Classification Techniques for Osteoporosis Detection from Analyzing Texture Features on Clinical Brain MRI data

Elena Rantou[1*], Anuraag Ravikumar[2] and Vasiliki N. Ikonomidou[2]

[1]U.S. Food and Drug Administration, Center of Drug Evaluation and Research

[2]Department of Bioengineering, Volgenau School of Engineering, George Mason University, Fairfax, VA

[*] This work does not represent the official position of U.S. Food and Drug Administration

Abstract

Osteoporosis is a diffuse skeletal disease that is characterized by bone mass reduction and changes in the microarchitecture of the bone which eventually result to fractures, pain and disability. Magnetic Resonance Imaging (MRI) techniques, can offer insight into the fat content of the marrow and pore structure, which has been associated with osteoporosis. While these techniques are promising, the cost of MRI exam makes them an unsustainable choice for screening. We address the question of whether regular clinical brain MRI exams, can be used to identify a population at risk of osteoporosis, and allow the physician to refer them for further screening. The data set includes an osteoporosis diagnosed and a control group. Important features (texture analysis characteristics) are identified by using robust randomization tests for the difference of two means. The ability of these features for detecting osteoporosis is investigated by using different classifiers. The use of different statistical criteria provides the means of selecting the best classifier according to its performance.

Key Words: Osteoporosis, classification, texture features

1. Introduction

Osteoporosis is a diffuse skeletal disease that is characterized by bone mass reduction and changes in the microarchitecture of the bone. It results in a loss of the mechanical strength of the bone and is clinically manifested by bone fractures due to reduced bone density, mostly in the hip and in the spine. It leads to 2 million fractures per year in the United States alone, with an estimated cost of $19 billion annually.

Currently, screening is mainly performed by bone mineral density measurements using dual energy X-ray absorptiometry (DXA) and quantitative ultrasound, which has been proposed as a more cost-effective and radiation-free surrogate to DXA.

Magnetic resonance imaging (MRI) techniques, can offer insight into the fat content of the marrow and pore structure, which has been associated with osteoporosis. In particular, T1-weighted MRI images have been shown to be able to detect osteoporosis, as they have the potential of detecting alterations in the structure of the trabecular bone.

2. Motivation

While all these techniques show the promise of MRI for osteoporosis diagnosis and monitoring, the cost of the MRI exam makes them an unsustainable choice for screening. However, MRI is prescribed routinely as a diagnostic procedure for a variety of neurological and orthopedic conditions, both of which affect patients in the age range of interest for osteoporosis screening. Based on this, we addressed the question of whether regular clinical brain MRI exams, already obtained as part of the normal clinical care for a concomitant, but unrelated condition, can be used to identify a population at risk for osteoporosis, and allow the physician to refer them for further testing using the established osteoporosis diagnostic techniques.

3. Data

Data used for this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). ADNI's medical history data set consists of 32 female subjects of whom 16 are patients (reported osteoporosis diagnosis-MEAN (SD) age: 74.6 (6.1)) and 16 are controls (did not report osteoporosis diagnosis- MEAN (SD) age: 78.2 (7.3)).

4. Imaging

MPRAGE (magnetization prepared rapid gradient echo sequence) images considered here are one of the most popular sequences for brain imaging in clinical and research studies. Such sequences capture high spatial resolution with whole brain coverage in short time. In an attempt to observe texture changes in the superior part of the cerebral spine, MPRAGE images will be used to detect alterations in the structure of the trabecular bone. The marrow, which fills the pores of the matrix, is a heterogeneous fluid-like substance, which is the origin of the MRI signal.

5. Texture analysis

Texture is an innate property of all surfaces and contains important information about its structural arrangement. Since the textural properties of images carry useful info for discrimination purposes, it is important to develop features for texture. Prior work has shown the usefulness of such textural features for categorizing images. Texture analysis refers to the statistical analysis of voxel intensities within a defined neighborhood, indicating patterns of small-scale variations in image intensity in a particular region.

Prior research (Herlidou et al.), detected texture analysis based changes in images of the calcaneus bone. However, a dedicated image of the calcaneus is a financially non-viable option for osteoporosis diagnosis. Brain MRIs are routinely prescribed for a variety of conditions, especially in middle-aged and elderly adults.

Texture features

All participants had a 3.0T brain MRI, including a MPRAGE sequence. The texture analysis characteristics were calculated by Haralick et al (17). Examples of these texture features are:

Angular Second Moment $$f_1 = \sum_{i}^{N_g} \sum_{j}^{N_g} \{p(i,j)\}^2$$

Contrast $$f_2 = \sum_{n=0}^{N_g} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}$$

Difference entropy $$f_{10} = -\sum_{i=0}^{N_g-1} p_{x-y}(i)\log\{p_{x-y}(i)\}$$

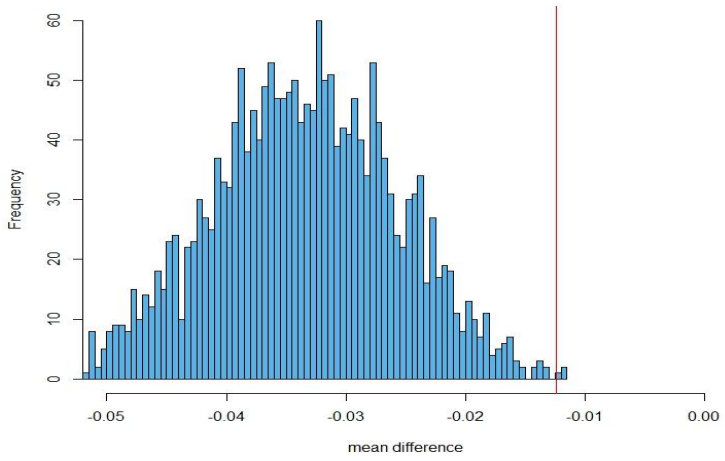
Where p(i, j) is the (i, j)th entry in a normalized gray-tone spatial-dependence matrix, p_x(i)=-$\sum_{j=1}^{N_g} p(i,j)$, N_g is the number of gray levels in the image, $p_{x+y}(k) = \sum_{i}^{N_g}\sum_{\substack{j \\ i+j=k}}^{N_g} p(i,j)$.

Average values of all texture features were calculated over the marked ROIs. Subsequently, the values were analyzed for group differences.

6. Methods

Due to the small sample size, sampling process and lack of prior knowledge as to the distribution, group differences were tested using two sample randomization tests, which are non-parametric. Under the assumption of indifference, all possible permutations of the observations are equally likely (18). For each variable, 5000 randomizations were used; the difference between the means of the two groups was calculated and the distribution of the differences was created. A p-value from this distribution is the probability that we would observe a result as extreme as the actual mean difference if there is no difference between the two groups. Therefore, p-values were calculated by determining how rare the observed mean difference is compared to the generated randomization distribution; after Bonferroni correction, alpha value was set to 0.0014. Such a randomization distribution, also indicating the p-value cutoff point (red vertical line) is shown in Figure 1.

Figure 1: Randomization test p-value



This process indicated 10 significant texture features shown in Table 1 below.

Table 1: Variable significance (feature selection)

Orientation for gray level co-occurrence matrix calculation

Texture Feature	0°	45°	90°	135°
ASM	0.016494416	0.015211355	0.014027766	0.010521874
Contrast	0.896022491	0.954948741	0.889214005	0.826708315
Correlation	0.42368053	0.42872054	0.39672375	0.41336468
Inv. Diff. Moment	**0.00115950**	**0.00114224**	0.00161861	**0.00078460**
Sum Average	0.26072260	0.26237919	0.76275894	0.76919203
Sum Entropy	0.00146901	**0.00131373**	**0.00132503**	0.00152634
Entropy	**0.00116655**	**0.00097130**	**0.00105667**	**0.00104268**
Diff. Variance	0.00347146	0.00247795	0.00218208	0.00184461
Diff. Entropy	0.00180069	0.00168697	0.00162634	**0.00125261**

Table 1: *p*-values from randomization tests (5000 repetitions) between participants with and without osteoporosis. Bonferroni-corrected statistically significant results in bold.

In order to estimate the potential of the proposed texture analysis methodology for osteoporosis detection, three classifiers were built. Those were linear discriminant analysis (LDA), Support Vector Machines (SVM) and a tree classification model. These classifiers were trained on the available data, and tested using a leave-two-out strategy, i.e. for every trial, the classifier was trained based on data from all but two participants, one patient and one control. This leave-one-out cross-validation approach provided an estimate for the test error rate. Excluded participants changed for every trial.

To study classifier performance we first consider the classifier as a mapping of instances to predicted images (21). For a classifier and an instance, there are four possible outcomes that are summarized in a confusion table, which includes true positives (*TP*), true negatives (*TN*), false positives (*FP*) and false negatives (*FN*). Out of those, one can define several performance metrics for a classifier. In this paper we have used the following metrics:

- Accuracy $AC = \frac{TP+TN}{P+N}$, where P corresponds to the total number of positives and N to the total number of negatives
- Precision $Pr = \frac{TP}{TP+FP}$
- Sensitivity $S_n = \frac{TP}{P}$
- Specificity $S_p = \frac{TN}{FN+TN}$ and
- Matthews Correlation Coefficient $MC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}}$

The MC (22) is mainly used in machine learning and expresses the type of association between the observed and the classified values. Its values range from -1 to 1, with a value of 1 corresponding to perfect agreement between the observed and the classified values. Similarly, a value of zero is equivalent to a classifier that randomly guesses and a value of -1 reveals total disagreement between prediction and observation.

7. Results

Table 1 shows the *p*-values resulting from the group comparison between patients and controls. Even after correction for multiple comparisons, several different texture features showed significant group differences or strong trends, indicating that the osteoporosis pathogenesis may affect them.

Table 2 compares the performance of the three classifiers, using the metrics discussed in the previous section. For the calculations in this table, a threshold value equal to 0.5 is used to identify true positive values (correctly classified patients). LDA outperforms the other two classifiers based on specificity and sensitivity. It is also the classifier which yields the smallest test error rate as shown in Table 3.

Table 2: Performance of classifiers

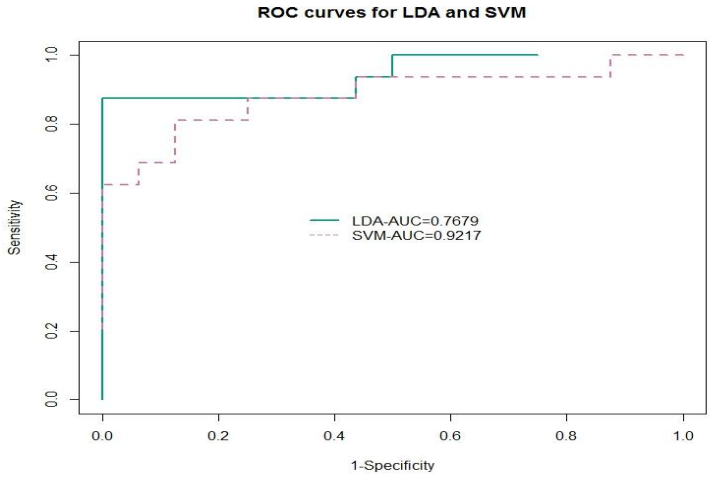
	LDA	SVM	Tree Model
Accuracy	0.8438	0.7813	0.8750
Precision	0.7895	0.8462	0.8750
Sensitivity	0.9375	0.6875	0.8750
Specificity	0.9231	0.7368	0.8750
MCC	0.7000	0.5727	0.7500

Table 3: Estimating the test error rate

Classifier	LOOCV
LDA	6.4375
SVM	7.2500
Tree Model	7.0000

Classification performance can be visually evaluated by a Receiver Operator Characteristic (ROC) curve shown in Figure 2. It also indicates high performance of the LDA classifier compared to SVM. To justify this statement in a quantitative way, the Area Under the Curve (AUC), value, is also calculated. Its value indicates the probability that the particular classifier will rank a positive case higher than a negative case.

Figure 2: ROC curves for LDA and SVM



8. Conclusions and discussion

In this study, we have shown that texture analysis derived features, can be used in order to distinguish patients with osteoporosis from an age-matched group of individuals without an osteoporosis diagnosis.

There are two possible mechanisms by which osteoporosis pathogenesis can influence texture on T_1-weighted MRI. The first one is changes in bone microarchitecture. Osteoporosis is related to both pore enlargement and changes in the trabeculae, with loss of horizontal trabeculae and thickening of the vertical ones.(5) Such changes in trabecular bone structure have been previously associated with aging and osteoporosis development,(26–28) and are known to relate to bone strength(29) and therefore fracture probability.

Furthermore, osteoporosis is known to change bone marrow composition. Marrow fat content is increased in both male (30) and female (31) patients with osteoporosis. This change in fat content will decrease the average T_1 value of the marrow, leading to changes in the appearance in the T_1-weighted MPRAGE images used in the study.

There are several shortcomings to the study that need to be addressed before any definite conclusions can be reached. The main drawback stems from the multicenter character of the ADNI study, which results in slightly different acquisition protocols as well as different hardware at different sites introducing extra variation.

Furthermore, there is a degree of uncertainty related to the health status, with respect to osteoporosis, of the study participants. The patient group was selected based on the mention of osteoporosis diagnosis during the medical screening interview. However, no information is available as to the severity of the disease or the presence of micro-fractures that may have interfered with the results of the study. Similarly, the fact that controls for the study did not report an osteoporosis diagnosis at the time of their enrollment in the ADNI study does not mean that they do not have osteoporosis. This is a shortcoming of a retrospective analysis based on limited data that can only be overcome with a prospective study including bone density measurements.

With the limitations taken into account, the study results support the initial hypothesis that osteoporosis-induced changes in the bone affect texture analysis based features in T_1 weighted MRI, and that they can be detected from T_1 weighted MPRAGE images of the brain, which include the superior part of the cerebral spine. These results confirm those by Herlidou et al (14) that detected texture analysis based

changes in T_1-weighted images of the calcaneus. However, whereas a dedicated MRI of the calcaneus is a financially non-viable option for osteoporosis diagnosis, brain MRIs are routinely prescribed for a variety of conditions. This is especially true in middle-aged and elderly adults, which are in increased risk for osteoporosis development, making it a financially reasonable initial screening test for this population if the MRI is already available. Further studies may be warranted to establish a relationship between bone density and textural characteristics of the bone on MPRAGE images.

Acknowledgements

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu).

References

1. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 2007;22:465–475. doi: 10.1359/jbmr.061113.

2. Looker AC, Borrud LG, Dawson-Hughes B, Shepherd JA, Wright NC. Osteoporosis or low bone mass at the femur neck or lumbar spine in older adults: United States, 2005-2008. NCHS Data Brief 2012:1–8.

3. Melton LJ 3rd, Chrischilles EA, Cooper C, Lane AW, Riggs BL. Perspective. How many women have osteoporosis? J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 1992;7:1005–1010. doi: 10.1002/jbmr.5650070902.

4. Cheng H, Gary LC, Curtis JR, Saag KG, Kilgore ML, Morrisey MA, Matthews R, Smith W, Yun H, Delzell E. Estimated prevalence and patterns of presumed osteoporosis among older Americans based on Medicare data. Osteoporos. Int. J. Establ. Result Coop. Eur. Found. Osteoporos. Natl. Osteoporos. Found. USA 2009;20:1507–1515. doi: 10.1007/s00198-009-0835-z.

5. Kumar V, Abbas AK, Aster JC, Fausto N. Robbins & Cotran Pathologic Basis of Disease: With STUDENT CONSULT Online Access, 8e. 8th ed. Saunders; 2009.

6. U.S. Preventive Services Task Force. Screening for osteoporosis: U.S. preventive services task force recommendation statement. Ann. Intern. Med. 2011;154:356–364. doi: 10.1059/0003-4819-154-5-201103010-00307.

7. Drake MT, Khosla S. Male osteoporosis. Endocrinol. Metab. Clin. North Am. 2012;41:629–641. doi: 10.1016/j.ecl.2012.05.001.

8. Tay W-L, Chui C-K, Ong S-H, Ng AC-M. Osteoporosis screening using areal bone mineral density estimation from diagnostic CT images. Acad. Radiol. 2012;19:1273–1282. doi: 10.1016/j.acra.2012.05.017.

9. Moris M, Peretz A, Tjeka R, Negaban N, Wouters M, Bergmann P. Quantitative ultrasound bone measurements: Normal values and comparison with bone mineral density by dual X-ray absorptiometry. Calcif. Tissue Int. 1995;57:6–10. doi: 10.1007/BF00298988.

10. Barkmann R. Quantitative Ultrasound. In: Grampp S, editor. Radiology of Osteoporosis. Medical Radiology. Springer Berlin Heidelberg; 2008. pp. 163–173.

11. Shen W, Gong X, Weiss J, Jin Y. Comparison among T1-Weighted Magnetic Resonance Imaging, Modified Dixon Method, and Magnetic Resonance Spectroscopy in Measuring Bone Marrow Fat. J. Obes. 2013;2013:298675. doi: 10.1155/2013/298675.

12. Manenti G, Capuani S, Fanucci E, Assako EP, Masala S, Sorge R, Iundusi R, Tarantino U, Simonetti G. Diffusion tensor imaging and magnetic resonance spectroscopy assessment of cancellous bone quality in femoral neck of healthy, osteopenic and osteoporotic subjects at 3T: Preliminary experience. Bone 2013;55:7–15. doi: 10.1016/j.bone.2013.03.004.

13. Li X, Kuo D, Schafer AL, Porzig A, Link TM, Black D, Schwartz AV. Quantification of vertebral bone marrow fat content using 3 Tesla MR spectroscopy: reproducibility, vertebral variation, and applications in osteoporosis. J. Magn. Reson. Imaging JMRI 2011;33:974–979. doi: 10.1002/jmri.22489.

14. Herlidou S, Grebe R, Grados F, Leuyer N, Fardellone P, Meyer M-E. Influence of age and osteoporosis on calcaneus trabecular bone structure: a preliminary in vivo MRI study by quantitative texture analysis. Magn. Reson. Imaging 2004;22:237–243. doi: 10.1016/j.mri.2003.07.007.

15. Tameem HZ, Selva LE, Sinha US. Texture Measure from Low Resolution MR Images to Determine Trabecular Bone Integrity in Osteoporosis. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007. EMBS 2007. ; 2007. pp. 2027–2030. doi: 10.1109/IEMBS.2007.4352717.

16. McAuliffe MJ, Lalonde FM, McGarry D, Gandler W, Csaky K, Trus BL. Medical Image Processing, Analysis and Visualization in clinical research. In: 14th IEEE Symposium on Computer-Based Medical Systems, 2001. CBMS 2001. Proceedings. ; pp. 381–386. doi: 10.1109/CBMS.2001.941749.

17. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. Syst. Man Cybern. IEEE Trans. On 1973;SMC-3:610–621.

18. Manly BFJ. Randomization, bootstrap and Monte Carlo methods in biology. Boca Raton, FL: Chapman & Hall/ CRC; 2007.

19. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilita. 1936.

20. Miller RG. Simultaneous statistical inference. New York: Springer-Verlag; 1981.

21. Fawcett T. An introduction to ROC analysis. Pattern Recognit. Lett. 2006;27:861–874. doi: 10.1016/j.patrec.2005.10.010.

22. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta BBA - Protein Struct. 1975;405:442–451. doi: 10.1016/0005-2795(75)90109-9.

23. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997;30:1145–1159. doi: 10.1016/S0031-3203(96)00142-2.

24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.

25. Amstutz HC, Sissons HA. The structure of the vertebral spongiosa. J Bone Jt. Surg Br 1969;51:540–550.

26. Gabet Y, Bab I. Microarchitectural changes in the aging skeleton. Curr. Osteoporos. Rep. 2011;9:177–183. doi: 10.1007/s11914-011-0072-1.

27. Milovanovic P, Djonic D, Marshall RP, Hahn M, Nikolic S, Zivkovic V, Amling M, Djuric M. Micro-structural basis for particular vulnerability of the superolateral neck trabecular bone in the postmenopausal women with hip fractures. Bone 2012;50:63–68. doi: 10.1016/j.bone.2011.09.044.

28. Milovanovic P, Potocnik J, Djonic D, Nikolic S, Zivkovic V, Djuric M, Rakocevic Z. Age-related deterioration in trabecular bone mechanical properties at material level: nanoindentation study of the femoral neck in women by using AFM. Exp. Gerontol. 2012;47:154–159. doi: 10.1016/j.exger.2011.11.011.

29. Issever AS, Link TM, Kentenich M, Rogalla P, Schwieger K, Huber MB, Burghardt AJ, Majumdar S, Diederichs G. Trabecular Bone Structure Analysis in the Osteoporotic Spine Using a Clinical In Vivo Setup for 64-Slice MDCT Imaging: Comparison to μCT Imaging and μFE Modeling. J. Bone Miner. Res. 2009;24:1628–1637. doi: 10.1359/jbmr.090311.

30. Griffith JF, Yeung DKW, Antonio GE, Lee FKH, Hong AWL, Wong SYS, Lau EMC, Leung PC. Vertebral bone mineral density, marrow perfusion, and fat content in healthy men and men with osteoporosis: dynamic contrast-enhanced MR imaging and MR spectroscopy. Radiology 2005;236:945–951. doi: 10.1148/radiol.2363041425.

31. Griffith JF, Yeung DKW, Antonio GE, Wong SYS, Kwok TCY, Woo J, Leung PC. Vertebral marrow fat content and diffusion and perfusion indexes in women with varying bone density: MR evaluation. Radiology 2006;241:831–838. doi: 10.1148/radiol.2413051858.

32. Amling M, Pösl M, Wening VJ, Ritzel H, Hahn M, Delling G. Structural heterogeneity within the axis: the main cause in the etiology of dens fractures. A histomorphometric analysis of 37 normal and osteoporotic autopsy cases. J. Neurosurg. 1995;83:330–335. doi: 10.3171/jns.1995.83.2.0330.

33. Ea H-K, Weber A-J, Yon F, Lioté F. Osteoporotic fracture of the dens revealed by cervical manipulation. Jt. Bone Spine Rev. Rhum. 2004;71:246–250. doi: 10.1016/S1297-319X(03)00113-1.

34. Platzer P, Thalhammer G, Oberleitner G, Schuster R, Vécsei V, Gaebler C. Surgical treatment of dens fractures in elderly patients. J. Bone Joint Surg. Am. 2007;89:1716–1722. doi: 10.2106/JBJS.F.00968.