frontiers | Frontiers in Endocrinology

TYPE Original Research
PUBLISHED 21 November 2023
DOI 10.3389/fendo.2023.1261088

OPEN ACCESS

EDITED BY
Shaolong Cao,
Biogen Idec, United States

REVIEWED BY
Li Huang,
Shenzhen Second People's Hospital, China
Shuangxi Ji,
University of Texas MD Anderson Cancer Center, United States

*CORRESPONDENCE
Weihua Zhou
whzhou@mtu.edu
Hong-Wen Deng
hdeng2@tulane.edu

†These authors have contributed equally to this work

RECEIVED 24 July 2023
ACCEPTED 30 October 2023
PUBLISHED 21 November 2023

CITATION
Zhao C, Keyak JH, Cao X, Sha Q, Wu L, Luo Z, Zhao L-J, Tian Q, Serou M, Qiu C, Su K-J, Shen H, Deng H-W and Zhou W (2023) Multi-view information fusion using multi-view variational autoencoder to predict proximal femoral fracture load. *Front. Endocrinol.* 14:1261088. doi: 10.3389/fendo.2023.1261088

Multi-view information fusion using multi-view variational autoencoder to predict proximal femoral fracture load

Chen Zhao[1†], Joyce H. Keyak[2†], Xuewei Cao[3], Qiuying Sha[3], Li Wu[4], Zhe Luo[4], Lan-Juan Zhao[4], Qing Tian[4], Michael Serou[5], Chuan Qiu[4], Kuan-Jui Su[4], Hui Shen[4], Hong-Wen Deng[4*] and Weihua Zhou[1,6*]

[1]Department of Applied Computing, Michigan Technological University, Houghton, MI, United States, [2]Department of Radiological Sciences, Department of Biomedical Engineering, Department of Mechanical and Aerospace Engineering, and Chao Family Comprehensive Cancer Center, University of California, Irvine, Irvine, CA, United States, [3]Department of Mathematical Sciences, Michigan Technological University, Houghton, MI, United States, [4]Division of Biomedical Informatics and Genomics, Tulane Center of Biomedical Informatics and Genomics, Deming Department of Medicine, Tulane University, New Orleans, LA, United States, [5]Department of Radiology, Deming Department of Medicine, School of Medicine, Tulane University, New Orleans, LA, United States, [6]Center for Biocomputing and Digital Health, Institute of Computing and Cybersystems, and Health Research Institute, Michigan Technological University, Houghton, MI, United States

Background: Hip fracture occurs when an applied force exceeds the force that the proximal femur can support (the fracture load or "strength") and can have devastating consequences with poor functional outcomes. Proximal femoral strengths for specific loading conditions can be computed by subject-specific finite element analysis (FEA) using quantitative computerized tomography (QCT) images. However, the radiation and availability of QCT limit its clinical usability. Alternative low-dose and widely available measurements, such as dual energy X-ray absorptiometry (DXA) and genetic factors, would be preferable for bone strength assessment. The aim of this paper is to design a deep learning-based model to predict proximal femoral strength using multi-view information fusion.

Results: We developed new models using multi-view variational autoencoder (MVAE) for feature representation learning and a product of expert (PoE) model for multi-view information fusion. We applied the proposed models to an in-house Louisiana Osteoporosis Study (LOS) cohort with 931 male subjects, including 345 African Americans and 586 Caucasians. We performed genome-wide association studies (GWAS) to select 256 genetic variants with the lowest p-values for each proximal femoral strength and integrated whole genome sequence (WGS) features and DXA-derived imaging features to predict proximal femoral strength. The best prediction model for fall fracture load was acquired by integrating WGS features and DXA-derived imaging features. The designed models achieved the mean absolute percentage error of 18.04%, 6.84% and 7.95% for predicting proximal femoral fracture loads using linear models of fall loading, nonlinear models of fall loading, and nonlinear models of stance loading, respectively.

Conclusion: The proposed models are capable of predicting proximal femoral strength using WGS features and DXA-derived imaging features. Though this tool is not a substitute for predicting FEA using QCT images, it would make improved assessment of hip fracture risk more widely available while avoiding the increased radiation exposure from QCT.

KEYWORDS

hip fracture, proximal femur, finite element analysis, deep learning, variational autoencoder

1 Introduction

The increasing elderly population and the rise in fracture incidence have made osteoporosis a considerable public health issue in U.S. Osteoporosis causes bones to become weak and brittle, leading to osteoporotic fractures. Osteoporosis affects about 18% of women and 6% of men globally (1). The economic burden of osteoporosis has been estimated at between $17 billion and $20.3 billion (2020 data) (2). Fracture of the proximal femur is a common and disastrous health outcome that limits previously functional elderly patients from living independently. Each year over 300,000 older people in the U.S. are hospitalized for hip fracture (3). The reported mortality rate is up to 20-24% in the first year after a hip fracture (4, 5), and a greater risk of dying may persist for at least 5 years (6). An inexpensive and accurate prognostic instrument for hip fracture risk assessment would enable individuals with a high risk for osteoporotic hip fracture to receive preventative treatment (7).

For the diagnosis of osteoporosis, areal bone mineral density (aBMD), assessed by dual energy X-ray absorptiometry (DXA), is the standard diagnostic clinical parameter (8, 9). Although DXA-derived aBMD correlates with bone weakness and fragility fracture (10), DXA is a 2D-projection technique that poorly accounts for 3D bone geometry and size (11), while bone geometry and bone size have strong genetic determination (12, 13). Further, efforts toward dissecting the genetic basis of osteoporosis using genome-wide association studies (GWASs) have been mainly focused on aBMD traits which have been widely studied, but GWAS results only explain part of the variance in hip fracture risk (14). Thus, both DXA-derived features and genetic factors provide limited information about skeletal factors on fracture risk. It has been shown that genetic determinants of aBMD, bone geometry and bone sizes are genetically correlated, sharing some commons genes (15).

Principles of physics dictate that hip fracture occurs when an applied force exceeds the force that the proximal femur can support. This force, the proximal femoral strength or fracture load, can be computed using subject-specific finite element analysis (FEA), which incorporates the biomechanically important features of the hip, i.e. the 3D bone geometry and distribution of bone density from quantitative computerized tomography (QCT) images (7, 16–19). Furthermore, FEA-computed proximal femoral strength is associated with incident hip fracture in men and women, and in men even after accounting for aBMD (18). QCT provides more accurate quantification of BMD in the lumbar spine and hip than DXA because QCT provides volumetric BMD (vBMD) while DXA calculates aBMD (20). In addition, QCT-based FEA describes the hip mechanical behavior and provides more information about bone quality and fracture risk than DXA (7, 16, 17, 19, 21, 22).

Although QCT-based FEA has shown significant value in the assessment of proximal femoral strength, radiation and availability of QCT limit its clinical usability. DXA images incurs much less radiation dose, but only describe aBMD, and proximal femoral shape and size in 2D. Yang et al. demonstrated that supplementing standard DXA-derived aBMD measurements with sophisticated femoral trabecular bone characterization from DXA significantly improved the performance of predicting hip fracture load (23). The aBMD measured by DXA is currently a standard clinical surrogate marker of bone strength to diagnose osteoporosis; however, integrating the heterogeneous distribution of bone material properties is more powerful for predicting bone strength (24).

However, for predicting bone strength, replacing 3D QCT with less robust 2D DXA data can potentially be compensated for by incorporating bone-strength related genetic variants (25). Genetic markers are important for identifying subjects at risk of hip fracture through effects on proximal femoral strength/structure. Using whole genome sequence (WGS) data, GWAS and large-scale collaborative studies have identified hundreds of genetic markers, explaining substantial proportions of population variation in osteoporotic traits (26), such as BMD (27–29) and fracture risk factors (30). Though aBMD is an important phenotype that is clinically relevant to osteoporotic hip fracture, it explains limited variance in hip strength (14, 31–34). Therefore, it is important to discover how genetics influence FEA-computed proximal femoral strength, thereby influencing hip fracture risk. Our hypothesis is that the DXA-derived imaging features and genetic features from WGS data could be incorporated to predict proximal femoral strength with high clinical applicability. Further, prediction models would help integrate the large number of high-dimensional inter-correlated complicated predictors from genetic data and image features to draw an overall conclusion regarding proximal femoral strength in individual patients.

In this paper, we propose a novel model, multi-view variational autoencoder with the product of expert (MVAE-PoE) for proximal

femoral strength prediction, as shown in Figure 1. The proposed MVAE-PoE incorporates variational autoencoder (VAE) to learn feature representation and employs the product of expert (PoE) for multi-view information fusion. A linear regression estimator is used to predict proximal femoral strength based on the extracted latent features. Extensive analyses were performed, leveraging the combination of whole genome sequence (WGS) data and DXA-derived imaging features.

2 Materials and methodology

The participant, intervention, comparison, and outcome (PICO) of this study is shown below.

- Population (P): A cohort of 931 male subjects, comprising 345 African Americans and 586 Caucasians, with available QCT images, WGS features, and DXA-derived imaging features, were included.
- Intervention (I): This is not a clinical intervention study. However, the goal is to develop a deep learning model for predicting proximal femoral strength integrating genetic information from WGS and imaging data from DXA so that the needs of intervention can be assessed.
- Comparison (C): The comparison would be between the predictive accuracy of the deep learning model when genetic information and DXA-derived imaging features are integrated versus the conventional FEA-computed proximal femoral strength using QCT.
- Outcome (O): The primary outcome is the accuracy of proximal femoral strength prediction using the proposed deep learning model, measured in terms of predictive performance metrics such as MSE, RMSE, MAPE and R^2-score.

2.1 Enrolled subjects and data generation

In this study, we propose a deep learning model to predict the proximal femoral strength calculated from QCT-based FEA by integrating WGS features and DXA-derived imaging features. The studied cohort was acquired from the LOS (35, 36). The LOS cohort is an ongoing research dataset (>17,000 subjects accumulated so far) with recruitment starting in 2011, aimed at investigating both environmental and genetic risk factors for osteoporosis and other musculoskeletal diseases (37, 38). All participants signed an informed-consent document before any data collection, and the study was approved by the Tulane University Institutional Review Board.

Peak BMD achieved and remained relatively stable at ages 20-50 years is most powerful in predicting BMD and risk to osteoporotic fractures later in life due to the relatively stable physiological and hormone status during this age period (39–42). A 10% increase in peak BMD would delay the onset of osteoporosis by 13 years (40). In comparison, a 10% increase in the age of menopause, or a 10% reduction in age-related bone loss would only delay the onset of osteoporosis by 2 years (40).

Therefore, in this study, we focus on a cohort of 931 male subjects, aged 20-50, consisting of 345 African Americans and 586 Caucasians, with available QCT images, WGS and DXA-derived features. The basic demographic information for the enrolled subjects is shown in Table 1.

2.2 QCT image acquisition and FEA for calculation of proximal femoral strength

The QCT scans (GE Discovery CT750 HD system; 2.5 mm-thick slices; pixel size, 0.695–0.986 mm; 512 × 512 matrix) were acquired at Tulane University Department of Radiology. For each

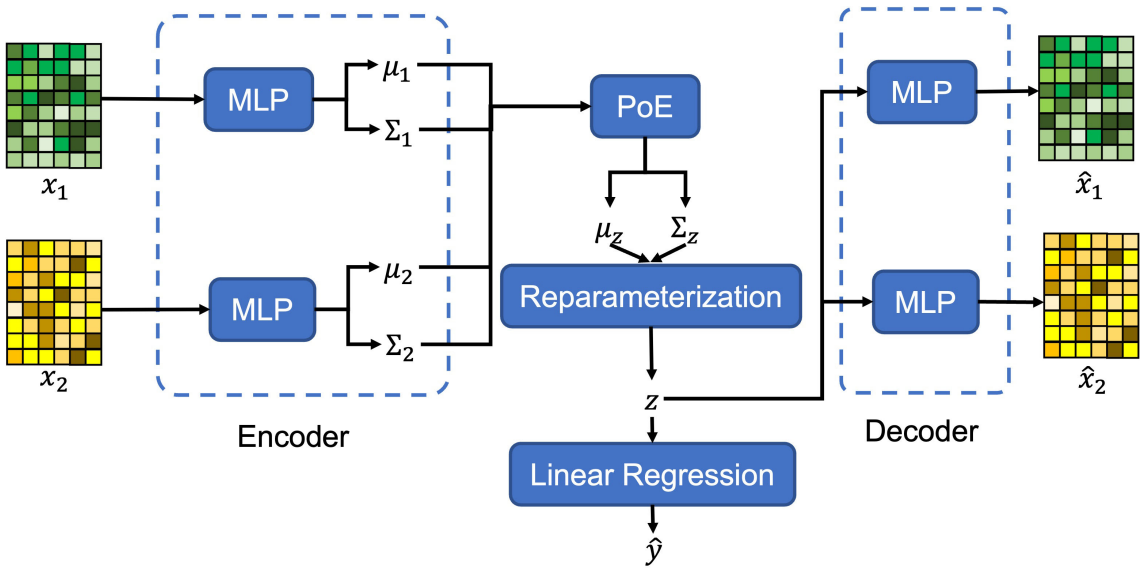

FIGURE 1
The graphical architecture of the proposed MVAE-PoE model.

10.3389/fendo.2023.1261088

TABLE 1 Demographic information for the enrolled subjects.

Race	Age (year, mean±SD)	Height (cm)	Weight (kg)
African-American	38.60±7.74 (min: 20 max: 51)	174.69±7.03 (min: 154.00, max: 190.80)	82.67±17.68 (min: 50.80, max: 135.20)
Caucasian	35.22±8.53 (min: 20, max: 51)	175.37±6.80 (min: 154.94, max: 198.00)	83.25±16.07 (min: 50.80, max: 135.40)
All	36.47±8.40 (min: 20, max: 51)	175.12±6.89 (min: 154.00, max: 198.00)	83.04±16.68 (min: 50.80, max: 135.40)

The ranges are illustrated in the parenthesis. SD, standard deviation.

QCT slice, contours of the left proximal femur were labeled by well-trained operators, in consultation with our experienced researcher (J.H.K.). We developed in-house software for automated annotation, containing a previously developed deep learning-based segmentation model (43) and a thresholding algorithm with edge tracing (44), in combination with manual visualization and correction. The deep learning model proposed by Zhao et al. (43) achieved a Dice similarity coefficient of 0.9888, indicating that only minor manual modifications are required for annotating new QCT images. Using the annotated contours, we then used linear and nonlinear FE models to estimate the strengths of the proximal femoral under two loading conditions, single-limb stance and loading from a fall onto the posterolateral aspect of the greater trochanter (19).

The nonlinear FEA models simulated mechanical testing of the femur in which displacement is incrementally applied to the femoral head (7, 16–19). The computed reaction force on the femoral head initially increases, reaches a peak value (the load capacity or fracture load), and then decreases. To achieve this mechanical behavior, the FEA models employ heterogeneous isotropic elastic moduli, yield strengths, and nonlinear post-yield properties. These properties are computed from the calibrated QCT density (ρ_{CHA}, g/cm^3) of each voxel in an element, which are then used to compute the ash density (ρ_{ASH}, g/cm^3) (ρ_{ASH}=0.0633 + 0.887 ρ_{CHA}), and ρ_{ASH} is used to compute mechanical properties. Each linear hexahedral finite element measures 2.5 mm on a side and the mechanical properties of the element are computed by averaging the values of each property over all voxels in the element, while accounting for the volume fraction of each voxel within the element. Together, these mechanical properties describe an idealized density-dependent nonlinear stress-strain curve for each element (16–19). Material yield is defined to occur when the von Mises stress exceeds the yield strength of the element. After yield, the plastic flow was modeled assuming a plastic strain-rate vector normal to the von Mises yield surface and isotropic hardening/softening. Displacement is applied incrementally to the femoral head, and the reaction force on the femoral head is computed at each increment as the distal end of the model is fully constrained. For the fall models, the surface of the greater trochanter opposite the loaded surface of the femoral head was constrained in the direction of the displacements while allowing motion transversely. The nonlinear FEA-computed proximal femoral fracture load was defined as the maximum FEA-computed force on the femoral head, i.e., the load capacity.

For phenotypes, we calculated three proximal femoral strengths under the two loading conditions: LF, NLF and NLS. The LF represents the load at the onset of fracture (7). To determine the LF, the factor of safety (FOS) at the centroid of each finite element in the model is calculated as the ratio of the yield strength of the finite element to the von Mises stress at the centroid of the element. The LF is defined as the force applied to the femoral head when the FOS values of 15 contiguous non-surface elements are equal to or less than 1.0 (7, 19). The NLF and NLS were calculated using the above-described method and represent the load capacity of the proximal femur (the maximum force of the femoral head can support) (19). The basic statistical information for the calculated proximal femoral strengths is shown in Table 2.

2.3 Whole genome sequence and GWAS for feature selection

The WGS of the human peripheral blood DNA were performed with an average read depth of 22× using a BGISEQ-500 sequencer (BGI Americas Corporation, Cambridge, MA, USA) of 350 bp paired-end reads (38). The aligned and cleaned WGS data were mapped to the human reference genome (GRCh38/hg38) using Burrows-Wheeler Aligner software (45). This process followed the recommended best practices for variant analysis with the Genome Analysis Toolkit (GATK) to guarantee precise variant identification (46). The HaplotypeCaller tool within GATK was employed to identify genomic variations, and we further enhanced the reliability of our variant calls through the application of the variant quality score recalibration method (46).

There were a total of 10,623,292 single nucleotide polymorphisms (SNPs) in the cohort with 935 subjects. For quality control, we removed genetic variants with missing rates larger than 5%, Hardy-Weinberg equilibrium exact test p-values less than 10^{-4}, and minor allele frequency (MAF) less than 5%.

TABLE 2 Statistical information for proximal femoral strengths under three loading conditions.

Race	LF (N, mean ±SD)	NLF (N, mean ±SD)	NLS (N, mean ±SD)
African American	2,587.16±831.52 (min: 875, max: 6,138)	4,353.82±570.49 (min: 2,555, max: 5,823)	21,414.04±4,350.22 (min: 11,084, max: 40,904)
Caucasian	2,162.70±782.18 (min: 716, max: 7,247)	4,281.34±562.30 (min: 2,683, max: 6,337)	19,275.38±3,954.22 (min: 10,092, max: 32,818)
All	2,319.99±826.24 (min: 716, max: 7,247)	4,308.20±566.13 (min: 2,555, max: 6,337)	20,067.90±4,231.25 (min: 10,092, max: 40,904)

The mean and standard deviation are illustrated in the parenthesis. N: Newton; SD, standard deviation; LF, linear fall fracture load; NLF, nonlinear fall fracture load; NLS, nonlinear stance fracture load.

Individuals with a missing rate larger than 20% were also excluded. Since subjects from two races were enrolled, principal component analysis (PCA) was applied to the genotypes and generated principal component scores (PCs) to perform population stratification or admixture (47). In addition, the age, weight, height, and first 10 PCs were used as covariates in GWAS (18, 35, 36).

The genome-wide association analyses were performed to test the association between each of three phenotypes and SNPs from WGS. Suppose that there are N subjects in the analyses. Let y_i be the value of the i-th subject for a phenotype and g_i be the genotype for the i-th subject, where g_i is the number of minor alleles that the subject carries at a SNP. We assume that there is a total of C covariates and the covariates for i-th subject are $\{v_1^{(i)}, v_2^{(i)}, \cdots, v_C^{(i)}\}$. For each SNP, the linear regression model is used to examine the effect of a SNP on a phenotype, as defined in Eq. 1.

$$E(y_i \mid g_i) = \alpha_0 + \alpha_1 v_1^{(i)} + \cdots + \alpha_C v_C^{(i)} + \beta g_i \tag{1}$$

where β is the effect size of the SNP on the phenotype after adjusting for the covariates. The aim is to test the null hypothesis that the SNP is not associated with the phenotype, which is equivalent to test H_0: $\beta = 0$. The score test statistic under this model is defined in Eq. 2.

$$T_{score} = UV^{-1}U \tag{2}$$

where $U = \sum_{i=1}^{N} \tilde{y}_i \tilde{g}_i$ and $V = \frac{1}{N}\sum_{i=1}^{N} \tilde{y}_i^2 \sum_{i=1}^{N} \tilde{g}_i^2$, $\tilde{y}_i$ and $\tilde{g}_i$ are the adjusted phenotype and genotype for i-th subject for the covariates, indicating that $\tilde{y}_i$ is the residual of y_i under the linear regression model defined in Eq. 3.

$$y_i = \alpha_0 + \alpha_1 v_1^{(i)} + \cdots + \alpha_C v_C^{(i)} + \varepsilon_i \tag{3}$$

where $\tilde{g}_i$ is the residual of g_i under the linear regression model defined in Eq. 4.

$$g_i = \alpha_0 + \alpha_1 v_1^{(i)} + \cdots + \alpha_C v_C^{(i)} + \tau_i \tag{4}$$

Under H_0, T_{score} follows a standard normal distribution (48).

2.4 DXA and DXA-derived imaging features

For each subject, aBMD (g/cm^2) at various skeletal sites (lumbar spine, hip, forearm, and total body) and body composition (fat/lean mass) were measured using a Hologic Discovery-A DXA (Hologic Inc., USA) by trained and certified research staff at Tulane Center for Biomedical Informatics and Genomics. To ensure quality assurance, the machine was calibrated daily using a phantom scan. The accuracy of BMD measurement was assessed by the coefficient of variation for repeated measurements, which was approximately 1.9% for femoral neck BMD (38). In addition, all the DXA images have been reanalyzed using the TBS iNsight software (Medimaps Group, Geneva, Switzerland) to obtain trabecular bone score (TBS). As a result, 196 DXA-derived imaging features were obtained and used as the imaging features in this study.

For quality control purposes: the DXA machine was calibrated daily, and long-term precision was monitored by phantoms with a coefficient of variation ≤0.7% for spine aBMD and a coefficient of variation ≤1.0% for hip aBMD (49). Mechanical malfunction, radiation quality, absorption coefficient, and tissue-equivalent materials were also checked and calibrated before the aBMD examination on a daily basis. The radiologist was licensed in the State of Louisiana and registered through the American Registry of Radiologic Technologists. The detailed DXA-derived imaging features are shown in Table S1.

2.5 Multi-view model for proximal strength prediction

For each Variational autoencoder (VAE), proposed by Kingma et al. (50), is a latent variable generative model which learns the deep representation of the input data. The goal of VAE is to maximize the marginal likelihood of the data (a.k.a evidence), which can be decomposed into a sum over marginal log-likelihoods of individual features, as illustrated in Eq. 5.

$$\log p_\theta\left(x^{(i)}\right) = D_{KL}\left(q_\phi(z \mid x^{(i)}) \,\|\, p_\theta(z \mid x^{(i)})\right) + \mathcal{L}\left(\theta, \phi; x^{(i)}\right) \tag{5}$$

where $x^{(i)}$ is the feature vector for i-th subject in the dataset $\{x^{(i)}\}_{i=1}^{N}$, N is the number of subjects, z is a random variable in the latent space, q_ϕ is the posterior approximation of z with the learnable parameters ϕ, p_θ is the ground truth posterior distribution of z with the intractable parameters θ, and $D_{KL}(\cdot \,\|\, \cdot)$ represents the Kullback–Leibler (KL) divergence between the approximated posterior distribution and the ground truth posterior distribution. Because of the non-negativity of the KL divergence, the log-likelihood $\log p_\theta(x^{(i)}) \geq \mathcal{L}(\theta, \phi; x^{(i)})$. If the approximated posterior distribution $q_\phi(z \mid x^{(i)})$ is identical to the ground truth posterior distribution $p_\theta(z \mid x^{(i)})$, then the $\log p_\theta(x^{(i)}) = \mathcal{L}(\theta, \phi; x^{(i)})$. Therefore, $\mathcal{L}(\theta, \phi; x^{(i)})$ is called the evidence lower bound (ELOB), which is defined by Eq. 6.

$$\begin{aligned}\mathcal{L}(\theta, \phi; x^{(i)}) &= \log p_\theta(x^{(i)}) - D_{KL}\left(q_\phi(z \mid x^{(i)}) \,\|\, p_\theta(z \mid x^{(i)})\right) \\ &= \mathbb{E}_{q_\phi(z \mid x^{(i)})}\left[\log p_\theta(x^{(i)} \mid z)\right] \\ &\quad - D_{KL}\left(q_\phi(z \mid x^{(i)}) \,\|\, p_\theta(z \mid x^{(i)})\right)\end{aligned} \tag{6}$$

Thus, minimizing the KL divergence is equivalent to maximizing the ELOB. To train the model explicitly and implement the loss function in a closed form, we parameterize the q_ϕ as a multivariate normal distribution (multivariate Gaussian distribution) with an approximately diagonal variance-covariance matrix. Then the analytical solution for the KL divergence is shown in Eq. 7.

$$\begin{aligned}&D_{KL}\left(q_\phi(z \mid x^{(i)}) \,\|\, p_\theta(z \mid x^{(i)})\right) \\ &= \frac{1}{2}\sum_{d=1}^{D}\left((\mu_d^{(i)})^2 + (\sigma_d^{(i)})^2 - \log((\sigma_d^{(i)})^2) - 1\right)\end{aligned} \tag{7}$$

10.3389/fendo.2023.1261088

where D is the number of the latent variables extracted by the VAE, and $\mu_d^{(i)}$ and $(\sigma_d^{(i)})^2$ are the approximate mean and variance of the posterior distribution of d-th latent variable for i-th subject.

We extend the VAE from single-view input into multi-view input fashion. Notably, as the fact that the product of Gaussian distributions is also a Gaussian distribution, we apply the PoE to generate the common latent space for the variation inference with an analytical solution. Suppose that under the multi-view setting, we have the data in M views, $x_1, x_2, \ldots x_M$. For the data in m-th view (m=1,…, M), a nonlinear function implemented by a neural network is employed as the encoder, denoted as $q_{\phi_m}(z_m \mid x_m^{(i)})$, where ϕ_m represents the learnable parameters of the nonlinear function for m-th view. For each encoder, we estimate the mean vector and the variance-covariance matrix of multivariate Gaussian distribution for the approximate posterior distribution, denoted as $\mu_m^{(i)}$ and $\Sigma_m^{(i)}$ for i-th subject, and we assume $\mu_m^{(i)} \in \mathbb{R}^D$ is a vector and $\Sigma_m^{(i)} \in \mathbb{R}^{D\times D}$ is a diagonal matrix where D is the dimension of the latent space. In our implementation, we employ multi-layer perceptron (MLP) as the encoder. To guarantee the positivity of the covariance, the output of the MLP is denoted as the $\log \Sigma_m^{(i)}$ first and then is converted to $\Sigma_m^{(i)}$ using the exponential function. Formally, the encoder is defined in Eq. 8.

$$\begin{aligned} q_{\phi_m}\left(z_m|x_m^{(i)}\right) &= \mathcal{N}\left(\mu_m^{(i)}, \Sigma_m^{(i)}\right) \\ &= \frac{1}{(2)^{D/2}\sqrt{\left|\Sigma_m^{(i)}\right|}} \exp\left(-\frac{1}{2}\left(z_m - \mu_m^{(i)}\right)^T \left(\Sigma_m^{(i)}\right)^{-1}\left(z_m - \mu_m^{(i)}\right)\right) \\ \mu_m^{(i)} &= MLP_m^{\mu}\left(x_m^{(i)}\right) \\ \Sigma_m^{(i)} &= \exp\left(MLP_m^{\Sigma}\left(x_m^{(i)}\right)\right) \end{aligned} \tag{8}$$

where z_m is the latent variable extracted by m-th view with the dimension of $D \times 1$. MLP_m^{μ} and MLP_m^{Σ} are the neural networks for calculating mean and covariance, respectively. Let $T_m^{(i)} = (\Sigma_m^{(i)})^{-1}$, then the multivariate Gaussian distribution for m-th view is rewritten as Eq. 9.

$$\begin{aligned} &q_{\phi_m}\left(z_m|x_m^{(i)}\right) \\ &= \frac{1}{(2\pi)^{D/2}\sqrt{\left|\Sigma_m^{(i)}\right|}} \exp\left(-\frac{1}{2} z_m^T T_m^{(i)} z_m + \left(\mu_m^{(i)}\right)^T T_m^{(i)} z_m + \Delta_m^{(i)}\right) \end{aligned} \tag{9}$$

where $\Delta_m^{(i)} = -\frac{1}{2}(\mu_m^{(i)})^T T_m^{(i)} \mu_m^{(i)} - \frac{D}{2}\log 2\pi + \frac{1}{2}\log|T_m^{(i)}|$.

A PoE models the target posterior distribution of the common latent variable from multi-view as the product of the individual posterior distribution of the latent variable from single-view. According to Eq. 9, $\Delta_m^{(i)}$ is not related to the latent variable z_m. Therefore, for the following analysis, $\Delta_m^{(i)}$ is considered as a constant. In our MVAE-PoE, the PoE generates the common latent variable z using Eq. 10.

$$q_\phi\left(z|x_1^{(i)} \cdots x_M^{(i)}\right) = \frac{1}{M}\prod_{m=1}^{M} q_{\phi_m}\left(z_m|x_m^{(i)}\right) \tag{10}$$

That is, the multivariate Gaussian distribution of the common latent variable is defined by the product of the multivariate Gaussian distribution of the latent variable extracted by m-th view. According to (51), the approximated posterior distribution of the common latent variable, z, is shown in Eq. 11.

$$\begin{aligned} q_\phi\left(z \mid x_1^{(i)} \cdots x_M^{(i)}\right) &= \mathcal{N}\left(\mu_z^{(i)},\ \Sigma_z^{(i)}\right), \\ \mu_z^{(i)} &= \left(\sum_{m=1}^{M} (\mu_m^{(i)})^T T_m^{(i)}\right)\left(\sum_{m=1}^{M} T_m^{(i)}\right)^{-1} \\ \Sigma_z^{(i)} &= \left(\sum_{m=1}^{M} T_m^{(i)}\right)^{-1} \end{aligned} \tag{11}$$

where $\mu_z^{(i)}$ and $\Sigma_z^{(i)}$ are the mean vector and variance-covariance matrix of the approximated posterior distribution of common latent variable for i-th subject. To make the neural network differentiable, we adopt the reparameterization trick (50, 52) to reparametrize the mean vector and the diagonal variance-covariance matrix of the multivariate Gaussian distribution, as shown in Eq. 12.

$$z^{(i)} = \mu_z^{(i)} + \left(\Sigma_z^{(i)}\right)^{1/2} \odot \epsilon_z \tag{12}$$

where $\epsilon_z \sim \mathcal{N}(0, I)$ and $\odot$ indicates the element-wise product. For each view, we employed MLP layers as the decoder to restore the features, which is denoted as MLP_m^{dec} for m-th view. The graphical architecture of the proposed MVAE-PoE model is shown in Figure 1.

2.6 Loss function

Using the multivariate Gaussian distribution, the ELOB for MVAE-PoE is derived in an explicit form, shown in Eq. 13.

$$\begin{aligned} \mathcal{L}(\theta, \phi; x_1, \cdots, x_M) &= \sum_{i=1}^{N}\sum_{m=1}^{M} \mathbb{E}_{z\sim q_\phi(z\mid x_m^{(i)})} \log p_\theta\left(x_m^{(i)}|z\right) \\ &- \sum_{i}^{N} D_{KL}\left(q_\phi\left(z|x_1^{(i)}, \cdots, x_M^{(i)}\right) \,\|\, p_\theta\left(z|x_1^{(i)}, \cdots, x_M^{(i)}\right)\right) \end{aligned} \tag{13}$$

The first term in the RHS of Eq. 13 is defined as the cross-entropy between the reconstructed data and the original input, and the second term in the RHS of Eq. 13 is the KL-divergence between the approximated posterior distribution and the true posterior distribution. The analytical form of the KL-divergence is the same as Eq. 3 since we employ the multivariate Gaussian distribution with an approximately diagonal variance-covariance as the ground truth. Thus, the close-form solution for the loss function is shown in Eq. 14.

$$\begin{aligned} \mathcal{L}(\theta, \phi; x_1, \cdots, x_M) &= \sum_{i=1}^{N}\sum_{m=1}^{M}\left(x_m^{(i)} \log\left(\hat{x}_m^{(i)}\right) + \left(1 - x_m^{(i)}\right)\log\left(1 - \hat{x}_m^{(i)}\right)\right) \\ &- \left(\frac{1}{2}\sum_{i=1}^{N}\sum_{d=1}^{D}\left(\left(\mu_d^{(i)}\right)^2 + \left(\sigma_d^{(i)}\right)^2 - \log\left(\sigma_d^{(i)}\right)^2 - 1\right)\right) \end{aligned} \tag{14}$$

where $\mu_d^{(i)}$ and $(\sigma_d^{(i)})^2$ are the approximate mean and variance of the posterior distribution of d-th latent variable for i-th subject, and

$\hat{x}_m^{(i)}$ represents the reconstructed feature vector for i-th subject from m-th view.

2.7 Model training and evaluation

20% of the subjects are randomly chosen as the test set, and the rest of the data are used as the training set. Predicting the proximal femoral strength is treated as a regression task. As shown in Figure 1, a linear regression model is employed to predict the proximal femoral strengths using the extracted latent variables, z.

For model evaluation, mean absolute error (MAE), mean absolute percentage error (MAPE), root mean squared error (RMSE) and R^2-score are employed. The definitions of MAE, MAPE, RMSE and R^2-score are shown in Eqs. 11-14.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{15}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \times 100\,\% \tag{16}$$

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(y_i - \hat{y}_i)^2} \tag{17}$$

$$R^2\text{-score} = 1\text{-}\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \sum_{i=1}^{N}y_i/N)^2} \tag{18}$$

where y_i is the ground truth of the proximal femoral strength and $\hat{y}_i$ is the model prediction. A lower MAE/MAPE/RMSE and a higher R^2 -score indicate better performance. According to Eqs. 15-18, 0 of MAE/MAPE/RMSE indicates the perfect match. According to Eq. 18, R^2 -score ranges from -∞ to 1, where 1 indicates the perfect match.

2.8 Interpretability of feature significance

Similar to (53), a leave-one-out technique is adopted to identify the feature significance in each view. A feature is significant if the performance of predicting proximal femoral strength decreases significantly when this feature is replaced by zero. By ranking the performance drops, the significance of the feature is obtained.

3 Results

3.1 Data processing results

We performed GWAS analysis for testing the association between each of the three types of proximal femoral strengths, including linear fall fracture load (LF), nonlinear fall fracture load (NLF), and nonlinear stance fracture load (NLS), and each of the single nucleotide polymorphisms (SNPs) after quality control. Manhattan plots of these three types of proximal femoral strengths are depicted in Figure 2. Since the sample size of the Louisiana Osteoporosis Study (LOS) cohort was relatively small for genetic association studies, we expanded our search space to look at a much wider landscape of associations by selecting top 256 SNPs with the lowest p-values to extract the WGS features that are associated with each phenotype. These identified SNPs were used as WGS features for the downstream task. Meanwhile, 196 DXA-derived imaging features were employed.

3.2 Model performance for proximal femoral strength prediction

We trained and tested the MVAE-PoE model using our workstation with a NVIDIA RTX 3090 GPU and an Intel core I9 CPU. The designed models were implemented using TensorFlow 2.5. We performed the grid search to optimize hyperparameters, and the searching space included the number of MLP layers in both encoder and decoder: 1, 2, 3; the dimension of common latent space: 32, 48, 64, 128 or 256; and the number of hidden units for each MLP layer: 32, 48, 64, 128, or 256. Table 3 shows the best performance achieved using our proposed MVAE-PoE model for predicting three proximal femoral strengths. Also, we performed experiments using the different combinations of these three views to test the effectiveness of information fusion. We plotted the model prediction and the ground truth of the three FEA-computed proximal femoral strengths with the best performance in Figure 3. In Figure 3, each subject is represented by a blue dot. The vertical axis is the predicted strength. The red dashed line indicates a perfect match, and the green dashed line is the linear regression result of the prediction.

Integrating information from two views significantly improved the performance of predicting proximal femoral strength. According to Table 3, the proposed MVAE-PoE model achieved its best performance for LF, NLF and NLS prediction using WGS features and DXA-derived imaging features. For example, the proposed model improved the R^2-score to 0.5569 compared with 0.4866 using DXA features alone for LF prediction.

DXA-derived imaging features are significantly more important than the WGS features in terms of the prediction performance. For LF, NLF and NLS, using DXA features alone, the designed models achieved the MAPEs of 18.04%, 6.84%, and 7.95%, respectively. Integrating DXA features with WGS features, the MAPEs were lowered by 1.62%, 0.55% and 0.46%, respectively. This finding was consistent with clinical practice that DXA-derived imaging features correlate with bone weakness and fragility fracture (10), with site-specific DXA explaining approximately 55% of the variability in predicting proximal femoral strengths (54).

Proximal femoral strengths depend on WGS features; however, using WGS features alone, the model does not generate satisfactory prediction results. For predicting LF, using only WGS features increased the MAPE from 18.04% to 24.47%. For predicting NLF, using only WGS increased the RMSE from 363.58 to 466.02.

10.3389/fendo.2023.1261088

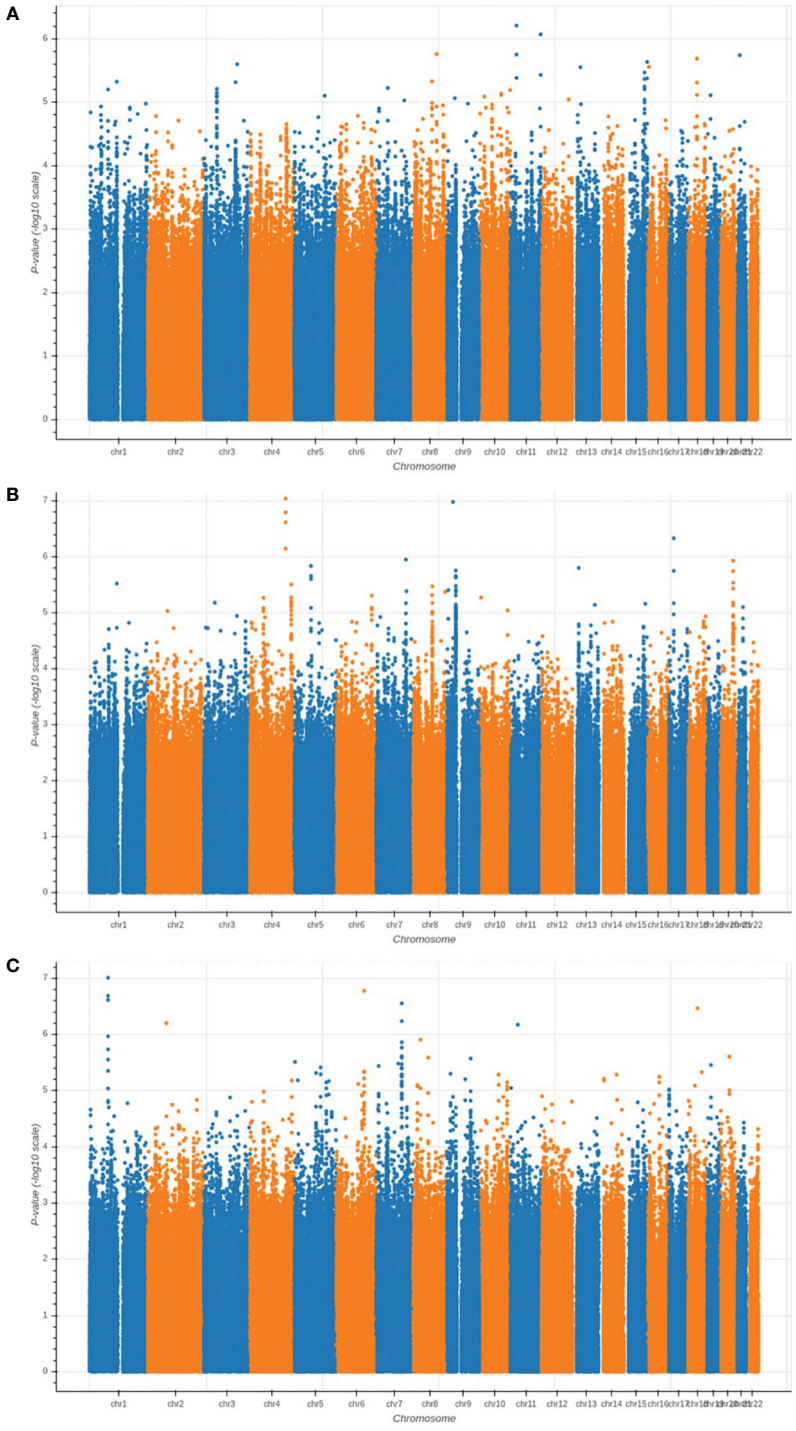

FIGURE 2
Manhattan plots of the GWAS results for **(A)** LF; **(B)** NLF; and **(C)** NLS. The horizontal axis represents the chromosome index and the positions of the SNPs; while the vertical axis represents the p-value of GWAS results for each SNP.

10.3389/fendo.2023.1261088

TABLE 3 Fine-tuned best performance for the prediction of three proximal femoral strengths.

Phenotype	WGS	DXA	Number of MLP layers	Number of hidden units	Dimension of latent space	R^2 – *score* ↑	RMSE ↓	MAE ↓	MAPE ↓
LF	✓	✓	3	48	128	**0.5569**	**468.77**	**355.57**	**18.04%**
		✓	2	256	48	0.4866	504.56	388.65	19.66%
	✓		3	128	128	0.3317	575.68	453.11	24.47%
NLF	✓	✓	2	128	128	**0.5726**	**363.58**	**284.32**	**6.84%**
		✓	2	64	48	0.4778	401.92	306.88	7.39%
	✓		2	128	128	0.2979	466.02	368.17	8.89%
NLS	✓	✓	3	256	48	**0.7107**	**1903.58**	**1441.42**	**7.95%**
		✓	3	32	48	0.6822	1995.16	1539.79	8.41%
	✓		2	256	64	0.2194	3126.87	2430.34	13.94%

The check marks in WGS features and DXA-derived features indicate that the corresponding view was used. The symbol ↑ indicates that higher is better and the symbol ↓ indicates that lower is better. If only one view was enrolled, then MVAE-PoE was degraded into a standard VAE model. For each type of the proximal femoral strength, the performance is sorted by R^2 -score. The bold values indicate the achieved best performance.

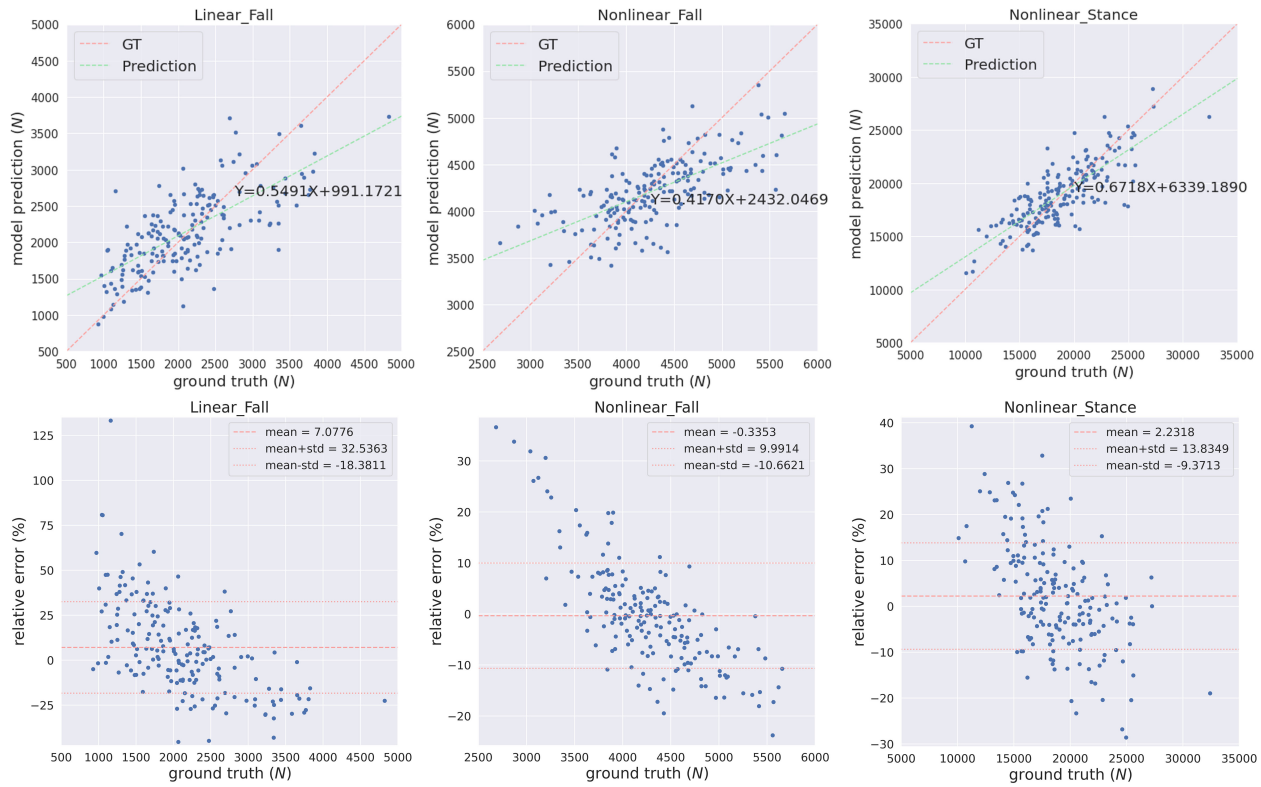

FIGURE 3
The predicted FEA-computed proximal femoral strengths and the ground truth (GT, in the horizontal axis) for LF, NLF and NLS. Each subject is represented by a blue dot. The vertical axis is the predicted strength. The red dashed line indicates a perfect match, and the green dashed line is the linear regression result of the prediction.

10.3389/fendo.2023.1261088

3.3 Performance comparison

We compared our MVAE-PoE model to other multi-view integration methods for prediction tasks. The tested models include:

- Multiview canonical correlation analysis (MCCA) (55). MCCA extends the canonical correlation analysis (CCA) into multi-view settings. CCA is a typical subspace learning algorithm, aiming at finding the pairs of projections from different views with the maximum correlations. For more than 2 views, MCCA optimizes the sum of pairwise correlations.
- Kernel CCA (KCCA) (56). KCCA is based on MCCA, however, it adds a centered Gram matrix to perform the nonlinear transformation on the input data.
- Kernel generalized CCA (KGCCA) (57). KGCCA extends KCCA with a priori-defined graph connections between different views.
- Sparse CCA (SCCA) (58): SCCA is a method for penalized CCA, which computes a rank-K approximation for a set of matrices and generates the sparse vectors for feature representation and interpretation.
- Multiview adversarial autoencoder (AAE) (59). One limitation of the variational autoencoder is that the prior distribution and posterior distribution are required to be pre-defined, and the KL-divergence is required to be differentiable. The AAE can use arbitrary priors to train the autoencoder.

For the above algorithms, a linear regression estimator was applied to perform the prediction task using the extracted latent variables. The overall performance comparison for predicting the three proximal femoral strengths is shown in Table 4. For the compared algorithms, the grid search was also performed to find the best hyperparameters.

Compared to other multi-view information extraction models, the proposed MVAE-PoE achieved the best performance for predicting all types of proximal femoral strengths. The MCCA, KCCA, KGCCA and SCCA are four machine learning-based methods and the AAE is a deep learning-based method. For the MCCA, KCCA, KGCCA and SCCA, we trained these models with different dimensions of latent variables; for KCCA and KGCCA, we further tested the linear, polynomial and radial basis function (RBF) kernels. Even with tremendous hyperparameter fine-tuning, these machine learning-based methods didn't generate better performance than the designed MVAE-PoE models. For the AAE model, we employed the same grid search settings. However, the achieved MAPEs were 20.68%, 7.22% and 10.27% for LF, NLF and NLS prediction, which indicated inferior performance than MVAE-PoE. The CCA-based methods have been commonly used in data fusion or integration; however, CCA-based methods treat the modalities as linearly and multivariately correlated without

TABLE 4 Comparison of hip fracture load prediction between existing multi-view information extraction algorithms and the proposed MVAE-PoE.

Phenotype	Method	R^2 –score ↑	RMSE ↓	MAE ↓	MAPE ↓
LF	MCCA	0.2901	596.34	440.39	22.93%
	KCCA	0.5212	489.77	391.58	20.43%
	KGCCA	0.4215	538.35	423.28	22.34%
	SCCA	0.5346	469.62	371.79	19.26%
	AAE	0.4563	519.26	403.58	20.68%
	MVAE-PoE	**0.5569**	**468.77**	**355.57**	**18.04%**
NLF	MCCA	0.2352	486.30	372.19	8.94%
	KCCA	0.4095	427.31	337.09	8.13%
	KGCCA	0.2541	480.23	365.32	8.76%
	SCCA	0.4758	402.60	309.34	7.48%
	AAE	0.5466	374.50	302.68	7.22%
	MVAE-PoE	**0.5726**	**363.58**	**284.32**	**6.84%**
NLS	MCCA	0.3980	2739.70	1872.63	10.14%
	KCCA	0.5958	2244.82	1700.97	9.38%
	KGCCA	0.5178	2451.87	1796.35	9.86%
	SCCA	0.6652	2043.08	1584.97	8.71%
	AAE	0.5998	2238.89	1820.15	10.27%
	MVAE-PoE	**0.7107**	**1903.58**	**1441.42**	**7.95%**

For each algorithm, the WGS features, and DXA-derived image features were used. Only the results with the best performance achieved by different algorithms are listed. The bold values represent the highest level of performance achieved.

considering the direction of the linear relationship (60). In this study, we demonstrate that MVAE-PoE enables more useful and generalizable representations by capturing the abstract relationship between the views for downstream tasks such as prediction tasks.

Our model, which has shown excellent performance in predicting proximal femoral strength by integrating information from multiple views, holds promise for other radiogenomics data analysis problems, such as cancer prediction. By combining radiological imaging features with genomic data, our model can uncover valuable insights into the development, progression, and treatment response of diseases. Leveraging the power of our model, we believe it can effectively analyze radiogenomics data to enhance prediction accuracy and contribute to advancements in personalized medicine.

4 Discussion

4.1 Feature importance analysis

We applied the leave-one-out method to determine the feature importance. The leave-one-out indicated that we replaced one specific feature by zero for each subject in the test set when evaluating this feature, and the replaced features are named as zero-filled features. We compared the MAE changes between using the raw features and the zero-filled features. If the MAE between the GT and the model prediction increased significantly, then the evaluated feature was a significant feature. For each model, we listed the top 15 most important features in Figure 4.

For the prediction of LF, according to Figure 4A, 10 of the top 15 most important features were WGS features, and 5 feature was a the DXA feature. The trochanter BMD (TROCH_BMD) was the most important DXA feature. Trochanteric BMD is associated with trochanteric fracture in the elderly and is among the best predictors of femoral strength (61). For NLF prediction, according to Figure 4B, one of the top 15 most significant features were DXA-derived features and the remaining 14 were WGS features. For NLS prediction, according to Figure 4C, 10 of the top 15 most significant features were DXA-derived features. This was consistent with the previous findings that DXA features explained approximately 55% of the proximal femoral strength while the proximal femoral strength was also influenced by genetics (25).

For each important WGS feature, we mapped the SNPs into the corresponding genes. The correspondingly associated clinical traits that were reported in the GWAS Catalog between each mapped gene and clinical traits are shown in Table 5. For LF, According to the meta-analysis using 339,224 subjects from 125 subjects including African Americans and Caucasians, TSPAN12 show a positive correlation with BMI (62). COX6C and CAPG showed a strong correlation with body-shape indices on subjects from UK Biobank datasets (63). For NLF, the detected most important genes, CAPG, also showed a strong correlation with body fat distribution (64). For NLS, ERBB4 was associated with obesity on subjects from UK Biobank (65), which contained 339,244 individuals.

4.2 Clinical application

Our proposed research not only addressed issues related to multi-view information fusion, but also leveraged the value of widely used DXA with information provided by genetic markers for predicting proximal femoral strength. Therefore, this study has the potential to significantly impact both research and clinical practice. In the AGES-

TABLE 5 Previously reported SNPs associated with bone-related clinical traits.

Phenotype	SNP	Nearest Gene	Clinical traits
LF	rs120932767	TSPAN12	BMI
	rs16897960	COX6C	Body-shape index
NLF	rs142460654	CAPG	Body-shape index, Body Fat
NLS	rs17334842	ERBB4	Obesity

The corresponding gene is listed.

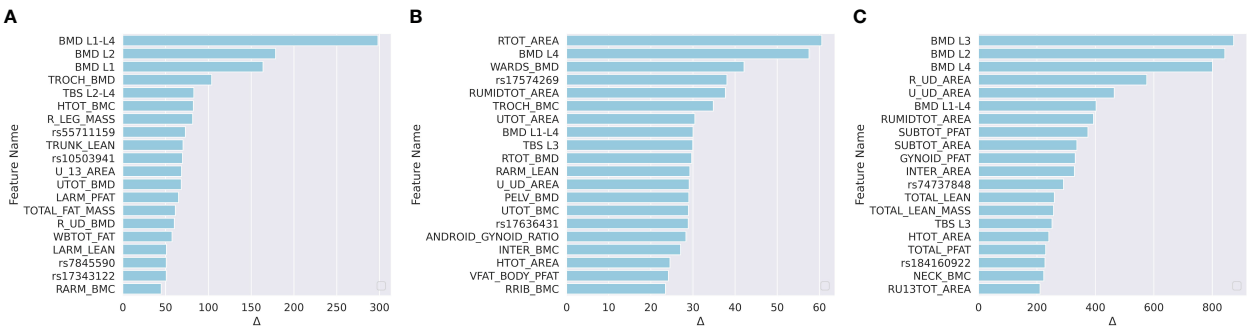

FIGURE 4
Ranked feature importance for **(A)** predicting LF; **(B)** predicting NLF; and **(C)** predicting NLS. Feature significance was determined by MAE changes between using raw features and zero-filled features. The vertical axis indicates the feature names, where WGS features are annotated by rsid, and the DXA are annotated by the abbreviations. Detailed explanations for the DXA features are shown in Table S1.

Reykjavik data set, Fleps et al. demonstrated that using FEA-computed hip fracture load to predict hip fracture was better than using total femoral aBMD only (22). In addition, genetic markers are important for identifying subjects at risk of hip fracture through effects on proximal femoral strength (25, 66, 10; 21). Due to their biological nature, genetic factors may also control nano-level bone mechanical properties and may further facilitate the prediction of bone strength and the assessment of hip fracture risk.

The most significant scientific impact of this study is the development and validation of the first comprehensive and accurate model for patient-specific assessment of predicting proximal femoral strength using multi-view information fusion by deep learning. Our multi-view deep learning-based model incorporates WGS features and DXA-derived imaging features, which are directly or indirectly related to proximal femoral strength and hip fracture. Deep learning-based techniques can automatically extract features and build accurate prediction models. Further, using the leave-one-out technique, the designed models are highly interpretable, leading to the identification of specific factors predictive of proximal femoral strengths.

The most practical clinical impact is the development and validation of an interpretable prediction model for proximal femoral strength using WGS features and DXA-derived image features, rather than using QCT. It is difficult to implement QCT-based femoral strength and hip fracture risk assessment in clinical practice due to the high radiation dosage and limited availability of QCT-based FEA. Lochmüller et al. suggested that clinical assessment of femoral fracture risk should preferably rely on femoral DXA (54). Our results suggest that there is a strong potential for using a combination of DXA and genetic markers to develop practical models for hip fracture risk assessment in the future.

4.3 Limitation

Our study primarily focused on a population aged between 20 and 51 years. While this age group provides valuable insights into proximal femoral strength prediction using DXA and WGS, it is important to acknowledge that exclusion of older subjects may constrain the generalizability of our findings to the elderly population. Hence, future research endeavors should consider incorporating a more diverse age range to enhance the applicability of our predictive model across a broader population spectrum.

Data availability statement

Data available on request due to privacy/ethical restrictions.

Ethics statement

All participants signed an informed-consent document before any data collection, and the study was approved by the Tulane University Institutional Review Board.

Author contributions

WZ: Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing. CZ: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. JK: Conceptualization, Software, Writing – original draft, Writing – review & editing. XC: Conceptualization, Writing – original draft, Writing – review & editing. QS: Methodology, Writing – original draft, Writing – review & editing. LW: Data curation, Validation, Writing – review & editing. ZL: Data curation, Resources, Writing – review & editing. LZ: Data curation, Resources, Writing – review & editing. QT: Data curation, Resources, Writing – review & editing. MS: Data curation, Resources, Writing – review & editing. CQ: Data curation, Resources, Writing – review & editing. KS: Data curation, Resources, Writing – review & editing. HS: Funding acquisition, Supervision, Writing – original draft, Writing – review & editing. HD: Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported in part by grants from the National Institutes of Health, USA (P20GM109036, R01AR069055, U19AG055373, R01AG061917, R01AR27065, R01AG028832, R01AR46197, R01AR064140 and M01RR00585) and NASA Johnson Space Center, USA contracts NNJ12HC91P and NNJ15HP23P. It was also supported in part by a seed grant from Michigan Technological University Institute of Computing and Cybersystems, a graduate fellowship from Michigan Technological University Health Research Institute and a graduate fellowship from Portage Health Foundation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fendo.2023.1261088/full#supplementary-material

References

1. Veronese N, Maggi S. Epidemiology and social costs of hip fracture. *Injury* (2018) 49:1458–60. doi: 10.1016/j.injury.2018.04.015

2. Razi AE, Hershman SH. *Vertebral Compression Fractures in Osteoporotic and Pathologic Bone: A Clinical Guide to Diagnosis and Management*. Cham: Springer International Publishing (2020).

3. Melton LJ, Kearns AE, Atkinson EJ, Bolander ME, Achenbach SJ, Huddleston JM, et al. Secular trends in hip fracture incidence and recurrence. *Osteoporos Int* (2009) 20:687–94. doi: 10.1007/s00198-008-0742-8

4. Cooper C, Atkinson EJ, Jacobsen SJ, O'Fallon WM, Melton LJ. Population-based study of survival after osteoporotic fractures. *Am J Epidemiol* (1993) 137:1001–5. doi: 10.1093/oxfordjournals.aje.a116756

5. Leibson CL, Tosteson ANA, Gabriel SE, Ransom JE, Melton LJ. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. *J Am Geriatr Soc* (2002) 50:1644–50. doi: 10.1046/j.1532-5415.2002.50455.x

6. Magaziner J, Lydick E, Hawkes W, Fox KM, Zimmerman SI, Epstein RS, et al. Excess mortality attributable to hip fracture in white women aged 70 years and older. *Am J Public Health* (1997) 87:1630–6. doi: 10.2105/ajph.87.10.1630

7. Keyak JH, Rossi SA, Jones KA, Skinner HB. Prediction of femoral fracture load using automated finite element modeling. *J Biomechanics* (1998) 31:125–33. doi: 10.1016/S0021-9290(97)00123-1

8. Blake GM, Fogelman I. The role of DXA bone density scans in the diagnosis and treatment of osteoporosis. *Postgrad Med J* (2007) 83:509–17. doi: 10.1136/pgmj.2007.057505

9. Genant HK, Engelke K, Prevrhal S. Advanced CT bone imaging in osteoporosis. *Rheumatol (Oxford)* (2008) 47 Suppl 4:iv9–16. doi: 10.1093/rheumatology/ken180

10. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* (1996) 312:1254–9. doi: 10.1136/bmj.312.7041.1254

11. Chang G, Honig S, Brown R, Deniz CM, Egol KA, Babb JS, et al. Finite element analysis applied to 3-T MR imaging of proximal femur microarchitecture: lower bone strength in patients with fragility fractures compared with control subjects. *Radiology* (2014) 272:464–74. doi: 10.1148/radiol.14131926

12. Ralston SH. Genetic control of susceptibility to osteoporosis. *J Clin Endocrinol Metab* (2002) 87:2460–6. doi: 10.1210/jcem.87.6.8621

13. Hernandez-de Sosa N, Athanasiadis G, Malouf J, Laiz A, Marin A, Herrera S, et al. Genetic contribution of femoral neck bone geometry to the risk of developing osteoporosis: A family-based study. *PloS One* (2016) 11:e0154833. doi: 10.1371/journal.pone.0154833

14. Lochmüller E-M, Zeller J-B, Kaiser D, Eckstein F, Landgraf J, Putz R, et al. Correlation of femoral and lumbar DXA and calcaneal ultrasound, measured in *situ* with intact soft tissues, with the in *vitro* failure loads of the proximal femur. *Osteoporosis Int* (1998) 8:591–8. doi: 10.1007/s001980050104

15. Mitchell JA, Cousminer DL, Zemel BS, Grant SF, Chesi A. Genetics of pediatric bone strength. *Bonekey Rep* (2016) 5:823. doi: 10.1038/bonekey.2016.50

16. Keyak JH. Improved prediction of proximal femoral fracture load using nonlinear finite element models. *Med Eng Phys* (2001) 23:165–73. doi: 10.1016/S1350-4533(01)00045-5

17. Keyak JH, Kaneko TS, Tehranzadeh J, Skinner HB. Predicting proximal femoral strength using structural engineering models. *Clin Orthopaedics Related Res* (2005) 437:219–28. doi: 10.1097/01.blo.0000164400.37905.22

18. Keyak JH, Sigurdsson S, Karlsdottir GS, Oskarsdottir D, Sigmarsdottir A, Kornak J, et al. Effect of finite element model loading condition on fracture risk assessment in men and women: The AGES-Reykjavik study. *Bone* (2013) 57:18–29. doi: 10.1016/j.bone.2013.07.028

19. Keyak JH, Kaneko TS, Khosla S, Amin S, Atkinson EJ, Lang TF, et al. Hip load capacity and yield load in men and women of all ages. *Bone* (2020) 137:115321. doi: 10.1016/j.bone.2020.115321

20. Link TM, Lang TF. Axial QCT: clinical applications and new developments. *J Clin Densitometry* (2014) 17:438–48. doi: 10.1016/j.jocd.2014.04.119

21. Skuladottir SS, Ramel A, Hjaltadottir I, Launer LJ, Cotch MF, Siggeirsdottir K, et al. Characteristics of incidence hip fracture cases in older adults participating in the longitudinal AGES-Reykjavik study. *Osteoporos Int* (2021) 32:243–50. doi: 10.1007/s00198-020-05567-x

22. Fleps I, Pálsson H, Baker A, Enns-Bray W, Bahaloo H, Danner M, et al. Finite element derived femoral strength is a better predictor of hip fracture risk than aBMD in the AGES Reykjavik study cohort. *Bone* (2022) 154:116219. doi: 10.1016/j.bone.2021.116219

23. Yang C-C, Nagarajan MB, Huber MB, Carballido-Gamio J, Bauer JS, Baum T, et al. Improving bone strength prediction in human proximal femur specimens through geometrical characterization of trabecular bone microarchitecture and support vector regression. *J Electron Imaging* (2014) 23:13013. doi: 10.1117/1.JEI.23.1.013013

24. Gong H, Zhang M, Jia S, Lv L. The relationship between orthopedic clinical imaging and bone strength prediction. *Med Novel Technol Devices* (2021) 9:100060. doi: 10.1016/j.medntd.2021.100060

25. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* (2007) 165:1076–87. doi: 10.1093/aje/kwk115

26. Richards JB, Zheng H-F, Spector TD. Genetics of osteoporosis from genome-wide association studies: advances and challenges. *Nat Rev Genet* (2012) 13:576–88. doi: 10.1038/nrg3228

27. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, et al. Multiple genetic loci for bone mineral density and fractures. *N Engl J Med* (2008) 358:2355–65. doi: 10.1056/NEJMoa0801197

28. Estrada K, Styrkarsdottir U, Evangelou E, Hsu Y-H, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* (2012) 44:491–501. doi: 10.1038/ng.2249

29. Al-Barghouthi BM, Mesner LD, Calabrese GM, et al. Systems genetics in diversity outbred mice inform BMD GWAS and identify determinants of bone strength. *Nat Commun* (2021) 12:3408. doi: 10.1038/s41467-021-23649-0

30. van Meurs JB, Rivadeneira F, Jhamai M, Hugens W, Hofman A, van Leeuwen JP, et al. Common genetic variation of the low-density lipoprotein receptor-related protein 5 and 6 genes determines fracture risk in elderly white men. *J Bone Miner Res* (2005) 21:141–50. doi: 10.1359/JBMR.050904

31. Cody DD, Gross GJ, Hou FJ, Spencer HJ, Goldstein SA, Fyhrie DP. Femoral strength is better predicted by finite element models than QCT and DXA. *J Biomech* (1999) 32:1013–20. doi: 10.1016/s0021-9290(99)00099-8

32. Cummings SR, Karpf DB, Harris F, Genant HK, Ensrud K, LaCroix AZ, et al. Improvement in spine bone density and reduction in risk of vertebral fractures during treatment with antiresorptive drugs. *Am J Med* (2002) 112:281–9. doi: 10.1016/S0002-9343(01)01124-X

33. Dufresne TE, Chmielewski PA, Manhart MD, Johnson TD, Borah B. Risedronate preserves bone architecture in early postmenopausal women in 1 year as measured by three-dimensional microcomputed tomography. *Calcified Tissue Int* (2003) 73:423–32. doi: 10.1007/s00223-002-2104-4

34. Melton LJ, Chrischilles EA, Cooper C, Lane AW, Riggs BL. Perspective how many women have osteoporosis? *J Bone Miner Res* (2009) 7:1005–10. doi: 10.1002/jbmr.5650070902

35. Yang T-L, Guo Y, Li J, Zhang L, Shen H, Li SM, et al. Gene-gene interaction between RBMS3 and ZNF516 influences bone mineral density. *J Bone Miner Res* (2013) 28:828–37. doi: 10.1002/jbmr.1788

36. Yang T-L, Shen H, Liu A, Dong S-S, Zhang L, Deng F-Y, et al. A road map for understanding molecular and genetic determinants of osteoporosis. *Nat Rev Endocrinol* (2020) 16:91–103. doi: 10.1038/s41574-019-0282-7

37. Qiu C, Yu F, Su K, Zhao Q, Zhang L, Xu C, et al. Multi-omics data integration for identifying osteoporosis biomarkers and their biological interaction and causal mechanisms. *iScience* (2020) 23:100847. doi: 10.1016/j.isci.2020.100847

38. Greenbaum J, Su K-J, Zhang X, Liu Y, Liu A, Zhao L-J, et al. A multiethnic whole genome sequencing study to identify novel loci for bone mineral density. *Hum Mol Genet* (2022) 31:1067–81. doi: 10.1093/hmg/ddab305

39. Hui SL, Slemenda CW, Johnston CC. The contribution of bone loss to postmenopausal osteoporosis. *Osteoporos Int* (1990) 1:30–4. doi: 10.1007/BF01880413

40. Hernandez CJ, Beaupré GS, Carter DR. A theoretical analysis of the relative influences of peak BMD, age-related bone loss and menopause on the development of osteoporosis. *Osteoporos Int* (2003) 14:843–7. doi: 10.1007/s00198-003-1454-8

41. Melton LJ, Atkinson EJ, Khosla S, Oberg AL, Riggs BL. Evaluation of a prediction model for long-term fracture risk. *J Bone Miner Res* (2005) 20:551–6. doi: 10.1359/JBMR.041206

42. Bonjour J-P, Chevalley T, Ferrari S, Rizzoli R. The importance and relevance of peak bone mass in the prevalence of osteoporosis. *Salud Publica Mex* (2009) 51 Suppl 1: S5–17. doi: 10.1590/s0036-36342009000700004

43. Zhao C, Keyak JH, Tang J, Kaneko TS, Khosla S, Amin S, et al. ST-V-Net: incorporating shape prior into convolutional neural networks for proximal femur segmentation. *Complex Intelligent Syst* (2023) 9:2747–58. doi: 10.1007/s40747-021-00427-5

44. Seitz P, Ruegsegger P. Fast contour detection algorithm for high precision quantitative CT. *IEEE Trans Med Imaging* (1983) 2:136–41. doi: 10.1109/TMI.1983.4307627

45. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324

46. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* (2010) 20:1297–303. doi: 10.1101/gr.107524.110

47. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* (2006) 38:904–9. doi: 10.1038/ng1847

48. Sha Q, Zhang Z, Zhang S. An improved score test for genetic association studies. *Genet Epidemiol* (2011) 35:350–9. doi: 10.1002/gepi.20583

49. Deng H-W, Shen H, Xu F-H, Deng H-Y, Conway T, Zhang H-T, et al. Tests of linkage and/or association of genes for vitamin D receptor, osteocalcin, and parathyroid hormone with bone mineral density. *J Bone Miner Res* (2002) 17:678–86. doi: 10.1359/jbmr.2002.17.4.678

50. Kingma DP, Welling M. Auto-encoding variational bayes(2014). Available at: http://arxiv.org/abs/1312.6114 (Accessed October 11, 2022).

51. Cao Y, Fleet DJ. *Generalized Product of Experts for Automatic and Principled Fusion of Gaussian Process Predictions* (2015). Available at: http://arxiv.org/abs/1410.7827 (Accessed August 3, 2022).

52. Maddison CJ, Mnih A, Teh YW. The concrete distribution: A continuous relaxation of discrete random variables(2017). Available at: http://arxiv.org/abs/1611.00712 (Accessed October 11, 2022).

53. Wang T, Shao W, Huang Z, Tang H, Zhang J, Ding Z, et al. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nat Commun* (2021) 12:3445. doi: 10.1038/s41467-021-23774-w

54. Lochmüller E-M, Müller R, Kuhn V, Lill CA, Eckstein F. Can novel clinical densitometric techniques replace or improve DXA in predicting bone strength in osteoporosis at the hip and other skeletal sites? *J Bone Miner Res* (2003) 18:906–12. doi: 10.1359/jbmr.2003.18.5.906

55. Kettenring JR. Canonical analysis of several sets of variables. *Biometrika* (1971) 58:433–51. doi: 10.1093/biomet/58.3.433

56. Arora R, Livescu K. Kernel CCA for multi-view learning of acoustic features using articulatory measurements. *Proc. Machine Learning in Speech and Language Processing (MLSLP 2012)* (2012) 34–7.

57. Tenenhaus A, Philippe C, Frouin V. Kernel generalized canonical correlation analysis. *Comput Stat Data Anal* (2015) 90:114–31. doi: 10.1016/j.csda.2015.04.004

58. Gao C, Ma Z, Zhou HH. Sparse CCA: Adaptive estimation and computational barriers. *Ann Statist* (2017) 45:2074–101. doi: 10.1214/16-AOS1519

59. Fan W, Ma Y, Xu H, Liu X, Wang J, Li Q, et al. (2020). Deep adversarial canonical correlation analysis, in: *Proceedings of the 2020 SIAM International Conference on Data Mining (SIAM)*, . pp. 352–60.

60. Zhuang X, Yang Z, Cordes D. A technical review of canonical correlation analysis for neuroscience applications. *Hum Brain Mapp* (2020) 41:3807–33. doi: 10.1002/hbm.25090

61. Cheng XG, Lowet G, Boonen S, Nicholson PHF, Brys P, Nijs J, et al. Assessment of the strength of proximal femur in *vitro*: relationship to femoral bone mineral density and femoral geometry. *Bone* (1997) 20:213–8. doi: 10.1016/S8756-3282(96)00383-3

62. 23andMe Research Team, Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, et al. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* (2019) 51:258–66. doi: 10.1038/s41588-018-0302-x

63. Christakoudi S, Evangelou E, Riboli E, Tsilidis KK. GWAS of allometric body-shape indices in UK Biobank identifies loci suggesting associations with morphogenesis, organogenesis, adrenal cell renewal and cancer. *Sci Rep* (2021) 11:10688. doi: 10.1038/s41598-021-89176-6

64. Rask-Andersen M, Karlsson T, Ek WE, Johansson Å. Genome-wide association study of body fat distribution identifies adiposity loci and sex-specific genetic effects. *Nat Commun* (2019) 10:339. doi: 10.1038/s41467-018-08000-4

65. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* (2015) 518:197–206. doi: 10.1038/nature14177

66. Perez-Castrillon JL, Olmos JM, Nan DN, Castillo J, Arozamena J, Montero A, et al. Polymorphisms of the WNT10B gene, bone mineral density, and fractures in postmenopausal women. *Calcif Tissue Int* (2009) 85:113–8. doi: 10.1007/s00223-009-9256-4