CLINICAL ANTHROPOLOGY IN MEDICINE

Pavel Bláha

Abstract:

Clinical anthropology represents one of the most important and prestigious fields of applied anthropology. Its significance consists in applications of anthropometric and anthroposcopic methods to medical research and clinical practice. It observes the growth, developmental, morphological and functional traits of individuals and the whole populations in dependence upon diseases and influences of various factors of external environment on human health. In the latter respect the scope of clinical anthropology is permeated with environmental anthropology and in a broader sense also with human ecology. Applications are based on elaboration of growth, developmental and morphological norms of characteristics of the human body and on definition of boundaries of their natural variability in dependence upon age, sex, ethnicity and regional origin.

The use of anthropological methods and anthropometric norms impinges upon most medical disciplines. These methods do not include only classical anthropometry, whose application still continues to be very important, but concern all metrical methods such as roentgenometry or some newly introduced 3D contact and non-contact morphometric approaches. The latter observe the same methodical and methodological rules. Methods of visual assessment employ various pattern templates, documentary records (imprints, photographs, etc.), constructional analyses and sophisticated approaches of data processing to objectivise results. They rely on standard norms that display an individual's position in respect to other members of a group and show how somatic characteristics reflect various developmental, medical, nutritional and social influences.

Fields of Application

Where and how are the above-mentioned anthropological methods applied in clinical practice? Generally in four principal fields: (1) in definition and quantification of deviations from normal conditions, (2) in diagnosis of disorders and prediction of their further development, (3) in planning of the extent of some forms of treatment, (4) in evaluation of the results of therapy. Each of these activities is associated with definite methodical rules, which have to be respected with the utmost care. In the first application this relates to the suitability and appropriateness of the control group, because correct conclusions need to exclude influences of personal mode of measurement as well as deviations due to family characteristics – see below. In the second application prediction cannot be based on growth curves of normal population(s), but on curves for a definite disorder or anomaly (e.g. a person with the Turner syndrome, or achondroplasia cannot grow in agreement with norms). Similarly, when diagnosing general characteristics such as biological age, it is not possible to use traits that form an

essential part of a disorder (e.g. skeletal age in achondroplasia, which represents disorders of ossification of the cartilage). As far as the third mentioned application is concerned, it is sometimes necessary to bear in mind the phenomena of regression after operations to preoperative situation. This depends on the extent of surgical procedure(s) and requires overshooting of an advancement. The results of the treatment should not be compared only with standard norms, but also with the original pretreatment state. We should not forget that in morphological characteristics the aesthetic impression and functional state are more inportant than the extent of residual deviation.

Anthropology can see the widest fields of its application in pediatrics, endocrinology and related disciplines during diagnosis, and in an objective observation and follow-up of treatment results of impaired growth and development in various diseases, inborn defects, obesity and endocrine or metabolic disorders. In neurology and neurosurgery it is helpful in diagnosing craniosynostoses and other cranial anomalies as well as in observing the recovery of cranial growth after operations. Similarly, in plastic and craniofacial surgery anthropometry contributes a lot to obtaining objective data about the extent of anomalies, and in planning the amount of reconstructive procedures. It is also indispensable for evaluating results of primary and corrective surgeries and on the basis of comparison of various procedures of therapy for optimisation of treatment. Its use is also common in orthopedics and traumatology when examining a pathological growth, impaired mobility and asymmetry of extremities, when planning the extent of surgical operations (e.g. prolongation of the extremity) and judging their adequacy and success in subsequent rehabilitation. In prosthetics it helps to determine the size and types of prostheses or epistheses. In orthodontics and jaw orthopedics it may become useful when we measure X-ray images in order to diagnose jaw disorders, or when we have to predict a prospective development of jaws. It is needed for determining (possible) growth patterns of the face so as to choose appropriate forms of treatment. In internal medicine anthropological methods serve for exploring the influence of various diseases, including diseases of civilisation, on somatic characteristics and on body composition of patients. It also helps to explain how body characteristics are related to the patient’s susceptibility to diseases.

The last application, of no less importance than the others, is in sport medicine where the impact of sport activities and physical training on the human organism (and, by way of contrast, somatic predispositions for sport achievements in various disciplines) are estimated. Anthropology is also needed in medical genetics and consulting when we observe light forms of anomalies in relatives, which are often manifested in the form of specific morphometric changes. It is indispensable for deter-mining the hereditary situation in a family. Psychiatry seems to be unrelated and too far-fetched in respect to anthropology but its methods may bring benefit when we find out physiognomic and dermatoglyphic stigmata of mental disorders. It goes without saying that this brief survey cannot enumerate all possible applications of anthropological methods in medicine and clinical practice.

And we can not forget possibilities of methods of physiological and molecular anthropology.

Specifics of Clinical Research

There are many problems associated with clinical research. Essential presupposition of correct results is the correct manner of choosing patients and types of research. Equally important are the methods of removing deformations due to an atypical com-position of examined series and methods of adjustment to the corresponding size of persons or to a constant interval between examinations.

In clinical practice it is common that increments are not obtained in regular intervals, but according to the patient's treatment schedule. Under certain conditions, however, we can arrange the data into shorter or longer intervals. With equal growth of given characteristics we convert the ten-month increment to a yearly value by dividing it by ten and multiplying by 12 (months). Similarly, 14-month increment divided by 14 and multiplied by 12 can be converted to a yearly value. However, in non-linear speed of growth we have to use the equation of the curve corresponding to the relevant model of growth.

Accuracy of measurement is particularly important in observing growth speed be-cause every error can have an impact on the two of the increments, the previous and the following. This can lead to an error diagnosis of growth impairement (false positive finding) and subject the particular individual to unnecessary assessments; even worse, it can mistakenly establish normal values (false negative finding) and hinder early diagnosis of growth defect. Unlike testing of findings in research of groups, in this case the second type of error is more serious. When growth speed is evaluated it is necessary to use stricter critical limits (10th and 90th percentile or even 25th percentile) be-cause slower long-term growth ultimately leads to a growth deficiency. It is also important that absolute values of some characteristics should be evaluated in relation to body size rather than age: for example, body weight compare to height (however, in newborns to the gestation age). Head circumference, which is an independent parameter, should in very small and very large individuals also be compared with the body height (stature age).

Choice of patients can also be a source of bias results. When morphological characteristics are used, we can more easily record more severe than less severe forms of malformation, which will present as a shift of characteristic distribution towards the more severe end of malformation spectrum (shift to the right with obliquity to the left). If more severe degree of malformation is lethal, the situation is reversed. In addition, distribution with two peaks can appear, which always shows evidence of inadequate composition of the series – inhomogeneous group, two mixed groups, incomplete selection, etc. This requires data control and if possible split of distribution (in mixed groups), clearance or supplementing of selection (in incomplete groups). Normal distribution of morphological characteristic is best guaranteed by selection according to non

phenotype symptoms of malformation (for example, levels of growth hormone).

Clinical research usually involves all patients who subsequently come during a certain period of time (consecutive patients) and match the criteria provided beforehand. As a control group, another group of patients is used (who may be treated by different methods, for example) or a group of healthy individuals, chosen by random stratified selection. A further type of selection is paired selection, when an individual from the research group pairs up with another individual who is, in terms of the characteristics that influence the research result (so-called covariant), identical with the research individual with the exception of the characteristic or factor whose effect we are observing (e.g. smoking, or a different type of treatment). Consequently, we link pairs of individuals with the same characteristics who differ only in the characteristics being studied. This process is called covariant matching. Characteristics that we consider covariants bring uncertainty, however, and can influence results because it is impossible to assure absolute conformity.

Even when all patient selection obligatory rules are followed, often the compared groups differ in size of the basic parameters that do not show dependency on the researched factors (are not covariant). The method which permits comparison in such cases is scaling to the same size of mentioned parameters. Body height is used for scaling in somatometry, in craniometry it is the presellar length of the skull base (N–S). If we divide the body height of the taller group by the body height of the less tall group we get a constant greater than 1. By multiplying all other parameters of the less tall group by this constant we transfer its values to the size corresponding to the taller group. Similarly we can transfer variables of the taller group to the size of the less tall group (body height of less tall group divided by body height of the taller group results in a constant lesser than 1, multiplying the parameters of the taller group by this constant transferes them). However, this procedure is only appropriate for isometric growth of adjusted parameters with respect to the basic (adjusting) parameter or with small differences between basic characteristic of both groups. Similarly, we can for a specific objective transfer female parameters (or deviations) to male parameters and vice versa. If the criteria for the use of this procedure are not fulfilled, the comparison must be performed with the use of z-score.

Clinical Trials

In clinical anthropology we distinguish two basic types of research – prospective and retrospective. A retrospective study allows data to be obtained quickly, but with less reliability of patient anamnestic data. Factors that had an effect in the past are tested, and the method is suitable for pilot studies and determination of factors that require increased attention. Prospective studies require long term monitoring but allow for control of all input conditions and continuous phenomenons that influence results. Therefore it is appropriate to use them in succesion to retrospective studies when we already have enough knowledge and when we choose characteristics suitable for monitoring. The retroprospective method is considered one that, based on detailed medical data, makes it possible to perform prospective monitoring of the past. One requirement in

retrospective studies is comparison with several control groups, particularly if these are not of own provenience.

However, the best organized research is known as the randomized control trial. A group of patients is randomly divided into two groups. This so called randomization does not provide the same composition of both groups, but it ensures that they differ randomly. Then, for example, each group is treated with another method, and after a certain time a test confirms if the treatment results of both groups differ randomly or significantly. The whole procedure is usually organized as a double-blind experiment; the organizer of the trial does not participate in treatment or evaluation. The people performing the trial do not know what medication they are administering or to which group the patient belongs. This allows for uniform approach to patients or, more precisely, differences are again random. Such experiments may not always be performed due to ethical reasons (please see below). Sometimes it is possible to substitute con-trolled trials by matching covariants; however, to achieve correspondence of all covariants is impossible, and disturbing covariants (characteristics) that influence results can occur.

Besides what are known as extensive procedures, in some situations it is possible to use an intensive procedure. This consists of a long-term follow up of the same patient (patients), when treatment can be modified or interrupted and results are continuously evaluated. The patient can be a control to himself. It is possible to follow patients with various physiological, biochemical and other characteristics and evaluate treatment results in respect to these characteristics and its interactions. Therefore more accurate observations are available than in, from the described aspects, heterogeneous groups. Many observations of one patient are more useful that one observation of many patients. When we analyze these statistical data, we have to remember that these are mutualy dependent.

However, confirmation of research results requires any other studies. As practical confirmations serve population experiments when effects of widely used measures resulting from research findings (for example, water fluoridation) can be verified. Its goal is to shift the whole distribution curve of risk factor in a population towards more positive values.

Diagnostic, Screening and Predictive Systems

An important position of anthropology as regards practical applications lies in the development of correct systems for diagnostics, screening and prediction, because their methodology is based on different principles.

Diagnostic systems are generally based on deviation +/– 2 SD from the mean value (or 3rd and 97th percentile in non-normal distributions), which excludes about 5 per cent of population in the given variable outside of normal. It was calculated that when 100 independent variables is evaluated, none of 100 individuals would have all variables within a norm. Despite this shortcoming, the deviation +/– 2 SD from the mean proves appropriate when applied to an impairement diagnostics of an individual who undergoes more medical examinations for confirmation or disproval of diagnosis.

It is a completely different situation if the same deviation is used for screening on the level of the whole population. For example, 5 thousand of 100 thousand newborns in Czech Republic in one year are excluded from the norm. However, in medium frequent impairements with frequency of 1:10 000 only 10 are born disabled. Thus system with such an extent of selection (5%) and with further necessary examinations fails organisationally and sometimes also financially. It is obvious that criteria of selection must be narrower and result from prevalence of impairement. The rate of the actually disabled within all those selected is particularly critical. The suitable extent of this ratio cannot be unambiguously established because it depends on impairement severity, its treatability, age changes and other factors. The above described principles are not always respected.

Systems for prediction of development of either healthy or disabled individuals can only be based on appropriate longitudinal studies. Correlation coefficients between predicted and predicating variables around 0.8 are necessary. These explain 64% of variability of the predicted variable. However, a good prediction provides only coefficients 0.9, explaining over 80% of variability. It is exceptional for such a high correlation to be found in biological systems, and for prediction it is necessary to use combinations of more predicting (independent) variables. We choose them from a correlation matrix of large amount of characteristics, so that they would prove high correlation with the predicted (dependent) variable and at the same time low correlation mutually with one another. Factor analysis helps to sort variables into groups according to their relation-ships. It is useful to search for combinations of the lowest number of the predicting variables with the highest correlation coefficient to the predicated parameter.

Principles of Correct Evaluation

In the introduction we mentioned some specific problems connected with the exclusion of personal mode of measurement as well as of deviations due to hereditary family characteristics on the results obtained. Then it is convenient to take measurements of healthy relatives, as a control group and calculate the mean z-score of the group. After subtracting the calculated figure from the z-score of the patient's trait, we correct the influence of characteristics inherited in the family (i.e. zpatient – zrelatives). Similarly we can correct influences of a different manner of examiner measurement. We examine a smaller control group of our own (with n greater than 20), in which we calculate z-score for evaluated traits. If we subtract these figures from the z-score of the patient's traits, counted against the same reference control, we will correct variations due to a different manner of our measurement (hence zpatient – zown control). Thus if we require accurate conclusions, it is necessary to compare data not only with norms from large population (i.e. reference control), measured by other research workers, but also with our own control measurements. Again the most suitable solution is offered by a check-up of relatives.

Methods of evaluation of deviations from norm with the help of SD or percentile methods and by means of profile patterns are quoted in numerous textbooks.

These sources mention also various ways of assessment of inherited growth potential and growth velocity of individual patients used in clinical practice. It is also possible to calculate similarity between profiles of morphograms of two patients or two relatives by means of correlation coefficients. Calculations are carried out by common statistical procedures: however, they do not compare two traits in many individuals but on the contrary they compare two individuals in respect of a greater number (n) of traits expressed by means of z-score (Fig. 1).

All of the above examples show the importance of familiarity with adequate methods of data processing in clinical anthropology.

All of the above examples show the importance of familiarity with adequate methods of data processing in clinical anthropology.

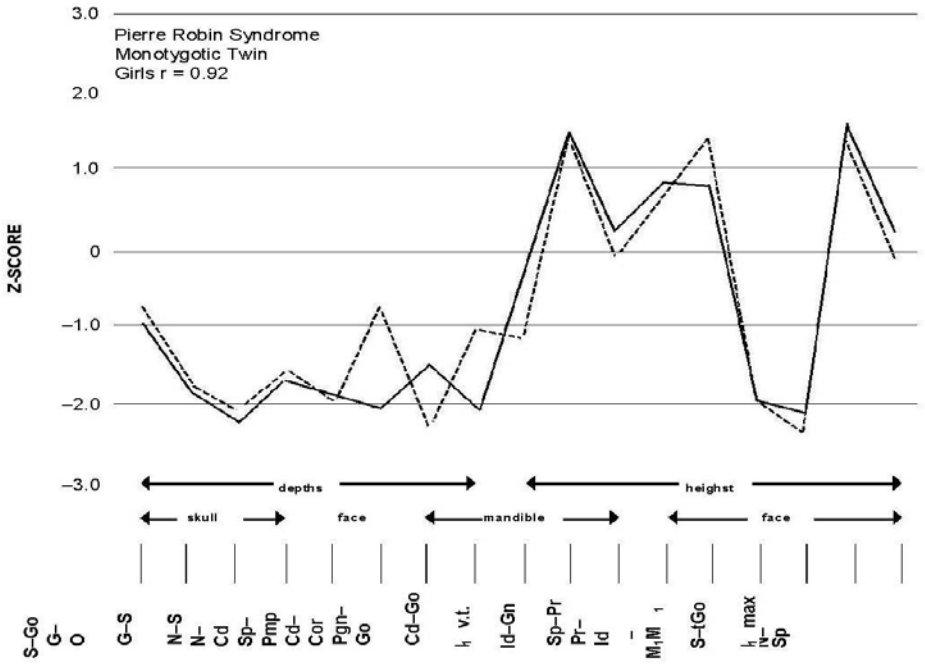

Figure 1 – High correlation in the profile pattern of roentgencephalometric measurements in monozygotic twin girls with Pierre Robin syndrome illustrates the same amount of impairement (After Cohen, 1985).

Accuracy of Evaluation

An indispensable condition of research is a sufficient reliability and accuracy (precision) of measurements and evaluation of qualitative traits. For an appraisal of sufficient accuracy and reliability of measurements it is common to use a calculation of standard error of measurement according to Dahlberg, denoted also as TEM (technical error of measurement). The sample of an investigated individuals (with n = 25) is measured twice under the same conditions. Measurements are carried out either by the same person for appraising the degree of individual agreement between different measurements (intra-observer error) or by two persons in order to obtain information about agreement between different examining individuals (interobserver error). Differences between both measurements are

raised to the second power, then they are all added together and divided by two, because in every measurement we can make an error (n = a number of repeatedly measured persons, not of different measurements). This is how we obtain their error variance and by extracting its square root we determine the standard error of measurement (SDe). If we divide this SDe, by the mean of the trait evaluated (correctly from both measurements), we will obtain the coefficient (Ve) of error variance that de-fines the accuracy of measurement and must not exceed the value of 5 per cent (we suggest 3 per cent). The more important parameter is reliability of measurement that is not related to the size of a trait but to its variance. The coefficient of reliability R will be obtained if we divide the difference between the total variance and the error variance, i.e. the real biological variance of a trait, by the total variance including also errors of measurement. This coefficient should exceed the value of 0.9 per cent. If it ranges below 0.8 per cent, the measurement makes no sense, because more than 20 per cent of variability of a given trait is caused by an error of measurement (Rx100 determines a percentage of the rate of biological variance to total variance and its subtraction from 100 specifies the rate of error variance caused by measurements). An appraisal of reliability may be performed also in a larger sample of observers, if each observer measures every proband and the subsequent calculation is done by two-way analysis of variance. These procedures are indispensable for unifying various manners of measuring by different examiners and they are also vital for calculations of coefficients of real correlation. As a result of errors in measurement, the latter is always greater than values calculated (the greater is the error of measurement, the greater is the difference).

An objective appraisal of qualitative traits naturally requires a sufficient agreement between two observers, or eventually between two observations done by the same examiner. It is carried out by calculating the constant kappa κ from a fourfold table that is depicted on Figure 2. Each field is occupied by data observed, on the right and at the bottom there are sums in rows and in the columns so called marginal frequencies. In the right lower corner there is the total number of individuals (n). The sum of individuals on the axis of agreement (the fields a +) divided by the total number of probands (n) deter-mines the degree of observed agreement. Then expected frequencies in each field are calculated by multiplying marginal frequencies divided by n: hence the field a = (a + b) × × (a + c) / n, and the field d = (c + d) × (b + d) / n. From these values the degree of expected agreement is calculated by adding them and dividing their sum by n (aexpected + + dexpected : n). Then the coefficient of inter-observer agreement will be determined accord-ing to the equation κ = observed agreement – expected agreement: 1 – expected agreement (where 1 symbolises total agreement). Most authors assume that the degree of agreement should exceed the value of 0.7. However, such a high degree of agreement does not necessarily guarantee correctness of evaluation, because both examiners can make errors.

Sensitivity and Specificity, Relative and Difference Risk, Odds Ratio

In the four-fold table (Fig. 2) we can also calculate the sensitivity and the specificity of our evaluation or test. The sensitivity of the test is a measure of the frequency of positive results in patients, and we can calculate it by dividing the number of patients with positive tests by the number of all individuals with the disease, or in the four-fold table a / (a + c). The higher the sensitivity, the more sensitive is the test; however, even 100% sensitivity does not mean that the individual with a positive test is suffering from the disease, because a certain percentage of tests can be positive even in healthy individuals – which is a measure of the specificity of the test. This characteristic shows the frequency of negative test results in healthy individuals. It can be obtained by dividing the number of healthy individuals with negative test by all healthy individuals, or d / (b + d). The higher the result, the more specific the test is; however again, a negative result does not mean that the individual is healthy, because negative results can also be found in individuals with the disease. Only a combination of both characteristics offers a measure of the validity of the evaluation. The rules here can be remembered by mnemonic techniques: SnNout = if sensitivity (Sn) is high, negative (N) test result excludes (out) the disease, SpPin = if specificity (Sp) is high, positive (P) test result confirms (in) the disease. We can use the high sensitivity of a test to exclude the disease (if the result is negative), and high specificity to confirm the disease (if the result is positive). These two data are important characteristics referred to in all legal tests. The false positive rate can be calculated as b/b+d and false negative rate as c / a + c.

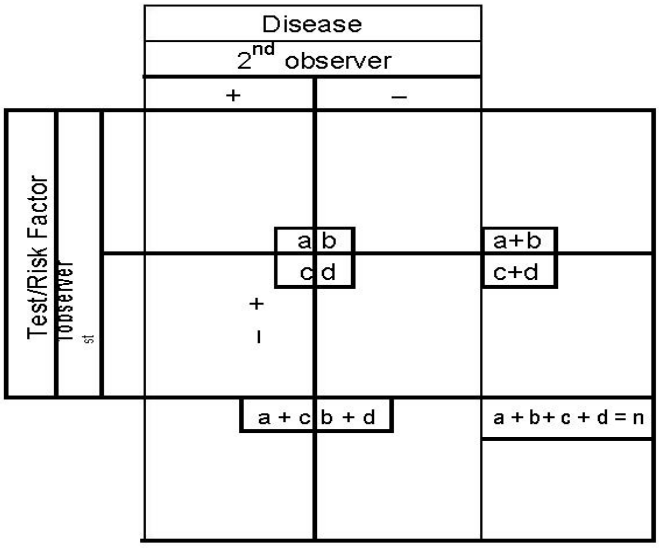

Figure 2 – Fourfold table for calculation of kappa coefficient (κ) of agreement between two observers, sensitivity and specificity, relative and difference risk as well as odds ratio.

In the four-fold table we can see the relationship a / (a + b). It is the positive predictive value of the test, which gives us number of diseased individuals (a) of those with positive test (a + b). We showed that this relationship is decisive for aplicability of screening tests and depends on the prevalence of disease or impairement (prevalence = a + c / n). The negative predictive value of a test that shows the number of non-impaired individuals of all individuals with a negative test result, i.e. d / (c + d), is not of great importance in this context.

From the four-fold table we can also calculate the relative risk RR = a / (a + b) : c / (c + d) and odds ratio OR = a/b : c/d. They show how many times greater the risk of impairment (e.g. disease or death) and the odds ratio is, given the presence of a certain risk factor (Fig. 2). Values greater than 1 show greater risk or odds, values lesser than 1 lesser risk or odds. OR (suitable for retrospective studies) compared to RR (suitable for prospective studies) does not depend on the frequency of the phenomenon investigated (e.g. a disease) within the population; however, the expression of relative risk is less precise. Difference (atributive) risk DR = (a / a + b) – (c / c + d) shows how many investigated phenomenons in absolute numbers are caused by the risk factor. Since the result value is per individual under risk it is necessary to multiply it by the number of persons under risk.

Ethics of Research

In conclusion it is necessary to draw attention to elementary ethical principles in screening patients. We must not perform research examinations in severely diseased persons or in situations when examinations could harm patients and expose them to excessive exertion. We should also avoid the danger of putting a patient into a disadvantageous and inconvenient condition in comparison to other patients (e.g. in randomised studies). Any research may be carried out only if questions of the study have no well-known answer and if these questions are not posed correctly, any research does not meet ethical principles either. These dangers presuppose a deep understanding of methods, methodologies and rules of clinical research. In every case examination presupposes that all patients or their legal representatives express an explicit approval and undersign a note about being informed about goals of examination.

References

Buschang, P. H., Tanguay, R., Demirjian, A. (1987): “Cephalometric reliability: A full ANOVA model for the estimation of true and error variance.” *The Angle Orthodontist 57*, pp. 168–175. Brattström, V., McWilliam, J., Semb, G. (1991): “Cephalometric scaling methodology in limited samples.” *European Journal of Orthodontics 13*, pp. 157–160.

Cohen, M. M. (1985): “Dysmorphic growth and development and the study of craniofacial syndromes.” *Journal of Craniofacial Genetics and Developmental Biology, Suppl. 1*, pp. 43–55.

Mascie-Taylor C. G. N. (1994): “Statistical issues in anthropometry.” In Ulijaszek, S. J., Mascie-Taylor, C. G. N. (EDS.):

Anthropometry: the individual and the population. Cambridge Univ. Press, Cambridge.

Sackett D. L. (1992): "*A primer on the precision and accuracy of the clinical examination*." JAMA 267, pp. 2638–2644.

Saksena S. S., Ramos-Arroyo M. A., Hodes M. E. (1987): "Craniofacial pattern profile (CFPP) evaluation of facial dysmorphology in a familial syndrome with corneal anesthesia and multiple congenital anomalies." *American Journal of Physical Anthropology 74*, pp. 465–471.

Šmahel Z. (2001): *Principy, teorie a metody auxologie*. Karolinum, Praha.

Soka, R. R., Rohlf F. J. (1994): *Biometry*. 3rd ed. W. H. Freeman, NewYork.

Ulijaszek S. J., Lourie J. A. (1994): "Intra- and inter-observer error in anthropometric measurement." In: Ulijaszek, S. J., Mascie-Taylor, C. G. N. (EDS.): *Anthropometry: the individual and the population*. Cambridge Univ. Press, Cambridge.

Ward R. E., Jamison P. L. (1991): "Measurement precision and reliability in craniofacial anthropometry: Implications and suggestions for clinical applications." *Journal of Craniofacial Genetics and Developmental Biology 11*, pp. 156–164.

Zar J. H. (1999): *Biostatistical Analysis*. 4th ed. Prentice Hall, London.

Contacts

Doc. RNDr. Pavel Bláha, PhD.
College of PE and Sport PALESTRA
Pilska 9, 190 00 Prague
Czech Republic
Tel.:+00420/281 932 013
blaha@palestra.cz