In a recent article entitled “Farewell to Paleodemography” Bocquet-Appel & Masset (1982) review and critically assess what they feel to be the principal methods and accomplishments of paleodemography. They are particularly concerned with the accuracy by which individual skeletons can be aged, and with the degree to which skeletal ages can be combined into meaningful demographic profiles. Their assessment is far from encouraging. Beginning first with the creation of demographic profiles, the authors compare several skeletal populations with their reference populations, that is, the populations by which the skeletons were aged. For example, the age distribution of skeletons from ancient Nubia (Swedlund & Armelagos, 1969) is compared to the age distribution of the McKern–Stewart sample of Korean War dead by whose pubic age criteria the Nubians were presumably aged. On the basis of such comparisons the authors conclude that skeletal populations are doomed to emerge as demographic copies of their reference populations.

The authors then proceed to the question of aging the individual skeleton. After considering various methods separately and combined, they estimate the highest correlation between chronological and skeletal age to be approximately 0.8. Based on a number of a priori criteria and assumptions, this level of accuracy is judged to be completely inadequate. The authors suggest, for example, that in order to insure a 95% level of accuracy in age classification for a population whose actual ages range over a 72-year span, only 1.4 age categories can be meaningfully created.

When then is the fate of paleodemography and those who continue to estimate ages and create age categories? The authors leave little doubt.

The scholars who persist in this course will only obtain artifacts; the information conveyed by the age indicators is so poor that the age distributions thus available can hardly reflect anything but random fluctuations and errors of method [emphasis ours] (Bocquet-Appel & Masset, 1982; p. 329).
The authors' farewell to paleodemography is clearly not a personal adieu but an obituary for the inquiry as a whole. Before the death becomes official, however, some reappraisal of the author's diagnosis seems warranted. We intend to consider carefully whether the authors, themselves, have adequately and accurately portrayed the methods, objectives and accomplishments of paleodemography. It is our position that the methods employed by paleodemographers are far from inadequate. It is also our position that the results of paleodemography lend absolutely no support to the contention that they represent nothing more than “random fluctuations and errors of method”. The authors' harsh appraisal must stand not only in light of their personal criteria but the criteria reasonably, conscientiously and successfully applied by others. We will then consider whether the authors have adequately considered and assessed the wider-ranging objectives and accomplishments of paleodemography.

2. The Relationship Between Skeletal Samples and Their Reference Populations

In order to determine whether skeletal samples are, as suggested by Bocquet-Appel & Masset, passive reflections of their reference populations, two skeletal samples from Sudanese Nubia were selected for analysis. The first represents 201 individuals excavated near Wadi Halfa and reported by Swedlund & Armelagos (1969). Although the reference population associated with this sample by Bocquet-Appel & Masset was the McKern–Stewart Korean War dead, the system most extensively used was that developed by Todd (1920) and later modified by Brooks (1955). The second sample consists of 162 individuals excavated at the site of Kulubnarti (Van Gerven et al., 1981). This sample was also aged using changes in the os pubis as described by Todd and later modified by Brooks.

Age categories for the two samples were arranged in five-year intervals from 15 through 55+. This was also done for Todd's original sample (Todd, 1920). In order to facilitate statistical comparison, the data was then cast into cumulative mortality curves (Figure 1). Statistical comparisons between each skeletal sample and the Todd reference sample were made using the Kolmogorov-Smirnov two-sample test for cumulative frequencies as discussed by Siegel (1958) and Lovejoy (1971).

As illustrated in Figure 1, differences between the Nubian samples and the Todd

Figure 1. A comparison of cumulative mortality between two ancient Nubian populations and their reference population based on Todd’s (1920) analysis of 306 American males. [Kulubnarti (- - -), Wadi Halfa (-----), Todd (-----)].
reference sample are extensive. With the exception of the first age category, all differences are significant at the 99% confidence level. Furthermore, the two skeletal samples are strikingly different from one another. These results in no way support the contention that skeletal samples correspond primarily to the age structure of their reference sample. The results do, on the other hand, correspond to previously observed differences among skeletal populations aged according to a common reference group. Green and co-workers (1974) for example, found important differences between skeletal samples from the site of Meinarti in Sudanese Nubia corresponding to periods of village growth and decline, they also found significant differences in life expectancy related to social status within the Meinarti population. It must be emphasized that all the Meinarti skeletal remains were aged in the same manner. Such differences, while making excellent biocultural sense, are difficult to explain if skeletal remains reflect only the demographic structure of their reference sample or random fluctuations and errors of method.

3. The Determination of Age at Death From Skeletal Remains

The second critical issue in Bocquet-Appel & Masset’s “Farewell to Paleodemography” concerns the ability of skeletal biologists to assign accurately ages at death to skeletons. While they accept the utility of age-dependent criteria applied to subadults, the authors are far more skeptical of age-related changes in the adult skeleton. It is suggested that unless the correlation between chronological age and skeletal age reaches 0.9 or better, the prospects for paleodemography are dim. It is our contention that this a priori assessment is of little practical value in determining the actual utility of skeletal age data.

In order to assess the impact of errors in age estimation, the sample of 306 males analysed by Todd (1920) for age changes in the os pubis were cast into a cumulative mortality curve based on their chronological ages. The sample was then cast into the same age categories based on pubic ages determined according to Todd without knowledge of chronological age. Pubic age values were derived from Todd’s ten stage system by averaging the oldest and youngest age associated with each individual’s stage assignment. All individuals whose oldest stage value was 10 (50+ years) were assigned to the 50-55+ age category. While this procedure may have resulted in our overaging some individuals classified by Todd as stage 9-10, it allowed us to avoid averaging his 50+ value as a fixed integer.

As indicated in Figure 2, the cumulative frequency distributions created by chronological and pubic age data are strikingly similar and are in fact tied through the first two age groups. A maximum difference of 9% occurs at age 40 and is statistically significant. Differences between all other age groups are not significant at the 95% confidence level. It is important to emphasize that the Todd system is seldom used in isolation. As Bocquet-Appel & Masset themselves demonstrate, the ability to assess age at death improves when multiple criteria are applied. Consequently, it is doubtful that skeletal age indicators produce nothing of interest for demographic purposes. Put another way, if the pubic ages derived from the Todd sample were our only basis for reconstructing their mortality profile, would it be reasonable to conclude that we had reconstructed nothing but “random fluctuations and errors of method”?

As a final point, it must be recognized that the age distribution cannot be assessed beyond the 50-55+ category. While only a small percentage (usually between 1 and 10%)
of skeletal remains fall in this interval, Bocquet-Appel & Masset contend that this constitutes a major source of demographic error. They, however, offer little in the way of concrete support for their contention that a large portion of adults in prehistoric societies lived well into their 60s, 70s and even 80s. They site mortality among the Dobe !Kung as exemplifying what they believe to be a typical mortality pattern with mode adult mortality

4. The Wider-Ranging Objectives and Accomplishments of Paleodemography

If, as Bocquet-Appel & Masset suggest, paleodemographic reconstructions provide nothing more than random fluctuations and errors of method, the consequences of such errors and fluctuations should be apparent in all aspects of paleodemographic analysis—not simply in the construction of mortality profiles. It is important to emphasize that physical anthropologists have not limited their use of demographic data to questions of mortality and survivorship. Researches into paleopathology as well as growth and development have made abundant use of mortality data. The accomplishments of this large body of research must also be given consideration before paleodemography is laid to rest.

Recent research into the age-related loss of skeletal tissue known as osteoporosis (Dewey et al., 1969; Van Gerven et al., 1969) clearly illustrates the value of age data derived from skeletal remains. In order to determine whether ancient human populations experienced osteoporotic bone changes, Dewey and co-workers measured cortical bone thickness at the proximal one-third of femurs from 203 Nubians excavated near Wadi Halfa. The individuals are from the sample a portion of which was discussed earlier and illustrated in Figure 1. As previously indicated, all ages at death were determined in the field using
changes in the os pubis. Observations of cortical thickness were not a factor in age determination. Measurements of cortical thickness were made on isolated femur sections prepared in the physical anthropology laboratories at the University of Utah without knowledge of the age or sex of any specimen. After cortical thickness values had been obtained for the sample, the age data was added and mean cortical thickness was determined for four age groups as illustrated in Figure 3.

Figure 3. Percentage reduction in average cortical thickness (osteoporosis) for males and females in (a) a modern American and (b) Ancient Nubian sample.

The results were then compared to the pattern of bone loss previously documented for a modern American sample (Bartley & Arnold, 1965) of known age using the same technique of cortical bone measurement. While the age groups used by Dewey do not correspond precisely to the modern series, the similarity between the ancient Nubian and modern American patterns of cortical bone loss is striking.* Both indicate a rapid rate of cortical thinning with advancing age among women with a much lower rate among men.

Before we can conclude, as Bocquet-Appel & Masset do, that skeletal age data is spurious, it must be explained how observations of bone loss, made independently of age, could produce such a coherent pattern. Indeed, if Bocquet-Appel & Masset are correct, such an analysis of osteoporosis should have produced no discernible change in cortical thickness by age—that is, the relationship should have emerged as a random one.

The successful application of a demographic approach to phenomena other than mortality has not been limited to osteoporosis. Lovejoy & Heiple (1981), for example, found a correlation of 0.97 between long-bone fractures and years at risk (determined from skeletal age) in a Late Woodland population from the Libben site in Northern Ohio. Beyond a simple measure of correlation, however, their demographic approach provided additional insight into the fracture process as it impacted different age groups. According to Lovejoy & Heiple (1981):

Rates of fracture show marked elevation in two periods of the life cycle: adolescence/young adulthood and old age. These are the two periods of the life cycle in which one would most likely expect such high rates, if accidental trauma were the primary cause (p. 538).

*In order to facilitate comparison, values beyond age 50 were combined by averaging in the modern sample.
These results also correspond closely to the pattern determined from a demographic survey of fractures in modern England and Wales (Buhr & Cooke, 1959) in which fractures associated with high levels of activity accumulated during the young years while other kinds of fractures accumulated during older age.

Further examples could be cited but the point would be the same. When age-related events and processes affecting skeletal remains are examined independently of age estimation, i.e. demographically, their patterns of occurrence make sense in light of modern skeletal biology. In no case do the results reflect the random fluctuations and errors of method predicted by Bocquet-Appel & Masset. The problem, therefore, seems to lie more with the authors' "farewell" and its underlying assumptions than with age estimation and paleodemography.

Before concluding, one final issue must be addressed. At the outset of their discussion, Bocquet-Appel & Masset acknowledged the accuracy with which age can be determined for subadult remains and accordingly exempt this segment of skeletal samples from their assessment. However, inasmuch as subadult remains regularly constitute over 50% of skeletal samples, and the authors conclude by dismissing all of paleodemography, some consideration of subadult remains and their value to paleodemography seems warranted. The Kulubnarti sample discussed earlier provides an excellent example of how valuable subadult materials can be to paleodemographic analysis. A comparison of probabilities of dying and mean life expectancies between early and later Christian cemeteries at the site has revealed a higher mortality rate among subadults (birth through 14 years) during early Christian times (Van Gerven et al., 1981). In order to determine whether this difference is due to differences in childhood stress and morbidity, probabilities of dying within and between the two cemeteries were compared to frequencies of cribra orbitalia. Cribra orbitalia is a lesion of the superior surface of the eye orbit observed in many Nubian populations that has been related to dietary iron deficiency as well as parasitic and bacterial infections (Carlson, et al., 1974; Van Gerven et al., 1981). In both cemeteries the frequency of cribra orbitalia shows a high correspondence to probabilities of dying from infancy through the early adult years. Of even greater significance, the higher probabilities of dying among subadults in the early Christian cemetery correspond to higher frequencies of the lesion.

These results support the contention that differences in childhood mortality from early to late Christian times are primarily due to differences in biological stress acting on these populations. It appears that the greater regional autonomy experienced by these people during the later Christian period was a positive influence on their biological well being. In terms of the issue at hand, it is particularly significant that subadult mortality is the most sensitive barometer of that biocultural change.

However, as with adult paleodemography, the analysis of subadults has not been limited to questions of mortality. For example, the analysis of bone growth and development, assessed independently of age determination, has been a major area of demographic inquiry.

Studies of ancient Amerindian (Johnston, 1962) as well as Nubian (Mahler, 1968; Hummert & Van Gerven, 1982) children have demonstrated important correspondences between prehistoric and modern growth patterns even though "...some degree of error of error is introduced by the very fact that the sample is skeletal. It does not represent the normal, healthy, population from which it was drawn" (Johnston, 1962, p. 249).

Mahler, for example, found that the velocity and symmetry of long-bone growth among
ancient Nubians was broadly similar to that of modern Americans studied by Maresh (1955) with the exception that Nubians had a later and stronger adolescent growth spurt. Mahler hypothesized that the stronger growth spurt observed for the Nubians reflected catch-up growth resulting from an inadequate childhood diet and reduced pre-adolescent growth.

Recently, Huss-Ashmore (1978) and Martin & Armelagos (1979) have investigated the nutritional basis for premature osteoporosis in juvenile and young adult Nubians and their results lend strong support to Mahler’s earlier hypothesis. When long-bone length was plotted against cortical thickness and midshaft width for the Nubian sample, it became apparent that Nubian juveniles maintained long-bone length at the expense of normal cortical bone development. Following research carried out by Garn et al. (1966) on modern children, Huss-Ashmore suggested that this decrease in cortical bone may be evidence for protein-energy malnutrition in ancient Nubia. Here, then, we have a major new hypothesis created from demographic research on subadult remains.

Studies such as this are not of simply passing importance, they represent a major arena of paleodemographic inquiry. Data produced by such research has provided important new insights into the dynamics of population adaptation as well as the biology of growth and development. And yet, if Bocquet-Appel & Masset are to be taken seriously, such research must either be dismissed or defined as non-demographic. In our view, neither solution is appropriate. We would prefer to acknowledge the real and important limitations within which all paleodemographers work, improve our methods when possible, and acknowledge the past, present and future value of paleodemographic inquiry.

5. Conclusions

In conclusion, for all of their grim projections and dire forecasts, Bocquet-Appel & Masset fail to account for one overriding fact; the results of paleodemographic research make sense. If demographic reconstructions are nothing more than passive reflections of their reference populations, important and bioculturally meaningful differences documented within and between skeletal populations should not exist, but they do exist. If age estimates combine to produce nothing more than random fluctuations and errors of method, such random fluctuations should be apparent in the study of age-dependent processes, but no such randomness is apparent. Most importantly, if paleodemography is to be dismissed, it should be found wanting in its entirety, but it is not. The authors ignore more than they consider. The study of subadult remains is entirely disregarded in their farewell as is the modern fluorescence in paleopathology and growth and development. These areas owe their current resurgence to a populational perspective provided by paleodemography.

What, then, is the status of paleodemography? The answer is clear. Rumors of its death have been greatly exaggerated.

References


