

Special Article - Biological Anthropology

Extirpating Inherently Biased Rote Approaches, Replacing them with Critical Thinking-Based Interpretation of Evidence

Rothschild BM*

Carnegie Museum, 4400 Forbes Ave, Pittsburgh, USA

*Corresponding author: Bruce M Rothschild, Carnegie Museum, 4400 Forbes Ave, Pittsburgh, PA 15213, USA

Received: April 02, 2021; Accepted: April 26, 2021;

Published: May 03, 2021

Abstract

Physical anthropology embraces a hunger for understanding the ecosphere in which we live, its impact on the life and health and our impact environmental signature. This has been pursued tenaciously, utilizing speculative approaches, with lesser attention to assuring adherence to fundamentals. The resulting perceptions of the environment and of the people therein, both contemporary and ancient, have been subject to a variety of biases. Although many are obvious and discussed in detail in this manuscript, it seems appropriate to question why they have persisted. What benefit does the biased individual gain? Certainly not advancement beyond circular reasoning, which itself reinforces the proponent and their philosophies. It certainly is easy to pursue studies and their promulgation by rote, minimizing cognitive effort expenditure. It is easier to pontificate a technique, than to pursue and assure its independent validation. It is easier to assume that students are performing correctly, without expending the rigor/time of/for actually testing fundamentals and assuring the validity of one's own techniques. It is easier to stalwartly defend the status quo that has defined one's life to date, than to subject it to potential modification and thus to consider critical thinking as an existential threat. Perhaps that explains apparent aversion to and attempts to block promulgation of evidence that application of scientific methodology to physical anthropology provides an opportunity for meaningful contributions beyond salvage work. The latter has value, but physical anthropology can offer much more. Extirpating the biases would be a major step in that direction and resurrect logos.

Keywords: Bias; Integrity; Scientific method; Evidence; Paleopathology

Introduction

Scientific methodology (logos) is predicated upon generating hypotheses and testing them, following where the collected data/evidence leads [1-9]. The data often generates new hypothesis. Once such a hypothesis is generated, there are two major components: Unbiased elicitation/extraction of pertinent data/evidence and analysis and interpretation of that evidence. The former may seem straight-forward; the latter, perhaps not so easily pursued. Seemingly, straight-forward data acquisition is itself actually quite susceptible to "mischief," compromising or even precluding interpretation. The current review examines some of the challenges that have compromised advancement of physical anthropologists from variably recording observations on the basis of a priori, speculation-based criteria to a rigorous scientifically based evidence accession approach that is amenable to interpretation. This is not the first time that modifications have been suggested to the approaches utilized in physical anthropology. Washburn (1951) emphasized the importance of transforming the field from reporting findings in individuals to population-oriented research, noting that this is especially true for understanding of disease. If physical anthropology is to meaningfully contribute (be relevant) to our understanding of ourselves and our kin (e.g., other primates), of our origins, and the effect/impact of our environmental interactions (e.g., health and disease, global

warming), it seems that attention to fundamentals and application of scientifically-based approaches by individuals whose skills have been vetted (independently verified) would facilitate that effort.

Avoidance of all speculation on the meaning of observations has been previously suggested (e.g., [10]), but even that approach is compromised because of the speculative assumptions involved in making the observations themselves. The Heisenberg [11] uncertainty principle, notation that the very process of observation alters that which is observed, may not directly apply to non-destructive analyses of inert materials (e.g., skeletons), but taphonomic processes inherent in specimen handling must be considered. However, there is an even more fundamental consideration: Observations are not made in isolation, but have context [12]. Such a simple physical aspect as work area lighting (when specimens are examined) may exaggerate or camouflage perception. Perhaps the most important challenge relates to the speculation that a given examiner's observations are valid; not just that he/she adhered to criteria that had been scientifically-vetted, but also that his/her skills in applying the criteria have themselves been independently validated. If underlying systematic bias (e.g., speculative criteria, insufficient skills in recognizing the element(s) necessary for fulfillment of the criteria) is not recognized, the results that are often perceived as consistent with preconceived notions actually lack accuracy.

The prepared mind often sees what it is prepared to see. "Recognizing someone's hand, for instance, significantly limits the possible interpretations of the object on that person's wrist. This a priori knowledge allows the visual system to sensitize the corresponding visual representations so that it is easier to recognize the surrounding objects when we attend to them. In fact, these context-driven predictions can allow us to choose not to attend to this object at all, if none of the possible identities 'suggested' by the context are of immediate interest" ([13], p. 617). Representing and processing objects in groups that tend to be found together might explain why recognition of an object that is highly associated with a certain context facilitates the recognition of other objects that share the same context" ([13], p. 617, citing [14,15]) and noting the same recognition performance effect from spatial relations. This has been referred to as expectation sets facilitating perception - context frames, schemata, scripts [13]. Thus, 'boundary extension' derives. This is a form of brain pre-processing memory distortion/distraction, in which reporting is not limited to what is physically present, but incorporates past experiences, often reporting what is not there, creating a "rush to judgement" rather than scientifically evidential assessment [15].

"Evidence-based data rests on certain philosophical assumptions: A singular truth, ascertainable through empirical enquiry; a linear logic of causality in which interventions have particular effect sizes; rigor defined primarily in methodological terms (especially, a hierarchy of preferred study designs and tools for detecting bias); and a deconstructive approach to problem-solving (the evidence base is built by answering focused questions, typically framed as 'PICO' - Population-Intervention-Comparison-Outcome)" [16]. It is the very rigor inherent to the scientific method that has been compromised by physical anthropology-related biases. Fundamentals have been assumed without verification/vetting. There are at least eight major components to such lapses.

Failure to Separate Speculation from Evidence

Speculation is not evidence. It is only an unproven hypothesis that has not been scientifically tested/vetted. Repetition of a speculative comment is not evidence. It simply imbues a mythology. As Douglas Verret noted (12 September 2018, personal communication), consensus is political, not scientific. Many believe that scientific consensus identifies a group in which there is total agreement that a statement is valid. Actually, such a consensus is often simply a group that agrees with and supports each other. It is not evidential.

Failure to Recognize Normal Variation

Biehler-Gomez et al. [17] assessed anthropologists who were three to ten years post degree, using a checklist, noting interobserver disagreement (20-59 %), especially in recognition of periosteal reaction, noting overdiagnosis (making diagnosis in absence of disease) in 71% of cases. The article by Baker et al. [18] exemplifies this ongoing challenge (see discussion below).

Insufficient Skills to Describe the Character Deviations from the Norm

Biehler-Gomez et al. ([17], p. 8) specifically note, "Description of bone lesions constitutes a specific skillset that is not entirely obtained

through the experience of a forensic pathologist or anthropologist or their specialty in forensic anthropology or bioarchaeology, despite their ability, range of knowledge and practice. The results of this study thus testify to the need for more specific and through training in the description of bone lesions to all practitioners on dry bone, regardless of their field of specialization or experience." This explains previous interobserver disagreement in studies by Bridges [19], Waldron and Rogers [20] and Baker et al. [18] among others. Botham's [21] analysis of what should be a relatively straight-forward issue, recognition of nonlethal cranial injuries, further documents this lapse. She clearly delineates overdiagnosis and failure to delineate sufficient recognition criteria, let alone adhere to them.

Failure to Use Evidence-Based Criteria

Zuckerman et al. ([22], p. 37) report transformation of "paleopathology from a descriptive enterprise to an interpretive, interrogative and independent one that is increasingly focused on ecological, epidemiological and, finally, evolutionary considerations". That lofty statement unfortunately fails to address the derivation of interpretations, still depending on speculative, not evidence-based diagnoses. They report (page 39) that "this new approach found fertile, though limited, ground in paleopathology." They ([22], p. 46) report Ortner's [23] notation of "lack of methodological and theoretical rigor," which unfortunately persists today, as exemplified by Baker et al. [18].

Zuckerman et al. ([24], p. 276) make the appropriate comment that lesions found in one disorder should be compared with those in "other conditions and in healthy controls to determine their specificity." That is correct, but their application is severely flawed. The study they criticized (Rothschild and Rothschild, 1995) as one that established a data base for recognizing disease as a population phenomenon was actually predicated specifically on study of one component [25]. That study actually described all changes present that are not routinely noted in healthy populations (both human and non-human). It was not purposed to distinguish among diseases, but rather established a database for such pursuits. That is the first step in describing the character of a disease: Identifying everything that seems at variance with previous experience examining thousands of skeletons, both human and non-human. However, there actually was one error in that study. The anthropologist coauthor [25] of the study was based in Guam, but failed to alert (nor correct in manuscript drafts) his coauthor that tibial flattening (not seen in population samples from other parts of the world) was actually characteristic of South Pacific populations [26]. Fault is admitted for assuming collaborator's diligent participation as a study coauthor [25]. Great effort IS subsequently exerted to independently vet collaborator input, with enhancing productivity [27,28].

However, there is a second, even more pertinent issue. Some (e.g., Brender Baker, Gillian Crane-Kramer, Lesley Gregoricka, Christine Lee, Shiela Lukehart, Charlotte Roberts, Ann Stodder, Ann Stone, Stevie Winingear) have denigrated those independently validated studies because "no clear description or images to demonstrate periosteal reaction" were provided ([18], p. 20). Rejecting evidence-based criteria, they instead provided a list of criteria that they speculate (without evidence) allow diagnosis of treponemal disease. Blunder is exemplified by their speculative criteria for recognition of

periosteal reaction (e.g., longitudinal striae), a phenomenon which has been tested and falsified as anything other than normal bone [29]. Conflating normal structures with those related to pathology is likely responsible for the alleged inability of individuals (e.g., Brender Baker, Gillian Crane-Kramer, Lesley Gregoricka, Christine Lee, Shiela Lukehart, Charlotte Roberts, Ann Stodder, Ann Stone and Stevie Winingear) to reproduce findings reported by teams that are perhaps more conscientious in their application scientific methodology in observational and diagnostic approaches.

Further, inclusion of pits as diagnostic markers of periosteal reaction by Brender Baker, Gillian Crane-Kramer, Lesley Gregoricka, Christine Lee, Shiela Lukehart, Charlotte Roberts, Ann Stodder, Ann Stone and Stevie Winingear (2019) exposes their fundamental lack of understanding of bone biology. The periosteum is the outer layer of the bone. That facilitates distinguishing periosteal reaction (which occurs on top of the periosteum, in contrast to taphonomic changes, which expose bone below the periosteal layer. Inclusion of pits (phenomena clearly exposing bone subjacent to the periosteum), as a diagnostic character for recognition of periosteal reaction, perhaps explains what appears to be frequent conflation of taphonomic changes with periosteal reaction by these individuals and those exposed to their speculations (e.g., their students). The criteria offered by Baker et al. [18] certainly explain variable reproducibility of findings between those who apply them and those who are more circumspect in attention to physiologic processes.

Zuckerman et al. ([24], p. 276) unfortunately seem to extrapolate their comparison (“with other conditions in healthy controls”) suggestion to application in DNA work. While complete genomes have specificity, segments do not. After all, there are segments of retroviral DNA in all of us [30]. If such segments happened to be chosen to compare organisms, we would all appear to be a virus. This illustrates a major difference between extrapolating by rote and critical thinking.

Part of the challenge has been the initially untested speculation [23,31], subsequently falsified [27], that “diseases cannot be expected to manifest in the same way in every environment or human population”. Harper and Armelagos [32] address the importance of paleoepidemiologic study, speculating on what diseases might have been present at different times, but without actually provided databased criteria for documenting their actual presence. Hahn et al. ([33], p. 2) note that the major cause of “errors may arise from inappropriate pre-analytics, which include all working steps prior to the actual measurement”.

Efforts to standardize osteological data are especially critical when access to collections will be subsequently compromised [34]. However, that standardization also needs to be vetted. As recording and localizing every detail is extremely time-consuming, short cuts are often invoked. A stage may be speculatively assigned to a phenomenon, without recognition that the stages may actually be manifestation of different derivation/pathophysiology (e.g., [35,36]). Specifics are lost, which cannot be resurrected, and the pertinent information, lost (e.g., [37-39]). This emphasizes the importance of vetting the classification system.

Gaps in Understanding of Pathophysiology

Assumption of the association of cribra orbitalia with anemia [40-

47], based on a speculative relationship to porotic hyperostosis, is one such example This is especially egregious given lack of correlation of the two phenomena (cribra orbitalia and porotic hyperostosis [48,49].

What can be learned from other fields? Sonnenschein and Soto ([10], p.1) decry ‘the lack of significant improvements in the understanding of carcinogenesis’. The consensus that has been achieved as to the elements “remained mostly irrelevant both to understanding carcinogenesis and to significantly benefiting the object of the whole effort, the cancer patient. There is a misunderstanding of the basic biological phenomenon, analogous to that concerning the relationship of porotic hyperostosis to iron deficiency anemia [40-47,49].

Insufficient Understanding of Statistical Methods and Their Premises

Part of the challenge to scientific studies has been the speculation [23,31], subsequently falsified [27], that “diseases cannot be expected to manifest in the same way in every environment or human population”. Harper and Armelagos [32] address the importance of paleoepidemiologic study, speculating on what diseases might have been present at different times, but without actually provided databased criteria for documenting their actual presence. Hahn et al. ([33], p. 2) note that the major cause of “errors may arise from inappropriate pre-analytics, which include all working steps prior to the actual measurement.”

In a study of post-graduate trainees comprehension of statistics, “38% could not apply the concept of specificity and sensitivity” ([50], p. 1). This is further exemplified by the otherwise excellent article by Plomp et al. [51]. Interested in the reproducibility of findings across populations, they selected a few individuals from a plethora of sites, rather than examining all from each site or utilizing a random number system to identify appropriate candidates. Thus, the studied individuals were not identified by a statistically valid method and therefore violated that statistical premise, rendering their statistics moot.

“Computer-discovered correlations” seemingly have “replaced understanding and guide prediction and action” ([52], p. 1). They note what they referred to as arbitrary correlations only due to size, not nature, emphasizing the importance of understanding the statistical technique and its premises.

Hahn et al. ([33], p. 2) note that sensitivity and specificity are common sense criteria, but there are other considerations: “measurement accuracy, accuracy expressed as systematic error, comparative precision expressed as random error, and repeatability, theoretical and practical limits of detection.” They further note (p.3) that “sample preparation usually starts with the correct choice of specimens” and “the requirement for “much training and experience”.

Failure of specific reporting that does not faithfully reflect the nature and range of findings distorts impressions [53]. This can result from “lack of understanding of methodological principles, parroting of common practices, a form of unconscious behavior, or an actual willingness to misread” ([53], p. 2613). Such spin often favors the author’s vested interest. Reporting methods, but not failure to comply with them, has been reported as beautification. This includes selective reporting of statistically significant results or cherry picking

choice of statistical test on the basis of which (e.g., listed by SAS) gives the most impressive results or those most compatible with the authors' preconceived notions or biases. Post hoc removal of outliers, modification of study group assignments or covariates also violate statistical premises [53-55]. Boutron and Ravaud ([53], p. 2617) recommend change the perception of spin from 'commonly accepted practice' to recognize it as a 'seriously detrimental research practice'.

That brings us to the subject of metaanalysis. They violate statistical assumptions of event/data independence, and are compromised by non-linearity and fractal effects (Succi and Coveny, 2019) and at time suggest biologically implausible/impossible events, especially when data entry is predicated upon speculative interpretations [56]. Collinearity is the term used to characterize non-independence of predictor variables [57]. Yazici ([58], p. 1) reported "46% of all metaanalysis publications have their conclusions changed by publications with falsified data and 32% of all analyses had a considerable change in the outcome," citing [59]. Only 3% are methodologically sound and non-redundant [58,60].

Biased, Selective Citation, Leading to Circular Logic

Falcone ([61], p. 122) comments on bias in citation practices, suggesting, "citation lies at the very heart of our gifting rituals." She further states (p. 124) "those gifts are not free and volitional... but that gifts are always a part of a complex system of obligations". Obligatory celebration of the work of one's academic advisors or colleagues contrasts with negative reciprocity, defined as trying to maximize utility at the expense of others. This is magnified by reviewers who express outrage that their publication(s) are not cited. This is highlighted by apparent preference to cite tertiary rather than primary sources. A tertiary source is typically a citation of another article, which itself describes the author(s)' perspective of the information provided by the primary source. Recalling the childhood game of "telephone," there is potential for much to be lost in the process, with reversal of findings and other misstatements. Assurance of the validity of the information can only be pursued by examination and vetting of the data and interpretations provided in primary sources. This is as important when the information supports one's own preconceived notions, as when it conflicts with them.

Snoddy et al. ([62], p. 89) noted, "Lack of awareness of best anthropological practices by scholars from other professional spheres can perpetuate a misunderstanding of the level of scientific study in our field", while failing to acknowledge the converse. They also noted that "palaeopathological methods have sometimes suffered from a kind of circular logic wherein older anthropological literature, which is no longer accurate, is used as the foundation for entire diagnostic schemes," with editorial acquiescence allowing its perpetuation.

Semantic Errors Resulting from Failure to Recognize that Different Fields Share a Common Lexicon, Often Quite Different in Meaning

Not only do the same terms differ significantly across fields of scientific endeavor (see glossary in [28], they also have historical context. This is exemplified by the 10,000 leprosaria that Pope Clement closed in 1508-1510. While the appellation led to the

presumption (e.g., [63]) that the 100 or so individuals buried in each of the associated cemeteries had leprosy, the authors failed to recognize that the term leprosy was historically applied to essentially any individual with a skin condition. They also seemed to have not done the math: 10,000 leprosaria with 100 individuals hospitalized per leprosaria equal a million burials. Suggestion that a million people had leprosy seems quite unreasonable. The evaluation of the Batavia leprosaria in Suriname by van Dissel et al. [64] revealed that neither archival, skeletal examination, nor DNA testing revealed any evidence of leprosy. So-called leprosaria were actually not repositories for leprosy. Further, the character and distribution of pathology reported by Møller-Christensen [63] are at variance with (different from) that observed in contemporary clinics and hospitals devoted to leprosy [65].

Conclusion

Snoddy et al. ([62], p. 89) noted, "Lack of awareness of best anthropological practices by scholars from other professional spheres can perpetuate a misunderstanding of the level of scientific study in our field", while failing to acknowledge the converse. They also noted, "Palaeopathological methods have sometimes suffered from a kind of circular logic wherein older anthropological literature, which is no longer clinically accurate, is used as the foundation for entire diagnostic schemes". These statements are repeated to emphasize that the concerns are not just those of the author of this manuscript. All testify to/document the need for more specific and more through training in the description of bone lesions for all practitioners of dry bone research, regardless of their field of specialization, experience or academic credentials. There is hope. Turn from the evil ways. Cease pursuing entrenched rote approaches at the expense of critical thinking. And, when applying statistics, review their premises to assure that data input and relationships fulfill those premises. And, consider reading Daniel Kahneman's *Thinking Fast and Slow* [66-85].

Research Highlights

- Speculation is only an unproven hypothesis that has not been scientifically tested/vetted.
- Inadequate attention to assurance of fundamentals (e.g., examination skills and understanding of physiology and statistics may result in reproducibility, at the expense of accuracy.
- If underlying systematic bias (e.g., speculative criteria, insufficient skills in recognizing the element(s) necessary for fulfillment of the criteria) is not recognized, the results that are often perceived as consistent with preconceived notions actually lack accuracy.

References

1. Ayala FJ. Darwin and the scientific method. *Proceedings of the National Academy of Science USA*. 2009; 106: 10033-10039.
2. Dodick J, Argamon S, Chase P. Understanding scientific methodology in the historical and experimental sciences via language analysis. *Science Education*. 2008; 18: 985-1004.
3. Hren D, Lukić IK, Marušić A, Vodopivec I, Vujaklija A, Hrabak M, et al. Teaching research methodology in medical schools: students' attitudes towards and knowledge about science. *Medical Education*. 2004; 38: 81-86.
4. Kell DB, Oliver SG. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-

- genomic era. *Bioessays*. 2004; 26: 99-105.
5. Kuhn TS. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press. 1962.
 6. McLelland CV. *The nature of science and the scientific method*. Boulder, Colorado: The Geological Society of America. 2006.
 7. Pérez DG, Torregrosa JM. A model for problem solving in accordance with scientific methodology. *European Journal of Science Education*. 1983; 5: 447-455.
 8. Wagensberg J. On the existence and uniqueness of the scientific method. *Biological Theory*. 2014; 9: 331-346.
 9. Willis BH, Beebe H, Lasserson DS. Philosophy of science and the diagnostic process. *Family Practice*. 2013; 30: 501-505.
 10. Sonnenschein C, Soto AM. Over a century of cancer research: Inconvenient truths and promising leads. *PLoS Biology*. 2020; 18: e3000670.
 11. Heisenberg W. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik*. 1925; 33: 879-893.
 12. Hayes SM, Nadel L, Ryan L. The Effect of scene context on episodic object recognition: Parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*. 2007; 17: 873-889.
 13. Bar M. Visual objects in context. *Nature Review Neuroscience*. 2004; 5: 617-629.
 14. Biederman I, Mezzanotte RJ, Rabinowitz JC. Scene perception detecting and judging objects undergoing relational violations. *Cognitive Psychology*. 1982; 14: 143-177.
 15. Bar M, Ullman S. Spatial context in recognition. *Perception*. 1996; 25: 343-352.
 16. Greenhalgh T. Will COVID-19 be evidence-based medicine's nemesis? *PLoS Medicine*. 2020; 17: e1003266.
 17. Biehler-Gomez L, Indra L, Martino F, Campobasso CP, Cattaneo C. Observer error in bone disease description: A cautionary note. *International Journal of Osteoarchaeology*. 2020: 1-9.
 18. Baker BJ, Crane-Kramer G, Dee MW, Gregoricka LA, Henneberg M, Lee C, et al. Advancing the understanding of treponemal disease in the past and present. *Yearbook of Physical Anthropology*. 2020; 171: 5-41.
 19. Bridges PS. The effect of variation in methodology on the outcome of osteoarthritic studies. *International Journal of Osteoarchaeology*. 1993; 3: 289-295.
 20. Waldron T, Rogers JU. Inter-observer variation in coding osteoarthritis in human skeletal remains. *International Journal of Osteoarchaeology*. 1991; 1: 49-56.
 21. Botham AD. Unthinking empiricism and the overdiagnosis of nonlethal cranial injuries: An interdisciplinary review of diagnostic criteria for healing, depressed cranial fractures. *Journal of Archeological Science: Reports*. 2019; 27: 101939.
 22. Zuckerman MK, Turner BL, Armelagos GJ. Evolutionary thought in paleopathology and the rise of the bicultural approach. In A. L. Grauer (Ed.), *A Companion to Paleopathology*. 2012: 34-57.
 23. Ortner D. Skeletal pathology: Probabilities, possibilities, and impossibilities. In J Verano, D Ubelaker, (Eds.). *Disease and Demography in the Americas* (pp. 5-13). Washington DC: Smithsonian Institution Press. 1992.
 24. Zuckerman MK, Harper KN, Armelagos GJ. Adapt or die: Three case studies in which the failure to adopt advances from other fields has compromised paleopathology. *International Journal of Osteoarchaeology*. 2016; 26: 375-383.
 25. Rothschild BM, Heathcote G. Characterization of the skeletal manifestations of the treponemal disease, yaws, as a population phenomenon. *Clinical Infectious Disease*. 1993; 17: 198-203.
 26. Rothschild BM. Odyssey in the evolution of a paleopathologist. *Fossil Record*. 2016; 20: 37-45.
 27. Rothschild BM, Martin LD. *Skeletal Impact of Disease*. Albuquerque: New Mexico Museum of Natural History Press. 2006.
 28. Rothschild BM, Schultze H-P, Peligrini R. *Herpetological Osteopathology: Annotated Bibliography of Amphibians and Reptiles*. Heidelberg, Germany: Springer-Verlag. 2012.
 29. Rothschild BM, Jellema L. Periosteal reaction recognition and specificity assessed by surface microscopy *International Journal of Osteoarchaeology*. 2020; 30: 355-361.
 30. Griffiths DJ. Endogenous retroviruses in the human genome sequence. *Genome Biology*. 2001; 2: reviews1017.1-reviews1017.5.
 31. Buikstra JE, Cook D. Paleopathology: An American account. *Annual Review of Anthropology*. 1980; 9: 433-470.
 32. Harper K, Armelagos G. The changing disease-scape in the third epidemiological transition. *International Journal of Environmental Research and Public Health*. 2010; 7: 675-697.
 33. Hahn A, Podbielski A, Meyer T, Zautner AE, Loderstädt U, Schwarz NG, et al. On detection thresholds- a review on diagnostic approaches in the infectious disease laboratory and the interpretation of the results. *Acta Tropica*. 2020; 205: 105377.
 34. Kakaliouras AM. When remains are "lost." Thoughts on collections, repatriation, and research in American physical anthropology. *Museum Journal*. 2014; 57: 213-223.
 35. Eamers K. Using osteohistology for ageing and sexing. In D. Ruscillo (Ed.), *Recent Advances in Ageing and Sexing Animal Bones*. 9th ICAZ Conference, Durham. 2002: 9-39.
 36. Milner GR, Boldsen JL. *Transition analysis age estimation: Skeletal Scoring Manual*. State College, Pennsylvania: Penn State University, State College, Pennsylvania. 2016.
 37. Imamura T, Saiki K, Okamoto K, Maeda J, Matsuo H, Wakebe T, et al. Characterization of individuals with sacroiliac joint bridging in a skeletal population: analysis of degenerative changes in spinal vertebrae. *BioMedical Research International*. 2014; 2014: 879645.
 38. Langley NR, Jantz LM, Ousley SD, Jantz RL, Milner G. *Data Collection Procedures for Forensic Skeletal Material 2.0*. Knoxville, Tennessee: University of Tennessee Forensic Anthropology Center: Knoxville, Tennessee. 2016.
 39. Ventades NG, Laza IM, Hervella M, de-la-Rúa C. A recording form for differential diagnosis of arthropathies. *International Journal of Paleopathology*. 2018; 20: 45-49.
 40. Fairgrieve SI, Molto E. Cribra orbitalia in two temporally distinct population samples from the Dakhleh Oasis, Egypt. *American Journal of Physical Anthropology*. 2000; 111: 319-331.
 41. Grauer AL. Pattern of anemia and infection from Medieval York, England. *American Journal of Physical Anthropology*. 1993; 91: 203-213.
 42. Huss-Ashmore R, Goodman AH, Armelagos GJ. Nutritional inference from paleopathology. *Advances in Archaeological Method and Theory*. 1982; 5: 395-474.
 43. McFadden C, Oxenham MF. A paleoepidemiological approach to the osteological paradox: Investigating stress, frailty and resilience through cribra orbitalia. *American Journal of Physical Anthropology*. 2000; 220: 1-13.
 44. Mittler DM, van Gerven DP. Developmental diachronic and demographic analyses of cribra orbitalia in the Medieval Christian populations of Kolumbarti. *American Journal of Physical Anthropology*. 1994; 93: 287-297.
 45. O'Donnell L Hill EC, Anderson AS, Edgar HJ. Cribra orbitalia and porotic hyperostosis are associated with respiratory infections in a contemporary mortality sample from New Mexico. *American Journal of Physical Anthropology*. 2020; 173: 721-733.
 46. Trancho G. Cribra orbitalia in the remains of ancient Nubians. *American Journal of Physical Anthropology*. 1987; 74: 511-520.
 47. Wapler U, Crubézy E, Schultz M. Is cribra orbitalia synonymous with anemia?

- Analysis of cranial pathology in Sudan. *American Journal of Physical Anthropology*. 2004; 123: 333-339.
48. Rothschild B, Zdilla M, Jellema L, Lambert H. Cribra orbitalia is a vascular phenomenon unrelated to marrow hyperplasia or anemia: Paradigm shift for cribra orbitalia. *Anatomical Record* (in press). 2020.
 49. Walker PL, Bathurst RR, Richman R, Gjerdrum T, Andrushko VA. The causes of porotic hyperostosis and cribra orbitalia: A reappraisal of the iron-deficiency-anemia hypothesis. *American Journal of Physical Anthropology*. 2009; 139: 109-125.
 50. Araoye I, He JK, Gilchrist S, Stubbs T, McGwin G, Ponce BA. A national survey of orthopaedic residents identifies deficiencies in the understanding of medical statistics. *Journal of Bone and Joint Surgery*. 2020; 102: e19.
 51. Plomp KA, Dobney K, Weston DA, Viðarsdóttir US, Collard M. 3D shape analyses of extant primate and fossil hominin vertebrae support the ancestral shape hypothesis for intervertebral disc herniation. *Evolutionary Biology*. 2019; 19: 226.
 52. Calude CS, Longo G. The deluge of spurious correlations in big data. *Foundations of Science*. 2017; 22: 595-612.
 53. Boutron I, Ravaut P. Misrepresentation and distortion of research in biomedical literature. *Proceedings of the National Academy of Science USA*. 2018; 115: 2613-2619.
 54. Schulz KF, Grimes DA. Multiplicity in randomized trials. II. Subgroup and interim analyses. *Lancet*. 2005; 365: 1657-1661.
 55. Taylor J, Tibshirani RJ. Statistical learning and selective inference. *Proceedings of the National Academy of Science USA*. 2015; 112: 7629-7634.
 56. Fonseca LL, Joyner C, Consortium TM, Galininski MR, Voit EO. A model of *Plasmodium vivax* concealment based on *Plasmodium cynomolgi* infections in *Macaca mulatta*. *Malaria Journal*. 2017; 16: 375.
 57. Dormann CF, Eliith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013; 36: 27-46.
 58. Yazici Y. Metaanalyses, network metanalyses and systematic reviews: The perpetual motion machine all over again. *Journal of Rheumatology*. 2020; 47: 1-3.
 59. Garmendia CA, Nassar Gorra L, Rodriguez AL, Trepka MJ, Veledar E, Madhivanan P. Evaluation of the inclusion of studies identified by the FDA as having falsified data in the results of meta-analyses: The example of the apixaban trials. *Journal of the American Medical Association Internal Medicine*. 2019; 179: 582-584.
 60. Ioannidis JP. The mass production of redundant, misleading, and conflicted systemic reviews and meta-analyses. *Milbank Quarterly*. 2016; 94: 485-514.
 61. Falcone JM. The Hau of theory: The Kept-Gift of theory itself in American anthropology. *Anthropological Humanism*. 2013; 38: 122-145.
 62. Snoddy AM, Beumont J, Buckley HR, Colombo A, Halcrow SE, Kinaston RL, et al. Sensationalism and speaking to the public: Scientific rigour and interdisciplinary collaborations in paleopathology. *International Journal of Paleopathology*. 2020; 28: 81-91.
 63. Møller-Christensen V. *Bone Changes in Leprosy*. Denmark: Munksgaard. 1958.
 64. Van Dissel JT, Pleeters T, Geluk A, Maat G, Menke HE, Tió-Coma M, et al. Archival, paleopathological and aDNA-based techniques in leprosy research and the case of Father Petrus Donders at the leprosarium 'Batavia', Suriname. *International Journal of Paleopathology*. 2019; 27: 1-8.
 65. Rothschild BM, Rothschild C. Skeletal manifestations of leprosy: Analysis of 137 patients from different clinical settings in the pre- and post-modern treatment eras. *Journal of Clinical Rheumatology*. 2001; 7: 228-237.
 66. Kahneman D. *Thinking Fast and Slow*. Farrar, Straus and Giroux, New York. 2011.
 67. Anonymous. *Wit & Wisdom*. The Week. 2019; 20: 23.
 68. Carvalho S. Anonymous peer review: Truth or trolling. 2020.
 69. Chaudhry S, Schroter S, Smith R, Morris J. Does declaration of competing interests affect readers' perceptions? A randomised trial. *British Medical Journal*. 2002; 325: 1391-1392.
 70. DeWitte SD, Stojanowski CM. The osteological paradox 20 years later: Past perspectives, future directions. *Journal of Archaeological Research*. 2015; 23: 397-450.
 71. Farke AA, Wolff ED, Tanke DH. Evidence of combat in Triceratops. *PLoS ONE*. 2009; 4: e4252.
 72. Goodman AH. On the interpretation of health from skeletal remains. *Current Anthropology*. 1993; 34: 281-288.
 73. Haffar S, Bazezachi F, Murad M. Peer review bias: A critical review. *Mayo Clinic Proceedings*. 2019; 94: 670-676.
 74. Kyle B, Reitsema LJ, Tyler J, Fabbri PF, Vassallo S. Examining the osteological paradox: Skeletal stress in mass graves versus civilians at the Greek colony of Himera (Sicily). *American Journal of Physical Anthropology*. 2018; 167: 161-172.
 75. Lee CJ, Sugimoto CR, Zhang G, Cronin B. Bias in peer review. *Journal of the American Society for Information Science Technology*. 2013; 64: 2-17.
 76. Leek JT, Taub MA, Pineda FJ. Cooperation between referees and authors increases peer review accuracy. *PLoS ONE*. 2011; 6: e26895.
 77. O'Connor E, Cousar M, Lentini J, Castillo M, Halm K, Zeffro T. Efficacy of double-blind peer review in an imaging subspecialty journal. *American Journal of Neuroradiology*. 2017; 38: 230-235.
 78. Pulverer B. Transparency showcases strength of peer review. *Nature*. 2010; 468: 29-31.
 79. Rogers J, Waldron T, Dieppe P, Watt I. Arthropathies in palaeopathology: The basis of classification according to most probable cause. *Journal of Archaeological Science*. 1987; 14: 179-193.
 80. Sanderg P, Sponheimer M, Lee-Thorp J, van Gerven D. Addressing the osteological paradox using high resolution stable isotope analysis. *American Journal of Physical Anthropology*. 2013; 157: 240.
 81. Silbiger NJ, Stubler AD. Unprofessional peer reviews disproportionately harm underrepresented groups in STEM. *Peer J*. 2019; 7: e8247.
 82. Tempelhoff J. Peer review in academic journals: A Pandora's box. *South African Journal of Science*. 2020; 116: 12-13.
 83. Wilcox C. Rude reviews are pervasive and sometimes harmful, study finds: Survey finds that "unprofessional" comments abound. *Science*. 2019; 366: 1433.
 84. Wood JW, Milner GR, Harpending HC, Weiss KM. The osteological paradox" Problem of inferring prehistoric health from skeletal. *Current Anthropology*. 1992; 33: 343-370.
 85. Wright LE, Yoder CJ. Recent progress in bioarchaeology: Approaches to the osteological paradox. *Journal of Archaeological Research*. 2003; 11: 43-70.