VERTEBRAL COMPRESSION FRACTURES: TOWARDS A STANDARD SCORING METHODOLOGY IN PALEOPATHOLOGY

Journal:	*International Journal of Osteoarchaeology*
Manuscript ID:	Draft
Wiley - Manuscript type:	Short Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Curate, Francisco; University of Coimbra, Research Centre for Anthropology and Health Silva, Tânia; University of Coimbra, Department of Life Sciences Cunha, Eugénia; University of Coimbra, Forensic Sciences Centre – National Institute of Legal Medicine
Keywords:	vertebral fractures, osteoporosis, scoring methods, reliability, paleopathology

VERTEBRAL COMPRESSION FRACTURES: TOWARDS A STANDARD SCORING METHODOLOGY IN PALEOPATHOLOGY

Francisco Curate[1,2,3*]; Tânia F. Silva[3]; Eugénia Cunha[2,3]

[1] Research Centre for Anthropology and Health – University of Coimbra, Portugal

[2] Forensic Sciences Centre – National Institute of Legal Medicine, Portugal

[3] Department of Life Sciences – University of Coimbra, Portugal

*corresponding author

Adress: CIAS – Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Apartado 3046

3001-401 Coimbra, Portugal

E-mail: fcurate@uc.pt

Other authors email:

TFS: taniadaniela_silva@hotmail.com

EC: cunhae@ci.uc.pt

Running title: Vertebral Compression Fractures In Paleopathology

ABSTRACT

Vertebral compression fractures are the most common osteoporotic fractures in postmenopausal women. Notwithstanding, its clinical diagnosis remains ambiguous. In paleopathological studies vertebral fractures and/or deformations are frequently disregarded. When observed, vertebral compression fractures are usually recorded without the support of quantifiable and comparable protocols. As such, a semi-quantitative method for vertebral compression fracture assessment (Genant et al., 1993) was applied to a large sample (N=198) from the Coimbra Identified Skeletal Collection, Portugal, and the reliability of the method was tested. Vertebral fracture scoring agreement was evaluated with the Kappa statistic and the percent of agreement. Intra-observer and inter-observer agreement are both appropriate. The Genant's semi-quantitative scoring methodology is easy to apply and highly

reproducible; as such, it should be adopted as the standard method to score vertebral fractures/deformations in any paleopathological investigation.

KEYWORDS vertebral fractures; osteoporosis; scoring methods; reliability; paleopathology.

INTRODUCTION

Osteoporosis (OP) is a metabolic pathological disorder characterized by the decrease in bone mass and quality and subsequent increase in fracture risk (NIH Consensus Development Panel, 2001). OP is essentially symptomless prior to bone fracture (Wylie, 2010), being classically associated with fractures in the proximal femur, the distal radius and the vertebral body (Johnell and Kanis, 2006).

Vertebral compression fractures and/or deformations are both the most common and underdiagnosed of the so-called osteoporotic fractures in postmenopausal women (Johnell and Kanis, 2006; Grados et al., 2009). The clinical diagnosis of vertebral compression fractures is ambiguous, inasmuch as there is not a consensual definition. They are frequently asymptomatic which translates in their underestimation in clinical practice (Delmas et al., 2005; Grados et al., 2009). Visual assessment is the most common method used in the clinical practice, but the results are exceedingly reliant on the knowledge of the observer (Ferrar et al., 2005; Olmez et al., 2005).

Descriptions of vertebral compression fractures in the paleopathological literature are not infrequent. Nevertheless, they commonly denote anecdotal cases (e.g., Foldes et al., 1995; Ortner, 2003; Reis et al., 2003; Sambrook et al., 1988; Strouhal et al., 2003), or refer to visual qualitative methods for the identification of vertebral fractures (e.g., Domett and Tayles, 2006; Hirata and Morimoto, 1994; Ives, 2007; Mays, 1996; Mays, 2006; Mays et al., 2006; Mensforth and Latimer, 1989; Snow, 1948). The «Spine Score» (Barnett and Nordin, 1960) has been employed for the definition of vertebral fractures in

archaeological populations (Gonzalez-Reimers et al., 2004). Other studies (e.g., Curate et al., 2009; Curate et al., 2013; Garcia, 2007) used Genant□s semi-quantitative method (Genant et al., 1993) for the evaluation of vertebral compression fractures.

Reproducible methods for the assessment of vertebral compression fractures, defined by unequivocal criteria, are to be favored in clinical and epidemiological settings, as well as in archaeological contexts. As such, this study aims to test the reliability of a semi-quantitative method for vertebral compression fractures and/or deformations assessment (presence/absence of fracture) in a skeletal sample from the Coimbra Identified Skeletal Collection.

MATERIALS AND METHODS

The sample studied comprised 196 individuals from the Coimbra Identified Skeletal Collection (Rocha, 1995), evenly distributed from both sexes, with an age-at-death ranging from 20 to 96 years old. The sample included individuals born between 1827 and 1914; and dead between 1910 and 1936. Individuals were typically blue-collar workers with low socioeconomic status. Only individuals with a complete vertebral column, without gross post-depositional and pathological modifications at the vertebral column were included in the sample.

Vertebral compression fractures and/or deformations were assessed macroscopically in the T4 through L4 vertebrae, with the Genant's scoring method (Genant et al., 1993). This semi-quantitative evaluation method is based on the vertebral shape (wedge, concave or crush) and on decreases in the anterior, posterior and/or middle vertebral heights (Figure 1), as follows:

1. Grade 0, no reduction;
2. Grade 1, minimal fracture, 20 – 25% decrease in any vertebral height;
3. Grade 2, moderate fracture, 25 – 40% decrease;
4. Grade 3, severe fracture, +40% decrease.

The first author (FC, Obs1), an experienced observer, evaluated 196 individuals in two different occasions. The second author (TFS, Obs2), an inexperienced observer, assessed 75 individuals after being clarified how to use the method and without the aid of an anatomical atlas. Both intra- and inter-observer variability in the assessment of vertebral fractures and/or deformations (presence/absence) were evaluated with the percent of agreement (%A; Watkins and Pacheco, 2000) and Cohen's Kappa (κ_c; Cohen, 1960). The percent of agreement is defined as:

$$\%A = (N - N'/ N)\times 100,$$

in which N corresponds to the total number of pairwise comparisons, and N' to the number of discordant pairs. Cohen's Kappa coefficient measures pairwise agreement for categorical variables, while correcting for projected chance agreement (Carletta, 1996; Rothwell, 2000). In the case of intra-observer reliability, agreement was assessed *per* subject, and not *per* vertebra. For inter-observer variability, agreement was estimated *per* subject and *per* vertebra. Bias index for the Kappa coefficient was also estimated (Sim and Wright, 2005).

All measurements (anterior, posterior and middle vertebral heights) were directly performed in the vertebrae, placed in lateral projection, with the aid of a digital outside caliper. Statistical analyses were achieved with IBM® SPSS® (version 19.0.0).

RESULTS

Both %A and κ_c suggest a remarkable level of intra-observer agreement between observations *per* individual. Inter-observer variability was somewhat higher but the measures of agreement between observers were also very satisfactory, both *per* individual and *per* vertebra. Bias index for the Kappa coefficient is very low (Table 1). Notwithstanding, while the inexperienced observer correctly identified all the actual vertebral fractures/deformations, it also incorrectly recorded grade 1 fractures/deformations in four individuals that were not affected. Also, when both observers recorded a fracture, the

attributed grade was consistently the same, except for one vertebral fracture/deformation (Obs1; grade 1 *vs.* Obs2; grade 2).

DISCUSSION

In paleopathological studies regarding trauma, vertebral fractures and/or deformations are often ignored. When observed, vertebral compression fractures are usually described without the assistance of quantifiable and reproducible protocols (Curate et al., 2011).

Genant's scoring method (Genant et al., 1993) displays a binary classification of vertebral fractures/deformations (present/absent), an evaluation of fracture severity (grades 0 to 3) and a visualization of vertebral shape after fracture (wedge, crush or concave). In this study, intra- and inter-observer reliability in the assessment of vertebral fracture presence was evaluated. Intra-observer agreement amongst observations was excellent, with a very high percent of agreement, and a Kappa coefficient (non chance agreement) reflecting an almost perfect agreement (Landis and Koch, 1977). Inter-observer agreement was also very high, with the κ_c statistic suggesting a lower, but still substantial agreement between observers. The literature on the subject supports these results (Genant et al., 1993; Grados et al., 2009; Li et al., 1995).

Bias for the Kappa coefficient is low, and disagreement between observations and observers is probably due to random error. Nonetheless, a negligible tendency for the inexperienced observer to record non-existent fractures was observed. The analysis of small isolated deformations in the vertebral column is occasionally complex (El Maghraoui et al., 2009). Also, while 20% reductions in any vertebral height have been proposed to define a minimal fracture/deformity, it is clear that borderline cases are difficult to interpret (Black et al., 1999) – especially in the case of untrained observers (Figure 2). Although the Genant et al. (1993) method is straightforward, it shows a learning curve, being dependent on training and experience (Grados et al., 2009).

Visual assessment of vertebral compression fractures is simple and useful for ruling out vertebral deformities due conditions other than osteoporosis. Nevertheless, reproducibility is very low (Grados et al., 2009; Jensen et al., 1984). Clinical and epidemiological trials with qualitative readings of vertebral compression fractures demonstrate the great variability in the identification of those fractures, which mainly corresponds to the interpretation of vertebral radiographs without standardized guidance, references to anatomical atlas and consensus readings by doctors and technicians (Black, 1999; Olmez et al., 2005). Quantitative morphometric methods (e.g., Eastell et al. 1991; McCloskey et al., 1993) are objective and reliable, being limited by a vast group of errors: false positives, positioning problems, measurement imprecisions (Grados et al., 2009; Weber et al., 1999). They also consider vertebral body heights in relation to contiguous vertebrae – making these methods unsuitable to evaluate compression fractures in isolated vertebrae.

The semi-quantitative method by Genant et al. (1993) is easy to apply, effective in ruling out vertebral compression fractures due to causes other than low bone mass, and highly reproducible. It is recommended by the «International Society for Clinical Densitometry» to diagnose vertebral fractures in the clinical setting (Schousboe et al., 2008). This study indicates that it is also an appropriate standard scoring method for vertebral compression fractures/deformities in paleopathological investigations: it is a practical, accessible and relatively fast technique, it may be implemented upon a complete vertebral column or an isolated vertebrae and mitigates many of the differences between observations and observers.

ACKNOWLEDGMENTS

Fundação para a Ciência e Tecnologia (grants #SFRH/BPD/74015/2010 and #PTDC/CSANT/120173/2010 – Paleoepidemiology of osteoporosis and osteoporotic fractures in Portugal since the Mesolithic: a transdisciplinary study).

REFERENCES

Barnett E, Nordin EC. 1960. The radiological diagnosis of osteoporosis: a new approach. *Clinical Radiology* **11**: 166-174.

Black D, Palermo L, Nevitt M, Genant H, Christensen L, Cummings S. 1999. Defining incident vertebral deformity: a prospective comparison of several approaches. *Journal of Bone and Mineral Research* **14**: 90-101.

Carletta J. 1996. Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics* **22**: 249-254.

Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**: 37-46.

Curate F. 2011. *O perímetro do declínio: osteoporose e fracturas de fragilidade em três amostras osteológicas identificadas portuguesas*. PhD Thesis, University of Coimbra: Coimbra.

Curate F, Albuquerque A, Correia J, Ferreira I, Pedroso de Lima J, Cunha E. 2013. A glimpse from the past: osteoporosis and osteoporotic fractures in a Portuguese identified skeletal sample. *Acta Reumatológica Portuguesa* **38**: 20-27.

Curate F, Piombino-Mascali D, Tavares A, Cunha E. 2009. Assottigliamento corticale del femore e fratture da fragilità ossea: uno studio della Collezione Scheletrica Identificata di Coimbra (Portogallo). *Archivio per l'Antropologia e la Etnologia* **139**: 129-146.

Delmas P, Genant H, Crans G, Stock J, Wong M, Siris E, Adachi J. 2003. Severity of prevalent vertebral fractures and the risk of subsequent vertebral and nonvertebral fractures: Results from the MORE trial. *Bone* **33**: 522-532.

Domett K, Tayles N. 2006. Adult fracture patterns in prehistoric Thailand: A biocultural interpretation. *International Journal of Osteoarchaeology* **16**: 185-199.

El Maghraoui A, Morjane F, Nouijai A, Achemlal L, Bezza A, Ghozlani I. 2009. Vertebral fracture assessment in Moroccan women: prevalence and risk factors. *Maturitas* **62**: 171-175.

Ferrar L, Jiang G, Adams J, Eastell R. 2005. Identification of vertebral fractures: an update. *Osteoporosis International* **16**: 717-728.

Foldes A, Moscovici A, Popovtzer M, Mogle P, Urman D, Zias J. 1995. Extreme osteoporosis in a Sixth Century skeleton from Negev Desert. *International Journal of Osteoarchaeology* **5**: 157-162.

Genant H, Wu C, Vankuijk C, Nevitt M. 1993. Vertebral fracture assessment using a semi-quantitative technique. *Journal of Bone and Mineral Research* **8**: 1137-1148.

Gonzalez-Reimers E, Mas-Pascual MA, Arnay-de-la-Rosa M, Velasco-Vázquez J, Santolaria-Fernández F, Machado-Calvo M. 2004. Noninvasive estimation of bone mass in ancient vertebrae. *American Journal of Physical Anthropology* **125**: 121-131.

Grados F, Fechtenbaum J, Flipon E, Kolta S, Roux C, Fardellone P. 2009. Radiographic methods for evaluating osteoporotic vertebral fractures. *Joint Bone Spine* **76**: 241-247.

Hirata K, Morimoto I. 1994. Vertebral osteoporosis in Late Edo Japanese. *Anthropological Science* **102**: 345-361.

Ives R. 2007. *An investigation of vitamin D deficiency, osteomalacia and age-related osteoporosis in six post-medieval urban collections*. PhD Thesis. University of Birmingham: Birmingham.

Jensen GF, McNair P, Boesen J, Hegedüs V. 1984. Validity in diagnosing osteoporosis. Observer variation in interpreting spinal radiographs. *European Journal of Radiology* **4**: 1-3.

Johnell O, Kanis J. 2006. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporosis International* **17**: 1726-1733.

Landis J, Koch G. 1977. The measurement of observer agreement for categorical data. *Biometrics* **33**: 159-174.

Li J, Wu CY, Jergas H, Genant HK. 1995. Diagnosing prevalent vertebral fractures: a comparison between quantitative morphometry and a standardized visual (semiquantitative) approach. In Genant HK, Jergas M, van Kuijk C (eds.). *Vertebral fracture in osteoporosis*. Radiology Research and Education Foundation: San Francisco; 271-280.

Mays SA. 1996. Age-dependent bone loss in a medieval population. *International Journal of Osteoarchaeology* **6**: 144-154.

Mays SA. 2006. Age-related cortical bone loss in women from a 3rd–4th century AD population from England. *American Journal of Physical Anthropology* **129**: 518-528.

Mays SA, Turner-Walker G, Syversen U. 2006. Osteoporosis in a population from medieval Norway. *American Journal of Physical Anthropology* **131**: 343-351.

Mensforth R, Latimer B. 1989. Hamann-Todd Collection aging studies: osteoporosis fracture syndrome. *American Journal of Physical Anthropology* **80**: 461-479.

NIH Consensus Development Panel. 2001. Osteoporosis prevention, diagnosis, and therapy. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy. *The Journal of the American Medical Association* ***285***: 785-795.

Olmez N, Kaya T, Gunaydin R, Vidinli BD, Erdogan N, Memis A. 2005. Intra- and interobserver variability of Kleerekoper's method in vertebral fracture assessment. *Clinical Rheumatology* **24**: 215-218.

Ortner D. 2003. *Identification of pathological conditions in human skeletal remains*. Academic Press: San Diego.

Reis M, Silva C, Cunha E. 2003. Multiple traumas in a medieval male from Serpa (Portugal). In Martín M, Rodríguez F (eds.). *¿Dónde estamos? Pasado, presente y futuro de la paleopatologia*. Universidad Autónoma de Madrid y Asociación Española de Paleopatología: Madrid; 490-495.

Rocha M. 1995. Les collections ostéologiques humaines identifies du Musée Anthropologique de l'Université de Coimbra. *Antropologia Portuguesa* **13**: 7-38.

Rothwell P. 2000. Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. *Journal of Neurology* **247**: 825-834.

Sambrook PN, Browne CD, Eisman JA, Bourke SJ. 1988. A case of crush fracture osteoporosis from Late Roman Pella in Jordan. *OSSA* **13**: 167-171.

Schousboe JT, Vokes T, Broy SB, Ferrar L, McKiernan F, Roux C, Binkley N. 2008. Vertebral fracture assessment: the 2007 ISCD official positions. *Journal of Clinical Densitometry* **11**: 92-108.

Snow CE. 1948. *Indian Knoll, Site Oh2, Ohio County, Kentucky*. University of

Kentucky Publications in Anthropology and Archaeology: Lexington.

Sim J, Wright CC. 2005. The Kappa Statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* **85**: 257-268.

Strouhal E, Nemecková A, Kouba M. 2003. Paleopathology of Iufaa and other persons found beside his shaft tomb at Abusir (Egypt). *International Journal of Osteoarchaeology* **13**: 331-338.

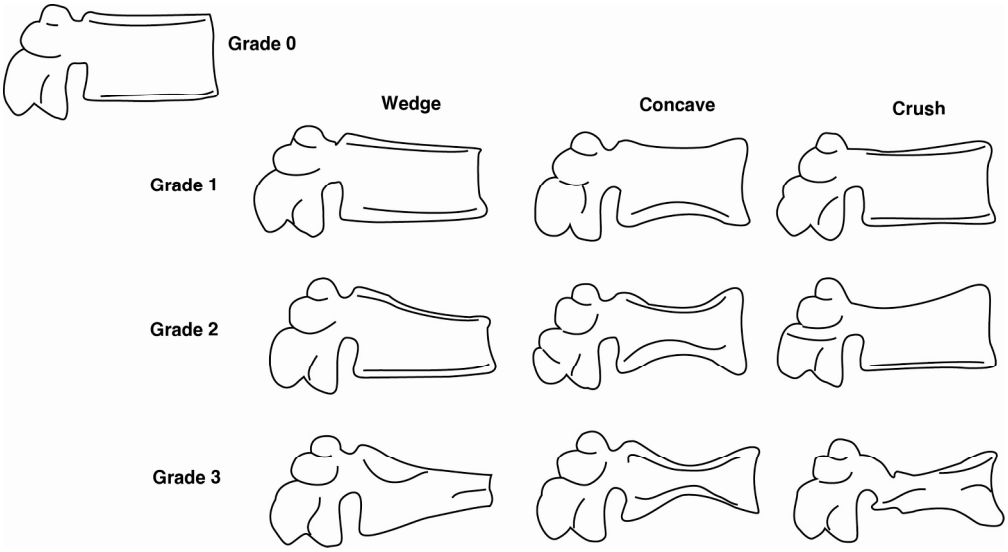
Watkins MW, Pacheco M. 2000. Interobserver agreement in behavioral research: importance and calculation. *Journal of Behavioral Education* **10**: 205-212.

Weber K, Lunt M, Gowin W, Lauermann T, Armbrecht G, Wieland E, Leb G, O'Neill T, Felsenberg D, Reeve J. 1999. Measurement imprecision in vertebral morphometry of spinal radiographs in the European Prospective Osteoporosis Study: consequences for the investigation of prevalent and incident deformities. *The British Journal of Radiology* **72**: 957-966.

Wylie CD. 2010. Setting a standard for a ''silent'' disease: defining osteoporosis in the 1980s and 1990s. *Studies in History and Philosophy of Biological and Biomedical Sciences* **41**: 376-385.

Table 1: Measures of agreement in the assessment of vertebral compression fractures with the Genant et al. (1993) method.

	N	%A (95% CI)	κ_c (95% CI)	Bias
Intra-observer (*per* individual)	196	97,3 (94,2 – 98,9)	0,899 (0,846 – 0,952)	0,005
Inter-observer (*per* individual)	75	94,7 (87,1 – 97,9)	0,688 (0,655 – 0,719)	0,053
Inter-observer (*per* vertebra)	975	99,5 (98,8 – 99,8)	0,703 (0,695 – 0,711)	0,003



Genant's semi-quantitative classification of vertebral compression fractures and/or deformations (adapted from Genant et al., 1993).
199x107mm (300 x 300 DPI)

Grade 1 wedge fracture/deformation, T12, male, 56 years (CISC).
69x49mm (300 x 300 DPI)