



INTERDISCIPLINARIA ARCHAEOLOGICA

NATURAL SCIENCES IN ARCHAEOLOGY

homepage: <http://www.iansa.eu>



Agreement in the Scoring of Dental Caries by Dentists and Anthropologists

Jiří Šneberger^{a*}, Anna Pankowská^b

^aMuseum of West Bohemia in Pilsen, Kopeckého sady 2, 301 00 Plzeň, Czech Republic

^bInstitute of Anthropology, Faculty of Arts, University of West Bohemia in Pilsen, Sedláčkova 15, 30125, Plzeň, Czech Republic

ARTICLE INFO

Article history:

Received: 20 April 2014

Accepted: 5 August 2014

Key words:

dental caries
assessment of agreement
Cohen Kappa
dentists
anthropologists
intra-observer
inter-observer
teeth identification

ABSTRACT

Correct scoring of caries lesions is a fundamental step in the bioarchaeological investigation of past populations. The present and the extension of caries lesions inform us about past subsistence and diet, human behaviour, social inequality and overall health status through time. We asked ourselves how relevant an interpretation of a past society's dependence on dental caries identification might be with an interest in demonstrating the quality of the primary data and the need for precise evaluation. We tested an assessment agreement of dental caries among observers from different subject fields. Two dentists and two anthropologists visually investigated 233 teeth and 3029 teeth surfaces from archaeological samples at the Plzeň "U Zvonu" cemetery dating to medieval and early modern times. We made use of Cohen kappa and Fleiss' kappa to calculate the inter-observer and intra-observer agreement. The results indicate that the overall prevalence of dental caries in the tested assemblage ranged from 1% to 6%, depending on the observer. The inter-observer agreement decreased from an average of $k=0.77$ (identification of teeth) to $k=0.40$ (evaluation of the extent of caries lesion). Assessment in the caries extension decreased more rapidly having been caused by the in all probability increasing demands of individual observation. We recorded homogeneity in the assessment of each observer in the intra-observer agreement with the exception of Dentist 2. We did not record a dependence on the subject field of the observer. The most probable cause of low inter and high intra-observer agreement could be in the methodological process of each observer, precision in the application of the methodology and the responsibility of the individual observer.

1. Introduction

Teeth obtained from an archaeological assemblage provide important information for the reconstruction of the behaviour of past populations (Larsen *et al.* 1991). Dental tissues are the best preserved part of human and non-human remains (Kohn *et al.* 1999; Lee *et al.* 2011; Wesolowski 2006) and often the sole source of information about past and recent populations using retrospective and forensic sciences (Bowers 2004; Larsen 1999; Scott, Turner 1997). Bioarchaeology and paleopathology are interested in the pathological conditions of dental tissues as dental caries (Scott, Turner 1988), *pulpitis* (Hillson 1996), peri-apical inflammation (Banerjee *et al.* 2000), enamel hypoplasia (Hillson 1997), abrasion, erosion and resorption (Deter 2009), ante-mortem loss of teeth (Hillson 2001; Waldron 2009), calculus and periodontitis

(Hillson 2005) or trauma (Buikstra, Ubelaker 1994). All of these pathologies can help establish a certain picture concerning lifestyle (Sheiham 2006), diet, food preparations, social status (Garcia-Closas *et al.* 1997; Piovesan *et al.* 2011; Reisine, Psoter 2001; Sheiham, Bönecker 2008), stress period during infancy (Kilian, Vlček 1989; Lovejoy 1985; Ubelaker 1987), sex differences (Keenleyside 2008; Lukacs 2008) and subsistence strategies (Nelson *et al.* 1999). On the basis of these and other similar bioarchaeological studies, anthropologists and archaeologists have created conceptions of past populations. We, however, are attempting to demonstrate that interpretations of prehistoric and historic populations can be inaccurate.

The most commonly investigated dental disorders in bioarchaeological studies are caries lesions. Dental caries serve as an indicator of different subsistence systems for a wide range of chronological periods in diverse cultures. The prevalence of caries reflects temporal and spatial trends in social stratigraphy, sex and gender inequality, diet and

*Corresponding author. E-mail: sneberger@seznam.cz

changes in food-preparation techniques and food habits. There is strong evidence to suggest that people of different social units consumed different food (Weiss 2009), that females have higher prevalence rates than males due to their reproductive physiology and advance susceptibility to oral disease, or that invention of agriculture had impaired effect on oral health. This is well documented in numerous transition processes.

Caries lesion is a disease manifested by focal demineralization of dental tissue caused by enzymes. The process of demineralization begins as a white-brown solid spot on the crown (Hillson 1996) and a yellow-brown soft spot on the root of the tooth (Nyvad, Fejerskov 1997). This is known as initial lesion or initial caries (Figure 1). The fact that most dental anthropologists do not take initial lesions into account in their evaluation is a serious issue (e.g. Hillson 1996; 2001; 2005; 2008). This is particularly problematic in dietary pattern studies. We always need to observe all kinds of pathologies. This would be the same as someone not assessing hairline fractures in a study engaged in traumas. There is additionally the issue of differentiating non-cavitated lesions from post-mortem changes due to diagenetic alteration. Distortion is still lower when we want to work with initial lesions than when we do not (Liebe-Harkort *et al.* 2010; 2011).



Figure 1. An example of the initial stage of the caries lesion. Photo by the authors: Using a Stereoscopic Zoom Microscope SMZ1500 (Nikon).

Caries lesion is represented in various stages, from slightly less visible demineralization to extensive cavitation or loss of the tooth crown (Figure 2). Caries lesions are the precisely localized destruction of enamel and dentine caused by bacterial production of acids in dental plaque (Selwitz *et al.* 2007). This process is primarily caused by bacteria from taxon *Streptococcus mutans*. Additional bacteria are *S. oralis*, *S. milleri*, *S. salivarius*, *Actinomyces neaslundii*, *A. viscosus* and *Lactobacillus acidophilus*. The fluctuating prevalence of dental caries in prehistory and history is due to three main factors: a) aggregation of complex oral bacterial flora (streptococcus) and individual susceptibility, b) exposure of the teeth surface (crown size, morphology, manner of subsistence, age, sex and the environment where the subjects lived), c) diet consistency and the manner in which the food is prepared (Larsen 1995; 1999; Larsen *et al.* 1991). It was previously believed that certain populations, such as Eskimos, were immune to caries lesions. Recent studies have shown that this is caused by the large amount of meat in the composition of the diet (Scott, Turner 1988). As a result of a diet mainly based on meat, hunters and gatherers are less affected by caries lesions than agriculturalists whose diet is based on grains and carbohydrates (e.g. Deter 2009). A tough diet (meat), however, in contrast to soft and sticky food inflicts the abrasion of teeth (e.g. Hillson 1996; 2005).

The great potential and importance of dental caries for bioarchaeological investigation is evident. The quality of the primary data and its precise evaluation is therefore essential. Detailed descriptions, using a magnifier and standard guidelines, are necessary for accurate assessment. A number of anthropological studies surprisingly differ in their assessment of the prevalence of dental caries (e.g. Čechová 1998; Pankowská 2009; Strouhal 1964). All the authors investigated Early Bronze Age populations from the Czech Republic (Čechová 1998; Pankowská 2009; Strouhal 1964). Čechová found that in 5–10% of the analysed individuals dental caries were present, Strouhal identified dental caries in 2.1–3.2% and Pankowská found them in 1.0–1.7% of the examined individuals. The difference between Pankowská and Strouhal is not as drastic as it is between Pankowská and Čechová and Strouhal and Čechová.



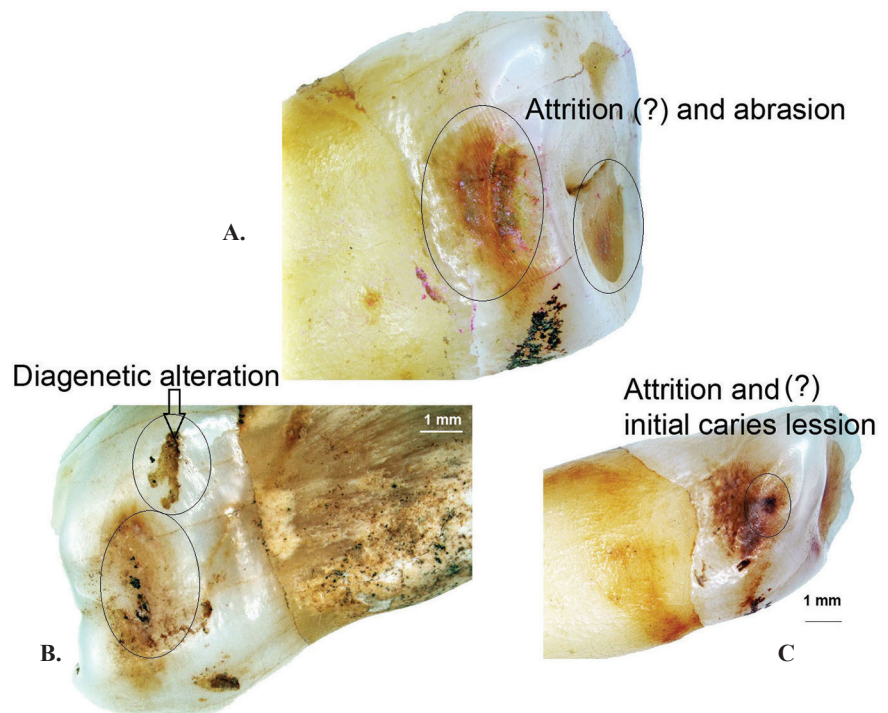
A.



B.

Figure 2. An example of A. visible demineralization of the tooth crown and B. partial crown destruction. Photo by the authors: Using a Stereoscopic Zoom Microscope SMZ1500 (Nikon).

Figure 3. Possible confusion of the dental caries. A. light brown spot produced in all probability by attrition and abrasion; B. light brown spot caused by diagenetic factors; C. possible attrition vs. initial dental caries lesion. Photo by the authors: Using a Stereoscopic Zoom Microscope SMZ1500 (Nikon).



There are two ways to look at these different interpretations. First, the dissimilarity between the amounts of caries lesions in the examined archaeological populations can be explained by bio-cultural factors (environmental factors, diet, economy, age, heterogeneity in mortality which causes intrinsic group factors such as total susceptibility, variation in social status, *etc.*). The second, more probable, interpretation is that individual researchers have had certain dissimilarity in their methods for evaluating dental caries which has led to mistakes and disagreement in assessment (Liebe-Harkort 2010).

This disagreement is a common problem in studies concerned with teeth from archaeological assemblages (Hillson 2001; 2008), which makes it difficult to compare different paleopathological and bio-archaeological studies. Those studies which include initial lesions tend to report a higher degree of caries than those confined to only frank cavities. Inconveniences in the evaluation of initial caries lesion are also reported in clinical dentistry (Ismail *et al.* 2007; Tam, McComb 2001; Basting, Serra 1999). Additional factors, however, impact the heterogeneity in the recording of dental caries even when the same methodology is used (Liebe-Harkort 2010). There are two factors which can impact the amount (over- or underestimate) of the identified dental defects: the experience of the observer and the state of preservation of the sample. Another issue is the effort to interpret ante-mortem teeth lost due to the influence of caries lesions (Larsen 1995).

Observer experience and branches of science are a crucial bias in disagreement concerning caries lesion evaluation. This kind of interest in dental caries stems from the completely different needs and concerns of various scientific disciplines. Dentists perceive lesions as an infectious disease

(comparison: Ismail 1997; Keenleyside 2008; Larsen 1995) of teeth which must be removed with a dental drill and filled with a filling. For anthropologists, however, caries lesions are a source of valuable information about the life of past populations and they are not allowed to apply these destructive analytical methods. The contexts of work with caries lesions are miscellaneous. The diagenetic change could be incorrectly identified as initial caries, or vice versa, because of dentists lacking experience with archaeological material. Dentistry works with certain chemical changes based on dental plaque which may cause discoloration of dental tissues. This discoloration can be easily distinguished from caries lesion and diagenetic changes. Diagenetic changes and initial lesions in archaeological assemblages are, in contrast, extremely similar (Figure 3). Dissimilarity such as this affects the perspectives of all observers from different branches of science.

In retrospective sciences such as physical anthropology and archaeology, researchers have problems with validity and reliability. This is particularly the case in morphoscopic assessment which is more subjective than morphometric assessment (Bruzek 1996). Bruzek *et al.* (1994) created four groups involving the most common reasons for errors in skeletal evaluations: the state of preservation; the definition of measurement points and the measurement itself; the accuracy of equipment and errors which can be made by an individual. Validity and reliability are possible through calculated intra- and inter-observer agreement (*e.g.* Alwas-Danowska *et al.* 2002; Garpel, Last 2005; Kiserund *et al.* 1999; Palmerim 1998). A number of studies have tried to highlight the difficulties and the importance of creating new reliable methods (Hillson 1996; 2001). Despite apparent inaccuracy and uncertainty, visual evaluation is still the most

commonly used method for assessment of caries lesions.

The aim of this paper is to assess the agreement in teeth and caries lesion identification between two groups of investigators: anthropologists and dentists. The question is how and why investigators will differ amongst one another depending on their subject field. We draw attention in this study to the risk of using accurate calculation based on extremely inaccurate data which can lead to an incorrect judgement in bioarchaeology.

2. Material

The material consists of 233 permanent isolated teeth (83 incisors, 45 canines, 51 premolars and 54 molars) obtained in an excavation of a medieval and modern hospital cemetery in Plzeň “U Zvonu”. The site was excavated in 2010/2011 as a rescue archaeological excavation on the site for a new building of the West Bohemian Gallery. It was conducted by the Museum of West Bohemia in Plzeň under the leadership of the archaeologists J. Orna and V. Dudková. Parts of the hospital church of Mary Magdalene and part of the contiguous cemetery were revealed throughout the excavation. All the teeth were visually sound or had a variety of occlusal, approximal, cemento-enamel junction (CEJ) or root caries lesions, both cavitated and non-cavitated.

3. Methods

All of the 233 teeth were evaluated by four specialists with varying levels of experience in visual diagnosis of carious lesions and teeth. The observers were comprised of two dentists: Dentist 1 (D1) and Dentist 2 (D2) and two anthropologists: Anthropologist 1 (A1) and Anthropologist 2 (A2). Both dentists had a range of experience with clinical dentistry which included work with dental caries. Dentist 1 was also familiar with dental paleopathology having participated in several anthropological studies. Anthropologist 1 had extensive experience with dental paleopathology. Anthropologist 2 was a beginner in dental anthropology.

Each researcher had two evaluations with a minimum interval of two weeks between each assessment. The study provided two consecutive steps:

1. Determining the type of tooth and its inclusion into the upper or lower jaw.
2. Evaluating all sides of the tooth for the presence and extent of caries lesion.

Each tooth was divided into three, individually assessed, segments: 1) crown; 2) root; 3) CEJ. The surfaces were evaluated in the following order: occlusal, buccal crown, mesial crown, oral crown, distal crown, buccal CEJ, mesial CEJ, oral CEJ, distal CEJ, buccal root, mesial root, oral root and distal root.

Visual inspection was undertaken in normal daylight with a combination of artificial lighting, without a magnifying glass or microscopes. A dental explorer was not allowed as this might damage the dental tissue. There was a requirement that each observer should be sufficiently relaxed and have a sufficient amount of time for evaluation of each tooth. These conditions were not strictly controlled and they were just recommended to avoid unnecessary mistakes.

The determination type of tooth was recorded by the letter “x” into the protocol where each tooth had own cell. The presence and extent of caries lesion was evaluated by the seven-grade and five-grade scale (modified after Hillson 2001 and Liebe-Harkort *et al.* 2010; 2011). The seven-grade scales were used for evaluation of the sides of the crown and CEJ of tooth. The five-grade scale was used, in contrast, for evaluation of the root. All points and scales are defined in Table 1.

Statistical methods

To express intra- and inter-observer agreement with standard references, kappa was calculated. In this study, Cohen un-weighted kappa (Cohen 1960), Cohen weighted kappa (Cohen 1968) and Fleiss kappa (Fleiss 1971) are presented.

The evaluation of the chosen statistical methods was divided into three sub-steps:

1. The agreement in determining the type of tooth and extent of caries lesion were calculated by Cohen weighted kappa.
2. Data gained from the inclusion tooth to the upper or lower jaw and evaluation of the presence of caries lesion were evaluated by Cohen un-weighted kappa.
3. Fleiss kappa was used for evaluation of agreement between all the observers together in case of determining the type of teeth and the presence of caries lesion on a particular tooth surface.

Table 1. Stages for evaluation of caries lesion. CEJ=Cemento-Enamel Junction.

Stages	Crown	Root	CEJ
0	Sound surface	Sound surface	Sound surface
1	Initial carie	Initial carie	Initial carie
2	One surface enamel carie	One surface dentinal carie	Crown and CEJ carie
3	One surface dentinal carie	One surface pulp carie	Crown and CEJ pulp carie
4	One surface pulp carie	Multiple surface dentinal caries	CJP and root carie
5	Multiple surface dentinal caries	Multiple surface pulp caries	CJP and root pulp carie
6	Multiple surface pulp caries		Crown, CJP and root carie
7	Absence of crown		Crown, CJP and root pulp carie

Table 2. Agreement by Landis, Koch (1977)

Kappa (k)	Strength of agreement
<0,00	Poor
0.01–0.20	Sight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Perfect

A table taken from Landis and Koch (1977) was used for expressing agreement of evaluation, where the authors divided agreement into five degrees and each level of conformity had word descriptions for easier understanding (Table 2). If the results had a positive value, they were statistically significant. This meant that the agreement was too high for it to be random (Landis, Koch 1977).

4. Results

A total of 233 teeth were examined. Each tooth had 13 surfaces associated into 3 units. This amounted to

3029 surfaces for visual evaluation carried out by one observer. The results were divided into five groups based on the type of evaluation and the pursued issues. The text only mentioned average results from all the observers. Detailed intra- and inter-observer results for all the observers are shown in Tables 3–7.

An evaluation dealing with the identification of the tooth type revealed a perfect consensus ($k=0.90$) in the intra-observer agreement and a substantial consensus ($k=0.77$) in the inter-observer agreement (Figure 4). The assessment of the inclusion teeth to the upper or lower jaw revealed a perfect consensus ($k=0.85$) in the intra-observer agreement and a substantial consensus ($k=0.66$) in the inter-observer agreement (Figure 5). Despite the same level of agreement, the k values declined. The lowest level of agreement and k values were in the evaluation agreement type of the tooth between all observers. The consensus had moderate values ($k=0.55$) in the inter-observer agreement.

Identification of the present caries lesions revealed a perfect consensus ($k=0.82$) in the intra-observer agreement and a moderate consensus ($k=0.55$) in the inter-observer agreement (Figure 6). Evaluating the extent of dental lesions revealed a substantial consensus ($k=0.78$) in the intra-

Table 3. Agreement in evaluation of the identification of a tooth. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2; Black colour: first evaluation of inter-observer agreement; Gray colour: second evaluation of inter-observer.

1 st evaluation					
intra-	D1	D2	A1	A2	
D1	<i>perfect (0.98)</i>	<i>substantial (0.71)</i>	<i>substantial (0.79)</i>	<i>substantial (0.79)</i>	D1
D2	<i>substantial (0.75)</i>	<i>substantial (0.80)</i>	<i>substantial (0.72)</i>	<i>substantial (0.74)</i>	D2
A1	<i>substantial (0.79)</i>	<i>substantial (0.76)</i>	<i>perfect (0.94)</i>	<i>perfect (0.88)</i>	A1
A2	<i>substantial (0.77)</i>	<i>substantial (0.74)</i>	<i>perfect (0.86)</i>	<i>perfect (0.89)</i>	A2
	D1	D2	A1	A2	intra-
2 nd evaluation					

Table 4. Agreement in evaluation of inclusion to the jaw. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2; Black colour: first evaluation of inter-observer agreement; Gray colour: second evaluation of inter-observer.

1 st evaluation					
intra-	D1	D2	A1	A2	
D1	<i>perfect (0.97)</i>	<i>moderate (0.51)</i>	<i>substantial (0.79)</i>	<i>substantial (0.72)</i>	D1
D2	<i>moderate (0.56)</i>	<i>moderate (0.68)</i>	<i>moderate (0.53)</i>	<i>moderate (0.60)</i>	D2
A1	<i>substantial (0.79)</i>	<i>substantial (0.63)</i>	<i>perfect (0.90)</i>	<i>substantial (0.79)</i>	A1
A2	<i>substantial (0.69)</i>	<i>moderate (0.57)</i>	<i>substantial (0.76)</i>	<i>perfect (0.83)</i>	A2
	D1	D2	A1	A2	intra-
2 nd evaluation					

Table 5. Agreement in evaluation of the presence of the caries lesion. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2; Black colour: first evaluation of inter-observer agreement; Gray colour: second evaluation of inter-observer.

1 st evaluation					
intra-	D1	D2	A1	A2	
D1	<i>perfect (0.98)</i>	<i>moderate (0.49)</i>	<i>substantial (0.73)</i>	<i>moderate (0.60)</i>	D1
D2	<i>moderate (0.56)</i>	<i>substantial (0.64)</i>	<i>fair (0.36)</i>	<i>moderate (0.51)</i>	D2
A1	<i>substantial (0.70)</i>	<i>moderate (0.44)</i>	<i>perfect (0.83)</i>	<i>substantial (0.63)</i>	A1
A2	<i>substantial (0.67)</i>	<i>substantial (0.66)</i>	<i>substantial (0.61)</i>	<i>perfect (0.84)</i>	A2
	D1	D2	A1	A2	intra-
2 nd evaluation					

Table 6. Agreement in evaluation of the extent of the caries lesion. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2; Black colour: first evaluation of inter-observer agreement; Gray colour: second evaluation of inter-observer.

1 st evaluation					
intra-	D1	D2	A1	A2	
D1	perfect (0.97)	moderate (0.45)	fair (0.33)	moderate (0.57)	D1
D2	moderate (0.56)	moderate (0.59)	slight (0.20)	moderate (0.55)	D2
A1	fair (0.33)	fair (0.25)	perfect (0.81)	fair (0.30)	A1
A2	moderate (0.55)	moderate (0.56)	fair (0.26)	substantial (0.75)	A2
	D1	D2	A1	A2	intra-
2 nd evaluation					

Table 7. Agreement between all the observers in evaluation of the identification of teeth and the presence of caries. k=kappa value; 95% CI = confidence interval; Kappa=strength of agreement.

	First evaluation			Second evaluation		
	k	95% CI	Kappa	k	95% CI	Kappa
Tooth Identification	0.53	0.52–0.55	moderate	0.54	0.52–0.55	moderate
Caries	0.36	0.34–0.38	fair	0.51	0.48–0.53	moderate

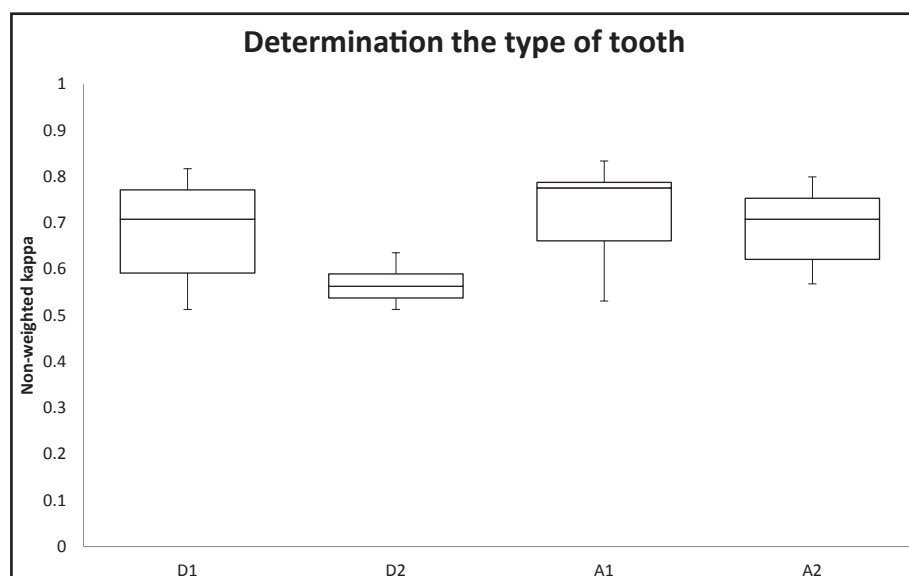


Figure 4. Agreement in determination of the type of tooth. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2.

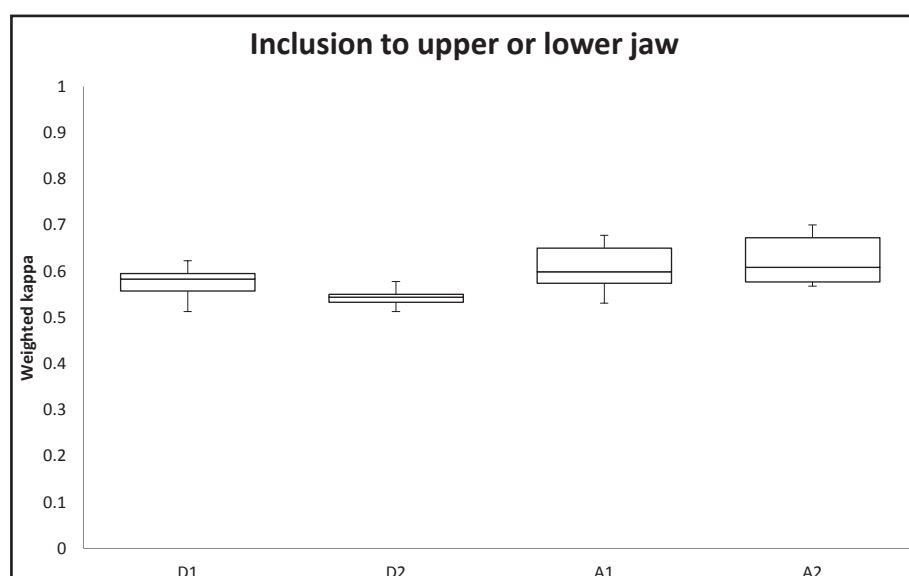


Figure 5. Agreement in the inclusion tooth to the upper or lower jaw. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2.

Figure 6. Agreement in the evaluation of the presence of the caries lesion. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2.

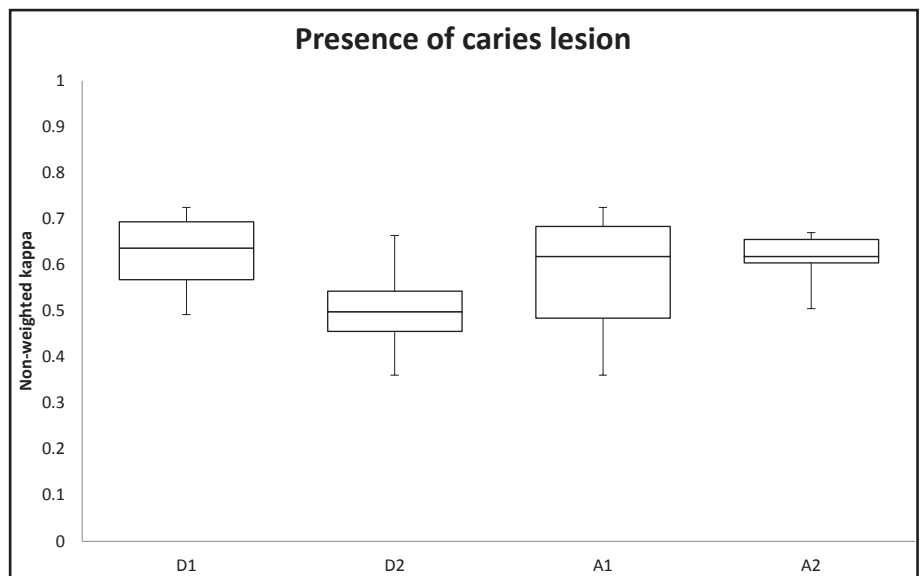
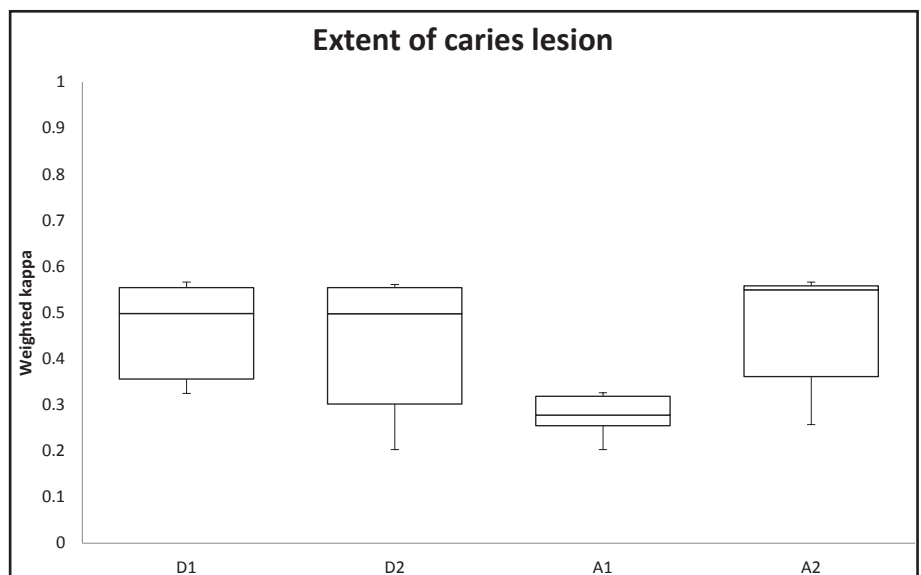


Figure 7. Agreement in the evaluation of the extent of the caries lesion. D1 and D2=Dentist 1 and 2; A1 and A2=Anthropologist 1 and 2.



observer agreement and a fair consensus ($k=0.40$) in the inter-observer agreement (Figure 7). The average agreement between all the observers in the evaluation of the present caries was moderate ($k=0.44$).

Caries lesion prevalence was low in all the observer evaluations. Dental caries were present in 1–6% of the 3029 assessed teeth surfaces based on the results of the observers. In the first evaluation, Dentist 1 identified 4%, Dentist 2 identified 5%, Anthropologist 1 identified 1% and Anthropologist 2 identified 6% of the teeth surfaces as decayed. In the second evaluation, Dentist 1 identified 4%, Dentist 2 identified 5%, Anthropologist 1 identified 2% and Anthropologist 2 identified 5% of the teeth surfaces as decayed. The percentages reveal the low amount of present caries lesions in the assessed assemblage.

In summary, the results indicate a solid agreement tendency in the intra-observer consensus with the exception of Dentist 2. His/her intra-observer agreement is $k=0.20$ lower than the other observers. With regard to the inter-observer agreement, Dentist 2 and Anthropologist 2 had the

lowest kappa results. Dentist 2 had the lowest kappa results in the parts of the study based on work with identification of teeth and Anthropologist 1 had the lowest kappa results in parts dealing with caries lesions. In contrast, the results reveal the lowest agreement with higher demands of evaluation. The obtained outcomes confirm the low reliability of the visual evaluation of the teeth and caries lesions. The study provides only one exception of good and useful results in the identification of teeth.

5. Discussion

In this study we tested the agreement between two dentists and two anthropologists in a visual evaluation of dental caries and teeth identification. 233 permanent teeth were used (3029 evaluated teeth surfaces) for this purpose.

Identification of present caries lesions obviously reveals a declining trend in agreement. This is most rapid in the inter-observation agreement where the consensus decreases

from an average $k=0.77$ to $k=0.55$. In assessing the extent of caries lesion, the declining trend continues. It is currently around $k=0.44$. This is in all probability caused by a steeper gradient increasing the demands of individual observation. Anthropologists are logically affected more because they have much less experience with caries lesions. For the sake of correctness, declining in agreement not only concerns anthropologists, but it concerns dentists as well. This is specifically Dentist 2, since his/her results were low. It is surprising since Dentist 2 has years of experience working with caries lesions and teeth. The most probable cause is the limited attention paid to the evaluation and a neglect of the methodological process. The greatest difference is apparent in the evaluation of the extent of caries lesions which needs the most time and attention. In contrast, Anthropologist 1 had *perfect* intra-observation agreement but *slight* and *fair* inter-observation agreement in the evaluation of the extent of caries lesions. There is also a methodological explanation. Anthropologist 1 disapproved initial caries lesion as a state of dental tissues which is in need of evaluation for obtaining a sufficient interpretation. These differences in the results and their causes revealed the necessity of using the same methodology and approach. The above-mentioned methodological problems only deplete, however, the 20–30% variability of evaluation error. The remaining evaluation error is in all probability caused by the unreliable and uncertain visual method.

The prediction of higher conformity was based on the logical fact of a similar experience in working with caries lesion and the intention of working with them. As was mentioned in the introduction, dentists need to detect dental decay and replace it with fillings. The most appropriate method in this situation is a dental probe. Dental probing is not allowed in evaluation of archaeological samples. Dentists are not accustomed to assessing caries lesions with only the use of the visual method and also are not familiar with diagenetic changes to dental tissue. Anthropologists have, in contrast, less experience with caries lesion than dentists, but are familiar with archaeological assemblages. The point of intention makes dentists more predisposed to determine more tooth surfaces as decayed. Anthropologists expect the presence of discoloration caused by the chemical process under the ground which can lead to determining initial caries as smooth surfaces only changed by chemicals. Liebe-Harkor *et al.* (2010; 2011) obtained the same conclusion, namely that observers from the same branch of science have a higher agreement evaluation. In the present study, the expected result was not accomplished. Surprisingly, a higher level of agreement was achieved between Dentist 1 and Anthropologist 1 and Anthropologist 2 (depending on the type of evaluation). The assumptions that dentists will overestimate and anthropologists underestimate caries lesions have not been met either. The discrepancy in the evaluation of caries lesion frequency was caused by Dentist 1 and Anthropologist 2. These two observers actually exchanged their roles with Dentist 1 having determined less caries lesions than Anthropologist 2. This fact was in all

probability caused by partial involvement with the opposite scientific branch.

Low dental caries prevalence was recorded in the assemblage. This could lead to a relatively high agreement in evaluation since a consensus in identification of a smooth tooth surface or healthy teeth is always higher than in assessment of decayed teeth. In any assemblage with a larger amount of carious teeth, it is probable that the agreement will be lower than in our study with the low number of present caries lesions.

The state of preservation and diagenetic change plays an important role in good determination. Diagenetic alteration of teeth causes problems in evaluation because they can look like caries lesions. This may lead observers to assess an altered surface as a caries lesion or vice versa. Due to these transformations, skeletons can be completely lost with the only aspect which survives being isolated teeth. Situations such as these reflect the need to be able to identify the type of tooth. Our results indicate insufficient agreement in determination of types. Although this is surprising, it is important to realise that it is extremely difficult at times to recognise types of teeth due to variations in morphology and diagenetic alteration (Pankowská *et al.* 2014).

The study provides unique results from assessing agreement among observers. This research is most appropriate in experimental sciences (Garpert, Last 2005; Kiserund *et al.* 1999; Sim, Wright 2005), technical disciplines (Alwas-Danowska *et al.* 2002; Keller *et al.* 2003), retrospective science and other natural science branches (Liebe-Harkort 2010; 2011; Palmeirim 1998). Analogous calculations should be included in all studies engaged in quantification of certain problems, provision of new methods, typological work or only for expression of the quality of our evaluation. These points are extremely important for future readers of our studies since it could obviously inform them about the potential in our new methods and interpretations. This study is consequently of appeal for all researchers who will read it in order to involve the calculation of agreement into their studies.

Generally, a suitable combination with good validity and reliability consists of visual, RTG and fluorescent methods. These methods could detect most dental caries types (cavity, initial caries, re-mineralized caries and conditions between) and should be helpful for distinguishing caries and diagenetic changes. There are no invariable directions as to which method is preferable in certain cases. The best approach is in all probability not to prioritize among these methods and in a synthesis part compare all data from all the used methods. We could consequently exclude certain types of caries and confirm others. The current separate method cannot cover all types of dental caries. It is therefore practically of no use to speak about validity and reliability when the employed method completely ignores a certain type of dental caries. Experience and knowledge are undoubtedly important parts of work with dental caries. In certain cases, however, not even experience and knowledge are sufficient for correct evaluation of the situation in the mouth. The best solution

is thereby a combination of optimal methods and adequate knowledge or collaboration with experienced colleagues.

6. Conclusion

The team of two dentists and two anthropologists performed an examination of the type of teeth and caries lesion of 233 teeth from an archaeological assemblage. The resulting values are positive in most intra-observation agreement with the exception of several evaluations performed by Dentist 2. In inter-observation, however, this is an acceptable consensus in the few cases of identification and the extent of caries lesions. We did not rate the researchers but instead were interested in demonstrating how visual macroscopic evaluation, a different methodology, observer experience and responsibility can lead to misclassification of dental caries and the identification of teeth. The results clearly indicate the low reliability and certainty of the visual method. It follows that an interpretation based entirely on the visual evaluation of teeth would most likely be distorted. Teeth bear valuable information about the lives of past populations. Loss or distortion of this information can lead to the creation of a completely incorrect picture of our ancestors. It is important to use more than just one type of diagnostic method with a well-elaborated methodology.

Acknowledgements

The authors would like to thank the participants for their time and effort. They would further like to thank Luke Lobo for his help with the statistical methods and Danny Angus for correction of the English. Our thanks are also due to the Museum of West Bohemia, namely to Jiří Orna and Veronika Dudková, for allowing us to study the skeletal sample Plzeň “U Zvonu”. This article was supported by SGS-2014-030 and NOTES CZ.1.07/2.3.00/20.0135.

References

- ALWAS-DANOWSKA, H. M., PLASSCHAERT, A. J. M., SULIBORSKI, S., VERDONSCHOT, E. H. 2002: Reliability and validity issues of laser fluorescence measurements in occlusal caries diagnosis. *Journal of Dentistry* 30, 129–134.
- BANERJEE, A., KIDD, E. A. M., WATSON, T. F. 2000: Scanning electron microscopic observations of human dentine after mechanical caries excavation. *Journal of Dentistry* 28, 179–186.
- BASTING, R. T., SERRA, M. C. 1999: Occlusal caries: Diagnosis and non-invasive treatments. *Restorative Dentistry* 30, 174–178.
- BOWERS, C.M. 2004: *Forensic dental evidence, An investigator's handbook*. Elsevier, London.
- BRUZEK, J. 1996: Interprétation biologique de séries archéologiques. Impact d'une diagnosesexuelle erronée à partir de simulation dans un échantillon de sexe connu. *L'identité des populations archéologiques* 16, 415–425.
- BRUZEK, J., MURAIL, P., HOUET, F., CLEUVENOT, E. 1994: Inter- and intra-observer errors in pelvic measurement and its implication for the methods of sex determination. *Antropologie* 32/3, 215–222.
- BUIKSTRA, J. E., UBELAKER, D. H. 1994: Standards For Data Collection From Human Skeletal Remains. *Arkansas Archaeological Survey Research Series* 44.
- ČECHOVÁ, L. 1998: *Dental Caries in the Period from the Bronze Age up to the Modern Times in Bohemia*. Karolium, Praha.
- COHEN, J. 1960: A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20/1, 37–46.
- COHEN, J. 1968: Weighted kappa: nominal scale agreement with provision for scale and disagreement or partial credit. *Psychological Bulletin* 70/4, 213–220.
- CUCINA, A., TIESLER, V., 2003: Dental caries and antemortem tooth loss in the Northern Peten area, Mexico: A biocultural perspective on social status differences among the Classic Maya. *American Journal of Physical Anthropology* 122, 1–10.
- DETER, C.A. 2009: Gradient of Occlusal Wear in Hunter-Gatherers and Agriculturalists. *American Journal of Physical Anthropology* 138, 247–254.
- FLEISS, J.L. 1971: Measuring Nominal Scale Agreement Among Many Raters. *Psychological Bulletin* 76/5, 378–382.
- GARCÍA-CLOSAS, R., GARCÍA-CLOSAS, M., SERRA-MAJEM L. 1997: A cross-sectional study of dental caries, intake of confectionery and foods rich in starch and sugars, and salivary counts of *Streptococcus mutans* in children in Spain. *American Journal of Clinical Nutrition* 66, 1257–1263.
- GARPET, R., LAST, J. 2005: The adult human occipital bone: Measurement variance and observer error. In: Zakrzewski, S.R., Clegg, M. (Eds.): *Proceedings of the Fifth Annual Conference of the British Association for Biological Anthropology and Osteoarchaeology*. BAR International Series 1383, 119–122.
- HILLSON, S. 1996: *Dental Anthropology*. Cambridge University Press, Cambridge.
- HILLSON, S. 1997: Relationship of Enamel Hypoplasia to the Pattern of Tooth Crown Growth: A Discussion. *American Journal of Physical Anthropology* 104, 89–103.
- HILLSON, S. 2001: Recording Dental Caries in Archaeological Human Remains. *International Journal of Osteoarchaeology* 11, 249–289.
- HILLSON, S. 2005: *Teeth*. Cambridge University Press, Cambridge.
- HILLSON, S. 2008: Dental Pathology. In: Katzenberg, M.A., Saunders, S.R. (Eds.): *Biological Anthropology of the Human Skeleton*, Second Edition. John Wiley & sons, 301–341.
- ISMAIL, A. I. 1997: Clinical diagnosis of precavitated carious lesions. *Community Dent Oral Epidemiol.* 25 (1), 13–23.
- KEENLEYSIDE, A. 2008: Dental Pathology and Diet at Apollonia, a Greek Colony on the Black Sea. *International Journal of Osteoarchaeology* 18, 262–279.
- KELLER, T. M., RAKE, A., MICHEL, S. C. A., SEIFERT, B., EFE, G., TREIBER, K., HUCH, R., MARINCEK, B., KUBIK-HUCH, R. A. 2003: Obstetric MR Pelvimetry: Reference Values and Evaluation of Inter- and Intra-observer Error and Intra-individual Variability. *Radiology* 227/1, 37–43.
- KILIAN, J., VLČEK, E. 1989: Age determination from teeth in the adult. In: Iscan, Y. M. (Eds.): *Age markers in the human skeleton*. Ch. Thomas Publisher, Springfield, 255–275.
- KISERUD, T., SAITO, T., OZAKI, T., RASMUSSEN, T., HANSON, M. A. 1999: Validation of diameter measurements by ultrasound: intra-observer and inter-observer variations assessed in vitro and in fetal sheep. *Ultrasound Obstet. Gynecology* 13, 52–57.
- KOHN, M. J., SCHOENINGER, M. J., BARKER, W. W. 1999: Altered states: Effects of diagenesis on fossil tooth chemistry. *Geochimica et Cosmochimica Acta* 68, 2737–2747.
- LANDIS, J. R., KOCH, G. G. 1977: The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33/1, 159–174.
- LARSEN, C. S. 1995: Biological Changes in Human Population with Agriculture. *Annual Review of Anthropology* 24, 185–213.
- LARSEN, C. S. 1999: *Bioarchaeology: Interpreting Behaviour from the Human Skeleton*. Cambridge University Press, Cambridge.
- LARSEN, C. S., SHAVIT, R., GRIFFIN, M. C. 1991: Dental Caries Evidence for Dietary Change: An Archaeological Context. In: Kelley, M. A., Larsen, C. S. (Eds.): *Advances in Dental Anthropology*. Wiley-Liss, New York, 179–202.
- LEE, S., LEE, U. Y., HAN, S. H., LEE, S. S. 2011: Forensic odontological examination of a 1500 year-old human remain in ancient Korea. *Journal of Forensic Odontostomatol* 29, 8–13.

- LIEBE-HARKORT, C., ÁSTVALDSDÓTTIR, Á., TRANAEUS, S. 2010: Quantification of Dental Caries by Osteologists and Odontologists – A Validity and Reliability Study. *International Journal of Osteoarchaeology* 20, 525–539.
- LIEBE-HARKORT, C., ÁSTVALDSDÓTTIR, Á., TRANAEUS, S. 2011: Visual and Radiographic Assessment of Dental Caries by Osteologists: A Validity and Reliability Study. *International Journal of Osteoarchaeology* 21, 55–65.
- LOVEJOY, C. O. 1985: Dental Wear in the Libben Population: Its Functional Pattern and Role in the Determination of Adult Skeletal Age at Death. *American Journal of Physical Anthropology* 68, 47–56.
- LUKACS, J. R. 1996: Sex Differences in Dental Caries Rates with the Origin of Agriculture in South Asia. *Current Anthropology* 37, 147–153.
- LUKACS, J. R. 2008: Fertility and Agriculture Accentuate Sex Differences in Dental Caries Rates. *Current Anthropology* 49, 901–914.
- LUKACS, J. R., LARGAESPADA, L. L., 2006: Explaining sex differences in dental caries prevalence: Saliva, hormones, and “life-history” etiologies. *American Journal of Human Biology* 18, 540–555.
- NELSON, G. C., LUKACS, J. R., YULE, P. 1999: Dates, Caries, and Early Tooth Loss During the Iron Age of Oman. *American Journal of Physical Anthropology* 108, 333–343.
- NYVAD, B., FEJERSKOV, O. 1997: Assessing the stage of caries lesion activity on the basis of clinical and microbiological examination. *Community Dentistry and Oral Epidemiology* 25, 69–75.
- PALMEIRIM, J. M. 1998: Analysis of skull measurements and measurers: Can we use data obtained by various observers. *Journal of Mammalogy* 79/3, 1021–1028.
- PANKOWSKA, A. 2009: Comparison of Health Status in Human Skeletal Remains Disposal in Settlements and Necropolises in the Early Bronze Age (In Central Moravia, Czech Republic). *Anthropologie, International Journal of Human Diversity and Evolution* 47/3, 215–228.
- PANKOWSKÁ, A., GALETA, P., ŠMOLÍKOVÁ, L., ŠNEBERGR, J., JURMAN, K. 2014: Hodnocení shody identifikace lidských stálých zubů v závislosti na jejich zachovalosti. *Česká Stomatologie* 114/1, 15–23.
- PIOVESAN, CH., MENDES, F. M., ANTUNES, J. L. F., ARDENGHI, T. M. 2011: Inequalities in the distribution of dental caries among 12-year-old Brazilian schoolchildren. *Brazilian Oral Research* 25, 69–75.
- REISINE, S. T., PSOTER, W. 2001: Socioeconomic Status and Selected Behavioral Determinants as Risk Factors for Dental Caries. *Journal of Dental Education* 2001, 1009–1016.
- SCOTT, G. R., TURNER, C. G. 1988: Dental Anthropology. *Annual Review of Anthropology* 17, 99–126.
- SCOTT, G. R., TURNER, C. G. 1997: *The anthropology of modern human teeth, Dental morphology and its variation in recent human populations*. Cambridge University Press, Cambridge.
- SELWITZ, R. H., ISMAIL, A. I., PITTS, N. B. 2007: Dental caries. *Lancet* 369, 51–59.
- SHEIHAM, A. 2006: Dental caries affects body weight, growth and quality of life in pre-school children. *British Dental Journal* 201, 625–626.
- SHEIHAM, A., BÖNECKER, M. 2008: Exploring the association of dental caries with social factors and nutritional status in Brazilian preschool children. *European Journal of Oral Sciences* 116, 37–43.
- SIM, J., WRIGHT, C. C. 2005: The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy* 85, 257–268.
- STROUHAL, E. 1964: Příspěvek k paleopatologii chrupu starší doby bronzové. In: Tesař, R. (Ed.): *Některé Stomatologické Problémy. Sborník k Šedesátinám prof. Škalouda*. Praha, 49–83.
- TAM, L. E., MCCOMB, D. 2001: Diagnosis of Occlusal Caries: Part II. Recent Diagnostic Technologies. *Pratique Clinique* 67, 459–463.
- TURNER, C. G. 1979: Dental anthropological indications of agriculture among the Jomon people of central Japan. X. Peopling of the Pacific. *American Journal of Physical Anthropology* 51, 619–635.
- UBELAKER, D. H. 1987: Estimating Age at Death from Immature Human Skeletons: An Overview. *Journal of Forensic Science* 32, 1254–1263.
- WALDRON, T. 2009: *Paleopathology*. Cambridge University Press, Cambridge.
- WALKER, P. L., HEWLETT, B. S. 1990: Dental Health Diet and Social Status among Central African Foragers and Farmers. *American Anthropologist* 92, 383–398.
- WEISS, E. 2009: *Bioarchaeological science. What we have learned from human skeletal remains*. Nova Science Publishers, New York.
- WESOŁOWSKI, V. 2006: Caries prevalence in skeletal series – Is it possible to compare? *Memórias do Instituto Oswaldo Cruz* 101, 139–145.