

Aging of the Spine: Characterizing genetic and physiological determinants of spinal curvature

Frances M. Wang¹, J. Graham Ruby¹, Anurag Sethi¹, Matthew Veras¹, Natalie Telis¹, Eugene Melamud^{1*}

Affiliations: ¹Calico Life Sciences LLC

*Corresponding Author Email: eugene@calicolabs.com

Running Title: Characterizing physiology and genetics of spinal curvature.

Abstract

Increased spinal curvature is one of the most recognizable aging traits in the human population. However, despite high prevalence, the etiology of this condition remains poorly understood. To gain better insight into the physiological, biochemical, and genetic risk factors involved, we developed a novel machine learning method to automatically derive thoracic kyphosis and lumbar lordosis angles from dual-energy X-ray absorptiometry (DXA) scans in the UK Biobank Imaging cohort. In 41,212 participants, we find that on average males and females gain 2.42° kyphotic and 1.48° lordotic angle per decade of life. Increased spinal curvature was strongly associated with decreased muscle mass and bone mineral density. Adiposity had opposing associations, with decreased kyphosis and increased lordosis. To gain further insight into the molecular mechanisms involved, we carried out a genome-wide association study and identified several risk loci associated with both traits. Using Mendelian randomization, we further show that genes fundamental to the maintenance of musculoskeletal function (COL11A1, PTHLH, ETFA, TWIST1) and cellular homeostasis such as RNA transcription and DNA repair (RAD9A, MMS22L, HIF1A, RAB28) are likely involved in increased spinal curvature.

Keywords: Kyphosis, Lordosis, Spine Curvature, Skeletomuscular Aging, Cobb Angles, Machine Learning, GWAS

Introduction

Increased spinal curvature is a frequently observed phenomenon during aging in humans and in other species.¹⁻³ Drift to higher thoracic kyphosis and lumbar lordotic angles starts in early ages, is gradual, and takes many decades to develop.^{4,5} These physiological declines are generally considered to be non-pathological unless kyphosis becomes so severe as to exceed more than 40° (hyperkyphosis), or more than 30° in lordosis regions (hyperlordosis).⁴ Hyperkyphosis is known to have a significant heritable component and affects females more than males.^{6,7,8} On the other hand, lumbar hyperlordosis is an understudied condition.⁹

At the extremes, acute injuries, disk fractures, and degenerative joint disorders are believed to be the main drivers of spinal hypercurvature.¹⁰ In the general population, increases in spine curvature develop largely in the absence of severe events. Prevalent aging conditions such as loss of bone mineral density, decline in muscle strength, and metabolic disorders have been suggested to play an important role.^{8,11} Over time, continued biomechanical stresses on the spine and aging-related declines in the repair capacity of cartilage and bone tissues can lead to the development of vertebral fractures, increased back pain, and consequently reduced mobility.^{12,5} Cumulative adverse consequences are an increased chance of falls, reduced health span, and increased mortality.^{13,14}

While the negative consequences of an aging spine on health outcomes are clear, the genetic risk factors and underlying molecular mechanisms of this

phenomenon remain poorly understood. Recent advances in the UK Biobank (UKBB) large-scale prospective imaging study create an opportunity to examine changes to spine curvature with age in a general population.¹⁵ We capitalized on this resource by developing machine learning models to assess spine curvature in an automated fashion across 41,212 dual-energy X-ray absorptiometry (DXA) scans. Using a fully automated machine learning (ML) pipeline, we were able to measure Cobb angles in both thoracic and lumbar regions, thus providing us with an opportunity to examine shared and unique risk factors in different regions of the spine.

Our analysis extends to the study of underlying pathological conditions, physiological, blood biochemistry, and genetic risk factors. We assessed contributions to spine angle from 415 diseases, 15 musculoskeletal traits, 59 biomarkers, and all common genetic variants. We further carry out Mendelian randomization (MR) to test for potential causal involvement of genes within these loci with changes in spinal curvature. To our knowledge, this is the first study to identify genetic risk factors involved in the development of kypholordotic angles.

Results

Cobb angle annotation by machine learning

We developed ML models to mimic a well-established Cobb angle measurement technique used in radiological practice.¹⁶ Briefly: a U-net segmentation model was trained to label the spine body, and a previously published¹⁷ object-detection model was

used to identify appropriate regions for Cobb angle estimation. Training and test data from those models were separate from the Cobb angle preliminary test and validation sets described below.

Cobb angles were estimated between tangent lines drawn through appropriate spine regions using Deming regression. In lateral DXA scan images, human annotators measured two sets of Cobb angles across 320 images: 4 independent annotators measuring T12-T5 kyphotic angle and 3 independent annotators measuring L4-L1 lordotic angle (**Figure 1A**). These images and their annotations were randomly split into a preliminary testing dataset (120 images), used to calibrate parameters of the ML & regression pipeline described above, and a validation dataset (200 images), used for final evaluation of each model's performance. There was a high degree of concordance across the human annotators ($r=0.66-0.91$, **Supplementary Figure 1**).

The ML-based estimation algorithms for both kyphosis and lordosis performed well on the 200-image validation dataset, closely matching the results of human annotators with high correlation values ($r=0.87$ and 0.86 for the kyphosis and lordosis models, respectively) (**Figure 1B and 1C**). The lordosis model has slightly lower accuracy, likely due to greater variability in the ability to define the location of the L4 and L5 junctions.

Epidemiology of Kyphosis and Lordosis

Among the 41,212 UKBB participants included in this study, 48.3% were male (**Supplementary Table 1**). The median age of DXA imaging subjects was 64

[interquartile interval (IQI) 58, 70] years (**Supplementary Table 1**). A total of 3,702 (9.0%) participants had hyperkyphosis (angle $<40^\circ$), and 3,109 (7.5%) had hyperlordosis (angle $<30^\circ$), of which 1,057 (2.6%) participants had both. Thoracic kyphosis Cobb angles were strongly correlated to lumbar lordosis Cobb angles ($r=0.49$, $p\text{-value}<0.001$, **Supplementary Fig. 2**). Kypholordotic angles were systematically higher in females compared to males at all ages (**Fig. 2A**). Both kyphotic and lordotic Cobb angles increased with age, at approximately the same rate: 2.28° (2.16° , 2.39°) of kyphotic angle and 1.32° (1.19° , 1.44°) of lordotic angle per decade of life (not adjusted for other factors).

Since both angles increase with age, we asked if they also increased as a function of each other. In the multivariable model, adjusted for age and sex, the kyphotic angle increased an additional $\sim 0.4^\circ$ for every degree of lordotic angle (Supplementary Figure 4). Consequently, with the increase of kypholordotic curvature with age, there was a graded increase in the prevalence of hyperkyphosis and hyperlordosis with age (**Fig. 2B**). By age 75+, the prevalence of hyperkyphosis and hyperlordosis was 15.4% and 10.4%, respectively. Participants with hyperkyphosis were 5.22 (4.89, 5.58) times more likely to also have hyperlordosis than those without hyperkyphosis and, relatedly, participants with hyperlordosis were 4.90 (4.61, 5.21) times more likely to also have hyperkyphosis. Those with higher observed kyphosis and lordosis angles tended to be older (**Fig. 2C**).

Disease Associations with Spine Curvature

Various pathological pre-existing conditions were associated with increased spine angles (**Fig. 3A**). First, among the strongest associations for kypholordotic angles were previous diagnoses of intervertebral disc disorders such as disc compression disorders (M41, M47, M51, G55) and osteoporosis (M81). For example, prior diagnosis of multiple sclerosis (MS) was highly associated with increased kyphotic angle by 2.24° (0.83° , 3.64°), potentially due to the effect of MS on posture.¹⁸ We also observed that a variety of gastric conditions including esophagitis, esophageal disease, and gastro-esophageal reflux disease (GERD) [K21, K22, K20] and, perhaps relatedly, diaphragmatic hernia (K44) were associated with increased kyphotic angles, also possibly due to the association of kyphosis with posture, reduced back muscle strength, and increased intra-abdominal pressure.¹⁹ Interestingly, we observed that metabolic disorders, such as diabetes, gout, and hyperlipidemia (E11, E14, M10, E78), were associated with decreased spine curvature. This effect was generally more pronounced in females, where diabetes was linked to reduction of kyphotic angle by 1.50° (0.71° , 2.30°) versus 0.63° (0.06° , 1.19°) in males (**Supplementary Fig. 3**). The association with these disorders with kypholordotic angles tended to be linear (**Supplementary Fig. 4**). More comorbid conditions were associated with kyphotic angles than lordotic angles, likely in part due to a higher precision in our machine learning-based estimates of kyphotic angles.

In serum biochemistry data, we observed that elevated levels of plasma alkaline phosphatase (ALP) were associated with significant increases in both kyphosis and

lordosis angles, supporting established literature that osteoporosis is an important contributing factor to increased spine curvature (**Fig. 3B**). Similar to the inverse associations we found between kypholordotic angles and metabolic comorbidities, in biochemistry data, elevation of metabolic biomarkers (hemoglobin A1C [HbA1c], urate, and triglycerides) were associated with significant reductions in kyphotic angles. This inverse association between biomarkers of metabolic disorders and kyphosis was consistent for both sexes (**Supplementary Fig. 5**).

Association of Musculoskeletal Traits with Spine Curvature

To evaluate how aging of the musculoskeletal system contributes to increased spinal curvature, we performed a phenotypic association scan with muscle, adiposity, and bone traits across the entire skeletal system in UKBB. Significant associations, adjusted for multiple hypothesis correction, are shown in **Fig. 4A**. We observed that bone mineral density (BMD) measures across the entire skeletal system were strongly associated with spine curvature. Both kyphotic and lordotic angles were consistently larger in those with lower BMD. For example, the kyphotic angle increased by 0.39° (0.31° , 0.47°), and the lordotic angle increased by 0.66° (0.58° , 0.73°) per standard deviation loss of arm BMD. This directional association was mostly consistent for both males and females across skeletal sites, with associations stronger in females than males (**Supplementary Fig. 6**).

Among muscle measurements, lower muscle mass in both the legs and the arms was also associated with increased spinal curvature, indicating the important role

skeletomuscular factors play in the development of pathology (**Fig. 4A**). On average angle increases by 0.91° (0.82° , 0.99°) kyphosis and 1.44° (1.36° , 1.52°) lordosis per standard deviation loss of arm muscle mass (**Supplementary Fig. 6**). Decreased grip strength was also strongly associated with increased spine angles, but association may be due to collinearity with muscle mass.

Interestingly, we observed that fat composition measures showed a differential effect between lordosis and kyphosis. Higher adiposity (total and visceral fat) was linked with increased lordotic angles but reduced kyphotic angles. The same differential effect was observed with body mass index (BMI) and was consistent for males and females (**Supplementary Fig. 4**).

Since physiological measurements are strongly associated and non-linearly dependent on each other, to identify most important predictors of spine curvature we used gradient boosting algorithm (GBM)²⁰. While predictors of angles at both sites were similar, their relative contribution (as measured by variable importance) was different (**Fig. 4B**). Consistent with single trait analysis, the top independent predictors of kyphosis were age, muscle mass, BMD, grip strength, and sex. Independent predictors of lordosis were similar to those for kyphosis, but sex, height, and BMI were stronger predictors of lordosis.

Genetic Risk Factors of Spine Curvature

Leveraging genotyping information, we next evaluated genetic risk factors associated with the development of spinal curvature. We carried out a genome wide association study (GWAS) for both kyphosis and lordosis angles within the participants

of Caucasian origin (n=33,413). Both traits had significant heritability h^2 (kyphosis: 0.34 [0.30, 0.38], lordosis: 0.26 [0.22, 0.30]). We found that there is a strong genetic correlation between kyphotic and lordotic angles ($R_g = 0.71$, $p\text{-value}=1.58e-48$) indicating that common genetic variation contributes to both traits.

The GWASes identified 8 genome-wide significant loci associated with kyphosis and 6 genome-wide significant loci associated with lordosis (**Supplementary Fig. 9**). Since we observed a strong genetic correlation between kyphosis and lordosis, to increase the statistical power, we carried out multi-trait analysis of GWAS (MTAG) to jointly model these traits.²¹ The results of MTAG were generally consistent with individual trait GWAS (**Fig. 5A**). Given greater power and precision of MTAG, we focus our downstream analyses of the MTAG summary statistics in the remaining sections.

Genetic Correlation

To examine if the genetic variation linked to spine angles shares genetics with other traits, we calculated genetic correlations with 10,170 and 10,138 GWAS catalog phenotypes for kyphosis and lordosis, respectively (**Fig. 5B, Supplementary Table 2**). Top traits associated with kyphosis were measures of muscle mass (appendicular lean mass), bone mineral density (heel, total body), and diseases related to the esophagus (esophageal disorder, Barrett's esophagus, hiatal hernia, GERD). The top genetic correlations of lordosis were similar to that of kyphosis but with stronger genetic associations with body type measures including height, BMI, waist circumference (**Supplementary Table 3**). In line with epidemiological association with GERD diagnosis, we observed a strong genetic correlation of both kyphosis and lordosis with

esophageal diseases and omeprazole usage (R_g between 0.10-0.25, all p-values < 1E-3), further suggesting that esophageal diseases share genetic risk factors with spine curvature traits.

To assess the similarity of demographic-adjusted genetic and phenotypic correlations, we computed an all-by-all correlation matrix for selected musculoskeletal traits and displayed it in a single heatmap (**Fig. 5C**). While most traits showed strong phenotypic correlation with one another, genetic correlations were less statistically powered and therefore more sparse. As expected, strong interdependence between all musculoskeletal traits can be observed with shared genetic origins. A mechanistic understanding of how these traits contribute to the development of spine curvature requires a deeper understanding of the loci involved.

Gene Prioritization

To better understand genetic risk factors we carried out two types of analysis (i) fine-mapping followed by colocalization with Genotype-Tissue Expression (GTEx) RNA splicing and RNA expression quantitative trait loci (eQTLs) (ii) MR with GTEx RNA expression as mediator of Cobb angle outcomes. Combining these multiple lines of support, we were able to prioritize genes with high likelihood of being involved in development of kypholordotic angles.

Fine-mapping identified 12 unique variants associated with kypholordotic angles (7 with kyphosis and 8 with lordosis) (**Table 1, Supplementary Table 4**). Of these, there were 3 overlapping fine-mapped lead single nucleotide polymorphisms (SNPs) shared by kyphosis and lordosis: rs881926 near SFRP4, rs9651832 near CCDC91, and

rs3783820 near SLC38A6. Variation near RAB28 and MMS22L genes showed a stronger association with kyphosis, while variation near COL11A1, OTOG, RAD9A, and ETFA genes showed a stronger association with lordosis. Full fine-mapping results can be found in **Supplementary Table 5 and 6**. Previous GWAS studies have implicated SNPs in these loci with multiple musculoskeletal traits such as height, heel bone mineral density, appendicular lean body mass, various measures of lung function, and spinal stenosis (**Supplementary Fig. 10**).

Using variants identified in fine-mapping with either lordosis or kyphosis, we ran colocalization analysis of Cobb angles with RNA expression data (eQTLs and splicing quantitative trait loci [sQTL]) across multiple tissues profiled in GTEx v8 (**Supplementary Table 7**). (See Methods). Although, for some genes we observed colocalization with both traits, overall colocalization found more association with lordosis (**Figure 5D**). Some of the fine-mapped SNPs, colocalized with RNA expression of multiple genes within the same loci. For example, we observed colocalization of kyphotic and lordotic angles with both PTHLH and CCDC9 on the chromosome 12 locus, along with both H1F1A and L3HYPDH on chromosome 14.

We followed up colocalization analysis with MR. In MR, genetic variants (instruments) are used to estimate the causal effect of an exposure (RNA expression) on an outcome (spinal curvature). The detailed description of methodology and its limitations can be found in the Methods. Given correlated RNAseq expression patterns observed in colocalization analysis, we can not rule out the possibility of horizontal pleiotropy, but we did find support for the potential mediation effect of multiple genes. We found that decreased expression of RAD9A (multiple tissues), ETFA (skeletal

muscle), COL11A (adipose tissue), ISL2 (esophagus), and increased expression of RNASEH2C (brain), and CCDC91/PTHLH and SCRAPER (multiple tissues) were significantly associated with increased Cobb angles (**Supplementary Table 8 and 9, Fig. 5D**).

Following MR-STROBE guidelines²², we evaluated sensitivity of our MR results. We repeated analyses using MR-Egger and median-weighted MR methods, and found that effect sizes were concordant across different estimation methods. MR-Egger didn't find that associations were statistically significant, but this was expected, given lower statistical power of this methodology. When multiple instruments were available, to avoid estimation bias we tested for heterogeneity of instruments. We did not find significant heterogeneity across instruments.

We summarized our findings in **Figure 5D**. Overall, a complex picture of genetic risk factors emerges, involving genes expressed in multiple tissues with fundamental roles in maintenance of tissues and cellular homeostasis. This included genes such as: COL11A1 (maintenance of connective tissue), RAB28 (intracellular trafficking), MMS22L (DNA double-strand break repair), TWIST1 (transcription factor involved in skeletal-vascular development), EPDR1 (glycoprotein related to neurogenesis), PLEKHA7 (key adherens junction component), RAD9A (DNA double-strand break repair), PTHLH (parathyroid like hormone), HIF1A (hypoxia response including angiogenesis and glycolysis), and ETFA (mitochondrial fatty acid beta-oxidation). Many genes also displayed non-specific tissue expression patterns.

Discussion

Increased spinal curvature is a universally recognized aging trait; however, the genetic and physiological etiology of this trait has not been well-characterized. In this study, we leveraged machine learning models to automatically quantify kypholordotic angles for 40,224 community-dwelling middle-aged to older adults in UKBB. Physiology, serum biochemistry, and genetics data were analyzed to build an integrated epidemiological and biological understanding of spinal curvature.

We observed that increased curvature in both the lumbar and thoracic spine were strongly intercorrelated and age-dependent (2.4° and 1.5° per decade of life). Previous investigations of the effect of aging on spine angles in males and females did not paint a consistent picture of sex-specific differences in kypholordotic angle and how they change with age.^{23,24} In our analysis, we see a consistent aging effect in both males and females, but with a different baseline of higher kypholordotic angles in females by the age of 40. Although we do not have measurements at earlier ages, likely, significant differences in spine curvature arise much earlier in life or due to developmental differences between sexes. While our cross-sectional analysis of aging effects could be biased by characteristics of the cohort; they are consistent with previous longitudinal assessment of aging effects.²⁵

As the population ages, both hyperkyphosis and hyperlordosis increased, reaching a substantial prevalence of 15.4% and 10.4%, respectively, by age 75 in the study. While the prevalence and assessment of hyperkyphosis has been well-characterized, the prevalence of hyperlordosis and its cause is limited.^{26,27} Previous

epidemiological studies of hyperkyphosis have implicated several potential pre-existing conditions, such as vertebral fractures and osteoporosis, that could contribute to the development of this pathology.²⁶ We have conducted a systematic evaluation of all common pre-existing conditions associated with Cobb angles. The majority of pre-existing conditions associations (multiple sclerosis, osteoporosis, and gastroesophageal diseases) were the same between kyphosis and lordosis, matching in both direction and effect sizes. In our analysis of biochemical factors, we likewise observed similar correspondence between kyphosis and lordosis. These results lead us to conclude that the increase in spinal curvature at two sites is driven by a similar set of underlying factors.

Our analysis of physiological measures suggested that the loss of muscle mass (and relatedly muscle strength) and decreased bone mineral density were the factors most strongly associated with increased spinal curvature. The slope of age-dependent angle increase was approximately the same for both males and females, even though females on average have lower muscle mass, lower BMD, and higher prevalence of osteoporosis. This suggests that the aging effect is not due to baseline differences in BMD and muscle measures and further highlights the important, complex relationship between spinal curvature with muscle mass and BMD with age.^{8,28–30}

Examining the effect of body adiposity measures on spine angles, we were surprised to find that increased adiposity (total fat mass, visceral fat, and BMI) was associated with decreased kyphotic angles, but increased lordotic angles. Among physiological measures considered, adiposity was the only measure that showed this opposing relationship between the two traits. This effect remained after adjusting for

age, sex, and respective spine angles. The increased lumbar lordosis may be due to the biomechanics of the spine, where the lumbar region compensates for increased adiposity in the trunk by moving the spine forward to adjust the center of gravity.³¹ Alternatively, decreased kyphotic angles were linked to a higher prevalence of metabolic disorders (type 2 diabetes, gout, elevated glucose, and HbA1c), but less so with lordotic angles. It is possible that metabolic disorders can result in an incidental increase in BMD, perhaps strengthening the upper spine and thus decreasing kyphotic angle, while more abdominal adiposity can increase the load on the lower spine.^{31–33}

Genetic analyses elucidated the complex origin of spine curvature traits. To our knowledge, this is the first GWAS of kyphosis and lordosis published to-date. We found that angles at both sites have substantial heritability (25-35%) and are strongly genetically correlated, indicating that common genetic drivers influence angles at both sites. Genetic correlation with other traits was consistent with phenotypic associations, with appendicular lean muscle mass and BMD among the strongest correlated traits. Interestingly, in both phenotypic and genetic associations, we observed an anti-correlation of height with spine angles (higher kyphosis/lordosis associated with smaller stature/height), but it was not clear to us if this was a trivial observation. However, this seems to indicate that there is at least in part shared genetics that determine these skeletally-related traits.

We also observed a strong genetic correlation of increased spine curvature with the development of gastroesophageal disease (GERD and Barrett's esophagus). This genetic correlation was consistent with increased use of proton-pump inhibitors (PPIs) such as omeprazole and GERD being among the strongest associated pre-existing

conditions, as well as future GERD diagnosis events. It is plausible that genetically driven decline in muscle mass is a common factor contributing to both conditions.³⁴ Increased incidence of GERD is also known to be associated with metabolic syndrome (eg. diabetes, obesity)^{35,36}, but we did not observe a significant genetic correlation between GERD and metabolic syndrome, indicating that metabolic syndrome is unlikely a common genetic risk factor between GERD and spine curvature. The link between these gastroesophageal diseases and spine curvature may also lie in the common use of PPIs for treatment. Chronic use of PPIs which has been associated with increased fracture risk and osteoporosis which may therefore lead to increased spinal curvature.^{37,38}

From genetic colocalization and MR, most of the alleles and implicated genes we identified were previously reported in GWASAtlas as associated with broad muscular-skeletal traits such as height, BMD, appendicular lean muscle mass, and maintenance of connective tissues.³⁹ Molecular functions and expression patterns of implicated genes did not point to any single cell type or specific tissue involvement, rather they indicated involvement of global cellular maintenance processes. For example, we found support for the hypothesis that expression of genes involved in DNA repair processes (RAD9A), and microtubule-dependent chromosome segregation (SCAPER).^{40,41} Potentially in line with the observation that fundamental DNA repair processes are playing a role, previous studies have demonstrated that overactive p53 activity, and partial loss of mitochondrial checkpoint protein BubR1 can lead to increased senescence and dramatic kypholordosis.^{42,43}

A potential causal association of parathyroid hormone (PTH)-like hormone (PTHLH) with the development of kyphosis is also interesting, as it specifically points to the role of calcium/phosphate homeostasis in the pathology. Rare variants in PTHLH are known to produce skeletal defects.⁴⁴ A number of other genes involved in calcium/phosphate homeostasis, such as Klotho and FGF23 have been shown in animal model systems to develop kyphosis, likely through BMD loss and development of osteoporosis.⁴⁵ As the PTHrP therapy is an approved treatment for osteoporosis, and is known to increase vertebral BMD, it might be possible to develop intervention strategies targeted to the kyphotic population with this genetic variation and potentially PTH deficiency in general.⁴⁶

Other genes such as ETFA, critical to normal mitochondria maintenance and metabolism, could be contributing to the development of kyphosis via skeletal muscle declines.^{47,48} It is plausible that reduced capacity to repair cellular machinery in muscle and connective tissues, results in damage to these tissues over the lifetime of individuals, and consequently leads to increased spinal curvature.^{49,50} These effects are potentially exacerbated in individuals who carry mutations in cellular maintenance pathways and become more apparent in spine curvature as a function of age.^{49,51,52}

Beyond genetic predisposition, our findings suggest that bone and muscle physiology are playing an important role in the progression of kyphosis. As such, interventions such as resistance training have the potential to improve both muscle mass and strength, as well as increase bone mineral density.⁵³⁻⁵⁵ Indeed physical therapy in hyper-kyphotic individuals was shown to improve kyphosis by ~3-5°.⁵⁶⁻⁵⁸ It is also plausible that if increased spinal curvature is detected early, and can be attributed

to osteoporosis, medications such as bisphosphonates could potentially offer additional benefits.^{25,59}

Overall, our study identifies that declines in muscle and bone maintenance processes are greatly involved in age-related increases in spinal curvature. We show that increased body adiposity has distinct and opposing effects on lumbar and kyphotic angles, and increased spine curvature is associated with the development of GERD. Beyond age, maintenance of muscle mass is the key driver for kyphosis, while fat composition measures are the strongest drivers of lordosis. Analysis of genetic risk alleles suggests that molecular mechanisms involved are shared across multiple tissues and involve maintenance of cellular homeostasis such as DNA repair, RNA biogenesis, and mitochondrial functions.

Limitations

There are several limitations of the current study. DXA scan used to quantify Cobb angles were conducted in the supine lateral position which could have a significant effect on spine angles as compared to an orthostatic measure.⁶⁰ In the analysis of lumbar lordosis, we were not able to reliably identify L5 vertebrae due to pelvis obstruction, thus all Cobb angles were computed from L1-L4. The association between aging and spinal curvature was based on cross-sectional data, and thus does not represent individual aging trajectories, but rather overall population trends. Participants of the UKBB imaging study are healthy volunteers recruited from the general community and thus tend to be healthier than the general population⁶¹; however, the cohort is well-suited to answer questions about various genetic, aging, and

pathological factors that contribute to increased spinal curvature. Additionally, the significant colocalization findings of a few SNPs (rs9651832, rs3783820, rs9651832, rs71457747, and rs12908334) with RNA expression (eQTL/sQTL) of multiple genes may suggest a transgression of potential horizontal pleiotropy in the genetic analyses, especially for MR. Another limitation to note for the colocalization and MR analyses is that many skeletal and connective tissues are unavailable in GTEx. Thus, the possibility of associations between genetic expression in these missing tissues and spine angles cannot be dismissed.

Acknowledgments

The authors would like to acknowledge Danny Park for his assistance with the bulk download of the EBI catalog. In addition, the authors would like to thank Elena Sorokin for downloading and processing GTEx v8 for colocalization and MR analyses. This research was supported by Calico Life Sciences LLC and conducted under UK Biobank Resource application number 18448.

References

1. Fon, G. T., Pitt, M. J. & Thies, A. C., Jr. Thoracic kyphosis: range in normal subjects. *AJR Am. J. Roentgenol.* **134**, 979–983 (1980).
2. Lin, R. M., Jou, I. M. & Yu, C. Y. Lumbar lordosis: normal adults. *J. Formos. Med. Assoc.* **91**, 329–333 (1992).
3. Benoist, M. Natural history of the aging spine. in *The Aging Spine 4–7* (Springer Berlin Heidelberg, 2005).
4. Ailon, T., Shaffrey, C. I., Lenke, L. G., Harrop, J. S. & Smith, J. S. Progressive Spinal Kyphosis in the Aging Population. *Neurosurgery* **77 Suppl 4**, S164–72 (2015).
5. Sparrey, C. J. *et al.* Etiology of lumbar lordosis and its pathophysiology: a review of the evolution of lumbar lordosis, and the mechanics and biology of lumbar degeneration. *Neurosurg. Focus* **36**, E1 (2014).
6. Roghani, T., Zavieh, M. K., Manshadi, F. D., King, N. & Katzman, W. Age-related hyperkyphosis: update of its potential causes and clinical impacts-narrative review. *Aging Clin. Exp. Res.* **29**, 567–577 (2017).
7. Yau, M. S. *et al.* Heritability of Thoracic Spine Curvature and Genetic Correlations With Other Spine Traits: The Framingham Study. *J. Bone Miner. Res.* **31**, 2077–2084 (2016).
8. Lorbergs, A. L. *et al.* A Longitudinal Study of Trunk Muscle Properties and Severity of Thoracic Kyphosis in Women and Men: The Framingham Study. *J. Gerontol. A Biol. Sci. Med. Sci.* **74**, 420–427 (2019).

9. Murray, K. J., Le Grande, M. R., Ortega de Mues, A. & Azari, M. F. Characterisation of the correlation between standing lordosis and degenerative joint disease in the lower lumbar spine in women and men: a radiographic study. *BMC Musculoskelet. Disord.* **18**, 330 (2017).
10. Katzman, W. *et al.* Association of spinal muscle composition and prevalence of hyperkyphosis in healthy community-dwelling older men and women. *J. Gerontol. A Biol. Sci. Med. Sci.* **67**, 191–195 (2012).
11. Patel, R. *et al.* Trabecular bone score and its association with Cobb angle kyphosis in older men: a cross-sectional study for the Osteoporotic Fractures in Men (MrOS) Study. *Osteoporos. Int.* **33**, 1171–1176 (2022).
12. Bartynski, W. S., Heller, M. T., Grahovac, S. Z., Rothfus, W. E. & Kurs-Lasky, M. Severe thoracic kyphosis in the older patient in the absence of vertebral fracture: association of extreme curve with age. *AJNR Am. J. Neuroradiol.* **26**, 2077–2085 (2005).
13. Koelé, M. C. *et al.* The association between hyperkyphosis and fall incidence among community-dwelling older adults. *Osteoporos. Int.* **33**, 403–411 (2022).
14. Kado, D. M., Huang, M.-H., Nguyen, C. B., Barrett-Connor, E. & Greendale, G. A. Hyperkyphotic posture and risk of injurious falls in older persons: the Rancho Bernardo Study. *J. Gerontol. A Biol. Sci. Med. Sci.* **62**, 652–657 (2007).
15. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
16. Vrtovec, T., Pernus, F. & Likar, B. A review of methods for quantitative evaluation of spinal curvature. *Eur. Spine J.* **18**, 593–607 (2009).

17. Sethi, A., Ruby, J. G., Veras, M. A., Telis, N. & Melamud, E. Genetics implicates overactive osteogenesis in the development of diffuse idiopathic skeletal hyperostosis. *Nat. Commun.* **14**, 2644 (2023).
18. Muhtaroglu, M. *et al.* Evaluation of thoracic kyphosis angle and respiratory functions in patients with multiple sclerosis. *Ann. Appl. Sport Sci.* **9**, 0–0 (2021).
19. Imagama, S. *et al.* Influence of lumbar kyphosis and back muscle strength on the symptoms of gastroesophageal reflux disease in middle-aged and elderly people. *Eur. Spine J.* **21**, 2149–2157 (2012).
20. Friedman, J. *et al.* glmnet: Lasso and elastic-net regularized generalized linear models. *Astrophysics Source Code Library* ascl:2308.011 Preprint at <https://ui.adsabs.harvard.edu/abs/2023ascl.soft08011F> (2023).
21. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
22. Skrivankova, V. W. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA* **326**, 1614–1621 (2021).
23. Milne, J. S. & Lauder, I. J. Age effects in kyphosis and lordosis in adults. *Ann. Hum. Biol.* **1**, 327–337 (1974).
24. Mohan, M. & Huynh, L. Sex Differences in the Spine. *Current Physical Medicine and Rehabilitation Reports* **7**, 246–252 (2019).
25. Kado, D. M. *et al.* Factors associated with kyphosis progression in older women: 15 years' experience in the study of osteoporotic fractures. *J. Bone Miner. Res.* **28**, 179–187 (2013).

26. Katzman, W. B., Wanek, L., Shepherd, J. A. & Sellmeyer, D. E. Age-related hyperkyphosis: its causes, consequences, and management. *J. Orthop. Sports Phys. Ther.* **40**, 352–360 (2010).
27. Been, E. & Kalichman, L. Lumbar lordosis. *Spine J.* **14**, 87–97 (2014).
28. Hicks, G. E. *et al.* Cross-sectional associations between trunk muscle composition, back pain, and physical function in the health, aging and body composition study. *J. Gerontol. A Biol. Sci. Med. Sci.* **60**, 882–887 (2005).
29. Katzman, W. B., Miller-Martinez, D., Marshall, L. M., Lane, N. E. & Kado, D. M. Kyphosis and paraspinal muscle composition in older men: a cross-sectional study for the Osteoporotic Fractures in Men (MrOS) research group. *BMC Musculoskelet. Disord.* **15**, 19 (2014).
30. Kaiser, J. *et al.* Correspondence between bone mineral density and intervertebral disc degeneration across age and sex. *Arch. Osteoporos.* **13**, 123 (2018).
31. Miranda, A. P. O. C., Penha, P. J., Pereira, L. G., Pessoa, W. C. & João, S. M. A. Influence of Sex and Body Mass Index on the Thoracic Kyphosis and Lumbar Lordosis. *J. Manipulative Physiol. Ther.* **45**, 508–514 (2022).
32. Ma, L. *et al.* Association between bone mineral density and type 2 diabetes mellitus: a meta-analysis of observational studies. *Eur. J. Epidemiol.* **27**, 319–332 (2012).
33. Muka, T. *et al.* The Association between Metabolic Syndrome, Bone Mineral Density, Hip Bone Geometry and Fracture Risk: The Rotterdam Study. *PLoS One* **10**, e0129116 (2015).
34. Kim, Y. M. *et al.* Association between skeletal muscle attenuation and gastroesophageal reflux disease: A health check-up cohort study. *Sci. Rep.* **9**,

- 20102 (2019).
35. Sun, X.-M., Tan, J.-C., Zhu, Y. & Lin, L. Association between diabetes mellitus and gastroesophageal reflux disease: A meta-analysis. *World J. Gastroenterol.* **21**, 3085–3092 (2015).
 36. El-Serag, H. The association between obesity and GERD: a review of the epidemiological evidence. *Dig. Dis. Sci.* **53**, 2307–2312 (2008).
 37. Fraser, L.-A. *et al.* The effect of proton pump inhibitors on fracture risk: report from the Canadian Multicenter Osteoporosis Study. *Osteoporos. Int.* **24**, 1161–1168 (2013).
 38. Center for Drug Evaluation & Research. FDA Drug Safety Communication: Possible increased risk of fractures of the hip, wrist, and spine with the use of proton pump inhibitors. *U.S. Food and Drug Administration*
<http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm213206.htm> (2019).
 39. Liu, X. *et al.* GWAS Atlas: an updated knowledgebase integrating more curated associations in plants and animals. *Nucleic Acids Res.* **51**, D969–D976 (2023).
 40. Höpfler, M. *et al.* Mechanism of ribosome-associated mRNA degradation during tubulin autoregulation. *Mol. Cell* **83**, 2290–2302.e13 (2023).
 41. Tong, H. L., Jiang, R. Y., Zhang, W. W. & Yan, Y. Q. MiR-2425-5p targets RAD9A and MYOG to regulate the proliferation and differentiation of bovine skeletal muscle-derived satellite cells. *Sci. Rep.* **7**, 418 (2017).
 42. Tyner, S. D. *et al.* p53 mutant mice that display early ageing-associated phenotypes. *Nature* **415**, 45–53 (2002).

43. Baker, D. J. *et al.* Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature* **479**, 232–236 (2011).
44. Klopocki, E. *et al.* Deletion and point mutations of PTHLH cause brachydactyly type E. *Am. J. Hum. Genet.* **86**, 434–439 (2010).
45. Kuro-o, M. *et al.* Mutation of the mouse klotho gene leads to a syndrome resembling ageing. *Nature* **390**, 45–51 (1997).
46. Miller, P. D. *et al.* Effect of Abaloparatide vs Placebo on New Vertebral Fractures in Postmenopausal Women With Osteoporosis: A Randomized Clinical Trial. *JAMA* **316**, 722–733 (2016).
47. Schiff, M., Froissart, R., Olsen, R. K. J., Acquaviva, C. & Vianey-Saban, C. Electron transfer flavoprotein deficiency: functional and molecular aspects. *Mol. Genet. Metab.* **88**, 153–158 (2006).
48. Fiorentino, T. V. *et al.* Pioglitazone corrects dysregulation of skeletal muscle mitochondrial proteins involved in ATP synthesis in type 2 diabetes. *Metabolism* **114**, 154416 (2021).
49. Bertram, C. & Hass, R. Cellular responses to reactive oxygen species-induced DNA damage and aging. *Biol. Chem.* **389**, 211–220 (2008).
50. Mognato, M. *DNA Repair: An Update.* (BoD – Books on Demand, 2019).
51. Almeida, M. & O'Brien, C. A. Basic biology of skeletal aging: role of stress response pathways. *J. Gerontol. A Biol. Sci. Med. Sci.* **68**, 1197–1208 (2013).
52. Pignolo, R. J., Law, S. F. & Chandra, A. Bone Aging, Cellular Senescence, and Osteoporosis. *JBMR Plus* **5**, e10488 (2021).
53. Katzman, W. B., Sellmeyer, D. E., Stewart, A. L., Wanek, L. & Hamel, K. A.

- Changes in flexed posture, musculoskeletal impairments, and physical performance after group exercise in community-dwelling older women. *Arch. Phys. Med. Rehabil.* **88**, 192–199 (2007).
54. Ball, J. M., Cagle, P., Johnson, B. E., Lucasey, C. & Lukert, B. P. Spinal extension exercises prevent natural progression of kyphosis. *Osteoporos. Int.* **20**, 481–489 (2009).
55. Bansal, S., Katzman, W. B. & Giangregorio, L. M. Exercise for improving age-related hyperkyphotic posture: a systematic review. *Arch. Phys. Med. Rehabil.* **95**, 129–140 (2014).
56. Itoi, E. & Sinaki, M. Effect of back-strengthening exercise on posture in healthy women 49 to 65 years of age. *Mayo Clin. Proc.* **69**, 1054–1059 (1994).
57. Katzman, W. B. *et al.* Sex differences in response to targeted kyphosis specific exercise and posture training in community-dwelling older adults: a randomized controlled trial. *BMC Musculoskelet. Disord.* **18**, 509 (2017).
58. Katzman, W. B. *et al.* Targeted spine strengthening exercise and posture training program to reduce hyperkyphosis in older adults: results from the study of hyperkyphosis, exercise, and function (SHEAF) randomized controlled trial. *Osteoporos. Int.* **28**, 2831–2841 (2017).
59. Tucci, J. R. *et al.* Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am. J. Med.* **101**, 488–501 (1996).
60. Kado, D. M. *et al.* Comparing a supine radiologic versus standing clinical measurement of kyphosis in older women: the Fracture Intervention Trial. *Spine* **31**, 463–467 (2006).

61. Keyes, K. M. & Westreich, D. UK Biobank, big data, and the consequences of non-representativeness. *Lancet* **393**, 1297 (2019).
62. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
63. Townsend, P. Deprivation*. *J. Soc. Policy* **16**, 125–146 (1987).
64. Wood, S. N. *Generalized Additive Models: An Introduction with R, Second Edition.* (CRC Press, 2017).
65. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).
66. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
67. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).
68. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
69. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–75, S1–3 (2012).
70. MHC region of the human genome - Genome Reference Consortium.
<https://www.ncbi.nlm.nih.gov/grc/human/regions/MHC>.
71. Wakefield, J. Bayes factors for genome-wide association studies: comparison with

- P-values. *Genet. Epidemiol.* **33**, 79–86 (2009).
72. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
73. Churchhouse, C. Rapid GWAS of thousands of phenotypes for 337,000 samples in the UK Biobank. *Neale lab* <http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank> (2017).
74. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **10**, e1004383 (2014).
75. Smith, G. D. & Ebrahim, S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
76. Didelez, V. & Sheehan, N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat. Methods Med. Res.* **16**, 309–330 (2007).
77. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
78. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **46**, 1985–1998 (2017).
79. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).

Methods

Study population: The UK Biobank

The UK Biobank (UKBB) recruited 503,000 community-based adult volunteers aged 40-69 from 2006-2010 in 22 study centers across the United Kingdom: England, Wales, and Scotland.⁶² In the current analysis, we evaluate the UKBB imaging subgroup of 40,224 participants who returned between 2014-2020 to undergo spinal dual-energy X-ray absorptiometry (DXA) assessment (Lunar iDXA densitometer; GE Healthcare, Chicago, Illinois). At the baseline visit, extensive demographic and biometric information were collected including age, race, smoking status, height, and weight via questionnaires and physical measurements. The Townsend Deprivation Index (TDI), a composite measure for socioeconomic status calculated from unemployment, car and home ownership, and household crowding status, was evaluated from self-reported metrics.⁶³ Additionally, blood was collected for serum biomarker and genomic analyses. Information on prior disease conditions were derived from baseline data and clinical data aggregated across healthcare providers: the Health Episode Statistics database from England, Patient Episode Database for Wales, and the Scottish Morbidity Record 01 for Scotland. The International Classification of Diseases, Tenth Revision (ICD10) was used.

This study was approved by the North West Multi-Centre Research Ethics Committee (MREC) and informed consent was collected from all participants. Further information on the study design of the UKBB has been previously published.⁶²

Machine learning

We constructed an image-processing pipeline that establishes a trace along the center of the vertebral column; then identifies an appropriate thoracic or lumbar section of the spine for analysis of kyphosis or lordosis, respectively; then measures the angle difference between lines tangential to the spinal traces at the top versus bottom of that section as a proxy for the Cobb angle. For the kyphotic angle, the thoracic spine approximately above and including vertebra T12 was analyzed; for the lordotic angle, the lumbar spine approximately between and including vertebrae T12 through L5 was analyzed. Full details of the method can be found in the Supplementary Method and GitHub repository (<https://github.com/graham-calico/KyphoLordoDxa.git>). Kyphosis and lordosis were measured using the python tool “spineCurve.py”. The following command-line options were invoked, for kyphosis: “--aug_flip --side_facing right”; and for lordosis: “--lumbar --aug_flip --side_facing right --aug_tilt 0.5”. Exact reproduction of the results in this manuscript should additionally invoke the “--legacy” flag. Further methods are described in the Supplementary Methods section.

Statistical Methods

For all analyses, two-tailed P-values were calculated. Significance was defined as a false discovery rate-corrected p-value (pFDR) less than 0.05. All epidemiological analyses were performed with R version 4.3.1.

We assessed participant characteristics at the time of the imaging visit,, stratified for presence and absence of hyperkyphosis ($>40^\circ$) and hyperlordosis ($>30^\circ$).

Continuous variables were summarized as medians and interquartile ranges (IQR), and categorical variables were presented as counts and percentages. The correlation between kyphotic and lordotic angles was displayed using locally smoothed regression with span=0.75 and the correlation coefficient (r) was calculated using Pearson product-moment correlation after confirming the linearity assumption could be reasonably met. Unadjusted Wald unconditional maximum likelihood estimation was used to calculate the risk ratio between prevalent hyperkyphosis and hyperlordosis.

To evaluate the association between age and kypholordotic angles, several analyses were conducted. First, the sex-stratified association between age and each kyphotic and lordotic angle was plotted using locally smoothed regression (span=0.75) with 95% confidence intervals (ggplot2). Second, participants were binned by age decile (e.g. 45-54, 55-64, etc.), and the proportion of individuals in each respective age category that had high kyphosis (40° to $50^\circ+$) and lordosis (30° to $40^\circ+$) angles by 5° intervals was graphed. Last, the mean age of participants in each 10° kyphosis by lordosis angle was presented as a heatmap.

To gain further insight into factors that might be contributing to increased spinal curvature, we used linear regression to evaluate the association of Cobb angles with 414 pre-existing medical conditions with prevalence >200 in our population, 59 clinical chemistry measures, and 15 body measure traits (adiposity, bone mineral density, muscle, lung traits) in UK Biobank. All associations were adjusted for age, age², sex, age*sex, smoking, BMI, and township deprivation index. FDR-adjusted significant associations between Cobb angle (either kyphosis or lordosis) and pre-existing conditions and clinical chemistry measures were presented. All 15 adiposity and

skeletal-muscular body measure associations were presented, regardless of significance since they were related to our main study interest of phenotypic body composition associations with kypholordotic angle. All these skeletal-muscular associations were adjusted for variables listed above except BMI was not included in the adjustment of measures highly collinear with BMI calculation [BMI, height, total fat mass, and visceral adipose tissue].

We expected that the associations of various musculoskeletal traits on spine curvature could be non-linear. Thus, to assess ranked independent predictors of Cobb angles (out of body measure traits, age, sex, and smoking status), we ran variable selection using a 10-fold cross-validated boosted generalized boosted regression model with interaction depth = 2 and modeled selected features using splines in the generalized additive model (GAM).⁶⁴ Variables with relative influence greater than 0 were selected for the model and relative variable influence percentages were presented. In secondary analyses, we reran all these linear regressions after stratifying by sex to examine potential effect modification by sex of these associations.

Genetics

UKBB imputed genotypes were used in all genetic analyses.¹⁵ Further details on their methods have been previously published (https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/impute_ukb_v1.pdf). We excluded SNPs with 1) minor allele frequency <1%, 2) low imputation quality with info value <0.9, 3) genotype missingness >10%, or 4) demonstrated deviation from Hardy-

Weinberg equilibrium (HWE $p < 1e^{-10}$). Participants were excluded from the genomics analyses if they were 1) not of European ancestry (field ID 22006), 2) had heterozygosity or genotype call rate outliers (field ID 22027), 3) demonstrated sex chromosome aneuploidy (field ID 22019), 4) No mismatch between genetic sex and self reported sex (field ID 22001), and 5) Not genetically related (field ID 22011). This yielded a total of 9,911,384 SNPs and 33,413 participants in the genetic analyses. Bonferroni correction was used to evaluate significance in all the genetics analyses and a threshold of $pFDR < 0.1$ was used. This study was conducted on a pre-existing dataset so was not preregistered.

Genetic Wide Association Study

For the initial genetic wide association study (GWAS), a whole-genome regression model was conducted on kyphosis and lordosis separately using REGENIE, a computationally efficient, machine learning-based parallel analysis approach.⁶⁵ In brief, REGENIE first uses cross-validated ridge regression by SNP block for dimension reduction of the genetic data. Then, a second cross-validated ridge regression for each trait is conducted to combine predictors from the first ridge regression into an overall prediction by trait, which is then decomposed by chromosome for a leave-one-chromosome-out (LOCO) scheme to decrease potential proximal contamination. These LOCO predictors are used as a covariate where each phenotype is tested by the set of imputed SNPs, also adjusting for genotype SNP chip (Illumina vs Affymetrix), sex, age, age^2 , $age * sex$, and recruitment center. We verified that the test statistics showed no

inflation compared to the expectation using the genomic control lambda coefficient (1.13 for kyphosis and 1.09 for lordosis) and the intercept (0.999, SD 0.009 for kyphosis and 1.000, SD 0.084 for lordosis) of linkage disequilibrium score regression.

Multi-trait Analysis of GWAS (MTAG)

The method for conducting MTAG from GWAS statistics has been previously described.²¹ MTAG enhances statistical power by leveraging the genetic correlation between traits to generate trait-specific estimates for each SNP. Based on linkage disequilibrium score regression (LDSC) estimates of genetic correlations between kyphosis and lordosis, a joint analysis of the two traits using the GWAS summary statistics was conducted using MTAG. Briefly, variants were restricted to those with minor allele frequency (MAF) > 0.01, and those that passed quality control (QC) in the individual trait GWAS. We verified that MTAG did not inflate the test statistics compared to the expectation using the genomic control lambda coefficient (1.14 for kyphosis and 1.11 for lordosis) and the intercept (0.998, SD 0.0085 for kyphosis and 0.9842, SD 0.0089 for lordosis) of LDSC.

Genetic Correlation and Heritability

To examine if the genetic variation linked to spine angles shares genetics with other traits, we calculated genetic correlations with GWAS catalog phenotypes for kyphosis and lordosis, respectively. All available summary statistics from the EBI GWAS catalog (www.ebi.ac.uk/gwas/) were downloaded on October 19, 2022.⁶⁶ Of the 25,475 summary statistics files, 20,021 could be successfully re-formatted by the munge step of

LDSC.⁶⁷ Of those, LDSC calculated a positive heritability for 14,843 summary statistics files. Of these traits, 10,170 and 10,138 for kyphosis and lordosis, respectively, yielded non-null SNP heritabilities and were used for downstream genetic correlation analysis.

In addition to traits available on the EBI GWAS catalog, we also estimated heritability of and genetic correlations between kypholordosis and musculoskeletal traits measured in the UKBB (adiposity, bone mineral content, muscle and lung traits). LDSC was calculated using the repository at <https://github.com/bulik/ldsc/>, version [aa33296](#). We estimated genetic correlation and heritability using default parameters and the `--rg` command and `--h2` command, respectively (example: `ldsc.py --rg kyphosis.sumstats.gz, lordosis.sumstats.gz --ref-ld-chr eur_w_ld_chr/ --w-ld-chr eur_w_ld_chr/ --out`). Allele polarization was pre-harmonized to match the reference file `w_hm3.snplist`. Analyses were adjusted for sex, age, age², age*sex, and recruitment center.

Fine-Mapping

GCTA was used for approximate conditional analysis.^{68,69} We considered all variants that passed QC (above) and were within 500kb of the locus index variant, except the major histocompatibility complex (MHC) region due to the complex of LD structure of the region (GRCh38::6:28,510,120-33,480,577).⁷⁰ For the LD calculation reference panel, we used genotypes from 5,000 randomly-selected, unrelated, Caucasian UKBB participants. For individual loci, variants with locus-wide evidence of association ($p_{\text{joint}} < 1E-6$) were considered conditionally independent. $p < 5E-8$ was used as the genome-wide association threshold for defining a locus.

Construction of genetic credible sets

For every distinct GWAS signal from, credible sets were calculated at a threshold of 95% probability of containing at least one variant with a true non-zero effect size.

First, the natural log approximate Bayes factor, Λ_j , was computed for the j th variant within the fine-mapping region:

$$\Lambda_j = \ln \left(\sqrt{\frac{V_j}{V_j + \omega}} \frac{\omega \beta^2}{2V_j(V_j + \omega)} \right) \quad (1)$$

where β_j and V_j are the estimated effect size and corresponding variance, respectively.⁷¹ The parameter ω denotes the prior variance in allelic effects and is estimated as $(0.15\sigma)^2$, where σ is the standard deviation of the phenotype. This standard deviation (σ) was estimated using the formula below:

$$2n_j f_j(1 - f_j) \sim \sigma^2 \frac{1}{\text{var}(\beta_j)} - 1 \quad (2)$$

Here, $\text{Var}(\beta_j)$ is the variance of the beta coefficients, f_j is minor allele frequency, n_j is sample size, and σ^2 is the coefficient of the regression which was used to estimate σ with $\sigma = \sqrt{\sigma^2}$.

In genetic loci with multiple distinct association signals, we used two techniques 1) exact conditional analysis adjusting for all other index variants in the fine-mapping region and 2) marginal analysis not adjusting for other index variants within the locus. In

genetic loci with only a single association signal, we used an unconditional meta-analysis.

Given l variants in the region, we then calculated the posterior probability, π_j , that the j th variant was driving the association using:

$$\pi_j = \frac{(1-\gamma)A_j}{1 + \sum_{k=0}^l A_k} \quad (3)$$

where γ denotes the prior probability for no association at this locus and k indexes the variants in the region (with $k=0$ indicating no association in the region). To account for the expected false discovery rate of 5%, we set $\gamma=0.05$ since a threshold of $p_{\text{marginal}} < 5 \times 10^{-8}$ was used to identify loci for fine-mapping.

For each signal, the 95% credible set was then constructed by (i) ranking all variants by their Bayes factor and (ii) including all ranked variants until their cumulative posterior probability exceeded 0.99.

PheWAS

Fine-mapped lead SNPs were annotated to potential functional genes using the Ensembl Variant Effect Predictor, Open Targets, and dbSNP databases. We also considered previous significant associations in the GWASAtlas catalog and summarized our findings in **Table 1**.

Colocalization

We performed colocalization tests to evaluate whether two traits (spine angle and gene expression) share the same underlying causal variants. For gene expression colocalizations, we used summary statistics from GTEx v8.⁷² For physiological and disease trait colocalizations, we used UKBB summary statistics of standardized quantitative and ICD10-categorized traits.⁷³ We identified UKBB phenotypes where the minimum p-value within the +/-500kb region around the locus tag SNP was $<5 \times 10^{-8}$. Colocalization analysis was conducted using the coloc R package with default priors and including all variants within 500kb of the index variant.⁷⁴ Two genetic signals were considered to have strong colocalization evidence if $PP3+PP4 \geq 0.99$ and $PP4/PP3 \geq 5$ and suggestive colocalization evidence if $PP3+PP4 \geq 0.8$ and $PP4/PP3 \geq 3$.⁷⁴

Mendelian Randomization Analysis

To further evaluate causal associations between the genes identified driving spinal curvature, we conducted MR analyses.⁷⁵ In MR analysis, we use cis-regulatory or splicing variants that affect RNA expression/function of implicated genes as instrumental variables and evaluate whether these expression and splicing changes drive changes in a phenotype of interest. Since genetic variants are theoretically randomly assigned when passed from parents to offspring, this method is expected to minimize the effect of confounding and avoid reverse causation.

The MR method relies on three important assumptions⁷⁶:

- (1) The genetic variant is associated with the exposure.
- (2) The genetic variant is not associated with any confounders.
- (3) The genetic variant is not associated with the outcome through any pathway other than via the exposure (no horizontal pleiotropy).

Since multiple tissues can affect spine curvature, we carried out two sample MR analyses across all 54 tissue types measured in GTEx v8.⁷² To identify independent SNP instruments for each exposure, we pruned GWAS-significant SNPs (p-value $< 5 \times 10^{-8}$) for each risk factor using a threshold of $r^2 < 0.01$ and LD window of 250 kb. The construction of reference panels for LD calculations was described in the fine-mapping section. When multiple independent eQTLs were observed for the same gene in a single tissue within GTEx, we evaluated the causal effect of a gene within each genome-wide significant locus on kyphosis and lordosis using inverse variance weighting (IVW). For genes with only a single independent eQTL association in a tissue, we used the Wald ratio test. A notable limitation of eQTL-based MR analysis is that bone and connective tissues are not available in the current version of GTEx (v8), and we cannot rule out the possibility that causal effects on spine angles are not due to gene expression in missing tissues.

For sensitivity analyses, the MR analysis was also reconducted using 1) MR-Egger regression and 2) weighted median-based and mode-based tests of the ordered Wald ratio estimates.^{77,78} Since weighted median and mode estimates assumes that the

estimates from pleiotropic variants are outliers, these methods tend to be less sensitive to the effect of pleiotropic variants.⁷⁹ On the other hand, MR-Egger involves a weighted linear regression to evaluate the marginal effect of each SNP to the outcome on the marginal effect of each SNP to the exposure. Thus, this method evaluates the overall directional pleiotropic contribution of weak instrumental SNPs on the risk estimate. As such, we used these methods as sensitivity analyses to confirm that there was a consistent directional effect between the variants and traits as from the IVW and Wald ratio results.⁷⁹

Ethics aspects

All UKBB data used in this study were anonymized and were accessed through project number 18448. The National Research Ethics Service Committee of North West-Haydock (REC reference: 11/NW/0382) approved the UKBB project. Signed informed consent was obtained electronically from all participants

<https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/Consent.pdf>.

Data availability

Data from the UK Biobank is available on request (<https://www.ukbiobank.ac.uk/>). Summary statistics are available from the GWAS catalog under accession number (TBD).

Code availability

Our open-source code can be accessed at <https://github.com/graham-calico/KyphoLordoDxa.git>.

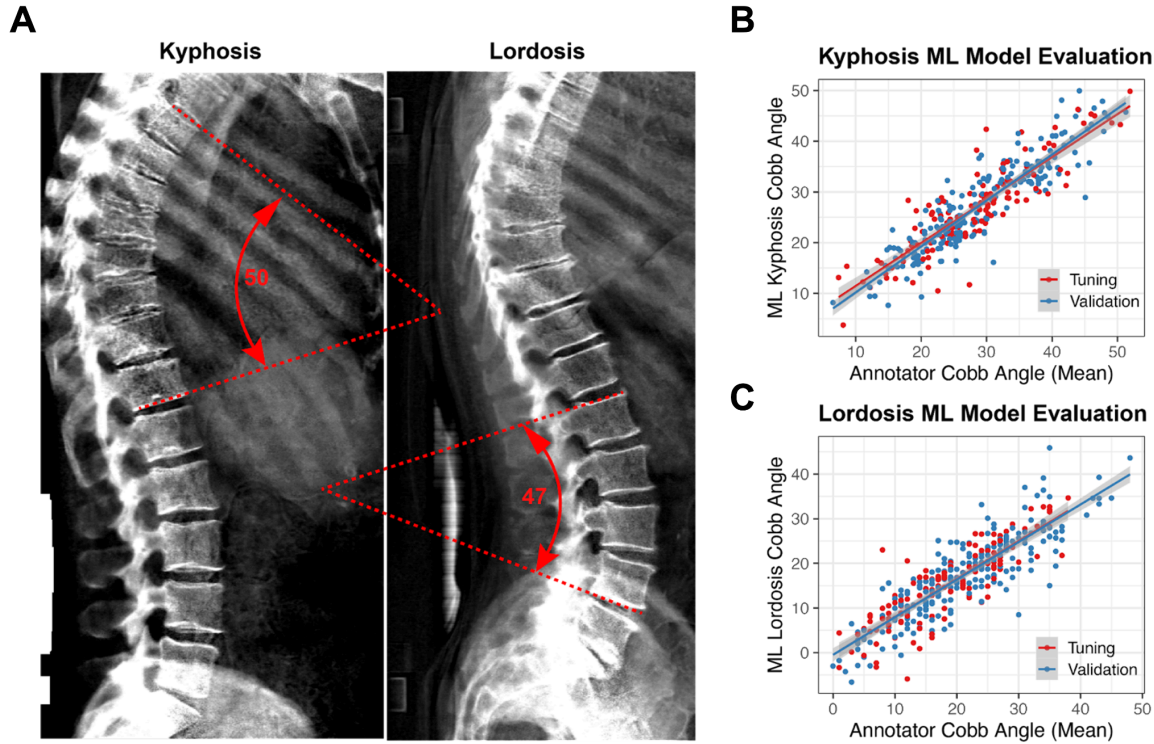


Figure 1: Machine learning model evaluation. A) Example of manually curated kyphotic (T5-T12) and lordotic (L1-L4) Cobb angles in lateral UK Biobank lateral DXA scan. (B-C) Correlations of Cobb angles as assessed by average annotator and machine learning derived kyphotic angles and lordotic angles in training (tuning) and validation sets.

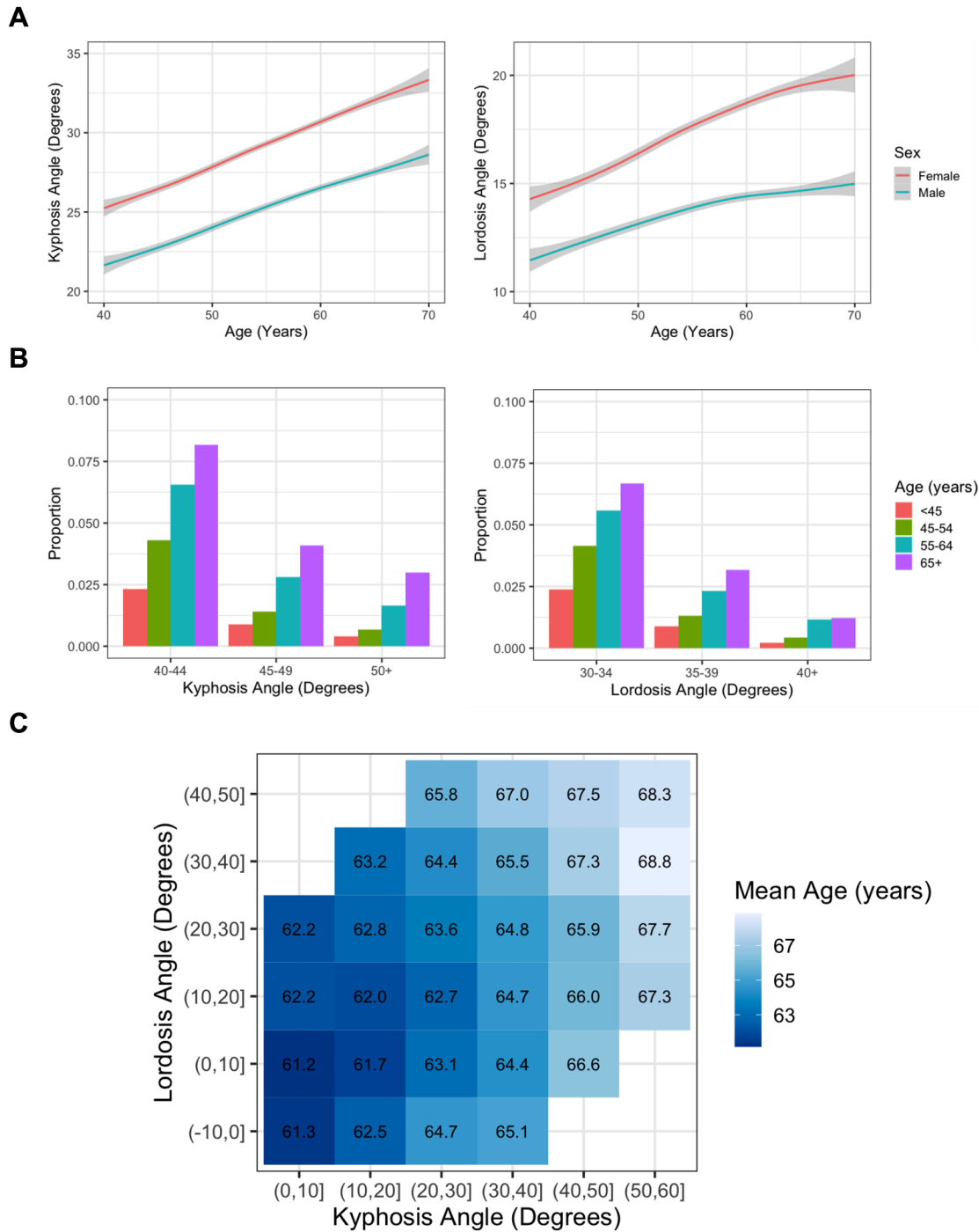


Figure 2. Distribution of kyphosis and lordosis. A) Kyphotic and lordotic angles as a function of age, stratified by sex. Spinal curvature increases with age in both sexes with approximately the same slope, but kyphotic and lordotic Cobb angles are greater in females. B) Prevalence of hyperkyphosis and hyperlordosis increases with age. C) Distribution of kyphotic and lordotic angle by age category and sex.

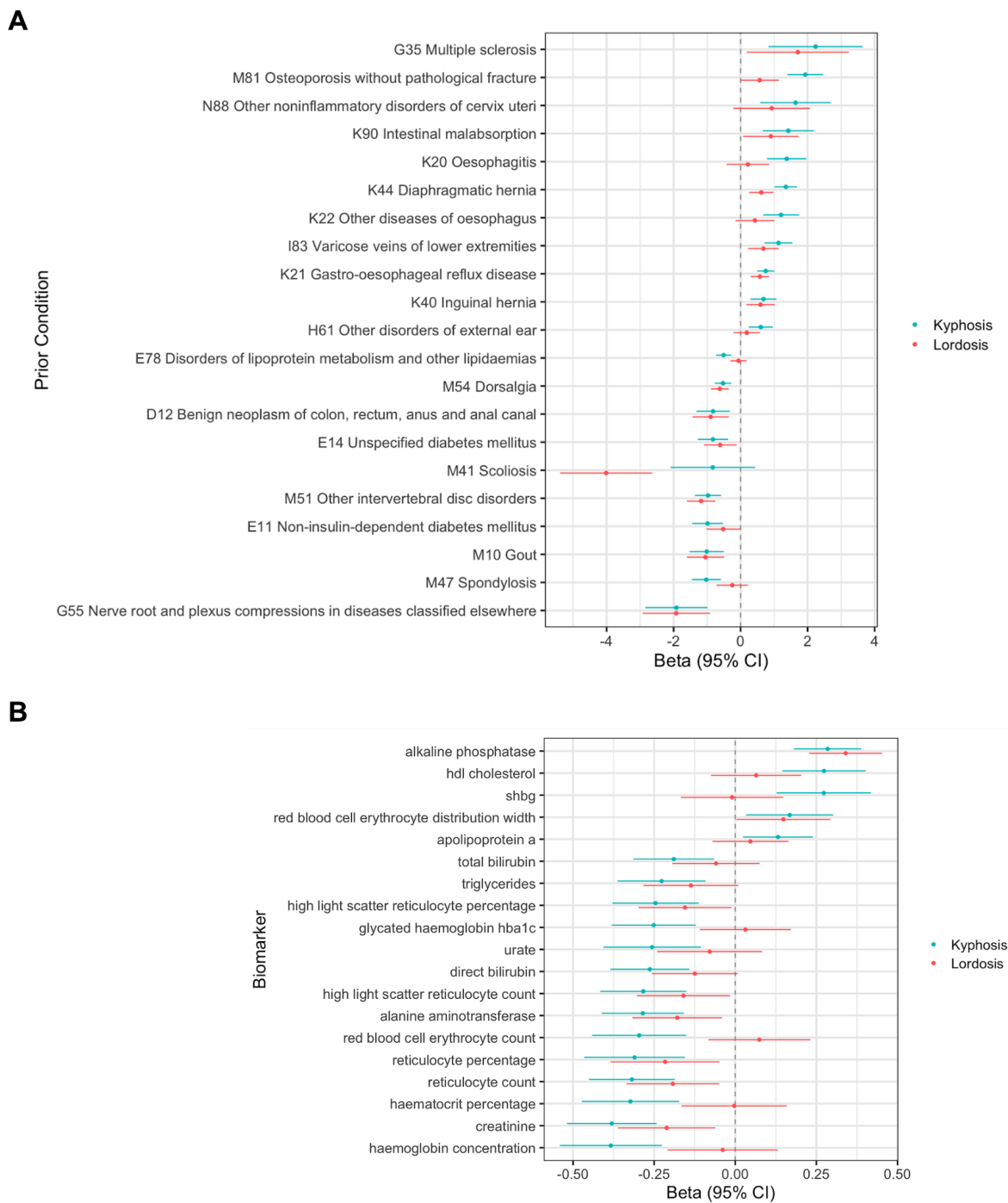


Figure 3. Spinal curvature associations with A) Pre-existing diagnosed conditions B) Biochemistry. All analyses were adjusted by age at visit, sex, BMI, smoking status, and Townsend deprivation index. Angle as a function of a disease or biomarker. Beta is a degree change in kypholordotic angle with the presence of a prior condition or per standard deviation increase in biochemical measure.

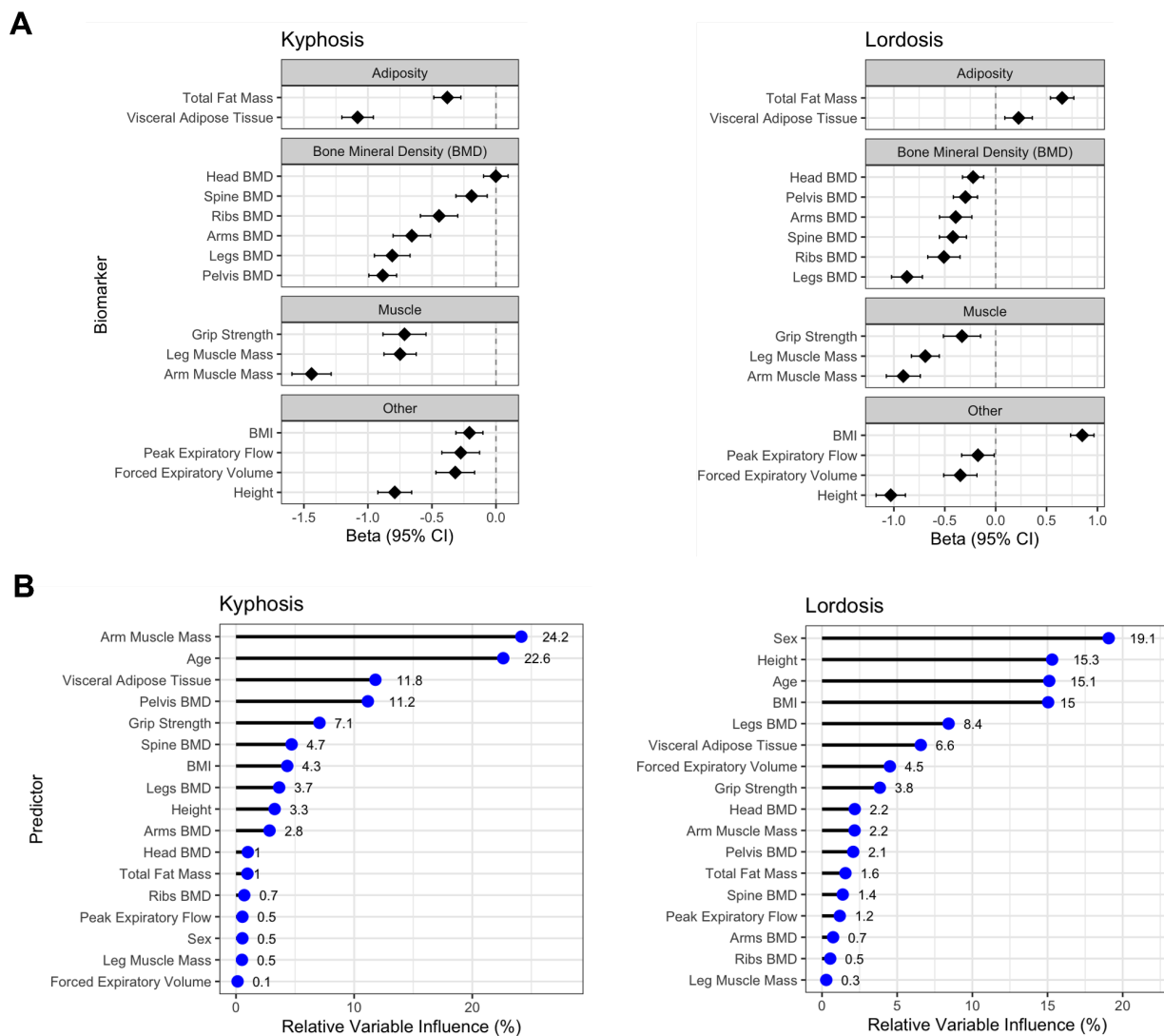


Figure 4: Physiology associations. A) Association of musculoskeletal traits with spine curvature. Analyses were adjusted by age at visit, sex, BMI, smoking status, and Townsend deprivation index, except BMI, height, total fat mass, and visceral adipose tissue were not adjusted for BMI. Beta is a degree change in angle per standard deviation increase in physiological measure. B) Importance of cross-validated GBM-selected variables for predicting spine curvature.

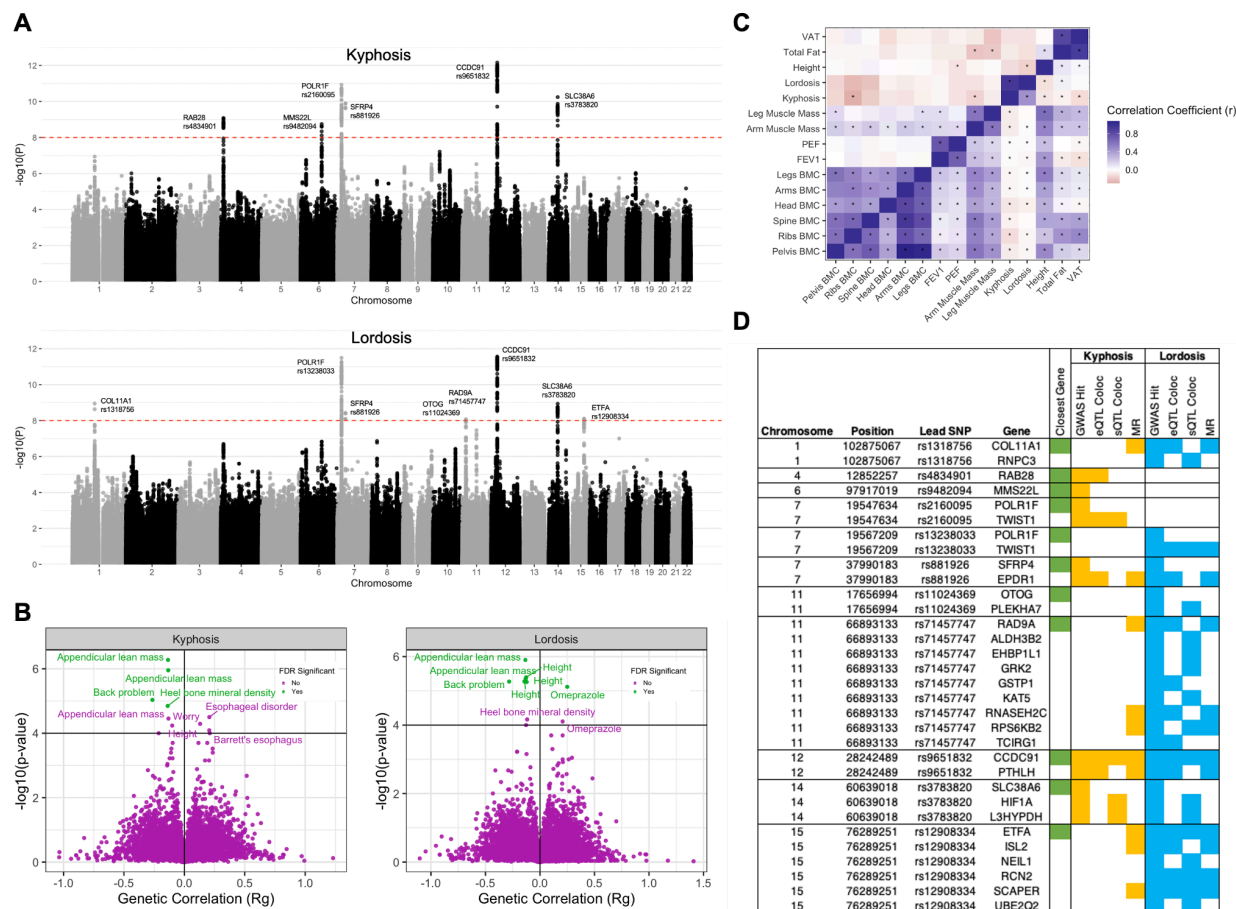


Figure 5. Genetics of spine curvature. A) GWAS manhattan plot with fine-mapped lead SNPs annotated B) GWAS catalog phenotypic correlations with kypholordotic angle. In green are significant associations where $pFDR < 0.05$. C) Comparison of genetic and phenotypic correlations. Displayed in the upper left triangle are genetic correlations for selected musculoskeletal traits. Displayed in the lower right triangle are phenotypic correlations of selected musculoskeletal traits. Both traits show similar patterns of genetic and phenotypic correlations, with greater kypholordotic angles anti-correlated with other traits. Significant correlations are marked with asterisks. D) Genetic summary figure of GWAS, colocalization (eQTL and sQTL), and MR associations by lead SNP and gene.

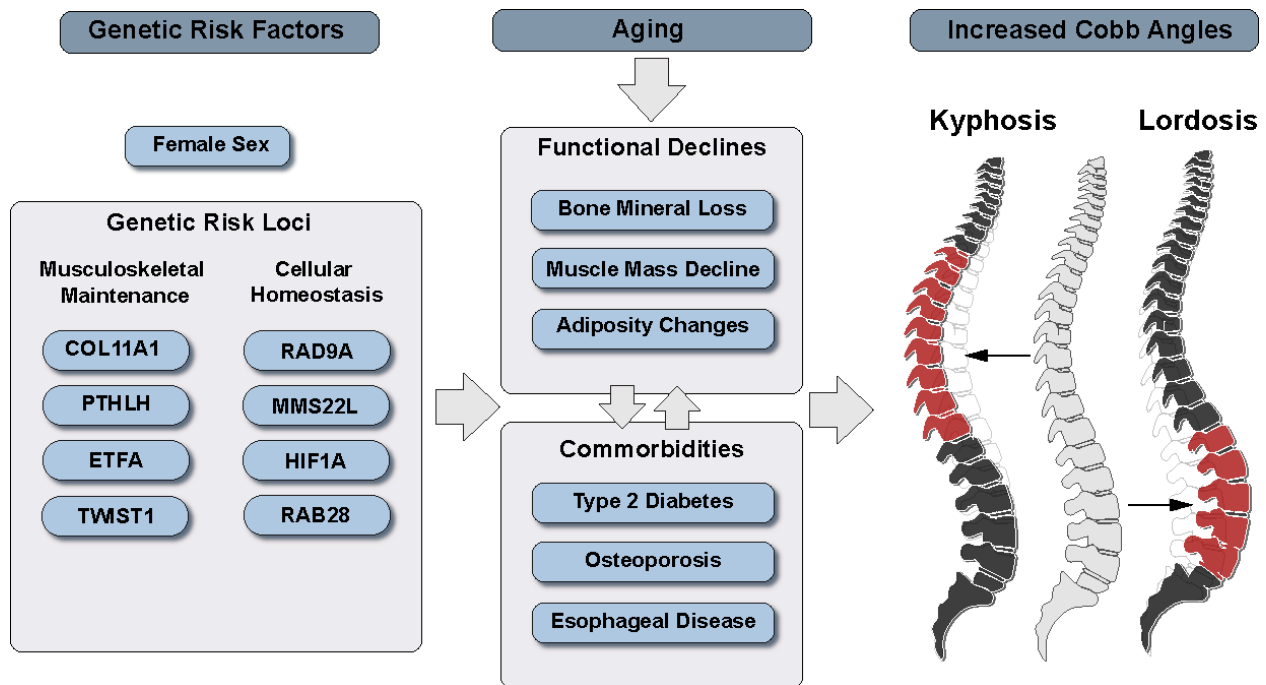


Figure 6: Multi-system physiology effect on spine curvature as a function of age.

Loss of bone mineral density, decreased muscle mass, and decreased lung capacity are associated with increased kypholordotic angles. Adiposity has opposing effects on the thoracic and lumbar spine: decreasing kyphosis and increasing lordosis. Aging is a common risk factor for declines in physiological function and presentation of comorbid conditions such as osteoporosis, type 2 diabetes, and esophageal diseases. Both traits are strongly genetically correlated and share common risk alleles such as PTHLH → BMD, COL11A1 → connective tissues, and ETFA → muscle mass, as well as a number of genes involved in DNA repair and cellular homeostasis processes (RAD9A, MMS22L, HIF1A, RAB28).

Table 1. Genetic associations between phenotypic traits and kypholordotic angle.

Trait	Nearest Gene	Lead SNP	Beta (95% CI)	P	PheWAS Associations
Lordosis	COL11A1	rs1318756	0.040 (0.027, 0.053)	1.13E-09	Height, waist-hip ratio, hip osteoarthritis, hip area and replacement, intervertebral disc disorder
Kyphosis	RAB28	rs4834901	0.050 (0.034, 0.066)	8.30E-10	Height, PEF, appendicular lean body and trunk mass, femoral neck area, spinal stenosis
Kyphosis	MMS22L	rs9482094	0.045 (0.030, 0.059)	1.80E-09	Cognitive performance, body mass index, plasma c-reactive protein, urinary sodium excretion, HDL-C
Kyphosis	POLR1F	rs2160095	0.051 (0.036, 0.065)	1.17E-11	Height, heel bone mineral density, spinal canal stenosis
Lordosis	POLR1F	rs13238033	0.047 (0.034, 0.060)	3.28E-12	Height, heel bone mineral density, FEV1, appendicular lean mass
Kyphosis	SFRP4	rs881926	0.063 (0.044, 0.082)	1.27E-10	Heel bone mineral density, height, appendicular lean body mass, FEV1, fasciitis
Lordosis	SFRP4	rs881926	0.053 (0.035, 0.071)	3.75E-09	See above
Lordosis	OTOG	rs11024369	0.041 (0.027, 0.054)	8.46E-09	No known significant associations
Lordosis	RAD9A	rs71457747	-0.083 (-0.113, -0.054)	3.44E-08	Height, trunk mass, basal metabolic rate, IGF1, serum cholesterol, phosphate, hip osteoarthritis and replacement
Kyphosis	CCDC91	rs9651832	-0.052 (-0.067, -0.038)	6.94E-13	FVC, FEV1, height, PEF, appendicular lean body mass
Lordosis	CCDC91	rs9651832	-0.047 (-0.060, -0.034)	2.78E-12	See above
Kyphosis	SLC38A6	rs3783820	-0.056 (-0.073, -0.039)	5.70E-11	Height, appendicular lean body and trunk mass, basal metabolic rate, IGF1, weight, urinary sodium excretion, blood pressure
Lordosis	SLC38A6	rs3783820	-0.048 (-0.064, -0.033)	1.16E-09	See above
Lordosis	ETFA	rs12908334	-0.038 (-0.051, -0.025)	7.80E-09	PEF, FEV1/FVC ratio, amino acid metabolism