Quantitative syndesmophyte measurement in ankylosing spondylitis using CT: longitudinal validity and sensitivity to change over 2 years

Sovira Tan[1], **Jianhua Yao**[2], **John A Flynn**[3], **Lawrence Yao**[2], and **Michael M Ward**[1]

[1]National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland, USA

[2]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda, Maryland, USA

[3]Johns Hopkins Medical Institutions, Baltimore, Maryland, USA

Abstract

Objectives—Accurate measurement of syndesmophyte development and growth in ankylosing spondylitis (AS) is needed for studies of biomarkers and of treatments to slow spinal fusion. We tested the longitudinal validity and sensitivity to change of quantitative measurement of syndesmophytes using CT.

Methods—We performed lumbar spine CT scans on 33 patients with AS at baseline, 1 year and 2 years. Volumes and heights of syndesmophytes were computed in four intervertebral disk spaces. We compared the computed changes to a physicians ratings of change ' based on CT scan inspection. Sensitivity to change of the computed measures was compared with that of the modified Stoke AS Spinal Score (radiography) and a scoring method based on MRI.

Results—At years 1 and 2, respectively 24 (73%) and 26 (79%) patients had syndesmophyte volume increases by CT. At years 1 and 2, the mean (SD) computed volume increases per patient were, respectively 87 (186) and 201 (366) mm^3. Computed volume changes were strongly associated with the physicians visual ratings of ' change (p<0.0002 and p<0.0001 for changes at years 1 and 2, respectively). The sensitivity to change over 1 year was higher for the CT volume measure (1.84) and the CT height measure (1.22) than either the MRI measure (0.50) or radiography (0.29).

Conclusions—CT-based syndesmophytes measurements had very good longitudinal validity and better sensitivity to change than radiography or MRI. This method shows promise for longitudinal clinical studies of syndesmophyte development and growth.

Ankylosing spondylitis (AS) is an inflammatory arthritis affecting primarily the sacroiliac joints and spine.[1] Growth of syndesmophytes at the intervertebral disk space (IDS) is a characteristic feature of AS. Because syndesmophytes represent progressive irreversible structural damage and are more easily detected than changes in the facet or sacroiliac joints, monitoring of their development has been a central focus of many studies. Studies of the pathogenesis of AS have tested associations of biomarkers and genetic polymorphisms with the extent and size of syndesmophytes.[2–8] Similarly, vertebral inflammation as seen on MRI has been examined for associations with the development of new syndesmophytes.[9–12] The impact of tumour necrosis factor-α inhibitors on the progression of syndesmophytes has been investigated, with implications for understanding the role of cytokines in the pathogenesis of AS as well as for clinical care.[13–15]

These studies used plain radiographs and semi-quantitative ratings as the method to detect and score syndesmophytes. The main limitations of this methodology are a consequence of the use of a two-dimensional (2D) technique to assess a 3D structure, with problems of projection, penetration and overlying shadows, resulting in poor visualisation of syndesmophytes. Semiquantitative rating methods also have limited sensitivity to change.[16 17] These problems are accentuated when the goal is to detect syndesmophyte growth, because growth is typically slow. Possibly as a result of these issues, much research has been inconclusive. Whether tumour necrosis factor-α antagonists influence spinal fusion remains unresolved.[13–15 18] Despite several studies, the relationship between inflammation and syndesmophyte development was recently characterised as 'enigmatic'.[19] Similarly, the search for biomarkers has produced few strong predictors of syndesmophyte growth.

With the aim of improving the assessment of syndesmophyte growth, we developed a computer algorithm measuring syndesmophytes on lumbar spine CT scans.[20 21] The algorithm exploits the complete 3D information of CT scans and assesses syndesmophytes along the entire vertebral rim in a fully quantitative way. The method has very good reliability and cross-sectional validity.[22] In this study, we assessed the longitudinal validity of the algorithm over 2 years, and compared its sensitivity to change to that of the modified Stoke AS Spine Score (mSASSS) and an MRI-based measure of chronic spine damage.

METHODS

Patients

We enrolled patients at the National Institutes of Health and Johns Hopkins Medical Institutions in this prospective longitudinal study. Inclusion criteria were age 18 years or older, diagnosis of AS by the modified New York criteria,[23] and a Bath AS Radiology Index (BASRI) Lumbar Spine Score of 0, 1, 2, or 3 (ie, excluding patients with completely fused lumbar spines).[24] We ensured representation of patients with different degrees of structural damage by enrolling at least five patients in each BASRI category. We excluded patients who were pregnant or had contraindications to MRI. The study protocol was approved by

the institutional review boards of both centres, and all patients provided written informed consent.

CT scanning

Patients were scanned at baseline, year 1 and year 2. They were scanned on either a Philips Brilliance 64 (slice thickness 1.5 mm) or a GE Lightspeed Ultra scanner (slice thickness 1.25 mm). For both scanners, voltage and current parameters were 120 kVp and 300 mAs respectively. Patients were scanned from T10 to L4, providing 4 IDSs for processing: T11–T12, T12–L1, L1–L2, L2–L3.

Radiography and MRI scanning

Radiographs of the lumbar spine were taken at baseline, year 1 and year 2. Patients underwent lumbar spine MRI scans at baseline and year 1 on either a 1.5 T Signa Excite (GE) or a 3.0 T Achieva (Philips). Sagittal T1-weighted and short tau inversion recovery (STIR) sequences were obtained.

CT quantitative image analysis

Our semiautomated computer algorithm quantitates syndesmophyte volumes and heights.[20][21] It detects syndesmophytes as any bone projecting from the periphery of the vertebral end-plates, as voxels lying between the two planes of the endplates. The algorithm reports the total volume of syndesmophyte and height of the tallest syndesmophyte at each IDS. Syndesmophyte heights were divided by the height of the IDS so that a height of 1 represented bridging. Syndesmophyte progression was assessed as the difference in measures between baseline and either year 1 or year 2. Measures in 4 IDSs were added to provide volume and height measures per patient.

CT reading

A reader (MMW) visually assessed the change in syndesmophyte volume and maximal height in all IDSs. To make visual comparisons, images of the baseline and follow-up scans, reformatted to the same resolution, were displayed simultaneously next to each other on the same monitor. The scans were 'registered' using the Insight Segmentation and Registration Toolkit (ITK) image processing software library[25] so that the scans were aligned with the two slices displayed next to each other always showing exactly the same part of the IDS and in the same orientation. The reader scrolled through all slices in the coronal, sagittal and axial viewing directions. From all the slices, the reader assessed the extent of syndesmophyte volume change. For height, the reader assessed change on the slice where he located the maximal syndesmophyte height. IDSs with bridging at baseline were excluded from the assessment of changes in height. A second reader (LY) scored a subset of eight patients (25% of the total) independently.

Radiography and MRI reading

The 4 IDSs processed by the algorithm were also scored on radiographs using mSASSS. Scores of 1 were excluded because these do not represent syndesmophytes. The radiographic grading was performed by two readers (MMW and LY). We refer to this mSASSS, reduced

to be comparable to computed measurements, as 'adapted mSASSS'. The adapted mSASSS scores of the two readers had excellent agreement (intraclass correlation of 0.94 and 0.98 for IDSs and patients, respectively). Additionally, one reader (MMW) scored lumbar and cervical radiographs using the standard mSASSS, which will be referred to as 'standard mSASSS', while 'mSASSS' without qualification refers to both. One reader (MMW) also scored the same 4 IDSs in the T1-weighted MRI scans using the AS spine MRI-chronic (ASspiMRI-c) grading system.[26] For all imaging modalities, readers were blinded to computed measurements but not to time sequence.

Statistical analysis

We used two methods to test the longitudinal validity of our CT algorithm. First, we first investigated if changes in the CT measures were exclusively or largely positive, as would be expected since syndesmophyte growth is progressive and irreversible. To assess this, we constructed cumulative probability plots of changes in syndesmophyte volume and height over 1 year and 2 years. If the CT measures were valid in detecting change, we would expect these to increase or remain unchanged for most patients, with little or no evidence of decreasing values over time. Change over 2 years would also be expected to be larger than change over 1 year. For comparison, we constructed similar plots for changes in mSASSS. Second, we tested if computed changes in syndesmophyte volume and height were associated with the physicians' visual ratings of change on CT. We used the stratified Kruskal–Wallis trend test to evaluate if computed volume and height changes increased with the physician's scores, accounting for non-independence of observations within patients.[27] We measured the sensitivity to change using standardised response means (SRM).[28] CIs for the SRM were derived using bootstrapping with 200 random samples and a sampling rate of 0.7. We used SAS software (V.9.2, Cary, North Carolina, USA) for data analysis. Prior to computing the SRM, a logarithm transformation was applied to all changes to constrain their range. Patients with no change were assigned a SRM of 0, and the final SRM was a weighted average of the SRMs of changed and unchanged patients.

RESULTS

Patients

Of the 33 patients, 28 (85%) were men and 5 (15%) women. The mean (SD) age was 45.5 (11.8) years and the mean duration of AS was 20.6 (12.4) years; 132 IDSs from the 33 patients were studied at baseline, year 1 and year 2. During the study, 29 (88%) patients used non-steroidal anti-inflammatory medications and 9 (27%) used tumour necrosis factor-α inhibitors.

At baseline, the algorithm detected syndesmophytes in 81 IDSs (61%) and the mean (SD) volume was 274 (388) mm^3. Twenty-four patients (73%) were found to have existing syndesmophytes at baseline, and the mean total volume per patient was 1095 (1278) mm^3. Sixteen patients (48%) patients had an adapted mSASSS>0 by reader 1, while 17 patients (52%) had an adapted mSASSS>0 by reader 2. The mean adapted mSASSS per patient was 3.9 (5.8) for reader 1 and 4.2 (5.6) for reader 2. The mean standard mSASSS was 17.2 (15.4).

Computed volume and height changes

At years 1 and 2, respectively, 24 (73%) and 26 (79%) of patients had increases in computed syndesmophyte volume by the algorithm. Figure 1 provides examples of syndesmophyte growth from baseline to year 1 and year 2. The method detected new syndesmophytes in IDSs without syndesmophytes at baseline, growth of existing syndesmophytes, and increased extent of fusion in IDSs which were already bridged at baseline.

Table 1 provides the mean computed volume and height changes. The mean 2-year volume increase was more than twice the mean 1-year volume increase. The mean 2-year maximal height was about twice the mean 1-year maximal height. Compared to the syndesmophyte volume at baseline, the mean 2-year volume increase per patient (201 mm^3) represents an 18% increase. The results, when stratified according to standard mSASSS score at baseline, show that the computed method detected changes among patients with low or higher mSASSS scores, and that measured syndesmophyte changes were larger with higher mSASSS scores.

We previously noted that an increase in syndesmophyte volume per patient of more than 3% represented a change greater than measurement error.[22] From baseline to year 1, 18 patients (55%) had an increase larger than 3%. From baseline to year 2, 23 patients (70%) had an increase larger than 3%. Additionally, two patients in whom the algorithm detected no syndesmophytes in all 4 IDSs at baseline developed new syndesmophytes at year 1, and three patients did so at year 2. For these patients, the rate of change cannot be computed because their baseline was 0.

Longitudinal validity study

The cumulative probability plots (figures 2 and 3) show that computed volumes increased for most IDSs and patients. Although a small number of IDSs exhibited negative volume changes, 17% at year 1 and 7% at year 2 (figure 2), when summed per patient, the total volume changes were nearly all positive or null (figure 3). For computed maximal height, the majority of patients (85% at years 1 and 2) also had positive or null change (figure 3). As expected, for computed volumes and heights, the increases from baseline to year 2 were visibly larger than from baseline to year 1. The two curves are clearly distinguishable, especially when changes are summed per patient. For mSASSS (either adapted or standard), the differences between the two curves are less appreciable, whether per IDS or per patient, because most scores did not change at either year 1 or year 2 (see figures 2 and 3 and online supplementary figure S2).

We next examined the association of changes in computed syndesmophyte measurements with the main reader's ratings of changes on CT scan (figure 4). Volume changes computed by the algorithm were significantly higher in IDS that were rated by the physician as having a larger volume increase (p<0.0002 from baseline to year 1, and p<0.0001 from baseline to year 2). The majority of IDS rated 0 by the physician also had no computed volume change. Changes in maximal height computed by the algorithm were also significantly associated with the physician's scores, (p=0.04 from baseline to year 1, and p=0 0.003 from baseline to year 2). The majority of IDSs rated 0 by the physician also had no computed maximal height

change. Few IDSs were rated as exhibiting any height change: 16 (12%) from baseline to year 1 and 22 (17%) from baseline to year 2. No IDS was rated 2 (change greater than one-half the IDS height) from baseline to year 1, and the algorithm did not find any height increase of more than one-half an IDS height either. From baseline to year 2, two IDSs were rated 2. In good agreement, the algorithm estimated the corresponding height increases to be 47% and 63% of the IDS heights. Changes in computed volume and heights also increased significantly (p=0.013 for volumes at years 1 and 2; p=0.046 for heights at years 1 and 2) with changes rated by the second physician reader (see online supplementary figure S3).

Sensitivity to change

Sensitivity to change was computed for all measures on a per patient basis (table 2). MRI was not performed at year 2, and was not done on three patients at year 1. From baseline to years 1 and 2, computed volume was by far the measure that was most sensitive to change.

DISCUSSION

For evaluative measures, it is important to know that they are correlated with reference measures at one point in time, and also whether they accurately capture changes in patients' status over time.[29] We have presented the first longitudinal study of a fully quantitative method for measuring syndesmophyte volumes and heights in CT scans. The method was previously shown to have very good reliability and cross-sectional validity.[22] By demonstrating largely unidirectional changes over time and strong associations with an independent measure of change in syndesmophyte size, the computed measures also have good longitudinal validity. Additionally, the sensitivity to change of our CT-based measures was superior to the adapted mSASSS, which is the directly comparable radiographic measure, the standard mSASSS, and the ASspiMRI-c. The cumulative probability plots show that, from baseline to year 1, 73% of patients had a volume increase compared to 18% for the adapted mSASSS (mean of two readers). The higher sensitivity to change of our computed measures reflects both the fully quantitative nature of the method and the improved visualisation of syndesmophytes using CT. Exploiting the 3D imaging capability of CT, we were able to quantitate syndesmophytes along the entire vertebral body rim. Although MRI also images the spine tomographically, cortical bone is poorly visualised on MRI because its water content is similar to surrounding tissues. Scoring systems based on MRI are semiquantitative, which also may limit their sensitivity to change. Our study showed that computed heights were less sensitive to change than computed volumes. This is likely because we measured total volume in the IDSs but only maximal syndesmophyte height in each IDS.

The strengths of this study include prospective evaluations over 2 years with no losses to follow-up, testing of measured syndesmophyte volume and height, and comparison of sensitivity to change with measures based on plain radiographs and MRI scans. Although our cohort was modest in size, it was large enough to demonstrate construct validity and provide estimates of sensitivity to change with narrow confidence limits. We were limited to two measures of longitudinal construct validity, because other possible reference standards, such as radiographs and spinal flexibility measures, did not change appreciably over 2 years.

For all imaging modalities (CT, MRI and radiographs), the readers were not blinded to the time order. We recognise that for a clinical study of treatment efficacy in which the endpoint is based on readings by human readers, blinding to time order would be important, because in this case it is the ability of the human reader to identify change that is under examination and which provides the data for inferences. However, in the present work, we are using the results of human readings as a reference to validate the results of the CT-based quantitative measure. The time sequence and the knowledge that syndesmophyte growth is progressive provide the only available ground truth. Here, the critical aspect is that readers were blinded to the method's measurements. CT-based quantitative measures were also computed with known time sequence as they would be in a clinical setting. For validation, it is necessary to compare these changes in quantitative measures with similar changes in human readings, also with known time sequence.

It should be stressed that the SRMs reported here provide useful relative comparisons among these measures, but should not be extrapolated beyond this study. We computed SRMs on the logarithm of the changes, because of the high heteroskedasticity of the changes, particularly for the unbounded volume measurements. The volume changes covered an extremely large range (from 1.8 mm^3 to about 2000 mm^3), as larger syndesmophytes at baseline had larger increases over time. The distribution of mSASSS score changes was also heteroskedastic but to a considerably lesser degree, with changes that ranged from 1 to 10. To make these differences more comparable, we applied a logarithm transformation to constrain their ranges and improve their fit in the SRM.

We only examined the lumbar spine to minimise radiation exposure. The average equivalent dose per patient was 8.01 mSv with the CT protocol we used, compared with 2.59 mSv for lateral radiographs of the cervical and lumbar spine. The question of radiation exposure has to be considered in close relation with the information obtained. Although the radiation exposure of CT is substantially higher than radiographs, each CT scan provides complete information on syndesmophytes, and none needed to be discarded because of poor visualisation. The advantages of low radiation exposure need to be weighed by the usefulness of the information gathered by that exposure. Finally, CT scanner technology is constantly improving, and reduction in radiation doses may allow expansion of the method to larger segments of the spine with little or no increase in radiation exposure.[30][31] The current evidence suggests that this measure is a useful and sensitive method to follow the course of syndesmophyte growth in AS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 2007; 369:1379–90. [PubMed: 17448825]

2. Maksymowych WP, Landewé R, Conner-Spady B, et al. Serum matrix metalloproteinase 3 is an independent predictor of structural damage progression in patients with ankylosing spondylitis. Arthritis Rheum. 2007; 56:1846–53. [PubMed: 17530713]

3. van Tubergen A, Ramiro S, van der Heijde D, et al. Development of new syndesmophytes and bridges in ankylosing spondylitis and their predictors: a longitudinal study. Ann Rheum Dis. 2012; 71:518–23. [PubMed: 21989544]

4. Haroon N, Maksymowych WP, Rahman P, et al. Radiographic severity of ankylosing spondylitis is associated with polymorphism of the large multifunctional peptidase 2 gene in the Spondyloarthritis Research Consortium of Canada cohort. Arthritis Rheum. 2012; 64:1119–26. [PubMed: 22034108]

5. Appel H, Ruiz-Heiland G, Listing J, et al. Altered skeletal expression of sclerostin and its link to radiographic progression in ankylosing spondylitis. Arthritis Rheum. 2009; 60:3257–62. [PubMed: 19877044]

6. Daoussis D, Liossis SN, Solomou EE, et al. Evidence that Dkk-1 is dysfunctional in ankylosing spondylitis. Arthritis Rheum. 2010; 62:150–8. [PubMed: 20039407]

7. Ruiz-Heiland G, Appel H, Poddubnyy D, et al. High level of functional dickkopf-1 predicts protection from syndesmophyte formation in patients with ankylosing spondylitis. Ann Rheum Dis. 2012; 71:572–4. [PubMed: 22186710]

8. Ward MM, Hendrey MR, Malley JD, et al. Clinical and immunogenetic prognostic factors for radiographic severity in ankylosing spondylitis. Arthritis Rheum. 2009; 61:859–66. [PubMed: 19565552]

9. Baraliakos X, Listing J, Rudwaleit M, et al. The relationship between inflammation and new bone formation in patients with ankylosing spondylitis. Arthritis Res Ther. 2008; 10:R104. [PubMed: 18761747]

10. Maksymowych WP, Chiowchanwisawakit P, Clare T, et al. Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. Arthritis Rheum. 2009; 60:93–102. [PubMed: 19116919]

11. Maksymowych WP, Morency N, Conner-Spady B, et al. Suppression of inflammation and effects on new bone formation in ankylosing spondylitis: evidence for a window of opportunity in disease modification. Ann Rheum Dis. 2013; 72:23–8. [PubMed: 22562977]

12. van der Heijde D, Machado P, Braun J, et al. MRI inflammation at the vertebral unit only marginally predicts new syndesmophyte formation: a multilevel analysis in patients with ankylosing spondylitis. Ann Rheum Dis. 2012; 71:369–73. [PubMed: 21979001]

13. van der Heijde D, Landewé R, Baraliakos X, et al. Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum. 2008; 58:3063–70. [PubMed: 18821688]

14. van der Heijde D, Landewé R, Einstein S, et al. Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis Rheum. 2008; 58:1324–31. [PubMed: 18438853]

15. van der Heijde D, Salonen D, Weissman BN, et al. Assessment of radiographic progression in the spines of patients with ankylosing spondylitis treated with adalimumab for up to 2 years. Arthritis Res Ther. 2009; 11:R127. [PubMed: 19703304]

16. Creemers MC, Franssen MJ, van't Hof MA, et al. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis. 2005; 64:127–9. [PubMed: 15051621]

17. Spoorenberg A, de Vlam K, van der Linden S, et al. Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol. 2004; 31:125–32. [PubMed: 14705231]

18. Baraliakos X, Listing J, Brandt J, et al. Radiographic progression in patients with ankylosing spondylitis after 4 yrs of treatment with the anti-TNF-alpha antibody infliximab. Rheumatology. 2007; 46:1450–3. [PubMed: 17623745]

19. Lories RJ, Dougados M. Inflammation and ankylosis: still an enigmatic relationship in spondyloarthritis. Ann Rheum Dis. 2012; 71:317–18. [PubMed: 22315213]

20. Tan S, Yao J, Ward MM, et al. Computer aided evaluation of ankylosing spondylitis using high-resolution CT. IEEE Trans Med Imaging. 2008; 27:1252–67. [PubMed: 18779065]

21. Tan S, Yao J, Yao L, et al. Improved precision of syndesmophyte measurement for the evaluation of ankylosing spondylitis using CT: a phantom and patient study. Phys Med Biol. 2012; 57:4683–704. [PubMed: 22750760]

22. Tan S, Yao J, Flynn JA, et al. Quantitative measurement of syndesmophyte volume and height in ankylosing spondylitis using CT. Ann Rheum Dis. 2014; 73:544–50. [PubMed: 23345598]

23. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum. 1984; 27:361–8. [PubMed: 6231933]

24. MacKay K, Mack C, Brophy S, et al. The Bath Ankylosing Spondylitis Radiology Index (BASRI): a new, validated approach to disease assessment. Arthritis Rheum. 1998; 41:2263–70. [PubMed: 9870884]

25. Ibanez, L.; Schroeder, W.; Ng, L., et al. The ITK software guide. Clifton Park, NY: Kitware Inc; 2003.

26. Braun J, Baraliakos X, Golder W, et al. Analysing chronic spinal changes in ankylosing spondylitis: a systematic comparison of conventional x rays with magnetic resonance imaging using established and new scoring systems. Ann Rheum Dis. 2004; 63:1046–55. [PubMed: 15066862]

27. Hothorn T, Hornik K, van de Wiel MA, et al. Implementing a Class of Permutation tests: the coin Package. J Stat Software. 2008; 28:1–23.

28. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care. 1990; 28:632–42. [PubMed: 2366602]

29. Ward MM. Evaluative laboratory testing. Assessing tests that assess disease activity. Arthritis Rheum. 1995; 38:1555–63. [PubMed: 7488275]

30. May MS, Wüst W, Brand M, et al. Dose reduction in abdominal computed tomography: intraindividual comparison of image quality of full-dose standard and half-dose iterative reconstructions with dual-source computed tomography. Invest Radiol. 2011; 46:465–70. [PubMed: 21467948]

31. Kalra MK, Woisetschläger M, Dahlström N, et al. Radiation dose reduction with sinogram affirmed iterative reconstruction technique for abdominal computed tomography. J Comput Assist Tomogr. 2012; 36:339–46. [PubMed: 22592621]
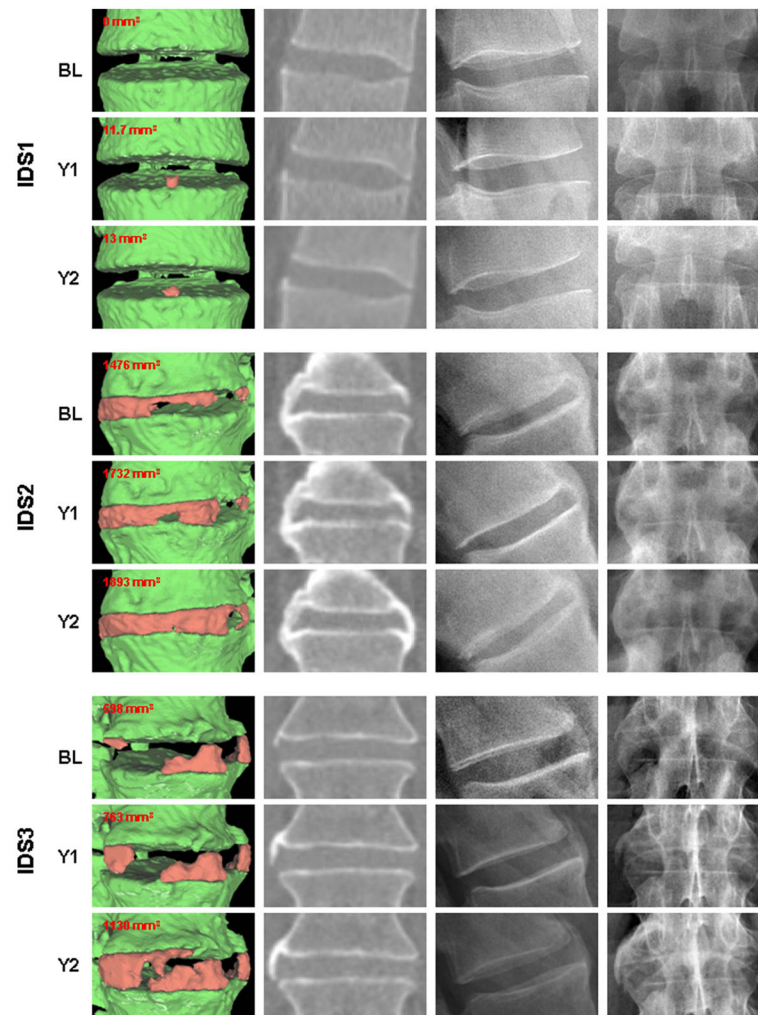
Figure 1.
Examples of syndesmophytes growth from baseline to year 2 detected by the computer algorithm. For each intervertebral disk space (IDS), from left to right, the first column shows three-dimensional surface reconstructions of the CT image with syndesmophytes in red and vertebral bodies in green, the second column shows a single slice from the CT scan, the third and fourth column show the corresponding lateral and antero-posterior (AP) radiographs, respectively. The three IDSs come from three different patients.
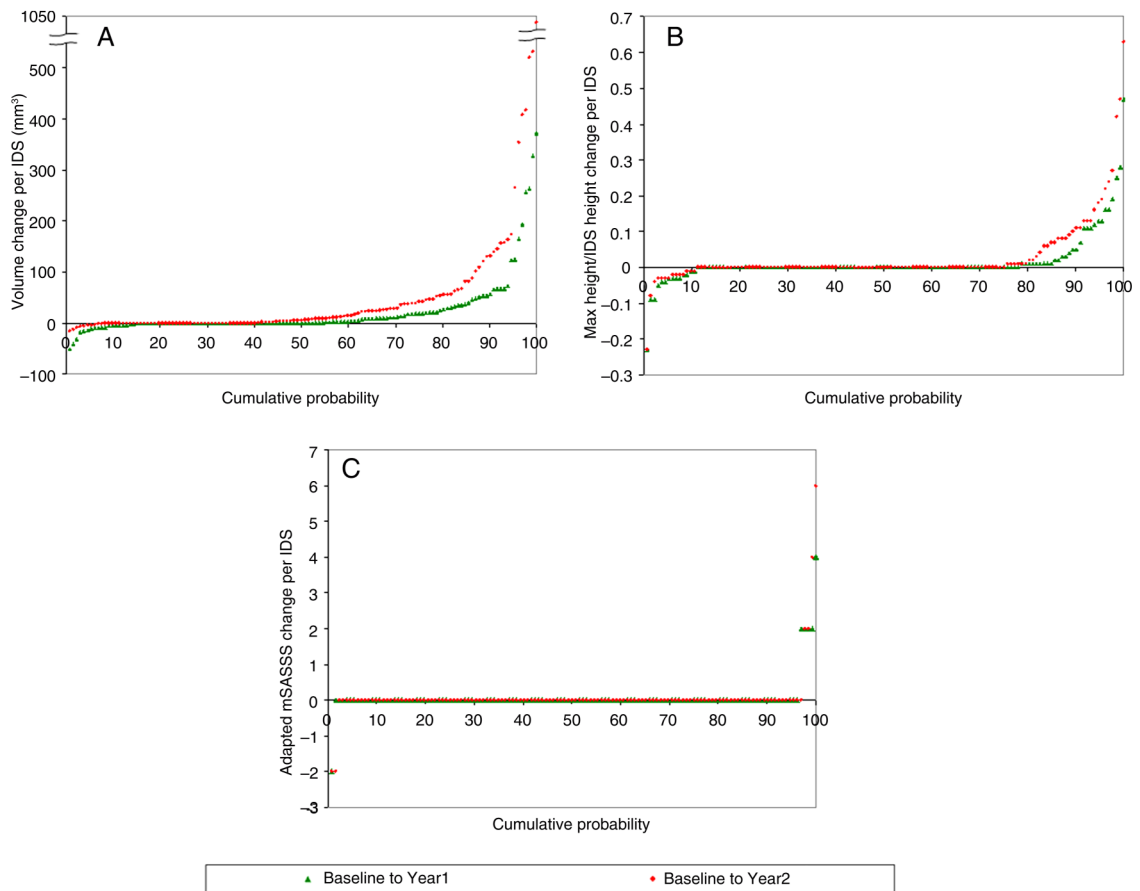
Figure 2.
Cumulative probability plots for changes per intervertebral disk space in (A) syndesmophyte computed volume, (B) syndesmophyte computed height and (C) adapted modified Stoke AS Spine Score (mSASSS) for one reader. The adapted mSASSS results for the other reader are similar (see online supplementary figure S1).
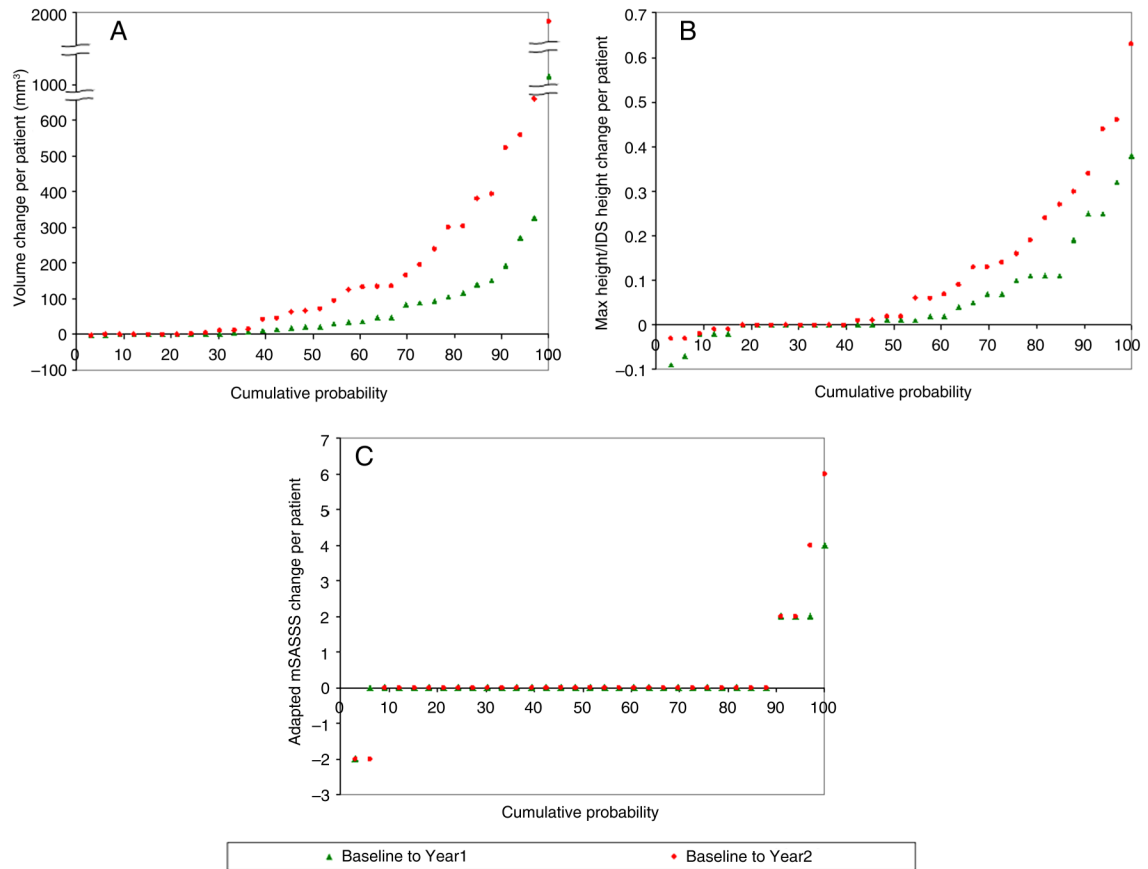
Figure 3.
Cumulative probability plots for changes per patient in (A) syndesmophyte computed volume, (B) syndesmophyte computed height and (C) adapted modified Stoke AS Spine Score (mSASSS). The adapted mSASSS results for one reader. The adapted mSASSS results for the other reader are similar (see online supplementary figure S1).
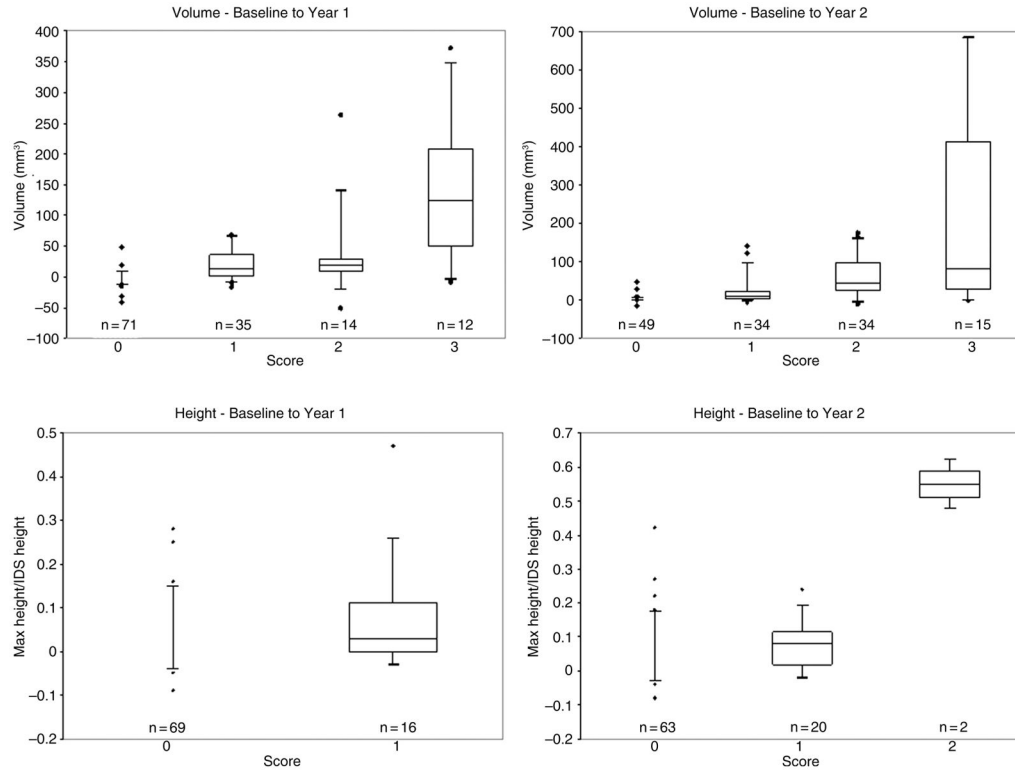
Figure 4.
Association of changes in computed syndesmophyte volume and height with the physician's scores of changes over time in the CT scans. Boxplots show the median and 25th and 75th quartiles, as well as the range and outliers. The first row of panels shows results for syndesmophyte volumes, and the second row shows results for heights. The first column of panels shows changes over 1 year, and the second shows changes over 2 years. N is the number of intervertebral disc spaces in each rating category. The 4-grade scoring system for volume change was: 0 for no change, 1 for small change, 2 for moderate change and 3 for large change. Height change was graded using a 3-grade scoring system: 0 for no change, 1 for an increase smaller than one-half the intervertebral disc height, and 2 for an increase larger than one-half the intervertebral disc height. Intervertebral disc spaces that were bridged at baseline were not scored for change in syndesmophyte height.

Table 1

Mean (SD) changes in syndesmophyte volume and height based on CT scans over 1 year and 2 years

	Per IDS				Per patient		
	n	Baseline to year 1	Baseline to year 2		n	Baseline to year 1	Baseline to year 2
Volume							
All	132	21.7 (60.1)	50.3 (127)		33	87.0 (186)	201 (366)
mSASSS 0–16	68	10.8 (37.7)	29.3 (86.7)		17	43.3 (89.3)	117 (204)
mSASSS 17–32	40	20.2 (43.7)	45.9 (72.0)		10	80.9 (79.0)	183 (180)
mSASSS 33–48	24	55.2 (108)	117 (237)		6	221 (397)	470 (745)
Height							
All	132	0.014 (0.068)	0.028 (0.094)		33	0.058 (0.11)	0.11 (0.16)
mSASSS 0–16	68	0.013 (0.080)	0.025 (0.11)		17	0.051 (0.12)	0.099 (0.18)
mSASSS 17–32	40	0.015 (0.053)	0.038 (0.087)		10	0.061 (0.10)	0.15 (0.16)
mSASSS 33–48	24	0.017 (0.049)	0.020 (0.050)		6	0.07 (0.10)	0.08 (0.12)

IDS, intervertebral disk space; mSASSS, modified Stoke AS Spine Score.

Table 2

Sensitivity to change for computed syndesmophyte volume and height from the CT scans, MRI using the ASspiMRI-c score, modified Stoke AS Spine Score (mSASSS) for each reader

	Baseline to year 1		Baseline to year 2	
	Mean	CI	Mean	CI
Log—computed volume	1.84	(1.80 to 1.87)	2.28	(2.24 to 2.32)
Log—computed height	1.22	(1.20 to 1.23)	1.09	(1.08 to 1.20)
Log—ASpiMRI-c	0.50	(0.49 to 0.52)	N/A	N/A
Log—adapted mSASSS, reader1	0.30	(0.29 to 0.31)	0.57	(0.56 to 0.58)
Log—adapted mSASSS, reader 2	0.28	(0.28 to 0.29)	0.27	(0.25 to 0.28)
Log—standard mSASSS	0.67	(0.65 to 0.68)	1.03	(1.01 to 1.05)