Skeletal Biology of Past Peoples: Research Methods
pages 189–224 © 1992 Wiley-Liss, Inc.

Chapter 11

Paleodemography: Problems and Techniques

Mary Jackes
Department of Anthropology, University of Alberta, Edmonton, Alberta, Canada T6G 2H4

INTRODUCTION

Paleodemographic studies based on human skeletal remains excavated by anthropologists and archaeologists seek to reconstruct basic biological and social facts of human life in the past—population structure, life expectancy, and mortality and fertility rates. The first step in such a study is to establish the distribution of ages at death, and by the 1970s this became a mandatory part of any anthropological report on a cemetery site.

Interest in paleodemography began slowly in the United States with the work of Hooton (1920, 1930), then picked up through the 1960s (Angel, 1968, 1969; Biraben, 1969; Churcher and Kenyon, 1960; Johnston and Snow, 1961; Swedlund and Armelagos, 1969). Paleodemography reached its apogee in central Europe with the publication of *The History of Human Life Span and Mortality* (Acsádi and Nemeskéri, 1970), and the 1970s saw paleodemography emerging as a topic for theses in several American graduate schools (e.g., Asch, 1976). While paleodemographic methods were studied in broad terms (Swedlund, 1975; Swedlund and Armelagos, 1976; Weiss, 1973, 1976), the techniques underlying the data on which the demographic parameters were based received little consideration.

Analysis of the large Libben sample led to the fullest expression of paleodemography based on methods developed in the United States between 1920 and the mid-1970s (Lovejoy et al., 1977). But it also contributed to increasingly trenchant questioning of this traditional paleodemography (Petersen, 1975; Howell, 1976). Howell's paper on Libben demography (1982) forcefully pointed out that the results of paleodemographic research were not in accord with the knowledge accumulated by demographers over 200 years.

In Europe, Masset arrived at a similar position through a consideration of age estimates (1971, 1973, 1976a,b,c). His work was continued by Bocquet-Appel (Bocquet, 1977); their jointly published doubts about the accuracy of age estimates (Bocquet and Masset, 1977; Bocquet-Appel and Masset, 1982) elicited a strong reaction on behalf of paleodemography in the United States (Van Gerven and Armelagos, 1983; Buikstra and Konigsberg, 1985; and Greene et al., 1986).

New histological techniques developed in forensic science during the 1960s and 1970s were expected to provide more accurate age estimates and during the 1980s they were tested on archaeological samples. At the same time, traditional methods of assessment continued to be refined (Katz and Suchey, 1986, 1989), and the methods used to age the Libben sample were published in great detail (e.g., Meindl et al., 1980, 1983, 1985; Meindl and Lovejoy, 1989). Such detail is desirable, a departure from the norm of the previous two decades, when aging techniques were generally accepted without question and, in the obligatory paragraph on "methods," were described in reports in the most general terms or not at all.

This chapter will summarize what has been learned over the last decade and propose methods of research that allow not so much a resurrection of paleodemography (cf. Bocquet-Appel and Masset, 1985), but its reincarnation in a humbler but wiser form. Paleodemographic methods must be suited to a foreseeable future of limited access to skeletal materials and of funding restraints.

With this in mind, there must be attempts to develop techniques that allow comparisons of data among sites and over time. Comparison alone can help identify sites that are biased by differential burial practices and/or by partial excavation. Comparison alone will allow researchers to differentiate among archaeological sites and see the relationship of archaeological mortality to historical, modern, and model mortality rates.

The primary concern in this chapter is with the techniques of adult age assessment, as well as a focus on the crux of the 1980s disagreements between European and North American paleodemographers, that is, on the basic question: How reliable are these age assessment techniques? If age estimates turn out to be inaccurate, there is no foundation on which to build "biological" paleodemography. Paleodemography will then become completely "archaeological," relying on such data as size and density of sites, houses, and hearths.

Throughout the discussion some of the techniques described are evaluated in the context of a sample of several hundred burials from Portugal dating from the Mesolithic and Neolithic (Jackes, 1988; Lubell and Jackes, 1988; Lubell et al., 1989).

MORPHOLOGICAL TECHNIQUES OF AGE ASSESSMENT

Pubic Symphyses

Todd (1920) first demonstrated the systematic changes on the face of the pubic symphysis. A new schema was drawn up by McKern and Stewart (1957) on the basis of their work on the young conscripts killed in the Korean War. Gilbert (1973) pointed out that the application of male standards to females was bound to introduce error, and Gilbert and McKern (1973) proposed a new system for aging the female os pubis. Work has continued on the Todd system (Brooks, 1955), and Suchey, after testing the accuracy of the 1973 female aging system of Gilbert and McKern and finding it wanting (Suchey, 1979), has refined the Todd schema in order to increase the accuracy of age estimates on male (Katz and Suchey, 1986) and female (Brooks and Suchey, 1990) pubes.

In Europe Nemeskéri independently developed a system of pubic symphysis age assessment (Nemeskéri et al., 1960). In seeking to apply it, European researchers noted that different results derived from U.S. and European techniques. Masset (1976c) pointed out that each method would give a different age for the same stage and, citing a study comparing the Todd and Nemeskéri methods on a Czech known-age sample, he concluded that each method is most applicable to a group with an age structure equivalent to that used to develop the method (Masset, 1976a: 334). Bocquet-Appel and Masset (1982) based their "Farewell to Paleodemography" on the opinion that the results of a technique depend on the age-at-death distribution of the sample used in developing the technique.

This was restated slightly by Jackes (1985) in demonstrating that each technique does in fact give different results. It was shown that, in cases where pubic symphyses alone are used in age estimates, it is sometimes possible to determine the technique used from the resulting age-at-death distribution. During the 1960s and 1970s the mean ages of the pubic symphysis stages were used for age assessment: the mean age was used as the age of the individual assessed at a given stage. In sites in which multiple aging techniques could not be used (e.g., ossuaries consisting of disarticulated remains), characteristic age-at-death curves were produced.

This replication of age-at-death distributions became clear in the comparison of North American sites of the period just before and after contact with Europeans, a period of extreme demographic change. Expected differences in mortality patterns were masked by the use of pubic symphysis age estimates (Jackes, 1986). Ten years earlier the interpre-

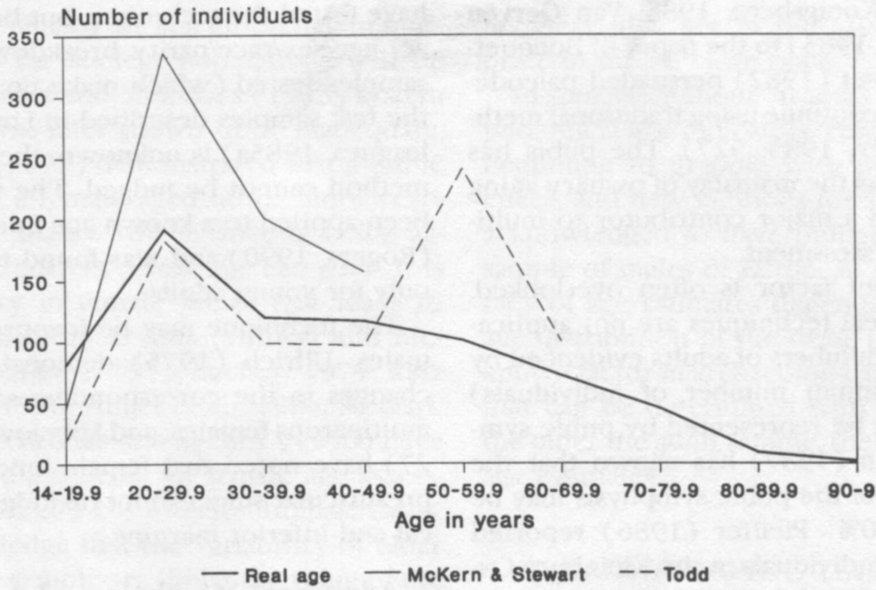


Fig. 1. Age estimations of males based on McKern and Stewart and Todd pubic symphysis techniques compared with known ages derived from data from Katz and Suchey (1986).

tation of this phenomenon of "similar results from dissimilar sites" had been quite different: "archeological data from different sites or different times at the same site produce similar demographic results, so we can have confidence that whatever we are doing, we are doing consistently" (Weiss, 1976: 357). And yet Acsádi and Nemeskéri (1970) had already shown that the European and American methods of age assessment gave fundamentally different pictures of past mortality.

We now have clearer evidence that different techniques produce different results. Katz and Suchey's work (1986) on a large sample of males of known age, clearly demonstrates the characteristic distribution of ages at death that results from the McKern and Stewart technique (see Fig. 1). We can now understand why, for many sites, "everyone is reported to have died before age 45 or 50" (Weiss, 1973: 59). This relative absence of the elderly is still accepted, although with acknowledgments that compensatory adjustments must be applied.

The Todd method also underages those over 60 (Fig. 1; see also Meindl et al., 1983, and Jackes, 1985). Nevertheless, Brooks (1955: 571) obtained quite high correlations of assessed and known age in her tests of Todd's system. However, the tests utilized a sample derived from Todd's original series selected so that "each pubic phase has approximately the same number of individual skeletons whose known age falls within that span." Meindl and colleagues (1983) also tested the Todd method. They obtained a correlation of .57 for males and females combined, on samples selected from the Todd collection to reflect the supposed age structure of archaeological populations (i.e., very few old people). Neither the Brooks nor the Meindl and colleagues test is thus independent of the age distributions imposed by the method, nor was either carried out on a sample other than the reference sample.

The Nemeskéri method results in assigning far too many individuals an age of 45–60 years (Jackes, 1985), and Bocquet et al. (1978) found a very low correlation with real age in testing the Nemeskéri pubic method on a known-age population. Brooks and Suchey (1990) advise against the use of the method, after testing it on 1,000 modern pubes.

Examination of the pubic symphysis was the basic method of choice for every skeletal biologist through much of the past three decades (e.g., Blakely, 1971: 45; Bender, 1979: 185). U.S. researchers (Meindl et al., 1983: 81; Lovejoy et al., 1985a: 12; Meindl et al., 1985) drew attention to the inaccuracy of pubic indicators, but the strong negative responses from others

(Buikstra and Konigsberg, 1985; Van Gerven and Armelagos, 1983) to the paper of Bocquet-Appel and Masset (1982) persuaded paleodemographers to continue using traditional methods (e.g., Storey, 1985: 527). The pubis has been regarded as the mainstay of ossuary aging techniques, and a major contributor to multiple-factor age assessment.

One important factor is often overlooked. Pubic symphyseal techniques are not applicable to all sites: numbers of adults evidenced by the MNI (minimum number of individuals) counts may not be represented by pubic symphyses. Waldron (1987) has shown that the rate of survival of the pubic symphysis may be only around 30%. Pfeiffer (1986) reported over 400 adult individuals in the Kleinburg Ossuary but only 51 adults (12.5%) were given ages by pubic symphyses (Jackes, 1985, 1986). Similarly, only 95 of some 312 adults (30%) could be aged on pubic symphyses in the Uxbridge Ossuary (Pfeiffer, 1986). Of the several hundred adults in recently analyzed Portuguese Mesolithic and Neolithic sites (e.g., Jackes, 1988; Lubell and Jackes, 1988; Lubell et al., 1989) only two or three were represented by analyzable symphyses. The exclusion of adults of indeterminate age from paleodemographic analyses has profound effects on the results obtained.

Auricular Surfaces

Lovejoy and colleagues (1985b) proposed an age-estimating technique based on the auricular surface of the ilium, which changes with age in specific ways that are highly correlated with changes in the pubic symphysis (Lovejoy et al., 1985a). The method requires a large series of clean and well-preserved auricular surfaces and controlled seriation. Both the Koyabashi (1967) form of this observation and the Lovejoy method, applied to Portuguese Mesolithic samples, were found to be unsatisfactory owing to the condition of the auricular surfaces and constant interobserver disagreements. However, a test on an independent known-age sample has suggested that the auricular surface is of more value than the pubis, producing interobserver correlations of 0.8 or more and a correlation with age of 0.6 (Bedford et al., 1989). Lovejoy and colleagues (Lovejoy, 1985b; Meindl and Lovejoy, 1989) have tested the technique, but because the exact age/sex/race/parity breakdown of the two samples tested (which make up 75%–82% of the test samples described in Lovejoy and colleagues, 1985a) is unknown, the utility of the method cannot be judged. The technique has been applied to a known age European sample (Rogers, 1990) and was found to be accurate only for young adults.

The technique may be inappropriate for females. Ullrich (1975) demonstrated marked changes in the corresponding sacral joints of multiparous females, and Lovejoy et al. (1985b: 27) have noted that female innominates with preauricular sulci cannot be judged on the apical and inferior margins.

Additional Morphological Age Criteria

Several other methods by which adult age can be estimated have been proposed: atrophy of the scapula (Graves, 1922), which is rarely useful for archaeological material, since the scapular blade is fragile; fusion of the sternum, judged of no value by Jit and Bakshi (1986); fusion of the cranial sutures, which is highly variable and the subject of considerable literature (e.g., McKern and Stewart, 1957; Nemeskéri et al., 1960; Meindl and Lovejoy, 1985; Masset, 1989); degenerative changes (Stewart, 1958), which are activity-, sex-, and population-specific in ways not fully documented; fusion of the maxillary suture, found to be of limited value by Gruspier and Mullen (1991); and the sternal end of some ribs (İşcan et al., 1984, 1985).

Only the last of these techniques appears potentially useful to skeletal biologists for aging large collections of skeletons. Unfortunately, it requires excellent preservation, identification and sex assignment of the third, fourth, or fifth rib, limiting its value in sites with disarticulated, mixed, or incomplete individuals. Furthermore, racial differences appear to be marked.

Sex, Race, and Morphological Age Assessment

Early work on pubic age estimation was restricted to males. Todd (1921) included females, but his sample was too small for analysis of variability. Stewart (1957: 18) stated that "until more reliable pubic age standards are available for females the sex difference in mortality curves for ancient populations will be suspect." Gilbert and McKern's (1973) system for females was later shown to be inadequate, and Suchey (1979) demonstrated that a pubic symphysis of a female aged 60 or more years had no more than a 10% chance of being assigned to the correct 5-year age category. It is the variability in female pubes that leads to these unsatisfactory results (Gilbert and McKern, 1973; Gilbert, 1973; Suchey, 1979; Katz and Suchey, 1986; Angel et al., 1986; Stewart, 1957). While Brooks and Suchey (1990) have now provided a schema for female age assessment based on 273 Los Angeles forensic cases, they acknowledge that the variability of older female pubic symphyses limits the value of the technique. It must be used in conjunction with other methods.

Sex has been a recognized problem in morphological studies, but racial differences have rarely been discussed in the context of morphological age assessments. Todd (1921) concluded that differences between black and white pubic symphyses existed but were unimportant. In further work on the same collection, Meindl and colleagues (1985: 33) acknowledged that their technique results in greater underaging of 60- to 69-year-old blacks than whites.

Buikstra and Konigsberg (1985) cited Hanihara and Susuki (1978; a test limited to individuals between 18 and 38) to prove that pubic symphysis age distributions are not population-specific. At the time, only one good test of this existed. Koyabashi (1967) tested the age estimation techniques on a sample of 142 adolescent and adult Japanese skeletons of known age; age changes in Japanese pubes seemed less obvious after age 35 than Todd or Brooks would suggest on the basis of U.S. blacks and whites.

Katz and Suchey (1989) have now shown that significant differences exist between white, black, and Mexican males in their Los Angeles pubes sample. The average age of a stage VI pubis in a white is 64 years (probability range 53–75 years), while an equivalent Mexican pubis would be about 47 years ($\pm 2\sigma = 39–55$ years). However, since only 2% of the Suchey sample of individuals over age 50 were Mexican, (none over 59 years), there is no clear idea yet of the age changes taking place in the Mexican pubis.

In sum, after about 70 years of studying morphological age indicators, researchers are only beginning to grapple with population differences, and new methods for aging females are acknowledged to have limitations. Even for a sample of males of European origin, the accuracy of age estimates depends directly on the age distribution of the dead, precisely the unknown factor in archaeological populations. All that can be determined with certainty is that the older the adult sample, the less accurate the age estimates.

RADIOGRAPHIC TECHNIQUES IN AGE ASSESSMENT

The cortical width and density of bone and the arrangement of trabeculae undergo changes with age that have been identified by radiographers and clinicians searching for techniques by which osteoporosis can be diagnosed or predicted. Anthropologists attempting to assess skeletal age in adults have recently used X-rays, especially to examine changes in trabeculae in femoral and humeral heads (summary in Sorg et al., 1989). Other techniques have either not been attempted on archaeological bone (calcaneus: Walker and Lovejoy, 1985) or have been tried and rejected (radius: Lovejoy et al., 1977, 1985a; clavicles: Jhamaria et al., 1983; Meiklejohn, personal communication).

Alteration in Femoral Trabeculae

The Singh Index, systematizing changes in the proximal femur, is used in clinical studies as a method of evaluating osteoporosis (Singh, 1972). Detailed literature on this index indicates that accurate ages cannot be assigned on the basis of femoral radiographs and that, generally, femoral scores have lower correlations with age than the index of second metacarpal cortical bone (e.g., Nordin et al., 1966). High correlations with age have been found ($r = .91$; Jhamaria et al., 1983), but the sample used could be regarded as biased. By contrast, Singh's data for female femora suggest a correlation with age of .6 (Singh, 1972, data extracted from Fig. 3, p. 65).

The first important use of radiography for anthropological age determination was by Nemeskéri and colleagues (1960; see also Acsádi and Nemeskéri, 1970), when they added age changes in the trabeculae of the proximal humerus and femur to pubic symphyseal face modification and ectocranial suture closure.

Bergot and Bocquet (1976) redefined the six stages of trabecular change. Subsequently Bocquet and his colleagues (1978) tested a method of estimating age by multiple regression, by using pubes, cranial sutures, and humeral and femoral radiographs on the known-age collection at the University of Coimbra, Portugal. The results based on humeri were very unsatisfactory, but at $r = .7$ for males and .6 for females, the trabecular stage of the proximal femur showed a more encouraging correlation with real age.

Walker and Lovejoy (1985) have proposed a fourth schema for analyzing femoral trabecular changes with age. They initially used an optical densitometer to avoid subjectivity in reading radiographs, but found it of no value. Their results are therefore based on a seriation of X-ray images, with emphasis on morphological changes in trabeculae and relative translucency of the various areas in the proximal femur as in other methods.

A series of inter- and intraobserver tests by five researchers on a large number of Portuguese Mesolithic and Neolithic femoral radiographs has been undertaken. Many radiographs are unreadable because of damage to the bone or heavy matrix in the medullary canal, and the lack of standardization enforced by differing burial conditions has been shown to bias results. Paired t tests demonstrated that interobserver differences are most significant when the Walker–Lovejoy technique is used (n = 73; 63% disagreement) and least with the Nemeskéri method (n = 80; 52% disagreement).

The Bergot–Bocquet and Nemeskéri methods generally correlate at higher levels, and the Walker–Lovejoy method consistently has lower correlations with other methods, irrespective of observer or site. Interobserver differences are less significant for Neolithic site radiographs, demonstrating the importance of X-ray clarity and bone preservation. Only the Walker–Lovejoy technique produces significant interobserver differences when these are tested by cultural period, and this is found only with the Neolithic material, indicating that it performs least well on the best preserved material.

No single method stands out as consistently performing better than the others, although the Walker–Lovejoy is in general the least satisfactory. The Walker–Lovejoy stages derive from seriation and are therefore often imprecisely defined in terms such as "more," "less," "greater," and "significant increase." This compounds the subjectivity inherent in reading radiographs (Bergot and Bocquet, 1976).

A number of problems must be resolved before a useful application of techniques of scoring femoral trabecular involution to archaeological material is possible. The comparability of grades among methods is questionable, and in several cases the descriptions of stages are ambiguous. Moreover, the stages may not be comparable between populations (e.g., in many of the Portuguese radiographs, subcapital rarefaction appears advanced over the formation of a clear Ward's Triangle, making Singh Grade 6 inappropriate as a description of observed changes).

On the basis of the work on Portuguese radiographs, this author proposes a scheme by which seven radiographic characteristics of the femoral head are deemed to pass through eight phases. Each area (the apex of the canal, Ward's Triangle, subcapital area, fovea capital area, the greater trochanter, the arciform bundle, the cephalic bundle) is observed in order to clarify whether all pass through the equivalent stage at the same time in all individuals. Age is assessed on the basis of clustering of the scores on each of the seven areas.

Of the 126 or so individuals represented by the Portuguese radiographs, few are taken to represent older people (determined on the basis of clustering trabecular stage with cortical variables, to be described below). This is almost inevitable, as less dense proximal femora do not survive. Thus radiography underrepresents the elderly on the basis of preservation, and because some elderly present "young" proximal femora, they will be underaged. Poor preservation of femoral heads suggests that it is worthwhile to consider another radiographic technique: measurement of cortical thickness.

Cortical Thickness

Cortical thickness is determined from radiographs for diagnostic purposes by two midshaft measurements, especially on femora: the Nordin Index (Barnett and Nordin, 1960) is the width of the femoral cortex expressed as a percentage of the shaft diameter. Other cortex widths are also measured, especially the second metacarpal (Barnett and Nordin, 1960; Morgan, 1973). Measurement of the cortex width in long bones has been applied by anthropologists to determine bone loss associated with disease or nutritional stress (summary in Macchiarelli, 1988).

Humeri have been suggested as preferable to femora, in that they are not weight-bearing (Smith et al., 1984). Walker and Lovejoy (1985) propose radiography of the clavicle, and their phases for the clavicle and for the proximal femora both refer to cortical thinning. Kaur and Jit (1990) have shown by direct measurement that the clavicular cortex is reduced markedly with age.

Measurement of the medial femoral cortex, just above the lesser trochanter, has been discussed by radiologists (e.g., Horseman et al., 1982). This area, the calcar femorale, extends to just below the lesser trochanter, where it is often complete in archaeological bone and is readily identifiable. Intraobserver error on the external measurement at this point is less than that for the midshaft external measurement in tests on Portuguese femora. Measurement of the medullary width at the subtrochanteric level is less accurate than simple measurement of the lower calcar femorale, judged on the basis of correlations with proximal femoral trabecular scores. One test has shown that the average difference between the subtrochanteric index (expressing the width of lower calcar femorale in relation to the external subtrochanteric diameter) on multiple X-ray films of the same bones is half that of the average difference for Nordin's Index. We can assume that Nordin's Index is much more sensitive to variations in radiographs. An index that expresses the breadth of the calcar femorale may be of value as an additional variable in analyses of the seven characteristics of proximal femoral trabeculae discussed above.

Intraobserver error and multiple interobserver error on radiographic indices have been tested on the Portuguese bones and found to be high—often significantly high. Direct measurement of the cortex on midfemoral bone samples is preferable. The ratio of bone sample cortical thickness to anterior posterior midfemoral diameter has low correlations with Nordin's Index (below .5 even on the Neolithic femora) based on multiple tests on the Portuguese femora. Radiographic cortical measurement error has been found to be high under the more ideal conditions of clinical trials: it seems reasonable to assume that Nordin's Index will have a very high error for archaeological material.

THE MULTIFACTORIAL APPROACH

Whenever complete individuals are available for analysis, multiple age assessment methods can be used. Nemeskéri et al. (1960) were the first to attempt a systematization of multiple aging techniques.

Acsádi and Nemeskéri (1970) have proposed that four age indicators be used in age estimation. First, the pubic symphysis is used to determine whether the individual is young, about 50 years, or old. On this basis, cranial sutures and trabeculae in the proximal humerus and femur are analyzed, and the pubis is reexamined. Age ranges for each point of observation are fixed, and age is determined by choosing the low, middle, or high end of each range, according to the initial information given by the symphysis pubis. The four ages thus derived are then averaged to give the final estimated age for the individual. Acsádi and Nemeskéri (1970: 123) claim that, on the basis of their study with individuals of known age, "actual, chronological age can be approximated fairly well." Accuracy is 80%–85% with a margin of error of ±2.5 years in individual cases and a lower error if ages of a sample are distributed (1970: 131; but see Sjøvold 1978: 111).

Concluding that three people died at age 18–20 requires drawing up the age-at-death distribution on the assumption that 1.0 (3 years/3 individuals) person died at each of ages 18, 19 and 20. But as Sjøvold (1978) pointed out, equal distribution is only one of a number of possible ways of combining 3 dead and 3 years; in fact it has a 0.1 probability, and for each individual the probability of death at a given age is only .33. Sjøvold, however, maintains that the confidence limits for methods of age determination are unknown and that a normal distribution of ages cannot be assumed for every age class of every age indicator, a contention which receives some support from Brooks and Suchey (1990: Fig. 4).

The accuracy of the method of Acsádi and Nemeskéri has been further tested on the Coimbra sample by Bocquet et al. (1978), who used six age indicators (including humeral and femoral cortical widths) on 194 males and 161 females. Of the six techniques, none showed high correlations with the actual population; the humeral cortical index for females, at −.688 (Boquet et al., 1978: 140) was the highest. The use of a theoretical population in which all age classes were equally represented provided correlations of .71 for both the male proximal femur and the female humeral cortical index. Combining all age indicators generates correlation coefficients of .829 for males and .789 for females. The combined method does not give results much better than the single indicator results, and Bocquet-Appel would maintain that only results in the $r = .9$ range can be used in paleodemography.

It is appropriate to enquire whether we know the true correlation of multiple techniques with age, since some have a linear relationship with age and others (e.g., cortical widths) do not. The method of computing ages in the face of this problem has not been discussed.

Another proposal for multiple techniques derives from the analysis of the Libben site (Lovejoy et al., 1977). Age estimates were based on pubis, proximal femur and distal radius radiography, dental attrition, cranial sutures, and auricular changes. Vertebral osteophytosis had been one of the original age indicators, but it was not used in the final assessments of age. The estimate of the reliability of the other six indicators and the weight assigned them was derived from a principal components analysis. Each indicator was weighted according to its contribution to the first principal component. The contribution to the first principal component, which accounted for 74% of the variance, was taken as an estimate of reliability. The weightings have not been published.

The accuracy of the technique was post facto tested on the basis of three samples chosen from among 500 specimens with "reliable records of age at death" (Meindl et al., 1980) in the Todd collection. Meindl et al. (1983) subsequently published an account of testing on two samples of 130 each. The skeletons were carefully chosen: in order "to mimic archaeological conditions two adult life tables were selected from the literature and two samples of 130 cases each . . . were chosen . . . to correspond to the two tables" (1983: 76–77). Sample 1 appears similar to Point of Pines (Bennett, 1973), while a distribution like that of Indian Knoll (Johnston and Snow, 1961) may have provided the basis for sample 2. Both sites were aged by McKern and Stewart pubic indicators and by dental attrition, although attrition coding methods differed between the two sites.

The composition of the two Todd Collection samples is shown in Figure 2. Sample 1 has fewer young adults and non-Europeans than sample 2. As "accuracy of age assessment is greatest in the younger age categories" (Lovejoy et al., 1985a: 9) and as pubic indicators are more likely to underage Europeans than Africans or Asians, it is predictable that correlations between indicators and known age will be higher for sample 2 than for sample 1. Mensforth and Latimer (1989) have shown that osteoporotic fracture rates are much lower among Todd Collection blacks than whites (e.g., among women in their fifties the black fracture rate is one quarter that of the rate for white women), so a young age given to a black on the basis of radiographs is most likely to be correct for a sample in which there are many young individuals. Again sample 2 should produce higher correlations between known age and estimated age than sample 1.

Methods and the results of age estimations of the two samples are discussed in Meindl et al. (1983) and Lovejoy et al. (1985a). Correlations between dental, auricular, and femoral ages and real age remain the same, but the correlation of the pubic symphysis, in its revised form, rises to .78 for sample 2.

The assumption outlined above, that accuracy would increase on the second test, holds true only for the pubis, and yet the results of the sample 2 tests have been published as im-

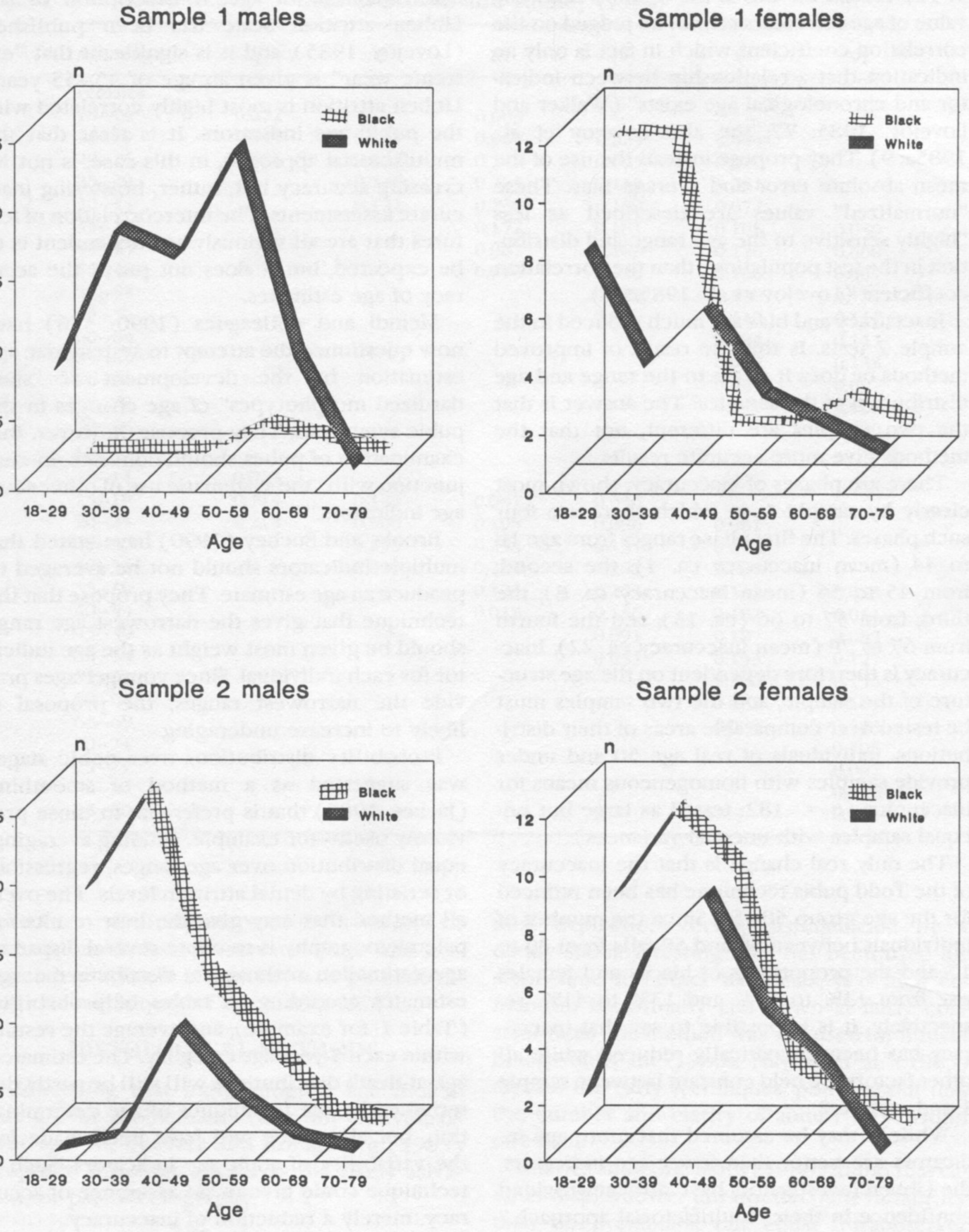


Fig. 2. Composition of the Hamann–Todd Collection samples used to test age indicators (Lovejoy et al., 1985a: Table 1).

provements over the methods used for sample 1. The reason for this is the opinion that "the value of age indicators cannot be judged on the correlation coefficient which in fact is only an indication that a relationship between indicator and chronological age exists" (Walker and Lovejoy, 1985: 77; see also Lovejoy et al., 1985a: 9). They propose instead the use of the mean absolute error and average bias. These "normalized" values are described as less "highly sensitive to the age range and distribution in the test population" than the correlation coefficient (Lovejoy et al., 1985a: 9).

Inaccuracy and bias are much reduced in the sample 2 tests. Is this the result of improved methods or does it relate to the range and age distribution of the samples? The answer is that the two samples are different, not that the methods give more accurate results.

There are phases of inaccuracy, shown most clearly by sample 1, in which there are four such phases. The first phase ranges from age 18 to 44 (mean inaccuracy ca. 4); the second, from 45 to 56 (mean inaccuracy ca. 8); the third, from 57 to 66 (ca. 13); and the fourth from 67 to 78 (mean inaccuracy ca. 22). Inaccuracy is therefore dependent on the age structure of the sample, and the two samples must be tested over comparable areas of their distributions: individuals of real age 50 and under provide samples with homogeneous means for inaccuracy ($p = .182$: tested as large but unequal samples with unequal variances).

The only real change is that the inaccuracy of the Todd pubis technique has been reduced for the age group 50–59. Since the number of individuals between 50 and 59 falls from 30 to 17 and the proportions of blacks and females rise from 13% to 47% and 13% to 41%, respectively, it is impossible to say that inaccuracy has been dramatically reduced while all other factors are held constant between sample 1 and sample 2.

While it may be assumed that more age indicators are better than fewer age indicators, the Libben investigators have not yet provided confidence in their "multifactorial approach." Apart from the pubic symphysis, which is not reliable, they emphasize the value of the auricular surface, which is highly correlated with the pubic symphysis and not universally applicable to archaeological material. They consistently note dental attrition as the most unbiased indicator of age. A description of the Libben attrition scale has been published (Lovejoy, 1985), and it is significant that "extreme wear" is given an age of 45–55 years. Libben attrition is most highly correlated with the pubic age indicators. It is clear that the multifactorial approach, in this case, is not increasing accuracy but, rather, bolstering inaccurate assessments. The intercorrelation of features that are all variously age-dependent is to be expected, but it does not prove the accuracy of age estimates.

Meindl and colleagues (1990: 356) have now questioned the attempt to systematize age estimation by the development of "standardized morphotypes" of age changes in the pubic symphysis. They propose, however, that examination of pubes should continue, in conjunction with "the systematic use of other adult age indicators."

Brooks and Suchey (1990) have stated that multiple indicators should not be averaged to produce an age estimate. They propose that the technique that gives the narrowest age range should be given most weight as the age indicator for each individual. Since younger ages provide the narrowest ranges, the proposal is likely to increase underaging.

Probability distributions over pubic stages was suggested as a method of smoothing (Jackes, 1985) that is preferable to those previously used—for example, running averaging, equal distribution over age ranges, regression, or seriating by dental attrition levels. The overall method that may give the best results for paleodemography is to chose several disparate age estimation techniques, distribute the age estimates according to tables of probability (Table 1, for example), and average the results within each 5-year age category. The estimated age-at-death distributions will still be partly determined by the techniques of age determination, but allowance will have been made for the variability of adult age indicators. Such a technique could give us no assurance of accuracy, merely a reduction of inaccuracy.

Finally, multiple techniques can be applied in a controlled way only in sites with complete skeletons and excellent preservation. Under the circumstances obtaining for most archaeological skeletal collections, published ages will

TABLE 1. Probability Distributions for Pubic Symphysis Stages Based on Data in Brooks and Suchey (1990)

Age range	I	II	III	IV	V	VI
Females						
10–14	0.022		0.003			
15–19	0.545	0.131	0.067	0.025	0.004	
20–24	0.386	0.346	0.148	0.065	0.030	
25–29		0.346	0.225	0.113	0.051	
30–34		0.131	0.236	0.159	0.077	
35–39			0.172	0.181	0.105	0.030
40–44			0.087	0.168	0.126	0.060
45–49			0.016	0.126	0.136	0.097
50–54				0.078	0.130	0.133
55–59				0.039	0.110	0.157
60–64					0.084	0.157
65–69					0.057	0.133
70–74					0.034	0.097
75–79					0.010	0.060
80–84						0.030
Males						
10–14	0.025					
15–19	0.714	0.149	0.067	0.030		
20–24	0.215	0.500	0.195	0.086	0.001	
25–29		0.295	0.294	0.151	0.043	
30–34		0.010	0.255	0.202	0.087	
35–39			0.125	0.203	0.141	0.018
40–44			0.018	0.156	0.182	0.051
45–49				0.091	0.187	0.087
50–54				0.035	0.153	0.127
55–59					0.100	0.155
60–64					0.052	0.161
65–69					0.008	0.143
70–74						0.106
75–79						0.067
80–84						0.036
85–90						0.003

have been variously assessed by a variety of techniques. Skeletal preservation, age- and sex-dependent though it may be, will be controlling our techniques of age assessment.

HISTOLOGICAL METHODS

Conceding that morphological age indicators are of problematic value for individuals over age 50, Buikstra and Konigsberg (1985) proposed histological data as the likely salvation for paleodemography.

Teeth

Cemental annulation. Gustafson (1950) first noted that the thickness of cementum increases with age in human beings. While age estimation based on cemental annulations, annual depositions being distinguished by a darkly staining resting line, has been used for some time for other mammals, in which cementum is normally laid down at fairly constant rates, the method was not used for human beings until the 1980s. Naylor et al. (1985) discuss the early techniques, pointing out that the number and clarity of annuli vary within one root.

Charles and colleagues (1986) found that canines display lower intratooth count variability than do premolars, but equal intraobserver variability. However, higher interobserver variability for canines than premolars led them to conclude that premolars give better results. They did not discuss the accuracy of age estimates for canines, and although they hypothesized that an excess of annuli may be counted in poorly defined areas of cementum when the cementum is thicker (as in canines), this was not evaluated.

Condon and coworkers (1986) found a correlation of premolar annuli with age of .7 for males and .95 for females. However, since the mean age of the females in their sample is lower than that of the males, their conclusion that cemental annuli are more accurate predictors of age in females than in males cannot be evaluated.

The results of a study by Miller and colleagues (1988) confirm the finding of Condon and coworkers (1986) that cemental annulation techniques can be used only for young individuals: they report a correlation of .78 between annuli counts and age for individuals under 35, but .01 for those over 35 years. Kay and Cant (1988) studied cemental annulation and attrition in *Macaca mulatta* lower first molars. Age could be predicted accurately for individuals under 10 years, but with increasing error beyond that, especially after age 14. Age was predicted more accurately in males than in females. Kay and Cant conclude that environmental factors may affect the formation of resting lines in cementum.

Neither the preferred tooth nor the technique by which annulations should be examined has been settled. Charles and colleagues (1986, 1989) stress the importance of 7μm decalcified stained sections, but recent attempts to produce these sections from archaeological material (from the Portuguese Neolithic ossuary cave, Casa da Moura, ca. 6877–5735 B.P.) have been unsuccessful. Decalcification and staining after embedding was also unsuccessful; the cementum is reduced to a wisp of discontinuous network, separated from the shredding dentin in a number of places—which was not unexpected, since cementum is much less mineralized than dentin (Furseth and Mjør, 1973).

The Portuguese Neolithic annuli are best visualized directly under polarized light with a lambda compensator wedge to enhance weak birefringence. The incremental lines are often obscured or fade out so that many counts must be abandoned. Both bacterial and fungal destruction are seen, occasionally resulting in almost complete obliteration of the incremental lines. Work on cementum has highlighted a factor that is of critical importance to paleodemography: poor bone preservation may make histological and chemical age assessment less valuable than morphological and radiological age assessment in archaeological samples.

Casa da Moura is in a limestone area and the deposits are therefore alkaline. While the chemistry of such sediments normally promotes tooth and bone preservation (Gordon and Buikstra, 1981), it may also destroy collagen (Protsch, 1986). Since the difference between annuli and their bounding resting lines is that the resting lines have less collagen and more mineral (Furseth and Mjør, 1973), collagen decomposition reduces the contrast between the two areas.

Literature on cementum diagenesis is minimal (see Hillson, 1986: 157, for summary). Sognnaes (1955) and Werelds (1961, 1962, 1967) report microbial damage in teeth and Beeley and Lunt (1980) document the loss of collagen and a decrease in the Ca/P ratio in dentin.

Mineralization increases with age (Fraysse and Kerebel, 1982), so microradiography rather than polarizing microscopy may assist in visualizing adult cemental annuli. The technique has been used successfully on herbivore teeth (Koike and Ohtaishi, 1985), but human annuli may require magnifications above those satisfactorily achieved with microradiographs. The use of reflecting microscopy with acid etched thick sections is now being attempted.

Root transparency and combined methods of age assessment. Gustafson (1950) proposed a technique of dental age assessment based on attrition, amount of root exposed above the gingiva, cemental thickness, root resorption, secondary dentin deposited in the pulp cavity, and transparency of the root. Miles (1963) tested root transparency on incisors and reports a correlation of .73 with known age. Philippas and Applebaum (1967, 1968) discussed age-dependent variations in maxillary lateral incisors and canines.

Burns and Maples (1976) warn that periodontal disease, dental pathology, and diet may alter the rate of dentin mineralization. Their tests of Gustafson's method give correlations of .76 with real age. Maples (1978) showed that root transparency was the best predictor of age, followed by secondary dentin, while multiple regression gives correlations with real age of better than .90 in all teeth except the first molars. The use of root transparency and secondary dentin alone provides fairly high correlations, especially for the central (.89) and lateral incisors and the second molar (.88).

Hillson (1986) and Costa (1986) provide short reviews of methods that have been used to observe root transparency, and Cook (1984) critically evaluates the work of Gustafson and of Johanson (1971) and several of their successors. Kilian and Vlček (1989) summarize Czech research on root transparency. Their own work on the anterior teeth of 116 adults of known age suggests that the relation between dental biological age and chronological age is linear in males, but not in females.

Although transparency of root dentin may be altered during diagenesis (Vlček and Mrklas, 1975), this author has observed transparency in Neolithic maxillary canines. The distribution of transparency is similar to that described by Philippas and Applebaum (1968), rather than the simple advance of transparency from the root tip to half way up the root; some authors (e.g., Bang and Ramm, 1970) have stressed the need to concentrate on the lower root area.

On the basis of the literature summarized above, the oldest tooth in our sample of sectioned Portuguese canines is estimated to be well over 80 years of age, and all ages for the canine sample accord with the sequence of wear levels established for the Portuguese Neolithic. The combined method has potential for archaeological age assessment, especially if annulation counts are used in conjunction with all other indicators. There is a need to define which teeth are best used for both root transparency and cemental annulation studies, taking into account that anterior teeth may be more affected by trauma and nondietary functions, and posterior teeth by pathology. Most importantly, the effect of heavy attrition must be studied (e.g., Neolithic canines display complete infilling of the pulp chamber at an early age), and tests must be undertaken to establish whether age estimates are equally reliable in populations in which dental attrition rates and pathology differ.

Cortical Bone Microstructure

Femoral cortical remodeling. Osteon counting techniques are well summarized in several review articles (Stout, 1989a,b; Ubelaker, 1986) and by Stout (this volume, Chapter 3). Some tests on modern forensic and dissecting room material have shown formidable accuracy: for example, Uytterschaut's (1985) test of the Ahlqvist and Damsten (1969) method produced a correlation with age of .959.

On the other hand, Bocquet-Appel et al. (1980) had poor results from the Coimbra sample. Their best results ($r = .7$) were based on the number of osteons and osteon fragments. They hypothesize that their results may be disappointing because they took 6-mm cores slightly above, rather than at, the femoral midshaft. However, the site of coring is within the range of error likely to occur on incomplete archaeological femora.

Lazenby (1984), finding both inadequate sample sizes and bias in the reference samples, doubts the accuracy of all cortical remodeling age estimates. Others (Bouvier and Ubelaker, 1977; Stout, 1978; Stout and Gehlert, 1980) have compared techniques, and their conclusion that Kerley's technique (1965) provides more accurate results is supported in part by Drusini (1987). Pfeiffer's (1985) analysis of the Kleinburg Ossuary demonstrated that estimates obtained by the Ahlqvist and Damsten method differ from those obtained by the Thompson (1979) method.

The Thompson method has been tested on an independent sample, though still largely from the same general population (forensic cases in the United States) as the original reference group, and been analyzed by the same researchers using the same microscopes: the definition of the structures observed and the optics involved are important considerations in tests of histomorphological age assessment techniques. In this test, Thompson (1981) and Thompson and Gunness-Hey (1981) found a correlation of .83 between real and estimated age. However, only 54% of individuals were placed in the correct 5-year age intervals, and the resulting differences in demographic parameters are considerable. Moreover, Americans, black or white, are more likely to be aged correctly than those representing disparate dietary regimes and/or environments.

Drusini (1987) has a number of criticisms of the Thompson technique and found it inaccurate in a test of Italian specimens. Similarly, the correlations between real age and age estimated by eight variables from the Thompson technique are quite low (.75 in females and .58 in males) for middle-aged to elderly Japanese (Narasaki, 1990).

Martin and colleagues (1981) argue that many environmental factors affect cortical remodeling rates (see also Stout, this volume, Chapter 2). Ericksen (1980), Richman and colleagues (1979), Thompson and Gunness-Hey (1981), and Thompson and colleagues (1984) have all reported population differences attributed to genetic and/or environmental factors (see also Pfeiffer, 1980, and Stout and Simmons, 1979). Stout (1983) implicates diet as a major variable. It seems unlikely that the differences between femora from Ledders (agricultural) and Gibson and Ray (preagricultural) can be attributed to differences in the age assessment techniques used by Buikstra (1976), but clarification is needed, since the bone formation rates for Ledders are consonant with an older age range. The differentiating factors for femoral cortical remodeling could be as fundamental as the level of physical activity (Ubelaker, 1974).

Consequently, the universal application of regression formulas based on North American forensic cases must be questioned. Ubelaker (1986:246) has applied Kerley's technique to 158 poor from the Dominican Republic and has shown that "the rate of age change in the San Domingo sample is somewhat different from the sample used in the original Kerley study." Population differences may also explain the unsatisfactory results in the Portuguese and Italian analyses (Bocquet-Appel et al., 1980; Drusini, 1987).

However, the problems are even more basic: the recognition of forming and resorbing osteons, of cement lines, of osteon fragments, even in some cases of the osteons themselves, since bone histomorphology is not maintained in all archaeological bone.

The histology of ancient bone has been studied by Baud (1987) and Garland and colleagues (Garland, 1987, 1989; Garland et al., 1987, 1988). Both have found that the attack on bone by microorganisms is of major importance. Garland et al. (1987) state that 100% of their sections contain exogenous material in bone spaces and 73% of archaeological bone may be destroyed by physical weathering or (in most cases) by microbial activity.

Stout (Stout and Teitlebaum, 1976) studied archaeological rib and found 39% of the Ledders sections were unreadable (Hanson and Buikstra, 1987). Differential preservation of the ribs of older and younger individuals must be considered in this context. Only Masset (1973) has considered differential preservation in any detail. Its importance has been clearly stated by Boddington and colleagues (1987: 4):

> The more porous the bone and the less dense the bone, the more susceptible it is to destruction. Porosity and density are factors which depend upon age, sex and health of the individual as well as varying between individual bones of the body. Hence, the nature and rate of decay is as much a product of the buried skeleton as the burial environment.

No research has focused on the relationship of diagenesis and age at death (but see Walker et al., 1988, on age-related preservation bias favoring young adults).

Research on the effect of the burial environment has barely begun. Archaeological bone may be exceptionally well preserved in dry environments (Martin et al., 1981; Cook et al., 1989), and perfect preservation is possible in permafrost (Amy et al., 1986). Ubelaker (1974) appears to have had a 100% success rate sectioning precontact American ossuary material, but well-preserved material is the exception, not the rule. Pfeiffer has made several attempts at osteon counting methods of age assessment. Her success rates ranged from 100% for a Late Archaic site (Pfeiffer, 1980), through 27%–37%, depending on the technique employed, for a late sixteenth century site (Pfeiffer, 1985), to 11% for an early nineteenth century site (Pfeiffer, 1989). Black (1979) abandoned cortical remodeling age assessment after producing no readable thin sections. Samson and Branigan (1987) also abandoned the attempt to count osteons in favor of quantifying porosity, as did Palmer (1987).

Histomorphological age estimation has been attempted on femora from Portuguese archaeological sites: two Mesolithic open-air estuarine middens dated 7929 and 7604 B.P. (calibrated weighted averages) and six Neolithic sites dated 6877–4638 B.P. (Lubell and Jackes, 1988). The Neolithic sites include deep limestone caves in the interior mountains and open sandy caves near the coast. One additional sample comes from a medieval rock grave burial.

Results of cortical remodeling techniques on the Portuguese sample. The Portuguese femoral sections exhibit physical, chemical and microbial destruction (Fig. 3). Unless thin sections are cleared with, for example, hydrogen peroxide (xylene is of no value here, cf. Stout, 1989a: 43), structures appear to be obscured by a dark granular congestion. In sections where this granulation is pronounced, fluorescence is increased and birefringence under polarized light is markedly reduced. In some cases, only a few birefringent areas can be discerned over a half section of femur. Histomorphology may be completely disrupted. While a few "ghost" areas of lamellae may be seen, most often near Haversian canals, the lamellae are often discontinuous and disintegrate into multiple circular or ovoid forms (Fig. 3A,B). Hackett (1981) has called these forms "lamellate foci," since they lie roughly oriented with the concentric lamellae surrounding the Haversian canals. The foci are surrounded by rims of bright (dense) reprecipitated bone 2–5μm wide (described by Hackett, 1981, as "cuffs"). At high magnifications it can be seen that within each larger focus there exist multiple small foci each just under 1μm in diameter (Fig. 3B). Lamellate foci are evidence that bone has undergone bacterial attack (Jackes, 1990).

Material from the earliest site (Moita do Sebastião) has been cleared with hydrogen peroxide, thus allowing observation in light microscopes. Of 31 sections, most show complete destruction of the morphology; only six retain any microstructure, and only in the midcortical third. The four femora that retain the most structure still have a great deal of lamellar bone and the least endosteal porosity; the mean number of empty spaces in the endosteal third of these sections range from 5.83 to 7.00. All other sections have a mean endosteal porosity count of 7.25 to 13.25 (Palmer, 1987: 150). Only these younger individuals allow examination of histomorphology and they could not provide age estimates by techniques that require examination of the periosteal third of a section.

Even when morphology has been partially maintained, the sections cannot be read. Cement lines do not show up clearly under transmitted, polarized, or fluorescent light and attempts at staining cement lines have been unsuccessful. Reduced birefringence, particularly of the outer lamellae of osteons, ensures that osteon area and osteon fragments cannot be observed accurately. In many sections every space is filled with calcite (see Fig. 3A) derived from the burial environment (Pate and Hutton, 1988). The calcite makes it impossible to examine the canal edges to see whether they are forming or resorbing. Calcite commonly fills cracks (Fig. 3A) that emanate from the periosteal border, thus precluding periosteal osteon counts. Attempts to remove the calcite with triammonium citrate result in almost complete destruction of the bone within a few hours.

It is no doubt valuable to continue research into techniques of preparing bone samples and observing microstructure. The extreme difficulty of preparing thin sections of archaeological bone has suggested the use of several alternatives, for example, reflected light microscopy (Bennike, 1990) and microradiography (Boivin and Baud, 1984; Garland, 1989; Stout and Simmons, 1979).

The reduced birefringence of osteons in archaeological specimens makes it impossible to distinguish among the various types of osteons visible under polarized light. Since the osteon types may have some association with age, this is particularly unfortunate for paleodemographers. But osteon type seems also to have a relationship with mineralization (Martin and Burr, 1989: 71), and degrees of mineralization are picked up by microradiography. Older osteons will be more mineralized; younger individuals with a higher rate of remodeling will have less mineralized osteons and older individuals with slower resorption and refilling rates should have more mineralized osteons. Detailed work on osteon types may help to answer Lazenby's major criticism, that by considering all osteons within a certain area, we

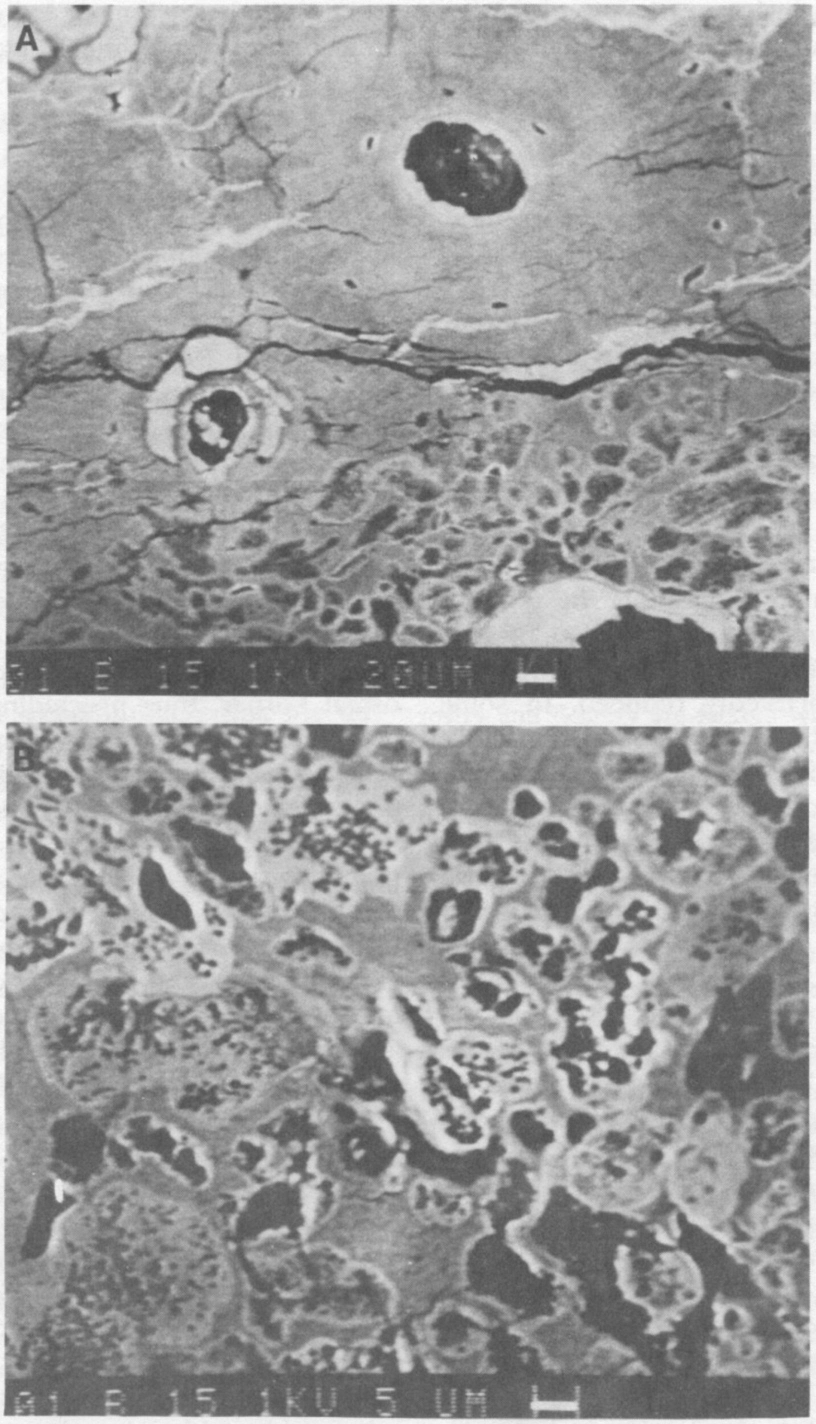


Fig. 3. **A:** Backscattered electron image of an undecalcified embedded endosteal midfemoral section from Cabeço da Arruda F, showing dessication cracks, calcite in spaces, and focal destruction in the lower third of the picture. ×210. **B:** Backscattered electron image of the midendosteal area of a femur from Caldeirão (P11/155), undecalcified and embedded, illustrating lamellate foci. 1,000×.

are including osteons formed years before, with the result that "our mean value will be more reflective of the younger age than of the true age at death" (Lazenby, 1984: 99).

Bone Density, Porosity, and Bone Mineral Content

Porosity of bone increases with age: there are more Haversian canals (there are more osteons per square millimeter and the osteons have a smaller area); the Haversian canal diameters are larger; there are more resorption spaces and incompletely refilled osteons, since the resorptive and refilling phases are slower in the elderly (Martin and Burr, 1989). But, while the volume of bone (excluding spaces) may decrease with age, the density and mineralization of bone may increase.

In archaeological bone, studies of porosity, density, and mineralization are almost impossible. Porosity has been studied in the Portuguese sample (Palmer, 1987) by fluorescing microscopy, but backscattered electron imaging demonstrates that not all calcite filled spaces can be visualized under fluorescing light. Image analysis with polarized light is possible, but calcite lies over bone as well as within spaces (Fig. 3A). Microbial activity introduces many false spaces into the bone that are surrounded by reprecipitated bone of increased density. Bone density values may be falsified when calcite is present in bone spaces and cannot be removed from a bone sample without destroying it. The calculation of density, as weight/volume bone, is an element in the Thompson technique, but it cannot be calculated for a bone that has undergone decomposition and in which many of the voids and cracks are filled with extraneous minerals. Portuguese midfemoral cortex samples have a mean bone density of 1.51 ($\sigma = .38$ $n = 105$ adults only), a value so low that it is unlikely to be a reliable indicator of in vivo bone density.

Bone Chemistry

There are changes in bone chemistry with age; a decrease of calcium, phosphate, and collagen and an increase of carbonate (Beattie, 1982; Lengyel and Nemeskéri, 1965; Lengyel, 1978), but there is little possibility that age trends can be discerned in archaeological bone.

Calcium and phosphorus. The Ca/P molar ratio of human bone mineral is 1.67, although microprobe analyses can show differences in Ca/P ratios across a single osteon (Ortner and von Endt, 1971). The ratio may be maintained throughout life (Mellors et al., 1966), or may increase after age 30 (Ascenzi, 1983). Deviations from a Ca/P ratio of 1.67 in archaeological bone can never be taken to indicate changes in bone chemistry that have resulted from age. Ratios in buried bone can be altered markedly (Sillen, 1989; Acsádi and Nemeskéri, 1970: 136–137; Lengyel, 1978; Salomon and Haas, 1967) and ratios in 1μm^2 areas of bone reprecipitated by bacterial action are particularly variable (Jackes, 1990).

It appears that bone apatite retains its chemical composition after death, although crystals may enlarge during diagenesis (Schoeninger et al., 1989; Susini et al., 1987; Jackes, 1990). Despite this, analyses that are not pinpointed on a few microns of bone will certainly find highly altered bone chemistry, solubility, and crystallinity because of the occurrence of brushite (Piepenbrink, 1986; Susini et al.,1987) and calcite in dead bone, and because of carbonate substitution.

Carbonate. Carbonates may substitute for phosphates in bone apatite during life (LeGeros et al., 1967), leading to altered crystal forms and increased solubility: the bones of the elderly will withstand diagenesis less well than those of young individuals.

While carbonate levels in bone may be related to age at death, it seems unlikely that the *in vivo* carbonate content of bone mineral could be measured accurately. As discussed, removal of diagenically introduced carbonates with triammonium citrate in one test resulted in the total destruction of samples. Baud (1987) has summarized information on the minerals that are formed by the reprecipitating activities of bacteria. These minerals are primarily calcite (i.e., pure $CaCO_3$). Groundwater carbonates may be added to bone carbonate without the intervention of bacteria (Pate and Hutton, 1988). At the same time, leaching in the soil may lead to a loss of apatite carbonates (Susini et al., 1987).

Collagen. Collagen is markedly reduced from the time growth is complete. However, collagen is also reduced in buried bone. The dense bright rims of lamellate foci lack collagen (Jackes, 1990), presumably utilized by bacteria. The mean percentage collagen of 22 Portuguese ribs sampled for stable isotope analyses and radiocarbon dating is 5.4% (range 2%–10%), compared with the 18%–20% expected in modern bones. Collagen levels are reduced even in areas with low soil temperatures (Nelson et al., 1986: 1945), and a loss of at least 60% collagen seems general (Lengyel and Nemeskéri, 1965; Lengyel, 1978; Pfeiffer, 1991).

Discussion. Bone decomposition depends on the burial microenvironment, the temperature, the humidity, and the pH of the deposits. Bone decomposition, as discussed above, also depends on the age at death of the individual. In spite of these considerations, Lengyel (1968) assumes that all femora from one level in a site will have equivalent bone decomposition, and that one can therefore reconstruct population characteristics and interpret deviations as individual differences, not diagenetic alterations. Pfeiffer (1991), on the other hand, has demonstrated the surprising variations in preservation of same-sex, same-age individuals, buried at the same time within one small and homogeneous area.

Lengyel has attempted chemical age assessment of 90 Mesolithic individuals from Vlasac (Nemeskéri, 1978). The remains were seriated on the basis of reduced phosphorus and increased carbonate and the ages were estimated by a regression formula that takes into account the difference between the Vlasac mean values and those of a sample of 700 autopsied individuals. Chemical ages were compared with morphological ages determined by Nemeskéri with the following results: in 54% ages were within the range of the morphological age; in 29.5% ages were contradictory. The published results do not allow us to determine which ages provided discrepant results.

Analyses of the micromorphology, density and chemistry of archaeological bone must start with the assumption that diagenesis has occurred. The hope that histomorphology will give us accurate age estimates for complete samples from a large variety of archaeological populations is vain; paleohistology and paleochemistry will not give paleodemography a new lease on life.

AGE DISTRIBUTIONS BASED ON DENTITION

Dental Wear

Wear is age-dependent, although factors such as type of diet and methods of food preparation will control the rate of attrition. Furthermore, wear is caused not only by the mastication of food but by individual idiosyncrasies (bruxism rather than age may determine attrition rates among modern European and North American adults) and by the use of teeth as tools. The normal pattern and rate of wear may be altered by malocclusion and pathology.

Opinions vary as to the importance of dental wear in age determination. Blakely (1971: 44) stated that "continued reliance on dental attrition as an accessory means of age estimation is a questionable practice," in contrast to the view that wear is "the best single indicator for determining age at death" (Lovejoy et al., 1985a: 12). Some of the many schemas upon which the observation of wear have been based are reviewed in Costa (1986), Hillson (1986), and Brothwell (1989). Seriation by dental wear (e.g., Lovejoy, 1985) can be followed by analysis of wear gradients (see e.g., Miles, 1963, 1978) and estimates of chronological age. Other researchers have developed quite different techniques (Molnar et al., 1983a,b; McKee and Molnar, 1988), which demonstrate, among other things, that different techniques can lead to different conclusions, for example with regard to sex differences in rates and patterns of wear.

Kieser and colleagues (1985) have shown that wear follows a linear progression up to 39 years of age, and then slows down (see also Hillson, 1986: 185). The fairly high correlations with age discerned by Lovejoy and colleagues (1985a: 9, $r = .7$) in studies of samples from the Todd Collection may result from the age distributions of those samples (Fig. 2). Molleson and Cohen (1990) demonstrated that attrition stages do not represent a series through which all dentitions pass in ordered and steady sequence. In fact, they suggest that the analysis of attrition stages produces underaging in archaeological samples. Testing of the Brothwell system showed that, while early stages are passed through rapidly, the later stages last for many years of an individual's life.

Lubell and colleagues (1989) and Lubell and Jackes (1988) have modified the dental wear grading categories of Smith (1984); the modified wear stages allowed consistent coding among three researchers. Nevertheless, the results of separate analyses of the Portuguese Mesolithic lower molar wear stages are discouraging. Attrition levels coded within a month of the time when the original schema was drawn up show no significant differences between observers. Observations made 1 year and 2 years later give a clear indication that observers altered criteria over time, yet the methods and detailed written criteria for scoring remained the same.

Other tests of interobserver and intraobserver error have been undertaken on large samples of loose upper and lower molars from the Portuguese Neolithic ossuary cave site of Casa da Moura. On the basis of random shifts of significance and nonsignificance among sets of teeth (from one molar to another, and from side to side), it appears that intraobserver error is as important as interobserver error. Even in higher wear categories in which distinctions were made on what seemed to be "objective" criteria (pattern of dentin exposures) agreement reached only around 50%.

All upper molars with no more than a pinpoint of dentin have now been further observed under a microscope. This study showed that even in the early stages of wear, when assessments are based on fairly objective categorizations (presence and absence of scratching, polishing, facetting, blunting, pinpoint dentin exposures), surprisingly little agreement is reached.

Other Dental Observations and Age Estimation

Attempts are under way to supplement wear scores with metrical data to reduce the subjectivity of wear assessment (Jackes, 1988). Kieser and colleagues (1985) showed the relation of crown height, occlusal slope, and interproximal facet width to age in a known age population. Measurements of cusp height are also possible (Tomenchuk and Mayhall, 1979; Molnar et al., 1983a), but cusp height is rapidly reduced in archaeological populations, and various crown height measurements have proven adequate in recent analyses.

Cementoenamel junction height above the alveolar margin (CEJ height). Tal (1984) analyzed CEJ heights on 100 mandibles of known age and found that the mean differences between the decades from those in their twenties to those in their forties were significant. Later age groups did not show significant differences, perhaps because of tooth loss.

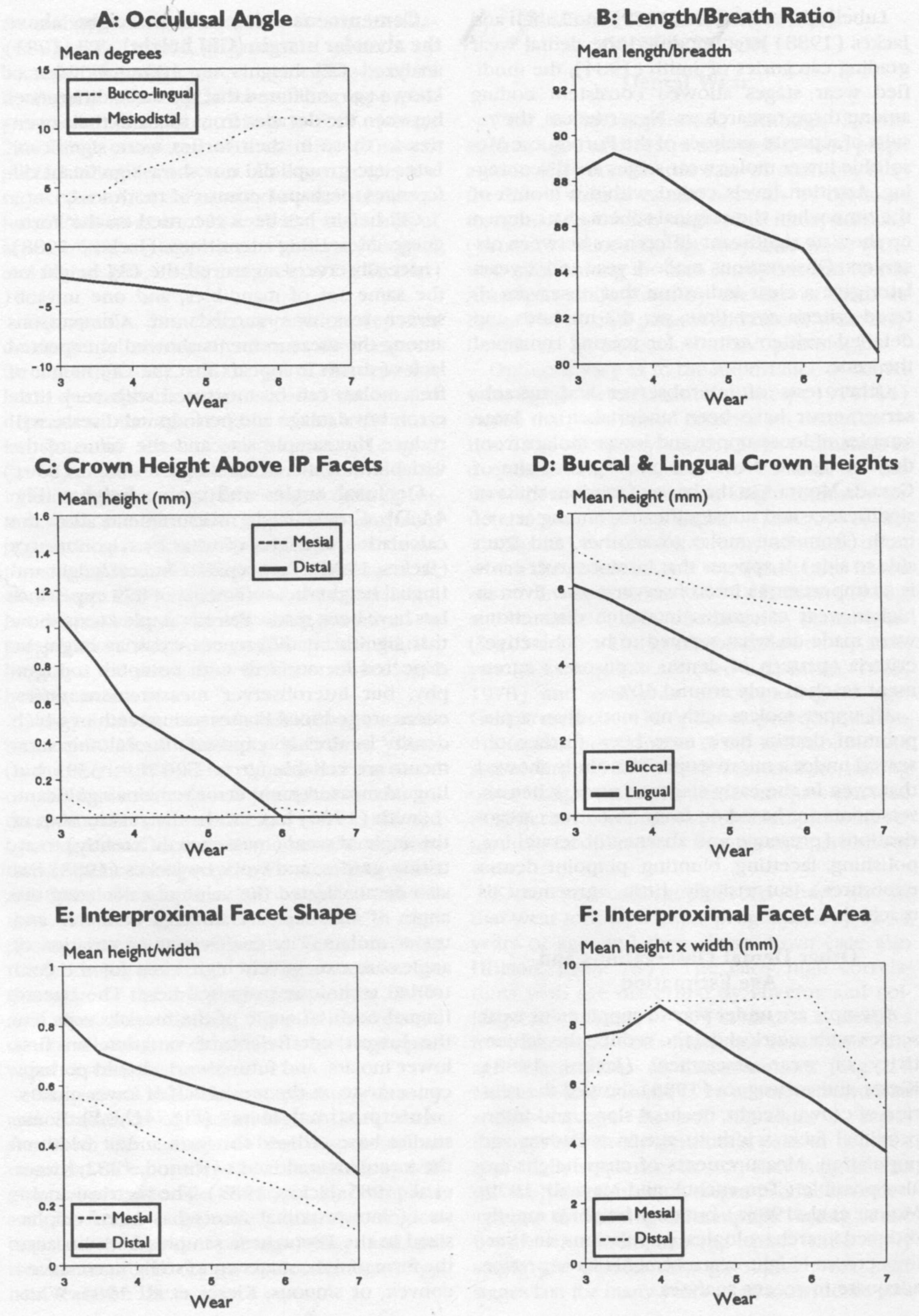
CEJ height has been recorded on the Portuguese Mesolithic dentitions (Jackes, 1988). Three observers measured the CEJ height on the same set of mandibles, and one intraobserver test was carried out. Comparisons among the measurements showed unexpected lack of error. It appears that the CEJ height of first molars can be measured with very little error; but damage and periodontal disease will reduce the sample size and the value of the variable in some archaeological samples.

Occlusal angles and crown heights (Fig. 4A,D). Crown height measurements allow the calculation of angles of wear by trigonometry (Jackes, 1988). Two separate buccal height and lingual height measurements of 879 upper molars have been made. Paired sample *t* tests show that significant differences exist, as might be expected for surfaces with complex topography, but interobserver measurement differences are reduced in measuring teeth in which dentin is already exposed. Buccal measurements are reliable ($p = .806$ $n = 339$), but lingual measurement error remains significant.

Smith (1984) has shown the relationship of the angle of wear (measured differently) to attrition grades, and work by Jackes (1988) has also demonstrated the value of calculating the angle of the occlusal surfaces of lower and upper molars. The coefficients of variation of angles are excessively high, even for the controlled technique proposed here. The buccolingual occlusal angle of the mesial cusps has the lowest coefficient of variation on first lower molars, and future work should perhaps concentrate on the mesial half of lower molars.

Interproximal facets (Fig. 4C,E,F). Some studies have utilized the form and/or width of the interproximal facets (Hinton, 1982; Kieser et al., 1985; Jackes, 1988). The metrical analysis of interproximal facets has been emphasized in the Portuguese samples both because the form of the contact area (straight, concave–convex, or sinuous; Kieser et al., 1985; Whit-

Occlusal


A: Occulusal Angle
Mean degrees
15
10
5
0
-5
-10
---- Bucco-lingual
Mesiodistal
3
4
5
6
7
Wear
B: Length/Breath Ratio
Mean length/breadth
92
90
88
86
84
82
80
3
4
5
6
7
Wear
C: Crown Height Above IP Facets
Mean height (mm)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
---- Mesial
Distal
3
4
5
6
7
Wear
D: Buccal & Lingual Crown Heights
Mean height (mm)
8
6
4
2
0
---- Buccal
Lingual
3
4
5
6
7
Wear
E: Interproximal Facet Shape
Mean height/width
1
0.8
0.6
0.4
0.2
0
---- Mesial
Distal
3
4
5
6
7
Wear
F: Interproximal Facet Area
Mean height x width (mm)
10
8
6
4
2
0
Mesial
---- Distal
3
4
5
6
7
Wear

taker et al., 1987) may not be an error-free observation and because work on loose molars allows more detailed analysis. The height of the facets is recorded as well as their width and the midpoint depth of the facet below the occlusal margin. In the expectation that facet width on upper molars would be inaccurate, two sets of measurements were made by the same individual, but the results of paired *t* tests were nonsignificant (mesial width, $p = .14$; distal width, $p = .47$; $n = 879$). A test on the depth of facets below the occlusal surface, using 277 first lower molars, gave equivalent results.

Length/breadth ratio and crown area (Fig. 4B). Reduction in the mesiodistal diameter of teeth is a result of interproximal wear (demonstrated for the Casa da Moura loose lower molars; Jackes, 1988). Upper molars show the same pattern (see also Mayhall, this volume, Chapter 4, for a discussion of dental measurement).

Tests on *In Situ* and Loose First Molars

A number of metrical variables can be used with or without wear scores to assign molars to categories. Of prime importance is the demonstration that variables can be grouped and that intercorrelations between variable groups are low. The lower molar variable groups generally sort out to express (1) the size and shape of the mesial facet; (2) the size and shape of the distal facet; (3) the buccolingual angles; (4) the length–breadth ratio of the crown or the crown area; (5) most crown and cusp heights and the height of the CEJ above the alveolar margin; and (6) the mesiodistal occlusal angles and the mesial crown heights. Variables from various groups can be combined in multivariate analyses to test the wear categories.

Subjectively, it appeared that the wear levels 2 and 3 (Smith, 1984) as used in the Portuguese project were inadequate for the range of wear at those levels, and this was confirmed objectively by discriminant analysis using the metrical variables defined above. It was demonstrated that wear level 3 performed unsatisfactorily; discriminant analysis indicated a low level of correct predictions of membership in wear level 3. Wear levels 2.5 and 3.5 were therefore added to the wear scores.

By using the means of three trial scorings of loose first lower molars as the grouping variable, 65% correct classification by discriminant analysis was achieved; the result was an improvement over the use of separate trial scores. In this test, 99% of the discrimination was achieved by the first two functions. The first expressed the mean of the crown heights at the four main molar cusps, the second the area of the mesial interproximal facet, the third the mesiodistal occlusal angle, and the fourth the buccolingual occlusal angle; the ratio of crown length to crown breadth weighed heavily on the fifth canonical discriminant function. Such analyses show that the wear of enamel and exposure of dentin form but one aspect of dental changes with increased age. Since, furthermore, wear is subject to observer error and the choice of factors defining a certain wear level is arbitrary, dental wear is not an ideal variable. Less subtle problems also arise: dental trauma and cupped wear (commonly over a millimeter deep in Neolithic lower molars) may lead to overestimation of attrition levels, and pathology may introduce variation.

For all these reasons, observations of dental wear should be balanced by the addition of dental measurements. When analyzed by wear categories, the coefficients of variation for many of these measurements are excessively high for biological data. Human dental wear may be so variable that measurements have little discriminating value, but it is more likely that the wear categories used to group the measurements are too coarse and inexact.

Cluster analysis (rather than analysis relying only on wear levels) can provide clear groupings of upper molars (Fig. 5) and leads to re-

Fig. 4. Upper first molar metrical changes plotted against a set of attrition scores. Sample ($n = 113$) consists of molars with both interproximal facets and dentin exposure. **A:** Buccolingual and mesiodistal occlusal slope. **B:** Ratio of length to breadth of the molar crown. **C:** Distance between the top of the facet and the occlusal surface of the tooth changes as the tooth wears down. **D:** Buccal and lingual crown heights. **E:** Shape of the facet alters as the facet height is lost and breadth increased. **F:** Interproximal facet area increases and then decreases as crown height is lost.

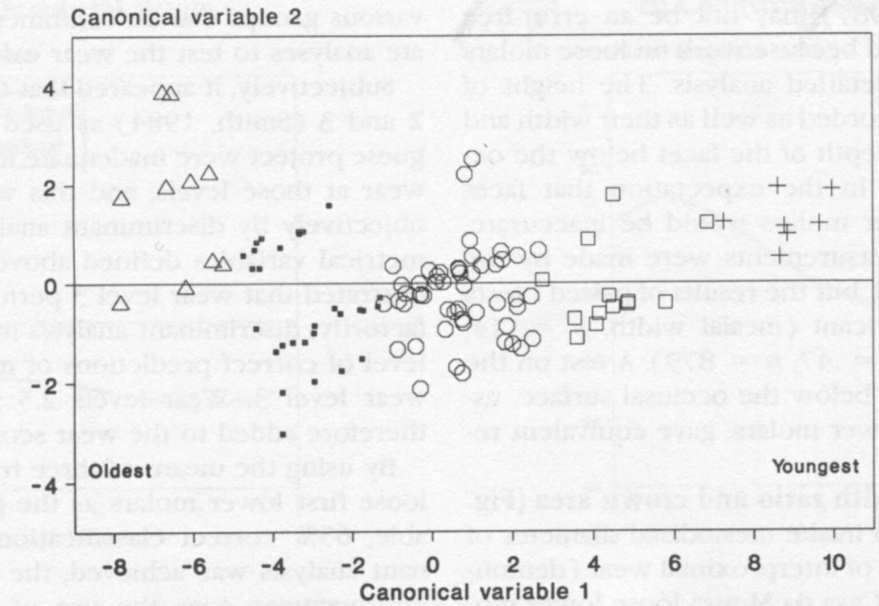


Fig. 5. Clustering of Casa da Moura upper first molars in which dentin is exposed, based on attrition score and three metrical variables (buccolingual angle, crown area, distal facet shape). Symbols indicate different clusters.

duced coefficients of variation for most of the metrical variables. However, the coefficients of variation for the angles and the facet variables remain beyond levels normally acceptable for biological data, and care should be taken that discrimination does not rely too heavily on these variables.

Sorting Teeth Into Age Categories

First, the sequence of crown and root calcification is identified (Moorrees et al., 1963; Anderson et al., 1976; Trodden, 1982) in conjunction with the early stages of occlusal wear and the appearance of interproximal attrition. Root formation scores are not error-free: tests on upper molars have shown marked differences between observers. Agreements reached only 47.6% ($n = 84$ incomplete roots, excluding cases considered broken in either or both observation runs). Agreement was highest for crown complete (59%) and initial root formation (100%), since these are the least subjective; but unequal sample sizes preclude valid analysis of each formation score for error.

In teeth with complete roots, the progress of occlusal wear and changes in the form of the interproximal facets up to the wear stage at which dentin is exposed are noted. Next, changes in the crown dimensions and occlusal angles, together with alterations in the interproximal facets, are identified. This permits the recognition of five levels of wear in which there is more than minimal dentin exposure.

Summary on Dentition

Dental wear alone will not provide a high correlation with real age (Kieser et al., 1985; Kay and Cant, 1988). Wear stages are arbitrarily defined and subject to inter- and intraobserver error. Wear is not a linear progression or, rather, dental anthropologists have not yet defined a series of wear stages that allows a linear progression with age to be identified on the basis of a large known-age, traditional diet sample. The Miles method has been tested on the known-age Lengua sample (Kieser et al., 1983), and although the correlation between real age and estimated age was high for mandibular samples (low for maxillary teeth), age estimates were less accurate in older adults (after age 27).

Detailed studies of enamel wear and dentin exposure, controlled by reference to cemental annulation and combined- method age estimates of individual teeth, and also incorporating analyses of metrical variables, may make age estimates more accurate. Grouping dentitions on the basis of multivariate analyses may

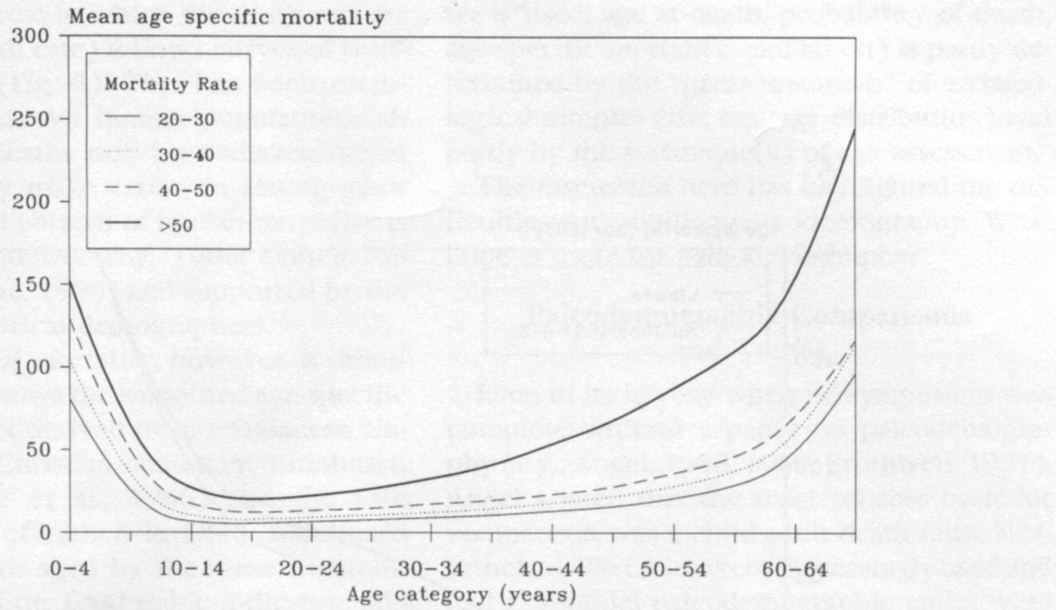


Fig. 6. Age-specific mortality curves for Finland, A.D. 1751 to 1868 based on data in Turpeinen (1979). The average age-specific mortalities are plotted for those years in which the mortality rates fell between 20 and 30, 30 and 40, 40 and 50, and 50 to 77.5.

well assist in assigning broad age categories that would be of value in discerning trends in rates of attrition and dental pathology. They may provide age categories for the study of cortical width, porosity, and degenerative changes. But the age estimates will not be accurate enough for full paleodemographic studies, since attrition is so strongly variable beyond young adulthood and paleodemographic analyses ideally require that ages at death be accurately ascribed to set 5-year age categories.

Furthermore, in populations with extreme dental pathology (e.g., North American maize agriculturalists), the dentition cannot be used to estimate adult age. Premortem tooth loss and high caries rates will make it impossible to observe many teeth, and the wear on the remaining teeth will not fall into the standard pattern. Similarly, in populations with extreme wear and/or extensive use of the teeth for functions other than eating, teeth will be worn too rapidly to provide the possibility of age estimations for older adults.

PALEODEMOGRAPHY

Paleodemography has been plagued by controversy over the last decade. The basis of the controversy is a very simple question: can anthropologists give accurate ages to the dead? Those who defend paleodemography state that the demographic curves generated by their data are meaningful (Van Gerven and Armelagos, 1983; Greene et al., 1986; Piontek and Weber, 1986). Buikstra and Konigsberg (1985) acknowledge that older adults may be underaged, but maintain that adult *q* values (probability of death curves) carry demographic information. The basis of this statement is a principal components analysis with unrotated component scores, specifically the second principal component, which explains only 18.5% of the total variance and to which q_{15} is the major contributor, that is, the probability of dying between 15 and 20 years of age.

Those who question the value of paleodemography state that the archaeological demographic parameters fall outside known values or curves and thus must be determined by inaccuracies in age assessment and biases in the data (Bocquet-Appel and Masset, 1982, 1985; Bocquet-Appel, 1986; Masset and Parzysz, 1985). Comparison of Figures 6 and 7 illustrates the point.

The age-specific mortality curves for Finland between 1751 and 1868 (Turpeinen, 1979) illustrate that human mortality (even in periods

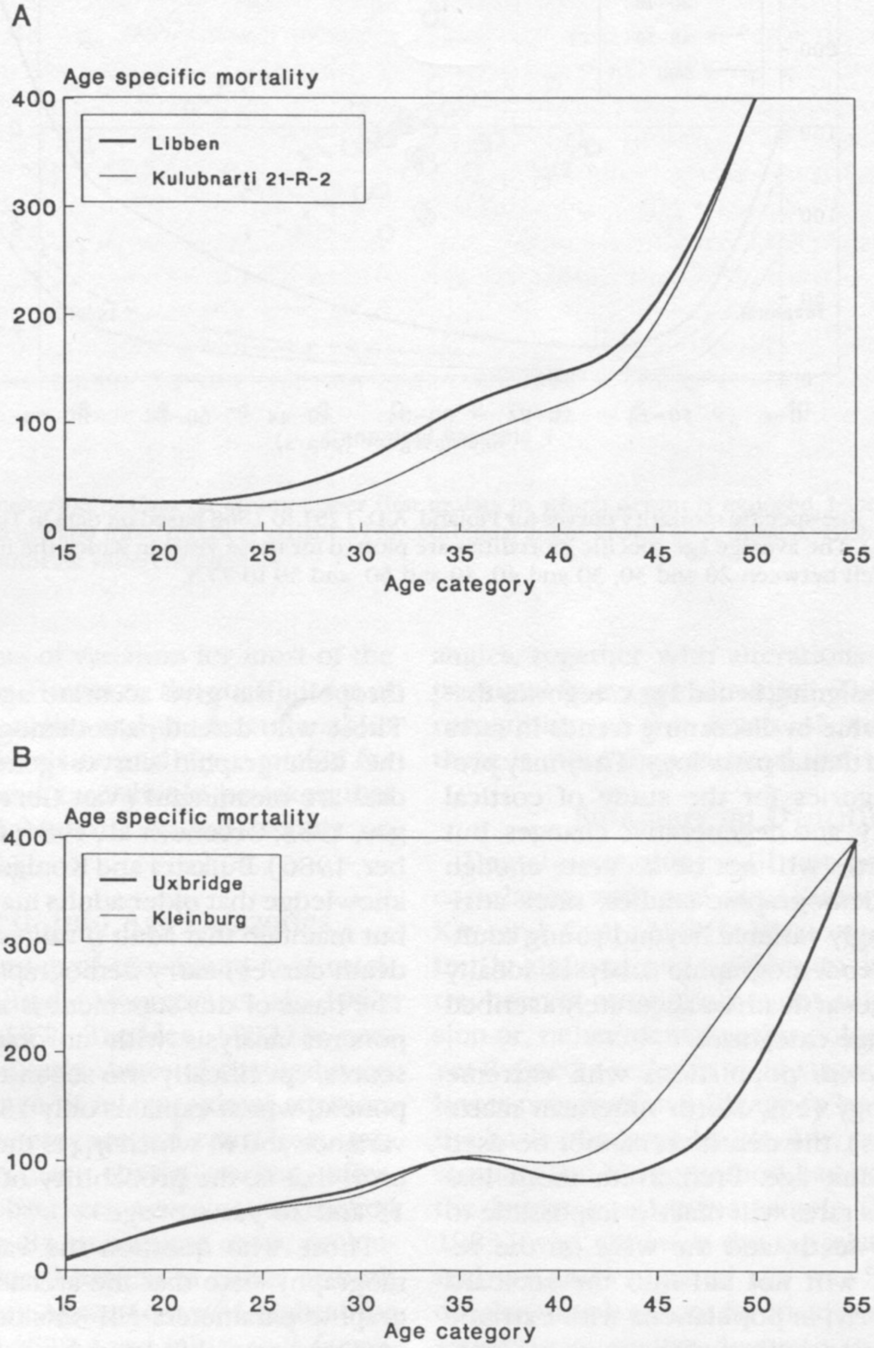


Fig. 7. Comparison of age-specific mortality curves. A: Libben and Kulubnarti aged by Todd pubic indicators smoothed on the basis of dental attrition scores. B: Uxbridge and Kleinburg aged by the McKern and Stewart method for males and the Gilbert and McKern method for females, and smoothed by equal distribution.

of extreme distress in which the death rate far exceeds the birth rate) follows curves of well-defined shape (Fig. 6). This has been established for all known human populations: although infant deaths may be underregistered and people may make errors in stating adult ages, the overall pattern of human mortality is clear (Coale and Demeny, 1966; United Nations, 1982; Gage, 1990) and supported by the findings of historical demographers.

Archaeological mortality, however, is different. Figure 7A shows the smoothed age-specific mortality curves derived from a Sudanese Nubian medieval Christian cemetery, Kulubnarti 21-R-2 (Greene et al., 1986) and the Late Woodland site of Libben in Ohio. Kulubnarti and Libben were aged by the same methods: major emphasis on Todd pubic indicators and dental attrition (the use of the Todd–Brooks method is confirmed by Greene et al., 1986: 197; Libben techniques are clearly stated in Lovejoy et al., 1977: 292 and note 10, and further outlined in later publications by the same team of researchers showing that the pubis and the dentition were most highly correlated with each other and with the assessed ages). The two curves are identical ($G = 8.4$, $p = .3$, 7 *df*). Figure 7B shows the age-at-death distributions derived from two precontact Ontario Huron ossuaries (see Jackes, 1986, for details). Both were given ages based on equal distribution of McKern–Stewart ages for males and Gilbert–McKern ages for females. Since Uxbridge showed a much higher level of skeletal tuberculosis (Pfeiffer, 1984) than the later site, Kleinburg (Jackes, 1977; Jackes, nd), identical age-specific mortality curves in young adulthood are best explained by identical age estimation techniques.

The Kleinburg curve is only one of a number of possible curves. Different age-at-death distributions will be generated by different treatments of the same 100 pubes (see Jackes, 1985). Different curves again are based on a sample of 112 femoral cores (Pfeiffer, 1984, since the Kleinburg site is an ossuary it is impossible to compare directly between osteonal and pubic ages or to use multifactorial age techniques). There is no basis for assuming that any curve best represents the true Kleinburg mortality. It is only known that the shape of the curve (whatever demographic parameter is used: age at death, probability of death, age-specific mortality, and so on) is partly determined by the "great unknown" of archaeological samples (the real age distribution) and partly by the technique(s) of age assessment.

The discussion here has highlighted the difficulties surrounding paleodemography. What hope is there for paleodemography?

Paleodemographic Comparisons and Trends

Even in its heyday when no symposium was complete without a paper on paleodemography (e.g., Angel, 1968, 1969; Brothwell, 1971), Angel argued that the most reliable basis for comparison was a child:adult death ratio. Nevertheless, life tables were increasingly used and in 1973 model paleodemographic tables were published (Weiss, 1973), which were intended to expand fragmentary data. They were used by Lovejoy and colleagues (1977) to calculate age-specific birth rates based on the Libben life expectancy and survivorship at age 15 (although the accuracy of survivorship depends on the completeness of the sample up to age 15). The Weiss tables are still used today (e.g., Corruccini et al., 1989).

There is little correspondence between the Weiss tables and other model or historical data. The lack of correspondence results from the way in which the 72 tables were constructed; the probability-of-death values fluctuate slightly according to a fixed pattern between 0 and 14 years, but not thereafter, in each of the nine tables per set of life expectancies at 15 years. The Weiss tables assume low mortality in infants and young children and high adolescent and early adult mortality—often characteristics of biased or misaged samples. The Weiss tables reflect this pattern simply because they derive most importantly from archaeological data. Predictive values calculated on the basis of some Weiss tables lie beyond the range of valid data, as determined by methods to be discussed below.

Anthropologists have also attempted to show trends by calculating either "the crude mortality rate" or "the crude birth rate": each of these values is actually 1 divided by the average age at death ($1/e_0$). The average age at death can be calculated through the use of a life table (as life expectancy at birth), or it can be derived by multiplying the number of dead in each age category (e.g., 5.0–9.9 years) by the midpoint of the age category (i.e., 7.5) and then summing the resulting figures for all age categories. The crude mortality or death rate, $1/e_o$ $1e_0$, has been cited for comparative purposes (Bass et al., 1971; Hassan, 1981; Owsley and Bass, 1979; Ubelaker, 1974; Weiss, 1973).

As the crude birth rate, $1/e_0$ has been used to show trends through time (Sattenspiel and Harpending, 1983). However, the high mortality level model life tables constructed by Coale and Demeny (1966) show that $1/e_0$ will provide an accurate crude birth rate only in populations that are either near zero growth or in decline. In populations with growth, $1/e_0$ will be higher than the birth rate.

Major Biases in Paleodemographic Data

Unfortunately, $1/e_0$ cannot be used for comparative purposes, for it has a number of inbuilt biases when derived from paleodemographic samples. The first bias is widely recognized: in many archaeological samples children, and especially infants (0–12 months), are not fully represented (for a further discussion see Saunders, this volume, Chapter 2). For North America, judging from the archaeological samples, it seems probable that Arikara villagers of South Dakota buried all infants or that most infants have been recovered during excavation (Bass et al., 1971; Owsley and Bass, 1979). At the opposite end of the pole, Huron ossuaries in southern Ontario provide fair certainty of complete excavation of all individuals within the sharply delimited burial area, but very few infants seem to have been buried in the general community burial area (Thwaites, 1898, Vol. 10: 273; Saunders and Spence, 1986). It is true that "changes in infant representation affect only infant life table values and survivorship values; life expectancy . . . [is] mathematically unaffected" (Moore et al., 1975: 60). However, life expectancy at birth (the demographic parameter on which comparison is usually based) is strongly affected.

The second bias results from the fact that average age at death (e_0) depends on the representation in older age categories. Assignment of skeletons to these categories in an age distribution has been accomplished by using running averages or by extending curves to some chosen final age. This is no more than a choice, one that is based on the preconceptions of the anthropologist. Van Gerven and Armelagos (1983), while contending that skeletal biologists can accurately estimate age from human archaeological bone, state that it is impossible to assign correct ages to those over 55 or so. By this view, which is widely accepted by those who use Todd or McKern and Stewart pubic indicators, there is no way of knowing whether the oldest individual in a group was 65 or 85. Opinions on the longevity of archaeological populations determine the final age category chosen. Since the choice of final age is arbitrary, Howell (1973) is perfectly correct in accusing paleodemographers of employing "a rubber yardstick." Life tables that end at 50, and in which the final age calculations are made on the assumption that all people die by 55 have reduced mean age-at-death values.

A third source of error in the calculation of the average age at death relates to the overall accuracy of estimating age from skeletons of individuals over the age of about 25, as discussed above. It is worth reemphasizing, since even recent literature (e.g., Paine, 1989) ignores the possibility of adult underaging and assumes that deviation from model tables is a result of cultural or postdepositional processes.

A fourth source of error can be identified in the calculation of e_0 and thus of $1/e_0$. Not all adult skeletons in an archaeological population can be given an estimated age. Beyond the statement that there are so many indeterminate adults older than age 25 or so, there is nothing to be done about estimating the age of a number of adults. Although there has been little work on this subject, much of the evidence covered in this chapter indicates that the older individuals in a sample are most likely to be given indeterminate age. Exclusion of these individuals from calculation of the average age at death will naturally depress the average age at death, giving the same result as does underrepresentation of adults. Such individuals have often been excluded from life table calculations. Equal distribution over all age categories, and, better yet, cumulative graphing, which is no doubt the best method of redistribution, are possible solutions. Distribution over adult age groups proportionate to the percentage membership of each age category (Biraben, 1969;

Asch, 1976) simply compounds age estimation errors.

The literature has emphasized the problem of infant underrepresentation, but in fact adult underrepresentation is far more important. Adults contribute more to the average age at death than do young children. It would require a ${}_5D_0$ value of 360 (that is, 360 children under 5 in a site) to equal the contribution of 20 adults in a "25–65" adult category, and if it were a "25–75" category, 400 individuals of 0–5 years would be necessary. This is despite the fact that early childhood deaths are here taken to be equally distributed between 0 and 4.9 years, as compensation for infant underrepresentation. In the same way, adult deaths are treated as though evenly distributed, to avoid the problem of left or right skewing of adult age-at-death distributions.

These four sources of error all have marked effects on e_0 and thus militate against the use of $1/e_0$ for comparative studies. Nevertheless, comparison is vital to the study of paleodemography. For example, an obvious question is How were American indigenous mortality and fertility rates altered by contact with Europeans and with the infectious diseases of the Old World? If this question cannot be answered on the basis of methods presently used in paleodemography, then one must either admit paleodemography to be a futile exercise (Bocquet-Appel and Masset, 1982), or work to change the methods.

The recent emphasis on the value of fertility rates (Sattenspiel and Harpending, 1983; Buikstra et al., 1986) must not be misconstrued to suggest that sites should be compared by using life table fertility rates (general or total fertility, or child:woman ratios). Sites at which infant underrepresentation and adult age estimate errors inflate the age classes between age 15 and age 45 (these are the age categories that determine life table fertility rates) result in erroneous fertility rates. It is occasionally possible to demonstrate the errors by comparison with ethnohistorical evidence. Jackes (1986) showed that Ossossané, a postcontact Huron site with a life table total fertility rate of 5.6 (total fertility = general fertility × 30), contradicted the ethnohistorical evidence of fertility rates below 4.4 (see also Engelbrecht, 1987). Adjusting the life table for infant underrepresentation simply increases total fertility to well over 6.0. Life table values (apart from the q values) are too dependent upon infant representation to be useful. Fertility estimates are too dependent on correct age estimates of females aged 15–45 to be considered.

Circumventing the Problems

There have been several proposals that avoid the problem of the underrepresentation of those below age 5 and the inaccurate aging of adults. The first was that of Bocquet and Masset (1977), who suggested that estimators could overcome some paleodemographic problems, specifically the estimator 5–14.99/20+ (termed the Juvenile:Adult Ratio or JA; Jackes, 1986). Jackes (1986, 1988) proposed comparing pre- and postcontact mortality in North America and the transition to horticulture in Portugal by basing the analysis on childhood q values (mean childhood mortality or MCM, the mean of ${}_5q_5$, ${}_5q_{10}$, and ${}_5q_{15}$). Although the q values of a life table are markedly affected by the representation of adults (i.e., indeterminate adults must be included in one broad adult age category, say 25–65, if q values are to be analyzed), they are not in any way determined by infant representation. Mortality based on q values (mortality quotients) from age 1 to age 15 is equally valuable, but since young children as well as infants may be underrepresented in archaeological sites, ages 5–25 are preferable.

MCM is highly correlated with the JA because both depend on the relative representation of those under and those over 20 in a buried population. Their correlation is .97, based on 52 presumably unbiased archaeological samples with $n > 100$, or .95 based on 17 samples from the literature on historical demography (France, Switzerland, Quebec, Poland). Model life table correlations (Coale and Demeny, 1966) are .997 (the relationship is linear).

Buikstra and colleagues (1986) proposed that the statistic 30+/5+, published by Coale and Demeny (1966: 39), be used as an estimate of fertility. This value does have good predictive power but the question of the source of adult age variations arises. Konigsberg and coworkers (1989) have acknowledged that a demographic parameter based on assignment of

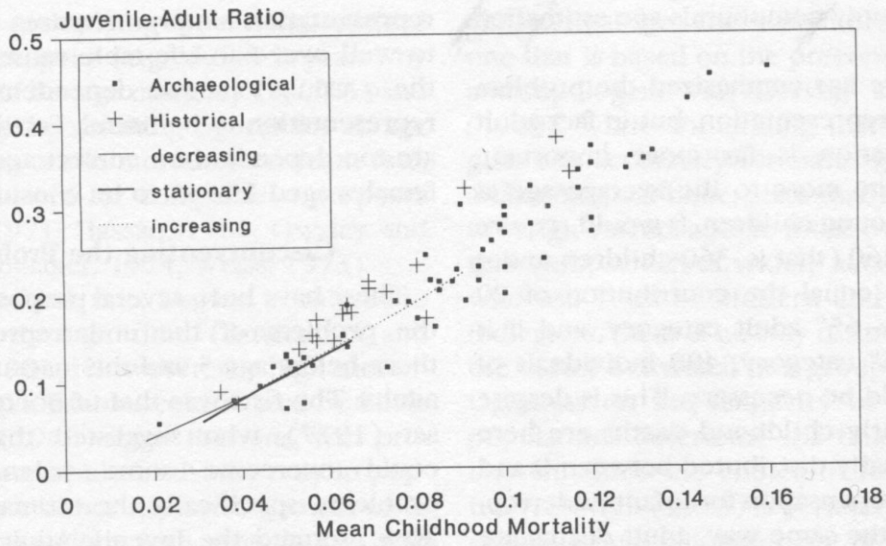


Fig. 8. Juvenile:Adult ratio plotted against mean childhood mortality for 60 archaeological sites compared with data drawn from 17 historical life tables. Trends shown are those based on Coale and Demeny (1966), West Model Life Tables 1–10, for populations that were decreasing, stationary, and increasing at low levels.

adult skeletal age to under 30 and over 30 categories is unsuitable for some sites and recommend use of 20+/5+ in the future. The correlation of MCM with log 20+/5+ is −.99986, the relationship being very slightly nonlinear.

The utility of the three proposed estimators is judged on the basis of 18 Model West tables from levels 1 to 10, for stationary, increasing
.01 (but only up to $r = .01$~~%~~), and declining populations (Coale and Demeny, 1966), calculated for pooled sexes ($r =$ the crude birth rate minus the crude death rate = the rate of natural increase). For stationary populations JA is the best predictor, correlated with log general fertility at .9999 (general fertility calculated from C_x values: infants/half the population aged 15–45 years). For increasing populations the best predictor is MCM, which is correlated with log general fertility at .9987. For declining populations the best predictor is log 20+/5+, which is correlated with log general fertility at −.9995. Overall, 20+/5+ and MCM will give equivalent results: $r = .9970$ with log general fertility and log total fertility.

The conjoint use of MCM and JA has been proposed (Jackes, 1986, 1988) because together they provide additional demographic information (Fig. 8). It can be seen which sites fall far from the line and might be considered biased in some way. For example, "high mortality sites" may actually be those in which there are more children than expected or fewer adults than expected. Sites that fall high on the line must be checked for possible sources of bias: (1) preferential burial of children, (2) incomplete excavation that has bypassed an area of adult burial, (3) exclusion of indeterminate adults from analysis. For example, the abnormal *q* curve presented by Kulubnarti 21-S-46 (Van Gerven et al., 1981) can be brought closer to other paleodemographic data simply by increasing the proportion of adults in the site. But this and other extreme outliers (beyond the range shown in Fig. 8: Kulubnarti 21-S-46, JA = 1.173 and Nea Nikomedia, JA = .683) could also be checked for evidence of extraordinarily rapid population increase. Analyses have concentrated on low levels of population increase, since extremely high rates of increase have been considered unlikely (Hassan, 1981).

High values may well represent populations that are expanding; included within the high value group of sites (JA above .3 and MCM above .12 on Fig. 8) are some that could represent increasing populations (e.g., Schild and Ledders; see Buikstra et al., 1986). Nevertheless, high values are also characteristic of sites that are known to have been in decline (Jackes, 1986).

It is possible to distinguish increasing from declining populations (and perhaps also biased populations) by considering not simply the level of childhood mortality, but the shape of mortality. For example, clustering of samples on the four q values from 5 to 20 indicates that sites of populations under stress (Larson, Grimsby: Jackes, 1983) may group separately from increasing populations. In other words, the shape of juvenile mortality quotient curves (*q* curves) can be considered, as well as the level of mortality, when one seeks to interpret paleodemographic evidence (Jackes, 1988).

Carrier's (1958) method of adjustment allows the calculation of life tables for nonstationary populations, and for this purpose standardized life tables have been proposed (Jackes, 1986; with figures published for *r*). The figures used give *r* values approximating those of Coale and Demeny (1966) very closely. The correction for nonstationary populations does not radically alter the shape of the mortality curves; it simply increases the calculated mortality quotients when $r < 0$ (population decline) and decreases the quotients when $r > 0$ (population increase). With an adjustment for population increase, the shape of the curve is altered only slightly, but the JA and MCM are reduced and a site like Ledders, which falls above the regression line, is brought to the line.

Ossossané, the Huron ossuary already referred to, provides good evidence for avoiding the assumption that predictions of fertility will be accurate once they have been adjusted for the rate of natural increase. The Ossossané life table (Katzenberg and White, 1979) total fertility rate of 5.6 has already been shown to be too high. If decline ($r = -.005$~~%~~) is adjusted for and total fertility predicted by regression based on 20+/5+, the result is an impossible 8.5. It can only be assumed that the ossuary does not contain all the dead. The addition of 150 adults to the life table would bring the JA and MCM into line, at high mortality in a declining population with a total fertility of 3.3. The picture would then accord with ethnohistorical evidence, including the possibility that not all adult males were buried in the ossuary.

-.005

It thus seems possible to determine whether a site is biased or nonstationary, and adjust accordingly in calculating mortality and fertility rates. Caution is necessary, however, since the usefulness of the estimators themselves may be limited. Masset and Parzysz (1985) have calculated the 33% confidence limits of JA (which they confusingly call "*r*") and have shown that, even in cemeteries containing 1,000 individuals aged 5–14 and over 19 years, fertility estimates cannot be considered accurate.

TABLE 2. Crude Birth Rate (CBR) Estimates for Barbados Slave Population

Basis for estimate[a]	CBR	SE
West 3, female, $r = 0$	40	—
Historical life table[b]	39.5	—
MCM (log CBR): Historical	39.49	0.007
JA (linear): Historical	50.37	1.11
JA (log CBR): Historical	51.88	0.007
20+/5+ (log CBR): Historical	40.37	0.006
JA derived from skeletons[c]	24.2	0.006

[a] JA—Juvenile:Adult ratio; MCM—Mean childhood mortality.
[b] Konigsberg et al. (1989).
[c] Corruccini et al. (1989).

Corruccini and colleagues (1989) have questioned fertility predictors of the type discussed above by comparison of Barbados slave historical records with skeletal evidence. The historical life table, as calculated by Konigsberg and colleagues (1989), gives a mean age at death and MCM that are in general accord with the pooled West level 3 at $r = 0$ age-at-death distribution (Table 2). The data provided by Corruccini and coworkers (1982, 1989) do not allow calculation of any values, so this second example of paired skeletal and ethnohistoric data cannot be analyzed. There must be full publication of ages at death and methods of age assessment to test bias in a skeletal sample. However, it is possible to demonstrate (Table 2) that estimators such as MCM can be completely accurate and that the Barbados skeletal series is probably incomplete and a poor choice upon which to base a trial of estimators.

SUMMARY

It has been shown that adult ages at death cannot yet be estimated with satisfactory accuracy from skeletal data.

Estimates based on morphological and radiographic methods do not appear to be reliable.

Chemical methods, although hardly tried, are probably too affected by diagenesis to be useful unless very uneconomic methods of point analysis are used. Histomorphology is possible only for material from sites with exceptional preservation.

On the other hand, it may be possible to produce higher correlations with real age by using a complex system based on dentition. Attrition, although shown to be variable and error-prone, can be made less subjective by the addition of metrical variables. In combination with histological techniques, teeth may provide age estimates where diagenesis, pathology, and trauma are limited.

Although the accuracy of age estimation may be improved, the degree of that accuracy will almost always be an unknown. Tests on modern samples cannot ensure accurate age assessment of past populations whose environment and life-style are imperfectly understood. Nevertheless, the minimum number of individuals count (MNI) can be based on teeth and extremely detailed analyses of sets of dentition of individuals aged 5–25 years can be undertaken. On the assumption that childhood age at death can be accurately estimated, childhood mortality and the proportion of children to adults in a cemetery sample can be used to assess the biases, to suggest whether the population is nonstationary, and to compare trends through time.

If the paleodemographic parameters discerned do not accord with model or historical data, the accuracy of the paleodemographic data should be questioned. One should not assume that the human experience of mortality and fertility has altered dramatically, from place to place and time to time (Howell, 1976; and see Roth, this volume, Chapter 10).

On the assumption that childhood mortality levels reveal something about a population, and that human mortality follows a discernible pattern irrespective of place and time, research in paleodemography can continue. But the limits of paleodemography are clear. If those limits are acknowledged, we can still trace human mortality and fertility through time.

ACKNOWLEDGMENTS

This research was done in association with Dr. David Lubell and was funded by the Social Sciences and Humanities Research Council of Canada (Grants 810-84-0030 and 410-86-2017) to Lubell, Jackes, and Dr. Christopher Meiklejohn (University of Winnipeg) and by the Central Research Fund, University of Alberta (to Lubell). I am grateful to Meiklejohn for sharing his data, and to Dr. Miguel Ramalho of the Geological Survey of Portugal, Lisbon, for allowing us to work on the skeletal collections under his care (Moita, Arruda, Casa da Moura, Melides, Furninha, and Fontainhas) and for facilitating our research. We thank our colleagues (Drs. Miguel Ramos, Henry Schwarcz, Gerd Weih, João Zilhão), students (Cidália Duarte, Kirsten Jackes, Pamela Mayne, Esther Palmer, Gary Tait), and technicians (Peter Black, Dennis Carmel, Malti Nikrad, José Severino Rodrigues, Paul Wagner). The long-term work of Carmel, Mayne, and Palmer has been especially important. Radiography was done in Portugal at the Santiago do Caçem Hospital, and by Paulo de Deus Machado at the Janares Clinic, Sintra.

REFERENCES

Acsádi G, Nemeskéri J (1970): "History of Human Life Span and Mortality." Budapest: Akademiai Kiadó.

Ahlqvist J, Damsten O (1969): A modification of Kerley's method for the microscopic determination of age in human bone. J Forensic Sci 14:205–212.

Amy R, Bhatnagar R, Damkjar E, Beattie O (1986): The last Franklin expedition: Report of a postmortem examination of a crew member. Can Med Assoc J 135:115–117.

Anderson DL, Thompson GW, Popovich F (1976): Age of attainment of mineralization stages of the permanent dentition. J Forensic Sci 21:191–200.

Angel JL (1968): Ecological aspects of palaeodemography. In Brothwell DR (ed): "The Skeletal Biology of Earlier Human Populations." London: Pergamon Press, pp 263–270.

Angel JL (1969): The bases of paleodemography. Am J Phys Anthropol 30:427–438.

Angel JL, Suchey JM, İşcan MY, Zimmerman MR (1986): Age at death from the skeleton and viscera. In Zimmerman MR, Angel JL (eds): "Dating and Age Determination in Biological Materials." London: Croom Helm, pp 179–220.

Ascenzi A (1983): Microscopic dissection and isolation of bone constituents. In Kunin AS, Simmons DJ (eds): "Skeletal Research," Vol 2. New York: Academic Press, pp 185–236.

Asch DL (1976): "The Middle Woodland Populations of the Lower Illinois Valley: A Study in Paleodemographic Methods." Northwestern University Archaeological Program, Scientific Papers No. 1.

Bang G, Ramm E (1970): Determination of age in humans from root dentine transparency. Acta Odontol Scand 28: 3–35.

Barnett E, Nordin BEC (1960): The radiological diagnosis of osteoporosis: A new approach. Clin Radiol 11:166–174.

Bass WM, Evans DR, Jantz RL (1971): "The Leavenworth Site Cemetery." University of Kansas Publications in Anthropology, No. 2.

Baud C-A (1987): Altérations osseuses post-mortem d'origine fongique ou bactérienne. In Duday H, Masset C (eds): "Anthropologie Physique et Archéologie: Méthodes d'étude des Sépultres." Paris: Centre National de la Recherche Scientifique, pp 135–146.

Beattie O (1982): An assessment of x-ray energy spectroscopy and bone trace element analysis for the determination of sex from fragmentary human skeletons. Can J Anthropol 2:205–215.

Bedford ME, Russell KF, Lovejoy CO (1989): The utility of the auricular surface aging technique (abstract). Am J Phys Anthropol 78:190–191.

Beeley JG, Lunt DA (1980): The nature of the biochemical changes in softened dentine from archaeological sites. J Arch Sci 7:371–377.

Bender SJ (1979): Paleodemographic analysis of a Late Woodland site in southwestern Michigan. Mid-Cont J Archaeol 4:183–208.

Bennett KA (1973): "The Indians of Point of Pines, Arizona." Anthropological Papers of the University of Arizona, No. 2.

Bennike P (1990): A simplified preparation of undecalcified bone for osteoanalyses (abstract). Populations of the Nordic Countries: Human Population Biology From the Present to the Mesolithic. Second Seminar of Nordic Physical Anthropology, Institute of Archaeology, University of Lund.

Bergot C, Bocquet JP (1976): Étude systématique de l'os spongieux et de l'os cortical du femur et de l'humerus en fonction de l'âge. Bull Mém Soc d'Anthropol (Paris): (Série XIII) 3:215–242.

Biraben JN (1969): Durée de la vie dans la population de Columnata (épipaléolithique oranais). Population 3:487–500.

Black JK III (1979): "The Biological and Social Analyses of a Mississippian Cemetery From Southeast Missouri." Anthropol Pap Mus Anthropol Univ Mich, No. 68.

Blakely RL (1971): Comparison of the mortality profiles of Archaic, Middle Woodland and Middle Mississippian populations. Am J Phys Anthropol 34:45–53.

Bocquet J-P (1977): Paléodémographie: Ce que nous apprend de la Nubie soudanaise. Annales: Économies, Sociétés, Civilisations 32:54–96.

Bocquet J-P (1978): Estimation methods of age at death in adult skeletons and demographic structure of the populations of the past. In Garralda MD, Grande RM (eds): "I Simposio de Antropologia Biologica de Espana." Madrid: Facultad de Biología Cuidad Universita, pp 37–47.

Bocquet J-P, Masset C (1977): Estimateurs en paléodémographie. L'Homme 7:65–70.

Bocquet J-P, Maia Neto MA, Tavares da Rocha MA, Xavier de Morais MH (1978): Estimation de l'âge au décès des squellettes d'adultes par régressions multiples. Contribuçoñes para o Estudo da Antropologia Portuguesa 10(3): 107–133.

Bocquet-Appel J-P (1986): Once upon a time: Palaeodemography. In Herrmann B (ed): "Innovative Trends in der prähistorischen Anthropologie." Mitt Berl Ges Anthropol Ethnol Rugesch 7:127–133.

Bocquet-Appel J-P, Masset C (1982): Farewell to palaeodemography. J Hum Evol 11:321–333.

Bocquet-Appel J-P, Masset C (1985): Palaeodemography: Resurrection or ghost. J Hum Evol 14:107–111.

Bocquet-Appel J-P, Tavares da Rocha MA, Xavier de Morais MH (1980): Peut-on estimer l'âge au décès à l'aide du remaniement osseux? Biom Hum (Paris) 15:51–56.

Boddington A, Garland AN, Janaway RC (1987): Flesh, bones, dust and society. In Boddington A, Garland AN, Janaway RC (eds): "Death, Decay and Reconstruction." Manchester: Manchester University Press, pp 3–9.

Boivin G, Baud CA (1984): Microradiographic methods for calcified tissues. In Dickson GR (eds): "Methods of Calcified Tissue Preparation." Amsterdam: Elsevier, pp 391–410.

Bouvier M, Ubelaker DH (1977): A comparison of two methods for the micrscopic determination of age at death. Am J Phys Anthropol 46:391–394.

Brooks ST (1955): Skeletal age at death: The reliability of cranial and pubic age indicators. Am J Phys Anthropol 13:567–598.

Brooks ST, Suchey JM (1990): Skeletal age determination based on the os pubis: A comparison of the Acsádi–Nemeskéri and Suchey–Brooks methods. Hum Evol 5:227–238.

Brothwell DR (1971): Palaeodemography. In Brass W (ed): "Biological Aspects of Demography." Society for the Study of Human Biology, Symposium No. 10, pp 111–130.

Brothwell D (1989): The relationship of tooth wear to aging. In İşcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 303–316.

Buikstra JE (1976): "Hopewell in the Lower Illinois Valley: A Regional Approach to the Study of Human Biological Variability and Prehistoric Behavior." Evanston, IL: Northwestern University Archaeological Program, Scientific Papers No. 2.

Buikstra J, Kongisberg LW (1985): Paleodemography: critiques and controversies. Am Anthropol 87:316–333.

Buikstra J, Konigsberg LW, Bullington J (1986): Fertility and the development of agriculture in the prehistoric midwest. Am Antiq 51:528–546.

Burns KR, Maples WR (1976): Estimation of age from individual adult teeth. J Forensic Sci 21:343–356.

Carrier, NH (1958): A note on the estimation of mortality and other population characteristics given death by age. Popul Stud (London) 12:149–163.

Charles DK, Condon K, Cheverud JM, Buikstra JE (1986): Cementum annulation and age determination in *Homo sapiens*. I. Tooth variability and observer error. Am J Phys Anthropol 71:311–320.

Charles DK, Condon K, Cheverud JM, Buikstra JE (1989): Estimating age at death from growth layer groups in cementum. In Işcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 277–301.

Churcher CS, Kenyon WA (1960): The Tabor Hill ossuaries: A study in Iroquios demography. Hum Biol 32:249–273.

Coale AJ, Demeny P (1966): "Regional Model Life Tables and Stable Populations." Princeton: Princeton University Press.

Condon K, Charles DK, Cheverud JM, Buikstra JE (1986): Cementum annulation and age determination in *Homo sapiens*. II. Estimates and accuracy. Am J Phys Anthropol 71:321–330.

Cook DC (1984): Adult age determination from the dentition. In Rathbun TA, Buikstra JE (eds): "Case Studies in Forensic Anthropology." Springfield, IL: Charles C Thomas, pp 307–328.

Cook M, Molto EL, Anderson C (1989): Fluorochrome labelling in Roman period skeletons from Dakhleh Oasis, Egypt. Am J Phys Anthropol 80:137–143.

Corruccini RS, Handler JS, Mutaw RJ, Lange FW (1982): Osteology of a slave burial population from Barbados, Indies. Am J Phys Anthropol 59:443–459.

Corruccini RS, Brandon EM, Handler JS (1989): Inferring fertility from relative mortality in historically controlled cemetery remains from Barbados. Am Antiq 54:609–614.

Costa RL (1986): Determination of age at death: Dentition analysis. In Zimmerman MR, Angel JL (eds): "Dating and Age Determination of Biological Materials." London: Croom Helm, pp 248–269.

Drusini A (1987): Refinements of two methods for the histomorphometric determination of age in human bone. Z Morphol Anthropol 77:167–176.

Engelbrecht W (1987): Factors maintaining low population density among the prehistoric New York Iroquois. Am Antiq 52:13–27.

Ericksen MF (1980): Patterns of microscopic bone remodeling in three aboriginal American populations. In Browman DL (ed): "Early Native Americans: Prehistoric Demography, Economy and Technology." The Hague: Mouton, pp 239–270.

Fraysse C, Kerebel B (1982): Aspects histologiques du vieillissement cementaire chez le chien. Bulletin du groupement international pour la recherche scientifique en stomatologie et odontologie 25:13–34.

Furseth R, Mjör IA (1973): Cementum. In Mjör IA, Pindborg JJ (eds): "Histology of the Human Tooth." Copenhagen: Munksgaard, pp 97–119.

Gage TB (1990): Variation and classification of human age patterns of mortality: Analysis using competing hazards models. Hum Biol 62:589–617.

Garland AN (1987): A histological study of archaeological bone decomposition. In Boddington A, Garland AN, Janaway RC (eds): "Death, Decay and Reconstruction." Manchester: Manchester University Press, pp 109–126.

Garland AN (1989): Microscopical analysis of fossil bone. Appl Geochem 4:215–229.

Garland AN, Freemont AJ, Stoddart RW, Jayson MIV (1987): Post-mortem histological changes and their implications for the study of exhumed human remains (abstract). Acta Anat 10:36.

Garland AN, Janaway RC, Roberts CA (1988): A study of the decay processes of human skeletal remains from the parish Church of Holy Trinity, Rothwell, Northamptonshire. Oxford J Archaeol 7:235–252.

Gilbert BM (1973): Misapplication to females of the standard for aging the male *os pubis*. Am J Phys Anthropol 38:39–40.

Gilbert BM, McKern TW (1973): A method for aging the female *os pubis*. Am J Phys Anthropol 38:31–38.

Gordon CC, Buikstra JE (1981): Soil pH, bone preservation, and sampling bias at mortuary sites. Am Antiq 46:566–571.

Graves WW (1922): Observations on age changes in the scapula. Am J Phys Anthropol 5:21–33.

Greene DL, Van Gerven DP, Armelagos GJ (1986): Life and death in ancient populations: Bones of contention in paleodemography. Hum Evol 1:193–207.

Gustafson G (1950): Age determinations of teeth. J Am Dent Assoc 41:45–54.

Hackett CJ (1981): Microscopical focal destruction (tunnels): In exhumed human bones. Med Sci Law 21:243–265.

Hanihara H, Suzuki T (1978): Estimation of age from the pubic symphysis by means of multiple regression analysis. Am J Phys Anthropol 48:233–239.

Hanson DB, Buikstra JE (1987): Histomorphological alteration in buried human bone from the Lower Illinois Valley: Implications for paleodietary research. J Archaeol Sci 14:549–563.

Hassan FA (1981): "Demographic Archaeology." New York: Academic Press.

Hillson S (1986): "Teeth." Cambridge: Cambridge University Press.

Hinton RJ (1982): Differences in interproximal and occlusal tooth wear among prehistoric Tennessee Indians: Implications for masticatory function. Am J Phys Anthropol 57:103–115.

Hooton EA (1920): "Indian Village Site and Cemetery near Madisonville, Ohio." Pap Peabody Mus, Archaeol Ethnol Harvard Univ 8(1).

Hooton EA (1930): "The Indians of Pecos Peublo: A Study of Their Skeletal Remains." New Haven: Yale University Press.

Horseman A, Nordin C, Simpson M, Speed R (1982): Cortical and trabecular bone status in elderly women with femoral neck fracture. Clin Orthop Related Res 166:143–151.

Howell N (1973): The feasibility of demographic studies in "anthropological" populations. In Crawford MH, Workman PL (eds): "Methods and Theories in Anthropological Genetics." Albuquerque, NM: University of New Mexico Press, pp 249–262.

Howell N (1976): Toward a uniformitarian theory of human paleodemography. In Ward RH, Weiss KM (eds): "Demographic Evolution and Human Populations." London: Academic Press, pp 25–40.

Howell N (1982): Village composition implied by a paleodemographic life table: The Libben site. Am J Phys Anthropol 59:263–269.

İşcan MY, Loth SR, Wright RK (1984): Age estimation from the rib phase analysis: White males. J Forensic Sci 29: 1094–1104.

İşcan MY, Loth SR, Wright RK (1985): Age estimation from the rib by phase analysis: White females. J Forensic Sci 30:853–863.

Jackes MK (1977): "The Huron Spine." University of Toronto: PhD thesis.

Jackes MK (1983): Osteological evidence of smallpox: A possible case from seventeenth century Ontario. Am J Phys Anthropol 60:75–81.

Jackes MK (1985): Pubic symphysis age distributions. Am J Phys Anthropol 68:281–299.

Jackes MK (1986): Mortality of Ontario archaeological populations. In Pfeiffer (ed): "The Skeletal Biology of Ontario Populations." Can J Anthropol 5:33–48.

Jackes MK (1988): Demographic change at the Mesolithic–Neolithic transition: Evidence from Portugal. Riv Antropol (suppl) 66:141–158.

Jackes MK (1990): Diagenetic change in prehistoric Portuguese human bone (4000 to 8000 BP). Eighteenth annual meeting of the Canadian Associan for Physical Anthropology, Banff, Alberta.

Jackes MK (nd): The Osteology of the Grimsby Cemetery. Unpublished manuscript.

Jhamaria NL, Lal KB, Udawat M, Banerji P, Kabra SG (1983): The trabecular pattern of the calcaneum as an index of osteoporosis. J Bone Joint Surg 65-B:195–198.

Jit I, Bakshi V (1986): Time of fusion of the human mesosternum with manubrium and xiphoid process. Indian J Med Res 83:322–331.

Johanson G (1971): Age determination from human teeth: A critical evaluation with special consideration of changes after fourteen years of age. Odontol Rev 22(Suppl 21): 1–126.

Johnston FE, Snow CE (1961): The reassessment of the age and sex of the Indian Knoll skeletal population: Demographic and methodological aspects. Am J Phys Anthropol 19:237–244.

Katz D, Suchey JM (1986): Age determination of the male os pubis. Am J Phys Anthropol 19:237–244.

Katz D, Suchey JM (1989): Race differences in pubic symphyseal aging patterns in the male. Am J Phys Anthropol 80:167–172.

Katzenberg MA, White R (1979): The *os coxae* from Ossossané ossuary: A paleodemographic analysis. Can Rev Phys Anthropol 1:10–28.

Kaur H, Jit I (1990): Age estimation from cortical index of the human clavicle in northwest Indians. Am J Phys Anthropol 83:297–305.

Kay RF, Cant JGH (1988): Age assessment using cementum annulus counts and tooth wear in a free-ranging population of *Macaca mulatta*. Am J Primatol 15:1–15.

Kerley ER (1965): The microscopic determination of age in human bone. Am J Phys Anthropol 23:149–163.

Kieser JA, Preston CB, Evans WG (1983): Skeletal age at death: An evaluation of the Miles method of ageing. J Archaeol Sci 10:9–12.

Kieser JA, Groeneveld HT, Preston CB (1985): Patterns of dental wear in the Lengua Indians of Paraguay. Am J Phys Anthropol 66:21–29.

Kilian J, Vlček E (1989): Age determination from teeth in the adult. In İşcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 255–275.

Koike H, Ohtaishi N (1985): Prehistoric hunting pressure estimated by the age composition of excavated sika deer (*Cervus nippon*) using the annual layer of tooth cement. J Archaeol Sci 12:443–456.

Konigsberg LW, Buikstra JE, Bullington J (1989): Paleodemographic correlates of fertility. Am Antiq 54:626–636.

Koyabashi K (1967): Trend in length of life based on human skeletons from prehistoric to modern times in Japan. J Fac Sci Univ Tokyo: Sect V, 3(Pt2):107–162.

Lazenby RA (1984): Inherent deficiencies in cortical bone microstructural age estimation techniques. Ossa 9/11: 95–103.

LeGeros RZ, Trautz OR, LeGeros JP, Klein E, Shirra WP (1967): Apatite crystallites: Effects of carbonate on morphology. Science 155:1409–1411.

Lengyel I (1968): Biochemical aspects of early skeletons. In Brothwell D (ed): "The Skeletal Biology of Earlier Human Populations." Oxford: Pergamon Press, pp 271–288.

Lengyel I (1978): Laboratory examination of the Vlasac human bone finds. Vlasac Mesolitiko Naselje na Djerdapu Tom II. Serbian Academy of Sciences and Arts, Monographs Vol. 512, Department of Historical Sciences, 5: 261–283.

Lengyel I, Nemeskéri J (1965): Investigation of the chemical composition of aged human bones belonging to recent and subfossil periods. In "International Conference on Gerontology, Separatum." Budapest: Akadémiai Kiadó, pp 141–146.

Lovejoy CO (1985): Dental wear in the Libben population: Its functional pattern and role in the determination of adult skeletal age at death. Am J Phys Anthropol 68:47–56.

Lovejoy CO, Meindl RS, Pryzbeck TR, Barton TS, Hemple KG, Kotting D (1977): Paleodemography of the Libben Site, Ottawa County, Ohio. Science 198:291–293.

Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ (1985a): Multifactorial determination of skeletal age of death: A method and blind tests of its accuracy. Am J Phys Anthropol 68:1–14.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP (1985b): Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. Am J Phys Anthropol 68: 15–28.

Lubell D, Jackes MK (1988): Portuguese Mesolithic–Neolithic subsistence and settlement. Riv Antropol 66 (Suppl):231–248.

Lubell D, Jackes MK, Meiklejohn C (1989): Archaeology and human biology of the Mesolithic–Neolithic transition in southern Portugal. In Bonsall C (ed): "The Mesolithic in Europe: Papers Presented at the Third International Symposium, Edinburgh 1985." Edinburgh: John Donald, pp 632–640.

Macchiarelli R (1988): Age-related rates and patterns of cortical bone involution in past human populations: A protohistorical Italian example. Riv Antropol 66(Suppl): 55–76.

Maples WR (1978): An improved technique using dental histology for estimation of adult age. J Forensic Sci 23: 764–770.

Martin RB, Burr DB (1989): "Structure, Function, and Adaptation of Compact Bone." New York: Raven Press.

Martin DL, Armelagos GJ, Mielke JH, Meindl RS (1981): Bone loss and dietary stress in an adult skeletal population from Sudanese Nubia. Bull Mém Soc Anthropol Paris (Ser 13) 8:307–319.

Masset C (1971): Erreurs systématiques dans la détermination de l'âge par les sutures crâniennes. Bull Mém Soc Anthropol Paris (Ser 12) 7:85–105.

Masset C (1973): Influence du sexe et de l'âge sur la conservation des os humains. In Sauter M (ed): "L'Homme, Hier et Aujourd'hui: Recueil d'études en Hommage à André Leroi-Gourhan." Paris: Cujas, pp 333–343.

Masset C (1976a): Sur quelques fâcheuses méthodes de détermination de l'âge des squelettes. Bull Mém Soc Anthropol Paris (Sér 13) 3:329–336.

Masset C (1976b): Sur des anomalies d'ordre démographique observées dans quelques sepultures néolithiques. In "IXe Congrès, Union Internationale des Sciences Préhistoriques et Protohistoriques, Nice." Prétirage, Thémes specialisés, pp 78–105.

Masset C (1976c): Sur la mortalité chez les anciens Indiens de l'Illinois. Curr Anthropol 17:128–131.

Masset C (1989): Age estimation on the basis of cranial sutures. In İşcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 71–105.

Masset C, Parzysz B (1985): Démographie des cimetières? Incertitude statistique des estimateurs en paléodémographie. L'Homme 94, XXV, pp 147–154.

McKee JK, Molnar S (1988): Measurements of tooth wear among Australian Aborigines: II. Intrapopulational variation in patterns of dental attrition. Am J Phys Anthropol 76:125–136.

McKern TW, Stewart TD (1957): Skeletal age changes in young American males. Quartermaster Research and Development Command Technical Report EP-45.

Meindl TW, Lovejoy CO (1985): Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. Am J Phys Anthropol 68:57–66.

Meindl RS, Lovejoy CO (1989): Age changes in the pelvis: Implications for paleodemography. In İşcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 137–168.

Meindl RS, Lovejoy CO, Mensforth RP (1980): Multifactorial determination of skeletal age at death: A double blind test on a population of known age (abstract). Am J Phys Anthropol 52:255.

Meindl RS, Lovejoy CO, Mensforth RP (1983): Skeletal age at death: Accuracy of determination and implications for human demography. Hum Biol 55:73–87.

Meindl RS, Lovejoy CO, Mensforth RP, Walker RA (1985): A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. Am J Phys Anthropol 68: 29–45.

Meindl RS, Russell KF, Lovejoy CO (1990): Reliability of age at death in the Hamann–Todd collection: Validity of subselection procedures used in blind tests of the summary age technique. Am J Phys Anthropol 83:349–357.

Mellors RC, Carroll KG, Solberg T (1966): Quantitative analysis of CA/P molar ratios in bone tissue with the electron probe. In "The Electron Microprobe." New York: Wiley, pp 834–840.

Mensforth RP, Latimer BM (1989): Hamann–Todd Collection aging studies: Osteoporosis fracture syndrome. Am J Phys Anthropol 80:461–479.

Miles AEW (1963): The dentition in the assessment of individual age in skeletal material. In Brothwell DR (ed): "Dental Anthropology." New York: Pergamon, pp 191–209.

Miles AEW (1978): Teeth as an indicator of age in man. In Butler PM, Joysey KA (eds): "Development, Function, and Evolution of Teeth." London: Academic Press, pp 455–462.

Miller CS, Dove SB, Cottone JA (1988): Failure of use of cemental annulations in teeth to determine the age of humans. J Forensic Sci 33:137–143.

Molleson TI, Cohen P (1990): The progression of dental attrition stages used for age assessment. J Archaeol Sci 17:363–371.

Molnar S, McKee JK, Molnar IM, Przybeck TR (1983a): Tooth wear rates among contemporary Australian Aborigines. J Dent Res 62:562–565.

Molnar S, McKee JK, Molnar I (1983b): Measurements of tooth wear among Australian Aborigines: 1. Serial loss of the enamel crown. Am J Phys Anthropol 61:51–65.

Moore JA, Swedlund AC, Armelagos GJ (1975): The use of life tables in paleodemography. In Swedlund AC (ed): "Population Studies in Archaeology and Biological Anthropology: A Symposium." Society of American Archeology: Memoir No. 30, pp 57–70.

Moorrees CFA, Fanning EA, Hunt EE (1963): Age variation of formation stages for ten permanent teeth. J Dent Res 42:1490–1502.

Morgan DB (1973): The metacarpal bone. Clin Radiol 24: 77–82.

Mullen G, Gruspier K (1991): Maxillary suture obliteration: A test of the Mann method. J Forensic Sci 36:512–519.

Narasaki S (1990): Estimation of age at death by femoral osteon remodeling: Application of Thompson's core technique to modern Japanese. J Anthropol Soc Nippon 98:29–38.

Naylor JW, Miller WG, Stokes GN, Stott GG (1985): Cemental annulation enhancement: A technique for age determination in man. Am J Phys Anthropol 68:197–200.

Nelson BK, DeNiro MJ, Schoeninger MJ, Paolo DJ, Hare PE (1986): Effects of diagenesis on strontium, carbon, nitrogen and oxygen concentration and isotopic composition of bone. Geochim Cosmochim Acta 50:1941–1949.

Nemeskéri J (1978): Demographic structure of the Vlasac epipaleolithic population. Vlasac Mesolitiko Naselje na Djerdapu Tom II. Serbian Academy of Sciences and Arts, Monographs Vol. 512, Department of Historical Sciences, 5:97–133.

Nemeskéri J, Harsányi L, Acsádi G (1960): Methoden zur diagnose des Iebensalters von skelettfunden. Anthropol Anz 24:70–95.

Nordin BEC, MacGregor J, Smith DA (1966): The incidence of osteoporosis in normal women, its relation to age and the menopause. Q J Med 35(137):25–38.

Ortner DJ, Von Endt DW (1971): Microscopic and electron microprobe characterization of the sclerotic lamellae in human osteons. Israel J Med Sci 7:480–482.

Owsley DW, Bass WM (1979): A demographic analysis of skeletons from the Larson Site, Walworth County, South Dakota. Am J Phys Anthropol 51:145–154.

Paine RR (1989): Model life table fitting by maximum likelihood estimation: A procedure to reconstruct paleodemographic characteristics from skeletal age distributions. Am J Phys Anthropol 79:51–61.

Palmer E (1987): Micromorphology of prehistoric human bone from the Mesolithic site of Moita do Sebastião, Portugal. Unpublished MA thesis, Department of Anthropology, University of Alberta.

Pate FD, Hutton JT (1988): The use of soil chemistry data to address postmortem diagenesis in bone mineral. J Archaeol Sci 15:729–739.

Petersen W (1975): A demographer's view of prehistoric demography. Curr Anthropol 16:227–246.

Pfeiffer S (1980): Bone remodeling age estimates compared with estimates by other techniques. Curr Anthropol 21: 793–794.

Pfeiffer S (1984): Palaeopathology in an Iroquoian ossuary, with special reference to tuberculosis. Am J Phys Anthropol 64:181–189.

Pfeiffer S (1985): Comparison of adult age estimation techniques using an ossuary sample. Can J Anthropol 4:13–17.

Pfeiffer S (1986): Morbidity and mortality in the Uxbridge ossuary. Can J Anthropol 5:23–31.

Pfeiffer S (1989): Characterization of archaeological bone decomposition in a sample of known length of interment (abstract). Am J Phys Anthropol 78:283.

Pfeiffer S (1991): An Exploration of Possible Relationships Between Bone Structure and Chemical Composition. First International Congress of Mummy Studies, Tenerife.

Philippas GG, Applebaum E (1967): Age changes in the permanent upper lateral incisor. J Dent Res 46:1002–1009.

Philippas GG, Applebaum E (1968): Age change in the permanent upper canine teeth. J Dent Res 47:411–417.

Piepenbrink H (1986): Two examples of biogenous dead bone decomposition and their consequences for taphonomic interpretation. J Archaeol Sci 13:417–430.

Piontek J, Weber A (1986): Controversy on paleodemography. Int J Anthropol 5:71–83.

Protsch RRR (1986): Radiocarbon dating of bone. In Zimmerman MR, Angel JL (eds): "Dating and Age Determination of Biological Materials." London: Croom Helm, pp 3–38.

Richman EA, Ortner DJ, Schulter-Ellis FP (1979): Differences in intracortical bone remodeling in three aboriginal American populations: Possible dietary factors. Calc Tissue Int 28:209–214.

Rogers T (1990): A test of the auricular surface method of estimating age-at-death and a discussion of its usefulness in the construction of paleodemographic lifetables. Eighteenth annual meeting of the Canadian Association of Physical Anthropology, Banff, Alberta.

Salomon CD, Haas N (1967): Histological and histochemical observations on undecalcified sections of ancient bones from excavations in Israel. Israel J Med Sci 3:747–754.

Samson C, Branigan K (1987): A new method for estimating age at death from fragmentary and weathered bone. In Boddington A, Garland AN, Janaway RC (eds): "Death, Decay and Reconstruction." Manchester: Manchester University Press, pp 101–108.

Sattenspiel L, Harpending H (1983): Stable populations and skeletal age. Am Antiq 48:489–498.

Saunders SR, Spence MW (1986): Dental and skeletal age determinations of Ontario Iroquois infant burials. Ontario Archaeol 46:45–54.

Schoeniger MJ, Moore KM, Murray ML, Kingston JD (1989): Detection of bone preservation in archaeological and fossil samples. Appl Geochem 4:281–292.

Sillen A (1989): Diagnosis of the inorganic phase of cortical bone. In Price TD (ed): "The Chemistry of Prehistoric Human Bone." Cambridge: Cambridge University Press, pp 211–229.

Singh IJ (1972): Femoral trabecular pattern of the upper end of the femur as an index of osteoporosis. J Bone Joint Surg 52-A:457–467.

Sjøvold T (1978): Inference concerning the age distribution of skeletal populations and some consequences for paleodemography. Budapest: Anthropol Kozlemenyek 22:99–114.

Smith BH (1984): Patterns of molar wear in hunter–gatherers and agriculturalists. Am J Phys Anthropol 63:39–56.

Smith P, Bloom RA, Berkowitz J (1984): Diachronic trends in humeral cortical thickness of Near Eastern populations. J Hum Evol 13:603–612.

Sognnaes RF (1955): Postmortem microscopic defects in the teeth of ancient man. Am Med Assoc Archaeol Pathol 59:559–570.

Sorg MH, Andrews RP, İşcan MY (1989): Radiographic aging of the adult. In İşcan MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 169–193.

Stewart (1957): Distortion of the pubic symphyseal surface in females and its effect on age determination. Am J Phys Anthropol 15:9–18.

Stewart TD (1958): The rate of development of vertebral osteoarthritis in American whites and its significance in skeletal age identification. The Leech (Johannesburg) 28: 114–151.

Storey R (1985): An estimate of mortality in a Pre-Columbian urban population. Am Anthropol 87:519–535.

Stout SD (1978): Histological structure and its preservation in ancient bone. Curr Anthropol 19:601–604.

Stout SD (1983): The application of histomorphometric analysis to ancient skeletal remains. Anthropos (Greece) 10:60–71.

Stout SD (1989a): Histomorphometric analysis of human skeletal remains. In İşcan MY, Kennedy KAR (eds): "Reconstruction of Life From the Skeleton." New York: Alan R. Liss, pp 41–52.

Stout SD (1989b): The use of cortical bone histology to estimate age at death. In Iścạn MY (ed): "Age Markers in the Human Skeleton." Springfield, IL: Charles C Thomas, pp 195–207.

Stout SD, Gehlert SJ (1980): The relative accuracy and reliability of histological aging methods. Forensic Sci Int 15:181–190.

Stout SD, Simmons D (1979): Use of histology in ancient bone research. Yearb Phys Anthropol 22:228–249.

Stout SD, Teitlebaum SL (1976): Histological analysis of undecalcified thin sections of archaeological bone. Am J Phys Anthropol 44:263–270.

Suchey JM (1979): Problems in the aging of females using the os pubis. Am J Phys Anthropol 44:263–270.

Susini A, Baud CA (1988): Altérations osseuses dues à l'action de la chaux. Journées Anthropol 24:69–75.

Susini A, Baud CA, Lacotte D (1987): Bone apatite crystals alterations in Neolithic skeletons and their relations to burial practices and soil weathering. Riv Antropol 66(Suppl):35–38.

Swedlund AC (1975): Population studies in archaeology and biological anthropology: A symposium. Soc Am Arch Mem 30 (issued as Am Antiq 40(2), Pt. 2).

Swedlund AC, Armelagos GJ (1969): Une recherche en paléo-demographie: La Nubie soudanaise. Annales: Economies, Sociétés, Civilisations 24:1287–1298.

Swedlund AC, Armelagos GJ (1976): "Demographic Anthropology." Dubuque, IA: Brown.

Tal H (1984): The prevalence and distribution of periodontal bone loss in dry mandibles of Bantu-speaking South African Blacks. Ossa 9/11:181–188.

Thompson DD (1979): The core technique in the determination of age at death in skeletons. J Forensic Sci 24: 902–915.

Thompson DD (1981): Microscopic determination of age at death in an autopsy series. J Forensic Sci 26:470–475.

Thompson DD, Gunness-Hey M (1981): Bone mineral–osteon analysis of Yupik–Inupiag skeletons. Am J Phys Anthropol 55:1–8.

Thompson DD, Laughlin SB, Laughlin WS, Merbs C (1984): Bone core analysis and vertebral pathologies Sadlermiut Eskimo skeletons. Ossa 9/11:189–193.

Thwaites RG (ed) (1898): "The Jesuit Relations and Allied Documents." Cleveland: Burrows.

Todd TW (1920): Age changes in the pubic bone, I: The male white pubis. Am J Phys Anthropol 3:285–334.

Todd TW (1921): Age changes in the pubic bone. II–IV. Am J Phys Anthropol 4:1–70.

Tomenchuk J, Mayhall JT (1979): A correlation of tooth wear and age among modern Igloolik Eskimos. Am J Phys Anthropol 51:67–77.

Trodden BJ (1982): A Radiographic Study of the Calcification and Eruption of the Permanent Teeth in Inuit and Indian Children. Archaeological Survey of Canada, Paper 112, Mercury Series. Ottawa: National Museums of Canada.

Turpeinen O (1979): Fertility and mortality in Finland since 1750. Popul Stud 33:101–114.

Ubelaker DH (1974): Reconstruction of demographic profiles from ossuary skeletal samples. A case study from the Tidewater Potomac. Smithsonian Contributions to Anthropology 18.

Ubelaker DH (1986): Estimation of age at death from histology of human bone. In Zimmerman, Angel JL (eds): "Dating and Age Determination of Biological Materials." London: Croom Helm, pp 240–247.

Ullrich H (1975): Estimation of fertility by means of pregnancy and childbirth alterations at the pubis, the ilium, and the sacrum. Int J Skel Res 2:23–39.

United Nations (1982): "Model Life Tables for Developing Countries." New York: United Nations.

Uytterschaut HT (1985): Determination of skeletal age by histological methods. Z Morphol Anthropol 75:331–340.

Van Gerven DP, Armelagos GJ (1983): "Farewell to paleodemography?" Rumours of its death have been exaggerated. J Hum Evol 12:353–360.

Vlček E, Mrklas L (1975): Modifications of the Gustafson method of determination of age according to teeth on prehistorical and historical osteological material. Scripta Med 48:203–208.

Waldron T (1987): The relative survival of the human skeleton: Implications for palaeopathology. In Boddington A, Garland AN, Janaway RC (eds): "Death, Decay and Reconstruction." Manchester: Manchester University Press, pp 55–64.

Walker RA, Lovejoy CD (1985): Radiographic changes in the clavicle and proximal femur and their use in the determination of skeletal age at death. Am J Phys Anthropol 68:67–78.

Walker PL, Johnson JR, Lambert PM (1988): Age and sex biases in the preservation of human skeletal remains. Am J Phys Anthropol 76:183–188.

Weiss KM (1973): Demographic models for anthropology. Mem Soc Am Archaeol No. 27.

Weiss KM (1976): Demographic theory and anthropological inference. Annu Rev Anthropol 5:351–381.

Werelds R-J (1961): Observations macroscopique et microscopique sur certaines altérations postmortem des dents. Bull Group Int Rech Sci Stomatol 4:7–60.

Werelds R-J (1962): Nouvelles observations sur les dégradations postmortem de la dentine et du cément des dents inhumées. Bull Group Int Rech Sci Stomatol 5:554–591.

Werelds R-J (1967): Du moment ou apparaissent dans les dents humaines les altérations postmortem en forme d'évidements canaliculaires: Présence de lesions dentaires identiques *in vivo* chez les poissons. Bull Group Int Rech Sci Stomatol 10:419–477.

Whittaker DK, Ryan S, Weeks K, Murphy WM (1987): Patterns of approximal wear in cheek teeth of a Romano-British population. Am J Phys Anthropol 73:389–396.