



This is a repository copy of *A plasma protein-based risk score to predict hip fractures*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/216499/>

Version: Published Version

Article:

Austin, T.R., Nethander, M. orcid.org/0000-0003-3688-906X, Fink, H.A. orcid.org/0000-0002-5985-5447 et al. (26 more authors) (2024) A plasma protein-based risk score to predict hip fractures. *Nature Aging*, 4. pp. 1064-1075. ISSN 2662-8465

<https://doi.org/10.1038/s43587-024-00639-7>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A plasma protein-based risk score to predict hip fractures

Received: 17 March 2024

Accepted: 1 May 2024

Published online: 27 May 2024

 Check for updates

Thomas R. Austin^{1,36}, Maria Nethander^{2,3,36}, Howard A. Fink^{4,5,36}, Anna E. Törnqvist², Diana I. Jalal^{6,7}, Petra Buzkova⁸, Joshua I. Barzilay⁹, Laura Carbone^{10,11}, Maiken E. Gabrielsen¹², Louise Grahnemo², Tianyuan Lu^{13,14,15}, Kristian Hveem^{12,16,17}, Christian Jonasson¹², Jorge R. Kizer^{18,19}, Arnulf Langhammer^{16,17}, Kenneth J. Mukamal²⁰, Robert E. Gerszten²⁰, Bruce M. Psaty^{1,21}, John A. Robbins²², Yan V. Sun²³, Anne Heidi Skogholt¹², John A. Kanis^{24,25}, Helena Johansson^{2,25}, Bjørn Olav Åsvold^{12,26}, Rodrigo J. Valderrabano²⁷, Jie Zheng^{28,29,30}, J. Brent Richards^{13,14,31,32,33,34}, Eivind Coward¹² & Claes Ohlsson^{2,35} ✉

As there are effective treatments to reduce hip fractures, identification of patients at high risk of hip fracture is important to inform efficient intervention strategies. To obtain a new tool for hip fracture prediction, we developed a protein-based risk score in the Cardiovascular Health Study using an aptamer-based proteomic platform. The proteomic risk score predicted incident hip fractures and improved hip fracture discrimination in two Trøndelag Health Study validation cohorts using the same aptamer-based platform. When transferred to an antibody-based proteomic platform in a UK Biobank validation cohort, the proteomic risk score was strongly associated with hip fractures (hazard ratio per s.d. increase, 1.64; 95% confidence interval 1.53–1.77). The proteomic risk score, but not available polygenic risk scores for fractures or bone mineral density, improved the C-index beyond the fracture risk assessment tool (FRAX), which integrates information from clinical risk factors (C-index, FRAX 0.735 versus FRAX + proteomic risk score 0.776). The developed proteomic risk score constitutes a new tool for stratifying patients according to hip fracture risk; however, its improvement in hip fracture discrimination is modest and its clinical utility beyond FRAX with information on femoral neck bone mineral density remains to be determined.

Osteoporosis is a disease related to aging, leading to an increased risk of fractures. Hip fractures are associated with significant disability and the highest mortality risk of all fracture types, with up to 20% of the patients dying in the first year following hip fracture^{1,2}. By 2050, the worldwide annual number of hip fractures is expected to reach 4.5–6.3 million^{2–6}.

As there are effective treatments to reduce hip fractures, identification of patients at high risk of hip fractures is important to inform

efficient intervention strategies. Current available hip fracture risk calculators, such as the commonly clinically used FRAX, are algorithms that integrate the fracture risks associated with clinical risk factors with or without information on bone mineral density (BMD). According to the national guidelines in many countries, FRAX should be used to aid in fracture risk prediction to select the individuals most likely to benefit from osteoporosis treatment⁷; however, fracture risk tools such as FRAX may be improved by the addition of new biomarkers of hip

fracture risk. Although it is reported that some polygenic risk scores (PRSs) for fractures and different BMD-related parameters predict hip fracture risk, none of the evaluated PRSs has yet been shown to improve hip fracture discrimination^{8–11}. Furthermore, a limitation with PRSs is that these do not transfer well between different ancestral groups^{12,13}.

Circulating proteins may be an alternative source of hip fracture biomarkers. Protein profiles are dynamic and may integrate information on genetic variations and environmental factors. They also reflect ongoing biological processes and may, thereby, reflect current health status and disease risk¹². Different platforms for large-scale measurements of circulating proteins are continuously being developed to include more proteins in their analyses, including the SomaScan aptamer-based platform and the Olink double antibody proximity extension platform. Protein-based risk scores derived from either of these two platforms have recently been shown to improve prediction of some diseases of major public health importance^{12,14}; however, to our knowledge, there is no report of a successful transfer of a proteomic risk score for any health condition developed using one large-scale proteomic platform to an independent validation cohort using a different proteomic platform¹². Therefore, the present study aimed to develop a proteomic risk score to predict hip fractures, validate its performance and clinical utility in several validation cohorts and compare its performance for hip fracture prediction when using alternative proteomic platforms.

Results

Summary of the study design

A proteomic risk score was developed in the Cardiovascular Health Study (CHS; proteomics determined using the aptamer-based SomaScan 5K platform; 3,171 participants, 456 incident hip fractures, 39% men, mean age 74.4 years; Fig. 1 and Supplementary Table 1). The proteomic risk score was validated in two Trøndelag Health Study (HUNT) cohorts (proteomics determined using SomaScan 5K platform (3,259 participants, 187 incident hip fractures, 61% men, mean age 64.5 years) or 7K platform (1,988 participants, 155 incident hip fractures, 54% men, mean age 63.9 years); Fig. 1 and Supplementary Table 1). In addition, the proteomic risk score was also validated in a subset of the UK Biobank (proteomics determined using the Olink double antibody proximity extension platform; 50,876 participants, 686 incident hip fractures, 46% men, mean age 57.0 years; Fig. 1 and Supplementary Table 1). The hip fracture prediction of the developed proteomic risk score was also compared to the hip fracture prediction of earlier developed PRSs for fractures (PRS-Fracture¹⁵), femoral neck-BMD (PRS-FN-BMD¹⁶) and the bone mass-related parameter speed of sound in the heel (PRS-gSOS⁸).

Development of the proteomic risk score in the CHS cohort

We developed three different proteomic risk scores for incident hip fractures in the CHS derivation cohort: one weighted risk score and two risk scores using the machine-learning techniques, LASSO (least absolute shrinkage and selection operator) and Elastic net, respectively (Methods and Supplementary Tables 2–4). Based on C-index improvements, the weighted proteomic risk score performed better for hip fracture prediction than the two proteomic risk scores derived by machine-learning techniques, when evaluated in two independent HUNT validation cohorts (Supplementary Tables 2–5). The weighted proteomic risk score includes 18 proteins that passed the Bonferroni-adjusted P value threshold of $P < 1.0 \times 10^{-5}$ for the association of the aptamer with incident hip fractures in the CHS cohort. The weights of the included proteins are the estimated β values from a Cox regression.

We next used Mendelian randomization (MR) to investigate if the proteins included in the proteomic risk score were causally related to hip fractures. We identified valid genetic instruments for 15 of the 18 circulating proteins included in the proteomic risk score¹⁷. None of these 15 circulating proteins displayed statistically significant evidence of a

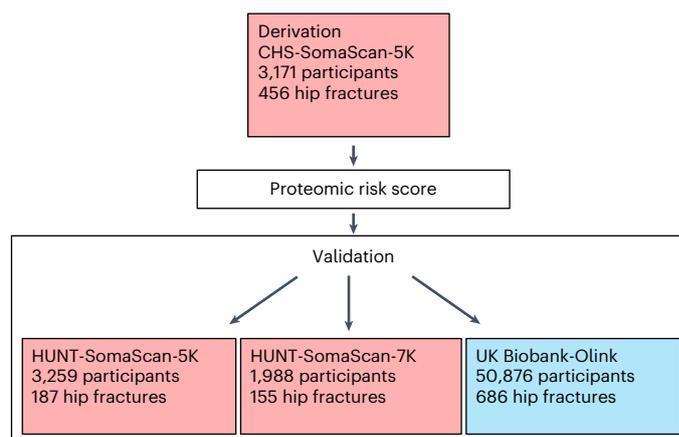


Fig. 1 | Overall design of the present study. A proteomic risk score was developed in the CHS (proteomics determined using the aptamer-based SomaScan 5K platform; 3,171 participants, 456 incident hip fractures, 39% men, mean age 74.4 years; Fig. 1 and Supplementary Table 1). The developed proteomic risk score was validated in two independent Trøndelag Health Study (HUNT) cohorts (proteomics determined using SomaScan 5K platform (3,259 participants, 187 incident hip fractures, 61% men, mean age 64.5 years) or 7K platform (1,988 participants, 155 incident hip fractures, 54% men, mean age 63.9 years)). In addition, the proteomic risk score was also validated in a subset of the UK Biobank (proteomics determined using the Olink double antibody proximity extension platform; all participants: 50,876 participants, 686 incident hip fractures, 46% men, mean age 57.0 years; randomly selected participants: 44,817 participants, 504 incident hip fractures, 46% men, mean age 56.7 years). Pink, cohorts (CHS and two HUNT cohorts) analyzed by aptamer-based SomaScan platform. Blue, cohort (UK Biobank) analyzed by the double antibody proximity extension Olink platform.

causal association with hip fractures (Supplementary Table 6), but the hip fracture genome-wide association studies (GWAS), used as source for the outcome analyses in the MR, was of limited size⁵.

Validation of the proteomic risk score in two HUNT cohorts

The effect sizes of this weighted proteomic risk score for hip fracture prediction were similar in the HUNT-SomaScan-5K validation cohort (Table 1), analyzed using the aptamer-based SomaScan 5K platform and in an independent HUNT validation cohort, analyzed using a later expanded version of the SomaScan platform (the SomaScan 7K platform; HUNT-SomaScan-7K cohort; Table 1, Fig. 2a and Supplementary Tables 1, 5 and 7). Next, we compared the hip fracture prediction for the proteomic risk score and available fracture/BMD-related PRSs in the two HUNT validation cohorts. Both for the two separate HUNT validation cohorts and when combined, the proteomic risk score (combined, hazard ratio (HR) 1.56; 95% confidence interval (CI) 1.36–1.79 per s.d. higher risk score) predicted hip fractures considerably more strongly than available PRSs for fractures (PRS-Fracture, HR 1.06; 95% CI 0.95–1.18), FN-BMD (PRS-FN-BMD, HR 1.20; 95% CI 1.07–1.34) and the bone mass-related parameter speed of sound in the heel (PRS-gSOS, HR 1.14; 95% CI 1.02–1.28; Table 1). In addition, the proteomic risk score, but not any of the PRSs, improved hip fracture discrimination as determined by C-index increase, starting from an age- and sex-adjusted base model, in the two HUNT validation cohorts (Supplementary Table 8).

Validation of the proteomic risk score in UK Biobank

Besides the large-scale aptamer-based proteomics platform used in the CHS derivation cohort and the two HUNT validation cohorts, the Olink double antibody proximity extension assay is an alternative large-scale proteomics platform. To avoid platform bias and enhance external validity and clinical utility, it is important to determine the transferability of proteomic risk scores between proteomic platforms.

Table 1 | Comparison of the incident hip fracture associations for the weighted proteomic risk score and polygenic risk scores in the two HUNT cohorts and in the UK Biobank

Predictor	Number of markers	Cox regression				
		HR	95% CI	P	n	n _{event}
HUNT-SomaScan-5K cohort						
Proteomic risk score	18	1.49	(1.24–1.80)	2.3 × 10 ⁻⁵	3,188	181
Proteomic risk score	13	1.44	(1.18–1.75)	2.4 × 10 ⁻⁴	3,188	181
PRS-Fracture	15	0.95	(0.82–1.10)	5.1 × 10 ⁻¹	3,188	181
PRS-gSOS	21,716	1.09	(0.94–1.26)	2.7 × 10 ⁻¹	3,188	181
PRS-FN-BMD	47	1.14	(0.98–1.32)	8.2 × 10 ⁻²	3,188	181
HUNT-SomaScan-7K cohort						
Proteomic risk score	18	1.66	(1.35–2.03)	9.6 × 10 ⁻⁷	1,939	153
Proteomic risk score	13	1.71	(1.38–2.12)	7.5 × 10 ⁻⁷	1,939	153
PRS-Fracture	15	1.20	(1.03–1.40)	2.2 × 10 ⁻²	1,939	153
PRS-gSOS	21,716	1.22	(1.04–1.44)	1.8 × 10 ⁻²	1,939	153
PRS-FN-BMD	47	1.28	(1.08–1.51)	3.4 × 10 ⁻³	1,939	153
HUNT-SomaScan combined cohorts						
Proteomic risk score	18	1.56	(1.36–1.79)	1.3 × 10 ⁻¹⁰	5,127	334
Proteomic risk score	13	1.56	(1.35–1.80)	1.5 × 10 ⁻⁹	5,127	334
PRS-Fracture	15	1.06	(0.95–1.18)	2.7 × 10 ⁻¹	5,127	334
PRS-gSOS	21,716	1.14	(1.02–1.28)	1.7 × 10 ⁻²	5,127	334
PRS-FN-BMD	47	1.20	(1.07–1.34)	1.1 × 10 ⁻³	5,127	334
UK Biobank-Olink, all participants						
Proteomic risk score	13	1.63	(1.52–1.76)	7.1 × 10 ⁻³⁹	50,450	679
PRS-Fracture	15	1.14	(1.05–1.23)	1.0 × 10 ⁻³	50,450	679
PRS-gSOS ^a	21,716	NA	NA	NA	NA	NA
PRS-FN-BMD	47	1.21	(1.12–1.31)	1.1 × 10 ⁻⁶	50,450	679
UK Biobank-Olink, randomly selected						
Proteomic risk score	13	1.64	(1.49–1.80)	1.6 × 10 ⁻⁴	44,428	501
PRS-Fracture	15	1.15	(1.05–1.26)	1.9 × 10 ⁻³	44,428	501
PRS-gSOS ^a	21,716	NA	NA	NA	NA	NA
PRS-FN-BMD	47	1.24	(1.13–1.35)	3.1 × 10 ⁻⁶	44,428	501

Base models were adjusted for age, sex and cohort specific factors. Associations with incident hip fractures were determined by Cox proportional regression models. A weighted proteomic risk score, including 18 proteins was developed in CHS. All 18 proteins were available in the two HUNT cohorts analyzed by the SomaScan platform, and 13 of these were available in the UK Biobank analyzed by the Olink platform. To test the replication between proteomic platforms, we also created a weighted proteomic score using the 13 proteins and evaluated its performance in the two HUNT cohorts and in the UK Biobank (using either all available participants or only the randomly selected participants). HRs and 95% CIs are given per s.d. higher risk score. Non-adjusted *P* values are derived using two-sided z-tests. For the analyses in Table 1, genetic analyses and proteomic analyses were required and the total number of participants and the number of incident hip fracture cases are given in Table 1. The results from the two HUNT-SomaScan cohorts were combined using fixed effects inverse-variance weighted meta-analysis. PRS-Fracture, weighted polygenic risk score based on independent GWAS significant signals for fractures at any bone site derived from Trajanoska et al.¹⁵. PRS-gSOS, polygenic risk score developed by the machine-learning technique LASSO for the bone mass-related parameter speed of sound in the heel determined by ultrasound in the UK Biobank, Lu et al.⁹. PRS-FN-BMD, weighted polygenic risk score based on independent GWAS significant signals for FN-BMD derived from Estrada et al.¹⁶. ^aThe PRS-gSOS was not feasible to use in the UK Biobank as this PRS was developed in the UK Biobank. NA, not available.

The Olink platform was used to analyze 1,463 proteins in a subset of the UK Biobank cohort. Of the 18 proteins used in the SomaLogic-derived proteomic risk score, 13 were available and also analyzed in the UK Biobank. The majority of the participants (44,817, 88%) with proteomics data in the UK Biobank used in our analyses were selected from a randomized subset within the UK Biobank (randomly selected participants), while the remaining participants (6,059, 12%) were selected by 13 biopharmaceutical companies¹⁸. In the two HUNT validation cohorts, a proteomic risk score restricted to these 13 proteins predicted hip fractures similarly as the proteomic risk score including all 18 proteins (combined, HR 1.56; 95% CI 1.35–1.80; Table 1). Notably, this proteomic risk score predicted hip fractures also in the UK Biobank, when all available participants (50,876) were included, with a HR of 1.63 (95% CI 1.52–1.76) per s.d. higher risk score (Table 1). A similar

effect estimate was observed when the analysis was restricted to the randomly selected participants in the UK Biobank (HR 1.64; 95% CI 1.49–1.80; Table 1). When meta-analyzing the results of the hip fracture prediction in the three validation cohorts (56,123 participants and 1,028 incident hip fracture cases), one s.d. higher proteomic risk score was associated with an HR of 1.63 (95% CI 1.52–1.74) for hip fractures (Fig. 2a and Supplementary Table 9).

The large dataset of the UK Biobank validation cohort, including a high number of incident hip fracture cases, enabled detailed analyses of hip fracture prediction. Sex- and age-stratified analyses in the UK Biobank revealed that the proteomic risk score was associated with incident hip fractures in both men and women and in both young and old participants (Supplementary Table 10). Analyses stratified on the median follow-up time of the hip fracture cases revealed that

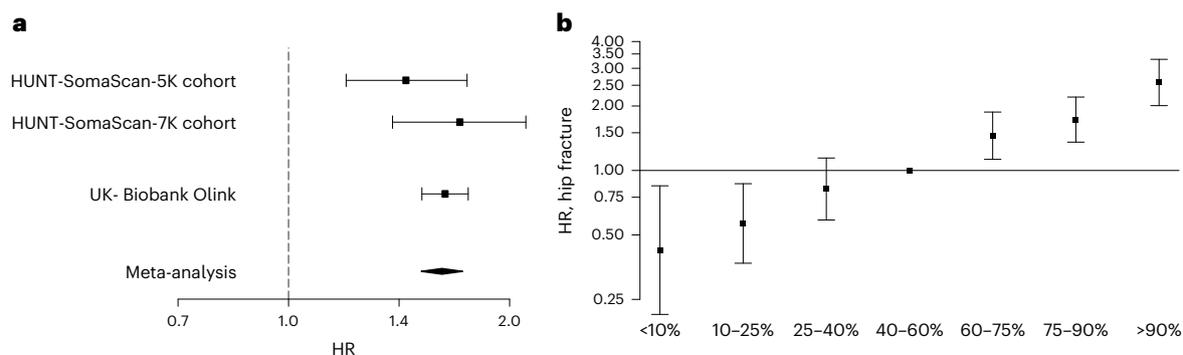


Fig. 2 | Association between the proteomic risk score and incident hip fractures. **a**, Association between the proteomic risk score and incident hip fractures in three separate validation cohorts. Base models are adjusted for age, sex and cohort-specific factors. Association with incident hip fractures is determined by Cox proportional regression models. Data are given as HRs and 95% CIs per s.d. higher risk score. The HUNT-SomaScan-5K cohort includes $n = 3,259$ participants and 187 incident hip fracture cases. The HUNT-SomaScan-7K cohort includes $n = 1,988$ participants and 155 incident hip fracture cases. The UK Biobank-Olink cohort includes $n = 50,876$ participants and 686 incident hip

fracture cases. The results from the proteomic risk score were combined using fixed effects inverse-variance weighted meta-analysis with a total of $n = 56,123$ participants and 1,028 incident hip fracture cases. **b**, Associations between seven total population percentile bins of the proteomic risk score and risk of incident hip fractures in the UK Biobank. Association with incident hip fractures is determined by Cox proportional regression models adjusted for age, sex, proteomic batch, ethnicity and UK Biobank center (50,876 participants and 686 incident hip fracture cases). Data are given as HRs and 95% CIs with the 40–60% bin as reference group.

the proteomic risk score predicted both early and late hip fractures (Supplementary Table 10). Furthermore, the associations between the proteomic risk score and incident hip fractures were essentially unaltered by adjustment for grip strength or self-reported falls at baseline and only modestly attenuated by adjustment for estimated BMD in the heel using ultrasound (Supplementary Table 10). Sensitivity analyses revealed that the strength of the association was not affected by adjustment for self-reported health rating at baseline or blood sample storage time (Supplementary Table 10). Furthermore, the association remained mainly unaffected by further adjustment for body mass index (BMI), smoking, diabetes and alcohol consumption (Supplementary Table 10). Most of the participants in the UK Biobank proteomic validation cohort were self-reported as white (94%; Supplementary Table 10) and the proteomic risk score predicted hip fractures in this large group. The group of self-reported non-white participants was heterogeneous and few had hip fractures (Supplementary Table 10). Exploratory analysis with limited statistical power suggested that the proteomic risk score also may predict incident hip fractures in non-white participants (Supplementary Table 10), but future larger studies are required to determine whether the developed proteomic risk score can be transferred to non-white groups.

In the UK Biobank, participants in the top 10% of proteomic risk score had an HR of 2.58 (95% CI 2.01–3.30) compared to participants with an average proteomic risk score (40–60%; Fig. 2b and Supplementary Table 11). Kaplan–Meier curves for non-hip fracture survival (Extended Data Fig. 1) and hip fracture survival probability curves (Extended Data Fig. 2) adjusted for age and sex illustrate the time dependency of the clear difference in hip fracture risk according to proteomic risk score quartiles. After 12 years, 288 participants in the highest proteomic risk score quartile but only 25 participants in the lowest proteomic risk score quartile had suffered an incident hip fracture (Extended Data Fig. 1).

Similarly, as observed in the two HUNT validation cohorts, the proteomic risk score predicted hip fracture risk more strongly than available PRSs in the UK Biobank validation cohort (Tables 1 and 2). The Pearson correlations for the proteomic risk score with PRS-Fracture ($r = 0.013$) and PRS-FN-BMD ($r = 0.036$) were low in the UK Biobank.

Finally, we determined whether the proteomic risk score added information for hip fracture prediction in the UK Biobank validation cohort beyond the clinically used fracture algorithm FRAX, which integrates information from clinical risk factors. The proteomic risk score improved fracture prediction in models adjusted for FRAX with or

without addition of estimated BMD (eBMD) (Supplementary Table 12). Discriminative analyses revealed that the proteomic risk score improved the C-index beyond FRAX, both with and without adding information on eBMD (Table 2). The proteomic risk score improved hip fracture discrimination beyond FRAX also when hip fracture discrimination was determined by the area under the curve (AUC) (Table 2 and Extended Data Fig. 3). Similar findings of improved hip fracture prediction beyond FRAX, as determined by C-index and AUC, were observed when the follow-up time was restricted to 10 years (Supplementary Table 13). In contrast, neither a PRS for fractures (PRS-Fracture) nor a PRS for FN-BMD (PRS-FN-BMD) improved hip fracture discrimination beyond FRAX (Table 2, Supplementary Table 13 and Extended Data Fig. 3). Furthermore, the proteomic risk score improved hip fracture reclassification indexes (integrated discrimination improvement, categorical net reclassification index (NRI) and categorical-free NRI) beyond a base model including FRAX or FRAX + eBMD (Supplementary Table 14). It should be emphasized that the combined NRI weights the events and non-events equally¹⁹. Therefore, one should rather consider the NRI results from events and non-events separately (shown in Supplementary Tables 14 and 15), showing that the improvements were mainly observed for events. Using an established clinical cutoff of 3% for hip fracture risk, above which pharmacological treatment is recommended by the Bone Health and Osteoporosis Foundation²⁰, the base FRAX model correctly identified 106 true incident hip fracture cases, while the addition of the developed proteomic risk score correctly identified 198 (+87%) true incident hip fracture cases in the UK Biobank cohort (Supplementary Table 15a). A similar magnitude of improvement in identification of true incident hip fractures was observed when the analyses were restricted to the randomly selected participants in the UK Biobank (+90%; Supplementary Table 15b). These improvements were due to improved sensitivity (Supplementary Table 15c,d).

In contrast, minor improvements of the different reclassification indexes were observed for PRS-FN-BMD when starting from a base FRAX model but not when starting from a base FRAX + eBMD model (Supplementary Table 14). PRS-Fracture did not improve fracture reclassification starting from a base model including FRAX or FRAX + eBMD (Supplementary Table 14).

Association with fractures at different bone sites

Next, we evaluated the performance of the developed hip fracture proteomic risk score for prediction of fractures at other bone sites

Table 2 | Hip fracture discrimination for the weighted proteomic risk score and polygenic risk scores beyond CRF-FRAX and CRF-FRAX + eBMD in the UK Biobank

Base model ^a	Predictor	C-index			AUC		
		Model			Model		
		Base	Base + predictor	P	Base	Base + predictor	P
UK Biobank-Olink, all participants							
FRAX-CRF	Proteomic risk score	0.735	0.776	5.4×10^{-9}	0.732	0.765	1.4×10^{-7}
FRAX-CRF	PRS-Fracture	0.735	0.736	5.2×10^{-1}	0.732	0.733	6.0×10^{-1}
FRAX-CRF	PRS-FN-BMD	0.735	0.734	7.2×10^{-1}	0.732	0.731	8.7×10^{-1}
FRAX -CRF + eBMD	Proteomic risk score	0.759	0.793	8.0×10^{-11}	0.753	0.782	9.7×10^{-10}
FRAX -CRF + eBMD	PRS-Fracture	0.759	0.759	4.5×10^{-1}	0.753	0.753	5.1×10^{-1}
FRAX -CRF + eBMD	PRS-FN-BMD	0.759	0.760	5.0×10^{-1}	0.753	0.754	4.6×10^{-1}
UK Biobank-Olink, randomly selected							
FRAX-CRF	Proteomic risk score	0.740	0.769	$2.9E \times 10^{-4}$	0.736	0.758	1.2×10^{-3}
FRAX-CRF	PRS-Fracture	0.740	0.742	4.3×10^{-1}	0.736	0.738	4.5×10^{-1}
FRAX-CRF	PRS-FN-BMD	0.740	0.738	7.0×10^{-1}	0.736	0.734	6.8×10^{-1}
FRAX -CRF + eBMD	Proteomic risk score	0.758	0.786	6.0×10^{-6}	0.752	0.775	1.8×10^{-5}
FRAX -CRF + eBMD	PRS-Fracture	0.758	0.759	1.7×10^{-1}	0.752	0.753	1.9×10^{-1}
FRAX -CRF + eBMD	PRS-FN-BMD	0.758	0.760	3.8×10^{-1}	0.752	0.754	4.0×10^{-1}

Fracture discrimination (C-index from Cox regression models and AUC from logistic regression models) beyond FRAX-CRF and/or eBMD using ultrasound in the UK Biobank. Comparisons of different AUCs were based on DeLong's test⁶⁶. Comparisons of different C-index were based on Kang et al.⁶⁷. Non-adjusted P values are derived using two-sided z-tests. Genetic analyses, proteomic analyses and eBMD were required (in total 49,087 participants and 663 incident hip fractures for all participants and 43,286 participants and 487 incident hip fractures in the randomly selected participants). Results for max available follow-up time are given in this table, whereas results for follow-up time restricted to 10 years are shown in Supplementary Table 13. FRAX-CRF, FRAX score for estimation of incident hip fracture risk using all available clinical risk factors in the UK Biobank. PRS-Fracture, weighted polygenic risk score based on independent GWAS significant signals for fractures at any bone site derived from Trajanoska et al.¹⁵. PRS-FN-BMD, weighted polygenic risk score based on independent GWAS significant signals for FN-BMD derived from Estrada et al.¹⁹. ^aBase model also adjusted for sex, proteomic batch, ethnicity and UK Biobank center.

besides hip fractures in the UK Biobank. Although, the proteomic risk score was associated with incident fractures at all bone sites evaluated, the association of the proteomic risk score was substantially stronger with hip fractures (HR 1.64, 95% CI 1.50–1.81) than with non-hip fractures such as forearm fractures (HR 1.18, 95% CI 1.08–1.29) and lower leg fractures (HR 1.20, 95% CI 1.08–1.33; Supplementary Table 16).

Discussion

Patients at high risk of fractures at the hip, the clinically most important fracture site, are important to identify early to start effective interventions. We developed a proteomic risk score that improved hip fracture prediction and discrimination in three independent validation cohorts, analyzed by two substantially different proteomic platforms. The developed proteomic risk score predicted hip fractures similarly in both young and old participants, and in both men and women. Finally, when added to FRAX with or without information on eBMD, the proteomic risk score, but not available PRSs, improved hip fracture discrimination. These results show that the proteomic risk score is strongly predictive of incident hip fractures in multiple independent populations.

Several PRSs have been developed that primarily predict different measures of BMD and some of these also predict hip fracture risk^{8–11}; however, none of these PRSs has been shown to improve hip fracture discrimination as determined by C-index or AUC beyond FRAX estimates, limiting their clinical utility^{8–11}. Similarly, none of the available fracture or BMD-related PRSs improved hip fracture discrimination in any of the three validation cohorts in the present study. In contrast, the proteomic risk score improved hip fracture discrimination, and the associations with incident hip fractures were substantially stronger for the proteomic risk score than for the different available fracture and BMD-related PRSs.

Similar to a recent study evaluating risk of coronary artery disease, weak correlation was found between available PRS and proteomic risk score in the present study evaluating hip fracture risk²¹. The low

correlation between proteomic risk score and PRS in both studies suggests that genetics and proteomics may contribute independent information for prediction of outcomes. In the recent study on coronary artery disease, genetics and proteomics added complementary information to the clinical risk factors for prediction of coronary artery disease²¹. In contrast, the developed proteomic risk score, but not available PRSs, improved hip fracture discrimination beyond clinical risk factors in the present study. As proteins integrate the effects of genes with effects caused by the environment, age, comorbidities, behaviors and drugs, the circulatory proteomic profile can provide information about health status and disease risk^{22,23}, most likely explaining why the proteomic risk score improved hip fracture discrimination more efficiently than available PRSs. We also note that current PRSs are based on relatively small samples of fractures, most of which occur at an age when the heritability of fracture is reduced²⁴. This is supported by previous studies demonstrating that the available PRSs display rather modest performances for hip fracture discrimination^{8–11}. One of the proteins, CD14, included in our proteomic risk score (Supplementary Table 2) has previously been reported to be associated with incident hip fractures in the MrOS cohort²⁵. We recently performed a large-scale meta-analysis of the association between circulating proteins, measured by the aptamer-based technique, and hip fracture risk and identified 23 hip fracture signals²⁶. Fifteen of these signals correspond to proteins included in the proteomic risk score developed in the present study.

As proteins integrate the effects of genes with effects caused by the environment, age, comorbidities, behaviors and drugs, the circulatory proteomic profile can provide information about health status and disease risk, most likely explaining why the proteomic risk score improved hip fracture discrimination more efficiently than available PRSs. We believe that many of the circulating proteins included in the identified proteomic risk score may be markers of current health status and the biological age, which in turn has an impact on hip fracture risk.

This notion is supported by a recent report demonstrating that four of the circulating proteins in our proteomic risk score (GDF15, MMP12, EGFR and WFDC2) also are included in a proteomic aging clock score that predicts accelerated biological aging and several age-related outcomes after adjusting for chronological age²⁷. Based on these findings, we propose that these four proteins predict hip fracture risk because they are general markers of biological aging and hip fracture risk increases when biological age increases. Future studies are warranted to determine what proportion of the plasma proteins included in the proteomic risk score are causally related to hip fractures and their underlying mechanism. Nevertheless, the clinical utility of a protein-based risk score for hip fracture prediction does not depend on whether the included proteins are causally related to hip fractures.

It is not only important to validate the performance of a developed proteomic risk score in independent validation cohorts, but also that it can be transferred to alternative proteomic platforms. Currently, there are two main proteomic platforms used in biomedical research: the SomaScan aptamer-based platform and the Olink double antibody proximity extension platform. We developed the proteomic risk score using the aptamer-based platform in CHS and successfully validated its performance to predict hip fractures in two independent HUNT cohorts, using two different versions of the SomaScan aptamer-based platform. Notably, the proteomic risk score was also validated in the UK Biobank where the circulating proteome was analyzed using the substantially different Olink double antibody proximity extension platform. To our knowledge, the present study is the first to report of a successful transfer of a proteomic risk score for disease prediction from one large-scale proteomic platform to an independent validation cohort analyzed by an alternative proteomic platform¹².

As clinically used hip fracture risk tools such as FRAX may be improved by the addition of validated new biomarkers, we determined the clinical utility of the proteomic risk score beyond FRAX with or without information on eBMD in the large UK Biobank validation cohort. The proteomic risk score significantly improved the hip fracture discrimination and reclassification indexes beyond a base model, including both FRAX without and with addition of eBMD. Based on these findings, we propose that the proteomic risk score is a biomarker candidate to be included as a new risk marker in future updates of FRAX²⁸. Notably, the proteins included in the proteomic risk score can either be analyzed by the SomaScan aptamer-based platform or by the Olink double antibody proximity extension platform using the same proteomic risk score as described in Supplementary Table 2, yielding similar performance for hip fracture prediction. This enhances the accessibility of the proteomic risk score for clinical use.

Among the three developed proteomic risk scores in CHS, the weighted proteomic risk score version, predicted hip fractures best when evaluated in two HUNT validation cohorts. The inferior performance of the two proteomic risk scores developed using the machine-learning techniques LASSO and Elastic net may be due to overfitting of these proteomic risk scores in the CHS derivation cohort.

The average baseline age and average age of incident hip fracture for the participants in the CHS derivation cohort were higher than for the participants in the three validation cohorts. The participants in the UK Biobank validation cohort were younger than the participants of the two HUNT validation cohorts. For the validation cohorts, the average ages at incident hip fractures were slightly lower compared to the average age of hip fracture cases in the general population²⁹. As the effect estimates for the association between the developed proteomic risk score and incident hip fractures were similar for the different evaluated cohorts, the proteomic risk score seems robust to use for hip fracture prediction within a wide age range. This notion is further supported by the similar effect estimates observed for the younger and older participants in our age-stratified analyses in UK Biobank; however, further studies are required to validate the performance of the proteomic risk score in older people.

Although the proteomic risk score was associated with incident fractures at all bone sites evaluated, the association of the proteomic risk score was substantially stronger with hip fractures than with non-hip fractures such as forearm fractures and fractures at the lower leg. This bone-site specificity is probably due to the fact that the proteomic risk score was developed to predict hip fractures and that the risk factors for fractures at different bone sites partly differ.

The strengths of the present study include the high number of incident hip fractures ($n = 456$ in the derivation cohort and $n = 1,028$ combined in the three validation cohorts) and the use of three independent validation cohorts analyzed by two different proteomic platforms.

The present study has several limitations. The group of non-white individuals in the UK Biobank was heterogeneous, defined using self-reported information and had few hip fractures. Although our exploratory analyses suggested that the proteomic risk score may predict incident hip fractures in non-white participants, future larger studies are required to determine if the developed proteomic risk score can be transferred to non-white groups. Thus, we call for caution on generalizing the findings beyond the populations studied in the present study. In addition, we did not adjust our analyses for kidney function, which could serve as a confounder in our analyses; however, we used median-normalized data across all aptamers, which attenuates kidney function associations and thus partly resembles adjustment for eGFR^{30,31}. Another limitation is that information on FN-BMD was not available for the included participants at the time of the proteomic baseline sample in the UK Biobank; however, information on eBMD in the heel, a strong predictor of hip fractures^{32,33}, was available for most of the included participants at the time of the baseline samples used for proteomic analyses. Therefore, eBMD was used as a hip fracture-related BMD measure in some of the models determining the clinical utility of the proteomic risk score beyond FRAX. Further studies in cohorts with information on circulating proteomics, FRAX estimates, and FN-BMD measures at baseline are warranted to determine the clinical utility of the developed proteomic risk score beyond FRAX with information on FN-BMD. Finally, it should be emphasized that the current costs for running the complete SomaScan or Olink proteomic assays, used in the present study, are substantial. Presently, it is not established that the magnitude of excess hip fracture risk predicted by the proteomic risk score is great enough to distinguish between individual patients for treatment decisions that will be cost-effective; however, future targeted hip fracture proteomics panels, based on the present findings, may be cheaper and add cost-effective information to future improved updates of FRAX²⁸.

In conclusion, the developed proteomic risk score enhanced hip fracture prediction and discrimination in three separate validation cohorts analyzed by two substantially different proteomic platforms. When added to FRAX with or without information on eBMD, the proteomic risk score, but not available PRSs, improved fracture discrimination. We propose that the developed proteomic risk score is a biomarker candidate to be included as a new risk marker in future updates of the fracture prediction tool FRAX²⁸. The developed proteomic risk score constitutes a new tool for stratifying patients according to hip fracture risk; however, its improvement in hip fracture discrimination is modest and its clinical utility beyond FRAX with information on FN-BMD remains to be determined.

Methods

Cohorts

The CHS was used to derive proteomic risk scores for hip fractures whereas two subcohorts within the HUNT study and a subsample of the UK Biobank were used for validation of the identified proteomic risk scores (Fig. 1).

The Cardiovascular Health Study

The CHS is a population-based longitudinal study of cardiovascular disease in older people (>65 years of age) recruited from four US

communities^{26,34}. Fasting EDTA-plasma was collected in 1992–1993 and was stored at -70°C until used for proteomic profiling²⁶. Incident hip fractures after the 1992–1993 study visit through 2015 were identified from hospital discharge diagnosis codes. The 1992–1993 exam was attended by 5,265 participants²⁶. Of those participants, all 3,171 with unthawed plasma available in 2020, had such plasma used for proteomic profiling using a SomaScan aptamer-based platform (5K SomaScan v.4.0 assay)^{35,36}. The CHS study was approved by institutional review boards at each of the four field centers and the Coordinating Center. The CHS is currently under a single institutional review board at the University of Washington (current approval no. MODCRO0000825). All CHS participants provided written informed consent. No compensation was provided to participants.

The Trøndelag Health Study

HUNT is a longitudinal health study in the Norwegian county of Trøndelag and it includes data from four visits between 1984 and 2019 (refs. 26,37,38).

The HUNT-SomaScan-5K cohort is a subcohort consisting of 3,259 participants from a HUNT cardiovascular project, including 1,270 participants with incident cardiovascular events and 1,989 participants without incident cardiovascular events²⁶. Therefore, we adjusted all HUNT-SomaScan-5K association analyses for incident cardiovascular disease (yes/no). The HUNT-SomaScan-7K cohort is a subcohort consisting of 1,988 new participants from two more recent studies, focusing on venous thromboembolism³⁹ and abdominal aortic aneurysm⁴⁰. We adjusted HUNT-SomaScan-7K association analyses for incident venous thromboembolism (802 cases) and incident abdominal aortic aneurysm (232 cases). The HUNT study has ethical approval from the Regional Committee for Medical and Health Research Ethics (REK Central Norway 2015/615) and informed consent was obtained from all participants²⁶. No compensation was provided to participants.

Previously unthawed, non-fasting, EDTA-plasma samples from the HUNT3 visit (2006–2008; stored at -80°C) were used for proteomic profiling. For the HUNT-SomaScan-5K cohort (using the same aptamer-based SomaScan platform as used in CHS) proteomic analyses were performed in 2017 (ref. 26), whereas the analyses for the HUNT-SomaScan-7K cohort (using the extended aptamer-based SomaScan 7K platform (v.4.1)) were performed in 2022.

The UK Biobank

The UK Biobank is a population-based cohort of approximately 500,000 participants aged 37–73 years. The participants were recruited between 2006 and 2010. Participant data include genome-wide genotyping, exome sequencing, whole-body magnetic resonance imaging, electronic health record linkage, blood and urine biomarkers and physical and anthropometric measurements¹⁸. For proteomic analyses, EDTA-plasma was collected and stored at -80°C until samples were analyzed. The plates were kept at -80°C and then sent to Olink (Uppsala, Sweden) for proteomic profiling. The UK Biobank Pharma Proteomics Project (UKB-PPP) is a collaboration between the UK Biobank and 13 biopharmaceutical companies¹⁸. Using baseline samples, 50,876 participants with successful proteomic analyses (starting in 2021) and available information on covariates (age, height, sex, weight and ethnicity) and not included in the COVID-19 case–control imaging study were included in the present proteomic hip fracture study. The majority (44,817, 88%) of the included participants were selected from a randomized subset within UK Biobank (randomly selected participants), while the remaining participants were selected by the 13 biopharmaceutical companies¹⁸. In sensitivity analyses, we restricted our analyses to the participants in the large randomized subsample, showing an essentially unchanged effect estimate for the association between the proteomic risk score and hip fractures in UK Biobank compared to the effect estimate when including all participants. Further details are available at <https://biobank.ndph.ox.ac.uk>. The UK Biobank has ethical

approval from the North West Multi-centre Research Ethics Committee (North West Research Ethics Committee, 11/NW/0382) and informed consent was obtained from all participants. No compensation was provided to participants. The present research was approved by the UK Biobank Research and Access Committee (application no. 51784).

FRAX estimates in the UK Biobank. FRAX estimates of the 10-year probability of experiencing a hip fracture for the participants in the UK Biobank were calculated by the international FRAX team (J. A. Kanis and H. Johansson) using the UK-specific FRAX tool (<https://www.sheffield.ac.uk/FRAX/>; v.1.4.4.) incorporating clinically relevant risk factors. Clinically relevant risk factors included in the FRAX algorithm were measured at the baseline visit for the UK Biobank such as age, sex, BMI (in kg m^{-2}), previous fractures (hip fractures and other osteoporotic fractures), current smoking, glucocorticoid use, rheumatoid arthritis and diagnosis of secondary causes of osteoporosis (type 1 diabetes, osteogenesis imperfecta in adults, untreated long-standing hyperthyroidism, hypogonadism or premature menopause, chronic malnutrition, chronic malabsorption and chronic liver disease). Participants with missing information on any FRAX clinical risk factor were considered free of the corresponding risk factors for the derivation of FRAX probability, as suggested by the FRAX model (<https://www.sheffield.ac.uk/FRAX/faq.aspx>). The relationships between risk factors and fracture risk in the FRAX model have been constructed using information derived from the primary data of population-based cohorts from around the world, including centers from North America, Europe, Asia and Australia, based on a series of meta-analyses^{41–47}. The FRAX algorithm has been externally validated with a similar geographic distribution with a follow-up in excess of 1 million patient-years and its construct summarized in a World Health Organization technical report^{48,49}. Since its launch in 2008, the FRAX model has proven to be well calibrated in diverse populations from Canada, Israel, Japan, Norway, Taiwan and the United Kingdom^{50–55}.

eBMD using ultrasound in the UK Biobank. Quantitative ultrasound of the heel was used to obtain a noninvasive eBMD that predicts fracture^{32,33}. A Sahara Clinical Bone Sonometer (Hologic Corporation) was used for quantitative assessment of calcanei in UK Biobank participants. Details of the complete protocol are publicly available on the UK Biobank website (www.ukbiobank.ac.uk/). eBMD (g cm^{-2}) was derived as a linear combination of speed of sound and bone ultrasound attenuation ($\text{eBMD} = 0.0025926 \times (\text{bone ultrasound attenuation} + \text{speed of sound}) - 3.687$)⁵⁶. Lower eBMD predicted high risk of incident hip fractures also in the present UK Biobank proteomics cohort (HR 1.76; 95% CI 1.60–1.93, per s.d. lower eBMD; the model was adjusted for age, sex, height, weight, ethnicity and assessment center).

Incident hip fractures

Hospitalizations were self-reported by CHS participants (every 6 months) and hospitalizations not reported by participants were identified from Medicare claims data²⁶. In the CHS, incident hip fractures were identified from hospital discharge International Classification of Diseases, Ninth Revision (ICD9) codes 820.xx, covering the time following the 1992–1993 CHS study visit through 2015. Pathologic fractures (ICD9 code 773.1x) and fractures caused by motor vehicle accidents (E810.xx–E825.xx) were excluded²⁶.

Hospital-based registries in the region were used to collect hip fracture data for the HUNT participants, covering the time interval from baseline (HUNT3 survey in 2006–2008) until March 2021. The following ICD10 codes S72.0, S72.1 and S72.2 or ICD9 code 820 were used for hip fracture definition²⁶.

In the UK Biobank, hip fractures were identified using ICD10 codes S72.0, S72.1 and S72.2 or ICD9 code 820 derived from registries covering the interval from baseline samples 2006–2010 until 31 October 2022.

Proteomics

SomaScan. The SomaScan 5K v.4.0 (CHS cohort and HUNT-SomaScan-5K cohort) and the SomaScan 7K v.4.1 (HUNT-SomaScan-7K cohort) aptamer-based assays were used to measure the concentrations of plasma proteins. The concentrations are given as relative fluorescent units. In brief, the aptamers from SomaScan are single-stranded DNA-based reagents called SOMAmers (slow off-rate modified aptamers). The negatively charged SOMAmers are designed to be complementary to the shape of the natively folded target proteins and bind the target protein tightly and specifically at a ratio of 1:1 (refs. 35,36). The method, which takes advantage of new chemically modified nucleotides, converts the measurement of protein levels into the measurement of nucleic acid levels assessed by a DNA oligo-array plate reader^{35,57}. The assay sensitivity has a median lower limit of detection in the femtomolar range, which is comparable to that of typical immunoassays^{14,58}. Results of these assays, reported in relative fluorescent units, are approximately proportional to plasma protein concentrations. Median intra- and inter-assay coefficients of variation for SomaScan v.4.0 and 4.1 are low at approximately 5% (refs. 14,31,58). The assays used for these analyses include 5,284 aptamers for the 5K platform and 7,596 aptamers for the extended 7K platform. We excluded aptamers marked as ‘deprecated’ (indicating a retired aptamer) and those marked as ‘non-human’ from the present analyses. Samples flagged by SomaLogic for poor quality assay were also removed. After these exclusions, 4,979 aptamers were available for analyses in the 5K platform and the same aptamers were also evaluated in the 7K platform.

Olink. The samples from UK Biobank were analyzed using the Olink Explore 1536 platform, measuring 1,472 protein analytes corresponding to 1,463 unique proteins. The Olink platform is considered a specific antibody-based assay. In brief, Olink uses a proximity extension assay technology where complimentary DNA-tags on matched pairs of antibodies hybridize when the antibodies are bound to the same target protein (antigen), enabling DNA amplification of the protein signal with a DNA polymerase. The PCR product is quantified and detected on a next generation sequencing readout^{59,60}. The Olink Explore platform consists of four panels focusing on inflammation, oncology, cardiometabolic and neurology proteins. Each panel has 12 internal controls and three proteins (CXCL8, IL-6 and TNF) are included in all four panels for quality assurance purposes. The performance of each protein assay is validated based on specificity, sensitivity, dynamic range, precision, scalability detectability and endogenous interference⁶¹. The median intra-individual coefficient of variation was 6.3%. The mean correlations across different panels for each of the three proteins (CXCL8, IL-6 and TNF) measured on all four protein panels varied between $r = 0.81$ and 0.96 .

Statistical analyses

Proteomic risk scores. *Weighted proteomic risk score.* From the Cox regression of the associations between each of the 4,979 aptamers and incident hip fractures, a weighted proteomic risk score was developed in the CHS derivation cohort. The proteomic risk score includes the 18 aptamers passing the Bonferroni-adjusted P value threshold of $P < 1.0 \times 10^{-5}$ in the CHS cohort (Supplementary Table 2). The weights are the estimated β values from the Cox regression.

LASSO proteomic risk score. We used LASSO with repeated data splitting to identify top aptamers (proteins) for inclusion in a risk score, using as few proteins as possible in the model. Protein prediction models were derived in CHS with the R package ‘glmnet’. For Cox regression analysis with LASSO penalty, a tenfold cross-validation was carried out for tuning parameter selection. The CHS data were randomly split 500 times, with 70% of the data in each split used as training data for model fitting and 30% of the data used for model testing. Age, sex and race

were forced into the model, while log-transformed and normalized protein values for all available aptamers were included as parameters. For each of the 500 splits, a set of proteins relevant for prediction were identified by the model based on the λ value, which gives the minimum mean cross-validated error in our model. Based on results from the 500 data splits, proteins were ranked by the frequency with which they were selected by the LASSO model.

The top 1–30 aptamers, ranked by their average selection frequency and coefficient estimates using LASSO, were carried into usual Cox regression models via repeated sample splitting. The data were split ten times, again into 70% training and 30% testing data, with C-index and AUC calculated for models containing age, sex, race and between 1–30 aptamers, by selection frequency in the previous step. A parsimonious model of top aptamers was chosen based on the number of aptamers included in the model that maximized AUC and C-index. The model that was found to be the most reliable protein model for prediction of fractures included 22 proteins and was validated in the HUNT cohorts (Supplementary Table 3).

Elastic net proteomic risk score. Using similar methods as for the LASSO machine-learning proteomic risk score, additional analyses were conducted using an elastic net (EN) penalty. This method is a hybrid of ridge regression and LASSO regularization which performs well in the setting of multicollinearity, in which parameters are highly correlated, as is seen in large proteomics datasets. In our EN analyses, an α value of 0.9 was chosen to allow for some multicollinearity while generating a parsimonious model, the top model, including 20 proteins (Supplementary Table 4).

Polygenic risk scores. We selected PRSs for fractures (PRS-Fracture¹⁵), FN-BMD (PRS-FN-BMD¹⁶) and the bone mass-related parameter speed of sound in the heel (PRS-gSOS⁸) based on the previously published largest GWAS^{8,15,16}, yielding the highest number of independent loci for each phenotype (see details below). The performances of PRS-FN-BMD and PRS-gSOS for fracture prediction have been published previously^{8–11}. There is also an alternative BMD-based PRS evaluating fracture prediction^{62–64}, which we have not used, as it is based on an early publication with a considerably smaller GWAS for FN-BMD⁶⁵.

PRS-Fracture The weighted PRS was based on 15 independent GWAS significant signals for fractures at any bone site derived from a previous fracture GWAS¹⁵.

PRS-FN-BMD The weighted PRS score was based on 47 independent GWAS significant signals for FN-BMD¹⁶.

PRS-gSOS The PRS was developed by the machine-learning technique LASSO using 21,717 genetic markers for ultrasound-derived speed of sound in the heel in the UK Biobank⁸.

Cox proportional hazards models. For hip fracture survival analyses, HRs, 95% CIs and significance levels were calculated using Cox proportional hazards models. Fractures were assessed from baseline to the diagnosis of fracture, death or the end of follow-up, whichever occurred first. The HR for a proteomic risk score or PRS was reported per one s.d. higher risk score. In addition, exploratory stratified analyses according to sex, age and ancestry were performed in the large UK Biobank cohort. Hip fracture risk in different total population percentile bins of the proteomic risk score, compared to participants with an average proteomic risk score (40–60%), was also determined.

Time-dependent analyses. Kaplan–Meier curves and hip fracture survival probability curves adjusted for age and sex were used to explore the time dependency of the difference in hip fracture risk according to proteomic risk score quartiles.

Fracture discrimination. A receiver operating characteristic (ROC) AUC was calculated using the roc.test function in the pROC R package.

The difference between AUCs were tested using DeLong's test, in the same R function (<https://cran.r-project.org/web/packages/pROC/index.html>)⁶⁶.

C-index and 95% CIs were calculated using the `rcorr.cens` function in the `Hmisc` R package (F. E. Harrel proposed the method and wrote the R package; <https://cran.r-project.org/web/packages/Hmisc/index.html>). Differences between C-index were tested using the `compareC` function in the R package `compareC` (<https://cran.r-project.org/web/packages/compareC/index.html>)⁶⁷.

Fracture reclassification. To evaluate the improvement in reclassification gained by adding a variable to a baseline predictor, the NRI using 3% predicted hip fracture threshold (NRI categorical; above which pharmacological treatment is recommended by the Bone Health and Osteoporosis Foundation²⁰), the integrated discrimination improvement and the category-free NRI were used^{68,69}.

All statistical computations were performed using R.

Association with fractures at different bone sites

We also evaluated the performance of the developed hip fracture proteomic risk score for prediction of fractures at other bone sites besides hip fractures in the UK Biobank. Fracture cases were identified using ICD10 and ICD9 codes (Supplementary Table 16) and included the following fracture groups: forearm fractures, hip fractures, major osteoporotic fractures (includes distal forearm fractures, hip fractures, vertebral fractures and upper arm fractures) and fractures of the lower leg.

Phase 2 proteins in the UK Biobank

During the revision of this manuscript, the phase 2 proteins from the Olink platform were released in a subsample of the participants included in the UK Biobank validation cohort. This enabled the inclusion of three additional proteins, resulting in a proteomic risk score with 16 proteins instead of 13 in this subsample of UK Biobank (Supplementary Table 2); however, the strengths of the association per s.d. increase in the proteomic risk score were very similar for the originally designed proteomic risk score of 13 and the new proteomic risk score of 16 in this subsample of the UK Biobank ($n = 39,551$, proteomic risk score of 13, HR 1.70, 95% CI 1.54–1.87; proteomic risk score of 16, HR 1.70, 95% CI 1.54–1.88). As the effect sizes for the hip fracture associations for proteomic risk score of 13 and proteomic risk score of 16 were similar and the sample size for proteomic risk score of 13 was substantially larger, the data in the article are presented using the originally designed proteomic risk score of 13.

Mendelian randomization

Genetic instruments for the plasma proteins included in the proteomic risk score were selected from a previously published GWAS on circulating proteins analyzed by the SomaScan aptamer-based technique in 35,559 Icelanders¹⁷. Correlated single-nucleotide polymorphisms were removed using LD-pruning with an r^2 threshold of 0.01. Outcome results were selected from a published GWAS on hip fractures (11,516 hip fracture cases)⁵. MR analyses were performed using the Mendelian-Randomization R package. Causal associations were estimated using either Wald ratio or inverse-variance weighted fixed effects depending on the number of valid genetic instruments.

Statistics and reproducibility

No statistical test was used to predetermine the sample sizes, but our sample sizes are similar to those reported in previous publications using proteomic data from the cohorts used in the present study^{18,26}. The numbers of included participants for each analysis are given in the table and figure legends. We used population-based longitudinal cohorts and no randomization was performed by us; however, in the UK Biobank the subcohort selected for the proteomic analyses was previously

randomly selected to be representative of the whole UK Biobank cohort as described in a previous publication¹⁸. To reduce potential selection bias, participants selected for the UK Biobank COVID-19 study were excluded from the present study evaluating proteomic risk scores¹⁸. The investigators were not blinded to allocation during experiments and outcome assessment. Data collection (incident hip fractures) and proteomic analyses were performed before the initiation of the present study. The statistical tests used are given in the legends of all tables and figures. The main finding that the weighted proteomic risk score predicted hip fracture risk was observed in three independent validation cohorts. Data distribution was assumed to be normal, but this was not formally tested. We followed the STARD guidelines.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Individual-level data from HUNT can be accessed by, or in collaboration with, a Norwegian principal investigator. Researchers can apply for HUNT data access from the HUNT Research Centre (<https://www.ntnu.edu/hunt>). To do this they must have obtained project approval from the Regional Committee for Medical and Health Research Ethics (REC)²⁶. Information on the application and conditions for data access in HUNT is available at <https://www.ntnu.edu/hunt/data>. Qualified investigators may access the CHS data by following the study policies described at https://chs-nhlbi.org/CHS_DistribPolicy. The authors are restricted from sharing CHS data as per the terms of their data-use agreement. Access to the UK Biobank data can be obtained by application to the UK Biobank (<https://www.ukbiobank.ac.uk/>). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability

Analyses have been performed using R v.4.1.1 (<https://cran.r-project.org/>) and the R packages `survival` (v.3.2-13) for Cox regressions; `glmnet` (v.4.1-8) for LASSO regression; `pROC` (v.1.18.5), `ROCR` (v.1.0-11), `PredictABEL` (v.1.2-4) for AUC and ROC analyses; `Hmisc` (v.5.1-2), `compareC` for C-index calculation and testing; and `survminer` (v.0.4.9), `rms` (v.6.8-0) and `metafor` (v.4.4-0) for plots. The FRAX tool is available at <https://www.sheffield.ac.uk/FRAX/> (v.1.4.4). Custom written R scripts are available at https://github.com/marianethander/Protein_Risk_Score and are included in the Supplementary Code file.

References

1. Rizkallah, M. et al. Comparison of morbidity and mortality of hip and vertebral fragility fractures: which one has the highest burden? *Osteoporos. Sarcopenia* **6**, 146–150 (2020).
2. Kanis, J. A. et al. SCOPE 2021: a new scorecard for osteoporosis in Europe. *Arch. Osteoporos.* **16**, 82 (2021).
3. Ferrari, S. et al. Unmet needs and current and future approaches for osteoporotic patients at high risk of hip fracture. *Arch. Osteoporos.* **11**, 37 (2016).
4. Johnell, O. & Kanis, J. A. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos. Int.* **17**, 1726–1733 (2006).
5. Nethander, M. et al. Assessment of the genetic and clinical determinants of hip fracture risk: genome-wide association and Mendelian randomization study. *Cell. Rep. Med.* **3**, 100776 (2022).
6. Kanis, J. A. et al. A systematic review of hip fracture incidence and probability of fracture worldwide. *Osteoporos. Int.* **23**, 2239–2256 (2012).
7. Kanis, J. A. et al. Adjusting conventional FRAX estimates of fracture probability according to the number of prior falls in the preceding year. *Osteoporos. Int.* **34**, 479–487 (2023).

8. Forgetta, V. et al. Development of a polygenic risk score to improve screening for fracture risk: a genetic risk prediction study. *PLoS Med.* **17**, e1003152 (2020).
9. Lu, T. et al. Improved prediction of fracture risk leveraging a genome-wide polygenic risk score. *Genome Med.* **13**, 16 (2021).
10. Nethander, M. et al. BMD-related genetic risk scores predict site-specific fractures as well as trabecular and cortical bone microstructure. *J. Clin. Endocrinol. Metab.* **105**, e1344–e1357 (2020).
11. Eriksson, J. et al. Limited clinical utility of a genetic risk score for the prediction of fracture risk in elderly subjects. *J. Bone Miner. Res.* **30**, 184–194 (2015).
12. Helgason, H. et al. Evaluation of large-scale proteomics for prediction of cardiovascular events. *JAMA* **330**, 725–735 (2023).
13. Li, Y. M. et al. Genetic risk factors identified in populations of European descent do not improve the prediction of osteoporotic fracture and bone mineral density in Chinese populations. *Sci. Rep.* **9**, 6086 (2019).
14. Lindbohm, J. V. et al. Plasma proteins, cognitive decline, and 20-year risk of dementia in the Whitehall II and Atherosclerosis Risk in Communities studies. *Alzheimers Dement.* **18**, 612–624 (2022).
15. Trajanoska, K. et al. Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *Brit. Med. J.* **362**, k3225 (2018).
16. Estrada, K. et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* **44**, 491–501 (2012).
17. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**, 1712–1721 (2021).
18. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
19. Pepe, M. S. Problems with risk reclassification methods for evaluating prediction models. *Am. J. Epidemiol.* **173**, 1327–1335 (2011).
20. LeBoff, M. S. et al. The clinician’s guide to prevention and treatment of osteoporosis. *Osteoporos. Int.* **33**, 2049–2102 (2022).
21. Moller, P. L. et al. Combining polygenic and proteomic risk scores with clinical risk factors to improve performance for diagnosing absence of coronary artery disease in patients with de novo chest pain. *Circ. Genom. Precis. Med.* **16**, 442–451 (2023).
22. Williams, S. A. et al. Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857 (2019).
23. Suhre, K., McCarthy, M. I. & Schwenk, J. M. Genetics meets proteomics: perspectives for large population-based studies. *Nat. Rev. Genet.* **22**, 19–37 (2021).
24. Michaëlsson, K., Melhus, H., Ferm, H., Ahlbom, A. & Pedersen, N. L. Genetic liability to fractures in the elderly. *Arch. Intern. Med.* **165**, 1825–1830 (2005).
25. Nielson, C. M. et al. Identification of hip BMD loss and fracture risk markers through population-based serum proteomics. *J. Bone Miner. Res.* **32**, 1559–1567 (2017).
26. Austin, T. R. et al. Large-scale circulating proteome association study (CPAS) meta-analysis identifies circulating proteins and pathways predicting incident hip fractures. *J. Bone Miner. Res.* **39**, 139–149 (2024).
27. Kuo, C.-L. et al. Proteomic aging clock (PAC) predicts age-related outcomes in middle-aged and older adults. Preprint at *MedRxiv* <https://doi.org/10.1101/2023.12.19.23300228> (2024).
28. Vandenput, L. et al. Update of the fracture risk prediction tool FRAX: a systematic review of potential cohorts and analysis plan. *Osteoporos. Int.* **33**, 2103–2136 (2022).
29. Lorentzon, M. & Cummings, S. R. Osteoporosis: the evolution of a diagnosis. *J. Intern. Med.* **277**, 650–661 (2015).
30. Lopez-Silva, C. et al. Comparison of aptamer-based and antibody-based assays for protein quantification in chronic kidney disease. *Clin. J. Am. Soc. Nephrol.* **17**, 350–360 (2022).
31. Candia, J., Daya, G. N., Tanaka, T., Ferrucci, L. & Walker, K. A. Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci. Rep.* **12**, 17147 (2022).
32. Bauer, D. C. et al. Quantitative ultrasound predicts hip and non-spine fracture in men: the MrOS study. *Osteoporos. Int.* **18**, 771–777 (2007).
33. Bauer, D. C. et al. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. Study of Osteoporotic Fractures Research Group. *Arch. Intern. Med.* **157**, 629–634 (1997).
34. Fried, L. P. et al. The Cardiovascular Health Study: design and rationale. *Ann. Epidemiol.* **1**, 263–276 (1991).
35. Gold, L. et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS ONE* **5**, e15004 (2010).
36. Austin, T. R. et al. Proteomics and Population Biology in the Cardiovascular Health Study (CHS): design of a study with mentored access and active data sharing. *Eur. J. Epidemiol.* **37**, 755–765 (2022).
37. Krokstad, S. et al. Cohort profile: the HUNT study, Norway. *Int. J. Epidemiol.* **42**, 968–977 (2013).
38. Åsvold, B. O. et al. Cohort profile update: the HUNT study, Norway. *Int. J. Epidemiol.* **52**, e80–e91 (2023).
39. Braekkan, S. K. et al. Venous thromboembolism and subsequent permanent work-related disability. *J. Thromb. Haemost.* **14**, 1978–1987 (2016).
40. Nyrønning, L. Å. et al. Symptoms of depression and risk of abdominal aortic aneurysm: a HUNT study. *J. Am. Heart Assoc.* **8**, e012535 (2019).
41. Kanis, J. A. et al. A family history of fracture and fracture risk: a meta-analysis. *Bone* **35**, 1029–1037 (2004).
42. De Laet, C. et al. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos. Int.* **16**, 1330–1338 (2005).
43. Kanis, J. A. et al. A meta-analysis of prior corticosteroid use and fracture risk. *J. Bone Miner. Res.* **19**, 893–899 (2004).
44. Kanis, J. A. et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone* **35**, 375–382 (2004).
45. Kanis, J. A. et al. Alcohol intake as a risk factor for fracture. *Osteoporos. Int.* **16**, 737–742 (2005).
46. Kanis, J. A. et al. Smoking and fracture risk: a meta-analysis. *Osteoporos. Int.* **16**, 155–162 (2005).
47. Johnell, O. et al. Predictive value of BMD for hip and other fractures. *J. Bone Miner. Res.* **20**, 1185–1194 (2005).
48. Kanis, J. A. et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos. Int.* **18**, 1033–1046 (2007).
49. Kanis, J.A. & WHO Collaborating Centre for Metabolic Bone Diseases. *Assessment of Osteoporosis at the Primary Health Care Level* (Univ. Sheffield Medical School, 2008).
50. Goldshtein, I., Gerber, Y., Ish-Shalom, S. & Leshno, M. Fracture risk assessment with FRAX using real-world data in a population-based cohort from Israel. *Am. J. Epidemiol.* **187**, 94–102 (2018).
51. Hippisley-Cox, J. & Coupland, C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *Brit. Med. J.* **339**, b4229 (2009).
52. Hoff, M. et al. Validation of FRAX and the impact of self-reported falls among elderly in a general population: the HUNT study, Norway. *Osteoporos. Int.* **28**, 2935–2944 (2017).
53. Leslie, W. D. et al. Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. *J. Bone Miner. Res.* **25**, 2350–2358 (2010).
54. Liu, I. T. et al. Validation of the Taiwan FRAX(R) calculator for the prediction of fracture risk. *Arch. Osteoporos.* **17**, 27 (2022).

55. Xu, G. et al. The accuracy of different FRAX tools in predicting fracture risk in Japan: a comparison study. *J. Orthop. Surg.* **28**, 2309499020917276 (2020).
56. Kemp, J. P. et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat. Genet.* **49**, 1468–1475 (2017).
57. Lollo, B., Steele, F. & Gold, L. Beyond antibodies: new affinity reagents to unlock the proteome. *Proteomics* **14**, 638–644 (2014).
58. Candia, J. et al. Assessment of variability in the SOMAscan assay. *Sci Rep.* **7**, 14248 (2017).
59. Assarsson, E. et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE* **9**, e95192 (2014).
60. Wik, L. et al. Proximity extension assay in combination with next-generation sequencing for high-throughput proteome-wide analysis. *Mol. Cell. Proteom.* **20**, 100168 (2021).
61. Zhong, W. et al. Next generation plasma proteome profiling to monitor health and disease. *Nat. Commun.* **12**, 2493 (2021).
62. Ho-Le, T. P. et al. Genetic prediction of lifetime risk of fracture. *J. Clin. Endocrinol. Metab.* **108**, e1403–e1412 (2023).
63. Ho-Le, T. P. et al. Assessing the clinical utility of genetic profiling in fracture risk prediction: a decision curve analysis. *Osteoporos. Int.* **32**, 271–280 (2021).
64. Ho-Le, T. P., Center, J. R., Eisman, J. A., Nguyen, H. T. & Nguyen, T. V. Prediction of bone mineral density and fragility fracture by genetic profiling. *J. Bone Miner. Res.* **32**, 285–293 (2017).
65. Stykarsdottir, U. et al. Multiple genetic loci for bone mineral density and fractures. *N. Engl. J. Med.* **358**, 2355–2365 (2008).
66. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
67. Kang, L., Chen, W., Petrick, N. A. & Gallas, B. D. Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach. *Stat. Med.* **34**, 685–703 (2015).
68. Pencina, M. J., D'Agostino, R. B. Sr. & Steyerberg, E. W. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med.* **30**, 11–21 (2011).
69. Pencina, M. J., D'Agostino, R. B. Sr., D'Agostino, R. B. Jr. & Vasan, R. S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat. Med.* **27**, 157–172 (2008).

Acknowledgements

We thank the population from the US communities for their contribution to the CHS. We thank the population of the County of Trøndelag and the staff at HUNT Research Center for their contribution to the HUNT study. We thank the population from the United Kingdom and researchers of the UK Biobank for their contribution to the UK Biobank study. We are grateful to the 13 participating companies (Alnylam Pharmaceuticals, Amgen, AstraZeneca, Biogen, Bristol-Myers Squibb, Calico, Genentech, Glaxo Smith Klein, Janssen Pharmaceuticals, Novo Nordisk, Pfizer, Regeneron and Takeda) for funding the UKB-PPP study. HUNT is a collaboration between the HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology NTNU), Trøndelag County Council, Central Norway Regional Health Authority and the Norwegian Institute of Public Health²⁶. The present study was supported by funding from the Swedish Research Council (2020-01392, C.O.); the Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (ALFGBG-720331 and ALFGBG-965235, C.O.); the Lundberg Foundation (LU2021-0096, C.O.); the Novo Nordisk Foundation (NNF 190C0055250, C.O.); and the Knut and Alice Wallenberg Foundation (KAW 2015.0317, C.O.). This research was also supported by contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222,

N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083 (J.A.R.), N01HC85086 and 75N92021D00006 (B.M.P.) and grants U01HL080295 (B.M.P.), U01HL130114 (B.M.P.) and R01HL144483 (B.M.P. and R.E.G.) from the National Heart, Lung and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke. Additional support was provided by R01AG023629 from the National Institute on Aging. A full list of principal CHS investigators and institutions can be found at [CHS-NHLBI.org](https://www.chs-nhlbi.org). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding sources had no role in the study design, data collection, data analysis, data interpretation, writing of the paper or in the decision to submit the paper for publication. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

The design of the study was performed in collaboration between T.R.A., H.A.F., B.M.P., E.C., J.B.R. and C.O. Proteomic analyses and assembly of the CHS cohort was conducted by T.R.A., H.A.F., D.I.J., P.B., J.I.B., L.C., J.R.K., K.J.M., R.E.G., B.M.P., J.A.R., Y.V.S., R.J.V. and J.Z. Proteomic analyses and assembly of the HUNT cohort were carried out by M.N., A.E.T., M.E.G., L.G., K.H., C.J., A.L., A.H.S., B.O.Å., E.C. and C.O. UK Biobank analyses were conducted by M.N., T.L., J.B.R. and C.O. The calculation of FRAX estimates was conducted by H.J. and J.A.K. T.R.A., M.N., H.A.F., A.E.T., E.C. and C.O. wrote the first draft of the manuscript. All authors contributed to subsequent drafts of the manuscript and made the decision to submit the manuscript for publication.

Funding

Open access funding provided by University of Gothenburg.

Competing interests

B.M.P. serves on the Yale Open Data Access Project funded by Johnson & Johnson, which had no impact on this paper. J.B.R. is founder and CEO of 5 Prime Sciences, which provides research services for biotech, pharma and venture capital companies for projects unrelated to this research. T.L. is an employee of 5 Prime Sciences. J.B.R. has served as an advisor to GSK and Deerfield Capital. The institution of J.B.R. has received investigator-initiated grant funding from Eli Lilly, GSK and Biogen for projects unrelated to this research. J.R.K. reports stock ownership in Abbott, AbbVie, Bristol-Myers Squibb, Johnson & Johnson, Medtronic, Merck and Pfizer. J.A.K. is a director of Osteoporosis Research, which maintains and develops FRAX. C.O. is an applicant on filed patent applications on the effect of probiotics on bone metabolism. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43587-024-00639-7>.

Correspondence and requests for materials should be addressed to Claes Ohlsson.

Peer review information *Nature Aging* thanks Robert Clarke for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

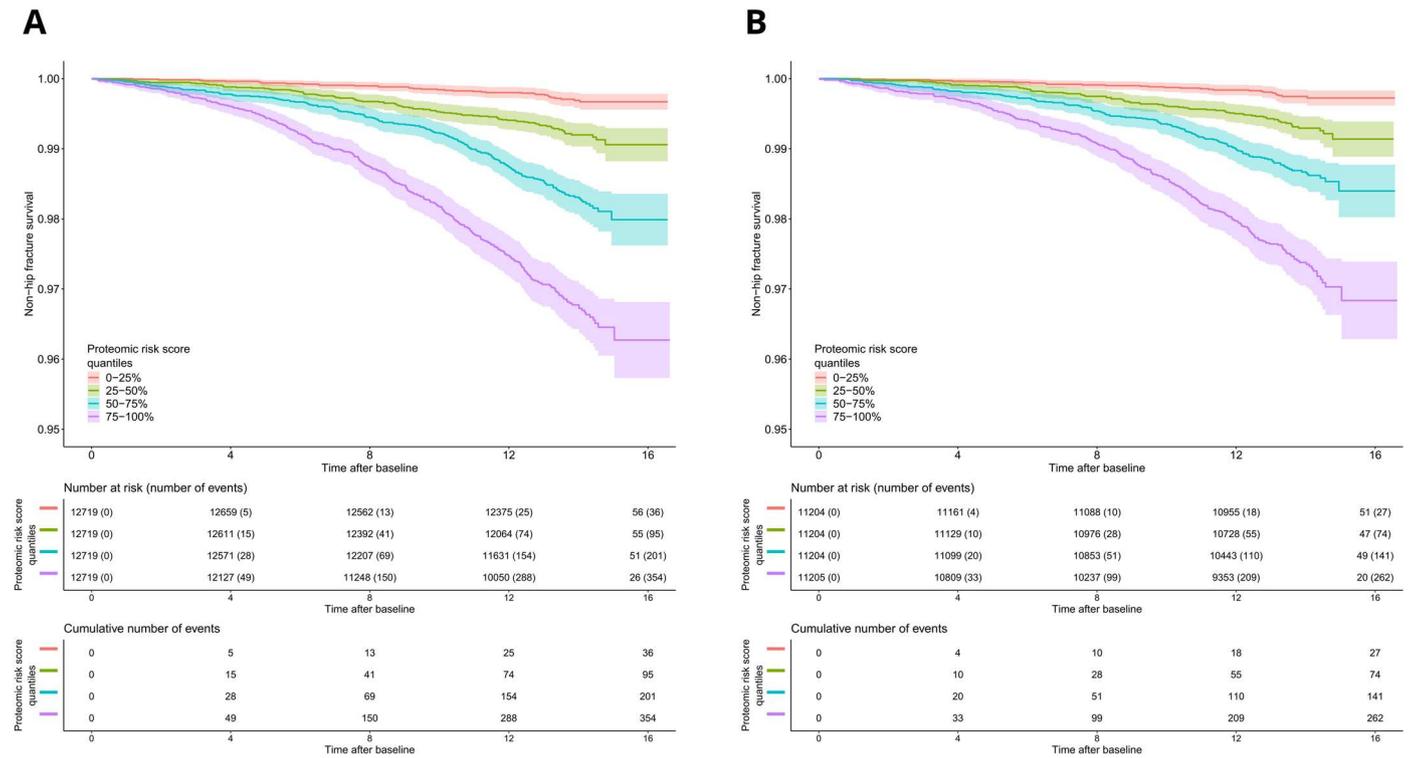
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless

indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

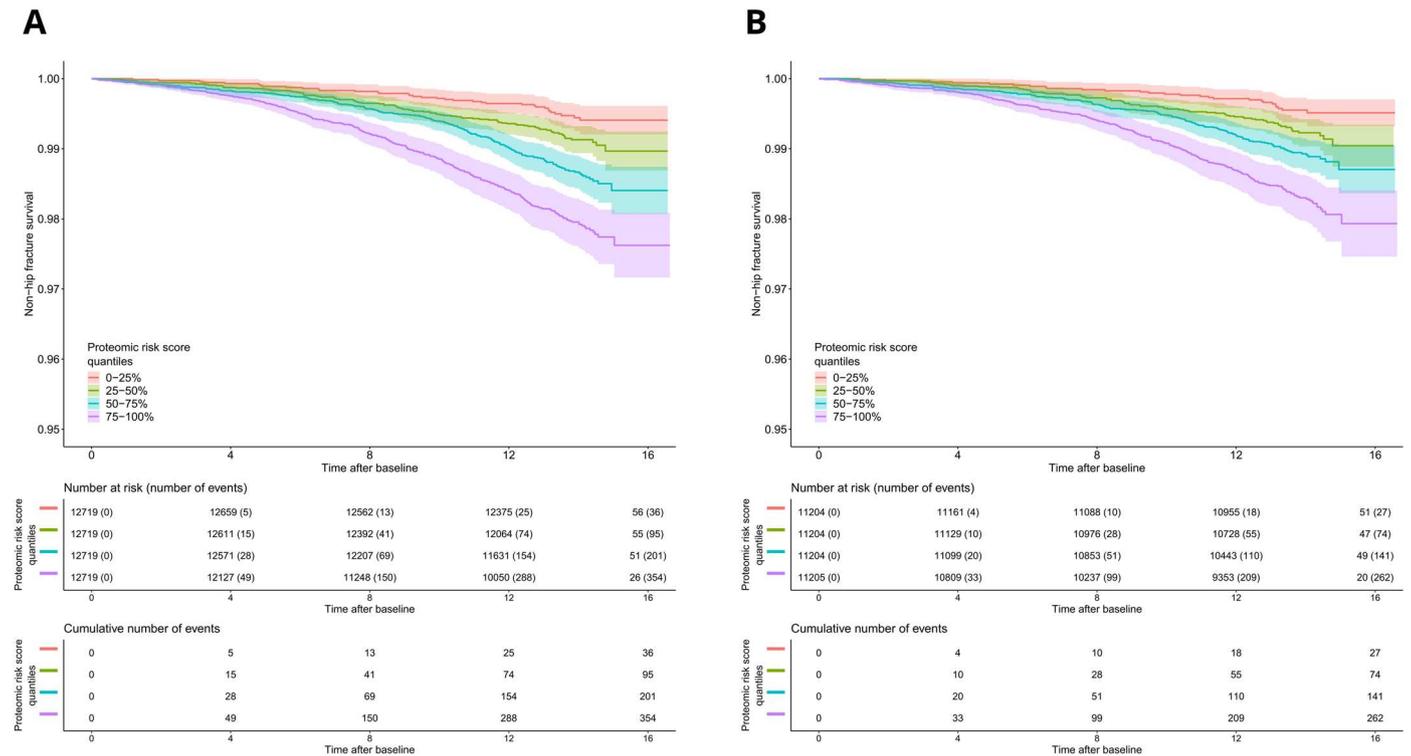
© The Author(s) 2024

¹Cardiovascular Health Research Unit, University of Washington, Seattle, WA, US. ²Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Osteoporosis Centre, Centre for Bone and Arthritis Research at the Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ³Bioinformatics and Data Center, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ⁴Geriatric Research Education and Clinical Center, VA Health Care System, Minneapolis, MN, US. ⁵Department of Medicine, University of Minnesota, Minneapolis, MN, US. ⁶Division of Nephrology, Department of Internal Medicine, Carver College of Medicine, Iowa City, IA, US. ⁷Iowa City VA Medical Center, Iowa City, IA, US. ⁸Department of Biostatistics, University of Washington, Seattle, WA, US. ⁹Division of Endocrinology, Kaiser Permanente of Georgia, Atlanta, GA, US. ¹⁰Charlie Norwood VAMC, Augusta, GA, US. ¹¹Division of Rheumatology, Department of Medicine, Medical College of Georgia, Augusta University, Augusta, GA, US. ¹²HUNT Center for Molecular and Clinical Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway. ¹³Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada. ¹⁴Quantitative Life Sciences Program, McGill University, Montreal, Quebec, Canada. ¹⁵Prime Sciences Inc, Montreal, Quebec, Canada. ¹⁶HUNT Research Centre, NTNU, Levanger, Norway. ¹⁷Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway. ¹⁸Cardiology Section, San Francisco VA Health Care System, San Francisco, CA, US. ¹⁹Department of Medicine, Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, US. ²⁰Department of Medicine, Beth Israel Deaconess Medical Center, Brookline, MA, US. ²¹Departments of Medicine, Epidemiology, and Health Systems and Population Health, University of Washington, Seattle, WA, US. ²²Department of Medicine, University of California, Davis, CA, US. ²³Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, US. ²⁴Centre for Metabolic Bone Diseases, University of Sheffield Medical School, Sheffield, UK. ²⁵Mary McKillop Institute for Health Research, Australian Catholic University, Melbourne, Victoria, Australia. ²⁶Department of Endocrinology, Clinic of Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. ²⁷Research Program in Men's Health, Aging and Metabolism, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, US. ²⁸Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. ²⁹Shanghai National Clinical Research Center for metabolic Diseases, Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commission of the PR China, Shanghai Key Laboratory for Endocrine Tumor, Shanghai Digital Medicine Innovation Center, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. ³⁰MRC Integrative Epidemiology Unit (IEU), Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK. ³¹Department of Human Genetics, McGill University, Montreal, Quebec, Canada. ³²Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada. ³³Department of Medicine, McGill University, Montreal, Quebec, Canada. ³⁴Department of Twin Research, King's College London, London, UK. ³⁵Region Västra Götaland, Sahlgrenska University Hospital, Department of Drug Treatment, Gothenburg, Sweden. ³⁶These authors contributed equally: Thomas R. Austin, Maria Nethander, Howard A. Fink. ✉ e-mail: claes.ohlsson@medic.gu.se



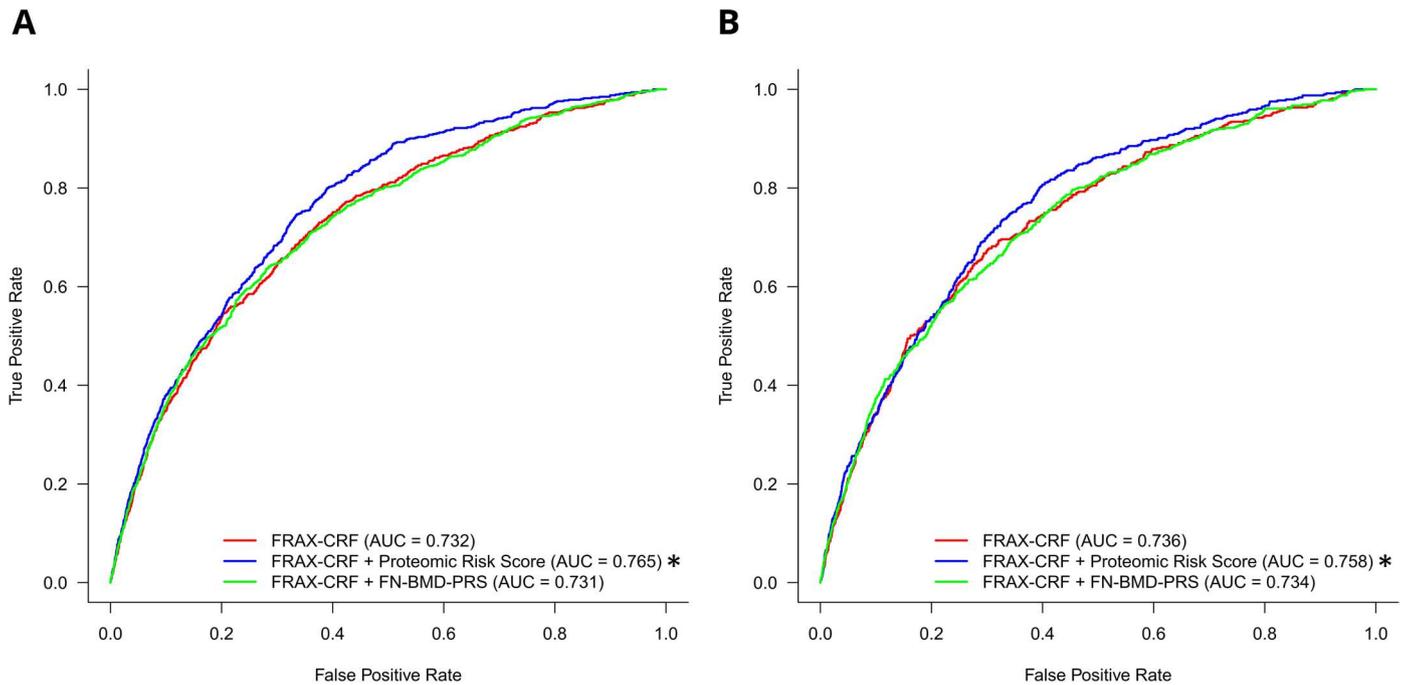
Extended Data Fig. 1 | Kaplan–Meier curves of hip fractures according to proteomic risk score quartiles in UK the Biobank. Data are given as point estimates of survival rate with 95% confidence intervals. **(A)** All participants with available proteomic analyses in the UK Biobank were included (50,876

participants and 686 incident hip fracture cases). **(B)** The randomly selected participants with available proteomic analyses in the UK Biobank were included (44,817 participants and 504 incident hip fracture cases).



Extended Data Fig. 2 | Hip fracture survival probability according to proteomic risk score quartiles adjusted for age, sex, and cohort specific factors in the UK Biobank. Data are given as point estimates of survival rate with 95% confidence intervals. **(A)** All participants with available proteomic analyses

in the UK Biobank were included (50,876 participants and 686 incident hip fracture cases). **(B)** The randomly selected participants with available proteomic analyses in the UK Biobank were included (44,817 participants and 504 incident hip fracture cases).



Extended Data Fig. 3 | Receiver operating characteristic curves and fracture discrimination (AUC from logistic regression models) using proteomic risk score or PRS-FN-BMD beyond FRAX-CRF in the UK Biobank. Genetic analyses, proteomic analyses and eBMD were required. **(A)** All available participants (in total 49,087 participants and 663 incident hip fractures). **(B)** Randomly selected participants (in total 43,286 participants and 487 incident hip fractures). The base model was also adjusted for sex, proteomic batch, ethnicity, and UK Biobank centre. PRS-FN-BMD = Weighted polygenic risk score based on independent GWAS significant signals for femoral neck bone mineral density (FN-BMD)

derived from Estrada et al.¹⁶. FRAX-CRF = FRAX score for estimation of incident hip fracture risk using all available clinical risk factors in the UK Biobank. The sensitivity (=true positive rate) and specificity (=1-false positive rate) are presented on the Y-axis and X-axis, respectively. *AUC for the model including FRAX-CRF + proteomic risk score (blue line) was significantly larger than the AUC for the model including only FRAX-CRF (red line). Two-sided DeLong's test gave **(A)** for all available participants $P = 1.4 \times 10^{-7}$ and **(B)** for the randomly selected participants $P = 1.2 \times 10^{-3}$ (ref. 18).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection The SomaScan 5K version 4.0 (CHS cohort and HUNT-SomaScan-5K cohort)
SomaScan 7K version 4.1 (HUNT-SomaScan-7K cohort)
Olink Explore 1536 platform.

Data analysis UKB RAP, R v4.1.1 (<https://cran.r-project.org/>)
The following R packages were used:
survival v3.2-13 for Cox regressions,
glmnet v4.1-8 for LASSO regression,
pROC v1.18.5, ROCR v1.0-11, PredictABEL v1.2-4 for AUC and ROC analyses,
Hmisc v5.1-2, compareC v1.3.2 for C-index calculation and testing,
metafor v4.4-0 for meta analysis
survminer v0.4.9, rms v6.8-0 and forestplot v3.1.3 for plots,
FRAX tool (<https://www.sheffield.ac.uk/FRAX/>; version 1.4.4.)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Individual level data from HUNT can be accessed by, or in collaboration with, a Norwegian principal investigator. Researchers can apply for HUNT data access from HUNT Research Centre (<https://www.ntnu.edu/hunt>). To do this they must have obtained project approval from the Regional Committee for Medical and Health Research Ethics (REC). Information on the application and conditions for data access is available at <https://www.ntnu.edu/hunt/data>.

Qualified investigators may access CHS data by following study policies as described here: https://chs-nhlbi.org/CHS_DistribPolicy. The authors are restricted from sharing CHS data per the terms of their data use agreement.

Access to UK biobank data can be obtained by application to UK biobank (<https://www.ukbiobank.ac.uk/>). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Since the risk of fracture differs between males and females, all analyses have been either adjusted for or stratified by self-reported sex. In CHS and UK Biobank, gender was determined by self-report at baseline examination. In HUNT, sex was used and determined by genetics. The data of the large well-powered UK Biobank are given gender stratified (Table S8). In addition, descriptives data are given gender (CHS and UK Biobank) or sex (HUNT) stratified in Table S1. All other association analyses are adjusted for gender (CHS and UK Biobank) or sex (HUNT).

Reporting on race, ethnicity, or other socially relevant groupings

All analyses including CHS or UKB participants have been adjusted for self-reported ethnicity.

Population characteristics

UK Biobank is a population-based cohort of approximately 500,000 participants aged 37-73 years. CHS is a population-based longitudinal study of coronary heart disease and stroke in adults aged 65 years and older. HUNT comprises data and samples obtained through four population surveys between 1984 and 2019 in 230,000 people from the Norwegian county of Trøndelag. See details in Supplementary Table 1 for the four included cohorts

Recruitment

UK Biobank participants were recruited between 2006 and 2010. CHS participants were recruited from four U.S. communities: Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Pittsburgh, Pennsylvania. At baseline (1989-1990), 5,201 individuals were enrolled. HUNT: All people living in the county of Nord-Trøndelag, Mid-Norway, that would pass the age of 20 during the period the field stations were in their municipality were eligible to participate in HUNT. Eligible participants, identified through the Norwegian National Population Register, were invited to HUNT by the HUNT research center, and if needed, they were reminded once. For details, see the methods section.

Ethics oversight

The CHS study was approved by IRBs at each of the 4 field centers and the Coordinating Center. CHS is currently under a single IRB at the University of Washington (current approval number MODCR00000825). All CHS participants provided written informed consent. The HUNT study has ethical approval from the Regional Committee for Medical and Health Research Ethics (REK Central Norway 2015/615) and informed consent was obtained from all participants. The UK Biobank has ethical approval from the North West Multi-centre Research Ethics Committee (North West Research Ethics Committee (11/NW/0382), and informed consent was obtained from all participants. The present research was approved by the UK Biobank Research and Access Committee (application number 51784).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the derivation of Proteomic risk scores, participants from CHS were included. In total 456 incident hip fracture cases and 2,715 controls. Validation cohorts included HUNT-SomaScan-5K (187 incident hip fracture cases and 3,072 controls), HUNT-SomaScan-7K (155 incident hip fracture cases and 1,833 controls) and UK-Biobank (686 incident hip fracture cases and 50,190 controls). Sample sizes for both discovery and replication were chosen on the basis of all data available at the time for analysis. No statistical method was used to predetermine sample size.
Data exclusions	To reduce potential selection bias, participants selected for the UKB Covid-19 study were excluded from this study as they were not randomized into the subset used in proteomic studies.
Replication	Validation of the proteomic risk score were performed in 3 cohorts (described above). All replications using the weighted proteomic risk score were successful while the replications in the HUNT validation cohort of the two proteomic risk scores based on machine learning were less convincing as described in the article.
Randomization	Not applicable (It is population-based longitudinal cohort study and not a randomized clinical trial)
Blinding	Not applicable (It is population-based longitudinal cohort study and not a randomized clinical trial)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA