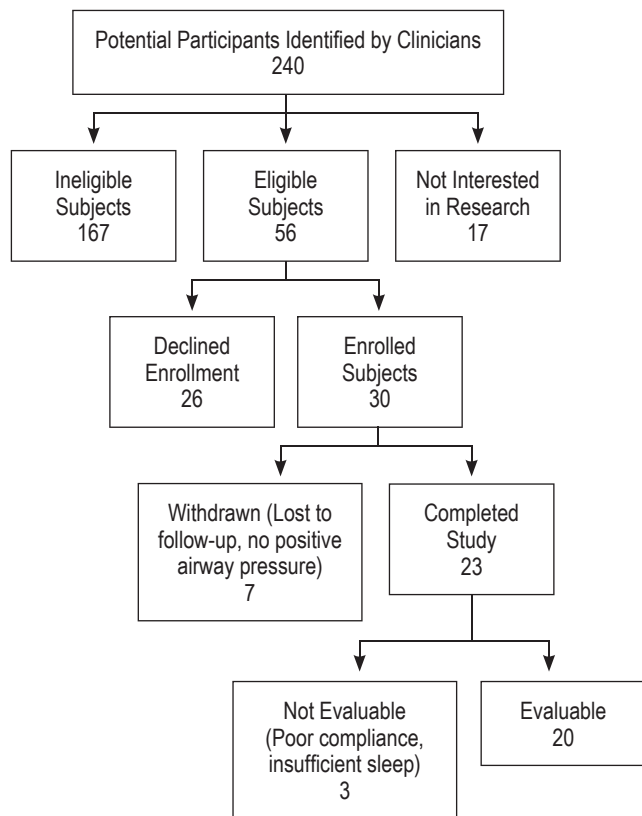


as the objective measures of facial change were considered secondary outcomes. The percentage of raters who identified the post-CPAP image as being more alert was calculated for each subject, and 50 was subtracted, so that 0 would indicate no consistently perceived change. A signed rank test was then used to determine whether the zero-scaled percentage of “correct” ratings differed from zero to an extent that was statistically improbable ($p < 0.05$). The same procedure was used to assess results for subjective ratings of youthfulness, attractiveness, and post-CPAP status. Post-CPAP changes in the 3 types of continuous outcome measures (surface tissue volume, skin color, and interpalpebral fissure size) also were assessed using signed rank tests. After analysis of the primary variable (subjective alert facial appearance), secondary analyses were not corrected for multiple comparisons because identification of potential associations was of higher interest, in this early-stage study, than was identification of pre- and post-CPAP differences that could be labeled as definitive. Finally, 3 types of variables were tested with Spearman correlation coefficients to assess which, if any, might predict subjective rater-determined improvement in appearance of alertness: these 3 types of variables were (1) objectively assessed pre- to post-CPAP changes in facial volume, color, and interpalpebral fissure size; (2) objective measures of sleep apnea severity and sleep from the baseline diagnostic polysomnograms; and (3) pre- to post-CPAP changes in subjective sleepiness ratings on the Epworth Sleepiness Scale. All analyses that involved subjective ratings or measures of color were repeated among the subset of subjects who were Caucasian. This was not attempted for other races because the sample size was insufficient. No previous data on subjective facial ratings or photogrammetry, among OSA patients, were available to inform the sample size for this novel study. A total sample size of 20 subjects was targeted. All analyses were performed with SAS, version 9.2 (SAS Institute Inc., Cary, NC).

RESULTS

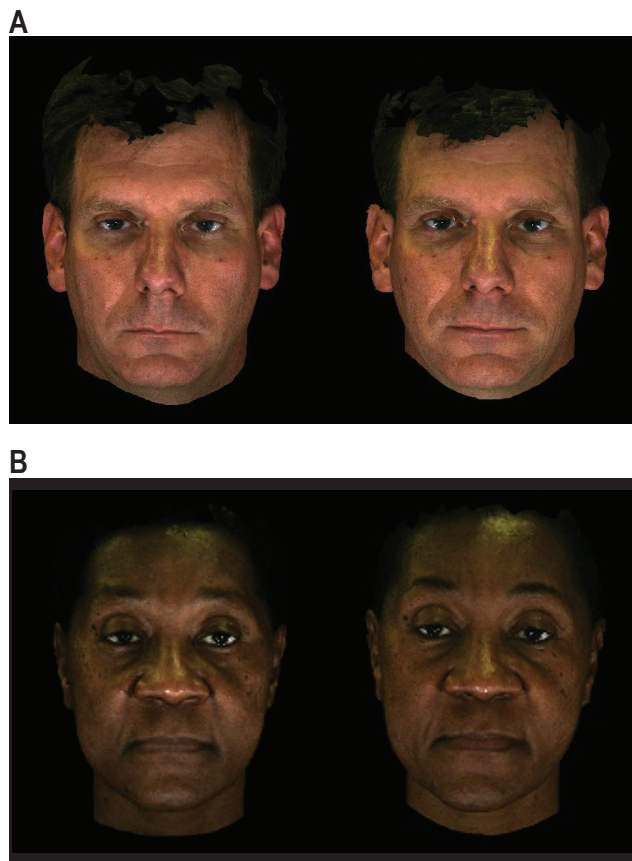
Subjects

Participants ($n = 20$) included 14 men and 6 women. The mean age (\pm SD) was 45 ± 11 years and the mean body mass

Figure 2—Identification of n = 20 participants for this research**Table 1**—Baseline diagnostic polysomnographic data from 20 patients

Variable	Mean ± SD
Total sleep time, min	370 ± 50
Sleep latency, min	20 ± 19
Sleep efficiency, %	83 ± 11
Arousal index, arousals/h of sleep	24 ± 19
Stage N1, %	14.8 ± 8.5
Stage N2, %	59.4 ± 10.3
Stage N3, %	8.5 ± 10.1
Stage R, %	17.4 ± 6.1
Apnea/hypopnea index, no./h of sleep	26 ± 21
Minimum oxygen saturation, %	84 ± 4
% of sleep time with oxygen saturation < 88%	3.3 ± 11.5

index was 31 ± 6 kg/m². Races represented included Caucasian (n = 16), African American (n = 3), and Asian American (n = 1). **Figure 2** outlines how these participants were identified. Baseline, full-night, laboratory-based polysomnographic findings are summarized in **Table 1**. The time between the photographs and Epworth scores obtained initially, and then again after at least 2 months of successful CPAP use, was 5.8 ± 1.8 months. The mean Epworth score was 13.5 ± 2.3 at baseline and 5.6 ± 2.7 after use of CPAP, resulting in a mean change of -7.0 ± 4.1 (paired T-test $p < 0.0001$). Adequate adherence with PAP was confirmed electronically for all but one subject whose data card failed to function properly. Among the 30 nights just

Figure 3—Frontal facial images taken before (left) and after (right) treatment with positive airway pressure for obstructive sleep apnea

For the first subject (A), each of 22 raters (100%) chose the post-treatment image (at right) as the one that showed a more alert appearance; 21 (95%) chose it as more youthful; 22 chose it as more attractive; and 22 chose it as the post-treatment image. For the second subject (B), 18 (82%) of the 22 raters chose the post-treatment image (at right) as the one that showed a more alert appearance; 19 (86%) chose it as more youthful; 20 (91%) chose it as more attractive; and 82% chose it as the post-treatment image. Each of these subjects had a baseline apnea/hypopnea index between 5 and 14 events per hour of sleep, and a minimum oxygen saturation between 75% and 85%.

prior to the post-PAP images, the percent of nights that showed ≥ 4 h of use ranged from 70% to 100% (mean $89\% \pm 10\%$). Mean hours of PAP use, on nights used, was 6.3 ± 0.6 .

Subjective Facial Ratings

For the 20 subjects, on average just over two-thirds of the 22 raters “correctly” identified the post-CPAP photograph rather than the pre-CPAP photograph as having a more alert appearance ($68\% \pm 26\%$, $p = 0.0053$, **Figure 3**). The mean percent of raters who “correctly” identified the more youthful appearing image was $64\% \pm 26\%$ ($p = 0.026$); the more attractive image, $67\% \pm 26\%$ ($p = 0.0068$); and the treated condition, $66\% \pm 27\%$ ($p = 0.015$). Subjects clearly did show some variability, and raters on average “erroneously” identified the *pre*-CPAP photograph as appearing more alert, youthful, attractive, or treated for 5, 7, 6, and 7 of the subjects, respectively. Raters often but not always showed consistency for their particular selection,

across all 4 ratings, meaning, for example, that some raters may have chosen the post-CPAP image as looking more alert, but the pre-CPAP image as looking more youthful. Investigators noted subjectively that post-PAP images appeared to contain less prominent forehead creases. For Caucasian subjects alone ($n = 16$), the rates of “correct” identification for alert, youthful, attractive, and treated images closely resembled those for the entire group (data not shown). Among the raters, the 10 lay community raters may have been somewhat more accurate in their assessments than were the 12 medical community raters (**Table 2**). The investigators themselves perceived differences in alertness between images (not quantified), but anecdotally could only identify one possible physical change with some consistency: forehead wrinkles often appeared to diminish after treatment with PAP.

Objective Facial Changes

The forehead region showed a change in volume (contraction) by an average of $-0.012 \pm 0.028 \text{ cm}^3$ per cm^2 of surface area ($p = 0.022$). Infraorbital regions showed no consistent change in volume after use of PAP (0.013 ± 0.058 , $p = 0.41$).

On average, luminosity (Y) decreased from before use of PAP to after use of PAP in infraorbital areas bilaterally, and the maxillary area on the right (**Table 3**). Redness (Cr, adjusted for luminosity) decreased under the eyes bilaterally and on the left cheek, and showed a trend in the same direction on the right cheek. Blueness (Cb, adjusted for luminosity) increased under the right eye, but did not show significant changes under the left eye or on the cheeks.

Among the Caucasian subjects alone ($n = 16$), luminosity (Y) decreased after use of PAP more clearly, and in all areas (all $p < 0.05$, **Table 3**). Redness similarly increased more prominently, and in both areas bilaterally. Blueness increased more prominently and under the eyes bilaterally, though it still did not change significantly on the cheeks.

The size of the interpalpebral fissure, assessed as the ratio of the interpalpebral distance to the intercanthal distance for the same eye, showed no consistent changes across subjects from before to after use of PAP ($p = 0.90$ for right eyes, $p = 0.73$ for left eyes).

Objective Photogrammetry Measures, Polysomnographic Results, and Epworth Sleepiness Scale Changes as Correlates of Subjective Ratings

Decreased percent of sleep time spent in stage N3 on the baseline polysomnogram showed at least a trend to predicted improved post-PAP ratings for alertness ($\rho = -0.41$, $p = 0.0754$), attractiveness ($\rho = -0.41$, $p = 0.0751$), and being treated ($\rho = -0.50$, $p = 0.0249$). Among the 16 Caucasian subjects alone, decreased N3 predicted post-PAP improvement in all ratings: alertness ($\rho = -0.50$, $p = 0.0490$), youthfulness ($\rho = -0.55$, $p = 0.0272$), attractiveness ($\rho = -0.53$, $p = 0.0348$), and being treated ($\rho = -0.51$, $p = 0.0457$). Among all 20 subjects, decreased right cheek redness from pre- to post-PAP also showed marginal associations with subjective ratings of improved alertness ($\rho = -0.44$, $p = 0.0537$) and being treated ($\rho = -0.48$, $p = 0.0325$). Decreased left cheek redness was similarly, but not significantly associated with improved alertness ($\rho = -0.35$, $p = 0.1314$) and rater perception of images as being treated ($\rho = -0.42$, $p = 0.0666$). However, among the 16 Caucasian subjects, changes in redness may have been more effective, in that decreased right cheek redness predicted or nearly predicted improvement in all 4 ratings: alertness ($\rho = -0.51$, $p = 0.0423$), youthfulness ($\rho = -0.46$, $p = 0.0765$), attractiveness ($\rho = -0.50$, $p = 0.0493$), and being treated ($\rho = -0.52$, $p = 0.0369$). Similarly, decreased left cheek redness nearly predicted improved

Table 2—Accuracy* of paired image assessments from medical vs. lay community raters

	Medical Raters (n = 12)		Lay Raters (n = 10)	
	Percent Correct	p-value**	Percent Correct	p-value**
Alertness	66 ± 31	0.0210	70 ± 24	0.0027
Youthfulness	60 ± 32	0.1599	68 ± 24	0.0056
Attractiveness	65 ± 32	0.0121	70 ± 23	0.0019
Post-Treatment	65 ± 31	0.0311	69 ± 24	0.0043

Percent correct values are mean ± SD. *Accuracy determined by concordance of image rated as more alert, youthful, attractive, or likely to be treated with known post-PAP status. **Signed rank test.

Table 3—Mean ± SD changes (post-PAP minus pre-PAP) and statistical significance (signed rank test) for luminosity and color in facial images

All Subjects (n = 20)	Luminosity (Y)	p	Blueness (Cb)	p	Redness (Cr)	p
Infraorbital						
Left	-6.03 ± 8.82	0.0073	0.67 ± 1.69	0.1536	-1.36 ± 2.27	0.0328
Right	-7.58 ± 8.54	0.0014	1.07 ± 1.50	0.0049	-2.36 ± 2.12	0.0002
Maxillary						
Left	-3.85 ± 9.23	0.1054	0.89 ± 2.24	0.1536	-1.05 ± 2.71	0.0328
Right	-5.76 ± 11.11	0.0362	0.73 ± 2.32	0.2943	-1.25 ± 2.65	0.0696
Caucasian Subjects (n = 16)						
Infraorbital						
Left	-8.26 ± 7.73	0.0008	0.98 ± 1.63	0.0443	-2.08 ± 1.83	0.0010
Right	-9.01 ± 8.04	0.0006	1.30 ± 1.43	0.0021	-3.06 ± 1.51	< 0.0001
Maxillary						
Left	-6.29 ± 7.72	0.0131	1.03 ± 2.41	0.1754	-1.75 ± 2.03	0.0034
Right	-7.93 ± 11.27	0.0076	0.85 ± 2.36	0.2522	-1.96 ± 2.15	0.0052

ratings for alertness ($\rho = -0.45$, $p = 0.0788$) and being treated ($\rho = -0.45$, $p = 0.0808$).

No other objective pre- to post-PAP photogrammetry changes (e.g., in infraorbital color) or baseline polysomnographic findings listed in **Table 1** (e.g., apnea-hypopnea index) successfully explained with significance or consistent trends the pre-to-post-PAP changes in facial image ratings. This did not change when the analysis was confined to Caucasian subjects. Among all subjects and Caucasians alone, neither the percent of nights with adherent use of PAP prior to the post-PAP images nor the changes in Epworth Sleepiness Scale scores explained changes in the 4 types of pre-to-post-PAP image ratings (all $p > 0.10$).

DISCUSSION

The results of this study show that patients with OSA are perceived to appear more alert, more youthful, and more attractive after adherent use of PAP. Specifically, after 20 adults with OSA and daytime sleepiness were treated, 22 human raters identified post-treatment facial images about twice as frequently as baseline images to look more alert, more youthful, more attractive, and more likely to reflect treatment. Objective data from three-dimensional photogrammetry showed contraction in forehead surface volume after PAP, but surprisingly, no consistent change in infraorbital surface volume. Redness, adjusted for luminosity, decreased in infraorbital and cheek regions bilaterally, especially when the analysis was confined to Caucasian subjects. The size of the interpalpebral fissures did not change with treatment. Baseline polysomnographic measures and pre-to-post-PAP changes in facial surface volume, color, interpalpebral fissures, and Epworth Sleepiness Scale scores for the most part failed to predict responses from the human raters. Two notable exceptions emerged, however. Improved ratings from pre- to post-PAP for alertness, youthfulness, attractiveness, and being treated were predicted by less deep NREM sleep on the baseline polysomnogram. Among Caucasian subjects, ratings were also predicted with significance or trends by decreased redness over the cheeks. In short, these findings may be the first to confirm, in a research setting, clinicians' anecdotal impressions that patients often look more refreshed or younger after treatment for OSA. Results also give some initial insight into objective cues that do and do not predict a more alert appearance.

Subjective raters masked to all clinical information clearly distinguished post-PAP images as looking more alert, and at the same time the OSA patients on average showed robust improvement in self-assessed sleepiness after PAP treatment. The observation was surprising, therefore, that improvement in subjects' sleepiness scores did not predict the alertness ratings. Conceivably, raters could have recognized facial features that change with PAP but do not reflect alertness. However, Epworth scores show little, if any, significant associations with apnea severity or objective sleepiness as assessed by multiple sleep latency tests.¹⁴ Also, clinical experience suggests that individuals who observe a patient often perceive the sleepiness better than the affected patient does.¹⁵ We speculate that measurement limitations of the Epworth are most likely to explain the lack of better correspondence among our

subjects between alertness ratings and improvement in self-assessed sleepiness.

Contraction of surface volume over the forehead, from pre- to post-PAP, potentially could reflect effects of nightly fluid volume shifts that occur in untreated OSA. Nocturnal fluid shifts are already known to increase neck circumference and decrease leg circumference,¹⁶ and nightly recurrence may promote development of floppy eyelids in OSA.¹⁷ Treatment of OSA may improve cardiac output, reduce fluid accumulation in lower extremities, reduce nightly fluid shift to the head and neck, and over time diminish the volume of subcutaneous facial tissue.

Interestingly, the investigators' own inspections of the pre- and post-treatment images suggested a reduction in prominence of forehead wrinkles, rather than volume. These changes conceivably could have arisen in association with the decrease in volume that was measured. Alternatively, consciously or not, patients may have felt less compelled to raise their eyebrows to keep sleepy eyelids open, after treatment for sleep apnea. Progressive contraction of the frontalis muscle is a known response to mild through moderate levels of drowsiness.¹⁸ Less contraction of the frontalis, or lowered muscle tone in association with better sleep on a chronic basis, could explain decreased wrinkles.

Lessened redness under the eyes and over the cheeks, after treatment with PAP, was an unexpected finding. However, OSA is increasingly understood as an inflammatory process with significant PAP-responsive effects on vascular endothelium and the immune system.¹⁹ Our findings suggest that in addition to local and systemic internal inflammation, OSA may be associated with dermatologic manifestations in the face, if not elsewhere.^{20,21}

In contrast, our data do not support the common notion that dark, swollen, baggy circles under the eyes are a main feature of sleepiness, at least in the context of the chronic disease model we chose. In fact, blueness tended to increase under the eyes, after treatment. The authors are unaware of previous reports in the biomedical literature that provide data for comparison.

The closest previous work in this area showed that 65 raters of photographs from 20 healthy young adults, before and after acute sleep deprivation, could distinguish images of sleep deprived subjects as appearing on average 19% more tired, 6% less healthy, and 4% less attractive.⁵ However, patients with sleep apnea were not studied, no objective sleep measures were obtained, photogrammetry was not used, and no objective photographic measures were taken to assess which might provide cues to the human raters. Other authors have used craniofacial dimensions derived from photographs, if not photogrammetry, to develop a reasonably successful prediction model for risk of OSA.^{22,23} Subsequent work suggested that some specific facial dimensions, now derived from magnetic resonance imaging, were associated with upper airway structural volumes.²⁴ None of these studies, however, examined changes in appearance after treatment of OSA or improvement in daytime sleepiness. Finally, other studies have used baseline photographs and then digital video to identify facial movements or changes in expression that could signal increasing sleepiness, in real time, as a warning for drowsy drivers.¹⁸ We cannot exclude the possibility that facial expression played some role in our human ratings and objective volumetric or interpalpebral measures, but we did

encourage participants to maintain neutral expressions during image acquisition.

One limitation of our observational study was that it was not a randomized controlled trial, which would be necessary to prove an effect of PAP on appearance. A modest sample size in this initial study did not allow many hypotheses, beyond those about the facial appearance of alertness, to be tested in a definitive manner. The volumetric, color, interpalpebral, apnea severity, and sleepiness findings will require confirmation in larger samples. Although efforts were made to keep photogrammetry equipment in the same position and settings through the duration of the protocol, an increase in luminosity on average with the post-PAP images raises the possibility that lighting may have changed inadvertently, or that seasonal effects on skin had some impact. Use of color measures that account for luminosity should have accounted for this potential confounding variable. Other potential confounders may have included variables such as weight change or newly adopted activities or sports, which were not assessed after the several months of PAP use. Static facial images in comparison to full personal interaction are unlikely to convey as much information about alertness, youthfulness, and attractiveness. Thus, the effects we found of untreated OSA and use of PAP on facial appearance may be magnified in real-life situations. However, recent research does suggest strong agreement between attractiveness ratings made on the basis of static images and videos.²⁵

In conclusion, the demonstration in this study that unbiased raters can discern an improved facial appearance—more alert and probably more youthful and attractive as well—after treatment for obstructive sleep apnea has significant implications for clinical practice. Adherence to PAP, as with daily medications, is a major challenge.²⁶ Clinicians routinely seek to motivate patients through education about cardiovascular, neurologic, psychiatric, quality of life, and other benefits that may accrue from successful treatment of OSA. Our findings suggest another strategy to augment use of PAP. Patients with OSA could be told that others around them are likely to perceive them to be more alert, more youthful, and more attractive if they use their PAP. Pre- and post-treatment pictures could be taken and shown to patients, who may well see improvement. Facial cues are probably the most important social communication tools after language.²⁷ Facial appearance affects perception of emotion, intention, health, productivity, and potential mate value; guides interactions and behavior; and could have impact on survival.²⁸ Patients may appreciate the implications of a more alert appearance and have more tangible, proximate, and personally motivating goals in mind when they make decisions each night about whether to use their PAP.

REFERENCES

- Jenkinson C, Stradling JR, Petersen S. Comparison of three measures of quality of life outcome in the evaluation of continuous positive airways pressure therapy for sleep apnoea. *J Sleep Res* 1997;6:199-204.
- Reimer MA, Flemons WW. Quality of life in sleep disorders. *Sleep Med Rev* 2003;7:335-49.
- Tefft BC. Prevalence of motor vehicle crashes involving drowsy drivers, United States, 1999-2008. *Accid Anal Prev* 2012;45:180-6.
- Codner MA, Wolffi JN, Anzarut A. Primary transcutaneous lower blepharoplasty with routine lateral canthal support: a comprehensive 10-year review. *Plast Reconstr Surg* 2008;121:241-50.

- Axelsson J, Sundelin T, Ingre M, Van Someren EJ, Olsson A, Lekander M. Beauty sleep: experimental study on the perceived health and attractiveness of sleep deprived people. *BMJ* 2010;341:c6614.
- Tzou CH, Frey M. Evolution of 3D surface imaging systems in facial plastic surgery. *Facial Plast Surg Clin North Am* 2011;19:591-602.
- Heike CL, Upson K, Stuhauig E, Weinberg SM. 3D digital stereophotogrammetry: a practical guide to facial image acquisition. *Head Face Med* 2010;6:18.
- American Academy of Sleep Medicine. *International classification of sleep disorders, 2nd ed.: diagnostic and coding manual*. Westchester, IL: American Academy of Sleep Medicine, 2005.
- Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991;14:540-5.
- Iber C, Ancoli-Israel S, Chesson A, Quan SF, for the American Academy of Sleep Medicine. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. Westchester, IL: American Academy of Sleep Medicine, 2007.
- Aldridge K, Boyadjiev SA, Capone GT, Deleon VB, Richtsmeier JT. Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet A* 2005;138A:247-53.
- Kau CH, Richmond S, Incrapera A, English J, Xia JJ. Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *Int J Med Robot* 2007;3:97-110.
- Kakumanu PM, Makrogiannis S, Bourbakis N. A survey of skin-color modeling and detection methods. *Pattern Recognit* 2007;40:1106-22.
- Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology* 1999;52:125-31.
- Dement WC, Carskadon MA, Richardson G. Excessive daytime sleepiness in the sleep apnea syndrome. In: Sleep apnea syndromes. Guilleminault C, Dement WC, eds. Alan R. Liss: New York. 1978. pp 23-46.
- Redolfi S, Yumino D, Ruttanaumpawan P, et al. Relationship between overnight rostral fluid shift and obstructive sleep apnea in nonobese men. *Am J Respir Crit Care Med* 2009;179:241-6.
- Acar M, Firat H, Acar U, Ardic S. Ocular surface assessment in patients with obstructive sleep apnea-hypopnea syndrome. *Sleep Breath* 2013;17:583-8.
- Hachisuka S, Ishida K, Enya T, Kamijo M. Facial expression measurement for detecting driver drowsiness. EPCE'11 Proceedings of the 9th international conference on engineering psychology and cognitive ergonomics. Springer-Verlag: Berlin. 2011:135-44.
- Oyama J, Yamamoto H, Maeda T, Ito A, Node K, Makino N. Continuous positive airway pressure therapy improves vascular dysfunction and decreases oxidative stress in patients with the metabolic syndrome and obstructive sleep apnea syndrome. *Clin Cardiol* 2012;35:231-6.
- Kahan V, Andersen ML, Tomimori J, Tufik S. Can poor sleep affect skin integrity? *Med Hypotheses* 2010;75:535-7.
- Kushida CA, Everson CA, Suthipinittharm P, et al. Sleep deprivation in the rat: VI. Skin changes. *Sleep* 1989;12:42-6.
- Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea—a novel quantitative photographic approach. *Sleep* 2009;32:37-45.
- Lee RW, Petocz P, Prvan T, Chan AS, Grunstein RR, Cistulli PA. Prediction of obstructive sleep apnea with craniofacial photographic analysis. *Sleep* 2009;32:46-52.
- Lee RW, Sutherland K, Chan AS, et al. Relationship between surface facial dimensions and upper airway structures in obstructive sleep apnea. *Sleep* 2010;33:1249-54.
- Rhodes G, Lie HC, Thevaraja N, et al. Facial attractiveness ratings from video-clips and static images tell the same story. *PLoS One* 2011;6:e26653.
- Sawyer AM, Gooneratne NS, Marcus CL, Ofer D, Richards KC, Weaver TE. A systematic review of CPAP adherence across age groups: clinical and empirical insights for developing CPAP adherence interventions. *Sleep Med Rev* 2011;15:343-56.
- Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci* 2000;4:223-33.
- Rhodes G. The evolutionary psychology of facial beauty. *Annu Rev Psychol* 2006;57:199-226.

ACKNOWLEDGMENTS

The authors are grateful for the generous support of the Jonathan A. Covault Memorial Foundation for Sleep Disorders Research, which made this study possible. The authors also thank the subjects who participated, the volunteers who rated facial images for this research, and Kenneth E. Guire, M.S., who provided advice on statistical analyses. Work for this study was performed at the University of Michigan, Ann Arbor, MI.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication January, 2013

Submitted in final revised form March, 2013

Accepted for publication April, 2013

Address correspondence to: Ronald D. Chervin, M.D., M.S., Michael S. Aldrich Sleep Disorders Laboratory, C728 Med Inn, Box 5845, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5845; Tel: (734) 647-9064; Fax: (734) 647-9065; E-mail: chervin@umich.edu

DISCLOSURE STATEMENT

This was not an industry supported study. Dr. Chervin serves on the board of directors for the AASM, ABSM, ASMF, and International Pediatric Sleep Association. He is on the advisory board of the non-profit Sweet Dreamzzz. He has received research support from Philips Respironics and Fisher Paykel and has consulted for the NIH, Proctor & Gamble, and Zansors. He serves as a section editor for UpToDate and receives remuneration. Dr. Burns has received research support from the NIH and the Department of Defense. The other authors have indicated no financial conflicts of interest.