Cortical Marrow Ratio in Plain X-rays of Femoral Neck Fractures

Viberg, Bjarke; Severin Gråe Harbo, Frederik; Ryg, Jesper; Overgaard, Soren; Ovesen, Ole; Lauritsen, Jens M

Published in:
Jacobs Journal of Orthopedics and Rheumatology

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Citation for pulished version (APA):
Viberg, B., Severin Gråe Harbo, F., Ryg, J., Overgaard, S., Ovesen, O., & Lauritsen, J. M. (2016). Cortical Marrow Ratio in Plain X-rays of Femoral Neck Fractures: Reliability and Relation to BMD. Jacobs Journal of Orthopedics and Rheumatology, 2(1), [019].

JACOBS PUBLISHERS

OPEN ACCESS

Jacobs Journal of Orthopedics and Rheumatology

Research Article

Cortical Marrow Ratio in Plain X-rays of Femoral Neck Fractures - Reliability and Relation to BMD

Bjarke Viberg[1,4*], Frederik SG Harbo[2], Jesper RYG[3,4], Søren Overgaard[1,4], Ole Ovesen[1,4], Jens Lauritsen[1,5]

[1]*Department of Orthopaedic Surgery and Traumatology, Odense University Hospital*

[2]*Department of Diagnostic Radiology, Odense University Hospital*

[3]*Department of Geriatric Medicine, Odense University Hospital*

[4]*Institute of Clinical Research, University of Southern Denmark*

[5]*Institute of Public Health, Department of Biostatistics, University of Southern Denmark*

**Corresponding author: Dr. Bjarke Viberg, Sdr. Boulevard 29, DK-5000 Odense C, Tel: +45 28669059; Fax: +45 65913039; Email: Bjarke.Viberg@rsyd.dk*

Received: 04-26-2016

Accepted: 11-09-2016

Published: 01-26-2016

Abstract

Objectives: The purpose of this study is to evaluate Cortical Marrow Ratio (CMR) for reliability and relation to low Bone Mineral Density (BMD) in patients with femoral neck fractures.

Methods: A total of 132 consecutive femoral neck fracture patients (median age 81.2 years, IQR 70.6-86.1) were assessed with DXA scans and digital hip x-rays. CMR was measured twice by two independent raters and analysed for reliability. CMR was then compared to BMD by means of a sensitivity/specificity analysis.

Results: Using total hip BMD, 47 patients had a T-score ≤ -2.5 with a median CMR of 1.61(IQR 1.44-1.74), and 85 patients had a T-score > -2.5 with a median CMR of 1.89 (IQR 1.77-2.11). The ICC was 0.87-0.98 for intra-rater and 0.86-0.90 for inter-rater reliability. CMR showed a correlation coefficient of 0.58-0.59, a sensitivity of 72.3-76.6 % and a specificity of 75.3-76.5 % in relation to low BMD.

Conclusions: CMR was found to be a highly reliable measure with acceptable sensitivity and specificity in relation to low BMD.

Keywords: Femoral Neck Fracture; BMD, X-Ray; Reliability; Sensitivity; Specificity

Abbreviations

BMD: Bone Mineral Density;
CI: Confidence Inteval;
DXA: Dual-energy X-ray Absorptiometry;
IQR: Inter Quartile Range;
ICC: Intraclass Correlation Coefficient

Background

The failure proportion for surgical treated femoral neck fractures is approximately 35 % in dislocated fractures [1] and 11 % in undisplaced fractures [2]. Predictors for failure should therefore be investigated and made feasible to implement in clinical practise. One possible predictor for failure of osteosynthesis is osteoporosis. Several experimental studies have shown that low BMD affects the strength of osteosynthesis [3,4] and have proposed a limit for high failure risk at 0.4 g/cm^3, [4,5]. However, clinical studies have yet proven an association between low BMD and failure [6] but a clinical study showed a tendency towards an association between failure in undisplaced femoral neck fractures and low BMD [7]. DXA scan has been recognized as the gold standard for diagnosing osteoporosis [8]. However, in a clinical setting it can be logistically very difficult to obtain a DXA scan before surgery and it may cause a surgical delay which could increase the in-hospital mortality [9].

A preoperatively investigation of low BMD should be fast and feasible and could be done by using the existing x-ray image used for diagnosing the fracture. The Singh Index [10] is the oldest and best known geometric measure for osteoporosis, but the Singh Index' reliability is either poor[11,12] or acceptable [13-15]. The major drawback of most studies using the Singh index is, however, that they only show poor to moderate correlation with BMD [11,12,14,16-20]. Several other geometrical measures have been suggested [14,17,21-25]. Of these, the canal bone ratio, the cortical thickness index, and the Dorr classification have shown good reliability by having an ICC above 0.8, but only canal bone ratio had a correlation with BMD above 0.7.

The canal bone ratio study [23] is on cadavers and uses a fixed measurement point which does not account for the morphological differences of small and large femora. In order to take account for the variability of femoral geometry CMR was developed. The purpose of this study is to evaluate CMR for reliability and relation to low BMD in patients with femoral neck fractures.

Methods

Subjects

Patients and BMD measurements were retrieved from a prospective consecutive cohort of patients with hip fractures (26), which included all hip fracture patients who were older than 45 years and treated at the Department of Orthopaedic Surgery and Traumatology, Odense University Hospital. 158 femoral neck fracture patients with DXA-scans had their x-rays retrospectively assessed. One patient was excluded due to an old fracture, three patients due to transferrals and 22 patients were excluded due to the femoral portion for measurement was not included. This left a total of 132 patients with femoral neck fracture and DXA-scans, who comprised the final study cohort. All patients were treated with closed reduction and IF using two Uppsala screws. All patients were postoperatively treated with Calcium and Vitamin D. The median time from operation to the DXA-scan was 80 days (IQR 42-142). All x-ray images from the cohort were evaluated to ensure correct fracture diagnosis. Any discrepancies in diagnosis were discussed and resolved.

Measurements

Preoperative x-rays of the patients in the final cohort were used. Based on a pilot study of 20 patients, the CMR was assessed by the following method:

1. A circle was drawn just below the lesser trochanter containing the femoral diameter. A second circle was drawn approximately two femoral diameters below the first circle. A line was drawn based on the centers of the two circles depicting the midline (Figure 1a).

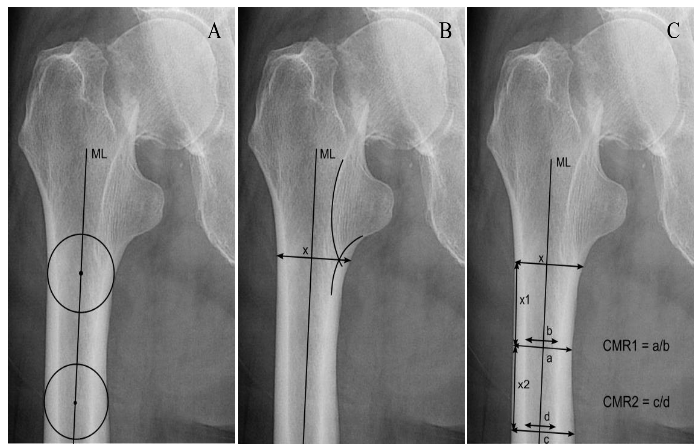

Figure 1. Illustrates the CMR measurement in steps.

A: depict two circles and draw a midline (ML). B: Measure the femoral diameter (x) perpendicular to ML below the lesser trochanter. C: Measure the outer (a,c) and inner cortex (b,d) at x*1 and x*2.

2. The circles were removed and a new line perpendicular to the midline was drawn placed at the crossing point between the dense trabecular structure of the lesser trochanter and the cortex below (Fig. 1b). The length of the line was then set to the width of the femur.

3. Two new lines were placed perpendicular to the midline at one and two x' distal (x = length of first line) to the first perpendicular line, and at these lines the femoral cortical and marrow diameter were measured (a-d in Fig. 1c).

CMR = femoral diameter / marrow diameter = a / b or c / d

Two independent observers analysed all x-rays twice (approximately 12 weeks apart) to obtain inter- and intra-rater reliability and agreement. Rater 1 was an orthopaedic resident and rater 2 was a radiology resident. For the purpose of inter-rater reliability and agreement, the first measurement of both raters was used. The BMD from the DXA-scans were compared to the first measurements of rater 1. The DXA-scanner was a Hologic Discovery and NHANES III was used as reference material [27]. The contralateral hip was scanned and low BMD level using total hip BMD was defined as a T-score ≤ -2.5 [28]. All x-rays were digital and measurements were carried out on a 21-inch screen or larger using Sectra AB's RIS/PACS x-ray viewing system. Edge detection and reinforcement was not used. The raters were blinded to each other's measurements and BMD results, which were merged with the CMR results after completion of the second CMR measurements.

Statistical analyses

The statistical software programme STATA 11 was used for the analyses. Data for group analyses was tested graphically and statistically (Shapiro Wilk) for distribution and was not normally distributed. Therefore the Wilcoxon rank-sum test (Mann-Whitney two-sample) was used for group analyses. Reliability is in this context defined by de Vet et al. [29] and based on the $ICC_{agreement}$ reliability parameter which was calculated with a multilevel mixed-effects linear regression technique [30,31]. For a graphical estimate of systematically bias and agreement, a software extension (SJ7-3: st0015_4) for STATA was downloaded to give the Bland-Altman plot. A sensitivity/specificity analysis (extension SJ-4-4: sbe36_2) was applied to evaluate CMR in relation to low BMD. ROC analysis was used to find optimal cut-off thresholds of sensitivity and specificity. The correlation coefficient (Spearman's rank) was calculated.

A retrospective power analysis for the reproducibility study using STATA' sampicc extension was done with a hypothesized ICC value of 0.9 and a null value of 0.8, which gives a power of 100 %. A post-hoc analysis on the sample size of the sensitivity/specificity analysis using precision 0.05, prevalence 0.36, and specificity = 0.76 gave a sample size of 438 [32]. The reporting of this article is done according to both GRASS and STARD guidelines [33,34].

Results

X-ray images were included for 132 patients (32 men and 100 women), median (IQR) 81.2 (70.6-86.1) years, and no age difference between the sexes (Wilcoxon rank-sum test: p<0.72) (Table I). A total of 47 patients (35.6 %) were found to have low BMD levels using total hip BMD with a median (IQR) BMD of 0.57 g/cm^2 (0.50-0.61).

Total hip BMD	T-score ≤ -2.5	T-score > -2.5
Number of patients	47	85
BMD, g/cm^2 (IQR)	0.57 (0.50-0.61)	0.74 (0.70-0.81)
CMR1 (IQR)	1.51 (1.35-1.63)	1.73 (1.62-1.42)
CMR2 (IQR)	1.61 (1.44-1.74)	1.89 (1.77-2.11)
Age (IQR)	82.8 (75.9-87.2)	80.0 (69.6-84.9)
Gender (male/female)	4/43	28/57

Table 1. Study population by total hip BMD (g/cm^2) status.

Reliability analysis

For CMR1 the median measurements were 1.81-1.84 and for CMR 2 they were 1.65-1.69 (Table II). The calculated reliability parameter ICC gave an $ICC_{agreement}$ of 0.87-0.98 for intra- and interrater of CMR1 and CMR2 (Table III).

	CMR1 (IQR)	CMR2 (IQR)
Rater 1, measurement 1	1.65 (1.51-1.80)	1.81 (1.58-2.01)
Rater 1, measurement 2	1.66 (1.51-1.80)	1.82 (1.60-2.00)
Rater 2, measurement 1	1.65 (1.53-1.78)	1.81 (1.58-1.98)
Rater 2, measurement 2	1.69 (1.55-1.80)	1.84 (1.63-2.01)

Table 2. Median CMR measurements.

	ICC CMR1 (CI)	ICC CMR2 (CI)
Intrarater for rater 1	0.98 (0.97;0.99)	0.98 (0.97;0.99)
Intrarater for rater 2	0.90 (0.86;0.93)	0.87 (0.83;0.91)
Interrater	0.90 (0.86;0.93)	0.86 (0.81;0.90)

Table 3. Intra- and interrater ICCagreement

No systematic difference (bias) between the raters were found, but there was an intra-rater difference for rater 2 in the CMR2 measurements of 0.05 (p<0.001, multilevel mixed-effects linear regression). A Bland-Altman plot was applied to the data, and there was a uniform distribution of the differences for the whole range of CMR1 and CMR2 values.

Sensitivity and specificity analysis

The ROC-analysis resulted in an optimal cut-off threshold value at 1.75 for CMR1 and 1.62 for CMR2. Table 4 shows the result yielding a sensitivity of 72.3-76.6 % and a specificity of 75.3-76.5 % for CMR1 and CMR2. The overall correlation coefficient was r=0.59 (p<0.0001) for CMR1 and r=0.58 (p<0.0001) for CMR2. Figure 2 shows a scatter plot of the CMR2 values in comparison to the T-score.

CMR	T-score≤ -2.5	T-score > -2.5	SN	SP	PPV	NPV
CMR1			72.3	75.3	61.8	83.1
Low BMD (CMR < 1.62)	34	13				
Normal BMD (CMR > 1.62)	21	64				
CMR2			76.6	76.5	64.3	85.5
Low BMD (CMR < 1.75)	36	11				
Normal BMD (CMR > 1.75)	20	65				

SN sensitivity, SP specificity, PPV positive predictive value, NPV negative predictive value.

Table 4. CMR in relation to total hip BMD.

SN sensitivity, SP specificity, PPV positive predictive value, NPV negative predictive value.

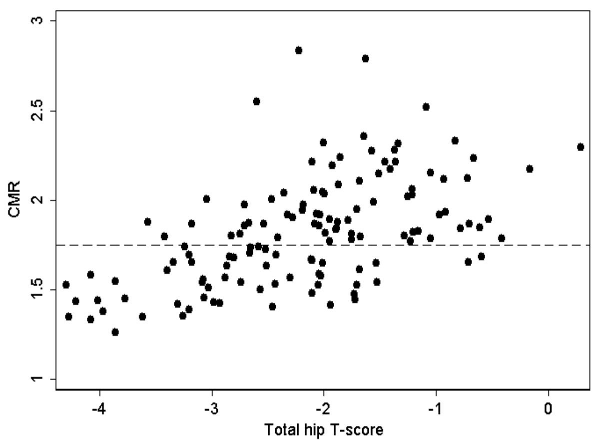

Figure 2. Scatter plot of CMR and total hip t-score.

Discussion

CMR has excellent intra- and inter-rater results with an ICC of 0.86-0.98. CMR have correlation coefficients of 0.58-0.59, an acceptable sensitivity percentage of 72.3-76.6 % and a specificity of 75.3-76.5 %.

Conceptually, the reliability expresses the ability to distinguish between patients, despite measurement errors. The reliability is therefore a characteristic of the performance of an instrument (CMR) and CMR is used for discriminative purposes [29]. ICCagreement takes the systematic difference between the measurements into account and is therefore used as the reliability parameter. This study showed excellent reliability with an inter- and intrarater ICC of 0.86-0.98 which only has been shown in one other paper [23]. It is not possible to use an agreement parameter such as the Bland-Altman plot because the BMD (range 0.4-1.1) and the CMR (1.3-2.8) results are not on the same scale. The correlation coefficient was r=0.58 and similar to other studies but should be interpreted with caution because high correlation does not imply close agreement since the correlation coefficient is blind to the possibility of bias [35]. Instead, a diagnostic precision analysis was applied for evaluation of classification of osteoporotic status.

Three studies [11,13,18] have analysed sensitivity and specificity of low BMD in relation to the Singh Index which gave results from 11 to 97 % but only sensitivity or specificity was high. Neither of them was on hip fracture patients. Only Bes et al. [13] got high sensitivity (71%) and specificity (93 %) but the analysis only included 5 patients. Sah et al. [14] showed acceptable results with sensitivity ranging from 62-85 % and specificity 58-84 % for the cortical thickness index and the Dorr classification. Even though the study has interesting results the study has a low number of participants (n=32) and the risk for a type 2 error is quite large. The present study have present study has similar results with sensitivity and specificity of 76 %.

The following limitations are noted: a major underpowering shown in the post-hoc analysis, and 14 % of the routine x-rays excluded due to insufficient femoral shaft length used for CMR measurement on the x-rays. There is a small systematic bias in the repeated measurement by rater 2, possibly due to the application of a different zooming level on the second measurement. In further studies specification of zooming should be specified for reading the digital x-rays.

n perspective low BMD should be investigated as a predictor for failure especially in the undisplaced femoral neck fracture. The DXA scan is the gold standard for diagnosing osteoporosis [8] but is not fast or feasible as a preoperative measure.

Therefore x-ray images could be an option for quick ascertainment of osteoporotic status in the immediate clinical setting before surgery. CMR has in contrast to other measurement a high reliability and acceptable sensitivity/specificity and should therefore with other BMD measurements be investigated in further studies.

Conclusions

CMR was found to be a reliable measure with intra- and interrater ICC between 0.86 and 0.98. CMR also have acceptable sensitivity of 72.3-76.6 % and specificity of 75.3-76.5 % in relation to low BMD.

Acknowledgement

The study was carried out on the department of Orthopaedic Surgery and Traumatology, Odense University Hospital.

References

1. Wang J, Jiang B, Marshall RJ. Arthroplasty or internal fixation for displaced femoral neck fractures: which is the optimal alternative for elderly patients? A meta-analysis. Int Orthop. 2009, 33: 1179-1187.

2. Gjertsen JE, Fevang JM, Matre K. Clinical outcome after undisplaced femoral neck fractures. Acta Orthop. 2011, 82: 268-274.

3. Bonnaire F, Zenker H, Lill C. Treatment strategies for proximal femur fractures in osteoporotic patients. Osteoporos Int. 2005, 16(Suppl 2): S93-S102.

4. Sjostedt A, Zetterberg C, Hansson T. Bone mineral content and fixation strength of femoral neck fractures. A cadaver study. Acta Orthop Scand. 1994, 65: 161-165.

5. Turner IG, Rice GN. Comparison of bone screw holding strength in healthy bovine and osteoporotic human cancellous bone. Clin Mater.1992, 9(2): 105-107.

6. Giannoudis P, Tzioupis C, Almalki T. Fracture healing in osteoporotic fractures: is it really different? A basic science perspective. Injury.2007, 38(Suppl 1): S90-S99.

7. Viberg B, Ryg J, Overgaard S. Low bone mineral density is not related to failure in femoral neck fracture patients treated with internal fixation. Acta Orthop. 2013, 85(1): 60-65.

8. Baim S, Leslie WD. Assessment of fracture risk. Curr Osteoporos Rep. 2012, 16(6): 581-589.

9. Moja L, Piatti A, Pecoraro V. Timing matters in hip fracture surgery: patients operated within 48 hours have better outcomes. A meta-analysis and meta-regression of over 190,000 patients. PLoS One. 2012, 7(10): e46175.

10. Singh M, Nagrath AR, Maini PS. Changes in trabecular pattern of the upper end of the femur as an index of osteoporosis. J Bone Joint Surg Am.1970, 52(3): 457-467.

11. Hauschild O, Ghanem N, Oberst M. Evaluation of Singh index for assessment of osteoporosis using digital radiography. Eur J Radiol. 2009, 71(1): 152-158.

12. Koot VC, Kesselaer SM, Clevers GJ. Evaluation of the Singh index for measuring osteoporosis. J Bone Joint Surg Br. 1996, 78: 831-834.

13. Bes C, Guven M, Akman B. Can bone quality be predicted accurately by Singh index in patients with rheumatoid arthritis? Clin Rheumatol. 2012, 31(1): 85-89.

14. Sah AP, Thornhill TS, Leboff MS. Correlation of plain radiographic indices of the hip with quantitative bone mineral density. Osteoporos Int. 2007, 18(8): 1119-1126.

15. Patel SH, Murphy KP. Fractures of the proximal femur: correlates of radiological evidence of osteoporosis. Skeletal Radiol. 2006, 35: 202-211.

16. Hubsch P, Kocanda H, Youssefzadeh S. Comparison of dual energy X-ray absorptiometry of the proximal femur with morphologic data. Acta Radiol. 1992, 33: 477-81.

17. Karlsson KM, Sernbo I, Obrant KJ. Femoral neck geometry and radiographic signs of osteoporosis as predictors of hip fracture. Bone. 1996, 18(4): 327-330.

18. Masud T, Jawed S, Doyle DV. A population study of the screening potential of assessment of trabecular pattern of the femoral neck (Singh index): the Chingford Study. Br J Radiol. 1995, 68(808): 389-393.

19. Smyth PP, Adams JE, Whitehouse RW. Application of computer texture analysis to the Singh Index. Br J Radiol. 1997, 70: 242-247.

20. Wachter NJ, Augat P, Mentzel M. Predictive value of bone mineral density and morphology determined by peripheral quantitative computed tomography for cancellous bone strength of the proximal femur. Bone. 2001, 28(1): 133-139.

21. El-Kaissi S, Pasco JA, Henry MJ. Femoral neck geometry and hip fracture risk: the Geelong osteoporosis study. Osteoporos Int. 2005, 16(10): 1299-1303.

22. Pulkkinen P, Partanen J, Jalovaara P. Combination of bone mineral density and upper femur geometry improves the prediction of hip fracture. Osteoporos Int. 2004, 15(4): 274-280.

23. Yeung Y, Chiu KY, Yau WP. Assessment of the proximal femoral morphology using plain radiograph-can it predict the bone quality? J Arthroplasty. 2006, 21(4): 508-513.

24. Dorr LD, Faugere MC, Mackel AM. Structural and cellular assessment of bone quality of proximal femur. Bone. 1993, 14(3): 231-242.

25. Gruen T. A simple assessment of bone quality prior to hip arthroplasty: cortical index revisited. Acta Orthop Belg.1997, 63(Suppl 1): 20-27.

26. Ryg J. The Frail Hip - A study on the risk of second hip fracture, prevalence of osteoporosis, and adherence to treatment in patients with recent hip fracture. Department of Endocrinology, Odense University Hospital. 2009, 186.

27. Looker AC, Wahner HW, Dunn WL. Updated data on proximal femur bone mineral levels of US adults. Osteoporos Int.1998, 8(5): 468-489.

28. Kanis JA, McCloskey EV, Johansson H. A reference standard for the description of osteoporosis. Bone. 2008, 42(3): 467-475.

29. de Vet HC, Terwee CB, Knol DL. When to use agreement versus reliability measures. J Clin Epidemiol. 2006, 59(10): 1033-1039.

30. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res.1999, 8(2): 135-160.

31. Carstensen B, Simpson J, Gurrin LC. Statistical models for assessing agreement in method comparison studies with replicate measurements. Int J Biostat. 2008, 4(1): Article 16.

32. Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med. 1996, 3(9): 895-900.

33. Bossuyt PM, Reitsma JB, Bruns DE. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med. 2003, 138(1): W1-W12.

34. Kottner J, Audige L, Brorson S. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011, 64(1): 96-106.

35. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol. 2003, 22(1): 85-93.