



Real-world evidence: new opportunities for osteoporosis research. Recommendations from a Working Group from the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO)

Rebecca J. Moon^{1,2} · Jean-Yves Reginster^{3,4} · Nasser M. Al-Daghri⁵ · Jotheeswaran A. Thiyagarajan⁶ · Charlotte Beaudart^{3,4} · Olivier Bruyère^{3,4} · Nansa Burlet⁴ · Manju Chandran⁷ · Mario Coelho da Silva⁸ · Philip G. Conaghan^{9,10} · Willard H. Dere¹¹ · Adolfo Diez-Perez¹² · Peyman Hadji^{13,14} · Philippe Halbout¹⁵ · Mickaël Hilgsmann¹⁶ · John A. Kanis^{17,18} · Eugene V. McCloskey^{19,20} · Sif Ormarsdottir²¹ · Daniel Prieto-Alhambra²² · Régis P. Radermecker²³ · René Rizzoli²⁴ · Yousef Al-Saleh^{25,26,27} · Stuart L. Silverman²⁸ · Lee S. Simon²⁹ · Friederike Thomasius¹³ · Tjeerd van Staa³⁰ · Andrea Laslop³¹ · Cyrus Cooper^{1,32,33} · Nicholas C. Harvey^{1,32}

Received: 3 March 2023 / Accepted: 28 April 2023 / Published online: 23 June 2023

© The Author(s) 2023

Abstract

Summary This narrative review summarises the recommendations of a Working Group of the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) for the conduct and reporting of real-world evidence studies with a focus on osteoporosis research.

Purpose Vast amounts of data are routinely generated at every healthcare contact and activity, and there is increasing recognition that these real-world data can be analysed to generate scientific evidence. Real-world evidence (RWE) is increasingly used to delineate the natural history of disease, assess real-life drug effectiveness, understand adverse events and in health economic analysis. The aim of this work was to understand the benefits and limitations of this type of data and outline approaches to ensure that transparent and high-quality evidence is generated.

Methods A ESCEO Working Group was convened in December 2022 to discuss the applicability of RWE to osteoporosis research and approaches to best practice.

Results This narrative review summarises the agreed recommendations for the conduct and reporting of RWE studies with a focus on osteoporosis research.

Conclusions It is imperative that research using real-world data is conducted to the highest standards with close attention to limitations and biases of these data, and with transparency at all stages of study design, data acquisition and curation, analysis and reporting to increase the trustworthiness of RWE study findings.

Keywords Drug effectiveness · Fracture · Osteoporosis · Real-world evidence · Epidemiology · Registry

Introduction

In 2019 in the European Union, UK and Switzerland, there was an estimated 32 million individuals living with osteoporosis and 4.3 million new fragility fractures during that year [1]. Despite the development of multiple effective primary and secondary preventive measures, ageing populations will mean an increasing burden of osteoporosis and associated

fragility fractures; it was estimated that, in 2010, globally, 158 million individuals aged over 50 years were at high risk of fracture, but by 2040, this number is expected to double [2]. As such, there is a need to accelerate drug discovery and approaches to optimising osteoporosis management to reduce fracture risk.

Traditionally, randomised controlled trials (RCT) have been considered the gold standard in establishing evidence-based treatment safety and efficacy, and in osteoporosis, they have also been used to demonstrate the benefit of

Extended author information available on the last page of the article

non-drug interventions including community screening to identify those at high risk of fracture [3, 4]. RCTs are an important source of evidence to assess causation between an intervention and an outcome; the process of randomisation ensures any observed differences in baseline characteristics between treatment groups are due to chance, and blinding, when used, prevents treatment allocation influencing behaviour and outcomes. However, these studies are costly and establishing long-term benefits of new interventions can take many years. Careful consideration of the generalisability of RCT findings is also required due to study inclusion and exclusion criteria that often exclude those with multimorbidity and polypharmacy and may limit participants to a specific sex or age group [5]. Indeed, often the target population for osteoporosis medication are excluded from the trials. Reyes et al. examined real-life users of alendronate in Sistema d'Informació per al Desenvolupament de l'Investigació en Atenció Primària (SIDIAP) Database from Catalonia (Spain) and the Danish Health Registries (DHR) from Denmark, and found that 56 and 63% of users, respectively, would have been excluded from the Fracture Intervention Trial (FIT) that established the efficacy of alendronate in reducing fractures in women with osteoporosis [6, 7]. Differences in healthcare provision between RCT sites and where the intervention will be used should be considered; for example, the number of medical consultations is often higher in a trial setting than during routine clinical care, and additionally, there is self-selection of those willing to participate in a clinical trial and often higher compliance. In particular, Black, Indigenous and people of colour (BIPOC) populations are often underrepresented in clinical trials [8]; yet, findings of RCTs are commonly extrapolated to the clinical care of more diverse populations than those in which the intervention was assessed. Despite 29% of osteoporotic fractures in the United States of America (USA) occurring in men [9], clinical trials assessing treatments for osteoporosis in men have typically been smaller and shorter in duration than comparable trials in women, with many using bone mineral density rather than fracture as the endpoint [10, 11]. Approval of these drugs in men has relied on the data collected in women and an assumption of a similar anti-fracture efficacy in both sexes [12].

Real-world evidence (RWE), which utilises data collected during routine healthcare, can complement research findings from RCTs and observational studies in which data are collected primarily for the purpose of research. RWE has more commonly been used for rare diseases, but its potential use in common health conditions is increasing to understand the natural history of disease, establish and compare the effectiveness of interventions and assess long-term drug safety in routine clinical care, guide regulatory and national reimbursement decisions, improve health and social care delivery, and understand patient experiences. Indeed, RWE

is of potential benefit in the field of osteoporosis, but it is imperative that the advantages and disadvantages of RWE are understood, and that the best practices for study conduct and reporting are followed to ensure high quality and transparent evidence is generated.

A European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) Working Group was convened in December 2022 to discuss the applicability of RWE to osteoporosis research and approaches that ensure the highest quality evidence is generated and reported for potential use in future drug regulatory applications and official guidelines. This group comprised 29 experts from across Europe, North America, the Middle East and Asia, with expertise including osteoporosis, rheumatology, gynaecology, epidemiology, pharmacoepidemiology, RWE, pharmacoconomics and regulatory affairs. This narrative review summarises the working group's agreed recommendations for the conduct and reporting of RWE studies with a focus on osteoporosis research. However, the overarching principles discussed here would also apply to other research fields.

What is real-world evidence?

RWE is derived from the analysis of real-world data (RWD). This is routinely collected data that relates to a patient's health or to the delivery of healthcare. Although the principles apply to broader epidemiological methods, the focus of RWD/RWE is predominantly on information derived from patients rather than free-living populations. The exposure of interest must be present in the dataset, which generally precludes the study of unlicensed medications in this context.

RWD can be derived from a variety of different sources, including, but not limited to, patient registries (including specific disease, drug, or medical device registries), prescription and dispensing data, insurance claim databases, health records (including retrospective chart reviews and the use of electronic health records) and patient-reported outcomes. Examples of patient-reported outcomes include wearables and biosensors (e.g. a smart watch measuring activity levels or sleep time, continuous glucose monitors or home pulse oximetry) and smart-phone application ("app")-based self-reporting, such as the ArthritisPower® app. The latter is a multipurpose digital platform enabling patients to track their disease and engage in educational materials, whilst additionally creating a disease registry and enabling RWD collection [13]. Some examples of other data sources that have been used and/or may be useful for future work in osteoporosis are shown in Table 1 [14–22]. A database of patient registries that could be utilised in RWE is maintained by The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) [23].

Table 1 Examples of real-world data sources from around the world

Real-world data type	Name	Country	Population	Strengths and limitations
Electronic health records	Clinical Practice Research Datalink (CPRD)	United Kingdom (UK)	Representative population of 2000 primary care practices in the UK	<ul style="list-style-type: none"> • Fracture recording has been validated [14] • Representative of the population [15] • Fracture recording has been validated [16]
	Sistema d'Informació per al Desenvolupament de l'Investigació en Atenció Primària (SIDIAF)	Catalonia, Spain	Computerised primary care records on 5.8 million patients from 279 primary care practices in Catalonia (representing 80% of the population)	<ul style="list-style-type: none"> • Nationwide coverage • Can be linked to other health and social registries through unique patient identifier (Civil Personal Register (CPR)) (e.g. hospitalisation data through National Registry of Patients)
Prescription dispensing records	Danish National Database of Reimbursed Prescription (DNDRP)	Denmark	Data on all reimbursed prescriptions redeemed at community pharmacies and hospital-based outpatient pharmacies since 2004, covering the entire Danish population [17]	<ul style="list-style-type: none"> • Adherence to medication is unknown • Captures more than 96% of individuals aged > 65 years • Limited to a single payer government insurance plan, may not be representative of the US population [18]
Insurance and administration database	Medicare Database	United States of America (USA)	Claims data for federal health insurance programme for those aged > 65 years, certain people < 65 years with disabilities, and people of any age with end-stage renal disease [18]	<ul style="list-style-type: none"> • Does not cover local clinic usage • May be biased towards those with underlying health conditions
	Medical Data Vision	Japan	Electronic health insurance claims from acute-care hospitals covering one-eighth of the population of Japan [19]	<ul style="list-style-type: none"> • Does not include primary care encounters • Automated data quality checking • Contains patient-identifiable data through unique CPR number allowing linkage to other databases
Clinical registry	Danish National Patient Registry	Denmark	Nationwide registry of all hospital contacts with primary and secondary diagnoses [20]	<ul style="list-style-type: none"> • Currently collects data from hospital-based service for local audit of care • Linkage to other RWD sources could be performed using national unique identifier (NHS number)
	The National Hip Fracture Database	UK	All patients admitted to hospital with a hip or femoral fracture in England, Wales and Northern Ireland since 2007 [21]	<ul style="list-style-type: none"> • Accuracy of self-reported data unknown • Patient-reported experiences
	Healthy Bones, Build Them for Life® Patient Registry (bonehealthandosteoporosis.org)	Worldwide	Anyone with osteoporosis or osteopenia can participate in the registry	

Table 1 (continued)

Real-world data type	Name	Country	Population	Strengths and limitations
Adverse event reporting	FDA Adverse Event Reporting System (FAERS)	USA	Voluntary reporting by healthcare professionals, consumers, and manufacturers of adverse events and medication errors [22]	<ul style="list-style-type: none"> • Not all events will be reported; some may be duplicate reported • Causality of adverse event to medication/device unknown
Patient-reported outcomes	ArthritisPower®	USA	Smartphone app reporting of symptoms	<ul style="list-style-type: none"> • Self-registration to the app/research study, therefore may not be a representative population • Reliance on self-reporting • Consent to link to health records

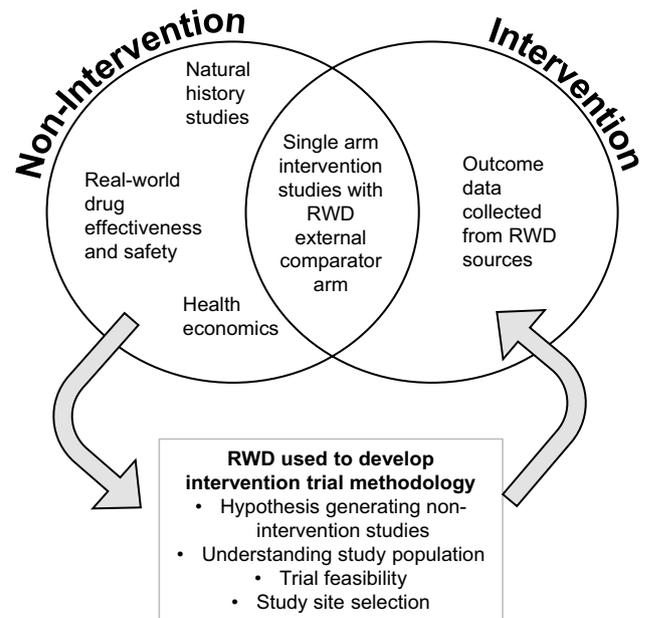


Fig. 1 Uses of real-world data in intervention and non-intervention (observational) studies

Data may be collected as part of routine clinical care and/or specifically for a research study, and can be quantitative or qualitative. Data from multiple sources can be linked using unique patient identifiers; Denmark and Sweden, for example, have excellent nationwide linked health and social data systems containing information on primary and secondary care and pharmacy dispensing that have enabled large nationwide population-based observational studies [24]. It is important that country-specific policies and laws with regard to data use and management are followed. Where required, consent to link data from electronic health records, laboratory data and claims data should be sought when recruiting patients to registries or self-reporting tools to optimise potential use in RWE studies and enable confirmation of any self-reported diagnoses to increase veracity in study findings.

RWE can be generated using many different study designs including non-intervention (observational) and intervention studies (Fig. 1). The broad study designs are not fundamentally different from that used for epidemiological observational surveys or clinical trials. As with all observational studies, the medical treatments received by the patient will have been decided by the healthcare professionals and patient subject to local policies and not by random allocation. Observational RWE however capitalises on the more detailed data that are now collected electronically and routinely, but in contrast to prospective observational studies in which data are collected on all participants at set time points, the timing of RWD (for example from diagnosis or treatment initiation) may vary between individual patients.

RWD can be incorporated into intervention studies in several ways, including in planning, as specific outcomes, or as an external comparator trial (Fig. 1). In a traditional RCT, RWD can be used in the planning of the study even if RWD are not collected with the purpose of analysis as trial outcomes. For example, RWD might be used in hypothesis-generating studies that are subsequently tested in an RCT, or RWD can be used to assess enrolment criteria, trial feasibility and support the selection of study sites by examining the impact of the proposed study's inclusion and exclusion criteria within the potential trial population. Alternatively, RWD on selected health outcomes and/or adverse events from sources such as electronic healthcare records, claims or administration databases may be integrated into an intervention study for pre-specified clinical outcomes. Finally, an external comparator group derived from RWD may be used to improve interpretability of a non-randomised single group intervention study. This would involve using a study design that overlaps between the traditional intervention and observational study. Data for the external comparator arm can be collected from either historical RWD or prospectively from a contemporaneous cohort of untreated patients. Consideration towards changes in clinical practice and standard care should be made when using historical RWD as the comparator. The possibility of differences in the trial and comparator group may exist due to confounding by indication if a contemporaneous group is used. This type of study is more appropriate when an RCT is unfeasible (e.g. due to low disease prevalence) or unethical, and is less likely to be used in post-menopausal osteoporosis. To date, RWE external comparator arm studies have most commonly been used in oncology and for rare diseases in children [25].

RWD has a number of advantages: representativeness of the true at-risk/treated population, often readily available data at relatively low cost allowing for more rapid analysis, and large study populations providing increased power to detect rare events. However, importantly, RWD may have inherent biases that should be considered when drawing causal inferences. Careful study design and analysis is needed to mitigate potential bias and ensure that high quality evidence is generated that could be accepted for regulatory approval.

RWE in osteoporosis research and practice

RWE approaches have been used widely across many fields. Here, we describe examples of how RWE has already been employed, and may be used in the future, to improve patient care, with examples from the field of osteoporosis.

Characterising the natural history of osteoporosis and related conditions

Natural history studies are epidemiological studies focusing on describing the prevalence, risk factors, clinical features and outcomes, burden and evolution of a disease. For example, studies using electronic health records, such as the Clinical Practice Research Datalink (CPRD) in the United Kingdom (UK), and insurance claims records have previously described the epidemiology of fracture [19, 26, 27] and trends in osteoporosis medication prescribing [28], adherence and discontinuation [29, 30].

With increasing availability and complexity of routinely collected electronic data, future studies might further characterise the at-risk population and be used to contextualise trial data within the target population to demonstrate transferability in clinical practice. Retrospective RWD can be used to identify risk factors, clinical outcomes or biomarkers occurring early in the evolution of the disease to enable earlier intervention and identify sub-populations who may derive the greatest benefit from earlier or new treatments, and in doing so, generate hypotheses for future intervention studies. RWD studies can be used to understand treatment adherence, associated determinants and clinical outcomes [31] and from this, help to develop pathways to promote adherence and optimise the patient care pathway.

Supporting regulatory decision-making for drugs and medical devices

RWE is playing an increasing role in informing regulatory decisions related to drugs and healthcare devices. This has in part been driven by the 21st Century Cures Act (2016) in the USA, which mandated that a programme be developed by the Food and Drug Administration (FDA) to evaluate the potential use of RWE. This multifaceted project hoped to expedite the process of drug authorization, particularly for new indications of previously approved drugs and biologics, and to support post-marketing surveillance of drug safety and efficacy. As a result, FDA first published their guidance "Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices" [32] in 2017 followed by a draft framework in 2018 [33], which outlined a framework for the use of RWE as valid scientific evidence to support FDA decisions. Similar frameworks have since been published by other regulatory bodies including the European Medicines Agency (EMA) [34] and the Pharmaceuticals and Medical Devices Agency in Japan [35]. The Therapeutic Goods Administration in Australia has also begun to review their approach to RWE in regulatory decisions [36].

Recent data have demonstrated that the integration of RWE in marketing authorization applications (MAAs) has become more common [37]. RWE was used to

support 39.9% of initial MAAs and 18.3% of extension of indication (EOI) applications submitted in 2018 and 2019 to the EMA [38]. The RWD used in these applications were mostly derived from disease and product registries, but electronic health records, claims, drug dispensing data and compassionate use programmes were also used as data sources. For most applications for both marketing authorization and EOI, RWE was used to support drug safety rather than drug effectiveness [38]. In further analysis, of the applications for which RWE was submitted to support efficacy, the RWE contributed to decision-making in five out of sixteen MAA and five out of ten EOI [39]. Overall RWE contributed to the pre-authorization decision-making on drug efficacy for approximately 3% of applications to the EMA during the 2-year period [39].

Although the use of RWE in MAA may be limited by the need for the drug to be authorised to enable RWD collection, there is a stronger role for its use in expanding licenced indications, for example across geographical regions, by patient demographics (e.g. sex, ethnicity), or for related clinical indications (e.g. postmenopausal compared with glucocorticoid induced osteoporosis). Within the field of osteoporosis, RWE was recently used to support a regional MAA leading to the approval of Prolia® (denosumab) for the treatment of post-menopausal osteoporosis by the National Medical Products Administration (NMPA) in the People's Republic of China in 2020. The MAA included data from the global clinical trial programme for Prolia® which established the efficacy and safety of this drug [40]. A RWE study from Taiwan and Hong Kong that demonstrated a favourable benefit: risk profile in ethnic Chinese women was employed to demonstrate that safety and efficacy data are comparable for a Chinese patient group to that observed in the RCT population. This was required to meet specific NMPA requirements. In this study, a population level claims database (Health Insurance Research Database) covering 99.9% of the population of Taiwan and a population-level clinical database (Clinical Data Analysis and Reporting System) which includes 80% of hospital admissions in Hong Kong were used. Fracture risk was compared in two groups of women, a treatment cohort who had Prolia® 60 mg subcutaneously every 6 months for up to 10 doses and a “control” group who discontinued Prolia® after a single dose. The relative risk of fracture reductions for the Prolia®-treated cohort compared to the control group was similar to that in the global RCT [40, 41]. Safety data were also collected using the same databases focusing on the incidence of hypocalcemia, atypical femoral fracture (AFF) and osteonecrosis of the jaw in women treated with Prolia® [42]. A similar approach supported the authorization of Eladynos (abaloparotide) by the EMA in October 2022 for the treatment of

osteoporosis in postmenopausal women at increased risk of fracture [43, 44].

Post-marketing safety monitoring, long-term effectiveness and comparator studies

Post-marketing drug surveillance has a vital role in monitoring drug safety over longer periods of follow-up than is undertaken in RCTs. A recent population-based cohort study in Ontario, Canada, utilised multiple linked registries of demographic information (The Registered Persons Database of Ontario), prescription dispensing records (Ontario Drug Benefits Program), comorbidities (Canadian Institute for Health Information's Discharge Abstract Database) and diagnostic and procedural information from hospital and emergency department visits (the National Ambulatory Care Reporting System Database) to illustrate that real-world patients prescribed bisphosphonates and denosumab are older, include a greater proportion of males and have higher prevalence of chronic kidney disease than the participants who took part in the RCTs that demonstrated the efficacy and safety of these medications [45]. This highlights the need to understand the effectiveness and side effects in the population who receive treatment. One such example is the work by Robinson et al. who assessed the safety of oral bisphosphonates in patients with moderate-severe chronic kidney disease [46]. The authors used routinely collected population-representative data from linked primary and secondary care records from the UK (CPRD and Hospital Episode Statistics (HES)) and Catalonia, Spain (SIDIAP, National Hospital Discharge Database and renal registry), in a case-control study including nearly 4000 new users of bisphosphonates and propensity score matching to a control group to reduce the risk of confounding. Bisphosphonate use was associated with an increased risk of progression of moderate to severe chronic kidney disease [46], thus providing important novel safety information in a group of patients who are commonly treated with bisphosphonates [45]. Similarly, RWE can also demonstrate clinical effectiveness in populations in which medications are actually prescribed, such as the work of O'Kelly et al. which showed that continued treatment with antiosteoporosis medications was associated with reductions in fracture rates using data for women aged over 50 years prescribed one or more antiosteoporosis medications in an anonymized German healthcare claims database representative of the German population [47].

RWE additionally has the potential to deepen our understanding of recognised adverse events and identify individuals who are at greater risk through stratification by demographic factors or comorbidity. Information on rare complications of antiosteoporosis drugs, including AFF

and osteonecrosis of the jaw, can be ascertained in more detail using RWD than from an RCT due to insufficient power in the clinical trials to capture detailed information on these rare adverse events. For example, data from the FDA Adverse Event Reporting System (FAERS), a voluntary healthcare professional submitted reporting system for adverse effects of drugs and medical devices, has been used to assess the risk of osteonecrosis of the jaw in patients on antiresorptive medications [48]. Multiple RWE studies were also used to support the removal of the FDA “box warning” from teriparatide. The initial concerns of increased risk of osteosarcoma in patients treated with teriparatide were based on pre-clinical animal studies, but RWE generated following authorization showed no increased risk in humans [49]. The supporting RWE studies utilised data from cancer patient registries [50], a voluntary patient drug registry [51] and Medicare and insurance claims databases [52].

RCTs typically compare a new drug treatment to either placebo or the currently accepted gold-standard treatment. Few trials are performed as comparator studies to identify best available treatment options. RWE can be used to address this evidence gap and identify the need for head-to-head trials, although care should be taken when comparing two treatment options that were available in different eras of clinical care provision. Ideally, comparisons should be made between treatments available at the same time for the same indication. In the field of osteoporosis, Cosman et al. compared the real-world effectiveness of abaloparatide and teriparatide on nonvertebral fracture incidence and cardiovascular outcomes following 18 months of treatment in anabolic-therapy-naïve women aged over 50 years. This was performed using retrospective data derived from an anonymised claims database in the USA with propensity score matching of 11,616 women treated with abaloparatide to a group treated with teriparatide. Non-inferiority of abaloparatide versus teriparatide on time to first non-vertebral fracture was demonstrated [44]. In contrast, Khalid et al. demonstrated that users of selective estrogen receptor modulators (SERMs) had lower risk of primary hip and major osteoporotic fracture compared to propensity-matched users of alendronate in CPRD and SIDIAP. Of note, the study cohort were at low risk of fracture, and the authors acknowledge that imbalances in unobserved confounders remain a possibility despite propensity matching [79]. In the same study, fracture risk in users of strontium ranelate and other oral bisphosphonates compared to users of alendronate replicated the findings of head-to-head RCTs of these drugs. These findings could therefore be used as the rationale for further comparator RCTs to optimise care pathways. Advanced approaches to management of RWD including use machine learning could improve fracture identification and drug comparison studies.

Health economics

RWE also has the potential to increase the quality and reliability of health economic evaluations. Cost-effectiveness analyses are assuming significant importance in policy decision-making and therefore are increasingly being conducted to assess the economic value of interventions such as antiosteoporosis medications [53] or fracture liaison services in the field of osteoporosis [54]. This helps facilitate the appropriate use of limited healthcare resources. Economic evaluations typically use models to characterise the natural history of disease and the effects of interventions, by combining epidemiological data, economic information (costs and quality of life) and efficacy/effectiveness data. By definition, cost-effectiveness analyses evaluate the effectiveness (the effect in a routine clinical care setting) of health interventions and not their efficacy (the effect in ideal trial conditions). However, most economic studies have used efficacy data derived from trials or meta-analyses. The incorporation of medication adherence has been recommended to better assess the real-life cost-effectiveness of interventions in clinical practice and recent studies have increasingly explored this [55, 56]. RWE may provide good estimates of treatment adherence, fracture probabilities, costs, quality of life and drug effectiveness, and is therefore crucial to design appropriate economic models that better reflect real-life settings. To date, however, very few real-world effectiveness studies have been used in economic evaluation and this represents an important area for further research.

Challenges and limitations of RWE

Similar to other clinical research, RWE studies need to adhere to principles for high quality evidence generation: transparency, data suitability to answer a specific research question and appropriate analysis to both minimise bias and characterise uncertainty. Recognising the challenges and limitations of using RWD is important to researchers conducting RWE studies and those appraising their findings.

Data quality

Using data from real-world sources can allow large quantities of data to be rapidly acquired but RWD usually do not have the quality assurance of data collected within a clinical trial or prospective observational study. In a trial setting, laboratory measurements are typically collected by a small number of trained fieldworkers following standard operating procedures with cross-calibration of instruments

allowing accuracy and precision. Clinical data are often collected through validated questionnaires or through completion of case report forms by the investigator though the provenance of the data used in the case report forms is not always described in detail. In contrast, the level of training, methods and instruments used for RWD are often unknown and will vary between healthcare providers leading to more heterogenous data. Blood pressure, for example, can be measured manually or electronically. Even computer-generated data, such as from dual-energy X-ray absorptiometry (DXA), is subject to measurement variation between instruments [57]. The level of coding of clinical data in electronic health records may vary between clinicians and/or could be related to reimbursement practices.

Covariate data are also limited by availability within the database used. This should be reviewed when considering database suitability to address a research question.

An ongoing FDA-funded project RCT-DUPLICATE is aiming to establish if non-randomised study approaches using healthcare databases can consistently match the results of published clinical trials and predict the results of ongoing trials to provide some confidence in the validity of RWE in the absence of RCT evidence [61] and work to understand and improve data quality using advanced RWE techniques is also in progress [62].

Risk of bias

Bias that threatens the internal validity of study can occur for several reasons. These are described below.

Recall and misclassification bias

Data derived from patient/caregiver questionnaires or interviews could be subject to recall bias when respondents selectively or inaccurately report historical events. The use of objectively defined and documented exposures and outcomes, for example prescriptions, fracture or death, may be preferable to subjective, patient-recalled and/or heterogeneously measured outcomes, but coding errors in drugs and diagnoses can result in misclassification bias. Furthermore, diagnostic codes, such as International Classification of Diseases (ICD) codes, have evolved over time, with more specific codes used in each edition, and coding of complications such as AFF and osteonecrosis of the jaw only recently being used. Even with an objective outcome, validation of the definition used and how this is derived from the data are required to prevent misclassification. When considering fracture as an outcome, several validation studies have been published for deriving fracture from diagnostic codes used in some of the more commonly used databases including Medicare [58], other US administrative claims databases [59], CPRD [14] and SIDIAP [16]. Konstantelos et al. recently

undertook a scoping review of how fracture was defined in studies of osteoporosis drugs using claims data from the USA or Canada [60]. Half of the 57 studies reviewed did not provide a citation for their fracture definition and half did not indicate specific data sources for the codes used for their outcome definitions. There was also marked variation in the definitions used. For example, amongst the 29 studies with a definition for hip fracture reported, twelve different definitions were used. Similarly, nine different definitions were used for vertebral fracture in fifteen studies [60]. Moreover, when considering hip fracture, some definitions included an inpatient diagnosis and procedural code, which would miss patients that are deemed unsuitable for or die before surgical intervention, whereas others include an inpatient or emergency department diagnosis, which may capture a larger number of cases. These differences in definitions limit comparisons between studies.

Missing data (information bias)

Missing data pose a considerable challenge in the analysis of RWD. Thought needs to be directed towards the cause of missingness, which can be classified as [63]:

- Missing completely at random (MCAR)—no systematic differences between those with and without values.
- Missing at random (MAR)—systematic differences between the missing values can be explained by differences in observed data.
- Missing not at random (MNAR)—systematic differences remain after observed data accounted for.

For example, missing data on a sit-to-stand test might occur because the test was not performed by individual clinicians or at certain centres (MCAR), because it is less commonly performed in younger patients (MAR) or because the patient was unable to participate in the test due to functional limitations (MNAR).

Complete case analysis, in which only those without missing data are included, can be performed, but can lead to selection bias, a lack of generalizability and a smaller sample size and study power. When data is MAR, this can be overcome using multiple imputation. This statistical approach allows individuals with incomplete data to be included in analyses. In brief, multiple copies of the dataset are created, with the missing values replaced by imputed values generated from predictive distribution based on the available non-missing data. Standard statistical methods are subsequently used to fit the model of interest to each of the imputed datasets, and the estimations, which will differ due to the variation introduced by the imputed missing values, averaged together to give an overall estimated association [63]. It is recommended

that specialist statistical knowledge is sought as multiple imputation is not appropriate for all missing data including data that is MNAR, where imputation will introduce further biases [63].

Healthy complier bias and confounding by indication

Confounding occurs when there are common causes for both the choice of intervention and outcome. This can be due to healthy-complier bias, in which healthier individuals are more likely to seek out and adhere to a treatment and may be at lower risk of an outcome. Conversely, confounding by indication (sometimes referred to as “channeling bias”) can occur when the decision to commence a treatment is influenced by both clinician and patient factors, such as disease severity, comorbidity, and expected outcomes and risks [64]. This results in an imbalance in the underlying risk profile between those who did and did not receive the treatment. This is particularly observed when treatment indications are narrow or restricted. For example, newer treatments might be limited to those with more severe disease. A good example, from the UK, would be teriparatide, which under guidance from NICE is permitted for use only in those who have a fracture on existing treatment and have low BMD [65]. If the risk profile is also an independent predictor of the clinical outcome being assessed, this leads to confounding by indication (Fig. 2). This can appear to strengthen, weaken or reverse a true effect, and will be particularly relevant to real-world drug

effectiveness comparison studies, where a treatment that is limited to more severe disease may appear to have a poorer outcome [66]. Confounding by indication can also be intertwined with time-lag bias, which occurs when two groups are compared without consideration of the underlying disease duration and how this might affect outcome. This most commonly occurs in studies comparing first- and subsequent-line drugs.

Propensity scoring can be used to control for confounding by indication and has been used in several of the previously mentioned studies that compared antiosteoporosis drug effectiveness and safety [44, 46]. In this approach, multivariable logistic regression models are used to estimate the probability of an individual being prescribed the treatment. Individuals with the same propensity score are considered to, on average, have the same likelihood of receiving a treatment. These scores can be used for case matching and/or in stratified analyses [64]. Presentation of cohort characteristics before and after propensity score matching should be reported to understand the treated population within the overall at-risk population. An important limitation is that propensity matching is restricted only to the measured matching variables and does not always overcome confounding.

Residual confounding

Residual confounding can still occur despite attempts to control for known covariates, including when propensity

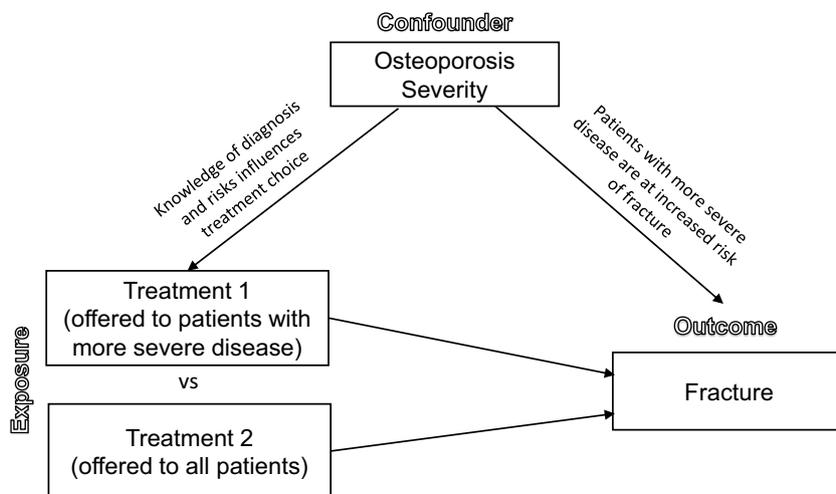


Fig. 2 Confounding by indication. An example in which osteoporosis severity influences treatment choice and fracture risk. In observational work, knowledge of underlying disease severity or other risk factors might influence treatment choices made by the clinician and/or patient, with a potentially more powerful intervention used in those with more severe disease/higher risk. These same decision influencing factors might also influence the likelihood of the outcome being studied resulting in confounding by indication. For example, in the

UK, teriparatide is only available for women with severe disease, who are by the same definitions used also at increased risk of fracture. Comparing fracture rate in women treated with teriparatide to those treated with other antiosteoporosis medications, would likely suggest that teriparatide had a higher fracture rate unless approaches to controlling confounding by indication, such as propensity matching, are used

scoring is used, either because covariates are unknown or not measured. Work by Robinson et al. has demonstrated the importance of consideration for residual confounding when drawing conclusions. Using data from CPRD, bisphosphonate use was associated with increased risk of hip, non-hip and osteoporotic fracture in patients with chronic kidney disease. However, when time to fracture was restricted to 180 days, a period in which the effect of bisphosphonates would not be expected to be apparent, the hazard ratios for fracture remained elevated, highlighting that the excess risk was likely not attributable to the bisphosphonates but may represent residual confounding [67].

Immortal time bias and negative time windows

Immortal time is when participants/study cases cannot experience an outcome during a period of the follow-up time. This can occur when study cases are allocated to a treatment group, but there is a delay in collecting a prescription. By design, participants allocated to the treatment group could not have died or fulfilled an outcome between the time of entering the cohort and the time of collecting the prescription, and as such this immortal time contributes to the treated group and leads to an underestimation of outcome events in that group. Careful study design to ensure exposure and data collection times are aligned or a time-varying exposure approach is used to prevent immortal time bias.

A negative time window is the period in which the effect of an intervention would not be expected to alter the outcome. Knowledge of the pharmacology of a drug is required to establish an appropriate window; for example, the effect of a bisphosphonate on fracture risk would not be expected immediately and thus any fracture within the first few months of treatment are not likely related to the bisphosphonate effects, whereas a beta-blocker would be expected to have a rapid effect on blood pressure. In the previously discussed study by O’Kelly et al. that demonstrated ongoing effectiveness of osteoporosis medications in clinical practice, fracture incidence in the early period (0–3 months) after treatment initiation was used as the control period, since an effect of the medications would not be expected in this period. This also allowed for the followed cases to act as their own control group [47].

Immeasurable time bias

Immeasurable time bias can occur in pharmacoepidemiologic studies when only primary care or community prescription dispensing records are used and drug dispensing during periods of hospitalisation is not considered [68]. Incorrect classification of hospitalised patients, who are typically at greater risk of death and other adverse events, as unexposed when hospital pharmacy records are unavailable

can overestimate drug benefits, and conversely, an apparent reduction in prescription dispensing during prolonged admissions or inaccurate knowledge as to when the drug was commenced during an inpatient stay may result in inaccurate follow-up periods. Adjustment for hospitalisation as a time-varying variable can be used to overcome this bias [69].

Managing “Big Data”

Use of real-world databases has the potential to generate substantial amounts of data in a relatively short timeframe. Ongoing projects, such as the OneSource Project that aims to automate the flow of electronic health records into external systems for RWE study usage [70] and the Data Analysis and Real World Interrogation Network (DARWIN EU®), will further facilitate ease of data access. “Big Data” has been defined as “extremely large datasets which may be complex, multi-dimensional, unstructured, heterogenous which are accumulating rapidly, and which may be analysed computationally to reveal patterns, trends, and associations” [71]. Increasingly multiple datasets are linked together to answer a single research question leading to more complex data. It is important that researchers using these databases understand how and where the data were generated and curated from the clinical encounter to the data as presented to them. This will require a close working relationship with those involved in the data curation and advanced approaches to analysis. Some data will need substantial preparation prior to use, for example extracting usable data from narrative clinical notes. This is likely to require the input of data scientists and technological approaches. Machine learning is being increasingly used in the analysis of complex datasets, and for example could be used to identify symptoms suggestive of fractures (such as acute back pain), but these approaches often lack transparency and reproducibility, and indeed the same algorithm can be run multiple times giving different answers [72].

Achieving best practice in the conduct and reporting of RWE

A narrative synthesis of the feedback from the Committee of Medicinal Products for Human Use (CHMP) appraisal on RWE to support claims in MAA and EOI to the EMA highlighted that the main issues with regards to the presented RWE included lack of pre-specified analysis plans, risk of confounding and selection bias, small sample size, missing data and lack of population representativeness [39]. Similar findings were also observed in FDA applications [37]. Drawing on these and the limitations and challenges of RWD discussed above, we recommend several principles for the conduct and reporting of RWE studies

with a focus on achieving transparency. These are summarised in Fig. 3.

Transparency in design and conduct: study registration and protocol publication

The planning of an RWE study should begin with a clearly defined research question or hypothesis based on a biologically plausible rationale. The “PICO” structure can be helpful in defining this: Population (e.g. demographics, disease), Intervention or variable of Interest (e.g. an intervention, exposure to a disease/variable, risk factor), Comparison (e.g. placebo, standard care, absence of risk factor), Outcome (e.g. risk of disease, event). The outcomes considered should be measured and reliable, and this may influence the choice of dataset.

Since 2005, it has been a requirement that clinical trials are registered prior to patient enrolment in a publicly available trial registry [73]. This approach was adopted to increase the transparency of trial conduct and reporting,

increase replicability of studies and improve study quality. Subsequently, research registries have expanded to include non-intervention studies and systematic reviews and meta-analyses, and registration of RWD studies should also be encouraged. The European Union Electronic Register of Post-Authorisation Studies (EU PAS Register (encepp.eu)) is a publicly available register specifically for non-intervention post-authorisation studies. RWE studies can also be registered in the clinicaltrials.gov database and the Open Science Framework (osf.io). Study protocols for RWE studies should include methods for data curation and analysis plans. The latter may include definitions for exposures and outcomes, inclusion and exclusion criteria, management of missing data, the approach to managing confounders, minimising bias and propensity scoring, and how the data will be analysed. Pre-specification of subgroup analysis with justification for the approach is also important. These practices aim to prevent “data dredging”, i.e. result-driven selection of study parameters and selective reporting towards positive

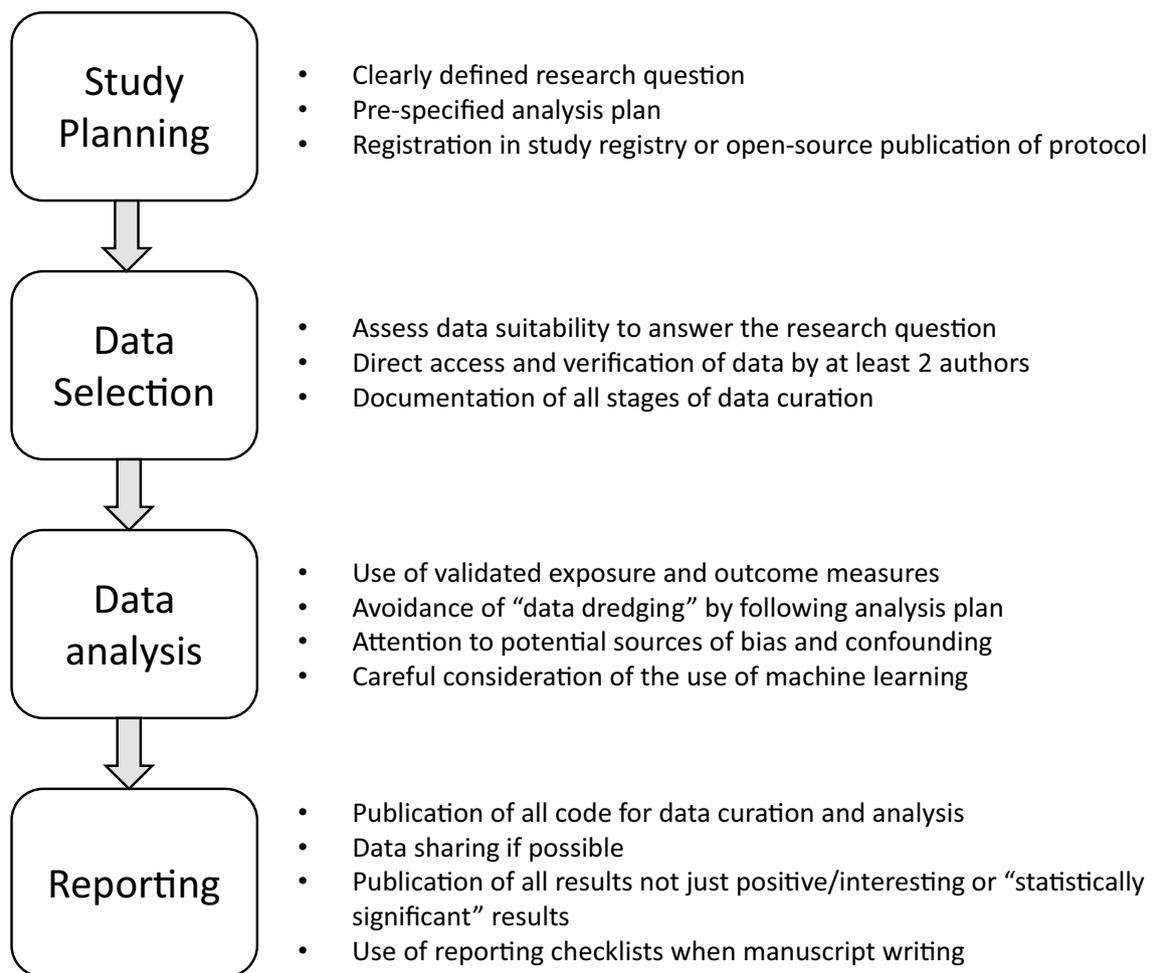


Fig. 3 Considerations in the planning, conduct and reporting of RWE studies in osteoporosis

or interesting findings, and thus increase confidence in published findings.

Transparency in data usage and analysis

Local data protection laws should, of course, be followed when accessing stored data. These vary widely across the world [74]. Consideration should also be given to protocol review by a research ethics review board to ensure that data usage complies with local ethical standards for healthcare research.

Considerations need to be given to the appropriateness of the data source to answer the specific research question, and an assessment of the completeness and accuracy of key study variables should be undertaken before proceeding to more detailed analysis. Structured frameworks for assessing the suitability of data sources to answer a research question are available [75].

Trustworthiness of a data source is increased if researchers using RWD know the origin of the data and how it has been transformed. Reporting of the statistical code used for data transformation, and cleaning and analysis in an open-source format should be encouraged to enable reproducibility of statistical methods in another dataset. A recent high-profile paper reporting a RWE study on hydroxychloroquine treatment in COVID-19 infection was retracted from publication after the authors were refused access to the dataset to evaluate the origination and completeness of the database and to replicate the analyses presented in the paper following concerns of data veracity [76]. Whilst considerations of confidentiality are important, and local data governance laws need to be followed, sharing of analytic code and data would have allowed for replication and confidence in the findings. As a result, *Lancet* journals have made modifications to the signed declarations by authors and peer review process for

manuscripts using real-world datasets. The author statement form will require that more than one author has directly accessed and verified the data reported in the manuscript, and that these authors are named in the contributors' statement [77].

Transparency in reporting

The minimum reporting requirements for studies using RWD data are outlined in the REporting of studies Conducted using Observational Routinely-collected Data (RECORD) and the REporting of studies Conducted using Observational Routinely-collected Data for non-interventional PharmacoEpidemiological research (RECORD-PE) checklists [78, 79]. This consistency in reporting ensures that key questions, such as why and how the research was conducted and whether the results reflect the prespecified research questions, are easily answered, allow for a clear assessment of study quality and validity, and facilitate the replication of methods and results.

Future research using real-world data in osteoporosis

Real-world data, when used in well-designed research studies, has the potential to answer many novel research questions in the field of osteoporosis. RWE will not replace RCTs. The latter are still required to demonstrate safety and efficacy of new treatments. RWE also cannot easily be used to study drugs that are not yet licenced as exposure data are not available. However, RWE could be used to expand osteoporosis drug horizons to specific patient groups, to establish new care pathways (for example by comparing long-term

Table 2 Examples of potential future uses of real-world evidence in osteoporosis research

Broad remit	Examples of potential future research
Natural history studies	<ul style="list-style-type: none"> • Establishing osteoporosis/fracture risk in specific sub-populations (e.g. Parkinson's disease, dementia) • Understanding treatment adherence to different types of osteoporosis medications (e.g. oral, intravenous, subcutaneous) • Establishing childhood and early-life risk factors for adult osteoporosis (e.g. obesity, childhood fracture, medication use) • Influence of external factors/phenomenon (e.g. COVID-19 pandemic) on fracture risk and/or prescribing habits
Post-marketing safety monitoring	<ul style="list-style-type: none"> • Assessment of potential rare side effects of medications • Establishing risk factors for atypical femoral fractures
Long-term effectiveness	<ul style="list-style-type: none"> • Effectiveness of osteoporosis medications in men, including comparison of different therapeutic agents • Elucidation of effects in specific co-morbidities (e.g. cardiovascular disease, chronic obstructive pulmonary disease, dementia, obesity) • Effectiveness of therapeutic interventions across ethnic groups
Health economics	<ul style="list-style-type: none"> • Integration of real-world effectiveness data into health economic analyses

drug effectiveness) and to assess potential rare side effects and expand safety data of already licenced drugs. Examples of potential uses of RWE in osteoporosis are given in Table 2.

Conclusions

The use of real-world data in research and decision-making is growing, and without doubt can be used to deepen our understanding of osteoporosis and fracture epidemiology, to understand the use, effectiveness and safety of interventions in clinical practice, and could potentially accelerate the approval of new medications. However, it is vital that this research is conducted to the highest standard with close attention to the limitations and biases of routinely collected observational data. As these data become increasingly complex, transparency is required at all stages of study design, data acquisition and curation, analysis and reporting to increase the trustworthiness of RWE study findings and increase its incorporation into regulatory and reimbursement decision-making and clinical practice guidelines.

Funding The ESCEO Working Group was funded by the ESCEO. The ESCEO receives unrestricted educational grants to support its educational and scientific activities from non-governmental organisations, not-for-profit organisations, non-commercial or corporate partners. The choice of topics, participants, content and agenda of the Working Groups as well as the writing, editing, submission and reviewing of the manuscript are the sole responsibility of the ESCEO, without any influence from third parties. PGC is supported in part by the UK National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care or the World Health Organisation.

Declarations

Ethics approval and consent to participate This narrative article contains no original data and thus issues of ethics, informed consent and patient confidentiality do not apply.

Conflict of interest RJM has received travel bursaries from Kyowa Kirin. OB has received consulting or lecture fees from Amgen, Aptissen, Biophytis, IBSA, Mylan, Novartis, Orifarm, Sanofi, UCB and Viatrix. MC has received honoraria for speaking and chairing engagements from AMGEN, DKSH and Kyowa Kirin. PGC has received speakers bureaus or consultancies from AbbVie, AstraZeneca, Bristol Myers Squibb, Eli Lilly, Galapagos, GlaxoSmithKline, Grunenthal, Merck, Novartis, Pfizer, Stryker and UCB, outside the submitted work. WHD is on the board of Directors for BioMarin, Seres, Mersana and Metagenomi, Chief Advisor to CEO at Angitia and is a former employee and shareholder at Eli Lilly and Amgen. MH has research grants paid to his institution from Amgen, Radius Health, and consulting fees from UCB. EVM reports research support and/or personal fees from Amgen, Eli Lilly, Fresenius Kabi, ObsEva and UCB, outside the submitted work. SO has received consultancy fees from Merck and Pfizer

not related to the current work or therapeutic field. RR has speaker bureau or advisory board for Abiogen, Naturex, Nestlé and ObsEva. YAS reports lecture honoraria from Amgen, Eli Lilly and SAJA. SLS has acted as a consultant for Radius and Amgen, received funding from Radius for clinical trial and received travel support from Radius and grant support from Amgen. LSS discloses associations with AstraZeneca, Pfizer, Rigel, Eupraxia, Biosplice, EMDSerono, Horizon, Direct, Lilly, Kaniska, Protalix, Chemomab, TLC, SpineThera, Kyoto, PPD, Galvani, Urica, Transcode, Boehringer Ingelheim, Bristol Myers Squibb, Priovant, Roivant, Ampio, Aura, Aurinia, GSK, Xalud, Neumentum, Neema, Amzell, Applied Bio, Aptinyx, Bexson, Bone Med, Bone Therapeutics, Cancer Prevention, Cerebral Therapeutics, Chemocentryx, Diffusion Bio, Elorac, Enalare, Foundry Therapeutics, Galapagos, Histogen, Gilead, Idera, Intravital, Ingel, Kiel Labs, Mesoblast, Mpathix, Minerva, Regenosine, Samus, Sana, StageBio, Theraly, Unity and Viridian. FT reports personal fees from Amgen, Eli Lilly, Fresenius, Stada and UCB outside the submitted work. CC reports personal fees from ABBH, Amgen, Eli Lilly, GSK, Medtronic, Merck, Novartis, Pfizer, Roche, Servier and Takeda, outside the submitted work. NCH reports personal fees, consultancy, lecture fees and honoraria from Alliance for Better Bone Health, AMGEN, MSD, Eli Lilly, Servier, Shire, Consilient Healthcare and Internis Pharma, outside the submitted work. JYR, NMA, JAT, CB, OB, NB, MCD, WHD, ADP, PH, PH, JAK, DPA, RPR, SLS, TVS and AL report no conflicts of interest related to this work.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial 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-nc/4.0/>.

References

1. Kanis JA, Norton N, Harvey NC, Jacobson T, Johansson H, Lorentzon M et al (2021) SCOPE 2021: a new scorecard for osteoporosis in Europe. *Arch Osteoporos* 16(1):82. <https://doi.org/10.1007/s11657-020-00871-9>
2. Odén A, McCloskey EV, Kanis JA, Harvey NC, Johansson H (2015) Burden of high fracture probability worldwide: secular increases 2010–2040. *Osteoporos Int* 26(9):2243–2248. <https://doi.org/10.1007/s00198-015-3154-6>
3. Söreskog E, Borgström F, Shepstone L, Clarke S, Cooper C, Harvey I et al (2020) Long-term cost-effectiveness of screening for fracture risk in a UK primary care setting: the SCOOP study. *Osteoporos Int* 31(8):1499–1506. <https://doi.org/10.1007/s00198-020-05372-6>
4. Chotiarnwong P, McCloskey EV, Harvey NC, Lorentzon M, Prieto-Alhambra D, Abrahamsen B et al (2022) Is it time to consider population screening for fracture risk in postmenopausal women? A position paper from the International Osteoporosis Foundation Epidemiology/Quality of Life Working Group. *Arch Osteoporos* 17(1):87. <https://doi.org/10.1007/s11657-022-01117-6>
5. Lai F, Pei L, Chen X, Li J (2021) Osteoporosis-related randomized clinical trials with middle-aged and older adults registered on the

- International Clinical Trials Registry Platform. *Front Endocrinol*. 12:702261. <https://doi.org/10.3389/fendo.2021.702261>
6. Reyes C, Pottegård A, Schwarz P, Javaid MK, Van Staa TP, Cooper C et al (2016) Real-life and RCT participants: alendronate users versus FITs' trial eligibility criterion. *Calcif Tissue Int* 99(3):243–249. <https://doi.org/10.1007/s00223-016-0141-7>
 7. Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA et al (1998) Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA* 280(24):2077–2082. <https://doi.org/10.1001/jama.280.24.2077>
 8. Smith Z, Botto E, Getz K (2022) Quantifying diversity and representation in pivotal trials leading to marketing authorization in Europe. *Ther Innovat Regulatory Sci* 56(5):795–804. <https://doi.org/10.1007/s43441-022-00421-0>
 9. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A (2007) Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *J Bone Miner Res* 22(3):465–475. <https://doi.org/10.1359/jbmr.061113>
 10. Vilaca T, Eastell R, Schini M (2022) Osteoporosis in men. *Lancet Diabetes Endocrinol* 10(4):273–283. [https://doi.org/10.1016/s2213-8587\(22\)00012-2](https://doi.org/10.1016/s2213-8587(22)00012-2)
 11. Nayak S, Greenspan SL (2017) Osteoporosis treatment efficacy for men: a systematic review and meta-analysis. *J Am Geriatr Soc* 65(3):490–495. <https://doi.org/10.1111/jgs.14668>
 12. Björnsdóttir S, Clarke BL, Mannstadt M, Langdahl BL (2022) Male osteoporosis-what are the causes, diagnostic challenges, and management. *Best Pract Res Clin Rheumatol*. 36(3):101766. <https://doi.org/10.1016/j.berh.2022.101766>
 13. Nowell WB, Curtis D, Thai M, Wiedmeyer C, Gavigan K, Venkatchalam S et al (2019) Digital interventions to build a patient registry for rheumatology research. *Rheum Dis Clin* 45(2):173–186. <https://doi.org/10.1016/j.rdc.2019.01.009>
 14. Van Staa TP, Abenhaim L, Cooper C, Zhang B, Leufkens HG (2000) The use of a large pharmacoepidemiological database to study exposure to oral corticosteroids and risk of fractures: validation of study population and results. *Pharmacoepidemiol Drug Saf* 9(5):359–366. [https://doi.org/10.1002/1099-1557\(200009/10\)9:5%3c359::aid-pds507%3e3.0.co;2-e](https://doi.org/10.1002/1099-1557(200009/10)9:5%3c359::aid-pds507%3e3.0.co;2-e)
 15. García-Gil Mdel M, Hermsilla E, Prieto-Alhambra D, Fina F, Rosell M, Ramos R et al (2011) Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). *Inform Prim Care*. 19(3):135–45. <https://doi.org/10.14236/jhi.v19i3.806>
 16. Martínez-Laguna D, Soria-Castro A, Carbonell-Abella C, Orozco-López P, Estrada-Laza P, Nogues X et al (2019) Validation of fragility fractures in primary care electronic medical records: a population-based study. *Reumatol Clin (Engl Ed)* 15(5):e1–e4. <https://doi.org/10.1016/j.reuma.2017.10.013>
 17. Johannesdóttir SA, Horváth-Puhó E, Ehrenstein V, Schmidt M, Pedersen L, Sørensen HT (2012) Existing data sources for clinical epidemiology: the Danish National Database of Reimbursed Prescriptions. *Clin Epidemiol* 4:303–313. <https://doi.org/10.2147/clep.S37587>
 18. Research Data Assistance Center (ResDAC): Getting started with CMS Data. <https://resdac.org/getting-started-cms-data#who-is-in-the-data> Accessed 15/12/2022.
 19. Fujiwara S, Buchanan-Hughes A, Ng A, Page J, Adachi K, Li H (2022) Real-world evaluation of osteoporotic fractures using the Japan Medical Data Vision database. *Osteoporos Int* 33(10):2205–2216. <https://doi.org/10.1007/s00198-022-06472-1>
 20. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT (2015) The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol* 7:449–490. <https://doi.org/10.2147/CLEP.S91125>
 21. Royal College of Physicians: National Hip Fracture Database (NHFD). <https://www.rcplondon.ac.uk/projects/national-hip-fracture-database-nhfd> Accessed 01/02/2023.
 22. US Food and Drug Administration: Questions and answers on FDA's Adverse Event Reporting System (FAERS). <https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers> Accessed 17/12/2022.
 23. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance: ENCePP Resources Database. <https://www.encepp.eu/encepp/resourcesDatabase.jsp> Accessed 15/12/2022.
 24. Vala CH, Lorentzon M, Sundh V, Johansson H, Lewerin C, Sten S et al (2020) Increased risk for hip fracture after death of a spouse-further support for bereavement frailty? *Osteoporos Int* 31(3):485–492. <https://doi.org/10.1007/s00198-019-05242-w>
 25. Rippin G, Ballarini N, Sanz H, Largent J, Quinten C, Pignatti F (2022) A review of causal inference for external comparator arm studies. *Drug Saf* 45(8):815–837. <https://doi.org/10.1007/s40264-022-01206-y>
 26. Moon RJ, Harvey NC, Curtis EM, de Vries F, van Staa T, Cooper C (2016) Ethnic and geographic variations in the epidemiology of childhood fractures in the United Kingdom. *Bone* 85:9–14. <https://doi.org/10.1016/j.bone.2016.01.015>
 27. Curtis EM, van der Velde R, Moon RJ, van den Bergh JP, Geusens P, de Vries F et al (2016) Epidemiology of fractures in the United Kingdom 1988–2012: variation with age, sex, geography, ethnicity and socioeconomic status. *Bone* 87:19–26. <https://doi.org/10.1016/j.bone.2016.03.006>
 28. Cortet B, Schott AM, Désaméricq G, Chauny JV, Samama P, Emery C et al (2022) Trends in postmenopausal osteoporosis treatment in France during the period 2007–2016: a nationwide claims database analysis. *Bone*. 154:116255. <https://doi.org/10.1016/j.bone.2021.116255>
 29. Reyes C, Tebe C, Martínez-Laguna D, Ali MS, Soria-Castro A, Carbonell C et al (2017) One and two-year persistence with different anti-osteoporosis medications: a retrospective cohort study. *Osteoporos Int* 28(10):2997–3004. <https://doi.org/10.1007/s00198-017-4144-7>
 30. Modi A, Sajjan S, Insinga R, Weaver J, Lewiecki EM, Harris ST (2017) Frequency of discontinuation of injectable osteoporosis therapies in US patients over 2 years. *Osteoporos Int* 28(4):1355–1363. <https://doi.org/10.1007/s00198-016-3886-y>
 31. Dalli LL, Olaiya MT, Kim J, Andrew NE, Cadilhac DA, Ung D et al (2023) Antihypertensive medication adherence and the risk of vascular events and falls after stroke: a real-world effectiveness study using linked registry data. *Hypertension* 80(1):182–191. <https://doi.org/10.1161/hypertensionaha.122.19883>
 32. US Food and Drug Administration (2017) Use of real world evidence to support regulatory decision making for medical devices
 33. U.S. Food & Drug Administration (2018) Framework for FDA's Real-World Evidence Program
 34. Arlett P, Kjør J, Broich K, Cooke E (2022) Real-world evidence in EU Medicines Regulation: enabling use and establishing value. *Clin Pharmacol Ther* 111(1):21–23. <https://doi.org/10.1002/cpt.2479>
 35. Pharmaceuticals and Medical Devices Agency: RWD WG. <https://www.pmda.go.jp/english/rs-sb-std/rs/0023.html> Accessed 13/12/2022.
 36. Therapeutic Goods Administration (Department of Health Australian Government) (2021) Real world evidence and patient reported outcomes in the regulatory context. Version 1.0 ed. Australia
 37. Purpura CA, Garry EM, Honig N, Case A, Rassen JA (2022) The role of real-world evidence in FDA-approved new drug and biologics license applications. *Clin Pharmacol Ther* 111(1):135–144. <https://doi.org/10.1002/cpt.2474>
 38. Flynn R, Plueschke K, Quinten C, Strassmann V, Duijnhoven RG, Gordillo-Marañón M et al (2022) Marketing authorization

- applications made to the European Medicines Agency in 2018–2019: What was the contribution of real-world evidence? *Clin Pharmacol Ther* 111(1):90–97. <https://doi.org/10.1002/cpt.2461>
39. Bakker E, Plueschke K, Jonker CJ, Kurz X, Starokozhko V, Mol PGM (2022) Contribution of real-world evidence in European Medicines Agency's regulatory decision making. *Clin Pharmacol Ther*. <https://doi.org/10.1002/cpt.2766>
 40. Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR et al (2009) Denosumab for prevention of fractures in postmenopausal women with osteoporosis. *N Engl J Med* 361(8):756–765. <https://doi.org/10.1056/NEJMoa0809493>
 41. Lai EC, Lin TC, Lange JL, Chen L, Wong ICK, Sing CW et al (2022) Effectiveness of denosumab for fracture prevention in real-world postmenopausal women with osteoporosis: a retrospective cohort study. *Osteoporos Int* 33(5):1155–1164. <https://doi.org/10.1007/s00198-021-06291-w>
 42. Storm NE, Chang W, Lin TC, Lange JL, Bradbury B, Critchlow CW et al (2022) A novel case study of the use of real-world evidence to support the registration of an osteoporosis product in China. *Ther Innov Regul Sci* 56(1):137–144. <https://doi.org/10.1007/s43441-021-00342-4>
 43. European Medicines Agency (2022) Eladynos (International non-proprietary name: abaloparatide) Assessment Report Amsterdam: European Medicines Agency
 44. Cosman F, Cooper C, Wang Y, Mitlak B, Varughese S, Williams SA (2022) Comparative effectiveness and cardiovascular safety of abaloparatide and teriparatide in postmenopausal women new to anabolic therapy: a US administrative claims database study. *Osteoporos Int* 33(8):1703–1714. <https://doi.org/10.1007/s00198-022-06413-y>
 45. Clemens KK, Jeyakumar N, Ouédraogo AM, Thain J, Khan T (2020) Bisphosphonate and denosumab initiation in older adults in Ontario, Canada: a population-based cohort study. *Arch Osteoporos* 15(1):133. <https://doi.org/10.1007/s11657-020-00796-3>
 46. Robinson DE, Ali MS, Pallares N, Tebé C, Elhoussein L, Abrahamson B et al (2021) Safety of oral bisphosphonates in moderate-to-severe chronic kidney disease: a binational cohort analysis. *J Bone Miner Res* 36(5):820–832. <https://doi.org/10.1002/jbmr.4235>
 47. O'Kelly J, Bartsch R, Kossack N, Borchert J, Pignot M, Hadji P (2022) Real-world effectiveness of osteoporosis treatments in Germany. *Arch Osteoporos* 17(1):119. <https://doi.org/10.1007/s11657-022-01156-z>
 48. Peng J, Wang H, Liu Z, Xu Z-L, Wang M-X, Chen Q-M et al (2022) Real-world study of antiresorptive-related osteonecrosis of jaw based on the US food and drug administration adverse event reporting system database. *Front Pharmacol*. 13:1017391. <https://doi.org/10.3389/fphar.2022.1017391>
 49. Krege JH, Gilsenan AW, Komacko JL, Kellier-Steele N (2022) Teriparatide and osteosarcoma risk: history, science, elimination of boxed warning, and other label updates. *JBMR Plus*. 6(9):e10665. <https://doi.org/10.1002/jbmr.410665>
 50. Gilsenan A, Midkiff K, Harris D, Kellier-Steele N, McSorley D, Andrews EB (2021) Teriparatide did not increase adult osteosarcoma incidence in a 15-year US postmarketing surveillance study. *J Bone Miner Res* 36(2):244–251. <https://doi.org/10.1002/jbmr.4188>
 51. Gilsenan A, Harris D, Reynolds M, McSorley D, Midkiff K, Jackson L et al (2021) Long-term cancer surveillance: results from the Forteo Patient Registry Surveillance Study. *Osteoporos Int* 32(4):645–651. <https://doi.org/10.1007/s00198-020-05718-0>
 52. Gilsenan A, Midkiff K, Harris D, McQuay L, Hunter S, Kellier-Steele N et al (2020) Assessing the incidence of osteosarcoma among teriparatide users based on Medicare Part D and US State Cancer Registry Data. *Pharmacoepidemiol Drug Saf* 29(12):1616–1626. <https://doi.org/10.1002/pds.5103>
 53. Li N, Cornelissen D, Silverman S, Pinto D, Si L, Kremer I et al (2021) An updated systematic review of cost-effectiveness analyses of drugs for osteoporosis. *Pharmacoeconomics* 39(2):181–209. <https://doi.org/10.1007/s40273-020-00965-9>
 54. Wu CH, Kao IJ, Hung WC, Lin SC, Liu HC, Hsieh MH et al (2018) Economic impact and cost-effectiveness of fracture liaison services: a systematic review of the literature. *Osteoporos Int* 29(6):1227–1242. <https://doi.org/10.1007/s00198-018-4411-2>
 55. Hilgsmann M, Boonen A, Rabenda V, Reginster JY (2012) The importance of integrating medication adherence into pharmaco-economic analyses: the example of osteoporosis. *Expert Rev Pharmacoecon Outcomes Res* 12(2):159–166. <https://doi.org/10.1586/erp.12.8>
 56. Hilgsmann M, Reginster JY, Tosteson ANA, Bukata SV, Saag KG, Gold DT et al (2019) Recommendations for the conduct of economic evaluations in osteoporosis: outcomes of an experts' consensus meeting organized by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO) and the US branch of the International Osteoporosis Foundation. *Osteoporos Int* 30(1):45–57. <https://doi.org/10.1007/s00198-018-4744-x>
 57. Shepherd JA, Fan B, Lu Y, Wu XP, Wacker WK, Ergun DL et al (2012) A multinational study to develop universal standardization of whole-body bone density and composition using GE Healthcare Lunar and Hologic DXA systems. *J Bone Miner Res* 27(10):2208–2216. <https://doi.org/10.1002/jbmr.1654>
 58. Ray WA, Griffin MR, Fought RL, Adams ML (1992) Identification of fractures from computerized medicare files. *J Clin Epidemiol* 45(7):703–714. [https://doi.org/10.1016/0895-4356\(92\)90047-Q](https://doi.org/10.1016/0895-4356(92)90047-Q)
 59. Curtis JR, Mudano AS, Solomon DH, Xi J, Melton ME, Saag KG (2009) Identification and validation of vertebral compression fractures using administrative claims data. *Medical Care*. 47(1)
 60. Konstantelos N, Rzepka AM, Burden AM, Cheung AM, Kim S, Grootendorst P et al (2022) Fracture definitions in observational osteoporosis drug effects studies that leverage healthcare administrative (claims) data: a scoping review. *Osteoporos Int* 33(9):1837–1844. <https://doi.org/10.1007/s00198-022-06395-x>
 61. Franklin JM, Paterno E, Desai RJ, Glynn RJ, Martin D, Quinto K et al (2021) Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE Initiative. *Circulation* 143(10):1002–1013. <https://doi.org/10.1161/circulationaha.120.051718>
 62. US Food and Drug Administration: FDA announces 4 grant awards for projects exploring the use of real-world data to generate real-world evidence in regulatory decision-making. <https://www.fda.gov/drugs/science-and-research-drugs/fda-announces-4-grant-awards-projects-exploring-use-real-world-data-generate-real-world-evidence> (2020). Accessed 21/02/2023
 63. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG et al (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 338:b2393-b. <https://doi.org/10.1136/bmj.b2393>
 64. Schneeweiss S, Avorn J (2005) A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 58(4):323–337. <https://doi.org/10.1016/j.jclinepi.2004.10.012>
 65. National Institute for Health and Care Excellence (2008) Raloxifene and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women (TA161). (Updated 2018).
 66. Kyriacou DN, Lewis RJ (2016) Confounding by indication in clinical research. *JAMA* 316(17):1818–1819. <https://doi.org/10.1001/jama.2016.16435>
 67. Robinson DE, Ali MS, Strauss VY, Elhoussein L, Abrahamson B, Arden NK et al (2021) Bisphosphonates to reduce bone fractures in stage 3B+ chronic kidney disease: a propensity score-matched

- cohort study. *Health Technol Assess* 25(17):1–106. <https://doi.org/10.3310/hta25170>
68. Suissa S (2008) Immeasurable time bias in observational studies of drug effects on mortality. *Am J Epidemiol* 168(3):329–335. <https://doi.org/10.1093/aje/kwn135>
 69. Oh I-S, Baek Y-H, Jeong HE, Filion KB, Shin J-Y (2020) Analytical approaches to minimizing immeasurable time bias in cohort studies. *Int J Epidemiol* 50(3):987–999. <https://doi.org/10.1093/ije/dyaa251>
 70. US Food and Drug Administration: Focus area: use of real-world evidence to support biologic, device, and drug development and regulatory decision-making. <https://www.fda.gov/science-research/focus-areas-regulatory-science-report/focus-area-use-real-world-evidence-support-biologic-device-and-drug-development-and-regulatory> Accessed 15/12/2022.
 71. European Medicines Agency: Big Data. <https://www.ema.europa.eu/en/about-us/how-we-work/big-data> Accessed 15/12/2022.
 72. Li Y, Sperrin M, Ashcroft DM, van Staa TP (2020) Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar. *BMJ*. 371:m3919. <https://doi.org/10.1136/bmj.m3919>
 73. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R et al (2004) Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *N Engl J Med* 351(12):1250–1251. <https://doi.org/10.1056/NEJMe048225>
 74. i-Sight Software: A practical guide to data privacy laws by country [2023]. <https://www.i-sight.com/resources/a-practical-guide-to-data-privacy-laws-by-country/> (2023). Accessed 20/04/2023.
 75. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J et al (2022) The structured process to identify fit-for-purpose data: a data feasibility assessment framework. *Clin Pharmacol Ther* 111(1):122–134. <https://doi.org/10.1002/cpt.2466>
 76. Mehra MR, Ruschitzka F, Patel AN (2020) Retraction-Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*. 395(10240):1820. [https://doi.org/10.1016/S0140-6736\(20\)31324-6](https://doi.org/10.1016/S0140-6736(20)31324-6)
 77. The Editors Of The Lancet Group (2020) Learning from a retraction. *Lancet* 396(10257):1056. [https://doi.org/10.1016/s0140-6736\(20\)31958-9](https://doi.org/10.1016/s0140-6736(20)31958-9)
 78. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I et al (2015) The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 12(10):e1001885. <https://doi.org/10.1371/journal.pmed.1001885>
 79. Langan SM, Schmidt SA, Wing K, Ehrenstein V, Nicholls SG, Filion KB et al (2018) The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 363:k3532-k. <https://doi.org/10.1136/bmj.k3532>
 80. Khalid S, Calderon-Larrañaga S, Hawley S, Ali MS, Judge A, Arden N, van Staa T, Cooper C, Javaid MK, Prieto-Alhambra D (2018) Comparative anti-fracture effectiveness of different oral anti-osteoporosis therapies based on “real-world” data: a meta-analysis of propensity-matched cohort findings from the UK Clinical Practice Research Database and the Catalan SIDIAP Database. *Clin Epidemiol* 10:1417–1431. <https://doi.org/10.2147/CLEP.S164112>

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

Authors and Affiliations

Rebecca J. Moon^{1,2}  · Jean-Yves Reginster^{3,4}  · Nasser M. Al-Daghri⁵  · Jotheeswaran A. Thiyagarajan⁶ · Charlotte Beaudart^{3,4}  · Olivier Bruyère^{3,4}  · Nansa Burlet⁴ · Manju Chandran⁷ · Mario Coelho da Silva⁸ · Philip G. Conaghan^{9,10}  · Willard H. Dere¹¹  · Adolfo Diez-Perez¹² · Peyman Hadji^{13,14}  · Philippe Halbout¹⁵ · Mickaël Hiligsmann¹⁶  · John A. Kanis^{17,18}  · Eugene V. McCloskey^{19,20}  · Sif Ormarsdottir²¹ · Daniel Prieto-Alhambra²² · Régis P. Radermecker²³  · René Rizzoli²⁴  · Yousef Al-Saleh^{25,26,27}  · Stuart L. Silverman²⁸ · Lee S. Simon²⁹  · Friederike Thomasius¹³ · Tjeerd van Staa³⁰  · Andrea Laslop³¹ · Cyrus Cooper^{1,32,33}  · Nicholas C. Harvey^{1,32} 

✉ Nicholas C. Harvey
nch@mrc.soton.ac.uk

¹ MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton SO16 6YD, UK

² Paediatric Endocrinology, University Hospital Southampton NHS Foundation Trust, Southampton, UK

³ WHO Collaborating Center for Epidemiology of Musculoskeletal Health and Ageing, Liège, Belgium

⁴ Division of Epidemiology, Public Health and Health Economics, University of Liège, Liège, Belgium

⁵ Biochemistry Department, College of Science, King Saud University, Riyadh, Saudi Arabia

⁶ Ageing and Health Unit, World Health Organisation, Geneva, Switzerland

⁷ Osteoporosis and Bone Metabolism Unit, Department of Endocrinology, Singapore General Hospital, Singapore, Singapore

⁸ Laboratory of Clinical and Therapeutic Pharmacology, Lisbon, Portugal

⁹ Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK

¹⁰ NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK

- 11 Department of Internal Medicine, Utah Center for Clinical and Translational Science, University of Utah School of Medicine, Salt Lake City, UT, USA
- 12 Department of Internal Medicine, Hospital del Mar-IMIM, Autonomous University of Barcelona and CIBERFES, Instituto Carlos III, Barcelona, Spain
- 13 Frankfurt Centre for Bone Health, Frankfurt, Germany
- 14 Philipps University of Marburg, Hesse, Germany
- 15 International Osteoporosis Foundation, Nyon, Switzerland
- 16 Department of Health Services Research, CAPHRI Care and Public Health Research Institute, Maastricht University, Maastricht, the Netherlands
- 17 Mary McKillop Institute for Health Research, Australian Catholic University, Melbourne, Australia
- 18 Centre for Metabolic Bone Diseases, University of Sheffield, Sheffield, UK
- 19 MRC Versus Arthritis Centre for Integrated Research in Musculoskeletal Ageing, University of Sheffield, Sheffield, UK
- 20 Mellanby Centre for Musculoskeletal Research, Department of Oncology & Metabolism, University of Sheffield, Sheffield, UK
- 21 Icelandic Medicines Agency, Reykjavik, Iceland
- 22 Pharmaco- and Device Epidemiology, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Oxford, UK
- 23 Department of Clinical Pharmacology, Diabetes, Nutrition and Metabolic Disorders, CHU Liege, Liege, Belgium
- 24 Geneva University Hospitals and Faculty of Medicine, Geneva, Switzerland
- 25 College of Medicine, King Saud Bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
- 26 King Abdullah International Medical Research Center, Riyadh, Saudi Arabia
- 27 Department of Medicine, King Abdulaziz Medical City, Riyadh, Ministry of National Guard-Health Affairs, Riyadh, Saudi Arabia
- 28 Cedars Sinai Medical Center, Los Angeles, USA
- 29 SDG LLC, Cambridge, MA, USA
- 30 Centre for Health Informatics, University of Manchester, Manchester, UK
- 31 Scientific Office, Austrian Medicines and Medical Devices Agency, Vienna, Austria
- 32 NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK
- 33 National Institute for Health Research (NIHR) Musculoskeletal Biomedical Research Centre, University of Oxford, Oxford, UK