

**Title:** Automated measurement of size of spinal curve in population-based cohorts: validation of a method based on total body dual energy X-ray absorptiometry scans

**Authors:** Amir Jamaludin<sup>1</sup>, Jeremy Fairbank<sup>2</sup>, Ian Harding<sup>3</sup>, Timor Kadir<sup>4</sup>, Andrew Zisserman<sup>1</sup>, Emma M Clark<sup>5</sup>

<sup>1</sup>Department of Engineering Science, University of Oxford; <sup>2</sup>Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Science, University of Oxford; <sup>3</sup>North Bristol NHS Trust, Bristol;

<sup>4</sup>Plexalis Ltd, Oxford; <sup>5</sup>Musculoskeletal Research Unit, Bristol Medical School, University of Bristol.

**Corresponding author:** Prof Emma Clark

[Emma.Clark@bristol.ac.uk](mailto:Emma.Clark@bristol.ac.uk)

0117 414 7841

#### **ORCIDi**

Amir Jamaludin	0000-0002-0096-5625
Jeremy Fairbank	0000-0002-0050-874X
Ian Harding	
Timor Kadir	0000-0002-4414-2401
Andrew Zisserman	0000-0002-8945-8573
Emma M Clark	0000-0001-8332-9052

Manuscript word count = 3961

## **Abstract (246)**

**Background:** Scoliosis is spinal curvature that may progress to require surgical stabilisation. Risk factors for progression are little understood due to lack of population-based research, since radiographs cannot be performed on entire populations due to high levels of radiation. To help address this, we have previously developed and validated a method for quantification of spinal curvature from total body dual energy X-ray absorptiometry (DXA) scans. The purpose of this study was to automate this quantification of spinal curve size from DXA scans using machine learning techniques.

**Methods:** To develop the automation of curve size, we utilised manually annotated scans from 7298 participants from the Avon Longitudinal Study of Parents and Children (ALSPAC) at age 9 and 5122 at age 15. To validate the automation we assessed (1) agreement between manual vs automation using the Bland-Altman limits of agreement, (2) reliability by calculating the coefficient of variation, and (3) clinical validity by running the automation on 4969 non-annotated scans at age 18 to assess the associations with physical activity, body composition, adipocyte function and backpain compared to previous literature.

**Results:** The mean difference between manual vs automated readings was less than one degree, and 90.4% of manual vs automated readings fell within 10 degrees. The coefficient of variation was 25.4%. Clinical validation showed the expected relationships between curve size and physical activity, adipocyte function, height and weight.

**Conclusion:** We have developed a reasonably accurate and valid automated method for quantifying spinal curvature from DXA scans for research purposes.

## **Keywords (4 to 6)**

Scoliosis

ALSPAC

Bristol DXA Scoliosis method

Machine learning

## **Key messages**

- Understanding of the causes of spinal curve progression (scoliosis) is hampered by lack of prospective population-based studies, driven mainly by ethical concerns over performing spinal radiographs in healthy populations because of radiation exposure
- We have previously developed an automated method of identification of scoliosis from total body DXA scans based on machine learning
- We have now developed an extension of this automation to quantify size of spinal curve, which has reasonable agreement with manual quantification and appears clinically valid
- This new automation may revolutionise scoliosis research by facilitating exploitation of large population-based research cohorts through quick insertion of the scoliosis phenotype

## INTRODUCTION

Scoliosis is a three-dimensional torsional rotation of the spine, usually measured and defined from standing spinal radiographs as a lateral curvature of the spine  $\geq 10^\circ$ [1]. The most common form is adolescent-onset idiopathic scoliosis (AIS) occurring between age 10 years and skeletal maturity[2]. It is not always a benign structural abnormality, as severe AIS may result in early degenerative joint disease[3], negative body image[4], psychosocial disturbances[5] and sometimes requires interventions in the form of physiotherapy[6], casting[7], bracing[8] or surgery[9] to prevent further deterioration.

Scoliosis often remains relatively stable during slow periods of growth and after growth has stopped[10]. However, during accelerated growth phases such as puberty the size of the scoliotic curve often increases i.e. progresses[10]. Scoliosis is conventionally quantified using the Cobb angle[11] on spinal radiographs, which measures the greatest angle at a particular region of the vertebral column, when measured from the superior endplate of a superior vertebra to the inferior endplate of an inferior vertebra. There is a well-known inter- and intra-observer variability in its measurement of around 5 degrees[12], and changes in patient position or muscle fatigue can change the shape of the spine, so there has been debate regarding what constitutes a real change in curve magnitude. The consensus is that 5 degrees is likely to represent a real change[13] although some authors have suggested it may be as high as 10 degrees[14].

As curve size increases, thresholds of actual values become important for consideration of a change in management. There is good evidence that bracing can benefit curves in skeletally immature patients[8] when they measure 25-45 degrees; beyond 40 degrees consideration is given to surgery[15]. These thresholds are part of a general assessment of individual patients where the degree of curve is only one factor in the management decision. Other considerations include symptoms such as pain or cosmesis, curve location and remaining growth.

Once an adolescent has been diagnosed with scoliosis, repeat imaging by spinal radiographs is often performed, unless the curve is small and growth has ended. The younger patients with larger curves at first presentation may warrant surveillance with spinal radiographs every 3 months initially, increasing to between 6 and 12 months if the curve progresses slowly or not at all during growth. During periods of rapid growth e.g. the peri-menarchal phase, it may also be necessary to measure more frequently depending on the curve magnitude. Surveillance usually finishes in adolescents a

year after the end of growth and thereafter spinal imaging is only considered if symptomatic in adulthood.

To help streamline these clinical pathways, the main goal of further epidemiological analysis of scoliosis is to identify predictors of spinal curve progression. This would potentially allow generation of a clinical prediction tool to identify people at low risk of curve progression, for example, who would then require less rigorous monitoring, and therefore less exposure to ionizing radiation during growth.

However, our understanding of the causes of curve progression is hampered by lack of prospective population-based studies, driven mainly by the serious ethical concerns over performing spinal radiographs in healthy populations because of the radiation exposure, equivalent to an entire year's background radiation[16]. There is increasing concern for adolescent females, the majority of those with AIS, as they may have an increased risk of breast and uterine carcinoma with increased radiation exposure[17].

To address this, we have developed and validated a method for identification of spinal curvature and measurement of curve size using the low radiation imaging technique of total body dual energy X-ray absorptiometry (DXA) scans for research purposes[18]: the DXA Scoliosis Method (DSM). As previously published[18], this manual method is reliable (substantial agreement was seen with a kappa of 0.75), repeatable (95% of repeat measures were within 5°, and there was no change in interobserver variability as curve size increased) and accurate (comparison with the gold standard of using the Cobb method on standing spinal radiographs was as expected). The DSM also produces valid estimates of prevalence of scoliosis, with the expected gender ratio[18].

The aim of this new work was to develop and validate a fully automatic quantification of the size of the spinal curve from DXA scans using machine learning techniques based on the ideas developed in the SpineNet software[19]. The intended purpose of this automated curve size measurement is to exploit population-based cohorts for research purposes. This will allow identification of clinically relevant predictors of curve progression or stabilisation for future testing of predictive ability. The availability of large datasets and increasingly powerful computational resources has made the development of such techniques feasible. Automated analysis using computer vision or machine learning techniques is crucial, as studies involving big data require rapid, accurate and consistent analyses of medical scans.

## METHODS

### **Study population**

ALSPAC is a geographically-based UK cohort that recruited pregnant women residing in Avon (South-west England), with an expected date of delivery between 1 April 1991 and 31 December 1992[20, 21]. A total of 14 541 pregnancies were enrolled, with 14 062 children born; see [www.alspac.bris.ac.uk](http://www.alspac.bris.ac.uk) for more information. The study website contains details of all the data that are available through a fully searchable data dictionary and variable search tool available at <http://www.bris.ac.uk/alspac/researchers/our-data/>. ALSPAC is a longitudinal cohort study so all participants are invited to research clinics every few years and which point all attendees have a new DXA scan. This study is based on 7333 participants who had DXA scans at the age 9 research clinic, 5122 who had DXA scans at the age 15 research clinic, and 4969 who had DXA scans at the age 18 research clinic.

### **Overall study design**

The DXA images were performed by trained technicians using a Lunar Prodigy (GE Healthcare, Madison, WI) and were obtained in a standard supine manner. The manually annotated DXA scans from age 9 and age 15 were combined, then randomly split into a training (n=9600), a validation (n=1200) and a test set (n=1200), and the automation was developed using the training and validation set. To automatically assess body positioning error, the previously published algorithm was used[22]. To assess reliability the test set was used, and agreement between manual and automated measurement was assessed using the Bland-Altman limits of agreement, and the coefficient of variation was calculated. To assess clinical validity of the new automated output, a scaling factor was used to turn the machine output into a number representing size of the curve in degrees. The automation was then run on the non-annotated age 18 scans which did not take part in the training, validation or test procedures. The association with physical activity, body composition, adipocyte function and back pain was assessed and compared to previous literature.

### **Measurement of angle size using the manual DSM**

As previously published[18], curve size was measured on DXA scans by a trained researcher using a modified Ferguson method[23] since DXA images obtained from older generations of scanners are low-resolution and individual end plates cannot be identified, and so the standard Cobb method cannot be used. To perform the modified Ferguson method the standard Lunar encore software was

used to draw lines and measure angles. First, a normal spine line (NSL) was drawn through the centre of the spine level with the first rib attachment, down to the centre of the spine at L5. Next, the apex of the curve was identified as the centre of the spinal column most translated from the NSL. Lines were then drawn from the apex of the curve to the NSL at the point where the centre of the spinal column first touched the NSL on return from the apex. In double or triple curves, the spine did not always return to the NSL before the next curve started, so the centre of the spinal column at the judged point of inflection was used as the end of the curvature. Also as previously published[18], the precision of the manual method was assessed through repeated angle measurement by two operators. In terms of inter-rater variation, 95% of repeated measurements were within 4° for operator 1 and 5° for operator 2. For interobserver variability, the mean difference between the two operators was 2° (SD 2°, range 0°–6°, mean curve size 13°). Of the angle measurements carried out by the two operators, 95 % lay within 4.6° of each other. No change was seen in interobserver variability as curve size increased and a possibility of a decreasing variability with increasing curve size.

#### **Development of the new software algorithm to automatically measure angle size using the training and validation set**

The automated curvature measurement algorithm is described in [22]. The method starts with a Convolutional Neural Network to segment several body parts visible in the DXA scan which include the head, spine, pelvis, pelvic cavity, and both legs. A curve representing the midpoint of the spine is then extracted from only the spine segment. The spinal curve is height normalised to 209 pixels for the curvature measurement. Then, curvature is directly measured on the spinal curve via an integral based curvature estimator using digital shapes. In short, the curvature value is defined as the ratio between two areas for a given circle centred on the spinal curve:  $\frac{\text{Left Area of the Circle}}{\text{Right Area of the Circle}}$ . The optimum radius was determined by looking at the fit of our curvature values against the human annotated angle measurements on the validation set. The best radius for our use case, at least in our dataset, is found to be 33 pixels. When the spine is completely straight, the ratio between the two areas is 1. To scale the curvature so that it starts from 0, 1 can be deducted from the original ratio:  $\frac{\text{Left Area of the Circle}}{\text{Right Area of the Circle}} - 1$ . The curve of the spine itself is estimated from segmentation maps using a method previously described[22]. Briefly, the curvatures of the spine were calculated for each row in the DXA scan and the maximum curvature in a given scan was used as the main curvature value. A scaling factor was identified by dividing the mean manual angle measurement by the automated output, and assumed the association between manual and automation was linear.

## **Statistical methods**

All statistical analyses were performed by EC using Stata 15. Participants with missing data during validity analyses were excluded from the final analysis. Agreement between manual and automated outputs was assessed first using the DXA scan data (test set) from participants aged 9 and 15 where both outputs were available. Assessment of clinical validity was made using the automated output from DXA scan data at age 18.

### **Agreement between manual and automated output using the test set**

To assess agreement, the Bland-Altman method was used on the test set to show the magnitude and distribution of differences in angle size measured by the manual DSM method and the automation at age 15 after excluding those with body positioning error (n=285). The mean difference in degrees was calculated, and the percentage within 5 degrees and 10 degrees was identified. As there were only a small number of large curves, the coefficient of variation was calculated over the range of 6 to 18 degrees.

### **Assessment of discrepancies**

Images where there were large discrepancies between manual angle measurement and automated output were reviewed by clinicians (IH, JF and EC). In some cases, there were errors with the DXA scans such as missing lower limbs or head such that the scaling was imperfect. To address this, prior to measuring the angle or curvature, the ratio of the leg length to the overall scan height were used to automatically exclude the erroneous scans. The leg length is estimated via segmentations of the legs, and the standard leg-to-body ratio is 0.5 while erroneous scans were seen to have ratios lower than 0.25.

### **Validity of final automated model**

The final model was then run on the non-manually annotated age 18 DXA scans. Basic descriptives of spinal shape were calculated. To assess the validity of the automated quantification of curve size, linear regression techniques were used to assess the association between the continuous measure of curve size at age 18 (initially for the whole cohort and then separately for males and females, excluding those with body positioning error), and continuous variables chosen due to associations previously identified with the manual measure of scoliosis using ALSPAC[24-26]. These variables were (1) objective measures of physical activity using an actigraph[24]; (2) measures of body composition and adipocyte function[25]; and (3) intensity of back pain[26]. For regression analyses, all continuous variables were converted to T scores (by subtracting the mean and dividing by the

standard deviation from the study data) to allow comparison of size of effect if necessary. It was hypothesized that similar directions of association to those previously found[24-26] would be identified between these continuous variables and curve size measured by the automation at age 18.

*Objective measures of physical activity:* As previously published[24], physical activity was measured over a 7 day period when ALSPAC participants were 11 years old, using an actigraph (model WAM 7164, MTI). Actigraph output was categorised using the same method described previously[27], into the daily number of minutes of moderate/vigorous activity, light activity or sedentariness. Previously published analyses[24] showed those children who did more moderate/vigorous physical activity at age 11 were 30% less likely to have developed scoliosis (as identified by the manual DSM method) by age 15 years i.e. a negative association. To assess the validity of the automated quantification of curve size, linear regression was used to quantify the association between the daily number of minutes of moderate/vigorous activity, light activity or sedentariness and the size of curve measured by the automation at age 18. Models were adjusted for the other objective measures of physical activity, and then additionally adjusted for age at time of physical activity measurement, fat mass and lean mass measured by total body DXA. It was hypothesized that a negative association would be identified between moderate/vigorous physical activity at age 11 and curve size at age 18.

*Measures of body composition and adipocyte function:* As previously published[25], body composition data included height, body weight and body mass index (BMI) measured at age 15 calculated as weight (in kg) divided by height squared (in m<sup>2</sup>). Total body fat mass (kg) and total body lean mass (kg) was measured at age 15 from the total body DXA scans. Previously published analyses[25] showed that per standard deviation (SD) increase in body weight, BMI, fat mass or lean mass there was a reduced risk of scoliosis i.e. a negative association. To assess validity of the automated quantification of curve size, linear regression was used to quantify the association between body weight, BMI, fat mass and lean mass, and the size of curve measured by the automation at age 18. Models were adjusted for age, and then additionally adjusted for height and the other anthropometry (fat or lean mass). It was hypothesized that a negative association would be identified between measures of body composition at age 15 and curve size at age 18. There is a well-recognised negative association between height and size of scoliotic curve.

In addition, measures of serum leptin and adiponectin were used as measures of adipocyte function at age 10 years to assess validity of the automated curve size measure. As previously published[25],



an inverse association was seen between leptin and presence of scoliosis at age 15, and a positive association seen between adiponectin and scoliosis. To assess validity of the automated quantification of curve size, linear regression was used to quantify the association between adipocyte function and the size of curve measured by the automation at age 18. Models utilised log transformation of leptin and adiponectin, and were adjusted for age, and then additionally adjusted for fat mass, lean mass and the other hormone measure (leptin or adiponectin). It was hypothesized that similar directions of associations would be seen between leptin and adiponectin and curve size at age 18.

*Intensity of back pain:* As previously published[28], a structured pain questionnaire assembled from domains and scales taken from questionnaires previously validated in UK populations was administered to ALSPAC participants at age 18. Previously published analyses[26] showed a positive association between size of curve from the manual measurements of scoliosis at age 15 and intensity of pain i.e. a positive association. To assess validity of the automated quantification of curve size, linear regression was used to quantify the association between the size of curve measured by the automation at age 18 and pain intensity at age 18. Pain intensity variables available were intensity at the present time, worst pain intensity and average pain intensity. As before, models were adjusted for age and BMI, and then additionally adjusted for depression. It was hypothesized that a positive association would be identified between pain intensity and curve size at age 18.

Finally, to illustrate 'dose-response' relationships, the continuous curve size variable was categorised into no curve, 6-10 degrees, 10-20 degrees and >20 degrees and histograms plotted to show mean physical activity, body composition variables and back pain according to curve size.

## RESULTS

### Agreement between manual identification and automated scoliosis output at age 15

Figure 1 shows the Bland-Altman plot for the difference between the manual and automated outputs for quantification of curve size at age 15 after excluding body positioning error. The mean size difference between the manual and automated output was 0.9 degrees (95%CI 0.5 to 1.3 degrees). 61.3% of readings fell within 5 degrees and 90.4% within 10 degrees. From the figure, there was some indication of a larger discrepancy for larger curve sizes. The coefficient of variation was 25.4%.

### Validity of automated scoliosis output at age 18: association with physical activity

The machine learning technique identified 8.9% of females and 5.1% of males as having a spinal curve  $\geq 10$  degrees. 31/3680 (0.8%) had curves  $\geq 20$  degrees. As hypothesized, a negative association was identified between the number of minutes of moderate/vigorous physical activity at age 11 and scoliosis curve size measured by machine learning techniques at age 18 (see Figure 2A), even after adjustment for age, sedentariness, light activity, fat mass and lean mass: per SD increase in moderate/vigorous physical activity, scoliosis curve size reduced by 0.07SD, 95%CI (-0.11 to -0.04),  $P < 0.001$ . Similar sizes of effect were seen separately for males and females (see Table 1).

### Validity of automated scoliosis output at age 18: association with body composition and adipocyte function

As hypothesized, a positive association was identified between serum adiponectin at age 10 and scoliosis curve size measured by machine learning techniques at age 18, even after adjusting for age, fat mass, lean mass and serum leptin: per SD increase in serum adiponectin, scoliosis curve size increased by 0.06SD, 95%CI (0.01 to 0.10),  $P = 0.018$ . Similar sizes of effect were seen separately for males and females (see Table 2).

As hypothesized, a negative association was identified between height and size of curve measured by machine learning techniques at age 18 (see Figure 2B), even after adjusting for age: per SD increase in height, scoliosis curve size decreased by 0.08SD (95%CI -0.11 to -0.04),  $P < 0.001$ . This association was mainly driven by the strong negative association seen in females (see Table 2). Similarly, a negative association was identified between lean mass and scoliotic curve even after adjustment for age, height and fat mass, again mainly driven by females (see Table 2).

In females alone, the expected negative associations were seen between body weight or BMI and size of curve (see Table 2). However, the opposite direction of association was seen in males, where per SD increase in total body fat mass, scoliosis curve size measured by machine learning at age 18 increased by 0.145SD (95%CI 0.093 to 0.197),  $P<0.001$  after adjustment for age, height and lean mass.

#### Validity of automated scoliosis output at age 18: association with pain intensity

In females only, positive associations were identified between scoliosis curve size measured by machine learning techniques at age 18 and current pain intensity, even after adjustment for age, BMI and depression: per SD increased in pain intensity, curve size increased by 0.09 (95%CI 0.01 to 0.17),  $P=0.035$ . No associations were seen in males, although the majority of associations between curve size and pain intensity were in a positive direction for both males and females (see Table 3).

## DISCUSSION

We have developed a fully automated method of measurement of the size of scoliotic curve from total body DXA scans for research purposes. This machine learning technique has reasonable agreement with manual methods for curve size quantification, and appears valid. It can be easily run on 100,000 scans, taking approximately 50 minutes in total, and is likely to revolutionise scoliosis research and allow identification of clinical and genetic predictors of curve initiation and progression through enabling the scoliosis phenotype to quickly be inserted into population-based research cohorts around the world.

The agreement between the automated output and manual measures is reasonable for an automated measure, with the mean size difference of around one degree. As expected, there was a larger discrepancy for larger curve sizes because of the small number of large curves in the manually annotated dataset. Nonetheless, expected dose-response relationships between size of spinal curvature at age 18 and physical activity, adipocyte function, height and lean mass were identified, indicating the output is still valid for larger curves.

Unexpectedly the direction of association between body composition and size of spinal curvature as measured by our machine learning techniques differed for males and females. As far as we are aware our analysis of spinal curvature at age 18 years is the largest analysis undertaken in males, and the only at this time point, so may be a novel finding. Potential explanations include differing impacts of scoliosis on physical functioning in males and females leading to differences in body composition and distribution of fat and muscle mass; or possibly different pathophysiological processes causing scoliosis in males and females. Interestingly, the association between serum leptin measured at age 10 and size of curves in males at age 18 was also in the opposite direction to that for females indicating this may be a true association. Alternatively it may indicate our automation is less valid for males, but seems less probable given the other expected associations seen. Our machine learning technique is based on an initial step of spinal segmentation, not a 'black box' unstructured analysis, so we know the output is based on spinal shape. In addition, to further understand any problems with the automated analysis in male scans, a random sample of 60 DXA scans from males aged 18 were assessed independently by clinicians (IH, JF and EC) and there was general agreement between the automation and clinician output. As expected, discrepancies occurred in the smallest sized curves, but were not more likely in scans from overweight males.

The benefits of our fully automated model compared to manual annotation of DXA scans is the vast reduction in time required to look at each spinal image, with the consequent large reduction in financial costs. It takes on average 0.03 second to analyse a single scan (when batch processed). This has resulted in the first feasible and low-radiation process for quantification of spinal curves in large populations for research purposes. Other no-radiation techniques are available such as EOS machines, but their use is limited by lack of availability. It is increasingly difficult to justify regular conventional spinal radiography because of the radiation risks, especially to adolescent females[17].

A potential limitation is that both the manual method and the automation described in this paper have been developed on DXA scans performed on a Lunar Prodigy machine produced by GE Healthcare. However, the Lunar Prodigy is an older generation of scanner and newer scanners such as the iDXA can produce higher quality images which may improve the accuracy of the automation. Other DXA manufacturers are available, (machines produced by GE Healthcare and Hologic comprise the majority), and it is currently unknown how our automation will perform on such images, although we are currently in the process of testing it on Hologic images and outputs are encouraging [un-published data]. However, the intended use of our automation is for research purposes in population-based cohort studies where the serial images are taken on the same machines. We do not intend to use our automation on repeat scans in individuals taken on machines by different manufacturers.

We are now in a position to use this fully automated method to insert the scoliosis phenotype into population-based research cohorts with total body DXA scans around the globe and quantify curve size. This will facilitate well-powered studies of the risk factors for progression of spinal curves and is likely to produce a step-change in our understanding of this little-researched disease. Next steps for developing the automation will include looking at repeated measures on individuals over time to look at curve progression and to train the automation on this progression, to train the automation to detect direction of curve and anatomical site of curve. Further research is required to better understand the differences in body composition identified between males and females with scoliosis.

# TABLES

**Table 1:** Association between objectively measured physical activity at age 11 and size of scoliotic curve at age 18 measured by machine learning techniques, separately for females and males. Beta is for SD change in curve size per SD change in exposure variable.

<b>FEMALES</b>	<b>N=1557</b>	<b>N=1311</b>
	<b>Association with curve size, adjusted for other activity</b>	<b>Association with curve size, additionally adjusted for age, fat mass and lean mass</b>
	<b>Beta, (95%CI), P value</b>	<b>Beta, (95%CI), P value</b>
<b>Sedentariness</b>	-0.005 (-0.040 to 0.031), P=0.791	-0.001 (-0.038 to 0.038), P=0.996
<b>Light activity</b>	-0.002 (-0.038 to 0.034), P=0.910	0.005 (-0.033 to 0.043), P=0.802
<b>Moderate/vigorous activity</b>	<b>-0.074 (-0.122 to -0.027), P=0.002</b>	<b>-0.082 (-0.133 to -0.032), P=0.001</b>
<b>MALES</b>	<b>N=1249</b>	<b>N=1118</b>
<b>Sedentariness</b>	0.012 (-0.025 to 0.050), P=0.521	0.013 (-0.037 to 0.053), P=0.527
<b>Light activity</b>	-0.023 (-0.062 to 0.016), P=0.253	-0.020 (-0.061 to 0.022), P=0.353
<b>Moderate/vigorous activity</b>	<b>-0.077 (-0.127 to -0.026), P=0.003</b>	<b>-0.060 (-0.113 to -0.007), P=0.027</b>

**Table 2:** Association between (A) measures of leptin and adiponectin at age 10 years, and (B) measures of body composition at age 15 with size of scoliotic curve at age 18 measured by machine learning techniques, separately for males and females. Beta is for SD change in curve size per SD change in exposure variable.

<b>FEMALES</b>		
	Association with curve size, adjusted for age N=967 Beta, (95%CI), P value	Association with curve size, additionally adjusted for fat mass, lean mass and other hormone level N=966 Beta, (95%CI), P value
Leptin (ng/mL)	-0.055 (-0.118 to 0.008), P=0.089	-0.007 (-0.020 to 0.006), P=0.286
Adiponectin (µg/mL)	0.022 (0.006 to 0.039), P=0.008	0.020 (0.003 to 0.036), P=0.020
	Association with curve size, adjusted for age N=1550 Beta, (95%CI), P value	Association with curve size, additionally adjusted for height and the other fat or lean mass variable N=1550 Beta, (95%CI), P value
Height (cm)	-0.022 (-0.071 to 0.023), P=0.375	
Body weight (kg)	-0.068 (-0.117 to -0.020), P=0.006	
BMI (kg/m <sup>2</sup> )	-0.068 (-0.116 to -0.019), P=0.006	
Total body fat mass (kg)	-0.035 (-0.080 to 0.011), P=0.137	-0.035 (-0.080 to 0.011), P=0.131
Total body lean mass (kg)	-0.059 (-0.108 to -0.010), P=0.018	-0.033 (-0.069 to 0.004), P=0.079
<b>MALES</b>		
	Association with curve size, adjusted for age N=916 Beta, (95%CI), P value	Association with curve size, additionally adjusted for fat mass, lean mass and other hormone level N=915 Beta, (95%CI), P value
Leptin (ng/mL)	0.093 (0.030 to 0.156), P=0.004	0.008 (-0.013 to 0.028), P=0.458
Adiponectin (µg/mL)	0.007 (-0.018 to 0.031), P=0.587	0.008 (-0.017 to 0.032), P=0.528
	Association with curve size, adjusted for age N=1357 Beta, (95%CI), P value	Association with curve size, additionally adjusted for height and the other fat or lean mass variable N=1357 Beta, (95%CI), P value
Height (cm)	0.018 (-0.035 to 0.071), P=0.502	
Body weight (kg)	0.104 (0.052 to 0.156), P<0.001	
BMI (kg/m <sup>2</sup> )	0.114 (0.062 to 0.166), P<0.001	
Total body fat mass (kg)	0.145 (0.093 to 0.197), P<0.001	0.145 (0.093 to 0.197), P<0.001
Total body lean mass (kg)	-0.005 (-0.057 to 0.048), P=0.868	-0.023 (-0.059 to 0.012), P=0.203

**Table 3:** Association between measures of pain intensity and size of scoliotic curve at age 18 measured by machine learning techniques, separately for males and females.

	Association with curve size, adjusted for age and BMI Beta, (95%CI), P value	Association with curve size, additionally adjusted for depression Beta, (95%CI), P value
<b>FEMALES</b>	<b>N=626</b>	<b>N=517</b>
Pain intensity at the present time	<b>0.089 (0.008 to 0.169), P=0.030</b>	<b>0.090 (0.006 to 0.174), P=0.035</b>
Worst pain intensity in past 6 months	0.033 (-0.047 to 0.114), P=0.418	0.050 (-0.034 to 0.135), P=0.244
Average pain intensity in past 6 months	0.066 (-0.015 to m0.147), P=0.111	0.069 (-0.015 to 0.154), P=0.106
<b>MALES</b>	<b>N=450</b>	<b>N=336</b>
Pain intensity at the present time	-0.057 (-0.145 to 0.316), P=0.207	-0.070 (-0.167 to 0.026), P=0.152
Worst pain intensity in past 6 months	0.076 (-0.013 to 0.164), P=0.093	0.077 (-0.023 to 0.176), P=0.129
Average pain intensity in past 6 months	0.036 (-0.053 to 0.124), P=0.425	0.029 (-0.068 to 0.126), P=0.551



## FIGURE LEGENDS

**Fig 1:** Graph produced using the Bland-Altman method to show the magnitude and distribution of differences in angle size (n=285) estimated by the manual and automated methods (automation minus manual reading). The two horizontal lines are the 95% limits of agreement illustrating the limits within which 95% of differences between the measurement methods are likely to occur.

**Fig 2:** Graphs showing (A) the mean number of minutes of moderate/vigorous physical activity performed at age 10 per category of scoliosis curve size as measured by machine learning techniques at age 18, and (B) the mean height in cm per category of scoliosis curve size as measured by machine learning techniques at age 18. Error bars are standard deviations.

### **ETHICS APPROVAL**

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Parental consent and child's assent were obtained for all measurements made. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

### **AUTHOR CONTRIBUTION**

EC and JF planned the overall objective, whilst AJ, TK and AZ planned and conducted the automation. AJ and EC wrote the first draft. EC, JF and IH provided clinical input. EC is responsible for the overall content as the guarantor of the paper. All authors revised the paper critically for intellectual content and approved the final version. All authors agree to be accountable for the work and to ensure that any questions relating to the accuracy and integrity of the paper are investigated and properly resolved.

### **DATA AVAILABILITY**

The informed consent obtained from ALSPAC participants does not allow the data to be made freely available through any third party maintained public repository. However, data used for this submission can be made available on request to the ALSPAC Executive. The ALSPAC data management plan describes in detail the policy regarding data sharing, which is through a system of managed open access. Full instructions for applying for data access can be found here: <http://www.bristol.ac.uk/alspac/researchers/access/>. The ALSPAC study website contains details of all the data that are available (<http://www.bristol.ac.uk/alspac/researchers/our-data/>).

### **FUNDING**

The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website <http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>. This research was specifically funded by the British Scoliosis Research Foundation. Funding for authors based at the University of Oxford was via an EPSRC Programme Grant Seebibyte (EP/M013774/1).

### **ACKNOWLEDGEMENTS**

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. This publication is the work of the authors, and EC will serve as guarantor for the contents of this paper, which do not reflect the views of the ALSPAC executive.

### **CONFLICTS OF INTEREST**

Amir Jamaludin has no conflicts of interest  
Jeremy Fairbank has no conflicts of interest  
Ian Harding has no conflicts of interest  
Timor Kadir has no conflicts of interest  
Andrew Zisserman has no conflicts of interest  
Emma M Clark has no conflicts of interest

## REFERENCES

1. Kane, W.J., *Scoliosis prevalence: a call for statement of terms*. Clinical Orthopaedics, 1977. **126**: p. 43-46.
2. Weinstein, S.L., et al., *Adolescent Idiopathic Scoliosis*. Lancet, 2008. **371**: p. 1527-1537.
3. Miller, N.H., *Cause and natural history of AIS*. Orthopedic Clinic of North America, 1999. **30**(3): p. 343-352.
4. Weinstein, S.L. and I.V. Ponseti, *Idiopathic scoliosis: Long term follow-up and prognosis in untreated patients*. Journal of Bone and Joint Surgery - American Volume, 1981. **63**: p. 702-711.
5. Fowles, J.V., D.S. Drummond, and D.S. L'Ecuyer, *Untreated scoliosis in the adult*. Clinical Orthopaedics, 1978. **134**: p. 212-217.
6. Romano, M., et al., *Exercises for AIS*. Spine, 2013. **38**(14): p. E883-E893.
7. D'Astous, J.L. and J.O. Sanders, *Casting and traction treatment methods for scoliosis*. Orthop Clin North Am, 2007. **38**(4): p. 477-484.
8. Weinstein, S.L., et al., *Effects of bracing in adolescents with idiopathic scoliosis*. New England Journal of Medicine, 2013. **369**: p. 1512-1521.
9. Beauchamp, E.C., R.C.E. Anderson, and M.G. Vitale, *Modern surgical management of early onset and AIS*. Neurosurgery, 2019. **84**: p. 291-304.
10. Weinstein, S.L., *The natural history of adolescent idiopathic scoliosis*. J Pediatr Orthop, 2019. **39**(6): p. S44-S46.
11. Cobb, J., *Outline for the study of scoliosis*. Am Acad Orthop Surg Instr Course Lect, 1948. **5**: p. 261-275.
12. Morrissey, R.T., et al., *Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error*. Journal of Bone and Joint Surgery - American Volume, 1990. **72**(3): p. 320-327.
13. Nachemson, A.L. and L.E. Peterson, *Effectiveness of treatment with a brace in girls who have AIS: a prospective, controlled study based on data from the Brace Study of the SRS*. Journal of Bone and Joint Surgery Am, 1995. **77**: p. 815-822.
14. Carman, D.L., R.H. Brown, and J.G. Birch, *Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation*. J Bone Joint Surg Am, 1990. **72A**: p. 328-333.
15. Bettany-Saltikov, J., et al., *Management of spinal deformities and evidence of treatment effectiveness*. Open Orthop J, 2017. **11**: p. 1521-1547.
16. Watson, S.J., et al., *Ionising radiation exposure of the UK population*. 2005 Review, 2005. **HPA-RPD-001**.
17. Brenner, A., et al., *Radiation risks of uterine cancer in atomic bomb survivors 1958-2009*. JNCI Cancer Spectrum, 2019. **2**: p. pky081.
18. Taylor, H.J., et al., *Identifying scoliosis in population-based cohorts: Development and validation of a novel method based on total body DXA scans*. Calcified Tissue International, 2013. **92**(6): p. 539-547.
19. Jamaludin, A., T. Kadir, and A. Zisserman, *SpineNet: Automated classification and evidence visualization in spinal MRIs*. Med Image Anal, 2017. **41**: p. 63-73.
20. Boyd, A., et al., *Cohort profile: The 'Children of the 90s' - the index offspring of the Avon Longitudinal Study of Parents and Children*. International Journal of Epidemiology, 2013. **42**(1): p. 111-127.
21. Fraser, A., et al., *Cohort profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort*. International Journal of Epidemiology, 2013. **42**: p. 97-110.
22. Jamaludin, A., et al., *Predicting scoliosis in DXA scans using intermediate representations*. 5th International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging. Post-proceedings, Springer's Lecture Notes in Computer Science, 2018. **11397**(15-27): p. DOI 10.1007/978-3-030-13736-6.

23. Ferguson, A.B., *The study and treatment of scoliosis*. Southern Medical Journal, 1930. **23**(2): p. 116-120.
24. Tobias, J.H., et al., *Association between physical activity and scoliosis: a prospective cohort study*. International Journal of Epidemiology, 2018: p. doi.org/10.1093/ije/dyy268.
25. Clark, E.M., et al., *Association between components of body composition and scoliosis: A prospective cohort study reporting differences identifiable before the onset of scoliosis*. Journal of Bone and Mineral Research, 2014. **29**(8): p. 1729-1736.
26. Clark, E.M., J.H. Tobias, and J. Fairbank, *The impact of small spinal curves in adolescents that have not presented to secondary care: a population-based cohort study*. Spine, 2016. **41**(10): p. E611-E617.
27. Tobias, J.H., et al., *Habitual levels of physical activity influence bone mass in 11-year-old children from the UK: Findings from a large population-based cohort*. Journal of Bone and Mineral Research, 2007. **22**(1): p. 101-109.
28. Deere, K., et al., *Obesity is a risk factor for musculoskeletal pain in adolescents: Findings from a population-based birth cohort*. Pain, 2012. **153**: p. 1932-1938.