# TESTING AN ALTERNATIVE APPROACH FOR STATURE DETERMINATION FROM FOOTPRINT MEASUREMENTS THAT ESCAPES RACIAL TYPOLOGY USING A SAMPLE OF ADULT EGYPTIANS AND MALAYSIANS

**Maram Atef [a], Somaya Madkour [a], Asmaa El-Banna [a], Ghada Abu-Sheasha [b], and Heba Abdel Samie Mohamed Hussein [a]**

[a] Department of Forensic Medicine & Clinical Toxicology, Faculty of Medicine, Alexandria University, Alexandria, Egypt;

[b] Department of Biomedical Informatics and Medical Statistics, Medical Research Institute, Alexandria University, Alexandria, Egypt

**Corresponding author: Heba Abdel-Samie Mohamed Hussein,** Assistant professor in Forensic Medicine and Clinical Toxicology, Faculty of Medicine, Alexandria University, Egypt

**E-mail:** husseinheba399@gmail.com **ORCID No:** 0000-0002-0022-4640

**Business phone number:** 00201002957439

**Submit Date** 2024-08-05

**Revise Date** 2024-09-18

**Accept Date** 2024-09-30

## ABSTRACT

**Background and Objectives:** A major issue with almost all previous methods of stature estimation is that before the method can be applied, an unknown individual must first be assigned to a certain sex or population. **Aim of this study:** The present study highlighted the differences in stature and footprint measurements between two populations of two different races and aimed to assess if escaping racial typology is possible and to determine the stature of unknown individuals before population determination while building their biological profiles. **Subjects and Methods:** The study tested this alternative approach using two divergent statistical methodologies, namely, random forest and multiple regression analysis. The study was performed on 200 volunteered adults, and they were distributed to two racial groups: 100 adult Egyptians and 100 adult Malaysians. The whole sample of each group was divided into two subgroups, namely, training and test subgroups. The training subgroup was used to build up a model, whereas the test subgroup was applied to validate the built model. **Results:** Overall, there were insignificant differences between the two populations regarding the errors in estimating stature in all developed models, although there were significant differences in the measurements of footprint and stature between the two populations. **Conclusion:** The present study concluded that the models of stature estimation developed from one ethnic population could be applied to another population with satisfactory predictive performances, and one equation can represent distinct parts of the world.

**Keywords:** stature; footprint; nonpopulation-specific; Malaysian; Egyptian

## INTRODUCTION

Like fingerprints, footprints can give information of great value in building the biological profiles of individuals **(Moorthy, 2017)**.

The individual characteristics of an adult foot as genetic inheritance, environmental differences, socioeconomic developments, ethnicities, and cultures can influence foot morphology. It was well proven that an adult foot is unique to each individual **(Delgado-Abellán et al., 2014; Nirenberg et al., 2017)**.

Apart from giving information regarding barefoot morphological features, the foot can also offer an idea regarding the body size of somebody, which can further help an investigator limit the pool of potential suspects **(Burrow et al., 2017; Shukla et al., 2017; Caplova et al., 2018)**.

The evaluations of stature can be the most useful point from which to initiate the process of identification in several cases where sequestered fleshed body fragments are discovered, particularly in the adult foot (where ethnicity and age indicators are mostly absent) **(Krishan et al., 2015)**.

In domestic scenes of crime, where criminals were known to leave footprint impressions, naked footprints may be recovered in the form of impressions. Several studies have recommended the analysis of the impressions of bare feet for developing countries, where the countryside population infrequently wears shoes because of socioeconomic, climatic, or religious reasons **(Webb et al., 2006; Byard & Payne-James, 2015)**.

Different previous studies tried to record the relationship between stature and measurements of footprint using linear as well as multiple regression statistical equations among various human populations. The main assumption for all these studies is that population-specific equations will improve the accuracy of the stature estimate, and ancestry must be determined first before an equation can be applied. Thus, it is well known that regional studies on stature estimations from footprints are desperately necessary **(Fawzy & Kamal, 2010; Kanchan et al., 2012; Hemy et al., 2013; Abledu et al., 2016)**.

In fact, ancestry determination is likely the most difficult part of the identification process. Migration, immigration, and admixture result in intrinsic diversity within each major genetic breeding population group and changes in the social construction of racial identity, thereby complicating the trials to relate biological traits to social labels. Consequently, it must be emphasized that attributing any distinctive feature or equation exclusively to a certain race is impracticable. Furthermore, in several cases, the parameters of a given group are based on assumptions, thereby making it difficult, if not impossible, to assign an unidentified individual to a certain group **(Moorthy, 2015; M'charek et al., 2020)**.

Given actual crime settings, where the perpetrator's identity is unknown, it is proposed that equations without sex or race indicators would be better for predicting stature **(Hemy et al., 2013)**.

Hence, the present study attempted to evaluate if it is possible to escape racial typology and build formulas that can represent diverse parts of the world away from the influence of genetics and environments. Additionally, the study evaluated the predictive performance of the nonpopulation-specific statistical models in predicting stature from footprints and how often they can provide the ranges of stature estimates that could be accepted in the forensic context using a randomly selected sample of Egyptians and Malaysians.

## SUBJECTS

Two hundred volunteers with a range of age of 18–30 years participated in the current cross-sectional study, and they were distributed on two groups according to race: 100 Egyptians (50 females plus 50 males) and 100 Malaysians (50 females plus 50 males). After taking the ethical approval, the informed consent was obtained from all the research contributors, and the basic demographic data (age, sex, and state of origin) were collected.

The study excluded subjects with apparent indicative abnormality of the feet or spine, history of injuries of the lower limbs, surgery of the foot or ankle, and edema of the lower limbs **(Krishan, 2008; Burrow, 2016)**.

## METHODS

### Footprint collection

The footprints were gathered from each subject by means of an inkless shoe print kit (Carolina, USA). The subjects were asked to put their foot on the preimpregnated stain-free shoeprint pad with light pressure and then to stamp on shoe print presensitized impression cards (incorporated in the kit) on a uniform surface, posing equivalent weight on both feet during standing. The right and left footprints were documented one by one **(Robbins, 1986)**.

### Analysis of footprints (Figure 1)

The designated longitudinal axis (DLA) and baseline were traced following Krishan **Robbins (1986)** to create a fixed axial alignment for measurements.

**The following measurements were taken in centimeters on the footprints (Robbins, 1986; Lohman et al., 1992; Fawzy & Kamal, 2010):**

- Maximum footprint length (MFL): the straight distance between the pterion to the maximal anteriorly projecting point of the big or the second toe, whichever is longer when the foot is fully stretched.
- Footprint length measurements (T1, T2, T3, T4, and T5 lengths): these measurements were measured from the maximum anterior point of every toe to the pterion starting from the big toe.
- Footprint breadth at the ball (FBB): the distance from the maximum lateral point to the maximum medial point of the borders of the forefoot area of the footprint.
- Footprint breadth at heel (FBH): the maximum length between the most protruding points on the medial surface and the lateral surface of the heel.
- Big toe pad width (BTW): the distance from the most medial to the most lateral points of the big toe.
- Big toe pad length (BTL): the maximum length of the toe pad.

**The following two indices were calculated from the measured parameters:**

- Heel/ball index of the footprint (HBI): the ratio of the FBH multiplied by 100 and divided by the FBB.
- Foot index (FI) from the footprint was calculated using the formula: FBB × 100/MFL.

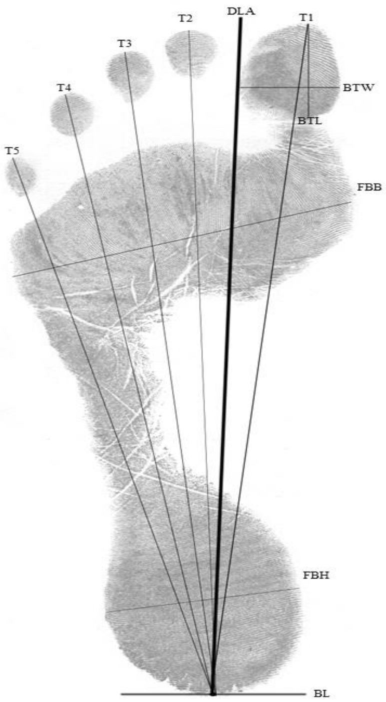


**Fig. (1):** Illustrative figure of footprint measurements showing the studied parameters: **T1 to T5**: diagonal footprint length measurements, **FBB**: foot breadth at the ball region, **FBH**: foot breadth at heel, **BTL**: big toe length, **BTW:** big toe width. **DLA**: diagonal longitudinal axis, **BL,** baseline.

**Measurement of stature (Breiman, 2001)**

The stature was measured in centimeters by requesting the subject to stand barefoot on the base panel of a typical metric stature measuring scale, both feet adjacent to each other and trunk straight along the vertical panel.

**Statistical analyses**

The effects of sex, race, and side on footprint measures were tested simultaneously using multiple linear models. Every participant had correlated observations [right and left], and the generalized estimating equations were utilized to adjust for such a correlation.

Four interaction terms were entered in every model, namely, and, to test for interaction (when the effect of one predictor on the outcome varies at the different levels of other predictors). If the interaction term does not significantly contribute to the model (p >.05), then only one coefficient is used to determine the effect of each relevant predictor on the outcome variable. Conversely, if the interaction term has a significant contribution to the model (p <.05), then two coefficients were applied to determine the effect of the relevant predictors on the outcome variable.

For the determination of stature from footprint measures, random forest (RF) was used: RF was introduced by **Cutler et al. (2012)** It is a flexible machine learning algorithm that could be used to calculate categorical and linear predictors. RF forms numerous decision trees and combines them together to obtain a more stable and precise prediction. For each tree, it picks a random number of contributors and a random number of predictors, and it is also used principally because it is not affected by unexplained variability in the data, generalizes well, has a framework for high-dimensional problems (small observation with a large number of predictors), and offers worthy measures of different visualization and importance **(R Core Team, 2013)**.

For the estimation of the performance of the model, the “hold out” method was utilized, where the whole sample was divided into two sets, namely, training and test sets. The models were built on training sets and validated on the test sets, and those that perform poorly on the training set will certainly perform poorly in the test set and will never generalize to the population. Contrarily, those that perform well on the training set might or might not generalize well to the population. The predictive performance of the models on the test set helps expect the conduct of the model in the population (generalizability of the model).

In our study, one training set was applied to build up the model, and it contained approximately 75% of one ethnicity group. Using the independent t-test, the predictive power of the model was compared between the two test sets: one with the same ethnicity as the training set and the other with a different ethnicity.

The measures of the model performance used include the mean error (**ME**), the mean absolute error (MAE), and the root mean square error (**RMSE**). The **ME** refers to the average of all the errors, and an “error” refers to the variance between the predicted and the actual values. For instance, if the model predicts the height of a participant at 170 cm and his actual at 167 cm, then the error is equal to 3 cm. **RMSD** can be defined as the square root of the average of squared errors. It is constantly non-negative; the lower the RMSD, the better the model power, and a value of 0 would indicate a perfect prediction. The **MAE** is the average of the absolute figures of the errors, and it is fundamentally easier to understand than the **RMSE**.

In this analysis, a full RF model was built, where all the predictors were included. Usually, the predictive performance of the model might improve by dropping unimportant predictors. Our strategy was to build the full model and compute the Gini-based variable importance. The Gini index estimates the error reduction that occurred when every variable was included in the model. The more important the variable, the more the reduction is in the error.

A new model (reduced model) was built using the most important five or six predictors. To check whether the predictive performance of the reduced model significantly decreased from that of the full model, the paired t-test was applied to compare the errors in the two models on the same training set. With no statistically significant difference between them, the reduced model was validated and compared on the test sets.

The RF models were built using R statistical software and the RF package **Liaw and Wiener (2002)** and the data were gathered, charted, and analyzed by means of SPSS software **(IBM Corp, 2012)**. The results of the significance test were cited as two-tailed probabilities and judged at the level of 5% **(Krishan & Sharma, 2007)**.

**RESULTS:**

**Demographic data**

The participants were divided into two groups: 100 Egyptian adults (50 females with a mean age of 22.14 ± 4.11 years and 50 males with a mean age of 22.28 ± 3.95 years) as well as 100 Malaysian adults (50 males with a mean age of 21.76 ± 2.05 years and 50 females with a mean age of 21.98 ± 1.48 years). The study revealed no significant differences concerning age between both sexes in the same population or between both populations.

**Stature (Table 1)**

The mean measured stature values were 172 cm ± 6.5 for Egyptian males and 161.5 cm ± 5.7 for Egyptian females. The mean stature values of the Malaysian sample were 170 cm ± 5.9 for males and 154.8 cm ± 4.9 for females.

In both studied populations, the males showed significantly larger measurements than the females (p =.001). When the two populations were compared, it was found that the mean values of stature are significantly higher in the Egyptians than in the Malaysians for males (p =.043) and females (p <.001).

**Table (1): Distribution of the studied groups according to the stature (in centimeters) (n = 200)**

| Stature (cm) | Egyptian | | Malaysian | |
|---|---|---|---|---|
| | Male (n = 50) | Female (n = 50) | Male (n = 50) | Female (n = 50) |
| **Min.–Max.** | 155.0–186.0 | 150.0–174.0 | 159.0–185.0 | 148.0–167.0 |
| **Mean ± SD** | 172.6 ± 6.5 | 161.5 ± 5.7 | 170.3 ± 5.9 | 154.8 ± 4.9 |
| **$t_1$ (p)** | $9.023^*$ (<0.001) | | $14.292^*$ (<0.001) | |
| **$t_2$ (p)** | $2.870^*$ (0.043) | | | |
| **$t_3$ (p)** | $6.322^*$ (<0.001) | | | |

Min, minimum; Max, maximum; t, Student's t-test; *, statistically significant at p ≤ 0.05

$t_1$ (p): Student's t-test between both sexes within each population.

$t_2$ (p): Student's t-test between Egyptian and Malaysian males.

$t_3$(p): Student's t-test between Egyptian and Malaysian females.

**Differences between footprint measurements by gender, ethnicity, and side by means of regression analysis (Table 2)**

In the current study, regression analysis was accomplished to compare the obtained values of the anthropometric measurements of footprints by sex, ethnicity, and side.

Regarding sex, all the footprint measurements of the males showed statistically significantly higher values than did the females (p <.001). However, no significant difference was observed between both sexes in the calculated indices (HBI and FI) (p =.826 and.378, respectively). This finding was consistent among both populations and on both feet.

Regarding ethnicity, all the measurements of footprint length in the Egyptians were significantly greater than in the Malaysians regardless of the sex or the side (p <.001).

Nevertheless, the breadth measurements of footprints revealed that FBB showed no statistically significant difference between the two studied populations (p =0.134), whereas the Egyptians tended to have significantly larger FBH measurements than did the Malaysians (p <0.001), with a higher difference having been observed in the right than in the left foot.

Additionally, the Egyptians had significantly higher HBI (p <0.001) and lower FI (p =0.002) than those of the Malaysians among both sexes and in both feet.

Regarding side, the right and left feet were overall comparable among both sexes of the two sampled populations, except for BTW.

**Stature determination using footprint dimensions**

The present study used two different statistical methods to test the accuracy of building models in predicting stature, namely, RF and multiple regression analysis.

Each statistical model was built on 74% of each population and was tested on two test sets: the remaining 26% of the same population and the whole other population to assess the applicability of the model.

Additionally, the full RF model ordered the various footprint measures according to their importance in the prediction of stature to build the reduced RF models and multiple regression models using the most significantly correlated variables, which were the maximum foot length and foot length measurements (T1, T2, T3, T4, and T5 lengths) in all the developed models.

**Table (2): Differences in footprint measurements by sex, ethnicity, and side**

| | Male vs female | | Egyptian vs Malaysian | | Lt foot vs Rt foot | |
|---|---|---|---|---|---|---|
| | **(B) [95% CI]** | **(p-Value)** | **(B) [95% CI]** | **(p-Value)** | **(B) [95% CI]** | **(p-Value)** |
| **MFL** | 2.35 [2.07, 2.64] | (<.001) | 0.86 [0.58, 1.15] | (<.001) | −0.01 [−0.07,0.04] | (.683) |
| **T1** | 2.35 [2.07, 2.62] | (<.001) | 0.99 [0.7, 1.26] | (<.001) | −0.02 [−0.08, 0.04] | (.494) |
| **T2** | 2.4 [2.12, 2.69] | (<.001) | 0.7 [0.418, 0.99] | (<.001) | 0.02 [−0.035, 0.08] | (.446) |
| **T3** | 2.37 [2.09, 2.64] | (<.001) | 0.7 [0.42, −0.98] | (<.001) | 0.03 [−0.03, 0.08] | (.356) |
| **T4** | 2.21 [1.95, 2.47] | (<.001) | 0.75 [0.49, 1.01] | (<.001) | 0.04 [−0.01, 0.088] | (.117) |
| **T5** | 2.026 [1.7, 2.8] | (<.001) | 0.71 [0.46, 0.97] | (<.001) | 0.04 [−0.017, 0.1] | (.160) |
| **FBB** | 0.86 [0.7, 1.01] | (<.001) | 0.113 [−0.04, 0.26] | (.134) | 0.02 [−0.022, 0.067] | (.32) |
| **FBH** | 0.481 [0.35, 0.61] | (<.001) | 0.405 [0.27, 0.55]$^{Rt}$ | (<.001) | −0.06 [−0.14, 0.016]$^{E}$ | (.118) |
| | | | 0.295 [0.16, 0.44]$^{Lt}$ | (<.001) | 0.05 [−0.009, 0.103]$^{M}$ | (.098) |
| **BTW** | 0.254 [0.18, 0.33]$^{Rt}$ | (<.001) | 0.042 [0.02, 0.107] | (.208) | 0.09[0.036,0.148]$^{Male}$ | (.001) |
| | 0.338 [0.26, 0.4]$^{Lt}$ | (<.001) | | | 0.01 [−0.04,0.056]$^{female}$ | (.743) |
| **BTL** | 0.408 [0.32, 0.5] | (<.001) | 0.077 [0.01, 0.17] | (.085) | −0.01 [−0.052, 0.027] | (.534) |
| **HBI** | 0.125 [−0.9, 1.24] | (.826) | 2.99 [1.87, 4.11] | (<.001) | −0.23 [−0.75, 0.296] | (.394) |
| **FI** | −0.26 [−0.8, 0.3] | (.378) | −0.94 [−1.5, −0.36] | (.002) | 0.12 [−0.079, 0.312] | (.242) |

**B** stands for the adjusted mean difference where the null value = 0; $^{M, E, Male, Female, Rt, Lt}$ stands for B calculated among Malaysians, Egyptians, males, females, right foot, and left foot respectively.
**MFL**, maximum footprint length; **T1 to T5**, length from pterion to each toe; **FBB**, footprint breadth at the ball; **FBH**, footprint breadth at heel; **BTW**, big toe width; **BTL**, big toe length**; HBI**, heel/ball index; **FI,** foot index

**Statistical models that were built on the Egyptians and applied on the Malaysians (Table 3)**

Training set (74% of the Egyptian sample) and validation set (26% of the Egyptian sample and 100% of the Malaysian sample)

**Stature determination using right footprint measurements:**

- The comparison of individual errors between full and reduced RF models on the Egyptian training sets showed no statistical significance, indicating that the reduced model was similar to the full model with fewer parameters. Hence, the reduced model was selected.
- The comparison of the mean errors between the Egyptian and Malaysian test sets in the reduced RF model exhibited no statistically significant difference.
- The multiple regression equation was built based on the measurements selected by the full RF model: using $R^2$, the predictive performance of the regression model was higher among the Malaysian test sets (76% vs. 64%). The comparison of errors between the Egyptian and

Malaysian test sets showed no statistically significant difference.

**Stature determination models using left footprint measurements**

- The comparison of mean individual errors between full and reduced models on the Egyptian training sets exhibited no statistically significant difference.
- The comparison of errors between the Egyptian and Malaysian test sets in the reduced RF model displayed no statistically significant difference.
- Table 3 shows the predictive power of the developed multiple regression equation. Using $R^2$, the predictive performance of the regression model was higher among the Malaysian test set (81% vs. 64%). The comparison of errors between the Egyptian and Malaysian test sets also showed no statistically significant difference.

**Table (3): Statistical models used to predict stature using footprint measurement building on the training set of the Egyptians and applied on the Malaysians**

| Model | Side | | Training set | First test set | Second test set | MD (95%CI) | t (p-Value) |
|---|---|---|---|---|---|---|---|
| | | | Egyptian | Egyptian | Malaysian | | |
| | | | (n = 74) | (n = 26) | (n = 100) | | |
| **Random forest (full model)** | | | | | | | |
| | **Right** | | | | | | |
| | | ME(SD) | 0.11(4.64) | 0.66(4.92) | −1.2 (5.2) | | 0.5(.608) |
| | | (Min, Max) | (−9.0, 8.9) | (−8.6, 10.1) | (−10.9,12.1) | | |
| | | RMSE | 4.6 | 4.9 | 5.3 | | |
| | | MAE | 3.8 | 4.1 | 4.4 | | |
| | **Left** | | | | | | |
| | | ME(SD) | 0.13(4.33) | −0.09(4.73) | −1.3 (4.76) | | 0.66(.506) |
| | | (Min, Max) | (−10.3, 8.9) | (−9.5, 8.03) | (−10.4,11.3) | | |
| | | RMSE | 4.3 | 4.6 | 4.9 | | |
| | | MAE | 3.45 | 3.95 | 3.97 | | |
| **Random forest (reduced model)** | | | | | | | |
| | Right | | | | | | |
| | | ME(SD) | 0.01(4.6) | 0.2(4.8) | −1.7(5.2) | (0.32, 1.7) | 4.14(.093) |
| | | (Min, Max) | (−9.5, 9.1) | (−9.7, 10.2) | (−11.2, 9.6) | | |
| | | RMSE | 4.6 | 4.7 | 5.4 | | |
| | | MAD | 3.7 | 3.9 | 4.5 | | |
| | Left | | | | | | |
| | | ME(SD) | 0.03(4.2) | 0.3(4.7) | −1.2(4.6) | (−1.07, 0.9) | 2.96(.357) |
| | | (Min, Max) | (−9.5, 9) | (−9.9, 9.1) | (−10.4, 7.1) | | |
| | | RMSE | 4.2 | 4.7 | 4.7 | | |
| | | MAD | 3.3 | 3.8 | 3.9 | | |
| **Multiple linear regression** | | | | | | | |
| | Right | | | | | | |
| | | $R^2$ | 69% | 66% | 72% | | |
| | | ME(SD) | 0(4.3) | 0.5 (4.9) | −0.2(5.1) | (−1.49, 2.88) | 0.6(.531) |
| | | (Min, Max) | (−7.1, 7) | (−9.2,10.3) | (−10,10.2) | | |
| | | RMSE | 4.25 | 4.9 | 5.03 | | |
| | | MAD | 3.4 | 3.8 | 4.2 | | |
| | Left | | | | | | |
| | | $R^2$ | 77% | 64% | 81% | | |
| | | ME(SD) | 0(3.96) | −0.18 (4.95) | −0.03(4.47) | (−1.87, 2.11) | 0.12(.91) |
| | | (Min, Max) | (−6.9, 7.2) | (−9.9, 10.1) | (−8.1, 8.9) | | |
| | | RMSE | 3.9 | 4.9 | 4.5 | | |
| | | MAD | 3.1 | 4.03 | 3.7 | | |

**$R^2$**, R-squared value; **ME**, mean estimated error; **SD**, standard deviation; **Min**, the minimum estimated error;

**Max**, the maximum estimated error; **RMSE**, root mean square error; **MAD**, median absolute deviation; **MD**, mean difference; **CI**, confidence interval

**Statistical models that were built on the Malaysians and applied on the Egyptians (Table 4)**

Training set (74% of the Malaysian sample) and validation set (26% of the Malaysian sample and 100% of the Egyptian sample)

**Stature determination models using right footprint measurements:**

- The comparison of individual errors between full and reduced models on the Malaysian training sets showed no statistical significance.
- The comparison of errors between the Malaysian and Egyptian test sets in the reduced RF model exhibited no statistically significant difference.
- The multiple regression model derived for stature estimation from the predictors selected displayed that the predictive performance was higher in th
- Egyptian test set than in the Malaysian test set (69% vs. 68%). The comparison of errors between the Malaysian and Egyptian test sets showed no statistically significant difference.

**Table (4): Statistical models used to predict stature using footprint measurements building on the training set of the Malaysians and applied on the Egyptians**

| Model | Side | | Training set | First test set | Second test set | MD (95%CI) | t (p-Value) |
|---|---|---|---|---|---|---|---|
| | | | Malaysian (n = 74) | Malaysian (n = 26) | Egyptian (n = 100) | | |
| **Random forest (full model)** | | | | | | | |
| | Right | | | | | | |
| | | ME (SD) | 0.19 (5.1) | 0.69 (5.1) | 0.55 (4.8) | | 0.27 (.789) |
| | | (Min, Max) | (−10.5,11) | (−8.7,10) | (−9, 11.6) | | |
| | | RMSE | 5.02 | 5.1 | 4.8 | | |
| | | MAD | 4.15 | 4.2 | 4.1 | | |
| | Left | | | | | | |
| | | M(SD) | 0.17 (4.4) | −1.11 (4.42) | 0.7 (4.4) | | 0.756 (.452) |
| | | (Min, Max) | (−7.6, 10) | (−8.2, 8.9) | (−10.1, 11.1) | | |
| | | RMSE | 4.3 | 4.5 | 4.5 | | |
| | | MAD | 3.6 | 3.7 | 3.6 | | |
| **Random forest (reduced model)** | | | | | | | |
| | Right | | | | | | |
| | | ME (SD) | 0.14 (5.1) | −0.6 (5.1) | 0.12 (5.05) | (−3, −0.6) | 1.521 (.532) |
| | | (Min, Max) | (−10,10.2) | (−11, 9.8) | (−9.3,10) | | |
| | | RMSE | 4.6 | 4.7 | 5 | | |
| | | MAD | 4.1 | 4.2 | 4.3 | | |
| | Left | | | | | | |
| | | ME (SD) | 0.05 (4.3) | −0.8 (4.4) | 0.12 (4.5) | (−3, −0.9) | 1.025 (.346) |
| | | (Min, Max) | (−8.8,9.3) | (−10, 7.1) | (−7,9.9) | | |
| | | RMSE | 4.3 | 4.5 | 4.5 | | |
| | | MAD | 3.5 | 3.7 | 3.6 | | |
| **Multiple linear regression** | | | | | | | |
| | Right | | | | | | |
| | | $R^2$ | 80% | 68% | 69% | | |
| | | ME (SD) | 0 (4.28) | −0.17 (5.3) | 0.2 (5.1) | (−2.6, 1.8) | −0.36 (.718) |
| | | (Min, Max) | (−7.7,7.3) | (−10.8,10) | (−10,10.3) | | |
| | | RMSE | 4.25 | 5.2 | 5.1 | | |
| | | MAD | 3.5 | 3.9 | 4.3 | | |
| | Left | | | | | | |
| | | $R^2$ | 85% | 75% | 72% | | |
| | | ME (SD) | 0 (3.7) | −0.8 (4.6) | 0.12 (4.9) | (−3.05,1) | −0.87 (.387) |
| | | (Min, Max) | (−6.5, 6.3) | (−10, 8.2) | (−9.9, 10) | | |

| | | | |
|---|---|---|---|
| RMSE | 3.7 | 4.6 | 4.9 |
| MAD | 3.01 | 3.8 | 4.1 |

**$R^2$**, R-squared value; **ME**, mean estimated error; **SD**, standard deviation; **Min**, the minimum estimated error; **Max**, the maximum estimated error; **RMSE**, root mean square error; **MAD**, median absolute deviation; **MD**, mean difference; **CI**, confidence interval.

**Stature determination models using left footprint measurements:**

- The comparison of individual errors between full and reduced models on the Malaysian training sets exhibited no statistically significant difference.
- The comparison of errors between the two test sets displayed no statistically significant difference.
- Using $R^2$, the predictive performance of the regression model was higher among the Malaysian test set (75% vs. 72%). The comparison of errors between the Malaysian and Egyptian test sets showed no statistically significant difference.

Overall, there are insignificant differences between the two populations regarding the errors in estimating stature in all the developed models **(Figure 2)**.

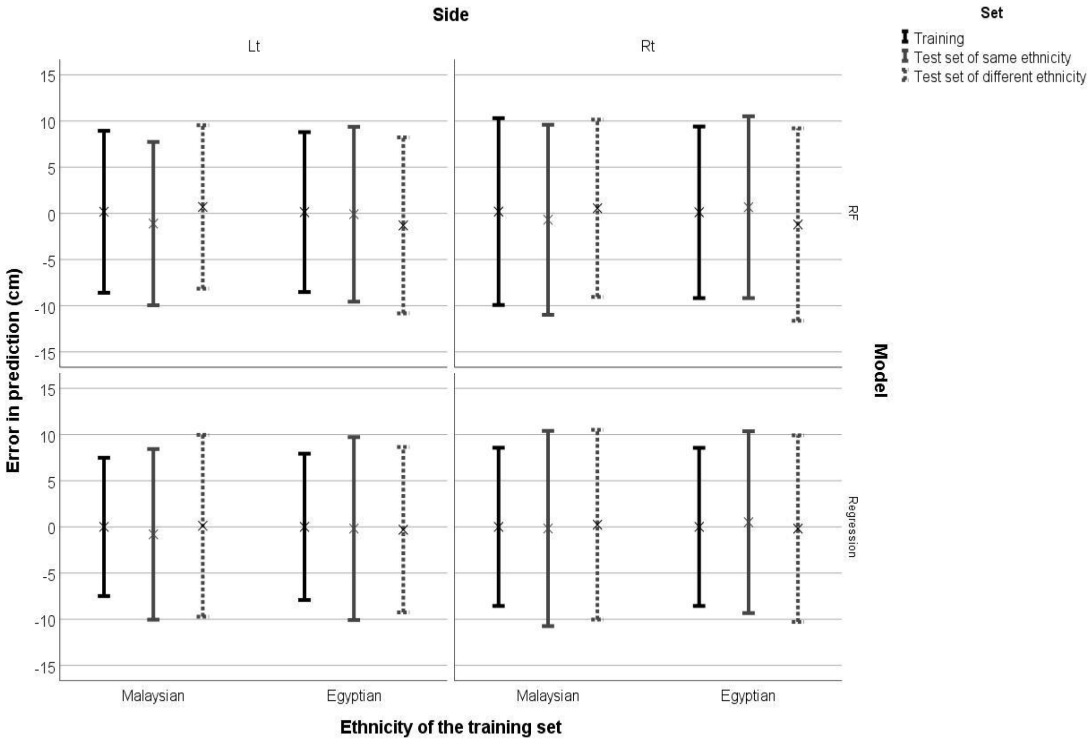


**Figure (2):** Error bar chart showing the distribution of errors in the training set (black line), two test sets (grey lines). Training set and first test sets are of the same ethnicity group, their error bar were represented by continuous lines. The second test set is of different ethnicity group [dashed grey error bar]. There are insignificant differences between the grey error bars.

## DISCUSSION

A footprint can give information even more than a fingerprint; nevertheless, it was reported that no field of detective science is so valuable and as ignored as the knack of tracing footsteps **(Webb et al., 2006)**.

The current study was designed to comprise the participants in the age group of 18–30 years, as it is the only group of age that warrants the slightest age-related influences on the body size of a person **(Krishan et al., 2015)**.

Conversely, the static footprints were collected from both the right and left feet of every contributor in the current study since numerous previous studies have recommended that the predilection of the left leg for standing and moving results in larger bone and muscle growth of the dominant leg in order to support the pressure, which leads to greater measurements **(Robbins, 1986; Fawzy & Kamal, 2010; Caplova et al., 2018)**. This was in contrast to the results of the current work that showed no significant differences between the measurements of both legs, except for big toe width measures. The current results are supported by the studies conducted by **Awais et al. (2018)** and **Shukla et al. (2017)**.

The current research studied the footprint measurements in two different populations (Egyptians and Malaysians), and it also highlighted whether sex and population differences influence a series of footprint measurements.

The current study revealed significant intersexual differences, and all footprint length parameters taken from the footprints were significantly greater in males than in females on both feet in the two studied populations. Additionally, the MFL of the males was nearly 2.35 cm greater than that of the females. Similar findings were reported in almost all previous studies conducted on the foot in different countries and on various populations.

By contrast, the comparative analysis of measurements between the two studied populations validated that overall, the Egyptians have significantly larger footprint measurements than those of Malaysians, except for FBB and big toe dimensions. The MFL in the Egyptians was significantly larger by roughly 0.86 cm, and the calculated footprint indices (HBI and FI) also showed significant differences between the two populations. This can be explained by the genetic factors that caused higher stature in Egyptians, which was shown clearly in the present study. Moreover, environmental and other conditions, such as routine barefooted walking and wearing of a certain kind of footwear, may influence the structure of the foot and lead to bodily and morphological disparities **(Delgado-Abellán et al., 2014; Hefner et al., 2014)**.

Regarding stature determination, the positive correlations of the measured height with the footprint measurements were previously described, and many regression equations were developed to predict height in different populations **(Fawzy & Kamal, 2010; Kanchan et al., 2012; Hemy et al., 2013; Abledu et al., 2016)**.

It is known in the forensic field that the models developed for the estimation of height from any racial population should be applied to the same one for best accuracy, but this rule has limitations as the identification of ethnicity is the greatest debatable inquiry that may face a forensic anthropologist when helping in identifying unidentified individuals. Subsequently, several problems with the full methodology for assigning individuals to a restricted number of ethnic groups have been described. Nowadays, there is a continuing effort to struggle away from racial stereotypes, converting the importance and validity of ethnic determination in forensic anthropology into a cross-disciplinary discussion. Anthropology claims that distinct races of human beings are not present **(Moorthy, 2015; M'charek et al., 2020)**.

Therefore, the main aim of the current study was to test if ancestry-specific equations could be applied on another ancestry with accepted accuracy, and this could be of help in cases when the investigator is not certain regarding ancestral affiliations. Nevertheless, it has to be remembered that the exact prediction of height from the footprint of a person may be an unattainable and needless aim and that there would constantly be an error of estimation a few centimeters.

Generally, forensic scientists categorize individuals into three main sets: Caucasoid, Negroid, and Mongoloid. Hence, the present study was conducted on two distinct populations (Egyptian and Malaysian) from two major different races (Caucasian and Mongoloid) and used two distinct statistical approaches, namely, RF and multiple regression analysis, to test the accuracy of building models in predicting stature.

The RF model is a robust algorithm, and it is a worthy and novel addition to the conventional statistics used in scientific anthropological research. Each tree is constructed using several randomness levels, and every tree is basically an independent model. Hence, the final model is not as liable to overfitting as in further classification methods. Generally, RF produces better outcomes, works well on a broad dataset, and can work with missing data by generating estimates for them. [28]

In spite of the presence of the differences according to sex, the correlations and prediction equations were established for both sexes (males and females), as it is a well-supported statement that equations for prediction, without indication of sex, are more consistent, particularly in actual practices, as it is not likely to identify the sex from the footprint visually.

Each model was built on 74% of one population and was tested on the remaining 26% of the same population and on the other population, and the results were presented separately by laterality as it was likely to visually differentiate the side of a footprint.

RF was built on the Egyptian sample with all the footprint measures as predictors, and the predictive power of the model was evaluated on two test sets: Egyptian and Malaysian test sets. Reduced model RF was built using maximum print length and footprint length measurements (T1, T2, T3, T4, and T5 lengths), as these measurements revealed the greatest correlation with stature. These findings are in accordance with several earlier researches, which have conveyed greater correlation coefficients with toe-heel length foot measurements than with other measurements, such as big toe breadth/length and breadth at ball/heel **(Robbins, 1986; Fawzy & Kamal, 2010)**.

The comparison of individual errors between full and reduced RF models showed no statistically significant difference, thereby indicating that the reduced model was similar to the full model with fewer parameters. Hence, the reduced model was selected as it can be applied in cases of partial footprints when not all the measurements are obtainable.

The present study affirmed that the predictive performance of the reduced RF model built on the Egyptians was almost the same among the Malaysian and Egyptian test sets. The comparison of errors between the Egyptian and Malaysian test sets displayed no statistically significant difference. Similar results were observed when using multiple regression analysis, and the predictive performance was overall the same in both populations with no statistically significant variation in the errors of stature prediction between two populations.

Moreover, when the equations for the valuation of stature from RF and multiple regression equations attained for the Malaysians were applied to the Egyptians, the rate of accuracy did not vary widely and there was no significant difference between errors.

Hence, despite the significant differences in footprint measures and the significant differences in stature between the two studied populations, there were no significant differences between the two populations regarding the individual errors in all of the developed models used to determine stature.

## CONCLUSION

The equations from measurements of footprint length for estimation of stature in the Egyptians and Malaysians showed accepted reliability and accuracy because the values of error were concluded to be not high. In contrast to other methods for estimating stature, the methods presented here could have satisfactory predictive performances if applied to individuals with ethnicities different from those of the individuals upon whom the model was trained.

Hence, previous studies might be more conservative in the generalization of their equations. From the practical point of view where race and sex are unknown, it is possible to predict stature from pooled equations.

## ABBREVIATIONS

| | |
|---|---|
| DLA | Designated longitudinal axis |
| MFL | Maximum footprint length |
| FBB | Footprint breadth at the ball |
| FBH | Footprint breadth at heel |
| BTW | Big toe pad width |
| BTL | Big toe pad length |
| HBI | Heel/ball index of the footprint |
| RF | Random forest |
| ME | Mean error |
| MAE | Mean absolute error |
| RMSE | The root mean square error |

**Statements and Declarations:**

**Funding**:

The authors received no financial support for this research, authorship, and/or publication of this article.

**Ethics approval and consent to participate:**

Approval was obtained from the Research Ethics Committee of Faculty of Medicine, Alexandria University Medicine (IRB number: 00012098, approval serial number: 0201044, dated 27/12/2017). Adherence to national and international regulations The Ethics Committee is constituted and operates according to ICH GCP Guidelines and applicable local and institutional regulations and guidelines that govern ethics committees' operation.

**Consent for publication:**

Not applicable as the research didn't include any individual person's data in any form.

**Availability of data and material**:

The authors confirm that the data supporting the findings of this study are available.

**Competing interests:**

None of the authors have any competing interests in the manuscript.

**Authors' contributions**: MA conducted the practical part of the research and helped in writing. SM, AE, and HA designed the study, interpreted the data, assisted in writing, and were in charge of the overall direction and planning. GA performed the statistical analysis.

**Acknowledgments:** None

## REFERENCES

**Abledu, J. K., Abledu, G. K., Offei, E. B., & Antwi, E. M. (2016).** Estimation of stature and body weight from footprint dimensions among a female population in Ghana. Australian Journal of Forensic Sciences, 48(2), 195-202.

**Awais, M., Naeem, F., Rasool, N., & Mahmood, S. (2018).** Identification of sex from footprint dimensions using machine learning: a study on population of Punjab in Pakistan. Egyptian Journal of Forensic Sciences, 8(1), 72.

**Breiman, L. (2001).** Random Forests. Machine Learning, 45(1), 5-32.

**Burrow, J. G. (2016).** The use of the Podotrack in forensic podiatry for collection and analysis of bare footprints using the Reel method of measurement. Science & justice : journal of the Forensic Science Society, 56(3), 216-222.

**Burrow, J. G., Kelly, H. D., & Francis, B. E. (2017).** Forensic podiatry–an overview. J Forensic Sci Crim Investig, 5(4), 1-8.

**Byard, R., & Payne-James, J. (2015).** Encyclopedia of forensic and legal medicine (2nd ed.). Oxford: Academic Press.

**Caplova, Z., Švábová, P., Fuchsová, M., Masnicová, S., Neščáková, E., Bodoriková, S. . . . & Beňuš, R. (2018).** Estimation of stature and body weight in Slovak adults using static footprints: A preliminary study. Legal medicine (Tokyo, Japan), 34, 7-16.

**Cutler, A., Cutler, D., & Stevens, J. (2012).** Random forests. In C. Zhang & Y. Ma (Eds.), Ensemble machine learning: methods and applications (p.p. 157-175). New York: Springer.

**Delgado-Abellán, L., Aguado, X., Jiménez-Ormeño, E., Mecerreyes, L., & Alegre, L. M. (2014).** Foot morphology in Spanish school children according to sex and age. Ergonomics, 57(5), 787-797.

**Fawzy, I. A., & Kamal, N. N. (2010).** Stature and body weight estimation from various footprint measurements among Egyptian population. Journal of forensic sciences, 55(4), 884-888.

**Hefner, J. T., Spradley, M. K., & Anderson, B. (2014).** Ancestry assessment using random forest modeling. Journal of forensic sciences, 59(3), 583-589.

**Hemy, N., Flavel, A., Ishak, N. I., & Franklin, D. (2013).** Estimation of stature using anthropometry of feet and footprints in a Western Australian population. Journal of forensic and legal medicine, 20(5), 435-441.

**IBM Corp. (2012).** IBM SPSS Statistics for windows, version 26.0. Armonk, NY: IBM Corp.

**Kanchan, T., Krishan, K., ShyamSundar, S., Aparna, K. R., & Jaiswal, S. (2012).** Analysis of footprint and its parts for stature estimation in Indian population. Foot (Edinburgh, Scotland), 22(3), 175-180.

**Krishan, K. (2008).** Estimation of stature from footprint and foot outline dimensions in Gujjars of North India. Forensic science international, 175(2-3), 93-101.

**Krishan, K., Kanchan, T., & DiMaggio, J. A. (2015).** Emergence of forensic podiatry--A novel sub-discipline of forensic sciences. Forensic science international, 255, 16-27.

**Krishan, K., & Sharma, A. (2007).** Estimation of stature from dimensions of hands and feet in a North Indian population. Journal of forensic and legal medicine, 14(6), 327-332.

**Liaw, A., & Wiener, M. (2002).** Classification and regression by randomForest. R news, 2(3), 18-22.

**Lohman, T., Roache, A., & Martorell, R. (1992).** Anthropometric standardization reference manual. Medicine & Science in Sports & Exercise. Washington, DC: Library of Congress.

**M'charek, A., Toom, V., & Jong, L. (2020).** The Trouble with Race in Forensic Identification. Science, Technology, & Human Values, 45(5), 804-828.

**Moorthy, T. N. (2015).** Estimation of Stature from Footprint Anthropometry Using Regression Analysis : A Study on the Bidayuh Population of East Malaysia. Arab Journal of Forensic Sciences and Forensic Medicine, 1, 13-21.

**Moorthy, T. N. (2017).** Neglected Physical Evidence during Crime Scene Investigation. Forensic Science & Addiction Research, 1, 1-2.

**Nirenberg, M. S., Krishan, K., & Kanchan, T. (2017).** A metric study of insole foot impressions in footwear of identical twins. Journal of forensic and legal medicine, 52, 116-121.

**R Core Team. (2013).** A language and environment for statistical computing.

Vienna, Austria: Foundation for Statistical Computing.

**Robbins, L. M. (1986).** Estimating height and weight from size of footprints. Journal of forensic sciences, 31(1), 143-152.

**Shukla, R., Lodha, D. A., & Das, S. (2017).** Stature Estimation from Footprint: A Study on Central Indian Population. European Journal of Forensic Sciences, 4, 1.

**Webb, D., Bernardo, D. V., & Hermenegildo, T. (2006).** Evaluating and improving footprint measurement: orientation and lengths. Anthropologie (1962-), 44(3), 269-280.

<u>الملخص العربي</u>

**اختبار نهج بديل لتحديد القامة من قياسات بصمة القدم غير معتمد علي التصنيف العرقي باستخدام عينة من المصريين والماليزيين البالغين.**

**مرام عاطف* سميه مدكور* أسماء البنا* غادة أبو شيعشع** هبه عبد السميع محمد حسين ***

*قسم الطب الشرعي و السموم الإكلينيكية كلية الطب جامعة الإسكندرية

**قسم المعلوماتية الطبية الحيوية والإحصاء الطبي، معهد البحوث الطبية، جامعة الإسكندرية، مصر

إن إحدى المشكلات الرئيسية في جميع الطرق السابقة تقريبًا لتقدير القامة هي أنه قبل تطبيق الطريقة، يجب أولاً تصنيف هذا الفرد غير المعروف لجنس أو مجموعة سكانية معينة. **الهدف من هذه الدراسة:** سلطت الدراسة الحالية الضوء على الاختلافات في قياسات الطول وبصمة القدم بين مجموعتين من عرقين مختلفين وتهدف إلى تقييم ما إذا كان الهروب من التصنيف العنصري ممكنًا وتحديد طول الأفراد غير المعروفين قبل تحديد عرقهم أثناء بناء ملفاتهم البيولوجية. **طريقة الدراسة:** اختبرت الدراسة هذا النهج البديل باستخدام منهجيتين إحصائيتين متباينتين، وهما الغابة العشوائية وتحليل الانحدار المتعدد. أجريت الدراسة على 200 متطوع بالغ، وتم توزيعهم على مجموعتين عرقيتين: 100 مصري بالغ و100 ماليزي بالغ. وقد تم تقسيم العينة الكاملة لكل مجموعة إلى مجموعتين فرعيتين هما مجموعتي التدريب والاختبار. تم استخدام مجموعة التدريب الفرعية لبناء معادلة، في حين تم تطبيق مجموعة الاختبار الفرعية للتحقق من صحة المعادلة المبنية. **النتائج:** بشكل عام، كانت هناك فروق ذات دلالة إحصائية بين المجموعتين فيما يتعلق بالأخطاء في تقدير الطول في جميع النماذج المتقدمة على الرغم من وجود اختلافات كبيرة في قياسات بصمة القدم والطول بين المجموعتين. **الخلاصة:** خلصت الدراسة الحالية إلى أن معادلات تقدير الطول التي تم عملها من مجموعة عرقية واحدة يمكن تطبيقها على مجموعة سكانية أخرى مع أداء تنبؤي مرضي، ويمكن لمعادلة واحدة أن تمثل أجزاء مختلفة من العالم.

**الكلمات المفتاحية:** الطول؛ بصمة القدم؛ غير خاص بالسكان؛ الماليزية؛ المصرية