Predicting sleep apnea from 3-dimensional face photography

Article *in* Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine · August 2019

CITATIONS
0

READS
689

8 authors, including:

Peter R Eastwood
University of Western Australia
217 PUBLICATIONS **4,851** CITATIONS
SEE PROFILE

Syed Zulqarnain Gilani
University of Western Australia
31 PUBLICATIONS **186** CITATIONS
SEE PROFILE

Jen Walsh
Sir Charles Gairdner Hospital
42 PUBLICATIONS **1,365** CITATIONS
SEE PROFILE

Kathleen J Maddison
University of Western Australia
20 PUBLICATIONS **583** CITATIONS
SEE PROFILE

Predicting sleep apnea from 3-dimensional face photography

Authors: [1,2]Peter Eastwood*, [3,4]Syed Zulqarnain Gilani*, [1,2]Nigel McArdle, [1,2]David Hillman, [1,2]Jennifer Walsh, [1,2]Kathleen Maddison, [5]Mithran Goonewardene, [3]Ajmal Mian.

*joint first author

Affiliations:

[1]Centre for Sleep Science, School of Human Sciences, University of Western Australia

[2]West Australian Sleep Disorders Research, Sir Charles Gairdner Hospital, Western Australia

[3]School of Computer Science and Software Engineering, University of Western Australia

[4]School of Science, Edith Cowan University.

[5]Oral Development and Behavioural Sciences, University of Western Australia

Corresponding Author:

Professor Peter Eastwood

Centre for Sleep Science

School of Human Sciences

University of Western Australia

35 Stirling Highway, Nedlands, WA 6009

Peter.Eastwood@uwa.edu.au

ABSTRACT

Study Objectives Craniofacial anatomy is recognised as an important predisposing factor in the pathogenesis of obstructive sleep apnea (OSA). Two-dimensional (2D) photography has shown that craniofacial features are related to the presence and severity of OSA. Three-dimensional (3D) has potential advantages over 2D imaging as it overcomes the limitation of representing a 3D structure (the face) in two dimensions. This study aimed to use 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 severity of OSA.

Methods 400 adults underwent overnight polysomnography and 3D face photography. 100 did not have OSA (apnea-hypopnea index, AHI <5 events/hr), 100 had mild OSA (5≤AHI<15 events/hr), 100 had moderate OSA (15≤AHI<30 events/hr) and 100 had severe OSA (AHI≥ 30 events/hr). Measurements of linear distances and angles, and geodesic distances were obtained between anatomical landmarks from the 3D photographs and their relationship to the presence and severity of OSA determined.

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, $p<0.01$). A maximum classification accuracy of 91% was achieved when linear and geodesic craniofacial measurements were combined into a single predictive algorithm.

Conclusions This study showed that geodesic measurements add value to the capacity to identify patients with OSA from 3D photographs of the face. These preliminary results also suggest that it might also be possible to predict the severity of an individual's OSA from such photographs.

Keywords

Obstructive sleep apnea, three dimensional photography, linear discriminant analysis, craniofacial anatomy, 3dMD, polysomnography

Statement of Significance

Many individuals with obstructive sleep apnoea (OSA) are currently undiagnosed 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 and permit determination of the combination of measurements that best correlates with the presence and severity of OSA. 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 study suggested that it might also be possible to predict the severity of an individual's OSA from such photographs. While 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.

INTRODUCTION

Obstructive sleep apnea (OSA) is a common disorder estimated to affect 15% of middle-aged males and 5% of middle-aged females[1]e. It is characterized by repetitive episodes of partial or complete upper airway obstruction which 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, the majority (75%) of cases 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) non-specific, resulting in a high rate of false positives[7]. Other low cost screening tools are needed[8-11].

Craniofacial anatomy is recognised as an important predisposing factor in OSA pathogenesis[12]. Studies using MRI have shown that mid-face and lower-face width are correlated to 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 clinical assessment

and their use is limited by cost in the case of MRI and the risks associated with exposure to ionising 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 anatomical 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 non-linear 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. While, similar to 2D photography, it can provide information regarding linear distances and angles, unlike 2D photography it can additionally provide information on facial contours (*i.e.,* geodesic distances) making it an ideal tool for assessing the role of craniofacial structure in the pathogenesis of OSA. Three-dimensional 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.

The present 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 were 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 analysed, presence and severity of OSA were determined, and 50 participants without OSA (apnea hypopnea index (AHI) <5 events/hr) and with good quality 3D face images (*e.g.,* without beards and missing data points) were identified.

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 analysed, 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 (EEG), electrooculogram (EOG) and chin electromyogram (EMG) 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 (ECG), respectively. Leg movements were monitored by EMG 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 (5≤AHI<15 events/hr), moderate (15≤AHI<30 events/hr) or severe (AHI≥ 30 events/hr).

At the time of the sleep study measurements were also obtained of each participant's height (stadiometer) and weight, 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 degree (ear to ear) and neck region 3D images using the technique of triangulation[24,25]. Obtaining the image required the subject to sit on a chair between two cameras, with their hair pulled back from their face. High-resolution images were captured within 1.5 milliseconds and simultaneous acquisition of geometry and colour-texture data were achieved by the synchronisation 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. The software allowed the face to be rotated to allow optimal visualisation and determination of each landmark. Landmark location was highly repeatable: the average ICC 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.

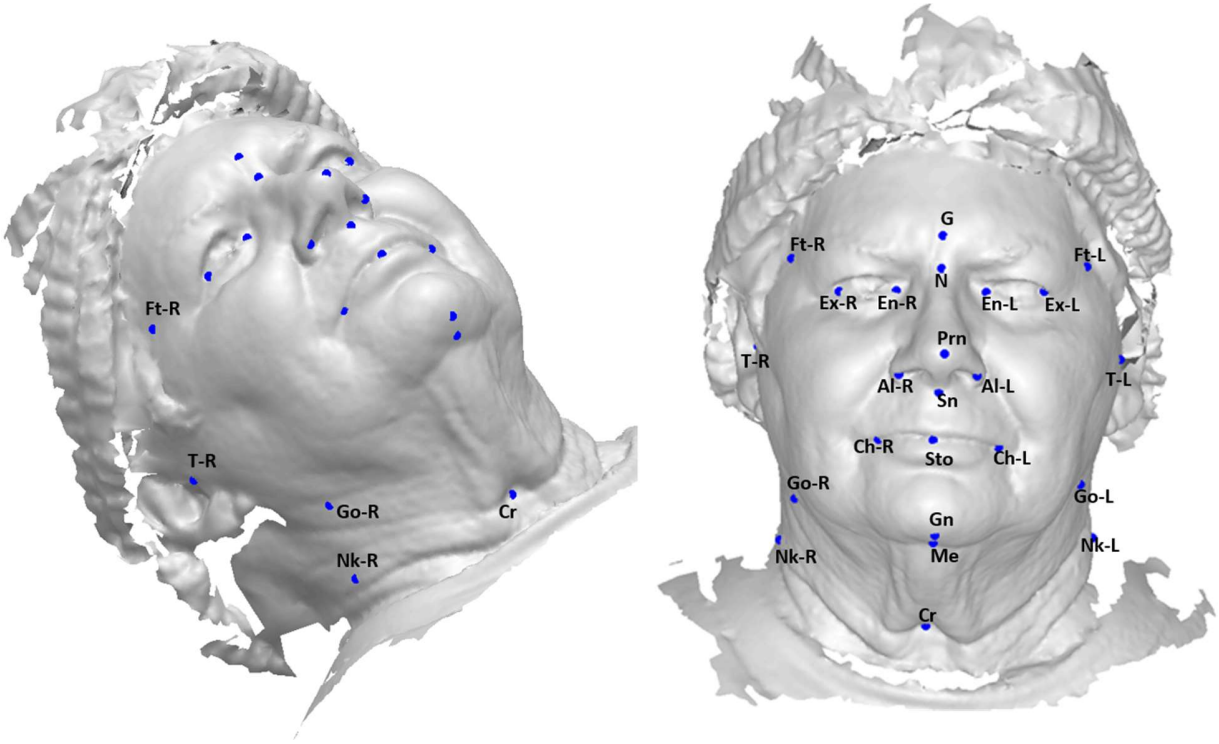
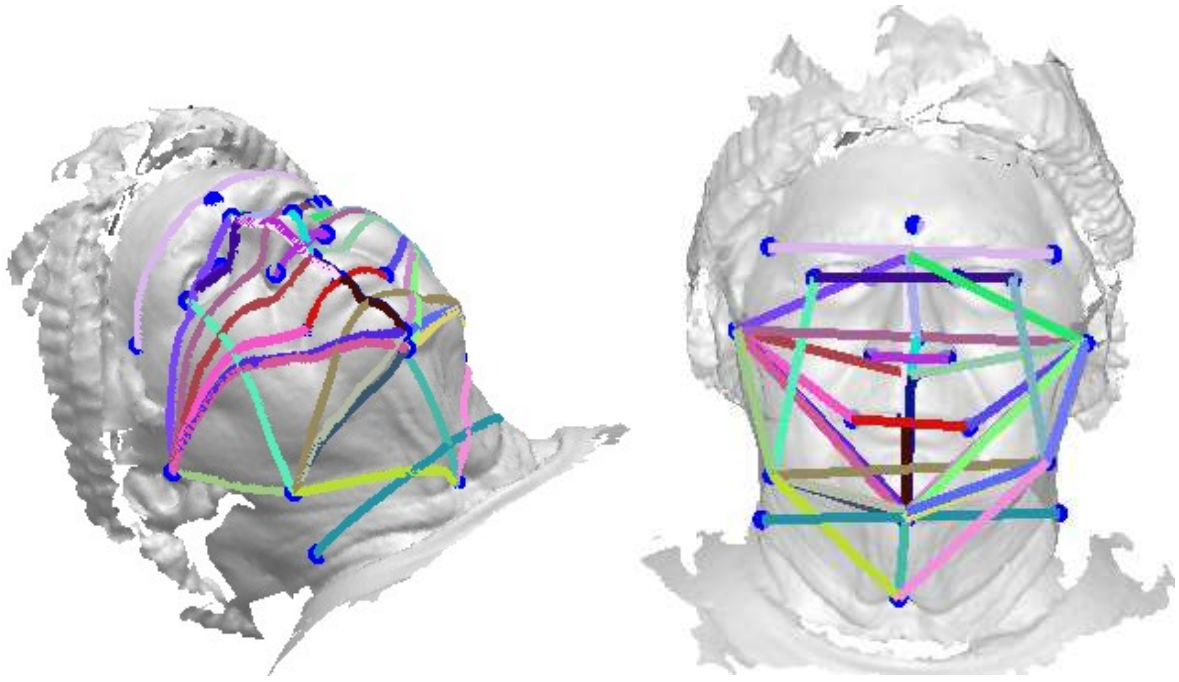


Figure 1. Annotated landmarks.
Al-L, Alare Left; *Al-R*, Alare Right; *Ch-L*, Chelion Left; *Ch-R*, Chelion Right; *Cr*, Cricoid; *En-L*, Endocanthion Left; *En-R*, Endocanthion Right; *Ex-L*, Exocanthion Left; *Ex-R*, Exocanthion Right; *Ft-L*, Frontotemporale Left; *Ft-R*, Frontotemporale Right; *G*, 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.

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] 25 pairs of points were selected and distances calculated for each in both the linear dimension (*i.e.,* direct Euclidean distance between the two points) and geodesic dimension (*i.e.,* the shortest distance between two points when following the contour of the face/skin) (Figure 2, Tables S1 and S2). Using linear dimensions only, angles were also determined between sets of three points (Figure S1, Table S3). These linear and geodesic distances and angles were termed 'features' and used in further analysis.



Geodesic Distances- profile view Euclidian Distances

Figure 2. Geodesic and Euclidian distances were determined between annotated landmarks.

Analysis

Data were divided into two classes, Controls and OSA, based on a polysomnography-derived threshold value of AHI≥5 events/hr 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 below) using AHI threshold values of 10, 15, 20, 25, 30, 35, 40, 45 and 50 events/hr.

The LDA algorithm was used to distinguish facial features between the two classes: Controls and OSA. The goal of this machine learning algorithm was to find, using all features, a one dimensional space where the distance between features 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 ten 'bins'. In each bin, 90% of the data (*e.g.*, features of 360 faces for a threshold of AHI=5) picked randomly were used to train the algorithm (*i.e.,* 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 (further details can be found in reference [27]). 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 Text S1 and S2.

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 (see supplementary Table 1 & 2).
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 (see supplementary Table 3).
7. Using the centre landmark of each 3-tuple as the centre vertex find the angle between the three landmarks of the 3-tuple. (see supplementary Figure 1)
8. Set threshold AHI<5 as Controls and AHI≥5 as OSA.
9. Create ten 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 (LDA) 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)

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 Operator 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 ANOVA with Tukey-Kramer corrections for multiple comparisons. All data are expressed as mean± standard deviation (SD) or mean±standard error (SE) and a p-value of 0.05 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, Text S2). Compared to those without OSA (AHI<5), those with OSA (AHI≥5) were older (47.4 ± 13.8 vs 54.5±15.8 yrs, $p<0.05$) and had increased AHI (2.6±1.28 vs 30.7±26.0 events/hr, $p<0.05$), BMI (26.9±5.2 vs 32.3±7.6 kg/m^2, $p<0.05$) and neck circumference (35.0±3.8 vs 39.7±4.8 cm, $p<0.05$) (Table 1).

Table 2. Subject Characteristics

OSA Severity	**Subjects**			**Age (yrs)**	**BMI (kg.m^{-2})**	**NC* (cm)**	**AHI (events.hr^{-1})**
	Males	**Females**	**Total**				
Controls (AHI<5)	25	75	100	47.4 ± 13.8	26.9 ± 5.2	35.0 ± 3.8	2.6 ± 1.3
Mild (5≤AHI<15)	39	61	100	50.2 ± 16.7	29.2 ± 6.8	37.7 ± 4.5	9.9 ± 2.9
Moderate (15≤AHI<30)	50	50	100	55.2 ± 14.3	32.3 ± 7.2	40.3 ± 4.3	22.6 ± 4.6
Severe (AHI ≥ 30)	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

BMI, body mass index; NC, Neck circumference; AHI, Apnea-hypopnea index. *Note that NC was obtained from only 322 participants (95 Controls; 84 Mild; 77 Moderate; 66 Severe). Values are Mean +SD.

The accuracy with which facial measurements, BMI and neck circumference could classify an individual as having OSA (AHI≥5) or not (AHI<5) is summarised in Table 3 and shown graphically in Figure 3 (the algorithm is reported in supplementary Text S3). The predictive accuracy of geodesic measurements alone was significantly greater than that for linear distances alone, being 89±0.01% and 86±0.01%, respectively ($p<0.01$). 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 91±0.02%) for measurements incorporating a combination of linear and geodesic distances.

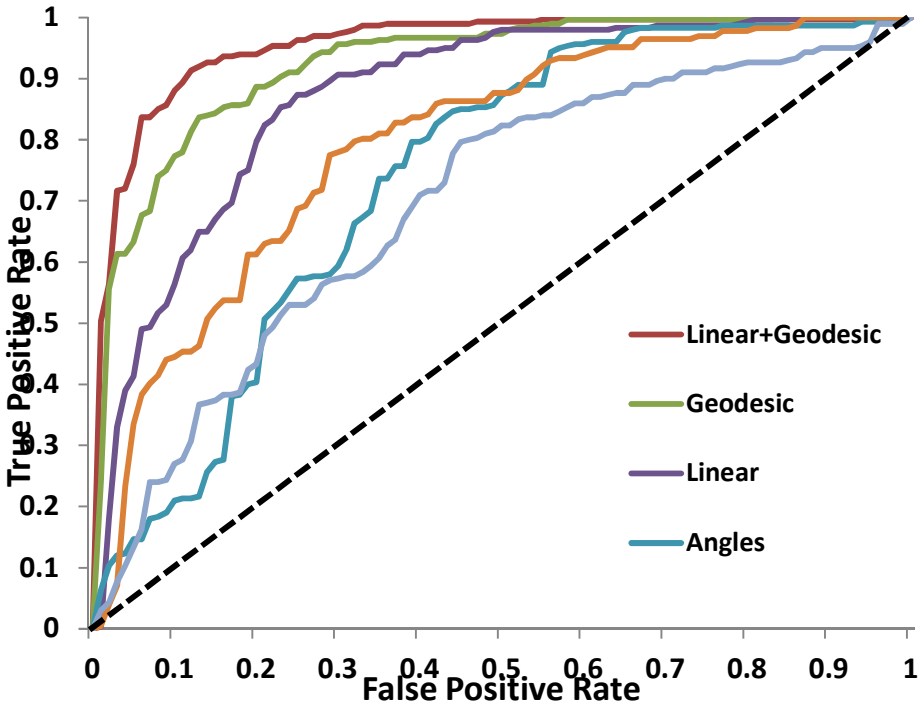


Figure 3. Receiver Operator Characteristic (ROC) curves showing the sensitivity (true positive rate) and specificity (false positive rate) for predicting Obstructive Sleep Apnea (AHI≥5) using

different landmarks or anthropometric measurements. NC, Neck circumference; BMI, Body Mass Index.

Table 3. Classification accuracy of individuals without OSA (AHI<5) and with OSA (AHI≥5) 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
Neck circumference	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

BMI, Body Mass Index; ROC, Receiver Operator Characteristic. Values are Mean ± SE.

Differences in landmark distances between those with and without OSA (based on a cut-off of 5 events/hr) are shown for geodesic distances, linear distances and angles in Tables S1, S2, and S3. Craniofacial features that differed between controls and OSA, for both geodesic and linear distances included: upper, mid and lower face depth; face width; mandibular length, width and posterior height; and neck width ($p<0.05$ for all). The most discriminatory angular measurements between the two groups (based on magnitude of difference and p-values) 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) ($p<0.05$ 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 (Tables S4 and S5). These included: upper and lower face depth; total and upper face height; lateral face height; face width; mandible width; and neck width ($p<0.05$ for all). Angular measurements tended to be less discriminatory between controls and different severities of OSA (Table S6).

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

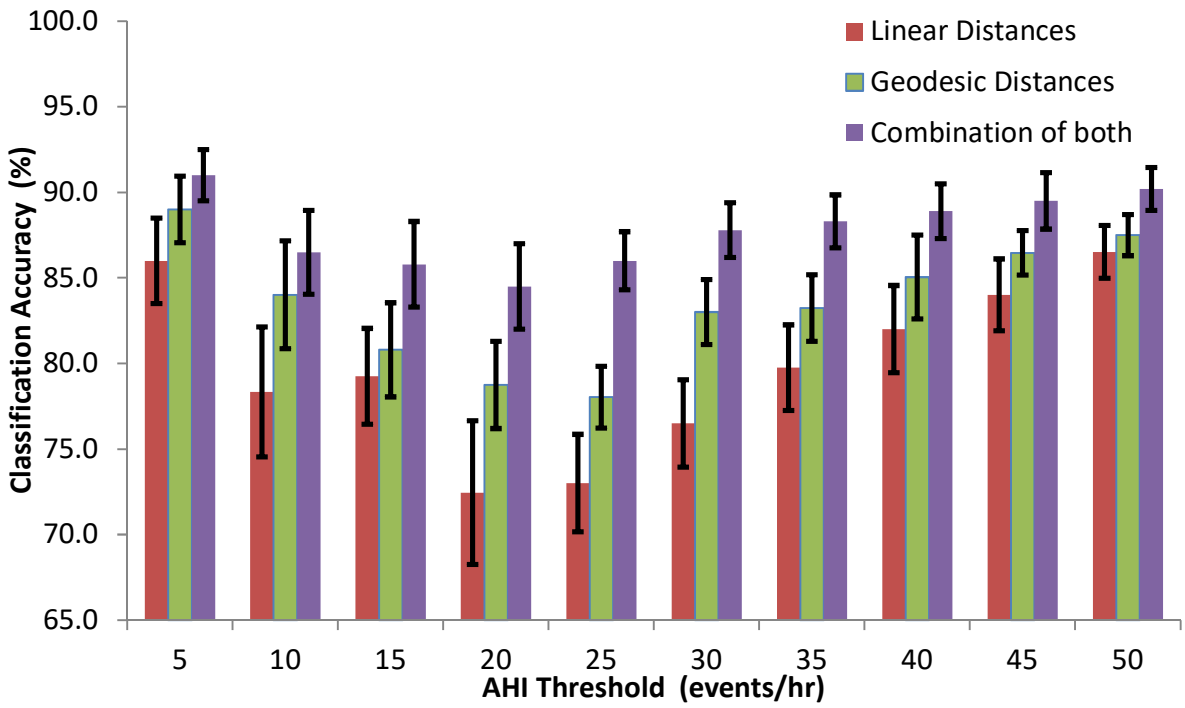


Figure 4. Classification accuracy for predicting Obstructive Sleep Apnea (OSA) when OSA was the algorithm was trained and tested at different AHI thresholds (between 5 and 50 events/hr) using linear distances (red column), geodesic distances (green bar) and combined geodesic and linear distances (purple bar). AHI, apnea hypopnea index. Error bars, ±SE.

DISCUSSION

The present 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.

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 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[12-15,26,28-30]. The techniques used in these previous studies have included MRI, CT, lateral cephalometry and 2D photography. While 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 ionising 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 non-linear nature of craniofacial anatomy, such as shape and contour.

Three-dimensional (3D) photography overcomes the limitation of representing the 3D structure of the face with 2D imaging as 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 equivalent measurements obtained from 3D CT[20]. The present 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/hr) based on this combination was 96%, 76%, 91% and 0.96, respectively.

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 76-79% of the cases, with sensitivities ranging from 73-85%, specificities from 28 to 70% and area under the ROC curve from 0.73 to 0.87[15,26,31]. When comparing between studies and between techniques an important consideration is the AHI cut-off used to define the presence or absence of OSA. In all of these previous studies a cut-off of 10 events/hr was used, whereas the present study used a cut-off of 5 events/hr. It was interesting to note that, in the present study, the sensitivity, specificity, accuracy and area under the ROC curve decreased to 90%, 76%, 86% and 0.92, respectively, when an AHI threshold of 10 events/hr 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/hr 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 controls and OSA tended to change similarly for both geodesic and linear distances, such that individuals with OSA were characterised by a 3-10% increase in measures of face depth; 10-13% increase in lateral face height; 5% increase in face width, 2-6% increase in mandibular length; 17-19% increase in posterior mandibular height; 4-5% increase in mandible width; and 10-13% increase in neck width. In terms of angular measurements, individuals with OSA had an 18% smaller maxillary-mandibular relationship angle, and a 10-12% 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 (*i.e.* 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 (*e.g.,* mandibular width, neck perimeter, mandibular length, facial width, binocular width) were similar to those that distinguished groups in the present study.

The data presented in this study suggests 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 0.97, respectively, the specificity was 0.76 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 the majority of cases are identified) and at least moderate specificity to ensure against too many false positives.

However, they are inadequate for a diagnostic test which must be both highly sensitive and highly specific. In their review of non-laboratory based devices to diagnose OSA, using an in-laboratory cut-off of 5 events/hr, Collop et al[32] 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 pre-test probability based on clinical assessment[33] 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 control, mild, moderate and severe OSA subjects, 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. While 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.

In conclusion, the main finding of this study was that geodesic measurements add value to the capacity to identify patients with OSA from 3D photographs of the face. These results suggest that it might also be possible to predict the severity of an individual's OSA from such photographs.

ACKNOWLEGEMENTS

The authors would like to 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. 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, and Edith Cowan 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). There were no conflicts of interest reported by the authors.

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. Teran-Santos J, Jimenez-Gomez 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.
22. Iber C A-IS, Chesson A, Quan S; for the American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. In. 1st ed: Westchester, IL: American Academy of Sleep Medicine, 2007.; 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. Balakrishnama S, Ganapathiraju A. Institute for Signal and Information Processing Linear Discriminant Analysis - a Brief Tutorial. 1998.
28. 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.
29. Remya KJ, Mathangi K, Mathangi DC, et al. Predictive value of craniofacial and anthropometric measures in obstructive sleep apnea (OSA). Cranio : the journal of craniomandibular practice. 2017; 35 (3): 162-167.
30. Cheung K, Ishman SL, Benke JR, et al. Prediction of obstructive sleep apnea using visual photographic analysis. J Clin Anesth. 2016; 32: 40-46.
31. 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. Computational and mathematical methods in medicine. 2015; 2015: 489761.
32. 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.
33. 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.