ABSTRACT

INTRODUCTION

The South African Police Service frequently relies on craniofacial approximation and superimposition to assist in identifying unknown deceased individuals. Standards to estimate lip height are however limited. Findings from this study share medical applications.

Aims and objectives:
Establish reliable standards for estimating lip height using dentoskeletal measurements.

Methods
Cone-beam CTs comprising 124 black and 39 white southern African adults were assessed. A series of dimensions were recorded using a DICOM viewer with an in-built measuring tool. Relationships between hard tissue structures (maxillary, mandibular and total central incisor heights, their corresponding root lengths, face height (N-Gn), and nose height (N-Sn)) and respective overlying soft tissues (upper, lower and total lip heights) were evaluated.

Results and conclusions
Statistically significant differences were observed between population, sex and age groups. A selection of regression equations to estimate lip height was calculated that included population, sex and approximate age (20-39 and 40+ years) for improved goodness-of-fit ($r^2$-value). Regression models using face height produced the strongest multiple correlation ($r$-value) and goodness-of-fit ($r^2$-value). Validation testing indicated that regression models often improved upon mean measurements, while offering a degree of individuality that mean values do not.

Keywords
Facial anthropology, craniofacial identification, craniofacial approximation, craniofacial superimposition, mouth morphology.

INTRODUCTION

High rates of violent deaths, illegal immigration and internal migration, including a shortage of identification documents, result in a high incidence of unidentified deceased in the South African medico-legal system. Furthermore, families may be unaware of missing or dead loved ones due to infrequent communication in poorer communities. The unidentified deceased presents a growing humanitarian and legal strain on the country, as only after formal identification can a police inquest or criminal case progress, family be notified, and other legal, religious and cultural requirements be addressed.1

Limited dental and DNA records are available for comparison, leading the South African Police Service (SAPS) to rely on craniofacial approximation (CFA) and craniofacial superimposition to assist in identifying possible matches. CFA recreates the likeness of an individual’s face from the features of their skull.2-4 Craniofacial superimposition overlays a number of antemortem images of a missing person with an unidentified skull, to assess their structural similarity.1-7

Although this paper focusses on forensic applications of the relationship between the skull and the overlying soft tissues, this relationship is relevant to several medical fields. Insight from reference data about hard and soft tissue associations could benefit surgeons and orthodontists treating dentofacial deformities; for whom obtaining harmonious facial characteristics and functionality are important considerations during diagnosis and treatment.5-10

There is an absence of appropriate standards for estimating lip height for southern African or even Sub-Saharan African individuals. Generally, little is known
about the placement of the mouth and lips during CFA, and yet they play a key role in the evaluation and recognition of the craniofacial complex. Previous anthropometric research on the mouth originated predominantly from the fields of dentistry and maxillofacial surgery.\(^5\)\(^,\)\(^1\)\(^,\)\(^1\)\(^5\)\(^,\)\(^1\)\(^5\) None of these studies were, however, performed in South Africa.

Lip height and mouth width dimensions were measured by Farkas et al.\(^1\)\(^5\) using sliding callipers on a North American white sample, and Ferrario et al.\(^1\)\(^5\) using optoelectronic equipment on a North Italian sample. They identified that males generally have greater oral dimensions than females, which is in agreement with three-dimensional studies on Czech, German, North American white and North American Latino individuals.\(^1\)\(^9\)\(^,\)\(^1\)\(^4\) However, no sexual dimorphism was identified by Wilkinson et al.\(^4\) on a European and Indian subcontinent population using direct caliper and photographic measurements, or by Ferrario et al.\(^1\)\(^1\) on a North Italian population using photographic measurements.

Lip height alters with age and can vary between population groups.\(^4\)\(^,\)\(^1\)\(^6\) In a study using a white Italian sample, Sforza et al. found that the vermilion surface area and height of the upper and lower lips progressively increased during juvenile development until late adolescence, and then decreased with ageing.\(^8\) The vermilion height to mouth width ratio was larger in females than in males, and decreased with age, the total lip height and lip volumes were however significantly larger in males than in females. In a study on an adult South African male sample by Schmidlin et al.,\(^1\) the upper and lower lip dimensions was found to similarly reduce with age, but the lip height dimensions were consistently greater than in Sforza et al.'s European sample.

Research focussing on the direct relationship between the lips and skull has been previously performed. Early observations by Gerasimov\(^1\)\(^6\) related lip height to the projection of the maxillary and mandibular incisors and their corresponding alveolar sockets. He associated small straight teeth to thin lips and orthognathism, and big prominent teeth to thick lips and prognathism. Angel\(^1\)\(^8\) agreed that lip height depended on the projection of the teeth, emphasising the impact of ancestry, and the strength of the incisive and buccinator muscles.

Gerasimov\(^1\)\(^6\), Gatif and Snow\(^1\)\(^9\) and Taylor\(^2\)\(^0\) all directly affiliated the vertical thickness of the mouth (referring to the thickest, pigmented part in the middle portion of the lips) to be equal to the vertical distance of the central incisors (from the upper cementoenamel junction to the lower cementoenamel junction).

George\(^2\)\(^1\), however, alternatively indicated that the upper lip is positioned parallel to the upper quarter mark of the maxillary central incisor, and the lower lip positioned parallel to the lower three-quarter mark of the mandibular central incisor. These existing methods for estimating lip height were, however, based on a predominantly European sample.

Wilkinson et al.\(^4\) constructed the only existing equation for calculating lip height using maximum height measurements of central incisors, with independent equations for European and Indian subcontinent individuals. These methods of lip height estimation (by Wilkinson et al.\(^4\), Taylor\(^2\)\(^0\) and George\(^2\)\(^1\)), have subsequently been tested on a central European sample by Malá and Velemínská.\(^2\) Wilkinson et al.'s\(^2\) equations offered the most accurate result for estimating upper and lower lip height (1.3 mm and 1.8 mm mean absolute error for upper and lower lip estimation, respectively). George\(^2\)\(^1\) was the most accurate for determining total lip thickness (3.4 mm mean absolute error).

This study aimed to develop reliable standards for estimating lip height using dento-skeletal measurements taken from a southern African sample. Statistical differences were assessed between population, sex and age groups, in order to determine whether universal or specific formulae should be used.

Relationships between dento-skeletal measurements and lip height were determined. The goodness-of-fit achieved using the available dento-skeletal measurements as compared to a mean model were also assessed.

**MATERIALS AND METHODS**

This study was retrospective, using clinical cone-beam computerised tomography (CBCT) scans of 124 black (72 male, 52 female; mean age 35 years) and 39 white (19 males, 20 females; mean age 36 years) southern African adults. Data were collected from the University of Pretoria, Oral and Dental Hospital (ethics clearance number: 212/2016).

Subjects were aged between 20 and 87 years, displayed neutral facial expressions and moderate bodyweight. Subjects with the following characteristics were excluded from the study: intrusive craniofacial trauma, congenital anomalies, extensive tooth loss, or surgery impeding the basic craniofacial appearance, particularly of the mouth.

Comparative hard and soft tissue dimensions were collected in OsiriX (DICOM viewer) (Figure 1, Table 1). Measurements included upper, lower and total lip vermilion height (1-3); maxillary, mandibular and total height of the central incisor crowns (A-C); corresponding root lengths of the incisors (D-E); skeletal nose height (N-Sn; F) and face height (N-Gn; G). Lateral dimensions (i.e. central incisor heights and root lengths) were taken from both left and right views.

One hundred subjects underwent three intra-observer repeat measurements following 7-day intervals, and 30 individuals were remeasured by another observer. Repeat measurements utilised the original set of CBCT scans, thus purely representing observer measurement error.

Detecting possible technological and biological variables/ errors using repeat scanning techniques was not feasible due to the retrospective nature and ethical considerations of this study. We are also conscious of the limited white sample, which is representative to the frequency of white patients visiting the dental hospital. The white data were maintained in this study to form a descriptive comparison for black and white southern Africans.
Repeatability was calculated using technical error of measurement (TEM) and relative TEM (r-TEM), using Microsoft Excel 2011 version 14.4.1. TEM was performed according to Dahlberg\textsuperscript{23} and Perini et al.\textsuperscript{24} for interobserver tests incorporating two observational repeats, and Langley et al.\textsuperscript{25} for intra-observer tests incorporating three observational repeat measurements.

The statistical package SPSS version 25 was used to analyse all other results. Wilcoxon signed rank tests were conducted to identify whether significant differences exist between the left and right maxillary and mandibular central incisors, and corresponding roots. The Wilcoxon signed rank test was also used to compare the central incisors and corresponding lip heights (to investigate Gerasimov\textsuperscript{16}, Gatiff and Snow\textsuperscript{19} and Taylor's\textsuperscript{20} theory for lip height estimation); the upper and lower lip heights were similarly compared. Differences between population and sex groups for the various dimensions were studied using Mann-Whitney U tests.

Age variation was investigated using both Kruskal Wallis H tests (assessing age groups by decades; 20-29 years, 30-39 years, etc.) and Mann-Whitney U tests (assessing two distinctly different age groups, as determined by the Kruskal Wallis H test). The influence of ancestry on the sample size and distribution was assessed by comparing both total and equal sample sizes for each test.

To determine if the collected hard tissue measurements can predict the associated soft tissue measurements, regression analyses were conducted. The independent variables (i.e. skeletal measurements, sex, age and population) that may contribute to the regression models were checked for inter-correlation with one another, using Spearman’s rank and point-biserial correlation tests.

CBCT scans have been proven to relate to the craniofacial complex with a high level of accuracy, thus the generated regression models did not require any subsequent adjustment to represent reality.\textsuperscript{26,27}

Regression models underwent validation testing using an independent hold-out sample comprising 30 cases. The composition of the hold-out sample is shown in Figure 2. Validation testing was only performed with a black sample, due to a shortage of available white individuals. The available sample was also weighted towards males and younger individuals.

Figure 1. Comparative hard and soft tissue dimensions taken for investigation:
1. Upper lip height.
2. Lower lip height.
3. Total occluded lips height.
A. Maxillary central incisor crown height.
B. Mandibular central incisor crown height.
C. Total occluded central incisors height.
D. Maxillary central incisor root length.
E. Mandibular central incisor root length.
F. Nose height.
G. Face height.

Figure 2. Distribution of sex and age groups within a black population - specific hold-out sample used for validation testing. Sample sizes vary according to the different dimensions obtainable within available scans.

Regression models underwent validation testing using an independent hold-out sample comprising 30 cases. The composition of the hold-out sample is shown in Figure 2. Validation testing was only performed with a black sample, due to a shortage of available white individuals. The available sample was also weighted towards males and younger individuals.

Although limited samples were available, we still found it necessary to check the efficacy of our research using a completely independent set of data; this has been proven to offer the most rigorous method of testing.\textsuperscript{26,29}

We encourage more independent validation testing in the future. Regression models were compared to the mean heights of upper, lower and total lips; population, sex and age were considered for all models and mean heights.
TABLE 1. Definitions of measurements taken.

| Measurement                        | Description                                                                                                                                 |
|-----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Upper lip vermilion height        | Taken from the point denoting the vermilion border of the upper lip in the midsagittal plane, to the most anterior point of contact between the upper and lower lips. (Upper lip - Stomion, Ls-St) |
| Lower lip vermilion height        | Taken from the most anterior point of contact between the upper and lower lips to the point denoting the vermilion border of the lower lip in the midsagittal plane. (Stomion - Lower lip, St-Li) |
| Total lip vermilion height        | Taken from the point denoting the vermilion border of the upper lip and lower lip in the midsagittal plane. (Upper lip - Lower lip, Ls-Li) |
| Maxillary central incisor root length | Longest apicocoronal distance, parallel to the long axis, between the most apical point of the maxillary central incisor root length. |
| Mandibular central incisor root length | Longest apicocoronal distance, parallel to the long axis, between the most apical point of the mandibular central incisor root length. |
| Total central incisor root length | Longest apicocoronal distance, parallel to the long axis, between the most apical point of the maxillary and mandibular central incisor root length. |
| Face height                       | Taken from the midline bony depression between the eyes and just below the glabella, where the frontal and two nasal bones meet, to the midpoint of the lower border of the mandible. (Nasion - Gnathion, N-Gn) |
| Nose height                       | Taken from the midline bony depression between the eyes and just below the glabella, where the frontal and two nasal bones meet, to the midline point at the base of the nasal spine. (Nasion - Subnasale, N-Sn) |

RESULTS

Adequate repeatability was obtained from intra- and inter-observer tests (Table 2). Mean intra-observer r-TEM was 1.85%; ranging from 0.69% (for nose height) and 0.70% (face height) to 2.76/3.03% (left/right mandibular central incisor root lengths).

Mean inter-observer r-TEM was 2.26%; ranging from 0.73% (nose height) and 0.74% (face height) to 4.05/4.06% (left/right mandibular central incisor root lengths).

The r-TEM percentages represent the level of deviation experienced between measurements, with the ideal result being zero. The given results could be a consequence of the relative size difference experienced between face height and nose height, in contrast to root length (i.e. 1 mm makes less difference to the greater dimensions experienced in face/nose height, compared to the smaller dimensions witnessed for root length). The comparatively limited reliability witnessed for root length measurements was however expected, due to the challenges experienced when attempting to distinguish the deep margins of the root in CBCT scans.

TABLE 2. Intra- and inter-observer errors for FSTT measured in this study.

| Dimensions                        | Intra-observer (n = 100; n* = 65, 3 repeat measurements) | Inter-observer (n = 30, 2 repeat measurements) |
|-----------------------------------|-----------------------------------------------------------|-----------------------------------------------|
|                                  | TEM (mm) | r-TEM (%) | Reliability | TEM (mm) | r-TEM (%) | Reliability |
| Upper Lip Height                 | 0.29     | 2.14      | 0.96        | 0.30     | 2.16      | 0.98        |
| Lower Lip Height                 | 0.28     | 2.21      | 0.96        | 0.34     | 2.65      | 0.97        |
| Total Lip Height                 | 0.55     | 2.20      | 0.98        | 0.67     | 2.21      | 0.98        |
| Maxillary Incisor Height (left)  | 0.13     | 1.23      | 0.98        | 0.15     | 1.34      | 0.98        |
| Maxillary Incisor Height (right) | 0.14     | 1.35      | 0.98        | 0.16     | 1.46      | 0.97        |
| Mandibular Incisor Height (left) | 0.13     | 1.50      | 0.98        | 0.15     | 1.71      | 0.97        |
| Mandibular Incisor Height (right)| 0.13     | 1.52      | 0.98        | 0.19     | 2.09      | 0.96        |
| Total Occluded Teeth Height      | 0.35     | 1.86      | 0.98        | 0.47     | 2.45      | 0.97        |
| Maxillary incisor root length (left) | 0.26     | 2.34      | 0.95        | 0.32     | 2.67      | 0.92        |
| Maxillary incisor root length (right) | 0.27     | 2.43      | 0.94        | 0.39     | 3.31      | 0.89        |
| Mandibular incisor root length (left) | 0.25     | 2.76      | 0.95        | 0.39     | 4.05      | 0.93        |
| Mandibular incisor root length (right) | 0.28     | 3.03      | 0.94        | 0.39     | 4.06      | 0.91        |
| Face Height*                     | 0.84     | 0.70      | 0.99        | 0.87     | 0.73      | 0.98        |
| Nose Height*                     | 0.35     | 0.69      | 0.99        | 0.37     | 0.74      | 0.99        |
| Minimum                          | 0.13     | 0.69      | 0.94        | 0.15     | 0.73      | 0.89        |
| Maximum                          | 0.84     | 3.03      | 0.99        | 0.87     | 4.06      | 0.99        |
| Mean                             | 0.30     | 1.85      | 0.97        | 0.36     | 2.26      | 0.96        |
The performance was however acceptable for a cautionary continuation of the investigation.

No significant differences between the left and right central maxillary/mandibular incisors/roots heights were found, and therefore only bony dimensions from the left side were used for further analyses. Utilising measurements taken from the left side is furthermore supported by the moderately improved performance of repeat measurements taken between the left and right aspects.

The calculated mean and standard deviation (SD) values for the retained dimensions are shown in Table 3.

Mann-Whitney U tests (Table 4) identified significant population differences in all soft tissue measurements.

Black individuals on average presented with a greater mean upper lip, lower lip, and total lip height compared to the white individuals. White individuals, however, had a significantly greater hard tissue nose height compared to the black individuals.

| Table 3. Mean (standard deviation) values for recorded lips, central incisors, roots, face (N-Gn) and nose (N-Sn) height dimensions, in a black and white southern African sample. |
|---|
| **Table 4** |
| n | Mean age (years) | Upper lip (mm) | Lower lip (mm) | Maxillary central incisor (left) (mm) | Mandibular central incisor (left) (mm) | Maxillary root (left) (mm) | Mandibular root (left) (mm) |
|---|---|---|---|---|---|---|---|
| Black | Female <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Male | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Total | Male | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| White | Female <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Male | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Total | Male | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Black | Total | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| White | Total | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Male | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Total | Male | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Black | Total | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| White | Total | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Male | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | <40 years | 39 (11) | 13.07 (3.52) | 12.70 (2.84) | 13.73 (2.79) | 19.31 (1.79) | 20.14 (1.97) |
| Total | Male | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Female | >40 years | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
| Total | Female | 11 (5) | 12.50 (3.10) | 12.40 (2.88) | 13.36 (2.77) | 19.17 (1.79) | 20.04 (1.97) |
Sexual dimorphism was more evident in black individuals than in white individuals (Tables 3 and 4). Black males had significantly greater mean dimensions compared to black females in the soft tissue structures (upper lip and total lip height) and hard tissues (maxillary and mandibular central incisor heights, maxillary central incisor root length). White males only presented a significantly greater face height and nose height compared to white females. The lips, incisors and roots were found to differ between the age groups 20-39 years (<40 years) and 40 years and above (≥40 years), with a notable decline in their dimensions with the progression in age (Table 5). The effect of age was most evident in the black sample, especially affecting black females (Tables 3 and 5).

The sample composition that included greater age variation for older black females (≥40 group mean age = 53 years, 13 years SD) compared to black males (≥40 group mean age = 50 years, 6 years SD) could have influenced this result. Black individuals typically presented with much greater lip heights than white individuals, but there was a similar decline in their dimension with age. Upper lip heights in the ≥40 years category were, on average, 1.86 mm and 1.03 mm thinner in black individuals compared to white individuals in the soft tissue structures (upper lip and total lip height). There was a similar 1.84 mm and 1.03 mm thinner in black individuals compared to white individuals in the soft tissue structures (upper lip and total lip height).

### Table 4. Mann-Whitney U results comparing, central incisors, roots, face (N-Gn) and nose (N-Sn) height dimensions in the various population and sex groups. The “Total” column presents results using all available data; the “Equal” column tests a uniform sample size.

| Population | Black Sex | Total | Equal | White Sex | Total | Equal |
|------------|-----------|-------|-------|-----------|-------|-------|
| n          | B:124, W:39 | B:39, W:39 | M:91, F:72 | M:72, F:72 | M:72, F:72 | M:52, F:52 | M:19, F:20 | M:19, F:19 |
| Upper lip  | 190.0**   | 106.0** | 2628.0* | 2190.0* | 1435.0* | 1017.0* | 232.5* | 227.0 |
| Lower lip  | 278.0**   | 127.0** | 2714.0* | 2210.0* | 1578.0* | 1092.5* | 205.0 | 204.0 |
| Maxillary incisor | 2116.0* | 704.0* | 224.0* | 1668.0* | 124.0* | 824.0* | 191.0 | 140.0 |
| Mandibular incisor | 2562.0* | 712.0* | 2317.0* | 1829.0* | 1203.0* | 912.0* | 171.0 | 155.0 |
| Maxillary root | 2977.0* | 947.0* | 3132.0 | 2355.0 | 1794.0 | 1293.0 | 159.0 | 152.0 |
| Mandibular root | 701.0* | 497.0* | 363.0* | 257.0 | 163.0* | 96.0 | 70.0 | 68.0 |
| Total lips | 672.0* | 94.0* | 1620.0 | 940.0 | 625.0 | 232.5 | 228.0 |
| Total incisors | 1388.5* | 597.0 | 1600.0* | 853.5* | 677.5 | 142.5 | 138.5 |
| Total incisors (Max. + Man.) | 1806.0 | 704.0 | 1354.0* | 1130.0* | 621.5* | 487.5* | 141.0 | 133.0 |
| Total roots (Max. + Man.) | 2226.0 | 916.5 | 1998.0 | 1538.0 | 779.0 | 234.5 | 193.0 | 147.0 |
| n          | B:63, W:39 | B:39, W:39 | M:31, F:43 | M:39, F:43 | M:42, F:42 | M:19, F:20 | M:19, F:19 |
| Face height | 968.0* | 623.0 | 482.5* | 322.5* | 218.0* | 109.0* | 62.5* | 59.5* |
| Nose height | 1844.0* | 1150.0* | 771.5* | 477.5* | 225.5* | 133.0* | 73.0* | 71.0* |

Key: B = Black, W = White, M = Male, F = Female, Max. + Man. = Maxillary + Mandibular

* value is significant at the 95% significance level (p < 0.05)
** value is significant at the 99% significance level (p < 0.01)

### Table 5. Significant Mann-Whitney U (test statistic) results, comparing individuals aged <40 years and ≥40 years, for height dimensions specific to lips, central incisors and roots. The “Total” column presents results using all available data; the “Equal” column tests a uniform sample size.

| Population | Black Males | Black Females | White Males | White Females |
|------------|-------------|---------------|-------------|---------------|
| n          | <40y:115    | <40y:48       | <40y:33     | <40y:33       |
|           | ≥40y:48     | ≥40y:33       | ≥40y:33     | ≥40y:33       |
| Upper lip  | 1663.0**    | 908.0        | 851.0*      | 381.5*        |
| Lower lip  | 1419.0**    | 725.0*       | 783.5*      | 293.5*        |
| Maxillary incisor | 1898.5* | 745.0* | 975.0* | 312.0* |
| Mandibular incisor | 1812.0* | 792.0* | 963.5* | 348.0* |
| Maxillary root | 2238.5 | 909.5 | 981.0* | 328.0* |
| Mandibular root | 1890.5** | 708.0* | 786.5* | 168.0 |
| Total lips | <40y:96     | <40y:43       | <40y:28     | <40y:28       |
|           | ≥40y:43     | ≥40y:28       | ≥40y:28     | ≥40y:28       |
| Total incisors (Max. + Man.) | 1537.5* | 773.0 | 735.0* | 325.0 |

Key: Max. + Man. = Maxillary + Mandibular

* value is significant at the 95% significance level (p < 0.05)
** value is significant at the 99% significance level (p < 0.01)
Maxillary incisor heights in the ≥40 category were, on average, smaller by 0.67 mm and 0.25 mm in black and white individuals respectively, and mandibular incisor heights, by 0.51 mm and 0.52 mm in black and white individuals respectively. Changes in root lengths were

Table 6. Best performing multiple linear regression equations for estimating upper, lower and total lip height (ordered according to their statistical strength), with validation test results (absolute mean inaccuracy represents the average result for the hold-out sample tested, maximum inaccuracy is the greatest inaccuracy demonstrated for a single individual).

| Description                        | Regression Equation                                                                 | F statistic, P-value | r²-value | Adjusted r²-value | SEE (mm) | Test sample (n) | Inaccuracy (mm) |
|------------------------------------|--------------------------------------------------------------------------------------|----------------------|----------|--------------------|----------|----------------|-----------------|
| Upper lip height                   | = 5.49 + (0.12 x Face Height) - (4.98 x Population) + (0.92 x Sex) - (1.96 x Age Group) | F4, 99 = 71.159, p<0.001 | 0.861    | 0.742              | 0.732    | 1.70           | 30              | 0.34             | 1.65 | 5.44 |
|                                    | = 20.22 + (0.02 x Nose Height) - (5.22 x Population) - (1.78 x Age Group)            | F3, 100 = 77.427, p<0.001 | 0.829    | 0.688              | 0.678    | 1.86           | 30              | 0.44             | 1.79 | 4.91 |
|                                    | = 15.16 + (0.45 x Maxillary Incisor Height) - (4.90 x Population) - (1.38 x Age Group) | F3, 159 = 78.848, p<0.001 | 0.773    | 0.598              | 0.590    | 1.96           | 30              | 1.05             | 1.95 | 5.67 |
|                                    | = 16.70 + (0.32 x Maxillary Root Length) - (0.00 x Population) - (1.48 x Age Group)  | F3, 159 = 77.514, p<0.001 | 0.771    | 0.594              | 0.586    | 1.97           | 30              | 0.99             | 1.91 | 5.51 |
| Lower lip height                   | = 12.06 + (0.07 x Face Height) - (4.36 x Population) - (2.63 x Age Group)            | F4, 100 = 74.236, p<0.001 | 0.831    | 0.690              | 0.681    | 1.81           | 30              | 0.24             | 1.42 | 4.41 |
|                                    | = 19.06 + (0.03 x Nose Height) - (4.63 x Population) - (0.57 x Age Group)            | F3, 100 = 66.354, p<0.001 | 0.816    | 0.666              | 0.656    | 1.88           | 30              | 0.51             | 1.56 | 4.23 |
|                                    | = 15.96 + (0.41 x Mandibular Incisor Height) - (4.52 x Population) - (1.92 x Age Group) | F3, 159 = 77.851, p<0.001 | 0.771    | 0.595              | 0.587    | 1.93           | 30              | 0.25             | 1.60 | 4.17 |
|                                    | = 19.36 + (0.06 x Mandibular Root Length) - (4.64 x Population) - (2.09 x Age Group) | F3, 159 = 74.444, p<0.001 | 0.763    | 0.582              | 0.574    | 1.96           | 30              | 0.18             | 1.63 | 4.08 |
| Total roots height                 | = 18.55 + (0.17 x Face Height) - (10.66 x Population) + (1.59 x Sex) - (4.45 x Age Group) | F4, 98 = 99.299, p<0.001 | 0.886    | 0.802              | 0.794    | 3.01           | 30              | 0.09             | 3.26 | 7.69 |
|                                    | = 41.75 - (0.01 x Nose Height) - (10.93 x Population) + (1.10 x Age Group)            | F4, 99 = 113.263, p<0.001 | 0.880    | 0.774              | 0.768    | 3.20           | 30              | 0.31             | 3.28 | 7.09 |
|                                    | = 31.89 + (0.38 x Total Occluded Incisors Height) - (10.47 x Population) - (2.91 x Age Group) | F3, 135 = 111.178, p<0.001 | 0.844    | 0.712              | 0.705    | 3.45           | 30              | 0.41             | 2.78 | 8.17 |
|                                    | = 30.21 + (0.43 x Total Incisors Height) - (10.75 x Population) - (2.90 x Age Group)  | F3, 135 = 109.192, p<0.001 | 0.842    | 0.708              | 0.702    | 3.47           | 30              | 0.68             | 2.73 | 6.71 |
|                                    | = 35.80 + (0.22 x Total Roots Height) - (11.04 x Population) - (3.35 x Age Group)     | F3, 134 = 103.059, p<0.001 | 0.835    | 0.698              | 0.691    | 3.54           | 30              | 0.50             | 2.48 | 6.45 |

Key: Population (Black = 1, White = 2), Sex (Male = 1, Female = 2), Age Group (<40 years = 1, ≥40 years = 2), Total Incisors Height (combined maxillary + mandibular central incisors height), Total Roots Height (combined maxillary + mandibular central roots height), SEE = Standard Error of Estimate

Table 7. Calculated upper, lower and total lip height mean measurements and validation test results.

| Description                        | Mean (mm) | Test sample (n) | Inaccuracy (mm) |
|------------------------------------|-----------|----------------|-----------------|
| Black (n=124)                      | 13.35     | 30             | 0.98            | 1.89 | 5.85 |
| Black Male (n=72)                  | 13.77     | 17             | 1.56            | 2.10 | 5.43 |
| Black Female (n=52)                | 12.78     | 13             | 0.26            | 1.37 | 3.23 |
| Black <40 years (n=91)             | 13.85     | 22             | 0.47            | 1.85 | 5.35 |
| Black ≥40 years (n=33)             | 11.99     | 8              | 2.30            | 2.40 | 5.81 |
| Black                             |           |                |                 |
| Black (n=124)                      | 12.63     | 30             | 0.18            | 1.59 | 3.18 |
| Black Male (n=72)                  | 12.99     | 17             | 0.05            | 1.62 | 3.54 |
| Black Female (n=52)                | 12.13     | 13             | 0.37            | 1.54 | 3.62 |
| Black <40 years (n=91)             | 13.16     | 22             | 0.20            | 1.56 | 3.71 |
| Black ≥40 years (n=33)             | 11.15     | 8              | 1.22            | 1.79 | 3.96 |

Table 8. Wilcoxon signed rank results comparing lip height to incisor height, including upper lip height to lower lip height.

| Description                        | Frequency (%) | Statistical Significance |
|------------------------------------|---------------|-------------------------|
| Upper Lip greater than Maxillary Incisor | 89            | p < 0.001; 8.580        |
| Lower Lip greater than Mandibular Incisor | 97            | p < 0.001; 9.598        |
| Total Lips greater than Total Occluded Incisors | 96            | p < 0.001; 4.849        |
| Upper Lip greater than Lower Lip     | 68            | p < 0.001; 4.038        |

| Description                        | Frequency (%) | Statistical Significance |
|------------------------------------|---------------|-------------------------|
| Upper Lip smaller than Maxillary Incisor | 92            | p < 0.001; 5.243        |
| Lower Lip smaller than Mandibular Incisor | 62            | p = 0.037; 2.081        |
| Total Lips smaller than Total Occluded Incisors | 82            | p < 0.001; 4.648        |
| Upper Lip greater than Lower Lip     | 54            | p = 0.240; 1.175        |

RESEARCH < 421
slightly more variable. The average maxillary root length in the ≥40 years category was 0.87 mm and 0.6 mm shorter in black and white individuals respectively, and the average mandibular root length was 0.88 mm and 0.17 mm shorter in black and white individuals respectively. This large variation between the population groups could partially be attributed to differences in sample sizes. Curve estimation, prior to regression analysis, confirmed all data had a linear relationship. All necessary assumptions for regression analysis were tested and met. Multiple linear regression that included population, sex and age group variables, significantly improved the goodness-of-fit for equations estimating upper, lower and total lip heights; without the demographic variables’ inclusion, the \( r^2 \)-values of generated equations were exceptionally weak, often below 0.100.

Table 6 lists a selection of equations with the best goodness-of-fit (\( r^2 \)-values) and validation test results for estimated lip heights; the equations independently utilise face height, nose height, central incisor height, and root length measurements. Notably, all independent variables used were not intercorrelated with one another. For comparison, Table 7 presents the validation results for a mean model approach to estimating upper, lower and total lip height.

Equations estimating upper lip height presented an \( r^2 \)-value ranging 0.594 to 0.742, with a standard error of the estimate (SEE) ranging from 1.70 mm to 1.97 mm; the equation utilising face height (including population and age group variables) performed best. This was evident in validation tests, with face height presenting an absolute mean inaccuracy of 1.65 mm, outperforming other tested equations (mean absolute inaccuracy ranging from 1.79 mm to 1.95 mm) and the calculated upper lip mean model value for black individuals (mean absolute inaccuracy, 1.89 mm).

The mean model value for black female upper lips, however, outperformed the regression equations during validation testing, with a 1.37 mm mean absolute inaccuracy (Table 7). In validation tests, the equations and mean models for upper lip height furthermore presented a consistent positive bias (calculated as the actual value minus estimated value), thus always underestimating the actual value.

Equations estimating lower lip height presented an \( r^2 \)-value ranging 0.582 to 0.690, with a SEE ranging from 1.81 mm to 1.96 mm. The equation utilising face height (including population and age group variables) performed best (mean absolute inaccuracy, 1.42 mm) against other tested equations (mean absolute inaccuracy ranging from 1.56 mm to 1.63 mm) and the generated lower lip mean model values for black individuals (mean absolute inaccuracy, 1.59 mm). During the validation testing of lower lip equations and mean models, results tended to demonstrate a positive bias, thus often underestimating the actual value.

Equations estimating total lip height presented an \( r^2 \)-value ranging 0.698 to 0.802, with quite a high SEE ranging from 3.01 mm to 3.54 mm. According to \( r^2 \) and SEE results, the equation utilising face height (including population, sex and age group variables) performed the best. However, validation tests identified the face height equation (that includes population, sex and age group variables) provided the most inaccurate results compared to all other equations (mean absolute inaccuracy, 3.26 mm).

The calculated maxillary plus mandibular root length equation (total roots length) utilising population and age group variables performed best (mean absolute inaccuracy, 2.48 mm). All tested equations outperformed the total lip mean model for black individuals in general (mean absolute inaccuracy, 3.34 mm).

Contrastingly, when considering sex and age, only the equation using total root length outperformed the mean models (Table 6, Table 7). A negative bias during validation testing was often presented for equations and mean models, thus they tended to overestimate the actual value.

General lip patterns identified, include the fact that black individuals tended to present a significantly greater upper, lower and total lip height relative to maxillary, mandibular and total incisor height, respectively. The upper lip also tended to be significantly greater than the lower lip (Table 8). In white individuals, the upper, lower and total lip height tended to be significantly smaller relative to the maxillary, mandibular and total incisor height, respectively. The upper lip also tended to be greater than the lower lip, but it was not statistically significant (Table 8).

**DISCUSSION**

Black southern African individuals notably presented statistically significant thicker lip heights in comparison to white southern African individuals of European origin.

This was similarly identified by Schmidlin et al. when investigating facial ageing patterns in black South African males, and comparing results to an Italian sample taken from Sforza et al. It is generally agreed that the prognathism typical to individuals of Sub-Saharan African origin is accompanied by thicker lips, but the exact relationship between prognathism and lip thickness is not clear. Thus, we cannot necessarily consider the correlation between these two variables as causal, as the correlation can likely be attributed to the fact that both features are commonly found in individuals of African origin.

The weak relationship between incisor and lip heights also suggested that the canon employed by Gerasimov is unreliable, and is especially inaccurate for individuals of Sub-Saharan African origin.

Thus, lip thickness cannot be assumed to be the same as the incisor enamel height or the gum line to gum line thickness. Wilkinson et al. similarly identified this when investigating incisor and lip height patterns in European
and Indian subcontinent groups; stating that factors such as population, prognathism, and age, can all play equally dominant roles in lip thickness - as similarly indicated in other existing research.13,15,17,18

Differences between the sexes in the black population included greater lip heights, central incisor heights and root lengths in males, which is comparable to several previous studies on other population groups.12-15 White males and females did not differ significantly regarding lip heights, central incisor heights and root lengths, which was similarly identified in only two existing studies.4,11 A clearer pattern of sex differences in white southern Africans might be detected using a larger sample size.

Age-related lip thinning and dental wear is well recognised.4,8,16,17,30-32 Similar to Schmidlin et al’s study,17 this study observed a similar decrease in lip heights for both black and white individuals. Thus, even though both groups experienced lip thinning with age, black individuals maintained considerably thicker lip heights compared to white individuals. Soft tissue changes are attributed to a reduction in skin elasticity, muscle tone and volume, which invariably causes the lips to thin and the prominence of the vermilion border to decrease.8,30,31

The observed dental wear is attributed to three factors: attrition, abrasion and erosion.30,32 Attrition is caused by mastication or grinding between opposing teeth. The effect of attrition can be exacerbated in those who eat a particularly fibrous diet, or habitually clench and grind their teeth. Abrasion is caused by food and foreign body contact (e.g. tooth brushing). Lastly, acid-based leaching and dissolution cause erosion. This can often be the result of frequent high-acidity fluid consumption (carbonated drinks, fruit juices).

The statistically significant reduction in root length is likely the result of root resorption, due to inflammation and/or orthodontic tooth movement over a lifespan.33-35 It is, however, possible that the age-related changes observed might be equally influenced by a degree of human variation accentuated by the sample composition.

This study was retrospective, with all CBCT scans on the hospital database being intended for orthodontic investigation and procedures. Many scans were thus localised (not complete craniofacial scans). Cases with clear pathology or trauma had to be excluded. Older individuals were also of limited representation. Thus obtaining adequate sample sizes and distributions was challenging.

If more scans were to become available, it would be ideal to test an equally representative sample (equal age, sex and population groups) to offer more informative results. This could, however, take an extensive timeframe to accumulate.

In direct response to concerns voiced by the SAPS, with regard to the lack of CFA standards for individuals of southern African or even Sub-Saharan African origin, this study provides a series of potential equations for estimating upper, lower and total lip height, which expands on the equations offered by Wilkinson et al.4 The equations utilising face height were statistically the strongest (according to r-values, r²-values, and SEE), and thus the height of the facial skeleton gives the best prediction of lip height. In the validation tests, however, utilising the total roots length (maxillary plus mandibular root lengths) equation with population and age group variables demonstrated a better solution to total lip height estimation.

Due to the weakened intra- and inter-observer reliability of root length measurements, it is, however, recommended that the total lip height regression equation be used as a secondary option to the independent lip height equations available. Maximum absolute errors accounted, which consider worst individual performance during validation tests, were quite large (Table 6 and 7), thus all methods of prediction need to be used with caution.

The generated equations incorporate both a black and white southern African sample. The limited sample sizes available for the white population, with validation tests only being performed on a black sample, indicates the equations are best suited to black southern African cases. The validation tests performed using a black sample verified that the equations generated in this study often performed better than the mean models. An exception was the upper lip mean model for black females. We do, however, suggest that equations demonstrate a preferable degree of individuality that mean values do not.

Acknowledgements

We are grateful for the financial assistance from the Leverhulme Trust (UK) (Grant Number: SAS-2017-005). Opinions and conclusions expressed are those of the authors and are not necessarily to be attributed to the Leverhulme Trust.

We thank the University of Pretoria for access to existing CBCT data. We are furthermore indebted to the anonymous patients that made this study possible.

References

1. Steyn M, L’Abbé EN, Myburgh J. Forensic Anthropology as practiced in South Africa. In: Blau S, Ubelaker DH, editors. Handbook of Forensic Anthropology and Archaeology. 2nd edition. New York: Routledge, Taylor & Francis Group. 2016; 151-65.
2. Gatiff BP. Facial sculpture on the skull for identification. Am J Forensic Med and Path. 1984; 5(4): 327-32.
3. Helmer R, Rohricht S, Petersen D, Mohr F. Assessment of the reliability of facial reconstruction. In: Iscan MY, Helmer RP, editors. Forensic analysis of the skull. New York: Wiley Liss Publications, 1993: 229-34.
4. Wilkinson C, Motwani M, Chiang E. The relationship between the soft tissues and the skeletal detail of the mouth. J Forensic Sci. 2003; 48(4): 728-32.
5. Aulsebrook W, Iscan MY, Slabbert J, Becker PJ. Superimposition and reconstruction in forensic facial identification: A survey. Forensic Sci Int. 1995; 75: 101-20.
6. Stephan CN. Craniofacial identification: Techniques of facial approximation and craniofacial superimposition. In: Blau S, Ubelaker DH, editors. Handbook of forensic anthropology and archaeology. California: Left Coast Press, 2006: 304-31.
7. Huete MI, Ibáñez O, Wilkinson C, Kahana T. Past, present, and future of craniofacial superimposition: Literature and international surveys. Legal Medicine 2015; 17(4): 267-78.
8. Storza C, Grandi G, Binelli M, Dolci C, De Menezes M, Ferrario VF. Age- and sex-related changes in three-dimensional lip morphology. Forensic Sci Int. 2010; 200: 182e1-e7.
9. Ferrario VF, Sforza C, Tartaglia GM, Sozzi D, Caru A. Three-dimensional lip morphometry in adults operated on for cleft lip and palate. Plastic Reconstr Surg 2003; 111: 2149-56.
10. Sawyer AR, See M, Nduka C. 3D stereophotogrammetry quantitative lip analysis. Aesthetic Plast Surg. 2009; 33: 497-504.
11. Ferrario VF, Sforza C, Miani A Jr, Tartaglia G. Craniofacial morphometry by photographic evaluations. Am J Orthod Dentofac Orthop. 1993; 103: 327-37.
12. Farkas LG, Hreczko TA, Katic MJ. Craniofacial norms in North American Caucasians from birth to young adulthood. In: Farkas LG, editor. Anthropometry of the head and face. 2nd edition. New York: Raven Press, 1994: 241-335.
13. Hajnis K, Farkas LG, Ngim RCK, Lee ST, Venkatadri G. Racial and ethnic morphometric differences in the craniofacial complex. In: Farkas LG, editor. Anthropometry of the head and face. 2nd ed. New York: Raven Press, 1994: 201-21.
14. Nanda RS, Ghosh J. Facial soft tissue harmony and growth in orthodontic treatment. Semin Orthod. 1995; 1: 67-81.
15. Ferrario VF, Sforza C, Serrao G. A three-dimensional quantitative analysis of lips in normal young adults. Cleft Palate Craniofacial J. 2000; 37(1): 48-54.
16. Gerasimov M. The face finder. London: Hutchinson & Co, 1971.
17. Schmidlin E, Steyn M, Houlton TMR, Briers N. Facial ageing in South African adult males. Forensic Sci Int. 2018; 289: 277-86.
18. Angel JL. Restoration of head and face for identification. In: Proceedings of the 30th annual meeting of the American Academy of Forensic Sciences, St. Louis, Missouri. 1978.
19. Gatilff BP, Snow CC. From skull to visage. J Biocommun. 1979; 6: 27-30.
20. Taylor KT. Forensic art and illustration. New York: CRC Press, 2001.
21. George RM. The lateral craniofacial method of facial reconstruction. J Forensic Sci. 1987; 32: 1305-30.
22. Malu PZ, Velemisnka J. Vertical lip position and thickness in facial reconstruction: A validation of commonly used methods for predicting the position and size of lips. J Forensic Sci. 2016; 61(4): 1046-54.
23. Dahllberg G. Statistical Methods for Medical and Biological Students. London: George Allen & Unwin, 1940.
24. Perini TA, de Oliveira GL, Ornellas JS, de Oliveira FP. Technical error of measurement in anthropometry. Rev Bras Med Esporte 2005; 11: 86-90.
25. Langley NR, Jantz LM, McNulty S, Maijanen H, Ousley SD, Jantz RL. Data for validation of osteometric methods in forensic anthropology. Data in Brief. 2018; 19: 21-8.
26. Kamburoglu K, Kolsuz E, Kurt H, Kiliç Ç, Özcan T, Paksoy CS. Accuracy of CBCT measurements of a human skull. J Digit Imaging 2011; 24(5): 787-93.
27. Garcia-Sanz V, Bellot-Arcis C, Hernández V, Serrano-Sánchez P, Guarnio J, Paredes-Gallardo V. Accuracy and reliability of cone-beam computed tomography for linear and volumetric mandibular condyle measurements. A human cadaver study. Sci Rep. 2017; 7(1): 11993.
28. Klepinger L, Giles E. Clarification or confusion: statistical interpretation in forensic anthropology. In: Reichs KJ (Ed.). Forensic Osteology: Advances in the Identification of Human Remains. Charles C Thomas, Springfield. 1998; 427-40.
29. Møller GR, Boldsen JL. Transition analysis: a validation study with known-age modern American skeletons. Am J Phys Anthropol. 2012; 148: 98-110.
30. Wilkinson C. Forensic facial reconstruction. Cambridge: Cambridge University Press, 2004.
31. Kaur M, Garg RK, Singla S. Analysis of facial soft tissue changes with aging and their effects on facial morphology: A forensic perspective. Egypt J Forensic Sci. 2015; 5: 46-56.
32. Sperber GH. Dental Wear: Attrition, Erosion, and Abrasion - A Palaeo-Odontological Approach. Dent J (Basel). 2017; 5(2): 19.