SECTION 4.
Stature and Metrical Skeletal Characteristics

This chapter will deal with the information which can be gained from the metrical analysis of skeletal remains. Measurement of the lengths of the long bones is most useful for the estimation of living stature of an individual. Measurements of the skull are used to calculate cranial indices which can be used in the comparison of skeletal populations. A few indices, such as the Meric and Cnemic, are calculated from long bone measurements.

All measurements taken in this study follow the methods described in Brothwell (1981).

4.1. Stature

4.1.1. Methods and Problems

The only living statistic which can be estimated with any accuracy from the skeleton is stature. According to Brothwell (1981:100), factors controlling this physical characteristic are c.90% genetic and only 10% environmental. This obviously has to be taken into account in the interpretation of mean stature estimates. Various regression formulae for calculating height have been compiled in the past, based on a number of different populations. For example, small groups of French skeletons were studied by Rollet (1888), Manouvrier (1892-3) and Pearson (1899). In 1898-1902 Hrdlicka (1939) measured the long bones of American whites and negroes, with known cadaver heights, and calculated long bone/stature ratios. Dupertius and Hadden (1951) also worked on American whites and negroes with known cadaver heights (Todd Collection). They tested the validity of Pearson’s formulae, which they found to give a consistently shorter stature than their own. Telkkä (1950) studied a small group of Finnish skeletons, mostly male, and calculated regression equations.

The most useful and extensive study to be carried out so far is that of Trotter and Gleser (1952, 1958, Trotter 1970). They used the skeletons of World War II dead, the Terry Collection, and later the Korean War dead, all of whom had a known living stature. Different formulae were calculated for the three major race types (white, negro and mongoloid), since it was found that the relationship of stature to length of long bones differed between them.

The method utilised is as follows. The maximum length of each complete long bone in the skeleton is measured (except for the tibia, for which the total length is used). The formula for the bone(s) with the least standard deviation is then chosen according to which bones are present. It is best to use the femur and tibia if these two bones are available. The long bones from the legs are undoubtedly of more value in this respect than those of the arms, since the former contribute more to stature than the latter.

Trotter and Gleser proposed a correction factor for individuals over the age of 30 years. The correction is to subtract 0.06cm for every year over the age of 30, and therefore an accurate age is required. This is not used with archaeological skeletal populations due to the difficulty of accurately determining age. The estimated living stature of an individual quoted in an archaeological skeletal report is taken to be the approximate greatest height attained by that individual during his or her lifetime.

Male and female skeletons require different formulae, due to the difference in bodily proportions between the two sexes. For this reason, if an individual skeleton cannot be sexed, it cannot be allocated an estimated height.

Although the Trotter and Gleser formulae were calculated from an American population, they have been used on various ancient European populations. This is because it is felt that they are more accurate than some other formulae which have been calculated from European populations. For example, Breitinger (1937) worked out formulae based on 2400 living males from Germany. Trotter (1970:71) states that in this case ‘The clear advantage of stature being measured on the living subject was unfortunately offset by the limited accuracy with which bones can be measured from bony prominences palpated through the skin’. Other earlier formulae (Pearson, Telkkä, Dupertius and Hadden, etc.) were in general calculated from skeletal groups numbering 200 or fewer individuals.

Huber (1968) points out that Trotter and Gleser measured bones in conditions varying from moist to dry, and bone lengths decrease slightly with drying. Assuming that limb bone proportions are the same in archaeological populations, stature will probably err on the short side, if at all, because of this. He also states that even if limb bone proportions are shown to be similar in modern and ancient populations, we know nothing about the possible relative changes in the trunk size.

L.H. Wells (1960) estimated the statures of some neolithic skeletons from West Kennet long barrow and Dark Age skeletons from S.E. Scotland using the formulae of Trotter and Gleser, Pearson, and Dupertius and Hadden. He found that both the 1952 and 1958 formulae of Trotter and Gleser gave widely discrepant estimates from different long bones of the same skeleton (a difference of as much as 27mm), whereas those from Pearson, and Dupertius and Hadden, were much closer (only 5mm and 14mm difference respectively). He says ‘Although all the discrepancies are well within the standard errors of estimate of the Trotter-Gleser formulae, it seems justifiable to conclude that Anglo-Saxons as a group had appreciably longer arms than modern White Americans, but were identical in mean limb proportions with the nineteenth century French series upon which the Pearson formulae were based' (1960:139). He suggests that this could be due to the more vigorous use of the upper limbs in the lifestyles of these populations when compared with modern populations.

Huber and Jowett (1973) have used the measurements taken by Trotter and Gleser and compared them with a population of early medieval Alamannic Germans. They found that bodily proportions of American whites and the medieval population were not significantly different, and concluded from this that it was reasonable to use the Trotter and Gleser formulae for such a group.

In his 1968 paper, Huber states that 'mean lengths of the long bones of the males from Weingarten [i.e. Alamanns] are no greater than those from any other early Medieval series from Northern Europe...and they are essentially the same as those of the Anglo-Saxons' (1968:80). He suggests that, as far as stature is concerned, they can be regarded as a homogeneous population. If this is the case, then the Trotter and Gleser formulae should be just as appropriate for estimating stature in the current study groups as it appears to be for the Alamanns, especially, as he points out later (1968:83), since 'the American white population was predominantly descended from the older Northern European and British populations, and...there is no reason to assume that the formulae for stature prediction do not apply to them'.

It should be noted that, at present, it is only possible to estimate the stature of adult skeletons. There has been no study on a known population of children, and since sexing is so difficult there may also be a problem here. Smith (1939) used diaphyseal lengths of foetal long bones to calculate foetal length, but the validity of this is questionable, and its use in archaeological populations is limited by the lack of foetal skeletons normally discovered. Since the main use of this method is to estimate the age of a skeleton, and given that the variability of height within a certain age group is likely to be fairly large, then it is doubtful whether stature by age can be estimated for children who are aged from the lengths of their long bones.

Steele and McKern (1969) and Steele (1970) suggest a method of estimating stature from fragmentary long bones (humerus, femur and tibia), based on 117 prehistoric American Indian skeletons, but since this only adds greatly to the error already involved in calculating stature it is not generally attempted. Its main use is in forensic anthropology, when the height is a useful criterion in identification.

Musgrave and Harneja (1978) have calculated regression formulae for estimating stature from metacarpal lengths, based on radiographs of the hands of 166 mainly white adults. They found a high correlation between stature and metacarpal length. However, if no long bones are present in an archaeological skeleton, it is doubtful whether there would be enough of the skeleton left to sex it confidently, or even if the metacarpals would have survived in a condition good enough to be measured.

4.1.2. Methods used in this Study

The Trotter and Gleser formulae are the most widely used today. In this study the 1970 American white formulae are used throughout (Wells' studies on the Jarrow and Monkwearmouth populations utilised the 1952 and 1958 formulae, but the statures have been recalculated for these two groups to make them more comparable with the others in this study). The 1970 formulae are actually the 1952 formulae, with the omission of those formulae involving a mixture of arm and leg bones, since these were felt by the authors to be less accurate. It is felt that the 1952 formulae are preferable to the 1958 formulae for male individuals for use with an ancient population, because they are based on an older group (from the Second World War and earlier, rather than the Korean War) and are therefore less affected by the demonstrable increase in height which has occurred during this century.

In this study only the complete long limb bones of adult male and female skeletons have been utilised, although broken or slightly eroded bones have been used if the majority of the bone was present. Since any estimation of stature can have an error of between 2 and 4cm when a bone is complete, it was felt that a slight inaccuracy in the measured length of the long bone would not greatly affect the estimated height.

Tables 4.1 and 4.2 show the numbers and percentages of the methods which were used for estimating stature at Jarrow, Monkwearmouth and The Hirsel.

Method MALES	HIR		MK		JA Sax.		JA Med.	
	N	%	N	%	N	%	N	%
Fe+Ti	33	53.2	17	40.5	5	26.3	14	43.8
Femur	16	25.8	9	21.4	8	42.1	8	25.0
Fibula	2	3.2	1	2.4	0	-	0	-
Tibia	3	4.8	7	16.7	1	5.3	5	15.6
Humerus	6	9.7	5	11.9	4	21.1	2	6.3
Radius	2	3.2	2	4.8	1	5.3	1	3.1
Ulna	0	-	1	2.4	0	-	2	6.3

Table 4.1.

Method FEMALES	HIR		MK		JA Sax.		JA Med.	
	N	%	N	%	N	%	N	%
Fe+Ti	37	64.9	10	55.6	3	25.0	16	42.1
Fibula	2	3.5	0	-	1	8.3	2	5.3
Tibia	2	3.5	4	22.2	1	8.3	7	18.4
Femur	11	19.3	3	16.7	4	33.3	7	18.4
Radius	2	3.5	1	5.6	1	8.3	4	10.5
Ulna	1	1.8	0	-	1	8.3	0	-
Humerus	2	3.5	0	-	1	8.3	2	5.3

Table 4.2.

The bones recorded under 'method' are in order of lowest to highest standard error for each sex. In almost every case the formula with the lowest error (Fe + Ti) has been used the most, so that the estimates of stature from these three sites should be fairly reliable.

4.1.3. Stature Estimates in the Study Populations

The average estimated statures in centimetres (from all bones) of the population groups in this study are as follows:

Site	Period	Sex	n	Mean	Range
NEM	Anglian	M	15	173.5	164.2 - 182.8
		F	14	163.7	148.3 - 176.1
BG	Saxon	M	35	171.8	162.5 - 179.6
		F	27	157.8	140.5 - 167.8
MK	Saxon	M	42	171.9	151.9 - 188.4
		F	19	159.5	145.9 - 169.2
JA	Saxon	M	19	171.0	160.9 - 184.4
		F	12	159.1	148.8 - 166.6
JA	Medieval	M	32	171.0	158.0 - 186.2
		F	38	159.7	152.2 - 168.0
HIR	9th-15th c.	M	62	167.7	154.4 - 177.2
		F	57	158.8	147.0 - 169.7
BF	Medieval	M	15	173.5	163.6 - 181.9
		F	8	162.5	154.6 - 176.6
GP	c.1100-1540	M	17	170.6	160.7 - 181.6
		F	13	162.7	153.0 - 170.6

Table 4.3.

The distribution in heights between the sexes is shown in figures 4.1 - 4.7. These bar charts show that there is a fairly similar spread of heights at all the sites, with the possible exception of Blackfriars. This last site had two male modes, possibly due to the small size of the sample rather than to any particular trend. Figure 4.8 shows the mean and range for each site graphically and by broad time period. It shows that all the means and ranges are within normal limits.

Table 4.4 shows the modes (in cm) of the various sites which are presented graphically in Figures 4.1-4.7, for ease of comparison. This shows that the sites are all fairly similar in general trend, with the exception of the Jarrow females and the Hirsel males, both of whom have a lower mode than the others.

Figure 4.1. Stature distributions at The Hirsel.

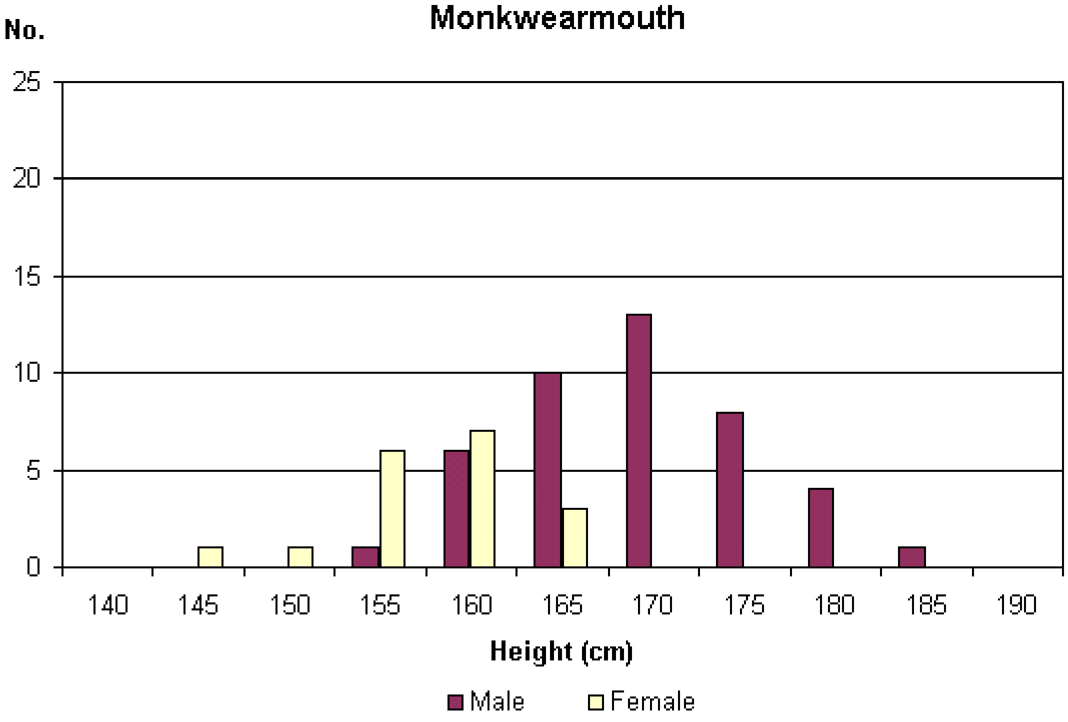


Figure 4.2. Stature distributions at Monkwearmouth.

Figure 4.3. Stature distributions at Jarrow.

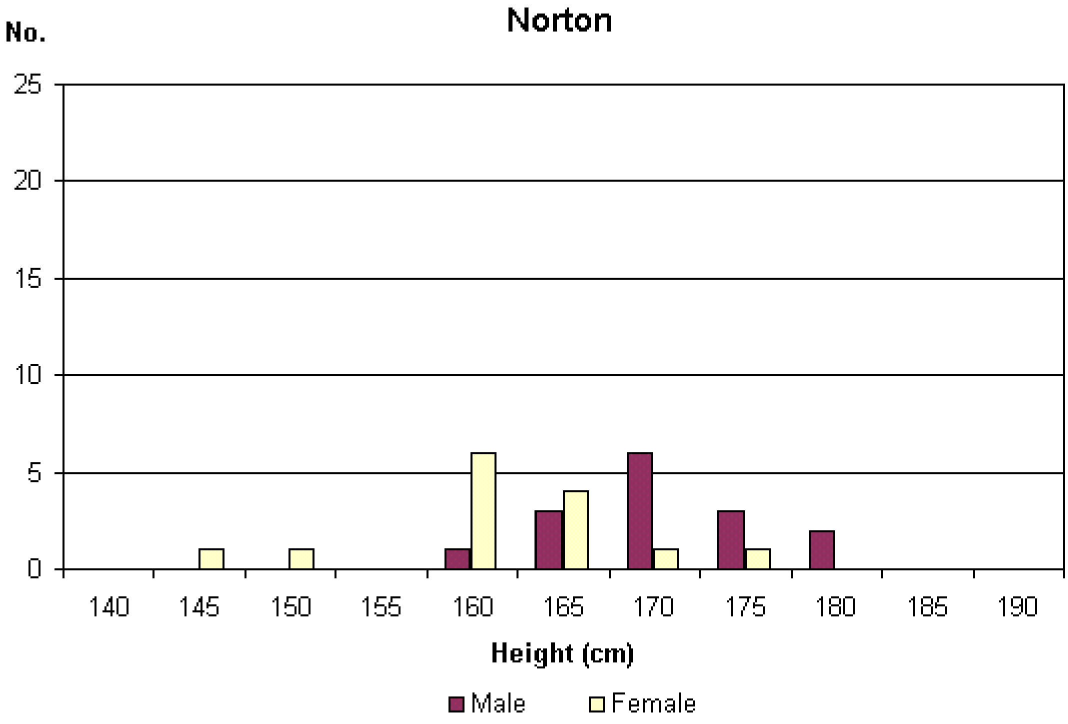


Figure 4.4. Stature distributions at Norton.

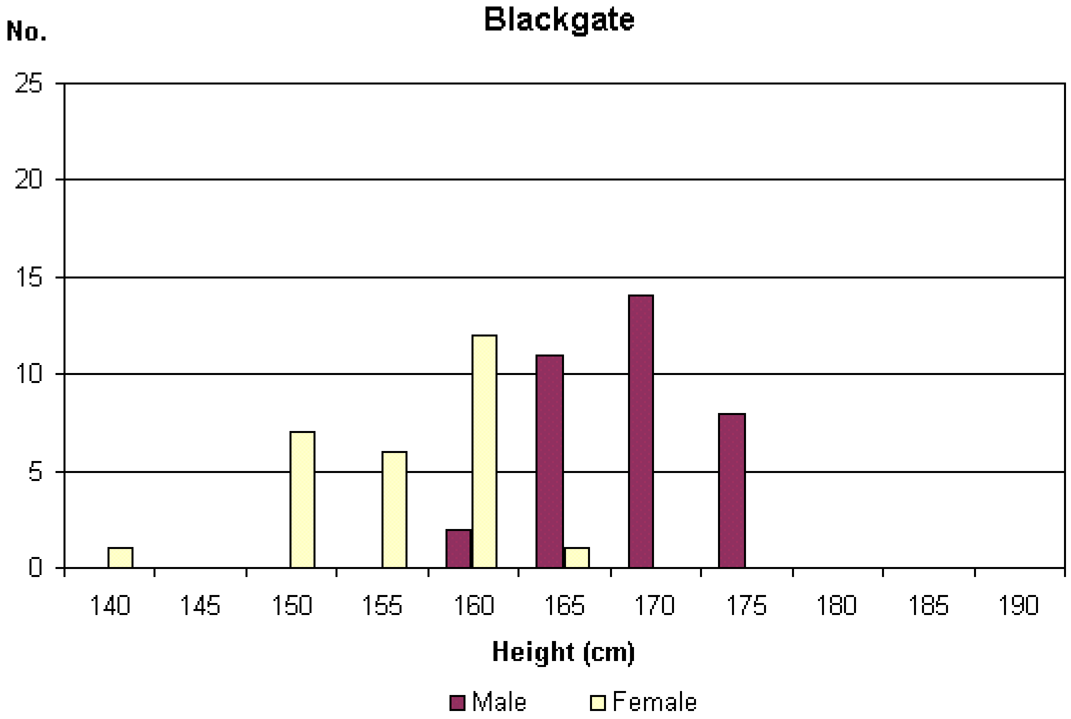


Figure 4.5. Stature distributions at Blackgate.

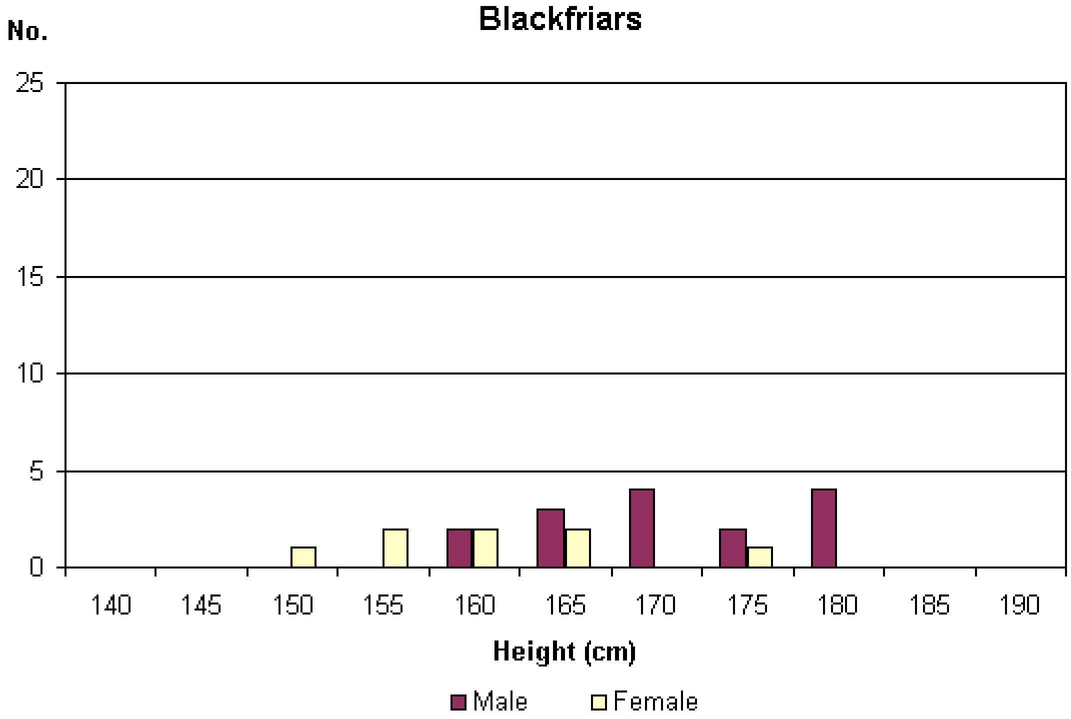


Figure 4.6. Stature distributions at Blackfriars.

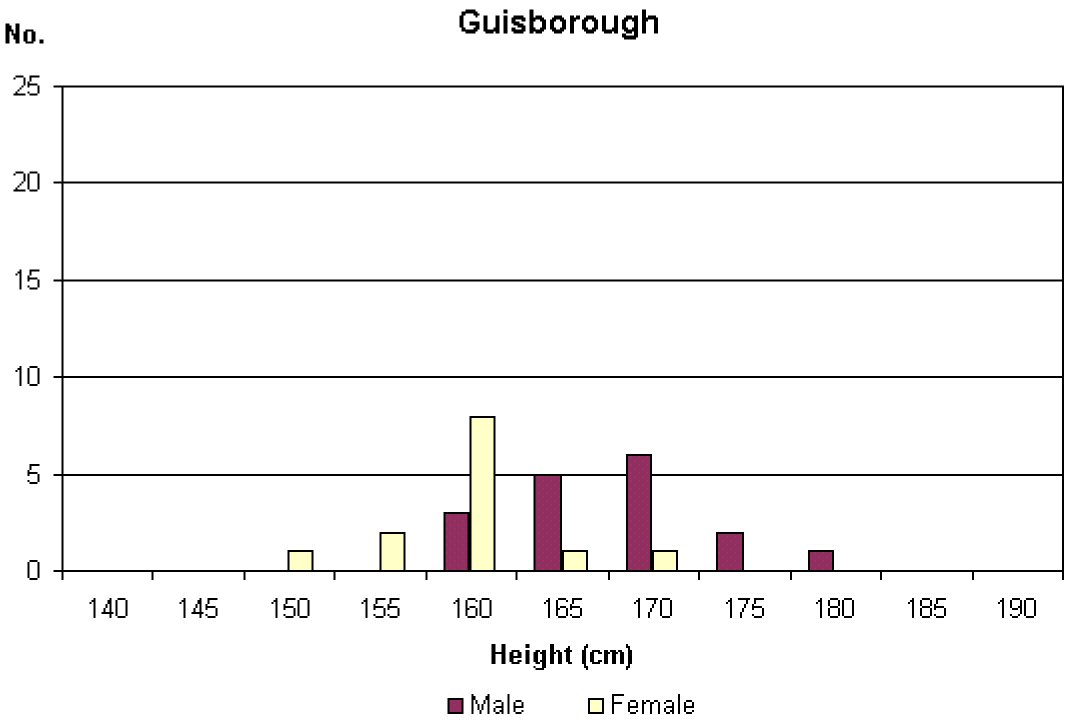


Figure 4.7. Stature distribution at Guisborough.

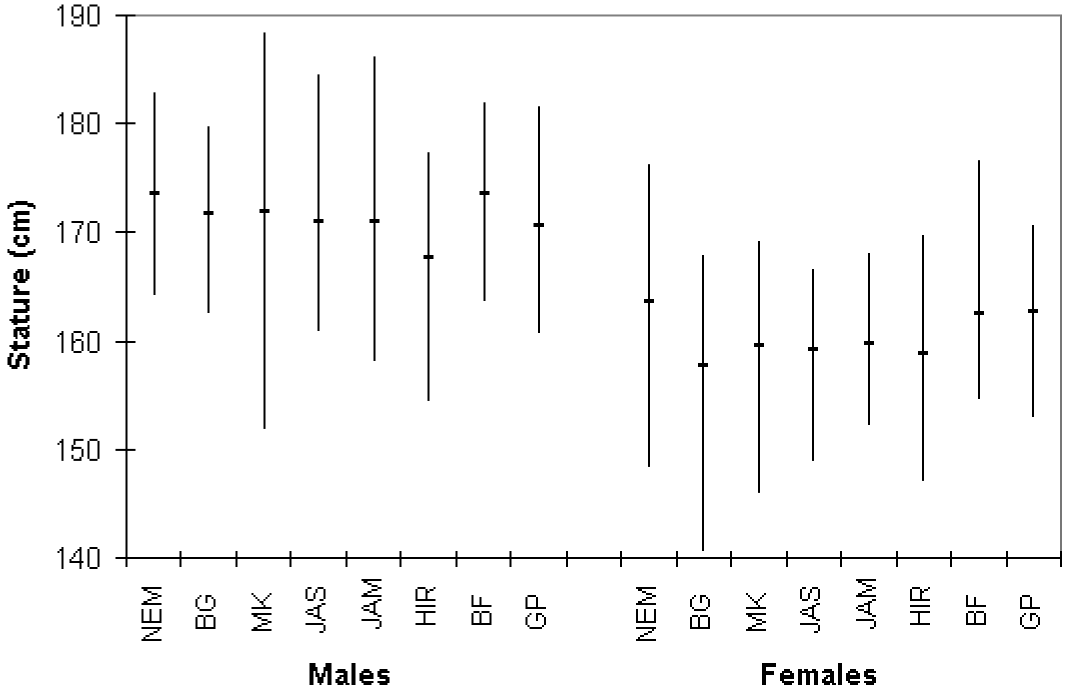


Figure 4.8. Means and ranges of stature by broad time period and site.

Site	Male	Female
HIR	165	160
MK	170	160
JA	170	155
NEM	170	160
BG	170	160
BF	170/180	160?
GP	170	160

Table 4.4.

It has been found, in all the populations in this study, that stature estimated for individuals with only arm bones is often noticeably greater than that of individuals for whom leg bone measurements can be used, especially in the females. This is in support of L.H. Wells' theory that the Anglo-Saxons and other early peoples had longer arms in proportion to their legs than do the modern Americans.

Tables 4.5 and 4.6 show the numbers, means and ranges of the statures (in cm) estimated from the leg bones only, for Jarrow, Monkwearmouth and The Hirsel. Table 4.5 includes those estimates based on the formula with the lowest error in both sexes (i.e. Femur + Tibia), and Table 4.6 includes estimates based on all the leg bone formulae. The results for all except the Jarrow males are very similar.

Site	Sex	N	Mean	Range
MK	M	17	171.8	160.5 - 183.3
	F	10	159.8	153.9 - 162.8
JA	M	19	169.9	160.8 - 183.1
	F	19	159.1	152.2 - 166.6
HIR	M	33	168.3	159.4 - 177.2
	F	37	158.9	149.3 - 166.1

Table 4.5.

Site	Sex	N	Mean	Range
MK	M	34	170.9	159.1 - 184.0
	F	17	159.9	145.9 - 169.2
JA	M	40	174.0	158.0 - 183.1
	F	41	159.3	148.8 - 168.0
HIR	M	54	167.8	155.2 - 177.2
	F	52	158.5	147.0 - 169.7

Table 4.6.

Mean statures were calculated from all the long bone types available at The Hirsel, in order to find out how great the variance is between the various estimates. The results are shown in Tables 4.7 (males) and 4.8 (females). Both sexes have a difference of 5.2cm (2") between the highest and lowest mean estimate. However, this is well within the standard errors of 2.99cm and 3.55 for the best regression formulae (Fe+Ti), suggesting that it is reasonable to use all stature estimates when calculating the mean, rather than having to limit the calculations to those skeletons which had intact femora and tibiae. In some skeletons the estimate was actually very close. Sk. 198 (male), for example, had three estimates of 173.9 (from Fe+Ti, Fem, and Tib) and one of 170.9 (Rad). This is not to say that the stature estimate for this skeleton is any more accurate than the others. It only suggests that it is closer to the American white population.

Formula	Mean	N	Range	s.d.
Fe + Ti	168.3	33	159.4 - 177.2	4.66
Femur	167.4	49	155.2 - 177.2	4.68
Fibula	166.6	19	162.1 - 170.8	3.03
Tibia	169.8	38	160.0 - 177.4	4.26
Humerus	170.5	37	154.4 - 181.3	5.68
Radius	169.8	38	154.5 - 179.2	5.50
Ulna	171.8	30	158.8 - 179.5	4.75

Table 4.7.

Formula	Mean	N	Range	s.d.
Fe + Ti	158.9	38	149.3 - 166.1	3.89
Fibula	157.5	16	150.1 - 162.8	3.56
Tibia	160.2	41	152.3 - 166.9	3.92
Femur	157.5	49	147.0 - 169.7	4.42
Radius	161.0	32	152.3 - 171.5	4.88
Ulna	162.7	23	155.3 - 171.3	4.29
Humerus	160.2	38	148.4 - 175.2	5.22

Table 4.8.

L.H. Wells (1960) found a variance of 27mm between stature estimates on the Humerus, Radius, Femur and Tibia of a male Anglo-Saxon Series, using Trotter and Gleser's formulae. Using his method of estimating mean stature from the mean long bone length, The Hirsel male population produced a variance of 35mm. Although this seems to give a better result than the mean calculated from estimates of stature derived from each individual skeleton, it is probably more accurate to produce a mean by the latter method.

As stated previously, Huber (1968) considers that Alamanns and Anglo-Saxons are very close in stature. He quotes a mean stature of 173.2cm for both (172.8 if Trotter's 1970 formulae are used). L.H. Wells quotes a similar figure of 172.3 (or 171.8 with the 1970 formulae). Both are higher than the majority of populations in this study, both Anglo-Saxon and Medieval. In Table 4.9, the mean lengths of long bones for Alamanns and Hirsel males are compared.

	Alamanns			The Hirsel		
Bone	N	Mean	s.d.	N	Mean	s.d.
Hum.	53	332	21.0	58	325	16.9
Rad.	30	249	14.9	53	241	13.7
Fem.	71	465	23.7	83	444	19.3
Tib.	48	377	22.5	37	361	17.9

Table 4.9.

This shows that the long bones of the Alamannic males were consistently longer than those of the Hirsel men. However, if the Trotter and Gleser formulae can be proved to be of use for Alamannic groups because the proportions of the limbs are similar to the American whites, then it is proportionality not actual size which is important. If the Humero-Radial length is divided by the Femoro-Tibial length and converted to a percentage, the Alamannic ratio is 69.0 and that of The Hirsel is 70.3. The sites in this study were combined to form two groups, Saxon (JA Sax, MK, BG and NEM) and Medieval (JA Med, BF, and GP). A ratio was calculated for the right limbs of each of these two groups to see if there was any great difference. The results, together with those of The Hirsel, the Alamanns, Pearson, Dupertius and Hadden, and Trotter and Gleser (combined series) are recorded in Table 4.10.

Group	Male	Female
Saxon	71.5	70.0
Medieval	69.9	67.2
The Hirsel	70.3	69.9
Alamanns	69.0	-
Pearson	70.5	68.6
Dupertius & Hadden	69.8	68.3
Trotter & Gleser	69.2	69.0

Table 4.10.

The results suggest a fairly similar proportionality within all the groups. The small differences account for the variance seen when estimating stature from one of the formulae with a greater standard error. As L.H. Wells suggested (1960), the upper limbs of Saxon men and women may be slightly longer in proportion to their legs than those of the Medieval period, although the difference is slight.

Wells also suggests that Teutonic migrations were producing a shift towards taller stature in Western Europe. Table 4.11 records the mean statures (in cm) of a few Anglo-Saxon series for comparison with those studied here.

Site	Author	Male	Female
North Elmham	C. Wells (1980)	172.1	157.5
Red Castle	C. Wells (1967)	169.7	158.1
Burgh Castle	Anderson (1989)	175.9	163.2
Nazeingbury	Putnam (1978)	175.3	168.2
Kingsworthy	Wells/Hawkes (1983)	173.6	161.3

Table 4.11.

These sites, all in the South-East of England, have a fairly high average stature. Most of the Saxon sites in this study are fairly close to the lowest two means, but The Hirsel is well below, and none of the populations reach anywhere near the mean heights attained by the Burgh Castle population. Even if Burgh Castle is exceptional, and the other sites are the norm for an Anglo-Saxon population (which seems likely), then the North-Eastern populations are still on the short side. Perhaps Northerners were less well-nourished than their southern counterparts in this period and were therefore not reaching their maximum potential height. The other alternative seems to be that these populations were more localised, and had a greater proportion of native peoples amongst them. However, it is dangerous to make assumptions about ethnic groups based on stature and long bone measurements alone. Cranial observations may provide more evidence (see Section 4.3), but it is unlikely that a distinction between environmental and genetic factors in these groups can be made based on present knowledge.

4.2. Indices Calculated from Long Bone Measurements

Although many indices have been invented by various workers in the past, and especially in the early days of physical anthropology, only a few are used regularly today. Ashley-Montagu (1951) lists four, namely the Radio-Humeral index (R/H x 100), the Pilastric index (taken at the midshaft of the femur, AP/ML x 100), the Meric and the Cnemic indices. Bass (1971) mentions a few more: the claviculo-humeral (useful for the indication of the relative development of the chest); the humero-radial (the same as Ashley-Montagu's radio-humeral); the robusticity of the clavicle, humerus and femur (to show the relative size and thickness of the shaft, and often used for sex determination); and of course, the platymeric and platycnemic indices. These last two are the most well-known and well-used indices in any osteological study, despite the fact that they are still not fully understood or explained. There is a growing feeling amongst a number of workers that such indices are merely measured because they are there.

The Meric index measures the antero-posterior flattening of the femoral shaft, and is taken just below the lesser trochanter (AP/ML x 100). The Cnemic is a similar measure of the medio-lateral flattening of the tibia, and is taken at the nutrient foramen (ML/AP x 100). They are usually classified into four categories each, as follows:

Meric Index		Cnemic Index	
Hyperplatymeric	x - 74.9	Hyperplatycnemic	x - 54.9
Platymeric	75.0 - 84.9	Platycnemic	55.0 - 62.9
Eumeric	85.0 - 99.9	Mesocnemic	63.0 - 69.9
Stenomeric	100.0 - x	Eurycnemic	70.0 - x

The larger the index, the broader the shaft of the bone in both cases.

Wells, in his report on the Jarrow skeletons (forthcoming), states that the fact that the two conditions of platymeria and platycnemia are more common in early and present-day primitive peoples than in advanced civilisations has caused them to be ascribed to the habit of squatting. He feels that this theory is difficult to sustain. As he says, 'in many populations femoral and tibial flattening vary independently of each other, and in known squatters both may be absent, or in non-squatters either may be found'. He also mentions a number of other theories concerning the conditions, such as the idea that platymeria is a response to unusual stresses on the femoral shaft, or that it is caused by various pathological processes, or that it is a physiological economization in the use of minerals for bone formation. Platycnemia has been claimed to be dependant on the degree of retroversion of the tibial head. Wells does not think that any of these theories are correct, and suggests a multifactorial origin for both conditions.

Lovejoy *et al* (1976) analysed the biomechanics of bone strength as applied to platycnemia. They state that 'higher cnemic indexes are more common among populations associated with neolithic and urban economies...[and] the triangular shape of the tibia is a more recent phenomenon' (1976:490). Like Wells, they discard the theory that a particular posture (i.e. squatting) could determine the form of the shaft, since 'the shape of an adult long bone results from a highly complex process of deposition and resorption, not simply by differential rates of growth'. Having studied the torsional strength of the tibia as a whole, they conclude that platycnemia is caused by a specific pattern of mechanical loading which is distinct from that producing eurycnemia. They suggest that a eurycnemic tibia is more adapted to all strain-inducing modes than the platycnemic, which is better equipped for more antero-posterior bending strain. However, what this means in terms of the archaeological and anthropological interpretation of the Cnemic index is unclear.

Andermann (1976) has studied the Cnemic index and found it to be greatly affected by the random variation of the position of the nutrient foramen. He studied 104 tibiae from the Dickson Mound collection of prehistoric American Indians, and concluded that a better measure of antero-posterior flattening could be taken at one-third the length of the tibia (proximal end). He found this index to be more consistent and comparable than either the cnemic index or the midshaft index, the latter being affected by biomechanical forces originating from the distal end of the shaft, and therefore of less use than the new index when considering the traits which influenced the original Cnemic index. However, as he himself admits, specimens which are incomplete or broken, for which the length cannot be measured, could not be used in the new index, since the measurement has to be taken at exactly one-third distance from the proximal end. It is also impossible to make comparisons with past work if the new index is used.

Lavelle (1974a) studied the femora of a number of British populations ranging from the bronze age to the present. He used measurements, indices and multivariate analysis. Both multivariate and simple statistics showed varying patterns of contrast between populations. After standardization of linear measurements against length, a progressive increase in size was seen from the bronze age to the present, and form was also seen to change by metrical analysis. Before standardization, however, there was little to choose between univariate and multivariate statistics as a method of biological distancing (see Section 4.3.1). Unfortunately he makes no conclusions about changes or otherwise in the meric index specifically.

4.2.1. Work on the Study Populations

Three long bone indices were calculated for the study populations, the Meric and Cnemic indices, and the index of femoral robusticity (Bass, 1971). This latter, as measured at The Hirsel, has been discussed in Section 3.2 on Sex.

An attempt was made to see if any correlation existed between the meric and cnemic indices in the adult population from The Hirsel. Scattergrams of one plotted against the other showed no specific trend, and the correlation coefficient calculated for the male L. meric against L. cnemic was very low (0.2375). There would appear to be very little relationship between the two, other than that determined by the sizes of the bones.

4.2.1.1. The Meric Index in the Study Populations

The means and ranges of the meric index (combined for left and right sides) at each of the study groups are recorded in Table 4.12.

	Male			Female		
Site	N	Mean	Range	N	Mean	Range
HIR	91	76.9	63.2-93.8	99	75.4	62.2-104.3
MK	47	75.9	64.1-87.5	28	72.5	62.9- 87.1
JA Sax	25	77.9	54.7-88.3	14	72.1	60.2- 83.0
JA Med	56	77.1	59.5-99.7	60	80.0	61.4- 93.4
NEM	37	72.1	60.5-83.3	31	72.3	60.0- 93.3
BG	53	76.8	67.5-91.4	51	73.6	62.9- 83.3
BF	31	82.3	71.1-93.3	22	87.1	74.2-104.3
GP	33	82.2	66.7-94.3	23	78.1	67.6- 90.0

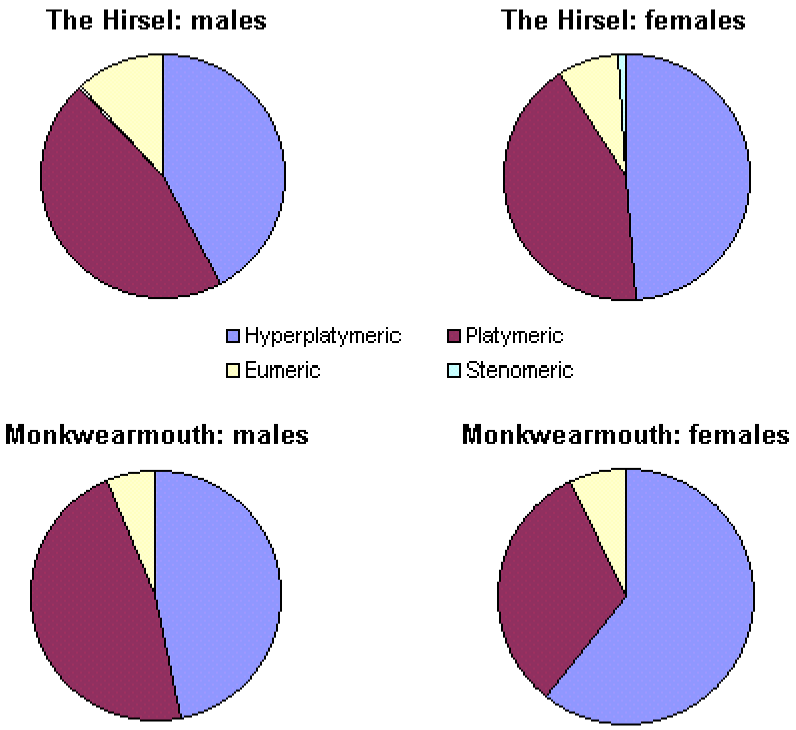
Table 4.12.

This suggests that the earlier populations had proportionately thinner femora than the later ones, and that at all but Medieval Jarrow and Blackfriars, the females had a smaller index than the males. Brothwell (1981) states that various authors have claimed that platymeria is more common in females, and more frequent in earlier peoples, and the figures from this study would seem to bear this out. He also suggests that the left femur is often more platymeric than the right. In these populations this is true of the majority of groups (JA Med, NEM, BF females, GP, BG and HIR females), but in all cases there was very little difference between the means of the two sides.

Almost all of the mean meric indices recorded in the table fall into the platymeric range. The females of Monkwearmouth and Saxon Jarrow and both sexes from Norton are in the hyperplatymeric group, and the Blackfriars females are in the eumeric category.

Figures 4.9 to 4.12 present the distributions over the categories at all the sites, in the form of pie charts. These show a marked similarity between both sexes from The Hirsel and Medieval Jarrow, and the Blackgate and Monkwearmouth males. The females from Norton and Guisborough are also fairly close to these. The females from Monkwearmouth, Saxon Jarrow and Blackgate, and the Norton males, seem to form another distinct group.

The males from the two medieval sites of Guisborough and Blackfriars have a similar distribution, but the Blackfriars females show a distribution different from any of the other groups, possibly due to the small size of the

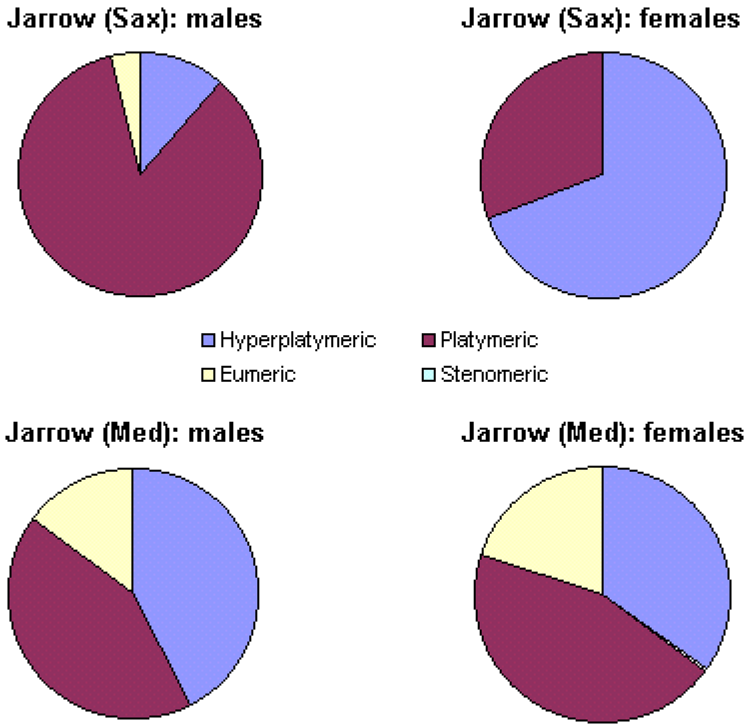

The Hirsel: males

The Hirsel: females

Monkwearmouth: males

Monkwearmouth: females

Figure 4.9. Meric index distribution: The Hirsel and Monkwearmouth.


Jarrow (Sax): males

Jarrow (Sax): females

Jarrow (Med): males

Jarrow (Med): females

Figure 4.10. Meric index distribution: Jarrow

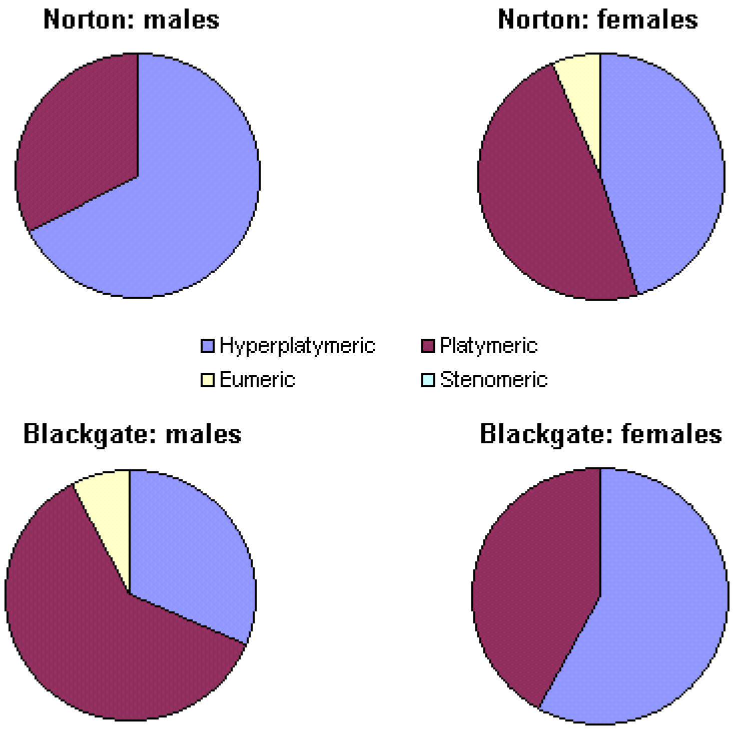


Figure 4.11. Meric index distribution: Norton and Blackgate.

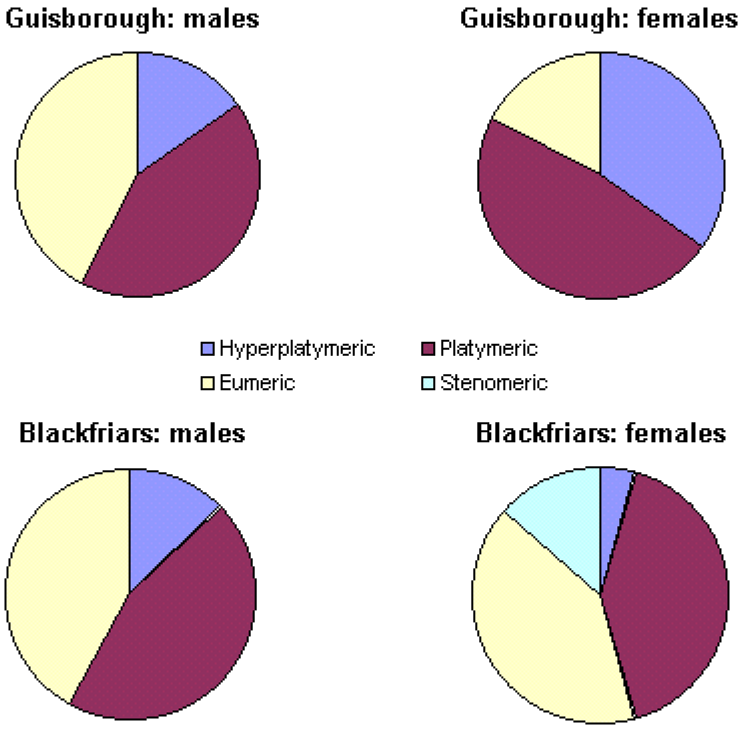


Figure 4.12. Meric index distribution: Blackfriars and Guisborough.

sample. The Saxon Jarrow males also have a strange distribution, with a large proportion of platymeric femora. If the Meric index does differ through time, which it certainly seems to at these sites, then the observed grouping of the Saxon females can be easily explained. The grouping of the Saxon males from Monkwearmouth and Blackgate with two medieval populations is less simple to understand, although it may be that the males were changing towards the medieval type at a greater rate than the females, or that they had a larger input into the genetic change in later periods than females. Since the reasons behind the flattening of the shaft of the femur have not been adequately explained it is difficult to reach any conclusions concerning these patterns.

4.2.1.2. The Cnemic Index in the Study Populations

The means and ranges of the Cnemic indices calculated for the study populations (for combined left and right sides) are recorded in Table 4.13.

	Males			Females		
Site	N	Mean	Range	N	Mean	Range
HIR	92	67.2	55.0-88.0	93	70.7	52.9-92.3 3
MK	46	66.3	52.5-78.9	25	70.4	60.7-91.9 3
JA Sax	22	67.4	54.7-87.5	17	70.7	56.6-81.6 3
JA Med	43	71.8	59.6-82.6	49	72.2	57.6-81.3 3
NEM	39	70.6	56.1-81.8	31	73.1	64.5-91.7 3
BG	46	66.4	57.5-82.4	28	69.4	55.3-80.6 3
BF	26	71.9	64.9-82.9	16	75.1	67.6-83.3 3
GP	32	68.9	56.1-85.3	20	69.1	62.5-80.0 3

Table 4.13.

In this case, the earlier sites have a slightly lower mean than the later in every case, except Norton. All the female means are greater than those of the males. All the group means fall into the Mesocnemic (HIR male, MK male, JA Sax male, BG and GP) and Eurycnemic (HIR female, MK female, JA Sax female, JA Med, NEM and BF) categories.

Figures 4.13 to 4.16 provide a graphic representation of the distribution of the indices into categories at each of the sites. There is a similarity between the distributions at The Hirsel and Saxon Jarrow, and Monkwearmouth and the males from Blackgate, Guisborough and Norton are also quite close. The Norton females show a similar pattern to the females from Medieval Jarrow, and the Guisborough and Monkwearmouth females are fairly close to each other. The Blackgate females and both sexes from Blackfriars do not correlate well with any of the other groups. In the case of the Cnemic index there does not appear to be much correlation with time period in the distribution patterns seen at these sites, but how this should be interpreted is unknown.

4.3. Cranial Measurements and Morphology

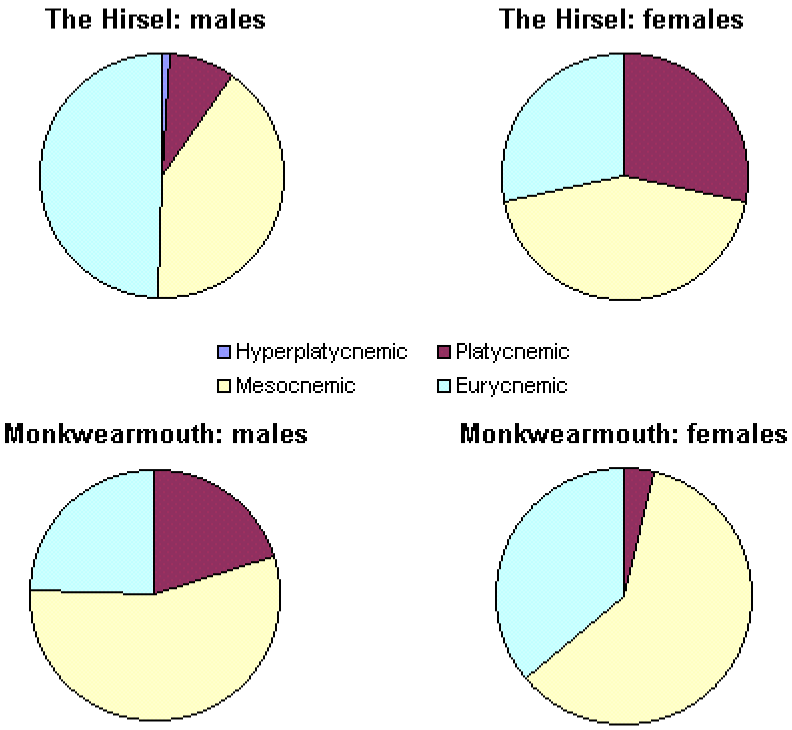
4.3.1. Techniques of Cranial Analysis in Current Use

For the purposes of most (British) osteological reports, the cranial measurements recommended by Brothwell (1981) are generally used. Indices are calculated from the main measurements, such as cranial length, breadth and height (for cephalic, height/length and height/ breadth). Krogman (1978), Ashley-Montagu (1951) and others give lists of the major indices and their category divisions. Other measurements are usually recorded in the hope that they will be useful for future research.

At the other end of the scale in craniometric research, particularly in America, and occasionally in Europe (e.g. Brothwell and Krzanowski, 1974; Tattersall, 1968a), complicated statistical methods are employed to compare biological distances between populations.

Hursh (1976) produced a survey of the techniques of measuring and analysing cranial form. As well as conventional methods of measurement with sliding and spreading callipers, he considers various analytical tools such as stereocontouring and even holography. He sees these ‘hi-tech’ procedures as the way forward in the field of analysis of cranial form, although he admits that they are obviously expensive, and that, in the case of stereocontouring, ‘the most serious question is what to do with the contour lines once you have them’! (1976:475).

As well as considering measurement techniques, Hursh summarises statistical methods in current use. Under the heading of ‘Univariate Measures’, he lists three problems associated with the use of ‘simple’ statistics. ‘First, as many will freely admit of themselves, statistics are not very well understood by a significant number of people in the field....Second, they are sometimes not complex enough to test the proposed model....Third, there may be a significant discrepancy between the implications of the statistical model and the assumptions of the evolutionarily directed culture of the contemporary biological scientist’ (1976:481). If univariate statistics are subject to misuse

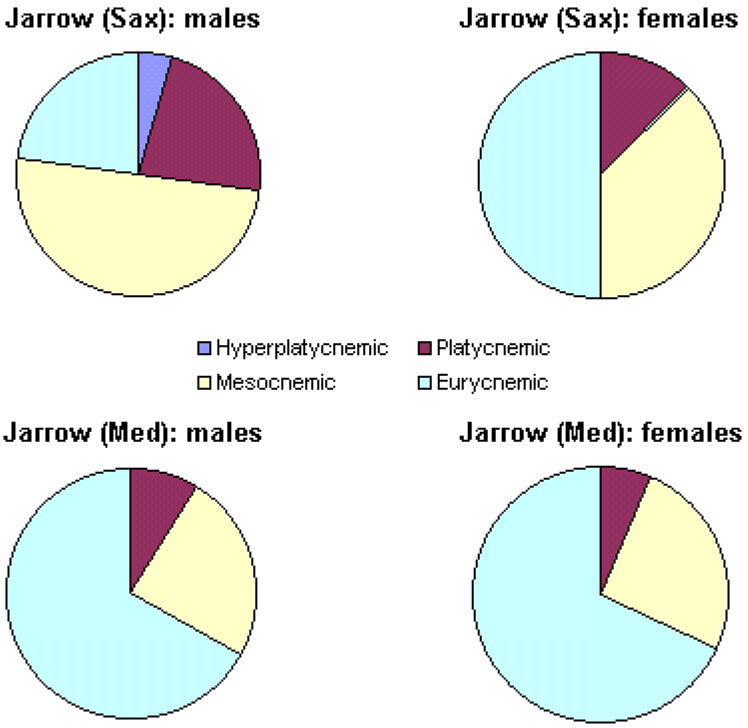

The Hirsel: males

The Hirsel: females

Monkwearmouth: males

Monkwearmouth: females

Figure 4.13. Cnemic index distribution: The Hirsel and Monkwearmouth.


Jarrow (Sax): males

Jarrow (Sax): females

Jarrow (Med): males

Jarrow (Med): females

Figure 4.14. Cnemic index distribution: Jarrow.

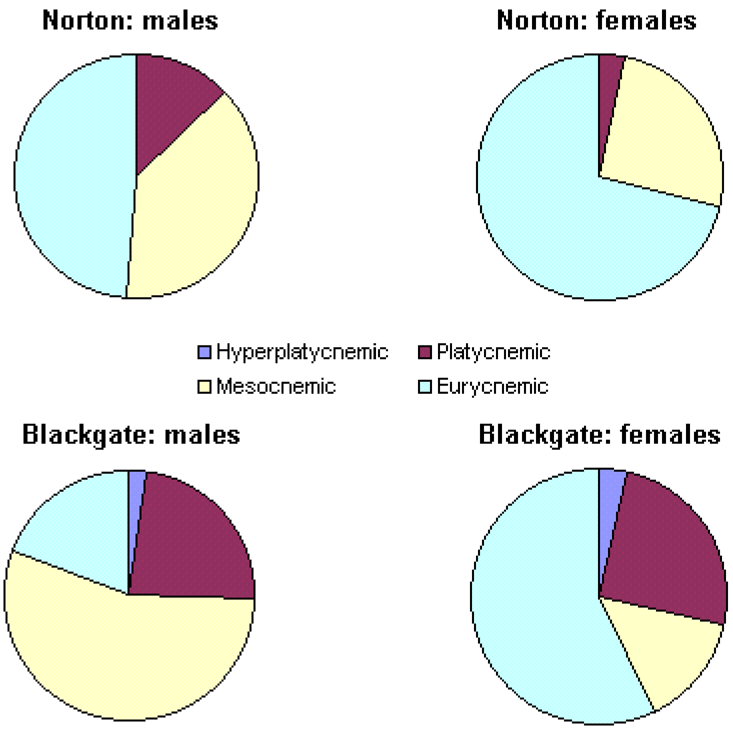


Figure 4.15. Cnemic index distribution: Norton and Blackgate.

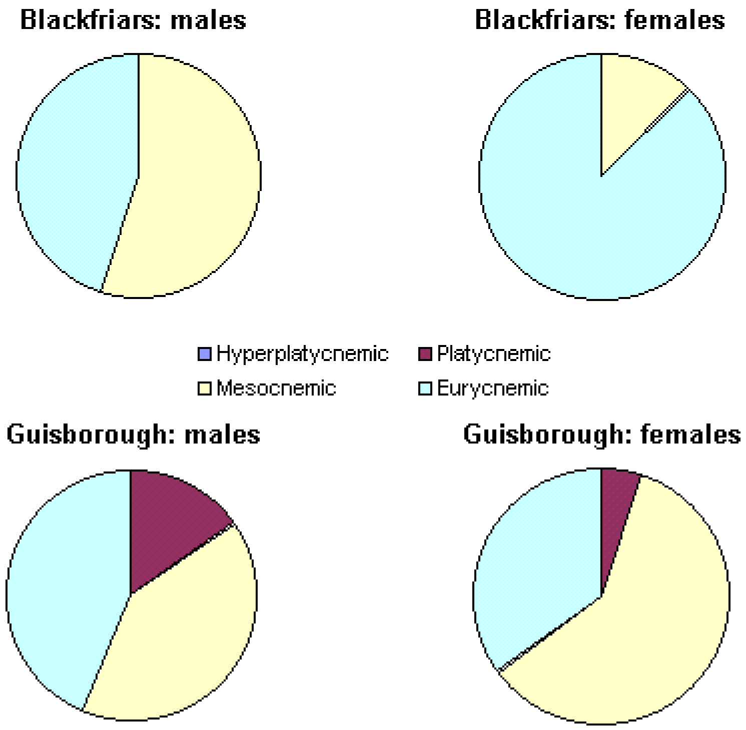


Figure 4.16. Cnemic index distribution: Blackfriars and Guisborough.

and error due to a lack of understanding, then it follows that the more complicated procedures of multivariate analysis will be even more incomprehensible to most osteologists.

Hardy and Van Gerven (1976) tested the effect of size variation on indices calculated from cranial measurements. They concluded from their results that ‘body size contributes substantially to morphological differences quantified from standard craniometric techniques’ (1976:82). Because of this, they recommend the use of principal components analysis followed by analysis of covariance to avoid the statistical problems of use of indices.

As early as 1923, Morant stated that ‘the cephalic index alone is quite incapable of discriminating between fundamental types or of distinguishing relationships between races which are known to be allied. Furthermore, no single character which has yet been suggested can fulfil either of these purposes and it is extremely unlikely that one will ever be found’ (1923:194). He used Pearson’s ‘Coefficient of Racial Likeness’ in the analysis of several population groups (e.g. Tibetans in the study of 1923). However, he also says that ‘it seems at present to be highly probable that differences in size are of relatively little importance; resemblance between the shapes of heads is the real criterion of relationship and this we are able to measure with angles and indices’ (1923:212).

A more recent study by Brown (1973) uses multivariate techniques to look at covariation in Australian Aboriginal skulls. She found it to be a useful method of craniometric research, since the collective analysis of a set of variables is more objective than analysis by conventional statistical techniques.

As mentioned earlier, Brothwell and Krzanowski (1974) have looked at a number of British skeletal groups using multivariate methods. At least 2000 skulls from 53 samples were used, varying from Neolithic to Medieval in date. The statistical tests tended to cluster the groups of similar time periods, and distance them from those of others, as would probably be expected. Brothwell says that some of these distinctions are probably biologically meaningful, and that there is some evidence for regional micro-evolution. Such an analysis may be useful when attempting to decide whether a group of skeletons are likely to belong to a certain period.

Jantz (1973) studied Arikara (American Indian) crania by multivariate methods. He also feels that variables should be considered together rather than individually. He suggests that many metrical variables are inherited to a large extent, even if ‘genetic and environmental aspects of morphological variation are still inadequately understood’ (1973:15). In his analysis he found that cranial length and breadth, the two variables used in the cephalic index, contributed very little to his canonical variates, and that variables from the face contributed the most. Thus, ‘the face tends to display more significant interpopulation variation than the cranial vault’ (1973:20). The reason for the predominant use of the cephalic index by most workers is that the face is unfortunately more susceptible to decay than the cranial vault, making it impossible to carry out any in-depth studies into facial indices in the average archaeological population.

Because of this, many workers in Europe have continued to use the cephalic index, due to its ease of calculation and the fact that it usually allows for a larger sample of skulls to be considered. Wiercinski (1974) studied brachycephalisation in various populations, mostly in Europe, and concluded that the process of increase in the cephalic index (brachycephalisation) was genetically rather than environmentally determined. Necrasov (1974) did a similar study on Rumanian populations, looking at the process of brachycephalisation through time and using it to suggest genetic affinities between skeletal groups. Alekseeva (1974) used some simple indices to differentiate between Slavs and Germans. His indices and measurements appear to show a reasonable difference between population groups.

Giles and Elliot (1962) have produced a set of discriminant functions for the identification of race from cranial measurements. This is of most use in forensic identification, since it is based on the differences between Whites, Negroes and American Indians. It may be possible to use a similar method to distinguish between closer populations in archaeological contexts, as Jantz (1973) and McKern and Munro (1959) attempted on American Indian groups. However, Hursh states that ‘discriminant function analysis will find differences even when they are not there. This does not actually mean that it creates differences, but that it is so good at detecting differences that it will be able to discriminate with high levels of accuracy on differences which are not attributable to causal origins, but rather to happenstance’ (1976:484). If this is the case, then it may not be a good idea to use the method on population groups which are very similar in time and space.

Utermohle *et al* (1983) have drawn attention to three other factors which might affect cranial measurements in both statistical analysis and simple comparisons of populations. They showed that there was a difference in measurements taken by different observers on the same set of skulls, that there was a difference between measurements taken at various time periods by the same observer on the same group of skulls, and that measurements were affected by varying levels of humidity. Although the differences in all these factors were at most about 3mm, they suggested that this would produce a large error when the measurements were used in multivariate statistics. Discriminant functions were calculated which could distinguish between measurements taken by the three observers to a reasonable degree. In their conclusion they state that ‘the potential inappropriateness of conclusions involving data collected by different observers is not a comforting prospect for a scientific discipline’ (1983:92). However, it is well known that in many branches of science errors are expected to occur most of the time, and these are generally taken into account in the final analysis.

4.3.2. Methods applied to the Study Populations

In the study of these population groups, craniometric techniques have been confined to the simple measurements and indices described by Brothwell (1981). There are three main reasons for this.

Firstly, Ubelaker (1978) suggests that a sample of 100 or more adults from each group being compared should be used in the estimation of biological distance by multivariate techniques. This would rule out all of the skeletal populations considered in the present study, since none of them has a large enough group of complete skulls.

Second, the more complex statistical techniques involve large and time consuming calculations, which, even if carried out by a computer, still need to be analysed by the observer. They are thus beyond the range of the current work, since they would need to have been done almost to the exclusion of the analysis of any other data. In other words, such a study is almost large enough for a thesis in itself.

Thirdly, it is not yet clear which methods would be most appropriate for small series, and the research involved to determine this is outside the scope of this study.

Although the craniometric study carried out on the study populations is of the simplest type, it was thought valid to include the data, since it is still comparable with other recent studies of British skeletal populations. Ubelaker states that ‘the potential of skeletal analysis for resolving archaeological problems involving biological hypotheses cannot be realized until the genetics of bone development is better documented’ (1978:88). Since this is undoubtedly the case, it seems unnecessary to rule out the possibility that cranial vault and face indices are able to provide useful information in this field.

The most recurrent theme in all of this work on statistical analysis of cranial measurements is that they can show a difference between populations. However, unless we are able to gain a better understanding about the biological background of these people, and learn more about the heritability of metrical traits, the results are very difficult to interpret. It is noticeable that, even after all the analysis has been carried out, most workers are only able to say that one population is closer to/more distant from another in their survey. It is equally possible to show this with even simple statistics. The problem which now has to be faced is that of obtaining possible biological or environmental causes for such distinctions.

4.3.3. Results of the Craniometric Analysis

The means and ranges of the cephalic index for all the populations are recorded in Table 4.14. Other indices were calculated on the cranial vault and face, but the sample sizes involved are so small that it is felt that they may give a misleading or biased picture. As can be seen from the table, the numbers involved in the calculation of the cephalic index at most of the sites were very small.

Site	Sex	N	Mean	Range
HIR	M	29	79.0	73.9 - 88.2
	F	32	77.9	71.8 - 86.0
MK	M	6	69.8	65.3 - 72.8
	F	8	72.7	66.6 - 79.9
JA Sax	M	5	75.3	70.4 - 79.8
	F	3	74.3	70.6 - 77.0
JA Med	M	7	78.7	72.2 - 82.4
	F	5	76.4	74.3 - 77.9
NEM	M	5	72.0	67.7 - 79.9
	F	8	74.0	68.8 - 76.1
BG	M	5	73.1	68.8 - 78.0
	F	3	75.0	72.0 - 76.7
BF	M	9	77.7	68.5 - 88.4
	F	4	82.5	80.7 - 83.3
GP	M	15	79.7	75.1 - 84.5
	F	7	76.1	72.6 - 79.4

Table 4.14.

It would seem to be fairly pointless to attempt to sort these groups into the categories of the cephalic index, but from the means there does seem to be a trend towards broad, rounded (brachycephalic) crania from the earlier to the later sites. This is shown graphically in Figure 4.17.

Figures 4.18-4.20 show the spread of the three main cranial indices at The Hirsel. Unfortunately, due to the small numbers of measurable crania at the other sites, it is not possible to make any conclusions about this data in comparison with that of the other groups in this study, other than to say that there are more brachycranial individuals in the later sites and more dolichocranial (long-headed) individuals in the earlier ones. At The Hirsel, there was very little difference between the sexes in the cephalic and height/breadth indices. The most noticeable difference was in the height/length index, where the greatest proportion of males fall into the mid-range category, whilst the majority of females are in the lowest group.

One other simple index was calculated for the males of these populations, to compare them with the European groups used by Alekseeva (1974) in his study of Slavs and Germans in the Middle Ages. He used an index based on the three major cranial dimensions to differentiate Germans and Western, Southern and Eastern Slavs. This is calculated as follows:

$$\frac{\text{Cranial Height}}{(\text{Length} + \text{Breadth})/2} \times 100$$

Unfortunately, his other three indices involve measurements which are only taken rarely, when preservation allows, and it was not possible to use them in this study. The results of the analysis are given in Table 4.15 below.

Group	Mean
Monkwearmouth	78.4
The Hirsel	79.1
Jarrow (Medieval)	79.6
Blackgate	80.1
South Germans	80.9
Middle Germans	81.4
Guisborough	81.5
Burgh Castle	81.9
West Scandinavia	81.9
Jarrow (Saxon)	82.0
Blackfriars	83.6

Table 4.15.

The results seem to indicate that the populations of Blackfriars and Saxon Jarrow were at the greatest distance from Monkwearmouth and Medieval Jarrow. This is very unlikely, since they are similar groups of a similar time period and belonging to a very small area. The reason for this discrepancy is probably the small sample sizes from Blackfriars and Saxon Jarrow, rather than any major morphological difference. The most reliable results are probably those from The Hirsel, Guisborough and Burgh Castle, since all are based on quite large samples. The difference of The Hirsel from the Germanic populations and the similarity of the latter two with Germanic and Scandinavian groups is quite striking. This index is probably quite a useful method of distinguishing between population groups, but should probably only be used to make final conclusions when larger sample sizes than these are available for study.

A similar study was carried out by Brothwell on the Bronze Age people of Yorkshire (1960b). As well as using the multivariate technique of Penrose distances, he also plotted various populations using the cephalic index against basi-bregmatic height. This produced a pattern in which the Bronze Age and Neolithic groups were all fairly close together. In Figure 4.21 the same technique is applied to the populations in this study, together with some of those listed in Table 4.15 from Alekseeva's study.

From this analysis it can be seen that the males from Saxon Jarrow (JAS) are the same as the South Germans (SG), that the Middle Germans (MG), Blackgate, Norton, West Scandinavians (WS) and Burgh Castle (BC) form a distinct group, Medieval Jarrow (JAM), Guisborough and Blackfriars form a looser group, and The Hirsel and Monkwearmouth seem to be very different from all the other groups. The females show a different pattern, with Jarrow and The Hirsel appearing fairly close, Blackfriars being at a distance, and the rest forming a fairly loose group. In both the males and the females, a horizontal dividing line can be drawn between the Saxon and Medieval groups, although in the females this division is less distinct.

Further analysis of the figures obtained in the metrical analysis of these sites will have to await a study by someone with a greater understanding of statistical techniques than the present author. However, considering the small number of cranial measurements available, it is unlikely that any complex statistical test would be valid on most, if not all, of these populations.

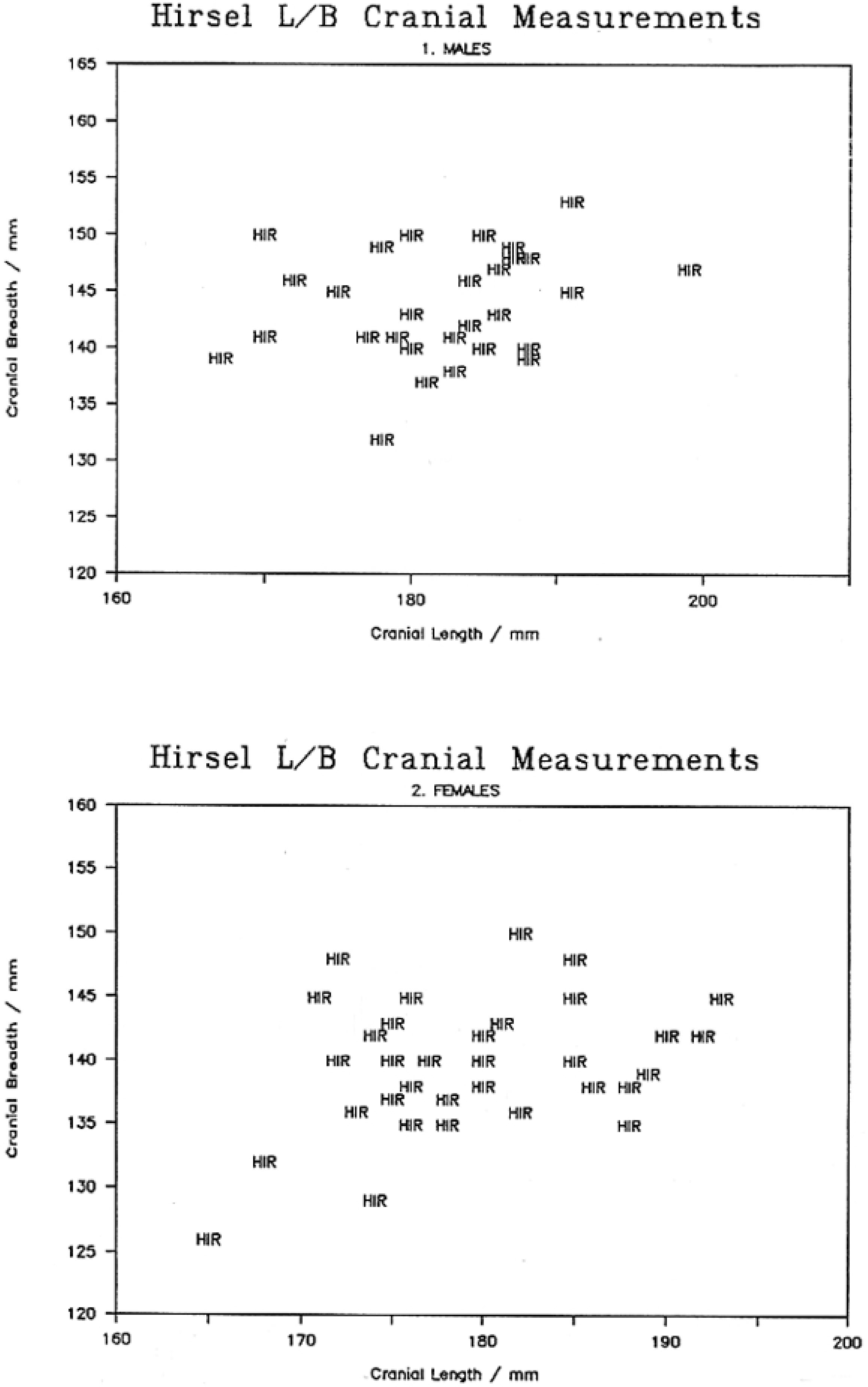
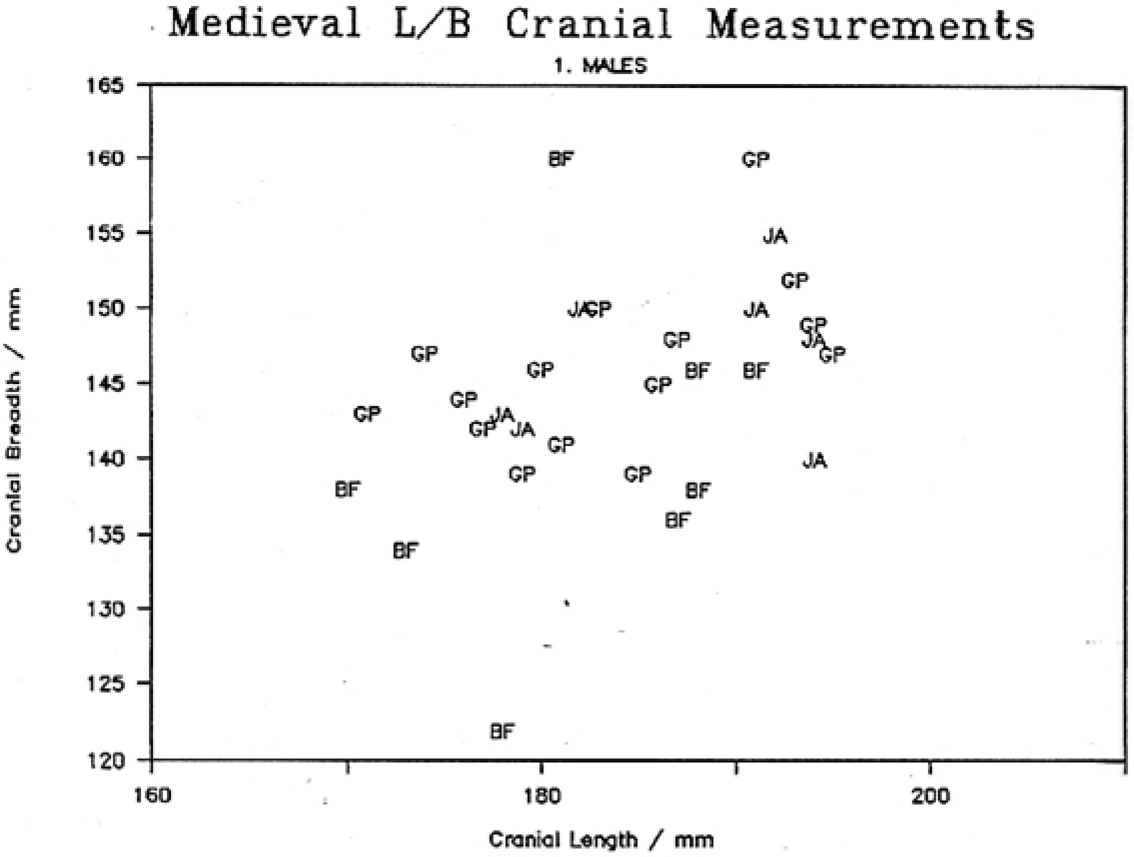
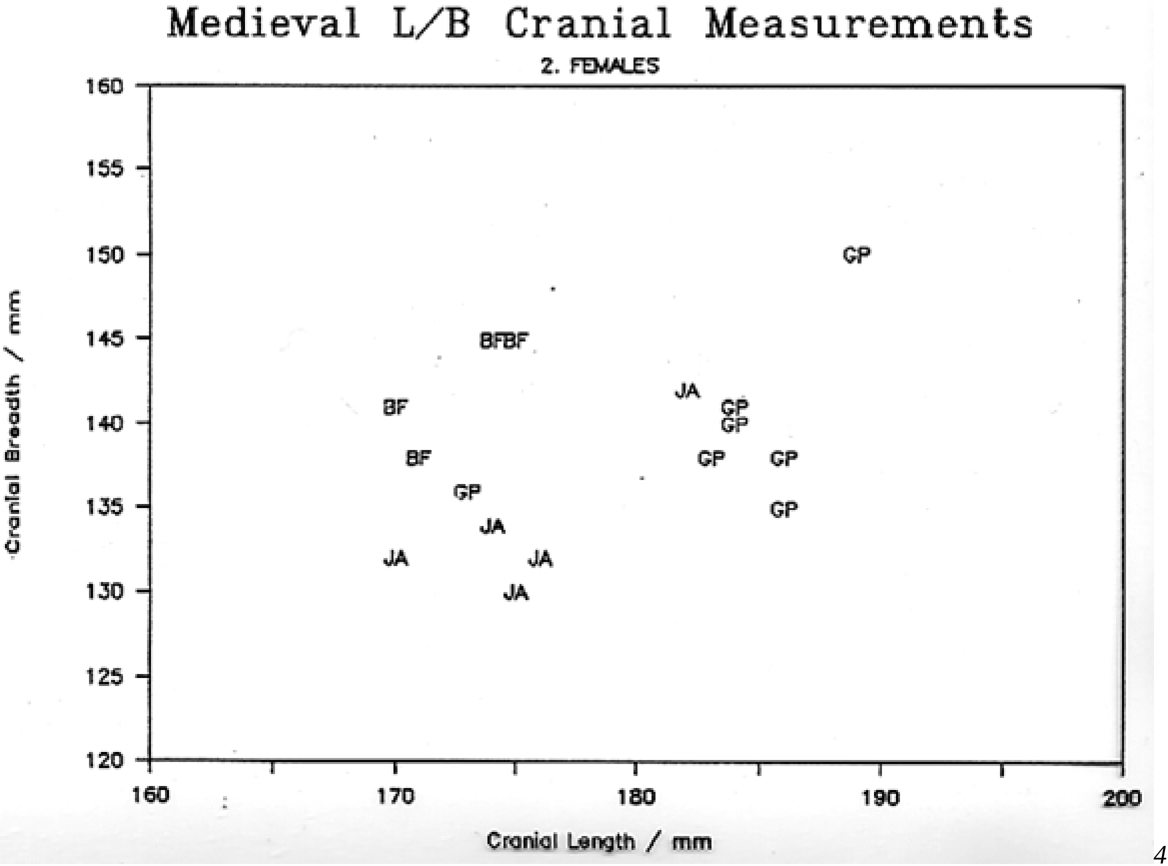
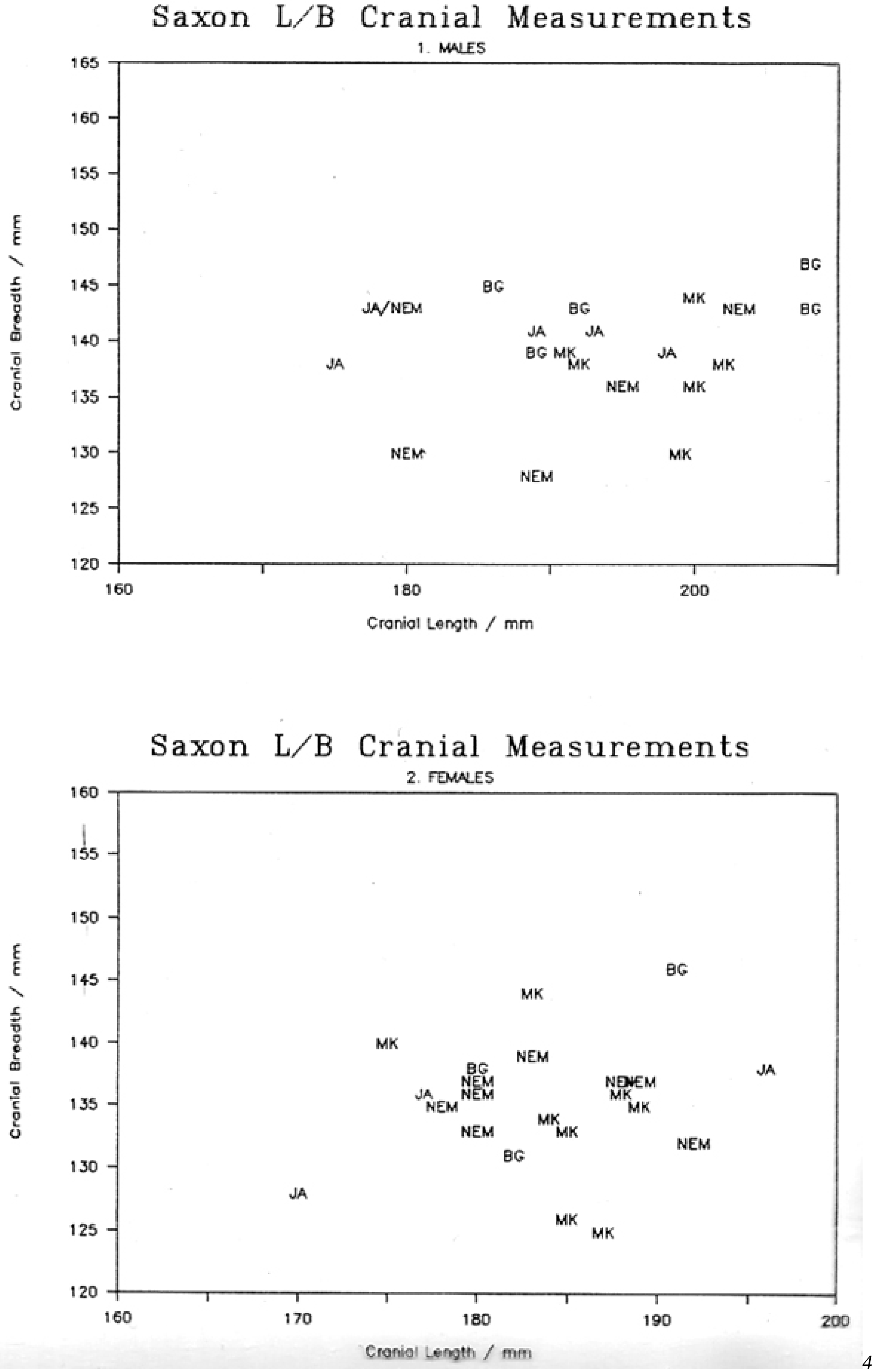


Figure 4.17a. Scattergraphs of L/B cranial measurements (The Hirsel).





4.17b.
Scattergraphs of L/B cranial measurements (Medieval sites).



4.17c.

Scattergraphs of L/B cranial measurements (Saxon sites).

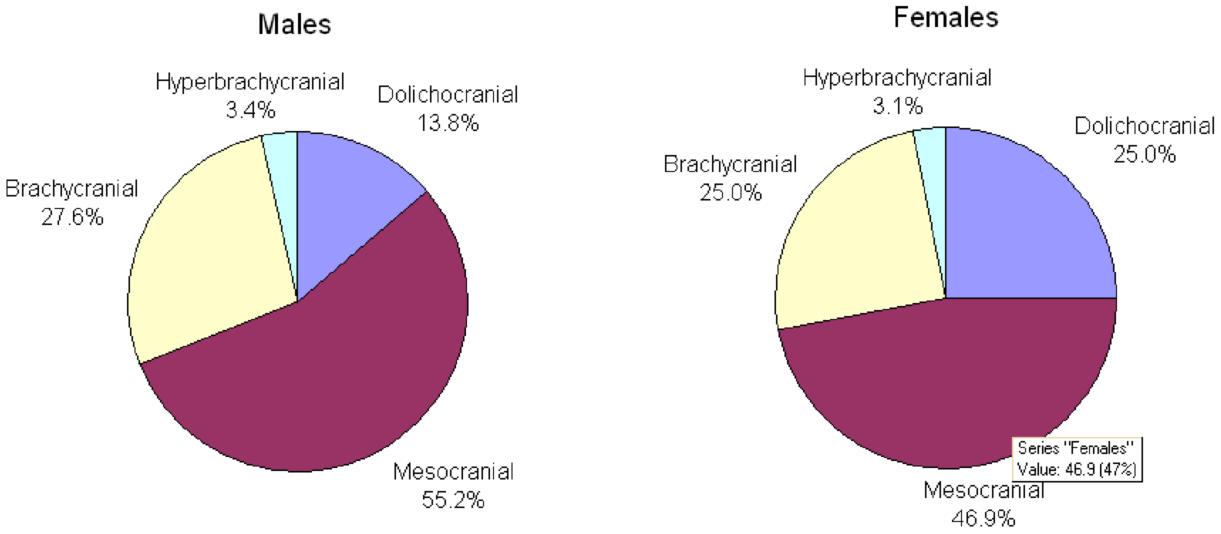


Figure 4.18. Cephalic index distribution at The Hirsel.

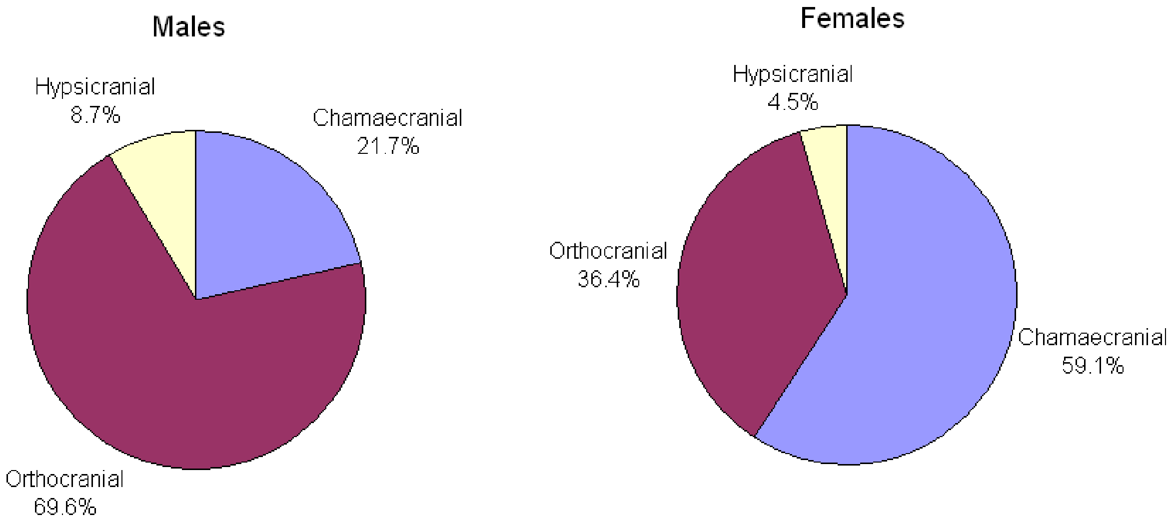


Figure 4.19. Height/length index at The Hirsel.

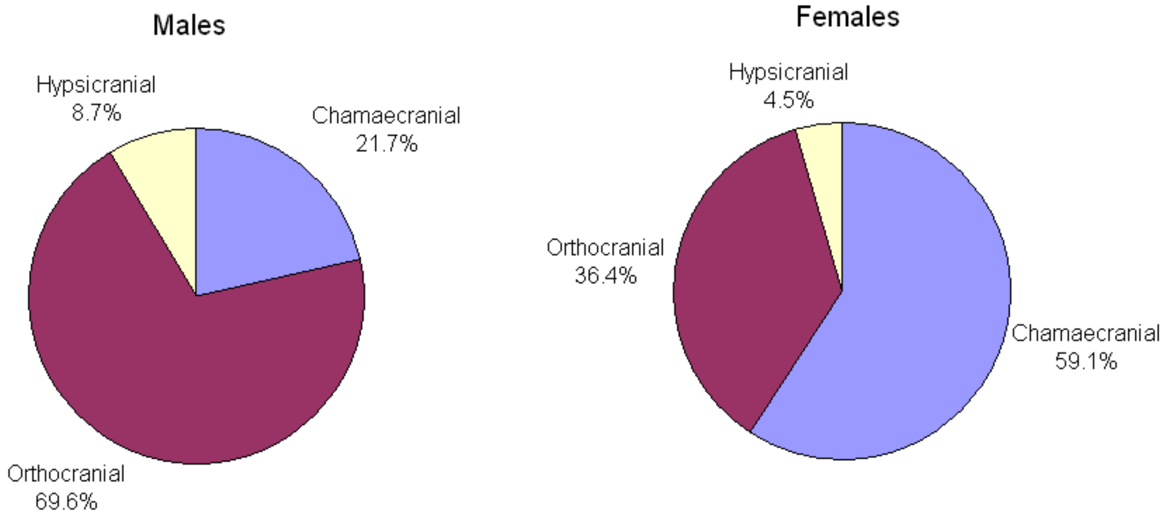


Figure 4.20. Height/breadth index at The Hirsel.

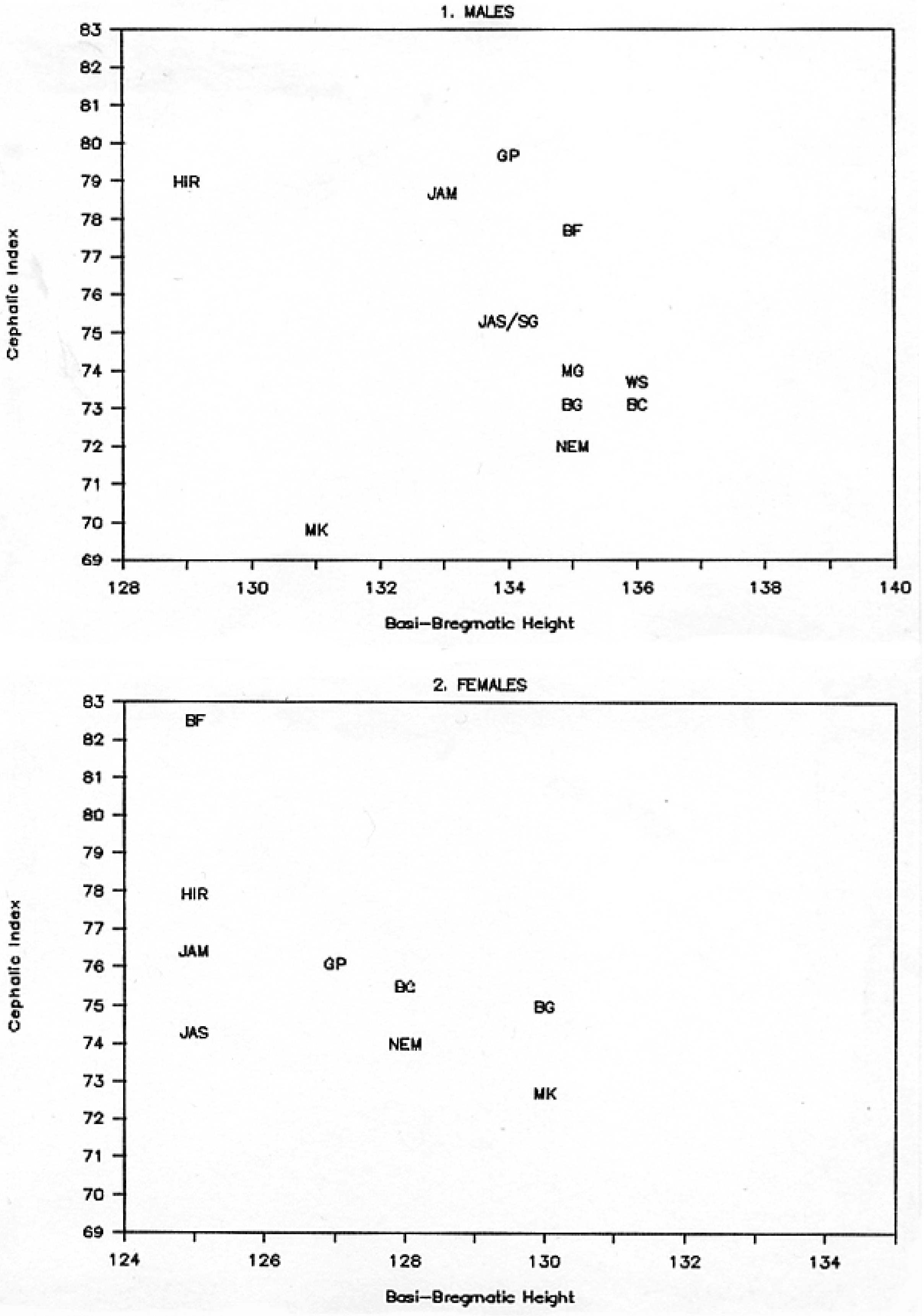


Figure 4.21. Cephalic index against vault height for various groups.