https://doi.org/10.5664/jcsm.8246

JCSM | Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Predicting sleep apnea from three-dimensional face photography

Peter Eastwood, PhD[1,2,*]; Syed Zulqarnain Gilani, PhD[3,4,*]; Nigel McArdle, PhD, MD[1,2]; David Hillman, PhD, MD[1,2]; Jennifer Walsh, PhD[1,2]; Kathleen Maddison, PhD[1,2]; Mithran Goonewardene, BDSc, MMedSc[5]; Ajmal Mian, PhD[3]

[1]Centre for Sleep Science, School of Human Sciences, University of Western Australia, Perth, Western Australia, Australia; [2]West Australian Sleep Disorders Research, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia; [3]School of Computer Science and Software Engineering, University of Western Australia, Perth, Western Australia, Australia; [4]School of Science, Edith Cowan University, Joondalup, Western Australia, Australia; [5]Oral Development and Behavioural Sciences, University of Western Australia, Perth, Western Australia, Australia; *Joint first author

Study Objectives: Craniofacial anatomy is recognized as an important predisposing factor in the pathogenesis of obstructive sleep apnea (OSA). This study used three-dimensional (3D) facial surface analysis of linear and geodesic (shortest line between points over a curved surface) distances to determine the combination of measurements that best predicts presence and severity of OSA.

Methods: 3D face photographs were obtained in 100 adults without OSA (apnea-hypopnea index [AHI] < 5 events/h), 100 with mild OSA (AHI 5 to < 15 events/h), 100 with moderate OSA (AHI 15 to < 30 events/h), and 100 with severe OSA (AHI ≥ 30 events/h). Measurements of linear distances and angles, and geodesic distances were obtained between 24 anatomical landmarks from the 3D photographs. The accuracy with which different combinations of measurements could classify an individual as having OSA or not was assessed using linear discriminant analyses and receiver operating characteristic analyses. These analyses were repeated using different AHI thresholds to define presence of OSA.

Results: Relative to linear measurements, geodesic measurements of craniofacial anatomy improved the ability to identify individuals with and without OSA (classification accuracy 86% and 89% respectively, $P < .01$). A maximum classification accuracy of 91% was achieved when linear and geodesic measurements were combined into a single predictive algorithm. Accuracy decreased when using AHI thresholds ≥ 10 events/h and ≥ 15 events/h to define OSA although greatest accuracy was always achieved using a combination of linear and geodesic distances.

Conclusions: This study suggests that 3D photographs of the face have predictive value for OSA and that geodesic measurements enhance this capacity.

Keywords: 3dMD, craniofacial anatomy, linear discriminant analysis, obstructive sleep apnea, polysomnography, Raine Study, three-dimensional photography

Citation: Eastwood P, Gilani SZ, McArdle N, et al. Predicting sleep apnea from three-dimensional face photography. *J Clin Sleep Med*. 2020;16(4): 493–502.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Obstructive sleep apnea (OSA) is currently undiagnosed in many individuals and simple, accurate screening tools are needed to predict those who have OSA. 3D facial photography is a quick method of capturing linear and nonlinear (geodesic) distances between different facial features in order to determine the combination of measurements that best correlates with the presence and severity of OSA.

Study Impact: The study found that OSA was able to be predicted with 91% accuracy when linear and geodesic craniofacial measurements from 3D photography were combined into a single predictive algorithm. The conclusions have not been validated in other populations with different age, race, and body mass index distributions.

INTRODUCTION

Obstructive sleep apnea (OSA) is a common disorder estimated to affect 15% of middle-aged males and 5% of middle-aged females.[1] It is characterized by repetitive episodes of partial or complete upper airway obstruction that are associated with hypoxemia, sympathetic activation, and sleep disruption. It is associated with burdensome symptoms and substantial medical comorbidities, including sleepiness-related accidents,[2] diabetes, cardiovascular diseases,[3] and depression.[4]

Studies demonstrate familial aggregations of OSA, suggesting powerful genetic predispositions, although these remain to be precisely characterized. Despite OSA being readily treatable, most cases (75%) remain unidentified,[5,6] because symptoms and signs are unrecognized, ignored, or misattributed to other causes. Current screening tools involve questionnaires, which while reasonably sensitive for OSA are (relatively) nonspecific, resulting in a high rate of false-positive results.[7] Other low-cost screening tools are needed.[8–11]

Craniofacial anatomy is recognized as an important predisposing factor in OSA pathogenesis.[12] Studies using magnetic resonance imaging (MRI) have shown that midface and lower-face width are correlated with OSA severity, supporting the notion that facial structure is important in the development of OSA.[13] Lateral cephalometry (x-ray) studies have demonstrated the importance of facial measures such as maxillary and mandibular length and intermaxillary space[14] in determining risk of OSA. Cephalometry and MRI allow accurate measurement of specific dimensions of the facial skeleton and upper airway. However, these techniques are not available for routine

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

clinical assessment and their use is limited by cost in the case of MRI and the risks associated with exposure to ionizing radiation in the case of cephalometry.

Two-dimensional (2D) photography has been used as an alternative to these more complex techniques and has several advantages. In particular, it is safe, inexpensive, portable, and easily accessible. Facial phenotypes assessed with 2D photography are closely correlated with upper airway anatomy as determined using MRI. Hence, 2D photography represents a reliable method for assessing both internal and external facial structures.[13] 2D photography captures several anatomic risk factors for OSA related to skeletal restriction, regional adiposity, and obesity. Such craniofacial risk factors include a wider and flatter mid and lower face, a shorter and retruded mandible, a smaller enclosed area within the mandible, and more soft tissues or fat deposition on the anterior neck.[12,15] Measurements such as these have been used to predict OSA severity with reasonable accuracy.[15,16] However, 2D photography cannot capture the nonlinear nature of craniofacial anatomy, such as shape and contour.

Three-dimensional (3D) photography overcomes this limitation of representing the 3D structure of the face with 2D imaging. Although similar to 2D photography it can provide information regarding linear distances and angles, unlike 2D photography it can additionally provide information on facial contours (ie, geodesic distances) making it an ideal tool for assessing the role of craniofacial structure in the pathogenesis of OSA. 3D photography allows assessment of the skeletal and soft tissues of the face and neck in a faster, cheaper, more readily available, and less invasive manner than MRI[17] and is already used in applications that range from measuring aesthetic facial parameters[18] to orthodontic diagnosis and evaluation of the craniofacial growth and development.[19] To date there has been only one study examining the potential role of 3D facial analysis to predict OSA in adults. In this study, Lin et al[20] showed in 36 male Asian patients with OSA that linear measurements of craniofacial distances, areas, angles, and volumes captured by 3D photography showed strong agreement with the same measurements obtained with 3D computed tomography (CT).

The current study aimed to use 3D facial surface analysis of linear and geodesic measurements to determine the combination of measurements that best correlates with the presence and severity of OSA. In particular, we sought to determine whether geodesic measurements increased the accuracy with which individuals with OSA could be identified from 3D photographs of the face.

METHODS

Participants

A sample of 50 middle-aged adults without OSA was recruited from participants in an ongoing community-based study of the prevalence of OSA between September and December 2015.[21] Sleep studies from sequential participants were analyzed, presence and severity of OSA were determined, and 50 participants without OSA (apnea-hypopnea index [AHI] < 5 events/h) with good-quality 3D face images (eg, without beards and missing data points) were identified.

In addition, a clinical sample of 350 middle-aged adults was recruited from the Sleep Clinic at the Western Australian Sleep Disorders Research Institute, Department of Pulmonary Physiology and Sleep Medicine at Sir Charles Gairdner Hospital between April 2015 and September 2016. Sleep studies from sequential participants were analyzed, severity of OSA determined, those with poor-quality 3D face images excluded, and participants allocated into one of 4 groups until the following sample sizes were reached: 50 without OSA, 100 with mild OSA, 100 with moderate OSA, and 100 with severe OSA.

Informed written consent was obtained from each participant and ethical approval for the study was obtained from the Human Research Ethics Committees at The University of Western Australia Human Research Ethics Committee (RA/4/1/7236) and Sir Charles Gairdner Hospital (No. 2014-059).

Sleep study

Standard overnight polysomnography was performed at one of two sites (Centre for Sleep Science, University of Western Australia or the Western Australian Sleep Disorders Research Institute, Sir Charles Gairdner Hospital).

Sleep studies at both sites were performed according to American Academy of Sleep Medicine (AASM) recommendations.[22] Briefly, electroencephalogram, electrooculogram, and chin electromyogram were measured using surface electrodes. Respiration was monitored with nasal prongs, an oronasal thermistor and thoracic and abdominal respiratory bands. Blood oxygen saturation (SaO_2) and heart rate were monitored continuously from a pulse oximeter on the index finger and electrocardiography, respectively. Leg movements were monitored by electromyography electrodes placed over the tibialis anterior muscle. A position sensor, microphone, and a live video feed via an infrared camera were used to monitor body position and snoring. A sleep technician monitored the recordings and video in each room for the duration of the study.

Data at both sites were acquired using Compumedics Grael (Compumedics, Victoria, Australia) system and scored by an experienced sleep technician using Profusion (PSG4) software according to the AASM 2012 (version 2.0) rules for the scoring of sleep and associated events.[23] The AHI was calculated as the total of all apneas and hypopneas divided by the total sleep time. Severity of OSA was defined as mild (AHI 5 to < 15 events/h), moderate (AHI 15 to < 30 events/h) or severe (AHI ≥ 30 events/h).

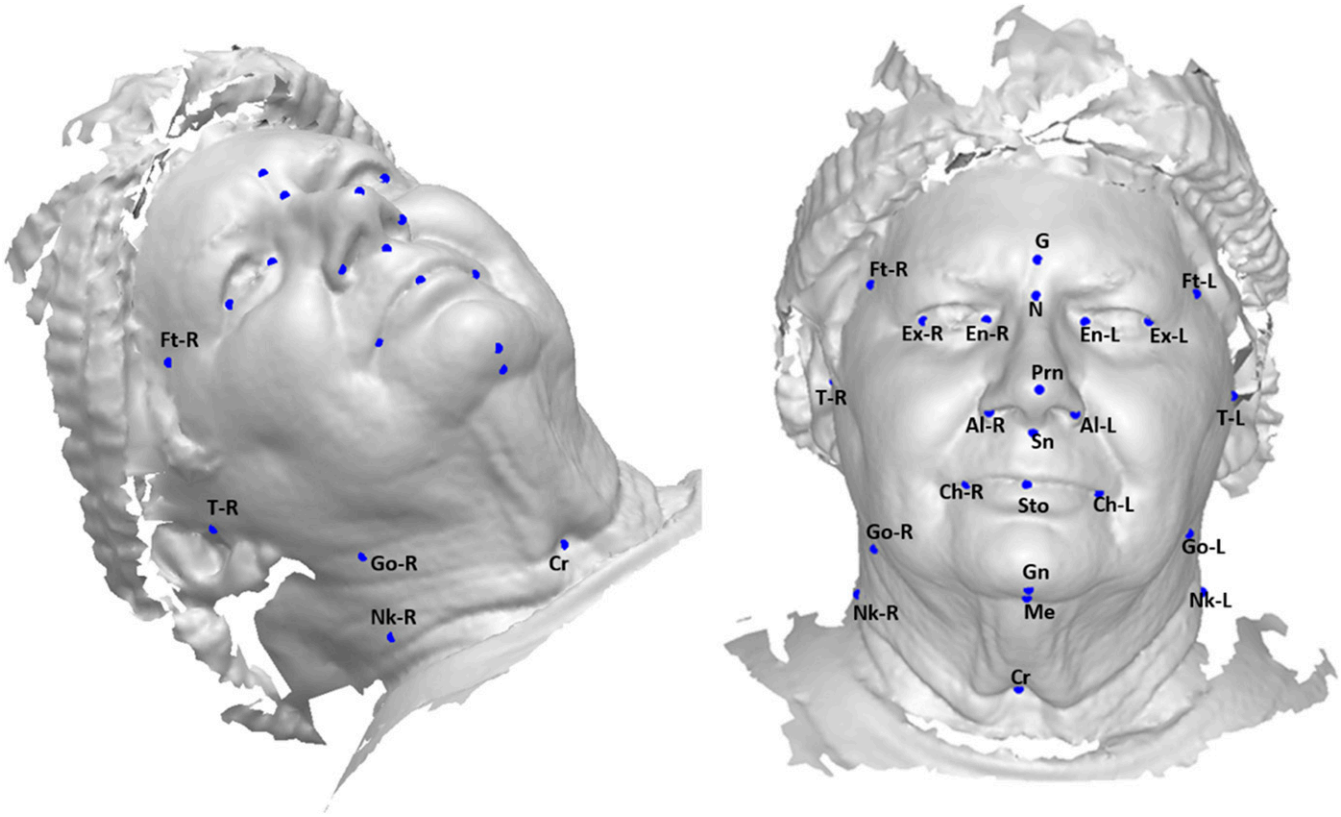
At the time of the sleep study, measurements were also obtained of each participant's height (stadiometer) and weight, body mass index (BMI) was determined (height/weight^2) and neck circumference measured using a tape measure positioned at the level of the cricoid.

3D photography

At both study sites, 3D photographs were taken using the 3dMD craniofacial scanner system (LCC, Atlanta, Georgia, USA). The scanner generates 180° (ear to ear) and neck region 3D images using the technique of triangulation.[24,25] Obtaining the image required the patient to sit on a chair between two cameras, with the hair pulled back from the face. High-resolution images were

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

Figure 1—Annotated craniofacial landmarks.



Al-L = alare left, *Al-R* = alare right, *Ch-L* = chelion left, *Ch-R* = chelion right, *Cr* = cricoid, *En-L* = endocanthion left, *En-R* = endocanthion right, *Ex-L* = exocanthion left, *Ex-R* = exocanthion right, *Ft-L* = frontotemporale left, *Ft-R* = frontotemporale right, *G* = glabella, *Gn* = gnathion, *Go-L* = gonion left, *Go-R* = gonion right, *Me* = menton, *N* = nasion, *Nk-L* = neck left, *Nk-R* = neck right, *Prn* = pronasale, *Sn* = subnasale, *Sto* = stomion, *T-L* = tragion left, *T-R* = tragion right. Permission to use this photograph has been provided by the participant.

captured within 1.5 milliseconds and simultaneous acquisition of geometry and color-texture data were achieved by the synchronization of the individual digital cameras.

Landmark annotation

A total of 24 landmarks were annotated by a single scorer who was blinded to OSA status, on the 3D image of each face (**Figure 1**) using custom software developed by the authors (AM, SG) in Matlab (R2018b) which allowed these landmarks to be objectively positioned. Two of these landmarks (left and right gonion) were physically marked on each individual's face at the time of scanning. Selection of landmarks were guided by earlier studies of Sutherland and Lee.[15,26,27] The software allowed the face to be rotated to allow optimal visualization and determination of each landmark. Landmark location was highly repeatable: the average intraclass correlation coefficient for the 22 landmarks was 0.99 when calculated from 20 faces (5 from each OSA category annotated twice each), excluding the gonions as they were physically marked on the face. The 3D coordinates of each landmark (in the *x*, *y*, and *z* axes) were stored for later analysis.

Feature extraction

The 24 landmarks allowed a total of 276 distances between two (paired) landmarks to be defined. For parsimony and guided by earlier studies[12,26,27] 25 pairs of points were selected and distances calculated for each in both the linear dimension (ie, direct euclidian distance between the two points) and geodesic dimension (ie, the shortest distance between two points when following the contour of the face/skin) (**Figure 2**, **Table S1** and **Table S2** in the supplemental material). Using linear dimensions only, angles were also determined between sets of three points (**Figure S1** and **Table S3** in the supplemental material). These linear and geodesic distances and angles were termed "features" and used in further analysis.

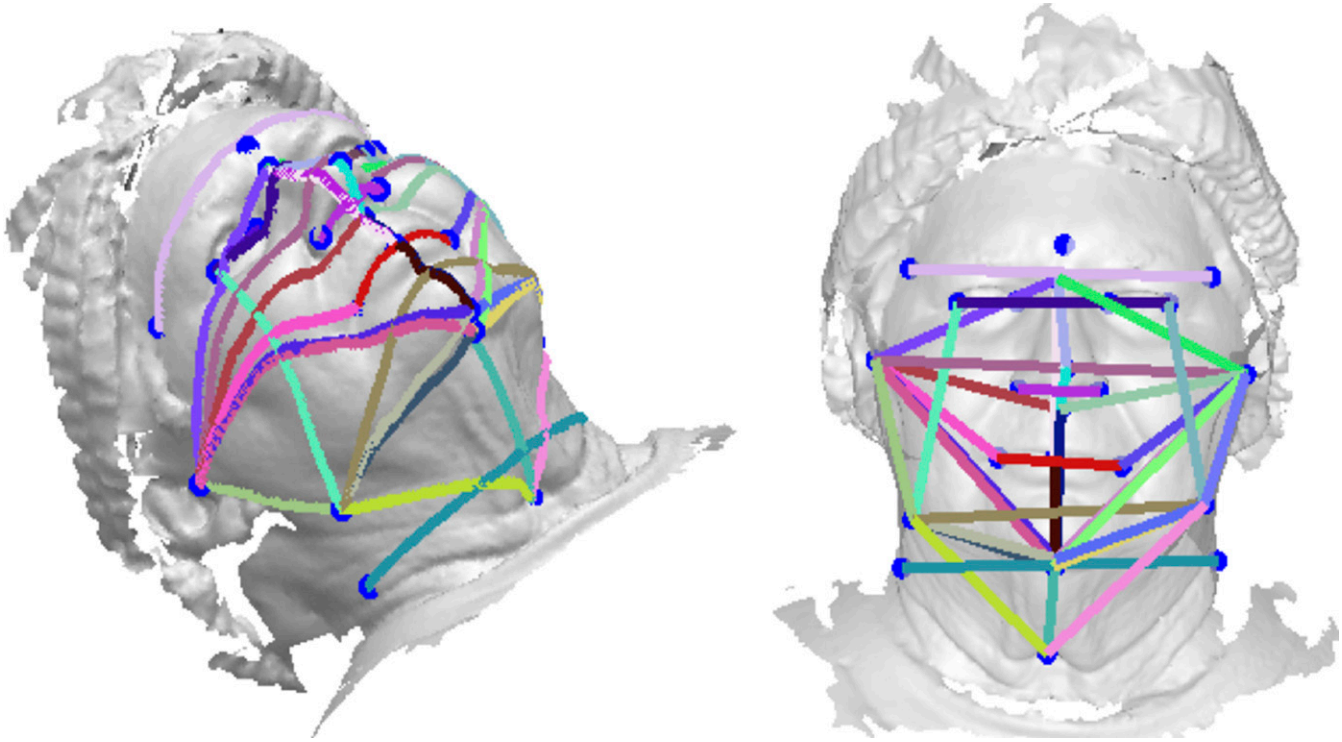
Analysis

Data were divided into two classes, control and OSA, based on a polysomnography-derived threshold value of AHI ≥ 5 events/h to define OSA (n = 100 controls and n = 300 OSA). Based on this threshold value, a linear discriminant analysis (LDA) algorithm was developed and trained using the 3D linear distances, geodesic distances, and angles of each face. The algorithm was then trained and tested to classify new unseen cases (see next paragraphs) using AHI threshold values of 10, 15, 20, 25, 30, 35, 40, 45, and 50 events/h.

The LDA algorithm was used to distinguish facial features between the two classes: control and OSA. The goal of this machine learning algorithm was to find, using all features, a one- dimensional space where the distance between features

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

Figure 2—Geodesic and euclidian distances were determined between annotated landmarks.



within the same class is minimal and the distance between the two classes is maximal. This is known as the LDA space. The data were divided into 10 "bins." In each bin, 90% of the data (eg, features of 360 faces for a threshold of AHI = 5 events/h) picked randomly were used to train the algorithm (ie, learning the LDA space). During testing the features of a face were projected in the LDA space and a class was assigned to it based on its distance from the mean of both classes.[28] During the training phase the LDA algorithm was provided with the labels of only the training data to "learn" the optimal space. The remaining 10% of data (the unseen cases) were used to test the algorithm. In this testing phase the algorithm was blinded to the actual label of the test face as it was required to assign a label to each face (based on the developed algorithm). The derived classification was then compared to the actual classification. OSA was assigned the positive class. Hence, faces correctly classified as controls or OSA were assigned true negative and true positive labels, respectively. Further details on this cross-validation methodology and algorithm development can be found in **Table 1** and the supplemental text.

The accuracy with which different combinations of measurements could classify an individual as having OSA or not was assessed by calculating sensitivity, specificity and accuracy and by performing receiver operating characteristic (ROC) analyses. Comparison of distances and angles between those without OSA and with OSA were assessed using *t* tests with Bonferroni corrections for multiple comparisons. Comparison of distances and angles between those without OSA and those with different severities of OSA were assessed using analysis of variance with Tukey-Kramer corrections for multiple comparisons. All data are expressed as mean ± standard deviation or mean ± standard error and a value of $P = .05$ was considered statistically significant. All analyses were undertaken using Matlab (R2018b).

RESULTS

Data from 400 participants (172 males), all of whom had 3D face scans, overnight laboratory-based sleep studies, and measurements of neck circumference and BMI, were used in this study to develop, train, and validate the predictive algorithm (**Table 1** and supplemental text). Ethnicity data were available on 155 of the participants of whom 72% were classified as Caucasian (both parents were Caucasian). Compared to those without OSA (AHI < 5 events/h), those with OSA (AHI ≥ 5 events/h) were older (47.4 ± 13.8 versus 54.5 ± 15.8 years, $P < .05$) and had increased AHI (2.6 ± 1.28 versus 30.7 ± 26.0 events/h, $P < .05$), BMI (26.9 ± 5.2 versus 32.3 ± 7.6 kg/m^2, $P < .05$) and neck circumference (35.0 ± 3.8 versus 39.7 ± 4.8 cm, $P < .05$) (**Table 2**).

The accuracy with which facial measurements, BMI, and neck circumference could classify an individual as having OSA (AHI ≥ 5 events/h) or not (AHI < 5 events/h) is summarized in **Table 3** and shown graphically in **Figure 3** (the algorithm is reported in the supplemental text). The predictive accuracy of geodesic measurements alone was significantly greater than that for linear distances alone, being 89 ± 1% and 86 ± 1%, respectively ($P < .01$). In general, the accuracy to predict the presence of OSA was lowest for simple anthropometric measures of BMI and neck circumference, was increased for measurements of linear or geodesic distances, and was maximal (with an accuracy of

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

Table 1—Algorithm.

1. Let $\mathbf{F}_j = [x_i, y_i, z_i]^T$, where j = 1,2,…,N, I = 1,2,….,(points on each face $\mathbf{F}_j$) and N = 400.
2. Manually annotate M = 24 landmarks on each face $\mathbf{F}_j$ and form the matrix landmarks, $\mathbf{L}_j = [x_i, y_i, z_i]^T$, where j = 1,2,…,N, i = 1,2,….,M, N = 400 and M = 24.
3. Select P = 25 pairs of landmarks on each face (Table S1 and Table S2).
4. Extract linear distance between the P landmarks on each face such that $x_p = \left(\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}\right)$, where x_1, y_1, z_1 are the coordinates of the first landmark in the pair and x_2, y_2, z_2 are the coordinates of the second landmark in the pair. P = 25.
5. Extract geodesic distance between the same P = 25 pairs of landmarks. Geodesic distance is the shortest surface distance between the coordinates of first point in the pair and the coordinates of second point in the pair.
6. Select ten 3-tuples of landmarks for calculating angles (Table S3).
7. Using the center landmark of each 3-tuple as the center vertex find the angle between the three landmarks of the 3-tuple. (Figure S1)
8. Set threshold AHI < 5 events/h as Controls and AHI ≥ 5 events/h as OSA.
9. Create 10 random folds with each fold having 360 faces for training and 40 unique faces for testing.
10. For each fold: a. Train linear discriminant analysis algorithm on the features of training faces and test for classification on features of test faces. b. Take the average of the classification accuracy for all ten folds and report results (Table 2)

AHI = apnea-hypopnea index, OSA = obstructive sleep apnea.

Table 2—Participant characteristics.

OSA Severity	Participants			Age (years)	BMI (kg/m^2)	NC[a] (cm)	AHI (events/h)
	Males	Females	Total				
Control (AHI < 5 events/h)	25	75	100	47.4 ± 13.8	26.9 ± 5.2	35.0 ± 3.8	2.6 ± 1.3
Mild (AHI 5 to < 15 events/h)	39	61	100	50.2 ± 16.7	29.2 ± 6.8	37.7 ± 4.5	9.9 ± 2.9
Moderate (AHI 15 to < 30 events/h)	50	50	100	55.2 ± 14.3	32.3 ± 7.2	40.3 ± 4.3	22.6 ± 4.6
Severe (AHI ≥ 30 events/h)	58	42	100	57.9 ± 15.6	35.4 ± 7.5	41.6 ± 4.7	59.7 ± 25.8
Total	172	228	400	52.7 ± 15.6	30.9 ± 7.4	38.3 ± 4.9	23.7 ± 25.6

[a]NC was obtained from only 322 participants (95 control; 84 mild; 77 moderate; 66 severe). Values are mean ± standard deviation. AHI = apnea-hypopnea index, BMI = body mass index, NC = neck circumference, OSA = obstructive sleep apnea.

91 ± 2%) for measurements incorporating a combination of linear and geodesic distances. The classification accuracy to predict the presence of OSA was not significantly different whether considering females alone or males alone (**Table 4**).

Differences in landmark distances between those with and without OSA (based on a cutoff of 5 events/h) are shown for geodesic distances, linear distances and angles in **Table S1**, **Table S2**, and **Table S3**. Craniofacial features that differed between the control and OSA groups, for both geodesic and linear distances included: upper, mid, and lower face depth; face width; mandibular length, width and posterior height; and neck width ($P < .05$ for all). The most discriminatory angular measurements between the two groups (based on magnitude of difference and values of P) were the mandibular width angle (right gonion-menton-left gonion); lower facial width angle (right tragion-menton- left tragion); and maxillary-mandibular relationship angle (subnasion-nasion-gnathion) ($P < .05$ for all).

Several of these geodesic and linear distances were different between control and mild, moderate, and severe OSA and between each of the different severities of OSA (**Table S4** and **Table S5** in the supplemental material). These included: upper and lower face depth; total and upper face height; lateral face height; face width; mandible width; and neck width ($P < .05$ for all). Angular measurements tended to be less discriminatory between controls and different severities of OSA (**Table S6** in the supplemental material).

The classification accuracy for predicting OSA was related to the AHI cutoff used to define the presence or absence of OSA. This is shown in **Figure 4**, which demonstrates that, relative to a cutoff of 5 events/h, accuracy decreases when using the two other most commonly used cutoffs of 10 and 15 events/h (note that the algorithm was retrained at each cutoff). However, regardless of the cutoff used the highest accuracy was achieved by using a combination of linear and geodesic distances.

DISCUSSION

The study was driven by the need to develop new, effective screening methods for OSA, given the high prevalence combined

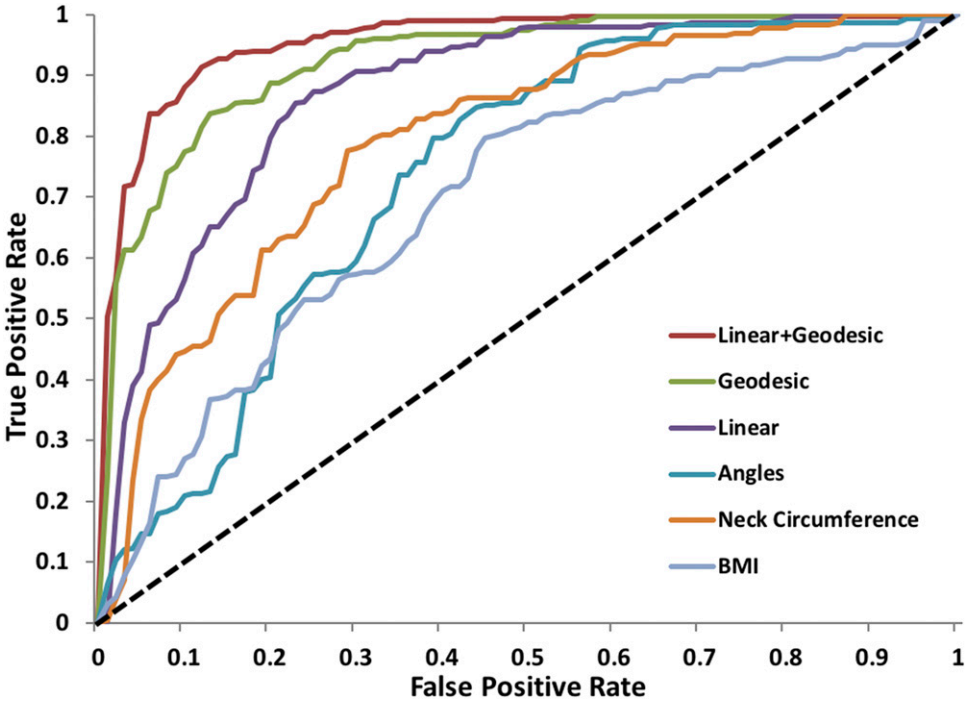
Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

Table 3—Classification accuracy of individuals without OSA (AHI < 5 events/h) and with OSA (AHI ≥ 5 events/h) using different features or combinations of features.

Feature	Sensitivity	Specificity	Accuracy	Area Under ROC Curve	Likelihood Ratio
Geodesic distances	0.96 ± 0.01	0.68 ± 0.04	0.89 ± 0.01	0.93	3.00
Linear distances	0.95 ± 0.01	0.51 ± 0.05	0.86 ± 0.01	0.88	1.94
Angles	0.95 ± 0.01	0.36 ± 0.06	0.82 ± 0.02	0.74	1.48
Combination of linear distances and geodesic distances	0.97 ± 0.01	0.76 ± 0.03	0.91 ± 0.02	0.96	4.04
Combination of linear distances, geodesic distances and angles	0.96 ± 0.01	0.68 ± 0.05	0.90 ± 0.01	0.93	3.00
NC	0.93 ± 0.01	0.40 ± 0.02	0.77 ± 0.01	0.78	1.55
BMI	0.97 ± 0.01	0.04 ± 0.01	0.75 ± 0.01	0.70	1.01
Combination of linear distances, geodesic distances, angles, NC, and BMI	0.96 ± 0.01	0.71 ± 0.04	0.90 ± 0.02	0.94	3.31

Values are mean ± standard error. AHI = apnea-hypopnea index, BMI = body mass index, NC = neck circumference, ROC = receiver operating characteristic, OSA = obstructive sleep apnea.

Figure 3—Receiver operating characteristic curves showing the sensitivity and specificity for predicting OSA (AHI ≥ 5 events/h) using different landmarks or anthropometric measurements.



AHI = apnea-hypopnea index, BMI = body mass index, NC = neck circumference, OSA = obstructive sleep apnea, true negative rate (specificity), true positive rate (sensitivity).

with low identification; the disappointingly low specificity of existing screening techniques (questionnaires) and the prospect that 3D craniofacial photography could improve that; and that 3D photography techniques are widely available with more reliably derived, less time-consuming measures than 2D facial photography, which requires formalized anteroposterior and lateral photographs. The current study showed that, relative to linear or anthropometric measurements, geodesic (nonlinear) measurements of craniofacial anatomy improved the ability to identify individuals with and without OSA. Maximal classification accuracy (91%) was found when geodesic and linear craniofacial measurements were combined into a single predictive algorithm. Such screening could represent the first inexpensive, widely available step along the diagnostic pathway for OSA.

The notion that surface facial dimensions can identify individuals with OSA is supported by findings from previous studies showing that dimensions of the facial skeleton and soft

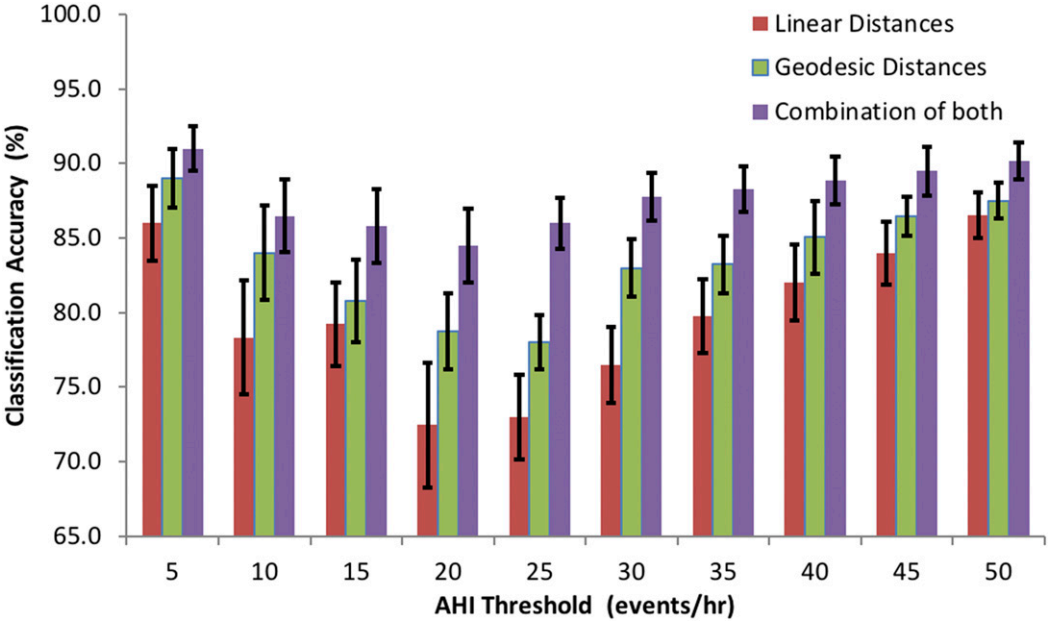
Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

Table 4—Classification accuracy of individuals without and with OSA using three AHI cutoffs (≥ 5, 10, and 15 events/h) using combination of linear and geodesic distances.

	Sensitivity	Specificity	Accuracy	Area Under ROC Curve	Likelihood Ratio
≥ 5 events/h					
Males	0.99 ± 0.01	0.72 ± 0.03	0.89 ± 0.01	0.82	3.52
Females	0.95 ± 0.01	0.79 ± 0.02	0.90 ± 0.01	0.93	4.47
All	0.97 ± 0.01	0.76 ± 0.03	0.91 ± 0.02	0.96	4.04
≥ 10 events/h					
Males	0.92 + 0.01	0.73 + 0.02	0.85 + 0.02	0.85	3.41
Females	0.87 + 0.02	0.79 + 0.02	0.86 + 0.02	0.89	4.19
All	0.90 + 0.02	0.77 + 0.03	0.87 + 0.02	0.92	3.91
≥ 15 events/h					
Males	0.93 + 0.01	0.71 + 0.02	0.85 + 0.02	0.85	3.24
Females	0.79 + 0.01	0.89 + 0.02	0.85 + 0.02	0.91	7.19
All	0.89 + 0.02	0.82 + 0.02	0.86 + 0.02	0.92	4.97

Values are mean ± standard error. AHI = apnea-hypopnea index, OSA = obstructive sleep apnea, ROC = receiver operating characteristic.

Figure 4—Classification accuracy for predicting OSA at different AHI thresholds using linear, geodesic, and combined distances.



Error bars, ± standard error. AHI = apnea-hypopnea index, OSA = obstructive sleep apnea.

tissues are related to upper airway dimensions[12,13] with airway narrowing an important predisposition to its collapse during sleep. Furthermore, many studies across different ethnic populations have shown that measurements of surface craniofacial dimensions are correlated with OSA severity.[12–15,26,29–31] The techniques used in these previous studies have included MRI, CT, lateral cephalometry, and 2D photography. Although cephalometry and MRI allow accurate measurement of specific dimensions of the facial skeleton and upper airway, these techniques are not suitable for routine clinical assessment and their use is limited by cost in the case of MRI and the risks associated with exposure to ionizing radiation in the case of cephalometry and CT. Craniofacial measurements from 2D photography have been used to predict OSA severity with reasonable accuracy[15,16]; however 2D photography is unable to effectively capture the nonlinear nature of craniofacial anatomy, such as shape and contour.

3D photography overcomes the limitation of representing the 3D structure of the face with 2D imaging because it has the ability to provide information regarding facial contours as well as the linear distances and angles available from 2D images. Furthermore, compared to 2D photography, 3D measurements of craniofacial distances and angles (obtained using the 3dMD system) have increased agreement with

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

equivalent measurements obtained from 3D CT.[20] The current study showed that the accuracy with which OSA can be predicted from craniofacial photographs is increased when information on facial contours is considered, regardless of the AHI threshold used to define the presence of OSA. Specifically, geodesic measurements alone were more accurate than linear measurements alone, with maximum accuracy derived from a combination of geodesic and linear measurements. The sensitivity, specificity, accuracy, and area under the ROC curve of the algorithm to predict OSA (AHI ≥ 5 events/h) based on this combination was 97%, 76%, 91%, and 0.96, respectively. Notably, these values were similar when considering the group overall, or males and females separately.

These measures of predictive capacity tend to be higher than previous studies using 2D facial photographic analysis, which have shown that correct OSA risk classification can be obtained in 76% to 79% of the cases, with sensitivities ranging from 73% to 85%, specificities from 28% to 70%, and area under the ROC curve from 0.73 to 0.87.[15,26,32] When comparing between studies and between techniques an important consideration is the AHI cutoff used to define the presence or absence of OSA. In all of these previous studies a cutoff of 10 events/h was used, whereas the current study used a cutoff of 5 events/h. It was interesting to note that, in the current study, the sensitivity, specificity, accuracy, and area under the ROC curve decreased to 90%, 77%, 87%, and 0.92, respectively, when an AHI threshold of 10 events/h was used, and to 89%, 82%, 86%, and 0.92, respectively, when an AHI threshold of 15 events/h was used, despite the algorithm being trained and then tested to classify new unseen cases at each threshold. This indicates that an AHI threshold of 5 events/h is where the accompanying facial morphology changes are maximal. It was notable, however, that regardless of the threshold used the maximum capacity to identify individuals with OSA was obtained when geodesic and linear craniofacial measurements are combined into a single predictive algorithm.

Craniofacial features that differed between the control and OSA groups tended to change similarly for both geodesic and linear distances, such that individuals with OSA were characterized by a 3% to 10% increase in measures of face depth; 10% to 13% increase in lateral face height; 5% increase in face width, 2% to 6% increase in mandibular length; 17% to 19% increase in posterior mandibular height; 4% to 5% increase in mandible width; and 10% to 13% increase in neck width. In terms of angular measurements, individuals with OSA had an 18% smaller maxillary-mandibular relationship angle, and a 10% to 12% decrease in angle between the menton and left/right mandible or tragion. It is difficult to compare these differences to those reported in previous studies using 2D measurements, as the distances and angles derived in the present study were calculated between points in 3D space (ie, each point having an *x*, *y*, and *z* coordinate), and are hence a composite of a 2D frontal view and a 2D lateral view. The only study using 3D craniofacial photography with which to compare our results was published recently by Lin et al.[20] Their analyses did not include geodesic distances; however, the linear measurements they found to be correlated with severity of OSA (eg, mandibular width, neck perimeter, mandibular length, facial width, binocular width) were similar to those that distinguished groups in the current study.

The data presented in this study suggest that 3D craniofacial photography has a potential role in screening for OSA in the general population, but not its diagnosis. Notably, even though the accuracy and sensitivity of the algorithm based on the combination of linear and geodesic distances were high at 91% and 97%, respectively, the specificity was 76% and the positive likelihood ratio (= sensitivity / (1 − specificity)) was only 4.04. These findings are consistent with a good screening test, which must have a high sensitivity (so that most cases are identified) and at least moderate specificity to ensure against too many false-positive results. However, they are inadequate for a diagnostic test that must be both highly sensitive and highly specific. In their review of nonlaboratory based devices to diagnose OSA, using an in-laboratory cutoff of 5 events/h, Collop et al[33] suggested that a useful diagnostic test should have a positive likelihood ratio > 5 and sensitivity of at least 0.825. However, it remains possible that its application in a population with increased pretest probability based on clinical assessment[34] could result in greater diagnostic accuracy, as might further refinements to the algorithm.

A further finding in this study was that there were systematic differences in several of the geodesic and linear distances between patients in the control, mild, moderate and severe OSA groups, although angular measurements tended to be less discriminatory. This finding suggests that, beyond identification of the presence of OSA, 3D photography may be helpful in predicting OSA severity. Although the internal validity of the data has been assessed in the methods used here, external validation in an independent sample of participants would be a desirable next step in evaluating this method for clinical use.

The method described in this study demonstrates the potential for 3D facial photographs to provide a rapid, simple, objective, and accurate method to identify individuals at high risk of OSA. Although it was developed using an expensive, high-resolution, research quality craniofacial scanner system, it might be possible to embed such a method into existing, cheap, off-the-shelf 3D photography systems. This could potentially provide a novel, accurate screening tool for widespread use in the general population to identify individuals at high risk of OSA. Such a screening tool is desirable given the high prevalence of OSA and its current low identification rates[5,35] and could represent a first, inexpensive, widely available step along the diagnostic pathway.

This study has several limitations. First, individuals with beards were excluded. This was necessary to obtain a full set of anatomic landmarks on all participants. However, beards are often present in those with maxillomandibular deficiency; thus, exclusion may skew the population. Second, the conclusions are limited to a middle-aged population. However, the prevalence of OSA is greatest and therefore the need for such screening tools is also greatest in this population. Third, the conclusions are limited to Caucasians. It is highly likely that different combinations of dimensions will be required to predict OSA in other ethnic groups given the well-known effect of ethnicity on craniofacial phenotype and OSA

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

anatomic risk factors.[36,37] Finally, the patients with OSA had a higher BMI, were older, and had a larger neck circumference than those without OSA. These features may affect facial parameters measured by 3D photography and could confound the results. However, as a real-world screening test our study shows that 3D photography was more predictive than BMI alone (**Figure 3**) in distinguishing patients with OSA from those without OSA.

In conclusion, the main findings of this study were that geodesic measurements add value to the capacity to identify patients with OSA from 3D photographs of the face and that a combination of linear and geodesic measures has a strong predictive value for the presence of OSA in the studied population.

ABBREVIATIONS

2D, two-dimensional
3D, three-dimensional
AASM, American Academy of Sleep Medicine
AHI, apnea-hypopnea index
BMI, body mass index
LDA, linear discriminant analysis
MRI, magnetic resonance imaging
OSA, obstructive sleep apnea
ROC, receiver operating characteristic
SaO_2, blood oxygen saturation

REFERENCES

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol*. 2013;177(9):1006–1014.
2. Terán-Santos J, Jiménez-Gómez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *N Engl J Med*. 1999;340(11):847–851.
3. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet*. 2005;365(9464):1046–1053.
4. Sforza E, de Saint Hilaire Z, Pelissolo A, Rochat T, Ibanez V. Personality, anxiety and mood traits in patients with sleep-related breathing disorders: effect of reduced daytime alertness. *Sleep Med*. 2002;3(2):139–145.
5. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep*. 1997;20(9):705–706.
6. Simpson L, Hillman DR, Cooper MN, et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. *Sleep Breath*. 2013;17(3):967–973.
7. Pereira EJ, Driver HS, Stewart SC, Fitzpatrick MF. Comparing a combination of validated questionnaires and level III portable monitor with polysomnography to diagnose and exclude sleep apnea. *J Clin Sleep Med*. 2013;9(12):1259–1266.
8. Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. *Sleep*. 2000;23(7):929–938.
9. Bouloukaki I, Kapsimalis F, Mermigkis C, et al. Prediction of obstructive sleep apnea syndrome in a large Greek population. *Sleep Breath*. 2011;15(4):657–664.
10. Sahin M, Bilgen C, Tasbakan MS, Midilli R, Basoglu OK. A clinical prediction formula for apnea-hypopnea index. *Int J Otolaryngol*. 2014;2014:438376.
11. Wilson G, Terpening Z, Wong K, et al. Screening for sleep apnoea in mild cognitive impairment: the utility of the multivariable apnoea prediction index. *Sleep Disord*. 2014;2014:945287.
12. Lee RW, Sutherland K, Chan AS, et al. Relationship between surface facial dimensions and upper airway structures in obstructive sleep apnea. *Sleep*. 2010;33(9):1249–1254.
13. Sutherland K, Schwab RJ, Maislin G, et al. Facial phenotyping by quantitative photography reflects craniofacial morphology measured on magnetic resonance imaging in Icelandic sleep apnea patients. *Sleep*. 2014;37(5):959–968.
14. Lowe AA, Fleetham JA, Adachi S, Ryan CF. Cephalometric and computed tomographic predictors of obstructive sleep apnea severity. *Am J Orthod Dentofacial Orthop*. 1995;107(6):589–595.
15. Lee RW, Petocz P, Prvan T, Chan AS, Grunstein RR, Cistulli PA. Prediction of obstructive sleep apnea with craniofacial photographic analysis. *Sleep*. 2009;32(1):46–52.
16. Perri RA, Kairaitis K, Cistulli P, Wheatley JR, Amis TC. Surface cephalometric and anthropometric variables in OSA patients: statistical models for the OSA phenotype. *Sleep Breath*. 2014;18(1):39–52.
17. 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(2):97–110.
18. 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.
19. Lane C, Harrell W Jr. Completing the 3-dimensional picture. *Am J Orthod Dentofacial Orthop*. 2008;133(4):612–620.
20. Lin SW, Sutherland K, Liao YF, et al. Three-dimensional photography for the evaluation of facial profiles in obstructive sleep apnoea. *Respirology*. 2018;23(6):618–625.
21. Straker L, Mountain J, Jacques A, et al. Cohort profile: the Western Australian Pregnancy Cohort (Raine) Study-Generation 2. *Int J Epidemiol*. 2017;46(5):1384–1385.
22. Iber C, Ancoli-Israel S, Chesson AL Jr, 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*. 1st ed. Westchester, IL: American Academy of Sleep Medicine; 2007.
23. Berry RB, Budhiraja R, Gottlieb DJ, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med*. 2012;8(5):597–619.
24. Weinberg SM, Kolar JC. Three-dimensional surface imaging: limitations and considerations from the anthropometric perspective. *J Craniofac Surg*. 2005;16(5):847–851.
25. Weinberg SM, Naidoo S, Govier DP, Martin RA, Kane AA, Marazita ML. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry. *J Craniofac Surg*. 2006;17(3):477–483.
26. Sutherland K, Lee RW, Petocz P, et al. Craniofacial phenotyping for prediction of obstructive sleep apnoea in a Chinese population. *Respirology*. 2016;21(6):1118–1125.
27. Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea–a novel quantitative photographic approach. *Sleep*. 2009;32(1):37–45.
28. Balakrishnama S, Ganapathiraju A; Institute for Signal and Information Processing. *Linear Discriminant Analysis - A Brief Tutorial*. Mississippi State, MS: Mississippi State University; 1998:1-8.
29. Amra B, Peimanfar A, Abdi E, et al. Relationship between craniofacial photographic analysis and severity of obstructive sleep apnea/hypopnea syndrome in Iranian patients. *J Res Med Sci*. 2015;20(1):62–65.
30. Remya KJ, Mathangi K, Mathangi DC, et al. Predictive value of craniofacial and anthropometric measures in obstructive sleep apnea (OSA). *Cranio*. 2017;35(3):162–167.
31. Cheung K, Ishman SL, Benke JR, et al. Prediction of obstructive sleep apnea using visual photographic analysis. *J Clin Anesth*. 2016;32:40–46.

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.

32. Espinoza-Cuadros F, Fernandez-Pozo R, Toledano DT, Alcazar-Ramirez JD, Lopez-Gonzalo E, Hernandez-Gomez LA. Speech signal and facial image processing for obstructive sleep apnea assessment. *Comput Math Methods Med*. 2015;2015:489761.
33. 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(5):531–548.
34. Myers KA, Mrkobrada M, Simel DL. Does this patient have obstructive sleep apnea?: The Rational Clinical Examination systematic review. *JAMA*. 2013;310(7):731–741.
35. Kapur V, Strohl KP, Redline S, Iber C, O'Connor G, Nieto J. Underdiagnosis of sleep apnea syndrome in U.S. communities. *Sleep Breath*. 2002;6(2):49–54.
36. Sutherland K, Lee RWW, Chan TO, Ng S, Hui DS, Cistulli PA. Craniofacial phenotyping in Chinese and Caucasian patients with sleep apnea: influence of ethnicity and sex. *J Clin Sleep Med*. 2018;14(7):1143–1151.
37. Sutherland K, Lee RW, Cistulli PA. Obesity and craniofacial structure as risk factors for obstructive sleep apnoea: impact of ethnicity. *Respirology*. 2012;17(2):213–222.

ACKNOWLEDGMENTS

The authors acknowledge the Raine Study participants and their families for their ongoing participation in the study and the Raine Study staff for their dedicated commitment to coordination and data collection.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication May 14, 2019
Submitted in final revised form July 27, 2019
Accepted for publication July 30, 2019
Address correspondence to: Professor Peter Eastwood, Centre for Sleep Science, School of Human Sciences, University of Western Australia, 35 Stirling Highway, Perth, WA 6009; Email: Peter.Eastwood@uwa.edu.au

DISCLOSURE STATEMENT

All authors have seen and approved this manuscript. Work for this study was performed at the University of Western Australia. The Raine Study was supported by the National Health and Medical Research Council, with additional funding for core management provided by the University of Western Australia, Raine Medical Research Foundation, Telethon Kids Institute, University of Western Australia Faculty of Medicine, Dentistry, and Health Sciences, Women and Infants Research Foundation, Curtin University, The University of Notre Dame, Edith Cowan University and Murdoch University. The 22-year follow-up was supported by the National Health and Medical Research Council (1021858, 1027449, and 1044840). Peter Eastwood was supported by a National Health and Medical Research Council Senior Research Fellowship (No. 1042341). The authors report no conflicts of interest.

Downloaded from jcsm.aasm.org by 65.21.193.86 on February 24, 2022. For personal use only. No other uses without permission.
Copyright 2022 American Academy of Sleep Medicine. All rights reserved.