

REVIEW

Open Access

Sex estimation using the human vertebra: a systematic review



Afiana Rohmani^{1,2}, Mohamed Swarhib Shafie² and Faridah Mohd Nor^{2*}

Abstract

Background: The vertebral column has been used in forensic studies for its weight-bearing function and relative density. Sex estimation is one of the essential elements in an anthropological examination, as it may narrow down the possibility of a match by half. Hence, it is crucial to derive the population-specific reference data in each vertebra for sex estimation. This systematic review explored the most sexually dimorphic vertebra by using the conventional anthropometric analysis.

Main body: An electronic comprehensive search was conducted using databases such as Scopus, Web of Science (WOS) and EBSCO Medline for relevant studies between 2008 and 2020. The main inclusion criteria were studies in English, and studies on sex estimation by morphometric analysis of vertebra by CT scan or dry bone. Only studies related to human adult age and vertebra were analysed. Literature search identified 84 potentially relevant articles, in which 19 articles had fulfilled the inclusion criteria. This review included studies on the cervical, thoracic and lumbar vertebrae in different populations.

Conclusion: The vertebral spine has demonstrated significant sexual dimorphism with variable prediction accuracies, whereby the body of a vertebra was found to be sexually dimorphic. It was shown that high accuracy of sex classification was provided by the second cervical, twelfth thoracic and first lumbar vertebrae, especially when they were used in combination.

Keywords: Systematic review, Sexual dimorphism, Vertebrae, Morphometric, Forensic

Background

Identification of unknown skeletal remains has been a challenge for forensic anthropologists, especially in mass disaster and advanced decomposition in human remains. Sex estimation is one of the most important elements in anthropological examination. It has impact on facilitating the identification of other skeletal parameters such as estimation of age at death, race, stature and relevance to studies of compounding biological factors such as pathological conditions, environmental and dietary habits (Cattaneo 2007; Marlow and Kozieradzka-Ogunmakin 2016).

It has been reported that the pelvic bone is the most accurate bone for estimation of sex (Bruzek and Murail 2006; Marlow and Kozieradzka-Ogunmakin 2016; Hora and Sládek 2018). By having sex estimation, a forensic scientist may be able to narrow down the possibility of a match by half (Robinson and Bidmos 2009). Accurate estimation of sex (100% accuracy) can be achieved based on the pelvis and skull combined (El Dine and El Shafei 2015). However, both pelvis and skull are not always preserved due to severe fragmentation and decomposition. Hence, exploration of sex differences in other bones is equally important to develop for future studies (Ramadan et al. 2017).

A vertebra comprises a combination of outer dense cortical bone and inner cancellous bone, that contribute to its weight-bearing function (Tan et al. 2004; Gülek et al. 2007; Yu et al. 2008; El Dine and El Shafei 2015).

* Correspondence: faridah.nor@ukm.edu.my; mnfaridah@gmail.com

²Forensic Unit, Department of Pathology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras, 56000 Kuala Lumpur, Malaysia

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Vertebral column has been used in forensic investigation due to its comparative thickness and capacity in resisting mechanical forces. In several cases, vertebral column was considered to be a well-preserved bone element (Hora and Sládek 2018; Padovan et al. 2019), as it usually survive in relatively harsh conditions such as bush-fire, mass disaster and other natural disruption (Blau and Briggs 2011).

Several studies have reported sexual dimorphism in different vertebrae based on vertebral size and shape, which contributed to the development of body stature and expanded development in vertebral transverse diameter (Taylor and Twomey 1984; Marlow and Kozieradzka-Ogunmakin 2016; Torimitsu et al. 2016). Individual bones have been studied for evaluation of sexual dimorphism including the cervical (Marlow and Pastor 2011; Bethard and Seet 2013; Amores et al. 2014; Torimitsu et al. 2016; Kaeswaren and Hackman 2019; Padovan et al. 2019; Rozendaal et al. 2020), thoracic (Yu et al. 2008; Hou et al. 2012; El Dine and El Shafei 2015)

and lumbar vertebrae by conventional metric and advanced morphological analyses (Zheng et al. 2012; Ostrofsky and Churchill 2015; El Dine and El Shafei 2015; Ramadan et al. 2017; Oura et al. 2018; Decker et al. 2019; Azofra-Monge and Alemán Aguilera 2020; Suwanlikhid et al. 2020).

The metric analysis is easy to manage by using the statistical methods (Pretorius et al. 2006). It involves collections of linear distances, angles and distance ratios from either dry bone or radiological images (Slice 2007; Omar et al. 2019). The morphological analysis is rather subjective as it explores two-dimensional (2D) or three-dimensional (3D) shapes for sex estimation in multiple bones (Teodoru-Raghina et al. 2017; Choong et al. 2020). In this review, studies on vertebrae are mainly focused on sexual dimorphism in different populations using the conventional metric analysis. This review summarised the morphometric parameters for sex estimation in different types of vertebrae by both radiographic method and dry bone measurements.

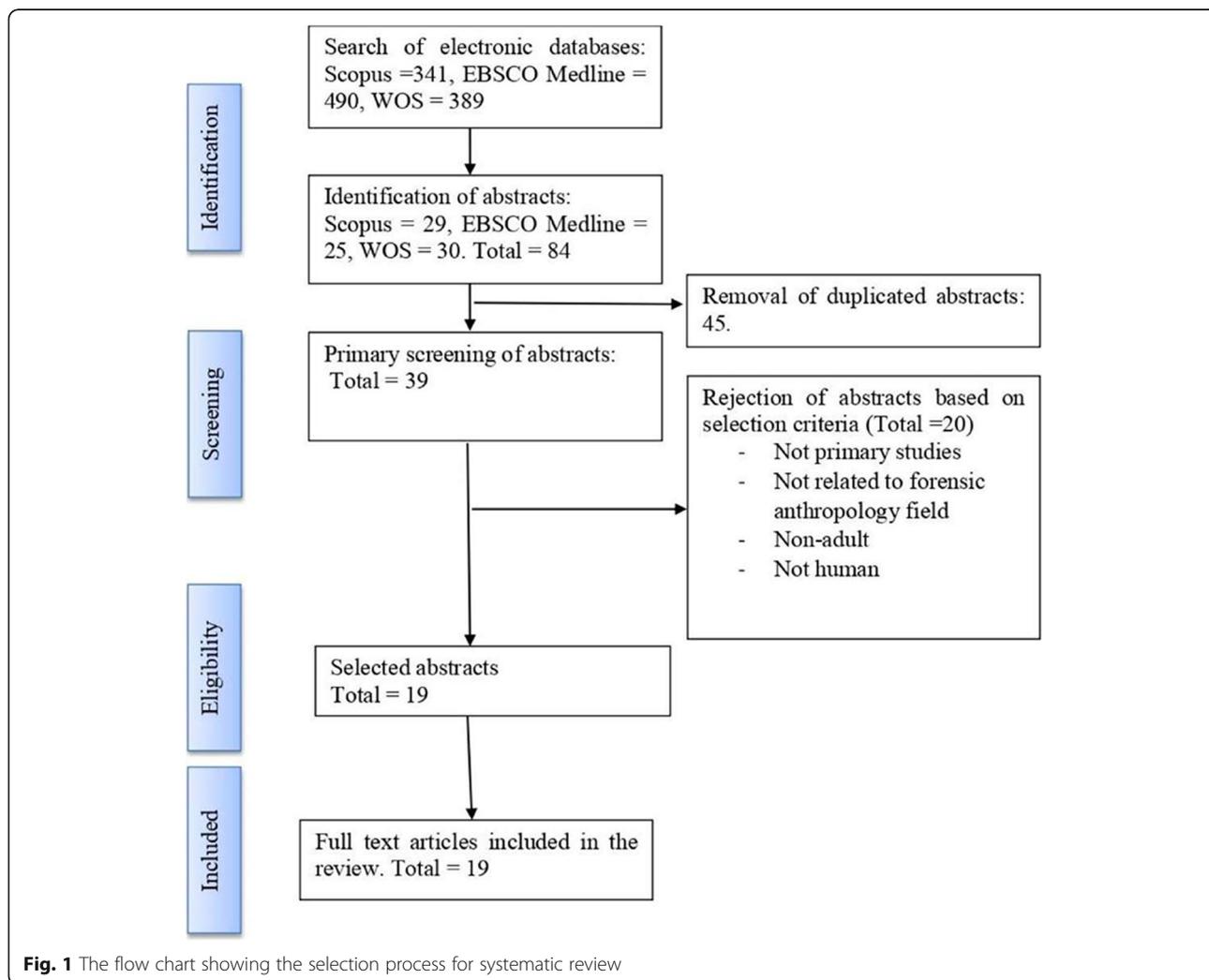


Fig. 1 The flow chart showing the selection process for systematic review

Main text

Methods

The systematic review search protocol PRISMA was adopted in this research study (Fig. 1). It has been registered in the Prospective Register of Systematic Reviews (PROSPERO) as the international database of registered systematic reviews (No: CRD42021252590).

Literature review

The focus of this review was to identify relevant research studies on sexual dimorphism in vertebral bones by morphometric analysis. Scopus, Web of Science (WOS) and EBSCO Medline were used to search for articles from the health science journals published between 2008 and 2020. The search strategy included a combination of three sets of words with truncation of an advanced search engine: (1) gender OR sex* (2) estimati* OR determin* OR assess* OR identificat* OR characterist* OR dimorphi* OR stud* (3) vertebr* OR spin* OR cervi* OR thora* OR lumba*.

Selection criteria

The results were selected from articles that were published in English including the abstracts. The studies selected were studies that performed morphometric analysis of human vertebrae for sexual estimation by computed tomography (CT) scan or dry bones. Articles such as reviews, news, editorials, letters or case reports were excluded from the review.

Data extraction and managing references

The articles were screened in three stages before they were included in the review. Firstly, title screening was done, and articles that were not suitable by the inclusion criteria were excluded. Next, the abstracts of the remaining articles were reviewed, and those that did not meet the inclusion criteria were excluded. Finally, the remaining articles were screened to exclude papers that were not within the scope of the literature. Duplicates were removed, and the remaining papers were selected again by at least two reviewers.

Before the data extraction phase, the reviewers will approve the full papers, that matched the inclusion criteria, and any discrepancies in opinions were discussed by the reviewers. Data extraction was performed independently for data validation by using the data collection form. The following data were recorded from these studies, i.e. title of the study along with the authors and year, types of vertebrae performed in the study, subjects or samples, study methods, results and remarks or conclusion of the study.

Inclusion and exclusion criteria

The inclusion criteria were primary studies, which included studies related to human adult age; studies related to vertebrae; studies by morphometric analysis such as calipers, CT scan or MRI; and studies using discriminant function analysis or regression analysis.

Results

The literature search identified 84 related articles. The reviewers evaluated the inclusion and exclusion criteria based on their titles and abstracts in all of the articles. Twenty papers were removed as they were non-relational types neither to single vertebral bone analysis, sexual dimorphism nor morphometric measurements, and these studies were not conducted within the field of forensic anthropology. The remaining 19 articles had matched with the inclusion and exclusion criteria, and hence were included in the review. The selection process of the systematic review was presented in a flow chart (Fig. 1).

Study characteristics

The descriptions of the articles were presented (Table 1). Briefly, all of the articles were published within the year 2008 until 2020, which comprised studies of vertebrae by 3D images or dry bones for estimation of sex. Based on bones types, seven studies were focused on cervical vertebrae (Marlow and Pastor 2011; Bethard and Seet 2013; Gama et al. 2015; Torimitsu et al. 2016; Kaeswaren and Hackman 2019; Padovan et al. 2019; Rozendaal et al. 2020), two studies on thoracic vertebrae (Yu et al. 2008; Hou et al. 2012) and seven studies on lumbar vertebrae (Zheng et al. 2012; Ostrofsky and Churchill 2015; Ramadan et al. 2017; Oura et al. 2018; Decker et al. 2019; Azofra-Monge and Alemán Aguilera 2020; Suwanlikhid et al. 2020). Two studies were done on thoracic and lumbar vertebrae (El Dine and El Shafei 2015; Garoufi et al. 2020), and one study on cervical and thoracic vertebrae (Amores et al. 2014).

In the review, nine studies used radiologic images to analyse the bones (Yu et al. 2008; Zheng et al. 2012; Hou et al. 2012; El Dine and El Shafei 2015; Torimitsu et al. 2016; Ramadan et al. 2017; Oura et al. 2018; Decker et al. 2019; Garoufi et al. 2020), whilst ten studies were conducted on dry bones (Marlow and Pastor 2011; Bethard and Seet 2013; Amores et al. 2014; Gama et al. 2015; Ostrofsky and Churchill 2015; Kaeswaren and Hackman 2019; Padovan et al. 2019; Azofra-Monge and Alemán Aguilera 2020; Suwanlikhid et al. 2020; Rozendaal et al. 2020). All of the studies used various experimental designs on several populations by conventional morphometric method, and were analysed by either discriminant function or regression analysis.

Table 1 The descriptive studies on sexual dimorphism in vertebral bone

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
1	Padovan et al. (2019)	C1 (atlas)	Brazil (110/81)	<ul style="list-style-type: none"> Four measurements were made in each atlas from samples of skeletal remains in the museum. The unpaired <i>t</i> test was used to test sexual dimorphism between the measurements. The stepwise-forward Wald method was used to obtain a logistic regression. 	<ul style="list-style-type: none"> The unpaired <i>t</i> test showed significant sex differences for all measurements ($p < 0.05$) Out of four variables, only two were selected by establishing the best model equation, i.e. anterior diameter of atlas and maximum transverse diameter of atlas. 	<ul style="list-style-type: none"> The atlas showed 81.2% accuracy in sex estimation.
2	Marlow and Pastor (2011)	C2	UK 153 known-sex individuals	<ul style="list-style-type: none"> Using Wescott's eight proposed sexually dimorphic dimensions plus an additional dimension width of the vertebral foramen (WVF), from samples of the museum collection of skeletal remains. The paired sample <i>t</i> test was used to determine significant differences between sexes. Utilize discriminant function analysis (DFA) to evaluate its efficacy for sex estimation. The stepwise technique chose a subset of factors based on squared partial correlation and the significance level from an analysis of covariance with the greatest number of separation capacity. 	<ul style="list-style-type: none"> The discriminant function by stepwise procedure, generated from the C2 vertebra measurements had the option to appraise sex with an accuracy of 83.3% with the option to arrange males and females with equivalent precision. By stepwise selection, maximum breadth across superior facets (SFB), maximum sagittal length (XSL), the width of the vertebral foramen (WVF), and dense sagittal diameter (DSD) form the discriminant functions. The dimension SFB is the best discriminator of sex. Testing the five discriminant functions equations by the Wescott's method resulted in an accuracy classification of 76.99%. The function of the present study has been tested with an overall classification accuracy of 83.3% ($p < 0.001$). 	<ul style="list-style-type: none"> The sex estimation method from Wescott's measurements of the C2 vertebra displayed significant discrimination between males and females with 83.3% accuracy.
3	Bethard and Seet (2013)	C2	American	<ul style="list-style-type: none"> Following the Wescott's measurements, five dimensions were measured from the skeletal remains collections. The Wescott's five discriminant functions were utilised to estimate sex in each case. 	<ul style="list-style-type: none"> The five discriminant functions by the Wescott's method were highly replicable, that yielded an overall accuracy of 80% or greater. The five discriminant functions by Wescott and Marlow & Pastor showed accuracy rates of 80% and more. 	<ul style="list-style-type: none"> The second cervical vertebra can provide a wealth of information in sex estimation.
4	Gama et al. (2015)	C2	Portuguese First sample = (99/91) (training set). Second sample = (24/23) (testing set)	<ul style="list-style-type: none"> Acquired 13 measurements from the skeletal collections by sliding calipers with an approximation of 0.5 mm. Performed a t-test (two-tailed) to analyse the differences in measurements between males and females. Using logistic regression model to construct estimation models. 	<ul style="list-style-type: none"> The most dimorphic dimensions were the LMA (11.18%) and DSMC (10.6%). The most predictive variables were LMA, DSMC, CMA and LMFS (right side). The resulting model identified sex in 89.7% of cases in the training set, whilst sex was accurately identified in 86.7% of cases in the testing set. 	<ul style="list-style-type: none"> The second cervical vertebra was useful for sex estimation with accuracies that ranged from 86.7 to 89.7%.
5	Torimitsu et al. (2016)	C2	Japanese (112/112)	<ul style="list-style-type: none"> Nine measurements were collected from cadavers by PMCT scanning and subsequent forensic autopsy was done. 	<ul style="list-style-type: none"> Measurements of DMFS and LMA on the C2 vertebra achieved expected cross-validated accuracies of 83.5% 	<ul style="list-style-type: none"> CT scan of C2 vertebra showed good estimation of sex with high accuracy rate.

Table 1 The descriptive studies on sexual dimorphism in vertebral bone (*Continued*)

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
6	Kaeswaren and Hackman (2019)	C2-C7	White Scottish (13/12)	<ul style="list-style-type: none"> ANOVA was utilised to examine mean differences between sex groups. Univariate discriminant function analysis (DFA) was performed on each variable to generate a formula for sex estimation. Using 25 human cadavers, three morphometric traits were measured for each cervical vertebra. A total of 150 cervical vertebrae were sampled. An independent (two-sampled) <i>t</i> test was conducted to establish sexual difference for all three measurements. Stepwise discriminant analysis was done to select the most dimorphic variable for sex estimation. 	<ul style="list-style-type: none"> and 83.1%, respectively. There were four variables (AS, DSD, DSMC, and DTMC), with correct prediction rates of 80%. A five-variable model with 92.9% accuracy rate was achieved. Discriminant functions utilised all three vertebral measurements i.e. the vertebral body height (CHT), anterior-posterior diameter of the vertebral foramen (CAP) and transverse diameter of the vertebral foramen (CTR) for sex estimation, with 77.3% - 100% accuracy for each vertebra. Sex was estimated with an accuracy of above 80% for each vertebra with C2 giving the most accurate sex estimation of 100% in all four combinations of vertebral measurements. Twenty-five discriminant functions out of 37 discriminant functions were significant with sex predicting accuracy greater than 80%. The CTR and CHT were found to contribute greatly towards biological sex variation. 	<ul style="list-style-type: none"> This study developed 25 multivariate discriminant functions that successfully classified individuals as male or female with accuracy greater than 80%.
7.	Rozendaal et al. (2020)	C1-C7	Europeans (157/138)	<ul style="list-style-type: none"> Samples were taken from the skeletal cemetery collections Three measurements were taken from each of the seven cervical vertebrae: Maximum cervical vertebrae body height (CHT), cervical anterior-posterior diameter (CAP) and cervical transverse diameter (CTR). Discriminant functions were generated for each cervical vertebra, using all three measurements, to establish whether sex could be estimated from a single vertebra. Multivariate discriminant functions were produced using all seven cervical vertebrae to investigate whether a combination of vertebrae may be used for sex estimation. The functions that achieved predicted accuracies of 80% or greater, were cross-validated on independent samples of 32 individuals from the skeletal collections. 	<ul style="list-style-type: none"> Results indicated that CAP measurement did not demonstrate sexual dimorphism, whilst CHT and CTR demonstrated significant difference between males and females ($p < 0.002$) (except for CTR of C1 and C2). Using combinations of all three measurements for sex estimation from a single vertebra, the accuracies ranged from 66.9 to 74% for males and 70.2–79.5% for females. This study produced seven discriminant function equations using 20 measurements from all seven cervical vertebrae, which achieved an overall accuracy rate of greater than 80%. The cross-validation test showed that among these functions, only four had achieved accuracies equal or greater than their predictive accuracies. The results indicated that C2 and C5 vertebrae were the most sexually dimorphic bones. 	<ul style="list-style-type: none"> The discriminant function equations achieved accuracy rates of 84.5% for cervical vertebrae (used in combination) in the European population.
8	Yu et al. (2008)	T12	Korean (52/50)	<ul style="list-style-type: none"> Used 33 linear measurements and two ratios by CT scan. An independent sample <i>t</i> test 	<ul style="list-style-type: none"> Twenty-three single variables with significant sex differences among 35 traits contributed to 	<ul style="list-style-type: none"> The 12th thoracic vertebra was sexually dimorphic with 90% accuracy rate in Korean

Table 1 The descriptive studies on sexual dimorphism in vertebral bone (Continued)

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
				<p>was performed to evaluate differences between the means of the parameters.</p> <ul style="list-style-type: none"> The respective discriminant equations were calculated for sex estimation using univariate, bivariate and stepwise methods of discriminant function analysis. Accuracies of the discriminant equations were obtained by cross-validation procedure. 	<p>correct sex classification rates ranging from 62.7 to 85.3%.</p> <ul style="list-style-type: none"> Three measurements on the vertebral body (sBDC, sBDcm, iBDcm, and iBDc) exhibited accuracies greater than 80%. The coronal diameter of the superior endplate of the vertebral body (sBDC), the ratio of the anterior to middle height of the body (Hm/HP), and the length of the left mammillary process and pedicle (IM&PL) predicted sex with 90% accuracy by DFA. 	<p>individuals.</p>
9.	Hou et al. (2012)	T12	Chinese (78/63)	<ul style="list-style-type: none"> Using 30 measurements from CT scan samples. The data were analysed by one-way analysis of variance (ANOVA). Univariate discriminant function analysis and stepwise discriminant function analysis were performed, respectively. A leave-one-out classification procedure was used to assess the validity of these functions. 	<ul style="list-style-type: none"> The accuracy of sex classification was between 56.4% and 90.1%. Variables such as sBDs, sBDsm, sVL, sBDC, iBDs, iBDsm, iVL, iBDC, mBDs, mBDc and BHp displayed 80% accuracy. The iVL had the highest accuracy rate of 90.1%. By stepwise discriminant function analysis, an equation with four variables, i.e. three linear measurements (superior maximum sagittal diameter of vertebral body endplate (sBDsm), inferior length of the whole vertebra (iVL), the distance between superior articular processes (sAD), and one ratio (the ratio of anterior to posterior height of the vertebral body (BHa/BHp) were obtained with 94.2% accuracy. 	<ul style="list-style-type: none"> The 12th thoracic vertebra was found to be sexually dimorphic in the North easterners in China with an accuracy rate of 94.2%.
10.	Amores et al. (2014)	C7 and T12	Mediterranean (61/60)	<ul style="list-style-type: none"> Using eight measurements from the skeletal collection samples. T-test was used to compare the data between sexes and evaluate the homogeneity of variance (<i>F</i> test). The effectiveness of the measurements was analysed by discriminant function analysis. The discriminant capacity of the selected variables was then evaluated using cross-validation procedure. 	<ul style="list-style-type: none"> For C7, discriminant function analysis selected the length of vertebral foramen (LVF), length of inferior surface of vertebra (LIVB), and width of inferior surface of vertebral body (WIVB) as having the highest discriminant power. For T12, the length of inferior surface of the vertebral body (LIBV) was selected for its discriminant capacity. The discriminant analysis yielded five functions, i.e. four for the 7th cervical and one for the 12th thoracic with 80% accuracy rate. 	<ul style="list-style-type: none"> The length of the vertebral bodies of the 7th cervical and 12th thoracic vertebrae offered the highest discriminant power for sex estimation. The C7 and T12 vertebrae showed a higher accuracy rate of approximately 80%.
11.	El Dine and El Shafei (2015)	T12 and L1	Egyptian (54/66)	<ul style="list-style-type: none"> Using 24 linear measurements and four ratios from the images of multi-slice computed tomography (MSCT) <i>t</i> test was used to establish the difference between sexes. Unstandardized coefficient. 	<ul style="list-style-type: none"> About 14 out of 24 linear measurements showed significant sex differences using T12 vertebra (predictive accuracy ranged from 49% to 85.5%), with three variables, i.e. lower endplate depth (EPDI), 	<ul style="list-style-type: none"> The T12 vertebra demonstrated a better sex estimation than L1 in the Egyptians. The accuracy increased when T12 and L1 were used in combination as sex predictors.

Table 1 The descriptive studies on sexual dimorphism in vertebral bone (*Continued*)

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
				<ul style="list-style-type: none"> Linear regression analysis was performed, in which each vertebral measurement was analysed for sex estimation. 	<ul style="list-style-type: none"> upper endplate width (EPWu), and superior vertebral length (VLs) as having more than 80% predictive accuracy rate. By using the L1 vertebra, only seven linear measurements and one ratio were sexually dimorphic (predictive accuracy ranged from 47% to 79%), and only the upper endplate depth (EPDu) showed an accuracy above 75% (79%). The accuracy of the T12 vertebra was 93.1%. The accuracy of the L1 vertebra was 68%. With a combination of T12 and L1 vertebrae, only five variables were used in the equation that predicted sex with high level of accuracy (96.3%). The study did not analyse the discriminant equation by cross-validation procedure. 	
12.	Zheng et al. 2012	L1	China (113/97)	<ul style="list-style-type: none"> About 29 linear measurements were taken from 3D CT models, and five aspect ratios were calculated from linear measurements. All measurements were significant by stepwise discriminant analysis ($p < 0.01$). Cross-validation of the discriminant function equations was performed to test the accuracy of the discriminant functions. 	<ul style="list-style-type: none"> About 25 traits demonstrated significant sexual dimorphism ranged from 57.1 to 86.6%. ($p < 0.01$) EPWu showed the highest predictive accuracy. Discriminant functions were upper endplate width (EPWu), left pedicle height (PHI), and middle endplate depth (EPDm) with predicted sex accuracy of 88.6%. 	<ul style="list-style-type: none"> The L1 vertebra may be used for sex estimation with 88.6% accuracy by discriminant function analysis.
13.	Ostrofsky and Churchill 2015	L1-L5	South Africa (47/51)	<ul style="list-style-type: none"> Samples comprised skeletal collection, and 11 measurements were taken to the nearest 0.1 mm with digital calipers. To compare male and female sample means, <i>t</i> test was applied (for variables normally distributed) and a nonparametric Wilcoxon ranks sum test was used (for data not normally distributed) Each variable that showed significant sex difference was subjected to univariate discriminant function analysis (DFA) to test its effectiveness for sex estimation. ($p < 0.01$) The leave-one-out cross-validation procedure was conducted to test the accuracy of the discriminant functions. 	<ul style="list-style-type: none"> The highest accuracy was obtained by the measurements of the vertebral body. Four variables of L1 and L2 vertebrae showed accuracies over 80%, i.e. vertebral body superior and dorso-ventral (BSDVD) and transverse diameters (BSTD). The discriminant functions predicted sex with accuracies over 80% for L1–L4, with the highest accuracy for L1 (87.1%). 	<ul style="list-style-type: none"> The lumbar vertebrae exhibited the greatest degree of sexual dimorphism, which may be used for sex estimation.
14.	Ramadan et al. 2017	L1	Egyptian (61/62)	<ul style="list-style-type: none"> About 15 linear measurements of L1 were taken by MSCT. An independent <i>t</i> test was applied to compare between 	<ul style="list-style-type: none"> Descriptive statistics showed significant differences between sexes for all measurements except the length of vertebral 	<ul style="list-style-type: none"> Sex could be estimated from L1 at 84.6% accuracy.

Table 1 The descriptive studies on sexual dimorphism in vertebral bone (*Continued*)

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
				<ul style="list-style-type: none"> different sexes. Correlation analysis was done followed by discriminant function analysis. 	<ul style="list-style-type: none"> foramen (LVF). The upper endplate width (EPWu) showed the highest accuracy. Additionally, sex could be predicted from L1 by discriminant analysis at an accuracy of 84.6%. The study did not analyse the discriminant equation for cross-validation procedure. 	
15.	Decker et al. (2019)	L1-L5	US-North American (76/78)	<ul style="list-style-type: none"> About 36 measurements were taken from abdominal CT scan comprising 30 linear measurements on vertebral body wedging angle, and five aspect ratios for each vertebra. A stepwise analysis method used the measurements to produce individual discriminant equations for L1 through to L5 and in combination studies of the lumbar vertebrae, and all accuracies were obtained by cross-validation process. 	<ul style="list-style-type: none"> L1 vertebra had 21 out of 29 significant measurements, with a predictive accuracy of 83.1%. L2 vertebra had 23 out of 29 significant measurements, with a predictive accuracy of 81.8%. L3 vertebra had 25 out of 29 significant measurements, with a predictive accuracy of 85.1%. L4 vertebra had 24 out of 29 significant measurements with a predictive accuracy of 82.5%. L5 vertebra had 23 out of 29 significant measurements with a predictive accuracy of 81.2%. The discriminant function for the five lumbar vertebrae had an overall accuracy rate ranged between 81.2 and 85.1% for sex estimation, with the highest accuracy attained by L3 vertebra. 	<ul style="list-style-type: none"> The study presented that the L1–L5 vertebrae can be used for sex estimation with accuracies ranging from 81.2 to 85.1% The accuracy rate rose to 92.2% when all five vertebrae were used in combination.
16.	Azofra-Monge and Alemán Aguilera (2020)	L1-L5	Spain (46/48)	<ul style="list-style-type: none"> Samples comprised identified adult individuals from the skeletal collections. Thirty three linear measurements were taken with digital calipers in millimetre. Gender differences between means were analysed by student's <i>t</i> test and non-parametric Mann-Whitney <i>U</i> test, and the binary logistic regression was generated. The equations with higher sex classification rates were selected and validation was performed on a random selection of 20% of the original sample. 	<ul style="list-style-type: none"> All measurements from the L1 and L2 vertebrae were greater in males than females, in which 11 variables were significant. ($p < 0.05$) The most sexually dimorphic vertebrae were L1 and L2. The total width (TW) was the variable selected in the equations for L1, L3 and L4 vertebrae. TW is the maximum distance between the ends of the transverse processes. 	<ul style="list-style-type: none"> The discriminant equations for sex estimation showed accuracy rates that ranged from 90.1 to 94.5% for L1, 85.4% to 89.4% for L2, 85.3% to 88.3% for L3, 85.3% to 88.2% for L4 and 80% to 85.3% for L5.
17.	Oura et al. (2018)	L4	Northern Finns Ages 20 = (147/228) Ages 30 = (147/228) Ages 46 = (618/745)	<ul style="list-style-type: none"> Samples were obtained from MRI scans. There were only three measurements from the vertebral body of L4 (width, depth and heights) as they were achievable in forensic context. A binary logistic regression was used for sex estimation and their accuracy was analysed. 	<ul style="list-style-type: none"> All the measurements were greater in men than in women. The multivariate regression analyses which included mean width, depth, and height of L4 vertebra, and yielded good sex accuracies in all age groups (86.4%, 87.7% and 82.8% at the ages of 20, 30 and 46, respectively). The classification accuracy for sex were consistently higher for females than males. 	<ul style="list-style-type: none"> The width, depth and height of the L4 vertebral body were found to be useful for sex estimation.

Table 1 The descriptive studies on sexual dimorphism in vertebral bone (*Continued*)

No.	Study	Types of bones	Sample population, number of subjects [male/female]	Methods	Results	Outcomes
18.	Suwanlikhid et al. (2020)	L1–L4	Thai (75/75)	<ul style="list-style-type: none"> • Samples of this study were dry bones that went through digital image processing after being captured by a camera. • Nine variables comprising the area, major and minor axes in total area, cortical and trabecular areas in each lumbar vertebra. • Quantitative variables and sex were analysed using discriminant function analysis. 	<ul style="list-style-type: none"> • The models based on the 46-year-old sample showed lower sex estimation accuracy than the corresponding 20- and 30-year-old sample models. • There were significant differences between sexes in most variables and males were generally greater than females in most measurement variables. ($p < 0.01$) • The upper endplate of the L1 vertebra had the most predictive precision (81.8%). Most of the upper lumbar vertebrae had higher precision than the lower lumbar vertebrae. 	<ul style="list-style-type: none"> • The lumbar vertebra can be used for stature and sex estimation in incomplete skeletal remains.
19.	Garoufi et al. (2020)	T1, T12, L1	European: Greek (119/96) and Danish (62/55)	<ul style="list-style-type: none"> • The samples comprised two modern contemporary European populations, i.e. the Greek sample of dry bones, and Danish sample of virtual vertebrae derived from CT 3D models. • Ten measurements were applied in each vertebra (T1, T12, L1). • The analysis of variance (ANOVA) was used to analyse sexual dimorphism in each variable, whilst the univariate discriminant function analysis (DFA) was performed to analyse sex estimation accuracy. 	<ul style="list-style-type: none"> • ANOVA showed that almost all of the measurements in three vertebrae (T1, T12, L1) were sexually dimorphic and the effect sizes were similar in two populations (Greek and Danish). The measurements, i.e. aVBH in T12 vertebra in Greek sample and aVBH in L1 vertebra in Danish sample did not yield any difference for sex estimation. • The discriminant function equations showed acceptable accuracies from cross-validation process with T1 showing the highest classification accuracy. These equations were based on three measurements, i.e. IEPW, mVL and mTD with 89.9% accuracy (Danish sample) and pVBH, mVL and mTD with 88.8% accuracy (Greek sample). 	<ul style="list-style-type: none"> • Among the three vertebrae, T1 was the most reliable for sex estimation in both populations reaching almost 90% cross-validated accuracy. • The maximum vertebral lengths (mVL) of the T1, T12, L1 vertebrae showed the highest degree of sexual dimorphism.

Cervical vertebrae for sex determination

A total of eight articles had focused on cervical vertebrae (Marlow and Pastor 2011; Bethard and Seet 2013; Amores et al. 2014; Gama et al. 2015; Torimitsu et al. 2016; Kaeswaren and Hackman 2019; Padovan et al. 2019; Rozendaal et al. 2020). Most metric studies of sexual dimorphism were conducted on the second cervical (C2) vertebra, except for one study that emphasised on the first cervical (C1) vertebra (Padovan et al. 2019). Whilst one study had utilised both 7th cervical vertebra (C7) and twelfth thoracic (T12) vertebrae (Amores et al. 2014), other studies had used all of the cervical vertebrae (C1–C7) for sex classification (Kaeswaren and Hackman 2019; Rozendaal et al. 2020).

The C2 vertebra, also known as the axis, was commonly employed for sex estimation, as it has well-

described morphological characteristics, and are mostly well-preserved even in adverse conditions (Gama et al. 2015). Besides being sexually dimorphic, C1 and C7 vertebrae were also found to be useful for sex estimation (Amores et al. 2014; Padovan et al. 2019).

An older publication by the Wescott's study (2000) was excluded from the review as it has undergone re-evaluation by many researchers in many sex estimation studies (Wescott 2000; Marlow and Pastor 2011; Bethard and Seet 2013). Continuous analysis and re-evaluation is vital to increase accuracy and precision of the existing methods. Marlow and Pastor (2011) had adopted and measured the Wescott's eight projected measurements and the width of vertebral foramen (WVF) on C2 vertebra in adults from the Spitalfields' anatomical collections held at the National History Museum, London. The

discriminant function analysis showed a high percentage of 83.3% for the classification of male and female, which was higher than that achieved by Wescott (2000). Previous conclusions by Wescott (2000) and Marlow and Pastor (2011) were re-evaluated on the modern American skeletal collections in Tennessee, in which C2 vertebra showed an accuracy rate of 86.7% (Bethard and Seet 2013).

Gama et al. (2015) had created a logistic regression model on C2 vertebra in the Portuguese population, and developed simple predictive model based on logistic regression models with accuracies, that ranged from 86.7 to 89.7%. In this study, about 13 dimensions of C2 vertebra were measured by adopting the Wescott's method (2000, which included the sagittal maximum body diameter (DSMC), maximum width of the axis (LMA), maximum width of the right superior facet (LMFSD) and maximum length of the axis (CMA).

Torimitsu et al. (2016) had conducted CT measurements on C2 vertebra to provide good classification of sex. In this study, two variables, i.e. maximum distance between superior articular facets (DMFS) and LMA, were demonstrated to have accuracy rates of 83.5% and 83.1%, respectively. Other dimensions of C2 vertebra such as the maximum length of the axis (XSL), maximum width of vertebral foramen (WVF), odontoid process sagittal diameter (DSD) and maximum distance between superior facets (SFB) were analysed for biological sex estimation (Marlow and Pastor 2011; Gama et al. 2015; Torimitsu et al. 2016).

Methods for sex estimation is important as many of the bones retrieved have often undergone severe fragmentation upon recovery. In the event of bone fragmentation due to advanced decomposition, the results from these studies have to be taken with caution. This may happen in conditions, whereby the vertebra and its parts such as the spinous process, transverse process and superior articular facets may be damaged, hence rendered immeasurable for research study or casework investigation. In corroboration with such circumstances, a study was conducted on wet disarticulated cervical vertebrae (C1–C7) from the white Scottish human cadavers (Kaeswaren and Hackman 2019), whilst another study was done on all of the cervical vertebrae from the cemetery collections of European ancestry (Rozendaal et al. 2020). Both studies revealed that both C2 and C5 vertebrae were sexually dimorphic with good sex predictors in two measured variables, i.e. the vertebral body height (CHT) and transverse diameter of vertebral foramen (CTR).

It is worthwhile to note that high classification rates of at least 80% were considered useful in sex estimation methods (Torimitsu et al. 2016). In many studies, measured variables of C2 vertebra including the length (Marlow and Pastor 2011; Bethard and Seet 2013; Gama

et al. 2015), width (Gama et al. 2015; Torimitsu et al. 2016), sagittal diameter of the vertebral body (Gama et al. 2015; Torimitsu et al. 2016), and the distance between superior articular facets were demonstrated to have high discriminant power with accuracies exceeding 80% (Marlow and Pastor 2011; Torimitsu et al. 2016). In other studies, the height of the vertebral body and transverse diameter of the foramen provided a high degree of sexual dimorphism (Kaeswaren and Hackman 2019; Rozendaal et al. 2020). Amores et al. (2014) have demonstrated good discriminant power with 80% accuracy on the measured variables of C7 vertebra from the Southern Spain laboratory skeletal collections, which comprised the length of the vertebral foramen (LVF), width and length of the inferior vertebral body (LIVB and WIVB).

Thoracic vertebrae for sex estimation

Under the subject heading of thoracic vertebrae, five articles were identified (Yu et al. 2008; Hou et al. 2012; Amores et al. 2014; El Dine and El Shafei 2015; Garoufi et al. 2020). Besides being a transitional vertebra, it has distinct morphological characteristics, which can be easily identified in disarticulated skeleton. Three articles had focused on a combination study of the twelfth thoracic (T12), C7, first lumbar (L1) and first thoracic (T1) vertebrae (Amores et al. 2014; El Dine and El Shafei 2015; Garoufi et al. 2020). In different populations, T12 vertebra was shown to be sexually dimorphic with an accuracy rate exceeding 83%. The T1 vertebra had the highest degree of sexual dimorphism (88.8% accuracy), followed by T12 vertebra (84.2%) (Garoufi et al. 2020).

CT scan analysis of the coronal diameter of superior endplate of the vertebral body (sBDc), ratio of the anterior to middle height of the vertebral body (BHm/BHp) and length of left mammillary process and pedicle (IM and PL) of T12 vertebra in the Korean population showed a high degree of sexual dimorphism with 90% accuracy rate (Yu et al. 2008).

Additionally, analysis of variables from CT images of T12 vertebra in the Chinese population showed a high degree of sexual dimorphism with 94.6% accuracy (Hou et al. 2012). These variables included superior maximum sagittal diameter of vertebral body endplate (sBDsm), inferior length of the whole vertebrae (iVL), the distance between superior articular processes (sAD) and one ratio (the ratio of anterior to posterior height of the vertebral body (BH_a/BH_p). Amores et al. (2014) have demonstrated that the length of the inferior surface of the vertebral body (LIVB) of T12 vertebra showed the highest dimorphism values with 80% accuracy.

El Dine and El Shafei (2015) performed a study by adopting the method by Yu et al. (2008), in which 24 linear measurements and 4 ratios by multi-slice computed

tomography (MSCT) were utilised on T12 vertebra in the Egyptian population. From the analysis, 14 measured variables have exhibited significant sex differences, and produced more than 80% predictive accuracy in three variables, i.e. the lower endplate depth (EPDI), upper endplate width (EPWu) and superior vertebral length (VLs). By regression analysis, this study had generated 93.1% accuracy for sex estimation, which was comparable to that by Yu et al. (2008) (90%).

Yu et al. (2008) conducted a study on T12 vertebra for sexual classification. In the study, the vertebral body endplate and sagittal length of the vertebra of T12 vertebra were found to be sexually dimorphic with accuracies exceeding 80%. The measured variables comprised sBDc, superior maximum coronal diameter of endplate of the vertebral body (sBDcm), coronal diameter of endplate on inferior plane (iBDc) or maximum coronal diameter of endplate on inferior plane (iBDcm). Additionally, coronal diameter of the superior endplate vertebral body (sBDc), ratio of the anterior to middle height of the body (BHm/BHp) and length of the mammillary process and pedicle of T12 vertebra (IM and PL) were also found to be sexually dimorphic with 90% accuracy rate (Yu et al. 2008).

In a study of T12 vertebra in the Chinese population by Hou et al. (2012), the vertebral body measurements (sBDs, sBDsm, iBDc, etc.) demonstrated accuracies exceeding 80%, with the sagittal length of the vertebra (iVL) showing 90% accuracy. By discriminant function analysis (DFA), a discriminant equation was produced with 94.2% accuracy, which was based on variables such as superior sagittal diameter of vertebral body endplate (sBDsm), ratio of anterior to posterior height of the vertebral body (BH_a/BH_p) and non-vertebral body measurements, i.e. the distance between superior articular process (sAD) and iVL. Besides, the variables such as vertebral sagittal lengths (iVL and sVL), which measured distances from the anterior edge of the vertebral body to posterior edge of vertebral spinous process at the inferior and superior planes were found to be highly accurate, with iVL yielding 90% highest accuracy rate (Hou et al. 2012).

El Dine and El Shafei (2015) conducted a study on T12 and first lumbar (L1) vertebrae in the Egyptians. Results revealed that sVL and iVL (vertebral sagittal length) showed 93.1% accuracy rate by multiple regression analysis. Also, the lower endplate depth (EPDI), upper endplate width (EPWu) and superior sagittal length vertebral (sVL) demonstrated an accuracy exceeding 80%.

Garoufi et al. (2020) performed a study on three vertebrae, i.e. T1, T12, L1 vertebrae in the Greek population, and a high degree of sexual dimorphism was reported in T1 vertebra (90% cross-validated accuracy) based on the

maximum vertebral length (mVL) and maximum distance (mTD). Accuracies ranging from 75 to 83% accuracy was demonstrated by T12 and L1 vertebrae (Garoufi et al. 2020).

Amores et al. (2014) had studied C7 and T12 vertebrae in the Mediterranean population, and demonstrated a high degree of sexual dimorphism with 80% accuracy rate by discriminant function analysis. The equations were based on the length of the inferior surface of the vertebral body and the width and length of the vertebral foramen of C7 vertebra and the length of the inferior surface of the vertebral body (LiBV) of T12 vertebra.

It can be concluded that the vertebral body and sagittal length of the thoracic vertebra played a crucial role for estimation of sex. The accuracy to predict sex correctly may be achieved when the vertebrae are complete and intact, particularly the sagittal length of the vertebra, that is formed by the body of a vertebra with spinous process.

Lumbar vertebrae for sex estimation

The first lumbar vertebra (L1) was shown to be sexually dimorphic by discriminant function analysis (Zheng et al. 2012; El Dine and El Shafei 2015; Ramadan et al. 2017), in which L1 had achieved the highest accuracy among all of the lumbar vertebrae (Ostrowsky and Churchill 2015; Decker et al. 2019; Azofra-Monge and Alemán Aguilera 2020; Suwanlikhid et al. 2020). Oura et al. (2018), on the other hand, had demonstrated that the fourth lumbar vertebra (L4) showed the highest degree of sexual dimorphism in the Northern Finn population.

Analysis of 29 linear measurements and 5 ratios from 34 traits on L1 vertebra in the Chinese population revealed accuracies ranging from 57.1 to 86.6% (Zheng et al. 2012). About five linear measurements associated with the vertebral body (EPWu 86.6%, EPDm 86.2%, EPWl 85.2%, EPDI 84.3% and EPDu 83.3%) gave an accuracy rate exceeding 80%, with EPWu showing the highest accuracy. The discriminant function produced was based on the upper endplate width (EPWu), middle endplate depth (EPDm) and left pedicle height (PHl) with an accuracy of 88.6%.

El Dine and El Shafei (2015) had studied on L1 and T12 vertebrae, and demonstrated varying degree of sexual dimorphism with accuracies ranging from 47 to 79% based on seven linear measurements and one ratio on L1 vertebra. By comparison, Zheng et al. (2012), who focused on L1 vertebra alone, had achieved higher accuracy rate (57.1–86.6%) compared to that by El Dine and El Shafei (2015) (79%) based on the upper endplate depth (EPDu). Also, Zheng et al. (2012) yielded higher accuracy rate (88.6%) based on three dimensions, i.e. upper endplate width (EPWu), left pedicle height (PHl) and middle endplate depth (EPDm) on L1 vertebra

compared to that by El Dine and El Shafei (2015) (68% accuracy), based on seven linear measurements and one ratio on L1 vertebra. By utilising both L1 and T12 vertebrae in combination, a high accuracy rate of 96.3% was demonstrated by El Dine and El Shafei (2015).

Ramadan et al. (2017) analysed 15 linear measurements on L1 vertebra in the Egyptians by CT scan, and by adopting the methods by Zheng et al. (2012) and El Dine and El Shafei (2015), results showed that EPWu demonstrated an accuracy of 84.6%, and nearly all of the measurements were significantly greater in males than females by discriminant function analysis.

Ostrowsky and Churchill (2015) performed physical osteological examination (POE) on all of the lumbar (L1–L5) vertebrae in the South Africans, and revealed that sex was correctly predicted in L1 until L4, with accuracies exceeding 80%. Measured variables such as the vertebral body superior dorso-ventral diameter (BSDVD) and body superior transverse diameter (BSTD) on L1 vertebra gave over 80% accuracy rate (87.1%) by discriminant function analysis (Ostrowsky and Churchill 2015), which was comparable to other studies based on measured variables, i.e. EPDm and EPWu/l (Zheng et al. 2012; Ramadan et al. 2017).

Studies on the lumbar vertebrae have also been shown to be useful for sex estimation (Decker et al. 2019; Azofra-Monge and Alemán Aguilera 2020; Suwanlikhid et al. 2020). Decker et al. (2019) had utilised abdominal CT scan on living patients in a modern adult population, whilst Azofra-Monge and Alemán Aguilera (2020) used dry bones samples from the laboratory skeletal collection in Spain. All of the studies displayed similar trends of accuracy by discriminant function analysis, which was comparable to a study by Ostrowsky and Churchill (2015). Similarly, Suwanlikhid et al. (2020) had analysed dry bones from the laboratory skeletal collections in Thailand and measured nine variables, i.e. area and surface axis on all of the lumbar vertebrae. They documented that precision rates were higher in the upper lumbar vertebrae than those in the lower lumbar vertebrae, which corroborated with the findings by Decker et al. (2019).

The body of a vertebra was found to be sexually dimorphic by discriminant function analysis, as evidenced from the measured variables, i.e. upper endplate depth (EPDU) and upper endplate width (EPWU) (Zheng et al. 2012; Ostrowsky and Churchill 2015; El Dine and El Shafei 2015; Ramadan et al. 2017; Suwanlikhid et al. 2020). Physical osteological examination (POE) on L1 vertebra in the South African population revealed that EPWU had attained an accuracy of 87.1% (Ostrowsky and Churchill 2015), whilst EPDU showed a lower accuracy of 81.8% (Suwanlikhid et al. 2020).

Decker et al. (2019) studied on 30 linear measurements, wedging angle and five aspect ratios on all of the lumbar vertebrae in the North American population, and revealed an overall accuracy of 81.2% to 85.1% by discriminant function analysis, with the highest accuracy achieved by L3 (85.1%). By using multilevel measurements, a higher accuracy of 92.2% was achieved (Decker et al. 2019).

Oura et al. (2018) study on the width, depth and heights of L4 vertebra in the Northern Finns by MRI scan, and showed that they were sexually dimorphic with accuracy rates exceeding 80% in all three age groups (20, 30 and 46 years). The L4 vertebra was chosen as it was easily accessible in both axial and sagittal slices in MRI scans, and results showed lower accuracy rate (82.8%) in the 45-year-olds sample population compared to that in the 20- or 30-year olds (86.4% and 87.7%, respectively) (Oura et al. 2018).

Discussion

On initial screening of the abstracts and full texts, 19 studies were identified (Fig. 1). Forensic anthropologists have relied heavily on the vertebrae for estimation of sex in forensic caseworks. From 24 vertebrae (excluding sacrum and coccygeal), C2, T12 and L1 vertebrae were predominantly utilised for estimation of sex in several populations. Although the cranium and pelvis were commonly used for sex estimation, acceptable accuracies have been shown by studies on vertebral bones. Rozenaal et al. (2020) had reported accurate sex estimation in a combination study of all of the cervical vertebrae (C1 until C7). A total of 25 discriminant functions were generated from each vertebra with 80% accuracy, and 100% accuracy was obtained when both 4th cervical vertebra (C4) and C2 vertebrae were combined (Kaeswaren and Hackman 2019). Similarly, the accuracy in sexing skeleton varied upon using L1 only, compared to using both L1 and T12 vertebrae, in which the accuracy rose from 68 to 93.3% (El Dine and El Shafei 2015). It was also demonstrated that the accuracy rose from 81.2 to 85.1% in individual study of each lumbar vertebra to about 92.2% when all of the lumbar vertebrae (L1–L5) were used in combination (Decker et al. 2019).

There was a significant variation in accuracy from different types of studies in different populations. The development of metric analysis for sex estimation are usually population-specific; hence, it can only be used for individuals from the same population (Bruzek and Murail 2006). Population variation and scientific study of inherited human variation may only enhance the accuracy and reliability of the method used and their outcomes (Kimmerle et al. 2008). For instance, L3 vertebra in the North American population showed an accuracy of 85.1% (Decker et al. 2019), whilst L3 vertebra in the

South African population showed an accuracy of 87.1% (Ostrofsky and Churchill 2015).

Discriminant function analysis (DFA) is a common statistical analytical technique applied in forensic anthropology (Gama et al. 2015). The discriminant function is generated by multiplying the coefficients with variables of the vertebral measurements. The discriminant score is obtained by having a value that will act as the cut-off point between males and females, also known as the sectioning point. If the scores are greater than the sectioning point, it will be predicted as male, whilst scores smaller than the sectioning point, will be predicted as female (Omar et al. 2021). However, some studies have used logistic regression analysis, instead of DFA (Gama et al. 2015; El Dine and El Shafei 2015; Oura et al. 2018; Azofra-Monge and Alemán Aguilera 2020). It is a discriminative statistical model analogous to DFA, whilst logistic regression analysis is considered more robust and flexible in terms of data assumptions (Gama et al. 2015; Kiales et al. 2020).

Discriminant analysis models should be subjected to evaluation and cross-validation analyses to estimate the accuracy of the classification, and to provide good estimation of the performance model. Cross-validation is one of the most practical methods to estimate the predictive model performance (Kuligowski et al. 2016). In some studies, the methods were created and tested against the sample population, but were not cross-referenced to other populations.

Some studies used cross-validation analysis to evaluate the discriminant capacity of their classification models (Yu et al. 2008; Zheng et al. 2012; Amores et al. 2014; Ostrofsky and Churchill 2015; Decker et al. 2019; Garoufi et al. 2020), other studies did not perform cross-validation analysis (El Dine and El Shafei 2015; Ramadan et al. 2017; Suwanikhid et al. 2020). Several studies used both training and test samples to validate their discriminant functions, whereby the training sample is used to develop sex prediction models, and the model built on the training sample is applied to the test sample for prediction. In one study, discriminant functions were validated in both the training samples (190 individuals) and test samples (47 individuals), whereby sex was correctly estimated in 89.7% and 86.7% of cases, respectively (Gama et al. 2015). However, the model showed gender-biased towards males in the training sample.

Two studies by Marlow and Pastor (2011) and Bethard and Seet (2013) had done re-evaluation of the discriminant functions for sex estimation on C2 vertebra on the Hamann-Todd and Terry osteological collections of black and white specimens by adopting the Wescott's method. Analysis on the contemporary Americans in the Tennessee skeletal collections showed 80% accuracy rate (Bethard and Seet 2013), and by re-evaluating the

Wescott's method on the eighteenth–nineteenth centuries European ancestry museum collection in London, UK, an overall classification accuracy of 76.99% was reached (Marlow and Pastor 2011).

Along with the dynamics of growth processes, vertebral anatomy can be directly affected by growth. The growth process can cause numerous changes in the bone structure and surrounding tissue. From the onset of puberty, the duration can be quite variable due to considerable variation in growth, whereby females tend to have a vertical growth spurt, whilst males tend to have a horizontal growth spurt (Taylor and Twomey 1984; Azofra-Monge and Alemán Aguilera 2020). The vertebrae are mostly predisposed to mechanical forces as they form the backbone of the body and with a small amount of rotation of the vertebral column, it may contribute to a change in size and shape of the vertebrae. Diet is an important factor that may influence bone balance and quality, which is particularly true in the elderly due to factors such as poor diet or energy balance. Whilst environmental factors and genes may affect growth hormone function and bone control development, other factors such as dietary pattern, daily physical activity and mechanical loading may contribute to changes in bone density, mass and strength and hence sexual variance (Torimitsu et al. 2016; Gilsanz et al. 2018; Muñoz-herandez et al. 2018).

Based on linear measurements of the vertebrae, males were found to have greater bone dimensions than females. Linear measurements represent bone size, whilst ratios represent bone shape, which is formed by a combination of several linear measurements (Hou et al. 2012). Studies on ratios of linear measurements have shown that accuracies of ratios were statistically lower than linear measurements (Yu et al. 2008; Zheng et al. 2012; Hou et al. 2012; El Dine and El Shafei 2015). Further analysis of sexual dimorphism may also be done by geometric morphometric analysis for the analysis of structural modularity and optimization of shape classification criteria.

There are two types of measurement techniques on bones such as CT-derived methods, and dry bone measurements. Techniques within the field of anthropology have relied heavily on metric variables on dry bones, which may be derived from either anthropological museum collections or laboratories, and this requires an extensive effort and resources to improve access to a good range of biological samples. In contrast, CT-derived methods have been demonstrated to be more easily available, minimally invasive and more efficient for research use and identification purposes, for both dead bodies and living human subjects. Besides, a difference of only 2 mm in measurement error was reported between dry bone measurement and imaging derived

method (Stull et al. 2014). This emphasizes that both methods are acceptable and feasible to be used by anthropologists as a tool for classification of sex.

Strength and limitation of the review

Many studies have demonstrated the impact of sex estimation formula on skeletal remains in forensic casework. From systematic review, 19 research articles have been identified. A critical review of these methods is highly relevant to decide which vertebrae and/or parts of the vertebra are important for identification of unknown subjects. Based on the accuracy score, the 12th thoracic vertebra showed the highest scores for sexual dimorphism. For further analysis, an advanced meta-analysis of the vertebrae may be performed to accurately reflect the methodological quality of these evidences in sex estimation studies over the past few years.

In this review, several limitations were identified. Although classification according to age groups is important to minimize the identification pool, different age groups were not performed in these studies, which perhaps may produce different effects on the results. Further, the review may have missed some important relevant studies, especially when only three search engines were used in the review.

Conclusions

It can be concluded that vertebral bones have provided good accuracies for sex estimation, and that most vertebral dimensions are population-specific. Although individual vertebra has been studied to evaluate sexual dimorphism, the percentage accuracy was found to increase with combination of studies of the vertebrae. Results also showed that for all vertebrae, the most sexually dimorphic area of the vertebra was the vertebral body. Further studies may be needed to determine sexual dimorphism in other areas and traits of the vertebrae in an advanced method for meta-analysis.

Abbreviations

BHa/BHp: The ratio of anterior to posterior height of vertebral body; BHm/BHp: The ratio of the anterior to middle height of the vertebral body; C1: First cervical vertebra; C2: Second cervical vertebra; C4: 4th Cervical vertebra; C7: 7th Cervical vertebra; CMA: Maximum length of the axis; CT: Computed tomography; DFA: Discriminant function analysis; DMFS: Maximum distance between the superior articular facets; EPDl: Lower endplate depth; EPWu: Upper endplate width; iBDC: Coronal diameter of endplate on inferior plane; iBDcm: Maximum coronal diameter of endplate on inferior plane; iVL: Inferior length of the whole vertebrae; L1: First lumbar vertebra; LMA: Maximum width of the axis; sAD: Distance between superior articular processes; sBDC: Superior coronal diameter of endplate of the vertebral body; sBDcm: Superior maximum coronal diameter of endplate of the vertebral body; sBDs: Sagittal diameter of endplate on superior plane; sBDsm: Superior maximum sagittal diameter of vertebral body endplate; sVL: Superior vertebral length; T12: 12th Thoracic vertebra; WWF: Width of vertebral foramen

Acknowledgements

We would like to thank the Faculty of Medicine, Universiti Kebangsaan Malaysia (UKM) for providing the resources in this systematic review.

Authors' contributions

AR had produced the main idea for the article, collaborated in performing the study, performed the data analysis and collaborated in writing the article. MS and FMN participated on the design of the study, collaborated in writing the article and helped to conceptualize the ideas. All authors have read and approved the final manuscript.

Funding

None.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Faculty of Medicine, University Muhammadiyah Semarang, Semarang, Indonesia. ²Forensic Unit, Department of Pathology, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras, 56000 Kuala Lumpur, Malaysia.

Received: 11 February 2021 Accepted: 31 August 2021

Published online: 16 September 2021

References

- Amores A, Botella MC, Alemán I (2014) Sexual dimorphism in the 7th cervical and 12th thoracic vertebrae from a Mediterranean population. *J Forensic Sci* 59(2):301–305. <https://doi.org/10.1111/1556-4029.12320>
- Azofra-Monge A, Alemán Aguilera I (2020) Morphometric research and sex estimation of lumbar vertebrae in a contemporary Spanish population. *Forensic Sci Med Pathol* 16(2):216–225. <https://doi.org/10.1007/s12024-020-00231-6>
- Bethard JD, Seet BL (2013) Sex Determination from the second cervical vertebra: a test of Wescott's method on a modern american sample. *J Forensic Sci* 58(1):101–103. <https://doi.org/10.1111/j.1556-4029.2012.02183.x>
- Blau S, Briggs CA (2011) The role of forensic anthropology in disaster victim identification (DVI). *Forensic Sci Int* 205(1–3):29–35. <https://doi.org/10.1016/j.forsciint.2010.07.038>
- Bruzek J, Murail P (2006) Methodology and reliability of sex determination from the skeleton. In: Schmitt A, Cunha E, Pinheiro J (eds) *Forensic anthropology and medicine*. Humana Press, Totowa, pp 225–242. https://doi.org/10.1007/978-1-59745-099-7_9
- Cattaneo C (2007) Forensic anthropology: developments of a classical discipline in the new millennium. *Forensic Sci Int* 165(2–3):185–193. <https://doi.org/10.1016/j.forsciint.2006.05.018>
- Choong CL, Alias A, Abas R, Wu YS, Shin JY, Gan QF, Thu KM, Choy KW (2020) Application of anthropometric measurements analysis for stature in human vertebral column: a systematic review. *Forensic Imaging* 20:200360. <https://doi.org/10.1016/j.fri.2020.200360>
- Decker SJ, Foley R, Hazelton JM, Ford JM (2019) 3D analysis of computed tomography (CT)-derived lumbar spine models for the estimation of sex. *Int J Legal Med* 133(5):1497–1506. <https://doi.org/10.1007/s00414-019-02001-8>
- El Dine FMMB, El Shafei MM (2015) Sex determination using anthropometric measurements from multi-slice computed tomography of the 12th thoracic and the first lumbar vertebrae among adult Egyptians. *Egypt J Forensic Sci* 5(3):82–89. <https://doi.org/10.1016/j.ejfs.2014.07.005>
- Gama I, Navega D, Cunha E (2015) Sex estimation using the second cervical vertebra: a morphometric analysis in a documented Portuguese skeletal

- sample. *Int J Legal Med* 129(2):365–372. <https://doi.org/10.1007/s00414-014-083-0>
- Garoufi N, Bertsatos A, Chovalopoulou ME, Villa C (2020) Forensic sex estimation using the vertebrae: an evaluation on two European populations. *Int J Legal Med* 37(6):2307–2318. <https://doi.org/10.1007/s00414-020-02430-w>
- Gilsanz V, Wren TA, Ponrartana S et al (2018) Sexual dimorphism and the origins of human spinal health. *Endocr Rev* 39(2):221–239. <https://doi.org/10.1210/er.2017-00147>
- Gülek B, Durgun B, Özer HT et al (2007) CT-based morphometric data of L3–L5 vertebrae: anatomic and surgical approach. *Neurosurg Q* 17(2):92–97. <https://doi.org/10.1097/WNQ.0b013e318032e0ca>
- Hora M, Sládek V (2018) Population specificity of sex estimation from vertebrae. *Forensic Sci Int* 291:279.e1. <https://doi.org/10.1016/j.forsciint.2018.08.015>
- Hou WB, Cheng KL, Tian SY, Lu YQ, Han YY, Lai Y, Li YQ (2012) Metric method for sex determination based on the 12th thoracic vertebra in contemporary north-easterners in China. *J Forensic Leg Med* 19(3):137–143. <https://doi.org/10.1016/j.jflm.2011.12.012>
- Kaeswaren Y, Hackman L (2019) Sexual dimorphism in the cervical vertebrae and its potential for sex estimation of human skeletal remains in a white scottish population. *Forensic Sci Int Reports* 1:100023. <https://doi.org/10.1016/j.fsir.2019.100023>
- Kimmerle EH, Jantz RL, Konigsberg LW, Baraybar JP (2008) Skeletal estimation and identification in American and East European Populations*. *J Forensic Sci* 53(3):524–532. <https://doi.org/10.1111/j.1556-4029.2008.00708.x>
- Klaes AR, Ousley SD, Passalacqua NV (2020) Statistical approaches to sex estimation. In: sex estimation of the human skeleton. Academic Press, London, pp 203–217. <https://doi.org/10.1016/B978-0-12-815767-1.00013-4>
- Kuligowski J, Pérez-guaita D, Quintás G (2016) Application of discriminant analysis and cross-validation on proteomics data. In: Jung K (ed) *Statistical Analysis in Proteomics*. Humana Press, New York, pp 175–184. https://doi.org/10.1007/978-1-4939-3106-4_11
- Marlow EJ, Kozieradzka-Ogunmakin I (2016) Metric sex estimation of ancient Egyptian skeletal remains. Part II : testing of new population-specific methods. *Bioarchaeology Near East* 10:27–46. <http://www.anthropology.uw.edu.pl/10/bne-10-02.pdf>
- Marlow EJ, Pastor RF (2011) Sex determination using the second cervical vertebra—a test of the method. *J Forensic Sci* 56(1):165–169. <https://doi.org/10.1111/j.1556-4029.2010.01543.x>
- Muñoz-hernandez V, Arenaza L, Gracia-marco L et al (2018) Influence of physical activity on bone mineral content and density in overweight and obese children with low adherence to the mediterranean dietary pattern. *Nutrients* 10(8):1075. <https://doi.org/10.3390/nu10081075>
- Omar N, Mohd Ali SH, Shafie MS, Nik Ismail NA, Hadi H, Ismail R, Mohd Nor F (2021) Sex estimation from reconstructed scapula models using discriminant function analysis in the Malaysian population. *Aust J Forensic Sci* 53(2):199–210. <https://doi.org/10.1080/00450618.2019.1661516>
- Omar N, Mohd Ali SH, Shafie MS, Nik Ismail NA, Hadi H, Mohd Nor F (2019) A preliminary study of sexual dimorphism of scapula By computed tomography in the Malaysian population. *Asian J Pharm Clin Res* 12(1):391. <https://doi.org/10.22159/ajpcr.2019.v12i1.30233>
- Ostrowsky KR, Churchill SE (2015) Sex determination by discriminant function analysis of lumbar vertebrae. *J Forensic Sci* 60(1):21–28. <https://doi.org/10.1111/1556-4029.12543>
- Oura P, Karppinen J, Niinimäki J, Junno JA (2018) Sex estimation from dimensions of the fourth lumbar vertebra in Northern Finns of 20, 30, and 46 years of age. *Forensic Sci Int* 290:350.e1–350.e6. <https://doi.org/10.1016/j.forsciint.2018.07.011>
- Padovan L, Ulbricht V, Groppo FC, Neto JSP, Andrade VM, Júnior LF (2019) Sexual dimorphism through the study of atlas vertebra in the Brazilian population. *J Forensic Dent Sci* 11(3):158–162. https://doi.org/10.4103/jfojfds_85_19
- Pretorius E, Steyn M, Scholtz Y (2006) Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *Am J Phys Anthropol* 129(1):64–70. <https://doi.org/10.1002/ajpa.20251>
- Ramadan N, Abd El-Salam MH, Hanon AF et al (2017) Identification of sex and age for Egyptians using computed tomography of the first lumbar vertebra. *Egypt J Forensic Sci* 7(1):1–8. <https://doi.org/10.1186/s41935-017-0025-7>
- Robinson MS, Bidmos MA (2009) The skull and humerus in the determination of sex: Reliability of discriminant function equations. *Forensic Sci Int* 186:86.e1–86.e5. <https://doi.org/10.1016/j.forsciint.2009.01.003>
- Rozendaal AS, Scott S, Peckmann TR, Meek S (2020) Estimating sex from the seven cervical vertebrae: an analysis of two European skeletal populations. *Forensic Sci Int* 306:110072. <https://doi.org/10.1016/j.forsciint.2019.11.0072>
- Slice DE (2007) Geometric morphometrics. *Annu Rev Anthropol* 36(1):261–281. <https://doi.org/10.1146/annurev.anthro.34.081804.120613>
- Stull KE, Tise ML, Ali Z, Fowler DR (2014) Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. *Forensic Sci Int* 238:133–140. <https://doi.org/10.1016/j.forsciint.2014.03.005>
- Suwanlikhid N, Prasitwattanaseree S, La Tegola L et al (2020) Sex and stature estimation from adult lumbar vertebrae in a Thai population based on image analysis. *Int J Morphol* 38(6):1651–1656. <https://doi.org/10.4067/S0717-9502020000601651>
- Tan SH, Teo EC, Chua HC (2004) Quantitative three-dimensional anatomy of cervical, thoracic and lumbar vertebrae of Chinese Singaporeans. *Eur Spine J* 13(2):137–146. <https://doi.org/10.1007/s00586-003-0586-z>
- Taylor JR, Twomey LT (1984) Sexual dimorphism in human vertebral body shape. *J Anat* 138:281–286. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1164068/>
- Teodoru-Raghina D, Perlea P, Marinescu M (2017) Forensic anthropology from skeletal remains to ct scans: a review on sexual dimorphism of human skull. *Forensic Anthropol* 25(3):287–292. <https://doi.org/10.4323/rjlm.2017.287>
- Torimitsu S, Makino Y, Saitoh H, Sakuma A, Ishii N, Yajima D, Inokuchi G, Motomura A, Chiba F, Yamaguchi R, Hashimoto M, Hoshioka Y, Iwase H (2016) Sexual determination based on multidetector computed tomographic measurements of the second cervical vertebra in a contemporary Japanese population. *Forensic Sci Int* 266:588.e1–588.e6. <https://doi.org/10.1016/j.forsciint.2016.04.010>
- Wescott DJ (2000) Sex variation in the second cervical vertebra. *J Forensic Sci* 45(2):462–466. <https://doi.org/10.1520/jfs14707j>
- Yu SB, Lee UY, Kwak DS, Ahn YW, Jin CZ, Zhao J, Sui HJ, Han SH (2008) Determination of sex for the 12th thoracic vertebra by morphometry of three-dimensional reconstructed vertebral models. *J Forensic Sci* 53(3):620–625. <https://doi.org/10.1111/j.1556-4029.2008.00701.x>
- Zheng WX, Cheng FB, Cheng KL, Tian Y, Lai Y, Zhang WS, Zheng YJ, Li YQ (2012) Sex assessment using measurements of the first lumbar vertebra. *Forensic Sci Int* 219:285.e1–285.e5. <https://doi.org/10.1016/j.forsciint.2011.11.022>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)