BMJ Open

The Irish dual-energy X-ray absorptiometry (DXA) Health Informatics Prediction (HIP) for Osteoporosis Project

Erjiang E [1], Tingyan Wang,[1,2] Lan Yang,[1,3] Mary Dempsey,[3] Attracta Brennan,[4] Ming Yu,[1] Wing P. Chan,[5] Bryan Whelan,[6,7] Carmel Silke,[6,7] Miriam O'Sullivan,[6,7] Bridie Rooney,[8] Aoife McPartland,[7] Gráinne O'Malley,[6,8] John J. Carey [6,9]

For numbered affiliations see end of article.

Correspondence to
Dr John J. Carey;
john.j.carey@nuigalway.ie

ABSTRACT

Purpose The purpose of the Irish dual-energy X-ray absorptiometry (DXA) Health Informatics Prediction (HIP) for Osteoporosis Project is to create a large retrospective cohort of adults in Ireland to examine the validity of DXA diagnostic classification, risk assessment tools and management strategies for osteoporosis and osteoporotic fractures for our population.

Participants The cohort includes 36 590 men and women aged 4–104 years who had a DXA scan between January 2000 and November 2018 at one of 3 centres in the West of Ireland.

Findings to date 36 590 patients had at least 1 DXA scan, 6868 (18.77%) had 2 scans and 3823 (10.45%) had 3 or more scans. There are 364 unique medical disorders, 186 unique medications and 46 DXA variables identified and available for analysis. The cohort includes 10 349 (28.3%) individuals who underwent a screening DXA scan without a clear fracture risk factor (other than age), and 9947 (27.2%) with prevalent fractures at 1 of 44 skeletal sites.

Future plans The Irish DXA HIP Project plans to assess current diagnostic classification and risk prediction algorithms for osteoporosis and fractures, identify the risk predictors for osteoporosis and develop novel, accurate and personalised risk prediction tools, by using the large multicentre longitudinal follow-up cohort. Furthermore, the dataset may be used to assess, and possibly support, multimorbidity management due to the large number of variables collected in this project.

Strengths and limitations of this study

► We have established a large anonymised database of >36 000 individuals with recorded demographics, DXA metrics, clinical and therapeutic variables, around one-fourth of whom have a prevalent fracture and one-fourth of whom have no identifiable major risk factor for osteoporosis.

► These data can be used to provide substantive exploration of the validity of current tools to identity people with low bone mineral density or at risk of fracture.

► Important limitations include the lack of data on under-represented populations, laboratory and other variables and validated prospective clinical variables and medications, and data from other regions in the country; we plan to address these in future studies.

INTRODUCTION

Today, non-communicable diseases (NCDs) represent the most common disorders worldwide, including musculoskeletal conditions like arthritis and osteoporosis.[1] Osteoporosis is a global health crisis currently affecting millions of men, women and children, with many more at risk.[2] Fragility fractures represent the clinical manifestation of osteoporosis. Predictions suggest that the annual incidence of fragility fractures will rise by 23.3% from 2017 to 2030 in six European countries, whose combined populations represent >60% of Europe's population.[2] Fractures result in substantial morbidity, increased mortality and associated social and healthcare costs.[3] More than 1 million quality-adjusted life-years are lost due to osteoporotic-related fractures in Europe each year.[2] In 2010, approximately 43 000 European deaths were fracture related while expenditure related to osteoporosis exceeded €37 billion.[4 5] Although inpatient hospital costs for osteoporotic fractures are similar to, or exceed those of other NCDs such as cancer and cardiovascular disease,[6 7] they receive considerably less attention. This may be one reason why the osteoporosis care provided to patients presenting with fracture is often narrowly focused and inadequate,[8] notwithstanding the fact that studies show that preventative treatment can reduce fracture risk by 33%–50%.[9] The accurate and timely identification of high-risk or high-cost patients is critical to facilitate effective care and reduce the burden of this disease.

The majority of osteoporotic fractures are related to genetics, patients' physical attributes, lifestyle, other illnesses and medications.[10–13] Many fractures are preventable by identifying those at risk before they fracture,

and then using proven, safe, cheap and effective interventions.[2] Bone mineral density (BMD) is the single best predictor of osteoporotic fracture in a postmenopausal woman without fracture.[12–19] BMD is typically measured non-invasively by dual-energy X-ray absorptiometry (DXA) machines. Many patients undergo DXA scanning to measure their BMD in order to estimate risk, classify them as osteoporotic (or not) and monitor the effects of interventions.[20 21] The combination of BMD with other factors (eg, smoking, blood pressure, blood sugar) enhances this assessment, and an array of risk tools have been developed, all with corresponding strengths and limitations.[22] However, even if a 'best tool' is available, a 'one-size-fits-all' approach with the 'best' algorithm has limitations in clinical practice; with a more personalised approach being preferable.[23]

WHO and its member states have committed to a Sustainable Development Goal, which aims to address NCDs and substantially reduce their burden by 2030.[24] Electronic health information (EHI) has great potential to improve the quality, efficiency and effectiveness of healthcare diagnosis, monitoring and provision while substantially reducing costs.[25 26] Large amounts of data are generated annually in healthcare for the provision of patient care, and for regulatory compliance, yet much of these data are underused, lying idle in large digital repositories.[25] Big data is defined as 'high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making and process automation'.[27] Advances in big data analytics and novel applications like artificial intelligence (AI) can advance this goal by improving disease detection and treatment at a population level, while also offering a personalised approach for individuals.[28] Presently, in osteoporosis care worldwide, the primary sources of EHI are DXA machines, hospital health systems, electronic medical records (EMRs) and hospital administrative systems.

Despite the fact that Irish people are recognised as one of the highest risk groups for fracture worldwide,[2] limited robust data exists on the burden of osteoporotic illness, the validity of fracture risk and appropriate classification and management criteria for the Irish population. In part way to addressing this limitation, the authors undertook a systematic review of studies for Ireland examining 'DXA', 'fracture risk assessment' and 'fracture risk prediction'. Two hundred three papers were identified, 39 of which evaluated risk prediction or DXA diagnostic criteria in their results. Most studies were small: four had >1000 subjects and only one had >10 000 subjects. None of the studies systematically evaluated the various DXA diagnostic criteria, seven papers compared other evaluation methods with DXA (ie, ultrasound, nail spectroscopy, CT, fracture risk tools), and seven papers evaluated various fracture risk factors or risk tools. One paper 'validated' the Fracture Risk Assessment Tool (FRAX) using a limited dataset from public hospital records on hip fractures. However, this paper did not include data on the accuracy of hip fracture diagnosis, other facilities, other fractures, DXA or other clinical information.[29] However, one study (n=90) did evaluate the ability of FRAX to correctly classify Irish people for treatment.[30]

COHORT DESCRIPTION

Patient and public involvement

This study was completed without participant involvement. A waiver of consent was granted by the Ethics Committee.

Source data

DXA scan data were collected at three hospital sites from four DXA machines. Each of these sites was located in a publicly funded academic teaching hospital: one in the Department of Geriatric Medicine at Sligo University Hospital, one in the Department of Rheumatology, Manorhamilton University Hospital and two in the Department of Rheumatology, Merlin Park, Galway University Hospital. One machine in Merlin Park hosted data imported from older scanners in 2005; scanners which had fewer software features. The source data from all four DXA machines (GE Lunar Prodigy Advance, Madison, Wisconsin, USA; software V.17.0) were extracted, validated, cleaned and merged into a single large anonymised data set.

Initial project plans included merging DXA data with EMRs used by the rheumatology and fracture liaison services, the Hospital Inpatient Enquiry System and the Electronic Discharge Summary systems, all of which were available electronically through the respective hospitals' information system. However, the Irish interpretation of the new General Data Protection Regulation (GDPR) legislation precluded such a measure without explicit patient consent. As this requirement would require significant extra staffing and funding to be feasible, only DXA scanner data are presented in this paper.

All patients were referred for a DXA scan by their doctor, including general practitioners, orthopaedic fracture clinics, osteoporosis services and other medical specialties, the majority of whom had an appropriate indication for a DXA scan.[31] All scans meet current national legislative requirements, and operate with best international practice as set forth by The International Society for Clinical Densitometry.[32] The DXA clinical team members are experienced, trained and International Society for Clinical Densitometry-certified in performing and interpreting bone densitometry; each clinic has similar standard operating procedures. Furthermore, each clinic is also the primary hub for osteoporosis and fracture liaison services in their respective hospitals. Patients who have been referred with a 'history of fracture' have this verified by the DXA technologist at the time of scanning, by reviewing the respective hospitals' electronic radiology systems. These systems are inaccessible to external facilities. Patients without an appropriate indication for a scan are returned to the referring physician, with either

a request for additional information or the referral being declined. The use of a protocol to consider DXA screening for younger 'healthier' patients (based on older criteria) would act as an early detection and treatment innervation to reduce fracture-related incidence in later years in the Irish population.

Cohort

All subjects who had at least one DXA scan on one of the four DXA scanners between 22 January 2000 and 1 November 2018 were included in the dataset. Also excluded were scans of DXA phantoms and scans, where demographics were not entered on the DXA machine. DXA scans with extreme values which appeared erroneous or could not be validated were also excluded. These included scans where the height <1 m or >2 m, weight <20 kg or more than than the machine limit (110–158 kg), body mass index >70 and a BMD <0.0 or >5.0. The data cleaning procedure was carried out prior to final GDPR legislation implementation when it was possible to identify and trace the scan using coded data. All data have since been completely anonymised such that it cannot be traced back to or used to identify any individual and/ or linked with any other data. This is in compliance with current GDPR legislation and the project's ethical approval requirements.

Control group

The control group was selected from the cohort. The control group is defined as those referred for a screening DXA scan and without having any other risk factor listed (other than age), including fracture. This control group, although not a random selection of healthy community volunteers, will serve in further comparative studies as the 'healthy control group'. The primary objectives are correct risk classification and diagnosis, and fracture risk assessment. Subsequent data analyses will include a comparison of those with fractures with those without, together with a comparison of different fracture types in the five major categories, as listed by the International Osteoporosis Foundation (ie, hip, spine/vertebral, humerus, forearm and other).

Variables

When a patient is referred for a DXA scan, important information such as clinical diagnoses and medications is collected and input to the DXA machine. All clinical information, including fracture sites, and medications for this project, has been independently reviewed and verified by a senior clinical investigator (JC) and listed as either a 'risk factor', 'protective factor', 'no effect' or 'osteoporosis or bone treatment'. Fractures are also grouped into one of five categories (ie, hip, spine/vertebral, humerus, forearm and other). Patient demographics include age, gender, race, height and weight. DXA variables include bone mineral content, bone area, BMD, T-score (National Health and Nutrition Examination Survey (NHANES) III/General Electric (GE)), Z-score

(NHANES III/GE, no weight adjustment), fat mass, lean mass, vertebral height, hip geometry, FRAX 10-year risk of major osteoporotic fracture and hip fracture.

Findings to date

A total of 37 053 subjects (19 years of data) from 4 DXA machines located at three hospital sites in Ireland were extracted. Following cleaning and validation of the data prior to data analysis, 36 590 unique subjects remained, each of whom had at least one DXA scan, ranging in age from 4 to 104 years. A total of 25 899 (70.78%) subjects had one scan only, 6868 (18.77%) subjects had two scans and 3823 (10.45%) patients had three or more scans. The majority of the scans were performed in Merlin Park University Hospital (53%), 31% were performed in Manorhamilton University Hospital and 16% in Sligo University Hospital. A summary of data completeness and availability is presented in table 1.

A total of 10 349 subjects who had been referred for a DXA scan had no obvious fracture risk factor (other than age in some cases) were selected to serve as the control group. A total of 9947 patients had at least 1 prior fracture, at 1 of 44 skeletal sites. Available for data analysis are 364 unique medical disorders, 186 unique medications and 46 DXA variables. There is considerable heterogeneity regarding comorbidities and medications. This likely reflects local practice and the referral populations being served.

Summary details for the female and male subjects are presented in tables 2 and 3, respectively. An unpaired Student's t-test was performed for comparison between groups for continuous variables, while a χ^2 test was performed for categorical variables. A p value <0.05 was deemed statistically significant. In terms of race and gender, almost 100% of the subjects are Caucasian (as expected), with 86% being female. The mean age of the female subjects is 60 years; 70% of the female subjects are postmenopausal women. Although 27% of the female subjects had a prior fracture, only 23% are taking calcium and/or vitamin D and <11% are on osteoporosis therapy. In general, male subjects are older (62.4 years vs 60.5 years), taller (172.3 cm vs 160.4 cm), heavier (80.8 kg vs 68.4 kg) and more likely to be corticosteroid users (24.5% vs 9.5%, p<0.001), although a similar proportion are smokers (9.9% vs 9.3%, p=0.196). While the male subjects have a higher mean femoral neck BMD (0.901 g/cm^2 vs 0.842 g/cm^2) and a lower 10-year fracture risk (11.03% vs 7.34%), all (p<0.001), they have a similar prevalence of fracture (27.3% vs 27.2%) and are less likely to be taking osteoporosis medication (7.3% vs 10.5%, p<0.01).

All analyses were performed using Minitab V.16.0, Coventry, UK, R V.3.5.1 and MS SQL Server Management Studio 18.

Future plans

Osteoporosis data for the Irish population are restricted in terms of national applicability and generality, with the majority of the studies limited to small heterogeneous

Table 1 Summary of data completeness and availability of demographic, clinical, biometric and DXA scan variables

Variables	MPH	MH	SH	Total
Number (%)	19 444 (53.1)	11 181 (30.6)	5965 (16.3)	36 590 (100)
Number of people we have age for (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Number of people we have gender (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Ethnicity, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Height, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Weight, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
BMI, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Menopausal status, women only, N (%)	17 142 (88.1)	9534 (85.3)	4869 (81.6)	31 545 (86.2)
Smoking status, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Status of family history of osteoporosis, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Status of prior fracture, N (%)	19 444 (100)	11 181 (100)	5965 (100)	36 590 (100)
Number of unique medical conditions	68	101	345	364
Number of unique medications	39	24	176	186
Number of scans				
Hip	23 931	31 906	14 628	70 465
Spine	24 434	19 179	8149	51 762
Forearm	947	1555	1000	3502
Whole body	122	336	127	585
LVA/DVA	8799	4133	534	13 466
Number of scan variables*	46	46	46	46

*Includes variables such as: bone mineral content, bone area, bone mineral density, T-score, Z-score, fat mass, lean mass, hip axis length, fracture risk score.
BMI, body mass index; DVA, dual vertebral assessment; DXA, dual-energy X-ray absorptiometry; LVA, lateral vertebral assessment; MH, Manorhamilton Hospital; MPH, Merlin Park Hospital; N (%), number (percentage); SH, Sligo Hospital.

studies. While a number of larger studies have been carried out, they have been published using limited information from national public hospital administrative data sets.[29 33 34] All studies which have been carried out to date include very limited information on individuals, do not provide information on patients admitted to other facilities or who were managed as an outpatient or never presented to a hospital, and are not linked to DXA data, medication use or to future events. Furthermore, some studies have questioned the accuracy of the FRAX tool for correctly identifying Irish people for fracture risk.[30 35] Multiple risk models have been proposed using data derived from a variety of sources for predicting low BMD, fractures and comorbidities.[11 36–49] The Health Informatics Prediction (HIP) Project 'big data' approach incorporates DXA material of almost 40 000 patients having approximately 10 000 fractures and a relatively large convenience control cohort spanning an almost 20-year period. The approach also includes the algorithmic risk score, as well as the variables used to calculate such scores. In addition, the project has access to an array of variables, including medications, comorbidities and known risk factors, variables which were not previously evaluated for Irish people. These data, uniquely offer for the first time, an opportunity to

study fragility fracture relationships on a larger scale in a more representative Irish population. These data also offer an opportunity to develop a model which could be replicated globally. In addition, because of the large number of variables, other medications and diseases, this cohort could be used in the future, to study other illnesses of interest including cardiovascular disease, diabetes mellitus and cancer.

In future papers, the authors plan to compare and contrast demographics, DXA biometrics and other clinical risk factors between those with low BMD (or who would be classified as 'osteoporosis' by various DXA criteria), those with prevalent fractures, those without prevalent risk factors, those who subsequently developed fractures and those who did not. We plan to assess the importance of available risk factors for our population, and how much they contribute. The performance of the various fracture risk tools, model assessment strategies and more novel AI methodologies for predicting those with low BMD and those with fractures using traditional variables, will also be evaluated. We will also be able to compare the results with those of other populations, and assess the importance of variables and models for specific fracture sites, in addition to comparing the various models at an aggregate 'one-size-fits-all' approach for

Table 2 Baseline summary of variables of female subjects from three centres

Variables	MPH	MH	SH	Total
Number (%)	17 142 (88.1)	9534 (85.3)	4869 (81.6)	31 545 (86.2)
Age (years), mean (SD)	60.5 (12.3)	60.0 (12.9)	61.7 (14.0)	60.5 (12.8)
Caucasian, N (%)	17 084 (99.66)	9505 (99.70)	4864 (99.90)	31 453 (99.71)
Height (cm), mean (SD)	160.3 (6.5)	160.5 (6.7)	160.3 (6.9)	160.4 (6.6)
Weight (kg), mean (SD)	68.8 (13.9)	67.9 (14.1)	68.0 (14.2)	68.4 (14.0)
BMI (kg/m^2), mean (SD)	26.8 (5.2)	26.3 (5.2)	26.5 (5.3)	26.6 (5.2)
Postmenopausal, N (%)	12 420 (72.5)	5856 (61.4)	4024 (82.6)	22 300 (70.7)
Hormone replacement, n (%)	1167 (6.8)	326 (3.4)	66 (1.4)	1559 (4.9)
Smokers, N (%)	2094 (12.2)	763 (8.0)	67 (1.4)	2924 (9.3)
Family history osteoporosis, N (%)	3563 (20.8)	956 (10.0)	60 (1.2)	4579 (14.5)
Corticosteroid use, N (%)	1744 (10.2)	1081 (11.3)	179 (3.7)	3004 (9.5)
Prior fracture, N (%)	4636 (27.0)	2460 (25.8)	1473 (30.2)	8569 (27.2)
Mean number comorbidities, mean (SD)*	1.55 (0.81)	1.43 (0.71)	1.77 (0.93)	1.6 (0.8)
Frail, N (%)	128 (0.8)	37 (0.4)	171 (3.5)	336 (1.1)
Prior falls, N (%)	93 (0.5)	97 (1.0)	114 (2.3)	304 (1.0)
Calcium and/or vitamin D use, N (%)	4492 (26.2)	1902 (20.0)	1048 (21.5)	7442 (23.6)
Osteoporosis medication use, N (%)†	1649 (9.6)	1034 (10.9)	633 (13.0)	3316 (10.5)
Number of other medications, mean (SD)‡	1.06 (0.24)	1.03 (0.16)	1.10 (0.33)	1.07 (0.26)
FN BMD (g/cm^2), mean (SD)	0.839 (0.139)	0.843 (0.138)	0.846 (0.139)	0.842 (0.139)
FN T-score, mean (SD)§	−1.1 (1.2)	−1.1 (1.1)	−1.1 (1.2)	−1.1 (1.2)
Whole body fat mass (kg), mean (SD)	19.02 (11.7)	29.21 (12.3)	19.7 (11.4)	26.3 (12.8)
Lean body mass (kg), mean (SD)	35.95 (7.1)	39.14 (5.5)	35.5 (6.3)	38.1 (6.0)
Absolute risk of major osteoporotic fracture, median (IQR)¶	11.1 (9.53)	10.6 (9.13)	11.2 (9.32)	11.03 (9.4)
Absolute risk of hip fracture, median (IQR)¶	1.69 (3.54)	1.65 (3.36)	1.91 (3.69)	1.73 (3.55)

*Refers to other recorded medical illnesses/conditions.

†Refers to non-oestrogen therapies: bisphosphonates, SERMs (including raloxifene, SERM, Evista, Tibolone), strontium, denosumab, parathyroid hormone (anabolic).

‡Refers to mean number of other recorded medications being taken by an individual patient.

§FN T-score calculated using NHANES III white female reference.

¶Refers to 10-year risk of MOF calculated from age and BMD, or FRAX. Older technology calculated 10-year risk based on age and BMD using a proprietary algorithm derived from a Swedish 10-year risk prediction model. All machines have been updated to provide a FRAX estimate in the past 5 years.

BMD, bone mineral density; FN, femoral neck; FRAX, Fracture Risk Assessment Tool; LVA, lateral vertebral assessment; MH, Manorhamilton Hospital; MOF, major osteoporotic fracture; MPH, Merlin Park Hospital; N (%), number (percentage); NHANES, National Health and Nutrition Examination Survey; SD, standard deviation; SERM, selective oestrogen receptor modulator; SH, Sligo Hospital.

populations, as well as a more personalised approach for different individuals with specific traits or characteristics.

Multidimensional analyses will be performed on subjects using online analytical processing to see the demographic characteristics and dynamic changes of a specific disease and the marginal effects of risk factors from multiple perspectives, including age, gender, region and effect of changes in BMD, height, weight and other biometrics over time. Data will be summarised using appropriate descriptive statistics, then compared using classification criteria, after which it will be evaluated by multiple fracture prediction algorithms using statistical, cluster and correlation analyses, coupled with classification and regression models. Finally, econometric and machine learning methods will be applied to the data to predict the risk of osteoporotic fractures and/or changes in BMD in patients. For example, we have one paper in press which is about detecting low BMD using machine learning techniques.[50] It is anticipated that the HIP framework (arising from the Irish DXA HIP for Osteoporosis Project) could also be applied to other geographic regions, and the prediction of other NCDs, such as cardiovascular disease, diabetes and cancer. These hypotheses will be explored at a later date.

STRENGTHS AND WEAKNESSES

This study has several strengths including a large data sample size from multiple centres where high-quality performance is a tradition and is embedded in the

Table 3 Baseline variables of male subjects from three centres

Variables	MPH	MH	SH	Total
Number (%)	2302 (11.9)	1647 (14.7)	1096 (18.4)	5045 (13.8)
Age (years), mean (SD)	61.0 (15.9)	62.7 (14.3)	64.8 (16.1)	62.4 (15.5)
Caucasian, N (%)	2287 (99.4)	1640 (99.6)	1093 (99.7)	5020 (99.5)
Height (cm), mean (SD)	172.6 (7.8)	172.2 (7.5)	171.9 (8.0)	172.3 (7.7)
Weight (kg), mean (SD)	81.3 (16.5)	80.7 (15.2)	79.8 (15.4)	80.8 (15.9)
BMI (kg/m^2), mean (SD)	27.3 (5.0)	27.2 (4.5)	26.9 (4.7)	27.2 (4.8)
Smokers/Tobacco users, N (%)	336 (14.6)	151 (9.2)	10 (0.9)	497 (9.9)
Family history of osteoporosis, N (%)	217 (9.4)	65 (3.9)	2 (0.2)	284 (5.6)
Corticosteroid use, N (%)	731 (31.8)	387 (23.5)	120 (10.9)	1238 (24.5)
Prior fracture, N (%)	733 (31.8)	380 (23.1)	265 (24.2)	1378 (27.3)
Number of medical diseases, mean (SD)	1.43 (0.69)	1.35 (0.59)	1.84 (0.94)	1.5 (0.8)
Frail, N (%)	20 (0.9)	19 (1.2)	106 (9.7)	145 (2.9)
Prior falls, N (%)	25 (1.1)	25 (1.5)	43 (3.9)	93 (1.8)
Taking calcium and/or vitamin D, N (%)	443 (19.2)	192 (11.7)	170 (15.5)	805 (16.0)
Taking osteoporosis medication, N (%)	172 (7.5)	95 (5.8)	101 (9.2)	368 (7.3)
Number of other medications, mean (SD)	1.03 (0.18)	1 (0)	1.20 (0.48)	1.11 (0.35)
Femoral neck BMD (g/cm^2), mean (SD)	0.898 (0.157)	0.904 (0.148)	0.899 (0.159)	0.901 (0.153)
FN T-score (white female NHANES III), decimal place OK, mean (SD)	−1.5 (1.3)	−1.3 (1.1)	−1.3 (1.2)	−1.3 (1.2)
Whole body fat mass (kg), mean (SD)	14.5 (10.8)	25.8 (10.0)	17.2 (9.7)	21.3 (11.1)
Lean body mass (kg), mean (SD)	44.0 (13.0)	55.0 (7.6)	48.9 (14.4)	51.4 (11.7)
Absolute risk of major osteoporotic fracture, median (IQR)	7.67 (6.22)	6.79 (4.88)	7.35 (5.69)	7.34 (5.66)
Absolute risk of hip fracture, median (IQR)	1.95 (3.45)	1.73 (2.5)	2.23 (3.57)	1.96 (1.96)

BMD, bone mineral density; FN, femoral neck; LVA, lateral vertebral assessment; MH, Manorhamilton Hospital; MPH, Merlin Park Hospital; N (%), number (percentage); NHANES, National Health and Nutrition Examination Survey; SD, standard deviation; SH, Sligo Hospital.

practice. Furthermore, there is a long follow-up period for many patients and a significant number of demographic, biometric and clinical variables. As a substantial portion of the cohort appear to be at 'low risk' for fracture at baseline, they can serve as a large convenience control group. Meanwhile, as >25% of the cohort also has a fracture, this will enable more robust analyses of fracture risk algorithms. Given that the prevalence of osteoporosis and low BMD depends on the criteria and calculation methods used,[51] this will require further exploration and explanation in an Irish context. Some patients in the dataset have had multiple scans and provide an opportunity to explore the marginal effect of risk factors such as changes in BMD with or without intervention and how such changes modify the outcome. Osteoporosis and the care of patients with fragility fractures is a global problem where current practice and the comprehensive utilisation of healthcare big data remains unsatisfactory,[15 52 53] as the majority of patients are neither diagnosed, investigated nor treated for their skeletal fragility following fracture.[52 54 55] Irish studies suggest that some patients with osteoporosis are similarly not treated, while others may be treated even in the absence of a DXA scan, fracture or

high fracture risk.[56–58] At first glance, it appears that many patients with fractures are not on treatment. There is an opportunity in this project to explore this further.

All research has its limitations. Some of the limitations of this study include the fact that >99% of the cohort are Caucasian, people from other parts of the country have not been included, those who have not had a DXA scan and patients who were scanned on other technologies are also excluded. While the current dataset has only 117 patients reporting other ethnicities, including 68 Asian, 18 black, 15 Hispanic and 16 'other'; if our results are replicated in other populations, it would suggest that they are robust and of greater interest and relevance. Furthermore, although we have information on prescribed medications, we do not have a complete list, doses or duration of therapy. We also acknowledge that considerable omissions can exist in the data processing stage. Although the data involved in our study have been carefully examined and verified, omissions can still occur due to the incomplete access to the entire patients' information records, such as scan images and records in other departments. It is possible that some healthy patients were misdiagnosed with fractures. It is also probable that some patients were

diagnosed with fracture but, only in a suspected condition. In addition, the data are reliant in many instances on patients' self-reporting for other risk factors and medication use, and/or the reports of their referring clinician all of which cannot be accurately verified. Since the introduction of GDPR legislation, we are no longer able to link DXA data to other datasets or back to the original sources, as all data have been anonymised. Ideally, we would be able to obtain consent from patients and track this information more carefully going forward, but in order to do so, it will require research funding or the allocation of staff resource to do this prospectively. However, this paper and other planned publications outlining the scope, validity and robustness of these measures and tools, will generate results and potential further hypotheses, which in turn will future support further collaborations for prospective cohort studies incorporating DXA technology for predicting fractures, cardiovascular disease and other morbidities.

Author affiliations
[1]Department of Industrial Engineering, Tsinghua University, Beijing, China
[2]Nuffield Department of Medicine, University of Oxford, Oxford, UK
[3]School of Engineering, National University of Ireland Galway, Galway, Ireland
[4]School of Computer Science, National University of Ireland Galway, Galway, Ireland
[5]Department of Radiology, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan
[6]School of Medicine, National University of Ireland Galway, Galway, Ireland
[7]Department of Rheumatology, Our Lady's University Hospital, Manorhamilton, Ireland
[8]Department of Geriatric Medicine, Sligo University Hospital, Sligo, Ireland
[9]Department of Rheumatology, Galway University Hospitals, Galway, Ireland

Collaborators All data for this project have been completely anonymised and are no longer traceable to source or identifiers coded or otherwise. All data are stored and handled confidentially. Currently, only a small group of the team has access to these data (EE, YL, JC) while the clinical team have access to the clinical systems but cannot link this to the study data. This Irish DXA HIP (Health Informatics Prediction) for Osteoporosis Project represents collaboration between the disciplines of clinical medicine and research, computer engineering, industrial engineering and big data science.

Contributors All authors were involved in the development of the programme, plans for analyses, have reviewed the results and participated in the writing and reviewing of the manuscript. The clinical members (JC, WPC, BW, CS, MO, BR, AM and GO'M) were involved in collection, the engineering and computer science members (EE, TW, LY, MD, AB and MY) in data cleaning, merging and analysis and five investigators (EE, LY, AB, MD and JC) at every stage to ensure the validity of data collection, processing and analyses.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Prior to data collection, formal approval was sought and obtained for each clinical site from the local Hospitals' Research Ethics Committee to process information from a retrospective cohort with the aim of exploring DXA validity for Ireland. There was considerable consideration around the use of, and access to, clinical systems, data validity, processing, storage and analytics, particularly in light of new legislation concerning data protection in the European Union, namely the General Data Protection Regulation (available at: https://gdpr-info.eu/https://gdpr-info.eu/).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Open access for published data. Original data set on file with authors.

ORCID iDs
Erjiang E http://orcid.org/0000-0001-7108-5740
John J. Carey http://orcid.org/0000-0001-7292-2328

REFERENCES

1 WHO. *Global health estimates 2016: disease burden by cause, age, sex, by country and by region, 2000-2016*. Geneva, 2018.
2 Foundation IO. IOF compendium of osteoporosis 2019.
3 Kanis JA, Cooper C, Rizzoli R, *et al*. Identification and management of patients at increased risk of osteoporotic fracture: outcomes of an ESCEO expert consensus meeting. *Osteoporos Int* 2017;28:2023–34.
4 Kanis JA. Diagnosis and clinical aspects of osteoporosis. In: *Pocket reference to osteoporosis*. Springer, 2019: 11–20.
5 International Osteoporosis Foundation. The global burden of osteoporosis: a Factsheet: international osteoporosis Foundation, 2014. Available: https://www.iofbonehealth.org/sites/default/files/media/PDFs/Fact%20Sheets/2014-factsheet-osteoporosis-A4.pdf [Accessed 12 Oct 2015].
6 Burden of illness for osteoporotic fractures compared with other serious diseases among postmenopausal women in the United States. *Mayo Clinic Proceedings: Elsevier*, 2015.
7 Piscitelli P, Iolascon G, Argentiero A, *et al*. Incidence and costs of hip fractures vs strokes and acute myocardial infarction in Italy: comparative analysis based on national hospitalization records. *Clin Interv Aging* 2012;7:575–83.
8 Richy F, Gourlay M, Ross PD, *et al*. Validation and comparative evaluation of the osteoporosis self-assessment tool (Ost) in a Caucasian population from Belgium. *Int J Med* 2004;97:39–46.
9 Hochberg M. Preventing fractures in postmenopausal women with osteoporosis. *Drugs Aging* 2000;17:317–30.
10 Rudman HA, Birrell F, Pearce MS, *et al*. Obesity, bone density relative to body weight and prevalent vertebral fracture at age 62 years: the Newcastle thousand families study. *Osteoporos Int* 2019;30:829–36.
11 Tian L, Yu X. Fat, sugar, and bone health: a complex relationship. *Nutrients* 2017;9:506.
12 Siris ES, Chen Y-T, Abbott TA, *et al*. Bone mineral density thresholds for pharmacological intervention to prevent fractures. *Arch Intern Med* 2004;164:1108–12.
13 Biver E, Durosier-Izart C, Chevalley T, *et al*. Evaluation of radius microstructure and areal bone mineral density improves fracture prediction in postmenopausal women. *J Bone Miner Res* 2018;33:328–37.
14 Black DM, Cauley JA, Wagman R, *et al*. The ability of a single BMD and fracture history assessment to predict fracture over 25 years in postmenopausal women: the study of osteoporotic fractures. *J Bone Miner Res* 2018;33:389–95.
15 Gourlay ML, Ritter VS, Fine JP, *et al*. Comparison of fracture risk assessment tools in older men without prior hip or spine fracture: the MROS study. *Arch Osteoporos* 2017;12:91.
16 Kanis JA, Bianchi G, Bilezikian JP, *et al*. Towards a diagnostic and therapeutic consensus in male osteoporosis. *Osteoporos Int* 2011;22:2789–98.
17 Langsetmo L, Leslie WD, Zhou W, *et al*. Using the same bone density reference database for men and women provides a simpler estimation of fracture risk. *J Bone Miner Res* 2010;25:2108–14.
18 Grgic O, Chung K, Shevroja E, *et al*. Fractures in school age children in relation to sex and ethnic background: the generation R study. *Bone* 2019;121:227–31.
19 Clark EM, Tobias JH, Ness AR. Association between bone density and fractures in children: a systematic review and meta-analysis. *Pediatrics* 2006;117:e291–7.
20 Carey JJ, Buehring B. Current imaging techniques in osteoporosis. *Clin Exp Rheumatol* 2018;36 (Suppl 114):115–26.
21 Carey JJ, Delaney MF. Utility of DXA for monitoring, technical aspects of DXA BMD measurement and precision testing. *Bone* 2017;104:44–53.

22 Rubin KH, Friis-Holmberg T, Hermann AP, *et al*. Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review. *J Bone Miner Res* 2013;28:1701–17.

23 Breda J, Wickramasinghe K, Peters DH, *et al*. One size does not fit all: implementation of interventions for non-communicable diseases. *BMJ* 2019;347:l6434.

24 WHO. Who global coordination mechanism on the prevention and control of NCDS 2016.

25 Mehta N, Pandit A. Concurrence of big data analytics and healthcare: a systematic review. *Int J Med Inform* 2018;114:57–65.

26 Zhang L, Wang H, Li Q, *et al*. Big data and medical research in China. *BMJ* 2018;360:j5910.

27 Gartner. Big Data. Available: https://www.gartner.com/it-glossary/big-data/ [Accessed 02 Mar 2017].

28 Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. *NPJ Digit Med* 2018;1:1–4.

29 McGowan B, Kanis JA, Johansson H, *et al*. Development and application of FRAX in the management of osteoporosis in Ireland. *Arch Osteoporos* 2013;8:146.

30 Brewer L, Mellon L, Duggan J. Ability of fracture risk assessment tool and national osteoporosis guideline group guidance to stratify people appropriately before fracture. *J Am Geriatr Soc* 2013;61:1633–4.

31 Mohammad A, Aamir MU, Mooney S, *et al*. Appropriateness of referrals to a tertiary referral centre for bone mineral density testing. *Ir J Med Sci* 2014;183:533–7.

32 Lewiecki EM, Binkley N, Morgan SL, *et al*. Best practices for dual-energy X-ray absorptiometry measurement and reporting: International Society for clinical densitometry guidance. *J Clin Densitom* 2016;19:127–40.

33 McGowan B, Casey MC, Silke C, *et al*. Hospitalisations for fracture and associated costs between 2000 and 2009 in Ireland: a trend analysis. *Osteoporos Int* 2013;24:849–57.

34 Kelly MA, McGowan B, McKenna MJ, *et al*. Emerging trends in hospitalisation for fragility fractures in Ireland. *Ir J Med Sci* 2018;187:601–8.

35 Beattie JR, Cummins NM, Caraher C, *et al*. Raman spectroscopic analysis of fingernail clippings can help differentiate between postmenopausal women who have and have not suffered a fracture. *Clin Med Insights Arthritis Musculoskelet Disord* 2016;9:CMAMD. S38493–16.

36 Crandall CJ, Larson J, LaCroix A, *et al*. Predicting fracture risk in younger postmenopausal women: comparison of the Garvan and FRAX risk calculators in the women's health Initiative study. *J Gen Intern Med* 2019;34:235–42.

37 Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *BMJ* 2009;339:b4229.

38 Kanis JA, Odén A, Johnell O, *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* 2007;18:1033–46.

39 Britton KA, Massaro JM, Murabito JM, *et al*. Body fat distribution, incident cardiovascular disease, cancer, and all-cause mortality. *J Am Coll Cardiol* 2013;62:921–5.

40 Compston JE, Flahive J, Hosmer DW, *et al*. Relationship of weight, height, and body mass index with fracture risk at different sites in postmenopausal women: the global longitudinal study of osteoporosis in women (glow). *J Bone Miner Res* 2014;29:487–93.

41 Prieto-Alhambra D, Premaor MO, Avilés FF, *et al*. Relationship between mortality and BMI after fracture: a population-based study of men and women aged ≥40 years. *J Bone Miner Res* 2014;29:1737–44.

42 Rudman HA, Birrell F, Pearce MS, *et al*. Obesity, bone density relative to body weight and prevalent vertebral fracture at age 62 years: the Newcastle thousand families study. *Osteoporos Int* 2019;30:829–36.

43 Schousboe JT, Taylor BC, Kiel DP, *et al*. Abdominal aortic calcification detected on lateral spine images from a bone densitometer predicts incident myocardial infarction or stroke in older women. *J Bone Miner Res* 2008;23:409–16.

44 Tankó LB, Christiansen C, Cox DA, *et al*. Relationship between osteoporosis and cardiovascular disease in postmenopausal women. *J Bone Miner Res* 2005;20:1912–20.

45 Vertebral fracture assessment-detected abdominal aortic calcification and cardiovascular disease in rheumatoid arthritis. *Seminars in arthritis and rheumatism*. Elsevier, 2014.

46 Bartosch P, McGuigan FE, Akesson KE. Progression of frailty and prevalence of osteoporosis in a community cohort of older women-a 10-year longitudinal study. *Osteoporos Int* 2018;29:2191–9.

47 Bauer DC*et al*. Factors associated with appendicular bone mass in older women. *Ann Intern Med* 1993;118:657–65.

48 Lim J, Kim J, Cheon S. A deep neural network-based method for early detection of osteoarthritis using statistical data. *Int J Environ Res Public Health* 2019;16:1281.

49 Rubin KH, Möller S, Holmberg T, *et al*. A new fracture risk assessment tool (FREM) based on public health registries. *J Bone Miner Res* 2018;33:1967–79.

50 Erjiang E, Wang T, Yang L, *et al*. Machine learning can improve clinical detection of low BMD: the DXA-HIP study. *J Clin Densitom* 2020. doi:10.1016/j.jocd.2020.10.004. [Epub ahead of print: 20 Oct 2020].

51 Carey JJ, Delaney MF, Love TE, *et al*. Dual-Energy X-ray absorptiometry diagnostic discordance between Z-scores and T-scores in young adults. *J Clin Densitom* 2009;12:11–16.

52 Curtis JR, Adachi JD, Saag KG. Bridging the osteoporosis quality chasm. *J Bone Miner Res* 2009;24:3–7.

53 Lewiecki EM, Leader D, Weiss R, *et al*. Challenges in osteoporosis awareness and management: results from a survey of US postmenopausal women. *J Drug Assess* 2019;8:25–31.

54 Cram P, Lix LM, Bohm E, *et al*. Hip fracture care in Manitoba, Canada and new York state, United States: an analysis of administrative data. *CMAJ Open* 2019;7:E55–62.

55 Lewiecki EM, Wright NC, Curtis JR, *et al*. Hip fracture trends in the United States, 2002 to 2015. *Osteoporos Int* 2018;29:717–22.

56 Queally JM, Kiernan C, Shaikh M, *et al*. Initiation of osteoporosis assessment in the fracture clinic results in improved osteoporosis management: a randomised controlled trial. *Osteoporos Int* 2013;24:1089–94.

57 Harlin SL, Harlin RD, Sherman TI, *et al*. Using a structured, computer-administered questionnaire for evaluating health-related QOL with chronic lower extremity wounds. *Ostomy Wound Manage* 2009;55:30.

58 Genant HK, Cooper C, Poor G, *et al*. Interim report and recommendations of the world Health organization Task-Force for osteoporosis. *Osteoporos Int* 1999;10:259–64.