# The implications of competing risks and direct treatment disutility in cardiovascular disease and osteoporotic fracture: risk prediction and cost effectiveness analysis

Bruce Guthrie,[1*] Gabriel Rogers,[2] Shona Livingstone,[3] Daniel R Morales,[3] Peter Donnan,[3] Sarah Davis,[4] Ji Hee Youn,[5] Rob Hainsworth,[2] Alexander Thompson[2] and Katherine Payne[2]

[1]Advanced Care Research Centre, Centre for Population Health Sciences, Usher Institute, The University of Edinburgh, Edinburgh, UK
[2]Manchester Centre for Health Economics, The University of Manchester, Manchester, UK
[3]Population Health and Genomics Division, University of Dundee, Dundee, UK
[4]School of Health and Related Research, The University of Sheffield, Sheffield, UK
[5]Open Health Group, Seoul, South Korea

[*]Corresponding author  bruce.guthrie@ed.ac.uk

## Disclosure of interests

## Scientific summary

The implications of competing risks and direct treatment disutility in cardiovascular disease and osteoporotic fracture: risk prediction and cost effectiveness analysis

Health and Social Care Delivery Research 2024; Vol. 12: No. 4
DOI: 10.3310/KLTR7714

NIHR Journals Library www.journalslibrary.nihr.ac.uk

# Scientific summary

## Background

Clinical guidelines help define and disseminate best practice. Guidelines increasingly use risk prediction tools to help target primary preventative treatments at people at highest risk. In National Institute for Health and Care Excellence (NICE) guidelines, the choice of risk threshold is commonly informed by model-based cost-effectiveness analyses (CEAs) for different levels of baseline risk. Risk prediction modelling and model-based CEA are therefore increasingly important for developing guidelines that recommend long-term preventative medicines, including primary prevention of cardiovascular disease (CVD) using statins and prevention of osteoporotic fracture using bisphosphonates.

### Risk prediction and competing mortality risk
Most risk prediction models do not account for competing mortality risk, which is when someone dies of another condition (e.g. lung cancer) before experiencing the event being predicted (e.g. CVD or fracture). This can lead to overprediction of event rates among older people and those with multimorbidity.

### Model-based cost-effectiveness analysis
Competing mortality risk is accounted for in model-based CEA, but whole-population estimates of competing mortality will not be correct at all levels of risk of CVD and fracture. Existing models also do not account for all harms, notably direct treatment disutility (DTD), which is the disutility arising from the hassles of taking treatments. Even small levels of DTD can be enough to outweigh relatively small lifetime benefits of primary prevention medication, but, to our knowledge, DTD impact has not been systematically estimated previously.

## Aim and objectives

The overall aim was to improve the evidence generated from risk prediction models and model-based CEAs to inform decision-making for selecting primary prevention treatments for CVD and osteoporotic fracture.

The prespecified objectives were to:

1.  externally validate the recommended risk prediction tools for primary prevention of CVD [QRISK®3 (ClinRisk Ltd, Leeds, UK)] and for osteoporotic fracture [QFracture-2012 (ClinRisk Ltd)]
2.  derive and internally validate new CVD and osteoporotic fracture risk prediction models accounting for competing risks of death
3.  externally validate the QRISK-Lifetime CVD risk prediction tool
4.  quantify the magnitude, variation and distribution of DTD in the general population and among people treated with statins or bisphosphonates
5.  examine the effect of accounting for competing risks and DTD on cost-effectiveness in the context of statins and bisphosphonates for the primary prevention of CVD and osteoporotic fracture, respectively.

The prediction modelling protocol was approved by the Clinical Practice Research Datalink (CPRD) Independent Scientific Advisory Committee (reference number 16_248). The Health Research Authority approved the DTD elicitation study (Integrated Research Application System: 220,492) and granted ethics approval (Research Ethics Committee: 17/NW/0124). A systematic review for CEA model parameters was registered with PROSPERO (CRD42021249959).

## Methods

### Objective 1 methods

For CVD modelling, CPRD GOLD data were used to define a cohort aged 25–84 years without CVD or prior statin prescription. The outcome was incident CVD. Multiple imputation was used to account for missing data. The performance of the published QRISK3–2017 model was evaluated in terms of discrimination (the ability of a tool to distinguish between those with and those without an event) and calibration (whether or not predicted risk is the same as observed risk) in the whole population, stratified by age and Charlson Comorbidity Index (CCI), and in subgroups with type 1 diabetes, type 2 diabetes and chronic kidney disease (CKD). Observed risk was estimated with and without accounting for competing risk (using Aalen–Johansen and Kaplan–Meier estimators, respectively).

For fracture modelling, the cohort was aged 30–99 years (prior fracture or bisphosphonate treatment were allowed) with follow-up to specified fracture, death from non-fracture causes, deregistration or end of study. Two outcomes were defined: major osteoporotic fracture (MOF) and hip fracture. QFracture-2012 performance was evaluated as for QRISK3.

For both cohorts, the earliest study entry date was 1 January 2004 and the end of the study was 31 March 2016.

### Objective 2 methods

Using the same data set as objective 1, participants were randomly allocated to derivation and test data sets in a 2 : 1 ratio. For CVD, two Fine–Gray models were derived in the derivation data set and internally validated in the test data set, alongside QRISK3. The competing mortality risk model (CRISK) accounted for competing mortality only, whereas the competing mortality risk model with Charlson Comorbidity Index (CRISK-CCI) also included the modified CCI as a predictor. Model performance was examined using discrimination and calibration. For fracture, separate Fine–Gray models (CFracture) were estimated for MOF and hip fracture.

### Objective 3 methods

The same data were used as for objective 1, but with the age range restricted to 30–84 years to match QRISK®-Lifetime (ClinRisk Ltd). As lifetime risk is not observed in this data set, model performance was evaluated at 10 years, and reclassification examined the characteristics of those recommended for treatment on the basis of a QRISK3 10-year risk of > 10%, a QRISK-Lifetime 10-year risk of > 10% and the QRISK-Lifetime highest risk, with thresholds chosen to recommend the same number of people for treatment as with QRISK3 > 10%.

### Objective 4 methods

Two groups of participants were recruited to studies to elicit DTD of preventative statins and bisphosphonates: people with direct experience of taking one of the medicines and a sample of the general population. We described the process of taking each medicine (one tablet per day for statins, one tablet per week taken on an empty stomach with a requirement to stay upright for at least 30 minutes for bisphosphonates). Elicitation used time trade-off (TTO) (primary analysis) and best–worst scaling (BWS) (exploratory analysis) surveys iteratively developed using think-aloud interviews with 19 patients, and online pilot studies.

### Objective 5 methods

For statins for the primary prevention of CVD, we modified the cohort-level decision-analytic model used in NICE's lipid modification guideline [NICE. Lipid Modification: Cardiovascular Risk Assessment and the Modification of Blood Lipids for the Primary and Secondary Prevention of Cardiovascular Disease. Clinical Guideline (CG181). Methods, Evidence and Recommendations. July 2014. URL: https://web.archive.org/web/20220201050407/https://www.nice.org.uk/guidance/cg181/evidence/lipid-modification-update-full-guideline-pdf-243786637 (accessed 12 October 2022)]. General updates

included rapid reviews to identify utility values and costs associated with CVD events, new regressions to predict baseline quality of life for people without CVD (based on Health Survey for England data) and type of first CVD event (based on data from objective 1), and inputs (costs, life expectancy) were updated to present-day values. For bisphosphonates for the prevention of fracture, we used the discrete-event simulation developed for NICE's Technology Appraisal 464 [NICE. Bisphosphonates for Treating Osteoporosis. Technology Appraisal Guidance (TA464). London: NICE; 2017].

For both models, we explored competing risk by parameterising probability of non-cause-specific death using relative survival models adjusting for predicted risk (QRISK3 or QFracture-2012). We incorporated DTD as elicited in objective 4 under three assumptions (lifelong, time limited, diminishing over time). We explored how these factors alone or in combination affect the estimated value of the preventative medicines in terms of cost per quality-adjusted life-year (QALY).

## Results

### Objectives 1 and 2: predicting cardiovascular disease

Discrimination of QRISK3 in the whole external validation cohort was excellent (Harrell's c = 0.865 for women, 0.834 for men), and comparable to the previous internal validation. However, discrimination was worse among people with more comorbidity, and was poor to moderate among older people (e.g. c = 0.611 for women and 0.585 for men aged 75–84 years). Calibration in the whole population, ignoring competing risks, was very good, with minor overprediction. There was larger overprediction among older people, which was considerable after accounting for competing risks.

Among people with type 1 diabetes, discrimination was excellent (c = 0.830 for women, 0.853 for men). There was evidence of overprediction at higher levels of predicted risk, which was larger after accounting for competing risks, although most overprediction happened well above the NICE 10% threshold for offering treatment. Discrimination among people with CKD was only moderate (women, c = 0.705; men, c = 0.671), but calibration was reasonable at recommended treatment thresholds.

The new competing risk model (CRISK-CCI) had similar discrimination to QRISK3 in the whole population (women, c = 0.864; men, c = 0819), with the same pattern of worse discrimination among older people and those with more comorbidity. Calibration was systematically better than QRISK3, although, as with QRISK3, there was overprediction in some subgroups with high predicted risk.

### Objectives 1 and 2: predicting fracture

Observed age-stratified incidences of both MOF and hip fracture were considerably higher in this study than in a previous external validation, which was partly explained by the use of hospital data in this study to ascertain fractures. Discrimination of QFracture-2012 in external validation was excellent among women (MOF, c = 0.813; hip fracture, c = 0.918) and good to excellent among men (MOF, c = 0.738; hip fracture, c = 0.888), similar to QFracture-2012 internal validation, but had poor to moderate discrimination among older people. Ignoring competing risks, QFracture-2012 showed serious underprediction in the whole population and in all subgroups of age and comorbidity, which was worse for hip fracture than for MOF. Accounting for competing risks reduced observed underprediction in the whole population, but there was very major overprediction among older people and at higher levels of predicted risk among people with more comorbidity.

The new competing risk model (CFracture) had similar discrimination to QFracture-2012 in the internal validation cohort (women: c = 0.813 for MOF, c = 0.914 for hip fracture; men: c = 0.734 for MOF, c = 0.883 for hip fracture). CFracture was better calibrated than QFracture-2012 but showed overprediction at higher levels of predicted risk for MOF (both sexes) and for hip fracture (among men). CFracture calibration was poor among people aged 85–99 years for both outcomes.

### Objective 3: predicting lifetime cardiovascular disease risk

Evaluated at 10 years' follow-up, QRISK-Lifetime had excellent discrimination (women, c = 0.844; men, c = 0.808) in the whole population, with the same pattern as QRISK3 and CRISK-CCI of worse discrimination among older people and those with high comorbidity. QRISK-Lifetime underpredicted 10-year risk among people at higher predicted risk, particularly older people, implying that estimated lifetime risk will be underpredicted. A total of 5.3% of participants were recommended for treatment by both QRISK3 and QRISK-Lifetime, and 27.4% by one or the other, but not both. Participants recommended for treatment by QRISK-Lifetime were younger than those recommended by QRISK3 (mean age: women, 50.5 vs. 71.3 years, respectively; men, 46.3 vs. 63.8 years, respectively), were much more likely to have a strong family history of CVD (women: 36.3% vs. 6.3%, respectively; men: 20.0% vs. 7.2%, respectively) and had many fewer observed events during the 10-year follow-up (women with a CVD event: 4.0% vs. 11.9%, respectively; men with a CVD event: 4.3% vs. 10.8%, respectively).

### Objective 4: direct treatment disutility elicitation

When measured by TTO, long-term statin use was associated with mean DTD of 0.034 among people willing to take statins; the equivalent number for bisphosphonates was significantly greater, at 0.067. The findings from the BWS experiment had face validity in that inconvenience influenced preferences. However, the estimated values for DTD are implausibly large.

Consistent with previous studies, these findings suggest three distinct preference phenotypes: some people would avoid taking the medicines at all costs, some people see no problem with them and some people are willing to trade length of life to avoid treatment. The first group are unlikely to initiate treatment and the second group do not anticipate DTD; in the third group, depending on the individual's strength of preference to avoid treatment and the magnitude of expected QALY gains from prevention, DTD may imply that a preventative medicine's negative characteristics outweigh its benefits.

### Objective 5: model-based cost-effectiveness analysis

General updates to the CVD model made high-intensity statins more cost-effective for primary prevention. Introducing accurate adjustment for competing risk of non-CVD death had the expected effect: more QALYs among people with below-average CVD risk for their sex and age (who experience lower rates of other-cause mortality) and fewer QALYs among people with above-average risk (whose non-CVD life expectancy is attenuated). However, the impact on incremental cost-effectiveness is minor, and statins remain almost universally cost-effective. Incorporating DTD has a more obvious effect, especially when we assume that it applies undiminished for as long as people take statins for primary prevention. Under that circumstance, the threshold at which expected long-term benefits outweigh DTD-related harm rises with age: for a 40-year-old, a 10-year risk of ≥ 8% would be enough to make treatment net beneficial whereas, for an 80-year-old, that figure rises to 38%.

The model assessing bisphosphonates for the primary prevention of osteoporotic fragility fracture shows that we overestimate value for money among people at the highest risk if we do not adjust for competing risk of non-fracture death. However, this generally affects only the magnitude of expected net benefit among people for whom some degree of benefit is expected. Even among people at highest risk of fracture, average QALY gains associated with bisphosphonates are small and swamped by DTD of any duration. Consequently, it is impossible to identify any group of people for whom oral bisphosphonates represent an effective use of NHS resources, if we assume population-level average DTD for everyone to whom the decision applies.

## Conclusions

### Implications for healthcare

Ignoring competing mortality in risk prediction overestimates the risk of CVD and fracture among older people and those with multimorbidity, which will lead to overestimation of the benefits of treatment.

This affects fracture risk prediction more than CVD because CVD is a more substantial proportion of total mortality. The QFracture-2012 prediction tool simultaneously underestimates fracture risk among people without high competing mortality risk, partly because it did not include fractures recorded only in hospital data in its derivation. CVD and fracture risk prediction are improved by accounting for competing mortality risks, and transparency of the tools would be improved by fully publishing the codes used to define events and predictors.

We have demonstrated an effective method of making accurate adjustment for competing risk of non-cause-specific death in decision-analytic CEAs. Although it made relatively little difference to the estimated cost-effectiveness of preventative interventions in the examples we explored, we have shown that it could potentially be important. Therefore, we recommend that modellers consider this issue when designing analyses of preventative treatments.

Although we have demonstrated that DTD exists and has the potential to alter the balance of benefits and harms for preventative treatments, we do not recommend that population-level average DTD is incorporated in base-case CEAs. Rather, we recommend that decision-makers review scenarios with and scenarios without DTD and highlight its possible impact, enabling prescribers to engage in shared decision-making that gives appropriate weight to individual preferences.

### Research recommendations

The excellent discrimination of QRISK3 and QFracture-2012 arises from including a very broad range of ages, but discrimination and calibration in subgroups are less good. Comparing models created in smaller age groups with whole-population models would be useful. Mortality is only one competing risk, and older people and those with multimorbidity are at risk of many different events. It is important to develop models that better account for multiple important events.

Cost-effectiveness analysis of statins for the primary prevention of CVD could usefully be further modified to (1) enable stratification according to specific coexisting long-term conditions, (2) account for likely adherence to statins in practice and (3) update secondary transitions reflecting the subsequent natural history of CVD among people experiencing events.

Future CEAs of bisphosphonates for the primary prevention of osteoporotic fragility fracture should explore different fracture risk prediction models, and use those based on demonstrable good ascertainment of fractures and accounting for competing mortality risk.

## Study registration

This study is registered as PROSPERO CRD42021249959.

## Funding

# Health and Social Care Delivery Research

> **Criteria for inclusion in the *Health and Social Care Delivery Research* journal**
> Reports are published in *Health and Social Care Delivery Research* (HSDR) if (1) they have resulted from work for the HSDR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

## HSDR programme

The HSDR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HSDR programme please visit the website at https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-and-social-care-delivery-research.htm.

## This report