

https://dspace.org/

Assessing tuberculosis in the skeleton with the use of decision tree analysis

Deona Botha; Rethabile Masiu; Maryna Steyn

E SCHWEIZERBARTSCHE VERLAGSBUCHHANDLUNG
https://hdl.handle.net/10539/42515
WOS

Downloaded from a DSpace repository. https://dspace.org/

Anthropol. Anz. 81/2 (2024), 233–239
J. Biol. Clin. Anthropol.
Published online 20 October 2023, published in print March 2024

Article

Assessing tuberculosis in the skeleton with the use of decision tree analysis

Deona Botha[1,*], Rethabile Masiu[2], Maryna Steyn[1]

[1] Human Variation and Identification Research Unit, School of Anatomical Sciences, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg, 2193, South Africa
[2] Basic Medical Sciences, School of Biomedical Sciences, Faculty of Health Sciences, University of the Free State, South Africa
* Corresponding author: deona.botha@wits.ac.za

With 1 figure and 2 tables

Abstract: Diagnosis of specific infectious diseases in the skeleton is often difficult and relies on expert opinion. Statistics is not often used as a tool to assist in such diagnoses, and therefore this study aimed at employing data mining and machine learning in the form of decision tree analysis to aid in recognizing tuberculosis (TB) in skeletal remains and find patterns of skeletal involvement. The sample included 387 modern South African individuals (n = 207 individuals known to have died of TB and n = 180 as a control group) which were scored for the presence or absence of 21 skeletal lesions documented to be associated with TB. A pruned decision tree classification analysis was done to detect significant patterns and associations between variables which produced a model with a moderate classification rate based on four of the variables. As expected, vertebral changes were selected first, followed by rib, acetabular and lastly cranial changes. As a proof of concept, it was shown that machine learning was able to identify patterns of changes in TB skeletons versus a control group. However, further investigation into the use of machine learning in assessing skeletal changes associated with specific diseases is needed.

Keywords: TB; diagnosis; decision trees; palaeopathology; skeletal lesions

Introduction

As is common knowledge for all palaeopathologists, it is extremely difficult to diagnose a specific disease from skeletal remains alone. Bone can only react to an insult in a limited number of ways (bone formation, bone destruction, or a combination of these) (Ortner 2003), and therefore different diseases may appear very similar as far as bony changes are concerned. Palaeopathologists therefore look at patterns of changes in bones affected, in order to draw up a differential diagnosis. The assessment of whether some observed feature is indeed pathological in nature is, of course, subjective and may rely heavily on the experience of the observer.

As palaeopathologists often have only small samples to work with, and frequently report on diseases in single skeletons or small series of skeletons from a specific site, it is rare to see studies where some form of statistical analyses is used to differentiate between diseases or aid in the compilation of a differential diagnosis. Establishing a differential diagnosis mostly relies only upon expert identification of specific lesions and changes to bone. In their study, Dangvard-Pederson et al. (2019) used a probabilistic approach to assess the probability of having tuberculosis (TB) when specific skeletal lesions are present. They assessed 18 different lesions in a control and a TB example and calculated various probabilities of having TB – as expected, vertebral and rib lesions were most frequently associated with TB, but the specificity and sensitivity of these lesions to specifically diagnose TB were low.

Recently, Masiu et al. (2023) followed this up with a similar study in a South African sample. Instead of using only a TB and a control sample, they also included a sample of individuals known to have died of lung disease other than TB. Amongst other things, it was found that vertebral lesions (on the ventral parts of the bodies of thoracic and lumbar vertebrae) occurred frequently in both the pulmonary disease group (58%) and the TB group (55%). This emphasizes the difficulty in trying to come to a possible diagnosis using skeletal indicators, and reiterates that the complete skeleton should be assessed and caution is advised to over-interpret single lesions. In the Masiu et al. (2023) study, intracranial lesions were added and although not frequent, showed some promise as to their association specifically with TB.

© 2023 E. Schweizerbart'sche Verlagsbuchhandlung, 70176 Stuttgart, Germany
DOI: 10.1127/anthranz/2023/1737
www.schweizerbart.de
0003-5548/2023/1737 $ 1.75

Please note that certain pages of this article have been removed in order to reduce the file size so that the PDF can be uploaded on the system (the system has a limit of 1MB for files and several PDF files are larger than this).

The first and last pages of each paper (with full bibliographic details and affiliations) are included.

If the entire unredacted paper is required, this can be emailed directly to whomever requires them by contacting Dr. Busisiwe Maseko on Busisiwe.Maseko@wits.ac.za

In this paper, we attempt to further explore the patterns of skeletal involvement and their potential to possibly diagnose a specific disease, particularly as pertaining to TB. Here we explore the possibility of using AI (artificial intelligence), in particular decision trees, to help see patterns (and possibly hidden patterns) of changes that may occur in TB. Decision tree analysis is a machine learning technique used in both supervised and unsupervised learning. They are used to make predictions or decisions based on a set of data. Decision trees are a type of algorithm that uses a tree-like structure to map out the possible outcomes of a decision (Quinlan 1986; Podgorelec et al. 2002). It works by first splitting the data into smaller subsets and then evaluating the data in each subset. The process continues until the tree finds a solution that is optimal based on the criteria. This process is repeated until a decision tree is created that can accurately predict or classify the data (Rokach & Maimon 2014). A decision tree's starting point is the root node, which is the top node of a decision tree. This node is used to represent the starting point of the decision tree and it contains all the data that is used to make decisions. All other nodes are connected to the root node in some way. The root node is followed by decision nodes, and ultimately leaf (terminal) nodes, which is used to determine the best outcome of a decision once the data has been processed (Podgorelec et al. 2002; Nikita & Nikitas 2020).

The advantages of using decision tree analysis include the ability to handle large amounts of data, to handle complex data, and to visualize the data in a way that is easy to understand. Decision tree analysis can also tolerate missing data and can identify important relationships between variables. Disadvantages include potential bias if the dataset is too small or if pre-processing of the data was inadequate (Shen 2011; Rokach & Maimon 2014). The use of decision trees for identifying certain patterns related to disease has the potential to provide researchers, and in particular palaeopathologists, with an additional way of assessing diseases of bone in cases where doubt exists.

The aim of this paper was thus to assess the use of decision trees to assist in differentiating between TB and non-TB skeletons in a sample of individuals with known causes of death, but more importantly, to rank the variables in terms of its usefulness in assigning a possible diagnosis of TB. As a proof of concept, an assessment was conducted to determine patterns of changes that are commonly known to be associated with a specific disease (in this case TB) that may potentially be usable in palaeopathology.

Materials and methods

Sample

The sample used in this study was selected from the Raymond A. Dart Collection of Human Skeletons (Dayal et al. 2009; Steyn 2019) and comprised of the remains of 387 individuals. This study used the data collected by Masiu et al. (2023) but only included the individuals documented to have had died of tuberculosis (n = 177) and the control group (n = 150). The pulmonary disease sample used by Masiu et al. was thus excluded, as many of these individuals probably had underlying TB at some point during their lifetime. In order to increase the sample size, 30 individuals were added to each of these two groups (TB and control). While the Masiu et al. (2023) study included only pre-antibiotic individuals, this study also included individuals who have died in the post-antibiotic period, here assumed to have been after 1950 (Pfuetze et al. 1955; Steyn et al. 2013). The total sample thus included 207 individuals diagnosed to have died of tuberculosis and 180 individuals that died of causes other than tuberculosis or pulmonary disease. As the Raymond A. Dart Collection of Human Skeletons comprises largely of unclaimed bodies from hospitals, it is assumed that these individuals would have been from a lower socio-economic background (Dayal et al. 2009; Steyn 2019). All information related to the cause of death for the complete sample was obtained from the Dart collection's records.

Methodology

All individuals were scored for the presence or absence of 21 tuberculosis-related lesions in various areas of the skeleton. A score of 0 was assigned in the absence of lesions, and a score of 1 was given if lesions could be observed macroscopically. Skeletal locations included the cranium, ribs, vertebrae, pelvis, humerus, radius, ulna and femur. A detailed description of the different variables is provided in Table 1, which were adapted from Masiu et al. (2023). In contrast to the study by Dangvard-Pederson et al. (2019), this study also included intracranial lesions that were found to show some potential to be associated with TB (Hershkovitz et al. 2002; Steyn et al. 2013; Steyn & Buskes 2016). Only individuals that were complete or almost complete were included to ensure that a maximum of two variables be unavailable for scoring. As complete assessments for inter- and intra-observer reliability for scoring of lesions was done in the Masiu et al. (2023) study, they were not repeated here.

Statical analysis

To ensure that the frequency of variables observed in pre- and post-antibiotic eras did not influence the statistical outcome due to differences in the occurrence of skeletal lesions, a chi-squared test between the two groups for all variables assessed was performed. This was done using IBM SPSS Statistics (version 28).

Statistical analysis for the decision tree was performed using the free WEKA 3.8.5 software (Witten et al. 2011). The software program (Waikato Environment for Knowledge Analysis) has various build-in statistical algorithms, and thus no coding was required. A decision tree was constructed using the classification function J48 (pruned tree). All variables were included in the analysis, from which the pruned

algorithm removed branches of the tree that did not add value to the final prediction. This process helped to reduce the risk of overfitting the data and improving the accuracy of the predictions. Pruned decision trees are used in order to optimize the model and reduce the number of features that the model is using by including only significant variables.

Results

The chi-squared tests showed that no statistical significance differences were present in the frequency of lesions observed between the pre- and post-antibiotic eras (with all *p*-values being ≥ 0.012). Therefore, the dataset was handled as a single entity and decision tree analysis conducted on the combined groups.

The pruned tree algorithm provided a decision tree that incorporated the most significant variables related to tuberculous lesions in the study population. The purpose of a pruned decision tree algorithm is to reduce the complexity of the tree by removing irrelevant or redundant nodes. Pruning can improve the accuracy of the tree by removing nodes that do not provide any additional information, making the tree more efficient and user-friendly. The classifier model, based on the current dataset, recognized only four of the 21 variables as being useful for diagnosing tuberculosis. A ten-fold cross-validation (k = 10) was included in the model to prevent overfitting of the data. Overfitting of data occurs when a model begins to memorize the data instead of learning from it. This can lead to inaccurate results since the model is not able to generalize from the training data (Lantz 2018). To avoid overfitting, it is important to cross-validate using appropriate techniques.

Overall, the statistical output of the classifier model proved average with moderate prediction accuracy. Classification results on the full dataset (387 instances) yielded a 58% correctly classified rate with a kappa value of 0.1431 (indicating fair agreement between observed and expected values). The mean absolute error (MAE) demonstrating the magnitude of difference between predicted and true values was 0.4819, indicating that the decision tree has a moderate accuracy and error prediction. The weighted average F-measure (a measure of the model's accuracy with 1.0 indicating perfect precision) was 0.575, suggesting that the accuracy of the model on this dataset is moderate. The confusion matrix (Table 2) shows that 118 individuals with tuberculosis and 84 control individuals were categorized correctly and accurately.

The decision tree (Fig. 1) includes the four variables/skeletal locations that could be positively associated with tuberculosis and displays a sequence in which these variables should be scored to obtain the most accurate diagnosis. These four variables were VEN2, RIB2, ACE1 and CRA (see Table 1). Throughout the model, a score of 1 assigned to a variable indicated that there is a relatively strong possibility that tuberculosis caused the development of the skeletal lesion in that area.

The root node is indicated as VEN2 (ventral part of the thoracic and lumbar vertebrae showing bone proliferation that covers at least half of the surface), suggesting that this variable is the most significant in skeletal tuberculosis in this sample. The next variable in line (first leaf node) to be assessed is RIB2, or bone proliferation on the visceral sides of the ribs that covers more than 5 cm or lytic lesions covering more than 5 mm. Clustered pitting of the articular

Table 2. Confusion matrix indicating the number of individuals classified per group.

Classified as →	Tuberculosis	Control
Tuberculosis	118	72
Control	77	84

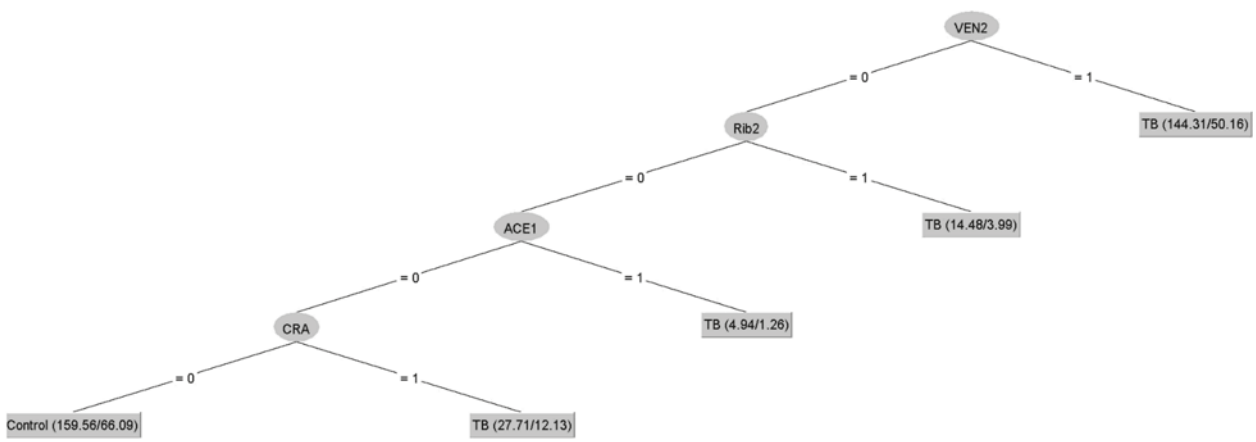

Fig. 1. Decision tree with root node (VEN2), followed by three leaf nodes (RIB2, ACE1 and CRA).

surface of the acetabulum (ACE1) were also found to be significantly related to tuberculosis and indicated as the second leaf node. Lastly, the third leaf node, where the tree terminates, shows that lesions in one or more of the cranial fossae may be indicative of tuberculosis.

It should be noted that the root node (vertebral bone proliferation) is the node best suited to split the data into the various decisions (Fig. 1) and has the strongest relationship with the other variables represented in the tree. According to this model, 144 individuals were correctly classified as having tuberculosis based on this variable (VEN2) alone, whereas 50 individuals were incorrectly classified. The number of individuals that could be classified by adding the subsequent variables decreases as one moves from the first (RIB2) to second (ACE1) leaf node. This shows that few individuals had lesions in all three of the first skeletal locations (vertebral bodies, ribs and acetabulum). It is possible that infection of the majority of individuals within the study group did not reach a stage where multiple skeletal areas were affected, most likely due to them succumbing to the disease. With the addition of lesions inside the cranium (the last variable to be added to the tree) the certainty of cases that could be classified as having tuberculosis increased. This suggests that the association between the first three variables and tuberculosis is strengthened by the presence of cranial lesions, although cranial involvement is not frequently seen in individuals within the study group. The termination node is reached when no further decision is needed, and in this case the presence of cranial lesions associated with the fossae acts as the final step in diagnosing tuberculosis in skeletal remains with accuracy.

The sequence of variables presented by the classifier model suggests that it is unlikely that cranial lesions on its own, and in the absence of lesions on the vertebrae, ribs or hip joint be associated with tuberculosis for this specific population group. It is recommended that all the variables included in the decision tree be used to make a diagnosis, starting with the root node in order to obtain the best outcome and most accurate result.

Discussion

In this study, we used decision trees in order to assess whether this kind of AI can potentially be usable to aid in diagnoses of skeletal pathology – in this case TB. The results are encouraging, in that the program was able to find patterns of skeletal involvement that confirms characteristic features that are generally known to palaeopathologists. As could be expected, the vertebral lesions, followed by rib lesions, showed the closest association with TB. It was somewhat surprising to see that the next best was the acetabulum, followed by lesions in the endocranial fossae that are not commonly associated with TB. Although previous studies have listed hip joint involvement to be the second most common skeletal location in tuberculosis (Roberts & Buikstra 2003; Ortner 2003), the involvement of the acetabulum as the only area affected was not frequently encountered in this specific sample. However, its association with other significant variables (i.e. affected vertebrae and ribs) was noteworthy. None of the other lesions added any new or usable information to the model.

While these patterns of changes (e.g., to vertebrae and ribs) are commonly known, it is useful to have their presence as indicators of TB supported by a different form of statistical analysis. It is also useful to be able to "rank" the skeletal indicators in the order of importance, and to realize that after changes in the endocranium, no other feature provided any additional support for a diagnosis of TB. Once again, the relative non-specificity of these lesions commonly associated with TB became clear, as the most informative of these (vertebrae) hovered between 50 and 60% in their specific association with TB (Dangvard-Pedersen et al. 2019; Masiu et al. 2023), similar to what was found in this study.

The epidemiology of tuberculosis has most likely changed from ancient to modern times (Murray & Lopez 1996; Mayer 2000). The addition of medical intervention and knowledge on the spread of tuberculosis in the 20th century undoubtedly aided in its prevention and management. With regards to the skeletal sample, antibiotics given to some of the individuals after 1950 in South Africa could have changed the patterns of skeletal expression of TB (Steyn et al. 2013; Steyn & Buskes 2016). As the pattern of skeletal involvement most likely differs between ancient and modern individuals, it may be the reason why skeletal features such as Potts' disease are not often encountered in modern skeletal collections. In this study no difference was observed between the pre- and post-antibiotic groups which may likely be contributed to both samples originating from a modern rather than an archaeological era. This suggests that the manifestation of skeletal lesions due to TB may be influenced by other factors in addition to antibiotic treatment. Another factor to consider is the evolution of both bacterium and host. Genetic evidence suggests that numerous strains of tuberculosis exist, which differ from one another in terms of aggressiveness and ability to spread (Braden et al. 2001). Milder forms of the infection as well as host resistance (McHenry et al. 2020) may possibly alter the levels of skeletal involvement. Although the factors contributing to the occurrence and severity of skeletal changes due to infection are still largely unknown, it is likely that the disease's epidemiology differs not only between populations, but also time periods.

It should of course be realized that any decision tree (and every other statistical test, for that matter) is only as good as the data that is used to build it. The larger the sample that is used for the building of the model, the better the results will be. While the current sample size (n = 387) is adequate to build a decision tree, the results will probably improve with a larger sample. Probably of more importance, though, is careful scrutiny of the individuals

that comprise the sample. The complete sample for this proof-of-concept study came from the Raymond A Dart Collection in the Gauteng Province, an area where TB was (and still is) extremely common. While the utmost care was taken to select a control group with no indication of TB or other lung diseases, the individual patient histories were not known and it is possible that they may have come into contact with TB during their lifetime (Van Schaik et al. 2014). It is thus important that similar studies are done in skeletons in other regions of the world to assess the usability of various statistical approaches.

Various authors (e.g., Buikstra et al. 2017; Klaus 2017) have called for increased rigor in paleopathology, especially when it comes to differential diagnosis, because of the overlapping nature of skeletal changes. While better descriptions of lesions, consistent use of terminology and interdisciplinarity can all improve the outcomes of paleopathological diagnoses, statistical analysis aiding in differential diagnoses should also be included. Decision trees may be one of the avenues to explore in this regard, and more research is needed as it holds the particular potential to observe hidden patterns that may not be visible to the naked eye in both modern and historical/archaeological skeletal remains. Methods such as this provide a ranking of lesions in terms of its association with TB, and may assist in assessing specific diseases in the skeleton.

References

Braden, C. R., Morlock, G. P., Woodley, C. L., Johnson, K. R., Colombel, A. C., Cave, M. D., … Crawford, J. T. (2001). Simultaneous infection with multiple strains of Mycobacterium tuberculosis. *Clinical Infectious Diseases, 33*(6), e42–e47. https://doi.org/10.1086/322635 PMID:11512106

Buikstra, J. E., Cook, D. C., & Bolhofner, K. L. (2017). Introduction: Scientific rigor in paleopathology. *International Journal of Paleopathology, 19*, 80–87. https://doi.org/10.1016/j.ijpp.2017.08.005 PMID:29198402

Dangvard Pedersen, D., Milner, G. R., Kolmos, H. J., & Boldsen, J. L. (2019). The association between skeletal lesions and tuberculosis diagnosis using a probabilistic approach. *International Journal of Paleopathology, 27*, 88–100. https://doi.org/10.1016/j.ijpp.2019.01.001 PMID:30661884

Dayal, M. R., Kegley, A. D. T., Štrkalj, G., Bidmos, M. A., & Kuykendall, K. L. (2009). The history and composition of the Raymond A. Dart Collection of Human Skeletons at the University of the Witwatersrand, Johannesburg, South Africa. *American Journal of Biological Anthropology, 140*(2), 324–335. https://doi.org/10.1002/ajpa.21072 PMID:19382178

Hershkovitz, I., Greenwald, C. M., Latimer, B., Jellema, L. M., Wish-Baratz, S., Eshed, V., … Rothschild, B. M. (2002). Serpens endocrania symmetrica (SES): A new term and a possible clue for identifying intrathoracic disease in skeletal populations. *American Journal of Physical Anthropology, 118*(3), 201–216. https://doi.org/10.1002/ajpa.10077 PMID:12115277

Klaus, H. D. (2017). Paleopathological rigor and differential diagnosis: Case studies involving terminology, description, and diagnostic frameworks for scurvy in skeletal remains. *International Journal of Paleopathology, 19*, 96–110. https://doi.org/10.1016/j.ijpp.2015.10.002 PMID:29198404

Lantz, B. (2018). Overfitting, Underfitting, and Model Selection. O'Reilly Media.

Masiu, R., Pedersen, D. D., Hill, L., & Steyn, M. (2023). The association between skeletal lesions and tuberculosis in a pre-antibiotic South African sample. *International Journal of Paleopathology, 40*, 20–32. https://doi.org/10.1016/j.ijpp.2022.11.001 PMID:36444782

Mayer, J. D. (2000). Geography, ecology and emerging infectious diseases. *Social Science & Medicine, 50*(7-8), 937–952. https://doi.org/10.1016/S0277-9536(99)00346-9 PMID:10714918

McHenry, M. L., Williams, S. M., & Stein, C. M. (2020). Genetics and evolution of tuberculosis pathogenesis: New perspectives and approaches. *Infection, Genetics and Evolution, 81*, 104204. https://doi.org/10.1016/j.meegid.2020.104204 PMID:31981609

Murray, C. J. L., & Lopez, A. D. (1996). *The Global Burden of Disease*. Cambridge: Harvard University Press.

Nikita, E., & Nikitas, P. (2020). Data Mining and Decision Trees. In Z. Obertova, A. Stewart, & C. Cattaneo (Eds.), *Statistics and Probability in Forensic Anthropology* (pp. 87–105). London: Academic Press; https://doi.org/10.1016/B978-0-12-815764-0.00012-5

Ortner, D. (2003). Identification of Pathological Conditions in Human Skeletal Remains (2nd ed.). USA: Academic Press.

Pfuetze, K. H., Pyle, M. M., Hinshaw, H. C., & Feldman, W. H. (1955). The first clinical trial of streptomycin in human tuberculosis. *American Review of Tuberculosis, 71*(5), 752–754. PMID:14361985

Podgorelec, V., Kokol, P., Stiglic, B., & Rozman, I. (2002). Decision trees: An overview and their use in medicine. *Journal of Medical Systems, 26*(5), 445–463. https://doi.org/10.1023/A:1016409317640 PMID:12182209

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81–106. https://doi.org/10.1007/BF00116251

Roberts, C. A., & Buikstra, J. E. (2003). *The Bioarchaeology of Tuberculosis*. Gainesville: University Press of Florida.

Rokach, L., & Maimon, O. (2014). *Data Mining with Decision Trees: Theory and Applications*. World Scientific: Series in Machine Perception Artificial Intelligence, Vol. 81. https://doi.org/10.1142/9097

Shen, Y. (2011). A decision tree approach to medical diagnosis. *Journal of Biomedical Informatics, 44*(5), 848–853. PMID: 21645638

Steyn, M. (2019). The Raymond A. Dart collection as a mirror of society: A case of Tuberculosis. *Adler Museum Bulletin, 45*(1), 1–6.

Steyn, M., & Buskes, J. (2016). Skeletal manifestations of tuberculosis in modern human remains. *Clinical Anatomy (New York, N.Y.), 29*(7), 854–861. https://doi.org/10.1002/ca.22688 PMID:26779783

Steyn, M., Scholtz, Y., Botha, D., & Pretorius, S. (2013). The changing face of tuberculosis: Trends in tuberculosis-associated skeletal changes. *Tuberculosis (Edinburgh, Scotland), 93*(4), 467–474. https://doi.org/10.1016/j.tube.2013.04.003 PMID: 23664272

van Schaik, K., Vinichenko, D., & Rühli, F. (2014). Health Is not always written in bone: Using a modern comorbidity index to assess disease load in paleopathology. *American Journal of Physical Anthropology, 154*(2), 215–221. https://doi.org/10.1002/ajpa.22494 PMID:24936606

Witten, I., Frank, E., & Hall, M. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Burlington, MA: Morgan Kaufmann Publishers.

Manuscript received: 9 May 2023
Revisions requested: 18 July 2023
Revised version received: 3 August 2023
Manuscript accepted: 23 August 2023