# Data Coding Scheme

## Introduction

The goal is to develop a numerical index of well being (adaptation) from skeletal data. Knowledge of both economic indicators (such as Gross National Product) and health status assessments suggest that useful indices can be developed from imperfect data. Our collection of skeletal data represent a large array of genetic, geographic, and temporal groups, which permit us to measure levels of health and to investigate their possible environmental causes.

Inevitably projects such as this face a trade-off between number of observations and the detail collected per observation. We emphasize number of observations (large number of individuals at a variety of sites) over detail (a complex coding scheme that records numerous features for each observation). We prefer this strategy because diverse sites enhance our comparative perspective. Moreover, even simple coding schemes, such as the one adopted here, can be very effective in representing important aspects of health. In short, we feel that the value of additional observations in a variety of ecological settings outweighs extraordinary detail on health for a smaller number of individuals at fewer sites.

## Site/Collection Identification

The given skeletal collection is identified by eight (8) alpha/numeric characters. These could be a site number (e.g., 23CG0234) or an abbreviated name (Cedargro).

## Individual Identification

Each individual is identified with a unique series of five (5) alpha/numeric characters (e.g., 00345 or 0956B).

## Inventory

Any statistical study of lesions must provide information about observable skeletal components. Thus, the coding scheme contains a category indicating that the appropriate skeletal components were available for observation. Any data available for a skeleton was recorded.

**Demography**

The age and sex distributions of a population reflect mortality and fertility, which are the ultimate measures of adaptive success in past human groups. Moreover, these distributions are useful for delineating differential effects of disease, nutrition, and stress generally by specific age, sex, and status groups.

Sex. It is assumed that the methods employed in the determination of sex are those used in standard analysis. A definite designation is one in which criteria from the pelvis, especially the pubis, is clear and used in conjunction with additional features. The pelvis can, at times, produce ambiguous sex information and thus result in an unknown or probable designation. A probable designation is employed when the pelvis cannot be used but cranial and postcranial attributes predominately indicate one sex or the other. Any designation using fewer than four nonpelvic criteria must remain in the unknown category.

The sex of the individual is designated with the single numeric code as follows:
1. Female, definite designation, the sex is certain; 2. Probable Female, possible designation, but the investigator is uncertain; 3. Male, definite designation, the sex is certain.; 4. Probable Male, possible designation, but the investigator is uncertain; 5. Sex is undetermined because the individual is less than 15 years of age and sex determination would be uncertain; 6. Unknown, the sex of the individual cannot be determined with any degree of reliability.

Age. There are a number of aging techniques and many investigators use one technique more effectively than other. Dental development is considered the best for the youngest ages while epiphyseal closure is used for the later growing years. Adult ages determined from the pubic symphysis and auricular surface of the pelvis are considered the most reliable. There are three columns for recording age: 1) summary age; 2) dental age for children; and 3) age range.

Summary age: The investigator is asked to designate a single year (i.e., 37) as the age at death. This is the best estimate which the investigator can make. In its simplest form without additional information this age is nothing more than the midpoint of the designated age range. Age is given in years using decimals to designate tenths of years (conversion of months). Individuals aged greater than 60 years, and for whom no further estimate is possible, are coded as 99. Summary age is recorded as a four (4) digit numeric field with one decimal place (e.g., 00.6 or 35.5).

Dental age: A dental age for children is required for the growth analysis. In many cases the dental age and the summary age are the same, however this category uses only dental development, whereas the summary age may include other sources of information. Dental age is recorded as a four (4) digit numeric field with one decimal place (e.g., 03.4).

Age range: It is preferable to use standardized age ranges (i.e. 50-54 years), but the data to be coded have already been collected and age ranges differ between investigators. We cannot expect after the fact concordance. The investigator provided the age range in years for that individual. For categories such as subadult or old adult, the investigator defined this category by providing an age range in years. If such ranges are being determined for this project, then standard five year intervals (i.e., 30-34 years) are employed. Age range is recorded as two numeric fields with two characters in each. The first represents the minimum age (e.g., 30) and the second the maximum age (e.g., 45).

Date of Birth: Where known the date of birth or decade (by using the midpoint, i.e., 1915) is recorded. Date of birth is a single field with four (4) numeric characters.

**Continental Ancestry**

Each skeleton is assigned to one of these populations of origin: 1. Native American; 2. European; 3. African; 4. Asian; 5. Mixed, any mixture of the above. This code is based the site context or from the historic literature. Although there are various degrees of mixture in some the groups, we cannot reliably estimate the mixture for an individual skeleton; 6. Unknown. Continental ancestry is a single character numeric field.

**Social Status**

An individual's position in a social hierarchy (i.e., social status) can influence access to both luxury items and/or necessities of life (e.g., adequate housing, nutrition, etc.). Thus, social status can have considerable influence on biological characteristics such as height, as well as determining resistance to disease (e.g., impairment of the immune system), physical work load (e.g., seen as increase degenerative joint disease) etc.

For some of the skeletal populations, there is clear evidence of social/economic distinctions evident from either the historical records or grave goods (i.e., differential access to luxury or scarce items). Each individual receives a three digit code. The first designates the stratification category of the society, the second indicates the number of strata, while the third indicates that person's place within the society.

Social Stratification Codes: 1. Denotes undifferentiated societies, reflecting the lack of significant differences in social stratification observable in the grave goods, and in archeological or historical evidence. If there is any doubt whether the presence of grave goods has any status meaning, these individuals or groups are placed here; 2. A ranked society is one in which there are social differences with groups or individuals having clear differential access to luxury or exotic goods. Membership in this category may be determined from either grave goods or historical records including archeological interpretation; 3. A class stratified society is one in which there are social differences with groups or individuals having clear differential access to wealth and/or subsistence resources. Assignment to this category may be determined from either grave goods or historical records including archeological interpretation.

Number of Social Stratification Code. The number of distinct social strata in the culture is designated by number. An undifferentiated society is coded as 1, a three strata society as 3, etc.

Individual Position Codes: 1. The person occupies the highest rank in a ranked society or is a member of an undifferentiated society; 2. The person occupies the second from the highest rank in a stratified society; 3. The person occupies the third from the highest rank in a stratified society; 4. The person occupies the fourth from the highest rank in a stratified society; etc.

Social Stratification Coding. The coding of social stratification uses a three (3) digit numeric code with the first designating the presence or kind of stratification in the society, the second the number of social strata, and the third the position of the individual in the social ranking. For example, a person in an undifferentiated society is coded as 111. A person in an historic class society could be coded as 331 (a class society, three classes, person belongs to highest class). If all that can be determined is that the society is ranked with only tendencies for high or low social status then the code might be 222 for a ranked society, two levels, and person belonging to the low social status.

**Growth and Heights**

Growth. Height for dental age is an excellent indicator of childhood nutrition and overall health. Maximum diaphyseal lengths of the femora (left is first then right if left is unavailable) are used to calculate growth statuses of juveniles using ages determined from dental development. Only the diaphyseal length (no epiphyses) are recorded here. This field contains three numeric characters for the measurement in millimeters.

Femur Length: Adult heights (stature) have been shown to be an excellent indicator of childhood nutrition and health. As we are using both skeletal heights and directly measured heights from historical sources, we decided to convert heights to bone lengths where necessary (this is done for data analysis). The bone to be recorded is the maximum length of the left femur (right if left is unavailable) recorded in millimeters. If necessary, femur length can be calculated from other bones using regression formulae. The femur length field contains three numeric characters and the lengths are recorded in millimeters.

Adult Heights: Heights are reported from historical documents on living populations (if available), and are estimated based on femoral length (adults) in skeletal samples. The left femur (maximum length) is preferred, then the right femur, but any long bone can be used for calculating heights. The various formulae for calculating stature from the maximum length of the femora or other bones are available in Krogman and Iscan (1986; American Whites and Blacks page 308, Mongoloids page 310, and Indigenes of Central Mexico pages 319-320) and Sciulli et al. (1990:275-280; Native Americans). The recorder calculated the heights. Heights, either calculated from long bones or derived from records are recorded in a 4 character numeric field in millimeters.

**Robusticity**

Femur: Robusticity provides information on body weight, past diet, current physiological health, work and activity patterns, and degree of mobility. The total subperiosteal area (TA) of the adult femur is most responsive to the combined effects of mechanical demand/physical activity and body weight, but activity is likely most important in behavioral interpretation. The anteroposterior (AP) and mediolateral (ML) diameters of the adult left femoral midshaft (right if the left is unavailable) is recorded in order to calculate TA. The formula for this calculation is (from Fresia et al. 1990):

TA = pi (Tap/2)(Tml/2)

where Tap = anteroposterior diameter at midshaft, and

Tml = mediolateral diameter at midshaft

Human populations vary widely in body size and, consequently, in femoral size. Therefore, it is absolutely essential that the measurement of TA be standardized when comparing populations. This can be done easily by dividing TA by femoral length to the third power in the following manner (Ruff et al. 1993):

TAstandardized = [pi (Tap/2)(Tml/2)]/max. length cubed

The anteroposterior (AP) and mediolateral (ML) diameters of the femur midshaft are recorded in millimeters in two fields each of two numeric characters.

Humerus: Robusticity data for the adult humerus is provided by the maximum length and the circumference of the midshaft. These two measurements in millimeters are entered into two fields the length with three numeric characters and the circumference two numeric characters.

**Enamel hypoplasias**

Enamel hypoplasias are excellent measures of childhood nutritional and morbidity stress, which complement growth rates and adult stature for reconstructing past health. Although they cannot be remodeled, they can be removed by wear and caries. Hypoplasias are reported only on the maxillary incisors and either the mandibular or maxillary canines for both deciduous and permanent teeth. The hypoplasias recorded are only linear groves that can be clearly seen with the unaided eye under good illumination.

Hypoplasias are recorded for 1) deciduous maxillary central incisor; 2) deciduous canine (maxillary or mandibular); 3) permanent maxillary central incisor; and 4) permanent canine (either maxillary or mandibular). There is one column for each of these teeth. Only systemic hypoplasias are recorded and the left teeth are used, but rights are reported if lefts are not available. The four teeth are scored as follows: 0 Not observable (no suitable teeth, incomplete development, or too worn, etc.); 1 No hypoplasia; One hypoplasia; 3 Two or more hypoplasias.

**Dental Disease**

Dental Caries. Caries results from a disease process, and without intervention it results in complete destruction and loss of the affected tooth. In most groups, dental caries is the primary cause of abscessing and loss of teeth. However, there are some groups where rapid wear leads to abscessing and tooth loss, while in others periodontal disease is the primary cause. The data are recorded for the permanent dentition only, as follows: 1. The total number of permanent teeth observed; 2. The total number of permanent teeth lost before death (antemortem); 3. The total number of teeth with lesions or restorations (i.e., fillings).

The data reported here are used to calculate individual (percent of carious teeth per mouth) and population (percent of total carious teeth per group) statistics on caries prevalence. There are three fields, each with two numeric characters.

Abscess: Abscesses can result from progressive caries or from tooth wear rapid enough to exceed the dentin's ability to fill the pulp chamber. In some cases the cause is not obvious (the loss is spontaneous). There is evidence that abscesses can be life threatening, or at the very least diminish resistance to disease and, even more than caries, affect dietary intake. Abscesses are recognized by a clear drainage passage leading from the tooth root(s) to the external surface of either maxilla or mandible. The data are recorded in two fields: 1) two numeric characters for the total number of sockets examined; and 2) one numeric character for the total number of abscesses.

**Anemia**

Anemia (as indicated by cribra orbitalia and porotic hyperostosis) can be caused by a variety of factors including an iron deficient diet, disease, and parasites. Scoring these conditions can be very complicated, but the information contained in the various skeletal expressions can be obtained by a simple scoring system. Cribra orbitalia and porotic hyperostosis are scored separately. To score as present or absent, at least one parietal and one orbit must be observable. Scattered fine pitting of parietals and occipital, some times called porotic pitting, is not scored as positive. There are two fields, each with one numeric character.

Cribra Orbitalia is scored as: 0 no orbits to be observed; 1 absent on at least one observable orbit; 2 presence of a lesion; 3 gross lesions with excessive expansion and large area of exposed diploe, which is the form associated with sickle cell disease and other severe forms of anemia.

Porotic Hyperostosis is scored as: 0 no parietals to be observed; 1 absent on at least one observable parietal; 2 presence of a lesion; gross lesions with excessive cranial expansion and huge areas of exposed diploe, which is the form associated with sickle cell disease and other severe forms of anemia.

**Auditory Exostosis**

It has been demonstrated that auditory exostoses (i.e., growth of extra bone occluding ear canal) are associated with swimming in cold water. These growths do impair hearing and are recorded as follows: 0 at least one auditory meatus not present for observation; 1 auditory meatus exhibits no exostosis; 2 exostosis present in one or both ears.

**Infection/Periosteal Reactions**

Infection of the bone (primarily by ubiquitous *Staphylococcus* or *Streptococcus* organisms) can be quite serious and debilitating because they are very difficult for the body's defense mechanisms to combat. All such infections result in pain, swelling with possible disfigurement, and interfere with normal activities. In addition, the infections are a burden on the individual's defense mechanism, which can result in reduced resistance to other disease processes.

Infectious lesions are complex to score because they can be isolated and minor, localized but chronic and debilitating, or the result of systemic disease. The skeletal sequelae of infection can exhibit the characteristics of active ongoing infection or the healing scars of past disease. Some periosteal reactions can result from trauma (bruising of the bone's periosteum) and these may be difficult to distinguish from infection. However, as most periosteal reactions are due to infection, they are scored as such unless the recorder has some reason to think otherwise. In order to ensure consistency of reporting, only major lesions of the major long bones are employed in the development of the index. Active and healed lesions are not differentiated. We focus almost exclusively on the tibia, which is the most common site for infectious lesions. There are two sets of scores, with the first being for the tibiae and the second for the remainder of the skeleton. Each of the fields contain one (1) numeric character.

Tibial scores: 0 no tibia(e) present for scoring; 1 no infectious lesions of the tibia(e) with at least one tibia available for observation; 2 slight, small discrete patch(s) of periosteal reaction involving less than one quarter of the tibia(e) surface on one or both tibiae; 3 moderate periosteal reaction involving less than one-half of the tibia(e) surface on one or both tibiae; 4 severe periosteal reaction involving more than one-half of the tibia(e) surface (osteomyelitis is scored here).

Remaining skeleton: 0 no periosteal reaction on any other bone than the tibiae; 1 periosteal reaction on any other bone(s) than the tibiae not caused by trauma; 2 evidence of systemic infection involving any of the bones (including the tibiae) of the skeleton. This would include specific diseases which include (but are not limited to) tuberculosis and syphilis.

**Degenerative Joint Disease**

Degenerative joint disease provides considerable information concerning activity patterns because chronic stress on the joints eventually damages the cartilaginous surfaces and, when sufficiently advanced, also the bone surface beneath. Within a given population, those individuals engaged in regularly occurring activities that produce chronic joint stress (e.g., rowing, running, etc.) will develop patterns (i.e., specific joints affected such as knees or elbows) of degenerative joint disease differing from that in the general population. In addition to differences in the pattern of joints affected, variation in the age at which the damage appears can be informative.

Degenerative diseases can be difficult to score consistently and yet can be very informative even when recorded with great simplicity. There are eight (8) fields, each with one numeric character: 1) shoulder and elbow; 2) hip and knee; 3) cervical; 4) thoracic; 5) lumbar vertebrae; 6) temporomandibular joint; 7) wrist; and 8) hand. The most severe manifestation from either the right or left side is scored.

Shoulder and Elbow are scored as one unit and if either joint is affected (score the most severely affected joint), it is scored as follows: 0 joints not available for observation; 1 joints show no sign of degenerative disease; 2 initial osteophyte or deterioration of the joint surfaces; 3 major osteophyte formation and/or destruction of the joint surface such as eburnation; 4 immobilization of the joint due only to degenerative disease; 5 systemic degenerative disease (e.g., rheumatoid arthritis, Alkaptonuria, etc.).

Hip and Knee are scored as one unit and if either joint is affected (score the most severely affected joint), it is scored as follows: 0 joints not available for observation; 1 joints show no sign of degenerative disease; 2 initial osteophyte or deterioration of the joint surfaces; 3 major osteophyte formation and/or destruction of the joint surface such as eburnation; 4 immobilization of the joint; 5 systemic degenerative disease.

Vertebrae are scored by type: cervical, thoracic, and lumbar. If four or more thoracic vertebrae are present, they are scored and if two or more cervical or lumbar are present, they are scored. Only the bodies of the vertebrae are scored for the most severe expression:

Cervical: 0 not observable; 1 no lesions on at least two observable vertebrae; 2 initial osteophyte formation along rim of the vertebral body(ies); 3 extensive osteophyte formation along rim of the vertebrae; 4 two or more vertebrae fused together

Thoracic: 0 not observable; 1 no lesions on at least four observable vertebrae; 2 initial osteophyte formation along rim of the vertebral body(ies); 3 extensive osteophyte formation along rim of the vertebrae; 4 two or more vertebrae fused together (keeping in mind that kyphosis from tuberculosis would be scored under infectious disease and not here).

Lumbar: 0 not observable; 1 no lesions on at least two observable vertebrae; 2 initial osteophyte formation along rim of the vertebral body(ies); 3 extensive osteophyte formation along rim of the vertebrae; 4 two or more vertebrae fused together.

Temporomandibular Joint. Deterioration of the temporomandibular joint (TMJ) can lead to difficulties in chewing, intense pain, and a large poorly understood array of psychosomatic diseases. Only extreme deterioration is recorded. This is recognized at the level that degenerative disease would be recorded on any other joint including osteophytes, eburnation, and joint surface deterioration: 0 TMJ not observable; 1 no deterioration; 2 joint deterioration.

Wrist and Hands: Radio-ulnar joint: 0 bones not observable or not recorded; 1 no degenerative disease of the joint; 2 degenerative disease of the joint.

Bones of the hand: 0 bones not observable or not recorded; 1 no degenerative disease of the joint; 2 degenerative disease of the joint.

**Trauma**

Trauma provides information about many aspects of society and the relationship of the people to the environment. Different activity patterns and terrains produce different patterns of trauma. For example, walking on rocks tends to produce higher frequencies of fractured ankles, while in less rugged terrain lower arm fractures tend to predominate. In addition, both intra and interpopulational violence can be documented by specific types and patterns of trauma such as parry fractures of the lower arm or depressed fractures of the cranium. The scoring of trauma focuses on the major bones of the limbs and the skull: humerus, radius, ulna, femur, tibia, fibula and skull. Unless it can be shown to be premortem or perimortem traumata (e.g., saw marks of an amputation or axe wound), they are only scored when there is some evidence of healing. Any form of surgery would be recorded as trauma. It is critical that postmortem modifications or damage not be recorded.

There are seven fields, each with one numeric character: 1) arm; 2) leg; 3) nasal bones; 4) face; 5) skull vault; 6) hands; 7) weapon wounds.

Arm: humerus, radius and ulna. If any bone shows trauma, it is scored as follows: 0 no long bones observable (must have humerus and at least one bone of the forearm to be scored); 1 not fractured; 2 healed fracture with acceptable alignment; 3 healed and poorly aligned; 4 healed with fusion of the joint; 5 healed fracture with alignment unknown.

Leg: femur, tibia, and fibula. If any bone shows trauma it is scored as follows: 0 no long bones observable (must have femur and tibia or fibula); 1 no fracture or other trauma; 2 healed fracture with acceptable alignment; 3 healed and poorly aligned with some loss of locomotor function; 4 healed with extreme loss of locomotor function such as loss of limb or complete fusion of the joint in the lower limb; 5 healed with alignment unknown.

Nasal and Nasal Process: 0 no bones to be observed; 1 no fracture; 2 healed fracture.

Face Other than Nasal: 0 no bones to be observed; 1 no fracture; 2 healed fracture.

Skull Vault: 0 no bones to be observed; 1 no fracture; 2 healed fracture.

Hand Fractures: 0 no bones to be observed or not recorded; 1 no fracture; 2 healed fracture(s).

Weapon Wounds to any Part of the Body and Head: 1 no weapon wounds; 2 weapon wound(s).

**References**

Fresia, Anne, Ruff, C.B. and Larsen, C.S. 1990 Temporal decline in bilateral asymmetry of the upper limb on the Georgia coast. In: Larsen, C.S. (ed) The Archaeology of Santa Catalina de Guale: 2. Biocultural Interpretations of a Population in Transition. Anthropological Papers of the American Museum of Natural History No. 68, pp. 121-132.

Krogman, Wilton M. and Iscan, Mehmet Y. 1986 The Human Skeleton in Forensic Medicine. Charles C. Thomas, Springfield IL.

Ruff, Christopher B., Trinkaus, E., Walker, A. and Larsen C.S. 1993 Postcranial robusticity in Homo. I: Temporal trends and mechanical interpretation. American Journal of Physical Anthropology 91:21-53.

Sciulli, Paul W., Schneider, K.N. and Mahaney, M.C. 1990 Stature estimation in prehistoric Native Americans of Ohio. American Journal of Physical Anthropology 83:275-280.