Forensic Science International 226 (2013) 299.e1–299.e7

Contents lists available at SciVerse ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Forensic Anthropology Population Data

# From toe to head: Use of robust regression methods in stature estimation based on foot remains

Adrián Pablos [a,b], Asier Gómez-Olivencia [a,c,*], Alfonso García-Pérez [d], Ignacio Martínez [a,e], Carlos Lorenzo [a,f,g], Juan Luis Arsuaga [a,b]

[a] Centro Mixto UCM-ISCIII de Investigación sobre Evolución y Comportamiento Humanos, Avda. Monforte de Lemos 5 (Pabellón 14), 28029 Madrid, Spain
[b] Departamento de Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain
[c] PAVE Research Group, Division of Biological Anthropology, Department of Archaeology and Anthropology, University of Cambridge, Pembroke Street, Cambridge CB2 3DZ, UK
[d] Departamento de Estadística, Investigación Operativa y Cálculo Numérico, Universidad Nacional de Educación a Distancia (UNED), Paseo Senda del Rey 9, 28040 Madrid, Spain
[e] Área de Paleontología, Departamento de Geografía y Geología, Universidad de Alcalá de Henares, 28871 Alcalá de Henares, Madrid, Spain
[f] Area de Prehistoria, Universitat Rovira i Virgili, Avinguda de Catalunya 35, 43002 Tarragona, Spain
[g] IPHES, Institut Català de Paleoecologia Humana i Evolució Social, Pça Imperial Tarraco, 1, 43005 Tarragona, Spain

ARTICLE INFO

*Article history:*
Received 27 July 2012
Received in revised form 1 November 2012
Accepted 7 January 2013
Available online 4 February 2013

*Keywords:*
Robust regression
Metatarsal
Talus
Calcaneus
Least trimmed squares

ABSTRACT

Stature estimation is a standard procedure in the fields of forensic and biological anthropology, bioarchaeology and paleoanthropology, in order to gain biological insights into the individuals/populations studied. The most accurate stature estimation method is based on anatomical reconstruction (i.e., the Fully method), followed by type I regression equations (e.g., ordinary least squares – OLS) based on long bones, preferably from the lower limb. In some cases, due to the fragmentary nature of the osseous material recovered, stature estimates have to rely on other elements, such as foot remains. In this study, we explore stature estimation based on different foot bones: the talus, calcaneus, and metatarsals 1–4 in Afro- and Euroamericans of both sexes. The approach undertaken in this study is novel for two reasons. First, individual estimates for each bone are provided, and tarsals and metatarsals are combined in order to obtain more accurate estimates. Second, robust statistical methods based on type I regression equations are used, namely least trimmed squares (LTS). Our results show that the best individual bones for estimating stature are the first and second metatarsal and both the talus and the calcaneus. The combination of a tarsal and a metatarsal bone slightly improves the accuracy of the stature estimate.


## 1. Introduction

Together with body size and age, stature is a major parameter used to characterize individuals and populations [1–4]. In studies of fossil or skeletal remains, stature is often estimated using regression formulae based on modern reference collections. The proportions of the different skeletal parts that comprise stature vary during growth. These proportions can also be affected by external factors (e.g., nutrition and physical activity) and internal factors (e.g., genotype), and diachronic changes in stature at the population level can be related to short- to long-term stress [2,3,5,6].

There is a growing literature on stature estimation from skeletal remains due to the importance of this parameter in a number of fields such as biological and forensic anthropology, bioarchaeology and paleoanthropology [1,5,7–18]. In forensic cases, stature is very valuable in helping to identify the person to whom the human remains belong [9,12]. In biological anthropology, bioarchaeology and paleoanthropology, stature is a measure of body size, which is related to the biology of the individual/population/species under consideration [6,19–21].

Two main approaches have been used to estimate living stature from skeletal remains: "anatomical" and "mathematical" methods [5,22]. Anatomical methods, such as the Fully method, involve summing the height of all elements directly involved in stature (i.e., talus and calcaneus, the tibia, femur, first sacral vertebra, the vertebral column and the cranium) and incorporating a soft tissue correction factor in order to estimate living stature [13]. More recently Raxter et al. [16] revised the Fully method for estimating stature using ordinary least squares (OLS) regression. The anatomical method provides the most accurate estimate, since it takes into account almost all skeletal components of stature [1,13,16]. However, the potential applicability diminishes in archaeo-paleontological and forensic contexts where well-preserved, complete skeletons are not always recovered.

* Corresponding author at: Division of Biological Anthropology, Department of Archaeology and Anthropology, University of Cambridge, Pembroke Street, Cambridge CB2 3DZ, UK. Tel.: +44 346273373.
*E-mail addresses:* asiergo@gmail.com, ag665@cam.ac.uk (A. Gómez-Olivencia).

0379-0738/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.forsciint.2013.01.009

The ''mathematical'' approach estimates stature based on the dimensions of one or more bones that show a linear relationship with stature (but that may or may not contribute to determining stature itself) using ratios or regression equations. This second approach has the advantage of estimating stature from more incomplete skeletons. There are however several methodological problems and different ''schools'' within this approach.

First, due to the differences in body proportions between populations such as the relative lengths of the limbs and trunk, population-specific regression equations have been proposed [1,6–8,15,23]. Moreover, sex specific equations are preferred due to sex differences in body proportions. However, in some cases, especially with fragmentary remains or in paleoanthropological contexts, the ancestry (or body proportions) of the individual/sample being studied is unknown [17]. Thus, in paleoanthropological cases, pooled sex and multi-population regression formulae may be preferred (see [11]).

The second problem comes from a mathematical point of view. While ratios have been used to estimate stature [24], usually regression equations are used; both model I (ordinary least squares – OLS) and model II (major axis – MA and reduced major axis – RMA). Several authors have discussed the benefits of using model II regression equations [17,25,26]. In the specific case of stature, Sjøvold [17] recommends the use of RMA, in which the loss of precision is very slight when compared to OLS under normal circumstances. In fact Sjøvold extends the RMA principle to several populations producing the 'line of organic correlation', which provides better estimates in smaller and taller populations than OLS. Despite the fact that model II regressions may in certain cases yield more accurate stature estimates, there is an intrinsic problem with this approach according to Smith [27]. This author argues that method selection should be based on whether the relationship is symmetric (type II) or asymmetric (type I). The comparison of two variables falls in the former while the estimation of one variable from the other one falls in the latter. Thus, stature estimation should employ type I regressions (OLS), regardless of the pattern of error in the data [27].

Finally, there is a drawback that affects regression equations regardless of whether they are derived by type I or II methods: outliers. Extreme values (or outlier individuals) may exercise substantial leverage on the regression line, which may alter its equation. Robust regression methods deal with these outliers, ''silencing'' their effect in the regression analysis [28–30] (see below).

Long bones have been extensively used for stature estimation when the Fully method is not feasible (e.g., [11,15,17,18]). Long bones, and especially those directly involved in stature (femur and tibia), are preferred for stature estimation, but in paleoanthropological, archaeological and forensic contexts it is likely that they will be found fragmented, which precludes any accurate measurement. In these contexts, smaller bones such as those from the foot may be more likely to be found intact and thus can offer a useful alternative for estimating stature [31–35].

Stature estimation methods based on foot remains have been successfully applied to different fossil hominins from sites where long bones are poorly preserved or absent, yielding results consistent with estimates from other anatomical regions. Examples include the metatarsals from the early Pleistocene site of Dmanisi [32,36]; the tarsal and metatarsal bones from the early Pleistocene site of Gran Dolina-TD6 [33,37]; the metatarsals from the late-Middle Pleistocene of Omo-Kibish I [38]; the upper Pleistocene Neandertal metatarsal from Valdegoba [34] and the metatarsal from the upper Pleistocene of Klasies River [39]. These studies produce stature estimates based on different foot bones, mainly talus, calcaneus and metatarsals.

The talus and calcaneus are both components of stature and are measured in the Fully method [13,16]. However, there is not, to our knowledge, a regression equation to calculate stature based solely on the measurement used in Fully's method (i.e., articulated talo-calcaneal height). Furthermore, due to the correlation between certain foot variables (especially foot length) and stature, other measurements from the foot that are not components of stature have also been used to derive regression equations. There is a growing literature on stature estimation methods based on the total length of the foot [40,41] and on individual foot bones such as talus [14], calcaneus [1,7,14] and metatarsals [8,10,12,42].

### 1.1. Objectives

Several papers have already demonstrated the feasibility of using foot bones to estimate stature. In the bioarchaeological and paleoanthropological record it is easier to find complete metatarsal and tarsal bones than other complete long bones [31–35]. From a statistical viewpoint it is more correct to estimate stature from complete bones rather than from estimates of their length (see [43]). The main objective of this paper is to generate regression equations for estimating stature based on the metatarsals, the talus and the calcaneus independently, and based on combinations of the individual tarsal bones and the metatarsals. The novelty of the present approach is the use of robust type I regression equations, namely least trimmed squares (LTS). The differences between this method and the most commonly used type I (OLS) and type II regression equations (RMA) will be discussed. Finally, recommendations for the best performing bones and methods to calculate stature based on the talus, calcaneus and metatarsals are provided.

## 2. Materials

A total of 564 foot bones belonging to 94 individuals from the Hamann-Todd Osteological Collection (HTOC-HTH) were measured. These measurements derive from a high resolution metrical study used for comparative purposes (see [33]). These skeletons belonged to men and women born between 1825 and 1910 and are curated in the Cleveland Museum of Natural History (CMNH – Cleveland, OH, USA) [44]. The study sample is composed of 48 males (25 Euroamericans and 23 Afroamericans) and 46 females (21 Euroamericans and 25 Afroamericans). Age at death ranged from 17 to 50 years. No elements with obvious or suspected pathology that might have adversely affected the results were used. The stature of each individual was taken from the CMNH files (see below).

The composition of the present sample is similar to that of Byers et al. [10] in which both Euroamerican and Afroamerican males and females are represented. While the sample used here contains fewer individuals, it contains similar proportion of Euroamericans and Afroamericans, while the sample of Byers et al. [10] was mainly composed of Euroamericans. The seven groups proposed by Byers et al. [10] to calculate each of the formulae have been selected, namely: (1) pooled sex and ancestry data; (2) pooled ancestry, all females; (3) pooled ancestry, all males; (4) Afroamerican females; (5) Afroamerican males; (6) Euroamerican females and (7) Euroamerican males.

For the 94 individuals the maximum length of the talus, calcaneus and metatarsals 1–4 was measured. In a preliminary study the fifth metatarsal was included, but an exploratory analysis proved that it is a poor predictor of stature compared to other metatarsals. Moreover, Byers et al. [10] demonstrated the limited feasibility of estimating stature from this bone.

### 2.1. Stature corrections

The statures of the individuals of the present sample are those at the time of death. Todd and Lindala [45] and Cobb [46] provide information on how they were recorded. Previously, Dupertius and Hadden [47] observed that the statures obtained from the Hamann-Todd cadavers were equivalent to living ones. Thus, no correction has been used here.

With respect to age-at-death, a correction to obtain the maximum stature following Byers et al. [10] (based on [48]) has been applied. While all individuals in the sample have a recorded cadaveric stature, we estimate living (forensic) stature by adding 0.6 mm per year for individuals older than 30 (following [9,10,48]). This has resulted in a very small difference, due to the age composition of our sample. However, other authors suggest using an age-correction factor for individuals older than 40 years old [49].