Hounsfield units measured in low dose CT reliably assess vertebral trabecular bone density changes over two years in axial spondyloarthritis

Marques, M.L.; Silva, N.P. da; Heijde, D. van der; Reijnierse, M.; Baraliakos, X.; Braun, J.; ... ; Ramiro, S.

Seminars in Arthritis and Rheumatism 58 (2023) 152144

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Hounsfield Units measured in low dose CT reliably assess vertebral trabecular bone density changes over two years in axial spondyloarthritis

Mary Lucy Marques [a,b,*], Nuno Pereira da Silva [c], Désirée van der Heijde [a], Monique Reijnierse [d], Xenofon Baraliakos [e], Juergen Braun [e], Floris van Gaalen [a], Sofia Ramiro [a,f]

[a] Department of Rheumatology, Leiden University Medical Center, Leiden, the Netherlands
[b] Department of Rheumatology, Coimbra University Hospital, Coimbra, Portugal
[c] Department of Radiology, Coimbra University Hospital, Coimbra, Portugal
[d] Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands
[e] Rheumazentrum Ruhrgebiet Herne, Ruhr-University Bochum, Germany
[f] Department of Rheumatology, Zuyderland Medical Center, Herleen, the Netherlands

ARTICLE INFO

Keywords:
Axial spondyloarthritis
Hounsfield units
Low dose CT
Reliability

ABSTRACT

Objectives: To describe low dose Computed Tomography (ldCT) Hounsfield Units (HU) two-year change-from-baseline values (expressing trabecular bone density changes) and analyse their inter-reader reliability per vertebra in radiographic axial spondyloarthritis (r-axSpA).

Methods: We used 49 patients with r-axSpA from the multicentre two-year Sensitive Imaging in Ankylosing Spondylitis (SIAS) study. LdCT HU were independently measured by two trained readers at baseline and two years. Mean (standard deviation, SD) for the change-from-baseline HU values were provided per vertebra by reader. Intraclass correlation coefficients (ICC; absolute agreement, two-way random effect), Bland-Altman plots and smallest detectable change (SDC) were obtained. Percentages of vertebrae in which readers agreed on the direction of change and on change >|SDC| were computed.

Results: Overall, 1,053 (98% of all possible) vertebrae were assessed by each reader both at baseline and two years. Over two years, HU mean change values varied from -23 to 28 and 29 for reader 1 and 2, respectively. Inter-reader reliability of the two-year change-from-baseline values per vertebra was excellent: ICC:0.91-0.99; SDC:6-10; Bland-Altman plots were homoscedastic, with negligible systematic error between readers. Readers agreed on the direction of change in 88-96% and on change >|SDC| in 58-94% of vertebrae, per vertebral level, from C3 to L5. Overall, similar results were obtained across all vertebrae.

Conclusion: LdCT measurement of HU is a reliable method to assess two-year changes in trabecular bone density at each vertebra from C3-L5. Being reliable across all vertebrae, this methodology can aid the study of trabecular bone density changes over time in r-axSpA, a disease affecting the whole spine.

Introduction

Radiographic axial spondyloarthritis (r-axSpA) is a chronic inflammatory disease of the sacroiliac joints and spine, characterized by inflammation and new bone formation both substantially contributing to the burden of disease [1,2]. New bone formation can lead to spinal or sacroiliac joint ankylosis, reduced mobility, and increased disability [2]. Paradoxically, bone formation coexists with bone loss, both contributing to the morbidity of the disease [3].

Osteoporosis with increased fracture risk, has been widely reported as a comorbidity in patients with r-axSpA [2,4,5]. Particularly, a higher prevalence of vertebral fractures (but not hip fractures) has been described in patients with r-axSpA when compared to healthy individuals [4,6]. Disease-specific aspects, such as local inflammation and reduced mobility, add to the traditional risk factors associated with systemic bone loss [6].

In r-axSpA, trabecular vertebral bone loss is hypothesized as a pathological phenomenon occurring locally throughout the whole spine, which has been suggested as a possible trigger for syndesmophyte-linked ankylosis [7]. Yet, important questions regarding the local vertebral

* Corresponding author at: Department of Rheumatology, Leiden University Medical Center; PO Box 9600, 2300 RC, Leiden, The Netherlands.
E-mail addresses: mary.lucy.marques@gmail.com (M.L. Marques), mail@dvanderheijde.nl (D. van der Heijde), m.reijnierse@lumc.nl (M. Reijnierse), Xenofon.Baraliakos@elisabethgruppe.de (X. Baraliakos), juergen.braun@elisabethgruppe.de (J. Braun), F.A.van_Gaalen@lumc.nl (F. van Gaalen).

https://doi.org/10.1016/j.semarthrit.2022.152144

Available online 30 November 2022

bone changes and their relationship with the disease- and treatment-related aspects are poorly understood [3]. Remarkably, while significant imaging advances have been observed in the assessment of inflammation and bone formation over the past years [8,9], local vertebral trabecular bone density changes remain challenging to assess in r-axSpA due to imaging limitations [6,10].

Conventionally, imaging techniques, namely Dual-energy X-ray absorptiometry (DXA), assess the patient's systemic bone loss status in comparison with data from healthy subjects matched by age and gender [11]. The possible underestimation of bone loss by the presence of bone formation is well-known in r-axSpA [4,12]. Moreover, most imaging techniques, even when able to exclude bone formation and, therefore, measuring trabecular bone only (e.g., quantitate Computed Tomography (qCT)), assess few vertebrae, mostly in the lumbar spine [10,13,14]. Since r-axSpA affects the whole spine, substantial data regarding local vertebral trabecular bone density changes is lacking.

Assessing local bone quality using CT scans with Hounsfield units (HU) quantification is possible and has been shown to correlate with DXA bone mineral density measurements [15–17]. HU assess the tissue density on CT, based on a defined scale of zero for water and -1,000 for air. Trabecular bone HU typically ranges from 200 up to 800 HU [16]. Modern imaging software programs allow HU to be calculated from a region of interest on CT scans without additional costs or radiation exposure [16,18].

Low dose CT (ldCT) has arisen as a promising tool to assess not only bone formation, but also vertebral trabecular bone density throughout the whole spine, using acceptable levels of ionizing radiation exposure. (21) Using ldCT scans, we have recently shown the excellent cross-sectional reliability of the trabecular HU measurements from C3 to L5 in patients with r-axSpA [19]. However, ldCT HU changes over time and their relation to measurement error have never been studied.

In the present study, we aimed to describe ldCT HU two-year change-from-baseline values and to analyse their inter-reader reliability per vertebra in patients with r-axSpA.

Methods

Study design and population

We used data from the Sensitive Imaging in Ankylosing Spondylitis (SIAS) study, a multicentre two-year prospective cohort of patients with r-axSpA recruited in Leiden (Netherlands) and Herne (Germany). Ethics approval was obtained at each centre. An informed consent was voluntarily signed by participants before enrolment, and coded data was used. The cohort was previously described in detail [9].

For the current analysis, we used baseline and two-year spine ldCT scans of 50 patients. This allowed us to include a maximum of 50 vertebrae per vertebral level (C3-L5), which was considered adequate according to a sample size calculation performed focusing on inter-reader reliability (for measurements of two readers, using pre-defined intra-class correlation coefficient (ICCs) of 0.80 or higher with a 95% CI±0.1) - Online supplementary text S1 [20].

Imaging assessments of bone density changes

In both centres, a standardized protocol was applied for ldCT imaging acquisition. The ldCT scanners used in Leiden and Herne were Aquilion one (Canon, Otawara, Japan), and Somatom Emotion 16 (Siemens, Erlangen, Germany), respectively. Patients were placed in the supine position, feet first. Helical CT scans were performed from the superior endplate of C2 to the inferior endplate of S1 using pre-defined settings: 60 mAs at 120 kVp and a pitch of 53/65 using automatic exposure control with SD 30 – Filter FC 18 – slice thickness 5 mm (min 10 mA, max 60mA). Reconstructions were obtained for three orthogonal planes: for the axial plane, with 1- and 3-mm slices, while for coronal and sagittal planes, with 2 mm slices. Imaging reconstructions were performed using iterative reconstruction algorithms for noise suppression (adaptive iterative dose reduction 3D enhanced, eAIDR 3D). The effective dose estimates for the whole spine were 3.8 (2.6) mSv and 4.7 (2.4) mSv per ldCT, respectively. Differences between 16- and 64-slice scanners, for a similar-sized patient (i.e., CTDIvol), were around 10%. The 64-slice scanner involved less ionizing radiation exposure, which in addition to the scanner settings and performance, can be explained by the area scanned (higher on the 16-slice scanner) [19].

Vertebral ldCT HU (continuous value for the whole vertebra) were assessed for each vertebra (C3 -L5) by two trained readers using OsiriX software (v6.5.1). The used methodology was validated in trauma patients and spine surgery candidates and was recently adapted for the whole spine of patients with r-axSpA [15–17,19]. The HU value obtained at each vertebra corresponded to the average image density within a manually selected region of interest centrally positioned in the vertebral trabecular bone, avoiding peripheral bone (periosteum and ectopic bone) (Fig. 1). Therefore, when we refer to bone density or HU hereafter, these terms pertain to vertebral trabecular bone.

First, baseline vertebral HU were independently measured by each reader to assess the feasibility and cross-sectional reliability of the HU methodology in r-axSpA patients. Given the excellent cross-sectional reliability [19], the two-year ldCT scans were subsequently assessed. The readers independently measured HU from the two time-point images blinded to the previous measurement values.

Statistical analysis

LdCT scans (C3-L5) from 50 patients were available. Only vertebrae from which HU were measured at baseline and two years by both readers could be included. Poor imaging quality precluded the assessment of HU at two years in all vertebrae from one patient. Thus, that patient was excluded.

Mean HU change values (*HU measurement at two-years – HU measurement at baseline*) by reader, and mean differences between readers' change values were provided per vertebra. The number and type of incident density abnormalities or artefacts were described per vertebra at each time-point, if reported by at least one reader.

Reliability and agreement of HU change values

Inter-reader reliability and agreement were assessed per vertebra (C3-L5). ICC single measurements, absolute agreement were used, applying two-way random effect models [21]. Agreement was assessed using Bland-Altman plots and the smallest detectable change (SDC = 1.96 x SDdifference /($\sqrt{k}$)) [22]. SDdifference is the standard deviation of the differences in HU change values between the two readers and k is the number of readers (n=2). The HU measurements are independently performed at each time point. Indeed, since the software automatically gives the HU in the manually selected region of interest, the comparison with previous images is not required. Therefore, the part of the SDC formula accounting for differences coming from two simultaneous measurements (dividing by $\sqrt{2}$) was not applicable here, the formula being adjusted accordingly.

The percentages of vertebrae in which readers agreed on the direction of change were computed. To be considered a significant bone density loss (negative change value: *HU measurement at baseline* > *HU measurement at two years*) or increase (positive change value: *HU measurement at baseline* < *HU measurement at two years*), the HU change value must surpass the SDC [22]. Therefore, the percentages of agreement on HU change values > |SDC|, i.e., above the absolute value of the SDC (>SDC and <-SDC) were also computed per vertebra.

As sensitivity analyses, the reliability and agreement were also tested for 1) each centre separately (external validity) and, 2) excluding vertebrae with incident density abnormalities or artefacts, therefore excluding potentially extreme HU change values due to these features.

Statistical analyses were performed using *STATA* software version 16.0.

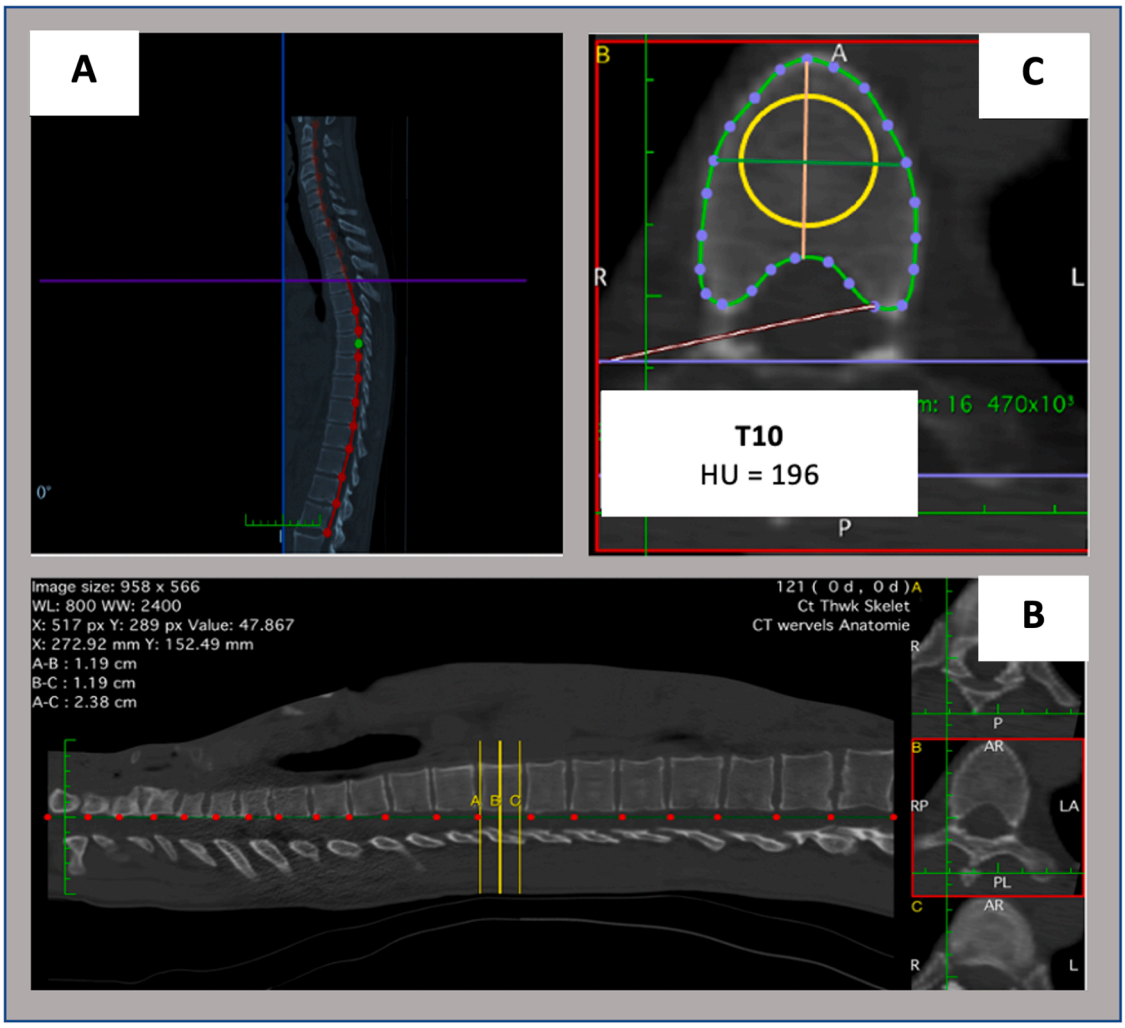

Fig. 1. Methodology of Hounsfield Units (HU) measurement. A: Using a three-dimensional curved-multiplanar reconstruction, the spinal curve was manually delimited. **B**: On the sagittal image, after identifying each vertebra (C3-L5), two lines of reference were manually positioned at the superior (yellow-line A) and inferior (yellow-line C) vertebral limits. Equidistant to A and C, the yellow-line B was automatically positioned by the software. **C:** HU measurements were taken from a single reconstructed cross-sectional slice at the level of yellow line B. After manually delimiting the vertebra, a circular region of interest was selected based on manually defined reference lines, having a diameter of 75% of the anteroposterior and transverse diameters. The vertebral average density within the sample region was displayed by the software in HU.

Results

Whole spine ldCT scans from 49 patients with r-axSpA (mean (SD) age of 49 (10) years; 88% male and 86% HLA-B27 positive) were included in the final analysis – 26 from Leiden and 23 from Herne. Demographic and clinical characteristics of the included patients were provided in Table 1.

Extreme imaging quality artefacts (not possible to identify the limits of the vertebra) precluded the assessment of 25 vertebrae, all in the cervical spine. Therefore, a total of 1,053 vertebrae (98% of all possible vertebrae) could be assessed for HU measurements at both time-points (220 cervical, 588 thoracic, and 245 lumbar vertebrae).

The HU mean (SD) change values varied from -23 (59) to 28 (70) and -23 (60) to 29 (70) for reader 1 and 2, respectively (Table 2). Negative change values were mostly observed in the thoracolumbar vertebrae rather than cervical vertebrae (Table 2). The HU values decreased from cranial to caudal vertebrae at both time-points. Data regarding vertebral HU values at each time-point by reader were provided in Supplementary Table S1.

Density abnormalities or artefacts were identified by at least one of the readers in 72 (7%) and 222 (21%) vertebrae, at baseline and two years, respectively. Imaging noise, photon starvation artefact (typical at the shoulder level) and sclerotic changes of the vertebral body (most prevalent at L5) were the most frequently reported at both time-points (Supplementary Tables S2 and S3).

Reliability and agreement of HU change values

ICCs for the HU two-year change-from-baseline values were between 0.91 to 0.99 (Table 2). The two readers agreed on the direction of the change in 88-96% of vertebrae, SDC varying from 6 to 10, with mean and median of 7 (Table 3). Bland-Altman plots showed homoscedasticity throughout the whole spine, with negligible systematic error between the readers (Supplementary Figs. S1–S22). Most vertebrae have shown HU changes beyond measurement error, with agreement of both readers on change values $>|SDC|$ obtained in 58-94% of the vertebra per vertebral level (Table 2).

In sensitivity analysis, the HU mean change values and direction of change varied per centre, though with the same pattern of higher prevalence of negative change values in the thoracolumbar vertebrae than in the cervical ones. For both centres, most vertebrae HU change-from-baseline values were beyond measurement error. In Leiden, the

Table 1
Baseline characteristics of patients with r-axSpA.

Assessment[§]	*N=49*
Male	43 (88)
Age, years	49.1 (9.9)
Body Mass Index, kg/m^2	26.6 (4.2)
HLA-B27 positive	42 (86)
ASDAS-CRP	2.6 (1.2)
TNFi treatment	12 (25)
NSAIDs treatment	32 (65)
Patients with syndesmophytes[#]	49 (100)
Patients with MRI BME[¥]	47 (96)
Cervical spine HU[*]	320 (104.7)
Thoracic spine HU[*]	197 (70.7)
Lumbar spine HU[*]	157 (63.5)

ASDAS, Ankylosing Spondylitis Disease Activity Score; **BME,** bone marrow edema; **CRP,** C-reactive protein levels; **HLA,** human leucocyte antigen; **HU,** Hounsfield units; **NSAIDs,** non-steroidal anti-inflammatory drugs; **TNFi,** Tumor necrosis factor inhibitors.

[§] Data is presented as mean (standard deviation) or no. (%).

[#] Defined as a patient with at least one quadrant that received a CT Syndesmophytes Score ≥1 (absolute agreement of two readers).

[¥] Defined as a patient with at least one quadrant with MRI BME (agreement of 2 out of 3 readers).

[*] Average of the two readers' measurements.

agreement in change values > |SDC| was obtained for 58-95% of vertebrae (Supplementary Table S4), and in Herne for 43-96% of vertebrae (Supplementary Table S5). The reliability remained excellent (Supplementary Tables S4 and S5).

The HU mean change values for vertebrae without incident density abnormalities or artefacts were identical to whole vertebrae data (Supplementary Table S6). After excluding vertebrae with these potential extreme change values, similar high percentages of agreement in changes >|SDC| were observed, with 56-93% of vertebrae changing HU values above measurement error (Supplementary Table S6).

Discussion

In the present study, we assessed the two-year change-from-baseline vertebral trabecular bone density as measured by ldCT HU in patients with r-axSpA. Excellent inter-reader reliability and agreement were found at the same vertebral level, from C3 to L5, most changes in HU occurring beyond measurement error.

The direction and magnitude of HU mean change values varied per vertebra and by centre. Notably, unlike the unidirectional nature of changes in syndesmophyte scoring, in which only progression is possible [9,23], bone density HU can decrease or increase over time. Lower HU values imply a lower ldCT attenuation, and therefore less-dense bones [16]. Negative change values meant bone density loss over two years, while positive change values represented bone density gain. The higher the HU change value, the greater magnitude of bone density change.

Despite the variability in change values observed by vertebra, readers agreed on the direction of the change in most vertebrae (88 to 96%). Inter-reader reliability was excellent. In addition to ICCs (0.91 to 0.99), whose results could be spuriously high due to the large spread of the HU change values [21], importantly, Bland-Altman plots showed homoscedasticity throughout the spine, and negligible systematic errors between readers [24].

For consistency, we used the mean/median SDC of seven as a cut-off to assess changes beyond measurement error in all vertebrae. Readers agreed on changes above the absolute value of SDC in most vertebrae (58-94%), with few changes in HU values attributable to measurement error [22]. Remarkably, sensitivity analysis performed in vertebrae without incident density abnormalities or artefacts has shown similar results, suggesting that significant bone density changes occurred despite the absence of these features.

Of note, features such as bone marrow edema (water-based), and fat deposition, which can interfere with CT attenuation, yielding lower HU values, are poorly identifiable in ldCT images [25]. Both bone marrow edema and fat deposition have been widely described in magnetic resonance imaging (MRI) of patients with axSpA, including in the SIAS cohort [26–28]. However, both bone marrow edema and fat deposition

Table 2
Two-year Hounsfield Units (HU) change values per reader, inter-reader reliability and agreement on changes beyond measurement error, from C3-L5.

Vertebra[§]	HU mean change values (SD)		Mean Difference (SD) between the readers	ICC	No agreement %	Agreement % Change >|7|[#] Change > 7[#]	Change < -7[#]	Total	Change <|7|[*]
	Reader 1	Reader 2							
C3	18 (56)	17 (56)	0.2 (5.0)	0.97	5	52	32	84	11
C4	18 (53)	17 (52)	0.3 (5.5)	0.98	2	59	32	91	7
C5	28 (70)	29 (70)	-0.7 (5.0)	0.99	5	59	32	91	5
C6	23 (62)	23 (62)	0.4 (5.8)	0.99	7	50	34	84	9
C7	-3 (60)	-2 (59)	-0.7 (6.9)	0.98	12	40	46	86	2
T1	-6 (87)	-6 (88)	0.6 (5.0)	0.98	4	45	49	94	2
T2	3 (45)	3 (45)	0.7 (4.0)	0.97	10	37	39	76	14
T3	1 (43)	2 (43)	0.6 (4.4)	0.95	12	27	31	58	31
T4	-2 (48)	-1 (47)	0.2 (5.3)	0.94	6	35	45	80	14
T5	0.02 (48)	-0.3 (48)	0.3 (5.1)	0.91	6	39	39	78	16
T6	-3 (44)	-3 (45)	-0.1 (4.6)	0.95	8	35	43	78	14
T7	-4 (43)	-4 (43)	-0.1 (4.5)	0.99	8	27	49	76	16
T8	-1 (38)	-1 (40)	0.1 (4.6)	0.98	10	33	41	74	16
T9	-9 (50)	-9 (50)	0.02 (5.4)	0.99	20	20	55	75	4
T10	0.2 (59)	1 (59)	-0.4 (5.2)	0.96	16	31	39	70	14
T11	-8 (53)	-7 (53)	-0.6 (4.3)	0.94	10	33	37	70	20
T12	-23 (59)	-23 (60)	-0.3 (4.4)	0.99	6	31	53	84	10
L1	-9 (33)	-7 (33)	-1.9 (6.3)	0.97	20	27	41	68	12
L2	1 (46)	2 (45)	-1.1 (4.6)	0.98	10	25	39	64	27
L3	-4 (35)	-2 (34)	-1.2 (6.3)	0.92	14	25	39	64	22
L4	-2 (24)	1 (22)	-1.5 (5.4)	0.91	12	29	37	66	22
L5	8 (43)	9 (43)	-1.0 (6.0)	0.97	14	43	31	74	12

[§] C3-C7: n=44; T1-L5: n=49.

[#] Percentages of agreement on change beyond measurement error, i.e., on HU change values above the absolute value of the smallest detectable change (>|SDC|= >SDC and <-SDC). For consistency, the mean/median SDC of 7 was used as a cut-off to assess changes beyond measurement error in all vertebrae.

[*] Percentages of agreement on HU change values within measurement error, i.e., below the absolute value of the smallest detectable change (<|SDC|= ≤SDC and ≥-SDC). **ICC –** intraclass correlation coefficient; **SD –** standard deviation.

Table 3
Number (%) of vertebrae with positive and negative two-year Hounsfield Units (HU) change values for each reader separately, percentage of agreement between readers, and smallest detectable change (SDC) per vertebra.

Vertebra[§]	Reader 1		Reader 2		No agreement (%)	Agreement %		SDC
	Positive	Negative	Positive	Negative		Positive	Negative	
C3	27 (61)	17 (39)	28 (63)	16 (36)	7	59	34	7
C4	29 (66)	15 (34)	28 (63)	16 (36)	7	61	32	8
C5	29 (65)	15 (34)	27 (61)	17 (39)	5	61	34	7
C6	22 (50)	22 (50)	25 (57)	19 (43)	7	50	43	8
C7	20 (45)	24 (55)	21 (48)	23 (52)	5	43	52	10
T1	24 (49)	25 (51)	23 (47)	26 (53)	2	47	51	7
T2	24 (49)	25 (51)	26 (53)	23 (47)	8	47	45	6
T3	22 (45)	27 (55)	22 (45)	27 (55)	12	39	49	6
T4	21 (43)	28 (57)	23 (47)	26 (53)	8	41	51	7
T5	23 (47)	26 (53)	24 (49)	25 (51)	10	43	47	7
T6	24 (48)	25 (51)	18 (37)	31 (63)	12	37	51	6
T7	18 (37)	31 (63)	19 (39)	30 (61)	10	33	57	6
T8	22 (45)	27 (55)	23 (47)	26 (53)	6	43	51	6
T9	13 (27)	36 (73)	14 (29)	35 (71)	6	25	69	7
T10	23 (47)	26 (53)	21 (43)	28 (57)	8	41	51	7
T11	25 (51)	24 (49)	21 (43)	28 (57)	8	43	49	6
T12	21 (43)	28 (57)	22 (45)	27 (55)	2	43	55	6
L1	24 (49)	25 (51)	23 (47)	26 (53)	6	45	49	9
L2	21 (43)	28 (57)	21 (43)	28 (57)	12	37	51	6
L3	20 (41)	29 (59)	20 (41)	29 (59)	12	35	53	9
L4	22 (45)	27 (55)	18 (37)	31 (63)	8	37	55	7
L5	28 (57)	21 (43)	28 (57)	21 (43)	4	55	41	8

Positive change value: HU measurement at two years > HU measurement at baseline.
Negative change value: HU measurement at two years < HU measurement at baseline.
[§] C3-C7: n=44; T1-L5: n=49.

have also been reported in the trabecular vertebral bone of patients with osteoporosis and compression vertebral fractures [29,30]. Vertebral trabecular bone loss may itself encompass much more than calcium loss.

In measuring HU at the centre of each vertebra from C3 to L5, we likely avoided the peripherally located axSpA-related bone marrow edema and fat deposition. Although these features can still be present centrally in the vertebrae, we would not expect drastic between-vertebrae differences (at each level) since the axSpA patients from the SIAS cohort had similarly advanced disease stages and a narrow age range [28].

New techniques such as DECT can detect water and fat content, subtracting calcium, and therefore allowing for a more accurate measurement of bone mineral density loss [31,32]. Several studies have shown the application of DECT for detecting bone marrow edema in trauma settings or osteoporosis [30,33]. However, data on arthritis, especially of the axial skeleton, is sparse and focusing on the sacroiliac joints [34–36]. DECT evaluation of bone marrow edema in the spine of patients with axSpA is lacking. Moreover, DECT depicts high contrast imaging at the cost of high ionizing radiation exposure (in line with the dose for standard CT), which makes it unacceptable for the assessment of the entire spine [25,33].

In all analyses, thoracic and lumbar vertebrae had more frequently negative two-year change-from-baseline values (bone density loss), contrasting with vertebrae from the cervical spine. The faster decrease of bone mass caudally than cranially within vertebrae was previously reported in normal spine histomorphometry data [37]. Using whole spine vertebrae from autopsies, structural differences between opposite sides of the spine (lumbar vs cervical) were shown to increase with age, with bone mass decreasing faster caudally than cranially within vertebrae [37].

In r-axSpA, the decrease in caudal HU change-from-baseline values shown over a relatively short follow-up (two years) may also be related with disease-specific features, such as local inflammation or impairment in mobility. In the same cohort, SIAS, thoracic and lumbar vertebrae were shown to have higher levels of bone marrow edema on MRI [27], and higher prevalence and incidence of syndesmophytes on ldCT scans [9,27]. Other factors, namely treatment with tumour necrosis factor inhibitors [38], may influence trabecular bone density. However, whether changes in trabecular bone density over time are associated with disease- or treatment-specific factors is beyond the scope of the present study.

Although not the focus of the present work, we showed for the first time that the gradient in HU values, decreasing from cranial to caudal vertebrae [19], persists over time (present at baseline and two-year measurements). This gradient is possibly related to inter-vertebral variability in dynamic forces, mobility, and size [37].

This study is not without limitations. The accuracy of HU measurements does not benefit from having a previous image, as occurs, for instance, in scoring of the syndesmophytes' progression. Therefore, the SDC formula required adjustments (not dividing by $\sqrt{2}$) [22]. Nevertheless, SDC values, which ranged between 6 and 10, were relatively small in comparison to the full range of HU values. Conversely, as HU can be independently measured, no re-measurement of previous images is needed when new time points become available. The HU values may be difficult to interpret by clinicians who are used to cut-off definitions for osteoporosis. However, it is important to note that this technique is intended for the local assessment of trabecular bone changes and not for the clinical diagnosis of osteoporosis [19]. Therefore, its validation against a gold standard for the diagnosis of osteoporosis, namely DXA or QCT, although previously performed in trauma subjects and spine surgery candidates [15–17], was not repeated in r-axSpA. Importantly, previously proposed HU cut-offs in relation to osteoporosis definitions have shown to be heterogeneous across studies, which imposes caution in their usage [17]. In addition, r-axSpA poses specific challenges. A possible validation against DXA would not be accurate as DXA measurements cannot avoid ectopic bone formation, while HU as measured by ldCT assess the trabecular central bone only, excluding ectopic bone formation. Moreover, a validation against QCT would be unfeasible for the whole spine due to unacceptable ionizing radiation exposure. Therefore, for the time being, HU values should be used to reliably compare bone density changes between vertebrae throughout the whole spine, for research purposes only.

This study has several strengths. First, the whole spine was assessed using low ionizing radiation exposure (the effective dose estimates were around 4 to 5 mSV per ldCT in SIAS but are currently estimated to be as low as 1.4 to 1.7 mSV per ldCT, using technical optimisation in the 64-

slice scanners), without noticeable imaging quality loss [19]. The methodology used to measure HU prevents artificially increased values due to peripheral ectopic bone formation, being easy to be used by others than radiologists within a limited number of hours of training. A comprehensive statistical methodology to test inter-reader reliability and agreement was used. Since the cross-sectional reliability of HU measurements was shown to be excellent [19], in the present study we assessed ICCs applying two-way random effects models [21]. The results mimicked those obtained in two-way mixed effects models (data not shown) but allow generalizability to any potential reader assessing the images [21]. The sensitivity analysis excluding vertebrae with incident density abnormalities and artefacts added robustness to the assessment of true changes in trabecular bone density. On the other hand, the sensitivity analysis stratified by centres from different countries (Leiden, Netherlands and Herne, Germany) added to the external validity of the results, limiting the concerns regarding HU values obtained from distinct automatic exposure control implementation and iterative reconstruction algorithms [18,39]. Moreover, the results coming from different CT scanners in different centres (Canon in Leiden, and Siemens in Herne), likely represent the most feasible approach for HU measurements in future studies (especially if multicentric). Of note, each patient was re-assessed using the same CT scanner and imaging acquisition settings, warranting less impact on the individual change-from-baseline HU values. This principle must be ensured in future studies aiming at assessing changes in HU over time. The HU measurements express the CT "bone density" and not the traditionally used calcium-related "bone mineral density", depicting the broad spectrum of changes occurring in the trabecular bone (e.g., age-related bone marrow changes) [28]. The continuous values of HU (as proposed by us) can be reliably compared between vertebrae throughout the whole spine in patients with axSpA for research purposes.

In summary, ldCT measurement of HU is a reliable method to assess changes in trabecular bone density over two years at each vertebra from C3 to L5. As reliable across all vertebrae, this innovative methodology can aid the comprehensive study of bone density changes over time in r-axSpA, a disease affecting the entire spine.

Author contributions

XB, JB, FAvG and MR performed the data collection for the SIAS cohort. MLM, DvdH, SR and FAvG developed the study design. MLM and NPdS independently performed the Hounsfield Units measurements. MLM prepared the dataset. MLM and SR performed the statistical analyses and synthesised the data. MLM has drafted the first version of the manuscript, and all authors have critically reviewed and agreed with the final version of the manuscript.

Funding

SIAS study was funded by the Dutch Rheumatism Association ("ReumaNederland"). MLM is supported by the Fundação para a Ciência e Tecnologia (FCT) grant SFRH/BD/143744/2019.

Data availability statement

The data that support the findings of this study are available from the corresponding authors on reasonable request.

Declaration of Competing Interest

The authors declare no conflicts of interest in relation to this study.

Acknowledgments

None

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2022.152144.

References

[1] Machado P, Landewe R, Braun J, et al. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. Ann Rheum Dis 2010;69:1465–70. https://doi.org/10.1136/ard.2009.124206.
[2] Navarro-Compán V, Sepriano A, El-Zorkany B, et al. Axial spondyloarthritis. Ann Rheum Dis 2021;80:1511–21. https://doi.org/10.1136/annrheumdis-2021-221035.
[3] Clunie G, Horwood N. Loss and gain of bone in spondyloarthritis: what drives these opposing clinical features? Ther Adv Musculoskelet Dis 2020;12. https://doi.org/10.1177/1759720X20969260. 1759720X20969260.
[4] Klingberg E, Lorentzon M, Mellstrom D, et al. Osteoporosis in ankylosing spondylitis - prevalence, risk factors and methods of assessment. Arthritis Res Ther 2012;14:R108. https://doi.org/10.1186/ar3833.
[5] Fitzgerald GE, O, Shea FD. The fascinating paradox of osteoporosis in axial spondyloarthropathy. J Rheumatol 2017;44:1767–76. https://doi.org/10.3899/jrheum.170051.
[6] Ramírez J, Nieto-González JC, Curbelo Rodríguez R, et al. Prevalence and risk factors for osteoporosis and fractures in axial spondyloarthritis: A systematic review and meta-analysis. Semin Arthritis Rheum 2018;48:44–52. https://doi.org/10.1016/j.semarthrit.2017.12.001.
[7] Lories RJ. Advances in understanding the pathophysiology of spondyloarthritis. Best Pract Res Clin Rheumatol 2018;32:331–41. https://doi.org/10.1016/j.berh.2018.12.001.
[8] Diekhoff T, Eshed I, Radny F, et al. Choose wisely: imaging for diagnosis of axial spondyloarthritis. Ann Rheum Dis 2022;81:237–42. https://doi.org/10.1136/annrheumdis-2021-220136.
[9] de Bruin F, de Koning A, van den Berg R, et al. Development of the CT Syndesmophyte Score (CTSS) in patients with ankylosing spondylitis: data from the SIAS cohort. Ann Rheum Dis 2018;77:371–7. https://doi.org/10.1136/annrheumdis-2017-212553.
[10] Marques ML, Ramiro S, Machado PM, et al. No relationship between bone mineral density and syndesmophyte formation at the same level in the lumbar spine of patients with radiographic axial Spondyloarthritis. RMD Open 2020;6. https://doi.org/10.1136/rmdopen-2020-001391.
[11] Bazzocchi A, Ponti F, Albisinni U, et al. DXA: Technical aspects and application. Eur J Radiol 2016;85:1481–92. https://doi.org/10.1016/j.ejrad.2016.04.004.
[12] Lim MJ, Kang KY. A Contemporary View of the Diagnosis of Osteoporosis in Patients With Axial Spondyloarthritis. Front Med (Lausanne) 2020;7. https://doi.org/10.3389/fmed.2020.569449.
[13] Lee SY, Song R, Yang HI, et al. The bone bridge significantly affects the decrease in bone mineral density measured with quantitative computed tomography in ankylosing spondylitis. PLoS One 2021;16:e0249578. https://doi.org/10.1371/journal.pone.0249578.
[14] Fauny M, Verhoeven F, Allado E, et al. Relationship between spinal structural damage on radiography and bone fragility on CT in ankylosing spondylitis patients. Sci Rep 2021;11:9342. https://doi.org/10.1038/s41598-021-88838-9.
[15] Patel Shaun P, Lee John J, Hecht Garin G, Holcombe Sven A, SCW and JAG. Normative vertebral hounsfield unit values and correlation with bone mineral density. J Clin Exper Orthopaed 2016;2:1–7.
[16] Schreiber JJ, Anderson PA, Rosas HG, et al. Hounsfield units for assessing bone mineral density and strength: a tool for osteoporosis management. J Bone Joint Surg Am 2011;93:1057–63. https://doi.org/10.2106/JBJS.J.00160.
[17] Ahern DP, McDonnell JM, Riffault M, et al. A meta-analysis of the diagnostic accuracy of Hounsfield units on computed topography relative to dual-energy X-ray absorptiometry for the diagnosis of osteoporosis in the spine surgery population. Spine J 2021;21:1738–49. https://doi.org/10.1016/j.spinee.2021.03.008.
[18] McKenney SE, Seibert JA, Lamba R, et al. Methods for CT automatic exposure control protocol translation between scanner platforms. J Am Coll Radiol 2014;11:285–91. https://doi.org/10.1016/j.jacr.2013.10.014.
[19] Marques ML, Pereira da Silva N, van der Heijde D, et al. Low-dose CT hounsfield units: a reliable methodology for assessing vertebral bone density in radiographic axial spondyloarthritis. RMD Open 2022;8:e002149. https://doi.org/10.1136/rmdopen-2021-002149.
[20] Giraudeau B, Mary JY. Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. Stat Med 2001;20:3205–14. https://doi.org/10.1002/sim.935.
[21] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
[22] Bruynesteyn K, Boers M, Kostense P, et al. Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis 2005;64:179–82. https://doi.org/10.1136/ard.2003.018457.
[23] de Koning A, de Bruin F, van den Berg R, et al. Low-dose CT detects more progression of bone formation in comparison to conventional radiography in

patients with ankylosing spondylitis: results from the SIAS cohort. Ann Rheum Dis 2018;77:293–9. https://doi.org/10.1136/annrheumdis-2017-211989.
[24] Gerke O. Reporting standards for a Bland–Altman agreement analysis: a review of methodological reviews. Diagnostics 2020;10:334. https://doi.org/10.3390/diagnostics10050334.
[25] Diekhoff T, Hermann KGA, Lambert RG. Future of low-dose computed tomography and dual-energy computed tomography in axial spondyloarthritis. Curr Rheumatol Rep 2022;24:198–205. https://doi.org/10.1007/s11926-022-01075-5.
[26] Machado PM, Baraliakos X, van der Heijde D, et al. MRI vertebral corner inflammation followed by fat deposition is the strongest contributor to the development of new bone at the same vertebral corner: a multilevel longitudinal analysis in patients with ankylosing spondylitis. Ann Rheum Dis 2016;75:1486–93. https://doi.org/10.1136/annrheumdis-2015-208011.
[27] Stal R, Baraliakos X, van der Heijde D, et al. Role of vertebral corner inflammation and fat deposition on MRI on syndesmophyte development detected on whole spine low-dose CT scan in radiographic axial spondyloarthritis. RMD Open 2022;8:e002250. https://doi.org/10.1136/rmdopen-2022-002250.
[28] Ahn GY, Koo BS, bin Joo K, et al. Use of quantitative vertebral bone marrow fat fraction to assess disease activity and chronicity in patients with ankylosing spondylitis. Kor J Radiol 2021;22:1671. https://doi.org/10.3348/kjr.2020.0953.
[29] Bäcker HC, Wu CH, Perka C, et al. Dual-energy computed tomography in spine fractures: a systematic review and meta-analysis. Int J Spine Surg 2021;15:525–35. https://doi.org/10.14444/8074.
[30] Diekhoff T, Engelhard N, Fuchs M, et al. Single-source dual-energy computed tomography for the assessment of bone marrow oedema in vertebral compression fractures: a prospective diagnostic accuracy study. Eur Radiol 2019;29:31–9. https://doi.org/10.1007/s00330-018-5568-y.
[31] Gruenewald LD, Koch V, Martin SS, et al. Diagnostic accuracy of quantitative dual-energy CT-based volumetric bone mineral density assessment for the prediction of osteoporosis-associated fractures. Eur Radiol 2022;32:3076–84. https://doi.org/10.1007/s00330-021-08323-9.
[32] Booz C, Noeske J, Albrecht MH, et al. Diagnostic accuracy of quantitative dual-energy CT-based bone mineral density assessment in comparison to Hounsfield unit measurements using dual x-ray absorptiometry as standard of reference. Eur J Radiol 2020;132:109321. https://doi.org/10.1016/j.ejrad.2020.109321.
[33] Foti G, Serra G, Iacono V, et al. Identification of traumatic bone marrow oedema: the pearls and pitfalls of Dual-Energy CT (DECT). Tomography 2021;7:424–33. https://doi.org/10.3390/tomography7030037.
[34] Carotti M, Benfaremo D, di Carlo M, et al. Dual-energy computed tomography for the detection of sacroiliac joints bone marrow oedema in patients with axial spondyloarthritis. Clin Exp Rheumatol 2021;39:1316–23. https://doi.org/10.55563/clinexprheumatol/xdlfzb.
[35] Wu H, Zhang G, Shi L, et al. Axial spondyloarthritis: dual-energy virtual noncalcium CT in the detection of bone marrow Edema in the Sacroiliac joints. Radiology 2019;290:157–64. https://doi.org/10.1148/radiol.2018181168.
[36] Chen M, Herregods N, Jaremko JL, et al. Bone marrow edema in sacroiliitis: detection with dual-energy CT. Eur Radiol 2020;30:3393–400. https://doi.org/10.1007/s00330-020-06670-7.
[37] Grote HJ, Amling M, Vogel M, et al. Intervertebral variation in trabecular microarchitecture throughout the normal spine in relation to age. Bone 1995;16:301–8. https://doi.org/10.1016/8756-3282(94)00042-5.
[38] Visvanathan S, van der Heijde D, Deodhar A, et al. Effects of infliximab on markers of inflammation and bone turnover and associations with bone mineral density in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:175–82. https://doi.org/10.1136/ard.2007.084426.
[39] Sookpeng S, Martin CJ, Cheebsumon P, et al. Practical experiences in the transfer of clinical protocols between CT scanners with different ATCM systems. J Radiolog Protect 2017;37:84–96. https://doi.org/10.1088/1361-6498/37/1/84.