Human Age Estimation From Skeletal and Dental Evidence

D. LUCY

PhD

1996

The theories and practice of adult human age estimation focusing on the application of non-parametric likelihood estimation techniques to continuous and discrete data types

David LUCY BSc

submitted for the degree of doctor of philosophy

Department of Archaeological Sciences

University of Bradford

1996
Abstract:

Keywords: Estimation, Bayesian, non-parametric, kernel density, likelihood, senescence, ageing, hard-tissue.

Anthropologists and forensic scientists have found adult human age estimation to be fraught with difficulties since the early years of the 20th Century.

This thesis examines the reasons for the problem from the perspective of the biological principles of ageing, and statistical procedures used by anthropologists to estimate age from hard tissue age indicators.

Initially the commonly employed statistical routine of regression analysis and 'regression like' analyses are examined in critical detail. It is proven that age estimates obtained using these techniques will always lead to 'biased' estimates, an observation which occurs in the anthropological literature on an anecdoteal basis, and the use of inverse regression is recommended instead of forward regression methods.

However, many age estimation variables are quantified on some discrete scale of measurement, and inverse regression techniques are only suited to continuous variables. So an approach is developed which uses Bayes' theorem to generate estimates and confidence intervals from discrete data, and which only assumes the conditional independence of age related changes.

The principles underlying the aforementioned approach are then elaborated and generalised using kernel density smoothing to be able to make estimates from continuous and mixed continuous/discrete data for age which is treated as continuous.
Acknowledgments

I owe a great debt of gratitude to Dr. Robert Aykroyd, Department of Statistics, School of Mathematics, University of Leeds, for all the hours over a period of nearly two years he has spent with me working on many of the ideas presented here. In particular the demonstration that forward regression leads to correlated errors with a slope of 1-\(r^2\), is mostly his work. However much of the other original work presented in this thesis could not have been accomplished without his inspired guidance. I should also like to thank my supervisor, Professor Mark Pollard, whose forbearance, encouragement and sage advice has contributed to the prompt completion of this thesis, and most of the supporting publications. I thank Professor Tore Solheim, Department of Oral Pathology and Section for Forensic Odontology, Oslo, Norway, for willingly providing what must be the world’s most complete, and largest reference set of dental age indicators, and for his many and valuable contributions for our joint publications. Other people whom I should also like to thank are the many staff and students of the Department of Archaeological Sciences, University of Bradford, for their many comments and pieces of advice. In particular I should like to acknowledge the occasional help of Dr. Charlotte Roberts, Paul Maclean, Mary Lewis, Siobhan Watts, Beverley Margerison, Dr. Randolph Haggerty, Dr. Paul Budd, Brett Scaife and Assumpta Vizcaino. Finally I acknowledge the help of Dr. David Hunt, Collections Manager of the Physical Anthropology Division, Department of Anthropology, National Museum of Natural History, Washington D.C. and Dr. Patty Stuart-Macadam of the Department of Anthropology of the University of Toronto, Canada, for providing concise information on their respective skeletal collections.
1Age and Archaeology

1.1 Why age estimation is important

There are two fundamental biological parameters which carry social significance, sex and age. The social significance associated with sex has come under scrutiny by historians and archaeologists only in more recent times (Ehrenberg, 1989; Engelstand, 1991; Gero and Conkey, 1991; Taylor, 1996). However, the study of age structures in past populations has been a long standing preoccupation.

As hard tissue is usually the only surviving direct source of biological information for past human groups which did not maintain written records, age estimation from skeletal manifestations of age has always played a crucial role in palaeodemography and human osteology. There is also the more immediate value in forensic cases where the skeletalised remains of an unknown individual, or individuals, are discovered and basic biological information is required to further the identification process. Although interest in the subject appears regularly in the forensic literature, too much stress should not be placed on its forensic importance as there are only about five or six cases in any given year in which pathologists in the USA have to use skeletal age estimation techniques (Reichs, 1995). There is no similar data for Europe.

Archaeologists and anthropologists have traditionally been interested in three basic demographic questions, how many people occupied a particular area at a particular time? What was the sex makeup of that population, and how old were people when they died? To examine these particular questions the life table has been used extensively (Acsádi and Nemeskéri, 1980), which at its most fundamental level is a list of age classes and the number of individuals who attain that age. Subsidiary information, such as population survivorship, the construction of a stationary population¹ (Drenhaus, 1978; Wood et al., 1994), and life expectancies at various ages are also calculable from the life table.

¹ A stationary population is one which the age structure is in dynamic equilibrium, and the total numbers in the population do not change.
Angel (1969) pointed to the problems of using demographic life tables for archaeological populations. These were that infants would be under-represented due to differential survival in the burial environment, cemetery populations are not from a single demographic cohort, or even from a limited period of time, but are made up of all individuals from an extended time period. The question of infant under-enumeration was dealt with by Moore et al. (1975), who claimed its effect on the life table was minimal, but found that violations of the stable population assumption could not be examined using the life table. Weiss (1975) investigated the effect of simulated demographic disturbance on the life-table, that is the violation of the stable population assumption, and found that fluctuations made little difference to the perceived demography, although this wouldn't apply to secular change.

Despite the problems with the demographic device of the life table, accurate ideas of age at death (Williams, 1992), or at least the mortality profile for a cemetery population (Boldsen, 1988; Siven, 1991) is essential to any attempt to construct the demography of an extinct population. Unfortunately anthropologists have always had reservations about their ability to estimate age accurately and precisely from human adult age indicators.

The controversy surrounding the demography of the Pecos Pueblo in New Mexico is a case in point. The site of Pecos Pueblo lies on the Pecos River, towards the extreme edge of the Sangre de Cristo mountain range, and was occupied from 1300 AD. It was excavated by A.V. Kidder between 1915 and 1929 who recovered the remains of some 2000 skeletons. These were subsequently examined by E.A. Hooton and T.W. Todd who produced an extensive and detailed report (Todd, 1927) which concluded that the inhabitants of the Pecos Pueblo had a mean age at death of 42 years. Further confirmatory work on the skeletal collection was conducted by Mobley (1980) who re-examined this material producing demographic data for each major period in the site’s occupation period, which were defined by pottery style. Mobley generally agreed with Todd’s age estimates and concluded that there was no evidence to suggest any major demographic discontinuity during the site’s occupation period. However, later authors (Palkovich, 1983; Ruff, 1981) argued that Todd and Mobley had overestimated the ages of adult individuals.
in the Pecos sample, suggesting that the mortality profile of this population was much like any other indigenous American mortality profile in that the mean age at death was 25 years, rather than the 42 years suggested by Todd and Mobley.

The site of Libben, Ohio, U.S.A, yielded 1327 skeletons of the indigenous Late Woodland people (800-1100 AD). Estimates of age were obtained from seven skeletal indicators from 1289 complete skeletons. The mean age at death was just over 15 years of age, with about 10% of the population surviving beyond the age of 40. No individuals were deemed to have survived beyond the age of 50. The authors attributed this mortality profile to the peoples of Libben being a population which were subject to an epidemic in that period (Lovejoy et al., 1977). Howell (1982) follows the implications of this extremely high mortality and says that the mortality represented in the Libben population is greater than for any recorded population at any point in history, and that there could not have been any more than two generations alive at any one instance. Howell (1982) treats the empirical mortality with caution, suggesting that it may be more a result of misleading age estimates.

The Neolithic population of Isbister on Orkney fare no better than the population of prehistoric Ohio. Hedges (1982) found that the mean age at death for 342 individuals was about 25 years, and apparently very few survived beyond the age of 50 years. As with the Libben population the social questions revolved around how many offspring had to be produced to sustain such a high turnover, and the passing on of specialist knowledge.

One of the more curious regularities in palaeodemography has been that skeletal populations always seem too young, especially when the estimates of mortality derived from hard tissue sources are compared to recorded ages for those populations for which records exist. Boddington (1987, p. 188-189) critically reviews a number of such studies and concludes that 'when archaeological data is compared to historical evidence the skeletal population has consistently higher mortality rates and the suspicion of under-ageing is hard to dispel.'

This has been highlighted by other historians who review the social consequences of age. In Medieval Europe age is linked to social obligation such as tax payment, service to fiefs, and rights such as inheritance. Shahar (1993) gives a comprehensive review of old age in Medieval and Renaissance Europe. In this study old age is considered to start at any age between from
40 to 70, even as young as 35 in one case. Many of the lower estimates are the work of individual writers who upon reaching a given age considered themselves old. Sahar thinks that a more reliable source of the start of old age is rendered by the laws pertaining to exemptions from social activities such as military service and taxation, which is almost universally in the range 60 to 70 years, an age range to which no-one survived if the archaeological data is to be believed.

The doubts anthropologists had about their ability to estimate age flared into a series of acrimonious arguments initiated by Bouquet-Appel and Masset (1982) and Bouquet-Appel (1986). Initially they argued from the discrepancy between mortality distributions seen from archaeological material and those seen in historical records, that age estimates from hard tissue were systematically biased. Then, from the basis of linear regression they argued that all age estimates were in fact nothing more than a passive reflection of the age structure of the reference sample from which the estimates had been made (a topic addressed in 4.1.4). They concluded that hard tissue derived age estimates 'can hardly reflect anything but random fluctuations and errors of method', and that ancient people had lived into what we today would consider old age.

A reply came in 1983 when Van Gerven and Armelagos (1983) pointed out that a number of anthropological studies had produced age estimates which reflected a mortality pattern which did not in any way correspond to the reference samples on which those estimates were based. They also argued that Bouquet-Appel and Masset (1982) produced 'little concrete evidence that ancient people had lived into their 60s, 70s and even 80s'.

This gloomy impression of the ages to which past peoples lived has been used by many population biologists examining longer term secular changes in the human lifespan. Deevey (1960) and Brothwell (1963; 1972) put longevity of all anatomically modern sapiens at about 35 years until the turn of the 19th century, although Deevey admits that age estimates for Palaeolithic humans were 'elaborately guessed at'. The uncertainty surrounding the accuracy of anthropological age estimates can have profound effects upon the ideas of evolutionary biologists. Cutler (1978) in his seminal paper on human lifespan evolution uses the same estimates as Deevey and is more or less forced into taking a view of human lifespan evolution which could be seen as contradicting conventional explanations of lifespan evolution.
1.2 Evolution and age estimation

There is a fundamental difference of opinion between two schools of thought in workers looking at skeletal remains and those who examine historical evidence. On the one hand past peoples lived for much less long than extant peoples whom we imagine live under similar conditions, and on the other that there is little difference in the two mortality profiles. In order to gain a further insight into which school seems the more credible it is necessary to see which, if either, of these divergent and incompatible opinions is the more commensurate with other biological models.

The presence of senescence\(^2\) and death requires an evolutionary explanation. The criticism is simple. In a world where each species of organism is to a greater or lesser degree adapted to live in its environment, how can senescence and eventual death exist? How is it that organisms which are so well adapted to their environment seem to go into decline and die? The criticism is easy to sustain, it is obvious that individuals which live a long time will always reproduce more for a given rate of reproduction than those which do not live as long. This will have the effect that long lived genotypes will soon displace short lived ones.

To the nineteenth century evolutionary biologists the fact that members of a biological species would seem to have a lifespan potential which was particular to that species suggested that lifespan was an evolved trait. For instance, elephants normally live longer than cats, who usually live longer than dormice. Admittedly some elephants have shorter lives than some cats, and some cats have shorter lives than some dormice, but on the whole there seems to be a species specific lifespan. This lifespan is more latterly known as the maximum lifespan potential, and is usually considered by modern biologists to be the age which the longest living 5% of individuals which can be said to belong to that species reach.

Even within closely related biological species the maximum lifespan potential can vary tremendously. For example, Turban shells live up to about 30 years, whilst closely related molluscs, nudibranchs only live for 90 days. Neither is size highly related to maximum lifespan potential. The giant clams

---

\(^2\) Senescence for the purposes of biological thought refers to an increasing probability of death occurring with age for any organism. The term can also apply to those changes in the organism which can lead to an increased probability of death occurring, a trivial example would be the occurrence of osteoporosis which inhibits an organisms ability to escape predation.
of the tropical Pacific can grow to about a metre across in 20 years or so, whereas the freshwater mussel grows to about 150 millimetres in length, and lives for about 100 years (Kirkwood, 1985).

The problem faced by the late nineteenth century evolutionary biologists was enormously complex. Of prime importance was the fact that there seemed to be three major sorts of senescence displayed by the biological world, and that the three types of senescent behaviour were not classifiable into other known biological categories.

The three major types of senescent behaviour displayed are: non-senescent, this applies to prokaryotic organisms, organisms whose cells lack mitochondria or chloroplasts. Non-senescent in this context means that the organisms show no increased probability of death with age, and as far as anyone can tell no dysfunctional change with age (Finch and Schneider, 1985).

Semelparous organisms live until first reproduction, after reproduction they go into rapid decline, swiftly followed by death. This applies to annual plants which grow from seed, they flower, reproduce and die fairly rapidly. Female octopuses are semelparous in that after spawning the female octopus grows listless, stops feeding, and dies in a matter of a week or so. This is most certainly as a result of programmed senescence, and in the case of the octopus is found to be the result of hormones secreted by a gland connected to the optic nerve. If this gland is removed, or disconnected, after spawning the cephalopod resumes feeding and lives for several months, although it cannot be persuaded to reproduce again (Kirkwood, 1985).

The third and final type of senescent behaviour is iteroparous. Iteroparous organisms have an extended reproductive period, breeding many times. They then usually have a short post-reproductive phase followed by death. Dysfunctional changes occur gradually starting at sexual maturity. Humans are iteroparous.

1.2.1 Early evolutionary theories of senescence

The first explicitly evolutionary theory of senescence was published in 1891 by August Weissmann, the originator of the germ/soma separation theory, 32 years after the first publication of Darwin’s ‘Origin of Species’. Weissmann said (cited in Medawar, 1952):

‘Death takes place because worn out tissue cannot forever renew itself. Worn out individuals are not only valueless to the species,'
but they are even harmful, for they take the place of those which are sound ... by the operation of natural selection, the life of a theoretically immortal individual would be shortened by the amount which one was useless to the species.'

Weismann’s theory is evolutionary in the sense it takes into account natural selection and the ability of the organism to reproduce, but makes mistakes by presupposing senescence to explain it, and placing natural selection as operating on species, not heritable traits.

Weismann's theory was superseded in 1910 by a population based theory by the geneticist R.A. Fisher. Fisher's theory supposed a stable population of non-senescent organisms, such as bacteria (Medawar (1952) later employed a population of laboratory test tubes as an illustration), and each individual has a finite probability per day of dying, and were we to count them at the end of each day, then we would find that the number of individuals at the end of each day was equal to the number of individuals the day before, multiplied by the probability per day of dying. A more visual representation of this process is given by Figure 1. Notice that if the individuals do not reproduce their numbers can only decrease, this tends to zero after some time.

Figure 1. Effect of random death on a cohort of immortal objects

As this population is nominally immortal and shows no signs of senescence it follows that each individual in it, irrespective of age, has an equal probability of reproduction. If we now focus on the numbers of offspring produced for each age category it can be seen that more offspring will be produced by young individuals simply because there are more of them. Older
ones are said to have less reproductive value in this instance. Fisher thought that as older individuals contributed so little to the population of these organisms that it would be easy to lose these individuals and not notice (see Figure 2). The problem here, as with Weismann’s theory, is that natural selection has to be seen as operating on the species, not the individual, and despite producing fewer new organisms there is still some small reproductive advantage in living a long time.

*Figure 2. Mortality and reproduction for an immortal population*

The first truly adequate evolutionary theory of senescence was by Medawar and Williams, and was based upon Fisher’s demographic model, but which answered many of the objections to it. Medawar and Williams proposed that were there such a phenotypic trait which could somehow give greater reproductive success early on in life then this reproductive advantage would outweigh even a very serious side effect, such as a shortened life. In effect any trait which killed an organism a long time after birth could easily be outweighed by fast reproduction early on, because the organism is unlikely to survive to old age due to random death throughout life (Medawar, 1952). This also explained why biologists had commonly used a Gompertz\(^3\) model, rather than a random decay model, to examine naturally occurring biological populations. This model is shown in Figure 3.

\(^3\) The exponential model suggested by Fisher is similar to any other exponential decay model i.e.: \(p_t = p_0e^{-\mu t}\) where \(p\) is the population remaining from a given cohort at any instant \(t\), and \(\mu\) is a constant and can be taken to be the probability of death at \(t\). Where the Gompertz model differs is \(\mu\) is a function of \(t\) which increases with \(t\), hence an increased probability of death with increasing \(t\).
This principle has been later termed antagonistic pleiotropy and has a certain amount of empirical support, the first evidence coming relatively quickly in 1969 from human populations. Hamilton and Mestler (1969) used the fact that about 15% of male inmates in psychiatric institutions in the United States were castrated for reasons of tractability. They found that the median lifespan for castrated males was some 14 years longer than for intact males, the difference for females being not so marked. The suggested reason for this by the evolutionary biologists was that sex hormones originating in the testis lead to greater reproductive success early in life, but ultimately an earlier death independent of mechanism.

Later work, reviewed by Albin (1988), showed that individuals who later displayed Huntingdon's disease, a neuro-disease which manifests itself in early middle age; had a significant (1.09 times as many offspring) reproductive advantage earlier on in life than individuals which did not later manifest Huntingdon's disease.

The very powerful combination of Fisher's statistical theory and Medawar and Walker's antagonistic pleiotropic mechanism not only gives an adequate basis for death to be introduced to a potentially immortal population, but an idea as to how maximum lifespan potential can change through evolutionary processes.

If we have a population of a known survivorship, and some indeterminate pressure which is killing members of that population suddenly is released then a greater number of members of that population will reach an older age. If more members are reaching an older age, then for a given
reproductive rate, ones which live longer will be at a reproductive advantage to those which don't, as in total they will have more offspring. Therefore there will be a steady rise in maximum lifespan potential until such a time where long-life confers no reproductive advantage.

If instead the pressure on the population grows greater, fewer members of that population will grow old and reap the benefits of an extended lifespan in terms of reproductive advantage. Instead the individuals at a reproductive advantage are those who posses traits which favour early reproductive success. In this case the maximum lifespan potential of the species will decrease. This dynamic is depicted in Figure 4.

*Figure 4. Maximum lifespan potential change for biological populations which display senescence*

1.2.2 Implication for archaeological populations

The implication for archaeological human populations is that were they to have lived for as short a time as some estimates have given, then there has to have been a dramatic increase in maximum lifespan potential in a relatively short space of time (since about 1600 AD. in Western Europe). This would actually accord with the discrepancy seen between historically recorded populations, and archaeological ones. However, all observed modern human groups have a maximum lifespan potential which is around one hundred years (Weiss, 1973), and the maximum lifespan potentials calculated from brain size for hominids (Sacher, 1975) and early anatomically modern
sapiens is about the same (Cutler, 1978). Any theory which attempted to explain the perceived shortness of human life in the archaeological record would simultaneously have to explain how the maximum lifespan potential has not changed significantly without contradicting the evolutionary model given above, which makes maximum lifespan potential critically dependent on the ages to which past populations have lived. Even if one were in some way able to argue that maximum lifespan potential been subject to secular change in the near past, one would have to explain how most human groups, despite being largely reproductively isolated (e.g. indigenous Australians), have evolved similar maximum lifespan potentials in the same time.

It is credible that singular historical events have resulted in human groups in specific geographical locations who did not live very long, and in fact localised spatial and temporal variability is expected, but to see this picture throughout all humanity raises too many questions which contradict current biological theory.

1.2.3 Premature female reproductive senescence

There is one way in which it is possible for human groups to maintain an unchanged maximum lifespan potential, yet appear to have been very short lived in the past. It involves premature female reproductive senescence, or menopause. Age specific fertility decline is a common trait amongst mammals, an elephant's fertility falls to about half its peak value in very old age. In some mammals fertility can fall of altogether in near the maximum lifespan potential, so the organism reaches a post-reproductive phase which can be as much as 10% of the maximum lifespan. Only in humans, some inbred strains of laboratory rats, and possibly some toothed whales, is the post-reproductive phase extended. In female humans this is very nearly half of the maximum lifespan potential.

The problem for an evolved maximum lifespan potential is that the model we have so far examined relies on individuals with extended lifespans possessing a greater reproductive advantage over those without extended lifespans, the extended lifespans being heritable. If all reproductive spans are the same regardless of maximum lifespan potential there is no mechanism to provide any advantage, or disadvantage, to a lifespan which extends past the reproductive span.

There are two main hypotheses for the presence of an extended post-reproductive phase in female humans. The first is that humans have
undergone a sudden, and recent, extension in maximum lifespan potential to its present value which is a long time past reproductive phase. The problem here is how could this evolve? To extend maximum lifespan potential it is required that an organism must be capable of reproduction to a point in time fairly close to the lifespan potential. Of course, there is always the evolutionary null argument that there is no disadvantage to an extended maximum lifespan potential, hence it could occur.

The other hypothesis is the 'grandmother hypothesis'. This states that an extended post-reproductive phase is due to human females investing in their offspring, and offspring's offspring. This in effect would give a reproductive advantage in that in most societies child mortality is high. Were some factor to occur which would heighten the probability of survival into the reproductive phase, such as grandmothers looking after grandchildren, then those grandmothers would, despite not actively reproducing themselves, be at a reproductive advantage. Using data from the Ache foragers of eastern Paraguay Hill and Hurtado (1991) looked for evidence of extended female lifespan contributing to the survival of successors. Their conclusions were that there was no unambiguous evidence for the proposition, nor against. However, there were several complicating factors, such as rapid population growth amongst the Ache over the last three decades or so, whereas the grandmother effect may be much greater under stable population conditions.

Where this could help to give biological anthropologists the flexibility needed to have the very low age estimates seen in the archaeological record is if premature reproductive senescence became a feature of the human life history relatively early on with anatomically modern humans. Were this so then the maximum lifespan potential would be nearly fixed, as there would be no mechanism by which it could change. This means that if at some point selective pressure were increased it would be possible to see low lifespans occurring globally without affecting the maximum lifespan potential. Of course the phenomenon of premature female reproductive senescence would still have to be explained in evolutionary terms, but no more so than it does already.
1.3 Conclusions

There have been considerable doubts as to whether adult human age estimation is accurate, or precise enough, for archaeologists and anthropologists to examine the demography of past peoples from the hard tissue evidence with any worthwhile results in either the specific historical incidence, or in general terms.

An approach as to whether the estimates for ages at death for archaeological populations can be made through a knowledge of the evolution of senescence, but this really fails to make any unambiguous resolution of the question. This is because the ideas concerning the evolution of senescence which have been developed over this century clearly give consistent accounts for all biological species except humans. To make any evolutionary theory of senescence complete it must account for all biological species, including humans. For this area of study to progress it is necessary to have knowledge of the secular changes in human lifespan to see whether evolutionary theory needs adjusting or even revising. As a prior condition it is absolutely essential to produce an accurate picture of past human lifespans.

The rest of this thesis is concerned with the ways in which age at death for adult humans is estimated from hard tissue evidence.
2 Age-related Changes in Adult Human Hard Tissue

2.1 Skeletal changes in the adult human

Many physiological parameters change with age in the adult human. Systolic blood pressure rises, near vision decreases, audio competence falls, finger dexterity and handgrip strength falls with age (Dean, 1988). However none of these are used as a basis to attempt to estimate human chronological age simply because estimates of adult human ages are only really required when that adult human has ceased to live, and none of the above mentioned physiological parameters are extant. As dead humans tend to be examined in various states of preservation, physiological parameters, such as skinfold thickness, tend not to be used. Instead there is a reductionist consensus amongst workers towards developing age estimation methods from tissue with the highest probability of intact survival, namely the calcified tissue of the human skeleton and dentition.

The most widely considered parts of the adult human skeleton for age estimation tend to be the non-moving joints as these show progressive change with age, yet have no mobile function, thus tend not to be subject to extrinsic factors as joint change due to functional stresses, damage and wear, and osteoarthritis.

2.1.1 The pubic symphysis

Age changes have been noted in the innominate by anatomists as early as 1777 (Bonn, cited in Todd, 1920). The main focus at the time being a close examination and anatomical description of the functional muscle attachments of the innominate. However, gross age-related changes in the pubic symphysis did not go unnoticed (Aeby, 1858; cited in Todd, 1920), and the observation that there was a progression towards a fusion of the pubis was made by Cleland in 1889 (cited in Todd, 1920).
Figure 5. Generalised appearance of the pubic symphysis with increasing age

Partially redrawn from Bass (1987)
Todd (1920) describes a sequence of ten stages summarised below:

- **phase 1** (age 18-19) the symphyseal face is heavily marked by a series of horizontal grooves with well defined groves with no delimiting margin, and no definition of extremities.

- **phase 2** (age 20-21) the definition of the ridges is beginning to disappear as the groves begin to fill with a fine textured bone.

- **phase 3** (age 22-25) a progression from phase 2 where there is continued infilling of the horizontal grooves. The most characteristic feature in this phase is the formation of a bevel on the ventral border resulting in the obliteration of the grooves towards the ventral border.

- **phase 4** (age 25-26) the ventral bevelled area continues to increase with a concomitant decrease in the well defined ridged area. Dorsal margin completely defined and the beginning of definition of the lower extremity.

- **phase 5** (age 27-30) the lower extremity continues to become more defined. The ridge and furrow system is no longer the most prominent feature, and there can be sporadic formation of the ventral rampart.

- **phase 6** (age 30-35) the extremities become ever more defined, and the formation of the ventral rampart takes place. The face develops a porous texture.

- **phase 7** (age 35-39) the symphyseal margins begin to develop some slight lipping and the deposits of material on the ventral rampart which serve as muscle attachment become to look more pronounced.

- **phase 8** (age 39-44) lipping of the extreme edges of the symphyseal face is slight but the face itself gets smoother.

- **phase 9** (age 45-50) dorsal margin becomes more lipped, whilst the ventral margin is sporadically lipped.

- **phase 10** (age 50+) symphyseal face becomes progressively more eroded and the ventral margin becomes less distinct. Generally there are fewer distinctive features.

Todd's main methodological technique seems to have been defining an archetype for each of his phases. The changes which Todd points to were defined by an inductive dialectic formed from a visual similarity between pubes, age groups defined on the basis of that similarity, and a selection of the 'most typical' being selected to define the phase. This is crucially different to the definition of an age change prior to the observation of this age change for any known age, as seen in Gustafson (e.g. Gustafson, 1950), is necessitated by the sheer complexity of sometimes very subtle morphological changes on the pubic surfaces.

Todd was probably more aware of the individual variation seen in pubic development than some later writers. For instance, the development of the ventral margin on some individuals reaches a terminal stage when only part formed. Todd's advice is that if some of the later morphological features are
present, then the ventral margin is to be ignored, and only the later features are to be taken into account. Todd was also aware that his age estimates would get very much less accurate with older target individuals.

Todd was also very much concerned with the effects of pathological conditions on the rate of development of the pubic symphysis. He points to two individuals, one of whom was an alcoholic, the other a sufferer of tuberculosis with marked skeletal change, both of whom exhibited very much accelerated development. Several specimens of retarded development are also cited such as an example of a rachitic dwarf.

What Todd missed was that according to his own anatomical assessment one of the distinguishing features for individuals aged thirty and above was the differentiation between the ventral rampart and pubis proper. The ventral rampart is, according to Cleland (cited in Todd, 1920) the attachment site for the gracilis muscle, thus will be to some extent functionally dependent. A later study suggests the ventral rampart is the development of an osseous connection which replaces a prior periosteal connection for the gracilis muscle (Budinoff and Tague, 1990), which again could be in some way functional.

Todd also acknowledged unavoidable problems with the sample of some 300 individuals from which his developmental sequence was induced, and to which unknown specimens should be compared (his reference sample). He cites inaccuracies in the official Ohio State records for some of the earlier additions to the Western Reserve reference collection, although these became more accurate later. Also there seems to have been a problem with rounding up of ages to the nearest five year category. Todd attempted to eliminate specimens which were dubious, although he admits some were bound to get through. However, the means by which this was done conveys some circularity in Todd's arguments. Todd (1920) states that specimens were eliminated on the basis of gross differences between stated age, and the expectation for the state of the pubic symphysis for the stated age. Exclusion of these specimens from the reference sample, were the stated age in fact correct, would have the effect of limiting the observed variability in the development of the pubic region given age, thus giving a misleading impression of the coherence of Todd’s developmental sequence.

Todd's survey of the development of the pubic symphysis of European-American males was extended to African-American males, European-American females and Afro-European-American females (Todd, 1921). From
90 specimens of male Afro-European-Americans Todd concluded that there was no difference in the sequence of development, or in the rate of development of the pubic symphysis, between European-Americans and Afro-European-Americans. More unexpectedly, from a sample of 47 European-American females, and 22 Afro-European-American females, Todd (1921) concluded that the sequence of development in pubic metamorphosis was exactly that seen in males, but only certain features were slightly different, such as a delay in the consolidation of the ventral aspect of the bone. Todd (1921) also clearly states that there is no change discernible related to parturition in females, although later research (Stewart, 1957) revealed that much of the variability found by Todd in female pubic bones was probably as the result of childbirth, and in particular one set of pits underlying the dorsal edge were thought to be caused by parturition.

Todd's work stood unaltered for nearly 30 years. The next development came from Brooks (1955) who utilised Todd's descriptions for a collection of unknown age skeletons from California. Brooks noted contradictions between ages estimated from the pubic symphysis and those obtained from cranial sutures. Brooks suggested on the basis of a re-examination of the Western Reserve sample that Todd's age brackets attributable to each phase should be modified. A major innovation was Brooks' development of model pubes made up from casts taken from pubis considered to display features 'typical' of each pubic phase. Brooks also suggested that Todd's extreme deviations from his idealised sample were probably just alternative expressions of an otherwise normal development in the pubic symphysis.

The next development in pubic symphyseal age estimation came in 1957 with work by McKern and Stewart. Their revision of Todd's system was based on a sample of 450 skeletons from young males killed in the Korean war. Like Brooks they devised a series of casts taken from moulds of specimens they thought best typified a particular phase. Their major innovation was identifying from Todd's descriptions nine separate and individual changes which occurred with age in the pubic symphysis. These they categorised into three variables which they then treated as being conditionally independent (McKern and Stewart, 1957). The abstracted variables were: the dorsal margin, the ventral rampart, and, the symphyseal rim. Each of these they categorised into five stages and cross-tabulated by age.
This way of examining pubic symphysis had the advantage that changes in the morphology did not have to happen in any particular sequence, as required by Todd. Thus as pointed out by McKern and Stewart (1957) any pubis which did not closely correspond to the descriptions could be more easily classified.

The reference sample employed by McKern and Stewart was composed of young males which could hardly be expected to represent the total age variability in any normal male population. However, it has to be borne in mind that the work was carried out explicitly by the American armed services with the intention to assist in the identification of soldiers killed on the battlefield, and as such cannot be reproached on the grounds of specificity to this narrow population.

Nemeskéri et al. (1960 cited in Acsádi and Nemeskéri, 1980) studied the pubes of 105 autopsied individuals aged between 23 and 93 years. There were 61 males and 44 females. This later formed one of the elements for Acsádi and Nemeskéri's (1980) complex age estimation method (see also Sjøvold, 1975). Their method closely follows that of Todd (1920, 1921) in that separate morphological traits were not separated as quasi-independent variables. However, Acsádi and Nemeskéri (1980) only used five separate developmental stages. They were later criticised by Brooks and Suchey (1990) for not using casts to ensure consistency, using an unspecified number of undocumented skeletons to supplement their reference sample and concentrating solely on early and late developmental features leading to many individuals falling between Acsádi and Nemeskéri's five stages.

Brooks (1955) when comparing 103 male and 82 female skeletons had shown that Todd (1921) underestimated sex differences in pubic development. This conclusion was largely agreed upon by Gilbert and McKern (1973) who expanded their system to include a female standard based upon 120 skeletons. This modification (Gilbert and McKern, 1973) was tested by Suchey (1979) on a sample of 11 female pubes using repeated measures by a team of physical anthropologists. The results were disappointing in terms of reproducibility, only 51% of estimates being within fifteen years of the documented age.

Meindl et al. (1985) evaluated the McKern and Stewart, and Todd, system using 64 male and 32 female documented specimens from the Hamman Todd collection (see 11.1.8). They found that Todd's system seemed to give the
better results and recommended that this should be used with few modifications.

Katz and Suchey (1986) also decided to abandon the approach pioneered by McKern and Stewart and go back to Todd's system of described archetypes, despite earlier work which broadly supported the McKern-Stewart system (Owings-Webb and Suchey, 1985). This was because the three components described by McKern and Stewart were interdependent; for instance the ventral arc of the pubic symphysis cannot develop without the face of the symphyseal joint changing. Katz and Suchey used pubic bones from 739 males from the autopsy cases of Los Angeles coroners courts, and made up a system comprising six phases which were groupings of Todd's original ten phases.

Many of the problems associated with age estimation using the pubic symphysis arise as a result of great natural variability and a certain inability on the part of humans to describe precisely complex changes in three dimensional surfaces. It is impossible to correct the former source of error, however the latter has been approached by both archaetypical seriation and the division into identifiable age changes. Neither has proved to be wholly satisfactory and it may be better in future work to classify pubic symphysis by some method of image recognition and neural network which would be able to classify these sorts of morphological changes reliably without the need to specify some preordained scheme of changes which is reliant on human perception to devise.

2.1.2 Cranial suture closure

During development the cranial vault consists of four bones, the occipital, two parietal bones, and a frontal bone. In some fully mature individuals these fuse completely to form a single dome. In most fully mature individuals there is some evidence for a continuous process of fusion of these bones. The interface between each bone is called a suture. In addition each bone is a sandwich of an outer table of hard compact bone, a spongy medullary layer, and an inner table composed of hard compact bone. Each of the tables can be considered as a separate bone, the sutures between the inner (endocranial) and outer tables (ectocranial) fusing semi-independently. The fusion, and eventual obliteration of these sutures, have been employed since the earliest years of anthropology as a means of providing information on human adult age.
Masset (1989) refers to Vesale as early as 1542 being the first recorded worker to notice a relationship between age and state of the cranial sutures. The first easily available reference in English is from Dwight (1890) who gives descriptions of 100 crania from adults of various ages, and pointed to a difference in opinion between the anatomists and medico-legal practitioners as to how variable the rate of closure is. It appears that even at this early date, age estimation based upon cranial suture closure was being used by the medico-legal establishment as a commonplace practice, so presumably had been in use for some time.

Parsons (1905) categorised 82 crania from the dissecting rooms of St. Thomas's Hospital by age into ten year categories. He then described the condition of both endocranial and ectocranial sutures for each case within its age class. Parsons came to the conclusions that were there no internal obliteration of the cranial sutures then the individual was probably under thirty years. After thirty years of age there would be partial obliteration of the coronal and sagittal endocranial sutures. All endocranial sutures in most
individuals would be obliterated after fifty years, and certainly after sixty years of age. Parsons also makes the observations that the lower half of the coronal endocranial suture is the first to fuse, which is quickly followed by the outside of the same suture. The sagittal suture closes internally in the region of the obelion, followed by the anterior portion, the posterior inch or so sometimes never closing fully; the lamboid suture closes later than either the sagittal or coronal sutures.

Todd and Lyon (1924) followed up Parson’s suggestion that endocranial suture closure was less erratic, and thus a better indicator of age than ectocranial suture closure. They used a sample of 307 European-Americans, and classified each suture into four stages dependent upon how much of the length of the suture had closed. They did not differentiate between contiguous suture closure and non-contiguous suture, calling, for instance, a suture which was half closed (stage 2) ‘half closed’ whether it was closed halfway along its length from one end, or half closed in isolated areas. They eliminated ‘uncharacteristic’ crania from their investigations, by which they meant crania which did not fit into the sequence of development observed in their reference sample. Rather than falling into the same circular argument seen in Todd’s earlier work on pubic symphysis fusion (1920; 1921) they tried to find some independent criteria by which specific crania could be deemed to be anomalous. To this end they were largely unsuccessful, but did find what they termed ‘an anthropoid strain in the pubic symphysis’ which was apparently a good indicator of anomalous crania, but this was only seen in some individuals in the third decade of life. Another indicator of anomalous endocranial suture closure was based on suggestion made by Bolk (cited in Todd and Lyon, 1924), although Todd and Lyon are not explicit about the exact nature of this marker. They found that after elimination of all anomalous crania they were left with a small sub-sample which closely fitted an induced pattern of endocranial suture closure. Generally Todd and Lyon agreed with Parsons (1905), in that the obelion fuses is mostly completely fused at the age of twenty-nine, the rest of the sagittal suture being fused by the age of thirty-five. The coronal suture would be fused by the age of thirty eight, and the lamboid suture the final one to close at about the forty-five. Todd and Lyon did not restrict their observations to the sutures of the upper cranial vault, they also looked at the masto-occipital suture, the spheno-temporal suture, and the squamous suture, arriving at the conclusion that there was little
fusion to be seen in any of what they termed the 'circum-meatal' sutures. The final sutures examined by Todd and Lyon were the sphenoparietal suture and the sphenofrontal suture. The sphenoid-parietal suture was found to fuse quickly until the age of twenty-nine, then slow down with gradual progress until forty-six, then slow right down until the age of sixty-five, when complete closure could be expected. The sphenofrontal suture was found to close slowly until the age of twenty-six when it closes rapidly until the age of thirty, when closure slows until complete fusion can be expected at the age of sixty-five.

It has to be said at this stage that the models induced by Todd and Lyon represent average closure patterns for a sample from a population. This to some degree ignores the fact that what they were actually examining from the sample appears to have been individuals with various suture closure stages. For instance, the statement of slow and steady suture closure would actually refer to a small number of sutures from that age group becoming fully fused relatively rapidly. A more appropriate model may be one of comparatively instantaneous fusion at a random point in any age class for any given individual. More interestingly Todd and Lyon found that delayed, or accelerated fusion in one suture would very often mean delayed, or accelerated, fusion in the other sutures.

Todd and Lyon (1925a) extended their work started on European-American males to African-American males. They found few differences in the rates of development from a sample of 120 African-American males to 307 European-American males, but found that almost three times as many African-American males had to be excluded from their study.

Brooks (1955) tested Todd’s ectocranial and endocranial age estimation procedures, and compared them to his (Todd’s) procedures for pubic symphyseal age estimation. Brooks used some 80 skeletons from the Western Reserve University Collection, and found that estimates derived from cranial suture closure gave estimates of age for individuals over 40 years which were significantly too low when compared to the documented ages for these individuals.

Powers (1962) used Todd’s cranial age estimation system on 111 crania from (mostly) young documented individuals. Powers found significant errors
in estimation, although it was noticed there was some evidence for rounding to the nearest five year interval amongst the documented specimens.
Acsádi and Nemeskéri (1980) modified Todd and Lyon's (1924) system for cranial suture age estimation. They divided the coronal and lamboid sutures into three equal parts for each side and awarded each segment a score between zero and four. Zero meant the suture was completely open, one was awarded to incipient closure, two to closure in progress, three to advanced closure, and four to complete obliteration. The sagittal suture was divided into four parts, the same scores being awarded as to the other two sutures. These observations were then summed to produce what Acsádi and Nemeskéri termed a cranial suture stage, estimates of age being made by regression models derived from a reference sample of 402 cadavers from the Budapest University's Institute of Forensic Medicine. They found it not necessary to produce separate models for males and females, and used only European individuals in their reference collection.

Perizonius (1984) used this cranial classification to give age estimates for 256 European individuals in the University of Amsterdam's skeletal collection. Like Acsádi and Nemeskéri, Perizonius found no significant differences between males and females. Perizonius found better estimates could be made if two models were employed; one for those under 50 years, and another for those over 50 years of age. The problem with this is that in order to know which model to use the age must be known, at least in terms of below or above 50. As this is unknown it is not possible to know which estimate to use in the first instance.

Meindl and Lovejoy (1985) refined Acsádi and Nemeskéri's technique, apparently making it more accurate and easier to replicate. They selected ten specific suture points from the cranium, and observed a single centimetre along each, awarding between zero and three points dependent on the degree of closure (zero completely open, three completely obliterated). These they measured for 236 crania from the Hamman-Todd Collection. Their best estimates could be made from only five of the original ten locations all of which were on the lateral anterior portion of the cranium.

The relationship between premature obliteration and pathology was investigated by Gladykowska-Rzeczycka (1979) who found that six individuals from a sample of 400 individuals from Czrna⁵ in Poland displayed unusual obliteration of the cranial sutures. An individual showing clear indications of

---

⁵ Czrna is a mixed period (Neolithic to Medieval) cemetery in Grodzisk, Poland.
scaphocephaly had complete obliteration of all cranial sutures, as did the remaining five skulls of children suffering from hydrocephalus. The problem here is that for the adult individual we have no idea how old they were, thus no idea of whether the suture obliteration is anomalous or not.

Zivanovic (1983) studied endocranial and ectocranial suture closure in 263 East-African Bantus and 309 European-African crania. Zivanovic found no differences between males and females, nor any differences between East-African Bantus and European-Africans in the rate of cranial suture closure. However it was found that a large number, particularly in the 41-50 age group, had asymmetric suture closure. The implication of this being that age estimates made from cranial sutures would be very unreliable were only half the cranial vault available.

Reichs (1989) describes a case of a male corpse where unusually early closure of the sagittal suture, both internally and externally, led to the whole cranial vault taking on the morphological appearance of a female. Reichs points out that had other evidence not confirmed the identity of the individual then assigning the skull to the post cranial skeleton would have proved impossible.

Cranial suture closure presents a lesser problem in terms of measurement than the pubic symphysis, in that the cranial sutures are roughly linear features on the human skull, so it is easy to judge the extent to which a suture is closed, and should be the case that inter-observer error will be small. The main problem could be that the cranial sutures do not fuse gradually. In the individual it is possible that a suture fuses in short bursts of activity, followed by quiescent periods which can continue for some unspecified period until the next spate of suture fusion. This would have the consequence that more refined systems of measurement, such as measuring the length of suture closure, or conversely the remaining length of open suture, will have little effect upon age estimation as a small increment in measurement will not necessarily relate to a small increment in time.
2.1.3 The auricular surface of the sacro-iliac joint

The joint between the sacrum and ilium can be seen to change in two ways with age. Firstly, there are structural changes to the surface morphology of the joint, second, the joint may fuse. Sashin (1930: cited in Iscan and Loth, 1989) first noted regular changes in the surface and the fact that males were far more vulnerable to fusion of the joint than females. Weiss (1954: cited in Iscan and Loth, 1989) noted the increase in height of the cranial sacral elevations until mid-life, although no systematic age estimation technique seems to have been formulated.

Figure 7. Generalised progress of the auricular surface with age

A system for estimating age from the sacroiliac auricular surface was devised in 1985 by Lovejoy et al. (1985a). They used 500 specimens from the
Hamman-Todd collection, 250 individuals from the Libben population, and 14 autopsy cases. They classified morphological change into eight categories particularly focusing on topography, marginal lipping and porosity. The method was supposed to be sex independent, although Lovejoy et al. were aware that certain features such as preauricular sulci could be very pronounced on females, and could lead to poor age estimates, and that the age changes were more difficult to interpret than those for the pubic symphysis.

Murray and Murray (1991) conducted independent tests of Lovejoy's auricular surface age tabulation using 189 individuals of mixed race and sex from the Terry collection. They found that an individual's sex had no effect on the error in estimation for that individual, but instead found that the individual's geographical origin did. Later findings showed that the main source of variation in auricular surface morphology was an individual's age, and that the reason geographical origin had seemed to affect the estimate so much was very different age profiles of African-Americans and European-Americans in the Terry collection. They concluded that the variability of auricular surface estimate was so high that Lovejoy's et al.'s method (1985a) could not give useful results on its own.

Again, as with pubic symphysis, the major problem is that age changes in the auricular surface are a complex series of three dimensional morphological changes, not amenable to simple reproducible measurement. The approach taken by most workers seems to have been one of establishing archetypes with which to compare unknown specimens, splitting into individual age-related traits has not as yet been attempted.

2.1.4 The sternal end of the fourth rib

Kerley (1970: cited in Iscan and Loth, 1989) noticed a continuing metamorphosis in the shape of the sternal end of human ribs throughout life due to continued ossification of the chondrocostal cartilage. Iscan, Loth and Wright (1984a; 1984b) devised a method of age estimation based upon assignment of the fourth right rib to one of eight phases. They based their system upon three components: the shape of the pit into the end of the rib endplate, the depth of that plate, and the configuration of the rim walls (Iscan

---

6 See 11.1.8
7 See 11.1.1
et al., 1984b). This was calibrated against a reference sample of 93 European-American male autopsy
cases. Component 1 was pit depth, and was measured in mm with a vernier calliper held orthogonal to the cranial-caudal axis of the rim walls. The cranial-caudal axis was avoided because occasionally long projections of bone form on these sides of the rib. Component 2 was pit shape. The pit starts as a shallow amorphous indentation in adolescence, grows increasingly v-shaped into early adulthood, gets u-shaped with fairly thick walls into middle age, the walls then get thinner by about the age of 55, finally the pit begins to loose some of its definition into late middle age. Components one and two are closely linked in that they are essentially an expression of the formation of a pit in the extreme endplate of the fourth rib. Component three (the rim shape) starts with a smooth regular rim in childhood, by 25 there appears the beginnings of slight scalloping and the formation of a more defined wall. At 50 the rim is becoming sharper and boney projections form which are most pronounced on the cranial and caudal margins, becoming more brittle and friable by 60 years of age.

Figure 8. Generalised progression of the sternal rib end with age

Partialy redrawn from Bass 1987
In another paper Iscan et al. (1984a) describe a different approach to categorising the same changes. In this paper they categorise the above components together into a series of nine phases, this time expanding their reference sample to 118 European-American male autopsy cases.

- **Phase 0:** smooth articulator surface with regular rim and rounded edges.
- **Phase 1:** the beginnings of an amorphous pit develop. In some cases scalloping may begin to appear on the rim.
- **Phase 2:** the pit in the rib end plate takes on a v-shaped formed by the anterior and posterior walls. The walls are thick and many manifest some scalloping.
- **Phase 3:** the pit deepens and starts to take on a u-shape. The walls are still fairly thick. Some scalloping may be apparent but the rim is becoming more irregular.
- **Phase 4:** Pit depth continues to increase. The walls become thinner, but the edges are still rounded. The rim is yet more irregular, and there is a noticeable decrease in bone density.
- **Phase 5:** Pit depth is little changed from Phase 4, but adopts a considerably widened u-shape. The walls show further thinning, and the edges begin to get sharp. There is an ongoing decrease in bone density.
- **Phase 6:** The pit is now a deep wide u-shape. There is further sharpening of the rim, and loss in bone density. Boney projections from the superior and inferior borders are frequent.
- **Phase 7:** There is little change in pit shape or size, the walls of which are thin and fragile. There is a marked decrease in bone density with significant deterioration in structural quality.
- **Phase 8:** In this terminal phase there is little change in pit size or shape. Sometimes the pit floor can have small projections of bone projecting from it. The walls are extremely thin and fragile, with regular boney projections. The bone as a whole is very light and shows great loss of density compared with its newly mature form.

Later the phase system was extended to European-American females (Iscan et al., 1985), when, as expected, there were found to be significant differences between male and female European-Americans.

Differences between African-American and European-Americans were examined with 49 African-American and 108 European-American male autopsy specimens, and, 14 African-American and 83 European-American autopsy specimens (Iscan et al., 1987). They found that African-American males progressed faster through the stages than European-American males, and few differences between European-American females and African-American, although the sample size in the latter case was very small. Subsequent work on African-American and European-American differences downplays the accelerated African-American development. Russell et al. (1993) came to the same general conclusions as Iscan et al. (1984a; 1984b) relating the development of the rib end to age, but concluded there were no
The major age change in the sternal rib end is the ongoing, and gradual, calcification of the chondrocostal cartilage, which is relatively easy to measure on an absolute continuous scale, and provides the major source of information about the stage of development of that rib. Because the age-related change is so amenable to measurement inter-observer classifications for any given rib tend to be highly consistent (Iscan and Loth, 1986a; Iscan and Loth, 1986b).

2.1.5 Bone density and internal structural changes

The internal trabecular structure of the epiphyseal ends of the long bones has been examined in relation to age. The general pattern throughout the long bones of the skeleton is one of dense and regular internal patterning after maturation of the bone, to a marked loss of internal bone mass and break-up of trabecular structure. Cortical thickness also lessens with increasing age. Describing these features for the proximal humerus and proximal femur Acsádi and Nemeskéri (1980) examined 105 skeletons of documented age and sex. They described changes observed from x-ray 

radiographs of both these sites in terms of six phases. For the proximal humerus:

- phase 1 (34-48). The apex of the medullary cavity is well below the surgical neck.
- phase 2 (50-55). Medullary cavity extended to about a quarter of the distance between the neck and epiphyseal line. Trabecular system looking more fragile.
- phase 3 (56-54). Medullary cavity may reach the epiphyseal line. Individual trabeculae may become thicker.
- phase 4 (54-58). Apex of medullary cavity extends beyond the epiphyseal line. Major trabecular system now shows small gaps.
- phase 5 (59-64). Gaps of up to five millimetres develop in the major tubercle. Trabecular structure is now fragmentary.
- phase 6 (58-65). Little change from phase 5, except trabecular system is now more rarefied.

The six stages in the spongiosa of the proximal femur are:
• phase 1 (31). Apex of medullary cavity well below lesser trocanter; trabecular texture thick, individual features difficult to distinguish.

• phase 2 (41-47). Apex of the medullary cavity reaches, or is beyond, the lesser trocanter. Trabecular structure begins to rarefy at the epiphyseal borders.

• phase 3 (50-55). Apex of the medullary cavity reaches the upper limit of the lesser tricanter. Trabecular structure in the medial part of the femoral neck is marked.

• phase 4 (54-58). Apex of the medullary cavity extends beyond the upper limit of the lesser tricanter. The diaphyseal-epiphyseal borders begin to grow less distinct.

• phase 5 (61-66). Few elements of the trabecular system remain in the neck. Cavities start to form in the trabecular system of the epiphyseal head of the greater tricanter.

• phase 6 (64-72). Large cavities have formed in the greater tricanter. The cavity in the medial neck has merged with the medullary cavity. Only small amounts of trabecular structure remain along the walls of the cortex, which in itself is thin and insubstantial.

In a more recent study Walker and Lovejoy (1985) examined changes in the internal structure of the proximal femur and clavicle from 130 individuals from the Hamman-Todd Collection using x-radiographs. They categorised bone density and trabecular patterns into eight phases for both bones.

Walker and Lovejoy realised that young females would look older than their male counterparts as at full maturation the cortex of the proximal femur and cortex of the clavicle would be thinner, so it was decided to treat females separately to males. However, there were no sex-specific morphological features. In a second examination of some 105 individuals (61 males, 44 females), it was found that the clavicle gave the best age estimates, with the male sample the clavicle was moderately better than the femur, but for the female sample the clavicle was very much better.

As bone density is usually measured by x-radiograph, or dual photon absorptiometry (Alhava and Kerjalainen, 1973; Bennike et al., 1993), the response is a function of the bone density and bone thickness. So long as the equipment is accurately calibrated, and a large enough volume from each bone is sampled, then inter-observer variability should be small compared with techniques which require a classification based upon morphological changes, where the experience and judgement of the observer play a significant role.
2.1.6 Discussion of skeletal age indicators

There are two main sources of error when looking at morphological criteria to estimate a continuous variable such as age. The first is the inherent variability between individuals of any given age, and is a function of the physical biology of the organism. The second is inter-observer error for the same entity, which in the case of pubic symphyseal, or auricular age estimation, can be as much as a factor in any misidentification of age as inter-individual error. In fact large inter-observer error can obscure inter-individual variability. The approach taken by Todd (1920), Meindl, Lovejoy Mensforth and Walker (1985), and Katz and Suchey (1986) of defining an archetype which, in effect, loosely maps onto an age group, relies upon all pubic symphyses developing the same features in the same way. Any pubis not manifesting those features considered typical, or displaying in an untypical sequence, cannot be successfully classified. The other approach is to separate out the various traits which are thought to be age-related, each trait being examined individually as with McKern and Stewart (1957), then atypical cases can be satisfactorily dealt with. However, as the material is being examined is a complex series of three dimensional morphological changes then having available a set of 'typical' casts, which reflect the features of the archetype should make the task of matching a morphological feature from an individual of unknown age to an order within the archetype range easier than evaluating each trait individually. This is possibly why McKern and Stewart's (1957) method for systematically evaluating individual traits seem to have been largely superseded by archetype-based classifications.

Not all skeletal age-related changes rely upon the correct classification of morphological traits. Certain features of the sternal rib end, such as pit depth, are amenable to direct length measurement (Iscan et al., 1984a; Iscan et al., 1984b), and recent approaches to cranial suture closure have been systematised in such a way as to make classification a very simple matter (Meindl and Lovejoy, 1985). Features such as bone density should be easy to consistently quantify with relatively inexpensive equipment, but emphasis has been given to the more difficult aspects of bone development such as trabecular organisation (Walker and Lovejoy, 1985).
2.2 Age-related changes in the teeth

The teeth are heavily calcified organs comprising four main tissues (Beynon, 1991):

- Dentine is a highly mineralised, quasi-vascular tissue having the processes of odontoblasts housed in its main structure. The dentine forms an armature comprising of tubules, a series of parallel tunnels running from the enamel to the pulpal chamber.

- The pulpal chamber, a cavity in the dentine which is filled with ground substance and the structures for the vascular supply of the tooth. Lining it is a layer composed of the cell bodies of the odontoblasts. Other cells inhabiting the pulp chamber are fibroblasts, undifferentiated ectomesenchymal cells and macrophages.

- Enamel is the hardest, most heavily mineralised tissue in the human body and covers the coronal half of the dentine to form the hard occlusal surface.

- Cementum covers the apical half of the dentine and is similar in composition, but not structure to the dentine. This forms a base for the attachment of the tooth into the alveolar bone.

*Figure 9. The main features of the human anterior tooth*

Partially redrawn from Bass (1987)
Although thought to be relatively quiescent organs, the teeth manifest many age-related changes, and as such have been employed to provide markers for adult human age estimation.

Adult age estimation using dental indicators mainly revolves around six indicators: occlusal attrition, secondary dentine formation, cementum build-up, recession of the periodontum, resorption of the calcified tissue surrounding the apical foramen and root dentine translucency (Altini, 1983; Costa, 1986; Kilian, 1989; Saunders, 1965; Xiaohu et al., 1992). Others are mentioned, such as colour changes, etching characteristics, microscopic surface changes, loss of water from dentine, dentine hardness (Bang, 1989); and changes in chemical composition of the enamel caused by adsorbed dietary ions (Noble, 1974).

Figure 10. The main age-related features seen in the human anterior tooth

(a) A=0  P=0  RDT=0  RR=0  C=0  SD=0
(b) A=1  SD=1  P=1  RDT=1  RR=1  C=1
(c) A=2  SD=2  P=2  RDT=2  RR=2  C=2

The age indicators marked are scored on the basis of Gustafson's (1950) descriptions and are:
- occlusal attrition = A
- secondary dentine = SD
- periodontal repositioning = P
- root resorption = RR
- cementum build-up = C
- root dentine translucency = RDT

Partially redrawn from Bass (1987)
2.2.1 Enamel and dentinal attrition

Of all the dentally based age markers attrition is the most commonly used. Attrition can be divided into two main types: occlusal attrition, ongoing wearing down of the hard biting surfaces of the dentition, and approximal attrition, the propensity of adjacent teeth in the dentition to wear against each other in the course of normal living (Whittaker, 1992). The former, occlusal attrition, has received most attention from forensic scientists and anthropologists.

Although occlusal attrition formed one of the six components of Gustafson’s dental age estimation technique (1950), the first formulation of a method based solely upon occlusal attrition was by Brothwell (1963; 1981). Brothwell used a reference population of a large, but unspecified number individuals from various British archaeological sites dating from the Neolithic to the Medieval, and produced a set of charts displaying the amount of exposed dentine for the three molar teeth. Workers who wished to estimate age for any given adult human simply examined the molar teeth, noted the area of exposed dentine, and compared this to Brothwell’s chart. Brothwell was keenly aware of the major criticism of this practice, namely, that the rate of attrition is utterly dependent upon contact with hard particulate matter in the diet, and that the ages for the reference sample had been estimated by examination of the pubic symphysis. Lavelle (1970), studied differences in occlusal attrition between 19th Century British, 9th Century British, West African and Oriental dentitions in a sample of 359 individuals, and a further 300 dentitions from a modern British population. The finding that the 19th Century British sample was significantly less worn than the rest, suggested that the main controlling factor was diet, although in individual cases conditions such as bruxism, and dental disease, could serve to accelerate occlusal attrition.

A very ingenious attempt was made to bypass this limitation by Miles (1963a; 1963b), who estimated ages for a sample of 32 Anglo-Saxon skulls from Breedon-on-the-Hill in Leicestershire. Miles started by examining adolescents from the sample. These individuals had wear displayed on the premolar teeth which erupt some accurately known time before the molar teeth. By observing the difference in wear between premolars and molars Miles was, with knowledge of the time between eruption, able to estimate a rate of enamel loss. By assuming a constant and uniform diet for the sample,
Miles was able to apply the previously estimated rate of wear to the rest of the sample, thus estimate ages for all 32 individuals. Miles was aware of drawbacks such as increased power of the mandibular musculature would tend to accelerate molar occlusal attrition, but found this effect cancelled out by posterior molars seeming to wear at a slower rate. Miles then went on to apply this technique to 150 skulls from the same site and period with credible results, although the true accuracy could not be known.

Takei (1970, 1984) used 1000 dentitions from modern Japanese people to devise a method of age estimation for application to the modern Japanese population. Takei's method consisted of classifying each tooth into one of four classes of occlusal attrition. From this Takei was able to arrive at an estimate of age as a displacement from the mean age of the reference population, each tooth giving a contribution to the overall age estimate. Takei took account of teeth which have occlusal restoration and wear at a lower rate, teeth which had caries on the occlusal surfaces, and teeth which were missing. Takei claimed a high correlation between known age and estimated age ($r = 0.81$).

Hongwei (1989) used 880 dentitions from a population based in the Peoples Republic of China to produce 10 equations, one for each age group, which gave age as a function of occlusal attrition for selected teeth. Occlusal attrition was classified into six stages, each tooth field having its own classification system, and each of the age categories employed a different set of teeth. Hongwei claimed a high degree of accuracy, but failed to address the problem of a certain circularity as it would be necessary to have accurate prior knowledge of age before the appropriate equation could be selected.

Dreier (1994) used a system based on molar occlusal attrition which featured 25 stages of wear for each tooth, using 107 dentitions, 76 from excavated archaeological sites, and 31 from unknown sources. Estimates of age derived from skeletal methods applied to the reference sample to calibrate his stages for molar wear. Dreier then applied the calibrated stages to 143 archaeological dentitions from Arikara, a North American Plains Indians site. Although Dreier made no claims about the accuracy of his age estimates the problems with Dreier's work are that most of the 25 stages identified are virtually indistinguishable from each other, and calibrating the occlusal attritional stages against age estimates derived from skeletal estimates.
makes the whole technique utterly dependent upon the accuracy of the skeletal age estimates.

In recent work which heralded a return to the methods pioneered by Miles (1962; 1963a), Mays et al. (1995), examined some 293 individuals from the Romano-British site of Poundbury, Dorset. Mays et al. found that first and second molar wear rates were approximately uniform throughout life in the Poundbury sample, but the third molar was highly variable. Mays et al. also found that mean crown heights were closely related to Brothwell's (1981) wear stages, however, Brothwell's stage 2 was found to be prolonged.

Occlusal attrition is fairly simple to observe in both intact and sectioned teeth, although no studies have been conducted to examine the issue of inter-observer error. A more pressing problem has always been the diet dependency of the rate of occlusal attrition. Miles' (1962; 1963a) offers a way around dietary dependence, but does assume a constant diet for any group of individuals one may wish to examine. It is quite likely that idiosyncratic variation in diet and masticatory activity within any human group would equal inter-group variability.

2.2.2 Secondary dentine formation

The continuing apposition of dentine after a tooth in the dentition has become fully mature has been observed since the 1930's (Fish 1932, cited in Philippas and Applebaum, 1967). Philippas and Applebaum (1967; 1968) describe two different types of secondary dentine, ‘regular’, and ‘irregular’, observed in maxillary lateral incisors and maxillary canine teeth. Regular secondary dentine they describe as forming on all walls of the pulp chamber, eventually occluding the whole pulp chamber. Irregular secondary dentine is described as forming rapidly under specific sites of damage and irritation, and as a response to functional pressure in the dentition. Irregular secondary dentine formed most rapidly on the lingual face of the coronal part of the pulp chamber. In the intervening years much work has been done on the causes of secondary dentine. Van Rensburg (1986; 1987) found no gradual build-up of secondary dentine in the absence of attrition or caries, or some other physical insult to the tooth organ. However, Wennberg et al. (1982) did find evidence for a gradual, non-pathological dentine build-up in mammals other than humans. Fahim and Messer (1986) found a marked secondary dentine response to caries in rats. In a major article Stanley et al. (1983) considered secondary dentine as one of three responses of the pulp to physical insult of
the tooth. They examined reactive sclerotic dentine, where the tubules underlying any dentine exposure become hyper-mineralised. Dead tracts, which are tubules over which a thin layer of reparative dentine has formed effectively sealing off the pulp from any penetrating bacteria; and reparative secondary dentine. Stanley et al. found reparative secondary dentine was a more likely response to attrition than either dentine sclerosis or dead tract formation. Dentinal sclerosis was the most likely response to caries and restorations. They also found that age-related dentine could suppress any reaction by the pulp to damage if it preceded the insult. Other effects which can have a marked effect on the rate of secondary dentine formation include whether an individual has been subject to renal transplantation. Nasstrom et al. (1993) found patients which had kidney transplants had about four times as much regular secondary dentine as individuals who had not. There was some question as to what was actually causing this but it was felt increased levels of corticosteroids were responsible for the increased calcification, and it therefore could be that any medical treatment where these substances were employed may produce the same effect.

Again, secondary dentine has mostly been examined in the context of being a single component of a multivariate age estimation system, as with Gustafson's method, but some attempts have been made to investigate the direct relationship between age and the amount of secondary dentine.

Moore (1970) used 250µm longitudinal labial-lingual sections from an unspecified, but reasonably large, number of individuals. Moore then measured the ratio of the pulp chamber diameter to the diameter of the tooth. This was then regressed on the known age in years for each individual, and found, as expected, to decrease with age, the correlation coefficient being -0.62. Moore was aware of the effect of lesions on the build-up of secondary dentine, and countered this effect by choosing teeth with no evidence for caries. Padayachee and van Wyk (1988) employed 50 teeth, and measured the amount of secondary dentine build-up at its greatest depth. This was then regressed against the known age, the correlation being 0.59. Padayachee and van Wyk were aware of the effect on the rate of development of secondary dentine of lesions and other physical insults to the tooth so they selected teeth with no sign of caries or other dental pathologies.

Solheim (1992) used a reference sample of 1000 teeth, consisting of approximately 100 of each anterior tooth locus, and aimed to examine the
number of ways in which secondary dentine could be measured. Solheim used longitudinal labial-lingual half sections and measured both incremental stages (i.e. as in Gustafson) and pulp chamber to root width ratios at the tooth cervix, the coronal quarter of the root, at the mid-point of the root, and, the apical quarter of the root. He took mean values of these measurements and regressed them against age for each anterior tooth in the dentition, producing correlations between 0.59 and 0.71, dependent on tooth. He also found that controlling for sex made no difference to the regression overall, but teeth from men were generally wider in the vestibulo-lingual direction. No differences were found from those teeth which were extracted due to caries, this was attributed to the fact that root dentine was under examination, and only coronal dentine would be affected by reparative secondary dentine. It was concluded that there was a fairly reliable build-up of age-related secondary dentine in the teeth of the human dentition.

Drusini (1993) in an attempt to use a non-destructive dental age estimation method employed low energy x-radiograms of teeth in the labial-lingual plane. Drusini used 68 premolars and 98 molars of known age and sex, and measured the ratio of the coronal pulp cavity length and the coronal length; that is the length from the occlusal enamel to the cervix, and from the top of the pulp chamber to the cervix. It was found that this measure of secondary dentine had a 0.88 correlation with age for premolars and 0.73 and 0.81 for female and male molars, respectively. This method was recommended for archaeological specimens where other age-related traits may have been destroyed by post-mortem degradation processes. Drusini avoided the problem of caries and other insults to the structure of the tooth by using only pristine examples.

Secondary dentine is difficult to quantify in that it is only visible as an area in sectioned teeth, or has to be quantified indirectly by measuring the decrease in pulp chamber diameter. In the former case the division between primary and secondary dentine may be obscured in some teeth.

2.2.3 Periodontal recession

There is a tendency throughout adult life for the attachment apparatus of the teeth to readjust to continuously adapting occlusal requirements. One of these changes has been called periodontal recession (Gustafson, 1950; Johanson, 1971), which is where the small dimple in the cementum surface
caused by the attachment of the periodontal ligament can be seen to gradually move down the root of the tooth with increasing age.

Solheim (1993b) examined this attribute in relation to age in a sample of 1000 teeth from a modern Scandinavian population. He measured the distance from the cemento-enamel junction to the most gingival periodontal fibres on the labial, lingual, mesial and distal surfaces, and found only a poor relationship to age. However, he included specimens which were known to have some degree of periodontal disease to simulate the effect of an unknown degree of periodontal affectation in a real population. Solheim concluded that despite a poor relationship to age periodontal recession may be a valuable supplementary piece of information for age estimation. Lamendin et al. (1992) used periodontal recession as one of two age estimation variables, the other being root dentine translucency. Lamendin et al. found just these two measures of age to be as accurate as any other technique, with the exception of younger individuals, when applied to a sample of 45 teeth from 24 forensic cases.

More recently it has been realised that recession of the periodontal ligament has an underlying cause which is the propensity of the human dentition to erupt continuously throughout adult life (Darling and Levers, 1975). Levers and Darling (1983) examined some 34 Romano-British, 116 Anglo-Saxon, and 80 Medieval mandibles and measured the distance from the inferior alveolar canal to the occlusal surface. They found that despite marked attrition to the occlusal surfaces throughout adult life, the distance from the inferior alveolar canal to the occlusal surface did not change significantly, concluding that teeth in the human dentition constantly erupted to functionally compensate for attrition. One of the problems with this work is that accurate estimates of age were not available for the sample.

Similar work was undertaken by Whittaker et al. (1985; 1982) on a sample of 550 Romano-British mandibles. It was thought that continuous eruption would complicate simple measurement of alveolar bone loss and indications of periodontal disease. More thoroughgoing work on continuous dental eruption was conducted by Whittaker et al. (1988; 1990) on 122 intact mandibles of known age at death from Spitalfields. Again it was concluded that teeth did gradually exfoliate throughout adult life (approximately 2mm between 17 and 50 years of age). However, in the case of the 18th and 19th Century specimens from Spitalfields little occlusal attrition was observed,
probably due to a diet lacking those hard elements which abrade enamel, so it was also concluded that continuous eruption was not so much a reactive response by the body to attrition, but a genetically programmed odontological process which went on regardless of other dental stressors.
2.2.4 Cementum build-up

A feature of change in the adult human tooth supporting structures, which must be in some way a consequence of continuous eruption (above), is the ongoing deposition of cementum around the root surface. Cementum build-up was a component part of the age estimation method formulated by Gustafson (1950), but has been observed in the human adult dentition since 1898 (Broomwell, cited in Lipsinisc et al., 1986). A more contentious point is whether human cementum displays annually based incremental structures observable in cross sections, as is the case with some other mammals such as the bear, caribou, moose, otter (Stott et al., 1982). Although it was thought that in many of these mammals it was an annual metabolic disturbance, caused by hibernation, which set up alternative metabolic pathways for the formation of cementum, it has been found that the common marmoset, which does not hibernate, and which is non-migratory, also displays annual changes in cementum (Stott et al., 1982). In a mainly methodological paper Naylor et al. (1985) warned that cementum annulations were difficult to count. They recommended thorough protocols for section taking, cleaning, and staining (Naylor et al., 1985), but also said that if sufficient care were taken cementum rings were observable, and highly related to age. Stott et al. (1982) used three teeth from known age cadavers. They examined 100 µm stained sections taken from across the root. By counting rings observable across the root, and by adding on an offset to account for the age of tooth formation, they obtained an age which differed from the known age by no more than three years except in the case of a single upper left pre-molar. Charles et al. (1986) used a sample of 42 modern mandibular canine and first premolar pairs. Their preferred technique was to demineralise, stain, and cut 7µm longitudinal sections. Thin demineralised sections gave results which were more reproducible, possibly due to the absence of superimposed layers seen in thick un-demineralised sections. It was also found that there was a high degree of correlation between the number of rings counted and age, but it is not reported whether there was a high degree of support for cementum lines being due to annual metabolic changes. Condon et al. (1986) applied the methodology devised by Charles et al. to a sample of 112 clinically extracted human canine teeth. They found that there was a direct correspondence between age and the number of cementum rings counted, but they also warn
against ring doubling, an observation that under certain conditions there can be displayed twice as many incremental features as appropriate to a tooth of that age. However, Lipsinic et al. (1986) used a sample of 31 maxillary premolars, cross sectioned, both decalcified and un-decalcified, and found very little correlation between the number of incremental lines and age. It was found that stained decalcified thin (5μm) sections gave the best conditions for counting the cementum lines. Miller et al. (1988) using 100 known age extracted teeth, from 100 individuals, found little correspondence between cementum rings and age, although they used thick (200-400μm) un-decalcified cross sections.

A way around the debate on annual incremental structures has been pursued by Solheim (1990), who measured cementum thickness throughout adult life. Solheim (1984) used a sample of 1000 extracted and autopsy half-sectioned teeth from a Scandinavian population. He measured thickness at the cemento-enamel junction, at the apical foramen, and, at two equidistant intermediate points down the root. Cementum thickness was found to approximately triple in thickness between the ages of 20 and 60 years at the point one third of the distance between the apical foramen and the cemento-enamel junction, with a correlation with age between 0.31 and 0.72. There seemed to be no significant affect on the rate of cementum formation of sex, or oral pathology, and it was concluded that cementum thickness could contribute significantly as a factor in any method of age estimation which used multiple indicators.

If it could be demonstrated conclusively that cementum did undergo some annual change in deposition as claimed by Condon et al. (1986) Charles et al. (1986) and Stott et al. (1982), and suitably repeatable protocols established for observing the incremental structure, then the problem of adult human age estimation would be virtually solved. However, there seems to be no biological basis for regular incremental structure formation (Miller et al., 1988), and difficulties with observation (Naylor et al., 1985). Nevertheless, as there is no mechanism by which cementum can form annual layers, it is interesting that it is possible to observe them at all, even if the cementum layers are not damaged by the extraction process.
2.2.5 Root dentine translucency

Root dentine translucency (age-related sclerotic dentine) is an easily observable age-related change in the root dentine of human teeth. It is the description given to the dentinal tubules disposition to gradually fill and take on the same refractive index and light transmission properties as the surrounding dentine, taking on a glassy appearance (Vasiliadis et al., 1983a; 1983b). This process commences at the apex of the tooth at about the age of twenty years, and continues up towards the coronal dentine throughout adult life, but probably never extends into the coronal dentine as tubules in this structure seem to remain constant into old age (Carrigan et al., 1984; Garberoglio and Brannstrom, 1976). The rate of formation of sclerotic dentine seems to be unaffected by the function of the tooth or external stimuli (Vasiliadis et al., 1983a). The mesial and distal parts of the root dentine are the first to commence sclerosis (Coughlan et al., 1985), giving a butterfly form to the topmost part of the translucent volume (Coughlan, 1987; Vasiliadis et al., 1983a). Scanning electron microscope studies have led to the conclusion that sclerotic dentine is a continuation of the process of peritubular dentine formation (Vasiliadis et al., 1983b) with a similar material infill to peritubular dentine (Vahl and Mierau, 1971), but the micrographs published by (Nalbandian et al., 1960) clearly show an arrangement of spicules occluding the sclerosed tubule suggesting sclerotic dentine is a separate process of dentine calcification by the dentine forming odontoblasts, although the observation of spicules was said to be a preparation artefact (Nalbandian et al., 1960), or even the remains of bacilli (Doberenz and Wyckoff, 1967). An interpretation based upon a separate process would be more in accord with observations from reparative sclerotic dentine (Mendis, 1985), and would fit with the observation that the mineral component of sclerotic dentine is of a different type to the surrounding peritubular dentine being calcium deficient, possibly tri-calcium phosphate or pyrophosphate (Collet et al., 1989; Moore and Leaver, 1974); as well as the fact that there is no significant correlation between the width of occluding tubules and age (Fleming and Altini, 1984).

Root dentine translucency was another component used by Gustafson and Johanson (1950; 1971) in their age estimation techniques. Johanson (1968) was the first in the odontological literature to attempt to estimate age by using root dentine translucency as an ordinal variable. Johnson used a sample of ninety-three anterior teeth, from twenty-seven males, aged
between 28 and 73. He measured ratios of transparent dentine to opaque dentine area, total root area, total root dentine area, and the width of the root at the cemento-enamel junction. There was found no significant correlation between any of these ratios and age. Bang and Ramm (1970) took a more direct approach in that they measured the minimum and maximum length of the translucent dentine in 400µm labial-lingual longitudinal sections from a sample of 1402 teeth taken from some 265 individuals, from a modern Scandinavian population. They found correlation coefficients between 0.90 (maxillary second left pre-molars) and 0.5 (mandibular molars, 1st root), correlation coefficients being about 0.65 for most teeth. Bang and Ramm found there to be no sex dependency, and that there seemed to be no difference were the overall length of the tooth taken into account. In agreement with Bang and Ramm's overall results, Lamendin and Cambray (1980) measured the length of root dentine transparency from a sample of 217 teeth. They found a correlation coefficient of 0.73 with age, although did not calculate separate age estimation equations for each tooth. Likewise, Wegener and Albrecht (1980) measured root dentine transparency for a sample of 601 teeth from 50 individuals. They did not treat each tooth type separately, but still obtained a correlation coefficient of 0.67. Solheim (1989) divided a sample of 1000 teeth from a Scandinavian population by tooth type, approximately 100 of each type, and measured the length of root dentine transparency. Solheim (1989) found there to be a high contralateral correlation in individuals who had contralateral teeth examined, the rate of root dentine transparency formation was linear, and unaffected by oral pathology. Correlations with age varied between 0.65 (maxillary lateral incisors) and 0.86 (mandibular fourth pre-molars) dependent on tooth (Solheim, 1989).

Root dentine translucency is usually observed by cutting sections from teeth. However, it is possible to estimate the length of age-related sclerotic dentine without damaging the tooth. Lamendin (1973) used radiographs to try to examine sclerotic dentine in the roots of human teeth, and, somewhat confusingly was able to verify a correlation with age of 0.73 found by other authors, whilst actually obtaining a correlation with age of 0.45. Drusini et al. (1991) measured root dentine translucency for 152 teeth from 134 individuals by using light transmitted through the tooth root by a high intensity light. They measured both the length of the sclerotic zone and the area, which were
both expressed in terms of their relative proportion of the total root length and area, respectively. It was found that the correlation with age was 0.58 for anterior teeth, and 0.84 for sclerotic length on premolars. Sclerotic area for pre-molars had a correlation of 0.81 with age. Lamendin et al. (1992) used root dentine translucency measured by light transmitted through the tooth root in conjunction with an estimate of periodontal recession. It was found that age estimates coincided with those from other anthropological age estimates for a sample of 91 individuals from a churchyard in Oslo, but, as the individuals were not of known age it was not possible to judge how effective the two combined dental criteria were.

Root dentine translucency has been applied to skeletons of unknown age from archaeological contexts. Bang (1993) used root dentine transparency to estimate the age of a prehistoric individual found in Norway using 400µm longitudinal labial-lingual sections of a left canine and left pre-molar from the mandible, and right canine from the maxilla. The age of the individual was estimated to be about 60 years, and the date taken from associated pollen and uncalibrated radiocarbon measurements was about 5000 BC. Beyer-Olsen et al. (1994) estimated the age of 248 skeletons excavated from the Medieval church of St. Olaf’s in Trondheim, Norway, using 400µm labial-lingual longitudinal sections. The estimated ages were not significantly different to those obtained by more conventional anthropological means. However, Lucy et al. (1995), when attempting to use root dentine translucency to estimate the ages of four individuals, from eight sections of six teeth, from the Medieval cemetery of the hospital of St James and St Mary Magdalene, Chichester (Lee and Magilton, 1989; Magilton and Lee 1989), found identifiable root dentine translucency in only one individual. This individual was a 50 year old male aged from sections of a right maxillary second molar, and a left maxillary central incisor. All the other seven teeth inspected had suffered extensive tunnelling by fungi (Dye et al., 1995; Lucy et al., 1995), suggesting that not all individuals from the archaeological record possessed teeth from which age estimates could be made by observations of root dentine translucency.

Root dentine in sectioned teeth is relatively easy to observe and measure both length, and area. It is debatable, however, whether accurate assessments of root dentine translucency can be made using non-destructive methods. For instance, the transmittance of light, or x-rays, will be dependent
on the absorption of the material in the photon path, which in itself will be heavily dependent on overall tooth dimensions. In the case of teeth which have been buried for some time dentine degradation will probably have a marked effect upon the transmittance properties of radiation. Were no transmittance to be noted in a tooth from an archaeological, or long buried forensic, context, then it need not be the case that there is little dentine translucence, implying a young individual, but merely degraded dentine structure with low transmittance properties. The problem for the biological anthropologist, or forensic pathologist, is that externally there might be no indication that the tooth had suffered fungal tunnelling, as teeth with very badly degraded dentine structures can have a pristine external appearance (Dye et al., 1995).

2.2.6 Cementum resorptive processes

Both Gustafson (1950) and Johanson (1971) employed as a factor in their age estimation techniques the ability of osteoblastic cells to resorb mineralised tissue (Boyde et al., 1984). Gustafson and Johanson examined the tendency for the very apex of the root to gradually disappear and quantified their findings in terms of arbitrary, but observable, stages of resorption. Both found resorptive changes in the root of the tooth to be the least correlated with age of the changes examined. Dalitz (1962), examining 146 extracted teeth from an European-Antipodean population, came to the conclusion that root resorption, quantified in the same way as in Gustafson’s work, was so poorly correlated with age that it could be discarded for the purpose of age estimation. Solheim and Kvaal (1993c) likewise found that root resorption from the apex of the tooth was very poorly related to age. However, Solheim decided to examine resorptive processes across the entire root surface which can be detected in the form of areas of pocketing on the surface of the cementum. It was concluded that although quantifying resorptive processes across the whole root surface was a better indication of age, the arbitrary classification of four stages of development were so subjective that this attribute would be better thought of as a component in a multi-factorial age estimation scheme.

2.2.7 Other dental age changes

Ten-Cate et al. (1977) found that the roots of teeth tend to get darker later in adult life which is thought to be due to the deposition of blood
products in the dentine structure (Solheim, 1988). Ten-Cate et al. found age estimates could be made for unknown specimens by comparison with known specimens grouped by age, although there were some doubts about the accuracy of these estimates. They then went on to measure colour on a continuous scale with colourimetry. The reported correlation of root colour with age is 0.9, but the authors do admit to some deletion of what were termed 'statistical outliers'; there was no evidence for either the sex of the individual, or any oral pathology affecting the colour (Ten-Cate et al., 1977, fig. 3). Solheim (1988) used 758 teeth of known age of extraction from a Scandinavian population, and measured the colour of the root using a dental colour classification system and colourimetry. Using the dental classification system the correlation with age varied between 0.59 (mandibular central incisors) and 0.84 (maxillary central incisors). Colourimetry revealed correlations with age which varied between 0.55 to 0.78. It was concluded that colour was more correlated to age than many dental age changes.

Certain bonds which are part of the proteinaceous component of dentine and cementum are known to emit fluorescent light when stimulated by green and ultra-violet light (Dye et al., 1995; Eastoe, 1967). This has been shown to be mainly due to protein cross-links (Sakura and Fujimoto, 1984), the number of which are thought to increase with age (Cerami et al., 1987). Kvaal and Solheim (1989) measured green-stimulated red fluorescence emitted from the dentine and cementum from 1000 teeth from a Scandinavian population. They found that cementum gave a larger amplitude of fluorescence than dentine when controlling for age, that the correlation of the intensity of emitted fluorescent light with age was 0.73 and 0.77 for dentine and cementum respectively, and that there was a significant correlation with dentine colour even when controlling for age. They found no effect of age, or sex, but did find that generally teeth which had been extracted from cadavers to fluoresce more strongly (Kvaal and Solheim, 1989).

The chemical environment for proton species attached to the phosphate ion in hydrated calcium phosphates, part of the mineral constituent of enamel, has been investigated in relation to age. Funduk et al. (1986) used proton nuclear magnetic resonance to examine these environments and the tiny amounts of non-bonded water present in enamel. There was found to be no differences between a group of teeth from 20 year old donors, and a group of 50 year old donors.
The presence of a globular structure in dentine has also been examined for age-related change. Globular dentine arises as a result of dentine maturation, where crystallites are seeded onto the protein matrix. The crystallites grow radially until impeded by a neighbouring crystallite. The resulting meta-crystals are termed calcospheres (Shellis, 1983), or globular dentine (Hillson, 1986, p.153). Mosca and Bonal (1983) found in two samples of 15 premolars, one from a group of individuals aged between 40 and 85, the other aged between 13 and 24, that the calcospheric superstructure became very scant, and that tubular irregularity became more common.

No attempt has been made to employ the above age-related changes for age estimation. Dentine fluorescence is easily quantified, but the observed differences between teeth from live donors and cadavers is not yet understood, giving rise to doubts about whether fluorescence is a stable quantity after extraction (Kvaal and Solheim, 1989); and changes in the globular dentine structure are extremely difficult to quantify (Mosca and Bonal, 1983).

2.3 Discussion of adult human hard tissue changes

Many parts of the adult human hard tissue have been exploited as age indicators. Some of these, such as root dentine translucency, auricular surface, or cranial sutures, are not apparently affected by activity or pathology. However some are at least partially pathological, for example bone density which can be affected by acknowledged conditions such as osteoporosis, and secondary dentine formation can be a response to occlusal attrition and various dentinal insults. Osteoporosis may be readily diagnosed from other elements of the skeleton (Johnston and Melton, 1990; Nordin et al., 1966), but other conditions which exacerbate secondary dentine formation such as the possible reaction to corticosteroids as used in renal failure (Nasstrom et al., 1993) are not so easily diagnosed from the hard tissue alone, and therefore may mislead biological anthropologists and forensic scientists when examining these age indicators.

All age indicators have a large inherent variability of any age indicator for chronological age, the most highly correlated being root dentine translucent length \( r = 0.80 \). This should necessitate a large reference sample for any age indicator. The largest single reference collection numbers some 1000 individuals for the work done by Tore Solheim since the early 1980’s.
Unfortunately Solheim found it necessary to divide this particular sample by tooth locus, so in reality there are only about 80 individuals represented. The next largest reference sample (see 11.1.8) consists of 764 individuals used by Lovejoy et al. (1985a) to develop the auricular surface as an age indicator. However, the 764 cited had 250 individuals from Libben (Ohio prehistoric site) included, so is in reality has only about 500 individuals. In fact nowhere in the literature has a large reference sample been used for age estimation, most numbering in the low hundreds, some even below fifty. The unavailability of large reference samples will continue to be a stumbling block for anthropologists and forensic scientists who wish to be able to estimate age at death of adult humans.
3 The Estimation of Age From Hard Tissue Changes

From chapter two it can be seen that many age estimation procedures have a stage where some numerical quantity is given to a morphological age change in human hard tissue. This quantity is usually at the ordinal scale or above, and the classification of the stages is usually dictated by ease of observation rather than actual change with age. The next stage in the procedure is, from the quantity placed upon the age change, to estimate the age of an individual based upon what is commonly called a reference sample. A reference sample is a set of observations about the same age changes, categorised by the same scheme, from a sample of individuals for which age is known from independent sources.

Implied in this procedure are two assumptions. The first is that all individuals in a population will go through the same age-related changes. This may seem like a small point, but, because samples for age estimation are always sampled cross-sectionally, i.e. a researcher samples from an extant population at a single moment, which assumes that the observed age changes in individuals for different ages represent some biological process though which all individuals in the population will eventually go. This is not fully justifiable since it is possible that individuals born in a certain year will go through a set of changes which are different to those born, for instance, ten years later or earlier. It could also be said that all individuals which comprise a reference sample are usually (particularly in the case of many skeletal age changes such as endo-cranial suture closure) dead, and therefore will not change further with age. We have no way of knowing whether these individuals would have continued to change with age had they lived.

The next assumption is that individuals in the target sample, those for which age needs to be known, are biologically similar to the reference sample in respect of the rate of age change. This assumption is to some extent reliant on the first assumption, it being true that if the same age changes did not occur, then they cannot possibly occur at the same rate of change.

Of course, because those who make inferences about age at death from the archaeological record are dealing with deceased human groups, it can never be known whether biological age parameters for the extinct group are the same as for more modern human groups. It can, however, be known if
ageing parameters are different. It can be shown that an age-related change operates at a markedly different pace in an extinct group by examining two or more age markers relative to each other, and to the same two age markers in extant populations. Then it can be known whether the rate of change for these age changes are different between the two populations, but not which one is different or by how much, although lack of difference does not by necessity imply similarity.

Given the truth of the assumptions above, the next stage is to compare the observable age-related change for the target sample to those of the reference sample, and, by invoking some means of calibration, finding ages for the target sample. Little attention has been given to this latter process in the anthropological and forensic literature, and to some extent has been taken for granted by scientists working in this field. This is the subject of this chapter and the rest of this thesis.

### 3.1 Univariate age estimation

#### 3.1.1 Categorisation by description

Appearing in the literature spanning for nearly a century are a number of different means of deducing age from observable biological parameters.

The early methods for estimating an age from a given observation would revolve around long lists of descriptions. The early work on cranial suture closure from Dwight (1890) and Parsons and Box (1905), gave case by case descriptions. Anyone wishing to estimate an age for an unknown individual would consult the descriptions and match the unknown to one of them. The disadvantage is that as cranial suture closure is essentially multivariate, in that each suture can close semi-independently, any crania not conforming to these descriptions could only be classified at the discretion of the investigator.

Todd and Lyon’s (1924; 1925b) consideration of each cranial suture as independent, and their subsequent classification of each one on a scale of one to four, would allow novel combinations to be classified. However, no instructions were given as to how to arrive at a best age given any particular set of suture closure scores.

Recently, age classification using a descriptive classification scheme to judge age from morphological changes in the pubic symphysis has enjoyed something of a revival. Meindl et al. (1985) and Lovejoy et al. (1985a) looked
at pubic symphyseal changes in a similar way to Todd and Lyon, but broadened their classifications so specimens unclassifiable in the Todd and Lyon system could be successfully classified.

A scheme which commonly appears in the anthropological and forensic literature for estimating age from a numerical score, is where the distribution of age for a given stage of an age change is estimated. For example, Gilbert and McKern (1973) quantified each of their components of the female Os pubis into six different stages, summed the scores from each component, and derived the mean age for each possible component sum. Their confidence interval was estimated from the standard deviation of ages about the score. Likewise Perizonius (1984) used the mean of ages for a given cranial index score, the cranial index score being derived from the sum of scores given to each quarter of the five major cranial vault sutures. Meindl et al. (1985) scored a form of cranial index from the sum of suture closure scores given to lateral anterior cranial sutures. Again, for each possible score they calculated the mean and standard deviation of ages appearing in their reference sample.

This type of procedure is analogous to a form of forward regression, where the score is assumed to be known without error, and age varies about the score awarded to a particular age change. The score is assumed to change monotonically with age. The problems associated with regression are covered in more depth in Chapter Four, but generally this simple approach circumvents some of the disadvantages of regression. Firstly, no particular model for change in score with age has to be assumed. Secondly, if the variation in the age change is dependent on age then there is no particular problem in age estimation, as each stage of the age change is considered separately. However, there is the assumption that ages for a given score conform to a gaussian distribution, which can be particularly misleading as a measure of confidence for age of an individual.

An extension to age categorisation by direct comparison is when more than one skeletal feature is used to classify an unknown skeleton into an age group. An approach which is commonly employed is to sum the numerical score from each individual age-related change, and then to treat this sum as though it were a numerical score from a single age-related change as before (Gilbert and McKern, 1973; Meindl and Lovejoy, 1985; Perizonius, 1984). Usually each individual score is given equal weight. Implicit in this procedure
is the assumption that each individual age-related skeletal change gives equal information about age. This is patently not true, at least in the case of Perizonius (1984), who cites correlation coefficients with age for the numerical score placed upon each individual skeletal change as being between 0.23 and 0.43. This information is unavailable in Meindl et al. (1985) and Gilbert and McKern (1973).

A variation on age estimation using description is where age for a given individual can be ascertained directly from observation with no intervening calculations. This is the case with cementum annulation ring counting, where the number of cementum rings, plus an offset to allow for the development of the tooth, will equal the number of years lived (Naylor et al., 1985; Stott et al., 1982). This has to be the most straightforward and unproblematic means of translating an observation into an age, even though the fact that humans deposit a single layer a year is the focus of some dispute (Lipsinic et al., 1986; Miller et al., 1988).

3.1.2 Regression models and single variables

Linear regression has often been the means by which age is estimated from a single observable which is either on the ordinal or continuous scale of measurement. Forward linear regression appears as a means by which to estimate age from measurements of the extent of root dentine translucency as a continuous measure in the work of Vasiliadis et al. (1983a). Solheim and Kvaal (1993c) also used forward regression analysis to estimate the relationship between the amount and distribution of resorption sites on the cementum surface of the root which they measured on an ordinal scale. Likewise, Ten-Cate et al. (1977) and Solheim (1988) used the colour of dentine as an age-related ordinal variable which was regressed against age. Although there are hidden assumptions regarding forward regression analysis, (see 4.1.1) linear regression was apparently successful in these applications, appropriately modelling the relationship between the observable variable and age.

More problematical is the use of non-linear forward regression modelling. Unless one variable is perfectly linearly correlated with another, a non-linear model will always show an increase in correlation coefficient over a linear one. Conventionally, unless there is some prior reason why a non-linear model should be applied, then a linear model is always preferred unless it is shown during the process of model testing to be inadequate. However, some workers
have taken the increase in correlation coefficient as an indication that any
derived non-linear model is somehow better. Bang and Ramm (1970) used
forward linear models to estimate age for root dentine translucency
measurements in cases where the length measurement was greater than nine
millimetres, but in cases where the translucent length was less than nine
millimetres they used a series of quadratic equations. They did not
demonstrate inadequacy of fit for linear models in the sub-nine millimetre
range, and, if the quadratic models given are plotted out across the range of
possible values for age, then it can easily be seen that most of their quadratic
models are close approximations to straight lines. However, later workers
employing the techniques of Bang and Ramm (1970) have continued to use
their quadratic models (Bang, 1993; Beyer-Olsen et al., 1994).

An inappropriate use of non-linear modelling can be found in Kerley and
Ubelaker (1978) who modelled bone osteon formation. In particular, their use
of cubic models to describe the relationship between primary osteons and
age, and, osteon fragments and age, seems problematic because the cubic
models described actually double back on themselves implying that the
number of cortical osteons can decrease with age, which is impossible given
the current conception of the biology of these skeletal features (Aiello and

3.2 Multifactorial age estimates

3.2.1 Multi-state classification

By the 1970s it was becoming clear that an age estimate based upon a
single skeletal trait was not going to give consistently accurate, or precise,
age estimates for human adults. An explicit attempt to use a more rational
means of estimating age from four skeletal indicators was outlined in the first
edition of Acsádi and Nemeskéri (1980), with a full tabulation of age change
and age being published in 1975 (Sjøvold, 1975) to provide a useable
estimation system for anthropologists. Based upon this work, a more
comprehensive article was published in 1980 by Ferembach et al. (1980).
Ferembach et al. used four skeletal age changes, classified into between four
and six ordinal stages. Cross-tabulations of all four skeletal age changes were
produced from a reference sample, the age attributable to an unknown
individual being the mean age of those in the reference sample with the same
combination of age markers, and the error the standard deviation of those
ages.

This method is analogous to forward regression analysis, where age is
considered to vary about an age marker which is known without error. As in
work cited earlier (e.g. Gilbert and McKern, 1973; Meindl and Lovejoy, 1985;
Perizonius, 1984) it would be more common to regard a score as being
dependent on age. Likewise it assumes the age distribution for any given
combination to be gaussian. Ferembach et al. (1980) cite errors for any given
state as being between 2.5 and 3 years, which would seem rather small, but
could be largely the effect of only having 105 skeletons in their reference
sample (Nemeskéri et al. 1960 cited in Acsádi and Neméskéri, 1980). As in a
cross-tabulation of four variables, each with only four states, there would be a
potential 256 possible states, even if it is accepted that there will be many of
the possible 256 states with no individuals corresponding to those states, it is
difficult to imagine only 105 cases giving an accurate estimate for any
combination of skeletal age changes given age. An advantage with this
approach is that there is no assumption made about the relative importance
of each age indicator, and cross-tabulation removes any dependency which
may exist between variables.

A method for deriving an estimate of age from multiple age indicators
was proposed by Lovejoy et al. (1985b). In this work it is difficult to make out
the analytical basis for assignment of age. As far as can be ascertained four
age indicators are noted for all the target sample (pubic symphysis, auricular
surface, cranial suture and sternal rib end), individual age markers being
seriated, and then divided into groups corresponding to a score for each age
indicator. The resultant age indicator scores are then subjected to principal
components analysis. From here on the technique is somewhat obscure, but,
either the first principal component is taken as being proportional to age, or,
the correlation matrix is used to assign weights to each age indicator from
which a final summary age is produced. Later publications referring to this
method (Bedford et al., 1993; Mensforth and Lovejoy, 1985; Saunders et al.,
1992; Walker and Lovejoy, 1985) tend to favour the latter interpretation of the
somewhat opaque description given by Lovejoy et al.

The use of seriation for ordering, and assigning stages to age indicators,
as stated by Lovejoy et al. (1985b), has the advantage that inter-observer
error should be minimised, and also the error due to a single observer varying in their attribution of state to given age change with time (although no further examination of this point was made). However, it does depend upon having a reasonably large reference sample. The use of principal component analysis to assign weights to variables removes the effects of inter-dependence between variables, and using weights overcomes the problem of each variable having the same contribution to any final estimated age without requiring a full cross-tabulation. No attempt was made by Lovejoy et al. to calculate confidence intervals or any measure of uncertainty for a point estimate.

3.2.2 Summary variable and uni-variate regression

Another approach which has been used frequently is to make a composite, or summary variable, from observations of separate variables, usually by summation, and then regression of the new composite variable against age. Gustafson (1947; 1950; 1955) summed points attributable to each of six age-related changes observable in the internal structure of teeth. Gustafson's estimate of age error (3.6 years) has long been held in suspicion. Maples and Rice (1979) recalculated Gustafson's error estimate from published data (Gustafson, 1950) and concluded that the estimate of error under forward regression conditions should be 7.02 years. Lucy and Pollard (1995) recalculated Gustafson's error using an inverse regression scheme and found that a more appropriate estimate of error should be 8.2 years. Finally, Aykroyd (1995) found there had been a mistake in Gustafson's original calculations of error, and managed to reproduce the original figure of 3.6 years, but concurred with Lucy and Pollard that an error of 8.2 years would be the most appropriate error measure for estimated ages.

Other attempts to use a composite variable have used age changes which are generally measured on a continuous scale. Drusini (1993) used a composite variable based on the ratio between the crown height and the coronal pulp height, both of which should decrease with age as the pulp fills with secondary dentine, and the enamel is subject to occlusal attrition. Similarly, Ito (1972; 1975), used a crown index which was the ratio of the whole coronal area, seen in section, to the coronal pulp chamber area. Schwarz et al. (1978) used a tooth index which was the ratio of the length from the root apex to coronal extent of the pulp chamber, to the root translucent length. The root translucent length increases with age, and the
apex to pulp chamber roof length decreases with age. Kashyap and Koteswara-Rao (1990) used a series of indices based upon ratios of various age-related dental parameters. For instance, occlusal attrition was scored as the ratio of the width of exposed dentine to the width of the dentine at the cemento-enamel junction. Secondary dentine was scored as the ratio of the length of the apex to pulp-chamber roof, to the length down the pulp chamber of any secondary dentine. Each of these was regressed against age, the final age for any individual being the mean of the four estimated ages.

All parametric regression procedures have associated with them a number of important requirements which are listed below:

1. Variables should give independent information about age. This means that the age-related changes cannot be dependent physically upon each other. The changes can be correlated with each other, and this would be expected given that they all change systematically with age, but increase in one variable cannot directly cause change in another.

2. Variables should vary continuously with age. A continuous variable has to be theoretically capable of adopting an infinite number of values, such as the linear measurement of root dentine translucency used by Bang and Ramm (1970).

3. The error distribution about the mean of any age-related variable for a given age is normal, which essentially means that the variables are univariate normal (although univariate normality does not necessarily imply the variables are multivariate normal). This means that any estimate of error for an unknown individual will also be normally distributed about the estimate of age. This is an additional constraint which may not be completely fulfilled when the data are examined. A corollary of this assumption is that the predicted variable should be continuous. This means that if some form of inverse regression to be considered necessary (Lucy and Pollard, 1995), where age change is regressed as the response variable against age as the controlling variable, then the age change variable should be continuous. Otherwise some other error model, such as a Poisson distribution, may be considered more appropriate.
4. When summation methods are used, such as by Gustafson (1947; 1950; 1955), it is assumed that all the variables are contributing the same amount of information about age. This fact that different age-related changes have different correlations with age strongly suggests that this is untrue.

Clearly, Gustafson's approach (1947; 1950; 1955) contravenes all four of the above conditions. Ito (1972; 1975), Drusini (1993), Kashyap and Koteswara-Rao (1990), and Schwarz et al. (1978) used variables which were continuous, but still assume that the observable quantity varies normally for a given age, the variables are independent, and, each variable gives equal amounts of information about the age of the individual. Because so many of the fundamental assumptions of regression analysis are broken does not mean that estimates of age will be wrong, or more inaccurate than other estimation procedures. It will mean that the estimates of error in any single instance will be inaccurate, and a non-Gaussian error may be more applicable in these circumstances.

3.2.3 Multiple regression

An alternative approach to deriving an estimate of age from age-related data from multiple age indicators is multiple regression. Johanson (1971) used Gustafson's variables, but attributed intermediate stages between Gustafson's four, to produce six possible states for each variable. Johanson (1971) then used multiple linear forward regression against age to produce a model which could then be used to give age estimates for unknown individuals. Maples (1978) used Gustafson's six age changes each categorised into four states. These were then subjected to multiple forward linear regression to produce a model which described age change with age. Maples found that significant improvements were made by discarding information about root resorption, and by producing separate models for each tooth locus. Thompson (1979; 1981) applied multiple forward linear regression to continuous data from osteon remodelling of cortical bone. Thompson produced separate models for each position from which bone specimens were taken, and separate models dependent upon the sex of the individual and whether there was any obvious pathology of the skeletal tissue. Solheim (1993a) used a mixture of continuous and ordinal variables from tooth sections and multiple regression. Solheim's measure of periodontal recession
was log transformed, which meant that there was no assumption of linear change with age. Likewise, Xiaohu et al. (1991) used log transformation and multiple forward regression on a set of dental age changes. Their age parameters were all scored on a continuous scale, but they also took into account dichotomous factors such as sex. Hanihara and Suzuki (1978) used linear forward multiple regression on stages of morphological changes in the pubic symphysis in a technique akin to that of McKern and Stewart (1957). They warn that their analysis does not apply to those aged above 38 years. The obvious problem here is that age is the estimated quantity and there is no independent means of judging whether an individual is over 38 years. Both Dreier (1994) and Hongwei and Jingtao (1989) used multiple forward regression to estimate age from occlusal attrition using tooth locus as a separate variable.

The requirements for the use of multiple regression are exactly the same as for univariate regression, with the exception that variables are weighted to account for their relationship with age. Johanson (1971), Maples (1978), Hanihara and Suzuki (1978), Hongwei and Jingtao (1989) and to some extent Solheim (1993a) and Xiaohu et al. (1991) used variables which are on an ordinal scale. Even Johanson's (1971) expansion of Gustafson's six variables from four stages to six stages comes nowhere near to approximating the nominally infinite number of stages required by a truly continuous variable. The 25 stages of Dreier (1994) arguably do approximate a continuous variable.

All authors regarded age for a given indicator in a given state as varying normally about some mean value corresponding to the observation. The estimates of error are taken from an estimate of this distribution. They also assume that each of their observed variables is independent of the other variables. The only workers to have used non-linear modelling are Solheim (1993a) and Xiaohu et al. (1991). For most age-related variables there is no evidence to support the use of anything but linear models. The only assumption not made when using multiple regression instead of univariate forward regression employing composite variable made by summing, or some other function of individual age-related observations, is that each variable gives the same information about age. Otherwise the full range of criteria applicable to univariate forward regression is equally valid for multiple regression. Additionally, if any variables are missing from a set of
observations then new regression models have to be calculated to compensate for the loss of information as seen in Johanson (1971). Otherwise some scheme of using a mean value to fill in for the missing value could be employed, but it is doubtful whether this would be fully satisfactory.

The use of multiple regression does not get around all the problems of univariate regression for age estimation, although it does allow each variable to contribute differing amounts of information to estimates of age. Also it is less amenable to being adapted to inverse methods, which will be discussed in the next Chapter.

### 3.3 Other multivariate methods

Takei (1970; 1984) used a method of age estimation from occlusal attrition which seemed to rely on a form of forward multiple linear regression. Takei used four stages of occlusal attrition for each tooth, and two dummy variables which related to whether the tooth had been restored, and whether it was missing or not. The result was a table of values for each tooth in the dentition, classified by the condition attributable to the teeth. The method was to sum the values appropriate to each tooth in an unknown dentition, with the sum being equated to the age attributable to that unknown individual. This has the advantage over multiple regression that missing values are accounted for, and also that other age-related factors, such as the progressive acquisition of caries and restorations, contribute to any estimate of age. However, Takei gives no estimate of expected error for any individual age estimate.

A most unusual use of multivariate analysis for age estimation is to be seen in the work of Johnson (1976). Johnson used linear discriminant function analysis to estimate the ages for 160 19th century individuals from Ashton-under-Lyme, Lancashire, UK. On the basis of cranial suture closure and dental occlusal attrition Johnson defined seven decadal age groups, noted cranial and dental age markers, and then defined a function to discriminate between the seven age groups. Johnson then used this discriminant function to assign age group to each individual in the excavated sample. The method is circular, in that the sample from which the discriminant function was defined was the same as the sample to which the function was applied. Johnson knew from parish records how many individuals belonged in each age group, but, except in 22 cases, not which individuals belonged in which age group. So the whole
assignment of age is based upon another circularity in that the defined function is reliant upon the initial assessment of which individuals belonged to each age group. Johnson describes as interesting the observation that the 22 individuals with positively known ages had estimates assigned to them which were marginally less accurate than those without positively known ages. Presumably the point of interest alluded to has to be one of how Johnson could possibly know this!
3.3.1 Direct age structure estimation

In recent years there has been a move in some quarters away from the traditional ‘forensic’ style of estimating age for each individual in a sample, and examining the age structure as a sum of individual ages, to estimating the age structure directly by comparison to a reference sample. Konigsberg and Frankenber (1992) used a maximum likelihood-based technique first developed by Kimura and Chikuni (1987) for the fisheries industry. The method revolves around comparing the distribution of age indicators in the reference sample to the distribution of the same indicators in the sample for which the age distribution is to be estimated (the target sample). The age distribution is then deduced iteratively using expectation maximisation for the target sample from that of the reference sample until convergence is reached (which when only small changes between each step are observed). Bouquet-Appel and Masset (1996) used a similar iterative approach, but recommended the use of proportional fitting technique instead of a maximum likelihood approach. The advantage with both of these methods is said to be that the age distribution of the reference sample is not reflected in the final estimated distribution for the target sample.

Although these iterative techniques for estimating the age structure for a sample of individuals of unknown age are promising, they are restricted to use in demography, as they cannot give estimates of age to specific individuals. Despite the rejection of ‘forensic’ based approaches by Bouquet-Appel and Masset (1996) for looking at age structure, there is still a need to be able to estimate the age at death for individuals. This is because many archaeological individuals are not found as part of a cemetery, but singly, as in barrow burials. Another point would be that anthropologists are also interested is the effect of disease, and in particular the mortality associated with disease whereby an estimate of age at death is useful in diagnosis. For this estimates of individual ages are necessary, and should be seen as complementary to the age structure of the overall population.

3.4 Summary and conclusions

The most reliable means of estimating age from an observation of any age-related physiological parameter has to be when there is a direct correspondence between some incremental structure in the hard tissue and
age such has been postulated in cementum annulation (Naylor et al., 1985; Stott et al., 1982). However, it appears that this simple correspondence may not exist (Miller et al., 1988). Direct descriptive methods (e.g. Dwight, 1890; Parsons and Box, 1905; Todd and Lyon, 1924; Todd and Lyon, 1925a) are certainly useful when examining the expected developmental sequence of an age-related observable, but are of limited value when a case deviates from the developmental nexus established by the reference sample. An alternative approach is to award a series of ordinal points, or measure a continuous change, for an age-related biological parameter, and use some statistical means to estimate age for unknown individuals. All the estimation methods outlined above have their drawbacks. All regression methods rely upon homoskedastic, continuous, normal, or multivariate normal, distribution of variables for a given age. These conditions have never been demonstrated to exist for any age-related variable. Summation methods (e.g. Gustafson, 1950) assume that each variable gives equal information about age. Cross-classification methods (e.g. Sjøvold, 1975) assume normal distribution of ages for any given age state, and again this has never been demonstrated to be the case. In fact, there has never been a flexible, rationally justifiable, and fully realised, statistical method for estimating age from biological parameters applied to human age estimation. The requirements for any such statistical method are that it must be able to include data from categorical, ordinal, and continuous variables in the same rational framework. It must make minimal assumptions about the nature of the data, and these assumptions must be explicitly realised. Also, it should be able to generate some estimate of confidence for any given estimate. The next chapter deals with a critique of regression models as applied to age estimation, and the following two chapters outline a new approach to statistical estimative procedures which goes towards fulfilling these requirements.
4The Theory of Regression Analysis as Applied to Age Estimation

As outlined in the last chapter, regression analysis has often been the means by which age estimates have been deduced from observations for age-related changes in the human adult hard tissue. The conventional means by which this has been achieved is by forward regression, where age is regressed as $y$ against age change as $x$, to produce an equation of the form $y = ax+b$, where $a$ is the slope of any linear model, and $b$ is the $y$ intercept.

Bedford et al. (1993), when testing the Gilbert-McKern age estimation system based on changes in the pubic symphysis, noticed a systematic error which tended to make young individuals slightly too old and old individuals much too young (their Table 2). Saunders et al. (1992), whilst testing the multivariate age estimation technique known as the 'complex method', found that age estimates from each separate age estimation method showed a similar systematic error, where young individuals were assigned ages too high, and old individuals assigned ages too low (their Table 1). Molleson (1993) also tested the 'complex method' on the skeletal material from Spitalfields, London, and, like Saunders et al. found the younger individuals were assigned ages too high and the older individuals too low. Wegener and Albrecht (1980) used an age estimation technique based upon root dentine translucency, and found that estimated ages were also too old for young individuals and too young for old individuals (their Figure 3). The same systematic discrepancy was also seen in an empirical study of different dental based age estimation methods by Solheim and Sundnes (1980).

Masset (1989, p.81-86) also noticed this phenomenon and called it 'the attraction of the middle'. He ascribed it to the skewed distribution ages in any reference sample from a human population. Masset tried solving the problem by varying the structure of the reference sample, but unfortunately the results were disappointing as there seemed to be no reduction in bias for estimates made for a documented sample.

The fact that this systematic error always takes on the same form and can be seen to occur irrespective of the age estimation technique employed suggests the cause lies deeply embedded in the methodology of age
estimation, and that at least part of the problem is the result of using conventional regression, or 'regression-like' models to produce age estimates.

4.1.1 Least squares regression

The most basic and commonly used modelling technique is least squares regression. The aim is to find a relationship between one or more controlling variables, sometimes known as independent variables, designated \( x_1, x_2, \ldots \), and a single response variable, \( y \), sometimes referred to as dependent variable. The simplest case is where there is a single controlling variable, \( x \), and the relationship is linear, i.e.;

\[
y = a + b x + \varepsilon
\]

where \( a \) and \( b \) are unknown parameters and \( \varepsilon \) is the random error associated with the measurement of the response. In order to calculate the line of best fit it is assumed that each value of \( x \) is exactly determined.

According to the model the error on the \( i^{th} \) point is:

\[
\varepsilon_i = y_i - \hat{a} - \hat{b} x_i
\]

therefore the sum of squares of all these errors is,

\[
S(a,b) = \sum \varepsilon_i^2 = \sum (y_i - \hat{a} - \hat{b} x_i)^2.
\]

In a least squares approach we define the 'best' fit to be one which minimises the function \( S(a,b) \), called the 'least squares criterion'. To determine this minimum value the values of \( a \) and \( b \) must be selected so that

\[
\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0.
\]

In this case

\[
\frac{\partial S}{\partial a} = -\sum 2(y_i - \hat{a} - \hat{b} x_i), \quad \text{and} \quad \frac{\partial S}{\partial b} = -\sum 2x_i(y_i - \hat{a} - \hat{b} x_i).
\]

Therefore

\[
\frac{\partial S}{\partial a} = -2 \sum y_i + 2 \hat{a} + 2 \hat{b} \sum x_i,
\]

71
and as a solution of

\[ \frac{\partial S}{\partial b} = -2 \sum x_i y_i + 2 \hat{a} \sum x_i + 2 \hat{b} \sum x^2. \]

So,

**Equation 1** \[ \frac{\partial S}{\partial a} = 0 \text{ when } \hat{a} n + \hat{b} \sum x = \sum y, \]

and

**Equation 2** \[ \frac{\partial S}{\partial b} = 0 \text{ when } \hat{a} \sum x + \hat{b} \sum x^2 \sum x_i y_i, \]

dividing Equation 1 by \( n \)

**Equation 3** \[ \bar{y} = \hat{a} + \hat{b} \bar{x}, \]

dividing Equation 2 by \( n \), and substituting for \( a \) using Equation 3

\[ \hat{b} \left( \frac{\sum x_i^2}{n} - \bar{x}^2 \right) = \frac{\sum x_i y_i}{n} - \bar{x} \bar{y}. \]

So,

\[ \hat{b} = \frac{S_{xy}}{S_{xx}}, \]

where \( S_{xy} \) is the covariance between \( x \) and \( y \), and \( S_{xx} \) is the variance of \( x \).

Estimates of age (\( y \)) are conventionally obtained by inserting values of the observed variable (\( x \)) into the derived expression (Equation 3). This form of least squares regression is depicted in Figure 11.
In Figure 11 for every $x$ value the corresponding $y$ values are normally distributed, emphasising the point that only the $y$ values have an associated error. The superimposed line is the function corresponding to the derived regression equation ($y = a + bx$) from these data.

If we now consider the residuals which are obtained by subtracting the estimated $y$ value from the observed $y$ value:

$$
\epsilon_i = y_i - \hat{y}_i
$$

Hence,

$$
\hat{\epsilon}_i = y_i - \bar{y} + \hat{b}(x_i - \bar{x}),
$$

and,

$$
\hat{\epsilon}_i = y_i - \bar{y} - \frac{S_{yi}}{S_{xx}}(x_i - \bar{x}).
$$

To examine any relationship between these residuals and the response variable, $y$, we shall regress the residuals on the response variable. The least squares approach is the same as above, with a change of variable names. Let the proposed linear relationship be $\epsilon = c + dy$ then, using least squares:
\[ \hat{c} = \bar{c} - \hat{a}\bar{y}, \quad \text{and} \quad \hat{d} = \frac{\sum (y_i - \bar{y})(\hat{\epsilon}_i - \bar{\epsilon})}{\sum (y_i - \bar{y})^2} = \frac{S_{\hat{c}}}{S_{xy}}. \]

By definition the mean of the new residuals equals zero. However, what we are really interested in is the slope of the regression line (d) and, more specifically, whether or not it is non-zero.

\[
\begin{vmatrix}
(y_i; y_i) & \sum (y_i - \bar{y}) & \sum (y_i - \bar{y})^2 \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\
& & \\ 
\end{vmatrix}
\]

\[
= \frac{\sum (y_i - \bar{y})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2} - \frac{\sum (y_i - \bar{y})}{\sum (y_i - \bar{y})^2} \frac{S_{xy}(x_i - \bar{x})}{S_{xx}},
\]

\[
= \frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} - \frac{S_{xy}}{S_{xx}} \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (y_i - \bar{y})^2},
\]

\[
= 1 - \frac{S_{xy}}{S_{xx}} \frac{S_{xy}}{S_{yy}},
\]

\[
= 1 - r_{xy}^2.
\]

Equation 4

where \( r_{xy} \) is the Pearson product-moment correlation coefficient between \( x \) and \( y \).

From this we can see that the graph of the residual against observed value will always have a positive slope since \( r^2 \) is always less than or equal to unity. When forward regression is used for age estimation this guarantees all young individuals will appear older than they really are, and all older individuals younger, whatever variables are used to construct the dataset, and that this effect is always inversely dependent upon the correlation coefficient between the age change and age.
Graphically this is represented in Figure 12, where the residuals from the points in the regression in Figure 11 are plotted against their respective $y$ values (ages). Notice the slope of the regression line as predicted by Equation 4.
Figure 12. Residual plot from the regression depicted in Figure 11.

As the correlation coefficient can only take values between minus one and plus one this means that the slope of the residual regression can only vary between zero and one. If there were perfect correlation, with correlation coefficient equal to one, then a plot of the residuals against the predicted variable would have zero slope. Conversely, were there no correlation, with a correlation coefficient of zero, then the slope would be equal to unity. Put simply, the poorer the correlation the greater the slope. With relatively low correlation between age indicators and age it is little wonder that anthropologists have been finding their residuals to be correlated with age. As all regression models are variants upon linear regression all regression models must suffer from the same sort of systematic error demonstrated above.

4.1.2'Regression-like' age estimates

The 'complex methods' used by Saunders et al. (1992) and Molleson (1993) both derive from the complex method first described by Acsádi and Nemeskéri (1980), and at first sight do not seem to revolve around regression analysis. Acsádi and Nemeskéri took four areas of the human skeleton which are known to change with age and identified between five and eight phases for each area. They then took the mean of the ages attributable to each phase from their reference sample as being the age attributable to that same phase being observed in any unknown skeleton. This gives four possible estimates of age which can be combined to produce an overall age estimate for that individual. The systematic error evident in the work of Saunders et al. (1992)
and Molleson (1993) stems from the second part of this procedure where the phase is considered to be exactly determined and the ages have an error associated with them. This is an exact analogue to the regression seen in Figure 11 where the distribution of residuals must lead to the same outcome as that seen in Figure 12, although the relationship between age and residual made in Equation 4 does not necessarily apply.

4.1.3 Inverse regression

Recently a number of authors have dealt with the use of forward regression models to estimate age, and a partial solution has been proposed (Konigsberg et al., 1994; Lucy and Pollard, 1995). As discussed above, to relate age and a numerical score placed on an age change, it has been usual to calculate a line of regression of age on score. It is suggested that this is the wrong way round, since the regression model assumes that all errors are in the y direction, that is in the known age; this is not the case. In this type of forward regression, we have to assume that any error in the x variable is small compared to that in the y variable, which strongly suggests that age should be chosen as the x variable, and age dependent variable as y. (It also suggests that errors associated with rounding the age to whole years should be avoided if possible.) This is an important point, because in general the line of regression of y on x is different from that of x on y, since the fit is obtained by minimising residuals in one direction only. Obviously, if the regression is carried out this way round, the equation needs rearranging before an estimate of age can be obtained from a measure of the observed variable.

$$\hat{x}_i = \frac{1}{b} (y_i - \hat{a}) .$$

Furthermore, if we now regress the residuals for estimation on the known age, that is derive a regression equation for known age minus estimated age against known age, we find that the slope is zero. Again, as in the case of forward regression, the approach is to define a new linear regression equation,

$$\epsilon = \hat{c} + \hat{d} y_i ,$$

where d is the slope of any linear relationship between the residual from the estimates, and the known age. What we are particularly interested in is whether d is non-zero for the inverse regression equation.
where, \( \hat{e}_i = x_i - \hat{x}_i \).

In this case, \( \hat{x}_i = \frac{y_i - \hat{a}}{b}, \) where: \( \hat{a} = y - \hat{b}\bar{x} \),

and, \( \hat{b} = \frac{S_{xy}}{S_{xx}} \).

So, \( \hat{x}_i = \frac{y_i - \bar{y} + \hat{b}\bar{x}}{b} \),
\( \hat{x}_i = \frac{S_{ss}}{S_{xy}} x_i - \frac{S_{ss}}{S_{xy}} \bar{y} + \bar{x} \),
\( \hat{x}_i = \bar{x} + \frac{S_{ss}}{S_{xy}} (y_i - \bar{y}) \).

hence, \( e_i = x_i - \bar{x} - \frac{S_{ss}}{S_{xy}} (y_i - \bar{y}) \).

\( \hat{d} = \frac{S_{xx}}{S_{sy}} \),
\( \hat{d} = \left( \frac{1}{S_{xx}} \right) \left[ \frac{1}{n} \sum (x_i - \bar{x})(e_i) \right] \) as \( \bar{e} = 0 \)

Multiplying this out:
\( \hat{d} = \left( \frac{1}{S_{xx}} \right) \left[ \frac{1}{n} \sum (x_i - \bar{x})(x_i - \bar{x}) - \frac{1}{n} \sum (x_i - \bar{x}) \left( \frac{S_{ss}}{S_{xy}} \right) (y_i - \bar{y}) \right] \),
\( \hat{d} = \left( \frac{1}{S_{xx}} \right) \left( \frac{S_{ss}}{S_{xy}} \right) \left[ \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y}) \right] \),
\( \hat{d} = \left( \frac{1}{S_{xx}} \right) S_{xx} \left[ \frac{S_{ss}}{S_{xy}} \right] S_{xy} \right] \),
\( \hat{d} = 0 \).

This means that estimates made by an inverse regression will not show the same systematic correlation seen with more conventional regression. For age estimation this means that estimates made about age for young individuals will no longer be systematically too old, and estimates for old individuals too young.
For inverse regression the estimate of error about any age is more complex than conventional regression. Furthermore, the error term associated with using such a calibration line to predict age values is not a constant, as is the case with conventional regression. As noted by Lucy and Pollard (1995) and Maples and Rice (1979), any errors on a regression line should form two parabolic curves about that regression line, i.e., the error term should grow larger as the predicted value gets further from the centroid of the regression line.

The problem posed is one of calculating the error about a calibration line derived from experimental data, and is the same as that faced by, for example, analytical chemists when trying to put error estimates on concentrations derived from a calibration curve. The basic method is outlined in Miller and Miller (1984, p.90-96) and consists of two parts: first is the calculation of the standard deviation of the regression of $y$ on $x$ ($s_{y|x}$) where $y$ is the value of the measured quantity (e.g., the point count for a given tooth), and $x$ is independent variable - the known age for the individual:

\[
s_{y|x} = \left[ \sum \frac{(y_i - \hat{y})^2}{n-2} \right]^{1/2}.
\]

Where: $y_i$ = the indicator score
\[
\hat{y} = \text{the estimated indicator score from the regression equation}
\]
\[
n = \text{number of pairs of observations used in the regression}
\]

Once this has been found the following approximate expression can be used to calculate the error associated with any individual age estimate ($S_{xo}$):

\[
S_{xo} = \frac{s_{y|x}}{b} \left[ 1 + \frac{1}{n} + \frac{(y_i - \bar{y})}{b^2 \sum (x_i - \bar{x})^2} \right]^{1/2}.
\]

It will be noted here that any estimate of $S_{xo}$ will always be greater than an estimate of $S_{y|x}$, indicating the greater efficiency of forward regression at the expense of including greater bias.

### 4.1.4 Inverse regression and a ‘farewell to palaeodemography’

Bouquet-Appel and Masset (1982) suggested that estimated ages are crucially dependent on the age structure of the reference sample of indicators and age (see Chapter 1). They argue that if there is a Gaussian partition of age between two subsets in a reference sample, and one derives a function to
estimate age by regression from each subset, then each function will be highly different. They illustrate this with their Figure 2 (i and ii). These figures depict regression functions derived from two different subsets, on them is marked functions A1, A2, B1 and B2. A1 and A2 are the regression lines $y$ on $x$ where the age indicator is $y$, B1 and B2 are what they term the inverse regression lines $y$ on $x$ where $y$ is age. Indeed functions B1 and B2 have very different intercepts with the $y$ axis, and as they rightly point out would lead to the effect whereby estimates of age were a reflection of the age structure of the reference sample, but A1 and A2 have very similar slopes and intercepts. Bouquet-Appel and Masset point out that A1 and A2 would be suitable for age estimation, but state that it is necessary to use functions which estimate $y$ as age on $x$ as indicator.

However, in the procedure termed inverse regression described above, $y$ is indicator and $x$ is age, which leads to the functions labelled A1 and A2 by Bouquet-Appel and Masset.

Figure 13 depicts both forward and inverse models for an age indicator, in this case root dentine translucency, against age.

*Figure 13. Four possible regression models from two subsets of the same reference sample*

In this case the reference sample has been artificially divided into those individuals aged above fifty years, and those below. The two inverse functions are very similar (equivalent to Bouquet-Appel and Masset’s A1 and A2), but it is
the forward models which are different (Bouquet-Appel and Masset’s $B_1$ and $B_2$). This is because a forward model attempts to estimate the variable which has a Gaussian partition and therefore makes estimates which reflect this partition. Inverse regression initially attempts to estimate the variable for which estimates are not required for which there is no Gaussian partition, so does not reflect one.

It must be concluded that although Bouquet-Appel and Masset (1982) were right to characterise conventional regression models as always giving estimates which are reliant on the distribution of the estimated variable, but this is because they used the wrong sort of regression.

### 4.2 Conclusions

The arguments outlined above highlight theoretical failings of regression models, and regression 'like' models, when applied to age-related hard tissue data. A major criticism is that most age-related changes are quantified on an ordinal scale of measurement, but ideally continuous regression models require age changes which can be measured on a continuous scale, although dummy variables can be used.

As shown above, forward regression will always make the estimated variable too high for low values of the controlling variable, and too low for high values of the controlling variable. This can be rectified by using an inverse approach to regression, but it is more complex, and if the age indicator is not continuous then a different error distribution to the gaussian distribution should be adopted. Furthermore, forward regression will produce estimates for unknown quantities which reflect the distribution of the estimated variable, as demonstrated by Bouquet-Appel and Masset (1982). Inverse regression does not.

Finally, many age estimation techniques use age observations from multiple independent sources. Multiple inverse regression is possible, but is a very difficult statistical procedure to perform.

The sort of regression models which would be suitable as a general form of regression for age estimation would be weighted multiple inverse regression. Technically this is possible, but very complex, and, it still does not address properly the problem of non-continuous age variables. The next chapter introduces an alternative approach to estimation from empirical data.
which makes none of the assumptions of regression analysis, and therefore is not subject to the same inherent problems as regression analysis.
5Age Estimation and Bayes’ Theorem

Regression and regression-like methods have been the mainstay of forensic and anthropological age estimation since the early decades of this century. However, most age indicators are scored as discrete variables, and although regression methods are to some degree robust to this, they are inherently unsuited to providing estimates when almost all the assumptions made (see previous Chapter) are violated. For some time non-parametric methods have been used in the estimation of age structure for a population. For example, Kimura and Chikuni (1987) describe the age-length-key for the estimation of the age structures for populations of fish, which uses a method of statistical analysis based upon a theory of probability called Bayes’ theorem.

5.1.1 Background to Bayes’ theorem

The Bayesian paradigm involves three important concepts - prior probability, likelihood and posterior probability (Phillips, 1973, p.57-60).

The prior probability is the initial assignment of the probability of any hypothesis being true before experimental evidence is considered. In age estimation this would be the probability of an individual belonging to a defined age group given no information (other than that the individual is similar to the members of a reference population). This probability is given the notation $Pr(A_i)$ which is the probability of the individual having an age which falls into age category $A_i$.

The likelihood is the conditional probability of a specific value of a variable being observed given the hypothesis is true. In our case this is the probability of an individual having a specific value of the indicator variable (an attrition score of 2, for example) given that the age category into which the individual falls is known. This is given the notation $Pr(I|A_i)$ which is the probability of an individual having indicator variable $I$ given they fall into age category $A_i$.

The posterior probability is the conditional probability of a hypothesis being true given the value of an observed indicator variable. In notation this is $Pr(A_i|I)$, the probability that the individual falls into age category $A_i$ given an observed indicator variable $I$. This is the final probability that an individual belongs to age category $A_i$ after taking into account both prior information.
from the reference population and observed evidence from the indicator variables.

Bayes' theorem states that the posterior probability is proportional to the prior probability multiplied by the likelihood. The constant of proportionality is given by the reciprocal of the sum over all age categories of the product of corresponding prior probabilities and likelihoods.

Using the notation described above, this can be written as:

\[
Pr(A|I) = \frac{Pr(A_i) \times Pr(I|A_i)}{\sum_i Pr(A_i) \times Pr(I|A_i)}.
\]

The principle is best illustrated with a simplified example. Suppose there were a reference population of 35 individuals, 17 of whom were young and 18 old. These individuals also had recorded for them two biological features which were thought to be related to whether they were young or old. These biological features could be either present or absent. These data are summarised in Table 1.

<table>
<thead>
<tr>
<th>Individual</th>
<th>Age</th>
<th>Indicator 1</th>
<th>Indicator 2</th>
<th>Individual</th>
<th>Age</th>
<th>Indicator 1</th>
<th>Indicator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>18</td>
<td>old</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>19</td>
<td>old</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>3</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>20</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>21</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>5</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>22</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>6</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>23</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>7</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>24</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>8</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>25</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>9</td>
<td>young</td>
<td>A</td>
<td>A</td>
<td>26</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>10</td>
<td>young</td>
<td>A</td>
<td>P</td>
<td>27</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>11</td>
<td>young</td>
<td>A</td>
<td>P</td>
<td>28</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>12</td>
<td>young</td>
<td>A</td>
<td>P</td>
<td>29</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>13</td>
<td>young</td>
<td>P</td>
<td>P</td>
<td>30</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>14</td>
<td>young</td>
<td>P</td>
<td>P</td>
<td>31</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>15</td>
<td>young</td>
<td>P</td>
<td>P</td>
<td>32</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>16</td>
<td>young</td>
<td>P</td>
<td>P</td>
<td>33</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>17</td>
<td>young</td>
<td>P</td>
<td>P</td>
<td>34</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual</th>
<th>Age</th>
<th>Indicator 1</th>
<th>Indicator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>old</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>old</td>
<td>A</td>
<td>P</td>
</tr>
<tr>
<td>20</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>21</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>22</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>23</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>24</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>25</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>26</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>27</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>28</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>29</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>30</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>31</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>32</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>33</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>34</td>
<td>old</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

*Where: P=present, A=absent*
If for the moment we concentrate solely on the first age indicator, and draw a cross-tabulation of age categories against this indicator we obtain Table 2.

**Table 2. Data for age and indicator 1**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Young</th>
<th>Old</th>
<th>Row Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Present</td>
<td>5</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>Column Totals</td>
<td>17</td>
<td>18</td>
<td>35</td>
</tr>
</tbody>
</table>

From the data in Table 2 it is possible to extract five different statements of probability for three distinct types of probability. First there is what is termed ‘joint’ probability (\(Pr(A, I)\)) which is the probability associated with selecting an individual at random from the sample and that individual being of a specific age and possessing a specific age indicator. For example: the probability of selecting an individual at random from the sample and that individual being young and possessing the age trait is 5 in 35. As can be seen this is merely the frequency in the appropriate cell in Table 2 divided by the sample total, and is given in Table 3.

**Table 3. ‘Joint’ probabilities for age and indicator state**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Young</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>12/35</td>
<td>2/35</td>
</tr>
<tr>
<td>Present</td>
<td>5/35</td>
<td>16/35</td>
</tr>
</tbody>
</table>

The next type of probability is the unconditional probability (sometimes termed marginal), which is the probability of selecting randomly an individual from the sample who is of a specific age, or the probability of selecting randomly an individual possessing a specific age indicator. For instance: the unconditional probability of an individual being old is 18/35, or the unconditional probability of an individual having the age-related trait is 21/35, and in notation terms is \(Pr(A)\) for the unconditional probability for an age, or \(Pr(I)\) for the unconditional probability of possessing the indicator. For this example unconditional probabilities are given in Table 4.
The final type of probability is the conditional probability which can be the probability of age given indicator state, or the probability of indicator state given age; these are denoted \( Pr(A|I) \) and \( Pr(I|A) \), respectively. For example: the conditional probability of any randomly selected individual having an indicator which is present, given the fact that they belong to the old age category (\( Pr(I=\text{present}|A=\text{old}) \)) is the number of individuals from the sample who have the age indicator divided by the number of individuals from the reference sample who fit in the old age category: i.e. 16/18. These conditional probabilities are given in Table 5.

Table 5. Conditional probabilities for age and indicator state

| Indicator | Age | \( Pr(A|I) \) | Age | \( Pr(I|A) \) |
|-----------|-----|---------------|-----|---------------|
|           | young | old | young | old |
| absent    | 6/7   | 1/7 | 12/17 | 1/9 |
| present   | 5/21  | 16/21 | 5/17 | 8/9 |

The probability of an individual belonging to the old age category given that the indicator variable is absent can be calculated from Equation 5,

\[
Pr(A=\text{old}|I=\text{absent}) = \frac{18/35 \times 2/18}{(18/35 \times 2/18) + (17/35 \times 12/17)},
\]

\[
Pr(A=\text{old}|I=\text{absent}) = \frac{18 \times 2}{(18 \times 2) + (17 \times 12)} = \frac{2}{14} = \frac{1}{7},
\]

which can be cross-checked against the entry in part one of Table 5. As the probability that any individual selected at random from the sample, for whom the age indicator is absent is 1/7, then if we assume that the sample is representative of a wider population, any individual from that population with the age indicator absent has a 1/7 probability of belonging in the old category; and therefore a 6/7 probability of belonging in the young age category if, the age categories are exclusive and exhaustive. ‘Exclusive’ means that each age category is separate, no individuals belonging to two
age categories, and exhaustive means that the individual appears in only one of the chosen categories with a probability of 1.

On its own, in a simple situation such as the illustration above, being able to calculate the probability that an individual belonging to a pre-defined age group using Equation 5 is of no great value, as the conditional probabilities for individuals belonging to age categories, given indicator state, can be calculated for all indicator states, as in Table 5, without recourse to Equation 5. However, a Bayesian interpretation becomes more useful as more indicator variables and more categories of age are employed.

5.1.2 Multivariate interpretations of Bayes' theorem

The above analysis can be applied to the situation where more than one variable has been observed, as is the case with Gustafson’s dental age traits, which requires the use of cross-tabulated data. In this case each permutation of age indicators is tabulated by age group, and the likelihoods calculated from the proportion of the reference population occurring with that particular permutation of age indicators for all age groups. An example can be seen in the work of Koningsberg and Frankenberg (1992, p.244), who were interested in assigning proportions of a sample population to age categories based upon a reference population. Again this is best illustrated by an example. If both indicators are now considered for the data in Table 1, and we draw up a cross tabulation of indicator 1 against indicator 2 controlling for age, then we get Table 6.

Table 6. Cross tabulation of indicators 1 and 2 controlling for age

<table>
<thead>
<tr>
<th>Indicator 1</th>
<th>Indicator 2</th>
<th>Indicator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>absent</td>
<td>present</td>
<td>present</td>
</tr>
<tr>
<td>absent</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>present</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>17</td>
</tr>
</tbody>
</table>

From Table 6 probabilities for the joint indicator events given age can be calculated \(Pr(I_1I_2|A)\), these form Table 7.
Table 7. Joint probabilities for indicators 1 and 2 given age

| Indicator 1 | Indicator 2 | Pr(I₁, I₂ | A_young) | Pr(I₁, I₂ | A_old) |
|-------------|-------------|-----------|-----------|----------|
| absent      | absent      | 9/17      | 1/18      |
|             | present     | 3/17      | 8/9       |
| present     | absent      | 0         | 3/17      |
|             | present     | 5/17      | 0         |

As with the example when only indicator 1 was considered, Equation 5 can be used to calculate the probability of an individual randomly selected from the sample being old or young given our knowledge of both indicator 1 and 2. If, for instance, there was an individual for whom indicator 1 was present, and indicator 2 also present, from Equation 5 we know:

\[
Pr(A = \text{young} | I₁ = \text{present}, I₂ = \text{present}) = \frac{17/35 \times 5/17}{(17/35 \times 5/17) + (18/35 \times 8/9)},
\]

\[
Pr(A = \text{young} | I₁ = \text{present}, I₂ = \text{present}) = \frac{17/35 \times 5/17}{17/35 + 18/35} = \frac{17}{35} \times \frac{5}{21}.
\]

Or, conversely, as the categories old and young are exhaustive and exclusive \(Pr(A=\text{old}|I₁=\text{present}, I₂=\text{present}) = 16/21\).

This approach was used by Koningsberg and Frankenberg (1992), but instead of estimating the age for any individual they adapted the technique to estimate the proportions of individuals in populations belonging to each age category. They demonstrated that unless an age indicator was perfectly correlated with chronological age, then estimates of the age structure for the target population would always resemble the age structure of the reference population. They solved this problem by applying a procedure derived from the expectation maximisation algorithm which gives the age structure of a population most likely to display the age indicators of the target sample. However, it is still true that if a traditional cross-tabulated approach is adopted the final probabilities assigned to an individual belonging to the given age categories may be very heavily influenced by the prior probabilities, which in turn would be nothing but a reflection of the reference population’s age structure (Koningsberg and Frankenberg, 1992).

When considering real data there is a problem presented by the relatively small numbers of individuals represented in reference samples. In the case of Gustafson’s (1950) data there are six ordinal variables, each with
four possible states. This gives $4^6 (=4096)$ possible permutations, the
distribution of which could not typically be estimated with even the largest
reference population which comprises only a few thousand individuals.

A more practical disadvantage is that of missing variables. To adequately
cover all possible combinations of Gustafson’s six variables would require 36
different tabulations. For the age estimation scheme envisaged by Acsádi and
Nemeskéri (Acsádi and Nemeskéri, 1980) which has only four variables,
Sjøvold (1975) tabulated every combination of variables which amounted to
sixteen tabulations.

These seemingly intractable problems can be overcome if an assumption
is made about the nature of the age indicators used. This assumption is that
all the variables are conditionally independent given age. Conditional
independence in this instance means that the variables may be correlated
with age, which is necessary for them to be age indicators, but not with each
other when age is controlled for. For instance in terms of the internal
structures of the tooth the formation of secondary dentine should not be
influenced by the formation of sclerotic dentine. Both get more pronounced
with age, but they are not directly dependent upon each other. If this
assumption can be made, and as pointed out above it is an implicit
assumption inherent in both the linear and multiple regression techniques
conventionally applied to these data, then there is an easy solution to the
problem of assigning probabilities for age to unknown individuals.

This solution consists of treating each variable as a separate univariate
modifier of the prior probability in order to produce a posterior probability
which acts as the prior probability for the next stage.

An example would be a continuation of the calculation made in the first
example to take into account the information given by indicator 2. From the
data in Table 6 we can produce a table like Table 5, but this time for indicator
2.

| Indicator | Age Pr(A|I) | Age Pr(I|A) |
|-----------|------------|------------|
|           | young      | old        | young      | old        |
| absent    | 9/10       | 1/10       | 9/17       | 1/18       |
| present   | 8/25       | 17/25      | 8/17       | 17/18      |

Table 8. Data for age and indicator 2
As we have already calculated the probability of any individual selected randomly from the sample being old, given that indicator 1’s state was absent, was 1/7, so we shall use this as the prior probability in Equation 5 to derive a posterior probability for that same individual being old, given that indicator 2 is in the present state. Similarly we know the prior probability of the individual being young to be 6/7. From Equation 5 and Table 8 we know,

\[
Pr(A = \text{old} | I_1 = \text{absent}, I_2 = \text{present}) = \frac{\frac{1}{7} \times 17/25}{\frac{1}{7} \times 17/25 + \frac{6}{7} \times 8/17} = \frac{289}{1489}.
\]

Hence the probability that the individual is young is 1200/1489, which is an increase from 1/7. We can say that our knowledge of the state of indicator two has modified our belief about the age of the individual. It is this explicit use of data to modify opinion which is at the core of a Bayesian approach. We first start by assuming that the population from which the individual came has the same distribution of ages as the reference sample. We then make an observation which we believe to be age-related about the individual, and use our knowledge about the same observation from the sample to modify our first opinion about that individual’s age. Then another observation is used to yet again modify this second opinion to produce our final opinion. In practice we can use as many age-related observations and as many age categories as we wish, which was the strategy adopted by Lucy et al. (1996) who classified individuals into ten year age categories based upon five dental age variables. If multiple age classifications are employed instead of the dichotomous age classification used in the example above, then the result is in effect a probability distribution of age for an individual. Estimates of confidence can be calculated by examination of this distribution. As the sum of the probability for all age classes is unity if the age classes are exclusive and exhaustive, then an estimate of age and confidence can be obtained from the 2.5th, 50th and 97.5th percentiles determined by linear interpolation.

The advantage with this scheme is that it gives a probability distribution, based upon a reference sample, for individuals of unknown age. Age estimates, and more importantly confidence intervals, will be unique to each individual, not some generic confidence interval as given by a regression model which would be common to all individuals. The production of a probability distribution carries a further advantage in that some combinations of age indicators are genuinely ambiguous, in that some sets of indicators make an individual look young, whereas other sets make them appear old.
This will produce probability distributions which may be bi-modal. Regression models will always give a single age estimate and assume a uni-modal distribution of confidence about that single age, which may not truly reflect the information given by the data.

If a particular age observation is missing for an individual, then it is easy to simply miss the variable out and only use those pieces of information which are available to the investigator. Using a full cross-tabulation scheme with some missing information is more difficult to handle. The differing amounts of information about age given by each age indicator (some being better than others e.g. root dentine translucency is more related to age than root resorption) is automatically taken into account; therefore those which are better indicators of age changing the probability distribution more than those which are not so good. Many age-related indicators have a distribution with age which is heteroscedastic, again this is automatically accounted for in this scheme. Ageing variables which give precise information in only one age range will only change the probability distribution in that age range.
Also, if we look at the full cross-tabulation of the data for the example used above (Table 7), then we notice that there is a probability of zero for indicator 1 being present and indicator 2 being absent. If we happen to encounter an individual with this particular combination of age indicators then a fully cross-tabulated approach cannot strictly give age estimates, or rather they would be exactly the same as the initial prior probabilities. This is because no individuals in the reference sample have an indicator configuration of present-absent for indicators 1 and 2, respectively, hence the reference sample gives us no information about indicator configurations for which there is no match. We could say that the absence of any individuals in the reference sample with a particular combination of age and indicator is not real, and is due to the small sample size, thus we could assign a small non-zero probability to that particular combination. However, if we assume conditional independence, then age can be estimated for combinations of indicator states which do not appear in the reference sample. This has most impact when dealing with schemes which have many variables and which are categorised into a relatively large number of states and age categories, and when the reference sample is relatively small. For example, were there to be five indicator variables and 10 age groups, and we needed a mean of approximately 10 individuals in each cell, then if we did not assume conditional independence we would need a minimum of 500 \((10 \times 10 \times 5)\) individuals in the reference sample. On the other hand, if we assume conditional independence in this instance, we only need a minimum of 100 individuals.

In summary the method described above satisfies almost all the requirements outlined in the previous chapter for a statistical method by which to estimate age from biological indicators. It provides simple and rationally justifiable estimates of age from ordinal and categorical data which do not have to conform to requirements such as being normally distributed about age, and homoscedastic. Confidence intervals are easy to calculate for each case and the distribution produced does not have to be uni-modal, which means there is no ‘loss’ of information from the original data. Altogether a Bayesian approach to age estimation is a powerful, simple estimative procedure which makes few assumptions about the nature of the data with which it is used, or the data upon which estimates are to be based.
6 Kernel Density Estimation of Continuous Distributions

6.1 The problems with discrete variables

Though the Bayesian estimation technique outlined previously is clearly an improvement on regression and regression-like methods, there are still a number of drawbacks with it. Age is treated as a categorical variable, whereas it is in fact a continuous variable, and only discrete ordinal or categorical indicator variables are considered. Our ideal estimation technique outlined in Chapter 4 would treat age as continuous, and be able to make estimates from indicator variables on both the discrete and continuous scale of measurement. One way in which age could be treated as a continuous variable would be to decrease the size of each age category. Moreover, in the example given by Lucy et al. (1996), each age indicator had five possible states, and the reference sample was divided into nine age categories, this means about 180 individuals would be needed to accurately approximate the conditional distribution of score as a function of age. To approximate a continuous distribution age would have to be divided into about 100 categories. This implies that a reference sample of nearly 2000 individuals would be required. As most anatomical reference collections number several hundred individuals at most, it is clearly unrealistic at the present moment to get a biologically homogenous reference sample of this size.

Another problem arises as a consequence of treating age as a categorical variable. The features of the probability density function described by a histogram are very sensitive to both the ‘bin width’ (the width in years of each category), and where the boundary between bins falls. An example is the distribution of known ages for a sample of 71 maxillary central incisors kindly provided by Dr. Tore Solheim (see 14 for a tabulation of these data).
Figures 14 to 16 are histograms of the age distribution for individuals in this sample. The only parameters which have been altered are the bin widths and bin start points. Figure 14 shows the age distribution starting at zero years and with each bin being ten years in width. The overall distribution could be described as bi-modal, with the gap between the two individual distributions being between 41 and 50 years of age. Figure 15 shows the
same data, but this time the bins are 15 years wide. This distribution could be described as uni-modal. Figure 16 is a histogram of exactly the same data, using the same 10 year bin width as Figure 14, but this time the bins start in the middle of each decade. The distribution now appears to show a bi-modality centred around the 56 to 65 year age category. Despite these being exactly the same data, the distributions they give would be described as very different and would lead to widely varying age estimates.

Although the problem of the selection of optimal histogram parameters has been addressed in the literature (e.g. Scott, 1979), all three problems listed above can be overcome quite simply by using smoothing to produce an estimate of a continuous distribution which links age with any single state of a given age indicator. There are two different general approaches which could be used, parametric and non-parametric. Parametric smoothing could be used if there was some function which described the relationship between any single state of a given age indicator and age. Data from a reference sample could be used to set the parameters of this function. However, as there is no a priori family of distributions with which to model this relationship then non-parametric smoothing is the preferred approach, since less rigid assumptions are needed about the underlying nature of the data. Again there are two approaches which could be taken; some form of interpolation of the empirical distribution produced by data classified into age categories, or the more ad hoc kernel density estimation of these distributions. Because there are really too few data in most reference samples to provide categorisation of age into more than nine categories, any smoothing based upon interpolation will retain to some degree the ‘stepped’ features of the original empirical distribution (unless the smoothing is very pronounced, in which case there is a very real chance of losing features which really are part of that distribution). For these reasons it is kernel density estimation which is applied here.
6.2 Kernel density principles

From Wertz (1978) a general expression for a distribution estimated from sample data using kernel density estimation is:

\[
\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)
\]

where \(x\) in this case is age, \(n\) is the number of individuals in the reference sample, \(h\) is a smoothing parameter (sometimes called the window width), and \(K\) is a probability density function is called the kernel. If the kernel is defined as a continuous probability density function, then it is guaranteed that \(f\) is also a continuous probability density function. The function estimated in Equation 6 can be more easily visualised as the sum of small kernels centred on the age of each individual from the reference sample (see Figure 17).

Figure 17. Two kernels combining to produce a probability density function

6.2.1 Selection of kernel parameters

The most common function for \(K\) is a Gaussian distribution with scale parameter \(h\). Where:

\[
K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x^2}{h^2}\right)}.
\]
This implies $K$ will be defined by parameter $h$ which is a smoothing parameter, the magnitude of which is proportional to the smoothness of the final probability density function. There are some interesting properties of $h$. If $K$ is Gaussian, and $h$ tends towards infinity, then the derived probability density function will tend towards being a uniform probability density. If $h$ is large compared to the range of $x$, then the final derived probability density function of $x$ will approximate a single gaussian probability density function centred on $\bar{x}$. Figure 18 illustrates how critical the selection of kernel parameter $h$ is to the estimated probability density function $f(x)$.

*Figure 18. The effect of various kernel parameters on the age distribution function from Solheim’s 1st mandibular incisors*

Figure 18 is basically a kernel-based reconstruction of Figures 14 to 16, and is constructed as a sum of kernels for each individual age in the dataset with different values being used for the smoothing parameter ($h$). It is apparent that these probability functions do not vary as much as the histograms in terms of their essential features. We imagine that for the kernel parameter $h=3$ years is not smoothed enough, the probability density function $f(x)$ is bumpy and quite irregular. By contrast, kernel parameter $h=11$ years is overly smooth displaying no hint of the bimodality seen in the histograms. A suitable optimal value would seem to be between $h=5$ years and $h=7$ years.

Silverman (1986: p. 44) recommends for the more experienced practitioner selection of a window width ($h$) based upon subjective choice,
since for most applications actually inspecting several estimated functions of \( x \) may give more insight into the underlying structure of the data itself. However, Silverman (1986: p. 48-74) outlines several numerical methods for estimating the optimal kernel window width which may be more suitable for age estimation, especially when more than one variable is considered as an age indicator. This is because if several density estimates are to be made, then it is difficult to manually inspect each distribution and maintain consistency of treatment between variables.

Silverman’s (1986: p. 38-44) suggestion for an automated method of estimating an optimal kernel window width is based upon the fact that generalised kernel estimation methods will have error components attributable to both bias and random error. A compromise estimate of window width which minimises both these quantities for mono-modal and multi-modal distributions is given as (Silverman, 1986: p. 47):

\[
\hat{h}_{opt} = 0.67 \cdot \min (R \text{ or } \delta) \cdot \frac{1}{\sqrt{n}},
\]

where \( \delta \) is the standard deviation of the observed sample, and \( R \) is the inter-quartile range. In the example above (Figure 18) with \( n=77 \), \( \delta \) being the smallest of \( \delta \) and \( R \) with a value of 18.81 years, \( h_{opt} = 5.3 \) years, which comparing with Figure 18, would seem to produce a smooth density estimate whilst retaining the features of the density function.

6.2.2 Selection of bin width

In practice it is impossible to compute fully continuous functions using conventional digitally-based computing, as an infinite memory capacity which could be operated on at infinite speed would be required. It would be possible to employ an analogue computer were the calculation of truly continuous distributions needed, but these devices can create their own artifactual features (Gleick, 1978, p.246). Some form of ‘discretization’ must be employed, where the continuous function is in effect considered as a smoothed histogram with many narrow bins. Several ways of estimating the optimal number of bins in a histogram exist (Scott, 1979), but in this case as the probability density estimate is an estimate of a continuous function, then
the smallest measurable unit of the continuous variable immediately suggests itself as a ‘given’ bin width (Wertz, 1978, p. 76). In most anthropological samples ages are usually known to the year, where any individual may appear at any point in that year with equal probability. There seems, therefore, little advantage to be gained by using more than one bin to represent one year for any estimate of a given probability density function.

6.2.3 Bounded probability density functions

Human adult age estimation conforms to the likelihood estimation assumptions of exclusivity and exhaustivity, that is the assumption that an adult human is of one age, and one age only, and, that probabilities are calculated for all possible ages for an adult human. This implies that the probability density function of age is bounded by zero and by a maximum lifespan potential of a human, any estimated probability which falls outside these two limits is simply not possible. If a kernel $K$ is selected such that $K$ is Gaussian (an asymptotic function), then it would be possible to ascribe probability outside the possible range of ages for an adult human. An approach which may be used to overcome this problem would be to use a function $K$ which could not ascribe any probability to values outside these bounds. Such a distribution could be a beta distribution with parameters $p$ and $q$, (Evans et al., 1992, p.31) where:

$$K = \frac{(p+q-1)!}{(p-1)!(q-1)!} x^{p-1}(1-x)^{q-1}, \quad 0 < x < 1.$$  

With $h$:

$$h = \sqrt{\frac{pq}{(p+q)^2(p+q+1)}}.$$  

An advantage of the beta distribution is that it can take on various forms dependent upon the parameters $p$ and $q$. Figure 19 shows five probability density functions based upon beta distributions:
The beta distribution has properties which make it bounded at either end of its range, but there are some disadvantages. The first is that as the distribution is for most values asymmetric, in which the mean does not coincide with the mode. If this kernel were used and an observation made at one of the extremes of the range, it would be tantamount to saying that the observation has a greater probability of being elsewhere than the point recorded. This could be overcome by using the mode as the central point for the kernel.

Human adulthood can begin at 15 years, but generally human adult age estimation techniques are only used for those aged above 25 years, so the zero point is not quite the hard boundary it might seem. Likewise, the maximum lifespan potential is considered to be in the region of 100-110 years, but only a relatively small number of individuals live much beyond 90. Given that the adult human lifespan is only loosely bounded, and that few individuals are represented by the extremes, then any probability attributed to the impossible regions by Gaussian kernel estimation is likely to be negligible (see Figure 18) thus there is no real reason to base kernels on the beta distribution.

6.2.4 Bayesian modification of kernel density functions

If $x$ is age, a continuous variable, and $y_i$ is a discrete or categorical indicator score which can take on any of $n$ values, so that there is: $y_1, y_2, \ldots, y_i,$
each of which range from 1 to \( n \). Let \( Pr(y_i|x) \) be the conditional probability of the observed value of \( y_i \) of indicator variable \( i \) occurring for a given age \( x \), and let \( Pr(y_i) \) be the unconditional probability of observing a particular value of \( i \) in the whole reference sample. Further, let \( f(y_i|x) \) be the conditional probability density function of any indicator variable \( i \) for any given \( x \), and \( f(x) \) the unconditional probability density function for any age \( x \). An appropriate form of Bayes’ theorem is:

\[
f(x|y) = \frac{Pr(y|x) f(x)}{Pr(y)}
\]

which, if conditional independence is assumed can be written:

\[
f(x|y) = \frac{Pr(y_1|x) Pr(y_2|x) \cdots Pr(y_n|x) f(x)}{Pr(y)}
\]

where 
\[
Pr(y) = \int Pr(y_1|x) Pr(y_2|x) \cdots Pr(y_n|x) f(x) dx.
\]

However, we do not have \( Pr(y|x) \), we have \( f(x|y_i) \), but:

\[
Pr(y_i|x) = \frac{f(x|y_i) Pr(y_i)}{f(x)}.
\]

Hence Equation 9 becomes:

\[
f(x|y) = \frac{f(x|y_1) Pr(y_1)}{f(x)} \frac{f(x|y_2) Pr(y_2)}{f(x)} \cdots \frac{f(x|y_n) Pr(y_n)}{f(x)} \frac{f(x)}{Pr(x)}
\]

where \( f(x) \) can be estimated in the same way as \( f(x|y_i) \) using kernel density functions.

A graphical example is given in Figure 20 where data from central maxillary incisors (14) has been used to generate a posterior probability distribution.
First the prior probability is calculated as a ‘sum of bumps’ as in Figure 18 with kernel parameter (5.3 years) calculated from Equation 7. Another probability distribution function of age is calculated from the distribution of ages with the same score for secondary dentine (2) in the reference sample. This is again a ‘sum of bumps’, the window width being calculated from Equation 7 (8.3 years). These two distributions are then combined according to Equation 10 to produce a posterior probability density function for age.

The advantage with this method of estimating age is that the estimated ages come out as continuous probability distribution from which there is no loss of information.
6.2.5 Kernel density estimation and continuous variables

The idea of estimating a probability density function using a smoothing technique, like kernel density estimation, can be used to generate posterior probability distributions for situations where both variables are continuous.

If a kernel is defined for one continuous variable, for example $x$, where $x$ is age:

$$f(x) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{x^2}{2h^2}},$$

and $h$ is defined in Equation 7. A similar function can be defined for the continuous response variable, for example root dentine translucency, as $y$:

$$f(y) = \frac{1}{\sqrt{2\pi h_1}} e^{-\frac{y^2}{2h_1^2}},$$

where $h_1$ is defined for $y$ in the same way $h$ is for $x$. A new function can be defined which is a function of both $x$ and $y$,

*Equation 11*

$$f(x,y) = \sum_{i=1}^{n} f(x-x_i) f(y-y_i).$$

Which will be a surface proportional to the joint probability of $x$ and $y$. Functions which are proportional to the conditional probabilities $Pr(x|y)$ and $Pr(y|x)$ can be recovered by orthogonal transepts through the joint probability surface.
Figure 21. Joint probability surface for root dentine translucency and age from 71 maxillary incisors.

Figure 21 shows a joint probability surface calculated from the data for age and root dentine translucency for 71 maxillary central incisors used in Figure 20, and Equation 11. This is produced by calculating kernel window widths for age and root dentine translucency for all individuals in the reference sample. Then a three dimensional kernel is calculated as the product of the two kernels. This is a three dimensional Gaussian figure which looks rather like an upturned champagne bucket. For each individual in the reference sample one of these figures is added to the joint probability surface producing the sort of three dimensional figure seen in Figure 21. The conditional probability of age for any root dentine translucency measurement is then calculated from a section parallel to the age axis through this figure. Figure 21 shows this line at a root dentine translucency measurement of 7mm. The probability distribution represented by the line is shown in Figure 22.
The conditional probability distribution represented in Figure 22 is equivalent to the probability distribution in Figure 18, and can be treated in exactly the same way to produce posterior probability distributions seen in Figure 20.

6.3 Summary

The use of kernel density methods to estimate conditional probability distributions allows age to be considered as a continuous variable whilst the variables from which age can be estimated can be continuous or discrete. There is no need to devise separate methods for each type of variable as both are treated in the same framework.

Estimates of confidence are arguably as important as any actual point estimate of age. For instance, if an individual were to be ascribed an age of 54 ± 60 years it would mean that we know very little about the age of that individual, however, 54 ± 10 would mean we have fairly good knowledge of that individual’s age. Because the ages which are estimated are (by definition) unknown, an age estimation technique which is precise is preferable to one which is accurate. In the example cited above the technique which gives an age estimate of 54 ± 60 may be more accurate, in that it may have been demonstrated that point estimates derived from it have a lower absolute error in tests upon documented samples, but for undocumented samples the accuracy can never be known. A more useful estimate of age would be that which could be demonstrated to have a reproducible proportion
of estimates fitting into the smallest confidence interval. In other words the more important parameter for an age estimation technique is the size of the confidence intervals. Many physical anthropologists and forensic scientists have been misled into thinking that absolute accuracy is the most important property of an age estimation method.

As stated before, all regression models give an estimate of confidence based upon the uncertainty in the model and the spread of points about that model, hence very similar error estimates are given for all point estimates which use that model. On the other hand the Bayesian estimation technique gives confidence intervals on a case by case basis depending entirely on the reference data, and with a distribution which is derived from the data, not an assumed gaussian distribution. This means that Bayesian estimation methods generally (although not always) give an estimate which is a high probability localised to a small window in the solution space.
7
Relative Performances of Estimation Procedures

7.1 Selection of performance parameters

Although, as pointed out above, accuracy is an age estimation parameter of secondary importance to narrow confidence intervals for archaeological samples, it is, however, important for forensic cases and for the odd archaeological specimen for which a historical age can be identified, and is therefore an appropriate measure of the effectiveness of any estimation technique. An idea of how accurate, or rather how inaccurate, age estimates across a sample are can be simply calculated by the mean of the difference between the estimates of age and real age for the sample. That is:

\[
\text{inaccuracy} = \frac{\sum |\hat{x}_i - x_i|}{n},
\]

when:
- \(\hat{x}_i\) = estimated age for \(i^{th}\) individual
- \(x_i\) = known age for \(i^{th}\) individual

In simple terms the greater the inaccuracy the worse the estimation for the sample.

The other important parameter is the confidence interval for any estimate of age. Here this is the calculated confidence interval for 95% of estimates, and it is important that 95% of the true ages fall within the estimated confidence interval. For regression-based methods confidence is uniform, or nearly uniform (this is not necessarily so, but to all practical intents is true for age indicators). For likelihood-based techniques the confidence interval is specific to the individual estimate, so initially a mean of confidence intervals shall be used, but in practice when real age estimation data is employed it may be more important to examine the confidence intervals in more detail.

7.1.1 Simulation

Bi-variate \(x, y\) data were simulated to approximate the sort of data types used in age estimation. When both axes are continuous the \(x\)’s were selected with a uniform distribution between 0 and 100 to simulate age in years. Then \(y\) values which were Gaussian for any \(x\) about the model \(y = 0.1x + 6\) to approximate the measurement of root dentine translucency in millimetres.
The Gaussian component of \( y \) values was selected using a Box-Müller approximation (Evans et al., 1992, p.118).

Equation 12

\[
y_1 = \left[ \frac{1}{10} x_i \sqrt{-2 \ln R_1} + \cos \left( 2 \pi R_2 \right) \sqrt{V} \right] + 6
\]

\[
y_2 = \left[ \frac{1}{10} x_i \sqrt{-2 \ln R_1} + \cos \left( 2 \pi R_2 \right) \sqrt{V} \right] + 6.
\]

Where \( R_1 \) and \( R_2 \) are uniform random between 0 and 1, and \( V \) is the variance of \( y \) about \( x \).

Where \( y \)'s were from a simulated discrete distribution, but still Gaussian for any \( x \), the same approximation was employed but only using the integer component. This gave a discrete scale between 0 and 8 to approximate Johanson’s (1971) dental age changes. As \( V \) was a scaling factor the relationship between \( x \) and \( y \) could be closely controlled to produce simulated distributions with varying correlation.

7.1.2 Point estimate selection

Selection of a point estimate for regression-based techniques is simply a matter of calculating the appropriate \( x \) value from the calibration relationship between \( x \) and \( y \). Point estimates from a continuous posterior probability distribution are more difficult as there can be a range of ‘best’ points to choose from. The mean of the continuous posterior probability distribution is merely the mean of \( x \) values. This measure of point estimate would be very poor as it would not take into account the actual distribution of probability on \( x \). The mode may be a useful measure, but is not satisfactory if the posterior distribution were bi-modal. This could be resolved by using the largest of the modal values as the point estimate. However, during the course of this investigation it has been observed that when a posterior distribution is bi-modal the true value of age often lies between the bi-modal values, corresponding to the median of the distribution. The point estimate of age will therefore be defined as the median of the posterior probability distribution, which carries implications for how biological age markers operate in the adult human and this is discussed in Chapter Eight.
7.2 Results of simulation

There is a notion that the more highly correlated one variable is with another the better it will be for estimating values of that variable. It was therefore decided that the best way of summarising the relative performances of likelihood estimation and classical regression-based estimation would be to examine the inaccuracy and confidence intervals across a range of simulated data sets of differing x,y correlations.

7.2.1 Both variables continuous

Figure 23 and Figure 24 show the relative performances of likelihood and regression-based approaches to estimation with simulated data of varying x,y correlation. They were calculated by adjusting the value of $V$ in Equation 12 between zero and one, extracting 400 x,y pairs, calculating the Pearson correlation coefficient between x and y, then using the y values to make estimates for x. The inaccuracy and confidence intervals were then plotted against the correlation coefficient. Although in the course of earlier work Lucy et al. (1996) had employed a ‘jack-knife’ resampling strategy (Efron, 1982) to minimise the effect of making estimates for cases which had been included in the reference sample, it was found that even with only 70 cases the difference between ‘jack-knifed’ estimates and those which had not been resampled was negligible. As the results (below) are purely comparative it was decided to be more expedient not to resample.

Figure 23. Inaccuracy of likelihood and regression estimation procedures against x,y correlation when both variables are continuous
Two things are immediately apparent. First despite the fact that this is simulated data which should satisfy every requirement of both linear regression and likelihood estimation, the correlation coefficient is not as good a predictor of either inaccuracy or confidence interval as one might expect. Secondly, likelihood estimation is more accurate, and gives smaller confidence intervals for correlation coefficients below about 0.9. For correlation coefficients above 0.9 regression-based techniques give the least inaccuracy and smallest confidence intervals.
7.2.2 Age variable continuous and indicator variable discrete

Discrete data were simulated as described above; Figures 25 and 26 summarise the results.

*Figure 25. Inaccuracy against correlation coefficient for when age is continuous and age indicator discrete*

![Graph showing inaccuracy against correlation coefficient for when age is continuous and age indicator discrete.](image)

*Figure 26. Confidence intervals against correlation coefficient for when age is continuous and indicator is discrete*

![Graph showing confidence intervals against correlation coefficient for when age is continuous and indicator is discrete.](image)

There is very little difference between the situations where both variables are continuous and one variable continuous and one discrete. For any data which is correlated at under 0.9 likelihood-based estimation schemes give greater accuracy with smaller confidence intervals, regression-based
calibrations giving better estimates when the variables are correlated at above 0.9.

A point to be noticed is that real variables considered to be age indicators usually have correlations with age in the region of 0.7 to 0.85, the confidence intervals for either approach do not fall below 50 years at best, and at worst are 120 years. It may be considered that a confidence interval of 50 years is valueless for age estimation as the entire range of possibility is in the order of 100 years, and this may be true with a regression approach which applies a nearly uniform error to estimates for all individuals. However, the confidence interval of 50 years is only a mean for likelihood-based estimates, in reality these can vary between 1 and 100 years dependent upon the individual. Therefore, some portion of the estimates obtained by likelihood estimation will have small confidence enough intervals to be useful to the anthropologist.

7.3 Data derived from dental examination

A large dataset (≈720) was available from tooth sections (see 14). This gave an ideal opportunity to examine the relative performances of likelihood and regression-based estimation with real variables which were mixed continuous and discrete. The dataset was divided into tooth locus comprising measurements from the ten anterior teeth as the different loci are thought to develop at significantly different rates, age indicator development from each contra-lateral pair being considered equivalent (Solheim and Sundnes, 1980). The discrete variables used were occlusal attrition, secondary dentine formation, and root surface roughness, which were measured in the same way as Johanson did, employing six different divisions. Root dentine translucency was the continuous age indicator. Further details on sectioning and data acquisition can be found in Solheim (1989; 1992; 1993c).

For these ten datasets age estimates and associated estimates of error were derived from linear forward multiple-regression, and kernel density likelihood estimation as described previously. The regression technique had to be forward as there is no commercially available statistics package which will perform multiple calibration. The same indicators of relative performance as used above were employed to give some idea of how well estimates of age could be made, although as this is real data how well any age indicator
correlates cannot be controlled in the same way as for the simulations. It was again considered to be expedient not to resample.

The age indicators, estimates and confidence intervals are listed in full in 14.

Table 9. Summary performance indicators for dental data

<table>
<thead>
<tr>
<th>Tooth locus</th>
<th>number in sample</th>
<th>multiple correlation coefficient</th>
<th>mean inaccuracy (years)</th>
<th>mean 95% confidence interval (years)</th>
<th>percentage of known ages falling within the confidence interval</th>
<th>mean inaccuracy (years)</th>
<th>mean 95% confidence interval (years)</th>
<th>percentage of known ages falling within the confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>central mandibular incisors</td>
<td>76</td>
<td>0.7377</td>
<td>8.4</td>
<td>37.8</td>
<td>96.0</td>
<td>10.0</td>
<td>53.8</td>
<td>96.0</td>
</tr>
<tr>
<td>lateral mandibular incisors</td>
<td>77</td>
<td>0.7850</td>
<td>7.9</td>
<td>36.6</td>
<td>97.3</td>
<td>9.2</td>
<td>45.8</td>
<td>97.3</td>
</tr>
<tr>
<td>mandibular canines</td>
<td>74</td>
<td>0.7100</td>
<td>8.4</td>
<td>36.3</td>
<td>95.8</td>
<td>9.7</td>
<td>53.4</td>
<td>94.5</td>
</tr>
<tr>
<td>mandibular first pre-molars</td>
<td>69</td>
<td>0.7720</td>
<td>8.7</td>
<td>35.2</td>
<td>94.1</td>
<td>8.6</td>
<td>48.2</td>
<td>94.1</td>
</tr>
<tr>
<td>mandibular second pre-molars</td>
<td>72</td>
<td>0.8330</td>
<td>6.6</td>
<td>37.9</td>
<td>98.6</td>
<td>8.5</td>
<td>45.3</td>
<td>95.8</td>
</tr>
<tr>
<td>central maxillary incisors</td>
<td>74</td>
<td>0.7910</td>
<td>6.5</td>
<td>33.5</td>
<td>94.5</td>
<td>7.8</td>
<td>42.0</td>
<td>97.2</td>
</tr>
<tr>
<td>lateral maxillary incisors</td>
<td>69</td>
<td>0.8400</td>
<td>6.0</td>
<td>36.0</td>
<td>100.0</td>
<td>7.7</td>
<td>41.7</td>
<td>97.1</td>
</tr>
<tr>
<td>maxillary canines</td>
<td>72</td>
<td>0.8532</td>
<td>5.7</td>
<td>36.3</td>
<td>97.2</td>
<td>7.1</td>
<td>36.9</td>
<td>94.1</td>
</tr>
<tr>
<td>maxillary first pre-molars</td>
<td>78</td>
<td>0.8243</td>
<td>7.4</td>
<td>33.2</td>
<td>96.2</td>
<td>8.5</td>
<td>46.0</td>
<td>96.0</td>
</tr>
<tr>
<td>maxillary second pre-molars</td>
<td>77</td>
<td>0.8004</td>
<td>7.8</td>
<td>36.1</td>
<td>97.4</td>
<td>7.9</td>
<td>42.6</td>
<td>97.4</td>
</tr>
</tbody>
</table>

An encouraging aspect of the results given in Table 9 is that in most cases more than 95% of estimates fall within the estimated 95% confidence interval for both estimation techniques. Even in those instances where less than 95% fall within the expected estimation interval it never fell below 94%.

Given the multiple correlation coefficient of approximately 0.80 across all samples used here, one would expect from the simulation studies above that the mean inaccuracy would be in the region of 15 years and 23 years for likelihood and regression-based methods, respectively, and the mean confidence interval around 70 years and 100 years for each technique. In fact the inaccuracy is about 7 years for likelihood estimates, and 8.5 years for regression-based estimates. Similarly the 95% confidence intervals are about 35 years and 47 years, respectively. This lends some credence to the belief that skeletal biologists have that examination of many age-related traits leads to better estimates of age. Why this should be is a mystery as the total amount of information available from which to estimate age is not of necessity greater from any combination of age-related indicators than from any single age indicator, unless of course the assumption of conditional independence is satisfied. The biological interpretations and implications of this condition are examined more thoroughly in Chapter 8.

The performance of likelihood estimation is significantly better than that of regression-based estimation. The inaccuracy is slightly less, but more importantly the 95% confidence intervals are, on average, considerably
smaller. As the confidence intervals for the real data are much smaller than for the simulations of single indicator this means that adult age estimation may have value for the skeletal biologist. A 95% confidence that any individual lies within a 30 year range, from a total range of approximately 100 years, is still not very satisfactory for anthropological or forensic purposes, but it is considerably more useful than the 80 year range derived from simulation.
8 Biological Aspects of Age Estimation

8.1 Are age indicators conditionally independent?

Many of the points raised in the preceding discussion (parts 4 to 7 inclusive) about estimating age for a given set of hard tissue features revolve around the notion of statistical independence of those skeletal features. As explained in Chapter 5 all age indicators for a given individual are by necessity correlated as they are dependent on age, and would not be age indicators if they did not change with age. However, statistical correlation is not the same as dependence. In age estimation we are interested in whether the different age indicators can be treated as independent sources of information about age, or whether age indicators are dependent on one another. In the latter case it would be incorrect to regard two indicators, one of which was dependent on the other, as containing two individual sets of information about age since as in reality there is only be one set of information. This aspect of age estimation is closely mirrored in the gerontological literature in the discussion of the proximal causation of senescence.

8.1.1 Proximal causes of senescence

The modern position on proximal causes of age have been dichotomised into two polar opposites for convenience by Wood et al. (1994), and much of the forthcoming discussion relies upon this simplification.

The opposites are the gerontological view, and the epidemiological view. The gerontological view holds that age, and eventual death, are the consequence of a single causal determinant which can be exogenous or endogenous. The gerontological position embodies the concept of ‘biological age’ as distinct from chronological age. The epidemiological view sees age-related processes as independent strands all operating separately to produce age-related phenomena. For an individual treating age-related phenomena from the epidemiological viewpoint the term ‘biological age’ has no meaning as there are many ‘biological ages’, one for each parameter which changes with age.

\[8\] Proximal being used in the sense described by Mayr (1982) to describe causal mechanisms as opposed to ‘ultimate’ causes which are evolutionary. An example would be bio-chemical causes of age change.
Currently there are about 14 distinct proximal mechanisms suggested for mammalian age changes (Morse and Rabinowitz, 1990), most of which can be interpreted from either the gerontological view, or the epidemiological view.

- **Clinker:** with ageing there is an accumulation in non-dividing cells of age pigments such as lipofuscin and ceroid (Cerami *et al.*, 1987), which cannot be metabolised by the cell, and which inhibit the affected cell's activity. However, although it is true these pigments do appear in old cells, any deleterious effect has not yet been demonstrated (Fujimore, 1989).

- **Thermal denaturation:** over time in the post-mitotic (no longer dividing) cells certain proteins are denatured, the idea being that metabolic processes are disrupted on a cellular level by defective proteins (Gershon, 1987). Unfortunately, it has been pointed out that denatured proteins are easily scavenged by the cell’s enzymes, leading to the observation that cells with high protein damage can recover (Rothstein, 1987).

- **Cellular loss:** loss of about 40% of post-mitotic cells from muscles, heart, brain and other vital organ systems has been held to be at least partly responsible for senescence at the organismal level; although not the liver which can regenerate indefinitely (Popper, 1987). Just why these cells die is unknown, but could be due to some of the other mechanisms listed here, but is unlikely to be due to the Hayflick limit as even aged individuals have only undergone something like thirty cell divisions (Hayflick, 1987).

- **Hormonal and enzymatic changes:** although there is a decrease in the activity of many enzymes and coenzymes with age, some enzyme systems, particularly hydrolytic enzymes found in the cell organelles called lysosomes actually increase. Lysosomes are involved with the destruction of infective microorganisms, and in the removal of the body’s own cells. Changes in sex hormones lead to reproductive senescence. A lowering of adrenocortical steroid levels lead to reduced protein production, with possible consequences for cellular loss, and implications for bulk bone loss in osteoporosis (Finch, 1987).

- **DNA deterioration:** considers random flaws in cellular DNA, caused either exogenously or endogenously over time, to be responsible for defective protein replication which prevents the cell carrying out its normal metabolic functions (Setlow, 1987). Although chromosomal damage has been shown to accumulate over time, any corresponding reduction in cellular function has not been demonstrated (Hanawalt, 1987).

- **Somatic mutation:** is where both mitotic and post-mitotic cells accumulate random mutations, the vast number of which are functionally deleterious. As the number of mutated cells increases tissue function decreases, leading to dysfunction on the organism. However, when cells divide quickly they throw off a number of mutations none of which can compete with the surrounding somatically normal cells, and quickly die off leaving only the normal cells. Hence mutated cells do not partake in ageing processes.

- **Radiation:** radiation reduces viable cell numbers, often killed cells are replaced by connective tissue. Background radiation is held by some to be responsible for many of the mechanisms outlined above. However, background radiation mainly affects mitotic cells, whereas many of the age changes are seen only in post-mitotic cells.

- **Changes in the immune system:** mammals have three immunological subsystems, antibodies, B-lymphocytes and T-lymphocytes. Antibodies and B-lymphocytes provide protection against external attack, T-lymphocytes
scavenge abnormal somatic cells, providing protection against tumours and other abnormal tissue. Mammals also produce auto-antibodies which can attack normal cells. With increased age there is a decrease in activity of B-lymphocytes and T-lymphocytes, and increase in the numbers of auto-antibodies which has been held to provide a mechanism for bulk bone loss (Siskind, 1987).

- **Programmed ageing**: the belief that senescence is under some sort of central control, and that this control is in some way genetically programmed so that there is a gene set for an ordered ageing process. The evidence for this is compelling in semelparous organisms, but not quite so in iterioparious organisms such as mammals (Martin, 1987).

- **Nutrition**: it has been suggested that the metabolic processes of those individuals on low calorie diets quickly lead to harmless end products, mainly carbon dioxide and water. On the other hand it is thought that individuals with high calorie diets produce many more intermediate compounds, such as aldehydes, which promote age changes such as collagen cross-linking. It is also thought that cumulative intakes of metal ions, such as copper and zinc, also lead to cross-linking compounds (Morse, 1988).

- **Rate of living**: this considers the mammalian body to only have so much energy, the higher the rate of living the greater the toll on what is thought to be a finite resource. The problem is that unlike the final imagoes of some insects which emerge with no mechanism by which to intake nutrient, there is no reason why mammals cannot carry on indefinitely. Some estimates of cell replacement suggest that the human body is replaced every seven years, although some non-dividing cells, such as nerve cells, are not.

- **Wear and tear**: this compares the mammalian organism to some inert mechanical device where age changes are simply due to accumulated metabolic insults throughout life. Again, as with the rate of living idea, higher metazoans have constant replacement of tissue throughout life (Russell, 1987).

- **Cross-linkage**: cross-links form between large metabolic molecules, such as proteins, with age. They can be created by exogenous factors such as radiation, or endogenous factors like other metabolic by-products. Cross-linked molecules take no further part in any metabolic processes, hence accumulate in cells, and form physical blocking mechanisms to other processes lowering the function of those cells, and leading to eventual cell death (Fujimore, 1989). Cross-linkage of the epidermal proteins, such as collagen and elastin which form the main connective tissues, leads to many of the age changes seen in mammalian skin (Rolandi et al., 1991).

- **Free radical formation**: free radicals are chemical species with odd electrons in low energy orbits. In mammalian systems these are mainly chemically highly active forms of oxygen and are formed by radiation and the organism’s own metabolic processes during respiration. Unsaturated fats which form a large component of cell walls are particularly vulnerable to lipid peroxidation by free radicals, and lead to eventual cell dysfunction. It is assumed that free radical damage continues throughout life, but that the organism can have some form of defence through production of free radical scavenging agents such as ascorbic acid (vitamin C), and super-oxide dimutase (SOD) (Harman, 1987). Oxygen free radicals have also been associated with genetic damage (Saul et al., 1987) which can lead to tumour formation (Pryor, 1987), and can provide a mechanism for many of the proximal causes of senescence and death given above (Armstrong et al., 1984).

---

9 Semelparious refers to organisms with a sharply defined reproductive period after which they die. Iterioparious organisms have an extended reproductive period, reproducing many times (see 1.2.1).
Most of the above proximal mechanisms, clinker, thermal denaturation, hormonal and enzymatic changes, DNA deterioration, somatic mutation, immune changes, cross linkage and free radical formation, can all lead to cellular loss, which can be in the next order down of proximal explanation. However, nutrition, wear and tear, radiation, programmed ageing and rate of living belong on the next tier up. It is this next level up which gives confusion between the gerontological and epidemiological positions. We have no idea whether a programmed ageing resides at the top of the causal hierarchy, or stochastic factors such as radiation, and there is no knowledge of how many further levels there are in this causal hierarchy. The only thing which is clear is that if the high level causal factors behind ageing have been correctly identified, and these are the only high level causal factors, then there are at least 25 equally plausible orders in which they can be placed. Illustrated are two flowcharts which best illustrate the extreme gerontological and epidemiological positions.

Figure 27. causal connections for a strong gerontological view of mammalian ageing

Figure 27 shows how the various biological causes of ageing relate to each other viewed from a strong gerontological view. The biological programme, which is held to control all the proximal mechanisms is at the top
of the causal chain. In this scheme all age-related changes are held to be expressions of this single variable (with the exception of those which are entirely due to exogenous factors).

Figure 28. Causal chain of mammalian age changes seen from the epidemiological view

In Figure 28 we see the same causal elements from the epidemiological viewpoint. In this case there are four causal factors at the top of the hierarchy, all those mechanisms further down the tree are dependent in some way upon these, and are an expression of these. In this case there is a strong argument for independence between the proximal causes in the first level of the causal order.

The order has ramifications for age estimation. If the picture of mammalian ageing seen in Figure 27 is correct, and an intrinsic age program is responsible for most age changes, then it would be incorrect to treat measurable age changes such as hormonal changes and oxygen free radical formation as independent expressions of age as they are all dependent on, and expressions of, the same biological process, namely the underlying ageing programme. In this case they could be treated as independent
measures of the biological programme, but not of chronological age itself. The case put forward in Figure 28 would suggest that hormonal changes and free radical formation would be expressions of an underlying ageing programme, and generalised wear and tear. This means that these two measurable quantities can be employed as independent expressions of age.

### 8.1.2 The biology of hard tissue age changes

The causal nexus surrounding age-related human hard tissue changes is parallel to that outlined above, and has implications for the use of mathematically-based estimation procedures. Basically, if the idea of a biological age, as separate and distinct from chronological age is biologically meaningful because of some underlying unitary cause of change, then many of the multivariate methods of age estimation will not have the assumption of independence fulfilled. If, on the other hand there is no unitary cause of hard tissue age change, and changes in these tissues are independent or fall into groups which are independent, then the assumptions of multivariate methods of age estimation are fulfilled.

Human age traits have often been cited as being either developmental for all ages up to young adult, and degenerative for older individuals (Bass, 1987, p.12; Iscan and Loth, 1989, p.23; Stewart, 1979, p.175). Age traits such as epiphyseal union are certainly the result of the normal programmed human growth cycle, and can therefore legitimately be called developmental. It is difficult to see how age changes such as cranial suture obliteration can bring about the loss of function to the organism entailed by the term ‘degenerative’. Of all the age-related changes outlined in this chapter only bulk loss of bone is clearly dysfunctional. Cranial fusion and auricular surface fusion if completed could in fact lead to a more robust structure. It could be said that completion of the pubic symphysis in child-bearing females could have adverse consequences during partuation, and were the sternal end of the ribs to entirely fuse to the sternum then the reduction in rib cage movement could arguably inhibit breathing (although most thoracic expansion relies on the diaphragm, and pubic fusion could again make a more robust bone for females in the post-reproductive phase of life). However, these changes are not exactly degenerative, the body does not get to a perfect state then start to loose function. Pubic fusion and sternal rib end ossification seem more to be the slow continuation of processes already in existence. Therefore, the common conception that most skeletal age changes are
degenerative is a meaningless value judgement applied to traits which are in no way dysfunctional, and probably has its origin in that these skeletal changes are associated with other changes in human physiology which are dysfunctional. The same is true of dental age changes. Attrition is clearly a function of wear and tear on the occlusal surfaces, so is dysfunctional and exogenous. However, regular secondary dentine is formed as a response to physical insult and as a normal age process, its role being adaptive (Stanley, 1983). Periodontal recession and cementum build-up are linked and have been seen as an eruption process to compensate for occlusal attrition (Whittaker, 1988; 1992; 1990). Root dentine translucency is another dental age change which has an adaptive component in that it prevents the passage of pathogens into the pulp complex (Fish, 1932). With the exception of occlusal and inter-proximal attrition no dental age changes are dysfunctional and in fact seem in part to be the human body’s response to various insults, or as part of normal development.

From the above argument it seems unlikely that human hard tissue changes are merely the result of day to day wear, but of some programmed response to external changes, and of other more obscure programmed responses. The problem is that there is no knowledge of what controls the age changes, how many control centres there are, and at what level these control centres operate.

The simplest model which can be imagined has a single control programme for all systematic hard tissue development, and a programme which responds to various bodily insults. This model is analogous to the gerontological view and would hold that age changes in the hard tissue of the human organism are largely under the control of a single unitary clock, and is depicted in Figure 29.
The implications for age estimation if this model is close to being a true model for age changes in human hard tissue are that all age changes are manifestations of three causes, two of which are environmentally determined. Any age which is estimated is only an estimate of a comparative stage in the programme which governs age changes, which will have a relationship to chronological age, but will not be an estimate of chronological age. In this instance it is impossible to estimate chronological age directly as knowledge of chronological age will always be mediated by the rate of the governing programme, which in itself may be a function of idiosyncratic variation between individual organisms and environmental factors.

An opposing model would be an analogue of the epidemiological view with many controlling programmes. In the biology of human hard tissue there is some support for this position. Teeth and bone, despite superficial similarities, are not the same. In the early evolution of higher organisms there seem to have been two skeletons, an exoskeleton and an endoskeleton. The last vestiges of the exoskeleton seems to be the teeth, the endoskeleton remaining the support structure. This explains the different ontogeny of bone and teeth (Hall, 1992). It also explains why different specialist cells, which are
descended from different primary cells (neural arch and mesenchymal cells for teeth and bone, respectively), are involved with the creation and maintenance of these tissues. If there are different programmes for different organ systems, then this would be the obvious place to look. There could also be a different controlling programme for each major bone system, so one could suggest the cranial vault had a separate programme to the innominate, which in its turn is separate to the rib cage. This position is illustrated in Figure 30.

*Figure 30. Causal hierarchy for human skeletal age changes from a strongly disparate view*

Here there are seven independent causes for observed hard tissue age changes, hence seven independent measures of chronological age, five of which are not environmentally determined. This means that each hard tissue age change is indirectly giving information about the chronological age of any individual.

For the former model (Figure 29) the concept of biological age has a definite meaning in terms of the development processes of an organism. In the latter (Figure 30) the idea of biological age may be useful, but it is by no means clear just what the age would refer to as there would be a multitude of these ages for any given individual.
8.1.3 Expectations rising from the independent and disparate views of hard tissue ageing processes

If the unitary view of hard tissue ageing processes is the more correct then it would be expected that, for a sample of individuals variation in measurable age marker would arise from sample variation between the different expressions of the underlying biological programme. This means that were an individual to have an age marker which was advanced for a given age, then we would expect all other markers for that individual to be also advanced for that age. In contrast, the independent view would predict that age markers are independent measures of chronological age, so advancement in one marker would not necessarily entail advancement in another (Wood et al., 1994, p. 34). This gives a basis for investigating which model is correct, or which combination of age indicators are independent. Partial correlation will tell us how one variable interacts with another, significant correlations obtained when controlling for age will give some basis by which to attribute causal relationships between age indicators.

8.1.4 Testing dependency between age indicators

Data from a collection of documented skeletal material from Amsterdam was kindly donated by Maja d’Hollosey (1989), for which various attributes of the pubic symphysis and auricular surface had been recorded (see 12.1.2). The measured attributes were summed for each individual to give an overall score for both pubic symphysis and auricular surface. The overall scores were then correlated with each other whilst controlling for age.

For dental tissues Gustafson’s data (1950) was employed (see 12.1.1). Partial correlations for all six variables were recorded whilst controlling for age, significance being assigned after Bonferroni\textsuperscript{10} correction.

After controlling for age the correlation between the developmental stages in the pubic symphyses and auricular surfaces was 0.46 (n=24, p=0.018), which is a significant correlation between these two variables.

The controlled correlation matrix is presented as Table 10.

\textsuperscript{10} Bonferroni corrections compensate for the probability of getting a significant result when testing for multiple hypothesis from a dataset (Milton and Arnold, 1995. P550-552).
Table 10. Partial correlations for Gustafson's (1950) dental indicators

<table>
<thead>
<tr>
<th>Correlation coefficients controlling for age, probability beneath (d.f. = 38)</th>
<th>Attrition</th>
<th>Cementum buildup</th>
<th>Peridontal movement</th>
<th>Root dentine translucency</th>
<th>Secondary dentine buildup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root resorption</td>
<td>-0.27</td>
<td>0.05</td>
<td>-0.12</td>
<td>-0.19</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>0.76</td>
<td>0.48</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td>Secondary dentine buildup</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.33</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.79</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root dentine translucency</td>
<td>0.05</td>
<td>-0.08</td>
<td>-0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.60</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peridontal movement</td>
<td>0.20</td>
<td>-0.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cementum buildup</td>
<td>-0.03</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As there are some 36 different combinations of correlations, 15 of which are being tested (6 are irrelevant and 15 duplicates), if we set the significance level at 0.95 (95%), then by chance alone we stand a 0.95^{15} (= 0.46) probability of assigning a significance to a correlation where there is none. To correct for this we shall allow a Bonferreri correction to give a 0.15 probability of assigning a significant correlation where there is none, which puts the level at which we shall assign significance at 0.01 (1%) (Milton and Arnold, 1995, p. 551). This means that none of the above correlations are significant, although the highest partial correlations are between secondary dentine, cementum build-up, and root resorption, which would be consistent with the view that these are somehow related processes.

8.2 Conclusions

The assumption of conditional independence, as pointed out in Chapter 4, is fundamental to all statistical methods which are useful for small reference samples for recovering age from a set of age indicators. Whether this assumption is justified, and at what level it is justified, will to some extent dictate how age for adult humans can be estimated, and is dependent upon an understanding of the underlying biology of the markers used for age estimation. For instance, the fact that there appears to be a very high correlation between changes in the pubic symphysis and auricular surface of the innominate suggest that there is a single biological mechanism which controls both of these changes. This means that age changes recorded in these skeletal features are measures of that process, not of chronological age.
However, the lack of any distinctive correlation seen in dental age changes would mean that there is no single process which controls development of the mature dentition, rather that there is a mosaic of sub-processes, all of which are independent measures of chronological age.

The implications of these results for age estimation are that dental age indicators could be used separately and independently, but data from the innominate would have to be pooled in some way to remove dependence. This suggests a modified version of Figure 30 as follows:

*Figure 31. Modified causal flowchart for skeletal and dental age indicators*

This (Figure 31) represents the causal connections between the age indicators considered here.

The above analysis is by no means definitive, being only an example of how some biological structures can be deduced from age indicators, and shows what effect our knowledge of these structures may have upon the way in which we estimate age. The analysis for the innominate is only based upon a reference sample of 27 individuals, and 40 for dental indicators. Neither sample contained information about both innominate and dental age indicators. Any future attempt should really be based upon a sample containing many more individuals (>1000) and contain data for all hard tissue age indicators commonly used for estimating age.
NConcluding Remarks

This thesis has examined i) how estimates of age affect our view of evolutionary processes, ii) why anthropologists’ estimates of age have until now been mostly incredible, iii) possible ways in which regression analysis has been used which can be shown to give unlikely age estimates, iv) several ways of estimating age using none of the traditional assumptions and v) how the biology of the ageing hard tissue of the human body affects how anthropologists can go about age estimation.

One of the problems is that many forensic scientists and biological anthropologists have regarded age estimation techniques as possessing their own essence, that is as systems which will produce age estimates which are somehow derived from nature without the assumptions and lines of thought usual in any scientific venture. This overly empirical approach leads to a methodology which does not focus critical attention on the underlying biological processes or statistical procedures. This thesis has attempted to redress this deficiency to some extent by emphasising throughout the fact (recently commented on by Bouquet-Appel and Masset, 1996) that all age estimation techniques are merely systematic ways of getting from observations which are classified through some ideal scheme, to making an inference based upon others of a similar type, and that in age estimation studies little attention has been given to the system itself. However, it would be unfair to entirely blame forensic scientists and biological anthropologists for what has been seen as a certain statistical conservatism (Konigsberg et al., 1994), since existing commercial statistical packages such as SPSS base their point estimated values upon forward regression.

9.1.1 The palaeodemographic paradox

There is one last biological point which has not been addressed, and it is one which is critical to the whole process of adult human age estimation. Up until now we have assumed that the only parameter which affects age change is age. We have given no thought to whether those areas of human hard tissue used for age estimation are plastic in the way other parts of the human skeleton are in that they react to the stresses and strains of everyday life. More importantly, this point affects archaeologists’ historical interpretation of the lifeways of past human populations. As discussed in Chapter One, the archaeological populations of both Isbister (Hedges, 1982), and Libben
(Howell, 1982) seemed to have populations which were unfeasibly young. The conclusion was that for these individuals to have died so young they must have led a particularly hazardous life. But what if the age indicators are somehow responsive to environmental strains, then the populations at these two sites could have been considerably older, yet lead particularly hazard free lives? This could lead to an analogue of the ‘osteological paradox’ promulgated by Wood et al. (1992), which could be stated thus: many of the age indicators are pathological and wear related; if a population was virtually free of these factors because it was suffering little hardship, then we see that skeletal population as being very young. We assume a high mortality, and consequently postulate an extremely hard life for them. Conversely, were a population to be under some stress which produced age markers which were advanced, then we would say that they were all very old when they died, and consequently we would say what an easy life they had. This is possibly one of the harder questions to examine, and is a consequence of archaeologists’ tendency to (usually by necessity) employ only one body of information from which to draw archaeological conclusions. Most modern reference populations for which age is known come from culturally homogenous Western populations. Even in those such as the Terry collection (see 11.1.1), which contain individual’s from different socio-economic groups, the socio-economic group does not by necessity relate to the factors which would give differences in the development rates of age indicators. The different socio-economic groups represented in the Terry Collection are broadly from different geographical origins, usually African American equates to low economic status, European American higher economic status. Any perceived difference between the rate of age indicator development would have to be partitioned between geographical origin and economic status. The question of to what degree age indicators are affected by conditions and activity throughout an individuals life could single-handedly negate any attempt to estimate adult human age, and underlines the necessity of actually understanding the biological processes of hard tissue age change for its own sake, rather than treating age indicators as a clock.
9.1.2 Can adult human age estimation ever be useful?

Whether it will ever be possible to estimate human adult age at death precisely, and with a level of accuracy which would be useful to anthropologists, is a debatable point, and one on which we may now have some insight. Bouquet-Appel and Masset (1982) stated that to make anthropologically useful estimates of age at death one would really require an indicator with a correlation with age of about 0.9. Inferring from the graphs in Chapter seven it can be seen that with a single age indicator a correlation of 0.9 would give us at best an average confidence interval of 50 years, and a mean inaccuracy of 15 years. Even extrapolating from the data presented in Chapter seven for multivariate age indicators if the indicators have a multiple correlation coefficient with age of 0.9 then the best we can expect would be a mean inaccuracy of about 5 years, and a mean confidence interval of nearly 30 years. In other words the most important parameter for unknown specimens, the width of the 95% confidence interval, never gets significantly below one third of the total expected lifespan. In view of this it may be that Bouquet-Appel and Masset (1982) were being less than gloomy about the future of age estimation for individuals from the archaeological record, and that in reality the aims of the forensic approach to adult human age estimation can never be fully realised.

9.1.3 Further work

It is obvious from the foregoing that no existing reference dataset of age-related data contains an adequate number of individuals. The most adequate are those samples of clinical origin from Scandinavia as employed by Bang and Ramm, and Solheim. This is a consequence of the relative scarcity of documented historical individuals and the legal framework in most Western nation states which disallows systematic work of this nature on recent cadavers. This situation may change with greater land pressure in centres such as London, leading to a demand for 18th and 19th century crypts to be cleared, such as occurred at Spitalfields (Molleson et al., 1993). This may act in the anthropologists’ favour, if it were seen as more useful to use the ihumations for serious scientific purposes than to merely exhume and dispose the human remains by unceremonious incineration.

There also needs to be a change of focus in the research aims of those looking at age at death from hard tissue indicators. The deeply empirical
approach which treats age indicators as a calibration device misses the point to some degree. Age indicators are active physiological parameters giving more information than just the age of an individual. They also give an insight into the underlying control mechanisms of human hard tissue. This change in focus needs to be extended to the evolution of the human lifespan. Our notions of how lifespan evolves has to work for all biological species, including humans. Our current understanding of the evolution of the human lifespan contradicts to some degree evolutionary theories which are compelling, and seem to work for all other biological species.

The final point is statistical. In this thesis a unified approach has been outlined to making estimates for a variable (continuous or discrete) from one or more variables, each of which can be continuous or discrete, which makes fewer assumptions than traditional regression-based approaches. Initial results are promising, but there are bound to be problems in the detail of how to apply these methods. A further problem lies in their dissemination and wider adoption, making it necessary for computer routines which match a commercial standard to be written and made available to the wider world outside the confines of anthropology.
References


List of Anthropological Collections in the West.

Update 4: 18th June 1995

11.1.1 The Terry Collection

The person to contact is Dr. David Hunt (MNHAN080@SIVM.SI.EDU) The Collections Manager of the Physical Anthropology Division, Department of Anthropology, National Museum of Natural History, The Smithsonian Institution, Washington D.C.

Dr. Hunt writes:

For the Terry Collection. Robert Terry began this collection in the early 1920's in response to the apparent lack of good osteological collections by which to teach skeletal biology and anatomy. The collections were made from morgues throughout St. Louis, Missouri of primarily indigents who's bodies were not claimed by family members, and thus became property of the state to be disposed of as seen fit. Those that were turned over to Terry were used for soft tissue dissection in cadaveral training at Washington University Medical School. When the bodies came into the lab, they came with a morgue record, the cadaver was autopsied, and a pathologist report made. Photographs of the cadaver were generally, but not always made, as well as plaster death masks approximately 850 of the individuals. Hair and Skin samples were also take for a good portion of the collection. All these records and samples are housed here at the museum.

A portion of the individuals were also willed bodies, some not all the people represent low social classes, but the majority were more impoverished. As you are aware, there is a good number of the individuals who came from TB sanatoriums, this is also another reason for the access of the collection, since for some periods of time, the states did not release TB victims to the families for burial but disposed of them to medical research for the TB problem.

When Terry retired in 1946, another well known skeletal biologist, Mildred Trotter, continued this collection on into the early 1960's until when she retired. We can than Mildred for her efforts in balancing out the demography of the collection by focusing on younger individuals and especially, white females, who were very much underrepresented. At the final termination of the project, nearly 1800 skeletons had been collected.

Anticipating her retirement, and knowing the disinterest that the Anatomy department had with the collection in the 1960's since they had begun to focus much more on brain morphology and function in the department, Mildred began correspondence with her long term friend, T. Dale Stewart concerning began correspondence with her long term friend, T. Dale Stewart concerning the possibilities of transferring the collection to the National Museum for permanent curation. In 1965 this transaction
was made between Washington University and the Department of Anthropology, and thus the collection is here now.

The collection now consists of 1732 specimens of known age, sex, ethnic origin, cause of death, pathological conditions, body measurements for many of the individuals as well as photographs, plaster death masks, hair and skin samples, pathologist reports and other ancillary information.

The demographic spread is:

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Males</td>
<td>463</td>
</tr>
<tr>
<td>Black Males</td>
<td>550</td>
</tr>
<tr>
<td>White Females</td>
<td>323</td>
</tr>
<tr>
<td>Black Females</td>
<td>390</td>
</tr>
<tr>
<td>Asiatic Males</td>
<td>5</td>
</tr>
<tr>
<td>Unknown origin</td>
<td>1</td>
</tr>
</tbody>
</table>

Age ranges from 16 yrs to 102 years with the highest proportions in the 45-50 and greater ages, but there are sufficient numbers in the lower age ranges for all groups except for a large number of young white females under 27 yrs. I understand there is a small collection of remnant skeletons still at Washington University, but these are not well maintained and not usually used or easily accessed.

11.1.2 The Hamilton Collection

Ontario, Canada. Comes from 19th Century burials? This one has been reburied but data has been carefully collected prior to the reburial. The contact is Dr. Shelley Saunders (Saunders@mcmail.cis.mcmaster.ca). Latest estimates of size are some 380 documented individuals.

11.1.3 St. Brides Church

About 200 individuals from 19th Century vault burials so all sexes and ages, teeth have been sectioned before and written up by Ivanhoe (1982; 1994). Contact is Dr. Louise Scheuer of the Royal Free Hospital, London. Technically the local clergy have absolute authority on this, but Dr. Scheuer usually administers requests.

Dr. Louise Scheuer  
Department of Anatomy and Cell Biology  
United Medical and Dental Schools of Guys and St. Thomas  
Royal Free Hospital  
London Bridge  
London  
SE1 9RT
11.1.4 Grant Collection

The keeper is:
Dr. Patty Stuart-Macadam
Department of Anthropology
University of Toronto
Toronto
Ontario
M5S1A1
Canada

e-mail pstuart@epas.utoronto.ca

Dr. Stuart-Macadam writes:
There are 202 individuals who had been received by the anatomy department between 1928 and the early 1950’s, largely from local hospitals and welfare institutions. Name, sex, age at death, and cause of death are known for individuals. There are 175 males, 27 females, most (147) are over 40 years of age and had been either transients, migrant workers, or recent immigrants without relatives. With one exception all were white, and on the basis of surnames and available records, of European origin.

11.1.5 Raymond Dart Collection

A collection of about 2000 skeletons in South Africa. Most of these were poverty stricken Black Africans. For some ‘immediate’ causes of death are known, but as most of the immediate causes of death are cardiac arrest this can be fairly useless. Of more value are a few with contributory causes of death e.g.: syphilis. The person to contact is:

Professor Maciej Henneberg
Chair of Anatomy and Human Biology
Department of Anatomy and Human Biology
University of Witwatersrand Medical School
Parktown 2193
South Africa

e-mail o55majs@witsvma.wits.ac.za
fax 643 4318
11.1.6 The Huntington Collection

The keeper is Dr. David Hunt (see under Terry Collection)

Dr. Hunt writes:

This collection was amassed by Dr. George Huntington during his tenure as an anatomy professor at the College of Physician and Surgeons, New York City from the 1880's to his later years in the 1920's. The collection consists of primarily European Immigrants and New York City residents who died in any number of the boroughs of the city, and from sanatoriums and poor house morgues. These individuals were property of the state and therefore able to retrieved by the college for cadaver training.

Because most of these individuals were used for teaching, not all the skeletal elements were preserved (as was done in the Terry Coll.) and thus the individuals are consistently not complete, especially missing ribs, sterna, crania, or crania that are complete, and in a good deal of the cases represented by only one side of the body. However, the preservation and documentation of the over 3600 individuals makes this collection very interesting and considerably valuable. All catalogued individuals have names, known sex, age, ethnic origin (and usually nationality), cause of death, and in some but not all cases heights, place of collection.

There are fascinating examples of trauma, infectious, tumours, and congenital diseases. Other problems with this collection at this time are the lack of access to the entire collection due to space problems at the museum. It is hoped that there will be more space allocated to Physical Anthropology in the near future to allow for the rehousing of portions of this collection from cold storage. The other deficit is that the collection is now stored by element rather than by individual, which makes individual study difficult, but research by element a boon since one can study a particular element in a series of several hundred with little effort.

The demography is primarily white males, (about 70%) and US black males making up most of the remainder, but there are representative females as well. More specifics on the demography in the near future.

11.1.7 The Cobb Collection

About 600 skeletons from African-Americans from the Washington DC area, now housed in Howard University, Washington DC.

The person to contact is:

Professor Blakey
Department of Sociology and Anthropology
PO Box 987
20059 USA.
11.1.8 The Hamman-Todd Collection (formerly known as the Western Reserve University Collection)

The keeper is:
Dr. Lyman Moore Jellema
Collections Manager
Department of Physical Anthropology
Cleveland Museum of Natural History
1 Wade Oval Drive
University Circle
Cleveland
Ohio
44106-1767 USA
(tel: 0101 216 231 4600)

Dr. Moore Jellema writes:

The Hamman-Todd collection is available for study. The human skeletal material consists of over 3000 cadaver derived human skeletons collected between 1912 and 1938 by Drs. C.A.Hamman and T.W.Todd. The majority of these specimens have documentation giving stated age, sex, race, height and weight. In addition we have soft tissue measurements, drawings of anomalies, x-rays, and photographs of the cadaver before dissection.

In addition there are non-human primate specimens. Dr. Moore Jellema writes:

The nonhuman primate material consists of over 900 specimens including 500 gorillas and chimpanzees, 20 orangs, 50 gibbons and representatives of every other primate genus. Since the majority of these specimens were wild shot between 1910 and 1938, the documentation is not as extensive as that of the human material. However we do have approximate age, sex, and collection location for many of the specimens.

11.1.9 Ferraz de Macedo Collection

Housed at the Museu Bocage of the Faculty of Sciences, Lisbon, Portugal. This is a collection of nineteenth century skeletons. Our sources tell us it was partly destroyed by fire in the 1970’s, but not how many individuals there are, nor how many are documented.
11.1.10 Eusébio Tamagnini Collection (Coimbra)

Apparently a very large collection of known age, sex, occupation and place of birth skeletons established by Eusébio Tamagnini some time after 1907 (again our sources do not tell us how large). As physical anthropology is still taught at the University of Coimbra, Portugal, one assumes the contact address is the Department of Physical Anthropology there.

11.1.11 The St. Barnabas Collection

The keeper is:

David Lucy
Department of Archaeological Sciences
University of Bradford
Bradford
West Yorkshire
U.K.
BD7 1DP
d.j.lucy@bradford.ac.uk

The St. Barnabas collection from the crypt of St. Barnabus, West Kensington, London, U.K. is composed of carefully sampled parts of the skeletons of 58 individuals. There are 24 male and 31 female and 3 unknown skeletons represented. All but the 3 unknown have names and date of birth and death attached. These individuals were buried between 1830 and 1853 and were from the upper middle classes of London at the time.

The sampling strategy was designed and executed by Dr. Sue MacLaughlan-Black of the Department of Anatomy, University of Aberdeen, Scotland, U.K. They were originally intended for DNA extraction but it was found that any DNA had been effectively destroyed by the decomposition of the oak coffins which contained the bodies. Because of the stringent requirements of DNA sampling the samples were taken with sterile tools, stored in double sterile, sealed bags. The samples have since been kept frozen and are available for destructive work.
12 Data Used for Causal Model Testing in Chapter Eight

12.1.1 Dental data

Gustafson’s (1950) data. Contains information on the age changes noted by Gustafson for 41 documented individuals as outlined in 2.

<table>
<thead>
<tr>
<th>Tooth number</th>
<th>Real Age (years)</th>
<th>Attrition</th>
<th>Periodontosis</th>
<th>Secondary Dentine</th>
<th>Cementum Apposition</th>
<th>Root Resorption</th>
<th>Transparency</th>
<th>Total points</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>39</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>45</td>
<td>23</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>33</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>35</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>30</td>
<td>37</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>38</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>38</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>42</td>
<td>39</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>21</td>
<td>39</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>23</td>
<td>45</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>25</td>
<td>45</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>44</td>
<td>48</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>38</td>
<td>48</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>41</td>
<td>48</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>22</td>
<td>49</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>26</td>
<td>49</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>17</td>
<td>50</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>19</td>
<td>51</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>18</td>
<td>51</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>46</td>
<td>52</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>43</td>
<td>52</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>40</td>
<td>52</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>15</td>
<td>53</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>24</td>
<td>55</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>47</td>
<td>55</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>29</td>
<td>64</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>64</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>65</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>16a</td>
<td>69</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>16b</td>
<td>69</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>
12.1.2 Innominate data

The data given here are from d’Hollosey (1989) and are observations from the auricular surface and pubic symphysis for 26 nominate bones. The morphological features for each surface are separated into traits in a similar way to McKern and Stewart (1957), but using a greater number of individual traits.

For the auricular surface these are:

- Grain and density: scores between one and five referring to the texture of the bone surface. One being a fine surface texture, five a rough texture.
- Microporosity: referring to the prevalence of small holes in the surface. Scored between one and three, one being under half the area possessing small holes, three being greater than half the area.
- Macroporosity: like microporosity, but refers to larger holes.
- Billowing: being the presence and coverage of regular undulating structures on the surface of the bone. Scored between one and five, one being where undulating structures are marked and extensive, five meaning a complete absence of this feature.
- Apex: the definition of the apical part of the auricular surface. Scored between one and four, one being the best defined, four being indistinct.
- Trabecular area: extent of the formation of osteophytes around the margins of the surface. Scored between one and four, one is a complete absence of bone forming activity, four where osteophyte formation is strong and clear.
- Transverse organisation: refers to an organisation running orthogonal to the billowing mentioned above. Scored on a scale of one to three, one refers to a strong pattern, three an absence of any transverse organisation.

For the pubic symphysis the traits are:

- Billowing: similar to that for the auricular surface, but scored between one and three, one being the most pronounced, three the least.
- Definition of the extremities: refers to how well the upper and lower edges of the symphysis are defined. Scored between one and three, one being not developed, three being very sharp.
- Ventral bevel: presence or absence of a defined ventral bevel. One being absent, two being present.
- Texture: referring to the surface of the bone. Scored between one and three, one being rough, three smooth.
- Dorsal margin: how well defined the dorsal margin of the pubic symphysis is. Scored between one and four, one refers to a complete lack of definition, four a highly defined structure.
- Dorsal osteophytes: whether there is bone forming activity upon the dorsal margin. Scored between one and three, one refers to whether there is osteophyte formation along less than half of the dorsal rim, three to most of the rim.
• Attachment of tendons: refers to boney indented tendon attachment. Scored between one and three, one meaning no tendon attachment, three strongly developed attachment points.

• Attachment of ligaments: similar to tendon attachment, but for ligament attachment.

• Erosion: the amount of surface pitting across the bone. Scored between one and three, one being an absence of pitting, three being abundant pitting.

• Trauma: the presence or absence of noticeable trauma to the surface of the bone.

<table>
<thead>
<tr>
<th>Individual attributes</th>
<th>Aricular surface scores</th>
<th>Pubic symphysis scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skeleton number</td>
<td>Real age</td>
<td>A</td>
</tr>
<tr>
<td>20702.00</td>
<td>50</td>
<td>1</td>
</tr>
<tr>
<td>21701.00</td>
<td>74</td>
<td>5</td>
</tr>
<tr>
<td>123701.02</td>
<td>83</td>
<td>3</td>
</tr>
<tr>
<td>123702.01</td>
<td>78</td>
<td>3</td>
</tr>
<tr>
<td>257701.00</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>293701.00</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>293702.00</td>
<td>32</td>
<td>2</td>
</tr>
<tr>
<td>293703.00</td>
<td>47</td>
<td>2</td>
</tr>
<tr>
<td>294702.00</td>
<td>69</td>
<td>4</td>
</tr>
<tr>
<td>344701.00</td>
<td>63</td>
<td>3</td>
</tr>
<tr>
<td>344702.00</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>428704.00</td>
<td>39</td>
<td>2</td>
</tr>
<tr>
<td>431701.00</td>
<td>50</td>
<td>4</td>
</tr>
<tr>
<td>438702.00</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>483702.00</td>
<td>59</td>
<td>4</td>
</tr>
<tr>
<td>489701.00</td>
<td>55</td>
<td>2</td>
</tr>
<tr>
<td>490704.00</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>493701.00</td>
<td>59</td>
<td>2</td>
</tr>
<tr>
<td>494704.00</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>502703.00</td>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td>528701.00</td>
<td>60</td>
<td>4</td>
</tr>
<tr>
<td>530706.00</td>
<td>82</td>
<td>4</td>
</tr>
<tr>
<td>532704.00</td>
<td>72</td>
<td>5</td>
</tr>
<tr>
<td>534701.00</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td>536702.00</td>
<td>39</td>
<td>4</td>
</tr>
<tr>
<td>536703.00</td>
<td>73</td>
<td>4</td>
</tr>
<tr>
<td>547701.00</td>
<td>72</td>
<td>3</td>
</tr>
</tbody>
</table>

Where:
- Grain and density = A
- Microporosity = B
- Macroporosity = C
- Billowing = D
- Apex = E
- Treburecular area = F
- Transverse organisation = G
- Billowing = H
- Definition of extremities = I
- Ventral bevel = J
- Texture = K
- Dorsal margin = L
- Dorsal osteophytes = M
- Attachments of tendons = N
- Attachment of ligaments = O
- Surface pitting = P
- Trauma = Q
### Summary Information For Chapters Two and Three

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Collection</th>
<th>Total</th>
<th>Indicator</th>
<th>Scored by</th>
<th>Statistical methods</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Todd</td>
<td>1920</td>
<td>Western Reserve</td>
<td>300</td>
<td>pubic symphysis</td>
<td>stages</td>
<td>direct</td>
<td>European American males</td>
</tr>
<tr>
<td></td>
<td>1921</td>
<td>Western Reserve</td>
<td>90</td>
<td>pubic symphysis</td>
<td>stages</td>
<td>direct</td>
<td>African American Males</td>
</tr>
<tr>
<td></td>
<td>1921</td>
<td>Western Reserve</td>
<td>69</td>
<td>pubic symphysis</td>
<td>stages</td>
<td>direct</td>
<td>European and African American Males</td>
</tr>
<tr>
<td>McKern and Stewart</td>
<td>1957</td>
<td>Korean war dead</td>
<td>450</td>
<td>pubic symphysis</td>
<td>components</td>
<td>direct</td>
<td>young males</td>
</tr>
<tr>
<td>Meindl, Lovejoy, Mensforth and Walker</td>
<td>1985</td>
<td>Hamman Todd</td>
<td>96</td>
<td>pubic symphysis</td>
<td>stages</td>
<td>PCA and regression</td>
<td>young males</td>
</tr>
<tr>
<td>Dwight</td>
<td>1890</td>
<td>unknown</td>
<td>100</td>
<td>cranial suture</td>
<td>description</td>
<td>direct</td>
<td>young males</td>
</tr>
<tr>
<td>Parsons</td>
<td>1905</td>
<td>dissecting room</td>
<td>82</td>
<td>cranial suture</td>
<td>description</td>
<td>direct</td>
<td>European-American</td>
</tr>
<tr>
<td>Todd and Lyon</td>
<td>1924</td>
<td>Western Reserve</td>
<td>307</td>
<td>cranial suture</td>
<td>stages</td>
<td>direct</td>
<td>mostly young individuals</td>
</tr>
<tr>
<td>Powers</td>
<td>1962</td>
<td>unknown</td>
<td>111</td>
<td>cranial suture</td>
<td>stages</td>
<td>direct</td>
<td>mostly young individuals</td>
</tr>
<tr>
<td>Perizonius</td>
<td>1984</td>
<td>University of Amsterdam skeletal collection</td>
<td>256</td>
<td>cranial suture</td>
<td>stages</td>
<td>regression</td>
<td>East Africans and European Africans</td>
</tr>
<tr>
<td>Meindl and Lovejoy</td>
<td>1985</td>
<td>Hamman Todd collection</td>
<td>236</td>
<td>cranial suture</td>
<td>stages of specific sites</td>
<td>regression</td>
<td>East Africans and European Africans</td>
</tr>
<tr>
<td>Zivanovic</td>
<td>1983</td>
<td>Hamman Todd collection</td>
<td>572</td>
<td>cranial suture</td>
<td>stages</td>
<td></td>
<td>East Africans and European Africans</td>
</tr>
<tr>
<td>Lovejoy, Meindl, Pryzbeck and Mensforth</td>
<td>1985</td>
<td>Hamman Todd, Libben and autopsy</td>
<td>764</td>
<td>auricular surface</td>
<td>stages</td>
<td></td>
<td>comparing unknown to known</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Collection/Site</td>
<td>Sample Size</td>
<td>Bone/Lesion/Structure</td>
<td>Analysis Method</td>
<td>Biological Sex</td>
<td>Reference Sample Notes</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>------------------------------------------</td>
<td>-------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>Murray and Murray</td>
<td>1991</td>
<td>Terry Collection</td>
<td>189</td>
<td>auricular surface</td>
<td>stages</td>
<td>mixed sex</td>
<td>African Americans and European Americans</td>
</tr>
<tr>
<td>Iscan, Loth and Wright</td>
<td>1984</td>
<td>autopsy</td>
<td>93</td>
<td>sternal rib end</td>
<td>stages</td>
<td>direct</td>
<td>scores</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>118</td>
<td>sternal rib end components</td>
<td></td>
<td></td>
<td>European American males</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mixed sex European Americans</td>
</tr>
<tr>
<td>Walker and Lovejoy</td>
<td>1984</td>
<td>Hamman Todd</td>
<td>130</td>
<td>bone density</td>
<td>stages</td>
<td>regression</td>
<td>mixed sex</td>
</tr>
<tr>
<td>Walker and Lovejoy</td>
<td>1985</td>
<td>Hamman Todd</td>
<td>105</td>
<td>bone density</td>
<td>stages</td>
<td>regression</td>
<td>treated males and females separately</td>
</tr>
<tr>
<td>Nemeskéri, Harsanyi and Acsádi</td>
<td>1960</td>
<td>autopsy cases</td>
<td>105</td>
<td>pubic symphysis</td>
<td>developmental stages</td>
<td>cross tabulation</td>
<td>separated males and females</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>402</td>
<td>cranial suture</td>
<td>developmental stages</td>
<td></td>
<td>tabulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>105</td>
<td>bone density</td>
<td>developmental stages</td>
<td></td>
<td>used x-radiographs on several long bone sites</td>
</tr>
<tr>
<td>Brooks</td>
<td>1955</td>
<td>Western Reserve</td>
<td>185</td>
<td>pubic symphysis</td>
<td>stages</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>80</td>
<td>cranial suture</td>
<td>stages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brothwell</td>
<td>1981</td>
<td>mixed archaeological</td>
<td></td>
<td>occlusal attrition</td>
<td>stages</td>
<td>direct</td>
<td>unknown reference sample</td>
</tr>
<tr>
<td>Takei</td>
<td>1970</td>
<td>modern Japanese</td>
<td>1000</td>
<td>occlusal attrition</td>
<td>stages</td>
<td>strange</td>
<td></td>
</tr>
<tr>
<td>Hongwei</td>
<td>1980</td>
<td>modern Chinese</td>
<td>880</td>
<td>occlusal attrition</td>
<td>stages</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Drier</td>
<td>1994</td>
<td>modern Chinese</td>
<td>107</td>
<td>occlusal attrition</td>
<td>stages</td>
<td>regression</td>
<td>included unknown specimens in the reference sample</td>
</tr>
<tr>
<td>Miles</td>
<td>1963</td>
<td>archaeological</td>
<td>32</td>
<td>occlusal attrition</td>
<td>stages</td>
<td></td>
<td>developed internal calibration</td>
</tr>
<tr>
<td>Mays, de-la-Rua and Molleson</td>
<td>1995</td>
<td>archaeological</td>
<td>293</td>
<td>occlusal attrition</td>
<td>stages</td>
<td>regression</td>
<td>solely archaeological samples internal calibration</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Sample Type</td>
<td>Age</td>
<td>Measurements</td>
<td>Method</td>
<td>Results Note</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>------</td>
<td>------------------------------</td>
<td>-----</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Moore</td>
<td>1970</td>
<td>modern extracted</td>
<td>unknown</td>
<td>secondary dentine</td>
<td>pulp chamber measurements</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Padayachee and van Wyk</td>
<td>1988</td>
<td>modern extracted</td>
<td>50</td>
<td>secondary dentine</td>
<td>dentine measurements</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Solheim</td>
<td>1992</td>
<td>modern extracted</td>
<td>1000</td>
<td>secondary dentine</td>
<td>pulp chamber measurements</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Drusini</td>
<td>1993</td>
<td>modern extracted and archaeological</td>
<td>166</td>
<td>secondary dentine</td>
<td>pulp chamber measurements</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Stott, Sis and Levy</td>
<td>1982</td>
<td>cadavers</td>
<td>3</td>
<td>cementum rings</td>
<td>ring counting</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>Charles, Condon, Cheverud and Buikstra</td>
<td>1986</td>
<td>modern extracted</td>
<td>42</td>
<td>cementum rings</td>
<td>ring counting</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>Condon, Charles, Cheverud, and Buikstra</td>
<td>1986</td>
<td>modern extracted</td>
<td>112</td>
<td>cementum rings</td>
<td>ring counting</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>Miller, Dove and Cottone</td>
<td>1988</td>
<td>modern extracted</td>
<td>100</td>
<td>cementum rings</td>
<td>ring counting</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>Solheim</td>
<td>1990</td>
<td>modern extracted</td>
<td>1000</td>
<td>cementum width</td>
<td>width measurement</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Bang and Ramm</td>
<td>1970</td>
<td>modern extracted</td>
<td>265</td>
<td>root dentine translucency</td>
<td>length measurement</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Lamendin and Cambray</td>
<td>1980</td>
<td>modern extracted</td>
<td>217</td>
<td>root dentine translucency</td>
<td>length measurement</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Wegner and Albrecht</td>
<td>1980</td>
<td>modern extracted</td>
<td>50</td>
<td>root dentine translucency</td>
<td>length measurement</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Solheim</td>
<td>1989</td>
<td>modern extracted</td>
<td>1000</td>
<td>root dentine translucency</td>
<td>length measurement</td>
<td>regression</td>
<td></td>
</tr>
<tr>
<td>Drusini, Calliari and Volpe</td>
<td>1991</td>
<td>modern extracted</td>
<td>134</td>
<td>root dentine translucency</td>
<td>length measurement</td>
<td>regression and estimated RDT</td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Sample Size</td>
<td>Parameter</td>
<td>Regression Type</td>
<td>Notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
<td>-------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dalitz</td>
<td>1962</td>
<td>146</td>
<td>cementum resorption scores</td>
<td>regression</td>
<td>by transmitted light poor correlation with age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solheim</td>
<td>1993</td>
<td>1000</td>
<td>resorption of the cement surface scores</td>
<td>regression</td>
<td>better than apical resorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solheim</td>
<td>1988</td>
<td>1000</td>
<td>root colour scores</td>
<td>regression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kvaal and Solheim</td>
<td>1989</td>
<td>1000</td>
<td>fluorescent properties of dentine intensity (continuous)</td>
<td>regression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gustafson</td>
<td>1950</td>
<td>42</td>
<td>secondary dentine occlusal attrition occlusal attrition periodontal recession recession root resorption root resorption cementum buildup root dentine translucency score</td>
<td>score regression</td>
<td>essentially the same as Gustafson (1950) but introduced half point scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johanson</td>
<td>1971</td>
<td>42</td>
<td>secondary dentine occlusal attrition occlusal attrition periodontal recession recession root resorption root resorption cementum buildup root dentine translucency score</td>
<td>score regression</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
14 Data Used for Comparative Estimations in Part 7