Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Bone Miner Res. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:

J Bone Miner Res. 2022 March ; 37(3): 428–439. doi:10.1002/jbmr.4494.

Bone Microarchitecture Phenotypes Identified in Older Adults Are Associated With Different Levels of Osteoporotic Fracture Risk

Danielle E Whittier[1], **Elizabeth J Samelson**[2,3], **Marian T Hannan**[2,3], **Lauren A Burt**[1], **David A Hanley**[1], **Emmanuel Biver**[4], **Pawel Szulc**[5], **Elisabeth Sornay-Rendu**[5], **Blandine Merle**[5], **Roland Chapurlat**[5], **Eric Lespessailles**[6,7], **Andy Kin On Wong**[8,9], **David Goltzman**[10], **Sundeep Khosla**[11], **Serge Ferrari**[4], **Mary L Bouxsein**[12,13,14], **Douglas P Kiel**[2,3], **Steven K Boyd**[1]

[1]McCaig Institute for Bone and Joint Health, Cumming School of Medicine, University of Calgary, Calgary, Canada

[2]Hinda and Arthur Marcus Institute for Aging Research, Hebrew Senior Life, Boston, MA, USA

[3]Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

[4]Division of Bone Diseases, Geneva University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland

[5]INSERM UMR1033, Université de Lyon, Hôpital Edouard Herriot, Lyon, France

[6]Regional Hospital of Orleans, PRIMMO, Orleans, France

[7]EA 4708-I3MTO, University of Orleans, Orleans, France

[8]Joint Department of Medical Imaging, University Health Network, Toronto, Canada

[9]Department of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[10]Department of Medicine, McGill University and McGill University Health Centre, Quebec, Canada

[11]Kogod Center on Aging and Division of Endocrinology, Mayo Clinic, Rochester, MN, USA

[12]Center for Advanced Orthopedic Studies, BIDMC, Boston, MA, USA

[13]Endocrine Unit, Massachusetts General Hospital, Boston, MA, USA

Address correspondence to: Steven K Boyd, PhD, McCaig Institute for Bone and Joint Health, University of Calgary, 3280 Hospital Drive NW, Calgary, AB T2N 4Z6, Canada. skboyd@ucalgary.ca.

Authors' roles: DEW and SKB conceptualized and designed the study. DEW completed all data analysis and prepared the manuscript, with supervision of SKB. LAB, DAH, PS, ESR, BM, RC, EL, EB, SF, AW, DG, and SKB contributed to participant data collection. MLB, DPK, EJS, and MTH contributed to data management and interpretation of results. All authors contributed to revisions of the final manuscript.

Disclosures

DPK and MTH received funding to their institution from Amgen and Radius. DPK received royalties for publication in UpToDate from Wolters Kluwer. He serves on a Scientific Advisory Board for Solarea Bio. MH received funding to her institution from Amgen. All other authors declare that they have no conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[14]Department of Orthopedic Surgery, Harvard Medical School, Boston, MA, USA

Abstract

Prevalence of osteoporosis is more than 50% in older adults, yet current clinical methods for diagnosis that rely on areal bone mineral density (aBMD) fail to detect most individuals who have a fragility fracture. Bone fragility can manifest in different forms, and a "one-size-fits-all" approach to diagnosis and management of osteoporosis may not be suitable. High-resolution peripheral quantitative computed tomography (HR-pQCT) provides additive information by capturing information about volumetric density and microarchitecture, but interpretation is challenging because of the complex interactions between the numerous properties measured. In this study, we propose that there are common combinations of bone properties, referred to as phenotypes, that are predisposed to different levels of fracture risk. Using HR-pQCT data from a multinational cohort (n = 5873, 71% female) between 40 and 96 years of age, we employed fuzzy c-means clustering, an unsupervised machine-learning method, to identify phenotypes of bone microarchitecture. Three clusters were identified, and using partial correlation analysis of HR-pQCT parameters, we characterized the clusters as low density, low volume, and healthy bone phenotypes. Most males were associated with the healthy bone phenotype, whereas females were more often associated with the low volume or low density bone phenotypes. Each phenotype had a significantly different cumulative hazard of major osteoporotic fracture (MOF) and of any incident osteoporotic fracture ($p < 0.05$). After adjustment for covariates (cohort, sex, and age), the low density followed by the low volume phenotype had the highest association with MOF (hazard ratio = 2.96 and 2.35, respectively), and significant associations were maintained when additionally adjusted for femoral neck aBMD (hazard ratio = 1.69 and 1.90, respectively). Further, within each phenotype, different imaging biomarkers of fracture were identified. These findings suggest that osteoporotic fracture risk is associated with bone phenotypes that capture key features of bone deterioration that are not distinguishable by aBMD.

Keywords

BONE; HIGH-RESOLUTION PERIPHERAL COMPUTED TOMOGRAPHY; PHENOTYPE; MACHINE LEARNING; OSTEOPOROSIS; FRACTURE RISK

Introduction

Osteoporosis is an age-related disease characterized by compromised bone strength, leading to increased fracture risk.[1,2] The lifetime risk of osteoporotic fragility fracture is very high, estimated to be 40% to 50% for women and 13% to 22% for men,[3,4] and because of an aging population, the number of osteoporotic fractures is expected to increase dramatically over the next two decades.[5-7] Fracture prevention is paramount, yet osteoporosis remains underdiagnosed and undertreated.[8] Current diagnosis relies on femoral neck (FN) or lumbar spine areal bone mineral density (aBMD), measured by dual-energy X-ray absorptiometry (DXA),[1,9] and although low aBMD is correlated with greater fracture risk,[10] it lacks sensitivity because it does account for information about bone structure that contributes to bone strength.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

High-resolution peripheral quantitative computed tomography (HR-pQCT) provides additive information about bone macrostructure and microarchitecture,[11] and our recent prospective study demonstrated that bone parameters measured using HR-pQCT provide additive predictive information to FN aBMD when assessing fracture risk.[12] However, interpretation is challenging because assessing HR-pQCT parameters individually overlooks the complex interactions between parameters and does not take into consideration that bone fragility may manifest in different forms. Previous studies have proposed that different trajectories of age-related bone loss related to bone shape exist, indicating that a "one-size-fits-all" approach to fracture risk assessment largely dependent on aBMD may not be suitable[13-15] Single-center studies have also explored using HRpQCT to identify clusters of bone microarchitecture properties associated with fracture, and although promising, these studies have been limited because of low sample size and lack of prospective fracture information.[16-18]

The identification of common combinations of bone properties, which we refer to as bone phenotypes, that are predisposed to a higher risk of fracture has potential to enhance our understanding how osteoporosis manifests across the population and improve patient-specific assessment of fracture risk. Thus, the objective of this study is to use cluster analysis with a large multicenter HR-pQCT data set to identify common phenotypes of bone microarchitecture in the population. Using the prospective fracture data of the cohort, we aim to determine whether certain bone phenotypes are associated with higher fracture risk. Further, we will investigate whether there are dominant imaging biomarkers of fracture risk that are unique to each phenotype.

Subjects and Methods

Participants

Female and male participants with individual-level data (n = 6836) were obtained from the Bone Microarchitecture International Consortium (BoMIC), the largest international HR-pQCT data set to date with prospective fracture information.[12] The BoMIC participants used in this study originate from seven international cohorts: the Framingham Study[19] and Mayo Clinic[20] in the United States; the Qualité Osseuse Lyon Orléans (QUALYOR),[21] Structure of Aging Men's Bones (STRAMBO),[22] and Os des Femmes de Lyon (OFELY)[23] in France; the Geneva Retirees Cohort (GERICO) in Switzerland;[24] and the Canadian Multicentre Osteoporosis Study (CaMos) in Canada.[25] The Swedish Osteoporotic Fractures in Men Study (MrOS)[26] cohort included in the original BoMIC data set was not used here because participant-level information was not available.[12] All individuals were aged 40 years or older and had undergone HR-pQCT and DXA imaging, anthropometric measurements, and completed standardized questionnaires regarding fracture and bone-related health history at a baseline study visit. Participants were followed until either an incident fracture occurred, death, last contact, or the cohort's study closing date.

All participants provided written informed consent for their respective cohort study, and the Institutional Review Board for Human Research at Hebrew SeniorLife approved of this study.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Medical imaging protocol

Femoral neck aBMD of participants was obtained using either Hologic (Hologic, Marlborough, MA, USA) or Lunar (GE Healthcare, Madison, WI, USA) DXA scanners, depending on the study site. Measurements of FN aBMD were standardized across DXA systems using previously reported equations.[27]

Participants were scanned at the distal radius and tibia using HR-pQCT (XtremeCT, Scanco Medical AG, Bruttisellen, Switzerland). Study sites followed the standard in vivo protocol for scan acquisition and analysis, and scan acquisition was performed by operators trained by the manufacturer.[28] Briefly, each participant had the non-dominant distal radius and one tibia scanned, and in the case of prior fractures, metal, or surgery, the contralateral side was scanned. The reference line was placed at the inflection point on the endplate of the distal radius and tibial plafond using an anteroposterior scout view. The scan region was set at a fixed offset of 9.5 mm and 22.5 mm proximal to the reference line for the radius and tibia, respectively, and extended 9.02 mm proximally, with an isotropic resolution of 82 μm (110 slices). Scans were inspected visually for motion artifacts and scored from 1 (no motion) to 5 (significant motion).[29] Any participants with one or more scans with a motion score of 4 or higher were excluded from analysis. The manufacturer's standard patient evaluation protocol (Image Processing Language, v6.0, Scanco Medical) was followed to obtain the resulting density and morphological parameters: total, cortical, and trabecular BMD (Tt.BMD, Ct.BMD, and Tb.BMD, respectively); total, cortical, and trabecular cross-sectional areas (Tt.Ar, Ct.Ar, and Tb.Ar, respectively); cortical thickness ($Ct.Th^{d}$); trabecular number (Tb.N); and trabecular inhomogeneity (Tb.1/N.SD). The superscript "d" denotes the parameters that are derived metrics in the standard analysis.[28] Cortical porosity and the direct measure of cortical thickness were not included because these require completion of the extended cortical analysis using the dual-threshold segmentation technique,[30] which was not available for all participants in the BoMIC cohort.

Failure load was estimated using linear micro finite-element (μFE) analysis. All μFE models were generated using axial compression conditions, with a 1% apparent strain applied. Failure load was defined as the load when 2% of the volume exceeded a critical strain of 0.7%.[31] The tissue modulus was set as 6829 MPa[32] and solved using custom software (FAIM, v8.0, Numerics88 Solutions Ltd., Calgary, Canada) for all cohorts except GERICO, which used a modulus of 10,000 MPa[31] and solved using software provided by the HR-pQCT manufacturer (Image Processing Language v5.16/FE v01.16, Scanco Medical AG). Results from the GERICO cohort were scaled linearly using a factor of 0.6829 to correct for the difference in tissue modulus.[33]

Unsupervised cluster analysis and validation

Bone phenotypes were identified using cluster analysis, an unsupervised machine-learning approach. Features selected for clustering included height (as a surrogate for long bone length) and HR-pQCT parameters from the radius and tibia capturing densities (Tt.BMD, Ct.BMD, and Tb.BMD), cross-sectional areas (Tt.Ar, Ct.Ar, and Tb.Ar), morphological parameters of the trabecular and cortical compartments (Tb.N, Tb.1/N.SD, and $Ct.Th^{d}$), and μFE-estimated failure load, resulting in a total of 21 features. All features were transformed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

to normalize the distribution, and the range scaled from 0 to 1. The whole BoMIC cohort, except for participants used for external validation, were transformed and scaled before clustering analysis.

The fuzzy c-means (FCM) algorithm was selected for bone phenotyping.[34,35] Other clustering methods were considered, but FCM was selected because it produces soft clusters, allowing for cluster to overlap in a way that is expected to be more representative of true phenotypic expression of bone microarchitecture. With FCM, each participant is assigned a membership coefficient to each cluster ranging from 0 (no association) to 1 (complete association). When necessary for statistical analysis, soft clusters can be discretized into hard clusters by assigning individuals to the cluster with which they have the highest membership coefficient.

The FCM model was run with a default fuzziness exponent of $m = 2$ and default squared Euclidean distance metric. The number of clusters to seed in the model was determined using agglomerative hierarchical clustering, where the centroid coordinates obtained from agglomerative clustering were used to initialize the FCM algorithm.[36] Agglomerative clustering was tested using Euclidean, maximum, Manhattan, Pearson, and Spearman distance metrics with Ward, average, and complete linkage methods.[37] Dendrograms of cluster agglomeration for each combination of distance and linkage were generated, and a subset of candidate clustering models were selected on visual inspection of the dendrograms for further internal validation and stability testing. Internal validation of the candidate models was performed using the Silhouette Index,[38] which ranges from −1 (data are poorly matched to the assigned clusters) to 1 (data are strongly matched to the assigned clusters). Cluster stability was assessed using resampling, where 80% of the data set was sub-sampled randomly 100 times and agglomerative clustering re-run.[39] The mean Jaccard Similarity Index across subsamples was computed, providing a range between 0 (no agreement across subsamples) and 1 (perfect agreement across subsamples). The hierarchical model with the highest Jaccard and Silhouette indices was chosen to initialize the FCM algorithm.

Internal validation of the final FCM model was performed by computing the Fuzzy Silhouette Index,[40] and stability of the final model was assessed by the Jaccard Similarity Index, using the same resampling method described for the candidate agglomerative models.

External validation was performed by assessing the reproducibility of classifying an individual into the same phenotype at two time points. Participants from the CaMos data set who had an additional study visit 5 years before the BoMIC baseline time point were used for external validation.[41] Soft cluster membership of the validation cohort at both time points was predicted independently using the fitted FCM cluster model. The agree-ability of soft cluster membership between time points was assessed by estimating the Pearson's (r) correlation coefficient of the cluster membership obtained at the two time points. Hard cluster classification was compared using a confusion matrix and estimation of overall accuracy and individual cluster sensitivity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Characterization of phenotypic clusters

Partial Spearman's (ρ) correlation coefficients between the cluster membership coefficients and individual bone parameters was performed to assess the distinguishing phenotypic characteristics of each cluster. Partial correlation coefficients were adjusted for membership with other clusters, and significance was set to $p < 0.05$, with adjustment for multiple comparisons using the Holm–Bonferroni method. The distribution of bone parameters and descriptive characteristics across phenotypes were visualized using density plots and heatmaps applied to plots of the cluster membership coefficients.

Fracture risk assessment

The primary outcome of interest was time to incident fracture at any osteoporotic fracture site, and the secondary outcome was time to incident fractures at only major osteoporotic fracture (MOF) sites. Osteoporotic fracture sites included in this study were the shoulder/collar, upper arm, elbow, upper forearm, wrist/lower forearm, hand, rib, spine, coccyx, pelvis, hip, lower thigh, knee, shin, ankle, foot, and heel, while MOF sites included only the forearm/wrist, upper arm, hip, and spine.

The cumulative hazard function for any osteoporotic fracture and MOF was determined for each phenotype using the Nelson–Aalen estimator.[42] Log-rank tests with post hoc testing was applied to determine if cumulative hazard curves of the phenotypes were significantly different. To account for confounding factors, we used Cox proportional hazard regression with adjustment for study cohort, sex, and age to estimate the hazard ratio (HR) and 95% confidence intervals (95% CI) for the association of each phenotype and risk of any osteoporotic fracture and MOF. The HRs were additionally adjusted for FN aBMD to evaluate whether phenotypes provided additive information to DXA in stratifying the population based on fracture risk. The proportional hazards assumption for all parameters used to estimate HR were confirmed using the Schoenfeld's residual correlation test.[43]

Finally, phenotype-specific imaging biomarkers were explored by calculating the HR and 95% CI for the association between individual bone parameters and incidence of any osteoporotic fracture and MOF within each phenotype. All HRs were standardized to per standard deviation (SD) decline and were estimated with and without adjustment for base covariates (study cohort, age, and sex) and compared with models with additional adjustment for FN aBMD.

All analysis was performed in R (v4.0.2) with specialized packages for performing clustering and validation (ppclust 1.1.0, cluster v2.1.2, fclust v2.1.1, and fpc v2.2-9) and fracture risk assessment (survival v3.2-11 and survminer v0.4.9).

Results

Unsupervised clustering and phenotype characteristics

A total of 5873 participants from the BoMIC cohort were selected after removal of missing data and scans with high motion scores (Supplemental Fig. S1). Participants were between the ages of 40 and 96 years at baseline and 70.9% females (Supplemental Table S1).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Unsupervised clustering was performed on the majority of the available participants (n = 5645), whereas the remaining participants (n = 228) were reserved for external validation of the final model. The agglomerative clustering approach identified three clusters of bone microarchitecture (Supplemental Fig. S2). The final fuzzy clustering model for bone phenotypes had an average Jaccard Similarity Index of 0.98 and Fuzzy Silhouette Index of 0.44, indicating excellent cluster stability and good within-cluster cohesion.

Partial correlations between cluster membership coefficients and bone parameters resulted in the generalized description of the three phenotypes as healthy, low density, and low volume (Fig. 1B; Table 1). The healthy phenotype had positive partial correlations with failure load, trabecular microarchitecture parameters, and cross-sectional areas (r = 0.09–0.35) and is characterized as having large bones with thick cortices and a well-connected trabecular network but not necessarily higher BMD (Fig. 2; Supplemental Fig. S3). The low density phenotype had similar overall cross-sectional area when compared with the healthy phenotype but had negative partial correlations with BMD, failure load, and all microarchitecture parameters (ρ = −0.14 to −0.58). The key characteristic of the low density phenotype was its low BMD, driven predominantly by a cortical thinning, followed by degradation of trabecular microarchitecture (Fig. 2; Supplemental Fig. S3). In contrast, the low volume phenotype had a positive association to cortical BMD and thickness (ρ = 0.04–0.25) but a negative correlation with failure load, trabecular microarchitecture, and cross-sectional areas (ρ = −0.05 to −0.54). This phenotype was subsequently characterized as having thicker and denser cortices but smaller overall bone size with deficits in the trabecular compartment that led to reduced bone strength (Fig. 2; Supplemental Fig. S3). Further, the age ranges of participants across phenotypes were highly comparable, but on average they were significantly different when compared using a two-sided analysis of variance with post hoc comparisons (adjusted $p < 0.05$). The low volume phenotype was the youngest on average (65.2 ± 8.2 years), followed by the healthy (68.0 ± 8.4 years) and low density (70.3 ± 8.1 years) phenotypes. The overall population distribution across the three phenotypes was balanced, but when stratified by sex, most males were classified in the healthy phenotype, whereas females were more often associated with the low volume or low density phenotypes (Fig. 1A).

Finally, the external validation cohort (n = 228), who had a second study visit 5 years prior, consisted of 67.1% female and ranged between 35 to 84 years of age at the earliest study visit. The descriptive characteristics and distribution of HR-pQCT parameters of the validation cohort were highly comparable to the rest of the BoMIC cohort (Fig 3A; Supplemental Fig. S3). {FIG3} The predicted hard cluster classification between time points agreed with 89.5% accuracy (Fig. 3B), and the cluster membership coefficients were strongly correlated between time points ($r \geq 0.92$) for all three phenotypes (Fig. 3C), indicating stable phenotype classification for most individuals over 5 years.

Fracture risk assessment

Participants were followed for a mean duration of 4.7 ± 2.4 years after imaging, during which time 10.4% sustained an incident fracture at any osteoporotic fracture site; approximately half of these occurred at a MOF site (Fig. 4).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The cumulative hazard of any incident osteoporotic fracture and MOF fracture over time was significantly different across the three phenotypes, with the low density phenotype having the highest rate of fracture, followed by the low volume phenotype (Fig. 5A). When the population was stratified by sex, the cumulative hazard functions across phenotypes remained similar for males and females (Supplemental Fig. S5), with the possible exception of males in the low volume phenotype, which was limited by sample size ($n = 79$).

The healthy phenotype was used as the reference for estimating the HR between fracture with the low density and low volume phenotypes (Fig. 5B). The low density phenotype had the highest association with any incident fracture before adjustment for covariates (HR = 3.28, 95% CI 2.60–4.14), which was reduced but remained significant after adjustment for base covariates and FN aBMD (adjusted HR = 1.57, 95% CI 1.12–2.20). In contrast, the low volume phenotype had a weaker association with any incident fracture before adjustment (HR = 1.99, 95% CI 1.56–2.55) but was less affected by adjustment for base covariates and FN aBMD (adjusted HR = 1.74, 95% CI 1.17–2.60). The association between fracture and phenotype was greater when considering only MOF sites, most notably for the low density phenotype (HR = 5.50, 95% CI 3.81–7.94) but also the low volume phenotype (HR = 2.70, 95% CI 1.82–4.00). When the population was stratified by sex, the HRs were higher in females when considering any fracture site (unadjusted HR = 2.29–4.01) or only MOF sites (unadjusted HR = 4.04–8.83), whereas HRs for males were lower or in some cases non-significant (Supplemental Fig. S5; Supplemental Table S2).

After stratifying the population by phenotype to explore imaging biomarkers, most bone parameters were associated with incident fracture in the low density phenotype. However, after adjustments, only two parameters remained significantly associated with MOF and an additional three when any fracture site was considered (Fig. 6; Supplemental Figs. S6 and S7). The strongest predictors of fracture were measures of BMD, specifically total and trabecular BMD (Tt.BMD and Tb.BMD, respectively) at the radius, and FN aBMD (adjusted HR = 1.46–1.71). Failure load at both the radius and tibia also had comparable association with incident fractures at any site (unadjusted HR = 1.46–1.47); however, these were no longer significant after additional adjustment for FN aBMD.

In contrast, the low volume phenotype had only seven bone parameters associated with incident fracture, all of which remained significant predictors after adjustments, except for FN aBMD. Failure load at the radius was the strongest predictor of incident fracture at any site and MOF sites (unadjusted HR = 2.10–2.51), and this association remained unchanged after adjustment for base covariates and FN aBMD (adjusted HR = 2.22–2.54).

The healthy bone phenotype had limited predictors of fracture. Before adjustments, five bone parameters were associated with any incident fracture and four with MOF. Only Tb.BMD at the radius was a consistent predictor of fracture at any site after adjustments for base covariates and FN aBMD (adjusted HR = 1.39, 95% CI 1.04–1.86) and MOF after adjustment for base covariates (adjusted HR = 1.65, 95% CI 1.04–2.60), but this did not remain significant after further adjustment for FN aBMD. These limited associations between bone parameters and fracture are likely due to the lower rate of fractures in this phenotype that has generally healthy bone characteristics.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Discussion

Using a large international prospective HR-pQCT cohort, three unique phenotypes of bone microarchitecture were identified in older adults, described as low density, low volume, and healthy bone, based on their anatomically different characteristics in bone microarchitecture. These phenotypes were associated with stratified risk of osteoporotic fracture, and within each phenotype unique bone imaging biomarkers were associated with within-phenotype fracture risk. Within the low density phenotype, declines in BMD parameters were the most strongly associated with higher fracture risk, whereas in the low volume phenotype, declines in estimated failure load were the most associated with higher fracture risk. The healthy bone phenotype had an overall low fracture risk with no distinct imaging biomarkers of increased fracture risk. These findings suggest that increased fragility is manifested differently across the population and that assessment of patient-specific bone health could benefit from assessment of bone phenotype.

The value of this approach is emphasized by the fact that FN aBMD did not distinguish fracture risk in the low volume phenotype, even though this group consisted of the highest proportion of females and second highest rate of incident fracture. These individuals had bones with thicker and denser cortices but smaller cross-sectional areas and reductions in trabecular bone quality, resulting in overall lower failure load. Femoral neck aBMD is limited to a two-dimensional measurement of bone density that does not account for cross-sectional area, and the thicker cortical bone likely masks deficits in the trabecular compartment. This affirms the known limitations of DXA and highlights that a single metric, such as FN aBMD, is not suitable for assessing fracture risk in all individuals. However, a common critique of HR-pQCT is that only peripheral skeletal sites can be measured, so predictive ability may be biased toward wrist fractures. To check that fracture associations were not driven solely by wrist fractures, the HRs for osteoporotic fractures at any fracture site were recomputed with exclusion of wrist fractures. The findings from this reanalysis were consistent with the results when wrist fractures were included (Supplemental Fig. S8). In addition, height and weight are often included as covariates in Cox proportional regression models of fracture risk, but we chose not to include them in this analysis. Height was not adjusted for because it was used as an input feature for clustering, and adjustment for weight was tested and found to trended toward a marginal increase in HRs, but this was not statistically significant.

External validation of the clustering model using two time points from a subset of participants demonstrated that phenotyping an individual was consistent over a 5-year period. Although this may not provide definitive precision of the phenotyping classification method, the purpose was to evaluate the stability of phenotypes over a clinically relevant period of time. The duration of 5 years is a significant amount of time in the context of clinical management of osteoporosis, especially when considering the validation cohort spanned the peri- and post-menopause age range during which accelerated age-related bone loss occurs in females.[2,41,44,45] Additional longitudinal cohort studies are needed to establish whether longer periods of aging, pharmaceutical interventions, comorbidities, or changes in lifestyle would result in greater shifts in phenotypes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Recent ex vivo studies have demonstrated that different age-related trajectories of bone strength exist for wide versus narrow bones at the radii and femoral neck.[13,14,46] These findings strengthen the premise that fundamentally different bio-mechanical mechanisms lead to reduced strength, even within a single sex or ethnicity. Differences found in wide versus narrow bones reflect the differences found here between the low density and low volume phenotype, which have distinctly different bone sizes and densities, but both present with elevated fracture risk relative to the healthy phenotype. Interestingly, the low density and low volume phenotypes identified here also draw comparisons to historical definitions of osteoporosis as two distinct manifestations, termed type I and type II involutional osteoporosis.[47,48] Further, the findings here are supported by smaller exploratory cohort studies that used cluster analysis to identify fracture phenotypes in males and females.[16-18] However, these studies were limited in sample size (n = 321–359), did not include validation of the clustering models, and were limited to only retrospective fracture history. An additional unique benefit to the current approach is that the use of soft clustering captures gradated change in an individual's phenotype.

There are limitations of this study that should be acknowledged. The analysis was limited to only the first osteoporotic fracture event after the baseline study visit. Prior fragility fractures and multiple incident fracture events were not accounted for and may have resulted in an underestimation of the association between phenotypes and fracture risk in our analyses. Because of sample size limitations, it was not investigated whether individual fracture sites of interest, such as the hip, were predominantly associated with a specific phenotype. Instead, all osteoporotic fracture sites were pooled together, with subanalyses of MOF sites. It is possible the frequency of fracture at specific sites varies across phenotypes, but this could not be explored in depth with this cohort. Another limitation of this study is the limited diversity of the population, as the cohort consisted predominantly of older female White adults from North America and Europe. Although the low density phenotype effectively distinguished males at high risk of fracture, the lower proportion of males limited the ability to determine whether the low volume phenotype provided added insight for fracture prediction in males. The finding that few males were classified to have a low volume phenotype may help elucidate why females are predisposed to a higher risk of fracture in terms of bone structure. However, the population distributions observed here may not reflect patterns of bone phenotypes in other ethnicities or age groups. It is likely that the population distribution across the phenotypes presented here would vary for younger populations and other ethnicities due to known differences in bone characteristics.[44,49-51] With this consideration, the descriptive names for the three phenotypes should be interpreted cautiously. The term "healthy bone" phenotype was chosen because of the above-average characteristics in bone properties observed for this group. However, these phenotypes are based on information captured by HR-pQCT and do not account for all attributes that determine bone health, such as tissue-level organization of the extracellular matrix. Along the same vein, atypical bone characteristics caused by diseases, such as osteogenesis imperfecta[52] or other rare bone diseases,[53] are not necessarily accounted for in these phenotypes and would warrant further investigation. Finally, males and females were deliberately pooled for phenotypic clustering in lieu of defining sex-specific clustering models because there is notable overlap in bone characteristics between

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sexes[44,54] and pooling sexes avoids duplication as observed in previous studies that separated the population on the basis of sex.[16,18] The distribution of males and females across phenotypes demonstrated that phenotypes intrinsically account for sex differences in bone characteristics, and all fracture risk analysis was compared with sexes pooled and stratified to compare the relationship between phenotypes and fracture were across both sexes.

The identification of bone microarchitecture phenotypes is a novel and innovative perspective on how HR-pQCT image data can be leveraged to assess bone health in an older adult population. Bone phenotypes strengthen the premise that bone fragility can manifest in different forms, and the complex interactions between structural and density properties should be taken into consideration to establish a patient-specific approach to managing osteoporosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health (R01AR061445 and AR027065), the National Heart, Lung, and Blood Institute Framingham Heart Study (N01-HC-25195, HHSN268201500001I), and research grants from the Investigator Initiated Studies Program of Merck Sharp & Dohme as well as Amgen. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, Merck, or of Amgen. Additional support was provided by the Friends of Hebrew SeniorLife in Boston, MA, USA; the Biomedical Engineering Graduate Program at the University of Calgary; and the Canadian Institutes of Health Research (#364554). Finally, we thank Philippe Wagner for providing support in recovering additional imaging data for this study.

Data Availability Statement

The population datasets analyzed during the current study are not publicly available to protect confidentiality and privacy of participants, but may be made available from the corresponding author on reasonable request and consent of co-authors bearing individual responsibility for their respective cohorts.

References

1. NIH-Panel. NIH consensus development panel on osteoporosis prevention, diagnosis, and therapy: osteoporosis prevention, diagnosis, and therapy. JAMA. 2001;285(6):785–795. [PubMed: 11176917]
2. Seeman E. Pathogenesis of bone fragility in women and men. Lancet. 2002;359(9320):1841–1850. [PubMed: 12044392]
3. Johnell O, Kanis J. Epidemiology of osteoporotic fractures. Osteoporos Int. 2005;16(Suppl 2):S3–S7. [PubMed: 15365697]
4. Lips P. Epidemiology and predictors of fractures associated with osteoporosis. Am J Med. 1997;103(2A):3S–8S discussion 8S–11S.
5. Tarride JE, Hopkins RB, Leslie WD, et al. The burden of illness of osteoporosis in Canada. Osteoporos Int. 2012;23(11):2591–2600. [PubMed: 22398854]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. J Bone Miner Res. 2007;22(3):465–475. [PubMed: 17144789]
7. Hopkins RB, Burke N, Von Keyserlingk C, et al. The current economic burden of illness of osteoporosis in Canada. Osteoporos Int. 2016;27(10):3023–3032. [PubMed: 27166680]
8. Khosla S, Shane E. A crisis in the treatment of osteoporosis. J Bone Miner Res. 2016;31(8):1485–1487. [PubMed: 27335158]
9. Fuggle NR, Curtis EM, Ward KA, Harvey NC, Dennison EM, Cooper C. Fracture prediction, imaging and screening in osteoporosis. Nat Rev Endocrinol. 2019;15(9):535–547. [PubMed: 31189982]
10. Cheng XG, Lowet G, Boonen S, et al. Assessment of the strength of proximal femur in vitro: relationship to femoral bone mineral density and femoral geometry. Bone. 1997;20(3):213–218. [PubMed: 9071471]
11. Boutroy S, Bouxsein ML, Munoz F, Delmas PD. In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. J Clin Endocrinol Metab. 2005;90(12):6508–6515. [PubMed: 16189253]
12. Samelson EJ, Broe KE, Xu H, et al. Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the Bone Microarchitecture International Consortium (BoMIC): a prospective study. Lancet Diabetes Endocrinol. 2019;7(1):34–43. [PubMed: 30503163]
13. Bigelow EM, Patton DM, Ward FS, et al. External bone size is a key determinant of strength-decline trajectories of aging male radii. J Bone Miner Res. 2019;34(5):825–837. [PubMed: 30715752]
14. Jepsen KJ, Kozminski A, Bigelow EM, et al. Femoral neck external size but not aBMD predicts structural and mass changes for women transitioning through menopause. J Bone Miner Res. 2017;32(6):1218–1228. [PubMed: 28084657]
15. Bolger MW, Romanowicz GE, Bigelow EMR, et al. External bone size identifies different strength-decline trajectories for the male human femora. J Struct Biol. 2020;212(3):107650. [PubMed: 33096230]
16. Edwards MH, Robinson DE, Ward KA, et al. Cluster analysis of bone microarchitecture from high resolution peripheral quantitative computed tomography demonstrates two separate phenotypes associated with high fracture risk in men and women. Bone. 2016;88:131–137. [PubMed: 27130873]
17. Litwic AE, Westbury LD, Robinson DE, Ward KA, Cooper C, Dennison EM. Bone phenotype assessed by HRpQCT and associations with fracture risk in the GLOW study. Calcif Tissue Int. 2018;102(1):14–22. [PubMed: 28913616]
18. Westbury LD, Shere C, Edwards MH, Cooper C, Dennison EM, Ward KA. Cluster analysis of finite element analysis and bone microarchitectural parameters identifies phenotypes with high fracture risk. Calcif Tissue Int. 2019;105(3):252–262. [PubMed: 31187198]
19. Samelson EJ, Demissie S, Cupples LA, et al. Diabetes and deficits in cortical bone density, microarchitecture, and bone size: Framingham HR-pQCT study. J Bone Miner Res. 2018;33(1):54–62. [PubMed: 28929525]
20. Amin S, Khosla S. Sex- and age-related differences in bone microarchitecture in men relative to women assessed by high-resolution peripheral quantitative computed tomography. J Osteoporos. 2012;2012:129760. [PubMed: 22496983]
21. Chapurlat R, Pialat JB, Merle B, et al. The QUALYOR (QUalite Osseuse LYon Orleans) study: a new cohort for noninvasive evaluation of bone quality in postmenopausal osteoporosis. Rationale and study design. Arch Osteoporos. 2017;13(1):2. [PubMed: 29282548]
22. Chaitou A, Boutroy S, Vilayphiou N, et al. Association between bone turnover rate and bone microarchitecture in men: the STRAMBO study. J Bone Miner Res. 2010;25(11):2313–2323. [PubMed: 20499368]
23. Sornay-Rendu E, Boutroy S, Duboeuf F, Chapurlat RD. Bone microarchitecture assessed by HR-pQCT as predictor of fracture risk in postmenopausal women: the OFELY study. J Bone Miner Res. 2017;32(6):1243–1251. [PubMed: 28276092]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

24. Biver E, Durosier-Izart C, Chevalley T, van Rietbergen B, Rizzoli R, Ferrari S. Evaluation of radius microstructure and areal bone mineral density improves fracture prediction in postmenopausal women. J Bone Miner Res. 2018;33(2):328–337. [PubMed: 28960489]
25. Nishiyama KK, Macdonald HM, Hanley DA, Boyd SK. Women with previous fragility fractures can be classified based on bone microarchitecture and finite element analysis measured with HR-pQCT. Osteoporos Int. 2013;24(5):1733–1740. [PubMed: 23179565]
26. Mellstrom D, Johnell O, Ljunggren O, et al. Free testosterone is an independent predictor of BMD and prevalent fractures in elderly men: MrOS Sweden. J Bone Miner Res. 2006;21(4):529–535. [PubMed: 16598372]
27. Fan B, Lu Y, Genant H, Fuerst T, Shepherd J. Does standardized BMD still remove differences between Hologic and GE-Lunar state-of-the-art DXA systems? Osteoporos Int. 2010;21(7):1227–1236. [PubMed: 19859644]
28. Whittier DE, Boyd SK, Burghardt AJ, et al. Guidelines for the assessment of bone density and microarchitecture in vivo using high-resolution peripheral quantitative computed tomography. Osteoporos Int. 2020;31(9):1607–1627. [PubMed: 32458029]
29. Pauchard Y, Liphardt AM, Macdonald HM, Hanley DA, Boyd SK. Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography. Bone. 2012;50(6):1304–1310. [PubMed: 22445540]
30. Buie HR, Campbell GM, Klinck RJ, MacNeil JA, Boyd SK. Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. Bone. 2007;41(4):505–515. [PubMed: 17693147]
31. Pistoia W, van Rietbergen B, Lochmuller EM, Lill CA, Eckstein F, Rüegsegger P. Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. Bone. 2002;30(6):842–848. [PubMed: 12052451]
32. MacNeil JA, Boyd SK. Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method. Bone. 2008;42(6):1203–1213. [PubMed: 18358799]
33. Whittier DE, Manske SL, Kiel DP, Bouxsein M, Boyd SK. Harmonizing finite element modelling for non-invasive strength estimation by high-resolution peripheral quantitative computed tomography. J Biomech. 2018;80:63–71. [PubMed: 30201250]
34. Bezdek JC, Ehrlich R, Full W. FCM: the fuzzy c-means clustering algorithm. Comput Geosci. 1984;10(2–3):191–203.
35. Ferraro MB, Giordani P. A toolbox for fuzzy clustering using the R programming language. Fuzzy Set Syst. 2015;279:1–16.
36. Milligan GW. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika. 1980;45(3):325–342.
37. Jain AK, Murty MN, Flynn PJ. Data clustering: a review. ACM Comput Surv. 1999;31(3):264–323.
38. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65.
39. Hennig C. Cluster-wise assessment of cluster stability. Comput Stat Data Anal. 2007;52:258–271.
40. Campello RJGB, Hruschka ER. A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Set Syst. 2006;157(21):2858–2875.
41. Burt LA, Hanley DA, Boyd SK. Cross-sectional versus longitudinal change in a prospective HR-pQCT study. J Bone Miner Res. 2017;32(7):1505–1513. [PubMed: 28294405]
42. Nelson W. Hazard plotting for incomplete failure data. J Qual Technol. 1969;1(1):27–52.
43. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239–241.
44. Burt LA, Liang Z, Sajobi TT, Hanley DA, Boyd SK. Sex- and site-specific normative data curves for HR-pQCT. J Bone Miner Res. 2016;31(11):2041–2047. [PubMed: 27192388]
45. Seeman E, Delmas PD. Bone quality—the material and structural basis of bone strength and fragility. N Engl J Med. 2006;354(21):2250–2261. [PubMed: 16723616]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

46. Jepsen KJ, Bigelow EM, Schlecht SH. Women build long bones with less cortical mass relative to body size and bone size compared with men. Clin Orthop Relat Res. 2015;473(8):2530–2539. [PubMed: 25690167]

47. Riggs BL, Khosla S, Melton LJ 3rd. A unitary model for involutional osteoporosis: estrogen deficiency causes both type I and type II osteoporosis in postmenopausal women and contributes to bone loss in aging men. J Bone Miner Res. 1998;13(5):763–773. [PubMed: 9610739]

48. Riggs BL, Melton LJ 3rd. Evidence for two distinct syndromes of involutional osteoporosis. Am J Med. 1983;75(6):899–901. [PubMed: 6650542]

49. Zhu TY, Yip BH, Hung VW, et al. Normative standards for HRpQCT parameters in Chinese men and women. J Bone Miner Res. 2018;33(10):1889–1899. [PubMed: 29893993]

50. Putman MS, Yu EW, Lee H, et al. Differences in skeletal microarchitecture and strength in African-American and white women. J Bone Miner Res. 2013;28(10):2177–2185. [PubMed: 23572415]

51. Gabel L, Macdonald HM, Nettlefold LA, McKay HA. Sex-, ethnic-, and age-specific centile curves for pQCT- and HR-pQCT-derived measures of bone structure and strength in adolescents and young adults. J Bone Miner Res. 2018;33(6):987–1000. [PubMed: 29394462]

52. Kocijan R, Muschitz C, Haschka J, et al. Bone structure assessed by HRpQCT, TBS and DXL in adult patients with different types of osteogenesis imperfecta. Osteoporos Int. 2015;26(10):2431–2440. [PubMed: 25956285]

53. Sidhu K, Ali B, Burt LA, Boyd SK, Khan A. Spectrum of microarchitectural bone disease in inborn errors of metabolism: a cross-sectional, observational study. Orphanet J Rare Dis. 2020;15(1):251. [PubMed: 32938479]

54. Macdonald HM, Nishiyama KK, Kang J, Hanley DA, Boyd SK. Age-related patterns of trabecular and cortical bone loss differ between sexes and skeletal sites: a population-based HR-pQCT study. J Bone Miner Res. 2011;26(1):50–62. [PubMed: 20593413]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

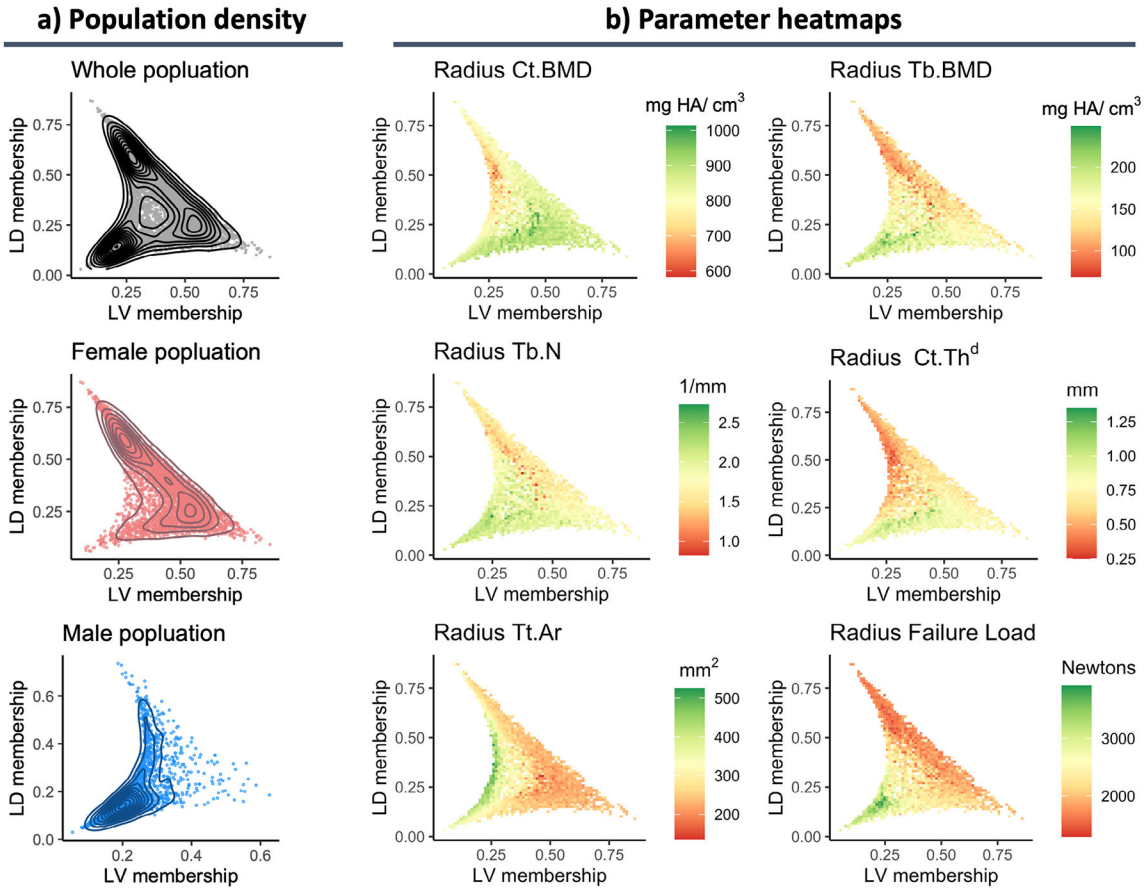

Fig. 1.

Characteristics of bone microarchitecture phenotypes. (*A*) Population distribution across phenotypes where each dot is an individual's membership coefficient with the low density (LD) versus low volume (LV) phenotype. In this plot, the origin would represent complete association to the healthy phenotype. The topology profiles represent the population density across the phenotypes. (*B*) Heatmaps applied to the same plots shown in *A* to visualize the relationship between several continuous parameters and phenotype membership coefficients. Heatmaps of all parameters are provided in Supplemental Fig. S2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

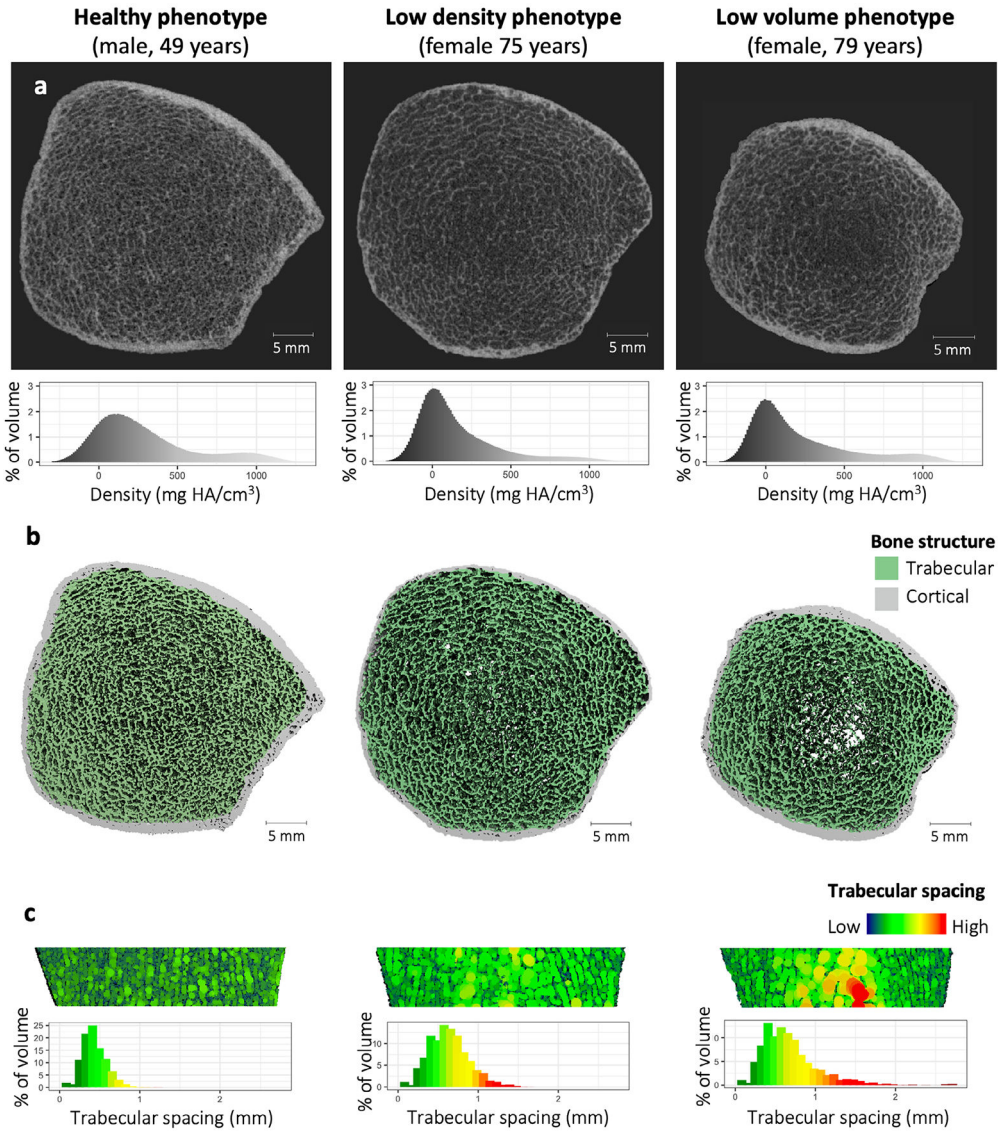

Fig. 2.

Examples of the three bone phenotypes. (*A*) Slice view of HR-pQCT images of the distal tibia from representative participants of each phenotype. The histograms show the distribution of the density values across the tibia scan region, reported in calibrated density units of mgHA/cm^3. (*B*) Three-dimensional reconstruction of bone microarchitecture of the same distal tibia scans in *A*, with trabecular and cortical bone color labeled. (*C*) Sagittal cut-plane of reconstructed images from *B*, with heatmaps representing spacing between adjacent trabecular struts. The histograms show the distribution of trabecular spacing across the trabecular compartment of the entire distal tibia scan region.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

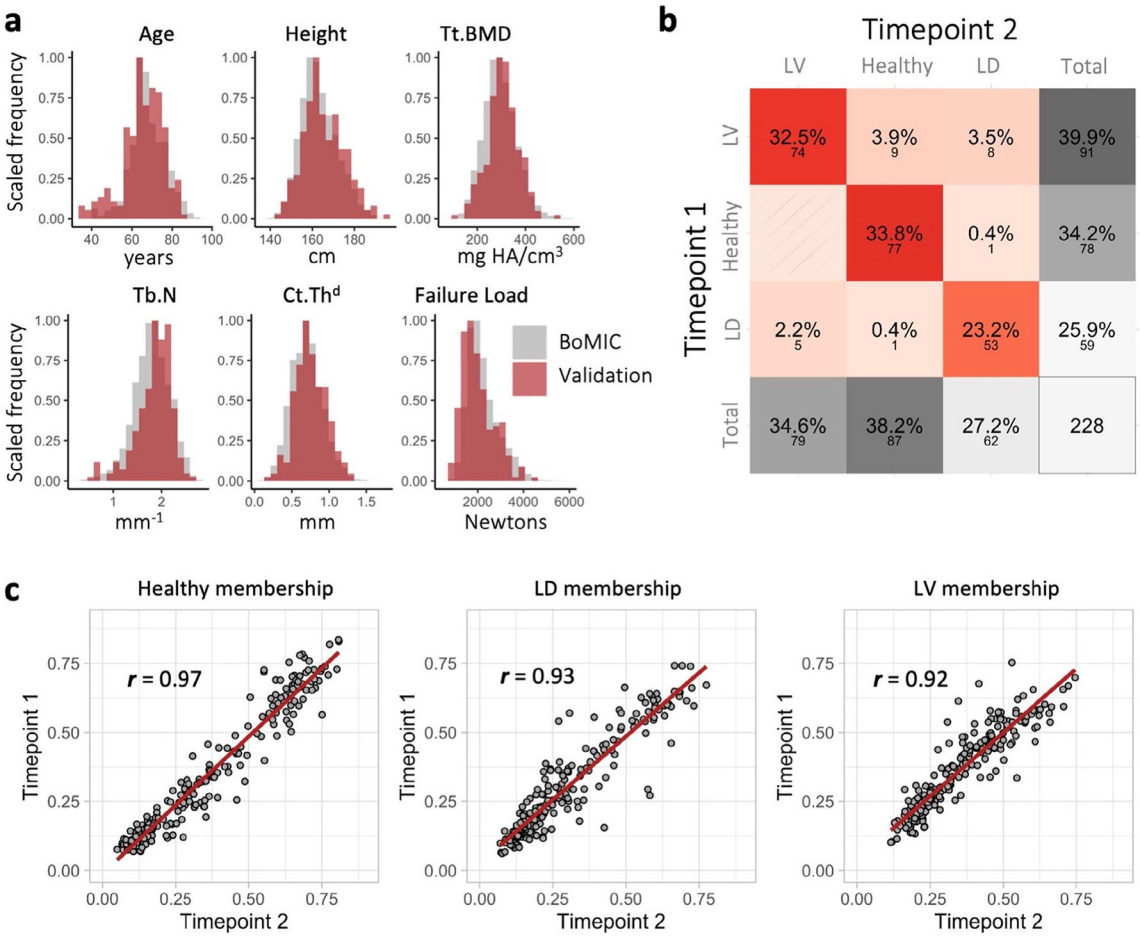

Fig. 3.

Summary of cluster external validation. (*A*) Histograms comparing the distributions of parameters for the whole BoMIC cohort (gray) to the validation cohort at the earliest time point (red). (*B*) Confusion matrix summarizing performance of hard cluster classification; large font represents the percentage and small font below the count of individuals in each classification outcome. (*C*) Correlation plots between the cluster memberships coefficients for the healthy, low density (LD), and low volume (LV) phenotypes predicted at the two time points. The regression lines are colored in red and Pearson correlation coefficients denoted by *r*.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

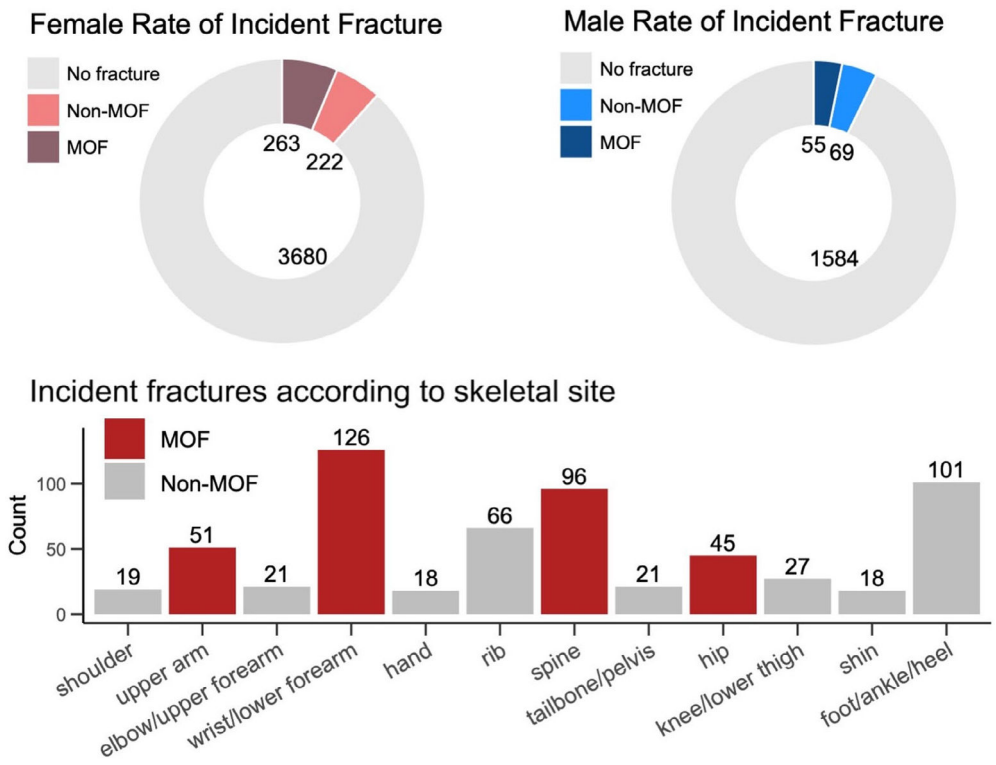

Fig. 4.
Summary of incident fracture events according to sex and fracture site.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

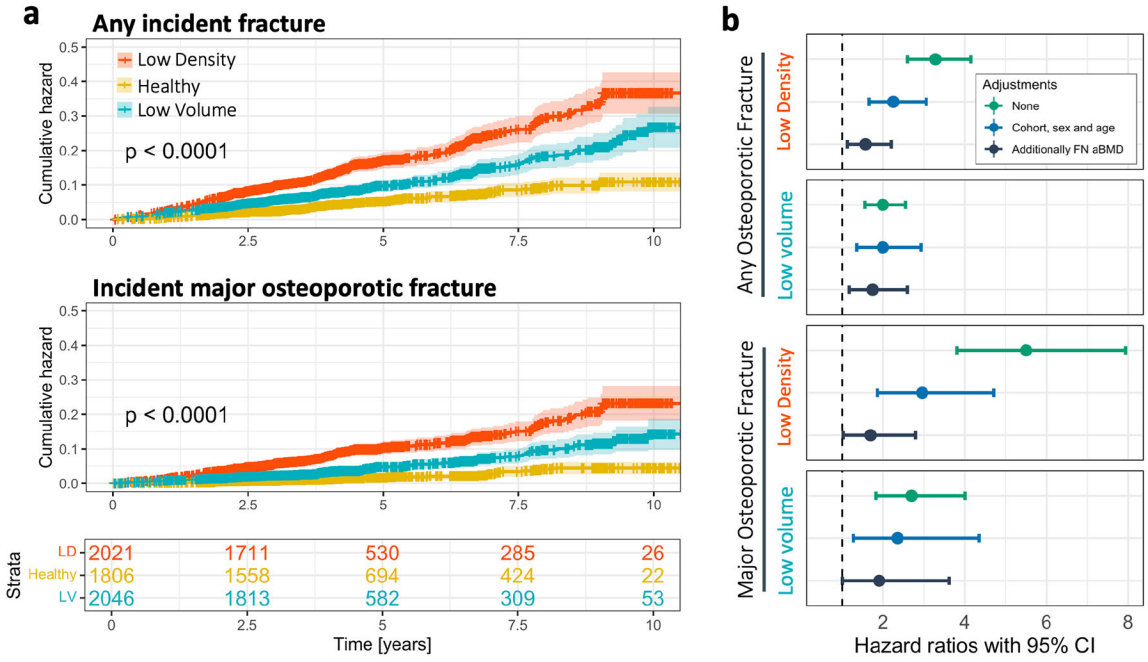

Fig. 5.

Association between incident fracture and phenotype. (*A*) Cumulative hazard functions determined using the Nelson–Aalen estimator, of any incident fracture and only major osteoporotic fracture. The shaded area in the plots represents the 95% confidence intervals (CI) of the cumulative hazard functions, and the censored observations are represented by the tick marks on the cumulative hazard curves. The *p* value represents the significance of the log-rank test performed comparing cumulative hazard functions. The bottom panel provides the proportion of the population remaining in each phenotype relative to follow-up time. (*B*) Hazard ratio (HR) and 95% CI for the association between phenotype and any incident fracture or only major osteoporotic fracture, with and without adjustments for covariates. The 95% CI are represented by the error bars. The HRs shown for the low density and low volume phenotypes are the risk relative to the healthy phenotype.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

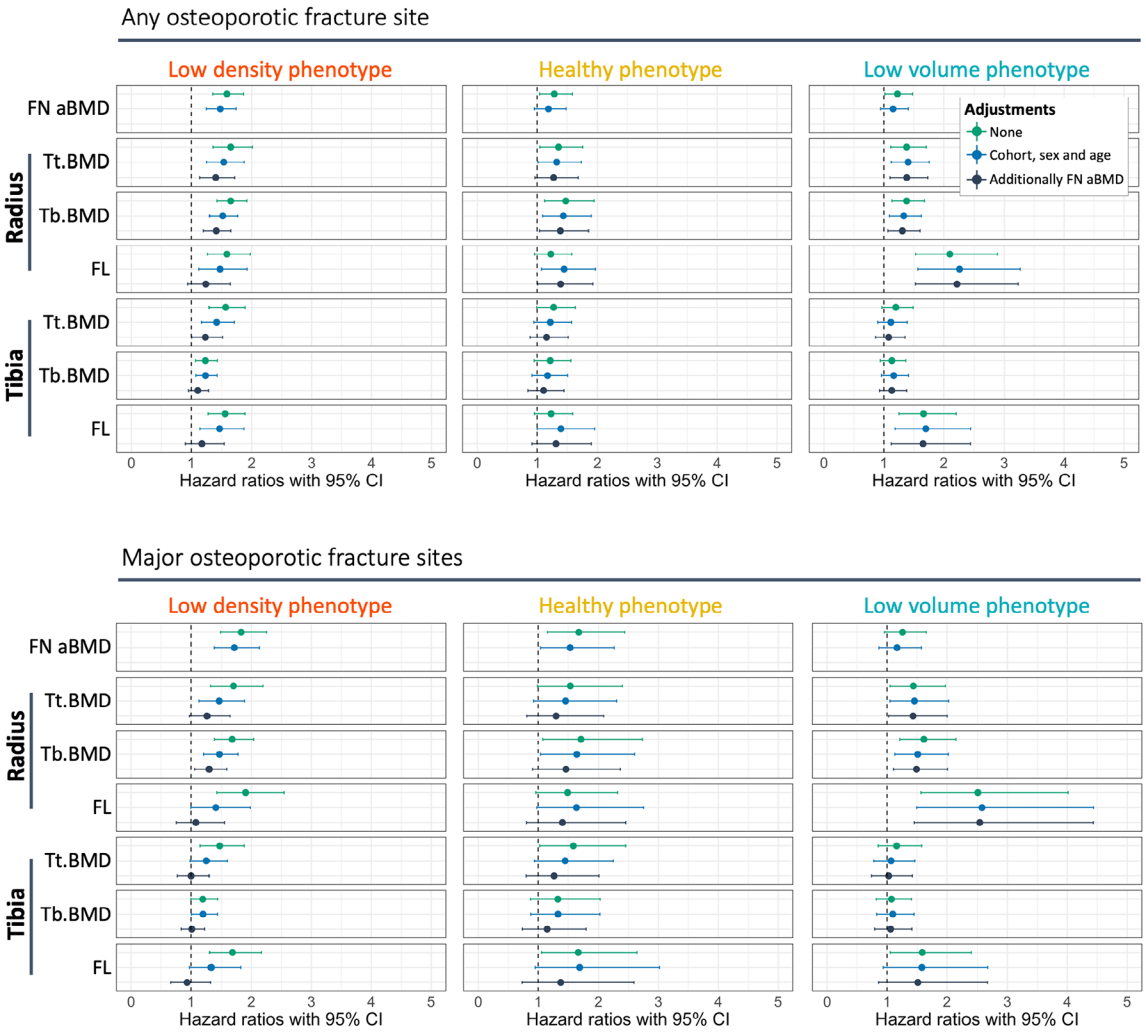

Fig. 6.

Within-phenotype associations between incident fracture and bone parameters. The hazard ratios (HR) and 95% confidence intervals (CI) are shown for the association between incident fracture of any site or only major osteoporotic fracture sites, relative to selected bone parameters. The HR are reported per standard deviation decrease of each bone parameter. A complete summary of HR and 95% CI for all parameters, with stratification by sex, are included in Supplemental Figs. S5 and S6 and Supplemental Table S3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Partial Spearman's ρ Correlation Coefficients for Association of Individual Parameters With Membership to Each Phenotype

	Spearman's ρ adjusted for membership with other clusters		
	Healthy phenotype	**Low density phenotype**	**Low volume phenotype**
Age	0.13	0.20	−0.41
Height	0.24	ns	−0.23
Weight	0.23	ns	−0.08
FN aBMD	0.13	−0.19	0.09
HR-pQCT—Radius			
Tt.BMD	−0.05	−0.58	0.28
Ct.BMD	−0.13	−0.43	0.25
Tb.BMD	0.15	−0.34	−0.04
Tb.N	0.20	−0.22	−0.05
Tb.1/N.SD[a]	0.15	−0.29	−0.04
Ct.Th[d]	−0.07	−0.51	0.15
Tt.Ar	0.34	0.19	−0.54
Ct.Ar	0.09	−0.47	−0.08
Tb.Ar	0.32	0.26	−0.54
Failure Load	0.31	−0.30	−0.18
HR-pQCT—Tibia			
Tt.BMD	0.04	−0.55	0.13
Ct.BMD	−0.06	−0.39	0.11
Tb.BMD	0.15	−0.26	−0.09
Tb.N	0.20	−0.14	−0.11
Tb.1/N.SD[a]	0.17	−0.18	−0.09
Ct.Th[d]	ns	−0.49	0.04
Tt.Ar	0.35	0.25	−0.51
Ct.Ar	0.11	−0.46	−0.14
Tb.Ar	0.32	0.32	−0.48
Failure load	0.29	−0.24	−0.26

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ns = not significant; FN = femoral neck; aBMD = areal bone mineral density; HR-pQCT = high-resolution peripheral quantitative computed tomography; Tt.BMD = total bone mineral density; Ct.BMD = cortical bone mineral density; Tb.BMD = trabecular bone mineral density; Tb.N = trabecular number; Tb.1/N.SD = trabecular inhomogeneity; Ct.Th[d] = cortical thickness; Tt.Ar = total cross-sectional area; Ct.Ar = cortical cross-sectional area; Tb.Ar = trabecular cross-sectional area.

All partial correlations shown are statistically significant ($p < 0.05$) after adjustment for multiple comparisons using the Holm–Bonferroni method.

[a]Correlation is inverted to such that a positive correlation reflects a beneficial association between phenotype and the parameter.

J Bone Miner Res. Author manuscript; available in PMC 2022 July 01.