Incorporating Novel Risk Markers into Established Risk Prediction Models

By Paula Dhiman

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

June 2015

Abstract

Introduction: Risk prediction models are used as part of formal risk assessment for disease and health events in UK primary care. They are used to identify patients at risk of disease, for whom preventative treatment can be offered. To improve the accuracy of risk prediction, new risk factors are being added to established risk prediction models. However, current methods used to evaluate the added value of these new risk factors have shown to be limited; insensitive to change; and lack interpretability and clinical relevance. These limitations can be addressed using health economic methodology to evaluate and compare interventions by means of their effectiveness and cost, but is yet to be used to evaluate and compare risk prediction models.

Objectives: Objective one of this thesis is to collate and critique current methods used to evaluate the added value of a new risk factor. Objective two is to refine current methodology to also assess cost impact when evaluating the added value of a new risk factor; using health economic evaluation methods. Objective three is to apply the refined methodology on two exemplar datasets; assessing the added value of family history (FH) of cardiovascular disease (CVD) to CVD risk assessment; and assessing the added value of bone mineral density (BMD) to fracture risk assessment.

Methods: Current validation measures used to describe the calibration and discrimination of a risk prediction model were used to populate the decision tree, with assigned costs of using the risk model, treatment, and the outcome. These were used to derive the incremental cost effectiveness ratio (ICER), where the measure of effect was a risk model performance based measure. The ICER was used to describe the added value of the new risk factor. A probabilistic sensitivity analysis was also performed to account for the uncertainty and derive 95% confidence intervals for the ICER, based on 10,000 iterations.

A range of £0-£100,000 was used for the willingness to pay, and was used to derive the incremental net monetary benefit (INB). This provided the probability the new risk model, with the new risk factor was cost effective. The cost-effectiveness analysis was performed in addition to the current approaches used in validation studies. This method was applied in two exemplar prospective cohort studies; adding FH to CVD risk prediction; and BMD to fracture risk prediction.

Results: A cost-effectiveness analysis using a decision tree framework was shown to be an effective way of evaluating the added value of the new risk factor. The cost-effectiveness analysis presented ICER estimates for the new risk factor; the ICER for adding FH to standard CVD risk factors was £799.91 (-£5,962.15 to £5,968.22) and £7,788.76 (-£42,760.16 to £48,962.39) per percentage unit increase in the Youden Index and Harrell's C-Index, respectively. Adding continuous BMD measurement to standard fracture risk factors produced an ICER of £367.25 (-£4,241.88 to £4,828.50) and £4,480.54 (-£22,816.84 to £22,970.55) per percentage unit increase in the Youden Index and Harrell's C-Index, respectively. As the Harrell's C-Index was not directly calculated from the decision tree, it needs to be interpreted with caution.

Uncertainty was explored with a probabilistic sensitivity analysis; analyses showed large amounts of parameter uncertainty were found, illustrated by wide 95% confidence intervals. This analysis also provided the cost effectiveness probability for the new risk factor. The maximum probability of adding FH of CVD to standard CVD risk factors being cost-effective is 0.7, with a minimum willingness to pay of £15,000 based on the improvement of the Youden Index discriminatory accuracy measure. The maximum probability of adding continuous BMD to standard fracture risk factors being cost-effective is 0.8, with a minimum willingness the pay of £32,500 based on the improvement of the Youden Index discriminatory accuracy measure.

Further, the CVD analysis found that treating low risk patients with statin therapy incorrectly was less costly (£788.40) than not treating them (£916.16). The cost of the unnecessary treatment was absorbed into the costs saved from avoiding the CVD event. It was also found that BMD used in a binary format to indicate osteoporotic patients, did not improve the discriminatory accuracy of standard fracture risk prediction (absolute decrease = 0.0048 (-0.62%)).

Conclusion: A cost-effectiveness analysis was a novel method to compare two risk prediction models; and to evaluate the added value of a new risk factor. It identifies the added value of a new risk factor encompassing the statistical and clinical improvement, and cost consequences when using the new risk factor in an established risk prediction model. Based on the added value of FH to CVD risk prediction, and BMD to fracture risk prediction, there is a good evidence base to add these risk factors into routine risk assessment of the respective conditions. Increased use of this method could help standardise risk prediction and increase comparability of risk prediction models within diseases; producing a league table approach to evaluate, appraise and identify beneficial new risk factors and better risk prediction models.

Published Papers

Dhiman P, Kai J, Horsfall L, Walters K, Qureshi N (2014). Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment. PLoS ONE 9(1): e81998. doi:10.1371/journal.pone.0081998.

Acknowledgements

In full gratitude, I would like to acknowledge the following individuals who encouraged, inspired, supported, assisted, and sacrificed themselves to help my pursuit of a higher education degree.

Firstly, I would like to thank all of my supervisors, Professor Nadeem Qureshi (University of Nottingham, Clinical Professor), Professor Min Yang (University Of Nottingham, Professor of Statistics), and Professor Richard Morris (University Of Bristol, Professor of Medical Statistics), for their efforts, advice, and guidance in producing this thesis, and for my growth as an independent researcher. In particular, I would like to thank Professor Nadeem Qureshi for his continued support and feedback whilst undertaking this PhD.

I would also like to thank Professor Tahir Masud (Nottingham University Hospitals NHS Trust, Geriatric Specialist) for his expertise in fracture risk prediction and help in gaining access to the AURORA dataset, which was used to evaluate the added value of bone mineral density in fracture risk prediction. I would also like to thank Professor Stig Anderson (Aalborg University Hospital, Clinical Associate Professor) for allowing access to his robustly collected AURORA dataset.

I would like to thank Professor Carol Coupland (University of Nottingham, Reader of Medical Statistics) for her statistical expertise and for externally reviewing the statistical methodology in this thesis. I would also like to thank Professor Matthew Jones (University of Nottingham, Assistant Professor of Health Economics) for his advice and expertise in health economics, and for his help in developing and reviewing the refined methodology. I would also like to thank Dr Tom Dent (PHG Foundation, Public Health Consultant) for his comments and advice for the literature review.

Lastly, but not the least, I would like to thank my family and friends for their continued support throughout this PhD. Simply put, I could not have done it without you!

Table of Contents

Abstract ..... **i**
Published Papers ..... **v**
Acknowledgements ..... **vii**
Table of Contents ..... **ix**
List of Figures ..... **xv**
List of Tables ..... **xxi**
List of Abbreviations ..... **xxix**
Chapter 1 - Introduction **1**
1.1 Rationale 1
1.1.1 Risk Prediction Models *1*
1.1.2 Improving Risk Prediction Models – New Risk Factors *2*
1.1.3 Methodology to Evaluate the Added Value of a New Risk Factor *3*
1.2 Aims and Objectives 4
1.2.1 Aim and Objectives *4*
1.3 Structure of thesis 5
1.4 Chapter Summary 6
Chapter 2 - Literature Review **7**
2.1 Objective and search strategy 7
2.2 Risk Prediction Models 7
2.3 New Risk Factors 9
2.4 Comparing Risk Prediction Models 10
2.5 Current Methodology 11
2.5.1 Guidelines *11*
2.5.2 Statistical Validation *12*
2.5.3 Clinical Validation *19*
2.6 Conclusion 22
2.7 Chapter Summary 23
Chapter 3 - Refined Methodology **25**
3.1 Rationale 25
3.2 Economic Evaluation in Health Care 26
3.2.1 Analyses *27*

3.2.2 Decision Analytical Framework *30*
3.3 Methods Guide and Reporting of Health Economic Evaluation 38
3.3.1 Methods – General *40*
3.3.2 Methods – Outcomes *42*
3.3.3 Methods – Resources and Costs *43*
3.3.4 Methods – Model-Based Economic Evaluations *48*
3.3.5 Methods – Analytical Methods *49*
3.4 Chapter Summary 57
Chapter 4 - Family History and Cardiovascular Disease Risk: Introduction and Methods 59
4.1 Introduction 59
4.1.1 The Burden of Cardiovascular Disease *59*
4.1.2 Prevention of Cardiovascular Disease in Primary Care *60*
4.1.3 Cardiovascular Risk Prediction *61*
4.1.4 Aim and Objectives *69*
4.2 Methods 70
4.2.1 Study Design *70*
4.2.2 Data Source *71*
4.2.3 Study population *73*
4.2.4 Study variables *74*
4.2.5 Statistical Analysis *89*
4.3 Chapter Summary 111
Chapter 5 - Family History and Cardiovascular Disease Risk: Results and Discussion (Current Methodology) 113
5.1 Results 113
5.1.1 Extraction *113*
5.1.2 Overall Cohort Characteristics *116*
5.1.3 Incidence of CVD *125*
5.1.4 Generalisability of the data *133*
5.1.5 Missing Data *137*
5.1.6 Derivation and Validation Cohorts *145*
5.1.7 Derivation of Models *157*
5.1.8 Assessment of Model Performance *178*
5.1.9 Sensitivity Analysis *189*
5.1.10 Section Summary *193*
5.2 Discussion 195
5.2.1 Summary of findings *195*

5.2.2 Strengths and Limitations *197*
5.3 Chapter Summary 201
Chapter 6 - Family History and Cardiovascular Disease Risk: Results and Discussion (Refined Methodology) **203**
6.1 Results 203
6.1.1 Cost-Effectiveness Analysis *203*
6.1.2 Sensitivity Analysis *222*
6.2 Discussion 229
6.2.1 Summary of findings *229*
6.2.2 Strengths and Limitations *230*
6.2.3 Clinical implications and Conclusion *233*
6.3 Chapter Summary 234
Chapter 7 - Bone Mineral Density and Fracture Risk: Introduction and Methods **235**
7.1 Introduction 235
7.1.1 The Burden of Osteoporosis *235*
7.1.2 Prevention of Fragility Fracture in Primary Care *236*
7.1.3 Fracture Risk Prediction *238*
7.1.4 Objective and Aims *245*
7.2 Methods 246
7.2.1 Study Design *246*
7.2.2 Data Source *247*
7.2.3 Study population *249*
7.2.4 Study variables *250*
7.2.5 Statistical Analysis *255*
7.3 Chapter Summary 268
Chapter 8 - Bone Mineral Density and Fracture Risk: Results and Discussion (Current Methodology) **271**
8.1 Results 271
8.1.1 Dataset *271*
8.1.2 Overall Cohort Characteristics *272*
8.1.3 Incidence of Fracture *278*
8.1.4 Generalizability of the data *283*
8.1.5 Missing Data *291*
8.1.6 Derivation and Validation Cohorts *292*
8.1.7 Derivation of Models *299*
8.1.8 Assessment of Model Performance *325*
8.1.9 Sensitivity Analysis *343*

8.1.10 Section Summary 344
8.2 Discussion 345
8.2.1 Summary of findings 345
8.2.2 Strengths and limitations 347
8.3 Chapter Summary 349
Chapter 9 - Bone Mineral Density and Fracture Risk: Results and Discussion (Refined Methodology) 351
9.1 Results 351
9.1.1 Cost-Effectiveness Analysis 351
9.1.2 Sensitivity Analysis 377
9.2 Discussion 380
9.2.1 Summary of findings 380
9.2.2 Strengths and Limitations 381
9.2.3 Clinical implications and Conclusion 383
9.3 Chapter Summary 384
Chapter 10 - Discussion 385
10.1 Summary of the findings 385
10.1.1 Objective 1: Review and Critique of Current Methodology 385
10.1.2 Objective 2: Refine Current Methodology using Health Economic Evaluation Methods 387
10.1.3 Objective 3: Apply the Refined Methodology to Two Exemplar Datasets 389
10.2 Strengths and Limitations 391
10.3 Implications 394
10.4 Recommendations for Future Research 396
10.5 Conclusion 397
References 399
Appendices 415
Appendix A. Search terms for Literature Review 415
Appendix B. Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment (93). 417
Appendix C. Read code and ICD code lists for Cardiovascular Disease 425
Appendix D. Read code list for Family History 439
Appendix E. Generalisability of extracted CPRD data to Health Survey for England 2011. 443
Appendix F. Regression output for missing data. 445
Appendix G. Schoenfeld residual plots for Proportional Hazards. 467

Appendix H. AIC, BIC, R^2, Likelihood Ratio values for each model and imputed dataset. 485
Appendix I. FH and CVD Study: Complete Case Analysis 487
Appendix J. Literature based Survival 497
Appendix K. Availability of bone mineral density in routinely collected data 501
Appendix L. Associations between fracture risk factors 509
Appendix M. Schoenfeld Residual Plots to test for proportional hazards for each risk factor. 515
Appendix N. Schoenfeld Residual Plots to test for proportional hazards for each risk model. 523

List of Figures

Figure 3.1 Visual representations of a Decision Tree structure. 33

Figure 3.2 Visual representation of the proposed decision tree for evaluating the added value of a risk factor. 36

Figure 3.3 Simplified visual representation of the proposed decision tree for evaluating the added value of a risk factor; combined cost of the treatment and outcome. 46

Figure 3.4 Cost-effectiveness plane. 52

Figure 3.5 Incremental Net Monetary Benefit (INB) as a function of the willingness to pay (λ). 54

Figure 3.6 Cost-Effectiveness Acceptability Curve. 56

Figure 4.1 Decision Tree to compare two CVD risk prediction models. 108

Figure 5.1 Selection of patients for the study cohort. 115

Figure 5.2 Comparison of mean values for variables by age group (10year) in the CPRD dataset compared with the HSE. 136

Figure 5.3 Schoenfeld residual plots for each risk factor in Model 1 (including standard risk factors only) using imputed dataset 1. 164

Figure 5.4 Schoenfeld residual plots for each risk factor in Model 2 (including standard risk factors and FH) using imputed dataset 1. 166

Figure 5.5 Schoenfeld residual plots for each risk factor in Model 3 (including standard and additional risk factors) using imputed dataset 1. 169

Figure 5.6 Schoenfeld residual plots for each risk factor in Model 4 (including standard and additional risk factors, with FH) using imputed dataset 1. 171

Figure 5.7 Cox-Snell Residual plots for each model using Imputation 1. 172

Figure 5.8 Log-displacement plots for each model using Imputation 1. 173

Figure 5.9 LMAX plots for each model using Imputation 1.....174
Figure 5.10 Kernel density plots for predicted risk for all models. 179
Figure 5.11 Predicted and observed risk by 10th of predicted risk for each risk prediction model in the derivation dataset.....180
Figure 5.12 Receiver Operating Characteristic Curve for each derived CVD risk model. 182
Figure 5.13 Reclassification plots comparing models with and without FH information. Reference line indicates the 10% risk threshold. 183
Figure 5.14 Calibration Plots of the 4 CVD risk prediction models derived using complete cases of the data.....190
Figure 5.15 Receiver Operating Characteristic (ROC) Curve, and Harrell's C-Index values.....191
Figure 5.16 Calibration Plots of the 4 CVD risk prediction models derived using imputed data and a literature based survival estimate (0.9551). 192
Figure 5.17 Receiver Operation Characteristic (ROC) Curve, and Harrell's C-Index values for models derived using the imputed data and literature based survival estimate (0.9551). 193
Figure 6.1 Decision tree to compare Model 1 (standard risk factors only) to Model 2 (standard risk factors with FH).....206
Figure 6.2 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 1 and Model 2. 209
Figure 6.3 Incremental net benefit, for Harrell's C-Index and Youden Index, comparing Model 1 and Model 2. 210
Figure 6.4 Decision tree to compare Model 3 (standard and additional risk factors only) to Model 4 (standard and additional risk factors with FH). 211

Figure 6.5 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 3 and Model 4. 212
Figure 6.6 Incremental net benefit, for Harrell's C-Index and Youden Index, comparing Model 3 and Model 4. 213
Figure 6.7 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 1 and Model 2. 218
Figure 6.8 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 1 and Model 2. 219
Figure 6.9 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis; comparing Model 3 and Model 4. 221
Figure 6.10 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 3 and Model 4. 222
Figure 7.1 Decision Tree to compare two fracture risk prediction models. 266
Figure 8.1 Selection of patients for the study cohort. 272
Figure 8.2 Comparison of AURORA and CPRD data characteristics, by age group. 287
Figure 8.3 Schoenfeld Residual plot for proportional hazards of previous fracture. 307
Figure 8.4 Kernel density plots for predicted risk for all models. 326
Figure 8.5 Kernel density plots of predicted risk comparing models with and without BMD information. 327
Figure 8.6 Predicted and Observed risk by 10^{th} of predicted risk for each risk prediction model in the derivation dataset. 328
Figure 8.7 Receiver Operating Characteristic curve for each model using the derivation dataset, with related Harrell's c-index (analogous to the AUC). 333

Figure 8.8 Reclassification plots comparing models with and without BMD measurement included. Reference lines indicate the 8.5% risk threshold.....334
Figure 8.9 Net Benefit for each treatment strategy using an 8.5% risk threshold. 343
Figure 9.1 Decision tree to compare Model 2 (standard risk factors only) to Model 3a (standard risk factors with binary BMD measurement). 354
Figure 9.2 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 3a. 358
Figure 9.3 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 3b. 360
Figure 9.4 Incremental net benefit for Harrell's C-Index and the Youden Index, comparing Model 2 with Model 3b. 361
Figure 9.5 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 4b. 363
Figure 9.6 Incremental net benefit for Harrell's C-Index and the Youden Index, comparing Model 2 with Model 4b. 364
Figure 9.7 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 3a. 369
Figure 9.8 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 3a. 370
Figure 9.9 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 3b. 372
Figure 9.10 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 3b. 373

Figure 9.11 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 4b. 375
Figure 9.12 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 4b. 376

List of Tables

Table 3.1 Table to identify patients as true/false positives and true/false negatives based on their observed and predicted outcomes. 35

Table 3.2 Values and their location to be inputted into the decision tree. 44

Table 4.1 Table of CVD risk factors by risk prediction model. 62

Table 4.2 Variables included in the JBS3, JBS3+, ASSIGN and Reynolds CVD risk prediction models. 75

Table 4.3 List of study variables. 77

Table 4.4 Criteria to class smokers as trivial, light, moderate, heavy and very heavy smokers based on the number of cigarettes smoked per day. 81

Table 4.5 Criteria to class drinkers as trivial, light, moderate, heavy and very heavy drinkers based on the number unit of alcohol consumed per week. 85

Table 4.6 Criteria to class stages of Chronic Kidney Disease, based on estimated glomerular filtration rate (mL/min/1.73m^2). 87

Table 4.7 Performance results from the model, compared to the observed outcome, using a pre-specified risk threshold. 96

Table 4.8 Cost of GPs as provided by the PSSRU (70). 102

Table 4.9 Lipid Modification Prescribing Proportions (26). 104

Table 4.10 Weighted cost of a CVD event (7). 105

Table 4.11 CVD event rates from the ASCOT LLA and CARDS study. 106

Table 4.12 Average cost of treated and untreated CVD events over 10 years follow up, by discount rate. 106

Table 4.13 Deterministic mean costs of risk model (with and without FH), treatment and fracture; with standard error, alpha, and beta for estimation of probabilistic cost derived from a Gamma Distribution. 110
Table 5.1 Baseline Characteristics of the overall sample derived from the CPRD database (N=302,591). 117
Table 5.2 Characteristics of the overall extracted cohort derived from CPRD, by FH status. Values are numbers (%), mean (SD), median [IQR]; and p-values. 121
Table 5.3 Characteristic of the overall sample derived from CPRD, by CVD outcome during the 10 year follow up period. Values are numbers (%), mean (SD), median [IQR]; and p-values.....126
Table 5.4 Crude incidence rates of the CVD outcome in the overall cohort (n=302,509). Values are the number of incident fracture cases, the total person years and crude incidence rates (95% Confidence Interval). 131
Table 5.5 Comparison of proportions of recorded data in each dataset for each comparable variable. Values are numbers (%) 134
Table 5.6 Comparison of characteristics for CPRD and HSE data. 135
Table 5.7 Missing data of data for standard, new, and additional CVD risk factors. 138
Table 5.8 Completeness and missing data of risk factors. 139
Table 5.9 Characteristics of the original and imputed datasets (with and without FH status). 143
Table 5.10 Baseline characteristics of the derivation and validation dataset. Values are numbers (%), mean (SD), median [IQR]. 146
Table 5.11 Crude CVD incidence rates for the derivation and validation datasets. 150
Table 5.12 Association of standard and new risk factor with CVD in the overall, derivation and validation datasets. 155

Table 5.13 Association between all risk factors included in the analysis to derive 10 year CVD risk prediction models. Values are beta coefficients (p-values) derived using logistic regression, unless otherwise stated.....158
Table 5.14 Unadjusted Cox Regression Analysis. 161
Table 5.15 Beta Coefficients and Hazard Ratios from Model 1 based on standard CVD risk factors. 163
Table 5.16 Beta Coefficients and Hazard Ratios from Model 2 based on standard CVD risk factors with FH. 165
Table 5.17 Beta Coefficients and Hazard Ratios from Model 3 based on standard and additional CVD risk factors.....167
Table 5.18 Beta Coefficients and Hazard Ratios from Model 3 based on standard and additional CVD risk factors with FH. 170
Table 5.19 AIC and BIC values, Royston's R^2 statistic, and Likelihood Ratio Tests to compare Model 1 to Model 2; and Model 3 to Model 4; using the Imputation 1 of the derivation dataset.....175
Table 5.20 Risk equations to calculate 10 year risk based on patient characteristics for each developed risk model (1-4). 176
Table 5.21 Descriptive information on discriminatory performance for each risk model, using a 10% risk cut off.....181
Table 5.22 Harrell's C-Index for each CVD risk prediction model. 182
Table 5.23 Risk reclassification table comparing Model 1 (standard CVD risk factors only) to Model 2 (standard CVD risk factors with FH information). 184
Table 5.24 Risk reclassification table comparing Model 1 with Model 2, by CVD outcome. 185

Table 5.25 Summary of Net Reclassification Index (NRI) and Integrated Discrimination Index (IDI) for all comparisons between model with and without FH information. 186
Table 5.26 Risk reclassification table comparing Model 3 (standard and additional CVD risk factors only) to Model 4 (standard and additional CVD risk factors with FH information).....187
Table 5.27 Risk reclassification table comparing Model 3 with Model 4, by CVD outcome. 188
Table 6.1 Summary of values inputted into the decision tree to compare Model 1 to Model 2, and Model 3 to Model 4. 205
Table 6.2 Output from each decision tree comparing models with FH information to models without. 207
Table 6.3 Results from the deterministic cost-effectiveness analysis evaluating the added value of FH information to CVD risk prediction. 208
Table 6.4 Deterministic probabilities used in the decision tree, for each risk model (with and without FH); with standard error, alpha, and beta for estimation of probabilistic probabilities derived from a Beta distribution.....214
Table 6.5 Results from the probabilistic cost-effectiveness analysis evaluating the added value of FH information to CVD risk prediction. 216
Table 6.6 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information using a 0% discount rate.....224
Table 6.7 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information derived using the complete cases.226
Table 6.8 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information derived using the imputed data with a literature based survival.....228

Table 7.1 List of risk factors used in FRAX and QFracture. 239
Table 7.2 List of study variables, including FRAX® risk factors. 250
Table 7.3 Osteoporosis Primary Prevention Drug Prescribing Proportions (161). 262
Table 7.4 Weighted cost of an osteoporotic fracture. Cost are from Stevenson et al (62), and fracture proportions are from Kanis et al (162). 263
Table 7.5 Fracture event rates from the Cumming et al (163). 264
Table 7.6 Average cost of treated and untreated fracture events over 4 years follow up, by discount rate. 264
Table 7.7 Deterministic mean costs of risk model (with and without BMD), treatment and fracture; with standard error, alpha, and beta for estimation of probabilistic cost derived from a Gamma distribution. 267
Table 8.1 Baseline Characteristics of the overall sample derived from the AURORA database (N=6,117). 273
Table 8.2 Characteristics of the overall sample derived from the AURORA database, by hip DXA based osteoporotic status. 276
Table 8.3 Characteristics of the overall sample derived from the AURORA database, by fracture outcome during the 4 year follow up period. 279
Table 8.4 Crude incidence rates of the fracture outcome in the overall cohort. 281
Table 8.5 Comparison of level of recording in the Danish AURORA and UK CPRD extracted datasets. 284
Table 8.6 Comparison of risk factors and outcome characteristics for both extracted datasets. 285
Table 8.7 Comparison of the crude fracture incidence rates between the AURORA and CPRD extracted datasets. 289

Table 8.8 Baseline characteristics of the derivation and validation datasets, including missing data. 293
Table 8.9 Crude fracture incidence rates for the derivation and validation datasets. 296
Table 8.10 Association between all risk factors to be included in fracture risk prediction models. Values are odd ratios, unless otherwise stated. 301
Table 8.11 Unadjusted Cox regression analysis 304
Table 8.12 Test of proportional hazard assumption from the unadjusted analysis. 306
Table 8.13 Adjusted Cox regression analysis using standard fracture risk factors, excluding BMD measurement, with previous fracture as a time varying covariate. 309
Table 8.14 Adjusted Cox regression analysis using standard fracture risk factors, including binary BMD measurement at the hip, with previous fracture as a time varying covariate. 311
Table 8.15 Adjusted Cox regression analysis using standard fracture risk factors, including continuous BMD measurement at the hip, with previous fracture as a time varying covariate. 313
Table 8.16 Results from data driven models 4a and 4b, based on models 3a and 3b 315
Table 8.17 Results from the proportional hazard assumption test. 317
Table 8.18 AIC and BIC values, Royston's R^2 statistic, and Likelihood Ratio Tests to compare Model 2 to Model 3a; Model 3b, Model 4a, and Model 4b; using the derivation dataset. 321
Table 8.19 Risk equations to calculate 4 year risk based on patient characteristics for each developed risk model (1a-4b) 323
Table 8.20 Descriptive performance related information for each risk model. Values include sensitivity, specificity, positive predicted value (PPV), negative predicted value (NPV) and the % correctly classified, using the 8.5% risk threshold). 330

Table 8.21 Harrell's C-Index for each developed risk model. 332
Table 8.22 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 3a (standard CVD risk factors with BMD measurement), using a clinical 8.5% risk cut off. 335
Table 8.23 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 3a (standard fracture risk factors with BMD measurement), using a clinical 8.5% risk cut off. 336
Table 8.24 Summary of Net Reclassification Index (NRI) and Integrated Discrimination Index (IDI) for all comparisons between developed fracture risk prediction models.....337
Table 8.25 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 3b (standard CVD risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off. 338
Table 8.26 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 3b (standard fracture risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off. 339
Table 8.27 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 4b (data driven standard CVD risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off. 341
Table 8.28 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 4b (data driven standard fracture risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off. 342
Table 9.1 Summary of values inputted into the decision tree to compare Model 2 to Models 3a, 3b, 4a, and 4b.....353
Table 9.2 Output from each decision tree comparing models with BMD measurement to models without. 355

Table 9.3 Results from the deterministic cost-effectiveness analysis evaluating the added value of BMD information to fracture risk prediction. 357
Table 9.4 Deterministic probabilities used in the decision tree, for each risk model (with and without FH); with standard error, alpha, and beta for estimation of probabilistic probabilities derived from a Beta distribution.....365
Table 9.5 Results from the probabilistic cost-effectiveness analysis evaluating the added value of BMD measurement to fracture risk prediction.....367
Table 9.6 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without BMD measurement using a 0% discount rate. 378

List of Abbreviations

AIC	Akaike Information Criteria
AUC	Area Under the Curve
AURORA	Aalborg University Hospital Record for Osteoporosis Risk Assessment
BIC	Bayesian Information Criteria
BMD	Bone Mineral Density
BMI	Body Mass Index
BNF	British National Formulary
CEAC	Cost Effectiveness Acceptability
CKD	Chronic Kidney Disease
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular Disease
DBP	Diastolic Blood Pressure
DCA	Decision Curve Analysis
Egfr	Estimated Glomerular Filtration Rate
FH	Family History
GP	General Practitioner
HbA1c	Glycated Haemoglobin

HDL	High-Density Lipoprotein
HES	Hospital Episode Statistics
HR	Hazard Ratio
HSE	Health Survey for England
ICD	International Statistical Classification of Diseases
ICER	Incremental Cost Effectiveness Ratio
IDI	Integrated Discrimination Index
IFG	Impaired Fasting Glucose
IGR	Impaired Glucose Regulation
IGT	Impaired Glucose Tolerance
IMD	Index of Multiple Deprivation
INB	Incremental Net Benefit
LDL	Low-Density Lipoprotein `
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NRI	Net Reclassification Index
ONS	Office for National Statistics
OR	Odds Ratio
PH	Proportional Hazard

PSSRU	Personal Social Service Research Unit
PY	Person Years
QALY	Quality Adjusted Life Year
QOF	Quality of Outcomes Framework
RA	Rheumatoid Arthritis
ROC	Receiver Operating Characteristic
SBP	Systolic Blood Pressure
SES	Socio-economic Status
THIN	The Health Improvement Network
TVC	Time-Varying Covariate

Chapter 1 - Introduction

The focus of this thesis is to evaluate and refine the methodology currently being used to assess the added value of a new risk factor to an established risk prediction model.

This chapter will provide a brief introduction to risk prediction models, new risk factors, and current methodology used to evaluate the added value of a new risk factor to a risk prediction model; whilst giving the rationale for this research.

1.1 Rationale

1.1.1 Risk Prediction Models

Risk prediction models are a key component in risk assessment in primary care (1-5). Risk prediction models are used as a targeted treatment strategy to identify patients 'at risk' of disease or a health event. These patients can then be provided with preventative treatment reducing the risk of disease and preventing the event from taking place (3, 6). Over time, this strategy has helped reduce disease burden, in terms of both mortality and morbidity, and contributed to cost savings (1, 3, 7); and remains a primary strategy to reduce risk.

However, the strengths of risk prediction models are dependent on how well and accurately the risk model predicts. The advantages of risk prediction models would be substantially reduced if they were poor performing, and were identifying patients 'at risk' incorrectly. This would result in increased missed diagnoses and over treatment. It is therefore essential to use validated risk prediction models, preferably recommended by national guidance, where their performance has been assessed and are considered worthy of being used in clinical practice.

1.1.2 Improving Risk Prediction Models – New Risk Factors

Improving future health outcomes, and avoiding use of poorly validated and out-dated risk prediction models, have been the primary reasons for updating risk models with emerging, and new risk factors (8, 9). As disease prevalence and the demographics of the target population changes, risk prediction models also need to be updated to prevent poor performance in the changing population. Adding new risk factors is a common way to help improve risk prediction by keeping models up to date and away from poor performance.

Incorporating new risk factors into established risk prediction models will have one of two effects; they will either improve risk prediction and hence, future health outcomes (add value); or worsen them (add no value). This needs to be assessed and firmly established before they are included into routine risk assessment used in clinical practice.

Lack of adequate assessment of the added value of new risk factors, has led to the development of many risk models, containing various risk factors for their respective diseases. Thus, more than one tool can be found for a single disease (8, 10-13); for example, cardiovascular disease (CVD) has many risk prediction models to calculate 10 year risk of CVD. Of these, three are used in a UK primary care setting; the National Institute for Health and Care Excellence currently recommends QRISK2 (3), The Joint British Societies for the prevention of cardiovascular disease uses a Framingham based risk model (14); and the SIGN guidance recommend using ASSIGN (15). Further, comparisons between differing risk models has been difficult due to inconsistency with reporting (16), which has led to confusion over which risk prediction model to use (17).

In addition, known risk factors for a disease are sometimes not scientifically embedded in the current risk model; but are used to alter risk estimates formally calculated from the

risk model. Consensus driven inflation factors have been used to account for the added value of an additional risk factor (17, 18). However, this uses the added value of a new risk factor independently from the standard risk factors included in the tool; and may over estimate risk of the disease. The added value of a new risk factor needs to be evaluated in addition to these standard risk factors for its true effect in a clinical setting.

1.1.3 Methodology to Evaluate the Added Value of a New Risk Factor

Methodologies to validate risk prediction models are used to evaluate the added value of new risk factors; these will be discussed in detail in Chapter 2. Briefly, the performance of a risk prediction model is based on its predictive ability and accuracy; this is assessed using the observed and predicted risk. Depending on how well the predicted and observed risk match shows how well the risk model is predicting. Further, pre-defined risk thresholds can be used to identify the proportion of patients the risk model classifies as 'at risk', correctly or incorrectly depending on the observed outcome. The changes in these measures are used to evaluate the added value of the new risk factor.

However, current validation methodology has shown to be limited for evaluating the added value of a new risk factor (19-22). New methodologies, specifically looking at the clinical improvement when comparing models with and without a new risk factor have been developed (23, 24). However, they also do not provide full evaluation of the effects of adding a new risk model, namely a cost evaluation.

The benefits of a risk prediction models are measured by the costs saved through prevention of the tool outcome. The efficacy of various preventative treatment and drug therapies has been evaluated with regards to expected costs saved (25, 26). Further, the evaluation of cost-effectiveness has been identified as the part of the criteria for assessing the added value of a new risk factor (27). Yet, the added cost of a new risk factor and its

effect on the cost of treatment is not accounted for in current methodology allowing for a full evaluation of the effects of a new risk factor.

Incorporating this element into methodology would provide a better approach of assessing the added value of a new risk fracture to an established risk prediction model. Firmly determining the added value and expected consequences of a new risk factor would then inform better decision making for national guidance and bring a standardised approach to using risk prediction models in clinical practice. New risk factors could be listed in order of their cost effectiveness in a League table style approach, which would help when recommending a single risk prediction model for routine risk assessment.

1.2 Aims and Objectives

1.2.1 Aim and Objectives

The aim of this thesis is to refine the methodology of assessing the incremental value of a new factor; to include a cost-effectiveness analysis of the new risk factor. This will be achieved by the following objectives:

1. Collate information of current methods used to assess the incremental value of novel risk factors. Evaluate these with respect to their application, strengths, and weaknesses.

2. Refine the identified methods to fully assess the clinical consequence when adding a new risk factor to an established risk prediction model; encompassing evaluation of statistical and clinical improvement, with cost implications.

3. Apply the refined methodology to two common exemplar conditions using generalisable data:

a. Assess the added value of family history of cardiovascular disease to cardiovascular disease risk prediction.

b. Assess the added value of bone mineral density to fracture risk prediction.

1.3 Structure of thesis

The aim and objectives of the thesis will be approached in the following structure:

- Objective 1 is covered in Chapter 2 which provides a literature review of the current methodology used to evaluate the incremental value of a new risk factor.

- Objective 2 is covered in Chapter 3 which describes the refined methodology to evaluate the added value of a new risk factor; combining statistical improvement, clinical improvement, with cost implications.

- The first part of Objective 3 is covered in Chapters 4, 5, and 6. Chapter 4 presents the introduction and methods for assessing the added value of family history information to cardiovascular disease risk prediction. Chapter 5 presents and discusses the results from the current methodology. Chapter 6 presents and discusses the results from the refined methodology proposed in Chapter 3.

- The second part of Objective 3 is covered in Chapters 7, 8, and 9. Chapter 7 presents the introduction and methods for assessing the added value of bone mineral density measurement to fracture risk prediction. Chapter 8 presents and discusses the results from the current methodology. Chapter 9 presents and discusses the results from the refined methodology proposed in Chapter 3.

- Chapter 10 summarises and discusses the findings of the thesis in its entirety.

1.4 Chapter Summary

This chapter has outlined the aim and objectives of this thesis (Section 1.2), and has also provided the structure of this work (Section 1.3). The rationale to assess the added value of a new risk factor is based on the importance of risk prediction models in clinical practice, by their potential effect in reducing future disease burden, which can be furthered by incorporating new and emerging risk factors. The methodology to assess the incremental value of the new risk factors has fallen short of providing a full evaluation of the new marker. Namely, it is lacking a cost evaluation component. Chapter 2 will expand on the current methodology used, evaluating its strengths and weaknesses in more detail.

Chapter 2 - Literature Review

The previous chapter provided an introduction into risk prediction models and gave the motivation for adding new risk factors to them. This provided an overview of the current issues faced when evaluating the added value of a new risk factor to an established risk prediction model. Based on the rationale given in the introduction, aims and objectives were set to answer the research question of how to evaluate the added value of a new risk factor.

The objective of this review is to collate and evaluate current methodology being used to assess the added value of new risk factors to existing risk prediction models. As well as providing an outline of these methods and metrics, this review will also highlight their strengths and limitations; highlighting the current issues faced and areas needing additional work.

2.1 Objective and search strategy

The review was conducted using MEDLINE to search for relevant papers. A full list of the search terms used to derive the papers can be found in Appendix A. In addition, MeSH titles from key papers were also searched for, such as “ROC Curve”, “Risk Assessment”, “Risk Factor”, “Biological Marker”, “Prognosis”, “Models, Statistical”, “Calibration”, “Area Under Curve”, “Predictive Value of Tests”, and “Sensitivity and Specificity”. References and citations of key papers were also searched for.

2.2 Risk Prediction Models

Risk prediction models are statistical models in which we derive risk of a future outcome, based on certain pre-selected risk factors. These risk factors can range from demographic information to test results and are included in the risk model based on

their association with the respective outcome. An outcome for a risk model includes disease, and health events (e.g. cardiovascular disease and fractures).

Risk models are of particular interest as they attempt to foresee the disease before it has occurred; and assign the appropriate risk to patients who may or may not develop the disease at some point in the time frame. They have many uses. They can be used to select patients suitable for randomised controlled trials (28); to improve the cost effectiveness of an intervention for disease prevention (29); to assist healthcare decision making (18); and assist the move toward personalised medicine (30).

The concept of risk prediction models, especially to assist in healthcare decision making, is extremely attractive and desirable for healthcare decision makers, General Practitioners (GPs), and clinicians. Early disease detection and prevention of disease is high priority for national guideline makers (3); the implementation of which, is showing reduced disease burden and cost (26).

Specifically, this has been achieved by using risk prediction models to derive patient risk of a disease given their profile to the risk factors of the model. In many cases, these risk models will allow for the patient to be placed in their appropriate risk group for the disease which often has a preventative treatment allocated to it (3, 6). For example, the QRISK2 cardiovascular (CVD) risk model is currently being recommended as part of national guidance (3), alongside a 10% risk threshold to identify and treat high risk patients of CVD using lipid modification drug therapy.

In light of these benefits of using risk prediction models for healthcare decision making, there is a constant effort to improve risk prediction; namely through the addition of new and emerging risk factors (8, 9, 31-33). Further motive for adding new risk factors and improving prediction is to reduce the burden and harm from poor

performing and outdated risk models, reducing over treatment of patients and missed diagnoses. A balance however, has been difficult to achieve.

2.3 New Risk Factors

Various new risk factors are emerging to improve risk prediction models. In particular, recent developments in genetic research, association with emerging genetic risk factors, and a general move toward personalised patient treatment and care have further promoted the improvement of risk models through incorporating new risk factors (14, 33, 34).

The aim of incorporating new risk factors into risk prediction is to improve the performance and accuracy of a risk prediction model, by a method of updating them. The measures used to describe the performance and accuracy for risk models are discussed in more detail below. Improving risk prediction would allow for further gains to be made by further reducing over treatment and missed diagnoses; thus reducing of future disease burden and cost.

However, it can be argued that just as evolving research recommends the addition of a new risk factor to update risk prediction and improve its benefits; there is also the option to replace an existing risk factor with a new one (27, 35), or to simply remove a risk factor that is now not providing any noticeable benefits. These avenues are rarely explored; highlighting that current methodology is not equipped to present information which would promote such a decision.

A cost effectiveness approach would account for this. It would evaluate the effects of adding, or removing a risk factor, and put it in relation to an expected cost. A new risk factor may cost less to measure, be more routinely collected, better measured, or all of

the above, than an already established risk factor; in such a case it may be worth considering replacing the established risk factor with a new one.

However, the attraction to the originality of developing a new and improved risk prediction model, with new risk factors, often outweighs the possibility of updating an established risk model. This has led to many risk models being developed for single diseases (8-12, 31, 36) and thus, difficulties in comparing them.

2.4 Comparing Risk Prediction Models

It is difficult to compare risk prediction models and identify superior models. This was illustrated by The National Institute for Health and Care Excellence (NICE) Lipid Modification: Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease guidance (17). In 2010, the 2008 version of the guidance was re-issued, changing the recommendations of solely using the Framingham risk prediction model, to using either the Framingham risk model or QRISK2 to calculate the 10 year risk of CVD in clinical practice. At this time, NICE also identified 'an urgent need to establish which score is most acceptable for use in the population of England and Wales'. This was due to the more than one model being available and the difficulties in determining which is better to use, highlighting the lack of certainty and confidence in this area. Since, the new guidance now recommends solely using the QRISK2 model (3).

Further, a report by the PHG Foundation (30), emphasised the lack of structure within the assessment of risk models, the difficulties as to which performance measure to use, under which circumstances, and how to interpret the changes in these performance measures to determine which tool ultimately performs better. Thus, an issue lies in the current methodology used to compare risk prediction models and

determine what a significant difference in model performance is, when adding a new risk factor to an established model (34).

2.5 Current Methodology

2.5.1 Guidelines

Recommendations were found for validating risk prediction models (37), and also for evaluating a new risk factor (20, 27). In particular, Hlatky et al (27) recommend using a cost-effectiveness analysis to identify the added gain in effect of the risk model in relation to the increased cost.

However, implementation has been poor. Further, reporting guidelines have only been recently developed (38) and have yet to standardise reporting for prediction models. A systematic review (39) comparing established risk prediction models for CVD showed comparing risk models was difficult due to lack of standardisation of reporting, limitations in methodology, and inconsistent use of performance measures.

Overall there is a need for a stronger methodological foundation and a systematic approach for assessing risk models, and hence the added value of a new risk factor (27, 34, 40); making reporting more consistent and methodology more informative.

Investigating the added value of a new risk factor in an external validation study, using a prospective cohort is a favoured approach (41). As the population in an external validation would differ from the sample used to develop the risk model, it provides less biased results of performance. Further, the sample would be followed up over time (42). Such a design also gives the option to have a longer follow up, which would be able to capture more ‘events’ in the study; allowing for more robust

assessment if the new risk factor is associated, with both long and short term risk of the outcome.

Many of the prerequisites required from the data and guidance given when validating a risk prediction model also apply when evaluating the added value of a new risk factor: to follow the standards of clinical research and to be protocol driven. It is advised that an inclusion and exclusion criteria for the research question be clearly defined; results and measurement be reproducible and comparable to the original model; and hypotheses be clearly stated. It would be especially helpful for authors to have knowledge as to what would be considered as good model predictive accuracy and base their hypotheses around previous estimates in their field. This would help conclude whether a significant improvement in the risk prediction model has been found with the addition of a new risk factor or not and whether their ambitions to improve a risk prediction model are realistic (42).

2.5.2 Statistical Validation

The initial requirement for a new risk factor is to have a statistically significant association with the outcome (27). This can be shown using basic statistical tests; and the magnitude of the effect can be shown via parameter estimates of the risk prediction model. However, on finding an association with statistically significant p-values or large parameter estimates does not imply clinical significance, improvement of the model, or evidence of incremental value (20), as such associations can be found easily if the sample size is large enough (43). The new risk factor must maintain this association after the existing standard risk factors are accounted for; thus giving the added value of the new risk factor.

The following measures can then be used to evaluate the improvement of the initial risk model when adding a new risk factor.

2.5.2.1 Global measures

Firstly, global measures of model improvement, such as the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) can be used to assess the incremental value of a new risk factor. Traditionally used as a stopping rule for developing a risk model, these measures are used to decide when to add a risk factor to a model. Both measures assess the fit of the model on the data, but penalize for complexity; the number of risk factors included in the model. Further detail regarding the application of these two measures can be found elsewhere (44).

These measures, however, are limited in the information they provide and should not solely be relied on; they provide no more than a statistical test of association between the new risk factor and the outcome (19). In addition, the Likelihood Ratio Test can also be used, but again gives no direct assessment of predictive accuracy (27).

2.5.2.2 Calibration and Discrimination

2.5.2.2.1 Calibration

Calibration is measure used to describe risk prediction models and how well the predicted risk matches the observed risk in the sample. It is measured by comparing the predicted frequency of events with the observed frequency, and is normally graphically presented, with a 45 degree line showing perfect calibration. A Hosmer-Lemeshow (45) test is also used to assess the difference between the predicted and observed events; its output is given as a p-value, where a higher p-value would indicate a better calibration of the model. This measure can only be reported when logistic regression has been used to develop the risk prediction models.

Calibration is key measure to describe the how well a risk model performs; however, it is less reported than discrimination measures (described below) (23, 46). Further, this measure has also been criticised due to its lack of interpretability, limited power; dependency on sample size; and sensitivity to how subgroups are formed (i.e. whether they are formed using equal increment of risk for deciles of risk) (19, 47, 48).

2.5.2.2.2 Discrimination

Discrimination is the most used and reported measure to assess predictive accuracy of a model. It is evaluated using a Receiver Operating Characteristic (ROC) Curve analysis (49), which plots sensitivity versus 1 minus specificity for every possible cut off value of the model.

Sensitivity is the proportion of true positives identified by the risk prediction model (proportion of correctly identified diseased patients) and specificity is the proportion of true negatives identified from the risk prediction model (proportion of incorrectly identifies diseased patients). This illustration helps define the risk threshold value (the maximum range of thresholds for declaring a person likely to get the disease) for the model; the point at which the sensitivity (true positive rate) of the test is maximised and, 1 minus specificity (false positive) is minimised. There is a trade-off between these values to derive the best cut off for the risk prediction model scores.

The most common measures calculated from a ROC curve analysis is the area under the curve (AUC), which is also presented as a C-statistic, Somers' D-statistic, and Harrell's C-Index; depending on the regression model being developed. For example, a risk model developed using survival analysis and Cox regression will use Harrell's C-Index to describe the AUC.

A larger AUC implies greater discriminatory accuracy of the risk model, which describes the ability of the risk prediction model to separate individuals who develop the outcome from individuals who do not. Specifically, the AUC is interpreted as the probability that out of two people, who respectively have, and have not observed the model outcome, the risk model will assign a higher risk estimate to the patient with the outcome. This is a well-known and popular method and the change is AUC is often reported to portray an improvement in a risk prediction model (39, 43).

However, the AUC value is difficult to interpret (21), and is often misunderstood for being a measure of predictive accuracy (50). It should be interpreted with further caution, as it presents the probability a higher risk is assigned to the individual with the observed outcome; which is not the probability that they are high risk, based on a pre-defined risk threshold. People may still be classed as low risk and be a missed diagnosis as a result of the risk prediction model.

Although the AUC is well known and is simple to calculate, small improvements with limited clinical meaning are often reported when evaluating the added value of a new risk factor (19, 21, 30). Further, it has proven to be difficult to show a substantial improvement in the model performance with the addition of a new risk factor using this metric, which is insensitive to change (19, 30, 47, 51, 52). This is shown in Example 2.1, which reports a 0.0011 absolute increase in the C-Index, when adding family history and high sensitivity C-reactive protein to an established CVD risk prediction model, as two new risk factors.

Pepe et al (40) show that even in the presence of large parameter estimates for the new marker (ranging from 2 to 3 times higher risk/odds of disease), significant improvement in the AUC of the model is hard to achieve.

Example 2.1:

Paper: C-Reactive Protein and Parental History Improve Global Cardiovascular Risk Prediction: The Reynolds Risk Score for Men. Ridker et al. [12]

End point/Measure: CHD

Traditional Model (M1): age; BP; smoking status; total cholesterol; HDL cholesterol

Model 2 (M2) Novel Markers: high sensitivity C-reactive protein (hsCRP); and parental history of Myocardial Infarction before age 60yrs.

Results:

Change in Bayes Information Criteria: 19157 (M2) – 19196 (M1) = -39

Change in C-Index: 0.700 (M2) – 0.689 (M1) = 0.011

17.8% of patients were reclassified using M2

20.2% of patients with risk between 5% and 20% were reclassified using M2.

Net Reclassification Index = 6.8%

Clinical Net Reclassification Index = 13.6%

Conclusion: A prediction model incorporating hsCRP and parental history significantly improves global cardiovascular risk prediction.

Siontis et al (39) compared established RPMs for CVD in a systematic review and found that only 18% of the studies included had a 5% relative difference between the AUC values. Thus, the question 'When does a difference make a difference?' has been posed (34).

It can be argued that a large proportion of the discriminatory accuracy is accounted for by the existing risk factors; age and gender are key risk factors which dominate many risk models (53, 54). Further, when a risk model is developed, it is with the aim that the final

model cannot be further improved upon (42, 55). Thus, the expectation from new risk factors for further increase performance may be unrealistic.

In this case, it would be worth looking to the methodology used to identify improvements in the performance of a risk prediction model, and put these values in clinical context (56). Evaluating costs in relation to the statistical improvement of the risk model would cater for this, allowing for better decision making.

Currently discrimination is measured by two components; sensitivity and specificity. With regard to improvement in discrimination, this is defined by increased sensitivity and specificity; however, this does not account for the possibility that it might be clinically, more effective to have a lower specificity and a higher false positive rate. For example, the CVD risk threshold has been lowered from 20% to 10% as research has shown it was both effective and cost effective to treat more patients, correctly or incorrectly (3).

Lipid treatment has shown to reduce the CVD event rate in both high and low risk patients (57, 58); thus over treatment provides benefit to the risk model. This would not have been known based on discriminatory measures alone; cost data with treatment efficacy data also need to be accounted for when assessing the added value of a new risk factor to identify the statistical improvement needed to truly improve the risk prediction model.

Further, discrimination and calibration measures are disjointed. Hence, a model with good discrimination can be poorly calibrated and vice versa. Therefore, when assessing performance of risk models, there is a trade-off between these measures (23); which is not ideal. This may contribute to the under reporting of calibration measures compared to measures of discrimination, in validation studies (39).

Despite these limitations, the AUC has become a standard measure to illustrate the effect of a new risk factor and is considered a good starting point for any analysis, it should not be forgone in light of new methods (20, 27, 47, 55). On attaining good discrimination and calibration in a risk model, it is still considered for use in clinical practice.

2.5.3 Clinical Validation

2.5.3.1 Reclassification Tables

In light of the above flaws, especially those regarding interpretability and clinical relevance, Cook et al proposed using Reclassification tables (23). They suggest that, if it can be shown that a model with a new risk factor can more accurately categorise patients into higher or lower risk groups, the model will have improved.

However, performing a reclassification analysis, needs to go beyond simple measures of total percentage of patients reclassified or the percentage of patients in the intermediate risk group reclassified to the lower or higher risk group (19, 43). Specifically, the movement of patients between risk categories may be correct or incorrect based on the observed outcome of the patient; this is not accounted for. Further, consideration also needs to be given to patients who are not reclassified.

Pencina et al (43) build on this method by introducing the Net Reclassification Index (NRI), which examines events and non-events separately. It considers upward reclassification amongst patients with the outcome and downward reclassification amongst patients without the outcome, to be beneficial with respect to the performance of the new model. Only patients who move between categories are included in the analysis; it is assumed for patients who do not move, the models perform the same.

Pencina et al also introduced the Integrated Discrimination Improvement (IDI) (43) which tests the extent to which the new risk factor correctly revises the predicted risk of patients who experience an event, upward; and the predicted risk of patients who do not, downward; eliminating the limitation and need of a pre-defined risk threshold.

The reclassification method is geared towards providing clinical understanding of the consequences of a change to a model compared with the current analysis. Advantages of using this method are that it incorporates the assessment of both discrimination and calibration (55); uses simple calculations; and that its basic results (event and non-event NRIs) are easy to interpret and can be used to visualise the impact of the new risk factor. Further, it has shown to give larger increments in the performance of a risk model compared the changes in the AUC, as shown in Example 2.1. Collectively, these reasons have influenced growth and popularity of this method; Tzoulaki et al (46) reported that out of 51 papers only 53% reported calibration, but 75% calculated the NRI from a reclassification table.

The key question Pencina et al (43) pose is how best to assess and quantify improvement in risk prediction offered by new risk factors. Statistical significance is not enough for this as it doesn't tackle the issue of clinical significance. However, although they introduce the NRI measure for this, they state that the AUC should still remain the first criterion when assessing a novel risk marker. This is reinforced by Janssens et al (34), who state that a reclassification analysis should follow the ROC curve analysis and is only warranted if an improvement is found in the AUC.

Within a reclassification analysis, the NRI is the most popular metric. It compares the classifications from two models for changes in the outcome, for a net calculation of movement in the 'right' direction. Using the NRI to quantify improvements in risk

prediction does however, have some limitations. The drawbacks of this method are that it is sensitive to chosen risk thresholds; in the number used, and the cut-off values for the risk categories (19, 59). The NRI has shown to increase as the number of risk categories increase (60).

Further, the NRI is often misinterpreted and misleading, and needs to be approached with caution. The NRI is the addition of the event and non-event NRIs, but the denominator for these values differ. Thus, the NRI does not represent a proportion as reported. It has a theoretical value range from -2 to 2, but its unit is yet to be determined (61).

Although the method presents more clinically relevant results with regards to the shift of patients between risk categories due to the new risk factor; this method is further limited as it does not incorporate cost data in relation to the identified benefits. Thus it does not fully incorporate the clinical consequences from the new risk factor.

2.5.3.2 Decision Curve Analysis

Another recently developed method which is used to evaluate the clinical usefulness of a new risk factor is a Decision Curve Analysis (24). The motivation for this method has also originated from the limitations of current measures, described above; in particular the lack of information of consequences.

The Decision Curve is plotted using the net benefit for each threshold of probability (similar to the cut off value in a ROC curve analysis); the net benefit is the net number of true positives gained by using a model with and without the new risk factor, without an increment in the false negatives.

Advantages of this method are that numerous risk models can be compared, and increased performance can be identified by an increase in the net benefit. Further, it claims to

encompass a fuller evaluation of risk prediction models and new risk factors with a decision analytical framework but without the need for additional cost data. However, a full evaluation on the effect of a new risk factor cannot be performed without cost data, and this cannot be substituted for. Hence, it can be argued that a DCA only partially assess the clinical usefulness and consequences of a new risk factor; and cost is yet to be explored.

Sourcing additional cost data has been identified as a limiting factor for not performing a full evaluation of the added value of a new risk factor. However, it should be recognised that publication of cost data has made it more available for use. Treatment costs of drug therapies can be found in the British National Formulary (BNF), costs for healthcare can be found using the Personal Social Services Research Unit (PSSRU) ; cost of diseases and events can be found in publications (62) and NICE costing reports (25, 26). Further, NICE costing reports can also be used for assumptions made when using a risk prediction model.

Therefore, sourcing additional cost information should not be viewed as a limitation anymore; and these sources, along with the scope of incorporating costs into the evaluation of the added value of a new risk factor, should not be overlooked. A cost-effectiveness analysis should be explored for providing a fuller and better evaluation of a new risk factor.

2.6 Conclusion

Risk prediction models have the potential to reduce disease burden and save costs by identifying patients ‘at risk’ of disease; and who would most benefit from preventative treatment. Adding new risk factors is a method to update and improve the performance of

a risk model, with respect to predictive and discriminatory accuracy. However, it is difficult to assess the value a new risk factor adds to an established risk prediction model.

Current methods to assess the added value of a new risk factor are limited. New methods have been developed but have failed to incorporate cost data; and thus fall short of providing a full analysis of the added value of a new risk factor. Limitations of sourcing and including cost data in current analyses no longer hold; and research to evaluate the cost impact as well and the statistically and clinical effectiveness of a new risk factor is viable.

2.7 Chapter Summary

This chapter reviewed the current methodology used to evaluate the added value of a new risk factor. It identified the common measures used to describe the improvement of a risk model when adding a new risk factor; providing an outline of the method with its strengths and weaknesses. Assessing the impact of the new risk factor in terms of cost has not been evaluated; the proposed methodology to incorporate this data in relation to the added effect is described in Chapter 3.

Chapter 3 - Refined Methodology

The aim of this chapter is to investigate and identify the relevant health economic methods available to evaluate the added value of a new risk factor to an established risk prediction model.

The first objective of this chapter is to summarise the rationale behind proposing a new methodology which incorporates costs, as well as measures of effect for evaluating new risk factors and their added value to established risk prediction models. This will be based on what was found in the literature review (Chapter 2).

The second objective will be to give a brief outline of current health economic methods; and describe the proposed cost-effectiveness analysis and decision tree framework to evaluate the added value of a new risk factor to an established risk prediction model. This will be described in line with current reporting standards and checklists for robustness.

3.1 Rationale

Chapter 2 presented the current methodology used for the evaluation of new risk factors in addition to the established standard risk factors for the disease outcome. They have shown current methodology to be inadequate in providing a full evaluation of the added value of a risk factor.

Current performance measures such as the AUC do not effectively compare risk models or describe the added value of a new risk marker; and provide no measure of the clinical impact a new risk factor may cause. Using these measures of effect alone has led to lack of effective improvement in model performance.

These limitations indicate the need for consideration and incorporation of cost as well as effects to assess the cost effectiveness of a new risk factor. By bridging current measures of model performance to clinical costs would provide more clinically relevant and interpretable estimates; allowing for better translation of results, improvement, and clinical impact. Therefore, depending on the improvement, a better informed decision can be made regarding the addition of a new risk factor to an established risk prediction model; to be used in routine risk assessment in primary care.

Decision analytical methods have been adopted in current methodology to provide more clinically relevant information. Decision Curve Analysis (DCA) was developed to provide more clinically relevant results about risk models and their potential improvement. However, like other methods, it fails to include and present information on the impact of costs the decision maker might expect. Therefore, a gap still remains in the methodology to account for costs resulting from using risk prediction models and hence to effectively evaluate if a new risk factor is cost effective or not, for routine risk assessment.

Economic evaluation in healthcare offers methods to account for the improvement in effects (statistical and clinical) of a risk model and associate this effect to a cost; giving more clinically relevant and interpretable estimates for the added value of the new risk factor.

3.2 Economic Evaluation in Health Care

Economic evaluation in health care is defined as the comparison of alternative options in terms of costs and consequences (63). By attaching costs to the consequence of any option, it immediately puts it into clinical context and adds value to consequence information alone; this is not currently addressed in studies evaluating the added value of a new risk factor.

The alternative options which are compared in a health economic evaluation tend to compare a new intervention to the standard practice. These can range from comparing different drugs, new technologies, and programmes or strategies. Hence, health economic evaluations can also be used to compare an established risk prediction model to a modified one (with a new risk factor); providing a more informative performance measure.

3.2.1 Analyses

Within health economic evaluations four different types of analyses can be used to measure costs and consequences. These are: a cost-minimisation analysis; cost-effectiveness analysis; cost-utility analysis; and cost benefit analysis. All of which involve a comparative analysis of different courses of action.

A cost-minimisation analysis is used typically when the consequences of the two comparative practices are assumed to be equal. This is a rare occurrence; and it is expected that both the costs and consequences will vary when adding a new risk factor to an established risk prediction model. Hence, this analysis would not be suitable to compare risk prediction models.

A cost-effectiveness analysis is a popular and better choice for providing a full economic evaluation, as treatment costs and consequences are evaluated together. Here, costs are attached to a single and common effect used to describe the difference between models; and is presented as the cost per unit of effect.

This analysis uses measures of effect on their natural scale, such as life years gained for patients; however, it can also be linked to other measures of effect for the patient, such as additional cases detected. This gives some flexibility in the choice of effect measures of interest; providing an optimal setting to assess the added value of a new risk factor, as the

effectiveness measure can be altered to a given risk prediction model performance measure.

A cost-utility analysis is a subgroup of a cost-effectiveness analysis; which also accounts for cost and consequences of each possible decision, but uses a preference based utility measure for the outcome. This analysis uses utility measures of effect on a generic scale. A utility measure is based on the preference a person has to be in a particular health state; the more they prefer their health state, the higher their utility. For example, two individuals may have the same outcome, but how they see this outcome and how it affects them may differ. The utility of an outcome in a cost utility analysis is commonly measured using quality adjusted life years (QALYs) gained, which adjusts the life year gained by individual preference of being in their health state. A QALY is typically derived using the EQ-5D/SF-60/Health Utilities Index.

A cost-benefit analysis is also similar to a cost-effective and cost-utility analysis, where consequences are valued in monetary terms and compared. However, two approaches can be used; a Welfarist approach which values the benefits of treatments with respect to the individuals' willingness to pay rather than against a fixed budget; or Grossman's Human Capital approach (64), which requires putting a value of human life which can be considered unethical. Further, a cost benefit analysis has a wider perspective which looks at benefits from a societal point of view, and requires more data collection. This is again unsuitable for evaluating the added value of a new risk factor.

3.2.1.1 Cost-Effectiveness Analysis for the Economic Evaluation of an Additional Risk Factor

Of the four possible analyses available to compare two risk prediction models and provide a full economic evaluation for a new risk factor, the cost-effectiveness and cost-utility

analyses are most appropriate. Both methods offer results considering the NHS and PSSRU perspective (65), and associate costs to the consequences in the results, which are presented as a single unit of effect.

Although both analyses are similar, and costs are evaluated the same way; differences occur in the measure of consequences. A cost-effectiveness analysis produces outcomes that are singular and programme specific (in this case disease specific); they are outcomes on their natural scale. It presents this information in the form of the cost per unit of effect; which could be used to describe the cost per increment in performance measure for two risk prediction models.

The cost-utility analysis, however, produces outcomes which are single or multiple, and not programme specific; they are outcomes on a generic scale, which are adjusted for individual preference for the outcome. This allows for more comparability across studies (and disease areas). This analysis also focuses on the quality of the health outcome/state, as well as the quantity; usually measured by quality adjusted life years (QALYs).

Overall, the choice between using a cost-effectiveness analysis or cost utility analysis has meant deciding between using a singular measure of effect regarding the quantity of the outcome in a cost-effectiveness analysis, or utility based effect measure including assessment of both the quantity and quality, to give the preference of the outcome in a cost utility analysis.

To assess the added value of a risk factor, an economical evaluation will be performed to compare performance of two risk prediction models; with and without the new risk factor. Thus, results will be comparable within disease types, dictated by the risk model of interest. The difference in cost and effect will then be used to value the new risk factor. For this, the primary interest is to compare popular performance measures, such as

discriminatory accuracy, on their natural scale. Preference based measures such as utilities, would be inappropriate to use for this.

The cost-effectiveness analysis has shown to be more applicable for the given scenario of comparing risk prediction models, than a cost-utility analysis. Further, it would not be limited in relation to comparability between diseases, as the same measures of performance can be used to evaluate the risk model. Using the cost-effectiveness method, for which the outcome attaches cost to a single effect, also allows variety in the measures of effect in the cost-effectiveness analysis. Once the expected cost for each risk model is derived, it can be attached to more than one measure of effect; or in this case, more than one risk model performance measure.

Results of the cost-effectiveness analysis will act as a bridge between the statistical improvement of risk models and their clinical impact; through assigning a cost to each unit of improvement. For example, taking the clinically irrelevant AUC value as a descriptive of model performance and measure of effect, the cost-effectiveness analysis would give the more clinically relevant result; the cost per increment in AUC. Hence, this proposed method is recommended to be used in addition to current methodology, as it uses this information to provide the cost-effectiveness overall measure of effect.

3.2.2 Decision Analytical Framework

The approaches described above show the various analysis techniques for comparing alternative options. In the case of comparing risk prediction models, with and without a new risk factor, the cost-effectiveness analysis is most appropriate.

Decision analytical modelling provides a framework for the relevant information, and is a method to present data for the economical evaluation. It uses a systematic approach to create a structure for the evidence and displays the cost and effects for each possible

option of the comparative analysis; where the structure reflects the choice and its consequences. This can then be used in conjunction with a cost-effectiveness analysis.

Using a decision analytical model works to identify an appropriate course of action from the specified alternative options (using a risk prediction model with and without the new risk factor). Decisions are based on expected values derived from the costs and probabilities presented in the decision analytical model. Hence, simple mathematical relations of probability (Box 3.1) can be used and applied to calculate the likelihood of each possible consequence from the decision model; and hence the expected cost and expected consequence for each option.

Box 3.1 Probability and its mathematical relations

Let A and B be events, where $P(A)$ = probability of event A occurring

Then:

$$\boldsymbol{P(A)} = \frac{number\ of\ times\ event\ A\ occured}{total\ number\ of\ events\ occured}$$

Using this information, the following definitions hold:

$$\boldsymbol{Joint\ Probability}: P(A\ and\ B) = P(A \cap B) = P(A) \times P(B)\ if\ events\ A\ and\ B\ are\ independent$$

$$\boldsymbol{Conditional\ Probability}: P(A\ given\ B) = P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$Events\ A\ and\ B\ are\ \boldsymbol{independent}\ if\ P(A) = P(A|B), that\ is, P(A)\ is\ not\ affected\ by\ event\ B$$

$$Events\ A\ and\ B\ are\ \boldsymbol{mutually\ exclusive}\ if\ P(A \cap B) = 0, hence\ P(A \cup B) = P(A) + P(B)$$

$$that\ is, events\ A\ and\ B\ cannot\ occur\ at\ the\ same\ time$$

Probabilities are a measure of the frequency of an event in a given sample, and can range from 0 (the event is impossible) to 1 (the event is definite). For example, if 2 out of a

sample of 10 patients have a heart attack, the probability of having a heart attack would be 0.2.

The probability provides the likelihood of an event taking place, which can then be used to derive the expected values for each option being compared; the mean value of the end point. The pathway of each option is populated with events and associated costs, based on the initial decision. Using this information (probabaility of the event and hence cost) the expected cost and consequences can be derived. Definitions of both are given in Box 3.2.

Box 3.2 The expected values from a decision analytical model

Expected Cost:

Cost of the intervention plus the sum of the therapy-specific costs resulting from the chosen pathway, weighted by the probability of a patients following that pathway.

Expected Consequence:

Outcome of the intervention plus the sum of the therapy-specific consequences resulting from the chosen pathway, weighted by the probability of a patient following that pathway.

The first step in decision analytical model is the Decision Tree structure, which presents an individual's prognosis based on the initial choice of comparators; the results of which are then compared to make a final decision. The decision tree uses branches and nodes to structure the pathway for each option, populated with costs and probabilities for each subsequent event. Figure 3.1 shows the visual representation for this structure, over a single time horizon.

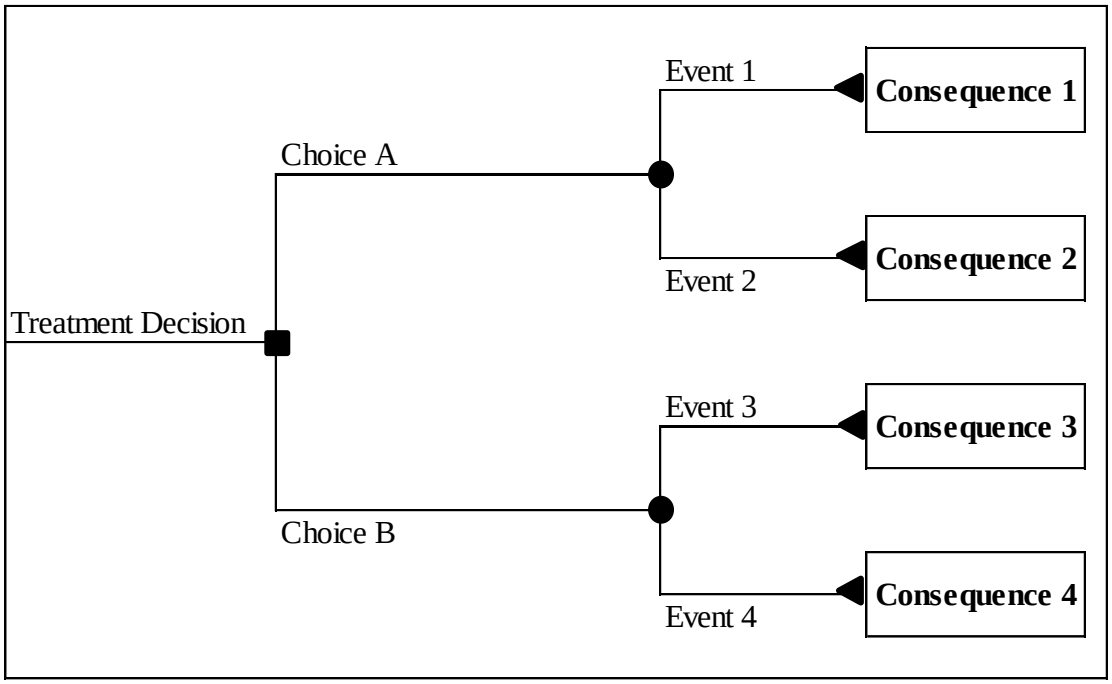


Figure 3.1 Visual representations of a Decision Tree structure.

3.2.2.1 Decision Tree

The decision tree is a simple structure to present the prognosis of each option. The branches at first represent the chosen option, and then become a pathway for each consequence of the initial choice.

The decision tree can be broken down into three parts, identified by its nodes. The initial square node is the decision node and is the start of the tree; it indicates the decision to be made between the options.

Once this decision is made, the next part of the tree is broken down by circular nodes, which are the event nodes. These nodes determine the unique pathway for the chosen option by the series of events which may take place as a result of the initial decision. Probabilities discussed above, show the likelihood of each event occurring. Moving left to right, the initial probability would be the probability of an event occurring. Subsequent probabilities would then be calculated using conditional probabilities based on the previous event occurring (Box 3.1).

This leads to the triangular node; the terminal node, for the final consequence of the initial decision. At this point the probabilities can be multiplied along each path to give the unique pathway probability, using joint probabilities (Box 3.1). Expected values are then the probability weighted summation of each pathway respective of costs or consequences.

The decision tree provides structure and clear visual representation to the evidence used to evaluate the added value of a new risk factor. Over a discrete time horizon, it visibly evaluates the costs and effects from the two risk prediction models. An extension of this method would be the Markov model which is used when modelling a large number of consequences over a longer or recurring (non-discrete) time period. In this case the decision tree may become 'bushy' and the Markov model would be used to model the different states (outcomes) and the probability of transitioning between those states. However, this is beyond the scope of this thesis.

3.2.2.2 Decision Tree for the Economic Evaluation of an Additional Risk Factor

The decision tree structure will be used to present information in costs and consequences which are needed for the cost-effectiveness analysis to evaluate the added value of a new risk factor to an established risk model. This will be performed as part of the validation study, in addition to current methodology used to assess and compare risk prediction models; with and without the new risk factor. This includes measures of calibration, discrimination.

As the observed outcome is known in a validation study, this information is used to assess the accuracy and performance of risk models, by its comparison to the predicted outcome. Using the disease prevalence and probabilities of the predicted disease prevalence conditioned on the observed outcome, the probability of being a true positive, false positive, true negative, and false negative will be derived. Thus each patient will be

labelled as a true positive, true negative, false positive, false negative dependent on their predicted and observed outcomes in the data, as shown in Table 3.1.

Table 3.1 Table to identify patients as true/false positives and true/false negatives based on their observed and predicted outcomes.

		Observed Outcome	
		✓	✗
Predicted Outcome	✓	True Positive	False Positive
	✗	False Negative	True Positive

This has therefore been incorporated into the proposed decision tree to evaluate the added value of a new risk factor, shown in Figure 3.2.

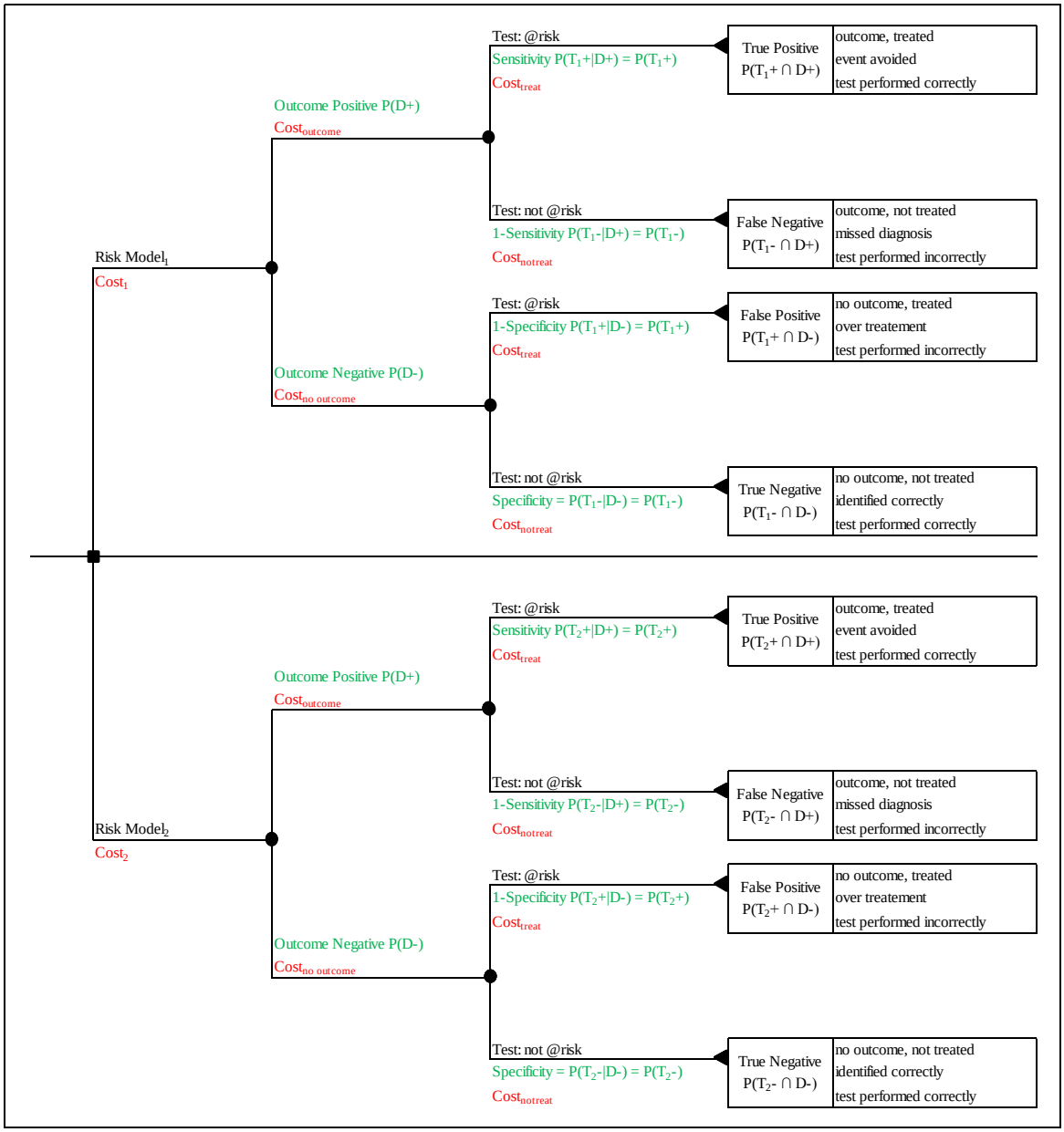


Figure 3.2 Visual representation of the proposed decision tree for evaluating the added value of a risk factor.

The proposed decision tree to evaluate the added value of a new risk factor begins by defining the square, decision node. The decision to be made here will be the choice between the risk prediction models, where Risk Model 1 will be the original established risk prediction model (without the new risk factor) and the Risk Model 2 will be the modified version (with the new risk factor).

Cost will be associated with each model, which will encompass the costs to collect all the data required for the included risk factors in each model. The key difference here will be the cost of the added risk factor, which will be incorporated into the cost of Risk Model 2.

The effects of the decision are then illustrated by the circular, event nodes. Two sets of event nodes occur in the decision tree; the first event node will be the observed outcome; and the second event node will be the predicted outcome. The predicted outcome will be based on a pre-specified risk threshold to indicate whether the patient is at risk of the outcome, or not.

A cost and probability for each event will be presented. The cost for the observed outcome event node will include cost of the outcome. The cost of the predicted event node will be accounted for by the risk model outcome specific treatment costs for the disease. This is based on the assumption that all patients predicted at risk of the outcome will be treated.

The probabilities for each event will be derived from the validation study data. Probability for the first event node (observed outcome) will reflect the probability of having the outcome P(D+), or not P(D-). The second set of event nodes will illustrate the whether the patient is at risk of the outcome or not, based on a pre-specified risk threshold. The pre-specified risk threshold for the disease will be derived from national clinical guidance. If this is not available, the opinion of clinical experts in the field will be sought.

The probabilities here are conditional on the previous event and are thus presented as (for each risk prediction model):

1. P(T+|D+): probability patient at high risk given they are disease positive
2. P(T-|D+): probability patient at low risk given they are disease positive

3. P(T+|D-): probability patient at high risk given they are disease negative

4. P(T-|D-): probability patient at low risk given they are disease negative

Values for 1 and 4 can also be interpreted as the sensitivity (true positive rate, conditioned on the observed outcome) and specificity (false positive, conditioned on the observed outcome) values for the risk prediction models. Values for 2 and 3 are 1 minus the respective sensitivity and specificity.

Finally, the triangular, terminal nodes will illustrate the end point of the decisions tree and will reflect the probability a patient is a true positive (TP, predicted 'at risk' and observed outcome), a false positive (FP, predicted 'at risk' and did not observe outcome), a false negative (FN, predicted not 'at risk' and observed outcome), and a true negative (TN, predicted and did not observe the outcome). These are the same outcomes presented in Table 3.1.

The cost and probability resulting from following each pathway to each terminal node will be calculated. These values will be used to then calculate the expected cost for each risk prediction model. Further details on values to be inputted into the model and the expected output are given below (Section 3.3.3, Box 3.5).

3.3 Methods Guide and Reporting of Health Economic Evaluation

The aim of the proposed methodology is to provide clinically useful estimates describing the added value of a new risk factor; thus aiding the decision to include a new risk factor in routine risk assessment. It is proposed to use a cost-effective analysis with a decision tree framework, in addition to current measures used to describe the added effect.

It is envisaged that this method would be used as part of future studies evaluating the added value of a new risk factor; providing further information to clinical guideline

makers, such as NICE. Therefore, it is important the developed methodology is robust and in accordance the current health economic methodological guidelines; and produces standardised results, comparable across studies, within diseases.

To create a validated method and standardised reporting approach, the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (66) was used to provide a robust framework to the proposed methodology, ensuring valid methodology was produced, encompassing all issues for consideration in health economic evaluation.

The CHEERS checklist is currently the gold standard for information to be included when reporting health economic evaluation studies; they consolidate existing reporting guidance; and provide the most up to date checklist. Fulfilling this checklist ensured, all health economic issues were considered in development of the proposed methodology. This checklist was cross-checked with the requirements used by NICE to appraise economic evaluations (65); and although this was not considered during the development of the CHEERS checklist, it includes all issues needed by NICE.

The CHEERS checklist covers all issues to be considered when reporting a health economic evaluation; from the title and abstract to the discussion. However, the methodological section was of particular interest, and was used to ensure the proposed methodology was in line with current guidelines and robust.

The full checklist can be found elsewhere (66); the relevant methodological section to be considered for the proposed analysis has been presented in Box 3.3.

Box 3.3 Methodological Issues for consideration for economic evaluation

Methods			
Target population and subgroups	4	Describe characteristics of the base-case population and subgroups analyzed including why they were chosen.	_____
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	_____
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	_____
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	_____
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	_____
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	_____
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	_____
Measurement of effectiveness	11a	*Single study–based estimates:* Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	_____
	11b	*Synthesis-based estimates:* Describe fully the methods used for the identification of included studies and synthesis of clinical effectiveness data.	_____
Measurement and valuation of preference-based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	_____
Estimating resources and costs	13a	*Single study–based economic evaluation:* Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	_____
	13b	*Model-based economic evaluation:* Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	_____
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	_____
Choice of model	15	Describe and give reasons for the specific type of decision-analytic model used. Providing a figure to show model structure is strongly recommended.	_____
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytic model.	_____
Analytic methods	17	Describe all analytic methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (e.g., half-cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	_____

Below is the information for each methodological point of the CHEERS checklist.

3.3.1 Methods – General

3.3.1.1 Target population, subgroups, and location

The target population used for the analysis will be dictated by the disease outcome of the risk prediction model and new risk factor being evaluated. The proposed analysis is recommended to be performed in addition to current methods using a validation dataset;

where overall characteristics of the study population will be described, as well as by any subgroups. If the model is developed by a particular subgroup, such as gender, the proposed analysis will be run conjointly and for each subgroup. The validation dataset will also provide information on the setting and location of the data, although the monetary impact of the effects of the new risk factor will be related to the UK system.

3.3.1.2 Study perspective and comparators

The perspective and comparators will be universal for each study and disease.

3.3.1.2.1 Comparators

The method is unique to comparing two risk prediction models; in this case where one includes a new risk factor and the other does not. Hence, the difference between the two models is analogous to the added value of the new risk factor.

3.3.1.2.2 Perspective

An analysis which incorporates costs in multifaceted; in particular, it introduces the choice of perspective from which the costs and benefits are taken; and needs to be clearly presented as it affects the results of the analysis. The perspectives an economic analysis can take are; societal, patient, and health service viewpoints. Of these the health service viewpoint is most commonly used (65). The societal perspective takes the widest viewpoint; and thus, is limited by the amount of data collection required. Costs from the patient perspective can be difficult to collect for observational studies; further, they can be subjective with high variability.

As risk prediction models are chosen and recommended by national guideline developers for healthcare, the health service viewpoint was adopted. In particular, according to NICE technology appraisal guidance (65), the NHS and PSSRU perspective is used when evaluating the added value of a new risk factor.

3.3.1.3 Time horizon

The time horizon used for the analysis will be disease and risk prediction model specific. For example, if the risk model is predicting 10 year risk estimates for the disease, such that it is interested in the outcome in a 10 year follow up period; a 10 year time horizon will be used for costs and effects in the cost-effectiveness analysis.

3.3.1.4 Discount rate

The evaluation of risk prediction models is a typical example of interventions requiring immediate cost for later effect; and thus requires a discount rate. For this, non-disease specific discount rates will be used and recommended for the cost-effectiveness analysis. Drummond (63) recommends using one of three possible discount rates; including a government discount rate; and the commonly used 3% and 5%.

As the perspective take for this analysis is that of the NHS/national guidance decision-maker, the government discount rate will be used. The most updated HM Treasury Green Book stipulated a 3.5% government discount rate (67). This was also consistent with the reference case described by NICE for economic evaluation (65).

A baseline value of 0% discount rate will also be used in a sensitivity analysis.

3.3.2 Methods – Outcomes

3.3.2.1 Choice of outcomes

The choice of outcome for the cost-effectiveness analysis will be the performance measures used in current methodology, performed as part of the validation study. This can range from increments in statistical performance measured by the calibration and discrimination (described by the AUC); or increments in clinical impact measured by the number of true cases and non-cases, number of outcome avoided.

The AUC and Youden Index (68) were used as the choice of outcome. The AUC was calculated as part of the validation study, independently from the decision tree; and the Youden Index was calculated from the decision tree (Figure 3.2), this is described in Box 3.4.

Box 3.4 Calculation of the Youden Index (68).

The Youden Index is used to assess the discriminatory accuracy of a test (model) and summarises the performance of the model given in Figure 3.2.

It assumes false positives to be as undesirable as false negative; and is the product of true diagnoses minus the product of the false diagnoses, out of the product of total number of patients observed with the outcome and total number of patients observed without the outcome. The formula for the Youden Index is given below:

$$\frac{(no.of\ true\ positives \times no.of\ true\ negatives) - (no.of\ false\ positives \times no.of\ false\ negatives)}{(total\ no.of\ diseased\ patients) \times (total\ no.of non\ diseased)}$$

3.3.2.2 Measurement of effectiveness

It is recommended the proposed cost effectiveness be performed in addition for current method in a single validation study. Hence estimates for the measures of effectiveness will be based on single studies; the respective validation study.

3.3.2.3 Measurement and valuation of preference-based outcomes

Preference based outcomes were not used. Reasons have been outlined in Section 3.2.1.1.

3.3.3 Methods – Resources and Costs

3.3.3.1 Estimating resource use and costs

Location for the inputted data for the decision tree has been outlined in Table 3.2.

Table 3.2 Values and their location to be inputted into the decision tree.

Input for Decision Tree	Location of data
Risk Model 1: P(D+)	Validation study
Risk Model 1: P(D-)	Validation study
Risk Model 2: P(D+)	Validation study
Risk Model 2: P(D-)	Validation study
Risk Model 1: P(T_1+|D+) = Sensitivity	Validation study
Risk Model 1: P(T_1+|D-) = 1-Specificity	Validation study
Risk Model 1: P(T_1-|D-) = Specificity	Validation study
Risk Model 1: P(T_1-|D+) = 1-Sensitivity	Validation study
Risk Model 2: P(T_2+|D+) = Sensitivity	Validation study
Risk Model 2: P(T_2+|D-) = 1-Specificity	Validation study
Risk Model 2: P(T_2-|D-) = Specificity	Validation study
Risk Model 2: P(T_2-|D+) = 1-Sensitivity	Validation study
Cost of Risk Model 1 ($Cost_1$)	Literature
Cost of Risk Model 2 ($Cost_2$)	Literature
Cost of Treatment	Literature
Cost of Outcome	Literature

T+ indicates the risk model predicted that the patient is at risk based on a pre-specified risk threshold.

T- indicates the risk model predicted that the patient is not at risk of the outcome based on a pre-specified risk threshold.

3.3.3.1.1 Probabilities

As the cost-effectiveness analysis will be performed in addition to current analyses used to evaluate the added value of a new risk factor, values from the single study, validation study will be used to populate the probabilities of the decision tree.

3.3.3.1.2 Costs

Costs are required for 3 stages of the decision tree, but will be disease dependent. Cost data will be used from current literature. This will include economic reports used by the national clinical guidance (NICE), reports from the PSSRU, the BNF, and current literature.

Overall, the first cost will be the cost of the risk prediction model. This encompasses the cost of data collection for all risk factors, including cost of the consultation with the GP.

The GP consultation will be cost by the average length of the consultation multiplied by the cost per minute of face to face contact with the GP. Data for this will be extracted from the Personal Social Services Research Unit. Cost for the additional marker, will be disease dependent, but will also be derived using the PSSRU, NICE economical evaluations, and other related studies.

The next stage of the decision tree will need the cost of the outcome, which will be derived from disease-specific clinical guidelines, other UK cost-related papers, and reports. The cost of the outcome will be the average cost of the outcome in the UK. If more than one definition exists for the outcome, a weighted average cost will be calculated for the decision tree model. For example, the definition for an osteoporotic fracture consists of fractures at the hip, spine, forearm, humerus and pelvis. In this instance, the prevalence of each fracture will be calculated and multiplied by their respective costs. The summation of these values will give the overall weighted mean cost of having an osteoporotic fracture.

The final stage of the decision tree will require the cost of being at high risk for the disease; hence the cost of the preventative treatment for the disease. For this, clinical guidelines will be used to identify the most probable course of treatment to prevent the outcome. If national guidance is not available, the opinion of clinical specialists will be sought.

It will be assumed that all patients at risk will be treated and the same compliance rate will be found using both risk prediction models. Based on this, a true positive patient can be classed as a correctly treated high risk patient; a false positive can be classed as an incorrectly treated low risk patient; a false negative can be classed as an incorrectly

untreated low risk patient; and a true negative can be classed as a correctly treated low risk patient.

The decision tree in Figure 3.2 can then be simplified to the decision tree presented in Figure 3.3.

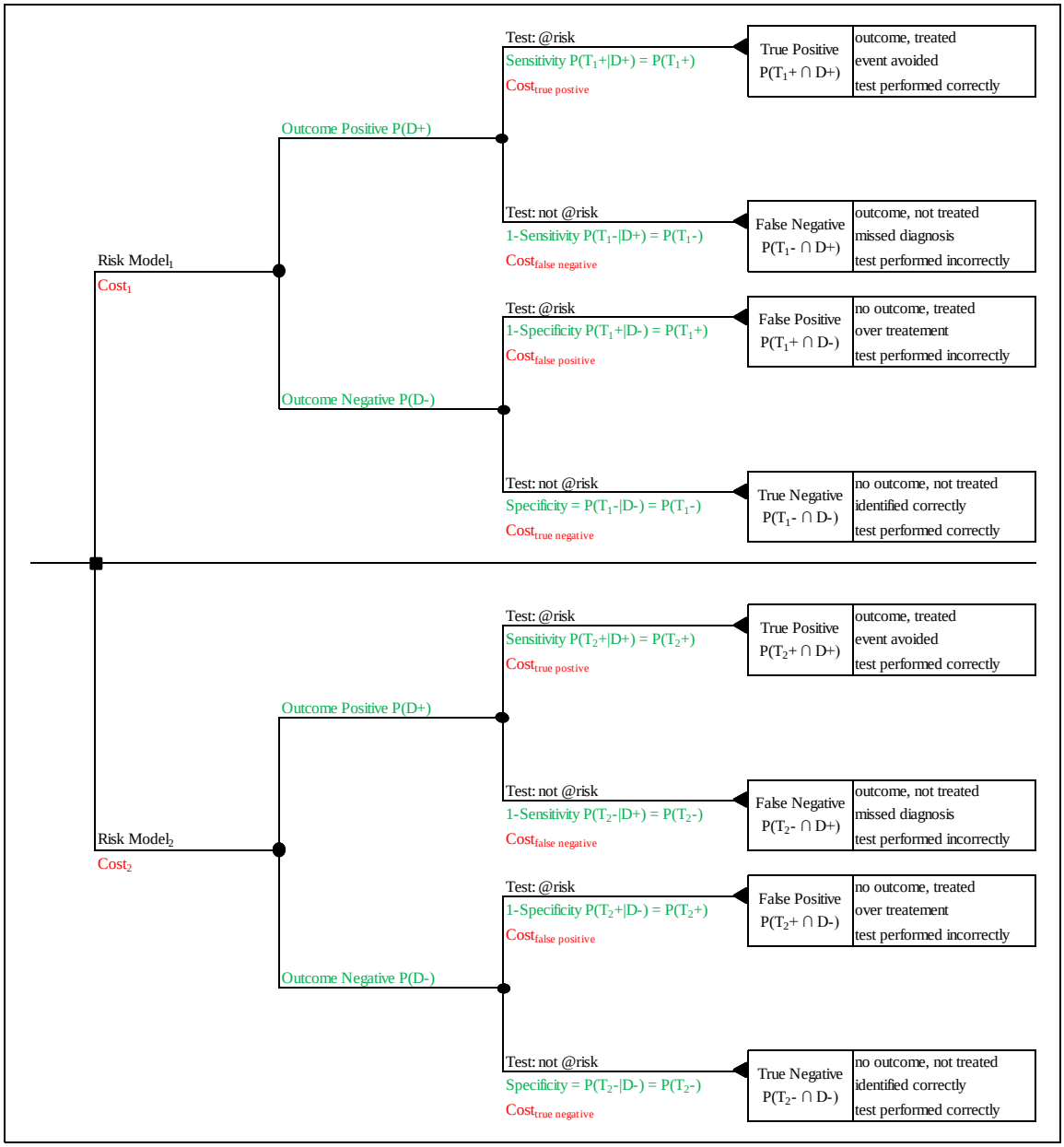


Figure 3.3 Simplified visual representation of the proposed decision tree for evaluating the added value of a risk factor; combined cost of the treatment and outcome.

In this decision tree, the cost of treatment and cost of outcome are combined using the treatment efficacy extracted from respective clinical trials. From these trials, the risk

prediction model outcome event rate will be derived in patients considered low risk who are treated; low risk untreated; high risk treated; and high risk untreated.

As it is unlikely participants in these trials will be identified as low or high risk based on a risk prediction model, study samples with high risk characteristics, such as personal history of a disease, and low risk characteristics can be used as a proxy indicator for their risk category.

Using this information, methods described by Fluerence et al (69) were used to convert the event rates for their respective follow up periods into per annum event rates; these were then converted into annual transition probabilities, depicting the annual probability of an event for the given time horizon of the cost effective analysis.

The annual transition probability was combined with the cost of treatment and the cost of the outcome to calculate the cost of a true positive (TP), false positive (FP), true negative (TN), and false negative (FP), as shown in Box 3.5.

Box 3.5 Equations to calculate the cost of true positive, false positive, false negative, and true negative.

$$Cost\ TP_n = (P(Outcome_{TP}) \times cost_{outcome}) + cost_{treat}$$

$$Cost\ FN_n = (P(Outcome_{FN}) \times cost_{outcome})$$

$$Cost\ FP_n = (P(Outcome_{FP}) \times cost_{outcome}) + cost_{treat}$$

$$Cost\ TN_n = (P(Outcome_{TN}) \times cost_{outcome})$$

Where:

$P(Outcome_{TP})$ is the probability of having the disease outcome if you are a true positive.

$P(Outcome_{FN})$ is the probability of having the disease outcome if you are a false negative.

$P(Outcome_{FP})$ is the probability of having the disease outcome if you are a false positive.

$P(Outcome_{TN})$ is the probability of having the disease outcome if you are a true negative.

3.3.3.1.3 Currency, price date, and conversion

Prices will be presented for the UK, and price date will be reflected by the respective validation study dates. For prices not given for the respective study dates will be inflated to the correct price year using price inflation factors given by the PSSRU (70).

3.3.4 Methods – Model-Based Economic Evaluations

3.3.4.1 Choice of model

A decision tree has been chosen to model the defined problem; comparing two risk models to evaluate the added value of a new risk factor. Reasons for this have been presented above in Section 3.2.2.2, with an outlined figure of the proposed decision tree in Figure 3.2.

3.3.4.2 Modelling assumptions

The main assumption used in the analysis is that every patient at high risk of the risk prediction model outcome is treated. Further, it is assumed all patients will be treated with the same treatment.

3.3.5 Methods – Analytical Methods

3.3.5.1 Analytical methods

Methods of analysis will be disease and data specific. Detailed methods are provided in Chapter 4, Section 4.2.5.9.1 and Chapter 7, Section 7.2.5.8.1 for the two exemplar studies.

Overall for the cost-effectiveness analysis, data will be used from the validation dataset to assess the performance of the risk models, and using their difference to evaluate the added value of a new risk factor. Data will first be described and then formatted to match the requirements of the risk factors included in the risk prediction models. Once the predicted risk is derived, it will be compared with the observed risk to derive probabilities to populate the proposed decision tree. The analysis will then assign costs to these to derive the expected cost from each risk prediction model.

The expected output from the decision tree (Figure 3.3) is described in Box 3.6.

Box 3.6 Method to derive the expected cost for each risk prediction model compared.

Let $P(T_n+)$ and $P(T_n-)$ be the probability of a patient is predicted at risk of the outcome, or not, respectively, using risks model $n = [1,2]$ and a pre-specified risk threshold. Risk Model 1 is the original model, and Risk Model 2 is the new model, with the new risk factor.

Let $P(D+)$ and $P(D-)$ be the probability of a patient being outcome positive and outcome negative, respectively.

Then the cost for each terminal node is:

$$Cost\ (T_n + \cap D+) = Cost_n + Cost\ TP_n$$

$$Cost\ (T_n - \cap D+) = Cost_n + Cost\ FN_n$$

$$Cost\ (T_n + \cap D-) = Cost_n + Cost\ FP_n$$

$$Cost\ (T_n - \cap D-) = Cost_n + Cost\ TN_n$$

And the probability of reaching the terminal node is:

$$P(TP_n) = P(D+) \times P(T_n + |D+) = P(D + \cap\ T_n+)$$

$$P(FN_n) = P(D+) \times P(T_n - |D+) = P(D + \cap\ T_n-)$$

$$P(FP_n) = P(D-) \times P(T_n + |D-) = P(D - \cap\ T_n+)$$

$$P(TN_n) = P(D-) \times P(T_n - |D-) = P(D - \cap\ T_n-)$$

Therefore the expected cost for a risk prediction model is:

$$Expected\ Cost_n = P(TP_n) \times Cost\ (T_n + \cap D+) + P(FN_n) \times Cost\ (T_n - \cap D+)$$

$$+ P(FP_n) \times Cost\ (T_n + \cap D-) + P(TN_n) \times Cost\ (T_n - \cap D-)$$

Using the information on the expected cost for each risk prediction model, the difference between models can be calculated. As the difference between the risk models would be the result of adding a new risk factor, this difference would also present the added value of this risk factor.

3.3.5.1.1 Incremental Cost-Effectiveness Ratio

Using information on expected costs for each risk prediction model in conjunction with the difference of various performance measures (effects) for risk prediction models, an incremental cost-effectiveness ratio (ICER) can be calculated Eq.3.1.

$$ICER = \frac{cost_{new\ model} - cost_{original\ model}}{effect_{new\ model} - effect_{original\ model}} \quad \text{Eq.3.1}$$

The ICER can be interpreted as the change in costs divided by the change in effects between two risk prediction models, which also translates to give the value of the additional risk factor. Based on this information the decision can be made, whether using the additional risk factor in routine assessment for the outcome is worthwhile through its cost-effectiveness.

Depending on the change in effect of interest, the ICER would give the cost per unit increase in the desired effect. This information is commonly presented graphically, using the cost effectiveness plane (Figure 3.4). The horizontal axis represents the difference in effect, and the vertical axis represents the difference in cost.

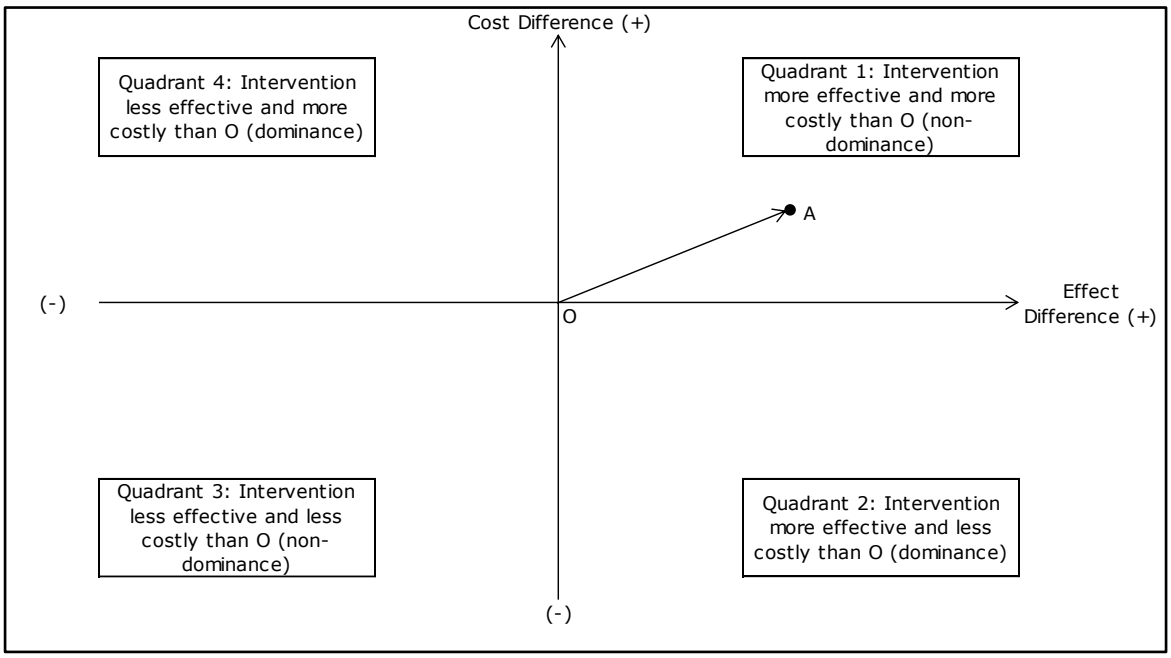


Figure 3.4 Cost-effectiveness plane.

The cost-effectiveness plane is a useful tool to describe the value of an additional risk factor and is commonly used in NICE economical reporting (65). Depending on which quadrant the change in cost and effect places this value (A), it describes if the modified risk model is more or less effective and costly that the original risk model. The slope from O to A gives the cost-effectiveness ratio.

Quadrants 2 and 4 are described as dominant, as the decision between options is obvious. For quadrant 2 the intervention is more effective and less costly than its comparator; and for quadrant 4 the intervention is less effective and more costly.

More thought needs to be given to values placed in quadrant 1 or 3, where the intervention is described as non-dominant. For quadrant 1 the intervention is more effective and more costly than its comparator; and for quadrant 3 the intervention is less effective and less costly.

A drawback of using an ICER to describe the added value of a new risk factor occurs when a negative ICER is calculated. In this case, it is difficult to know whether this is a

result of lessened effect or increased cost. It is therefore useful to also present the net benefit.

3.3.5.1.2 Net Benefit

An extension to the ICER is the net monetary benefit, which accounts for the additional factor affecting the decision of adding a new risk factor to an established risk prediction model; and its worth. The net monetary benefit is a function derived from the ICER; and introduces the willingness to pay (WTP). The WTP is the maximum amount the decision maker would be prepared to pay for the added effect of the new risk factor; which acts a threshold for the ICER, such that values below the WTP would be acceptable, and values above would not. Acceptable ICER values can therefore be illustrated by choosing an ICER which is below a pre-specified WTP threshold, λ (Eq.3.2).

$$ICER = \frac{\Delta cost}{\Delta effect} < \lambda \qquad \text{Eq.3.2}$$

Using this formula, it can be re-arranged to give the incremental net monetary benefit (INB), Eq.3.3:

$$INB = \lambda(\Delta effect) - \Delta cost \qquad \text{Eq.3.3}$$

The INB is the increase in effectiveness ($\Delta effect$) multiplied by the amount the decision maker is willing to pay (λ) per unit of effectiveness, less of the increase in cost ($\Delta cost$). Therefore an intervention is considered cost effective if (Eq.3.4):

$$INB > 0 \qquad \text{Eq.3.4}$$

As the willingness to pay for a unit of effectiveness is rarely known, the INB can be plotted for a range of values, shown in Figure 3.5.

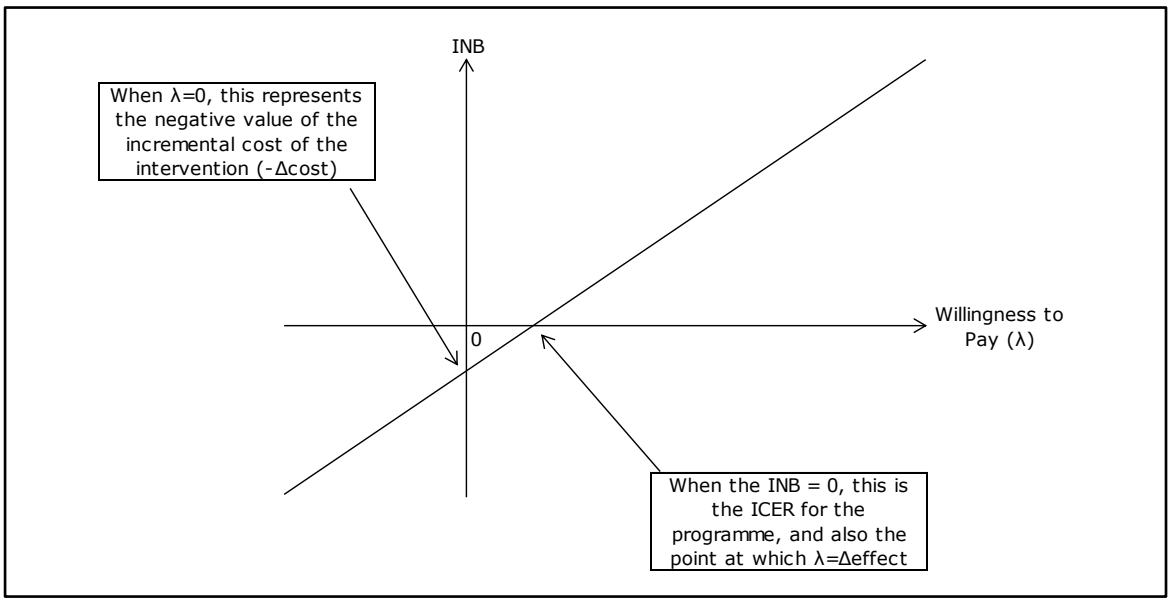


Figure 3.5 Incremental Net Monetary Benefit (INB) as a function of the willingness to pay (λ).

3.3.5.2 Variability, Heterogeneity, and Uncertainty

3.3.5.2.1 Variability

Variability in a decision analytical model differs from its uncertainty. Variability in a model occurs when the inputted parameters, such as costs and probabilities, vary between patients or locations.

In this case, the decision model is run for each subgroup to account for the variability. However, it is anticipated that variability in the outcome will be accounted for in the risk prediction model.

3.3.5.2.2 Heterogeneity

Where variability accounts for difference in the populations; heterogeneity measures the inter-patient differences (patient characteristics). This may affect the inputted parameters and is accounted for in the model methodological uncertainty, which is dealt within the next section.

3.3.5.2.3 Uncertainty

There are 4 types of uncertainty in a decision model; relating to methodology (structure); input parameters; extrapolation; and generalisation. All of these uncertainties can be dealt with by sensitivity analyses.

The simplest way to perform a sensitivity analysis on a decision model is using a one-way analysis. In this, the input for each parameter is varied one at a time to investigate the impact on the study results. However, it is rarely the case where the uncertainty would depend on individual parameters.

A multi-way sensitivity analysis recognises this, and investigates the combined variability of several factors inputted into the model. However, this may result in a large number of combinations of the individual parameters, to fully account for the uncertainty.

A scenario sensitivity analysis focuses on certain scenarios in which certain parameters would be varied. Standard scenarios are used, such as best and worst cases, and best guess scenarios, along with any other the analyst deems suitable to the situation. However, a large amount of uncertainty may remain.

A probabilistic approach is a favourable for a sensitivity analysis, and is used within the proposed analysis to account for uncertainty in the model. In this method, probability distributions are assigned to parameters, instead of estimates, to fully reflect the uncertainty in the model. Samples are then drawn randomly from the distributions to generate the empirical distribution for the ultimate effect size; the cost-effectiveness ratio. Beta distributions will be used for probability estimates as values are limited between 0 and 1; and Gamma-distributions will be used for cost estimates, as it is positively skewed and does not produce negative costs.

This is an iterative method, where samples are drawn for a pre-specified amount; and is undertaken by a Monte Carlo Simulation. In this case, the model is run a large number of times, each time with a different randomly drawn sample. The estimates (mean/expected values) are calculated for each sample, leading to an equal number of sets of expected costs and effects; reflecting the combined parameter uncertainty. This method would also provide a 95% confidence interval for the ICER estimates.

The results from the probabilistic sensitivity analysis (PSA) can be displayed using a scatterplot on the cost-effectiveness plane in Figure 3.4, with estimated ICER and its confidence interval. Based on the net benefit, this can be displayed on cost-effectiveness acceptability (CEAC) curves to provide a fuller picture of the parameter uncertainty (Figure 3.6).

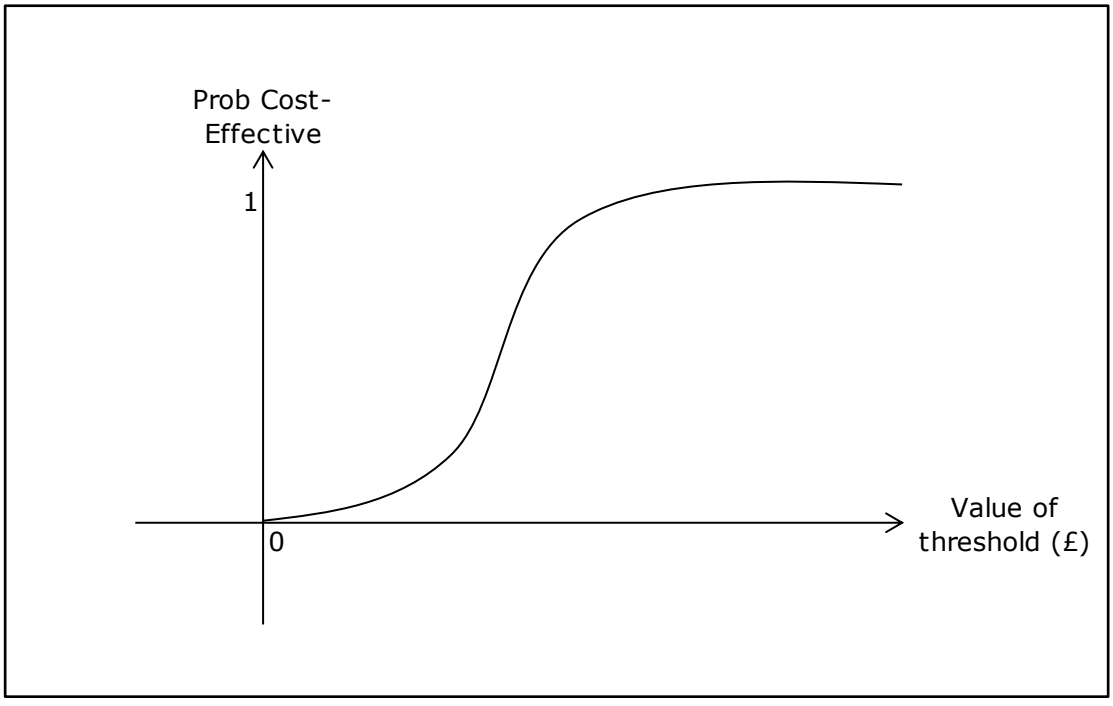


Figure 3.6 Cost-Effectiveness Acceptability Curve.

The CEAC curve allows the willingness to pay, to vary, and provides the probability an intervention (new risk model) is cost-effective at a chosen threshold value (willingness to

pay). It would also present the maximum probability the new risk factor is cost-effective. This is calculated as the proportion of iterations from the PSA which satisfy the INB condition (Eq.3.4), for each value of the WTP; the maximum probability would be the probability value at which the curve plateaus, such any increase in the WTP would not make the new risk factor more cost-effective.

The traditional CEAC varies the willingness to pay, but keeps the change in cost and effect constant. It can therefore be used to identify the minimum willingness to pay for a specified change in cost and effect.

3.4 Chapter Summary

This chapter outlines the rationale for the need of refined methodology. It accounts for the criticism of current methods described in Chapter 2, and incorporates cost implications into the evaluation of additional risk factors to established risk prediction models.

Health economical methods are outlined with regard to the analyses and framework. The proposed refined methodology is then detailed, in line with the CHEERS checklist and NICE reference case for robustness. This will also help provide comparable and standardised results.

This method will be applied using two examples of risk prediction; the impact of adding family history of cardiovascular disease risk prediction; and bone mineral density score to fracture risk prediction.

Chapter 4 - Family History and Cardiovascular Disease Risk: Introduction and Methods

The previous chapter describes a refined methodology to evaluate the added value of new risk factors to established risk prediction models. In this chapter, the added value of family history (FH) of cardiovascular disease (CVD) to CVD risk prediction is evaluated using the current and refined methodology.

There are two sections to this chapter. The first section gives an introduction to CVD risk prediction; it explains the rationale behind improving cardiovascular risk prediction, and describes the current practice of cardiovascular risk assessment in UK primary care. It also identifies FH of CVD as an important risk factor for CVD risk prediction.

The second section describes the methodology used to evaluate the added value of FH of CVD to CVD risk prediction. This outlines the study design, data source, and study population. Details are then given on the analysis which describes the current, and refined methodology.

4.1 Introduction

4.1.1 The Burden of Cardiovascular Disease

Cardiovascular disease (CVD) includes coronary heart disease (CHD) and stroke, and is the main cause of death, as well as a leading cause of premature death and major cause of disability in England and Wales (71). It accounted for 180,000 deaths in 2010, of which 45% were due to CHD and 28% due to stroke. Further, CVD accounted for 2% of premature deaths in men and 18% of premature deaths in women (72). Specifically, reports have shown that 1 in 5 men and 1 in 10 women will die from CHD (71).

Reports have also shown a 45% drop in death rates from CHD for people aged 65 and under, in the last 10 years (1998-2008) (71); however the number CVD related deaths is set to increase to 23.3 million by 2030, leaving CVD to be the single leading cause of death (73). This, coincided with costs of £8.7 billion to the UK healthcare system in 2009 and £19 billion per year to the UK economy, explains the continued effort for the prevention of CVD.

4.1.2 Prevention of Cardiovascular Disease in Primary Care

Prevention strategies for CVD in primary care consist of primarily two interventions. The first focuses on promoting healthier lifestyles in the general population (for example smoking cessation, increased physical activity and dietary changes). The second uses targeted lipid modification drug therapy, such as statins, to reduce CVD risk. These are assigned to patients based on their risk of CVD.

In practice, CVD risk prediction models are used to derive risk estimates for each patient; these combine pre-specified CVD risk factors already recorded in primary care electronic healthcare records. Based on the patient risk assessment and the risk threshold, individual interventions are prescribed to reduce risk, providing lifestyle advice to low risk patients and further drug treatment to patients at high risk.

The National Institute for Health and Clinical Excellence (NICE) clinical guidelines for lipid modification (3) lists primary prevention of CVD as a key priority for implementation, including systematic strategies to identify and treat people at high risk of CVD. Currently, a 10% risk threshold is recommended for 10 year CVD risk to offer statin drug therapy (3). Based on both the cost and clinical effectiveness evidence the treatment threshold was lowered from 20% to 10% in NICE clinical guidance.

Although risk prediction models are used as part of a full CVD risk assessment and are advantageous to help identify patients at high risk of CVD, how well they perform in identifying the correct patients for preventative treatment is subjective to the risk prediction model used and risk threshold used to treat them.

Further, where more than one model is available to risk assessment, it has been shown to be difficult to recommend a single risk model for CVD risk prediction. The 2008 NICE clinical guidance for lipid modification was re-issued in 2010 (17); this amended the recommendation of solely using the Framingham CVD prediction model for CVD risk assessment. Guidance was updated to also recommend the UK based QRISK2 risk model for CVD risk assessment which provided better prediction accuracy for the UK population (74). Since 2013, NICE now recommend using only QRISK2 for CVD risk prediction (3).

4.1.3 Cardiovascular Risk Prediction

Numerous models have been developed for CVD risk prediction, these include the Framingham Risk Score (JBS3) (14), QRISK2® (11), ASSIGN (8), and the Reynolds Risk Score (31). In the UK, the most popularly used CVD risk models are the Framingham Risk Score (JBS3), QRISK2® and ASSIGN.

4.1.3.1 Cardiovascular Risk Factors

Cardiovascular risk prediction models often differ with respect to the risk factors they include; new and other additional risk factors can affect the performance of the model. Table 4.1 shows the variation in risk factors which are included in each CVD risk model.

Among the numerous CVD risk factors, age, gender, systolic blood pressure (BP), total cholesterol, high density lipoprotein (HDL) cholesterol and smoking status are considered as the ‘standard’ risk factors for CVD. These have shown to account for about 75% of

CVD events; 13% of global CVD deaths are attributable to raised blood pressure, and 9% are due to smoking (75).

Table 4.1 Table of CVD risk factors by risk prediction model.

Risk Score	JBS3 (Framingham)	QRISK2	ASSIGN	Reynolds
Standard Risk Factors	✓	✓	✓	✓
Additional variables	Ethnicity Deprivation (Townsend Score) Impaired glucose Diabetes Blood Pressure treatment Chronic Kidney Disease Atrial Fibrillation Non fasting glucose Triglyceride Levels Rheumatoid Arthritis Left Ventricular Hypertrophy Family History of CVD Alcohol consumption Physical Activity HbA1c BMI Liver Enzymes	Ethnicity Body mass Index Family history of CHD Deprivation (Townsend score) Treated hypertension Rheumatoid Arthritis Chronic Renal Disease Type 2 Diabetes Atrial Fibrillation	Diabetes Family history of CHD Index of Multiple Deprivation Score (IMD)	High sensitivity C-Reactive Protein Family History of CHD

Age and gender are both strong risk factors of CVD, but are both non modifiable risk factors; and are used to identify patients at high risk. The number of CVD events

increases with age and is also higher in men than women; a lower age-adjusted CVD death rates was found among males than females (76). Guidelines use these risk factors to estimate the overall risk of CVD rather than to tailor preventative treatment.

Modifiable CVD risk factors, outside their normal range have shown to account for up to 90% of population attributable risks for Myocardial Infarction in men, and 94% in women (77). This includes high BP, increased smoking and abnormal lipid levels (high total cholesterol and/or low HDL cholesterol). Prevention strategies of CVD, hence, focus on improving these risk factors, through smoking cessation, lifestyle changes, and antihypertensive and lipid modification treatment as a primary method to reduce risk and CVD burden.

4.1.3.2 Cardiovascular Risk Prediction Models

The primary strategy for CVD prevention in primary care is to use risk prediction models to identify individuals at high risk and treat them accordingly to reduce their risk. Based on the single risk estimate derived from the model, it is used with the pre-specified risk threshold (currently 10%) to identify those at high risk of CVD and those who would most benefit from the preventative treatment.

Although there are many risk prediction models available to use, in the UK the Framingham Risk Score, QRISK2 and ASSIGN are most commonly used. The choice of using these models is based on their performance to identify high risk patients in the UK population. Methods and metrics used to assess the performance are described in Chapter 2.

4.1.3.2.1 The Framingham Risk Score

The Framingham Risk Score presents the 10 year risk of CVD, which is the likelihood of a person having a CVD event within 10 years of using the model. This model is derived

using data collected prospectively for 20 years, from residents of Framingham in the USA. In the UK it is adopted by the Joint British Societies (JBS3) guidelines for prevention of CVD in clinical practice (14).

Clinically, the model is easy to use, and its baseline model only requires data for the 'standard' CVD risk factors. Accounting for these risk factors alone provides the model with a discriminatory accuracy of 0.790 (3).

However, to account for additional risk factors, they have been either driven by evidence or by consensus, which may vary according to guidance. For example, BMI is embedded in CVD risk prediction and is accounted for in the calculated risk estimate. However, in JBS2 (18), to account for a positive family history of premature CVD (<55 years for men and <65 year for women) in a first degree relative, the risk estimate calculated by the risk model was increased by a factor of 'about 1.3'; further the NICE lipid modification guidelines increases the risk estimate by a factor of 2 if the patient has more than one relative in their family history of CVD (17). Similarly, as the model has not been developed to account for different ethnic backgrounds (used a predominantly white cohort); the CVD risk estimate was inflated by 1.4 to account for South Asian ethnicities (18).

The Framingham model can be criticised for being derived from a non UK population, which questions its generalisability to this setting. It has shown to overestimate the absolute risk of CVD in lower risk populations and underestimate the risk in high risk populations; and would generally overestimate risk in the current UK population. This is further highlighted by its poor calibration to a UK population (74, 78).

4.1.3.2.2 QRISK2®

The QRISK2 has been developed in the UK and uses data from QRESEARCH, a large anonymised primary care database of electronic healthcare records. This database holds data on over 10 million patients, collected over a 17 year period, from 529 general practices using the EMIS GP system.

The QRISK2 model has shown to be a better predictor of CVD; with better calibration and discrimination than the Framingham risk score, both in internal and external validations (74, 78). However, it faces limitations of using routinely collected primary care data (79). Data may not be fully validated in routinely collected medical data, and incomplete data also limits the data. Methods such as using complete cases or multiple imputation can be employed to address incomplete data; however, the former may produce biased results and the latter may violate modelling assumptions (80). Evidence suggests 22-38% of data is missing for the modifiable 'standard' CVD risk factors in primary care databases (81). Of these, height, weight and BP were found to be missing at random and, smoking was not.

QRISK2 includes substantially more CVD risk factors in its baseline model, including FH information, thus more data collection and resource is required. This may affect the uptake of the model compared to the simpler Framingham risk model. The increased number of risk factors also increases the amount of missing data, providing further limitations.

Further, discrepancies can be found between models in risk factors measuring the same information. For example, QRISK2 includes chronic renal disease, which is identified using read codes and is presented as a binary measurement. The JBS3 includes chronic kidney disease (CKD), which is identified using the estimated glomerular filtration rate

(egfr) calculated from serum creatinine and is thus a continuous measure. Although using binary risk factors increases ease of use, there is a potential loss of information from not using continuous measures where applicable.

It can also be argued that by using binary measures for chronic renal disease, which assumes the disease is not present if the pre-specified read code is not found, would lead to an underestimation of missing data for the model.

4.1.3.2.3 ASSIGN

The ASSIGN (8) CVD risk prediction model was developed using the Scottish Heart Health Extended Cohort (SHHEC) and uses similar risk factors to the Framingham risk score. It is an extension of the SIGN risk score, adding social deprivation (Scottish Index of Multiple Deprivation (SIMD)) and family history of premature CVD and shows marginally better calibration and discrimination than the Framingham risk score. Further it better reflects the variation in CVD events according to socio-economic status in Scotland (82). However, similar to the Framingham risk model it also lacks generalisability to the UK population and is limited as it is derived from the Scottish population (83).

4.1.3.3 Additional/New Risk Factors

Standard risk factors have shown to account for a large number of CVD events; the addition of new risk factors is used to further increase the number of correctly identified CVD events and improve the predictive ability of the model. This has produced many risk prediction models; however the improvement gained from the new risk factor needs to be better quantified.

The additional risk factors currently included in established risk models, beyond the 'standard' risk factors of CVD have been presented in Table 4.1. However, more risk

factors are constantly being identified, and claim to improve CVD risk prediction. This includes potential genetic information being researched in genetic studies (84, 85).

4.1.3.3.1 Family History and Cardiovascular Disease

Family history is a known non-modifiable additional risk factor for CVD risk prediction models. It is accounted for in CVD risk prediction by adding it to the standard CVD risk factors in CVD risk prediction; and if it is not embedded within the risk model; it has been used to inflate the calculated CVD risk estimate (17, 18).

Improving CVD risk prediction is beneficial as it clinically identifies more patients at high risk of a CVD event (8, 9, 86-88). Although FH is a known risk factor for many chronic diseases, including CHD (89-92), its prognostic value in CVD risk prediction for individual patients is not firmly or explicitly established.

Varying methods have been used when accounting for additional FH information. Guidelines have adopted consensus driven inflation factors which are used on the predicted CVD risk estimates derived from risk models with standard CVD risk factors without FH. This does not fully account for the medical evidence assessing the added predictive accuracy of using FH information in CVD risk assessment.

Family history of coronary heart disease is included in the QRISK2 risk model, but the extent of FH recording in GP databases has shown to be poor (93). Briefly, in a cross-section of 1,504,535 patients extracted from The Health Improvement Network (THIN), a routinely collected GP database, 19% had a positive or negative family history of CHD recorded in their records (93). This translated to only 9% of patients having a positive FH recorded and fewer still with additional FH (age of onset and relative affected) information recorded in their GP records. Further details have been provided in Appendix B. As well as affecting model development, lack of FH data collection may affect the

feasibility of using this risk factor in the model; as data would not be available to populate the CVD risk model and fully make use of its added benefit. Hence, better understanding the value of FH of CVD to CVD risk prediction would promote better data collection in routine risk assessment in primary care.

Studies have found up to a threefold increase of coronary heart disease (CHD) events among patients where a FH of CHD was found (94-97). Further, the strength of association increased with the number of relatives found, in particular if the relative was affected at an early age (98, 99). However, the independent effect of FH information needs to be shown on the predictive accuracy of identifying CVD events when added to standard CVD risk factors, as well as through the strength of it univariate association with CVD.

Studies have shown the added value of FH information to CVD risk models (8, 9, 86-88), and have found up to a 7.3% (87) relative increase in the area under the Receiver Operating Characteristic (ROC) curve (AUC). However, it has been difficult to put these increments into clinical context to assess the clinical impact of FH on CVD risk prediction. Further, studies do not look at the added value of FH information independently of other new risk factors being added to CVD risk prediction (8, 9). This makes it difficult to of identify improvement solely due to the addition FH information.

More clarity is therefore needed in the assessment of the added value of FH information in CVD risk assessment. Its association with CVD needs to go beyond univariate analyses, and it needs to be assessed in addition to the standard CVD risk factors. In addition, the independent added value of FH information needs to be assessed, separate from other novel risk factors, to uniquely identify its added value. Further, better translation of statistical improvement of CVD risk prediction to clinical context is needed.

A cost-effectiveness analysis with a decision tree would provide better information of the improvement of CVD risk models through the addition of FH information. Using current performance measures for risk prediction models; this can be used with cost data to fully evaluate the clinical consequence of adding FH to CVD risk prediction. Based on this, better informed decisions can be made about incorporating FH information in CVD risk models rather than accounting for it through the opinion of guideline developers.

4.1.4 Aim and Objectives

The aim of this study is to assess the value of FH information in addition to the standard CVD risk factors used in the Framingham risk model.

The objectives to achieve this aim are:

- Recalibrate the Framingham Risk Score on a UK population, with and without family history using routinely collected data (Chapter 5).
- Internally validate both models and assess the improvement in predictive ability of addition of family history, using current measures (Chapter 5).
- Assess the added value of family history as a CVD risk factor using a decision tree based cost-effectiveness analysis described in Chapter 3 (Chapter 6).

4.2 Methods

The previous section described CVD risk assessment in primary care; the compulsory (standard) CVD risk factors; and the optional (additional) risk factors which are currently being included in CVD risk assessment via CVD risk prediction models. Family history information is also an additional risk factor and its added value to CVD risk prediction is yet to be determined.

This section outlines the methodology used to identify the improvement FH information will add to standard and/or additional CVD risk factors in a CVD risk prediction model setting. It begins by describing the design of the study, the source of the data, and the study population. The extraction of the data used is then described, leading to the proposed statistical and cost-effectiveness analysis to quantify improvement.

4.2.1 Study Design

A prospective cohort study was designed to evaluate medical tests. The 'tests' being evaluated were CVD risk prediction models, with and without family history of CVD as an additional risk factor.

The initial extracted cohort from was split into two datasets; two-thirds of the data was allocated to the derivation (training) dataset of the study; and the remaining third was allocated to the validation (test) dataset.

The derivation dataset was used to create a CVD risk prediction model based on the established CVD risk prediction models using both 'standard' and additional CVD risk factors (as outlined in Table 4.1), with and without family history. These models were then 'tested' using the validation dataset. This assessed their performance which allowed for comparisons between models to be made and the impact of family history information on CVD risk assessment to be shown.

4.2.2 Data Source

Data for the study was extracted from the Clinical Practice Research Datalink (CPRD), formerly known as the General Practice Research Database (GPRD). This database is a large, validated primary care research database with information on over 12 million patients and 650 practices. It collects information of patient demographics, clinical events, referrals, tests, prescriptions and deaths. It also includes anthropometric data such as BMI and also self-reported lifestyle information such as smoking status and alcohol consumption.

4.2.2.1 Data Files

Data for each consultation with the general practitioner (GP) is stored in appropriate files, of which the clinical, additional clinical details, test, and therapy files were used in this study

The ***clinical file*** stores medical history events; this includes symptoms, signs and diagnoses which can be used to identify clinical diagnoses made by the GP; and deaths. Often, more than one record is held per patient per consultation in this file.

Further information regarding clinical events recorded in the clinical file are found in the ***additional clinical details file,*** where additional data can be entered in structured data areas. Other information, such as lifestyle data and anthropological measurements can also be found in this file.

The ***test file*** contains data on tests requested by the GP, including the type of test and the result. It follows a similar format to the additional clinical details file in providing structured data areas for the test information. Data for the result can denote either qualitative text based results (e.g. ‘normal’ or ‘abnormal’) or quantitative results involving numerical values.

The ***therapy file*** contains data on all prescriptions given to the patient, relating to drugs and appliances issued by the GP.

4.2.2.2 Coding

Clinical events, diagnoses and procedures described predominantly in the clinical, additional clinical details and test files, are coded using the Read Code hierarchy. Read Codes map onto the International Statistical Classification of Diseases, 10th Revision (ICD-10) and are used as standard practice by GPs to standardise recording of patient medical activity.

Read Codes are often complimented by entity codes used in the additional clinical details and test files. Entity codes refer to the structured data area in the GP computer software and may be associated with up to 7 data fields, to content of which can range from raw values or encoded values representing dates, Read Codes or text.

Prescriptions recorded in the therapy file are coded using the Multilex product code system which links to the British National Formulary (BNF), identifying the chapter and section of the drug.

4.2.2.3 Linkage of Data

Data from CPRD can also be linked to other patient level datasets. This includes the Hospital Episode Statistics (HES), the Office of National Statistics (ONS), and Socio-economic Status (SES). Approximately 55% of UK and 70% of English practices were covered by this linkage at the time of extraction.

The HES data provides information of all admissions and day case management of patients in NHS hospitals in England. There are three levels of HES data available from CPRD, of which the integrated HES data was used for the study. This provided the reason

of hospitalisation (using the ICD-10 hierarchy) and discharge date of the patient, and was used to identify patients with a CVD event during the follow up of the study.

The ONS data provides information on mortality and includes the official date and cause of death, using ICD-10 codes. This data was used to identify patients who had died during the follow up from any cause, and also patients who died from CVD related events.

The SES data provides the Townsend Score and Index of Multiple Deprivation (IMD) for patients. The former is based on patient postcode and the latter is based on the practice postcode which the patient is registered with. Townsend Score was used for this study

4.2.3 Study population

The baseline date for this study was 1st Jan 2001, and patients were followed up for 10 years (1st Jan 2011). They were selected for the study using the criteria below.

4.2.3.1 Practice Selection

All practices contributing to the CPRD database were eligible for the study. However, only those providing data to CPRD for the defined study period were included. Practices were excluded from the study if their data was not deemed research quality after the baseline study date and/or the date their data was last collected by CPRD was after the study end date.

4.2.3.2 Cohort Selection

From the selected practices using the criteria above, patients were eligible for the study if they were aged between 30-79 years and have a FH of CVD recorded (either 'no FH of CVD' or 'FH of CVD') at the study baseline date. Patients must have also have been registered with their practice for at least 1 year and hence have provided at least 1 year of data.

Patients were excluded if they were found not to be contributing to the CPRD database during the study period. Specifically, if they had registered with the practice after the study end date (01/01/2011) or the last CPRD data collection date for the practice; and if they had transferred out of the practice on or before the study start date (01/01/2001) or before the date the practice data was deemed research quality.

Patients were also excluded if they had any pre-existing history of CVD (including Myocardial Infarction, Angina and Peripheral Vascular Disease), Cerebrovascular Disease (excluding subarachnoid haemorrhage and vascular dementia), or had received any hypertension or lipid treatment before baseline.

Patients reporting a CVD event with a missing event date were also removed from the study.

4.2.4 Study variables

Variables for this study are based on existing CVD risk prediction models. We were particularly interested in the Joint British Societies' guidelines (JBS3), which is based on the American Framingham risk prediction model. Data permitting, risk factors included in other CVD risk prediction models were also investigated: the JBS3 including other recognised CVD risk factors (JBS3+); the Scottish ASSIGN; and American Reynolds risk scores. A summary of each model is given in Table 4.2, which also outlines the additional CVD risk factors considered for each model.

Table 4.2 Variables included in the JBS3, JBS3+, ASSIGN and Reynolds CVD risk prediction models.

Risk Score	JBS3*/ Framingham	JBS3+	ASSIGN	Reynolds
Standard Risk Factors**	✓	✓	✓	✓
Additional variables		Ethnicity Impaired glucose Deprivation (Townsend Score) Non fasting glucose Triglyceride Levels Family History Diabetes Physical activity CKD*** Alcohol intake HbA1c LVH*** BMI*** Liver Enzymes**** Hypertensive treatment Atrial Fibrillation Rheumatoid Arthritis	Diabetes Family History IMD***	hsCRP*** Family History

***Primary algorithm of interest**

****Standard risk factors (in all 4 models) are: Age; Gender; Smoking Status; Systolic BP; Total Cholesterol; and HDL Cholesterol**

*****CKD=Chronic Kidney Disease; LVH=Left Ventricular Hypertrophy, BMI=Body Mass Index, IMD=Index of Multiple Deprivation Score, hsCRP= high sensitivity C-reactive protein**

******Liver Enzymes are aspartate aminotransferase (AST), alanine aminotransferase (ALT) and gamma-glutamyl transpeptidase (GGT)**

The variables to be used in this study and their format are given in Table 4.3. Each patient had their own unique identifier which linked them across the different data files described above in CPRD. Details of age and gender were recorded for each patient in the study and data was searched for, for the other variables.

Read Code lists were developed to identify to outcome (CVD) and other study variables, such as diabetes, smoking status and anthropological measurements. These were then complimented with entity code lists to search the structured data fields in CPRD (in both the additional clinical details file and test file).

Read Code lists were developed using recognised primary care database methodology (100) and were informed by the Quality of Outcomes Framework (QOF) business rules for a standardised approach. If QOF rules were not found, Read Codes were scrutinised by three assessors for robustness.

Data on study variables was searched within a 5 year window before the baseline date (01/01/1996 – 01/01/2001) and the most recent value before baseline was used for the analysis.

Table 4.3 List of study variables.

Variable List	Format
Patient ID	Continuous
Baseline study date	Date – 01/01/2001
Follow-up end date	Date – 01/01/2011
Standard Risk Factors	
Age	Continuous (years)
Age	Categorised (10 year age bands)
Sex	Dichotomous (male/female)
Systolic blood pressure – treated	Continuous (mmHg)
Systolic blood pressure – untreated	Continuous (mmHg)
Total cholesterol	Continuous (mmol/l or mg/dL)
HDL cholesterol	Continuous (mmol/l or mg/dL)
Smoking status	Dichotomous (yes/no)
Outcome Data	
Outcome: CVD event	Dichotomous (yes/no)
CVD event date	Date
Additional Risk Factors	
BMI	Continuous (kg/m2)
Ethnicity	Categorical
Deprivation	Categorical – Townsend score (1= less deprived-5=more deprived)
Physical Activity	Continuous (amount of moderate exercise/week)
Alcohol	Continuous (units/week)
Diabetes	Dichotomous (yes/no)
LVH	Dichotomous (yes/no)
Impaired Glucose Regulation	Dichotomous (yes/no)
Non fasting glucose	Continuous (mmol/l)
Fasting glucose test	Continuous (mmol/l)
Triglyceride levels	Continuous (mmol/l)
hsCRP	Continuous – default (20mg/l)
CRP	Continuous (mg/l)
HbA1c (glycated haemoglobin)	Continuous (mmol/mol)
ALT	Continuous (IU/L)
AST	Continuous (IU/L)
GGT	Continuous (IU/L)
CKD (Stage 3-5)	Dichotomous (yes, no)
Estimated Glomerular Filtration Rate	Continuous (umol/L)
New Risk Factors	
Family history	Dichotomous (yes/no)
Detailed family history	Categorical (FH/FH+age/FH+relative/FH+age+relative)

4.2.4.1 Outcome

The outcome for the study was a CVD event, including stroke as well as coronary heart disease conditions. It consisted of Read codes indicating coronary heart disease, stroke and TIA conditions; excluding peripheral vascular disease from CHD conditions. The code list was then cross-checked and validated using other published CVD Read Code lists (101).

To support the developed CVD Read Code list, an ICD-10 code list was also developed for CVD. Using the CVD Read Code list, Read Codes were mapped to their respective ICD-10 codes. The respective code lists can be found in Appendix C.

Using the Read Code list, patients were identified to have a CVD event if at least one of these Read Codes were present in their medical records (clinical file) during the 10 year follow-up period (1st January 2001 to 1st January 2011).

Similarly the ICD-10 CVD code list was used to identify patients incurring a CVD event in their HES or ONS records within this time period. Additional patients were also classed as outcome positive if they were put on the CVD or stroke register within the follow-up period. These patients were identified using the entity codes 16 and 17 presented in the additional clinical file, respectively.

Patients classed as having a CVD event, without an event date were dropped from the study.

For the statistical analysis, outcome data from each of these sources (clinical, additional, HES and ONS) were aggregated into a single binary variable indicating the presence or absence of a CVD event, using earliest record of a CVD event during the follow up period.

4.2.4.2 Derivation of Standard Risk Factors

Data for each risk factor was extracted from CPRD using all the available files. Description of how data was extracted for the standard, additional, and new risk factors are provided below.

The standard risk factors for the study were age, gender, blood pressure, total cholesterol, HDL cholesterol and smoking status, which were common amongst the 4 CVD risk prediction models outlined in Table 4.2.

4.2.4.2.1 Age

Using the patient date of birth, their age was calculated at the baseline date of the study, given in years. Age was treated as a continuous variable, in the analysis, but was also categorised into 10 year age groups for descriptive purposes.

4.2.4.2.2 Gender

Gender was specified for each patient; they were classed as either male or female in a binary variable.

4.2.4.2.3 Blood Pressure

Systolic (SBP) and diastolic (DBP) blood pressure was calculated taking the mean of the two most recent recordings taken before baseline, which were within 5 years of each other. If two values could not be found within 5 years of each other, the most recent value close to baseline was taken.

Systolic blood pressure was used as the standard CVD risk factor in the risk prediction model and was presented for patients receiving hypertensive treatment at baseline and patients who were not; thus using a hypertensive treatment indicator variable for the analysis (0,1).

Both variables were given in mm Hg and were treated and continuous variables for the analysis. Outliers were removed using methods described below (Outliers).

4.2.4.2.4 Cholesterol

Values for total cholesterol, high (HDL) and low (LDL) density lipoprotein were taken from the database and were presented in mmol/L units. Values given in mg/dL units were converted to mmol/L using the following conversion equation:

$$1\,mg/dL = 0.02586\ mmol/L \qquad \text{Eq.4.1}$$

Cholesterol values with units which could not be converted to mmol/L were excluded from the analysis. However, cholesterol values where no unit was specified in the data were assumed to be mmol/L.

All cholesterol related variables were used as continuous variable in the analysis. Outliers were removed using methods described below (Outliers).

4.2.4.2.5 Smoking status

Smoking data was recorded in the clinical file and the additional clinical details file. Both files were merged by the given additional detail identification number. Smoking status was first derived based on the medical Read codes recorded in the clinical file, and then again based on entity codes recorded in the additional clinical details file. The created variables identified patients as smokers, ex-smokers and non-smokers. Smokers were then categorised into trivial, light, moderate, heavy and very heavy smokers; based on the number of cigarettes smoked per day. Table 4.4 presents the criteria used to categorise smokers.

Table 4.4 Criteria to class smokers as trivial, light, moderate, heavy and very heavy smokers based on the number of cigarettes smoked per day.

Smoking Level	Number of Cigarettes per day
Trivial	Less than 1
Light	1-9
Moderate	10-19
Heavy	20-39
Very Heavy	40 or more

As both smoking statuses were recorded in the same consultation, they were crossed checked against one another. In the case of any discrepancies between the data, the worst case scenario or record giving the maximum amount of information on smoking status was taken; creating a single smoking status variable. Using this information, a binary variable was created assigning 0 to non-smokers and ex-smokers and 1 to current smokers.

If smoking status was not found in either file, but smoking amount was given and was more than 0, the patient was assumed to be a smoker.

Further, codes indicating health promotion of smoking or smoking advice, assumed the patient was a smoker. Passive smokers with no indication of smoking were assumed to be non-smokers.

Outliers were removed using methods described below (Outliers).

4.2.4.3 Derivation of Additional Risk Factors

Each risk prediction model listed in Table 4.2 also had optional additional variables. This ranged from deprivation scores to liver enzymes, and in some cases included family history. Data permitting, these variables were also included in the study to assess the value of family history in addition to the fuller set of risk factors for each risk model.

Additional variables were chosen to be included in the analysis based on the strength of their association with the CVD outcome, both in the unadjusted univariate and adjusted multivariate analysis. The strength of the association was based on a significance level of 0.05.

4.2.4.3.1 Impaired Glucose Regulation (IGR)

Patients were identified with impaired glucose regulation based on impaired glucose tolerance and impaired fasting glucose. Values for both conditions are used to represent the abnormal glucose regulation; between normal and diabetes. However, both conditions can be mutually exclusive or occur together. Thus, values for both conditions were used to identify IGR.

Patients were identified to have an IGT diagnosis using developed Read Code lists. Entity codes were used to derive fasting blood glucose and two-hour blood glucose values. These were also used as part of an oral glucose tolerance test; if the patients fasting glucose value was below 7mmol/l and the two-hour blood glucose value was between 7.8 and 11.1mmol/l, they were diagnosed with IGR for the study.

Further, if the patient had impaired fasting glucose (fasting blood glucose between 6.1 and 6.9mmol/l) on two separate occasions, they were also diagnosed with IGR for the study.

A sensitivity IGT variable was also created using HbA1c values and random blood glucose values. These were again identified using entity codes and data recorded in the test file. If the HbA1c value was between 42 and 47 mmol/mol or the random blood glucose value was between 7.7 and 11.1mmol/l, the patient was diagnosed with IGR.

IGR was presented as a binary variable, indicating its diagnosis for each patient. If this was not recorded, the patient was assumed to not have IGR.

4.2.4.3.2 Fasting and Non-Fasting blood glucose

Fasting and non-fasting blood glucose values were presented as continuous variables with mmol/l unit. Outliers were removed using methods described below (Outliers).

4.2.4.3.3 Triglyceride levels

Triglyceride values were derived from the test file using entity codes. This was treated as a continuous variable in the analysis. Outliers were identified based on the distance between the observed value and the mean triglyceride level, as per the methodology outlined below (Outliers).

4.2.4.3.4 Diabetes

Diabetic patients were identified by diagnoses recorded in the clinical file using a Read code list. Records were also searched in the additional clinical details files using entity codes indicating the patient was on the diabetic register (entity code=18) or giving their current diabetes status (entity code=26). Diabetics were then identified based on their HbA1c and random blood glucose test results; a patient was classed as diabetic if their HbA1c value was greater than or equal to 48 mmol/mol or random blood glucose was greater than or equal to 11.1 mmol/L.

4.2.4.3.5 Deprivation

Patient's deprivation score was derived from linked data. The level of deprivation was defined using the Townsend Index, derived from the patient's postcode and combining 4 census variables: car ownership, household size, owner occupation, employment status. Each postal enumeration district (around 150 households) is assigned a Townsend Deprivation score. These are divided into national quintiles and in the CPRD database patients are assigned a quintile score according to the enumeration district in which they live. This score ranges from 1–5, where a higher score indicates greater deprivation.

4.2.4.3.6 Ethnicity

Ethnicity for each patient was derived from the clinical files, based on Read codes. Further information was found in the linked HES files. In the case where data from clinical and HES files did not match, the HES data was used and considered the 'gold standard' and used due to its standardised collection and completeness. Patients were identified as White, White, other/mixed, Indian, Pakistani, Bangladeshi, Asian, other/mixed, Black Caribbean, Black African, Black, other/mixed, Chinese, Chinese, other/mixed, European, European, other/mixed, Other/mixed, and Ethnic NOS. However, for the analysis this was collapsed to a lower categorical variable; with White, South Asian, Black and Other/Mixed groups.

4.2.4.3.7 Left Ventricular Hypertrophy

Patients with LVH were identified using Read Codes, and were diagnosed based on ECG results.

4.2.4.3.8 Physical Activity

Physical activity for each patient was primarily derived using the additional clinical details file and related entity codes. This provided the type of exercise for each patient; inactive (0 times per week), gentle (1 time per week), moderate (2-3 times per week) and vigorous (more than 3 time per week) exercise. Additional information was sought from the clinical file which presented results from the GPPAQ and FITT physical activity assessments.

4.2.4.3.9 Alcohol Consumption

Alcohol consumption and drinking status was derived using a similar approach to smoking status. Both the clinical and additional clinical details files were used to create a variable based on the Read codes and another based on the entity codes. Patients were

classed as tee-total; ex-drinker or current drinker. This variable was then dichotomised to a binary variable indicating non/past (0) and current drinkers (1).

Current drinkers were then classed as trivial; light; moderate; heavy; or very heavy drinkers based on their alcohol consumption (alcohol units) per week. Further details of this classification are given in Table 4.5.

Table 4.5 Criteria to class drinkers as trivial, light, moderate, heavy and very heavy drinkers based on the number unit of alcohol consumed per week.

Alcohol Level	Units of alcohol consumed per week
Trivial	Less than 1
Light	1-2
Moderate	3-6
Heavy	7-9
Very Heavy	Above 9

If a patients drinking status was not recorded but their alcohol consumption was given, they were classed as a current drinker. Similarly, if alcohol health promotion or alcohol advice was indicated, patients were assumed to be current drinkers.

Outliers were removed using methods described below (Outliers).

4.2.4.3.10 Body Mass Index (BMI)

Height, weight and BMI were derived using the additional clinical file using entity codes. If any one of these measurements were missing, the other two would be used to calculate the third measurement.

Height was measured in metres (m) and weight was measured in kilograms (kg). Both variables were treated as continuous for the analysis. Outliers were removed using methods described below (Outliers).

4.2.4.3.11 Glycated Haemoglobin (HbA1c)

HbA1c was derived using entity codes in test file. The units of the values were either from the Diabetic Control and Complication Trial (DCCT) or the International Federation of Clinical Chemistry (IFCC) and were given in % or mmol/mol, respectively. For the analysis all values were presented in mmol/mol.

Outliers were removed using methods described below (Outliers).

4.2.4.3.12 High Sensitivity C-reactive Protein (hsCRP)

Values for hsCRP are not recorded in CPRD; however CRP is and was identified using the test file and entity codes.

Outliers were removed using methods described below (Outliers).

4.2.4.3.13 Liver Enzymes

The liver enzymes used in the study were ALT, AST and GGT. Data for these were found in the test file and identified using entity codes. The units for these values were iU/L, and outliers were removed using methods described below (Outliers).

4.2.4.3.14 Chronic Kidney Disease

Chronic kidney disease is included in the study as both a continuous and binary measure. Both forms of the risk factor was derived from the estimated glomerular filtration rate (egfr) which was calculated from the serum creatinine measurement found in the test file and identified using entity codes. The units for serum creatinine were umol/L, and outliers were removed using methods described below (Outliers).

Estimated glomerular filtration rate was calculated based on age, gender, ethnicity, and serum creatinine measurement using Eq.4.2:

$$egfr = 175 \times \left(\frac{serum\ creatinine}{88.4}\right)^{-1.154} \times (age^{-0.203}) \times 0.742(if\ female) \times 1.210(if\ black) \quad \text{Eq.4.2}$$

Age was calculated at time serum creatinine was taken using the patient birth date and date of serum creatinine test; it is expressed in years. Patients were classed as black if they were Black-Caribbean, Black-African, or Black-mixed. Units of the egfr were mL/min/1.73m^2.

The egfr was dichotomised according to guidelines to indicate CKD (Stage 3, 4, and 5), given in Table 4.6.

Table 4.6 Criteria to class stages of Chronic Kidney Disease, based on estimated glomerular filtration rate (mL/min/1.73m^2).

CKD	Estimated Glomerular Filtration Rate (egfr)
Stage 1	egfr≥90
Stage 2	60≤egfr≤89
Stage 3	30≤egfr≤59
Stage 4	15≤egfr≤29
Stage 5	egfr<15

4.2.4.4 Derivation of New Risk Factors – Family History

Family history of CVD was identified using the selected Read codes (Appendix D) which included codes with details on CVD, relative and age of onset. Read codes were also included if they only specified a FH or relative, without details on disease. This was combined with the CVD Read code list to identify more FH of CVD patients. The additional clinical file was also used with entity codes to further identify more patients and family history details.

All patients included in the study have a family history recorded, however the variable used in the analysis illustrates whether this is a positive or negative family history. I.e. ‘no family history’ has to be recorded for the family history to be negative. This is presented

as a binary variable. For a positive family, the minimum amount of information required was family history of CVD, regardless of any additional details given.

Additional details were searched for in patients with a positive family history to create a detailed FH risk factor. Data was searched for either the age or relationship (relative) in the observation. This created a categorical variable giving the level of information recorded with the family history. Patients were categorised into the following:

1. Family history of CVD
2. Family history of CVD with age of onset given
3. Family history of CVD with relative specified
4. Family history of CVD with age of onset and relative given

4.2.4.5 Outliers

Outliers were identified and removed for all continuous variables in the analysis; using two recognised methods (102).

The first method applied of acceptable ranges from an external source of data. Raw data from The Health Survey for England 2011 (103, 104) was used to identify age- and gender- specific ranges for height, weight, BMI, total cholesterol, HDL cholesterol, systolic BP, diastolic BP, and HbA1c. Age was grouped into 10-year age bands: 16-24; 25-34; 35-44; 45-54; 55-64; 65-74; ≥75. Ten-percent was added and subtracted to the maximum and minimum values of the range, respectively; data within this study were recorded as missing if the value fell outside of these adjusted age- and gender- specific ranges. Data for other continuous variables included in this study were not found.

The second method was applied on all variables for which data could not be found in the Health Survey for England 2011. These were: smoking amount; alcohol consumption amount; fasting and non-fasting blood glucose values; CRP; and liver enzyme values. This method used the standard deviation from the mean to identify outliers in the data. Looking at the distribution for each variable, any measurement more extreme than the mean ±5 times the standard deviation was recorded as missing in the data.

4.2.5 Statistical Analysis

4.2.5.1 Descriptive Analysis

The extraction of the dataset from CPRD was first described in line with the inclusion and exclusion criteria of the study. This was presented in a CONSORT style diagram. The characteristics of the extracted cohort were then described; as a complete dataset. A descriptive analysis was performed, reporting number (%), mean (SD) and median (IQR) for categorical, normal continuous and non-normal continuous variables, respectively. Missing values were also presented.

The overall cohort was then described using the same methods, presenting information by patient family history status (positive/negative). Statistical tests, such as chi-squared tests for binary data, and t-tests to compare the means of normally distributed continuous variables between groups (the Mann-Whitney U test will be used for non-normal continuous data) were used to test the significance of the association. A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

The incidence of CVD during the follow up period was then described by comparing the baseline characteristics of patients who suffered a CVD event during follow up to those who did not. Associations were again tested for using the method described above. Crude

incidence rates of CVD in the data were also used to describe patterns of CVD in the data. This was presented by all categorical risk factors, using 10 year age categories for age.

4.2.5.2 Generalisability of the data

The extracted cohort was compared to the Health Survey for England 2011 (103, 104) population to evaluate its generalisability to the UK population. Comparable variables in the HSE dataset were age, gender, total cholesterol, HDL cholesterol, blood pressure (SBP and DBP), height, weight, BMI, current smoking status and HbA1c values.

Firstly, the amount of data available in each of these variables was presented and compared; overall, and by gender. Mean values for variables in continuous format were derived from both datasets and plotted for each 10 year age group; proportions were used for binary variables, such and smoking and CVD event.

Specific comparisons were made, to mean values and proportions reported for individuals in the age group containing the mean age of the extracted CPRD patients. P-values were reported from chi-square tests for categorical data, t-tests for continuous normally distributed data, and Wilcoxon Rank Sum test for non-normally distributed data. A significance level of 0.05 was used.

To check the generalisability of the incidence rate of CVD event in the study, it was compared to QRISK2 validation study (74) which used a similar cohort to assess the performance of QRISK2 for 10 year CVD risk prediction.

Patients for this study were extracted from a The Health Improvement Network (THIN) database and consisted of patients without a pre-existing history of CVD, aged 30-85 years, and with no prescription history of statins. The outcome measure used was comparable to the CPRD study; the first diagnosis of CVD (including myocardial

infarction, angina, coronary heart disease, stroke, and transient ischemic heart disease) recorded in the GP database.

4.2.5.3 Missing Data

Missing values were presented in the descriptive analysis, however were imputed for the CVD risk factors of the study in the main analysis.

Missing data of standard CVD risk factors were first described and patterns of missing data were checked. Indicator variables to describe missing data of each standard risk factor were created, on which a logistic regression analysis was performed. Based on these findings, missing data was described as missing completely at random (MCAR), indicating the complete cases are a random sample (probability a value is missing is independent of the observed and unobserved data); missing at random (MAR), indicating the complete cases are not a random sample (probability a value is missing is dependent on the observed data); or missing not at random (MNAR), indicating again the complete cases are not a random sample (probability a value is missing is dependent on the unobserved data).

Missing data was also classed as either monotone or arbitrary. A monotone missing data pattern would assume that if a variable is present for a patient, then all preceding variables for that patient would also not be missing; indicating some order/pattern to the missing data. Otherwise data was said to be non-monotone. A dataset with either type of missing data pattern is said to be arbitrary.

Pair-wise correlations of missing data in each of the variables to be imputed were checked against other risk factors in the study. Associations between missing data of incomplete risk factors were investigated against other risk factors of the study using logistic

regression. This was to better identify the missing data for each variable and use the related risk factor in the regression model to impute data.

4.2.5.4 Multiple Imputation

Missing data was imputed using the entire extracted cohort, before it was split into the derivation and validation datasets. Ten imputed datasets were created; each dataset was analysed using standard methods and the results were combined to produce estimates and confidence intervals incorporating uncertainty from missing data using Rubin's Rules (105).

Data was imputed using the `mi impute` command in Stata (106) for standard risk factors, additional CVD risk factors showing statistical significant associations with the CVD outcome and with less than 50% missing. Logistic, ordinal logistic, and multinomial logistic regression was used to impute for the incomplete binary, ordered categorical and unordered categorical risk factors, respectively; and linear regression was used to impute for incomplete continuous risk factors. Predictors for the imputation also included age, diabetes and hypertensive treatment at baseline and the outcome CVD variable. Imputations were created by gender.

4.2.5.5 Derivation and Validation

The extracted cohort was randomly split into two datasets. This was achieved by randomly assigning a value between 0 and 1 to each patient, using a uniform distribution. Patients with a value below and equal to 0.33 were included in the validation dataset and patients with values above 0.33 were used in the derivation dataset. By using the 0.33 threshold, the data was split, using two thirds of the data as the derivation dataset and a third as the validation dataset (9, 11, 31, 83).

A descriptive analysis was again performed on these datasets to compare their characteristics; missing values were also presented.

4.2.5.6 Derivation of Models

Using the training dataset, the primary analysis created four CVD risk prediction models based on the standard, additional and new (FH) risk factors:

- Model 1: standard CVD risk factors
- Model 2: standard CVD risk factors and family history
- Model 3: standard CVD risk factors with additional variables
- Model 4: standard CVD risk factors with additional variables and family history

4.2.5.6.1 Exploration of the data

Using the non-imputed data, associations and potential confounding was investigated between risk factors, using logistic, ordinal logistic, multinomial logistic and linear regression; for binary, ordinal categorical, unordered categorical and continuous risk factors, respectively. A Bonferroni adjusted significance level of 0.004 was used (0.05/12).

4.2.5.6.2 Model Building

The imputed datasets were used to build each model. Cox proportional hazards regression was used to derive these models to account for the time to event data and also to account for patient dropouts through censoring.

Results from the models were described, presenting the beta coefficient with p-value for each risk factor and the hazard ratio with a 95% confidence interval.

An unadjusted analysis was first performed to identify the strength of association for each risk factor individually, with the CVD outcome. Standard risk factors, FH, and additional risk factors showing a significant association ($p<0.05$) with CVD were then used to derive Models 1-4 in the multivariate analysis.

As part of the internal validation, using the derivation dataset only, the fit of each model was described using the Akaike Information Criteria (AIC) and Bayes' Information Criteria (BIC); where lower values show improvement in the fit of the model.

The fit of nested models (model 1 and 2, and model 3 and 4) were also compared using Likelihood Ratio tests. A p-value less than 0.05 would reject the null hypothesis that the difference in the likelihood ratio for each model was equal to 0; an insignificant improvement in model fit was found.

Regression diagnostics were assessed for each model of the analysis. Specifically the proportional hazards (PH) assumption was checked for each risk model. Log-log plots of survival for each categorical risk factor in the model were checked; where parallel lines between categories indicated no violation of the PH assumption. Post-estimated Kaplan-Meier curves plotted the observed and predicted survival for each categorical risk factor in the model; the closer these lines were, the less likely it was that (PH) assumption was being violated. A proportional hazards test using Schoenfeld residuals was used to test the assumption for all variables in the model and also present an overall p-value for the model. A p-value less than 0.05 indicated the assumption was not being met by the model/risk factor.

Finally, time varying covariates were also added to each model to check for PH assumption; a p value less than 0.05 would indicate a violation of the assumption. If the

PH assumption was not met, each risk factor found to violate this assumption was added to the model as a time varying covariate.

Cox-Snell residuals were used to assess the overall goodness of fit of the model Martingales residuals were used to assess the functional form of each risk factor; and Deviance residuals were used to identify outliers. Outliers and influential observations were also assessed using dfbeta values, likelihood displacement values, and LMAX values; these were plotted against analysis time. Outliers and influential observations were highlighted in the study and a sensitivity analysis was performed removing these observations. This was checked for all imputed datasets individually.

4.2.5.7 Creating the Risk Prediction Model

Once all 4 models were finalised, their beta coefficients were used to create 4 risk prediction models and calculate risk of CVD for each patient, using Eq. 4.3:

$$\widehat{risk} = 1 - S_0(t)^{exp(\sum_{i=1}^{p} \beta_i X_i - \sum_{i=1}^{p} \beta_i \bar{X}_i)} \qquad \text{Eq.4.3}$$

Where $\mathbf{S_0(t)}$ is the baseline survival rate at follow up time, **t**(for this example, a follow up time of 10 years will be used); beta ($\boldsymbol{\beta_i}$) are the regression coefficients for each included risk factor in the model (**i**); $\mathbf{X_i}$ is the observed data value for each risk factor; $\mathbf{\bar{X}_i}$ is the corresponding mean for each risk factor; and **p** is the total number of risk factors included in the model.

4.2.5.8 Validation of the Models – Current Methodology

The derived risk models were validated internally using the validation dataset. The metrics used to describe their performance; assessed the calibration, discrimination, reclassification, and clinical usefulness of each model; comparing models 1 and 2, and

models 3 and 4 to assess the additional value of FH information. The difference between performance measures were used to quantify the improvement.

4.2.5.8.1 Calibration

Calibration is a measure of how well the predicted risk agrees with observed risk in the data. When the mean predicted risk within subgroups of the data matches the proportion of cases in the data, the model is considered well calibrated. This was assessed visually using calibration plots; plotting the mean predicted risk and observed risk of CVD at 10 years was also calculated, using the 10 year Kaplan-Meier estimate. The ratio of the predicted to the observed CVD risk was then plotted against each tenth of predicted risk. Ideally these observations should overlay on one another to indicate good calibration.

4.2.5.8.2 Discrimination

Discrimination measures how well the model differentiates between those who go on to experience the outcome or not. This was based on using the NICE recommended clinical risk threshold of 10% (3). Using this threshold and knowing the observed outcome for each patient, the performance of the model can be tabulated as shown in Table 4.7.

Table 4.7 Performance results from the model, compared to the observed outcome, using a pre-specified risk threshold.

		Predicted Outcome	
		High Risk	**Low Risk**
Observed Outcome	**Disease**	a (true positive)	b (false negative)
	No Disease	c (false positive)	d (true negative)

The components to assess discrimination were first derived from the models using values given in Table 4.7. Sensitivity and specificity were used to describe the true positive and

negative rates from using each model, respectively. Explicitly, the sensitivity describes the number of patients the test identifies correctly; i.e. patients who are test positive and are observed to be outcome positive. Specificity describes the number of patients who tested negative from the model and do not experience the outcome within the follow up time. Formulae for sensitivity and specificity are given below [Eq.4.4, Eq.4.5].

$$Sensitivity = \frac{no.of\ true\ postives}{total\ no.of\ diseased\ patients} \quad \text{Eq.4.4}$$

$$Specificity = \frac{no.of\ true\ negatives}{total\ no.of\ \ non\ diseased\ patients} \quad \text{Eq.4.5}$$

The positive (PPV) and negative (NPV) predicted values described the proportion of outcome positive and negative patients out of patients classed as outcome positive and negative from the model. I.e. the PPV is the proportion of outcome positive patients given they have tested positive from the risk model. Similarly, the NPV it the proportion of patients who do not have the outcome during the follow up and test negative from the model. Formulae for the PPV and NPV are given below [Eq.4.6 Eq.4.7].

$$PPV = \frac{no.of\ true\ postives}{total\ no.of\ high\ risk\ patients\ predicted\ by\ the\ model} \quad \text{Eq.4.6}$$

$$NPV = \frac{no.of\ true\ negatives}{total\ no.of\ low\ risk\ patients\ predicted\ by\ the\ model} \quad \text{Eq.4.7}$$

The sensitivity and specificity was then used to construct the Receiver Operating Characteristic (ROC) curve, which plots the sensitivity (true positive rate) against 1-specificity (false positive rate) across all possible risk thresholds for each model. The area under each curve (AUC) was presented by Harrell's C-index (107). Both are rank order statistics used for interpretation for pairs of patients with and without the outcome. It gives the probability that using these pairs; the model assigns a higher risk to patients with

the outcome, than the patient without. Hence, larger values in both indicate better discrimination.

4.2.5.8.3 Reclassification

The ability of FH information to reclassify patients from their respective CVD risk categories was assessed using Reclassification Tables (108, 109). For this, patients were classed into low and high risk groups based on the NICE recommended 10% risk threshold. A patient with risk below 10% were considered low risk of CVD, and patients with risk above or equal to 10% were classed as high risk of CVD. Clinically, high risk patients would be given treatment to prevent the prospective outcome.

Using these patient risk classifications, the ability of FH to reclassify patients between risk groups was assessed by comparing Model 1 (standard CVD risk factors only) to Model 2 (standard CVD risk factors with FH); and Model 3 (standard and additional CVD risk factors only) to Model 4 (standard and additional CVD risk factors with FH). This was presented using reclassification tables; a cross tabulation to compare the risk classification from 2 models; explicitly shows the shift of patients between their risk categories. From these, the total number (%) reclassified was calculated, followed by the Net Reclassification Index (NRI); a metric to compare the classification from 2 models by outcome, providing a net calculation of the changes in the right direction. For example, if a patient who does have the outcome is reclassified to a lower risk group, or a patient with the outcome is reclassified to a higher risk group, the reclassification would be considered in the right direction. The formula used to calculate the NRI is given below [Eq.4.8]:

$$NRI = \big(P(up|event) - P(down|event)\big) + \big(P(down|nonevent) - P(up|nonevent)\big) \quad \text{Eq.4.8}$$

where

$$P(up|event) = \frac{number\ of\ events\ moving\ up}{number\ of\ events}$$

$$P(down|event) = \frac{number\ of\ events\ moving\ down}{number\ of\ events}$$

$$P(down|nonevent) = \frac{number\ of\ nonevents\ moving\ down}{number\ of\ nonevents}$$

$$P(up|nonevent) = \frac{number\ of\ nonevents\ moving\ up}{number\ of\ nonevents}$$

A category-less form of the NRI was also calculated; the Integrated Discrimination Improvement (IDI) integrates the NRI over all possible cut offs for the risk of CVD and is also the difference in discrimination slopes between models with and without the FH risk factor. The formula used for this calculation is presented below [Eq.4.9].

$$IDI = \left(\bar{P}(new|event) - \bar{P}(new|nonevent)\right) + \left(\bar{P}(old|nonevent) - \bar{P}(old|event)\right) \quad \text{Eq.4.9}$$

where $\bar{P}(new|event)$, $\bar{P}(new|nonevent)$, $\bar{P}(old|nonevent)$ *and* $\bar{P}(old|event)$ are the mean predicted risk of the new and old model of an event for patients who do and do not experience the event, respectively.

4.2.5.8.4 Decision Curve Analysis

The novel method of a Decision Curve Analysis (24) was also adopted and extended to assess the clinical usefulness of each model and to assess the added value of FH information, given a pre-specified risk threshold (10%). This analysis will be used to calculate the Net Benefit (NB) for the new (with FH) and original (without FH); which is the net number of true positives gained by using a model with FH (Model 1 and 3) compared to a model without (Model 3 and 4) at the 10% risk threshold. This will be presented using the Decision Curve Analysis (DCA) plot. An increase in NB indicates a better performing model.

Explicitly, the NB [Eq.4.10] would calculate the percentage of true positives identified by the model without an increase in the number of false positive.

$$NB = \left(\frac{number\ of\ true\ positives}{total\ sample} - \frac{number\ of\ false\ postives}{total\ sample}\right)\left(\frac{p_t}{1-p_t}\right) \qquad \text{Eq.4.10}$$

where n is the number of patients and p_t is the pre-specified risk threshold.

4.2.5.9 Validation of Models – Refined Methodology

4.2.5.9.1 Cost-effectiveness analysis - Refined Methodology

The final part of validating the risk models and evaluating the added value of FH information will be a Cost-effectiveness analysis (CEA) using a Decision Tree; as outlined in Chapter 3, Section 3.3.

To ensure a robust cost-effectiveness analysis the CHEERS checklist (66) was used alongside the NICE Reference Case (65); this consolidates all health economic reporting guidelines. It was used to develop the refined methodology in Chapter 3; and was used to create detailed, study specific methodology.

4.2.5.9.1.1 General Methods

The target population were taken from a primary care setting and have been described in Section 4.2.3. The risk prediction models were not developed for any subgroup.

The perspective taken for the analysis was from the UK based National Health Service (NHS) and the Personal Social Services Research Unit (PSSRU) (70); and the comparators for risk prediction models with (Model 2 and 4) and without (Model 1 and 3) FH information.

The time horizon used for this study reflects the follow-up period used for the developing the risk prediction model; 10 years. Costs were discounted in the primary analysis using a

3.5% discount rate (67); the government discount rate. A sensitivity analysis was also performed using a 0% discount rate.

4.2.5.9.1.2 Outcomes

The outcomes of choice for the cost-effectiveness analysis were performance related measures. In particular, Harrell's C-Index (AUC) (48, 49) and the Youden Index (68), common measures of discriminatory accuracy were used. The AUC was derived from current validation methods described earlier (Section 4.2.5.8.2), and the Youden Index was calculated from the output of the decision tree, as described in Chapter 3, Section 3.3.2.1.

The measurements of effectiveness were based on a single study; the respective validation study. Preference based outcomes were not used, reasons for this are provided in Chapter 3, Section 3.2.1.1.

4.2.5.9.1.3 Costs

Three sets of costs were needed for the analysis; for the risk prediction model; the CVD outcome; and the treatment. Costs were UK costs, for the costing year of 2013-2014. All three sets of costs were derived from the NICE Lipid Modification Economic Evaluation and Costing Report (26), the PSSRU 2014 (70), and the Prevention of Cardiovascular Disease Costing Report (7). Further details are provided below.

4.2.5.9.1.3.1 Cost of Risk Prediction Model

The process of assessing the risk of a patient is outlined as follow:

- Initial consultation with GP - includes collection of age, gender, current smoking status, blood pressure. Additional time would be allowed to collecting basic and

detailed CVD FH information. The GP is also assumed to request a lipid profiling blood test sample during this consultation.

- Lipid Profiling Blood Test - Patient goes to different part of the GP surgery/health centre, where a nurse takes the blood sample.

- Follow up GP consultation - Patient returns to GP to discuss results and, if required, preventative treatment for CVD.

The cost of using a risk prediction model and collecting its data was assumed to take one GP appointment and a lipid profile. Unit costs and average consultation lengths associated with practice nurses and GPs was extracted from PSSRU Unit Costs of Health and Social Care 2014 (70). An outline is given in Table 4.8.

Table 4.8 Cost of GPs as provided by the PSSRU (70).

	GP
Average Length of Surgery Consultation	11.7 minutes
Cost per minute of patient contact	£3.90
Cost per hour of face to face contact	£234

Thus one GP appointment would cost £45.63. A cost of £2.00 was taken from the NICE lipid modification costing report (26), for a lipid profile per patient. In total, the cost of using the risk prediction model collecting only standard risk factors was £47.63 per patient. Collecting additional, detailed FH information was assumed to add an additional half of a consultation, costing an additional £22.82 (total cost: £70.45).

Cost of additional risk factors was assumed to be absorbed into existing costs of using the risk prediction model, thus the cost for Model 3 (including standard and additional risk factors) was assumed to be the same as Model 1 (standard risk factors alone); and the cost of Model 4 (standard, additional and FH risk factor) was assumed to be the same as

Model 2 (standard and FH risk factor). If however, the additional risk factor required a blood test, this was included as an additional lipid profile cost of £2.00.

The NICE lipid modification report assumed any additional GP appointments would be consumed in existing resources (26). This was also assumed for this study.

4.2.5.9.1.3.2 Cost of treatment

After the assessment of CVD risk, patients were identified as low or high risk of CVD based on the 10% risk threshold. In line with current guidance (26), it was assumed all patients at high risk of CVD were treated with lipid modification drug therapy.

Currently, five statin therapies are prescribed to prevent CVD events. These are Atorvastatin, Fluvastatin, Pravastatin, Rosuvastatin, and Simvastatin. Of these, Atorvastian has been recommended as the primary prevention treatment for CVD (3). As per the NICE lipid modification recommendations, Table 4.9 shows the estimated prescribing proportions for each of these therapies used in the NICE lipid modification costing report (26), as well as the costs, which were taken from the Electronic Drug Tariff (110).

Table 4.9 Lipid Modification Prescribing Proportions (26).

Lipid Modification	Weighted Cost of Drug (£)	Prescribing Proportion (%)	Weighted Prescription Cost (£)
Atorvastatin	18.48	60	11.09
Fluvastatin	35.79	0.5	0.18
Pravastatin	20.02	5.5	1.10
Rosuvastatin	268.46	3	8.05
Simvastatin	13.32	31	4.13
		Weighted Cost of Treatment	**24.55**

Based on the estimated prescribing proportions and the weighted cost of each drug therapy, the weighted cost of treatment was estimated as £24.55 per patient per annum.

4.2.5.9.1.3.3 Cost of CVD Event

The outcome of a CVD event included Myocardial Infarction, Transient Ischaemic Attack, Stroke, Heart Failure, and Angina. The weighted cost for each outcome is presented in Table 4.10 and is derived from the NICE Prevention of CVD Costing Report (7).

Table 4.10 Weighted cost of a CVD event (7).

HRG description	Unit cost of event (£)	*Number of events	Proportion of total %	Weighted Outcome Cost (£)
Angina (people aged over 69)	3,475	1951	18.62	647.05
Angina (people aged under 70)	1,475	739	7.05	103.99
MI without complications	5,055	2657	25.36	1,281.95
MI with complications	11,459	1033	9.86	1,129.86
TIA (people aged over 69)	1,339	645	6.16	82.48
TIA (people aged under 70)	1,339	222	2.12	28.39
Stroke (people aged over 69)	4,348	2594	24.76	1,076.56
Stroke (people aged under 70	4,348	635	6.06	263.49
			Weighted Cost of treatment	**4,613.76**

***The number of CVD events is an annual distribution of events avoided taken from CG67 Lipid modification, source: Primary CVD events table 49, Ward S, Lloyd Jones M, Pandor A et al (2007) (111). This has been applied as a distribution for calculating the weighted average cost of events. For MI with or without complications the average non elective mandatory tariff for 2010 has been used as this relates to surgical procedures for MI.**

The weighted cost of a CVD event was estimated to be £4,613.76.

The cost of treatment and the CVD event was then combined to derive the cost of an untreated event from a low risk patient; the cost of a treated event from a low risk patient; the cost of an untreated event from a high risk patient; the cost of a treated event from a high patient.

To calculate the cost of each of these scenarios, the CVD event rate was extracted from two clinical trials, the ASCOT LLA Study (57) and the CARDS Study (58), investigating the effect of 10mg of Atorvastatin vs placebo in patients with and without a history of CVD (used as a proxy for high and low risk of CVD), respectively. The event rates from these studies are provided in Table 4.11.

Table 4.11 CVD event rates from the ASCOT LLA and CARDS study.

	No. CVD	Total No. Participants	CVD Event Rate*
Untreated Low risk (CARDS)	127	1412	0.09 (0.08 to 0.11)
Untreated High risk (ASCOT)	640	5137	0.12 (0.12 to 0.13)
Treated Low risk (CARDS)	83	1429	0.06 (0.05 to 0.07)
Treated High risk (ASCOT)	489	5168	0.09 (0.09 to 0.10)

***event rate is relative for the individual study follow-up period; 3.3 years for ASCOT Study; 3.9 years for CARDS Study.**

Using methods described in Fleurence et al (69), the CVD annum rate and annual transition probabilities were calculated. These were then combined with the weighted CVD cost and cost of treatment to give the average cost of an untreated CVD event in low risk patients; a treated CVD event in low risk patients; an untreated event in high risk patients; and a treated event in high risk patients over a 10 year follow up. These costs are provided in Table 4.12.

Table 4.12 Average cost of treated and untreated CVD events over 10 years follow up, by discount rate.

	0% discounted Average Cost (£)	3.5% discounted Average Cost (£)
Un-treated Low Risk	1,101.16	916.16
Untreated High Risk	1,823.30	1,516.37
Treated Low Risk	947.99	788.40
Treated High Risk	1,614.52	1,342.73

4.2.5.9.1.4 Model-based Economic Evaluations

A decision tree was chosen to compare two risk prediction models. Reasons for this have been described in Chapter 3, Section 3.2.2.1.

For the analysis it has been assumed all patients at risk are treated. Further, the treatment is according to the estimated prescribing proportions found in the NICE lipid modification guidelines.

4.2.5.9.1.5 Deterministic Analysis

4.2.5.9.1.5.1 Decision Tree

The proposed analysis has been described in full in Chapter 3, Section 3.3.5. Figure 4.1 presents the format for the decision tree used to evaluate the added value of FH to CVD risk prediction; including standard, and standard and additional risk factors. A 10% risk threshold was used to identify patients at high risk.

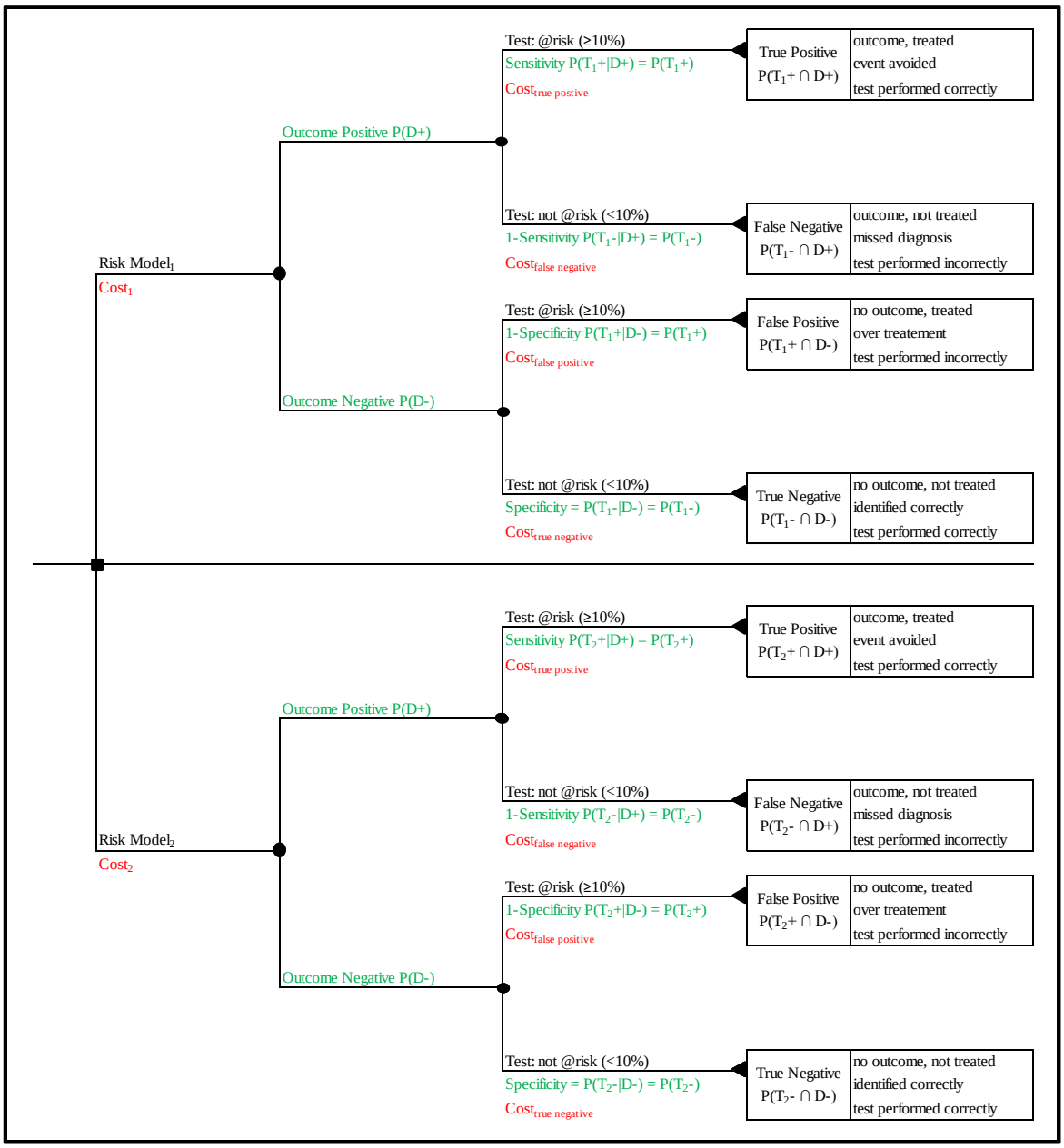


Figure 4.1 Decision Tree to compare two CVD risk prediction models.

4.2.5.9.1.5.2 Input and Output of the Decision Tree

The probabilities for each event node were calculated from the current performance measures in the validation dataset; methods for this are outlined in Section 4.2.5.8.2. Namely, the proportion of patients classed at high $P(T_i+)$ and low $P(T_i-)$ risk, and the sensitivity $P(T_i + \mid D_i+)$ and specificity $P(T_i - \mid D_i-)$ was used for each model, where i is the derived model number ($i = 1,2,3,4$).

The output from the decision tree was the expected cost of each risk prediction model; which was based on the cost and probabilities associated with the performance of each model. This was calculated using Chapter 3, Section 3.3.5.1.

4.2.5.9.1.5.3 Incremental Cost-Effectiveness Ratio and Net Monetary Benefit

The output of the decision tree; the expected cost of using a risk model with and without FH information, was combined with the measures of discriminatory accuracy (Youden Index and Harrell's C-Index), using the Incremental Cost Effectiveness Ratio (ICER). This calculated the change in cost, divided by the change in effect between models with and without FH information.

Two ICERs were calculated as part of the analysis; the Youden Index and Harrell's C-Index. Both are summary measures derived from the ROC curve and describe discriminatory accuracy of a risk model. Both ICERs were interpreted as the cost per increment of discriminatory accuracy gained. These values were visually presented using the cost-effectiveness plane.

In addition, the ICER will be extended to calculate the incremental net monetary benefit (INB). As no performance related willingness to pay has been found, this will be presented over a range of thresholds (£0-£10,000).

4.2.5.9.1.6 Probabilistic Sensitivity Analysis (PSA)

Although the above deterministic cost-effectiveness analysis presented the ICER, an effect size for the added value of FH information to standard CVD risk assessment, it did not quantify the uncertainty.

A PSA analysis was also performed in addition to the deterministic analysis above. Based on the estimate for probabilities used in the decision tree and their standard errors, if

available, Beta distributions for probability values, and Gamma distributions for cost, were applied to the decision tree presented in Figure 4.1. If standard errors were not given, 10% of the mean was used to estimate this (112-114). Table 4.13 provides the information required for each cost value used in the PSA; values for the probabilities are presented in Chapter 6, Section 6.1.1.2.

Table 4.13 Deterministic mean costs of risk model (with and without FH), treatment and fracture; with standard error, alpha, and beta for estimation of probabilistic cost derived from a Gamma Distribution.

	Mean Cost (£)	SE	alpha	beta
Annual cost treated CVD	4,613.76	461.376	100	46.1376
Annual cost untreated CVD	4,613.76	461.376	100	46.1376
Treatment cost	24.55	2.455	100	0.2455
Cost of Model without FH	47.63	4.763	100	0.4763
Cost Model with FH	70.45	7.045	100	0.7045

Ten-thousand iterations were produced and the mean ICER over all iterations was presented with a 95% confidence interval.

4.2.5.9.1.6.1 Net Monetary Benefit

Further analysis was also performed as part of the PSA. An intervention is deemed as cost effective if the ICER is less than the amount a person is willing to pay; thus the willingness to pay acts as a threshold to the ICER. The incremental net monetary benefit (INB) accounts for both the ICER and the willingness to pay, such that if the INB is above 0 for a given willingness to pay, the intervention is considered cost-effective. The derivation of the INB formula is presented in Chapter 3, Section 3.3.5.1.2.

Similar to the ICER, two INBs were calculated. The first used the Youden Index as a measure of effect and the second used Harrell's C-Index value for each model. No

recommendations are available for the willingness to pay threshold and the value per incremental measure of effect, for which the intervention would be deemed cost effective. Hence the INB was plotted over a range of possible values for the willingness to pay (£0-£100.000); producing a Cost Effectiveness Acceptability (CEAC) Curve for both measures of effect. This presented the probability of the intervention (new risk model – with FH) being cost effective, compared to standard care (original risk model – without FH)

All analysis took place using Stata v12.1 (115) and MS Excel.

4.3 Chapter Summary

This chapter gives and introduction to CVD risk assessment in UK primary care, and outlines the rationale for adding FH information to standard and standard with additional CVD risk factors. It describes the methodology to evaluate the added value of FH information to CVD risk prediction. This includes the study design, methods to develop the CVD risk models, current methods to validate the models and evaluate the added value of FH information, and refined methods to further evaluate the clinical impact of the adding FH information to routine CVD risk assessment.

The next chapter presents the results from the data extraction; descriptive analysis; model development; and validation of the models, using current methodology.

Chapter 5 - Family History and Cardiovascular Disease Risk: Results and Discussion (Current Methodology)

The previous chapter gave the rationale and described the methodology for evaluating the addition of FH information to CVD risk prediction. The analysis was described in two sections; using current methods; and refined methods.

This chapter presents the results from carrying out the current methods described. This chapter begins with outlining the extraction of the data; describing its baseline characteristics; and assessing its generalisability to the UK. Data is then imputed using multiple imputation techniques before performing an unadjusted and adjusted analysis to derive the CVD risk prediction models with and without FH information.

Finally, the derived risk prediction models are validated using current performance measures; describing calibration, discrimination, reclassification, and net benefit derived from the decision curve analysis.

5.1 Results

5.1.1 Extraction

In September 2012, there were 12,288,790 acceptable patients contributing to CPRD registered with 652 acceptable practices. Of these, 637,212 patients did not have a history of CVD and had a Read Code to identify their CVD family history status; positive (‘family history of CVD’) or negative (‘no family history of CVD’).

Patients had to have had at least 1 year of follow up data and were excluded if they were no longer providing data to CPRD, or belonged to a practice that was not providing data to CPRD during the study period. They were also excluded if there did not meet the inclusion criteria for the study; not aged 30-79 years at baseline; had a pre-existing CVD condition; on lipid or hypertension treatment before baseline (identified if they had more than one prescription for treatment); or diagnosed with familial dyslipidaemia before baseline.

After accounting for the exclusion criteria of the study, 334,621 patients were removed from the study and 302,591 were eligible for the study from 378 practices. Figure 5.1 shows the selection of the cohort.

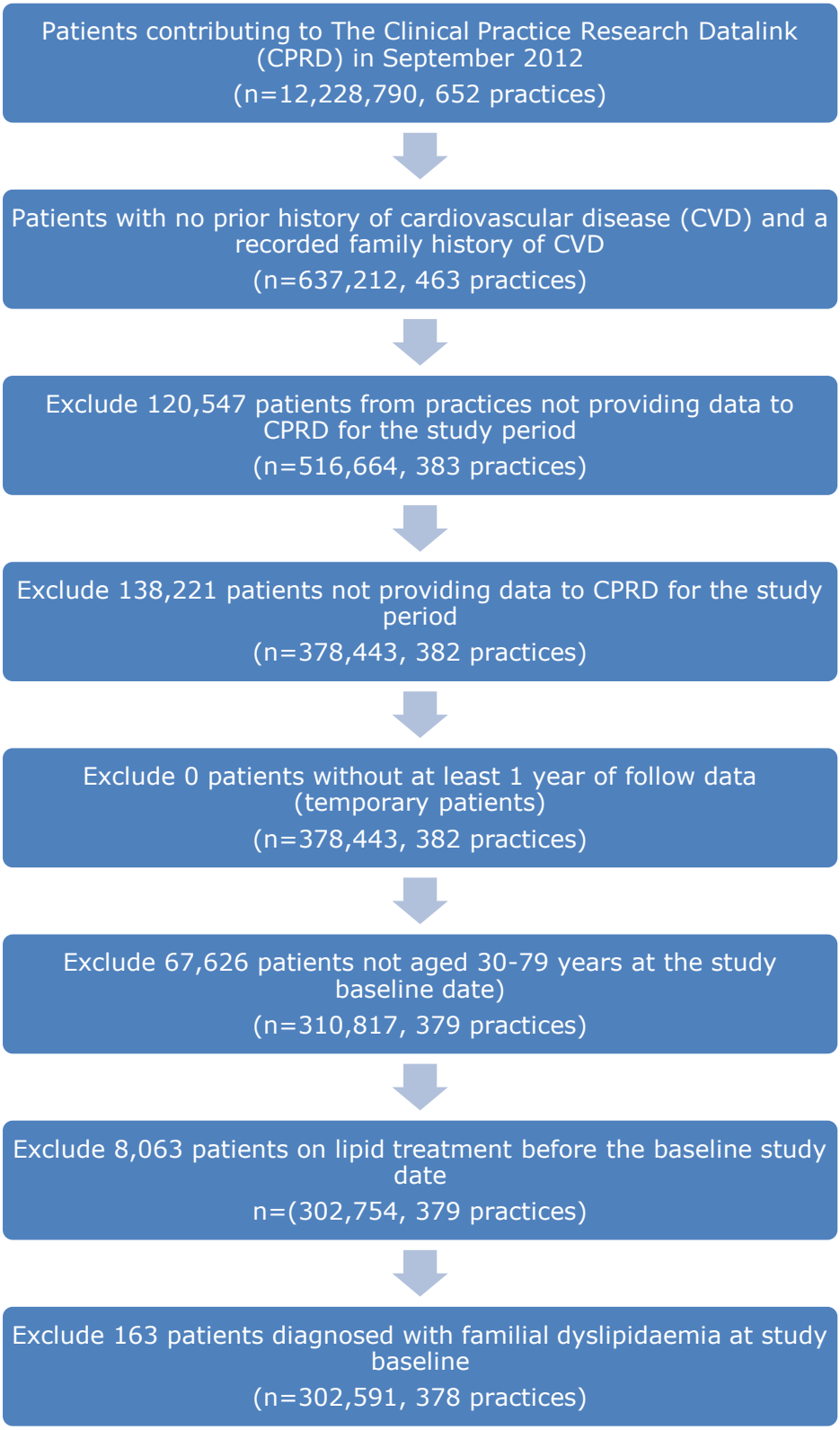


Figure 5.1 Selection of patients for the study cohort.

5.1.2 Overall Cohort Characteristics

Baseline characteristics of the overall extracted dataset, consisting of 302,591 patients, are presented in Table 5.1. Of the predominantly white (39.97%) total sample, 39.7% had a positive status of family history of CVD, of which 18.1% had age specified and 81.8% had no additional information recorded.

The total sample had a median age of 48 [IQR: 38-58] and consisted of more females than males (58.5%); fewer deprived patients (5.6%); and a higher proportion of non-smokers to smokers (47.5% vs. 42.1%, respectively). Smokers were largely moderate smokers with a median of 10 [IQR: 5-20] cigarettes smoked per day.

The overall cohort had a mean BMI of 26.2 (SD=5.0), with 27.5% categorised as having normal BMI. Total cholesterol had a mean of 5.7 mmol/l (SD=1.1) in the dataset, and HDL cholesterol had a mean of 1.4 mmol/l (SD=0.4). Very few patients were diagnosed with Left ventricular Hypertrophy (<1%), Impaired Glucose Regulation (<1%) and diabetes (1.8%); however 15.48% were receiving hypertensive treatment at baseline. A mean SBP of 129.7 mmHg (SD=17.6) was found among patients.

Approximately 3% of patients were classed with chronic kidney disease (Stage 3, 4, and 5); a median serum creatinine of 86 umol/L [75-98] was found, which translated to a median egfr of 67.9 mL/min/1.73m^2 [59.1-77.9].

Large amounts of missing data (>50%) was found for exercise status, and data requiring blood tests. Missing data is described in further detail below.

Table 5.1 Baseline Characteristics of the overall sample derived from the CPRD database (N=302,591).

Characteristic		No	%
Family History	Negative	182,469	60.30
	Positive	120,122	39.70
Family History, detailed	Negative	182,469	60.30
	Positive	98,318	32.49
	Positive, with age	21,792	7.20
	Positive, with relative	10	0.00
	Positive, with age and relative	2	0.00
Gender	Female	176,921	58.47
	Male	125,670	41.53
Ethnicity	White	120,940	39.97
	S. Asian	1,733	0.57
	Black	680	0.22
	Other/Mixed	31,192	10.31
	Missing	148,046	48.93
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	57,281	18.93
	2	47,442	15.68
	3	39,892	13.18
	4	27,650	9.14
	5	16,909	5.59
	Missing	113,417	37.48
BMI	Underweight	2,824	0.93
	Normal	83,173	27.49
	Overweight	67,152	22.19
	Obese	34,985	11.56
	Missing	114,457	37.83

[Table continues on the next page]

Characteristic cont.		No	%
Smoking Status	Non-Smoker	143,773	47.51
	Ex-Smoker	26,404	8.73
	Current Smoker	127,472	42.13
	Missing	4,942	1.63
Alcohol Consumption	T-total	14,977	4.95
	Ex-Drinker	384	0.13
	Drinks alcohol	154,930	51.20
	Missing	132,300	43.72
Exercise Status	Inactive	14,148	4.68
	Active	111,400	36.82
	Missing	177,043	58.51
Diabetic	No	297,197	98.22
	Yes	5,394	1.78
Impaired Glucose Regulation	No	302,351	99.92
	Yes	240	0.08
Hypertensive Treatment	No	255,753	84.52
	Yes	46,838	15.48
Left Ventricular Hypertrophy	No	302,442	99.95
	Yes	149	0.05
Chronic Kidney Disease Stage	1	2,715	0.90
	2	20,151	6.66
	3	7,817	2.58
	4	72	0.02
	5	1	0.00
	Missing	271,835	89.84

[Table continues on the next page]

Characteristic cont.		Mean	SD
Height (m)		1.68	0.10
	Missing	114,457	37.83
Weight (kg)		74.06	16.07
	Missing	114,424	37.81
BMI		26.23	4.97
	Missing	114,457	37.83
SBP (mmHg)		129.67	17.63
	Missing	62,072	20.51
DBP (mmHg)		78.73	9.46
	Missing	62,072	20.51
Total Cholesterol (mmol/L)		5.66	1.06
	Missing	260,310	86.03
HDL Cholesterol (mmol/L)		1.41	0.41
	Missing	288,095	95.21
LDL Cholesterol (mmol/L)		3.65	0.98
	Missing	292,482	96.66
HbA1c (mmol/mol)		57.68	18.66
	Missing	300,062	99.16
Random Blood Glucose (mmol/L)		5.89	2.72
	Missing	300,652	99.36
Fasting Blood Glucose (mmol/L)		5.89	2.16
	Missing	300,448	99.29
Albumin (g/L)		42.13	3.74
	Missing	277,425	91.68
Platelets (10*9/L)		256.52	69.36
	Missing	261,714	86.49

[Table continues on the next page]

Characteristic cont.		Median	IQR
Age (years)		48	[38-58]
Cigarettes per day*		10	[5-20]
	Missing	54,896	43.07
Alcohol units per week**		6	[2-12]
	Missing	35,907	23.18
CRP (mg/L)		5	[2-10]
	Missing	299,644	99.03
Triglyceride value (mmol/l)		1.5	[1-2.1]
	Missing	283,111	93.56
ALT value (u/L)		24	[17-34]
	Missing	286,018	94.52
AST value (u/L)		22	[18-27]
	Missing	288,255	95.26
GGT value (u/L)		25	[16-48]
	Missing	289,948	95.82
Serum Creatinine (umol/L)		86	[75-98]
	Missing	270,731	89.47
Estimated Glomerular Filtration Rate (mL/min/1.73m^2)	Missing	67.9 270,731	[59.1-77.9] 89.47

***Uses denominator of current smokers (n=127,472)**
****Uses denominator of patients who drink alcohol (n=154,930)**

Characteristics of the extracted cohort were also described by FH status, looking at the difference between the baseline characteristics of patients with a positive FH recorded or a negative FH recorded. Details of this are presented in Table 5.2. A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

Table 5.2 Characteristics of the overall extracted cohort derived from CPRD, by FH status. Values are numbers (%), mean (SD), median [IQR]; and p-values.

Characteristic		No FH (n=182,469)		FH (n=120,122)		p-value
Gender	Female	104,437	57.24	72,484	60.34	<0.001*
	Male	78,032	42.76	47,638	39.66	
Ethnicity	White	71,845	39.37	49,095	40.87	
	S. Asian	1,116	0.61	617	0.51	
	Black	467	0.26	213	0.18	<0.001*
	Other/Mixed	18,847	10.33	12,345	10.28	
	Missing	90,194	49.43	57,852	48.16	
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	35,482	19.45	21,799	18.15	
	2	28,833	15.80	18,609	15.49	
	3	24,814	13.60	15,078	12.55	
	4	16,138	8.84	11,512	9.58	<0.001*
	5	8,951	4.91	7,958	6.62	
	Missing	68,251	37.4	45,166	37.60	
BMI	Underweight	1,733	0.95	1,091	0.91	
	Normal	49,604	27.18	33,569	27.95	
	Overweight	37,784	20.71	29,368	24.45	<0.001*
	Obese	18,531	10.16	16,454	13.70	
	Missing	74,817	41.00	39,640	33.00	
Smoking Status	Non-Smoker	93,358	51.16	50,415	41.97	
	Ex-Smoker	15,925	8.73	10,479	8.72	
	Current Smoker	70,652	38.72	56,820	47.30	<0.001*
	Missing	2,534	1.39	2,408	2.00	
Alcohol Consumption	T-total	9,571	5.25	50,415	41.97	
	Ex-Drinker	219	0.12	10,479	8.72	
	Drinks alcohol	86,683	47.51	5,820	47.30	<0.001*
	Missing	85,996	47.13	2,408	2.00	

[Table continues on the next page]

Characteristic cont.		No FH (n=182,469)		FH (n=120,122)		p-value
Exercise Status	Inactive	8,235	4.51	5,913	4.92	
	Active	63,818	34.97	47,582	39.61	<0.001*
	Missing	110,416	60.51	66,627	55.47	
Diabetic	No	179,309	98.27	117,888	98.14	
	Yes	3,160	1.73	2,234	1.86	0.009
Impaired Glucose Intolerance	No	182,338	99.93	120,013	99.91	
	Yes	131	0.07	109	0.09	0.070
Hypertensive Treatment	No	158,025	86.60	97,728	81.36	
	Yes	24,444	13.40	22,394	18.64	<0.001*
Left Ventricular Hypertrophy	No	182,399	99.96	120,043	99.93	
	Yes	70	0.04	79	0.07	<0.001*
Chronic Kidney Disease Stage	1	1,534	0.84	1,181	0.98	
	2	11,033	6.05	9,118	7.59	
	3	4,231	2.32	3,586	2.99	
	4	47	0.03	25	0.02	<0.001*
	5	1	0.00	0	0.00	
	Missing	165,623	90.77	106,212	88.42	
		Mean	**SD**	**Mean**	**SD**	**p-value**
Height (m)		1.68	0.10	1.67	0.10	
	Missing	74,817	41.00	39,640	33.00	<0.001*
Weight (kg)		73.65	15.91	74.61	16.26	
	Missing	74,801	40.99	39,623	32.99	<0.001*
BMI		26.00	4.85	26.54	5.10	
	Missing	74,817	41.00	39,640	33.00	<0.001*

[Table continues on the next page]

Characteristic cont.		Mean	SD	Mean	SD	p-value
SBP (mmHg)		128.81	17.60	130.89	17.61	<0.001*
	Missing	41,822	22.92	20,250	16.86	
DBP (mmHg)		78.19	9.39	79.50	9.51	<0.001*
	Missing	41,822	22.92	20,250	16.86	
Total Cholesterol (mmol/L)		5.64	1.06	5.67	1.06	<0.001*
	Missing	163,160	89.42	97,150	80.88	
HDL Cholesterol (mmol/L)		1.42	0.42	1.41	0.41	0.603
	Missing	176,085	96.50	112,010	93.25	
LDL Cholesterol (mmol/L)		3.64	0.99	3.66	0.97	0.398
	Missing	177,956	97.53	114,526	95.34	
HbA1c (mmol/mol)		58.21	18.51	56.96	18.85	0.096
	Missing	181,014	99.20	119,048	99.11	
Random Blood Glucose (mmol/L)		5.93	2.73	5.82	2.70	0.424
	Missing	181,275	99.35	119,377	99.38	
Fasting Blood Glucose (mmol/L)		6.03	2.28	5.72	2.01	<0.001*
	Missing	181,303	99.36	119,145	99.19	
Albumin (g/L)		42.07	3.82	42.20	3.64	0.008
	Missing	168,088	92.12	109,337	91.02	
Platelets (10*9/L)		255.05	69.64	258.48	68.95	<0.001*
	Missing	159,139	87.21	102,575	85.39	
		Median	**[IQR]**	**Median**	**[IQR]**	**p-value**
Age (years)		47	[38-58]	48	[40-57]	<0.001*
Cigarettes per day**		10	[5-20]	10	[4-20]	<0.001*
	Missing	27,812	39.36	27,084	47.67	
Alcohol units per week***		6	[2-12]	6	[2-12]	<0.001*
	Missing	19,287	22.25	16,620	24.35	

[Table continues on the next page]

Characteristic cont.		Median	[IQR]	Median	[IQR]	p-value
CRP (mg/L)		5	[2-10]	6	[2-10]	0.035
	Missing	180,829	99.10	118,815	98.91	
Triglyceride value (mmol/L)		1.5	[1-2.2]	1.4	[1-2.1]	<0.001*
	Missing	174,056	95.39	109,055	90.79	
ALT value (u/L)		24	[17-34]	24	[17-35]	0.579
	Missing	173,067	94.85	112,951	94.03	
AST value (u/L)		22	[18-27]	22	[18-27]	0.504
	Missing	174,279	95.51	113,976	94.88	
GGT value (u/L)		25	[16-47]	26.5	[17-50]	<0.001*
	Missing	174,920	95.86	115,028	95.76	
Serum Creatinine (umol/L)		86	[76-98]	86	[75-97]	0.003
	Missing	165,017	90.44	105,704	88.00	
Egfr (mL/min/1.73m^2)		68.0	[59.2-78.1]	67.9	[59.0-77.7]	0.424
	Missing	165,023	90.44	105,708	88.00	

*** Statistically significant differences between patients with and without FH.**
****Uses denominator of current smokers for each prospective group (no FH n= 70,652; FH n=56,820).**
*****Uses denominator of patients who drink alcohol for each prospective group (no FH n= 86,683; FH n=5,820).**

Table 5.2 shows the baseline characteristics of patients by their CVD family history status; if this is positive or negative. Statistically significant differences are found between all characteristics excluding HDL cholesterol (p=0.603), LDL cholesterol (p=0.398), IGR (p=0.070), HbA1c (p=0.096), and random blood glucose (p=0.424).

In particular, a lower proportion of males were found with a positive FH (39.7%, p<0.001) and patients with a negative FH had a lower median age (47 years, p<0.001). A higher proportion of smokers (47.3%, p<0.001) were found with patients with a positive FH of CVD, who also had a higher SBP (130.9 mmHg, p<0.001), and total cholesterol (5.67 mmol/l, p<0.001). More patients with a positive FH status were also found on hypertensive treatment at baseline (18.6%, p<0.001).

A higher proportion of white patients were found with a positive FH status (40.9%, $p<0.001$), as well as a higher proportion of patients with a more deprived Townsend score (6.6%, $p<0.001$).

5.1.3 Incidence of CVD

Overall, 5,847 (1.9%) patients had a CVD event during the follow up period. Table 5.3 shows the characteristics of patients who incurred a CVD event and patients who did not. A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

Table 5.3 Characteristic of the overall sample derived from CPRD, by CVD outcome during the 10 year follow up period. Values are numbers (%), mean (SD), median [IQR]; and p-values.

Characteristic		no CVD (n=296,744)		CVD (n=5,847)		p-value
Family History	Negative	179,346	60.44	3,123	53.41	<0.001*
	Positive	117,398	39.56	2,724	46.59	
Gender	Female	174,167	58.69	2,754	47.10	<0.001*
	Male	122,577	41.31	3,093	52.90	
Ethnicity	White	116,445	39.24	4,495	76.88	
	S. Asian	1,695	0.57	38	0.65	
	Black	663	0.22	17	0.29	<0.001*
	Other/Mixed	30,388	10.24	804	13.75	
	Missing	147,553	49.72	493	8.43	
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	55,987	18.87	1,294	22.13	
	2	46,167	15.56	1,275	21.81	
	3	38,699	13.04	1,193	20.40	<0.001*
	4	26,610	8.97	1,040	17.79	
	5	16,128	5.43	781	13.36	
	Missing	113,153	38.13	264	4.52	
BMI	Underweight	2,753	0.93	71	1.21	
	Normal	82,016	27.64	1,157	19.79	
	Overweight	65,690	22.14	1,462	25.00	<0.001*
	Obese	33,968	11.45	1,017	17.39	
	Missing	112,317	37.85	2,140	36.60	
Smoking Status	Non-Smoker	141,490	47.68	2,283	39.05	
	Ex-Smoker	25,745	8.68	659	11.27	
	Current Smoker	124,670	42.01	2,802	47.92	<0.001*
	Missing	4,839	1.63	103	1.76	

[Table continues on the next page]

Characteristic cont.		no CVD (n=296,744)		CVD (n=5,847)		p-value
Alcohol Consumption	T-total	14,655	4.94	322	5.51	
	Ex-Drinker	371	0.13	13	0.22	<0.001*
	Drinks alcohol	152,073	51.25	2,857	48.86	
	Missing	129,645	43.69	2,655	45.41	
Exercise Status	Inactive	13,800	4.65	348	5.95	
	Active	109,468	36.89	1,932	33.04	<0.001*
	Missing	173,476	58.46	3,567	61.01	
Diabetic	No	291,714	98.3	5,483	93.77	<0.001*
	Yes	5,030	1.70	364	6.23	
Impaired Glucose Intolerance	No	296,514	99.92	5,837	99.83	0.0119
	Yes	230	0.08	10	0.17	
Hypertensive Treatment	No	252,082	84.95	3,671	62.78	<0.001*
	Yes	44,662	15.05	2,176	37.22	
Left Ventricular Hypertrophy	No	296,602	99.95	5,840	99.88	0.0142
	Yes	142	0.05	7	0.12	
Chronic Kidney Disease Stage	1	2,641	0.89	74	1.27	
	2	19,475	6.56	676	11.56	
	3	7,336	2.47	481	8.23	<0.001*
	4	61	0.02	11	0.19	
	5	1	0.00	0	0.00	
	Missing	267,230	90.05	4,605	78.76	
		Mean	**SD**	**Mean**	**SD**	**p-value**
Height (m)		1.68	0.10	1.68	0.10	0.0789
	Missing	112,317	37.85	2,140	36.60	
Weight (kg)		73.97	16.02	78.34	18.00	<0.001*
	Missing	112,284	37.84	2,140	36.60	

[Table continues on the next page]

Characteristic cont.		Mean	SD	Mean	SD	p-value
BMI		26.2	4.95	27.64	5.52	<0.001*
	Missing	112,317	37.85	2,140	36.60	
SBP (mmHg)		129.42	17.52	141.98	18.46	<0.001*
	Missing	61,160	20.61	912	15.60	
DBP (mmHg)		78.66	9.46	82.3	9.05	<0.001*
	Missing	61,160	20.61	912	15.60	
Total Cholesterol (mmol/L)		5.65	1.06	5.82	1.08	<0.001*
	Missing	255,849	86.22	4,461	76.30	
HDL Cholesterol (mmol/L)		1.42	0.41	1.39	0.42	0.1906
	Missing	282,706	95.27	5,389	92.17	
LDL Cholesterol (mmol/L)		3.64	0.98	3.85	0.94	<0.001*
	Missing	286,960	96.70	5,522	94.44	
HbA1c (mmol/mol)		57.57	18.77	58.89	17.35	0.3438
	Missing	294,410	99.21	5,652	96.66	
Random Blood Glucose (mmol/L)		5.84	2.65	7.34	4.12	<0.001*
	Missing	294,862	99.37	5,790	99.03	
Fasting Blood Glucose (mmol/L)		5.85	2.14	6.71	2.57	<0.001*
	Missing	294,695	99.31	5,753	98.39	
Albumin (g/L)		42.15	3.73	41.54	3.89	<0.001*
	Missing	272,508	91.83	4,917	84.09	
Platelets (10*9/L)		256.59	69.21	254.28	74.05	0.2493
	Missing	257,102	86.64	4,612	78.88	
		Median	**[IQR]**	**Median**	**[IQR]**	**p-value**
Age (years)		47	[38-57]	63	[53-72]	<0.001*
Cigarettes per day**		10	[5-20]	10	[5-20]	<0.001*
	Missing	53,679	43.06	1,217	43.43	

[Table continues on the next page]

Characteristic cont.		Median	[IQR]	Median	[IQR]	p-value
Alcohol units per week***		6	[2-12]	6	[2-12]	0.0914
	Missing	35,076	23.07	831	29.09	
CRP (mg/L)		5	[2-10]	8	[3-14]	<0.001*
	Missing	293,916	99.05	5,728	97.96	
Triglyceride value (mmol/L)		2	[1-2.1]	2	[1.1-2.2]	0.0014*
	Missing	277,891	93.65	5,220	89.28	
ALT value (u/L)		24	[17-34]	22	[16-34]	0.0612
	Missing	280,836	94.64	5,182	88.63	
AST value (u/L)		22	[18-27]	23	[19-30]	<0.001*
	Missing	282,866	95.32	5,389	92.17	
GGT value (u/L)		25	[16-47]	36	[21-72]	<0.001*
	Missing	284,528	95.88	5,420	92.70	
Serum Creatinine (umol/L)		86	[75-97]	91	[79-104]	<0.001*
	Missing	266,161	89.69	4,560	77.99	
Egfr (mL/min/1.73m^2)		68.1	[59.3-78.1]	63.5	[54.0-73.4]	<0.001*
	Missing	266,171	89.70	4,560	77.99	

*** Statistically significant differences between patients with and without CVD.**
****Uses denominator of current smokers for each prospective group (no CVD n= 124,670; CVD n=2,802).**
*****Uses denominator of patients who drink alcohol for each prospective group (no CVD n= 152,073; CVD n=2,857).**

A higher proportion of white (76.9%, p<0.001), aged (median: 63 years, p<0.001), males (52.9%, p<0.001) were found to have a CVD event during the follow up period; with an increased Townsend deprivation score (31.2%, p<0.001). A statistically significant higher proportion of patients also had a positive FH status (46.6%, p<0.001) in the CVD event group.

Statistically significant differences were found between all standard risk factors, expect for HDL cholesterol, which remained similar across CVD event groups (p=0.191).

Increased mean total cholesterol (5.82 mmol/l (SD=1.1), p<0.001), mean SBP (141.98 mmHg (SD=18.5), p<0.001), and increased proportion of current smokers (47.9%, p<0.001) was found for patients with a CVD event. Patients with a CVD event also displayed lower levels of estimated glomerular filtration rate (median: 63.5 [54.0-73.4], p<0.001).

Patients with a CVD event also reported statistically significant increase in the proportion of diabetics (6.2%, p<0.001), patients on hypertensive treatment at baseline (37.2%, p<0.001), mean weight (78.3kg (SD=18.0), p<0.001), and mean BMI (27.6 (SD=5.5), p<0.001). Fewer patients were considered 'active' regarding their exercise status (33.0%, p<0.001).

Of the 5,847 CVD events found in the extracted sample (302,591), 5,844 (1.93%) were eligible for the analysis out of 302,509 patients; 82 patients were excluded from the analysis has they entered and exited the study on the same day. Crude incidence rates for the CVD outcome are presented in Table 5.4.

Table 5.4 Crude incidence rates of the CVD outcome in the overall cohort (n=302,509). Values are the number of incident fracture cases, the total person years and crude incidence rates (95% Confidence Interval).

Risk Factor		No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)
Family History	Negative	3,122	1,665,422.4	18.75 (18.1 to 19.42)
	Positive	2,722	1,063,897.0	25.59 (24.64 to 26.56)
Age Category	25-34	96	399,694.4	2.40 (1.97 to 2.93)
	35-44	527	852,501.7	6.18 (5.68 to 6.73)
	45-54	1,027	694,148.9	14.80 (13.92 to 15.73)
	55-64	1,487	470,551.3	31.60 (30.04 to 33.25)
	65-74	1,755	242,292.9	72.43 (69.12 to 75.9)
	75+	952	70,130.3	135.75 (127.39 to 144.65)
Gender	Female	2,753	1,615,386.2	17.04 (16.42 to 17.69)
	Male	3,091	1,113,933.3	27.75 (26.79 to 28.74)
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	1,294	523,978.3	24.7 (23.39 to 26.08)
	2	1,274	428,845.0	29.71 (28.12 to 31.38)
	3	1,192	360,406.1	33.07 (31.25 to 35.01)
	4	1,040	246,980.3	42.11 (39.63 to 44.75)
	5	781	148,077.2	52.74 (49.17 to 56.57)
Ethnicity	White	4,493	1,070,764.5	41.96 (40.75 to 43.21)
	S. Asian	38	15,185.1	25.02 (18.21 to 34.39)
	Black	17	6,076.0	27.98 (17.39 to 45.01)
	Other/Mixed	804	278,396.8	28.88 (26.95 to 30.95)
BMI Category	Underweight	71	25,866.3	27.45 (21.75 to 34.64)
	Normal	1,157	777,109.3	14.89 (14.05 to 15.77)
	Overweight	1,460	594,583.2	24.56 (23.33 to 25.85)
	Obese	1,016	292,684.9	34.71 (32.64 to 36.91)

[Table continues on the next page]

Risk Factor cont.		No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)
Exercise Status	Inactive	348	125,448.2	27.74 (24.97 to 30.81)
	Active	1,929	1,012,196.1	19.06 (18.23 to 19.93)
Exercise Status, detail	Inactive	348	125,448.2	27.74 (24.97 to 30.81)
	Gentle	827	318,101.7	26.00 (24.29 to 27.83)
	Moderate	894	545,047.2	16.40 (15.36 to 17.51)
	Vigorous	109	107,167.1	10.17 (8.43 to 12.27)
Smoking Status	Non/Ex	2,940	1,542,101.8	19.06 (18.39 to 19.77)
	Current	2,801	1,142,984.2	24.51 (23.62 to 25.43)
Alcohol Consumption	T-total/Ex	335	135,870.5	24.66 (22.15 to 27.44)
	Current	2,854	1,405,064.2	20.31 (19.58 to 21.07)
Diabetes	No	5,480	2,703,734.6	20.27 (19.74 to 20.81)
	Yes	364	25,584.8	142.27 (128.38 to 157.66)
LVH	No	5,837	2,728,326.7	21.39 (20.85 to 21.95)
	Yes	7	992.7	70.51 (33.62 to 147.91)
IGR	No	5,834	2,727,938.0	21.39 (20.84 to 21.94)
	Yes	10	1,381.4	72.39 (38.95 to 134.54)
Hypertensive treatment	No	3,668	2,368,753.5	15.48 (14.99 to 15.99)
	Yes	2,176	360,565.9	60.35 (57.87 to 62.94)
Chronic Kidney Disease Stage	1	74	22,801.5	32.45 (25.84 to 40.76)
	2	676	167,974.9	40.24 (37.32 to 43.40)
	3	481	59,245.7	81.19 (74.25 to 88.78)
	4	11	316.9	347.13 (192.24 to 626.81)
	5	0	10.0	-
Total		**5,844**	**2,729,319.4**	**21.41 (20.87 to 21.97)**

The total crude incidence rate for a CVD event was 21.41 per 10,000 person years (py) (20.87 to 21.97). Increasing crude incidence rates for CVD were found with increasing age and increasing deprivation; where patients aged 75 year or above, or scored 5 on their Townsend deprivation score had an crude incidence rate of 135.75 (127.39 to 144.65) and 52.74 (49.17 to 56.57) per 10,000 py, respectively. High CVD incidence rates were also found amongst diabetics (142.27 per 10,000 py (128.38 to 157.66)), patients on hypertensive treatment at baseline (60.35 per 10,000 py (57.87 to 62.94)), and patients with CKD (82.59 per 10,000 py (75.60 to 90.22)).

Patients with a positive FH status had higher crude CVD incidence rate (25.59 per 10,000 py (24.64 to 26.56)) compared to patients with a negative family history of CVD (18.75 per 10,000 py (18.1 to 19.42)).

Lower CVD incidence rates were found with patients with 'normal' BMI (18.5≤BMI<25) and patients from an ethnic minority background.

5.1.4 Generalisability of the data

5.1.4.1 Characteristics

The extracted cohort from CPRD was compared to the Health Survey for England (HSE) 2011 (104) population to evaluate its generalisability to the UK population. Comparable variables in the dataset were age, gender, total cholesterol, HDL cholesterol, blood pressure, height, weight, BMI, current smoking status and HbA1c values. The HSE–2011 data comprised of 10,617 individuals. Of these 6,790 were aged between 25 and 74.

Varying levels of missing data was found between the samples, shown in Table 5.5. The HSE data had better recording for cholesterol (total and HDL), height, weight, BMI and HbA1c values.

Table 5.5 Comparison of proportions of recorded data in each dataset for each comparable variable. Values are numbers (%)

Variable	CPRD			HSE		
	Male % (n=125,670)	Female % (n=176,921)	Overall % (n=302,591)	Male % (3,027)	Female % (n=3,763)	Overall % (n=6,790)
Total Cholesterol	15.94	12.57	13.97	44.41	39.84	42.37
HDL Cholesterol	5.37	4.38	4.79	44.41	39.84	43.37
Systolic Blood Pressure	70.97	85.54	79.49	57.53	54.91	56.36
Diastolic Blood Pressure	70.97	85.54	79.49	57.53	54.91	56.36
Height	56.09	66.5	62.17	98.91	98.84	98.88
Weight	56.10	66.51	62.19	93.04	96.00	94.36
Body Mass Index	56.09	66.50	62.17	84.31	82.33	83.21
Smoking Status	97.97	98.65	98.37	76.84	76.83	76.83
HbA1c	1.07	0.67	0.84	48.84	47.74	48.35

Table 5.6 shows a summary of the comparable characteristics found in both datasets, and found, the HSE population had a median age of 48 years (IQR: 38-61), and consisted of more females (55.4%) than males (44.6%). A higher proportion of smokers (55%) were found in the HSE dataset compared to the extracted CPRD sample (43%); which also had an increased mean weight (74.1 kg (SD=16.1), height (1.69 m (SD=10.1) and BMI (27.7 (SD=5.3).

Conversely, patients in the HSE data had a lower total cholesterol, higher HDL cholesterol, lower blood pressure, and lower HbA1c.

Table 5.6 Comparison of characteristics for CPRD and HSE data.

Characteristics		CPRD (n=302,591)		HSE (n=6,790)		p-value
Median Age (years), [IQR]		48	[38-58]	48	[38-61]	0.747
		No	**%***	**No**	**%***	
Gender (%)	Female	176,921	58.47	3,763	55.42	<0.001
	Male	125,670	41.53	3,027	44.58	
Smoking Status (%)	Non/Ex-Smoker	170,177	56.24	3,625	53.39	<0.001
	Current Smoker	127,472	42.13	1,592	23.45	
		Mean	**SD**	**Mean**	**SD**	
Weight (kg)		74.06	16.07	76.80	16.97	<0.001
Height (m)		1.68	0.10	1.69	10.06	0.935
BMI		26.23	4.97	27.70	5.33	<0.001
Total Cholesterol (mmol/l)		5.66	1.06	5.50	1.05	<0.001
HDL Cholesterol (mmol/l)		1.41	0.41	1.53	0.45	<0.001
SBP (mmHg)		129.67	17.63	125.96	16.50	<0.001
DBP (mmHg)		78.73	9.46	74.17	10.43	<0.001
HbA1c (mmol/mol)		57.68	18.66	39.09	8.63	<0.001

***percentages do not add to 100% to account for missing data**

Mean values for total cholesterol and SBP found in CPRD were comparable to the HSE data for each 10 year group (Figure 5.2). However, mean values were parallel but lower for HDL cholesterol, and higher for DBP. The CPRD sample had proportionally lower weight and shorter height for all age groups than the HSE sample, and hence also had proportionally lower mean BMI values. A higher proportion of smokers and a higher mean HbA1c were also found in the CPRD database, which followed a similar trend to the HSE data across the age categories.

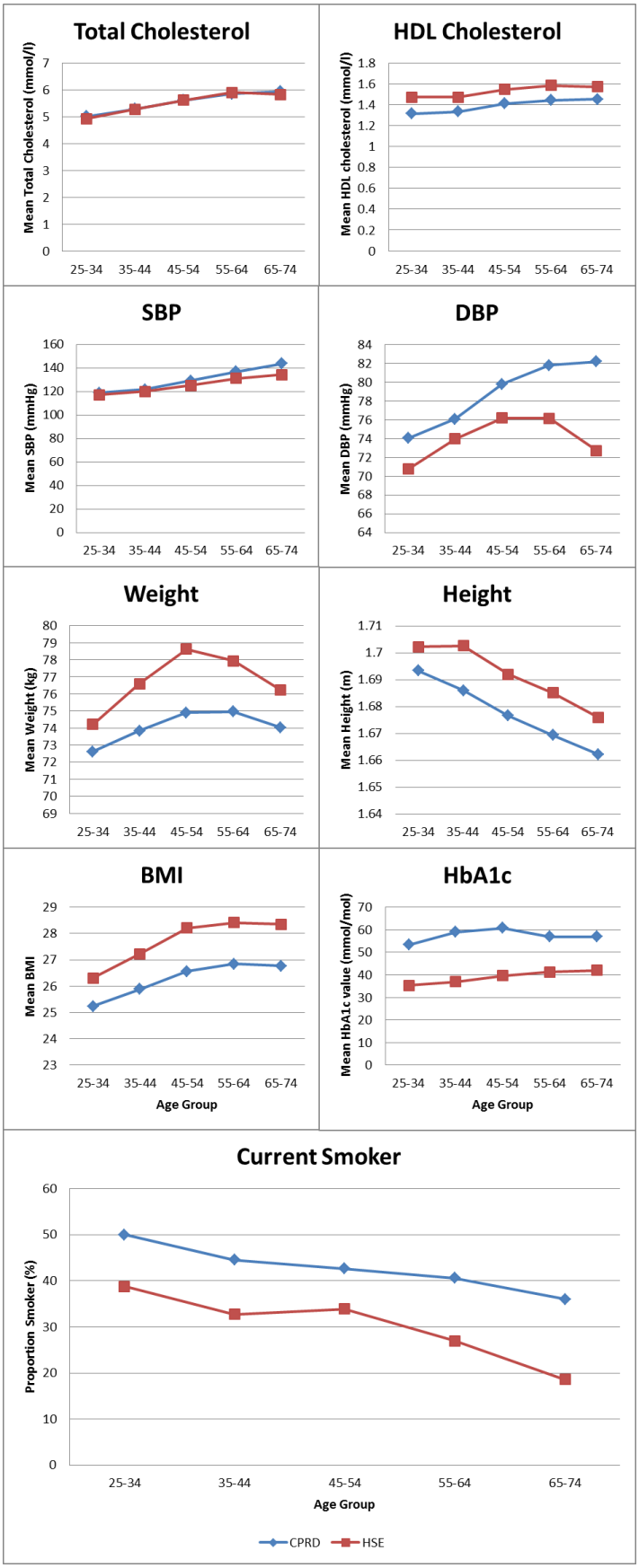


Figure 5.2 Comparison of mean values for variables by age group (10year) in the CPRD dataset compared with the HSE.

Results remained consistent for males and females for all characteristics; however larger differences between CPRD and HSE data were found in the mean SBP for males aged 55-64 and 65-74 (Appendix E).

5.1.4.2 CVD event rate

The HSE data reported prevalent CVD events which were not comparable to the CVD incidence rate calculated from the CPRD sample. To assess the generalisability of the CVD event rate from the extracted CPRD dataset, it was compared to the QRISK2 validation study (74) which used a similar cohort to assess the performance of QRISK2 for 10 year CVD risk prediction.

A total of 2,084,445 patients were extracted from the THIN database and used in the analysis; of which 93,564 (4.48%) incident cases of CVD were recorded by the GP computer systems. The THIN sample had a total 11,862,381 person years contributing to their analysis, which gave them a crude CVD incidence rate of 78.87 CVD events per 10,000 py.

This is a higher event rate compared to the 1.93% event rate found in the CPRD extracted sample which had a crude incidence rate of 21.41 (20.87 to 21.97) CVD events per 10,000 py.

5.1.5 Missing Data

Missing values were investigated amongst standard, additional and the new risk factor. Missing data of these risk factors in the CPRD data is given in Table 5.7.

Table 5.7 Missing data of data for standard, new, and additional CVD risk factors.

Risk Factors		No	%
Standard Risk Factors	Age	-	-
	Gender	-	-
	Smoker	4,942	1.63
	Total Cholesterol	260,310	86.03
	HDL Cholesterol	288,095	95.21
	SBP	62,072	20.51
New Risk Factor	FH	-	-
Additional Risk Factors	Townsend Deprivation Score	133,417	44.09
	Ethnicity	148,046	48.93
	Physical Activity - binary	177,043	58.51
	Physical Activity ordinal	181,711	60.05
	Alcohol Consumption	132,300	43.72
	BMI	114,457	37.83
	Diabetes	-	-
	Impaired Glucose Regulation	-	-
	Left Ventricular Hypertrophy	-	-
	Random Blood Glucose Value	300,652	99.36
	Fasting Blood Glucose Value	300,448	99.29
	Triglyceride Value	283,111	93.56
	HbA1c Value	300,062	99.16
	C-Reactive Protein	299,644	99.03
	Liver Enzyme: ALT	286,018	94.52
	Liver Enzyme: AST	288,255	95.26
	Liver Enzyme: GGT	289,948	95.82
	CKD – binary	271,835	89.84
	CKD – egfr	270,731	89.47

Overall only 4.3% of the total sample had complete data for the standard CVD risk factors of the study; and 0.7% had complete data for the standard, new, and additional risk factors for the study (Table 5.8).

Table 5.8 Completeness and missing data of risk factors.

	Standard Risk Factors		+ New and Additional Risk Factors	
	No	**%**	**No**	**%**
Complete data on all risk factors	13,146	4.34	2,009	0.66
1 missing risk factor	27,490	9.08	6,784	2.24
2 missing risk factors	199,672	65.99	17,042	5.63
3 missing risk factors	60,547	20.01	63,717	21.06
4 missing risk factors	1,736	0.57	56,061	18.53
5 missing risk factors	-	-	67,578	22.33
6 missing risk factors	-	-	43,071	14.23
7 missing risk factors	-	-	28,660	9.47
8 missing risk factors	-	-	17,107	5.65
9 missing risk factors	-	-	562	0.19

Data was complete for age, gender, FH, diabetes, impaired glucose regulation, and left ventricular hypertrophy. Thus the maximum number of missing risk factors amongst the standard risk factors was 4, and was 9 for standard, new, and additional risk factors.

Large amounts of missing data (>50%) was found for total cholesterol, HDL cholesterol, physical activity, the number of units of alcohol drank per week, random and fasting blood glucose, triglyceride, HbA1c, C-reactive protein, values for the 3 liver enzymes, and CKD (binary and egfr).

The most common 2 missing standard risk factors were total cholesterol and HDL cholesterol, accounting for (66%); 20% had missing values for SBP, total cholesterol, and HDL cholesterol. Thus a monotone pattern was not found between the standard risk factors; and missing data was arbitrary.

Patients aged 45-54 and 55-64 years had the most complete data, accounting for 27.25% and 29.62% of the total number with complete data (13,146), respectively. Patients with 4 risk factors missing were largely accounted for by those aged 35-44 years (34.39%). This remained consistent with missing standard and additional risk factors.

The relationship of missingness for each incomplete risk factor (smoking status , total cholesterol, HDL cholesterol, SBP, Townsend deprivation score, ethnicity, alcohol intake (dichotomised format) and BMI) was assessed with one another and also using the complete risk factors (age, gender, FH, CVD outcome, diabetes, IGR, LVH) as possible predictors of their missingness.

5.1.5.1 Correlation

Small to moderate correlations were found for missing data of total cholesterol, HDL cholesterol, and systolic blood pressure. Specifically, correlations were found between the missingness of total cholesterol with age (-0.1559), family history (-0.1205), and SBP (-0.1796). The missingness of HDL cholesterol was correlated with total cholesterol (-0.1401) and SBP (-0.1128); and the missingness of SBP was correlated with gender (-0.1778).

5.1.5.2 Regression

Associations were reported between the missingness of each incomplete standard and additional risk factor with less than 50% missing data (smoking status, total cholesterol,

HDL cholesterol, SBP, Townsend Deprivation Score, Ethnicity, BMI, and alcohol consumption). These are presented in Appendix F.

The unadjusted analysis show age, gender to be strong predictors for the missingness of smoking status, total cholesterol, HDL cholesterol, SBP, Townsend Deprivation Score, Ethnicity, BMI, and alcohol consumption ($p<0.01$). In addition, alcohol consumption was associated with the missingness of smoking status ($p<0.001$); increased deprivation, occurrence of a CVD event, and positive FH status were a strong predictors for the missingness of total cholesterol ($p<0.001$); total cholesterol showed statistically significant association with the missingness of HDL ($p<0.001$); positive FH status and occurrence of a CVD event was associated with the missingness of SBP; and deprivation was associated to the missingness of ethnicity.

Most of the associations were not statistically significant in the adjusted analysis of the missing data with observed data; but indicated that assuming the data was missing at random was plausible.

Associations were also investigated between missing data with unobserved data of other risk factors. This showed statistically significant associations in both the unadjusted and adjusted analyses. In particular, unobserved data for SBP was associated to missing data of all other risk factors ($p<0.001$); and unobserved data for Townsend score was associated to missing data of all other risk factors, but BMI ($p<0.001$). Unobserved data for smoking status was associated to alcohol consumption.

5.1.5.3 Multiple Imputation

Missing data was associated with both the observed and unobserved data of other risk factors; and also described a non-monotone (arbitrary) pattern of missing data.

Whereas it is difficult to determine whether the data is missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR); data can be distinguished between MCAR and MAR. It was assumed data was MAR based on the associations of missing data with observed data; the plausibility of this assumption was further strengthened by including variables to both predict the missing risk factor and predict if it is missing (identified from the above regression analyses).

Multiple imputations by chained equations were used to impute values for incomplete standard risk factors, and incomplete additional risk factors with less than 50% of the data missing. Thus, physical activity, the number of units of alcohol drank per week, random and fasting blood glucose, triglyceride, HbA1c, C-reactive protein, values for the 3 liver enzymes, and CKD (binary and egfr) were not imputed for and were removed from the analysis due to insufficient data.

Values were imputed for smoking status, SBP, total and HDL cholesterol, BMI, alcohol consumption, ethnicity, and Townsend deprivation score. The regression models used to impute the data were adjusted for each incomplete risk factor imputed for, age, hypertensive treatment at baseline, CVD event during follow up, and diabetes. Ten imputation datasets were created using imputation models were developed by gender; and with and without adjusting for FH status.

Table 5.9 shows the characteristics of the original and imputed data, with and without FH status. No difference was found in mean imputed values when adjusting for FH status or not. Therefore, imputation models adjusting for FH status was used in the main analysis.

Table 5.9 Characteristics of the original and imputed datasets (with and without FH status).

Imputed Risk Factors		Original Dataset			Mean Imputed with FH		Mean Imputed without FH	
		Missing no. (%)	Mean	SD	Mean	SD	Mean	SD
Total Cholesterol (mmol/L)		260,310 (86.03)	5.66	1.06	5.53	0.0097	5.54	0.0052
HDL Cholesterol (mmol/L)		288,095 (95.21)	1.41	0.41	1.45	0.0052	1.44	0.0032
SBP (mmHg)	Overall	62,072 (20.51)	129.67	17.63	129.45	0.0212	129.46	0.0149
	Treated	4,111 (8.78)	141.84	18.2	141.56	0.0183	141.57	0.0251
	Non treated	57,961 (22.66)	127.04	16.36	127.23	0.0227	127.25	0.0170
BMI		114,457(37.83)	26.23	4.97	26.25	0.0032	26.26	0.0047
		Missing no. (%)	No	%	No	%	No	%
Smoker	No	4,942 (1.63)	170,177	57.26	172,961	57.16	173,011	57.18
	Yes		127,472	42.74	129,630	42.84	129,580	42.82
Consume alcohol	No	132,300 (43.72)	15,361	8.93	29,842	9.86	29,626	9.79
	Yes		154,930	91.07	272,750	90.14	272,966	90.21

[Table continues on the next page]

Imputed Risk Factors cont.		Missing no. (%)	No	%	No	%	No	%
Ethnicity	White	148,046 (48.93)	120,940	78.11	235,921	77.97	235,956	77.98
	S. Asian		1,733	1.12	3,448	1.14	3,432	1.13
	Black		680	0.45	1,305	0.43	1,311	0.43
	Other/Mixed		31,192	20.33	61,917	20.46	61,893	20.45
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	113,417 (37.48)	57,281	30.27	91,178	30.13	91,138	30.12
	2		47,442	25.20	75,654	25.00	75,761	25.04
	3		39,892	21.14	64,055	21.17	64,015	21.16
	4		27,650	14.49	44,473	14.70	44,431	14.68
	5		16,909	8.90	27,231	9.00	27,246	9.00

5.1.6 Derivation and Validation Cohorts

The overall cohort was spilt into the derivation and validation datasets; where two-thirds of the data was assigned to creating the 10 year CVD risk models and the remaining one-third was assigned to validating them.

The derivation dataset consisted of 202,770 (67%) randomly selected patients, and the validation dataset included the remaining 99,821 (33%) patients.

The models to be developed using the derivation dataset are:

- Model 1: standard CVD risk factors (without FH)
- Model 2: standard CVD risk factors (with FH)
- Model 3: standard CVD risk factors with additional recognised CVD risk factors (without FH)
- Model 4: standard CVD risk factors with additional recognised CVD risk factors (with FH)

Table 5.10 shows and compares the baseline characteristics for each dataset, evaluating their comparability. It shows similar proportions of recording of data and missing values between datasets.

Table 5.10 Baseline characteristics of the derivation and validation dataset. Values are numbers (%), mean (SD), median [IQR].

Characteristic		Derivation (n=202,770)		Validation (n=99,821)	
Family History	Negative	122,094	60.21	60,375	60.48
	Positive	80,676	39.79	39,446	39.52
Gender	Female	118,328	58.36	58,593	58.70
	Male	84,442	41.64	41,228	41.30
Ethnicity	White	80,786	39.84	40,154	40.23
	S. Asian	1,157	0.57	576	0.58
	Black	463	0.23	217	0.22
	Other/Mixed	21,025	10.37	10,167	10.19
	Missing	99,339	48.99	48,707	48.79
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	38,396	18.94	18,885	18.92
	2	31959	15.76	15,483	15.51
	3	26,819	13.23	13,073	13.1
	4	18,381	9.06	9,269	9.29
	5	11,286	5.57	5,623	5.63
	Missing	75,929	37.45	37,488	37.56
BMI	Underweight	1,924	0.95	900	0.90
	Normal	55,691	27.47	27,482	27.53
	Overweight	45,035	22.21	22,117	22.16
	Obese	23,439	11.56	11,546	11.57
	Missing	76,681	37.82	37,776	37.84
Smoking Status	Non Smoker	96,417	47.55	47,356	47.44
	Ex-Smoker	17,784	8.77	8,620	8.64
	Current Smoker	85,246	42.04	42,226	42.30
	Missing	3,323	1.64	1,619	1.62

[Table continues on the next page]

Characteristic cont.		Derivation (n=202,770)		Validation (n=99,821)	
Alcohol Consumption	T-total	9,921	4.89	5,056	5.07
	Ex Drinker	258	0.13	126	0.13
	Drinks alcohol	103,882	51.23	51,048	51.14
	Missing	88,709	43.75	43,591	43.67
Exercise Status	Inactive	9,402	4.64	4,746	4.75
	Active	74,729	36.85	36,671	36.74
	Missing	118,639	58.51	58,404	58.51
Diabetic	No	199,161	98.22	98,036	98.21
	Yes	3,609	1.78	1,785	1.79
Impaired Glucose Intolerance	No	202,611	99.92	99,740	99.92
	Yes	159	0.08	81	0.08
Hypertensive Treatment	No	171,530	84.59	84,223	84.37
	Yes	31,240	15.41	15,598	15.63
Left Ventricular Hypertrophy	No	202,670	99.95	99,772	99.95
	Yes	100	0.05	49	0.05
Chronic Kidney Disease Stage	1	1,805	0.89	910	0.91
	2	13,445	6.63	6706	6.72
	3	5,173	2.55	2644	2.65
	4	46	0.02	26	0.03
	5	1	0	0	0.00
	Missing	182,300	89.90	89,535	89.70
		Mean	**SD**	**Mean**	**SD**
Height (m)		1.68	0.10	1.68	0.10
	Missing	76,681	37.82	37,776	37.84
Weight (kg)		74.06	16.08	74.07	16.04
	Missing	76,657	37.80	37,767	37.83

[Table continues on the next page]

Characteristic cont.		Mean	SD	Mean	SD
BMI		26.22	4.94	26.25	5.03
	Missing	76,681	37.82	37,776	37.84
SBP (mmHg)		129.66	17.62	129.71	17.66
	Missing	41,695	20.56	20,377	20.41
DBP (mmHg)		78.72	9.48	78.75	9.44
	Missing	41,695	20.56	20,377	20.41
Total Cholesterol (mmol/l)		5.66	1.06	5.66	1.06
	Missing	174,603	86.11	85,707	85.86
HDL Cholesterol (mmol/l)		1.42	0.41	1.41	0.41
	Missing	193,089	95.23	95,006	95.18
LDL Cholesterol (mmol/l)		3.64	0.98	3.67	0.98
	Missing	196,030	96.68	96,452	96.62
HbA1c (mmol/mol)		57.76	19.04	57.51	17.87
	Missing	201,067	99.16	98,995	99.17
Random Blood Glucose (mmol/L)		5.90	2.79	5.86	2.56
	Missing	201,469	99.36	99,183	99.36
Fasting Blood Glucose (mmol/L)		5.92	2.24	5.84	2.00
	Missing	201,341	99.3	99,107	99.28
Albumin (g/L)		42.13	3.72	42.12	3.78
	Missing	185,979	91.72	91,446	91.61
Platelets (10*9/L)		256.04	69.55	257.48	68.98
	Missing	175,460	86.53	86,254	86.41
		Median	**[IQR]**	**Median**	**[IQR]**
Age (years)		48	[38-58]	48	[38-58]
Cigarettes per day*		10	[5-20]	10	[5-20]
	Missing	36,585	42.92	18,311	43.36

[Table continues on the next page]

Characteristic cont.		Median	[IQR]	Median	[IQR]
Alcohol units per week**		6	[2-12]	6	[2-12]
	Missing	76,680	73.82	37,832	74.11
CRP (mg/L)		5	[2-10]	5	[2-10]
	Missing	200,799	99.03	98,845	99.02
Triglyceride value (mmol/L)		1.5	[1-2.1]	1.5	[1-2.1]
	Missing	189,710	93.56	93,401	93.57
ALT value (u/L)		24	[17-34]	24	[17-34]
	Missing	191,732	94.56	94,286	94.46
AST value (u/L)		22	[18-27]	22	[18-27]
	Missing	193,286	95.32	94,969	95.14
GGT value (u/L)		26	[16-48]	25	[16-48]
	Missing	194,293	95.82	95,655	95.83
Serum Creatinine (umol/L)		86	[75-98]	86	[76-98]
	Missing	181,558	89.54	89,163	89.32
Egfr (mL/min/1.73m^2)		68.0	[59.2-77.9]	67.8	[59.0-77.8]
	Missing	181,563	89.54	89,168	89.33

***Uses denominator of current smokers for each prospective group (derivation n= 85,246; validation n=42,226).**
****Uses denominator of patients who drink alcohol for each prospective group (derivation n= 152,073; validation n=51,048).**

The proportion of incident CVD events in the derivation and validation datasets was 1.92% (n=3,894) and 1.96% (n=1,953), respectively.

Crude CVD incidence rates, excluding censored patients can be found in Table 5.11, where trends are comparable between datasets and also the overall cohort.

Table 5.11 Crude CVD incidence rates for the derivation and validation datasets.

Risk Factor		Derivation			Validation		
		No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)
Family History	Negative	2,093	1,114,205.6	18.78 (18 to 19.61)	1,029	551,216.8	18.67 (17.56 to 19.84)
	Positive	1,799	714,612.9	25.17 (24.04 to 26.37)	923	349,284.1	26.43 (24.77 to 28.19)
Age Category	25-34	71	267,977.4	2.65 (2.1 to 3.34)	25	131,717.0	1.9 (1.28 to 2.81)
	35-44	357	570,918.2	6.25 (5.64 to 6.94)	170	281,583.4	6.04 (5.19 to 7.02)
	45-54	681	466,300.3	14.6 (13.55 to 15.74)	346	227,848.5	15.19 (13.67 to 16.87)
	55-64	970	315,003.5	30.79 (28.92 to 32.79)	517	155,547.8	33.24 (30.49 to 36.23)
	65-74	1,184	161,602.7	73.27 (69.21 to 77.56)	571	80,690.28	70.76 (65.19 to 76.81)
	75+	629	47,016.4	133.78 (123.73 to 144.66)	323	23,113.9	139.74 (125.3 to 155.84)
Gender	Female	1,835	1,080,835.5	16.98 (16.22 to 17.77)	918	534,550.7	17.17 (16.1 to 18.32)
	Male	2,057	747,983.0	27.5 (26.34 to 28.72)	1,034	365,950.3	28.26 (26.58 to 30.03)

[Table continues on the next page]

Risk Factor cont.		Derivation			Validation		
		No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	827	351,226.2	23.55 (21.99 to 25.21)	467	172,752.1	27.03 (24.69 to 29.60)
	2	853	288,853.0	29.53 (27.61 to 31.58)	421	139,992.1	30.07 (27.33 to 33.09)
	3	796	241,976.6	32.9 (30.69 to 35.26)	396	118,429.5	33.44 (30.3 to 36.90)
	4	709	164,118.5	43.2 (40.13 to 46.5)	331	82,861.8	39.95 (35.87 to 44.49)
	5	529	98,698.0	53.6 (49.22 to 58.37)	252	49,379.2	51.03 (45.11 to 57.74)
Ethnicity	White	3,003	714,939.0	42 (40.53 to 43.53)	1,490	355,825.5	41.87 (39.8 to 44.06)
	S. Asian	29	10,169.1	28.52 (19.82 to 41.04)	9	5,015.9	17.94 (9.34 to 34.48)
	Black	14	4,116.5	34.01 (20.14 to 57.42)	3	1,959.4	15.31 (4.94 to 47.47)
	Other/Mixed	516	187,722.6	27.49 (25.22 to 29.96)	288	90,674.2	31.76 (28.3 to 35.65)
BMI Category	Underweight	52	17,607.4	29.53 (22.5 to 38.76)	19	8,259.0	23.01 (14.67 to 36.07)
	Normal	775	520,052.9	14.9 (13.89 to 15.99)	382	257,056.4	14.86 (13.44 to 16.43)
	Overweight	951	398,865.4	23.84 (22.37 to 25.41)	509	195,717.8	26.01 (23.84 to 28.37)
	Obese	679	196,436.3	34.57 (32.06 to 37.27)	337	96,248.6	35.01 (31.47 to 38.96)

[Table continues on the next page]

Risk Factor cont.		Derivation			Validation		
		No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)
Exercise Status	Inactive	228	83,437.2	27.33 (24.00 to 31.11)	120	42,011.0	28.56 (23.88 to 34.16)
	Active	1,300	678,517.9	19.16 (18.15 to 20.23)	629	333,678.3	18.85 (17.43 to 20.38)
Exercise Status, detail	Inactive	228	83,437.2	27.33 (24.0 to 31.11)	120	42,011.0	28.56 (23.88 to 34.16)
	Gentle	547	213,811.7	25.58 (23.53 to 27.82)	280	104,290.0	26.85 (23.88 to 30.18)
	Moderate	607	365,295.8	16.62 (15.35 to 17.99)	287	179,751.4	15.97 (14.22 to 17.92)
	Vigorous	76	71,692.2	10.6 (8.47 to 13.27)	33	35,474.8	9.30 (6.61 to 13.08)
Smoking Status	Non/Ex	1,934	1,034,771.3	18.69 (17.88 to 19.54)	1,006	507,330.5	19.83 (18.64 to 21.09)
	Current	1,891	764,273.9	24.74 (23.65 to 25.88)	910	378,710.3	24.03 (22.52 to 25.64)
Alcohol Consumption	T-total/Ex	225	90,172.4	24.95 (21.9 to 28.44)	110	45,698.2	24.07 (19.97 to 29.02)
	Current	1,899	941,707.9	20.17 (19.28 to 21.09)	955	463,356.3	20.61 (19.34 to 21.96)
Diabetes	No	3,630	1,811,823.4	20.04 (19.39 to 20.7)	1,850	891,911.2	20.74 (19.82 to 21.71)
	Yes	262	16,995.0	154.16 (136.58 to 174.01)	102	8,589.8	118.75 (97.8 to 144.18)
LVH	No	3,888	1,828,117.6	21.27 (20.61 to 21.95)	1,949	900,209.1	21.65 (20.71 to 22.63)
	Yes	4	700.9	57.07 (21.42 to 152.07)	3	291.9	102.79 (33.15 to 318.7)

[Table continues on the next page]

Risk Factor cont.		Derivation			Validation		
		No of incident cases	**Total Person years**	**Crude Incidence Rate per 10000 person years (95% CI)**	**No of incident cases**	**Total Person years**	**Crude Incidence Rate per 10000 person years (95% CI)**
IGR	No	3,886	1,827,923.3	21.26 (20.6 to 21.94)	1,948	900,014.8	21.64 (20.7 to 22.63)
	Yes	6	895.2	67.02 (30.11 to 149.19)	4	486.2	82.27 (30.88 to 219.20)
Hypertensive treatment	No	2,454	1,588,547.1	15.45 (14.85 to 16.07)	1,214	780,206.5	15.56 (14.71 to 16.46)
	Yes	1,438	240,271.4	59.85 (56.83 to 63.02)	738	120,294.5	61.35 (57.08 to 65.94)
Chronic Kidney Disease Stage	1	43	15,222.1	28.25 (20.95 to 38.09)	31	7,579.5	40.90 (28.76 to 58.16)
	2	452	111,561.1	40.52 (36.95 to 44.43)	224	56,413.8	39.71 (34.83 to 45.26)
	3	319	39,302.1	81.17 (72.73 to 90.58)	162	19,943.6	81.23 (69.64 to 94.75)
	4	7	196.4	356.47 (169.94 to 747.72)	4	120.5	331.91 (124.57 to 884.34)
	5	0	10.0	-	-	-	-
Total		**3,892**	**1,828,818.5**	**21.28 (20.62 to 21.96)**	**1,952**	**900,501.0**	**21.68 (20.74 to 22.66)**

After removing patients who entered and exited the study of the same day, patients contributed 1,828,818.5 person years of observation to the derivation dataset; and 900,501.0 person years to the validation dataset. This gave comparable crude incidence rates of 21.28 (20.62 to 21.96) CVD events per 10,000 py and 21.68 (20.74 to 22.66) CVD events per 10.000 py, respectively.

Both datasets show a similar pattern of CVD incidence rate by FH status, gender, age category, and social deprivation; whereby CVD incidence rate is higher amongst patients with a positive FH of CVD, who are male, from an older age group, and are more deprived.

Table 5.12 shows the association of each standard risk factor and FH with CVD, overall and in the derivation and validation datasets. Statistically significant ($p<0.05$) associations were found for all risk factors except HDL cholesterol; this remained consistent for both the derivation and validation datasets. A statistically significant association was also found with the new risk factor, FH of CVD ($p<0.001$).

Overall, patients with a CVD event were older, male, current smokers, have a higher total cholesterol, lower HDL cholesterol, higher overall SBP and positive FH of CVD.

Table 5.12 Association of standard and new risk factor with CVD in the overall, derivation and validation datasets.

Overall		**no CVD (n=296,744)**		**CVD(n=5,847)**		**p value**
Family History	Negative	179,346	60.44	3,123	53.41	<0.001
	Positive	117,398	39.56	2,724	46.59	
Gender	Female	174,167	58.69	2,754	47.1	<0.001
	Male	122,577	41.31	3,093	52.9	
Smoker	No	167,235	56.36	2,942	50.32	
	Yes	124,670	42.01	2,802	47.92	<0.001
	Missing	4,839	1.63	103	1.76	
Median Age (years)		47	[38-57]	63	[53-72]	<0.001
Mean Total Cholesterol (mmol/L)		5.65	1.06	5.82	1.08	<0.001
	Missing	255,849	86.22	4,461	76.3	
Mean HDL Cholesterol (mmol/L)		1.42	0.41	1.39	0.42	0.1906
	Missing	282,706	95.27	5,389	92.17	
Mean SBP (mmHg)	Overall	129.42	17.52	141.98	18.46	
	Treated	141.5	18.18	148.57	17.47	<0.001
	Non treated	126.89	16.29	137.3	17.7	
	Missing	61,160	20.61	912	15.6	
Derivation		**no CVD (n=198,876)**		**CVD (n=3,894)**		**p value**
Family History	Negative	120,000	60.34	2,094	53.78	<0.001
	Positive	78,876	39.66	1,800	46.22	
Gender	Female	116,493	58.58	1,835	47.12	<0.001
	Male	82,383	41.42	2,059	52.88	
Smoker	No	112,266	56.45	1,935	49.69	
	Yes	83,354	41.91	1,892	48.59	<0.001
	Missing	3,256	1.64	67	1.72	
Median Age (years)		47	[38-57]	63	[53-72]	<0.001

[Table continues on the next page]

Derivation cont.		**no CVD (n=198,876)**		**CVD (n=3,894)**		**p value**
Mean Total Cholesterol (mmol/L)		5.65	1.05	5.81	1.09	<0.001
	Missing	171,637	86.3	2,966	76.17	
Mean HDL Cholesterol (mmol/L)		1.42	0.41	1.41	0.44	0.6477
	Missing	189,493	95.28	3,596	92.35	
Mean SBP (mmHg)	Overall	129.4	17.51	141.76	18.51	
	Treated	141.5	18.19	148.56	17.58	<0.001
	Non treated	126.89	16.28	136.99	17.65	
	Missing	41,083	20.66	612	15.72	
Validation		**no CVD (n=97,868)**		**CVD (n=1,953)**		**p value**
Family History	Negative	59,346	60.64	1,029	52.69	<0.001
	Positive	38,522	39.36	924	47.31	
Gender	Female	57,674	58.93	919	47.06	<0.001
	Male	40,194	41.07	1,034	52.94	
Smoker	No	54,969	56.17	1,007	51.56	
	Yes	41,316	42.22	910	46.59	<0.001
	Missing	1,583	1.62	36	1.84	
Median Age		47	[38-57]	63	[53-72]	<0.001
Mean Total Cholesterol (mmol/L)		5.65	1.06	5.83	1.04	<0.001
	Missing	84,212	86.05	1,495	76.55	
Mean HDL Cholesterol (mmol/L)		1.41	0.41	1.36	0.37	0.1128
	Missing	93,213	95.24	1,793	91.81	
Mean SBP (mmHg)	Overall	129.44	17.55	142.42	18.34	
	Treated	141.52	18.14	148.58	17.26	<0.001
	Non treated	126.89	16.31	137.92	17.79	
	Missing	20,077	20.51	300	15.36	

5.1.7 Derivation of Models

5.1.7.1 Exploration of risk factors

Associations between risk factors were assessed for potential confounding using the derivation dataset. The effect sizes of association based on linear regression for continuous risk factors (age, total cholesterol, HDL cholesterol, SBP, and BMI), and logistic regression for categorical risk factors (smoking status, alcohol consumption, Townsend deprivation score, and ethnicity) are presented in Table 5.13. Age was the only non-normal continuous variable, and so was log-transformed as the outcome for analysis. A Bonferroni adjusted significance level of 0.004 was used (0.05/24).

Statistically significant associations were found between standard and additional risk factors, and the new risk factor; these have been highlighted in Table 5.13. Age was a strong predictor for all risk factors and was found to be higher amongst females, non-smokers, patients who do not currently consume alcohol, lowered deprivation score, increased total and HDL cholesterol, increased SBP, increased BMI, patients with a white ethnicity, and patients with a positive FH status ($p<0.001$).

Other strong predictors for a positive FH status were gender, smoking status, total cholesterol, SBP, BMI, and alcohol consumption ($p<0.001$). The odds of having a positive family history increased with increased age, in females, current smokers, increasing cholesterol, increasing SBP, increasing BMI, and patients who consume alcohol. The odds of a positive FH status also decreased in South Asian and Black patients compared to White patients, and patients with a lower deprivation score.

Table 5.13 Association between all risk factors included in the analysis to derive 10 year CVD risk prediction models. Values are beta coefficients (p-values) derived using logistic regression, unless otherwise stated.

Predictors	Outcome													
	Log Age*	Gender (males)	Current Smoker (yes)	Consume Alcohol (yes)	Total Cholesterol (mmol/L)**	HDL Cholesterol (mmol/L)**	SBP (mmHg)**	Diabetes (yes)	BMI**	Townsend Deprivation Score***	Ethnicity***			FH Status (positive)
											Ethnicity (S. Asian)	Ethnicity (Black)	Ethnicity (Other/Mixed)	
Age (years)	-	-0.003 (<0.001)	-0.013 (<0.001)	-0.010 (<0.001)	0.022 (<0.001)	0.004 (<0.001)	0.703 (<0.001)	0.058 (<0.001)	0.037 (<0.001)	-0.002 (<0.001)	-0.037 (<0.001)	-0.023 (<0.001)	-0.009 (<0.001)	0.004 (<0.001)
Gender (males)	-0.010 (<0.001)	-	0.099 (<0.001)	0.617 (<0.001)	-0.095 (<0.001)	-0.293 (<0.001)	4.462 (<0.001)	0.470 (<0.001)	0.625 (<0.001)	-0.008 (0.403)	-0.124 (0.045)	-0.172 (0.079)	0.069 (<0.001)	-0.122 (<0.001)
Current Smoker (yes)	-0.043 (<0.001)	0.099 (<0.001)	-	1.138 (<0.001)	0.002 (0.859)	-0.039 (<0.001)	-1.748 (<0.001)	0.242 (<0.001)	-0.445 (<0.001)	0.368 (<0.001)	-0.405 (<0.001)	0.011 (0.91)	0.0340 (0.033)	0.363 (<0.001)
Consume Alcohol (yes)	-0.031 (<0.001)	0.617 (<0.001)	1.138 (<0.001)	-	0.018 (0.576)	0.051 (0.007)	-0.356 (0.056)	-0.582 (<0.001)	-0.367 (<0.001)	-0.613 (<0.001)	-2.158 (<0.001)	-0.913 (<0.001)	-0.138 (<0.001)	0.308 (<0.001)
Total Cholesterol (mmol/L)	0.054 (<0.001)	-0.086 (<0.001)	0.002 (0.859)	0.016 (0.576)	-	0.04 (<0.001)	2.892 (<0.001)	-0.187 (<0.001)	0.422 (<0.001)	0.045 (<0.001)	-0.357 (<0.001)	-0.184 (0.082)	-0.031 (0.099)	0.031 (0.007)
HDL Cholesterol (mmol/L)	0.066 (<0.001)	-2.095 (<0.001)	-0.23 (<0.001)	0.317 (0.007)	0.278 (<0.001)	-	-0.548 (0.21)	-0.898 (<0.001)	-2.944 (<0.001)	-0.308 (<0.001)	-2.016 (<0.001)	-0.52 (0.285)	0.069 (0.388)	-0.039 (0.431)
SBP(mmHg)	0.007 (<0.001)	0.014 (<0.001)	-0.006 (<0.001)	-0.001 (0.056)	0.01 (<0.001)	0 (0.21)	-	0.033 (<0.001)	0.085 (<0.001)	0.002 (<0.001)	-0.026 (<0.001)	-0.006 (0.03)	-0.006 (<0.001)	0.007 (<0.001)
Diabetes (yes)	0.189 (<0.001)	0.470 (<0.001)	0.242 (<0.001)	-0.582 (<0.001)	-0.204 (<0.001)	-0.134 (<0.001)	11.529 (<0.001)	-	3.416 (<0.001)	0.29 (<0.001)	0.97 (<0.001)	0.591 (0.017)	-0.062 (0.266)	0.08 (0.019)
BMI	0.005 (<0.001)	0.025 (<0.001)	-0.018 (<0.001)	-0.016 (<0.001)	0.019 (<0.001)	-0.02 (<0.001)	1.066 (<0.001)	0.101 (<0.001)	-	0.027 (<0.001)	-0.026 (<0.001)	0.033 (<0.001)	-0.016 (<0.001)	0.023 (<0.001)

[Table continues on the next page]

Predictors cont.	Outcome													
	Log Age*	Gender (males)	Current Smoker (yes)	Consume Alcohol (yes)	Total Cholesterol (mmol/L)**	HDL Cholesterol (mmol/L)**	SBP**	Diabetes (yes)	BMI**	Townsend Deprivation Score***	Ethnicity***			FH Status (positive)
											Ethnicity (S. Asian)	Ethnicity (Black)	Ethnicity (Other/Mixed)	
Townsend deprivation score (2)	0.003 (0.044)	-0.002 (0.898)	0.049 (0.002)	-0.260 (<0.001)	0.059 (0.004)	-0.006 (0.662)	0.906 (<0.001)	0.169 (0.006)	0.192 (<0.001)	-	-0.386 (0.003)	0.077 (0.694)	0.002 (0.909)	0.038 (0.013)
Townsend deprivation score (3)	-0.007 (<0.001)	-0.022 (0.176)	0.225 (<0.001)	-0.476 (<0.001)	0.057 (0.009)	-0.032 (0.035)	0.572 (<0.001)	0.253 (<0.001)	0.454 (<0.001)	-	0.115 (0.334)	0.549 (0.002)	0.028 (0.216)	-0.015 (0.353)
Townsend deprivation score (4)	-0.015 (<0.001)	-0.019 (0.295)	0.442 (<0.001)	-0.765 (<0.001)	0.091 (<0.001)	-0.082 (<0.001)	0.673 (<0.001)	0.390 (<0.001)	0.832 (<0.001)	-	1.097 (<0.001)	1.570 (<0.001)	0.002 (0.947)	0.145 (<0.001)
Townsend deprivation score (5)	-0.022 (<0.001)	0.008 (0.729)	0.708 (<0.001)	-1.049 (<0.001)	0.057 (0.056)	-0.082 (<0.001)	1.340 (<0.001)	0.504 (<0.001)	1.158 (<0.001)	-	1.213 (<0.001)	1.623 (<0.001)	-0.143 (<0.001)	0.354 (<0.001)
Ethnicity (S. Asian)	-0.104 (<0.001)	-0.124 (0.045)	-0.405 (<0.001)	-2.158 (<0.001)	-0.38 (<0.001)	-0.258 (<0.001)	-7.513 (<0.001)	0.970 (<0.001)	-0.633 (<0.001)	1.009 (<0.001)	-	-	-	-0.226 (<0.001)
Ethnicity (Black)	-0.065 (<0.001)	-0.172 (0.079)	0.011 (0.91)	-0.913 (<0.001)	-0.201 (0.081)	-0.084 (0.286)	-1.961 (0.029)	0.591 (0.017)	0.945 (<0.001)	1.199 (<0.001)	-	-	-	-0.404 (<0.001)
Ethnicity (Other/Mixed)	-0.029 (<0.001)	0.069 (<0.001)	0.034 (0.033)	-0.138 (<0.001)	-0.034 (0.099)	0.013 (0.386)	-1.782 (<0.001)	-0.062 (0.266)	-0.388 (<0.001)	-0.031 (0.027)	-	-	-	-0.025 (0.109)
FH Status (positive)	0.018 (<0.001)	-0.122 (<0.001)	0.363 (<0.001)	0.308 (<0.001)	0.034 (0.007)	-0.007 (0.431)	2.132 (<0.001)	0.080 (0.019)	0.552 (<0.001)	0.131 (<0.001)	-0.226 (<0.001)	-0.400 (<0.001)	-0.025 (0.109)	-

*Outcome is log-transformed age; analysed using linear regression

**Analysed using linear regression

*** Ordinal logistic regression used for Townsend Deprivation Score; Multinomial logistic regression used for ethnicity.

5.1.7.2 Model Building

Models were developed using the imputed dataset described above (n=202,770). Data was imputed for all incomplete standard risk factors (total cholesterol, HDL cholesterol, smoking status, and SBP) and incomplete additional risk factors statistically significantly associated with CVD (shown in Table 5.3) and with less than 50% missing data (alcohol consumption, BMI, Townsend deprivation score, and ethnicity). Other additional risk factors were excluded from the analysis, due to insufficient data.

5.1.7.2.1 Unadjusted analysis

The association of each risk factor with the CVD outcome was assessed. The results of the unadjusted Cox regression models are presented in Table 5.14.

Table 5.14 Unadjusted Cox Regression Analysis.

Risk Factor		Imputed Unadjusted Analysis				
		Beta	SE	p-value	HR	95% CI
Age (years)		0.08	0.00	<0.001	1.09	(1.09 to 1.09)
Gender	Female	Ref				
	Male	0.49	0.03	<0.001	1.62	(1.53 to 1.73)
Smoking Status	No	Ref				
	Yes	0.28	0.03	<0.001	1.32	(1.24 to 1.41)
Total Cholesterol (mmol/L)		0.30	0.02	<0.001	1.35	(1.29 to 1.41)
HDL Cholesterol (mmol/L)		-0.09	0.16	0.567	0.91	(0.64 to 1.29)
SBP (mmHg)	Not treated	0.03	0.00	<0.001	1.03	(1.03 to 1.03)
	Treated	0.04	0.00	<0.001	1.04	(1.03 to 1.04)
Family History	No	Ref				
	Yes	0.30	0.03	<0.001	1.34	(1.26 to 1.43)
Ethnicity	White	Ref				
	S. Asian	-0.42	0.19	0.023	0.66	(0.46 to 0.94)
	Black	-0.20	0.27	0.455	0.82	(0.48 to 1.39)
	Other/Mixed	-0.43	0.05	<0.001	0.65	(0.59 to 0.71)
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	Ref				
	2	0.23	0.05	<0.001	1.26	(1.14 to 1.38)
	3	0.33	0.05	<0.001	1.39	(1.26 to 1.53)
	4	0.60	0.05	<0.001	1.81	(1.64 to 2.00)
	5	0.80	0.06	<0.001	2.23	(2.00 to 2.49)
Diabetes	No	Ref				
	Yes	2.11	0.06	<0.001	8.28	(7.30 to 9.39)
BMI		0.05	0.00	<0.001	1.06	(1.05 to 1.06)
Consumes Alcohol	No	Ref				
	Yes	-0.23	0.06	<0.001	0.79	(0.70 to 0.90)

The unadjusted analysis presented in Table 5.14 shows increased hazard with increasing age, for male, smokers, and patients with a positive FH status. Patients with a positive FH of CVD have a 34% higher hazard of CVD (HR = 1.34 (95% CI: 1.26 to 1.43)). Hazard of CVD increased by 9% per year increase in age (HR = 1.09 (95% CI: 1.09 to 1.09)); 62% in males (HR = 1.62 (95% CI: 1.53 to 1.73)) and 32% in smokers (HR = 1.32 (95% CI: 1.24 to 1.41)). However, patients who consumed alcohol showed 21% lower hazard of CVD (HR = 0.79 (95% CI: 0.70 to 0.90)).

Of the other additional risk factors, increased hazard of CVD was found with increased deprivation score diabetes and BMI; patients with a Townsend deprivation score of 5 had over 2 times higher hazard of CVD compared to patients with a score of 1. Lower hazard was found in non-white ethnicities, but this did not show statistical significance ($p>0.05$).

The unadjusted analysis also showed increased hazard of CVD per increment in SBP; a higher hazard was found with patients taking hypertensive treatment at baseline (HR = 1.04 (95% CI: 1.03 to 1.04)) compared to those who were not (HR = 1.03 (95% CI: 1.03 to 1.03)).

5.1.7.2.2 Adjusted Analyses

5.1.7.2.2.1 Model 1 – Standard CVD Risk Factors

Model 1 of the analysis includes the standard risk factors of CVD only, based on the Framingham CVD risk prediction model. The output of this model is presented in Table 5.15.

Table 5.15 Beta Coefficients and Hazard Ratios from Model 1 based on standard CVD risk factors.

Risk Factor		MI Adjusted Analysis (Model 1:SRF only)				
		Beta	SE	p value	HR	95% CI
Age (years)		0.07	0.00	<0.001	1.08	(1.07 to 1.08)
Gender	Female	Ref				
	Male	0.61	0.06	<0.001	1.84	(1.62 to 2.09)
Smoker	No	Ref				
	Yes	0.47	0.03	<0.001	1.60	(1.50 to 1.70)
Total Cholesterol (mmol/L)		0.04	0.03	0.167	1.04	(0.98 to 1.10)
HDL Cholesterol (mmol/L)		-0.01	0.18	0.951	0.99	(0.66 to 1.49)
SBP (mmHg)	Not treated	0.01	0.00	<0.001	1.01	(1.01 to 1.01)
	Treated	0.01	0.00	<0.001	1.01	(1.01 to 1.02)

The adjusted analysis including standard CVD risk factors continues to show increased hazard of CVD with increased age, males, smokers, increased total cholesterol, and SBP (treated and untreated). A 1% lower hazard was found per unit increase in HDL cholesterol; however this was not statistically significant.

Total cholesterol continued to show increased hazard per unit increase but reduced to a 4% (HR =1.04 (95% CI: 0.98 to 1.10) increased hazard of CVD after the adjustment of other standard CVD risk factors; and became a statistically insignificantly result.

Although both treated and untreated SBP continued to who increased hazard of CVD per increment in SBP; no difference was found with treated and non-treated patients. Both sets of patients a 1% (HR=1.01 (95% CI: 1.01 to 1.02), HR=1.01 (95% CI: 1.01 to 1.01), respectively).

5.1.7.2.2.1.1 Proportional Hazards

The proportional hazards assumption was checked for Model 1 using Schoenfeld Residuals; this assumption was checked for each imputed dataset separately. Figure 5.3 shows the Schoenfeld residual plots for each risk factor of Model 1 using imputed dataset 1. Similar plots for the other imputed datasets can be found in Appendix G.

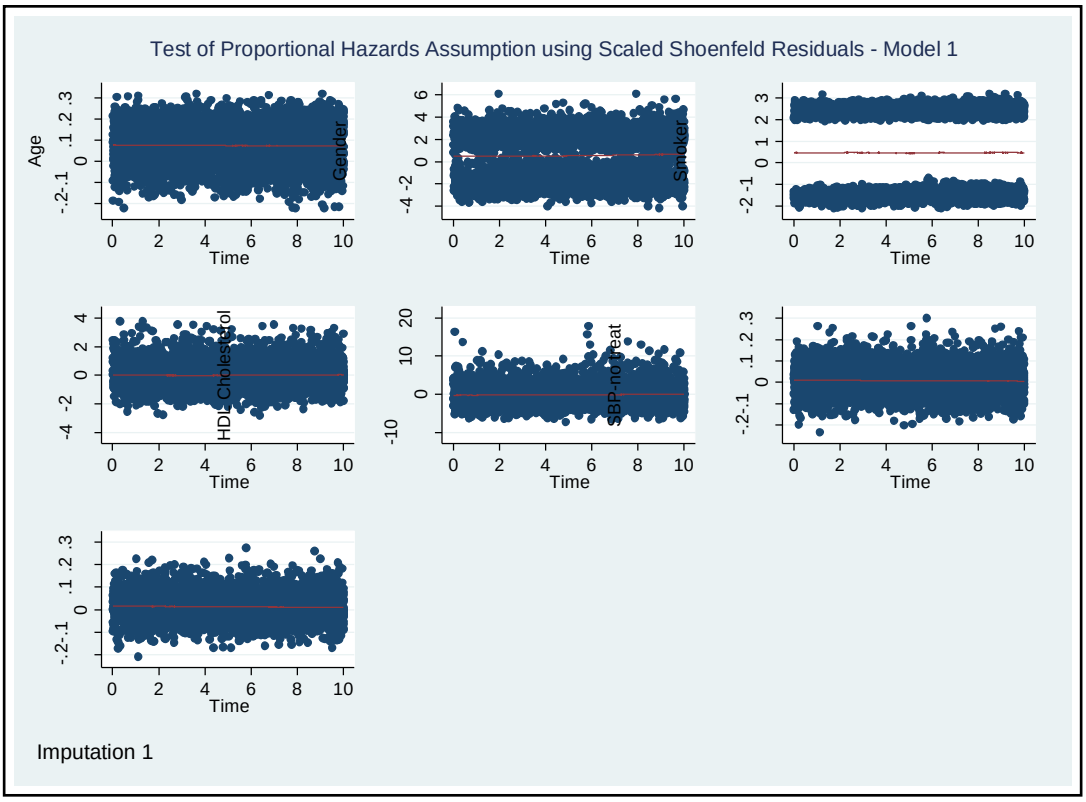


Figure 5.3 Schoenfeld residual plots for each risk factor in Model 1 (including standard risk factors only) using imputed dataset 1.

Model 1 did not violate the proportional hazards assumption; no trend was found in the residuals for each risk factor which deviated greatly from 0.

5.1.7.2.2.2 Model 2 – Standard CVD Risk Factors with FH

Model 2 of the analysis adds FH information to the standard CVD risk factors. The output of this model is presented in Table 5.16.

Table 5.16 Beta Coefficients and Hazard Ratios from Model 2 based on standard CVD risk factors with FH.

Risk Factor		MI Adjusted Analysis				
		Beta	**SE**	**p value**	**HR**	**95% CI**
Age (years)		0.08	0.00	<0.001	1.08	(1.07 to 1.08)
Gender	Female	Ref				
	Male	0.62	0.06	<0.001	1.87	(1.64 to 2.13)
Smoker	No	Ref				
	Yes	0.45	0.03	<0.001	1.57	(1.47 to 1.67)
Total Cholesterol (mmol/L)		0.03	0.03	0.210	1.04	(0.98 to 1.10)
HDL Cholesterol (mmol/L)		0.00	0.18	0.979	1.00	(0.66 to 1.50)
SBP (mmHg)	Not treated	0.01	0.00	<0.001	1.01	(1.01 to 1.01)
	Treated	0.01	0.00	<0.001	1.01	(1.01 to 1.01)
Family History	Negative	Ref				
	Positive	0.26	0.03	<0.001	1.29	(1.21 to 1.38)

Adding FH information into the risk model has little effect in the associations of the standard risk factors and their effect size. FH information itself showed a statistically significant relationship with hazard of CVD; after adjusting for the standard CVD risk factors patients with a positive FH of CVD had 29% (HR = 1.29 (95% CI: 1.21 to 1.38) higher hazard of a CVD event vs. patients with a negative FH of CVD.

Other risk factors in the model maintained the strength of their associations with the hazard on CVD; total and HDL cholesterol remained insignificant in the model. Further, no change was in the hazard of CVD per unit increase of HDL cholesterol (HR=1.00 (95% CI: 0.66 to 1.50).

5.1.7.2.2.2.1 Proportional Hazards

The proportional hazards assumption was checked for Model 2 using Schoenfeld Residuals; this assumption was checked for each imputed dataset separately. Figure 5.4 shows the Schoenfeld residual plots for each risk factor of Model 2 using imputed dataset 1. Similar plots for the other imputed datasets can be found in Appendix G.

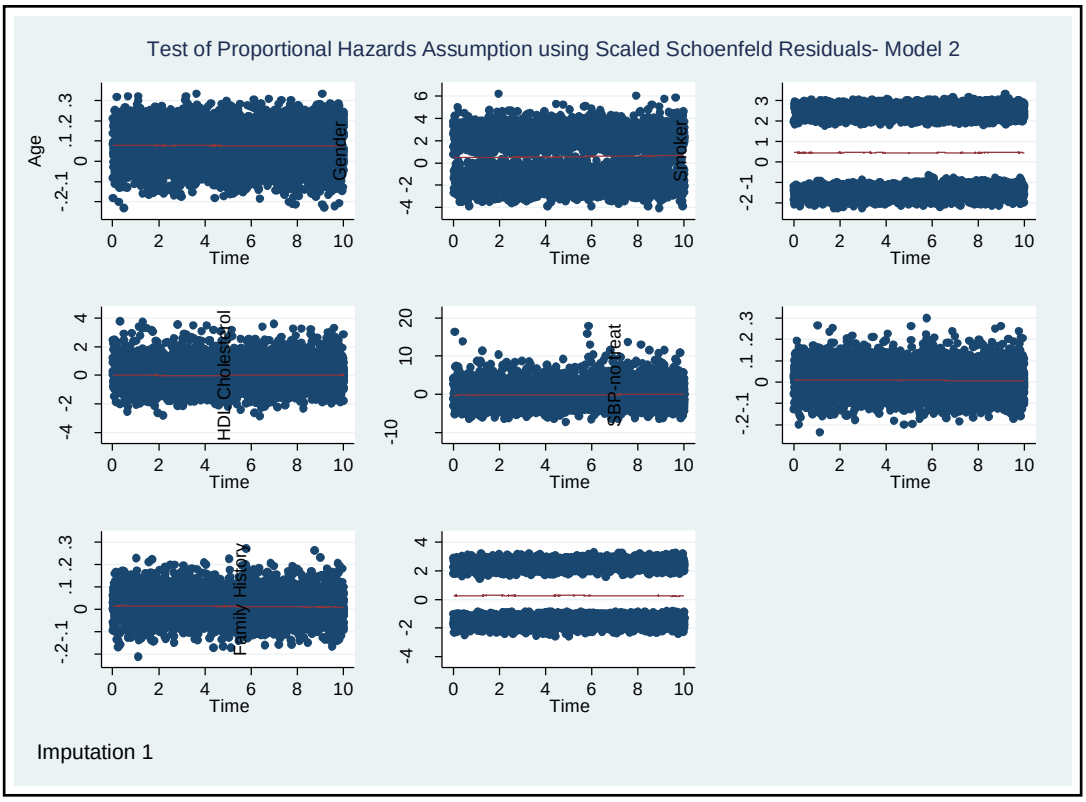


Figure 5.4 Schoenfeld residual plots for each risk factor in Model 2 (including standard risk factors and FH) using imputed dataset 1.

Model 2 did not violate the proportional hazards assumption; no trend was found in the residuals for each risk factor that deviated greatly from 0.

5.1.7.2.2.3 Model 3 - Standard and Additional CVD Risk Factors

Based on the unadjusted analysis of the additional recognised CVD risk factors, diabetes, Townsend deprivation score, BMI, and alcohol consumption showed statistically significant associations with the hazard of CVD. These were used to create Model 3 and

were included in the model if they showed statistically significant associations with CVD after adjustment of the standard risk factors, based on a significance level of $p<0.05$.

Diabetes, deprivation score, BMI and alcohol consumption were all included into the model. The output from this model is presented in Table 5.17.

Table 5.17 Beta Coefficients and Hazard Ratios from Model 3 based on standard and additional CVD risk factors.

Risk Factor		MI Adjusted Analysis				
		Beta	SE	p value	HR	95% CI
Age (years)		0.07	0	<0.001	1.08	(1.07 to 1.08)
Gender	Female	Ref				
	Male	0.65	0.06	<0.001	1.92	(1.68 to 2.19)
Smoker	No	Ref				
	Yes	0.44	0.03	<0.001	1.55	(1.45 to 1.66)
Total Cholesterol (mmol/L)		0.04	0.03	0.163	1.04	(0.98 to 1.10)
HDL Cholesterol (mmol/L)		0.16	0.19	0.411	1.18	(0.77 to 1.80)
SBP (mmHg)	Not treated	0.01	0	<0.001	1.01	(1.00 to 1.01)
	Treated	0.01	0	<0.001	1.01	(1.01 to 1.01)
Diabetes	No	Ref				
	Yes	1.03	0.07	<0.001	2.79	(2.44 to 3.19)
BMI		0.03	0.01	<0.001	1.03	(1.02 to 1.04)
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	Ref				
	2	0.18	0.05	<0.001	1.20	(1.09 to 1.32)
	3	0.27	0.05	<0.001	1.32	(1.19 to 1.45)
	4	0.52	0.05	<0.001	1.69	(1.52 to 1.87)
	5	0.71	0.06	<0.001	2.03	(1.80 to 2.29)
Alcohol Consumption	No	Ref				
	Yes	-0.14	0.06	0.03	0.87	(0.77 to 0.99)

The results of the Cox regression analysis presented in Table 5.17, show similar results for the standard CVD risk factors used in Model 1 (Table 5.15) except HDL cholesterol.

Adding diabetes, Townsend deprivation score, BMI, and alcohol consumption did not however change the hazard of increasing age, males, smokers, total cholesterol, and treated and untreated SBP. However, an 18% increased risk can be found per unit increase of HDL cholesterol (HR=1.18 (95% CI: 0.77 to 1.80)); which remains statistically insignificant.

Compared to the unadjusted analysis, similar effect sizes were found with the additional risk factors. The hazard of CVD increased in patients with diabetes, increasing BMI, and deprivation score; but decreased in patients consuming alcohol. Patients with diabetes had an almost 3 fold increase in hazard of CVD (HR=2.79 (95% CI: 2.44 to 1.32)); per unit increase of BMI increased hazard by 3% (HR=1.03 (95% CI: 1.02 to 1.04)); patients with a Townsend deprivation score of 5 had 2 times higher hazard than patients with a score of 1 (HR=2.03 (95% CI: 1.80 to 2.29)); and patients who consumed alcohol had 13% lower hazard of CVD compared to patients who did not consume alcohol or had given up (HR=0.87 (95% CI: 0.77 to 0.99)).

5.1.7.2.2.3.1 Proportional Hazards

The proportional hazards assumption was checked for Model 3 using Schoenfeld Residuals; this assumption was checked for each imputed dataset separately. Figure 5.5 shows the Schoenfeld residual plots for each risk factor of Model 3 using imputed dataset 1. Similar plots for the other imputed datasets can be found in Appendix G.

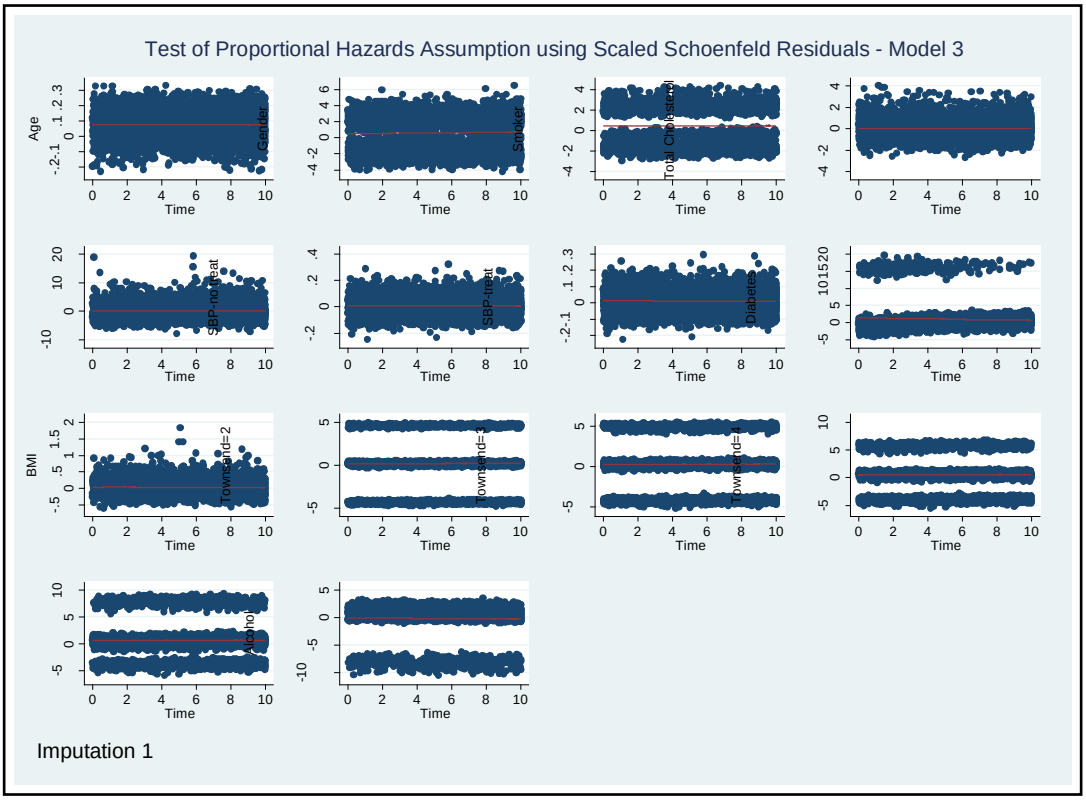


Figure 5.5 Schoenfeld residual plots for each risk factor in Model 3 (including standard and additional risk factors) using imputed dataset 1.

Model 3 did not violate the proportional hazards assumption; no trend was found in the residuals for each risk factor that deviated greatly from 0.

5.1.7.2.2.4 Model 4 - Standard and Additional CVD Risk Factors with FH

Family history information was added to Model 3 which included standard and additional CVD risk factors. The results from the Cox regression for Model 4 are presented in Table 5.18.

Table 5.18 Beta Coefficients and Hazard Ratios from Model 3 based on standard and additional CVD risk factors with FH.

Risk Factor		MI Adjusted Analysis				
		Beta	SE	p value	HR	95% CI
Age (years)		0.07	0	<0.001	1.08	(1.07 to 1.08)
Gender	Female	Ref				
	Male	0.67	0.06	<0.001	1.95	(1.71 to 2.22)
Smoker	No	Ref				
	Yes	0.42	0.03	<0.001	1.53	(1.43 to 1.63)
Total Cholesterol (mmol/L)		0.04	0.03	0.195	1.04	(0.98 to 1.10)
HDL Cholesterol (mmol/L)		0.17	0.19	0.399	1.18	(0.77 to 1.81)
SBP (mmHg)	Not treated	0.01	0	<0.001	1.01	(1.00 to 1.01)
	Treated	0.01	0	<0.001	1.01	(1.01 to 1.01)
Diabetes	No	Ref				
	Yes	1.04	0.07	<0.001	2.82	(2.47 to 3.22)
BMI		0.03	0.01	<0.001	1.03	(1.02 to 1.04)
Townsend Deprivation Score (1=less deprived; 5=more deprived)	1	Ref				
	2	0.18	0.05	<0.001	1.20	(1.09 to 1.32)
	3	0.27	0.05	<0.001	1.32	(1.19 to 1.45)
	4	0.52	0.05	<0.001	1.68	(1.51 to 1.86)
	5	0.70	0.06	<0.001	2.01	(1.79 to 2.26)
Alcohol Consumption	No	Ref				
	Yes	-0.15	0.06	0.016	0.86	(0.76 to 0.97)
Family History	Negative	Ref				
	Positive	0.25	0.03	<0.001	1.29	(1.21 to 1.37)

Adding FH information to standard and additional risk factors produced similar results to Model 3. All risk factors showed statistically significant association with CVD, except

total and HDL cholesterol; increased hazard of CVD was found per increment of HDL cholesterol.

Patients with a positive FH had 29% higher hazard of CVD compared to patients with a negative FH (HR=1.29 (95% CI: 1.21 to 1.37)).

5.1.7.2.2.4.1 Proportional Hazards

The proportional hazards assumption was checked for Model 4 using Schoenfeld Residuals; this assumption was checked for each imputed dataset separately. Figure 5.6 shows the Schoenfeld residual plots for each risk factor of Model 4 using imputed dataset 1. Similar plots for the other imputed datasets can be found in Appendix G.

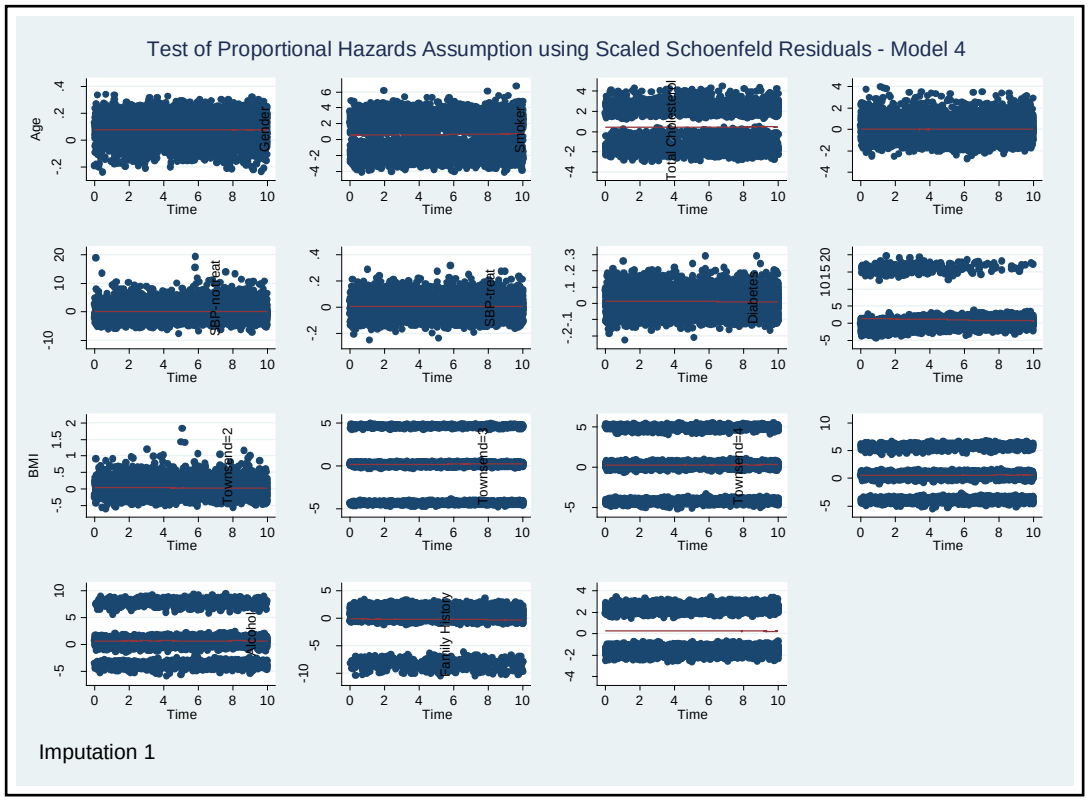


Figure 5.6 Schoenfeld residual plots for each risk factor in Model 4 (including standard and additional risk factors, with FH) using imputed dataset 1.

Model 4 did not violate the proportional hazards assumption; no trend was found in the residuals for each risk factor that deviated greatly from 0.

5.1.7.2.3 Residuals – Goodness of Fit, Outliers, and Influential Observations

Cox Snell residuals were checked to assess the goodness of fit of each model. DFBETA, log-displacement values, and LMAX value were used to identify outliers and influential observations.

Figure 5.7 shows the Cox Snell residual plots for each model, using the first imputed dataset; and does not show improvement of fit when adding FH to the standard risk factors (Model 1) or standard and additional CVD risk factors (Model 3). Similar results were found for other imputed datasets.

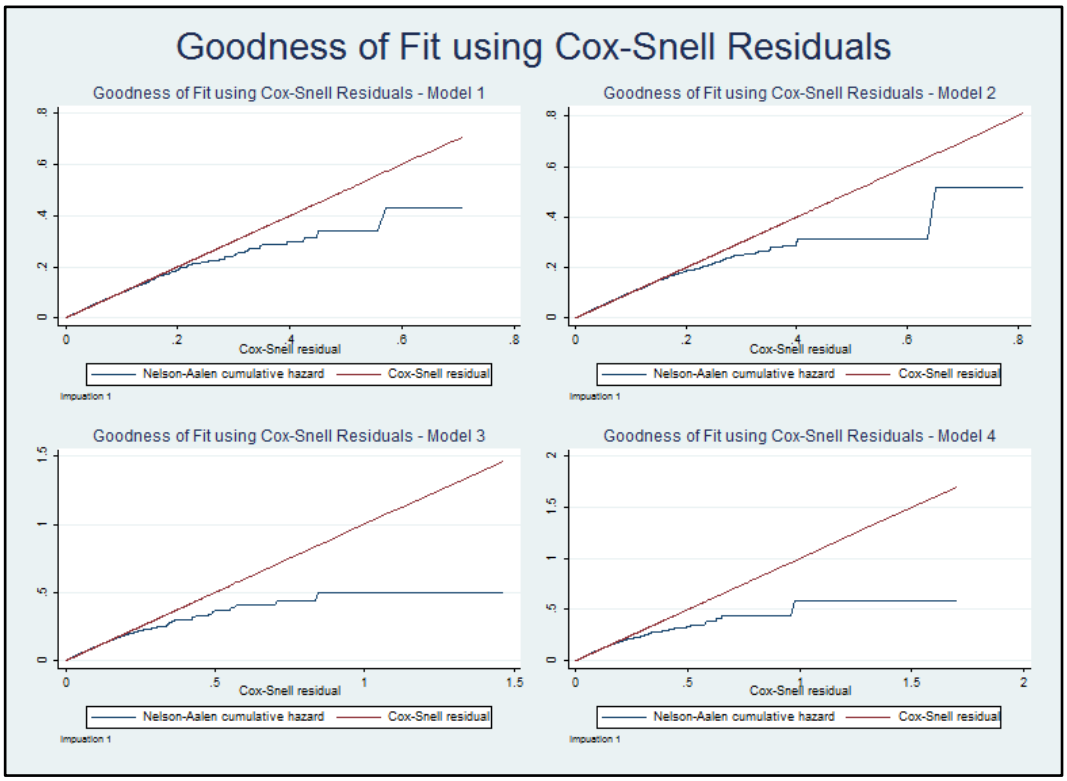


Figure 5.7 Cox-Snell Residual plots for each model using Imputation 1.

Figure 5.8 and Figure 5.9 show the DFBETA residual, log-displacement and LMAX value plots. Recurring outliers and influential observations were found in both figures, which were based on Imputation dataset 1; similar results were found for other imputed

datasets. Identified outliers were patients with risk factors for CVD, but did not have the event during follow up; and remained in the analysis.

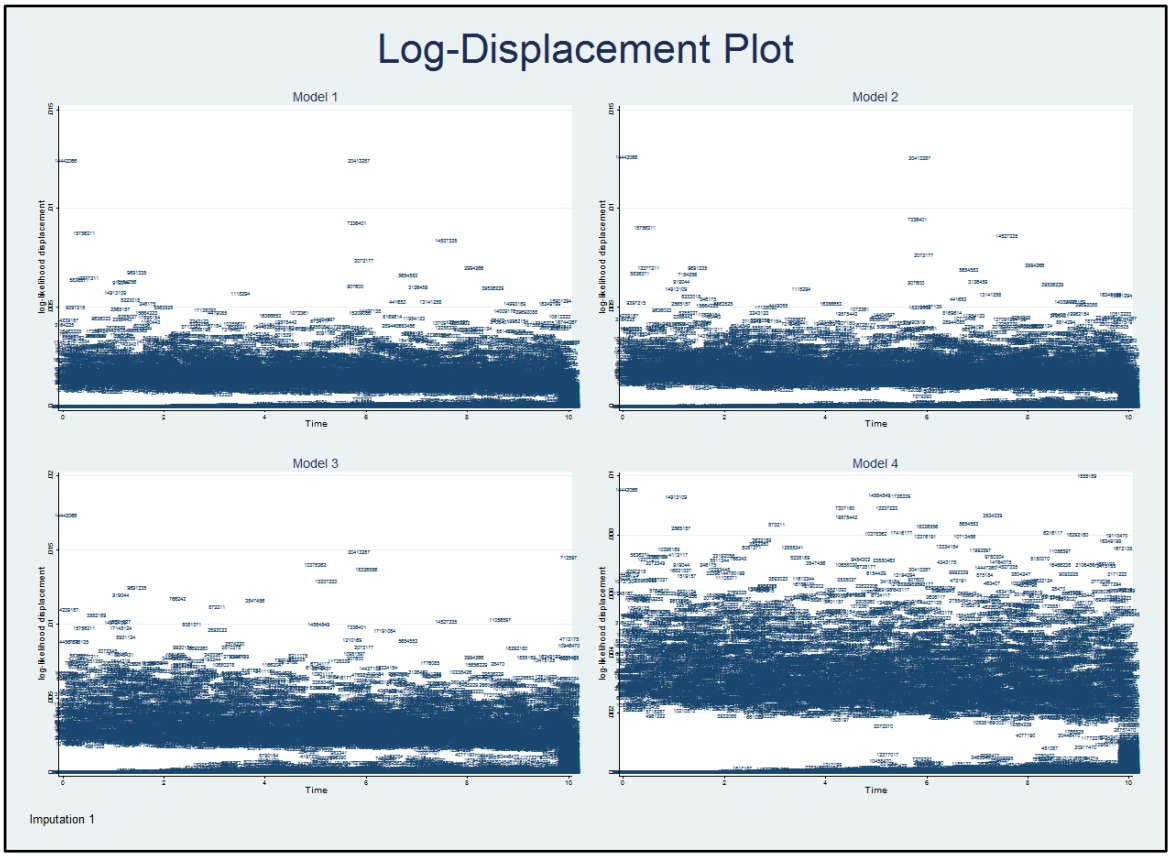


Figure 5.8 Log-displacement plots for each model using Imputation 1.

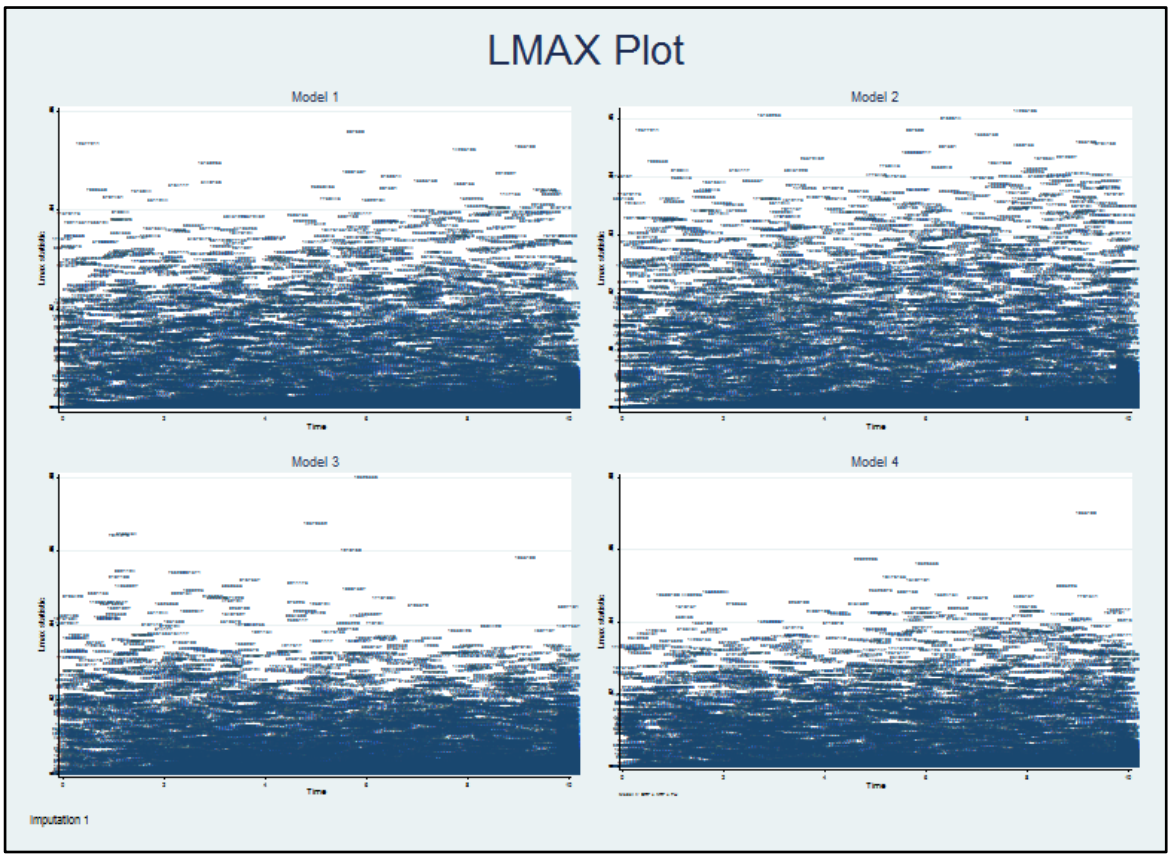


Figure 5.9 LMAX plots for each model using Imputation 1.

5.1.7.2.4 Model Fit

The AIC, BIC, likelihood ratios, and Royston's modified R^2 statistic were reported for all models. To assess the change in the likelihood ratio's a Likelihood Ratio test was used. The results of these tests for Imputation 1 are presented in Table 5.19.

Table 5.19 AIC and BIC values, Royston's R^2 statistic, and Likelihood Ratio Tests to compare Model 1 to Model 2; and Model 3 to Model 4; using the Imputation 1 of the derivation dataset.

Model	AIC	BIC	R2	LR	df	ΔLR	p (df)
Model 1	88601.98	88673.51	66.75%	5679.20	7	62.53	<0.0001 (1)
Model 2	88541.44	88623.20	67.21%	5741.73	8		
Model 3	88113.48	88256.55	70.31%	6181.69	14	60.56	<0.0001 (1)
Model 4	88054.92	88208.21	70.72%	6242.25	15		

Table 5.19 shows a reduction in both the AIC and BIC values when adding FH information to CVD risk prediction. The R^2 statistic also shows a higher proportion of variation accounted for in models with FH information; an increase of 0.46% was found adding FH to standard risk factors; and an increase of 0.41% was found when adding FH to standard and additional CVD risk factors.

Similar results were found for the other imputed datasets, presented in Appendix H.

5.1.7.3 Final CVD Risk Prediction Models

In total, 4 CVD models were developed; the beta coefficients of which were used to create 4 CVD risk prediction models. The beta coefficients for each model are presented in Table 5.15-5.18; the respective risk prediction models are presented in Table 5.20. A 10 year observed survival rate of 0.9788 was used.

Table 5.20 Risk equations to calculate 10 year risk based on patient characteristics for each developed risk model (1-4).

General Risk Equation to calculate 10 year risk	$1 - 0.9788^{exp(\sum_{i=1}^{p} \beta_i X_i - \sum_{i=1}^{p} \beta_i \bar{X}_i)}$
Risk Model	**Equation**
Model 1	where $\sum_{i=1}^{p} \beta_i X_i =$ age*0.0744807+gender*0.6092344+smoking status*0.4675567+ total cholesterol*0.0385686+HDL cholesterol*-0.0115373+ not treated SBP*0.0079478+treat SBP*0.012793 where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ mean age*0.0744807+mean gender*0.6092344+ mean smoking status*0.4675567+ mean total cholesterol*0.0385686+mean HDL cholesterol*-0.0115373+ mean not treated SBP*0.0079478+mean treat SBP*0.012793
Model 2	where $\sum_{i=1}^{p} \beta_i X_i =$ age*0.0754907+gender*0.6243609+smoking status*0.4484851+ total cholesterol*0.0347755+HDL cholesterol*-0.0049942+ not treated SBP*0.0077431+treat SBP*0.0124276+family history*0.2577314 where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ mean age*0.0754907+mean gender*0.6243609+ mean smoking status*0.4484851+mean total cholesterol*0.0347755+ mean HDL cholesterol*-0.0049942+mean not treated SBP*0.0077431+ mean treat SBP*0.0124276+mean family history*0.2577314

[Table continues on the next page]

Risk Model cont.	Equation
Model 3	where $\sum_{i=1}^{p} \boldsymbol{\beta}_i \boldsymbol{X}_i =$ age*0.0736218+gender*0.6513517+smoking status*0.4410497+ total cholesterol*0.0397922+HDL cholesterol*0.1632083+ not treated SBP*0.0060994+treat SBP*0.0103684+ diabetes*1.02692+BMI*0.0293807+Townsend=2*0.1843656+ Townsend=3*0.2740791+Townsend=4*0.522959+ Townsend=5*0.7078391+consumes alcohol*-0.1394517 where $\sum_{i=1}^{p} \boldsymbol{\beta}_i \bar{\boldsymbol{X}}_i =$ mean age*0.0736218+mean gender*0.6513517+ mean smoking status*0.4410497+mean total cholesterol*0.0397922+ mean HDL cholesterol*0.1632083+mean not treated SBP*0.0060994+ mean treat SBP*0.0103684+mean diabetes*1.02692+mean BMI*0.0293807 +mean Townsend=2*0.1843656+mean Townsend=3*0.2740791+ mean Townsend=4*0.522959+mean Townsend=5*0.7078391+ mean consumes alcohol*-0.1394517 ***[Table continues on the next page]***

Risk Model cont.	Equation
Model 4	where $\sum_{i=1}^{p} \boldsymbol{\beta_i X_i} =$ age*0.074564+gender*0.6663224+smoking status*0.4248647+ total cholesterol*0.0368726+HDL cholesterol*0.1674674+ not treated SBP*0.005948+treat SBP*0.0100666+diabetes*1.036887+ BMI*0.0287288+Townsend=2*0.1829155+Townsend=3*0.2746468+ Townsend=4*0.5175504+Townsend=5*0.6983048+ consumes alcohol*-0.154752+family history*0.2522272 where $\sum_{i=1}^{p} \boldsymbol{\beta_i \bar{X}_i} =$ mean age*0.074564+mean gender*0.6663224+ mean smoking status*0.4248647+mean total cholesterol*0.0368726+ mean HDL cholesterol*0.1674674+mean not treated SBP*0.005948+ mean treat SBP*0.0100666+mean diabetes*1.036887+mean BMI*0.0287288+ mean Townsend=2*0.1829155+mean Townsend=3*0.2746468+ mean Townsend=4*0.5175504+mean Townsend=5*0.6983048+ mean consumes alcohol*-0.154752+mean family history*0.2522272

5.1.8 Assessment of Model Performance

Performance of each model was checked individually in an internal validation, using the validation dataset; and comparisons were made between Model 1 (standard CVD risk factors only) and Model 2 (standard CVD risk factors with FH); and Model 3 (standard and additional CVD risk factor only) and Model 4 (standard and additional CVD risk factors with FH).

5.1.8.1 Description of Predicted Risk

The predicted risks from the derived models were plotted using Kernel density plots for patients with and without a CVD event during follow up. Figure 5.10 shows similar distributions of predicted risk for patients with and without an event during follow up, which is not affected with the addition to FH information to CVD risk prediction.

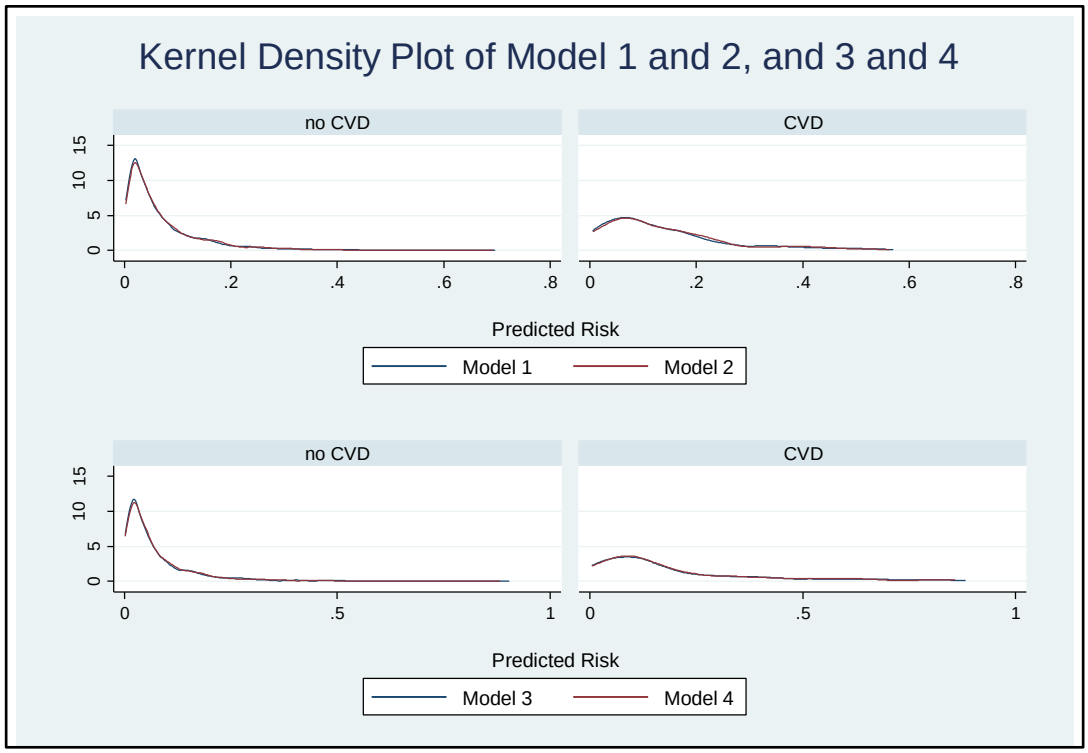


Figure 5.10 Kernel density plots for predicted risk for all models.

5.1.8.2 Calibration

The goodness on fit of each model was assessed by describing the calibration for each model to the observed data using calibration plots.

Visual representation of observed and predicted outcomes in presented in Figure 5.11; plotting the observed and predicted risk event rates for each tenth of predicted risk.

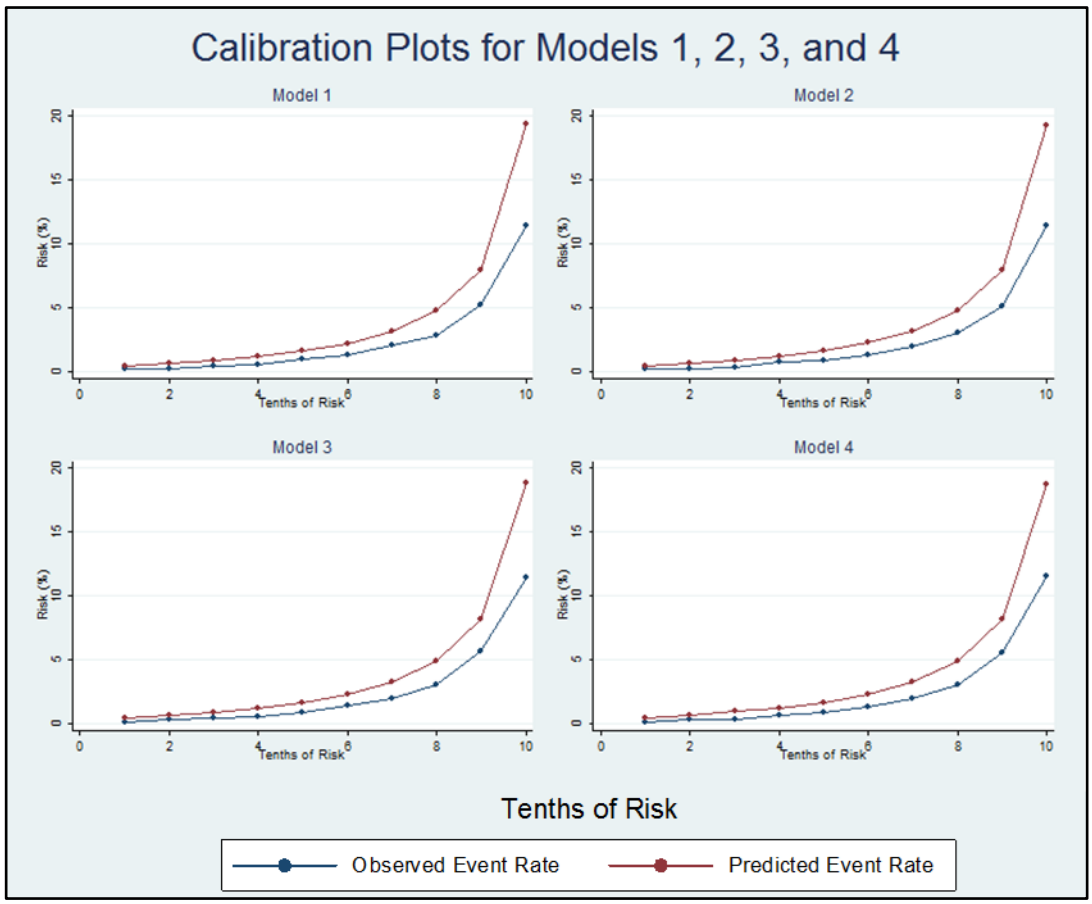


Figure 5.11 Predicted and observed risk by 10th of predicted risk for each risk prediction model in the derivation dataset.

Figure 5.11 shows good calibration of the models to the derivation dataset and show no improvement with the addition of FH information.

5.1.8.3 Discrimination

5.1.8.3.1 Sensitivity, Specificity, Positive, and Negative Predicted Values

The discriminatory accuracy of the models was first by investigating the sensitivity, specificity, positive, and negative predicted values. These have been presented in Table 5.21.

Table 5.21 Descriptive information on discriminatory performance for each risk model, using a 10% risk cut off.

N=99,821	Model 1	Model 2	Model 3	Model 4
Mean Predicted Risk (SD)	4.21% (6.2)	4.26% (6.28)	4.40% (6.99)	4.45% (7.07)
Range of Risk in Sample	0-74	0-78	0-100	0-100
10% Cut Off				
Sensitivity	43.32%	44.39%	43.22%	44.44%
Specificity	89.66%	89.47%	89.31%	89.09%
PPV	7.72%	7.76%	7.46%	7.52%
NPV	98.75%	98.77%	98.75%	98.77%
% correctly classified	88.76%	88.58%	88.41%	88.21%

Small increases were found between all measures of performance, presented in Table 5.21. Overall a higher mean predicted risk is found in models including FH information (Models 2 and 4) compared to those without (Models 1 and 3); and a higher specificity value was found than sensitivity.

The sensitivity marginally increases with the addition of FH and the specificity decreases. The biggest increase in sensitivity was found when adding FH information to standard and additional CVD risk factors; which increased from 43.22% to 44.44%.

5.1.8.3.2 Receiver Operating Characteristic (ROC) Curve

The ability of the models to discriminate between CVD events and non-events was described using ROC curves, and the area under the ROC curve using Harrell's C-Index (analogous to the AUC).

Comparisons were made between Model 1 and Model 2, adding FH information to standard CVD risk factors; and Model 3 and Model 4, adding FH information to standard and additional CVD risk factors. Table 5.22 shows Harrell's C-Index for each model.

Table 5.22 Harrell's C-Index for each CVD risk prediction model.

	Harrell's C-Index (95% CI)	Change in C-Index (% change)
Model 1	0.8152 (0.8063 to 0.8241)	0.0022 (0.27%)
Model 2	0.8174 (0.8086 to 0.8262)	
Model 3	0.8234 (0.8149 to 0.832)	0.0019 (0.23%)
Model 4	0.8253 (0.8168 to 0.8338)	

Small increments were found in Harrell's C-Index when adding FH information to standard and standard and additional CVD risk factors. Similar increments were found in both comparisons; the Harrell's C-Index for Model 1 increased by 0.3% and by 0.2% for Model 3, respectively.

Figure 5.12 shows the ROC curve for each model, compared to the 0.5 reference line.

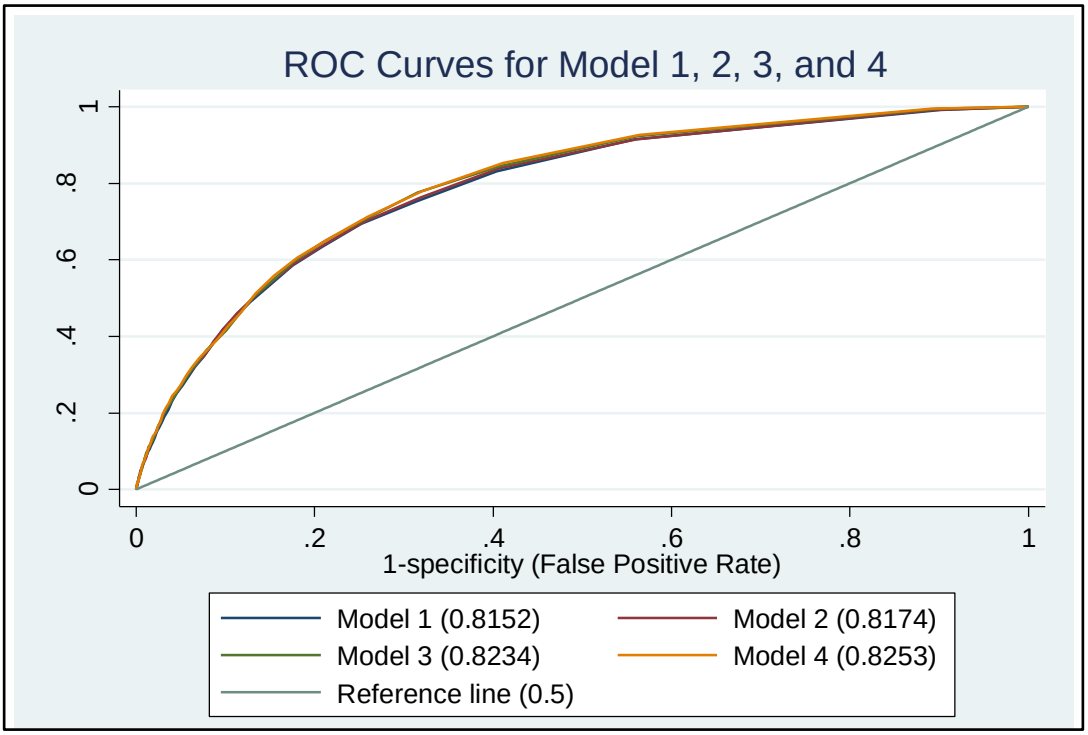


Figure 5.12 Receiver Operating Characteristic Curve for each derived CVD risk model.

All models perform better than chance, but also perform similarly to one another based on discriminatory accuracy.

5.1.8.4 Reclassification

Reclassification plots and table were used to assess the shift of patients between categories of risk with the addition of FH information. Model 1 was compared to Model 2, and Model 3 was compared to Model 4 to evaluate the value of FH in addition to standard and standard and additional CVD risk factors, respectively.

Figure 5.13 shows the reclassification plots for these comparisons.

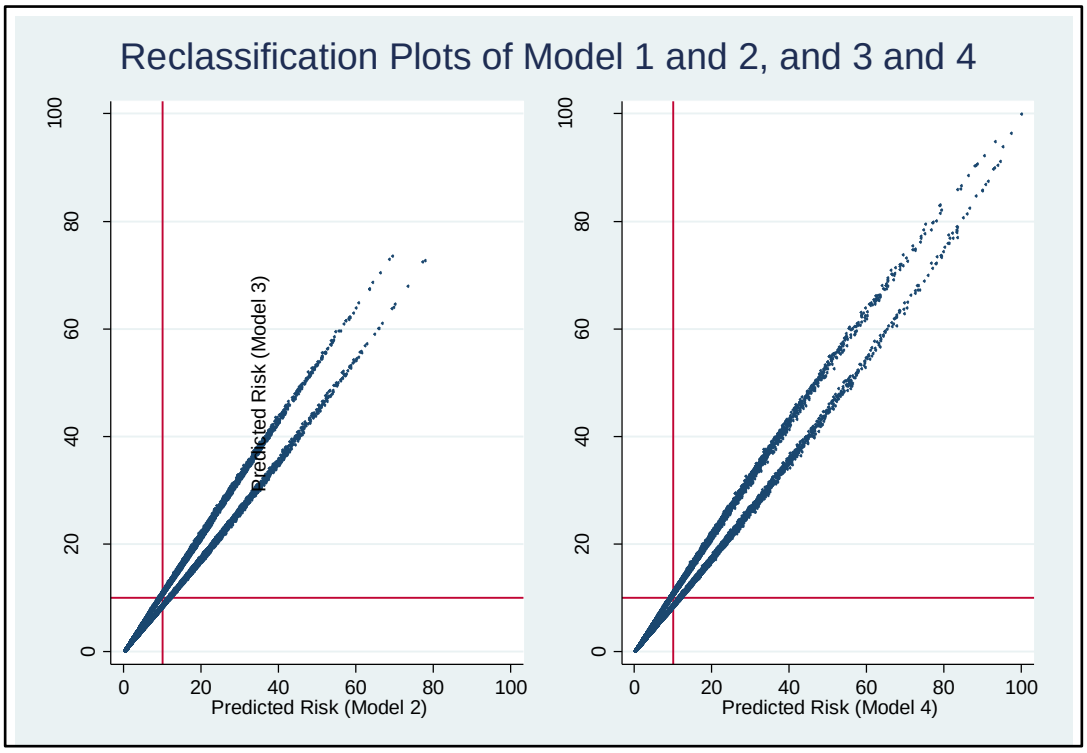


Figure 5.13 Reclassification plots comparing models with and without FH information. Reference line indicates the 10% risk threshold.

Figure 5.13 shows the agreement between risk models with and without FH information, with a small shift in patients to a different risk category, based on the 10% risk threshold.

5.1.8.4.1 Risk Reclassification – Model 1 vs Model 2

A reclassification table comparing standard CVD risk factors with and without FH information is presented in Table 5.23. It shows that a total of 1,683 (1.69%) of patient were reclassified in to a different risk category based on their FH status. Of these 949 (56.39%) patients were reclassified into a higher risk category, and 734 (43.61%) were reclassified into a lower risk category.

Table 5.23 Risk reclassification table comparing Model 1 (standard CVD risk factors only) to Model 2 (standard CVD risk factors with FH information).

			with FH			
			<10%	≥10%	Total	Total No. (%) Reclassified
without FH	<10%	No	87,910	949	88,859	949 (1.07%)
		%	98.93%	1.07%		
		No. Events	1,058	49		
		No. Non events	86,852	900		
		Observed Event Rate	1.20%	5.16%		
	≥10%	No	734	10,228	10,962	734 (6.7%)
		%	6.70%	93.30%		
		No. Events	28	818		
		No. Non events	706	9,410		
		Observed Event Rate	3.81%	8.00%		
	Total		88,644	11,177	99,821	1,683 (1.69%)

Table 5.24 shows the reclassification based on the additional of FH information, by CVD event outcome.

Table 5.24 Risk reclassification table comparing Model 1 with Model 2, by CVD outcome.

No CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	86,852 98.97%	900 1.03%	87,752
	≥10%	706 6.98%	9,410 93.02%	10,116
	Total	87,558	10,310	97,868
CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	1,058 95.57%	49 4.43%	1,107
	≥10%	28 3.31%	818 96.69%	846
	Total	1,086	867	1,953

Of the 1,683 patients reclassified, 49 (2.91%) were correctly reclassified into a higher risk group and 41.95% were reclassified correctly in to a lower risk category. However, over half of the reclassified patients were shifted to an incorrect risk category based on their FH information.

Very small changes were found in the proportion of patients recommended for treatment (10.98% to 11.20%) and those who were not (89.02% to 88.80%). Further, the proportions of missed diagnoses and unnecessarily treated within these groups remained the same; 1% and 92%, respectively.

Table 5.25 Summary of Net Reclassification Index (NRI) and Integrated Discrimination Index (IDI) for all comparisons between model with and without FH information.

Comparison	Event NRI	Non-Event NRI	Overall NRI	IDI
Model 1 vs Model 2	1.08%	-0.20%	0.0088	0.0019
Model 3 vs Model 4	1.23%	-0.22%	0.0101	0.0021

Table 5.25 shows an NRI and IDI of 0.0088 and 0.0019 were found when comparing Model 1 with Model 2; adding FH to standard CVD risk factors, respectively.

5.1.8.4.2 Risk Reclassification – Model 3 vs Model 4

A reclassification table comparing standard and additional CVD risk factors, with and without FH information is presented in Table 5.26. It shows that a total of 1,609 (1.61%) of patients were reclassified in to a different risk category based on their FH status. Of these 925 (57.49%) patients were reclassified into a higher risk category, and 684 (42.51%) were reclassified into a lower risk category.

Table 5.26 Risk reclassification table comparing Model 3 (standard and additional CVD risk factors only) to Model 4 (standard and additional CVD risk factors with FH information)

			with FH			
			<10%	≥10%	Total	Total No. (%) Reclassified
without FH	<10%	No	87,587	925	88,512	925 (1.05%)
		%	98.95%	1.05%		
		No. Events	1,052	57		
		No. Non events	86,535	868		
		Observed Event Rate	1.20%	6.16%		
	≥10%	No	684	10,625	11,309	684 (6.05%)
		%	6.05%	93.95%		
		No. Events	33	811		
		No. Non events	651	9,814		
		Observed Event Rate	4.82%	7.63%		
	Total		88,271	11,550	99,821	1609 (1.61%)

Table 5.27 shows the reclassification based on the additional of FH information, by CVD event outcome.

Table 5.27 Risk reclassification table comparing Model 3 with Model 4, by CVD outcome.

No CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	86,535 99.01%	868 0.99%	87,403
	≥10%	651 6.22%	9,814 93.78%	10,465
	Total	87,186	10,682	97,868
CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	1,052 94.86%	57 5.14%	1,109
	≥10%	33 3.91%	811 96.09%	844
	Total	1,085	868	1,953

Of the 1,609 patients reclassified, 57 (3.5%) were correctly reclassified into a higher risk group and 55% were reclassified correctly in to a lower risk category. Less than half of the reclassified patients were shifted to an incorrect risk category based on their FH information.

Very small changes were found in the proportion of patients recommended for treatment (11.3% to 11.6%) and those who were not (88.7% to 88.4%). Further, the proportions of missed diagnoses and unnecessarily treated within these groups remained the same similar; 1% and 92%, respectively.

Table 5.25 shows an NRI and IDI of 0.0101 and 0.0021 were found when comparing Model 3 with Model 4; adding FH to standard and additional CVD risk factors, respectively.

5.1.8.5 Decision Curve Analysis

A Decision Curve Analysis (DCA) allows new risk factors to be assessed by comparing risk models and identifying which model leads to the highest net benefit (NB) at a given risk threshold; the biggest increase in the proportion of true positives without a change in the proportion of false positives. A pre-specified risk threshold of 10% was used.

No change was found between any of the risk models, and a net benefit of -0.003 was found for all models at the 10% risk threshold. No change was found on addition of FH information.

5.1.9 Sensitivity Analysis

Sensitivity analyses were performed using complete cases of the data, and also using a literature based survival estimate for CVD (74).

5.1.9.1 Complete Case Analysis

The complete case analysis used 8,775 patients in the derivation dataset and 4,621 patients in the validation dataset. Four risk models were created which included the same risk factors; however alcohol consumption was not included in Model 3 and 4, which included standard and additional CVD risk factors without and with FH information, respectively.

Family history showed a statistically significant association with CVD in the univariate analysis; patients with a positive FH status had 34% higher hazard of CVD compared to patients with a negative FH (HR=1.34 (95% CI: 1.26 to 1.43)). In addition to the standard and standard and additional CVD risk factor a lower hazard of CVD was found amongst patients with a positive FH.

Good calibration was found in all risk models, and improvement was shown in models with FH information (Figure 5.14).

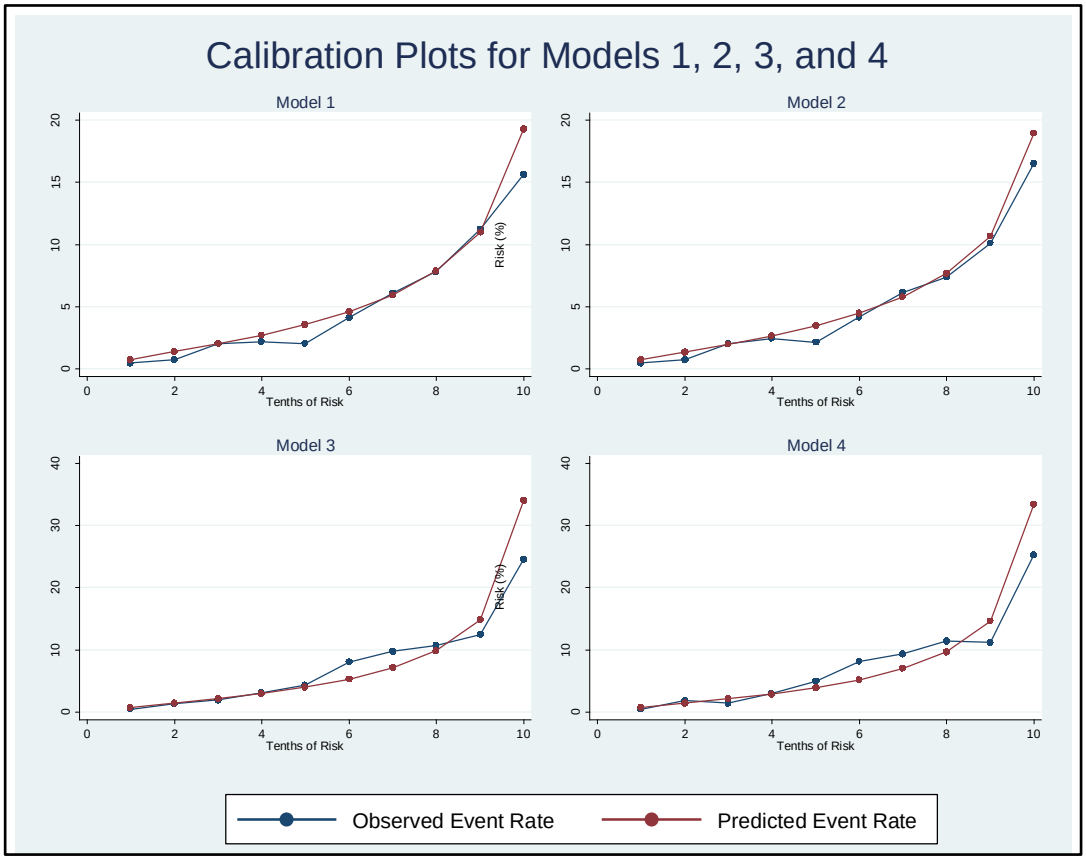


Figure 5.14 Calibration Plots of the 4 CVD risk prediction models derived using complete cases of the data.

However, Harrell's C-Index reduced in models in FH information (Figure 5.15); a 0.5% decrease was found when comparing standard risk factors with (Model 2) and without FH (Model 1), and also when comparing standard and additional risk factors with (Model 4) and without FH (Model 3). Model 3, which included standard and additional CVD risk factors, the highest discriminatory accuracy.

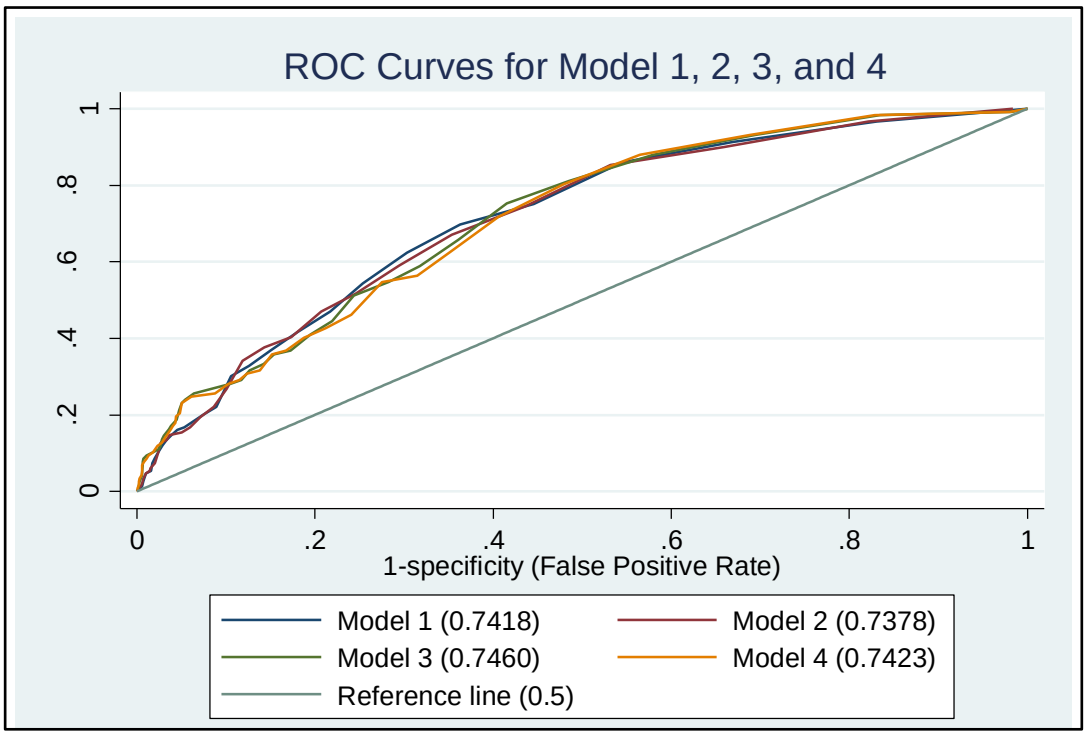


Figure 5.15 Receiver Operating Characteristic (ROC) Curve, and Harrell's C-Index values.

Reclassification tables showed fewer patients were recommended for treatment in models using FH information, however the proportion of missed diagnoses and unnecessary treatments remained respectively similar.

Full results from the analysis can be found in Appendix I.

5.1.9.2 Literature Based Survival

A literature based survival estimate was also used to create the 4 risk prediction models using the beta estimates from the main analysis. An estimate of 0.9551 was used from Collins et al (74).

Poor calibration was found in all risk models, which remained similar in all models (Figure 5.16).

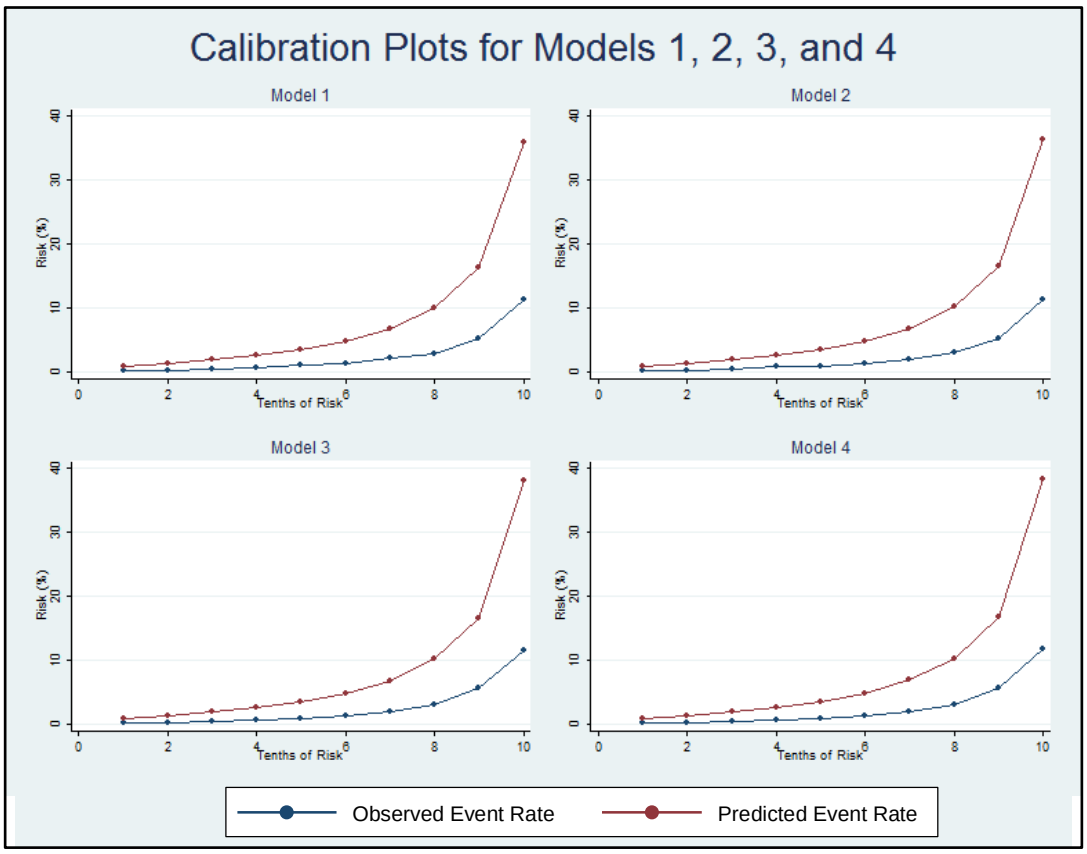


Figure 5.16 Calibration Plots of the 4 CVD risk prediction models derived using imputed data and a literature based survival estimate (0.9551).

Harrell's C-Index increased by small amounts in models with FH information (Figure 5.17); a 0.3% change was found when comparing standard risk factors with (Model 2) and without FH (Model 1); and a 0.2% change when comparing standard and additional risk factors with (Model 4) and without FH (Model 3). Model 4 had the highest discriminatory accuracy (0.8253).

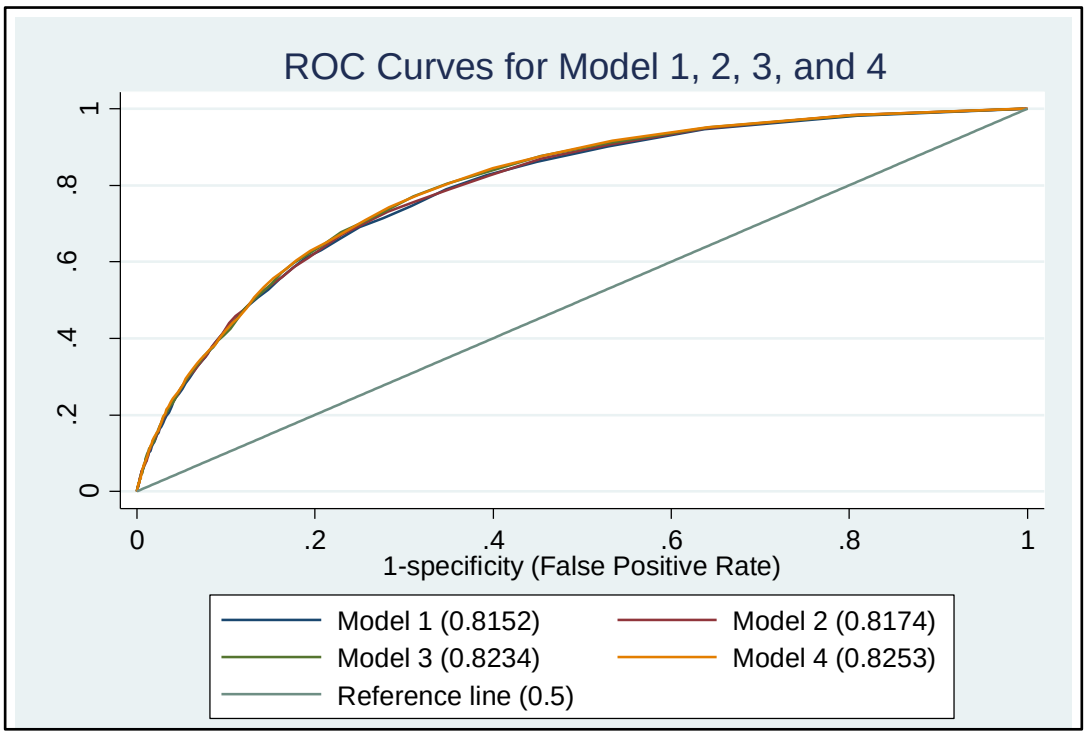


Figure 5.17 Receiver Operation Characteristic (ROC) Curve, and Harrell's C-Index values for models derived using the imputed data and literature based survival estimate (0.9551).

Reclassification tables showed a marginally higher proportion of patients were recommended for treatment in models using FH information, however the proportion of missed diagnoses and unnecessary treatments remained respectively similar.

Full results from the analysis can be found in Appendix J.

5.1.10 Section Summary

This section presented the results from the data extraction from CPRD; described the baseline characteristics of the cohort and investigated trends in the incidence of CVD. Cox regression was used to model the data and create four CVD risk prediction models; including standard CVD risk factors only (Model 1); standard CVD risk factors with FH (Model 2); standard and additional CVD risk factors only (Model 3); and standard and additional CVD risk factors with FH (Model 4).

The derived models were then validated using current methodology and performance measures describing calibration, discrimination, and reclassification. The next section discusses these findings.

5.2 Discussion

The added value of family history (FH) of cardiovascular disease (CVD) is not explicitly established in primary care; yet it is considered an additional risk factor for CVD and is included in CVD risk assessment. It has been accounted for by either consensus driven risk inflation factors (18) or by including it in a CVD risk prediction model (11). Further, current measures used to describe the improvement of FH in addition to established CVD risk prediction models have shown to be poor at quantifying its added value (23), and does not evaluate the clinical implications it may have; a new methodology is needed.

The aim of this study was to identify and improve quantify the added value FH information contributes to CVD risk prediction, in addition to the standard and additional CVD risk factors.

This was achieved by recalibrating a Framingham based risk prediction model (10) using routinely collected primary care data. Models were created including the standard CVD risk factors, with and without FH information; and including standard and additional CVD risk factors, with and without FH information. Models were then validated using current and refined methodology. The previous section presented the results from the current methodology; in this section the results from the current methods are discussed.

5.2.1 Summary of findings

A prospective cohort was extracted from CPRD, a routinely collected primary care database, including patients if they had FH of CVD recorded (either positive or negative) in their medical records. This led to 2% of the acceptable patients contributing to CPRD being included in the study, compared to 19% extracted from The Health Improvement Network (THIN) in a similar extraction of patients (93). Further, 40% of patients extracted from CPRD had a positive FH of CVD; and of which 18% had additional details

for age of onset or relative specified. This was lower compared to the THIN extraction which searched free text data as well as medical records for FH information; thus allowing detection of more FH information.

The extracted cohort was compared to the Health Survey for England dataset 2011 (103), and showed to be a slightly unhealthier sample; patients had a higher SBP, higher total cholesterol with lower HDL cholesterol, and higher HbA1c value. However, this was not translated through to the CVD event rate during follow up, which was found to be lower compared to other studies (74). It was found that although the sample had more positive and modifiable CVD risk factors, it also had a lower mean age, consisting of fewer males and smokers; which were found to be stronger CVD risk factors.

Family history of CVD showed a strong relationship with CVD throughout the analysis; and CVD incidence rates were higher in patients with a positive FH recorded. Univariate Cox regression reported a 34% higher hazard ratio of CVD in patients with a positive FH recorded; this stayed consistent when adding FH to standard CVD risk factors and standard and additional CVD risk factors.

Despite the strong association shown in the univariate and multivariate regression, small improvements were found in the global fit between models with and without FH. Further, when validating the models, no change was found between the distributions of risk amongst the patients in model including FH information, from those without. Good calibration was found in all models; but no significant improvement was found when adding FH information. Small improvements were again found in discriminatory accuracy; 0.3% relative increase was found in Harrell's C-Index when adding FH information to standard CVD risk factors.

The increments in the discriminatory accuracy in particular, were lower compared to other studies evaluating the added value of FH information to CVD risk prediction (8, 9). However, direct comparison was difficult; other studies, evaluating the added value of FH to CVD risk prediction, did not investigate this independently from other new risk factors at the time. The ASSIGN risk score assessed the added value of family history and social deprivation, comparing it to the Framingham risk score; they reported a 1.4% and 3.2% relative increase in the AUC of models for women and men, respectively (8). The Reynolds risk score assessed the added value of FH (parental history of myocardial infarction) with high sensitivity C-reactive protein (9) and reported a 0.4-0.5% relative increase in the AUC of the new model.

Results of the study were consistent throughout the analysis, and the added value of FH information did not vary when adding it to standard or standard and additional CVD risk factors. Its added value was shown to be independent from other CVD risk factors. A small proportion of patients shifted to a higher or lower risk category when adding FH information to CVD risk prediction. Reclassification tables showed that 1.7% and 1.6% of patients were reclassified into a higher or lower risk category when including FH information to standard and standard with additional CVD risk factors, respectively. Small increments were found in the net reclassification index (NRI) for both comparisons made, based on the 10% risk cut off.

5.2.2 Strengths and Limitations

The added value of FH information has not been evaluated independently from other novel CVD risk factors. The ASSIGN risk scores assess it in addition to social deprivations, and the Reynold risk score assesses it in addition to high sensitivity C-reactive protein. Thus, this is the first study to evaluate the independent added value of FH information to CVD risk prediction. Further, the added value of FH information is

assessed in addition to standard with additional CVD risk factors, as well as standard CVD risk factors alone.

Detailed FH information was also derived from the CPRD database to assess its added value in CVD risk prediction. However, limited additional FH information was found regarding the age of onset and relative affected. The added value of detailed FH could therefore not be assessed, but was presented in the descriptive analysis.

The extracted dataset only included patients with a FH of CVD recorded; positive or negative. As shown, when assessing the generalisability of the data, this produced a biased sample, which had more CVD risk factors compared to the Health Survey for England. However, this did not translate to the CVD event rate found in the data which was comparably lower than other CVD risk model validation studies due to the lower mean age, fewer males and smokers also found.

Assuming patients without a recorded FH of CVD to have a negative family history was felt to be unreasonable at this time. Reports show an increasing trend in the treatment of CVD during 1998-2011 (71), and therefore an increased prevalence of CVD. Thus, it can be expected that this would translate to an increased recording of FH of CVD during this time. However, a decreasing availability of FH of CHD was found during the respective period (Appendix K). This shows there is a significant proportion of patients would have an unrecorded positive FH of CVD.

A consort diagram was presented to describe the selection of patients from CPRD. It describes the inclusion and exclusion criteria for the study; however it fails to provide a breakdown of the patients initially extracted from CPRD, and hence the number of patients excluded with a prior history of CVD and excluded with a missing FH.

Information describing this was unavailable as data was externally supplied by CPRD based on the specified inclusion and exclusion Read codes.

In hindsight, it would have been useful to also extract patients with missing FH information and assess the CVD event rate amongst these patients. The study using the THIN database to assess the availability of FH information showed FH of CHD in particular it was more likely to be missing with increasing deprivation.

A robust study was performed, which sourced CVD data from linked Hospital Episode Statistics (HES) data and Office for National Statistics (ONS) based on Read and ICD codes for CVD. Identified records were combined and the earliest CVD event was used in the analysis; leading to more accurate estimates.

The study also used established risk factors for CVD to predict its risk. To increase generalisability of the results, risk factors were included in the analysis in the same format and units used in other models. However, it was found that hsCRP was not recorded in CPRD; but CRP was recorded and included in the analysis. This may limit analysis and generalisability of these results which may not be able to discriminate as well at lower levels of CRP. This however does not affect the adjusted analysis as CRP was not included due to insufficient data.

Further, routinely collected data from UK primary care was used to derive the CVD risk models. QRISK2 has shown the advantages of using primary care data for risk prediction, especially by it improved calibration on a UK population (74) compared to the Framingham and ASSIGN risk models. However, the simplified Framingham risk score was used when developing the models, which provided a better foundation for evaluating the added value of FH information to standard CVD risk factors, and standard and additional CVD risk factors.

A limitation of using routinely collected data from primary care for this study was missing data. A large amount of missing data was found for CVD risk factors, including the standard CVD risk factors for cholesterol. A complete case analysis would have limited to analysis for only 4% of the extracted data; less than 1% of the entire CPRD database. Further, complete case analyses have shown to provide biased results when data is not missing completely at random (116-118), which is a rare occurrence.

Data was assumed to be missing at random; missing data was explored and found to be related to the observed data for other risk factors. This allowed for the more favoured and commonly practised approach, multiple imputation (119-121), to be used. However, missing data was also found to be associated to the unobserved data. To increase the plausibility of the missing at random assumption predictors for the incomplete risk factor and predictors for the data to be missing were included. All variables associated to the missing data, identified through the correlation and regression analyses were included in the multiple imputation models, as well as the CVD outcome and FH.

The generalisability of extracted cohort was compared to the Health Survey for England dataset and also the THIN dataset for the CVD event rate. However, the THIN database also uses the VISION software and includes practices registered with CPRD. Therefore there may an overlap of patients in both extracted samples. This makes the CVD event rates less comparable as patients may be double counted.

Another limitation for the study was performing an internal validation, which uses the data from the same population to develop and validate the CVD risk models. This may lead to over optimistic results of the performance of the risk models. To account for this limitation, a commonly practised method which randomly assigns patients to the

derivation and validation datasets was used; further a similar 1:3 ratio was also used to split the data (11, 31, 83).

Finally, current measures were used to evaluate the added value of FH information. However, these have been criticised for being insensitive to improvement of the risk model when adding new risk factors (23), and not clinically relevant. This chapter has also found small changes in performance measures when adding FH information. This limitation has been accounted for by using refined methodology to make current measures more clinically applicable; and is presented in the next chapter.

5.3 Chapter Summary

This chapter presented and discussed the results of the current methodology used to evaluate the added value of FH information to CVD risk prediction. This included a description of the extraction, the characteristics and generalisability of the sample, and the development of the risk models.

The next chapter will present and discuss the results of the refined methodology used to evaluate the added value of FH information to CVD risk prediction.

Chapter 6 - Family History and Cardiovascular Disease Risk: Results and Discussion (Refined Methodology)

The previous chapter presented the extracted cohort from CPRD; it described the baseline characteristics; and also looked at the trends in incidence rates of CVD. Missing data was imputed for the analysis and 4 CVD risk prediction models were created using the derivation dataset: 1) standard CVD risk factors only; 2) standard CVD risk factors with FH information; 3) standard and additional risk factors only; and 4) standard and additional CVD risk factors with FH information. These models were then validated using the validation dataset. Current methodology was used to assess each model and evaluate the added value of FH information; these included measures of discrimination and reclassification.

In this chapter, evaluation of added value of FH to CVD prediction is continued; using the refined methodology presented in Chapter 3; a cost effective analysis using a decision tree framework is performed to assess the added value of FH information to CVD risk prediction.

6.1 Results

6.1.1 Cost-Effectiveness Analysis

A decision tree was used for the cost-effectiveness analysis, which compared risk models with FH information to models without. Model 1, with standard CVD risk factors only, was compared to Model 2, which added FH. Model 3, with standard and additional risk factors was compared to Model 4, which added FH. The decision tree used for this

approach is presented in Chapter 3, Section 3.3.3, Figure 3.3, and was used for both the deterministic and probabilistic cost effectiveness analyses.

The Youden Index and Harrell's C-Index, measures of discriminatory accuracy, were used as measures for the performance of each model, and the change between these measures were used to describe the added value FH information.

6.1.1.1 Deterministic Cost-Effectiveness Analysis

A deterministic cost-effectiveness analysis was performed for both comparisons described above. This used the sensitivity and specificity for each model using a 10% risk cut off. A summary of values used in the decision tree to compare Model 1 to Model 2, and Model 3 to Model 4 are presented in Table 6.1. Values assume the base case for the analysis; described in Chapter 3. In their respective comparisons, Model 1 and Model 3 are termed as the original models (without FH), and Model 2 and Model 4 are the modified models (with FH).

Table 6.1 Summary of values inputted into the decision tree to compare Model 1 to Model 2, and Model 3 to Model 4.

Inputs	Model 1 vs. Model 2 Estimates	Model 3 vs. Model 4 Estimates
P(D+)	0.0196 (0.0004)	
P(D-)	0.9804 (0.0004)	
Original Model (without FH)	**Model 1**	**Model 3**
P(T+)	0.1098 (0.0010)	0.1133 (0.0010)
P(T-)	0.8902 (0.0010)	0.8867 (0.0010)
Sensitivity: P(T+|D+)	0.4332 (0.0112)	0.4322 (0.0112)
P(T+|D-)	0.1034 (0.0010)	0.1069 (0.0010)
Specificity: P(T-|D-)	0.8966 (0.0010)	0.8931 (0.0010)
P(T-|D+)	0.5668 (0.0112)	0.5678 (0.0112)
Modified Model (with FH)	**Model 2**	**Model 4**
P(T+)	0.1120 (0.0010)	0.1157 (0.0010)
P(T-)	0.8880 (0.0010)	0.8843 (0.0010)
Sensitivity: P(T+|D+)	0.4439 (0.0112)	0.4444 (0.0112)
P(T+|D-)	0.1053 (0.0010)	0.1091 (0.0010)
Specificity: P(T-|D-)	0.8947 (0.0010)	0.8909 (0.0010)
P(T-|D+)	0.5561 (0.0112)	0.5556 (0.0112)
Cost of Original Model (without FH)	£47.63	
Cost of Modified Model (with FH)	£70.45	
Cost of Low Risk - Untreated	£916.16	
Cost of Low Risk - Treated	£788.40	
Cost of High Risk - Untreated	£1,516.37	
Cost of High Risk - Treated	£1,342.73	

6.1.1.1.1 Model 1 vs. Model 2

Figure 6.1 shows an example of the decision tree comparing Model 1 to Model 2.

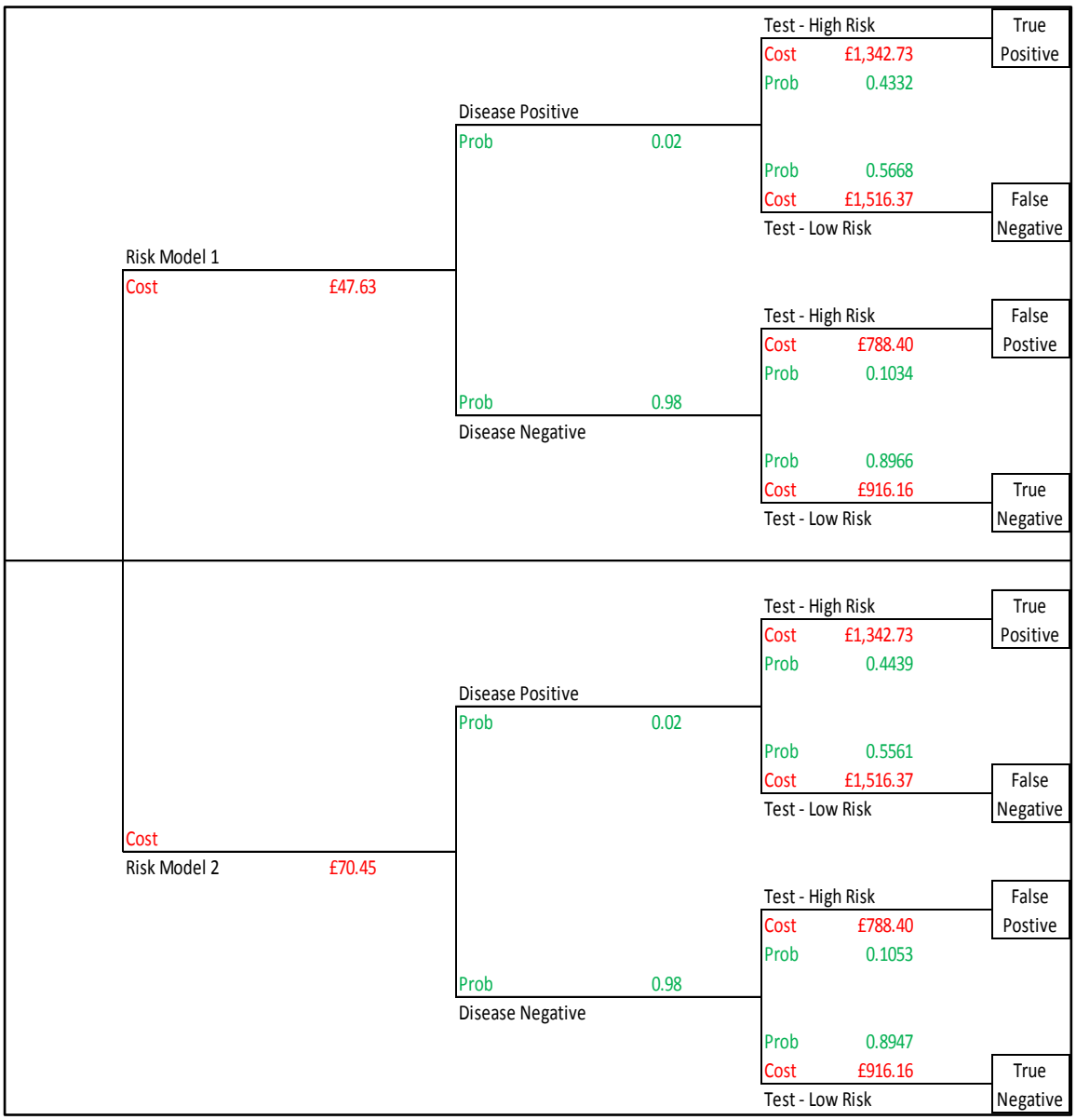


Figure 6.1 Decision tree to compare Model 1 (standard risk factors only) to Model 2 (standard risk factors with FH).

Comparing Model 1 and Model 2, the expected cost of calculating CVD risk based on standard risk factors only is £961.12 per person. Adding FH information to this increases the expected cost per person to £983.65. Table 6.2 provides a summary of the output produced by the decision tree.

Table 6.2 Output from each decision tree comparing models with FH information to models without.

		Model 1	Model 2	Model 3	Model 4
True Positive $P(D+ \cap T_n+)$	Probability	0.0085	0.0087	0.0085	0.0087
	Proportion, per 1000	8	9	8	9
	Cost	£1,390.36	£1,413.18	£1,390.36	£1,413.18
False Positive $P(D- \cap T_n+)$	Probability	0.1013	0.1033	0.1048	0.1070
	Proportion, per 1000	101	103	105	107
	Cost	£836.03	£858.85	£836.03	£858.85
False Negative $P(D+ \cap T_n-)$	Probability	0.0111	0.0109	0.0111	0.0109
	Proportion, per 1000	11	10	11	11
	Cost	£1,564.00	£1,586.82	£1,564.00	£1,586.82
True Negative $P(D- \cap T_n-)$	Probability	0.8791	0.8772	0.8756	0.8734
	Proportion, per 1000	879	877	876	873
	Cost	£963.79	£986.61	£963.79	£986.61
	Expected Cost of Test	£961.12	£983.65	£960.68	£983.18
	Expected Youden Index	0.3298	0.3386	0.3252	0.3353

Adding family history to standard CVD risk factors does not change the proportion of patients who would be recommended treatment based on their risk and 10% treatment threshold. Using Model 1 or 2, 11% of patients would be considered high risk on CVD and would recommended preventative treatment. Of these, a 1% absolute difference was found between patients would be treated unnecessarily; 93% patients would be treated unnecessarily using Model 1 compared to 92% with the addition of FH using Model 2.

The output from the decision tree also shows that the cost of a false positive is lower (£836.03) than a false negative (£1,564.00); it would therefore the better to over treat patients than miss any diagnoses from a cost-effectiveness point of view.

Conversely, adding family history to standard CVD risk factors marginally decreases the number of patients who would be considered low risk based on their CVD risk factors.

Similarly, adding family history to standard CVD risk factors does not change the proportion of patients who would be considered low risk. Using Model 1 or 2, 89% of patients would be considered low risk and would not be recommended preventative treatment. Of these patients, 1% would be classed as a missed diagnosis for both models, with and without FH information.

6.1.1.1.1.1 Incremental Cost-Effectiveness Ratio (ICER)

Based on Table 6.2, the incremental cost of using Model 2 is £22.54 for a 3% increase was found in the Youden Index. Harrell's C-Index for each model has been calculated in Chapter 5 Section 5.1.8.3.2; a C-Index of 0.8152 (SE: 0.0045) and 0.8174 (SE: 0.0045) was found for Model 1 and Model 2, respectively, giving a 0.27% increase. A summary of the Incremental costs and effect are presented in Table 6.3.

Table 6.3 Results from the deterministic cost-effectiveness analysis evaluating the added value of FH information to CVD risk prediction.

Comparison	Change in Cost	% Change in effect (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Model 1 vs. Model 2	£22.54	2.66%	£846.36	0.27%	£8,283.42
Model 2 vs. Model 3	£22.51	3.10%	£726.59	0.23%	£9,812.99

An ICER of £846.36 per percentage increase in Youden Index was found; and an ICER of £8,283.42 per percentage increase in Harrell's C-Index. These have been plotted on a cost-effectiveness plane (Figure 6.2). Both measures lie in the first quadrant of the graph, indicating non dominance of the Model 2 (with FH); for both discriminatory accuracy

measures, Model 2 is more effective and more costly than Model 1 (adding FH produces a more effective model, at an increased cost).

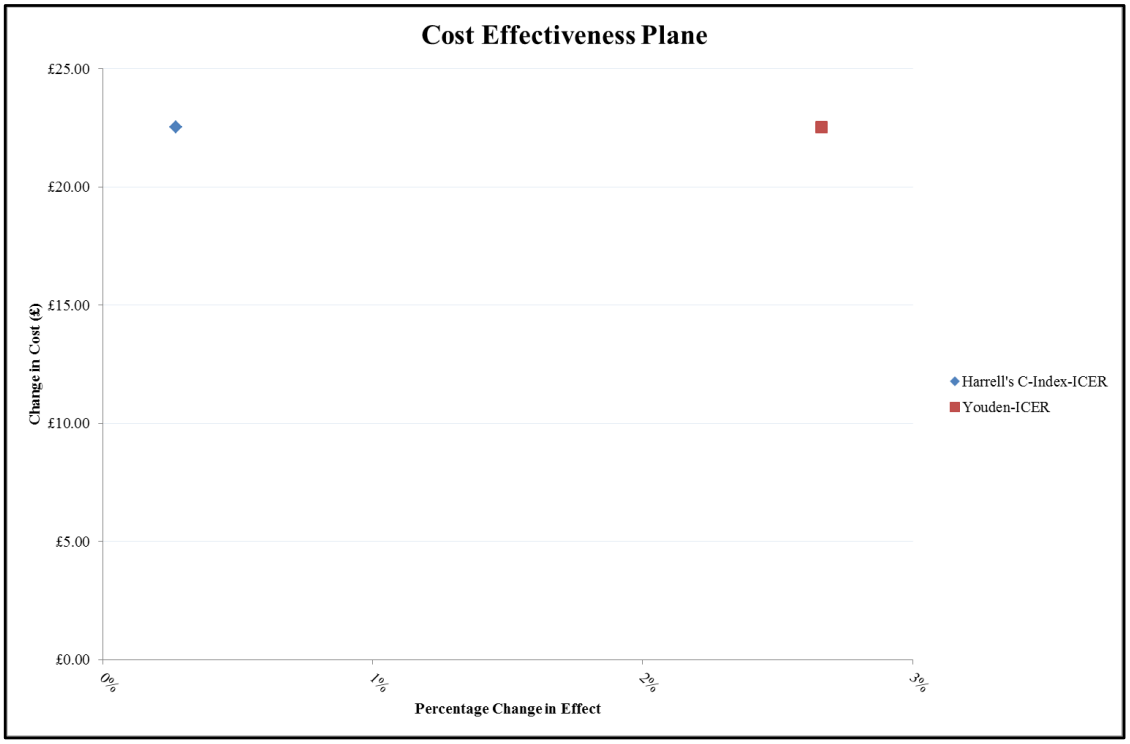


Figure 6.2 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 1 and Model 2.

6.1.1.1.1.2 Incremental Net Monetary Benefit (INB)

Using the ICER values calculated above, the INBs were derived for each outcome measure which were used to show the point at which adding FH is cost effective; thus satisfying the condition $\Delta C/\Delta E<\lambda$. This introduced the willingness to pay (λ) which was varied to give the INB for each willingness to pay value and outcome measure (Youden Index and Harrell's C-Index).

Figure 6.3 shows the INB for both measures of effect, using a £0 to £10,000 range for the willingness to pay; the INB crosses 0 for each measure of effect at the ICER value given above. The INB also increases as the willingness to pay increases for both measures of

effect. Thus, adding FH to standard CVD risk factors becomes more cost effective as the willingness to pay increases.

Further, Figure 6.3 shows a higher willingness to pay is required for improvement in Harrell's C-Index (at least £8,000) compared to the Youden Index (at least £1,000); anything below these amounts would not show FH to be cost-effective.

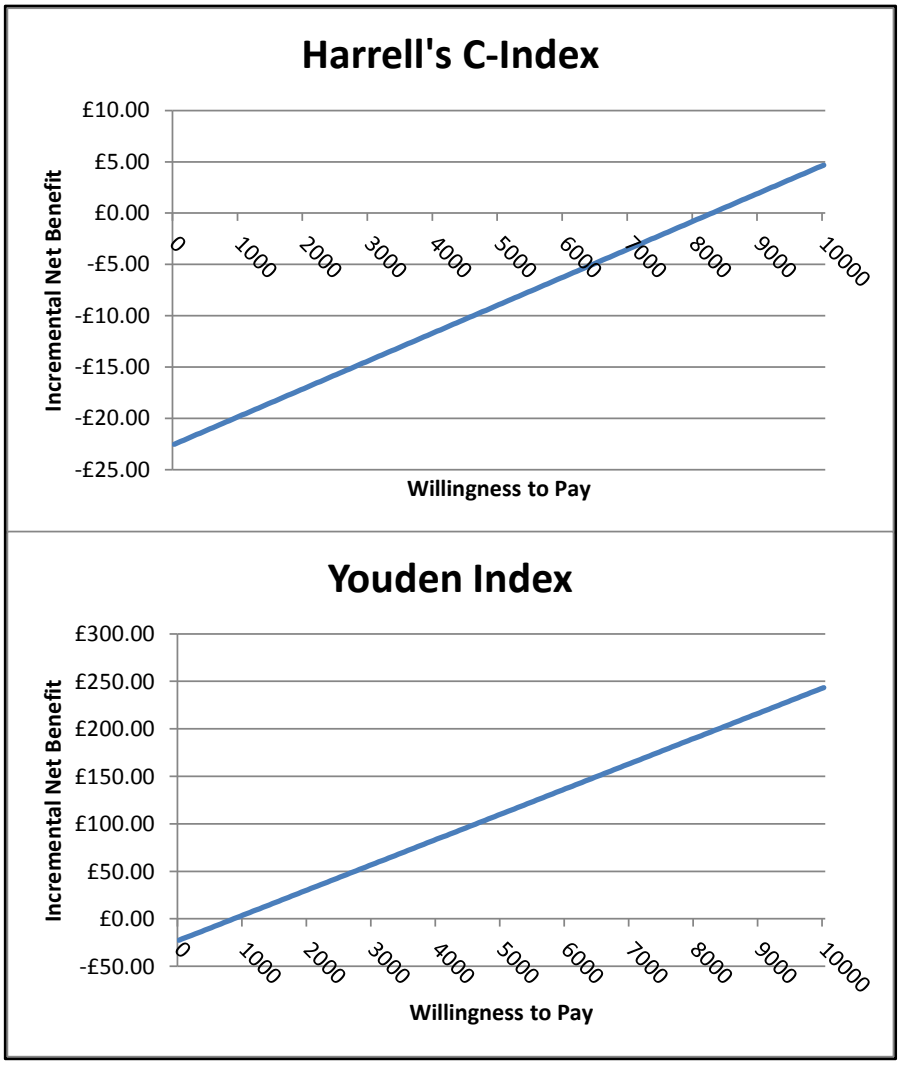


Figure 6.3 Incremental net benefit, for Harrell's C-Index and Youden Index, comparing Model 1 and Model 2.

6.1.1.1.2 Model 3 vs. Model 4

Figure 6.4 shows an example of the decision tree comparing Model 3 to Model 4.

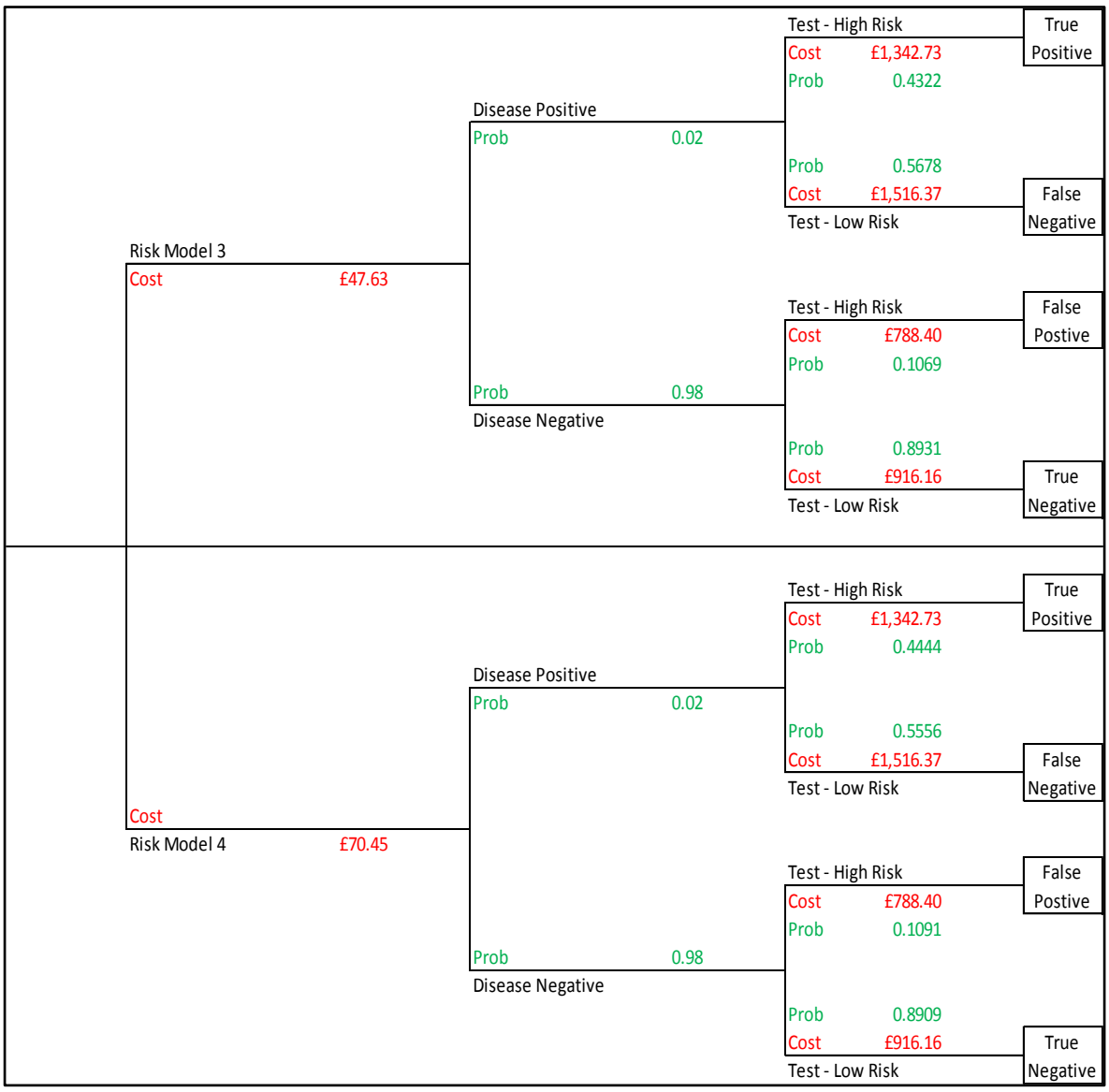


Figure 6.4 Decision tree to compare Model 3 (standard and additional risk factors only) to Model 4 (standard and additional risk factors with FH).

Comparing Model 3 and Model 4, the expected cost of calculating CVD risk based on standard and additional risk factors only is £960.68 per person. Adding FH information to this increases the expected cost per person to £983.18 (Table 6.2).

Results for the proportion of patients recommended for treatment or considered low risk, correctly and incorrectly, were analogous to the Model 1 and Model 2 comparison.

6.1.1.1.2.1 Incremental Cost-Effectiveness Ratio (ICER)

Based on Table 6.2, the incremental cost of using Model 4 is £22.50 and a 3% increase was found in the Youden Index. Harrell's C-Index for each model has been calculated in Chapter 5, Section 5.1.8.3.2; a C-Index of 0.8234 (SE:0.0044) and 0.8253 (SE: 0.0043) was found for Model 3 and Model 4, respectively, giving a 0.23% increase.

An ICER of £726.59 per percentage increase in Youden Index was found; and an ICER of £9,812.99 per percentage increase in Harrell's C-Index. These have been plotted on a cost-effectiveness plane (Figure 6.5). Both measures lie in the first quadrant of the graph, indicating non dominance of the Model 4 (with FH); for both discriminatory accuracy measures, Model 4 is more effective and more costly than Model 3.

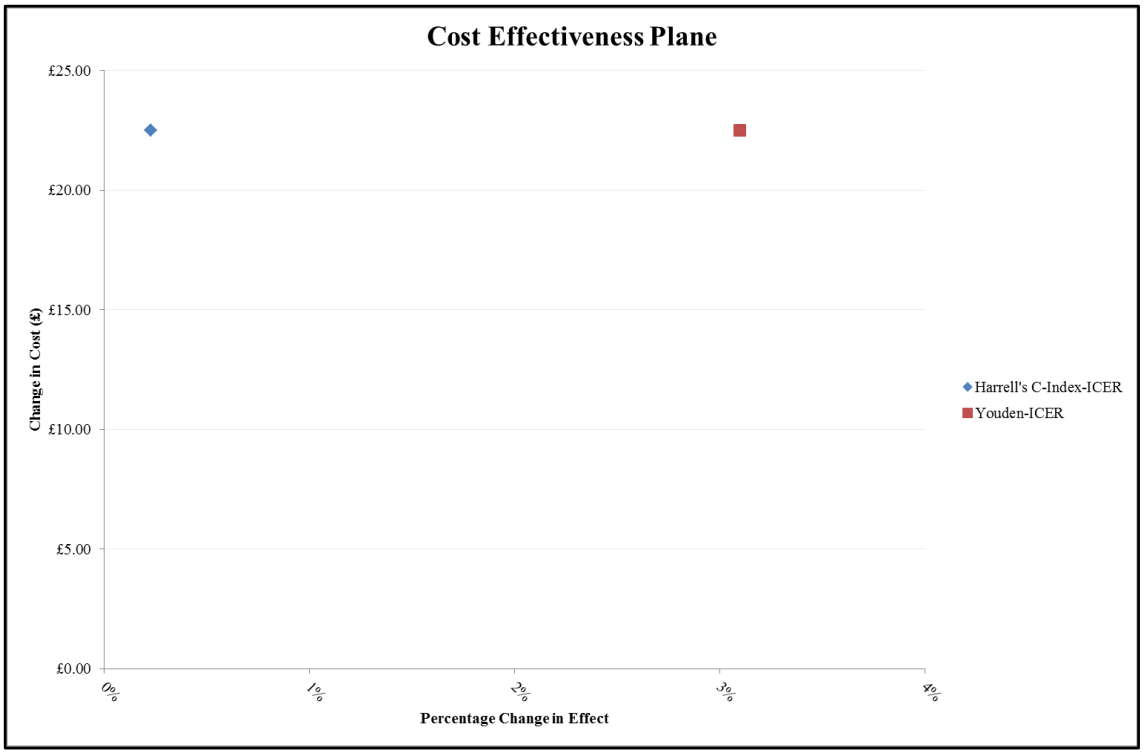


Figure 6.5 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 3 and Model 4.

6.1.1.1.2.2 Incremental Net Benefit (INB)

The INB was calculated using the ICER calculated above, and varying the willingness to pay for the improvement between risk models using the Youden Index and Harrell's C-Index.

Figure 6.6 shows the INB for both measures of effect, using a £0 to £10,000 range for the willingness to pay; the INB increases with the willingness to pay increases for both measures of effect. A higher willingness to pay is required for improvement in Harrell's C-Index.

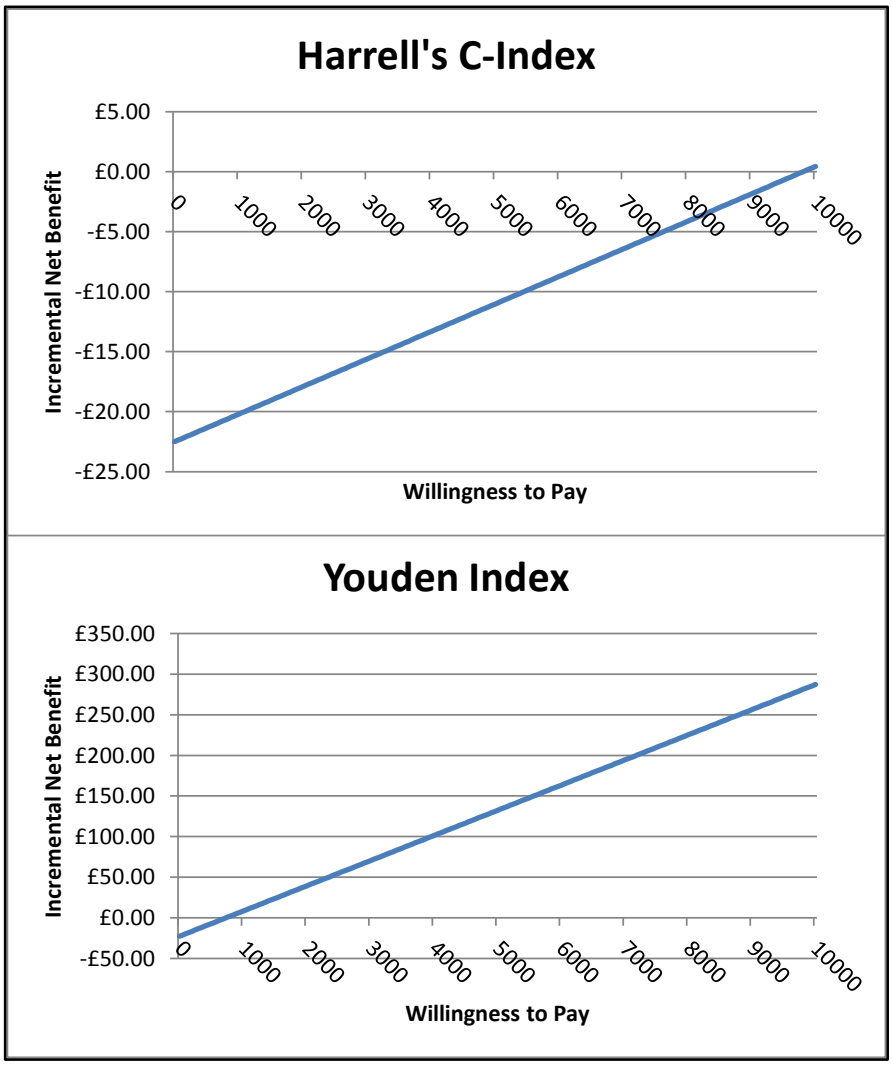


Figure 6.6 Incremental net benefit, for Harrell's C-Index and Youden Index, comparing Model 3 and Model 4.

6.1.1.2 Probabilistic Sensitivity Analysis

As well as the deterministic cost-effectiveness analysis, a probabilistic analysis was also performed to measure the uncertainty of each ICER. This was performed using the Youden Index and Harrell's C-index as measures of effect; and were compared for all comparisons described above.

Using the input values for the decision tree presented in Table 6.1, the estimates and their standard errors (if available) were used for the probability distribution. A beta distribution was used for proportion estimates and the gamma distribution was used for costs. The alpha and beta values for costs and their gamma distributions are presented in Chapter 4, Section 4.2.5.9.1.6. Table 6.4 gives the alpha and beta values for probabilities used in the decision tree and their beta distributions.

Table 6.4 Deterministic probabilities used in the decision tree, for each risk model (with and without FH); with standard error, alpha, and beta for estimation of probabilistic probabilities derived from a Beta distribution.

Probability	Value	SE	Alpha	Beta
P(D+)	0.0196	0.0004	1952.69	97852.52
P(D-)	0.9804	0.0004	97852.52	1952.69
Model 1				
P(T+)	0.1098	0.0010	10962.05	88859.37
P(T-)	0.8902	0.0010	88859.37	10962.05
Sensitivity: P(T+|D+)	0.4332	0.0112	845.13	1105.86
P(T+|D-)	0.1034	0.0010	10116.56	87756.91
Specificity: P(T-|D-)	0.8966	0.0010	87756.91	10116.56
P(T-|D+)	0.5668	0.0112	1105.86	845.13

[Table continues on the next page]

Probability cont.	Value	SE	Alpha	Beta
Model 2				
P(T+)	0.1120	0.0010	11175.87	88635.08
P(T-)	0.8880	0.0010	88635.08	11175.87
Sensitivity: P(T+|D+)	0.4439	0.0112	866.11	1084.89
P(T+|D-)	0.1053	0.0010	10310.58	87562.92
Specificity: P(T-|D-)	0.8947	0.0010	87562.92	10310.58
P(T-|D+)	0.5561	0.0112	1084.89	866.11
Model 3				
P(T+)	0.1133	0.0010	11308.51	88508.17
P(T-)	0.8867	0.0010	88508.17	11308.51
Sensitivity: P(T+|D+)	0.4322	0.0112	843.14	1107.86
P(T+|D-)	0.1069	0.0010	10465.02	87403.17
Specificity: P(T-|D-)	0.8931	0.0010	87403.17	10465.02
P(T-|D+)	0.5678	0.0112	1107.86	843.14
Model 4				
P(T+)	0.1157	0.0010	11550.68	88276.21
P(T-)	0.8843	0.0010	88276.21	11550.68
Sensitivity: P(T+|D+)	0.4444	0.0112	867.11	1083.89
P(T+|D-)	0.1091	0.0010	10680.93	87177.30
Specificity: P(T-|D-)	0.8909	0.0010	87177.30	10680.93
P(T-|D+)	0.5556	0.0112	1083.89	867.11

Ten-thousand iterations were produced and the mean ICER over all iterations was presented with a 95% confidence interval, using the Youden Index and C-Index as measures of effect. A summary of the probabilistic output is given in Table 6.5.

Table 6.5 Results from the probabilistic cost-effectiveness analysis evaluating the added value of FH information to CVD risk prediction.

Comparison	Change in Cost (95% CI)	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Model 1 vs. Model 2	£22.48 (£5.74 to £39.21)	2.81% (-6.83% to 12.45%)	£799.91 (-£5,962.15 to £5,968.22)	0.29% (-1.26% to 1.84%)	£7,788.76 (-£42,760.16 to £48,962.39)
Model 2 vs. Model 3	£22.43 (£5.81 to £39.04)	3.26% (-6.45% to 12.98%)	£687.02 (-£6,028.59 to £5,583.11)	0.23% (-1.23% to 1.70%)	£9,589.09 (-£40,780.40 to £42,952.87)

6.1.1.2.1 Model 1 vs. Model 2

The probabilistic analysis presented 10,000 scenarios of the change in effect and changes in costs in relation to the specified distributions; each scenario had a unique ICER. These were plotted on the cost effectiveness plane for both measures of effect (Figure 6.7).

Results from the probabilistic analysis are consistent with the deterministic cost-effectiveness analysis. Based on the change in Youden Index and Harrell's C-Index, the mean ICER was £799.91 (-£5,962.15 to £5,968.22) and £7,788.76 (95% CI: -£42,760.16 to £48,962.39), respectively.

A cost-effectiveness plane for both measures are presented in Figure 6.7, and show over half of the iterations from the probabilistic sensitivity analysis fall in the increased cost and increased effect quadrant. Based on this, adding FH to CVD risk prediction is cost-effective over 50% of the time, when compared to CVD risk prediction without FH.

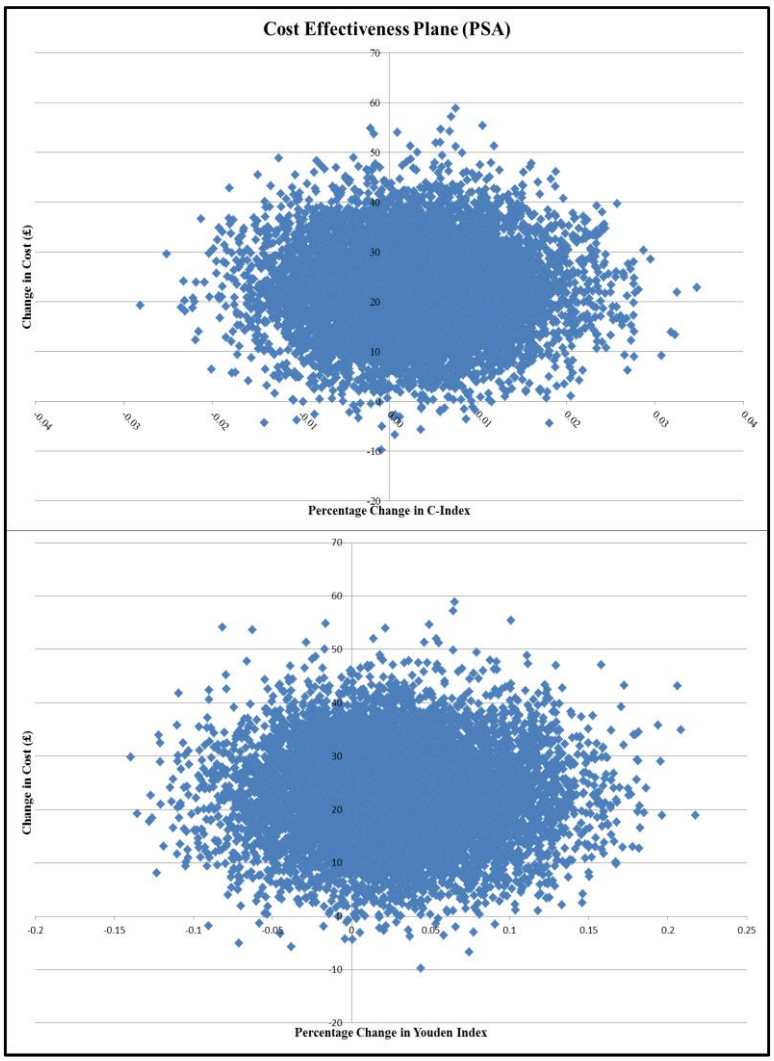


Figure 6.7 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 1 and Model 2.

6.1.1.2.1.1 Cost Effectiveness Acceptability Curve

Figure 6.7 plots the ICER for each scenario generated by the probabilistic sensitivity analysis; dependent on the proportion of ICERs which lie in the cost-effective quadrants (quadrant 1 and quadrant 2), the probability of adding FH being cost-effective can be calculated based on the willingness to pay.

The willingness to pay was varied from £0-£100,000 and the probability of adding FH to standard CVD risk factors being cost-effective was calculated and plotted on the cost-effectiveness acceptability curve (Figure 6.8).

Figure 6.8 shows that when the willingness to pay is £0, a small number of iterations can be found to show FH is cost effective; this is considerably higher when based on the Youden Index (0.43) compared to Harrell's C-Index (0.02). Further, the CEAC curves do not reach a cost-effectiveness probability of 1; this indicates a proportion of iterations from the probabilistic analysis do not show FH as being cost-effective.

Given the maximum willingness to pay of £100,000 per percentage increment of the Youden Index and Harrell's C-Index, the probability Model 2 (with FH information) is cost-effective to Model 1 (with only standard CVD risk factors) is 0.71 and 0.63, respectively. Thus, given the data, there is maximum, 71% and 63% chance that the additional cost of FH information is cost-effective at £100,000 per percentage increment of the Youden and Harrell's C-Index.

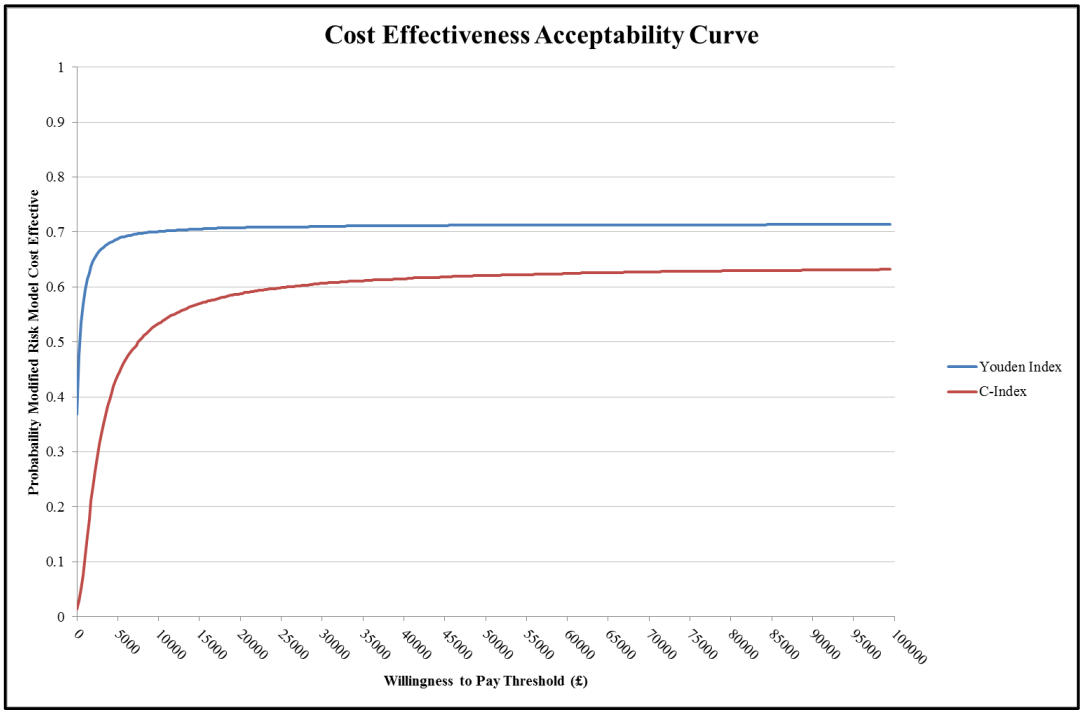


Figure 6.8 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 1 and Model 2.

Further, Figure 6.8 shows that a minimum willingness to pay needed for a maximum 0.71 probability of adding FH to be cost effective is approximately £15,000 per percentage increment in the Youden Index. However, a minimum willingness to pay for a maximum 0.63 probability of adding FH information to be cost effective is approximately £62,000 per percentage increment in Harrell's C-Index. Values below each of these willingness's to pay shows a considerable drop in the probability of adding FH to CVD risk prediction being cost-effective; however, values above do not provide a large increment in the probability of cost-effectiveness.

6.1.1.2.2 Model 3 vs. Model 4

Results from the probabilistic analysis are consistent with the deterministic cost-effectiveness analysis; but a lower ICER was found based on the Youden Index (£687.02 (95% CI: -£6,028.59 to £5,583.11)) and a higher ICER was found with Harrell's C-Index (£9,589.09 (95% CI: -£40,780.40 to £42,952.87)). A cost-effectiveness plane for both measures are presented in Figure 5.9, and show over half of the iterations from the probabilistic sensitivity analysis fall in the increased cost and increased quadrant.

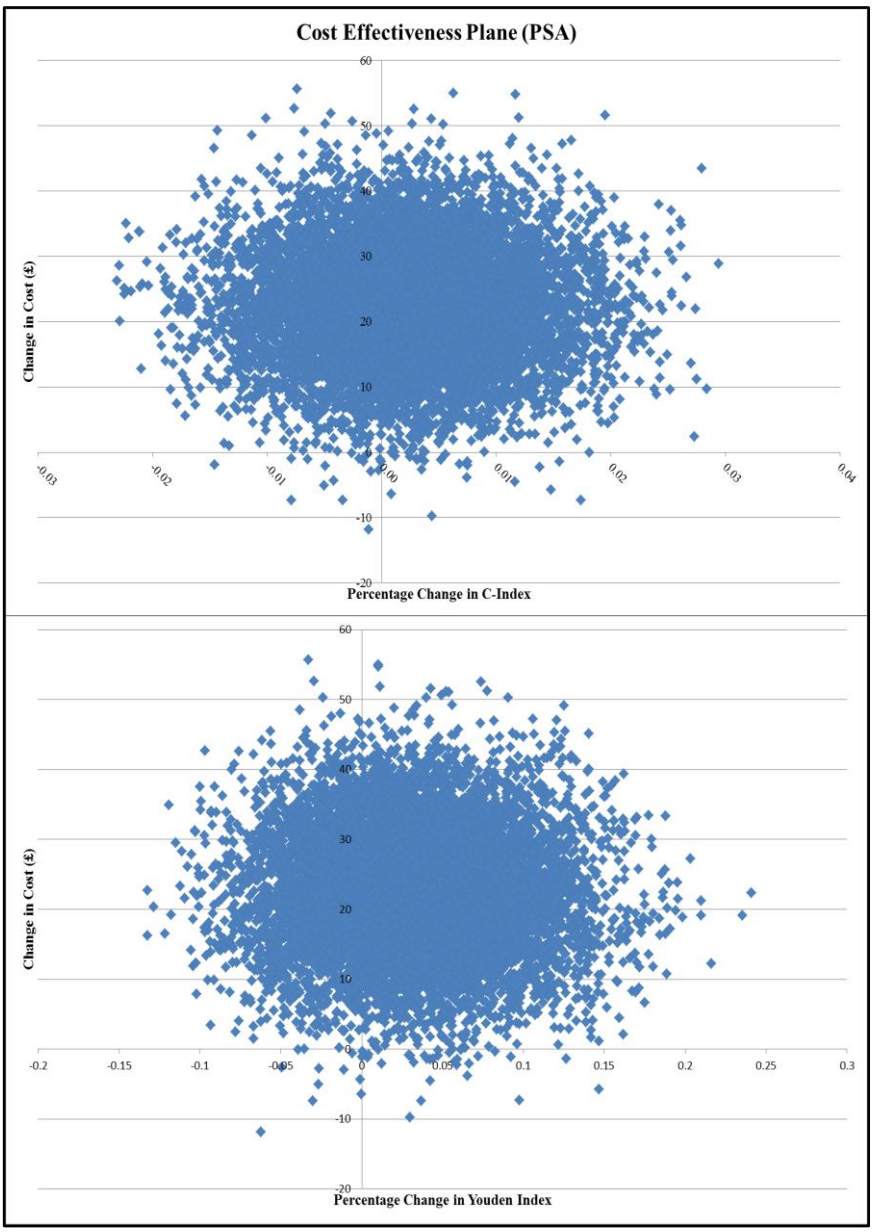


Figure 6.9 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis; comparing Model 3 and Model 4.

6.1.1.2.2.1 Cost Effectiveness Acceptability Curve

A cost effectiveness acceptability curve, varying the willingness to pay for the added effect is presented in Figure 6.10. It shows that given the maximum willingness to pay of £100,000 per percentage increment of the Youden Index and Harrell's C-Index, the probability Model 4 (with FH information) is cost-effective to Model 3 (with only standard and additional CVD risk factors) is 0.73 and 0.61, respectively.

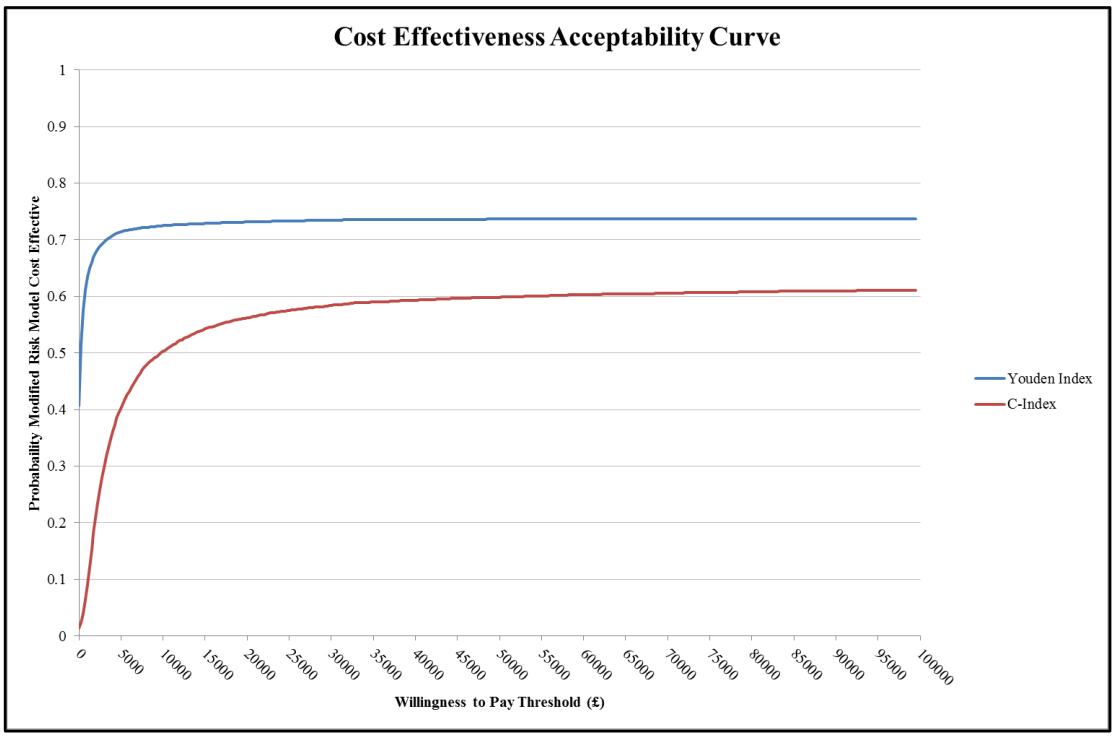


Figure 6.10 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 3 and Model 4.

Further, Figure 6.10 shows that a minimum willingness to pay needed for a maximum 0.73 probability of adding FH to be cost effective is approximately £32,000 per percentage increment in the Youden Index. However, a minimum willingness to pay for a maximum 0.61 probability of adding FH information to be cost effective is approximately £69,000 per percentage increment in Harrell's C-Index.

6.1.2 Sensitivity Analysis

6.1.2.1 Discounting

A sensitivity analysis was performed for each cost-effectiveness analysis, which did not discount the cost in the analysis. This was performed for analyses comparing Model 1 with Model 2, and Model 3 with Model 4.

Table 6.6 presents a summary of the output for each comparison; and results from both the determinist and probabilistic analysis. Small changes were found in the ICERs for

both discriminatory accuracy measures, in both comparisons. No change was found in the results of the cost effectiveness acceptability curve.

Table 6.6 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information using a 0% discount rate.

Analysis	Comparison	Change in Cost	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Deterministic	Model 1 vs. Model 2	£22.48	2.66	£845.28	0.27	£8,260.07
	Model 2 vs. Model 3	£22.44	3.10	£724.50	0.23	£9,782.16
Probabilistic	Model 1 vs. Model 2	£22.61 (£5.95 to £39.28)	2.80 (-6.99 to 12.58)	£809.01 (-£5,988.49 to £6,874.34)	0.26 (-1.27 to 1.80)	£8568.76 (-£41,669.48 to £46,192.06)
	Model 2 vs. Model 3	£22.44 (£5.76 to £39.11)	3.21 (-6.62 to 13.04)	£698.15 (-£6,345.30 to £6,058.21)	0.2 (-1.24 to 1.68)	£10,213.50 (-£52,142.67 to £48,144.68)

6.1.2.2 Complete Case Analysis

Using the cohort of only the complete cases, the decision curve analysis and cost effectiveness analyses were re-run.

No change was found in the results of the decision curve analysis; when comparing Model 1 with Model 2, and Model 3 with Model 4.

The cost-effectiveness analysis for both comparisons, adding FH information to CVD risk prediction, is reported in Table 6.7. Negative ICERs were reported when adding FH information to both standard, and standard with additional CVD risk factors. Both comparisons showed an increased cost for a decreased effect measured by the Youden Index and Harrell's C-Index.

The cost effectiveness acceptability curve showed adding FH to standard CVD risk factors had a less than 50% chance of being cost effective with a willingness to pay of £100,000 per percentage increment in measures of both effect.

Further, based on the percentage change in Youden Index, adding FH information (Model 4) to standard and additional risk factors (Model 3) had a 0 probability of being cost effective.

Table 6.7 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information derived using the complete cases.

Analysis	Comparison	Change in Cost	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Deterministic	Model 1 vs. Model 2	£24.11	-4.41	-£546.80	-0.54	-£4,445.62
	Model 2 vs. Model 3	£24.73	-95.30	-£25.95	-0.49	-£5,062.13
Probabilistic	Model 1 vs. Model 2	£24.18 (£6.87 to £41.49)	-1.43 (-51.66 to 48.81)	-£1,696.23 (-£1,455.74 to £1,544.19)	-0.52 (-£7.88 to £6.85)	-4,681.52 (-£11,061.44 to £9,849.47)
	Model 2 vs. Model 3	24.73 (£6.7 to £42.76)	-95.02 (-105.73 to -84.31)	-26.03 (-£46.77 to -£7.48)	-0.39 (-£8.41 to £7.63)	-6,364.58 (-£8,537.77 to £8,897.06)

6.1.2.3 Literature based Survival

Using CVD risk prediction models created using a literature based survival of 0.9551 (74), the decision curve analysis and cost effectiveness analyses were re-run.

No change was found in the results of the decision curve analysis; when comparing Model 1 with Model 2, and Model 3 with Model 4.

The cost-effectiveness analysis for both comparisons, adding FH information to CVD risk prediction, is reported in Table 6.8.

Similar results were found when comparing Model 1 and 2; however, adding FH information to standard and additional risk factors produced a negative ICER for the percentage change in the Youden Index. Thus, for in increased cost there is a decreased effect.

The cost effectiveness acceptability curve showed less than 50% chance of adding FH information to standard and addition risk factor being cost-effective with a willingness to pay of £100,000 per percentage increment in the Youden Index. Based on Harrell's C-Index, however, the maximum probability of adding FH information being cost effective was 0.6; over a range of £0 to £100,000 willingness to pay.

Table 6.8 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without FH information derived using the imputed data with a literature based survival.

Analysis	Comparison	Change in Cost	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Deterministic	Model 1 vs. Model 2	£22.40	1.41	£1,588.18	0.27	£8,233.32
	Model 2 vs. Model 3	£22.34	-0.26	-£8,459.66	0.23	£9,741.06
Probabilistic	Model 1 vs. Model 2	£22.33 (£5.64 to £39.01)	1.51 (-5.31 to 8.32)	£1,480.74 (-£9,875.08 to £8,265.88)	0.28 (-1.26 to 1.81)	£7,989.23 (-£45,407.15 to £47,600.13)
	Model 2 vs. Model 3	£22.28 (£5.82 to £38.74)	-0.23 (-6.70 to 6.25)	-9,874.65 (-£10,594.22 to £10,272.45)	0.23 (-1.23 to 1.70)	£9,531.26 (-£48,302.26 to £44.132.71)

6.2 Discussion

This section discusses the added value of FH information when evaluated using the refined methodology. It combines the findings of the results using current methodology to provide more clinically relevant information with regards to the addition of FH to CVD risk prediction.

6.2.1 Summary of findings

Adding family history to CVD risk prediction increased the expected cost of the original risk model of each comparison; the cost increased by at least £22.50 based on the cost of treatment, cost of CVD event and cost of collecting the data use the risk model. For this cost, the probability of being a true or false positive increased, and the probability of being a true or false negative decreased; although small changes were found.

This study also found treated low risk CVD patients (false positive) were less costly than untreated low risk patients (true negative); the cost of treatment was absorbed in the lowered event rate in treated low risk patients. Thus, based on the efficacy of statin therapy, over treatment is more beneficial than perfect discriminatory accuracy or correct treatment. In this case, achieving perfect discrimination would not be effective or cost-effective; opting for increased sensitivity and lower specificity. This was influenced by the CVD event rate in clinical trials for patients at low risk patients given placebo compared to low risk patients given statin therapy (58).

This finding is consistent with finding from the NICE clinical guidelines for lipid modification (3). In 2014, NICE lipid modification guidelines recommended reducing the risk threshold for statin therapy to 10% based on both clinical and cost-effectiveness data.

The results for the refined methodology also found positive incremental cost-effectiveness ratios (ICERs) for the change in the expected Youden Index and Harrell's C-Index. This

provided that for an increased cost for the addition of FH information to both standard and standard with additional CVD risk factors, there is an improvement in the performance of the model based on the discriminatory accuracy of the model.

Further, based on the Youden Index, adding FH to standard and additional risk factors compared to standard CVD risk factors alone lowered the ICER further through an increase in this measure of effect. The Harrell's C-Index however, showed an increased ICER, with a reduced effect. In both cases an increased willingness to pay is needed for an increased incremental net benefit (INB) and for the additional of FH to be considered cost effective. The maximum probability, addition of FH information to be considered cost effective, was 70% based on the Youden Index and 60% for Harrell's C-Index.

Performing the probabilistic sensitivity analysis captured the uncertainty around the ICER estimates for both measures of discriminatory accuracy. It showed wide confidence intervals reflecting large amounts of uncertainty in the derived estimates. Although the cost effectiveness plane for each comparison and measure of effect showed a large proportion of cases where adding FH would increase both cost and effect, a few cases were also found to provide increased discriminatory accuracy for a reduction of cost; and be considered cost saving.

6.2.2 Strengths and Limitations

Current literature has evaluated the added value of FH using current model validation methods (8, 9). The effect of FH information on the expected cost of using CVD risk prediction has not been investigated. This study is the first to provide such information with respect to CVD risk prediction models, and making change in performance measures more clinically relevant.

Further, the refined methodology also provided information to help define what an improved performance in a risk model would be in relation to effectiveness and cost-effectiveness. By incorporating treatment cost and treatment efficacy data into the analysis, this identified that false positives were less costly than true negatives, and the cost of unnecessary treatment was absorbed into the cost saved from the CVD event avoided. This effect is only shown when incorporating costs effectively in the evaluation of new markers. Thus, solely improving the statistical performance of a risk model using current methods may produce a less cost-effective model.

However, caution is needed as the analysis is dependent on the treatment efficacy, as this information is needed for patients considered low and high risk of CVD. In this study, two clinical trials were used for this purpose; both were comparable and investigated the effects of 10mg of Atorvastatin vs. a placebo. These trials also allowed for the effects of Atorvastatin to be investigated in high and low risk patients; a sample of patients with a previous CVD event which was used as a proxy for high CVD risk patients, and a sample with no pre-existing history of CVD which was used as a proxy for low risk CVD patients.

Current guidance recommends prescribing 20mg of Atorvastatin as the primary prevention strategy for high risk CVD patients; hence using clinical trials comparing the effects of 20mg of Atorvastatin vs. placebo would be preferable. However, none could be found at the time of this study, which also provided treatment effects on low and high CVD risk patients.

It is also recognised and advised that meta-analyses would be more beneficial to use for data regarding the treatment effects of Atorvastatin compared to individual clinical trial data. These were searched for, however in this case, none were found presenting

information on the primary prevention drug for high CVD risk (Atorvastatin) versus placebo, and in patients considered at low or high risk of CVD.

Further, the decision model assumes that all patients at high risk are treated with preventative treatment. It does not account for the varying compliance rate patients may have depending on the treatment. However, as the decision tree is designed to look at the difference between models with respect to the number of patients treated or not, the same compliance rate to treatment, if treated, is assumed for both risk models; thus balancing this bias.

The refined methodology also accounted for both calibration and discrimination in its summary measure. The disassociation of current measures can limit interpretation, as well calibrated risk prediction models may not have good discrimination, and vice versa. However, this was accounted for in the decision tree which calculates the expected cost for each test based on the disease prevalence in the validation dataset and disease prevalence used to develop the risk prediction models.

A further strength of the study is that the refined method is performed in addition to the current validation methods, and where possible, estimates were derived from the validation study (sensitivity, specificity etc.). Thus, estimates were single study based and more robust.

The study did not factor the cost of collecting possible missing data in GP records required for the use of the risk model. However, the costings did include the collection of every risk factor assuming no information was available. Where it is recognised this may not coincide exactly with clinical practice, which may have some data already recorded and some data missing, it is assumed the total cost of using the risk model would not change or vary over the time points at which data was collected.

Limitations of the study with respect to the refined methodology are reflected by the large amount of uncertainty in estimates of the ICERs. Larger confidence intervals were found for the ICERs calculated from the change in Harrell's C-Index, which was calculated independently from the decision tree. To minimise this limitation, 10,000 iterations of the decision tree were performed, and where possible data for the decision tree was derived from the validation study with respective standard errors.

Small standard errors were found with the estimates of sensitivity and specificity, which were approximately 2.5% of the estimate used. However, larger standard errors were used with estimates of cost which were derived from the literature. As the standard error was not reported for these costs, it was assumed the standard error for all costs in the data was 10% of the estimate. This follows current practice (112-114), but also increases the parameter uncertainty in the model.

Further, the study calculates ICERs for each measure of effect and also presents information on the probability that adding FH information is cost effective based on the willingness to pay. Current NICE guidance uses a £20,000-£30,000 per quality adjusted life year (QALY) to identify if an intervention is cost-effective. However, QALYs have not been used as the measure of effect, and instead discriminatory performance measures are used. This cannot be translatable to the threshold used by NICE. However, it is proposed to use this information to create a league table for new risk factors, allowing them to be compared and better risk prediction models to be identified.

6.2.3 Clinical implications and Conclusion

Both the results from the current and refined methods, showed added benefit of including FH information to CVD risk prediction. Incorporating this information statistically into

the CVD risk prediction model increases the risk of CVD and the likelihood of treating patients correctly and incorrectly; both of which are shown to be cost saving.

This would feed into national guidance and hence clinical practice. Thus the implication of this study would see improved risk prediction translate to improved prevention of future CVD.

6.3 Chapter Summary

This chapter presented and discussed the results of the refined methodology used to evaluate the added value of FH information to CVD risk prediction. This included a cost-effectiveness analysis, adding FH information to standard and standard and additional CVD risk factors.

The next chapter will apply a similar analysis to evaluate the added value of bone mineral density to fracture risk prediction.

Chapter 7 - Bone Mineral Density and Fracture Risk: Introduction and Methods

The previous chapter investigated the added value of family history of cardiovascular disease to cardiovascular disease risk prediction using current and the refined methodology. In this chapter, the added value of bone mineral density to fracture risk prediction is evaluated.

There are two sections to this chapter. The first section gives an introduction to fragility fractures; it explains the rationale behind improving fracture risk prediction, and describes the current practice of fracture risk assessment in UK primary care. It also identifies bone mineral density (BMD) as an important risk factor for fracture risk prediction.

The next section of this chapter describes the methodology used to evaluate the added value of BMD to fracture risk prediction. This outlines the study design, data source, and study population. Details are then given on the analysis which describes the current and refined methodology.

7.1 Introduction

7.1.1 The Burden of Osteoporosis

Osteoporosis is a disease of the bones, which causes them to become weak and fragile; and hence they become more likely to break/fracture. Osteoporosis results in an increased risk of fragility fractures. Fragility fractures are defined as fractures resulting from forces which would not normally result in a fracture, for example, falling from standing height or lower. Such fractures are most common at osteoporotic sites; the spine; hip; and wrist.

Fragility fractures are a major and increasing cause of morbidity; approx. 300,000 of such fractures are presented in UK hospitals, per year (122). With an expected rise in incidence parallel to the world's ageing population, the potential cost resulting from these fractures is also projected to rise to £2.2 billion in 2025 (123).

Osteoporosis causes over 8.9 million fractures worldwide of which 4.5 million occur in the USA and Europe, and account for 2.8 million disability adjusted life years (124). Further, 1.2 million disability adjusted life years are accounted for by hip fractures.

Rises in mortality and morbidity can also be expected due to pain and disability from a fragility fracture. Hip fracture in particular causes significant pain and loss of function, resulting in more hospital care. A hip fracture will normally require surgical repair or replacement; 20% of hip fractures are fatal and 50% result in disability, leaving only 30% to fully recover (125).

In 1990, there were 1.7 million hip fractures worldwide; this is projected to increase to 6 million by 2050 (126). The total cost resulting from osteoporosis in England and Wales is approximately £614 million (127), annually; a large proportion of which is due to hip fractures and approximately 2% from vertebrae fractures.

7.1.2 Prevention of Fragility Fracture in Primary Care

With treatment available to prevent fragility fractures at osteoporotic sites, prevention is a high priority. This is reflected in available primary care prevention schemes and recommendations for a targeted approach from guideline developers (2).

However the challenge, as with most prevention strategies, is to accurately identify people at risk who will best respond to the treatment available. Currently, an opportunistic approach has been used as a high fracture risk case finding strategy (2), rather than a

systematic, targeted screening policy to identify these patients. However, research recommendations have been made to evaluate strategies to better identify patients at risk as part of formal fracture risk assessment.

A strategy employed by other disease areas (1, 3-5) is to use pre-existing patient information found in computerised GP medical records to calculate fracture risk for patients. Hence, case finding in primary care has led to using fracture risk prediction models which combine fracture risk factors into a single equation, calculating the absolute fracture risk estimates for patients. This can be used with a pre-specified risk treatment threshold to class patients at high or low risk of fracture; where high risk patients could be treated to prevent the outcome.

Currently, a consensus agreed treatment threshold of 20% can be used with 10 year fracture risk estimates to identify patients at high risk (128, 129). Once risk patients at high risk of fracture have been identified, preventative treatment can then be prescribed to lower their risk. Initial strategies can include the assessment of falls risk and its prevention, maintenance of mobility and correction of nutritional deficiencies such as vitamin D, calcium, and protein.

Specific therapeutic options available in primary care to reduce fracture risk target treatment of osteoporosis. Bone sparing drug treatments, such as bisphosphonates are used to lower osteoporotic fracture risk. Currently, Alendronate is the primary prescribed therapy to reduce fracture risk (2, 6).

Less consideration is given to lifestyle changes and advice, compared to diseases such as CVD, as more risk factors for fracture risk are non-modifiable and cannot be altered or treated. For example, risk factors such as, previous fracture, parental hip fracture although are strongly associated with fracture, are untreatable.

7.1.3 Fracture Risk Prediction

There are 3 main fracture risk prediction models: the UK FRAX (13); the UK based QFracture (12); and the Australian based Garvan model (130). In the UK, the FRAX and QFracture are the most commonly known and used, but include different risk factors to predict fracture risk.

7.1.3.1 Fracture Risk Factors

Fracture risk factors are a combination of modifiable and non-modifiable risk factors which vary by risk prediction model. Risk factors included in the UK based risk prediction models are presented in Table 7.1.

Table 7.1 List of risk factors used in FRAX and QFracture.

Risk Factors	FRAX	QFracture
Age	30-84 years	40-90 years
Gender	✓	✓
BMI	✓	✓
Weight	✓	✓
Height	✓	✓
Previous Fracture	✓	✓
Parental Hip Fracture	✓	✓
Smoking Status	✓	✓
Alcohol Consumption	✓	✓
Hormone Replacement Therapy	×	✓
Menopausal Symptoms	×	✓
Endocrine Disorders	×	✓
Glucocorticoid use	✓	✓
Secondary Osteoporosis*	✓	×
Asthma	×	✓
CVD	×	✓
History of Falls	×	✓
Gastrointestinal Malabsorption	×	✓
Chronic Liver Disease	×	✓
Rheumatoid Arthritis	✓	✓
Type 2 Diabetes	×	✓
Tricyclic Antidepressants	×	✓
Bone Mineral Density	✓ (optional)	×

***Secondary Osteoporosis includes type 1 diabetes, osteogenesis, hypothyroidism, malnutrition, chronic liver disease, and menopause.**

The non-modifiable risk factors such as age, gender, and those indicating the presence of a disease are primarily used to accurately identify high fracture risk patients. The modifiable risk factors such as alcohol consumption and BMI on the other hand can also be used to treat the patient and lower their risk through lifestyle advice.

As with many diseases, age and gender are significant risk factors for fragility fractures, independent of BMD measurements (131). The frequency of hip fractures specifically has shown to increase exponentially with age (132); and a 1:2 ratio of men to women has also been found in ages above 65 years (132, 133).

Hormonal differences are the main cause of the varied incidences of fractures between males and females. These differences are accounted for in the model by including risk factors such as premature menopause, secondary amenorrhea and hyperthyroidism in addition to gender, and have been used to identify secondary osteoporosis (13).

Other risk factors include body mass index (BMI), where a lower BMI and bodyweight is associated to an increased risk of fracture. Previous fracture is also a well-established risk factor, and has shown to increase risk of a subsequent fracture. Specifically, a prior fracture is associated with a 2 fold relative increase in the subsequent risk of fracture in women, and 3.5 fold relative increase in men (134). The risk of another fracture is particularly high in the first year after the event (135-137).

Use of glucocorticoids (for more than 3 months) is another well-known risk factor for fracture as its adverse effects are known to increase bone fragility. Studies have shown up to a 4 fold increase in the risk of osteoporotic fracture at the traditional sites by taking this medication; and effects were cumulative (138). Similarly, family history also appears to be a risk factor for future fractures and has been recognised in current clinical guidelines for the assessment of osteoporosis (139).

7.1.3.2 Fracture Risk Prediction Models

Fracture risk prediction models predict the 10 year risk of a major osteoporotic fracture. The definition of a major osteoporotic fracture includes fractures at the spine, hip, forearm, pelvis, and proximal humerus (arm). Risk estimates derived from these models identify those at high risk of fractures. This is often based on a 20% consensus driven risk threshold (128, 129).

7.1.3.2.1 FRAX®

The FRAX risk prediction model predicts 10 year risk of a major osteoporotic fracture. It was derived from data collected from 9 prospective population based cohort studies. These are the Rotterdam Study, the European Prospective Osteoporosis Study (EPOS), the Canadian Multicentre Osteoporosis Study (CaMos), and cohorts from Rochester, Sheffield, Dubbo, Hiroshima; and two from Gothenburg (139-145).

The FRAX risk model has therefore not been developed using a UK population; instead it has been re-calibrated for the UK population and is frequently used in UK primary care. However, it has shown to overestimate risk for males and females in the UK, with poorer discrimination compared to the UK based QFracture risk model (146, 147).

Further, although the FRAX risk model has been externally validated (148), it has not been independently evaluated and its model has not been published. Therefore comparisons between models have been difficult (149).

The included risk factors in the FRAX risk model are presented in Table 7.1. Bone mineral density has been highlighted as an optional fracture risk factor, as although it is embedded in the FRAX risk prediction model, it is optionally used in UK primary care (2).

7.1.3.2.2 QFracture®

QFracture also predicts the 10 year risk of a major osteoporotic fracture. It has been derived and internally validated using the QResearch database (v2012). QResearch is a large, primary care database, containing the electronic health records over 13 million patients registered from over 620 UK practices using the Egton Medical Information System (EMIS) computer system. The practices and patient included in this database have

been shown to be nationally representative and also similar to other primary care databases (150).

The original algorithm was published in 2009, used a cohort aged between 30 and 85 years and defined a major osteoporotic fracture to be a fracture at the spine, hip and forearm. It has since been updated in 2012, using a cohort aged between 30 and 100 years and includes the proximal humerus fractures in its outcome definition. The updated model also made further developments, as recommended by NICE guidance (151). This included adding other risk factors (e.g. previous fragility fracture, ethnicity, epilepsy and use of anticonvulsants, and care home residency). However, BMD was not included and remains a strong difference between QFracture and FRAX, in addition to other additional risk factors included in QFracture.

Although the updated model has not been externally validated, the model algorithm has been released and the initial model showed good calibration and discrimination. Results were comparable from both the internal (152) and external validation (149), which also showed better performance in identifying hip fractures compared to other major osteoporotic fractures.

The strengths of using routinely collected data for model development include being able to update models to reflect changes in the population, improvements in data quality, and the ability to easily incorporate new risk factors. However, this model is also limited.

Adding large numbers of new risk factors to models increases the complexity of the model, creating a less parsimonious model. Further, due to the high amount of missing data found in routinely collected primary care databases, including a large number of risk factors increases the likelihood of incomplete data in records, and further limit data collection. Hippisley-Cox et al, describe multiple imputation methods to overcome

incomplete data; however incompleteness would remain an issue when calculating risk using computerised medical records in clinical practice.

Incomplete data in a risk model would increase the inaccuracy of the risk estimate. The requirement of further data collection may affect its likelihood of being used in clinical practice due to additional costs; something which is not currently being measured in validation studies.

7.1.3.3 Additional Risk Factors - BMD

The World Health Organisation (WHO) defines Osteoporosis as being characterised by low bone density and structural deterioration of bone tissue leading to bone fragility and increased susceptibility of fractures. Diagnostically, WHO uses bone mineral density (BMD) to define Osteoporosis; where a t-score value, derived from a DXA scan taken at the femoral neck, below -2.5 standard deviation (SD) from the population average is defined as osteoporotic (128).

Bone mineral density measurement has been shown to improve fracture risk prediction (153-155), and reports a 2 times higher risk of fracture per SD reduction in BMD t-score. It is also directly related to therapeutic options for prevention of fracture, such as bisphosphonates; and thus is a modifiable fracture risk factor. For example, Alendronate is the primary prevention treatment for fractures and is used to slow down bone loss. Yet, the treatable BMD measurement is optionally included in routine fracture risk assessment.

Clinical guidance for FRAX and QFracture has shown it is not compulsory to include BMD measurement at the femoral neck when calculating risk of fracture in the UK. The National Institute for Health and Care Excellence (NICE) recommends research assessing the added value of BMD to fracture risk and to increase its evidence base as a risk factor

for fracture (156). However, the feasibility of evaluating the BMD risk factor in General Practice databases first needs to be assessed.

Lack of recording of BMD in primary care databases has proven to be a common limitation with FRAX and QFracture in the UK, which fail to fully utilise its added value to standard fracture risk factors. The extent to which data on BMD scores from DXA scans is unavailable in routinely collected primary care data is presented in Appendix K. A sample of 60,658 patients was extracted from the Clinical Practice Research Datalink; aged 40-90; not on any osteoporotic treatment; and with complete data for age, gender, BMI, smoking status, and alcohol consumption. Of these patients, 5 (0.01%) had an osteoporosis recording based on BMD measurements taken by a DXA scan.

The recording of BMD in primary care databases is therefore a strong limitation when deciding to include it into routine primary care fracture risk assessment as its data would need to be additionally sourced; in addition to the standard risk factors. National clinical guidance has recommended increasing the evidence base for BMD as a fracture risk factor and evaluating its added cost and clinical effectiveness as part of routine fracture assessment. This would first need to take place using prospective cohort studies not using routinely collected data.

The first step for its evaluation and inclusion in routine fracture risk assessment in UK primary care would be to effectively identify its added value to current fracture risk prediction; statistically (calibration and discrimination), clinically (reclassification), and in a cost-effectiveness analysis. Using this information would better inform guidance decisions to collect BMD information more robustly in routinely collected primary care databases of computerised medical records.

This would then allow for further evaluation studies of the added value of BMD measurement in the UK primary care setting.

7.1.4 Objective and Aims

The aim of this study is to assess the value of BMD measurement in addition to the standard fracture risk factors used in the FRAX® risk model using a robustly collected prospective cohort.

The objectives to achieve this aim are:

- Recalibrate the FRAX® Risk Score, with and without BMD measurement (both in a continuous t-score and binary osteoporotic status format (Chapter 8).
- Internally validate all models and assess the improvement in predictive ability by the addition of BMD measurement, using current measures (Chapter 8).
- Assess the added value of BMD measurement as a fracture risk factor using a decision tree based cost-effectiveness analysis described in Chapter 3 (Chapter 9).

7.2 Methods

The previous section described fracture risk assessment in primary care. It described compulsory (standard) and optional (additional) risk factors which are being included in fracture risk assessment via fracture risk prediction models; determining the added value of bone mineral density (BMD) information to these factors is of particular interest.

This section outlines the methodology used to identify the improvement BMD measurement will add to standard and/or additional fracture risk factors in a fracture risk prediction model setting. It begins by describing the design of the study, the source of the data, and the study population. The extraction of the data used is then described, leading to the proposed statistical and cost-effectiveness analysis to quantify improvement.

7.2.1 Study Design

A prospective cohort study was designed to evaluate medical tests. The 'tests' being evaluated were fracture risk prediction models, with and without BMD measurement as an additional risk factor.

The initial extracted cohort was split into two datasets; two-thirds of the data was allocated to the derivation dataset of the study; and the remaining third was allocated to the validation dataset.

The derivation dataset was used to create a FRAX based fracture risk prediction model, using its established risk factors, with and without BMD measurement.

These models were then 'tested' using the validation dataset. This assessed each of their performance which allowed for comparisons between models to be made and the impact of BMD on fracture risk assessment to be shown.

7.2.2 Data Source

Baseline data from the Aalborg University Hospital Record for Osteoporosis Risk Assessment (AURORA) (157, 158) was used for this study, and was located in Aalborg, Denmark. It was initially collected to assess the age discrimination of osteoporosis screening, and thus had rigorous and validated data on all fracture risk factors used in the FRAX® algorithm tabulated below (Table 7.2).

To follow up patients of this dataset, the National Patient Registry of Denmark (159) was also used and linked to the AURORA dataset. This database collects information of patients in Danish hospitals.

7.2.2.1 Data Collection of AURORA

7.2.2.1.1 Referral

The data was collected, taking advantage of the re-organisation of osteoporosis screening in North Denmark in January 2010. Along with the increased DXA capacity, a new scheme for referral and reply was launched. As patients presented themselves for their referred DXA scan, a short self-reported questionnaire detailing fracture risk factors was administered; improving the quality of the reply sent to the referring doctor.

Patients were referred by their doctor if they presented with at least one fracture risk factor (low BMI, previous fracture, parental hip fracture, smoking status, alcohol consumption, glucocorticoid use, rheumatoid arthritis, secondary, osteoporosis) or if they were aged 80 years or above.

7.2.2.1.2 DXA Scan and Anthropometric Measurements

On referral by their doctor, patients were sent an acknowledgement of the received referral form to the hospital, an appointment for a DXA scan, with a short questionnaire.

The DXA scan would be performed and other anthropometric measurements would be taken. On completion of the referral appointment, DXA scan results, data on osteoporosis risk factors, as well as anthropometric measurements had been collected in the AURORA database.

The DXA scan was performed using a Hologic scanner and recorded bone mineral density as a t-score value from the total hip and L1-L4 of the spine.

Height was measured on a Harpender stadiometer and recorded in metres.

7.2.2.1.3 Questionnaire

A short self-reported questionnaire was provided with the DXA scan appointment letter. The questionnaire systematically included questions for age; gender; previous fracture (giving options for spine, distal forearm, hip, and other); year of fracture and description of the underlying mechanism; family history of a fragility fracture (spine, forearm, or hip) with a description of the affected family member; smoking status; alcohol intake; corticosteroid intake; other disease (if yes, description); treatment for osteoporosis (if yes, drug name and duration); sun exposure; dairy intake; and menopause (if yes, age of onset).

7.2.2.1.4 Validation of results

Data was reviewed for each patient referred to all hospitals in the region on North Denmark by the same, single, endocrinology consultant. Digital X-rays and laboratory results relevant to bones (including chest X-rays) were all reviewed by the said consultant.

Further, the short self-reported questionnaire was presented to the technical staff at the DXA scan appointment; staff would check the questionnaire for any missing data and ambiguity, and encourage completion.

7.2.2.2 Data collection of Danish National Patient Registry

Data collection of the Danish National Patient Registry (159) began in 1977 and collects data of patients in all Danish hospitals. Originally, only inpatient data was collected, but in 1995 it became mandatory to also record outpatient data.

Data for each hospital record was separated into numerous data files, each linked with a record number. Of these files, two files were used for this study; 't_adm' and 't_diag'. The former contained all the admission data for the hospital record and also the primary diagnosis; these were used to extract the primary diagnosis and also date of the diagnosis. The latter data file, included further diagnosis details of the patient, and included details on secondary diagnoses, complications, referred diagnoses, temporary diagnoses and appendix diagnoses. It also included the mechanism, if any, of the diagnosis.

7.2.3 Study population

Data collection for the AURORA dataset began on 1st January 2010 and data was collected for 3 years (up to 31st December 2012). The baseline date for the study is therefore 1st Jan 2010. The follow up date for patients is 1st January 2014; which will allow for 2-4 year follow up period.

In line with the age range given with the FRAX risk score, patients must be aged between 40-90 years and will be excluded if they are on any osteoporotic preventing drug or taking any bone sparing drugs for more than 1 year prior to baseline.

7.2.4 Study variables

Variables for this study are based on existing fracture risk prediction models; namely FRAX. A summary of this risk model and its risk factors are given in Table 7.2.

Table 7.2 List of study variables, including FRAX® risk factors.

Variable List	**Format**
Patient ID	Continuous
Baseline study date	Date – 01/01/2004
Follow-up end date	Date – 31/12/2014
FRAX Risk Factors (Standard Risk Factors)	
Age	Continuous (years)
Sex	Dichotomous (male/female)
Height	Continuous (cm)
Weight	Continuous (kg)
Previous Fracture (in adult life, occurring spontaneously, or from trauma which otherwise would not result in a fracture.	Dichotomous (yes/no)
Parental Hip Fracture	Dichotomous (yes/no)
Current smoker	Dichotomous (yes/no)
Glucocorticoid use (currently exposed/has been for more than 3 months)	Continuous (mmol/l or mg/dL)
Rheumatoid Arthritis	Dichotomous (yes/no)
Secondary Osteoporosis*	Dichotomous (yes/no)
Alcohol consumption (more than 3 units per day)	Dichotomous (yes/no)
Outcome Data	
Outcome: Fracture at an osteoporotic site	Dichotomous (yes/no)
Fracture event date	Date
Additional Risk Factors	
Bone Mineral Density	Categorical (osteoporotic/otherwise – normal/osteopenic)
Bone Mineral Density	Continuous (t-score value – SD)

***Secondary Osteoporosis includes type 1 diabetes; osteogenesis imperfecta in adults; untreated, long standing hyperthyroidism; hypogonadism; premature menopause (<45 years); chronic malnutrition; malabsorption; and chronic liver disease.**

Data for all risk factors presented in Table 7.2 have been recorded in the AURORA dataset, except for rheumatoid arthritis and secondary osteoporosis. However linkage to the Danish National Patient Registry allowed for these diagnoses before study baseline to

be extracted for the patients contributing to the AURORA dataset. Details on this extraction are given below (Derivation of Standard Risk Factors).

7.2.4.1 Outcome

The outcome for this study is a fragility fracture or fracture at an osteoporotic site (clinical spine, hip, forearm, pelvis and proximal humerus) within the follow-up period, 1st Jan 2010 to 1st January 2014. A comprehensive ICD code list was developed for each of these fractures and also fractures not otherwise specified (NOS) but were not due to an external force.

To ensure systematic code selection for each code list, recognised database methodology (160) was used and applied using Stata statistical software. The ICD code dictionary was searched for relevant ICD code stems for the outcome, for example some codes under the stems M (Diseases of the musculoskeletal system and connective tissue), and S and T (Injury, poisoning and certain other consequences of external causes), are associated with fractures. Duplicate and irrelevant codes were then removed.

The ICD code list was also searched for terms such as “fracture”, “fragility”, “femur”, “spine”, and “forearm” etc. The developed ICD code list was then amended to match Danish ICD codes by prefixing a “D” on the UK based ICD code. For example, the UK based ICD code S72 indicates ‘fracture of femur’, and would be coded as DS72 to link to Danish records.

The code list was then linked to the Danish National Registry dataset to search for patients with the outcome code during the follow up period.

To identify a fragility fracture, ideally both the fracture and mechanism need to be known, as they are specifically non traumatic events. Both pieces of information were found in

the linked data and fractures resulting from traffic, work and sports accidents were excluded from the study.

7.2.4.2 Derivation of Standard Risk Factors

The standard risk factors and their required format for the analysis for the study are outlined in Table 7.2.

7.2.4.2.1 Age

Age was taken at the DXA scan appointment. It was recorded in years and treated as a continuous variable in the analysis. For descriptive purposes, age was also categorised into 10 year age bands.

7.2.4.2.2 Gender

Gender was specified for each patient; they were classed as either male or female.

7.2.4.2.3 Body Mass Index (BMI)

Height and weight was taken at the DXA scan appointment. Height was measured on a Harpender stadiometer and recorded in metres (m). Weight was recorded in kilograms (kg). Using this information the patients BMI was calculated using the formula:

$$BMI = \frac{weight\ (kg)}{height\ (m)^2} \qquad \text{Eq.7.1}$$

7.2.4.2.4 Previous Fracture

Both the self-reporting questionnaire and the Danish National Patient Registry were used to identify previous fractures. Options given in the questionnaire were fractures at the spine, hip forearm and other fractures.

Previous fractures in the Patient Registry were identified using the developed fracture ICD code list. Fractures before the baseline date were used to indicate a previous

fracture. Further, this also allowed calculating the number of previous fractures the patient may have had.

Both variables were aggregated to produce a single, binary previous fracture variable. If patients reported a previous fracture, the number of previous fractures was also presented. If a previous fracture was reported in both the questionnaire and the patient registry, only the patient registry data was included in the calculation of the number of previous fractures, to avoid double counting.

7.2.4.2.5 Parental Hip Fracture

Parental fractures were taken in the short questionnaire at the DXA scan. This included fractures at the hip, spine or forearm. However, according to the FRAX risk score, only parental hip fractures were used for the study.

7.2.4.2.6 Smoking status

Smoking status was recorded and used as a binary variable. No additional descriptive information on the amount smoked was provided.

7.2.4.2.7 Alcohol Consumption

Alcohol consumption was also recorded in the AURORA dataset as a binary variable. This indicated patients who had a daily intake of more than 3 units of alcohol per day. This equated to and was described to the patient as having more than 2 beers, or more than 2 glasses of wine, or more than 4cl of liquor.

7.2.4.2.8 Osteoporosis causing medication

Glucocorticoid use was indicated through taking more than 5mg of corticosteroids (for example Prednisone); this was presented in a binary variable.

7.2.4.2.9 Rheumatoid Arthritis

Rheumatoid Arthritis (RA) was not recorded in the AURORA dataset, but was extracted from the Danish National Patient Registry in a similar manner to the outcome.

An ICD code list was created for Rheumatoid Arthritis, which was then linked to the Danish National Patient Registry administration and diagnosis files. Records were restricted to before the study baseline date, to identify existing patients diagnosed with RA.

7.2.4.2.10 Secondary Osteoporosis

Secondary osteoporosis is defined as osteoporosis that is drug induced or caused by certain conditions. FRAX® includes glucocorticoid use as a risk factor and hence includes disorders such as: type 1 diabetes; osteogenesis imperfecta in adults; untreated, long standing hyperthyroidism; hypogonadism; premature menopause (<45 years); chronic malnutrition; malabsorption; and chronic liver disease to indicate secondary osteoporosis.

Of these conditions, data on bone affecting diseases and drugs, and premature menopause were recorded in the AURORA dataset. For the other conditions, data was extracted from the Danish National Patient Registry, using methods similarly outlined for Rheumatoid Arthritis. That is, ICD codes lists were created for each condition and used to extract records of patients before the baseline date. This was then used to identify patients with secondary osteoporosis; those who had at least one condition.

7.2.4.3 New Risk Factor – Bone Mineral Density

Bone mineral density was explicitly recorded in the AURORA dataset. It was taken at the DXA scan appointment, which was performed on a Hologic scanner. Bone mineral density was recorded as a T-score taken at the total hip and L1-L4 of the spine. For

descriptive purposes patients were classed as osteoporotic, osteopenic, or normal based on their score using the following guidelines (128):

- Osteoporotic (T-score≤-2.5)
- Osteopenic (-2.5<T-score<-1.0)
- Normal (T-score≥-1.0)

A binary BMD risk factor was also created to indicate osteoporotic patients, to be used in the analysis (osteoporotic, or otherwise – normal/osteopenic).

7.2.5 Statistical Analysis

The statistical analysis for adding BMD to fracture risk prediction models is similar to the statistical analysis performed when adding FH to CVD risk prediction models. A full description of the statistical analysis can be found in Chapter 4, Section 4.2.5.

Key differences between the analytical methods have been described below.

7.2.5.1 Descriptive Analysis

The extracted AURORA dataset used for the study was described in line with the inclusion and exclusion criteria; it was presented using a CONSORT style diagram.

The characteristics of the sample was described overall and also by BMD status (osteoporotic or otherwise – normal/osteopenic). The incidence of fracture in the dataset was presented, and the baseline characteristics were compared between patients with and without a fracture during follow-up. A Bonferroni adjusted significance level of 0.002 was used (0.05/12).

Crude incidence rates were also calculated for each categorical standard risk factor in the study, including categorical BMD measurement.

Further details can be found in Chapter 4, Section 4.2.5.1.

7.2.5.2 Generalisability of the data

The extracted sample was compared to an extracted sample from the Clinical Practice Research Datalink (CPRD), using the same inclusion and exclusion criteria. All risk standard factors (gender, alcohol consumption, height, weight, BMI, smoking status, previous fracture, secondary osteoporosis, glucocorticoid use, and rheumatoid arthritis) were compared. P-values were presented from chi-square tests for categorical data, t-tests for normally distributed continuous data, and Wilcoxon Rank Sum tests for non-normally distributed continuous data. A significance level of 0.05 was used.

The level of recording for each of these variables was presented, as well as the trends for each risk factor over 10 year age groups. Finally, the fracture incidence rates were compared between datasets.

Further details can be found in Chapter 4, Section 4.2.5.2.

7.2.5.3 Missing Data

Missing data was presented in the descriptive analysis. Patterns of missing data were checked; logistic regression and pair-wise correlations were used to assess the association of risk factors with the missing data.

Missing data for was excluded for the analysis. A sensitivity analysis was performed, including the missing values as a subgroup for each variable.

Further details on the correlation and regression of missing data can be found in Chapter 4, Section 4.2.5.3.

7.2.5.4 Derivation and Validation

The extracted sample from the AURORA dataset was splits into two dataset; two-thirds was used to derive the fracture risk prediction model (training dataset); and one-third was used to validate them (test dataset).

Baseline characteristics were compared between datasets; missing values were also presented.

Further details can be found in Chapter 4, Section 4.2.5.5.

7.2.5.5 Derivation of the Models

Using the training dataset, the primary analysis created four fracture risk prediction models based on the standard and new risk factor:

- Model 1: BMD only
- Model 2: Standard risk factors
- Model 3: Standard fracture risk factors with BMD
- Model 4: Data driven fracture risk prediction model, including BMD

These were created using both the binary (osteoporosis indicator) and continuous form of BMD measurement (t-score). In total 7 fracture risk prediction models were created.

7.2.5.5.1 Exploration of the data

The relationship between standard risk factors, as well as with the new BMD risk factors was explored. Further details can be found in Chapter 4, Section 4.2.5.6.1.

7.2.5.5.2 Model Building

Cox proportional hazards regression was used to derive these models to account for the time to event data and also to account for patient dropouts through censoring.

Regression diagnostics were assessed for each model of the analysis. Model assumptions were also checked as well as any outliers.

Further details can be found in Chapter 4, Section 4.2.5.6.2.

7.2.5.6 Creating the Risk Prediction Model

Beta coefficients were used to create the fracture risk prediction models. Further details can be found in Chapter 4, Section 4.2.5.7.

7.2.5.7 Validation of the Models – Current Methodology

The derived risk prediction models were internally validated using the derivation dataset and validation dataset.

Validation of the fracture risk models consisted of the assessment of 5 parts. First, the fit of the model was assessed. The calibration of the model to the respective datasets was then analysed; followed by the assessment of the models ability to discriminate between patients with and without a fracture during follow-up. The ability of the models with BMD measurement to reclassify patients into their correct risk groups was assessed; which leads to the decision curve analysis (DCA).

Methods used to assess the validity of the models are described in full detail in Chapter 4, Section 4.2.5.8.

Comparisons were made between Model 2 and Model 3; and Model 2 and Model 4; using binary and continuous forms of BMD measurement as the new risk factor.

A pre-specified risk threshold was used for the analysis. The 10 year risk threshold used for fracture risk is 20% (128, 129); this was reduced to give a 4-year risk threshold adjusted for the follow up period for this study.

A 20% risk threshold assumes an 80% non-fracture rate at 10 years. Thus, the annual survival rate based on this risk threshold is 0.978 ($0.8^{1/10}$), giving an annual risk threshold of 2.2% ($(1 - 0.978) \times 100$). The 4-year risk threshold was therefore, 8.5% ($(1 - 0.978^4) \times 100$) and was used to identify patients at high or low risk of a fracture over 4-years.

7.2.5.7.1 Decision Curve Analysis

A decision curve analysis was performed to identify the change in net benefit between models with and without BMD at the 8.5% risk threshold. Further details of this analysis can be found in Chapter 4, Section 4.2.5.8.4.

7.2.5.8 Validation of the Models – Refined Methodology

7.2.5.8.1 Cost-effectiveness analysis – Refined Methodology

The final part of validating the fracture risk prediction models and evaluating the added value of BMD as a fracture risk factor will be a cost-effectiveness analysis using a decision tree framework, as discussed in Chapter 3, Section 3.3.3.

To ensure a robust cost-effectiveness analysis the CHEERS checklist (66) was used alongside the NICE Reference Case (65); this consolidates all health economic reporting guidelines. It was used to develop the refined methodology in Chapter 4; and was used to create a detailed, study specific methodology for assessing the added value of BMD to fracture risk prediction.

7.2.5.8.2 General Methods

The target population were taken from a Danish primary care setting; patients were referred from primary care based on their fracture risk factors to an osteoporosis clinic to have their bone mineral density measured by a DXA scan. The study sample has been

described in Section 7.2.3. The risk prediction models were not developed for any subgroup.

The perspective taken for the analysis was from the UK based National Health Service (NHS) and the Personal Social Services Research Unit (PSSRU) (70); and the comparators for risk prediction models with (Model 3 and 4) and without (Model 2) BMD measurement.

The time horizon used for this study reflects the follow-up period used for the developing the risk prediction model; 4 years. Costs were discounted in the primary analysis using a 3.5% (67) discount rate; the government discount rate. A sensitivity analysis was also performed using a 0% discount rate.

7.2.5.8.3 Outcomes

The outcomes of choice for the cost-effectiveness analysis were performance related measures. In particular, the Youden Index and Harrell's C-Index, measures of discriminatory accuracy were used. These were derived from current validation methods described earlier (Section 7.2.5.7).

The measurements of effectiveness were based on a single study; the respective validation study. Preference based outcomes were not used, reasons for this are provided in Chapter 3, Section 3.2.1.1.

7.2.5.8.4 Costs

Three sets of costs were needed for the analysis; for the risk prediction model; the fracture outcome; and the treatment. Costs were UK costs, for the costing year of 2013-2014. All three sets of costs were derived from the NICE Osteoporosis: fragility fracture risk

costing report (25), the NICE Osteoporosis primary prevention costing template and report (161), Stevenson et al (62), and Kanis et al (162).

7.2.5.8.4.1 Cost of Risk Prediction Model

The process of assessing the risk of a patient is outlined as follow:

- Initial consultation with GP - includes collection of standard fracture risk factors (age, gender, height, weight, previous fracture, parental hip fracture, current smoker, alcohol consumption, glucocorticoid use, rheumatoid arthritis, and secondary osteoporosis) and referral for a DXA scan.

- DXA Scan - Patient is referred for a DXA scan and attends appointment, where technical staff perform DXA scan.

- Follow up GP consultation - Patient returns to GP to discuss results and, if required, preventative treatment for CVD.

The cost of using a risk prediction model and collecting its data was assumed to take one GP appointment. Unit costs and average consultation lengths associated with practice nurses and GPs was extracted from PSSRU Unit Costs of Health and Social Care 2014 (70). These have been provided in Table 4.7.

The cost of using the risk prediction model collecting only standard risk factors was hence £45.63 per patient. The cost of a DXA scan was £69.00 (25) (total cost: £114.63).

Cost of a follow up GP appointment was assumed to be absorbed into existing costs of using the risk prediction model.

7.2.5.8.4.2 Cost of treatment

After the assessment of fracture risk, patients were identified as low or high risk of fracture based on the 8.5% risk threshold. In line with current guidance, it was assumed all patients at high risk of fracture were treated with bisphosphonate drug therapy.

Currently, bisphosphonates, Raloxifene, Strontium Ranelate, and Teriparatide are prescribed to prevent fragility fracture events in Primary and Secondary Care; of which bisphosphonates and Strontium Ranelate are mainly prescribed in Primary care. Bisphosphonates include Alendronate; Etidronate; and Residronate. Of these, Alendronate has been recommended as the primary prevention treatment for fracture (6).

As per the NICE Osteoporosis primary prevention costing template (161), Table 7.3 shows the estimated current and future prescribing proportions for each of these therapies, as well as the respective costs. For this study future prescribing proportions were used. The unit cost used to calculate the weighted average cost is based on the Prescription Pricing Authority (PPA) NHS Electronic Drug Tariff (110), assessed in 2008; and drug costs are based on the British National Formulary (BNF) edition 54.

Table 7.3 Osteoporosis Primary Prevention Drug Prescribing Proportions (161).

Drug Therapy	Weighted Cost of Drug (£)	Future Prescribing Proportion (%)	Weighted Prescription Cost (£)
Alendronate	24.39	81.2	19.80
Etidronate	84.48	1.7	1.47
Risedronate	263.15	15.6	41.16
Strontium Ranelate	333.71	1.5	4.87
		Weighted Cost of Treatment	**67.29**

Based on the estimated prescribing proportions and the weighted cost of each drug therapy, the weighted cost of treatment was estimated as £67.29 per patient per annum.

7.2.5.8.4.3 Cost of osteoporotic fracture

The outcome of an osteoporotic fracture included fractures at the hip/femur, spine (vertebrae), forearm, pelvis, and proximal humerus. The weighted cost for each outcome is presented in Table 7.4 and is derived using costs from Stevenson et al (62), weighted by the proportion of osteoporotic fractures given in Kanis et al (162).

Table 7.4 Weighted cost of an osteoporotic fracture. Cost are from Stevenson et al (62), and fracture proportions are from Kanis et al (162).

Fracture Site	Unit cost of event (£)	*Number of events	Proportion of total %	Weighted Outcome Cost (£)
Hip	10,760	18,870	40.2	4,328.50
Femur	13,771	1,406	3.0	412.77
Vertebra	1,706	9,781	20.9	355.73
Distal Forearm	537	8,183	17.4	93.68
Pelvis	9,236	2,241	4.8	441.24
Proximal Humerus	1,112	6,427	13.7	152.36
		Weighted Cost of treatment		**5,784.27**

The weighted cost of an osteoporotic fracture was estimated to be £5,784.27.

The cost of treatment and the outcome fracture was then combined to derive the cost of an untreated event from a low risk patient; the cost of a treated event from a low risk patient; the cost of an untreated event from a high risk patient; the cost of a treated event from a high risk patient.

To calculate the cost of each of these scenarios, the fracture event rate was extracted from a clinical trial investigating the effect of 5mg/d Alendronate vs placebo in female patients

aged 54-81 years, with low bone mineral density and without a vertebral fracture (163). Event rates for fracture were presented by categorical t-score for BMD measurement and were used as a proxy to identify low risk (not osteoporotic: -2.5<t-score<1.6) and high risk (osteoporotic: t-score≤-2.5) patients; these are presented in Table 7.5.

Table 7.5 Fracture event rates from the Cumming et al (163).

	No. Fractures	Total No. Participants	Fracture Event Rate*
Untreated Low risk	153	1406	0.11 (0.09 to 0.13)
Untreated High risk	159	812	0.20 (0.17 to 0.23)
Treated Low risk	165	1395	0.12 (0.10 to 0.14)
Treated High risk	107	819	0.13 (0.11 to 0.16)

***event rate calculated for the study follow-up rate; 4.2 years.**

Using methods described in Fleurence et al (69), the fracture annum rate and annual transition probabilities were calculated. These were then combined with the weighted fracture cost and cost of treatment to give the average cost of an untreated fracture event in low risk patients; a treated fracture event in low risk patients; an untreated event in high risk patients; and a treated event in high risk patients over a 4 year follow up. These costs are provided in Table 7.6.

Table 7.6 Average cost of treated and untreated fracture events over 4 years follow up, by discount rate.

	0% discounted Average Cost (£)	3.5% discounted Average Cost (£)
Un-treated Low Risk	626.05	574.87
Untreated High Risk	1169.89	1074.27
Treated Low Risk	952.32	874.49
Treated High Risk	1027.72	943.72

7.2.5.8.5 Model-based Economic Evaluations

A decision tree was chosen to compare two risk prediction models. Reasons for this have been described in Chapter 3, Section 3.2.2.1.

For the analysis it has been assumed all patients at risk are treated. Further, the treatment is according to the estimated prescribing proportions found in the NICE osteoporosis primary prevention guidelines (161).

7.2.5.8.6 Analysis

7.2.5.8.6.1 Decision Tree

The proposed analysis has been described in full in Chapter 3, Section 3.3.5. Figure 7.1 presents the format for the decision tree used to evaluate the added value of BMD to fracture risk prediction; including standard fracture risk factors. An 8.5% risk threshold was used to identify patients at high risk.

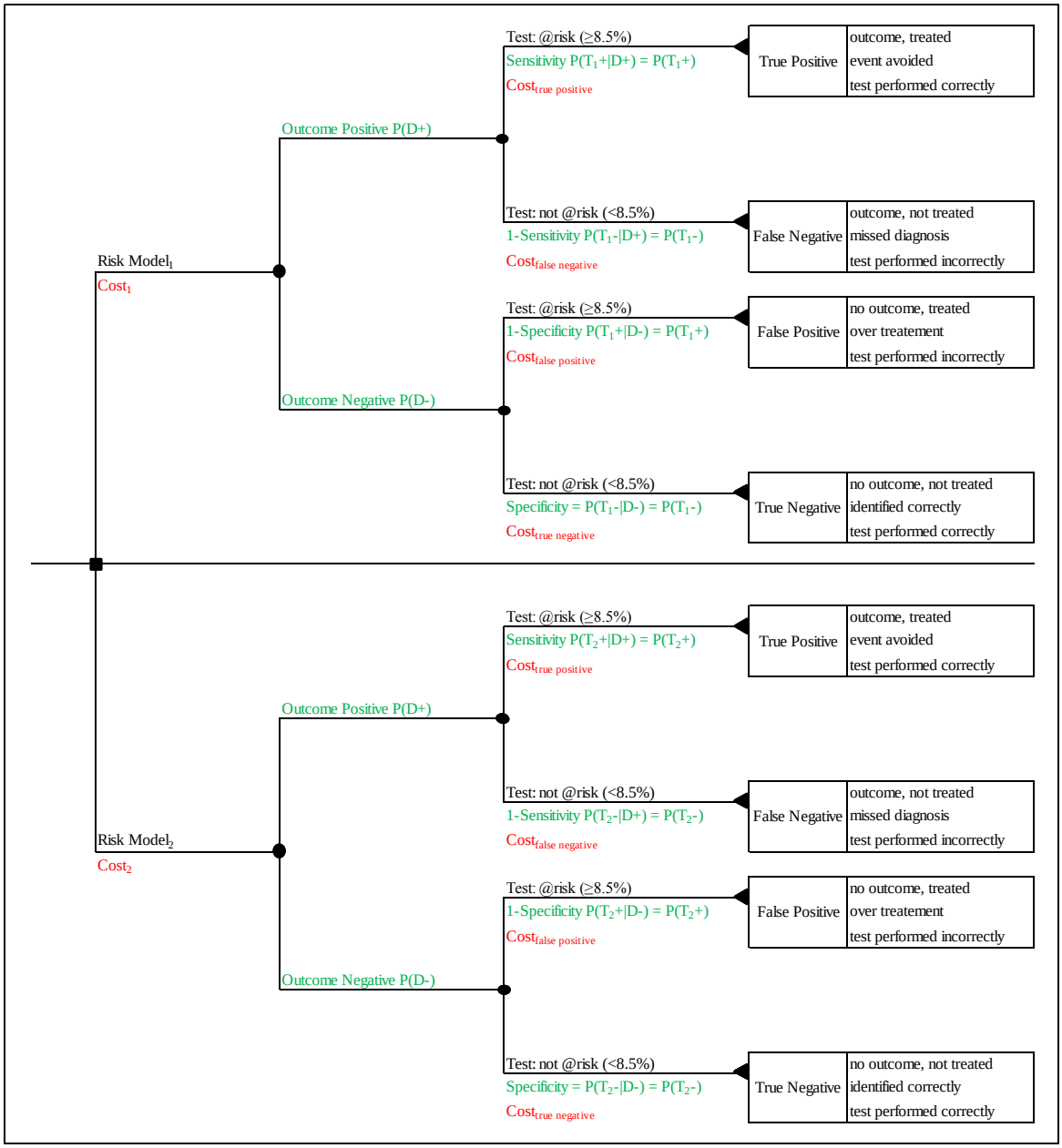


Figure 7.1 Decision Tree to compare two fracture risk prediction models.

7.2.5.8.6.2 Input and Output of the Decision Tree

The probabilities for each event node were calculated from the current performance measures in the validation dataset; methods for this are outlined in Chapter 4, Section 4.2.5.8.2.

The output from the decision tree was the expected cost of each risk prediction model; which was based on the cost and probabilities associated with the performance of each model. This was calculated using Box 3.6 in Chapter 3, Section 3.3.5.1.

7.2.5.8.6.3 Incremental Cost-Effectiveness Ratio and Net Monetary Benefit

The output of the decision tree; the expected cost of using a risk model with and without BMD measurement, was combined with the measures of effect described above, using the Incremental Cost Effectiveness Ratio (ICER). Full details of the methods can be found in Chapter 4, Section 4.2.5.9.1.5.3.

7.2.5.8.6.4 Probabilistic Sensitivity Analysis (PSA)

Although the above deterministic cost-effectiveness analysis presented the ICER, an effect size for the added value of BMD measurement to standard fracture risk assessment, it did not quantify the uncertainty.

A PSA analysis was also performed in addition to the deterministic analysis above. Full details have been provided in Chapter 4, Section 4.2.5.9.1.6.

Table 7.7 provides the information required for each cost value used in the PSA, values for the probabilities are presented in Chapter 9, Section 9.1.1.2.

Table 7.7 Deterministic mean costs of risk model (with and without BMD), treatment and fracture; with standard error, alpha, and beta for estimation of probabilistic cost derived from a Gamma distribution.

	Mean Cost (£)	SE	alpha	beta
Annual cost treated Fracture	5,784.27	578.427	100	57.8427
Annual cost untreated Fracture	5,784.27	578.427	100	57.8427
Treatment cost	67.29	6.729	100	0.6729
Cost of Model without BMD	45.63	4.563	100	0.4563
Cost Model with BMD	114.63	11.463	100	1.1463

Ten-thousand iterations were produced and the mean ICER over all iterations was presented with a 95% confidence interval.

7.2.5.8.6.4.1 Net Monetary Benefit

Further analysis was also performed as part of the PSA. An intervention is deemed as cost effective if the ICER is less than the amount a person is willing to pay; thus the willingness to pay acts as a threshold to the ICER. The INB accounts for both the ICER and the willingness to pay, such that if the INB is above 0 for a given willingness to pay, the intervention is cost-effective. The derivation of the INB formula is presented in Chapter 3, Section 3.3.5.1.2.

Similar to the ICER, two INBs were calculated. The first used the Youden Index as a measure of effect and the second used the AUC (Harrell's C-Index) value for each model. No recommendations are available for the willingness to pay threshold and the value per incremental measure of effect, for which the intervention would be deemed cost effective. Hence the INB was plotted over a range of possible values for the willingness to pay; producing a Cost Effectiveness Acceptability (CEAC) Curve for both measures of effect. This presented the probability of the intervention (new risk model – with BMD) being cost effective, compared to standard care (original risk model – without BMD); over a range of willingness to pay thresholds (£0-£100,000).

All analysis took place using Stata v12.1(115) and MS Excel.

7.3 Chapter Summary

This chapter gives and introduction to fracture risk assessment in UK primary care, and outlines the rationale for adding BMD measurement to standard fracture risk factors. It describes the methodology to evaluate the added value of BMD measurement to fracture risk prediction. This includes the study design, methods to develop the fracture risk

models, current methods to validate the models and evaluate the added value of BMD measurement, and refined methods to further evaluate the clinical impact of the adding BMD measurement routine fracture risk assessment.

The next chapter presents the results from the data extraction; descriptive analysis; model development; and current validation of the models.

Chapter 8 - Bone Mineral Density and Fracture Risk: Results and Discussion (Current Methodology)

The previous chapter gave the rationale and described the methodology for evaluating the addition of BMD measurement to fracture risk prediction. The analysis was described in two sections; using current methods; and refined methods.

This chapter presents the results from carrying out the current methods described. This chapter begins with outlining the extraction of the data; describing its baseline characteristics; and assessing its generalisability to the UK. Missing data is excluded from the analysis before performing an unadjusted and adjusted analysis to derive the fracture risk prediction models with and without bone mineral density measurement.

Finally, the derived risk prediction models are validated using current performance measures; describing calibration, discrimination, reclassification, and the net benefit derived from the decision curve analysis (DCA).

8.1 Results

8.1.1 Dataset

During the two year collection period for the AURORA dataset (1st Jan 2010 – 31st Dec 2012), 7,912 patients were referred by their doctor to the hospital for a DXA scan, and sent a short questionnaire. Of these patients 1,795 did not fit the inclusion criteria for the study and were excluded.

The final sample included patients aged 40-70, with a recorded DXA at the femoral neck, given as a t-score value. Patients in the study had also not been on any osteoporotic drug therapy for more than 1 year prior to baseline.

A CONSORT diagram of how this sample was reached is given in Figure 8.1.

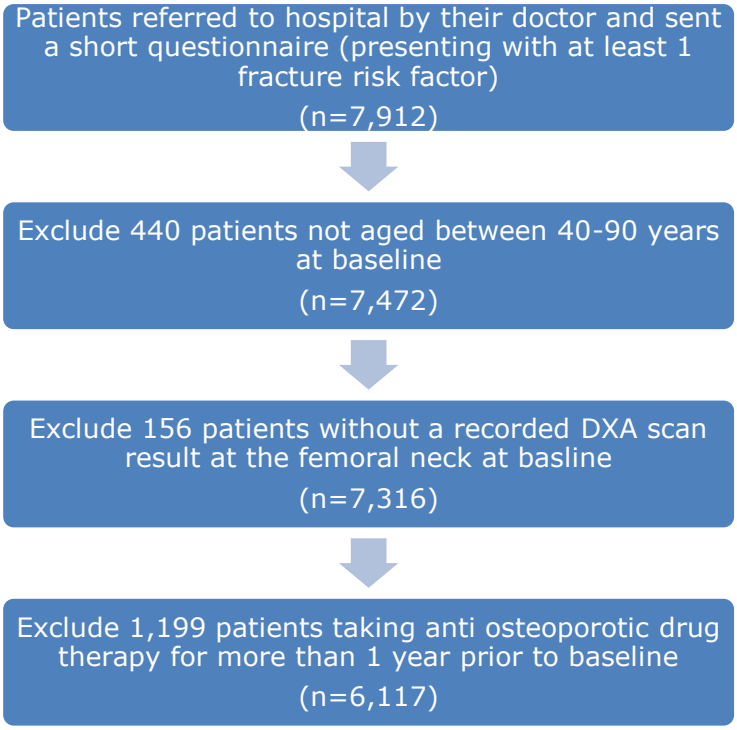


Figure 8.1 Selection of patients for the study cohort.

8.1.2 Overall Cohort Characteristics

Baseline characteristics of the overall extracted dataset are presented in Table 8.1. Of the predominantly female (79.6%) total sample, 10% were osteoporotic based on their hip DXA scan (t-score≤-2.5). A mean t-score of -1.14 (SD: 1.09) was found among all patients, but was significantly lower for females (mean= -1.19 (SD: 1.10)) than males (mean=-0.97 (SD: 0.98)) ($p<0.0001$).

The mean age among patients was 63 years (SD: 10.93), who had a mean height and weight of 1.65m (SD: 0.08) and 72.15kg (SD: 0.96), respectively. Seventy-seven percent were non-smokers and 94.8% consumed less than 3 unit of alcohol per day.

Twenty-nine percent of patients had a previous fracture at an osteoporotic site, of which 24.4% had more than one fracture. Glucocorticoids were consumed for more than 3 months by 13.1% of the total cohort; and 10.3% had been diagnosed with Rheumatoid Arthritis. Of the total cohort, 16.2% were diagnosed with secondary osteoporosis. Of these 990 patients, 74.2% of females had menopause and of which 21.2% had premature menopause (<45 years).

Table 8.1 Baseline Characteristics of the overall sample derived from the AURORA database (N=6,117).

Characteristic		No	%
Gender	Female	4,868	79.58
	Male	1,249	20.42
Osteoporotic (Hip DXA)	No	5,503	89.96
	Yes	614	10.04
Osteoporotic Status (based on UK guidelines)	Normal	2,813	45.99
	Osteopenic	2,690	43.98
	Osteoporotic	614	10.04
Previous Fracture	No	4,358	71.24
	Yes	1,759	28.76
No of Previous Fractures*	1 fracture	1,329	75.55
	2-4 fractures	392	22.29
	5+ fractures	38	2.16
Parental History of Hip Fracture	No	4,114	67.26
	Yes	2,003	32.74

[*Table continues on the next page]*

Characteristic			No	%
Current Smoking Status	Other (non/ex)		4,711	77.01
	Smoker		1,406	22.99
Alcohol Consumption	<3 units per day		5,798	94.79
	>3 units per day		319	5.21
Glucocorticoid Use	No		5,318	86.94
	Yes		799	13.06
Rheumatoid Arthritis	No		5,487	89.70
	Yes		630	10.30
Other Bone Affecting Disease	No		3,521	57.56
	Yes		2,596	42.44
Secondary Osteoporosis	No		5,127	83.82
	Yes		990	16.18
	By disease			
	Type 1 diabetes	No	5,991	97.94
		Yes	126	2.06
	Osteogenesis	No	6,117	100
		Yes	0	0
	Hypothyroidism	No	6,112	99.92
		Yes	5	0.08
	Malnutrition	No	6,113	99.93
		Yes	4	0.07
	Chronic Liver Disease	No	5,985	97.84
		Yes	132	2.16
	Menopause (Females only) **	No	1,258	25.84
		Yes	3,610	74.16
	Premature Menopause (<45years)***	No	2,845	78.81
		Yes	765	21.19

[*Table continues on the next page]*

Characteristic cont.		Mean	SD
Age (years)		62.94	10.93
Weight (kg)		72.15	15.60
	Missing (n, %)	59	0.96
Height (m)		1.65	0.08
	Missing (n, %)	192	3.14
BMI		26.39	5.07
	Missing (n, %)	198	3.24
Hip DXA T-score		-1.14	1.09

[*] **proportion out of patients with previous fracture**

[**] **proportion out of number of females in dataset**

[***] **proportion out of number of females who have had menopause**

Characteristics of the cohort were also described by osteoporotic status based on a t-score of less than or equal to -2.5 from the hip DXA scan (Table 8.2). A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

Table 8.2 Characteristics of the overall sample derived from the AURORA database, by hip DXA based osteoporotic status.

Characteristic			Normal/Osteopenic (n=5,503)		Osteoporotic (n=614)		
			No	%	No	%	p-value
Gender	Female		4,331	78.70	537	87.46	<0.001*
	Male		1,172	21.30	77	12.54	
Previous Fracture	No		4,037	73.36	321	52.28	<0.001*
	Yes		1,466	26.64	293	47.72	
No of Previous Fractures**	1 fracture		1,143	77.97	186	63.48	
	2-4 fractures		299	20.40	93	31.74	<0.001*
	5+ fractures		24	1.64	14	4.78	
Parental History of Hip Fracture	No		3,629	65.95	485	78.99	<0.001*
	Yes		1,874	34.05	129	21.01	
Current Smoking Status	Other (non/ex)		4,295	78.05	416	67.75	<0.001*
	Smoker		1,208	21.95	198	32.25	
Alcohol Consumption	<3 units per day		5,219	94.84	579	94.3	0.569
	>3 units per day		284	5.16	35	5.70	
Glucocorticoid Use	No		4,792	87.08	526	85.67	0.325
	Yes		711	12.92	88	14.33	
Rheumatoid Arthritis	No		4,935	89.68	552	89.9	0.863
	Yes		568	10.32	62	10.10	
Other Bone Affecting Disease	No		3,195	58.06	326	53.09	0.018
	Yes		2,308	41.94	288	46.91	
Secondary Osteoporosis	No		552	89.90	3195	58.06	0.179
	Yes		62	10.10	2308	41.94	
	By disease						
	Type 1 diabetes	No	5,387	97.89	604	98.37	0.428
		Yes	116	2.11	10	1.63	

[Table continues on the next page]

Characteristic cont.			Normal/Osteopenic (n=5,503)		Osteoporotic (n=614)		
			No	%	No	%	p-value
	By disease						
	Osteogenesis	No	5,503	100	614	100	-
		Yes	0	0	0	0	
	Hypothyroidism	No	5,498	99.91	614	100	0.970*****
		Yes	5	0.09	0	0	
	Malnutrition	No	5,500	99.95	613	99.84	0.970*****
		Yes	3	0.05	1	0.16	
	Chronic Liver Disease	No	5,384	97.84	601	97.88	0.941
		Yes	119	2.16	13	2.12	
	Menopause (Females only)***	No	1,111	25.65	147	27.37	0.390
		Yes	3,220	74.35	390	72.63	
	Premature Menopause (<45years)****	No	2,549	79.16	296	75.9	0.136
		Yes	671	20.84	94	24.1	
			Mean	**SD**	**Mean**	**SD**	**p-value**
Age (years)			62.07	10.62	70.75	10.61	<0.001*
Weight			73.64	15.3	58.55	11.03	<0.001*
	Missing (n, %)		38	0.69	18	2.93	
Height			1.66	0.08	1.60	0.08	<0.001*
	Missing (n, %)		135	2.45	56	9.12	
BMI			26.76	5.03	22.86	4.02	<0.001*
	Missing (n, %)		139	2.53	56	9.12	
Hip DXA T-score			-0.93	0.91	-3.05	0.54	<0.001*

* **Statistically significant differences between osteoporotic and normal/osteopenic patients.**

** **proportion out of patients with previous fracture**

*** **proportion out of respective number of females in the group**

**** **proportion out of respective number of females who have had menopause in the group**

***** **Fishers Exact test**

Statistically significant differences were found for all characteristics ($p<0.001$) but alcohol consumption ($p=0.569$), glucocorticoid use ($p=0.325$), rheumatoid arthritis ($p=0.863$) and secondary osteoporosis ($p=0.179$), which did not show associations with diagnosis of osteoporosis based on a DXA scan at the hip.

Osteoporotic patients consisted of a higher proportion of females, patients with a previous fracture, with a parental history of hip fracture, and a lower BMI. Fewer patients had gone through menopause, however of those females who had, a higher proportion of premature menopause was found, although this was not statistically significant ($p=0.136$).

8.1.3 Incidence of Fracture

Overall 471 (7.7%) fractures occurred during the follow up period of the study. None of these fractures were specifically classed as fragility fractures, however 19.3% were considered pathological/spontaneous; 68.4% were fractures at an osteoporotic site; and 12.3% were not otherwise specified. All fractures did not occur due to an external force. The highest proportion fractures at an osteoporotic site were at the forearm (26.8%) and hip (18.3%). Table 8.3 shows the characteristics of the overall cohort, by fracture outcome, during the 4 year follow-up period. A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

Table 8.3 Characteristics of the overall sample derived from the AURORA database, by fracture outcome during the 4 year follow up period.

Characteristic		No Fracture (n=5,646)		Fracture (n=471)		
		No	%	No	%	p-value
Gender	Female	4,471	79.19	397	84.29	0.008
	Male	1,175	20.81	74	15.71	
Osteoporotic Status,	No	5,133	90.91	370	78.56	<0.001*
based on Hip DXA)	Yes	513	9.09	101	21.44	
Osteoporotic Status, based on Hip DXA (UK) – detailed	Normal	2,696	47.75	117	24.84	
	Osteopenic	2,437	43.16	253	53.72	<0.001*
	Osteoporotic	513	9.09	101	21.44	
Previous Fracture	No	4,151	73.52	207	43.95	<0.001*
	Yes	1,495	26.48	264	56.05	
No of Previous Fractures**	1 fracture	1,169	78.19	160	60.61	
	2-4 fractures	302	20.20	90	34.09	<0.001*
	5+ fractures	24	1.61	14	5.30	
Parental History of Hip Fracture	No	3,774	66.84	340	72.19	0.018
	Yes	1,872	33.16	131	27.81	
Current Smoking Status	Other (non/ex)	4,365	77.31	346	73.46	0.056
	Smoker	1,281	22.69	125	26.54	
Alcohol Consumption	<3 units per day	5,363	94.99	435	92.36	0.014
	>3 units per day	283	5.01	36	7.64	
Glucocorticoid Use	No	4,904	86.86	414	87.9	0.520
	Yes	742	13.14	57	12.1	
Rheumatoid Arthritis	No	5,057	89.57	430	91.30	0.236
	Yes	589	10.43	41	8.70	
Secondary Osteoporosis	No	4,741	83.97	386	81.95	0.253
	Yes	905	16.03	85	18.05	

[Table continues on the next page]

Characteristic cont.		Mean	SD	Mean	SD	p-value
Age (years)		62.63	10.86	66.68	11.18	<0.001*
Weight (kg)		72.39	15.62	69.37	15.01	<0.001*
	Missing (n, %)	54	0.96	5	1.06	
Height (m)		1.65	0.08	1.63	0.08	<0.001*
	Missing (n, %)	169	2.99	23	4.88	
BMI		26.42	5.08	25.98	4.98	0.076
	Missing (n, %)	175	3.1	23	4.88	
Hip DXA T-score		-1.09	1.08	-1.76	1.01	<0.001*

***Statistically significant differences between patients with and without a fracture during follow-up.**

****proportion out of patients with previous fracture**

Patients who suffered a fracture during follow-up consisted of more females (84.3%, p=0.008) and had a higher mean age compared to patients without a fracture (66 years versus 62 years, p<0.001). A lower mean t-score of -1.76 (SD=1.01) was also found among patients who had a fracture during follow up (p<0.001).

A higher proportion of patients with a fracture also had had a previous fracture; consumed more than 3 units of alcohol per day, and had a lower height and weight (p<0.001). Conversely, only 27.8% of patients with a fracture had a parental history of hip fracture, compared to 33.2% in the no fracture group (p=0.018).

Statistically significant differences were not found in glucocorticoid use (more than 3 months) (p=0.520), rheumatoid arthritis (p=0.236), secondary osteoporosis (p=0.253), and BMI (p=0.076). A higher proportion of smokers was found in patients with a fracture during follow up, however this had borderline statistically significance (p=0.056).

Crude incidence rates were also calculated for the fracture outcome, and are presented in Table 8.4. Of the 471 fractured patients, 468 were included in the analysis, as 3 entered and exited the study on the same day.

Table 8.4 Crude incidence rates of the fracture outcome in the overall cohort.

Risk Factor		No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)
Age Category	40-49	29	1727.7	167.86 (116.65 to 241.55)
	50-59	103	3845.7	267.83 (220.79 to 324.89)
	60-69	135	4516.9	298.88 (252.49 to 353.80)
	70-79	132	2864.7	460.77 (388.51 to 546.48)
	80-89	68	1000.1	679.96 (536.12 to 862.39)
	90-99	1	26.6	376.51 (53.04 to 2672.85)
Osteoporotic - Hip	No	368	12641.0	291.12 (262.84 to 322.43)
	Yes	100	1340.6	745.92 (613.15 to 907.42)
Osteoporotic - Spine	No	302	10501.4	287.58 (256.91 to 321.92)
	Yes	158	3238.8	487.84 (417.41 to 570.16)
Gender	Female	395	11097.3	355.94 (322.52 to 392.83)
	Male	73	2884.3	253.09 (201.21 to 318.35)
Parental History Hip Fracture	No	338	9390.2	359.95 (323.55 to 400.44)
	Yes	130	4591.4	283.14 (238.42 to 336.24)
Current Smoker	No	343	10792.2	317.82 (285.91 to 353.30)
	Yes	125	3189.4	391.92 (328.90 to 467.02)
Alcohol Consumption more than 3 units per day	No	433	13275.6	326.16 (296.84 to 358.38)
	Yes	35	706.0	495.75 (355.94 to 690.46)
Glucocorticoid Use (3 months)	No	411	12177.6	337.51 (306.40 to 371.77)
	Yes	57	1804.1	315.95 (243.71 to 409.61)

[Table continues on the next page]

Risk Factor cont.		No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)
BMI - low (<18.5)	No	430	13243.1	324.70 (295.41 to 356.88)
	Yes	15	311.0	482.31 (290.77 to 800.02)
Rheumatoid Arthritis	No	428	12592.9	339.87 (309.15 to 373.65)
	Yes	40	1388.7	288.04 (211.28 to 392.68)
Secondary Osteoporosis	No	384	11700.0	328.21 (296.97 to 362.73)
	Yes	84	2281.7	368.15 (297.27 to 455.94)
Previous Fracture	No	207	10151.4	203.91 (177.94 to 233.67)
	Yes	261	3830.2	681.42 (603.57 to 769.32)
Previous Fracture, detail	None	207	10151.4	203.91 (177.94 to 233.67)
	1 fracture	157	2969.3	528.74 (452.17 to 618.26)
	2-4 fractures	90	794.1	1133.35 (921.80 to 1393.44)
	5+ fractures	14	66.8	2096.73 (1241.79 to 3540.26)
Total		**468**	**13,981.6**	**334.73 (305.73 to 366.47)**

The total crude fracture incidence rate was 334.73 (305.73 to 366.47) fractures per 10,000 person years (py). The highest incidence rates were found in increasing age categories (excluding age 90-99 years), and also with increasing number of previous fractures; where patients aged 80-89 years, or with 5 or more previous fractures had a crude fracture incidence rate of 679.96 (536.12 to 862.39) and 2096.73 (1241.79 to 3540.26) per 10,000 py, respectively.

Higher crude fracture incidence rate were also found with females, patients diagnosed as osteoporotic based on DXA scan results taken at the hip and spine, current smokers, patients consuming more than 3 units of alcohol per day, patients not taking glucocorticoids for 3 months, patients with a low BMI (<18.5), and secondary osteoporosis.

Conversely, lower incidence rates were found with patients diagnosed with rheumatoid arthritis and with a parental history of hip fracture.

8.1.4 Generalizability of the data

To assess the generalizability of the AURORA dataset to the UK population, the extracted cohort was compared to a sample extracted from CPRD.

The sample extracted from CPRD used a baseline date of 1st Jan 2004, for which patients were followed up for 10 years (1st Dec 2014). Patients were excluded if they were not aged between 40-90 years and if they were on any osteoporotic treatment or taking any bone sparing drugs prior to baseline. Further, patients had to have been registered with their practice for at least a year also have complete data on the 'core' fracture risk factors (age, gender, height, weight, smoking status, and alcohol consumption) within 12 months prior to baseline.

Based on this defined extraction, 60,658 patients were eligible for this study from CPRD, and its characteristics were compared to the AURORA dataset. Table 8.5 shows the level of recording for all the fracture risk factors.

Table 8.5 Comparison of level of recording in the Danish AURORA and UK CPRD extracted datasets.

Characteristics	AURORA	CPRD
Total Sample	6,117	60,658
Age	100	100
Gender	100	100
Height	96.86	72.86
Weight	99.04	72.70
BMI	96.76	72.42
Current Smoker	100	86.27
Alcohol Consumption	100	100
Parental history of hip fracture	100	100
Previous Fracture	100	100
Rheumatoid Arthritis	100	100
Glucocorticoid Use	100	100
Secondary Osteoporosis	100	100
Outcome Fracture	100	100

[*]Values for CPRD assume negative response ('No') for missing data for previous fracture, parental hip fracture, alcohol consumption, glucocorticoid use, rheumatoid arthritis, secondary osteoporosis; and fracture during follow up.

Similar levels for recording were found in both datasets, with small amounts of missing values for all risk factors. The fracture outcome was recorded in all patients for both samples. Varying and higher levels of missing data was found in CPRD for height, weight, BMI, smoking status and alcohol consumption.

Table 8.6 shows a summary of the risk factor and outcome characteristics for both datasets.

Table 8.6 Comparison of risk factors and outcome characteristics for both extracted datasets.

Characteristics		AURORA (n=6,117)		CPRD (n=60,658)		
		Mean	SD	Mean	SD	p-value
Age (years)		62.94	10.93	56.09	13.08	<0.001
Height (m)		1.65	0.08	1.68	0.10	<0.001
Weight (kg)		72.15	15.60	74.89	16.32	<0.001
BMI		26.39	5.07	26.32	4.96	0.294
		No	%	No	%	
Gender	Female	4,868	79.58	31,646	52.17	<0.001
	Male	1,249	20.42	29,012	47.83	
Previous Fracture	No	4,358	71.24	54,504	89.85	<0.001
	Yes	1,759	28.76	6,154	10.15	
Parental History of Hip Fracture	No	4,114	67.26	60,658	100.00	-
	Yes	2,003	32.74	0	0.00	
Current Smoking Status	No	4,711	77.01	20,490	33.78	<0.001
	Yes	1,406	22.99	31,842	52.49	
Alcohol Consumption greater than 3 units per day	No	5,798	94.79	44,851	73.94	<0.001
	Yes	319	5.21	15,807	26.06	
Glucocorticoid Use	No	5,318	86.94	59,019	97.30	<0.001
	Yes	799	13.06	1,639	2.70	
Rheumatoid Arthritis	No	5,487	89.70	60,151	99.16	<0.001
	Yes	630	10.30	507	0.84	
Secondary Osteoporosis	No	5,126	83.80	56,678	93.44	<0.001
	Yes	991	16.20	3,980	6.56	
Fracture during follow up	No	5,646	92.30	56,516	93.17	0.481
	Yes	471	7.70	4,142	6.83	

[*]**Values for CPRD assume negative response ('No') for missing data for previous fracture, parental hip fracture, alcohol consumption, glucocorticoid use, rheumatoid arthritis, secondary osteoporosis; and fracture during follow up.**

A higher mean age was found in the AURORA dataset compared to CPRD (62.94 years (SD: 10.93) vs. 56.09 years (SD: 13.08), p-value<0.001); and although patients from both datasets had similar BMI results; patients from the AURORA dataset were slightly shorter (1.65m (SD: 0.08) vs. 1.68m (SD: 0.10) and lighter (72.15kg (SD: 15.60) vs. 74.89kg (SD: 16.32)) than patients extracted from CPRD (p<0.001).

Patients in the AURORA dataset contained fewer males than the CPRD extraction, which had a more even spread of gender (47.83% males, p-value<0.001). A higher prevalence rate of previous fractures was also found in AURORA compared to the CPRD dataset, in which no patients had a parental history of hip fracture. In addition to this, a higher proportion of patients were taking glucocorticoids for more than 3 months, diagnosed with rheumatoid arthritis and diagnosed with secondary osteoporosis (p<0.001).

Conversely, lifestyle behaviours showed a lower proportion of current smokers and patients consuming more than 3 units of alcohol per day (p-value<0.001). Overall, the AURORA dataset had a higher proportion of incident fractures over a follow-up period of 4 years (7.7%) compared to the 6.8% of incident fractures found in CPRD, in a follow-up period of 10 years.

Figure 8.2 shows the mean values for each comparable risk factor in the AURORA and CPRD datasets.

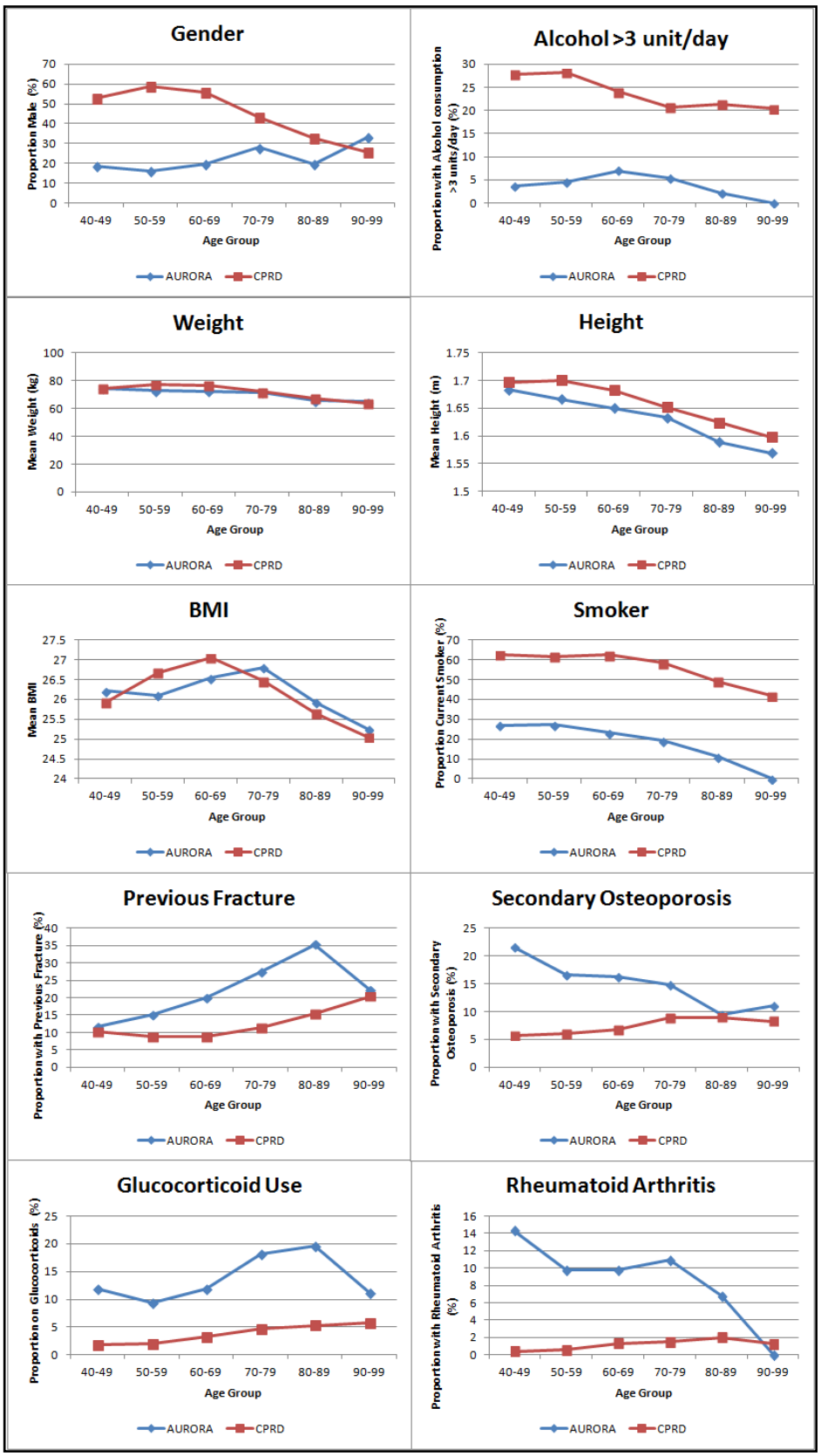


Figure 8.2 Comparison of AURORA and CPRD data characteristics, by age group.

Similar trends were found in each dataset for height, weight, and BMI across the age categories, although the highest mean BMI was found in patients aged 60-69 in CPRD and 70-79 in AURORA. A higher proportion of males was also found in all age groups, expect for patient aged 90-99 years.

The proportion of current smokers and patients consuming more than 3 units of alcohol per day remained consistent between both datasets over the age categories; elderly patients had a lower proportion of current smokers and patients consuming more than 3 units of alcohol per day in both datasets. However, the proportion of current smokers and patients consuming more than 3 units of alcohol in CPRD also remained consistently high.

Less similarity in trends of the other risk factors was found between the two datasets. Higher proportions of patients with previous fractures, diagnosed with secondary osteoporosis, diagnosed with rheumatoid arthritis, and who took glucocorticoids for more than 3 months, was found in the AURORA dataset. This remained consistent over the age groups, with the exception of patients diagnosed with rheumatoid arthritis, aged 90-99, for which more patients were found in CPRD.

Table 8.7 shows a comparison of the crude fracture incidence rates between the AURORA and CPRD extracted datasets.

Table 8.7 Comparison of the crude fracture incidence rates between the AURORA and CPRD extracted datasets.

Risk Factor		AURORA (n=6,117)			CPRD (n= 60,658)		
		No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10,000 person years (95% CI)
Osteoporosis, based on hip DXA	No	368	12641.0	291.12 (262.84 to 322.43)	-	-	-
	Yes	100	1340.6	745.92 (613.15 to 907.42)	-	-	-
Gender	Female	395	11097.3	355.94 (322.52 to 392.83)	1107	98671.5	112.19 (105.77 to 119.00)
	Male	73	2884.3	253.09 (201.21 to 318.35)	574	126284.3	45.45 (41.88 to 49.33)
Age Group	40-49	29	1727.7	167.86 (116.65 to 241.55)	298	96653.2	30.83 (27.52 to 34.54)
	50-59	103	3845.7	267.83 (220.79 to 324.89)	250	60586.3	41.26 (36.45 to 46.71)
	60-69	135	4516.9	298.88 (252.49 to 353.80)	303	35368.4	85.67 (76.55 to 95.88)
	70-79	132	2864.7	460.77 (388.51 to 546.48)	423	22054.0	191.8 (174.37 to 210.98)
	80-89	68	1000.1	679.96 (536.12 to 862.39)	393	9972.8	394.07 (356.97 to 435.02)
	90-99	1	26.6	376.51 (53.04 to 2672.85)	14	321.1	436.01 (258.23 to 736.19)
Previous Fracture	No	207	10151.4	203.91 (177.94 to 233.67)	1414	204242.1	69.23 (65.72 to 72.94)
	Yes	261	3830.2	681.42 (603.57 to 769.32)	267	20713.7	128.9 (114.33 to 145.33)

[Table continues on the next page]

Risk Factor cont.		AURORA (n=6,117)			CPRD (n= 60,658)		
		No of incident cases	**Total Person years**	**Crude Incidence Rate per 10,000 person years (95% CI)**	**No of incident cases**	**Total Person years**	**Crude Incidence Rate per 10,000 person years (95% CI)**
Parental History of Hip Fracture	No	338	9390.2	359.95 (323.55 to 400.44)	1681	224955.8	74.73 (71.24 to 78.38)
	Yes	130	4591.4	283.14 (238.42 to 336.24)	-	-	-
Current Smoking Status	No	343	10792.2	317.82 (285.91 to 353.30)	527	67986.0	77.52 (71.17 to 84.42)
	Yes	125	3189.4	391.92 (328.90 to 467.02)	873	124329.0	70.22 (65.71 to 75.03)
Alcohols Consumption greater than 3 units per day	No	433	13275.6	326.16 (296.84 to 358.38)	1312	173375.2	75.67 (71.69 to 79.88)
	Yes	35	706.0	495.75 (355.94 to 690.46)	369	51580.6	71.54 (64.6 to 79.22)
Glucocorticoid Use	No	411	12177.6	337.51 (306.40 to 371.77)	1632	219728.7	74.27 (70.76 to 77.97)
	Yes	57	1804.1	315.95 (243.71 to 409.61)	49	5227.1	93.74 (70.85 to 124.03)
Rheumatoid Arthritis	No	428	12592.9	339.87 (309.15 to 373.65)	1654	223476.7	74.01 (70.53 to 77.67)
	Yes	40	1388.7	288.04 (211.28 to 392.68)	27	1479.1	182.55 (125.19 to 266.19)
Secondary Osteoporosis	No	384	11698.7	328.24 (297.00 to 362.77)	1506	211619.4	71.17 (67.66 to 74.85)
	Yes	84	2282.9	367.95 (297.11 to 455.69)	175	13336.4	131.22 (113.15 to 152.18)
Total		**468**	**13981.6**	**334.73 (305.73 to 366.47)**	**1681**	**224955.8**	**74.73 (71.24 to 78.38)**

Of the 4,161 patients who suffered a fracture during the follow up period in the CPRD dataset, after censoring, less than half were eligible for the analysis (1,681 patients, 40.40%). For the AURORA dataset, almost all incidence fractures were eligible for the analysis (99.36%). Further, based on the define BMD readcodes, the CPRD extraction failed to capture any patients with a hip based DXA scan result; thus crude incidence rates could not be presented and compared with the AURORA dataset for this risk factor.

Accounting for censored patients, the AURORA dataset presented a higher overall crude fracture incidence rate compared to CPRD. Similar trends in the incidence rates were found between AURORA and CPRD; which were higher for females than males; increased with increasing age categories; higher among patients with a previous fracture; and higher with patients diagnosed with secondary osteoporosis.

However, crude incidence rates did not agree for current smoking status, consumption of more than 3 units of alcohol per day, glucocorticoid use, and rheumatoid arthritis; where CPRD displayed a higher fracture incidence rate for non-smokers, patients who drink less than 3 units of alcohol per day, taking glucocorticoid for more than 3 months, and with a diagnosis of rheumatoid arthritis.

8.1.5 Missing Data

The completeness of the AURORA data is given Table 8.5, which shows all risk factors are complete except for BMI, due to the missing data of height and weight. In total, 198 (3.24%) patients have incomplete data for all the FRAX based risk factors; including the new risk factor, hip DXA t-score.

Due to the small proportion of missing data (<5%) which was only found in BMI risk factor, missing data were excluded from the main analysis. A sensitivity analysis was performed including the missing data as a subgroup.

8.1.6 Derivation and Validation Cohorts

The overall cohort was split into the derivation and validation datasets; where two-thirds of the data was assigned to creating the risk models and the remaining one-third was assigned to validating them.

Two – thirds of the data was used to create 4 fracture risk prediction models:

- Model 1: BMD only
- Model 2: Standard risk factors
- Model 3: Standard fracture risk factors with BMD
- Model 4: Data driven fracture risk prediction model, including BMD

These were created using both the binary (indicating osteoporosis) and continuous form of BMD measurement (t-score). In total 7 fracture risk prediction models were created.

The derivation dataset contained 4,093 (66.9%) patients (including patients with missing data), on which the models were created. The remaining third of the overall cohort consisted of 2,024 patients (including missing data).

Table 8.8 shows the baseline characteristics for each dataset, and also found that no statistically significant differences were found in the risk factors, between the datasets. A Bonferroni adjusted significance level of 0.002 was used (0.05/24).

Table 8.8 Baseline characteristics of the derivation and validation datasets, including missing data.

Characteristic		Derivation (n=4,093)		Validation (n=2,024)	
		No	%	No	%
Gender	Female	3,266	79.79	1,602	79.15
	Male	827	20.21	422	20.85
Outcome Fracture	No	3,775	92.23	1,871	92.44
	Yes	318	7.77	153	7.56
Outcome Fracture Type*	Fragility	0	0	0	0
	Pathological/Spontaneous	62	19.50	29	18.95
	Femur/Hip	57	17.92	29	18.95
	Spine	25	7.86	14	9.15
	Forearm	86	27.04	40	26.14
	Humerus	43	13.52	22	14.38
	Pelvis	4	1.26	2	1.31
	NOS	41	12.89	17	11.11
Osteoporotic (Hip DXA)	No	3,683	89.98	1,820	89.92
	Yes	410	10.02	204	10.08
Osteoporotic Status (based on UK guidelines)	Normal	1,886	46.08	927	45.80
	Osteopenic	1,797	43.90	893	44.12
	Osteoporotic	410	10.02	204	10.08
Previous Fracture	No	2935	71.71	1,423	70.31
	Yes	1158	28.29	601	29.69
No of Previous Fractures	1 fracture	862	74.44	467	77.70
	2-4 fractures	270	23.32	122	20.30
	5+ fractures	26	2.25	12	2.00
Parental History of Hip Fracture	No	2,755	67.31	1,359	67.14
	Yes	1,338	32.69	665	32.86

[Table continues on the next page]

Characteristic cont.		Derivation (n=4,093)		Validation (n=2,024)	
		No	%	No	%
Current Smoking Status	Other (non/ex)	3,182	77.74	1,529	75.54
	Smoker	911	22.26	495	24.46
Alcohol Consumption	<3units per day	3,875	94.67	1,923	95.01
	>3 units per day	218	5.33	101	4.99
Glucocorticoid Use	No	3,577	87.39	1,741	86.02
	Yes	516	12.61	283	13.98
Rheumatoid Arthritis	No	3,686	90.06	1,801	88.98
	Yes	407	9.94	223	11.02
Other Bone Affecting Disease	No	2,382	58.20	1,139	56.27
	Yes	1,711	41.80	885	43.73
Secondary Osteoporosis	No	3,438	84.00	1,689	83.45
	Yes	655	16.00	335	16.55
	By disease				
	Type 1 diabetes	4,010	97.97	1,981	97.88
		83	2.03	43	2.12
	Osteogenesis	4,093	100	2,024	100
		0	0	0	0
	Hypothyroidism	4,089	99.90	2,023	99.95
		4	0.10	1	0.05
	Malnutrition	4,090	99.93	2,023	99.95
		3	0.07	1	0.05
	Chronic Liver Disease	4,006	97.87	1,979	97.78
		87	2.13	45	2.22
	Menopause (Females only) **	853	20.84	405	20.01
		2,413	58.95	1,197	59.14

[Table continues on the next page]

Characteristic cont.		Derivation (n=4,093)		Validation (n=2,024)	
		No	%	No	%
	By disease				
	Premature Menopause (<45years)***	1904	46.52	941	46.49
		509	12.44	256	12.65
		Mean	**SD**	**Mean**	**SD**
Age (years)		62.91	10.92	62.99	10.96
Weight (kg)		72.12	15.47	72.23	15.86
	Missing (n, %)	47	1.15	12	0.59
Height (m)		1.65	0.08	1.65	0.08
	Missing (n, %)	131	3.20	61	3.01
BMI		26.39	5.04	26.39	5.13
	Missing (n, %)	135	3.30	63	3.11
Hip DXA T-score		-1.13	1.09	-1.16	1.08

[*]out of patients with a fracture

[**]proportion out of respective number of females

[***]proportion out of respective number of females with menopause

The proportion of fractures in the derivation and validation datasets were 7.77% (n=318) and 7.56% (n=153), respectively. After censoring, this reduced to 316 and 152, respectively (inclusive of missing data). Crude fracture incidence rates, based in this are found in Table 8.9, where similar trends were found between datasets and are comparable to the overall crude incidence rates presented in Table 8.4.

For the main analysis (excluding missing data) 300 fractures (7.33%) were found out of 4,090 patients in the derivation dataset and 145 fractures (7.40%) were found out of 1960 patients in the validation dataset.

Table 8.9 Crude fracture incidence rates for the derivation and validation datasets.

Risk Factor		Derivation			Validation		
		No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)
Age Category	40-49	17	1169.9	145.31 (90.33 to 233.74)	12	557.8	215.15 (122.19 to 378.84)
	50-59	70	2534.7	276.17 (218.49 to 349.07)	33	1311.0	251.71 (178.95 to 354.06)
	60-69	93	3062.9	303.63 (247.79 to 372.06)	42	1453.9	288.87 (213.48 to 390.88)
	70-79	83	1906.0	435.46 (351.17 to 539.98)	49	958.7	511.11 (386.29 to 676.26)
	80-89	52	652.7	796.75 (607.13 to 1045.59)	16	347.4	460.56 (282.15 to 751.77)
	90-99	1	26.6	376.51 (53.04 to 2672.85)	-	-	-
Osteoporotic - Hip	No	245	8475.9	289.05 (255.03 to 327.61)	123	4165.1	295.31 (247.48 to 352.40)
	Yes	71	876.8	809.73 (641.68 to 1021.78)	29	463.8	625.28 (434.52 to 899.78)
Osteoporotic - Spine	No	191	7025.8	271.86 (235.91 to 313.28)	111	3475.6	319.37 (265.16 to 384.67)
	Yes	119	2149.6	553.59 (462.55 to 662.55)	39	1089.1	358.08 (261.63 to 490.10)
Gender	Female	266	7417.5	358.61 (318.01 to 404.40)	129	3679.8	350.56 (295.00 to 416.59)
	Male	50	1935.3	258.36 (195.82 to 340.88)	23	949.0	242.36 (161.05 to 364.71)

[Table continues on the next page]

Risk Factor cont.		Derivation			Validation		
		No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)	No of incident cases	Total Person years	Crude Incidence Rate per 10000 person years (95% CI)
Parental History Hip Fracture	No	220	6281.5	350.24 (306.88 to 399.71)	118	3108.7	379.58 (316.92 to 454.64)
	Yes	96	3071.3	312.57 (255.90 to 381.79)	34	1520.2	223.66 (159.81 to 313.02)
Current Smoker	No	240	7279.1	329.71 (290.53 to 374.18)	103	3513.1	293.19 (241.70 to 355.65)
	Yes	76	2073.7	366.5 (292.71 to 458.90)	49	1115.8	439.16 (331.91 to 581.07)
Alcohol Consumption more than 3 units per day	No	293	8875.9	330.11 (294.39 to 370.15)	140	4399.7	318.2 (269.63 to 375.53)
	Yes	23	476.9	482.33 (320.52 to 725.83)	12	229.2	523.66 (297.39 to 922.09)
Glucocorticoid Use (3 months)	No	279	8184.5	340.89 (303.15 to 383.33)	132	3993.1	330.57 (278.73 to 392.06)
	Yes	37	1168.3	316.7 (229.47 to 437.11)	20	635.8	314.57 (202.95 to 487.59)
Menopause	No	68	1962.8	346.44 (273.15 to 439.39)	29	928.9	312.19 (216.94 to 449.24)
	Yes	198	5454.7	362.99 (315.79 to 417.24)	100	2750.9	363.52 (298.82 to 442.23)
Premature Menopause (<45 years)	No	280	8175.4	342.49 (304.64 to 385.05)	127	4032.1	314.97 (264.69 to 374.81)
	Yes	36	1177.4	305.76 (220.56 to 423.89)	25	596.8	418.92 (283.07 to 619.97)
BMI - low (<18.5)	No	289	8876.1	325.59 (290.14 to 365.38)	141	4367.0	322.87 (273.75 to 380.82)
	Yes	11	187.0	588.16 (325.72 to 1062.04)	4	124.0	322.63 (121.09 to 859.62)

[Table continues on the next page]

Risk Factor cont.		Derivation			Validation		
		No of incident cases	**Total Person years**	**Crude Incidence Rate per 10000 person years (95% CI)**	**No of incident cases**	**Total Person years**	**Crude Incidence Rate per 10000 person years (95% CI)**
Rheumatoid Arthritis	No	289	8457.8	341.70 (304.49 to 383.45)	139	4135.1	336.15 (284.66 to 396.94)
	Yes	27	895.0	301.69 (206.90 to 439.93)	13	493.8	263.29 (152.88 to 453.43)
Secondary Osteoporosis	No	262	7846.5	333.9 (295.83 to 376.89)	122	3853.4	316.60 (265.12 to 378.07)
	Yes	54	1506.2	358.51 (274.58 to 468.10)	30	775.4	386.89 (270.51 to 553.34)
Previous Fracture	No	144	6832.0	210.77 (179.01 to 248.17)	63	3319.4	189.79 (148.26 to 242.95)
	Yes	172	2520.8	682.32 (587.61 to 792.31)	89	1309.4	679.69 (552.18 to 836.64)
Previous Fracture, detail	None	144	6832.0	210.77 (179.01 to 248.17)	63	3319.4	189.79 (148.26 to 242.95)
	1 fracture	105	1919.6	546.99 (451.76 to 662.29)	52	1049.8	495.36 (377.47 to 650.07)
	2-4 fractures	57	557.9	1021.62 (788.04 to 1324.45)	33	236.2	1397.28 (993.37 to 1965.44)
	5+ fractures	10	43.2	2311.27 (1243.59 to 4295.60)	4	23.5	1701.81 (638.72 to 4534.31)
Total		**316**	**9352.8**	**337.87 (302.60 to 377.25)**	**152**	**4628.8**	**328.38 (280.11 to 384.96)**

8.1.7 Derivation of Models

8.1.7.1 Exploration of risk factors

Before creating the fracture risk prediction models, associations between risk factors were assessed for potential confounding. The effect sizes of association based on linear regression for continuous risk factors (age, BMI and t-score), and logistic regression for binary risk factors are presented in Table 8.10.

Statistically significant associations were found with each risk factor ($p<0.004$), and have been highlighted in Table 8.10. Age and gender showed strong confounding effects with osteoporotic status based on hip DXA t-score (t-score<-2.5) and continuous hip DXA t-score; per year increase in age and males had 8% ($p<0.001$) higher and 57% ($p<0.001$) lower odds of being osteoporotic (t-score <-2.5). Specifically, per year increase in age reduced hip DXA t-score by 0.03 ($p<0.001$), and males had an increased t-score by 0.21 ($p<0.001$).

In addition to males; increased BMI, positive parental history of hip fracture, and diagnosis of rheumatoid arthritis also showed statistically significant associations with increased hip DXA t-score ($p<0.001$). Conversely, patients with a previous fracture, smokers, and diagnosis of secondary osteoporosis were associated with a reduction of hip DXA t-score.

Osteoporotic status (t-score<-2.5) and continuous hip DXA t-score also showed statistically significant associations with other risk factors. These were with age, gender, BMI, previous fracture, parental hip fracture, and smoking status ($p<0.001$). Continuous hip DXA t-score also showed association with rheumatoid arthritis but not with secondary osteoporosis ($p<0.003$ and 0.044, respectively).

A further breakdown of associations between risk factors is presented in Appendix L.

Table 8.10 Association between all risk factors to be included in fracture risk prediction models. Values are odd ratios, unless otherwise stated.

	Outcome											
Predictors	**Age (years)***	**Gender (males)**	**BMI***	**Previous Fracture (yes)**	**Parental History of Hip Fracture (yes)**	**Current Smoker (yes)**	**Glucocorticoid Use (yes)**	**Alcohol Consumption (>3 units/day) (yes)**	**Rheumatoid Arthritis (yes)**	**Secondary Osteoporosis (yes)**	**Osteoporotic (yes)**	**Hip DXA t-score***
Age (years)	-	1.01 (<0.001)	0.01 (0.336)	1.04 (<0.001)	0.96 (<0.001)	0.98 (<0.001)	1.02 (<0.001)	1.01 (0.215)	0.99 (0.077)	0.98 (<0.001)	1.08 (<0.001)	-0.03 (<0.001)
Gender (males)	1.64 (<0.001)	-	0.64 (0.002)	0.78 (0.008)	0.46 (<0.001)	0.97 (0.754)	3.18 (<0.001)	3.49 (<0.001)	1.54 (<0.001)	0.24 (<0.001)	0.43 (<0.001)	0.21 (<0.001)
BMI	0.03 (0.336)	1.02 (0.002)	-	1 (0.839)	0.99 (0.052)	0.94 (<0.001)	1.02 (0.008)	0.96 (0.008)	1.03 (0.009)	1 (0.683)	0.81 (<0.001)	0.09 (<0.001)
Previous Fracture (yes)	4.9 (<0.001)	0.78 (0.008)	-0.04 (0.839)	-	0.68 (<0.001)	1.05 (0.588)	0.56 (<0.001)	1.48 (0.009)	0.57 (<0.001)	1.13 (0.207)	2.9 (<0.001)	-0.58 (<0.001)
Parental History of Hip Fracture (yes)	-4.87 (<0.001)	0.46 (<0.001)	-0.33 (0.052)	0.68 (<0.001)	-	0.97 (0.703)	0.28 (<0.001)	0.8 (0.166)	0.37 (<0.001)	0.96 (0.623)	0.48 (<0.001)	0.15 (<0.001)
Current Smoker (yes)	-2.54 (<0.001)	0.97 (0.754)	-1.56 (<0.001)	1.05 (0.588)	0.97 (0.703)	-	0.85 (0.178)	3.22 (<0.001)	0.98 (0.874)	1.76 (<0.001)	1.58 (<0.001)	-0.27 (<0.001)

[Table continues on the next page]

	Outcome											
Predictors cont.	**Age (years)***	**Gender (males)**	**BMI***	**Previous Fracture (yes)**	**Parental History of Hip Fracture (yes)**	**Current Smoker (yes)**	**Glucocorticoid Use (yes)**	**Alcohol Consumption (>3 units/day) (yes)**	**Rheumatoid Arthritis (yes)**	**Secondary Osteoporosis (yes)**	**Osteoporotic (yes)**	**Hip DXA t-score***
Glucocorticoid Use (yes)	2.81 (<0.001)	3.18 (<0.001)	0.65 (0.008)	0.56 (<0.001)	0.28 (<0.001)	0.85 (0.178)	-	1.37 (0.11)	1.53 (0.003)	0.73 (0.026)	1.19 (0.265)	0 (0.953)
Alcohol Consumption (>3 units/day) (yes)	0.96 (0.215)	3.49 (<0.001)	-0.96 (0.008)	1.48 (0.009)	0.8 (0.166)	3.22 (<0.001)	1.37 (0.11)	-	0.44 (0.014)	1.15 (0.448)	1.29 (0.257)	-0.09 (0.241)
Rheumatoid Arthritis (yes)	-1.01 (0.077)	1.54 (<0.001)	0.7 (0.009)	0.57 (<0.001)	0.37 (<0.001)	0.98 (0.874)	1.53 (0.003)	0.44 (0.014)	-	0.75 (0.069)	0.96 (0.825)	0.17 (0.003)
Secondary Osteoporosis (yes)	-2.04 (<0.001)	0.24 (<0.001)	-0.09 (0.683)	1.13 (0.207)	0.96 (0.623)	1.76 (<0.001)	0.73 (0.026)	1.15 (0.448)	0.75 (0.069)	-	1.19 (0.223)	-0.09 (0.044)
Osteoporotic (yes)	8.47 (<0.001)	0.43 (<0.001)	-3.87 (<0.001)	2.9 (<0.001)	0.48 (<0.001)	1.58 (<0.001)	1.19 (0.265)	1.29 (0.257)	0.96 (0.825)	1.19 (0.223)	-	-
Hip DXA t-score	-3.4 (<0.001)	1.2 (<0.001)	2.07 (<0.001)	0.58 (<0.001)	1.14 (<0.001)	0.79 (<0.001)	1 (0.953)	0.92 (0.24)	1.16 (0.003)	0.92 (0.044)	-	-

8.1.7.2 Model Building

Five risk models were created using the derivation dataset, which consisted of 4,093 patients and 318 fractures during a 4 year follow up period. These models have been outlined above.

Excluding missing data, and accounting for censoring, 3,955 patients were included in the analysis, or which 300 had a fracture during follow-up (7.59%). The total time at risk was 9058.1 person years; the crude fracture incidence rate was 331.2 fractures per 10,000 person years.

8.1.7.2.1 Unadjusted Analysis

The association of each risk factor with the outcome was assessed. The results of the unadjusted Cox regression models are presented in Table 8.11.

Table 8.11 Unadjusted Cox regression analysis

Risk Factor			Beta Coef	p value	Hazard Ratio (95% CI)
Standard risk factors	Age (years)		0.03	<0.001	1.03 (1.02 to 1.04)
	Gender	Female	Ref		
		Male	-0.32	0.048	0.73 (0.53 to 1.00)
	BMI		-0.03	0.032	0.97 (0.95 to 1.00)
	Previous Fracture	No	Ref		
		Yes	1.22	<0.001	3.38 (2.69 to 4.24)
	Parental History of Hip Fracture	No	Ref		
		Yes	-0.14	0.282	0.87 (0.68 to 1.12)
	Current Smoker	No	Ref		
		Yes	0.13	0.384	1.12 (0.86 to 1.46)
	Alcohol Consumption (>3 units/day	No	Ref		
		Yes	0.41	0.065	1.50 (0.97 to 2.32)
	Glucocorticoid Use	No	Ref		
		Yes	-0.06	0.734	0.94 (0.66 to 1.34)
	Rheumatoid Arthritis	No	Ref		
		Yes	-0.13	0.539	0.88 (0.59 to 1.32)
	Secondary Osteoporosis	No	Ref		
		Yes	0.04	0.815	1.04 (0.77 to 1.40)
New risk factors	Osteoporotic (based on Hip DXA)	No	Ref		
		Yes	1.02	<0.001	2.79 (2.11 to 3.67)
	Hip DXA t-score		-0.59	<0.001	0.55 (0.50 to 0.61)

The unadjusted analysis shows increased hazards of fracture by 3% (HR =1.03 (95% CI: 1.02 to 1.04) per year increase in age. Males had 27% lower hazard of an osteoporotic fracture compared to females (HR =0.73 (95% CI: 0.53 to 1.00); and per unit increase in

BMI lowered the hazard of fracture by 3% (HR =0.97 (95% CI: 0.95 to 1.00). However, these were borderline statistically significant.

Previous fracture shows a strong association with a new fracture in the follow up period, with 3 times higher hazard of an osteoporotic fracture in patients with a previous fracture (HR=3.38 (95% CI: 2.69 to 4.24). Conversely, lower hazard was found with glucocorticoid use, rheumatoid arthritis, parental history of hip fracture and dichotomised hip based DXA t-score, indicating osteoporotic patients (t-score<-2.5).

Patients who had taken glucocorticoids for more than 3 months had 6% (HR=0.94 (95% CI: 0.66 to 1.34) lower hazard of fracture; patients diagnosed with rheumatoid arthritis had 12% (HR =0.88 (95% CI: 0.59 to 1.32) lower hazard of fracture; and patients with a parental history of hip fracture had 13% (HR=0.87 (95% CI: 0.68 to 1.12) lower hazard of fracture during follow-up. However, these were not statistically significant.

Patients classed as osteoporotic based on their hip DXA t-score, displayed an increased hazard of fracture during follow up which was statistically significant; compared to patients with normal/osteopenic hip DXA t-scores, osteoporotic patients have almost 3 times higher hazard of fracture (HR=2.79 (95% CI: 2.11 to 3.67). Using continuous hip DXA t-scores also yielded a similar outcome; where a 45% lower hazard (HR=0.55 (95% CI: 0.50 to 0.61)) per increment in t-score was found (increasing t-scores indicate more normal BMD in patients, as per the NOF classification: Osteoporotic (t-score≤-2.5); Osteopenic/Normal (t-score>-2.5)).

8.1.7.2.1.1 Proportional Hazards Assumption

The proportional hazards assumption was tested for each risk factor based on the unadjusted analysis. Table 8.12 shows the Chi-squared p-values testing if this assumption was met.

Table 8.12 Test of proportional hazard assumption from the unadjusted analysis.

Risk Factor	p-value ($\chi 2$)
Age (years)	0.5533
Gender	0.7264
BMI	0.7706
Previous Fracture	0.0004
Parental History of Hip Fracture	0.9124
Current Smoker	0.9734
Alcohol Consumption (>3 units/day)	0.9052
Glucocorticoid Use	0.3539
Rheumatoid Arthritis	0.8816
Secondary Osteoporosis	0.5458
Osteoporotic (based on Hip DXA)	0.8846
Hip DXA t-score	0.6146

Table 8.12 shows the proportional hazards assumption was met but all risk factors except previous fracture. Proportional hazards plots using Schoenfeld residuals were used to visually assess this assumption for all risk factors (Appendix M). Previous fracture showed to violate this assumption (Figure 8.3).

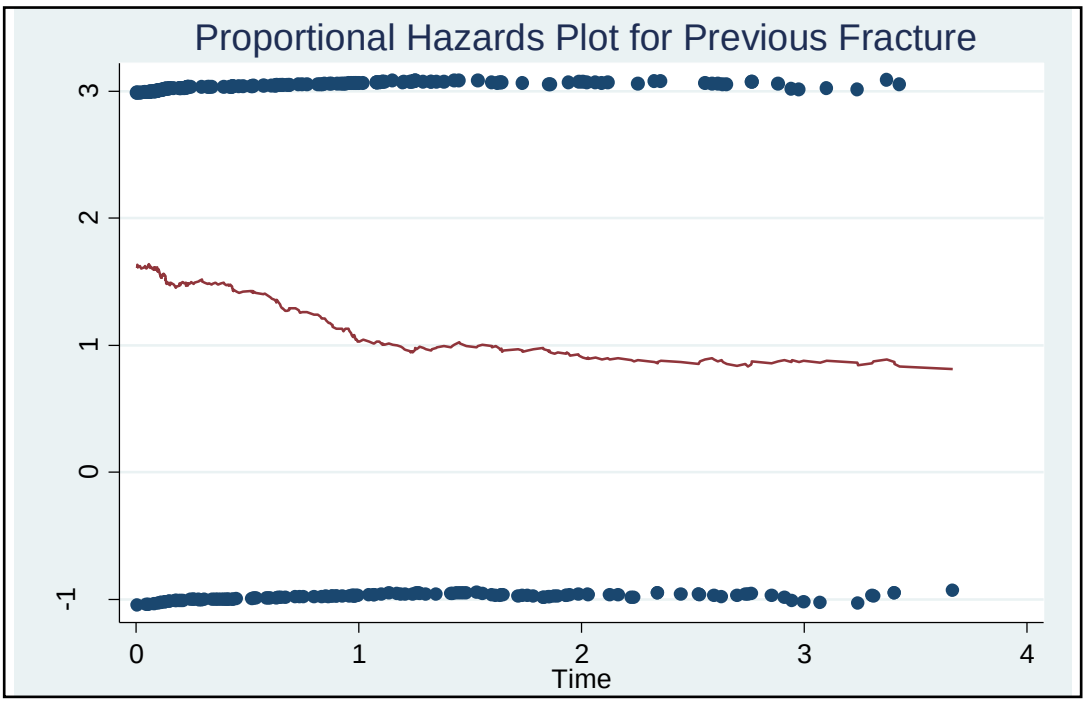


Figure 8.3 Schoenfeld Residual plot for proportional hazards of previous fracture.

Figure 8.3 also shows a reduction in the hazard of fracture based on the previous fracture risk factor, over analysis time. It can also be noticed that the reduction in hazard occurs in the first year of follow up.

Interaction terms for all risk factors with analysis time were included in the unadjusted analysis to also assess for proportional hazards. All showed insignificant association with fracture, except for previous fracture. The fixed effect of this risk factor showed a 5.5 times higher hazard ratio in patients with a previous fracture compared to those without (HR = 5.56 (95% CI: 3.86 to 8.02)). The time varying effect of previous fracture showed reduced hazard of fracture over the analysis time (HR = 0.63 (95% CI: 0.49 to 0.82). Hence, an interaction of previous fracture with analysis time is included in the following adjusted analyses.

8.1.7.2.2 Adjusted Analysis

8.1.7.2.2.1 Model 1a and 1b – BMD measurement only (categorical and continuous)

Model 1a and 1b have been presented in the unadjusted analysis (Table 8.11). Using a categorical based BMD score, indicating patients as osteoporotic (t-score≤-2.5) or normal/osteopenic (t-score>-2.5) or using BMD as a continuous t-score, gave a similar result. Using dichotomised BMD score showed an increased hazard of fracture among osteoporotic patients (HR = 2.79 (95% CI: 2.11 to 3.67)). Using continuous BMD t-score values showed a decrease of 45% (HR = 0.55 (95% CI: 0.50 to 0.61)) in hazard per increment in t-score; where increasing t-scores indicate a more normal BMD in patients, as per the NOF classification (128): Osteoporotic (t-score≤-2.5); Osteopenic/Normal (t-score>-2.5)).

8.1.7.2.2.2 Model 2 – standard risk factors only

Model 2 included the standard risk factors for fracture, excluding BMD measurement. As previous fracture does not meet the proportional hazards assumption, an interaction between previous fracture and analysis time was included in the model. The output of this model is presented in Table 8.13.

Table 8.13 Adjusted Cox regression analysis using standard fracture risk factors, excluding BMD measurement, with previous fracture as a time varying covariate.

Model 2: Standard Risk Factors only				
Risk Factor		**Adjusted Beta Coef**	**p value**	**Adjusted Hazard Ratio (95% CI)**
Age (years)		0.02	<0.001	1.02 (1.01 to 1.04)
Gender	Female	Ref		
	Male	-0.28	0.089	0.75 (0.54 to 1.04)
BMI		-0.02	0.068	0.98 (0.95 to 1.00)
Previous Fracture	No	Ref		
	Yes	1.59	<0.001	4.88 (3.37 to 7.08)
Parental History of Hip Fracture	No	Ref		
	Yes	0.08	0.562	1.08 (0.83 to 1.40)
Current Smoker	No	Ref		
	Yes	0.11	0.416	1.12 (0.85 to 1.47)
Alcohol Consumption (>3 units/day)	No	Ref		
	Yes	0.35	0.129	1.41 (0.90 to 2.21)
Glucocorticoid Use	No	Ref		
	Yes	0.08	0.674	1.08 (0.75 to 1.55)
Rheumatoid Arthritis	No	Ref		
	Yes	0.09	0.652	1.10 (0.73 to 1.65)
Secondary Osteoporosis	No	Ref		
	Yes	-0.01	0.965	0.99 (0.73 to 1.35)
Previous Fracture (TVC)		-0.45	0.001	0.64 (0.49 to 0.83)

Table 8.15 shows age and previous fracture as the strongest predictors for fracture during follow up amongst the standard fracture risk factors (p<0.001). The hazard of an osteoporotic fracture increased by 2% (HR=1.02 (95% CI: 1.01 to 1.04) per year increase of age, and was almost 5 times higher in patients with a previous fracture (HR=4.88 (95%

CI: 3.37 to 7.08). This remains consistent with the results of the unadjusted analysis. Further, the interaction term of previous fracture with time shows the hazard of fracture reduces over time with patients with a previous fracture (β=-0.4535, $p<0.001$).

Adjusting for standard risk factors, the hazard of patients who have taken glucocorticoids for more than 3 months, are diagnosed with rheumatoid arthritis, or had a parental history of hip fracture, now increase. Remaining statistically insignificant, patients taking glucocorticoids for 3 months or more have 9% (HR=1.09 (95% CI: 0.76 to 1.56) higher hazard than those who had not; patients diagnosed with rheumatoid arthritis had 10% (HR=1.10 (95% CI: 0.73 to 1.65) higher hazards; and patients with parental history of hip fracture had 8% (HR=1.08 (95% CI: 0.84 to 1.40) higher hazard of fracture.

Further, after adjusting for other standard risk factors, the effect of secondary osteoporosis reduces to a 1% lower hazard among patients diagnosed with secondary osteoporosis (HR = 0.99 (95% CI: 0.73 to 1.35).

8.1.7.2.2.3 Model 3a – standard risk factors with categorical BMD

Model 3a includes the standard fracture risk factors and interaction used in Model 2, with addition of BMD measurement in a binary format (dichotomised at t-score≤-2.5). The output of this model is presented in Table 8.14.

Table 8.14 Adjusted Cox regression analysis using standard fracture risk factors, including binary BMD measurement at the hip, with previous fracture as a time varying covariate.

Model 3a: Standard Risk Factors + BMD (categorical)				
Risk Factor		**Adjusted Beta Coef**	**p value**	**Adjusted Hazard Ratio (95% CI)**
Age (years)		0.02	0.002	1.02 (1.01 to 1.03)
Gender	Female	Ref		
	Male	-0.23	0.171	0.80 (0.57 to 1.10)
BMI		-0.01	0.368	0.99 (0.96 to 1.01)
Previous Fracture	No	Ref		
	Yes	1.54	<0.001	4.67 (3.21 to 6.78)
Parental History of Hip Fracture	No	Ref		
	Yes	0.09	0.484	1.10 (0.85 to 1.42)
Current Smoker	No	Ref		
	Yes	0.07	0.602	1.08 (0.82 to 1.42)
Alcohol Consumption (>3 units/day)	No	Ref		
	Yes	0.37	0.110	1.44 (0.92 to 2.25)
Glucocorticoid Use	No	Ref		
	Yes	0.05	0.782	1.05 (0.73 to 1.51)
Rheumatoid Arthritis	No	Ref		
	Yes	0.09	0.681	1.09 (0.72 to 1.64)
Secondary Osteoporosis	No	Ref		
	Yes	-0.03	0.827	0.97 (0.71 to 1.32)
Osteoporotic	No	Ref		
	Yes	0.56	<0.001	1.75 (1.28 to 2.38)
Previous Fracture (TVC)		-0.45	0.001	0.64 (0.49 to 0.83)

The addition of osteoporotic status, based on the hip DXA BMD measurement, as a fracture risk factor shows a statistically significant association with the outcome.

Osteoporotic patients had a 75% higher hazard of osteoporotic fracture than patients with a normal/osteopenic hip BMD (HR=1.75 (95% CI: 1.28 to 2.38)).

Age and previous fracture remained as the strong risk factors of fracture during follow-up, with similar estimates of effect. The remaining risk factors also had similar effects compared to Model 2 and continued to be statistically insignificant in the model.

Adding BMD as a binary risk factor reduced the effect of secondary osteoporosis further; a 3% reduction in hazard of fracture was found in patients with secondary osteoporosis compared to patients without (HR = 0.97 (95% CI: 0.71 to 1.32)); though the effect remained statistically insignificant.

8.1.7.2.2.4 Model 3b – standard risk factors with continuous BMD

Model 3b includes the standard risk factors for fracture and the interaction used in Model 2, with the addition of BMD in a continuous form (t-score results from hip DXA scans). The output of this model is presented in Table 8.15.

Table 8.15 Adjusted Cox regression analysis using standard fracture risk factors, including continuous BMD measurement at the hip, with previous fracture as a time varying covariate.

Model 3b: Standard Risk Factors + BMD (continuous)				
Risk Factor		**Adjusted Beta Coef**	**p value**	**Adjusted Hazard Ratio (95% CI)**
Age (years)		0.01	0.239	1.01 (1.00 to 1.02)
Gender	Female	Ref		
	Male	-0.16	0.335	0.85 (0.61 to 1.18)
BMI		0.03	0.048	1.03 (1.00 to 1.05)
Previous Fracture	No	Ref		
	Yes	1.39	<0.001	4.02 (2.76 to 5.84)
Parental History of Hip Fracture	No	Ref		
	Yes	0.10	0.447	1.11 (0.85 to 1.43)
Current Smoker	No	Ref		
	Yes	0.02	0.891	1.02 (0.77 to 1.34)
Glucocorticoid Use	No	Ref		
	Yes	0.04	0.838	1.04 (0.72 to 1.49)
Alcohol Consumption (>3 units/day)	No	Ref		
	Yes	0.38	0.099	1.46 (0.93 to 2.28)
Rheumatoid Arthritis	No	Ref		
	Yes	0.11	0.598	1.12 (0.74 to 1.68)
Secondary Osteoporosis	No	Ref		
	Yes	-0.09	0.557	0.91 (0.67 to 1.24)
Hip DXA T-score		-0.51	<0.001	0.60 (0.52 to 0.69)
Previous Fracture (TVC)		-0.44	0.001	0.64 (0.50 to 0.84)

Adding continuous hip DXA based BMD to the standard fracture risk factors also showed a statistically significant association with the outcome. A 40% lower hazard was found per increment in the BMD t-score (HR=0.60 (95% CI: 0.52 to 0.69)). This remained consistent with the unadjusted analysis of hip DXA t-score.

Previous fracture continues to be a strong predictor for fracture with a slightly lower hazard ratio than in Model 2 (HR=4.02 (95% CI: 2.76 to 5.84)). Conversely, adding continuous hip BMD measurement reduces the effect of age and causes it to become statistically insignificant (p=0.0072). A 1% increase in hazard of fracture was found per year increment in age (HR=1.01 (95% CI: 1.00 to 1.02)).

Small changes in effect sizes were found on addition of continuous hip DXA t-score for some risk factors; males had a 15% (HR=0.85 (95% CI: 0.61 to 1.18)) lower hazard, compared to females; patients with parental history of hips fractures had 11% (HR=1.11 (95% CI: 0.85 to 1.43)) increase hazard; and patients consuming more than 3 units of alcohol a day had 46% (HR=1.46 (95% CI:0.93 to 2.28)) higher hazard of fracture.

A change in direction of effect was again found for secondary osteoporosis and also BMI; compared to the unadjusted analysis and also the adjusted analysis Model 2 for BMI. Although both results remained statistically insignificant, and 9% reduced hazard was found for patient with secondary osteoporosis (HR=0.91 (95% CI: 0.67 to 1.24)) and 3% increase per unit increase of BMI (HR=1.03 (95% CI: 1.00 to 1.05)).

8.1.7.2.2.5 Model 4a and 4b – standard risk factors with categorical and continuous BMD

Based on the change in direction of effect from the unadjusted analysis and model 2 for secondary osteoporosis and BMI in Models 3a and 3b; two data driven models have been also been created. To remove the confounding effects in Model 3a and 3b on addition of categorical or continuous BMD measurement, secondary osteoporosis was removed from Model 3a to create Model 4a; and secondary osteoporosis, BMI and smoking status were removed from Model 3b to create Model 4b. Output from these models is presented in Table 8.16.

Table 8.16 Results from data driven models 4a and 4b, based on models 3a and 3b

Model 4a: Standard Risk Factors + BMD (categorical)				
Risk Factor		**Adjusted Beta Coef**	**p value**	**Adjusted Hazard Ratio (95% CI)**
Age (years)		0.02	0.001	1.02 (1.01 to 1.03)
Gender	Female	Ref		
	Male	-0.22	0.177	0.80 (0.58 to 1.11)
BMI		-0.01	0.365	0.99 (0.96 to 1.01)
Previous Fracture	No	Ref		
	Yes	1.54	<0.001	4.66 (3.21 to 6.77)
Parental History of Hip Fracture	No	Ref		
	Yes	0.09	0.479	1.10 (0.85 to 1.42)
Current Smoker	No	Ref		
	Yes	0.07	0.618	1.07 (0.82 to 1.41)
Glucocorticoid Use	No	Ref		
	Yes	0.05	0.782	1.05 (0.73 to 1.51)
Alcohol Consumption (>3 units/day)	No	Ref		
	Yes	0.36	0.111	1.44 (0.92 to 2.25)
Rheumatoid Arthritis	No	Ref		
	Yes	0.09	0.679	1.09 (0.73 to 1.64)
Osteoporotic	No	Ref		
	Yes	0.56	0.001	1.74 (1.28 to 2.38)
Previous Fracture (TVC)		-0.45	0.001	0.64 (0.49 to 0.83)
Model 4b: Standard Risk Factors + BMD (continuous)				
Age (years)		0.01	0.105	1.01 (1.00 to 1.02)
Gender	Female	Ref		
	Male	-0.15	0.382	0.86 (0.62 to 1.20)
Previous Fracture	No	Ref		
	Yes	1.41	<0.001	4.09 (2.82 to 5.95)

[Table continues on the next page]

Model 4b: Standard Risk Factors + BMD (continuous) cont.				
Risk Factor cont.		**Adjusted Beta Coef**	**p value**	**Adjusted Hazard Ratio (95% CI)**
Parental History of Hip Fracture	No		Ref	
	Yes	0.09	0.481	1.10 (0.85 to 1.42)
Glucocorticoid Use	No		Ref	
	Yes	0.04	0.823	1.04 (0.73 to 1.49)
Alcohol Consumption (>3 units/day)	No		Ref	
	Yes	0.36	0.112	1.43 (0.92 to 2.22)
Rheumatoid Arthritis	No		Ref	
	Yes	0.11	0.582	1.12 (0.75 to 1.69)
Hip DXA T-score	No	-0.44	<0.001	0.64 (0.57 to 0.72)
Previous Fracture (TVC)		-0.44	0.001	0.64 (0.50 to 0.84)

Model 4a now includes all standard risk factors, excluding secondary osteoporosis, and Model 4b has excluded BMI, smoking status, and secondary osteoporosis. Minimal change was found in the effect sizes of the included risk factors compared to Model 3a and 3b, respectively.

8.1.7.2.3 Assessment of Proportional Hazards Assumption

The proportional hazards assumptions was tested and met for each model, based on Schoenfeld residuals. Results for each predictor in each model is presented in Table 8.17, as well as results for the global proportional hazards test for each model.

Table 8.17 Results from the proportional hazard assumption test.

Model	Risk Factor	p-value
Model 1a	Osteoporotic	0.6146
	Global	0.6146
Model 1b	Hip DXA t-score	0.8846
	Global	0.8846
Model 2	Age	0.1565
	Gender	0.4050
	Previous Fracture	0.1659
	Parental History of Hip Fracture	0.9437
	Current Smoker	0.6894
	Alcohol Consumption	0.8373
	Glucocorticoid Use	0.5516
	Rheumatoid Arthritis	0.9821
	Secondary Osteoporosis	0.6338
	BMI	0.8070
	Previous Fracture (time varying)	0.1229
	Global	0.8832
Model 3a	Age	0.1399
	Gender	0.4187
	Previous Fracture	0.1755
	Parental History of Hip Fracture	0.9009
	Current Smoker	0.6045
	Alcohol Consumption	0.8625
	Glucocorticoid Use	0.5376
	Rheumatoid Arthritis	0.9793
	Secondary Osteoporosis	0.6813
	BMI	0.9102

[Table continues on the next page]

Model cont.	Risk Factor	p-value
Model 3a cont.	Osteoporotic	0.7475
	Previous Fracture (time varying)	0.1278
	Global	0.9167
Model 3b	Age	0.2853
	Gender	0.4821
	Previous Fracture	0.1591
	Parental History of Hip Fracture	0.8763
	Current Smoker	0.5705
	Alcohol Consumption	0.8836
	Glucocorticoid Use	0.5118
	Rheumatoid Arthritis	0.9262
	Secondary Osteoporosis	0.6449
	BMI	0.6736
	Hip DXA t-score	0.6639
	Previous Fracture (time varying)	0.1193
	Global	0.9140
Model 4a	Age	0.1302
	Gender	0.4459
	Previous Fracture	0.1720
	Parental History of Hip Fracture	0.9137
	Current Smoker	0.6346
	Alcohol Consumption	0.8695
	Glucocorticoid Use	0.5372
	Rheumatoid Arthritis	0.9750
	BMI	0.9164
	Osteoporotic	0.7327
	Previous Fracture (time varying)	0.1270

[Table continues on the next page]

Model cont.	Risk Factor	p-value
Model 4a cont.	Global	0.8855
Model 4b	Age	0.2302
	Gender	0.5066
	Previous Fracture	0.1582
	Parental History of Hip Fracture	0.9075
	Alcohol Consumption	0.7944
	Glucocorticoid Use	0.5254
	Rheumatoid Arthritis	0.9125
	Hip DXA t-score	0.7491
	Previous Fracture (time varying)	0.1194
	Global	0.7803

Table 8.17 shows that the proportional hazard assumption was not violated for any of the 7 risk models. Visual representations of Schoenfeld residuals to assess proportional hazards can be found in Appendix N.

8.1.7.2.4 Residuals – Goodness of fit and Outliers

Cox-Snell Residuals were checked for assess the goodness of fit for each variable. Plots of these residuals for each model can be found in Appendix N, which show a good fit for each model.

Influential points were also investigated using scatter plots of DFBETA, log-likelihood displacement values, and the LMAX statistic against the analysis time. Small proportions of patients showed to be influential, and were identified as patients at extreme risk of fracture but only a few had the outcome. These patients had increased age, higher proportions of females, smokers, and patients with a previous fracture and parental history

of hip fracture. These patients were recurring for each model and risk factor and were left in the analysis. Plots of these residuals can be found in Appendix N.

8.1.7.2.5 Model Fit

The AIC, BIC, and likelihood ratios were reported for all models. To assess the change in the likelihood ratios a Likelihood Ratio test was used. However, the likelihood ratio was tested between nested models; Model 2 with Model 3a and Model 3b. The results of these tests are presented in Table 8.18.

Table 8.18 AIC and BIC values, Royston's R^2 statistic, and Likelihood Ratio Tests to compare Model 2 to Model 3a; Model 3b, Model 4a, and Model 4b; using the derivation dataset.

Models	AIC	BIC	LR (df)	R^2 (%)	Comparison	Δ AIC	Δ BIC	Δ LR	p-value
Model 1a	4770.34	4776.62	42.89 (1)	8.54	Model 1a vs. Model 2	85.02	-36.72	-	-
Model 1b	4693.18	4699.46	120.05 (1)	23.02	Model 1b vs. Model 2	7.86	-113.88	-	-
Model 2	4685.32	4813.34	147.92 (11)	27.92	-	-	-	-	-
Model 3a	4675.79	4815.45	159.44 (12)	29.89	Model 2 vs. Model 3a	9.53	-2.11	-11.52	<0.001
Model 3b	4630.94	4770.6	204.3 (12)	37.23	Model 2 vs. Model 3b	54.38	42.74	-56.38	<0.001
Model 4a	4673.84	4801.86	159.39 (11)	29.88	Model 2 vs. Model 4a	11.48	11.48	-	-
Model 4b	4629.01	4733.75	200.23 (9)	36.58	Model 2 vs. Model 4b	56.31	79.59	-	-

Table 8.18 shows reduction in the AIC and BIC values between the single BMD measurement models (1a, 1b) and standard risk factor only model (2). In comparison to the standard risk factor only model, these values reduce further on addition of BMD measurement; categorical or continuous. The data driven models (4a and 4b) also show this effect.

Model 2 with the standard risk factors was the only model to be nested in models 3a and 3b, including the additional BMD risk factor. The likelihood ratio tests showed statistically significant improvement of fit of models with the addition of BMD measurement, categorical or continuous, to the standard fracture risk factors ($p=0.0007$, $p<0.0001$, respectively).

The R^2 values also show a higher amount of variation being accounted for in models with BMD measurement compared to models without. Further, models including BMD measurement as a continuous score accounted for more variation (R^2 for Model 3b is 37.23%) than models using BMD measurement as a binary risk factor (R^2 for Model 3a is 29.89%).

8.1.7.3 Final Fracture Risk Prediction Models

In total, 7 fracture models have been developed; the beta coefficients of which will be used to create 7 fracture risk prediction models. The beta coefficients for each model are presented in Table 8.11 (Model 1a and Model 1b), and Tables 8.13-8.16 (Model 2, 3a, 3b, 4a and 4b). These have been used with a 4-year survival rate of 0.89 to produce 7 fracture risk prediction models presented in Table 8.19.

Table 8.19 Risk equations to calculate 4 year risk based on patient characteristics for each developed risk model (1a-4b)

General Risk Equation to calculate 4 year risk	$1 - 0.89^{exp(\sum_{i=1}^{p} \beta_i X_i - \sum_{i=1}^{p} \beta_i \bar{X}_i)}$
Risk Model	**Equation**
Model 1a	where $\sum_{i=1}^{p} \beta_i X_i$ = 1.024482*osteoporosis where $\sum_{i=1}^{p} \beta_i \bar{X}_i$ = 1.024482*mean osteoporosis
Model 1b	where $\sum_{i=1}^{p} \beta_i X_i$ = -0.5912826*t-score where $\sum_{i=1}^{p} \beta_i \bar{X}_i$ = -0.5912826*mean t-score
Model 2	where $\sum_{i=1}^{p} \beta_i X_i$ = 0.0237745*age+-0.2826461*gender+-0.0225011*BMI+1.585278*previous fracture+ 0.0762559*parental hip fracture+0.1138883*smoking status+0.0773898*glucocorticoid use+ 0.3465287*alcohol consumption+0.0936966*rheumatoid arthritis+ -0.0069432*secondary osteoporosis+-0.4535108*(previous fracture*time) where $\sum_{i=1}^{p} \beta_i \bar{X}_i$ = 0.0237745*mean age+-0.2826461*mean gender+-0.0225011*mean BMI+ 1.585278*mean previous fracture+0.0762559*mean parental hip fracture+ 0.1138883*mean smoking status+0.0773898*mean glucocorticoid use+ 0.3465287*mean alcohol consumption+0.0936966*mean rheumatoid arthritis+ -0.0069432*mean secondary osteoporosis+-0.4535108*mean (previous fracture*time)
Model 3a	where $\sum_{i=1}^{p} \beta_i X_i$ = 0.0186827*age+-0.228784*gender+-0.0113651*BMI+1.540559*previous fracture+ 0.092011*parental hip fracture+0.0732564*smoking status+0.0508706*glucocorticoid use+ 0.3649544*alcohol consumption+0.0854353*rheumatoid arthritis+ -0.0346885*secondary osteoporosis+0.5568944*osteoporosis+-0.4481145*(previous fracture*time)

[Table continues on the next page]

General Risk Equation to calculate 4 year risk	$1 - 0.89^{exp(\sum_{i=1}^{p} \beta_i X_i - \sum_{i=1}^{p} \beta_i \bar{X}_i)}$
Risk Model cont.	**Equation**
Model 3a cont.	where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ 0.0186827*mean age+-0.228784*mean gender+-0.0113651*mean BMI+ 1.540559*mean previous fracture+0.092011*mean parental hip fracture+ 0.0732564*mean smoking status+0.0508706*mean glucocorticoid use+ 0.3649544*mean alcohol consumption+0.0854353*mean rheumatoid arthritis+ -0.0346885*mean secondary osteoporosis+0.5568944*mean osteoporosis+ -0.4481145*mean (previous fracture*time)
Model 3b	where $\sum_{i=1}^{p} \beta_i X_i =$ 0.0071931*age+-0.1615582*gender+0.0268478*BMI+1.39069*previous fracture+ 0.1000272*parental hip fracture+0.0192416*smoking status+0.0374944*glucocorticoid use+ 0.3774416*alcohol consumption+0.1097646*rheumatoid arthritis+ -0.0932063*secondary osteoporosis+-0.5110986*t-score+-0.4404955*(previous fracture*time) where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ 0.0071931*mean age+-0.1615582*mean gender+0.0268478*mean BMI+ 1.39069*mean previous fracture+0.1000272*mean parental hip fracture+ 0.0192416*mean smoking status+0.0374944*mean glucocorticoid use+ 0.3774416*mean alcohol consumption+0.1097646*mean rheumatoid arthritis+ -0.0932063*mean secondary osteoporosis+-0.5110986*mean t-score+ -0.4404955*mean (previous fracture*time)
Model 4a	where $\sum_{i=1}^{p} \beta_i X_i =$ 0.0187859*age+-0.2242276*gender+-0.0114226*BMI+1.539469*previous fracture+ 0.0930207*parental hip fracture+0.0697597*smoking status+0.0509313*glucocorticoid use+ 0.3635054*alcohol consumption+0.0859078*rheumatoid arthritis+0.5549807*osteoporosis+ -0.4478827*(previous fracture*time)

[Table continues on the next page]

General Risk Equation to calculate 4 year risk	$1 - 0.89^{exp(\sum_{i=1}^{p} \beta_i X_i - \sum_{i=1}^{p} \beta_i \bar{X}_i)}$
Risk Model cont.	**Equation**
Model 4a cont.	where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ 0.0187859*mean age+-0.2242276*mean gender+-0.0114226*mean BMI+ 1.539469*mean previous fracture+0.0930207*mean parental hip fracture+ 0.0697597*mean smoking status+0.0509313*mean glucocorticoid use+ 0.3635054*mean alcohol consumption+0.0859078*mean rheumatoid arthritis+ 0.5549807*mean osteoporosis+-0.4478827*mean (previous fracture*time)
Model 4b	where $\sum_{i=1}^{p} \beta_i X_i =$ 0.0095458*age+-0.1452714*gender+1.409116*previous fracture+0.0923278*parental hip fracture+ 0.0412411*glucocorticoid use+0.3579932*alcohol consumption+ 0.1144666*rheumatoid arthritis+ -0.4434596*t-score+-0.4403983*(previous fracture*time) where $\sum_{i=1}^{p} \beta_i \bar{X}_i =$ 0.0095458*mean age+-0.1452714*mean gender+1.409116*mean previous fracture+ 0.0923278*mean parental hip fracture+0.0412411*mean glucocorticoid use+ 0.3579932*mean alcohol consumption+0.1144666*mean rheumatoid arthritis+ -0.4434596*mean t-score+-0.4403983*mean (previous fracture*time)

8.1.8 Assessment of Model Performance

Performance of each model was checked individually in an internal validation, using the validation dataset. Comparisons were made between Model 2 with Model 3a and Model 3b; Model 2 and Model 4a and Model 4b to assess the added value of BMD (categorical or continuous) to fracture risk prediction.

8.1.8.1 Description of Risk

The predicted risk was plotted using kernel density plots for patients with and without a fracture during the follow up period. Figure 8.4 shows an increased density of lower risk scores for patients without a fracture than those with a fracture for all models.

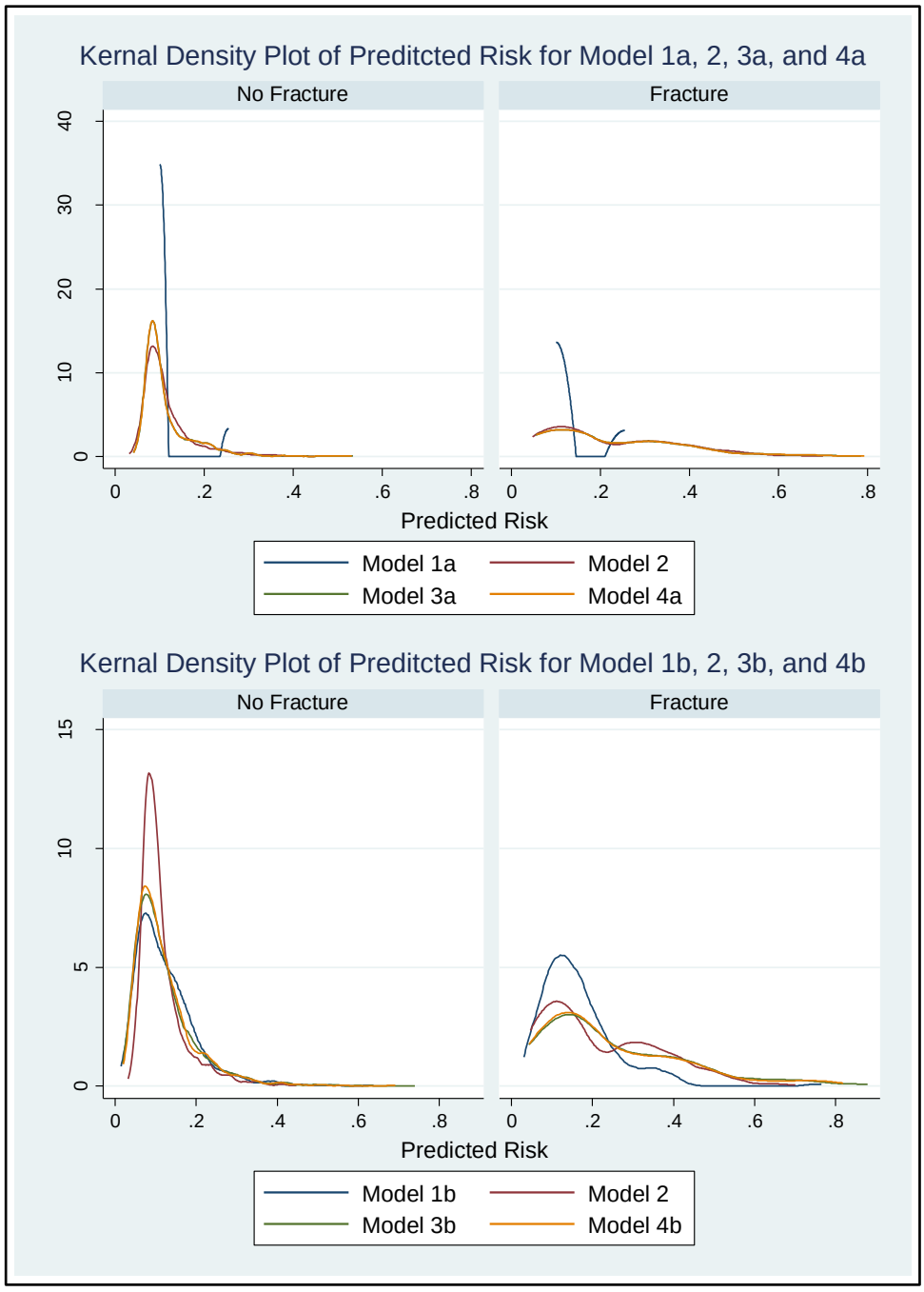


Figure 8.4 Kernel density plots for predicted risk for all models.

Investigating this further, Figure 8.5 shows models using the categorical BMD risk factor in model 3a and 4a, have an increased density of lower risk values in patients without fracture compared to model 2, but show similar density of risk scores for patients with fracture.

More disparity is shown however, when comparing models with the continuous BMD risk factor. A decreased density of lower risk values is found in patients without a fracture; and an increased density of higher risk scores are found in patients with a fracture.

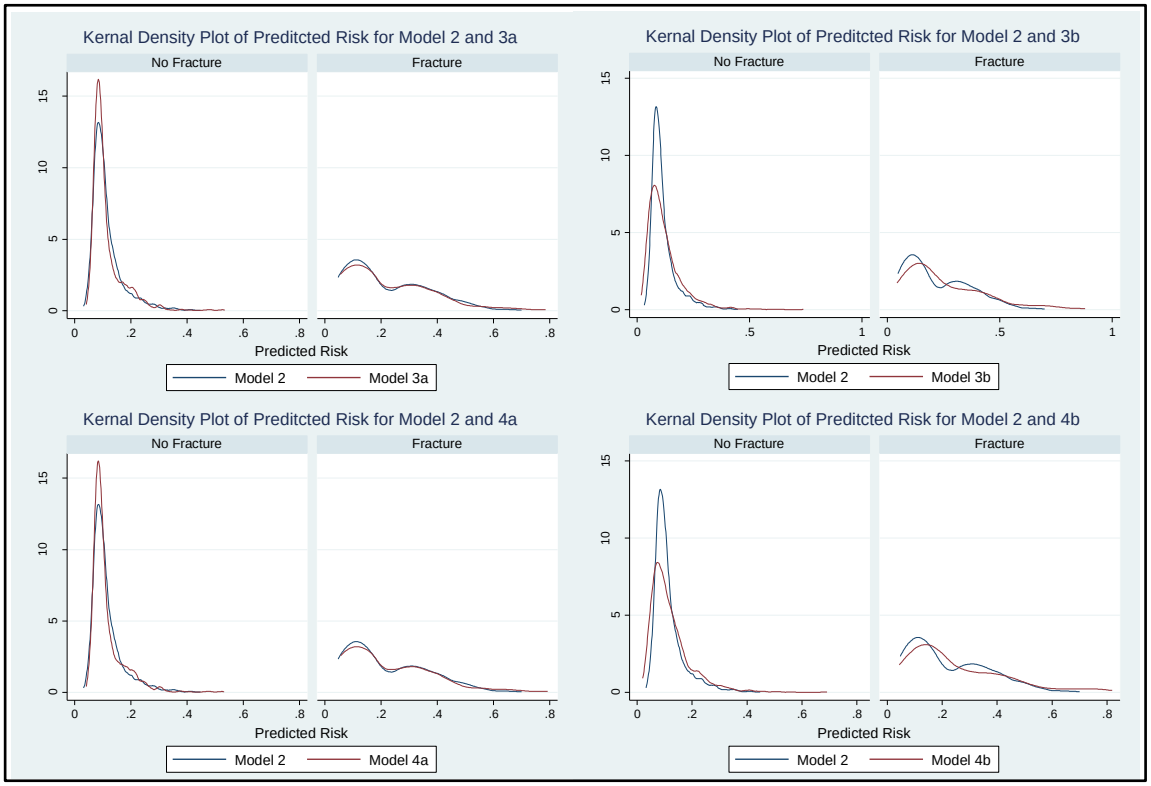


Figure 8.5 Kernel density plots of predicted risk comparing models with and without BMD information.

8.1.8.2 Calibration

8.1.8.2.1 Calibration Plots

Visual representation of observed and predicted outcomes is presented in Figure 8.6; plotting the observed and predicted event rates for each tenth of predicted risk.

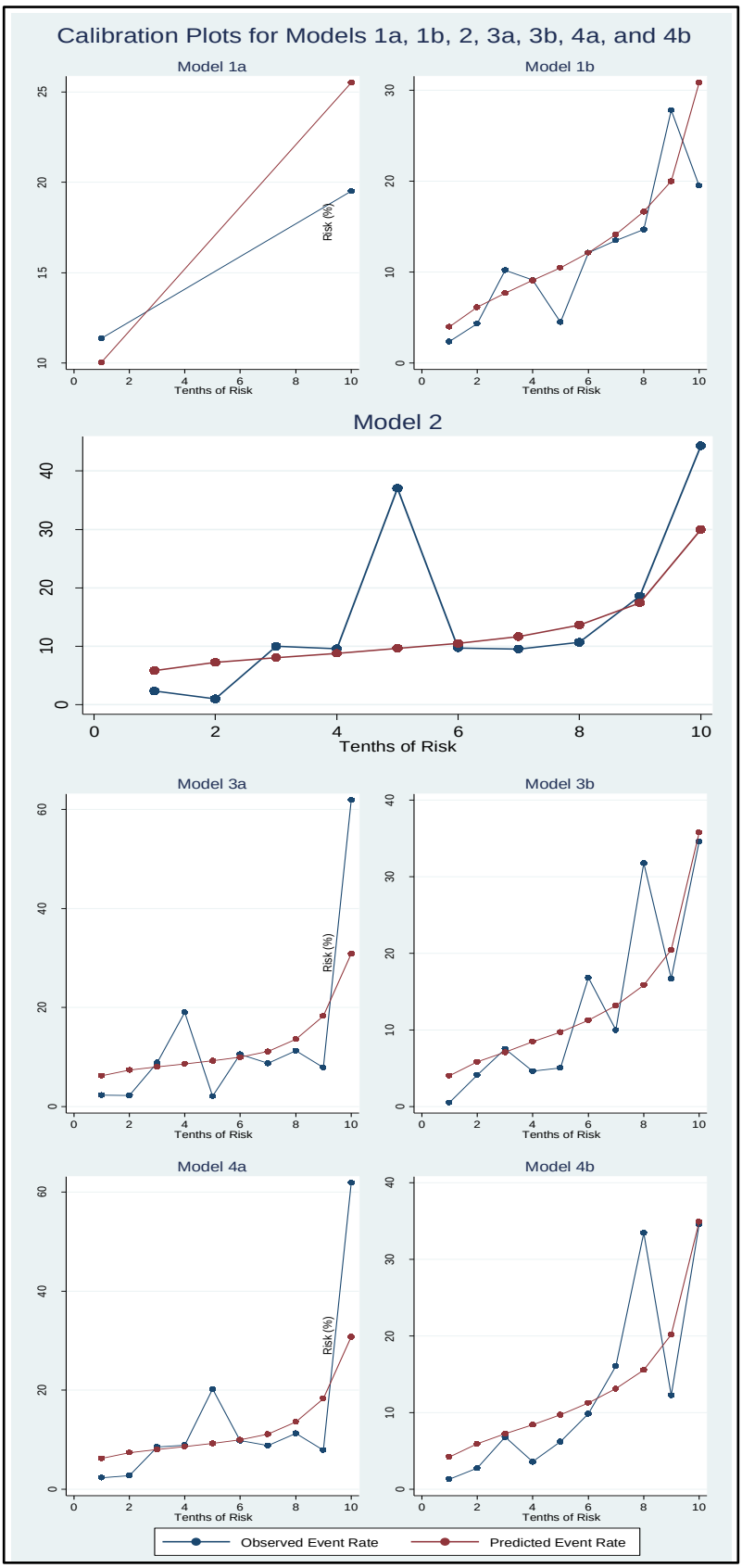


Figure 8.6 Predicted and Observed risk by 10th of predicted risk for each risk prediction model in the derivation dataset.

Figure 8.6 shows average calibration of the developed models. The calibration does not improve greatly as model risk factors are added; however, better improvement in calibration is seen when adding BMD measurement in a continuous format compared to a binary BMD risk factor.

8.1.8.3 Discrimination

8.1.8.3.1 Sensitivity, Specificity and Positive and Negative Predicted Values

The ability to identify fracture events from non-fracture events was first described by investigating the sensitivity and specificity rates, positive, and negative predicted values, for each risk model using pre-specified 8.5% risk threshold. The results from these metrics are presented in Table 8.20.

Table 8.20 Descriptive performance related information for each risk model. Values include sensitivity, specificity, positive predicted value (PPV), negative predicted value (NPV) and the % correctly classified, using the 8.5% risk threshold).

	Model 1a	Model 1b	Model 2	Model 3a	Model 3b	Model 4a	Model 4b
Mean Predicted Risk (SD)	11.5% (4.54)	12.84% (7.9)	12.27% (7.22)	12.33% (7.75)	13.17% (9.64)	12.32% (7.74)	13.06% (9.36)
Range of Risk in Sample	10-26	1-76	3-70	4-79	2-88	4-79	2-82
8.5% Cut Off							
Sensitivity	100.00%	85.62%	90.34%	86.90%	89.66%	86.21%	89.66%
Specificity	0.00%	34.10%	33.28%	35.37%	37.36%	35.26%	38.02%
Positive Predicted Value	-	9.46%	9.76%	9.70%	10.26%	9.62%	10.36%
Negative Predicted Value	-	96.72%	97.73%	97.13%	97.84%	96.97%	97.87%
% Correctly Classified	7.45%	37.94%	37.50%	39.18%	41.22%	39.03%	41.84%

An increasing trend in the mean risk and range of risk can be found between risk models with the increasing information, where mean and range of risk is highest for Model 3b, with all standard risk factors and continuous BMD measurement.

Based on the 8.5% risk threshold, the total amount of correctly classified patients was similar amongst all models, however was highest for Model 4b (42%); a data driven model with continuous BMD measurement. Adding binary BMD measurement for osteoporosis (Model 3a and 4a) reduced the sensitivity by an average of 4% and increased the specificity by 2%. However, continuous BMD measurement marginally reduced the sensitivity, providing an average 4% increase in specificity.

Minor changes were found in positive and negative predicted values on addition of BMD measurement to fracture risk prediction (binary and continuous).

8.1.8.3.2 Receiver Operating Characteristic (ROC) Curve

The ROC curve analysis was used to further assess the discriminatory accuracy of each model. The seven derived models were compared to the ROC reference line at 0.5, indicating no discriminatory ability; and were also compared to each other, with and without BMD information. Table 8.21 shows Harrell's C-Index values for each model.

Table 8.21 Harrell's C-Index for each developed risk model.

Model	Harrell's C-Index	Change in Harrell's C-Index (% change)
Model 1a	0.5446 (0.5129 to 0.5763)	0.2194 (40.28%)
Model 1b	0.6353 (0.5914 to 0.6792)	0.1287 (20.26%)
Model 2	0.7640 (0.7181 to 0.8099)	-
Model 3a	0.7592 (0.7124 to 0.8061)	-0.0048 (-0.62%)
Model 3b	0.7728 (0.7317 to 0.8139)	0.0088 (1.15%)
Model 4a	0.7587 (0.7118 to 0.8056)	-0.0053 (-0.69%)
Model 4b	0.7707 (0.7294 to 0.8120)	0.0067 (0.88%)

Small changes were found in Harrell's C-Index when adding either binary or continuous BMD measurement to standard fracture risk factors. Adding binary BMD measurement (Model 3a and 4a) to standard fracture risk factors lowered the discriminatory accuracy of model (Model 2).

Harrell's C-Index for all model have been visually presented in Figure 8.7.

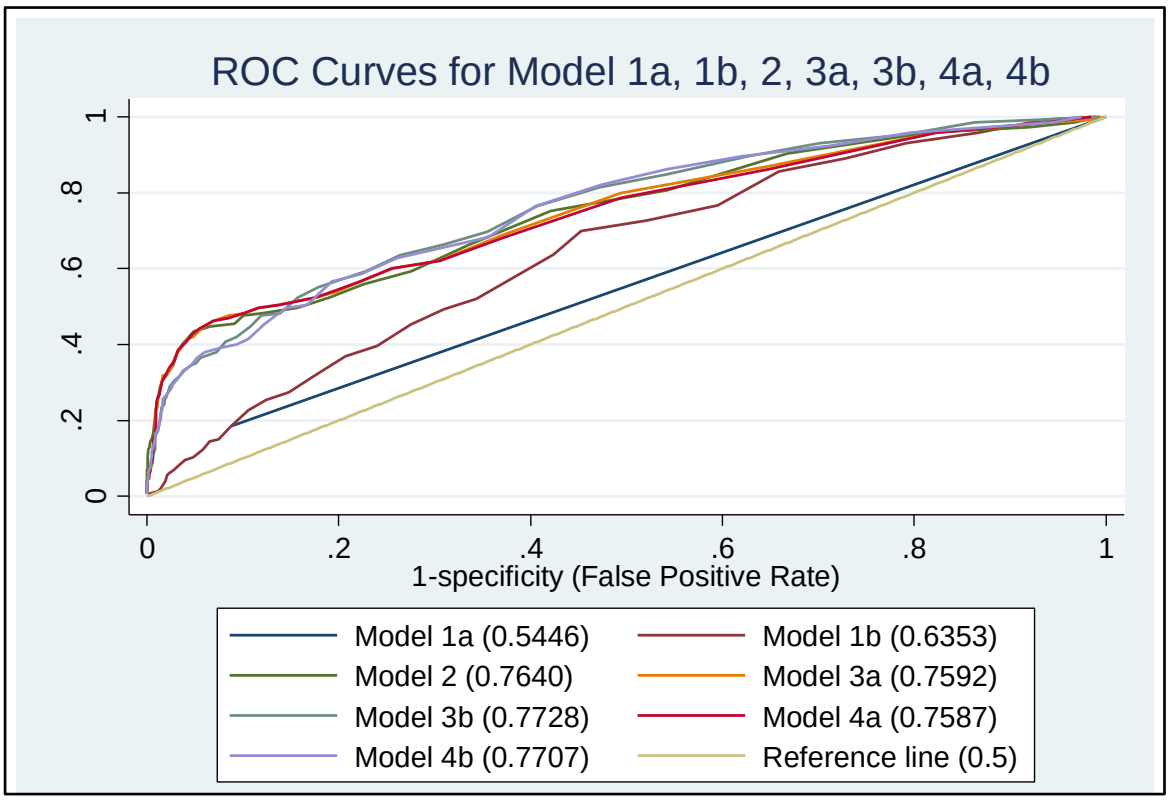


Figure 8.7 Receiver Operating Characteristic curve for each model using the derivation dataset, with related Harrell's c-index (analogous to the AUC).

Although all models performed better than chance (C-Index = 0.5) when identifying fracture events in the validation dataset; better discrimination was found with models adding continuous BMD measurement to standard fracture risk factors. The maximum 1.15% increase was found in Harrell's C-Index when adding continuous BMD measurement (Model 3b) to standard fracture risk factors (Model 2).

8.1.8.4 Reclassification

Reclassifications tables and plot were used to assess the shift of patients between categories of risk with the addition of the FH risk factor. Model 2 (without BMD measurement) was compared to Model 3a, 3b, 4a, and 4b. The pre-specified risk threshold for fracture risk was used to calculate the Net Reclassification Index (NRI) and continuous risk scores predicted by each model was used to calculate the Integrated Discrimination Improvement (IDI) (a category-less version of the NRI).

Figure 8.8 shows the reclassification plots comparing model 2 with all models including BMD measurement.

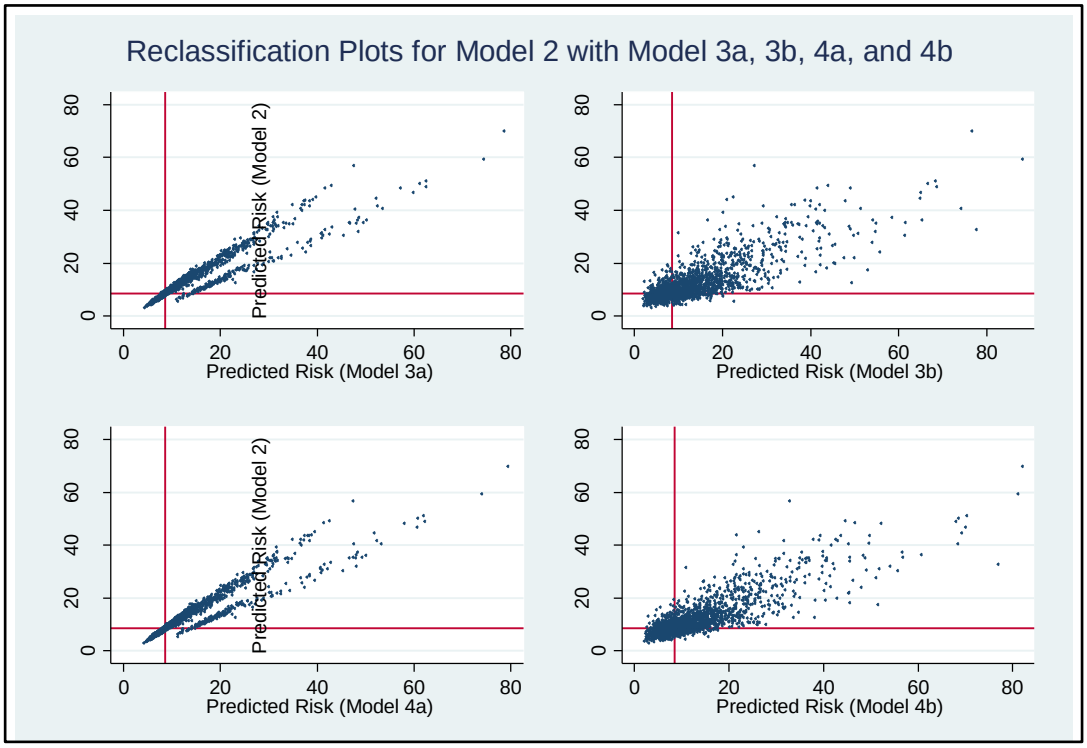


Figure 8.8 Reclassification plots comparing models with and without BMD measurement included. Reference lines indicate the 8.5% risk threshold.

Figure 8.8 shows the agreement between risk models with and without BMD measurement, with a small shift in patients to a different risk category, based on the 8.5% risk threshold. Most reclassification of patients takes place when comparing Model 2 without BMD, with models including BMD as a continuous measurement (Model 3b and 4b).

8.1.8.4.1 Comparing risk with and without BMD (binary)

8.1.8.4.1.1 Risk Reclassification – Model 2 vs. Model 3a

The reclassification table comparing Model 2 with Model 3a is given in Table 8.22. This shows that when using BMD measurement (binary) in addition to the standard fracture risk factors, 5.56% (n=109) of patients are reclassified into a lower or higher risk

category. Of the 109 patients reclassified, 33 (30%) were reclassified into a higher risk group, and 76 (70%) into a lower risk group when using BMD measurement in the risk prediction model.

Of patients classed as low risk using only standard fracture risk factors, 5% are classified into high risk using added BMD measurement; and of patients classed as high risk, 6% are reclassified into the high risk category.

Table 8.22 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 3a (standard CVD risk factors with BMD measurement), using a clinical 8.5% risk cut off.

			with BMD			
			<8.5%	≥8.5%	Total	Total No.(%) Reclassified
without BMD	<8.5%	No	585	33	618	33 (5.34%)
		%	94.66%	5.34%		
		No. Events	14	0		
		No. Non events	571	33		
		Observed Event Rate	2.39%	0.00%		
	≥8.5%	No	76	1,266	1,342	76 (5.66%)
		%	5.66%	94.34%		
		No. Events	5	126		
		No. Non events	71	1,140		
		Observed Event Rate	6.58%	9.95%		
	Total		661	1,299	1,960	109 (5.56%)

Table 8.23 shows a further breakdown of this, which presents the reclassification in patients with and without a fracture, based on the addition of using categorical BMD measurement.

Table 8.23 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 3a (standard fracture risk factors with BMD measurement), using a clinical 8.5% risk cut off.

No Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	571 94.54%	33 5.46%	604
	≥8.5%	71 5.86%	1140 94.14%	1211
	Total	642	1173	1815
Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	14 1%	0 0%	14
	≥8.5%	5 3.82%	126 96.18%	131
	Total	19	126	145

Table 8.23 shows that of the 109 patients reclassified, none were reclassified correctly into a higher risk group; but 65% were reclassified correctly into a lower risk category.

Small changes were found in the proportion of patients recommended and not recommend for treatment based on their binary BMD measurement; 68.74% would be recommend for treatment based standard fracture risk factors only and 66.28% would be recommended using BMD measurement. Similar proportions of unnecessary treatments were found using either model.

Table 8.24 shows the NRI information for this comparison. The NRI comparing model 2 with model 3a is -0.01, and the integrated discrimination index (IDI) is 0.0047.

Table 8.24 Summary of Net Reclassification Index (NRI) and Integrated Discrimination Index (IDI) for all comparisons between developed fracture risk prediction models.

Comparison	Event NRI	Non-Event NRI	Overall NRI	IDI
Model 2 vs. Model 3a	-3.45%	2.09%	-0.01	0.0047
Model 2 vs. Model 3b	-0.69%	4.08%	0.03	0.0214
Model 2 vs. Model 4a	-4.14%	1.98%	-0.02	0.0048
Model 2 vs. Model 4b	-0.69%	4.74%	0.04	0.0181

8.1.8.4.1.2 Risk Reclassification – Model 2 vs. Model 4a

Similar results were found when comparing Model 2 with the data driven Model 4a, including binary BMD measurement.

8.1.8.4.2 Comparing risk with and without BMD (continuous)

8.1.8.4.2.1 Risk Reclassification – Model 2 vs. Model 3b

The reclassification table comparing Model 2 with Model 3b is given in Table 8.25. This shows that when using continuous BMD measurement (t-score) in addition to the standard fracture risk factors, 36.99% (n=529) of patients are reclassified into a lower or higher risk category. Of the 529 patients reclassified, 227 (43%) were reclassified into a higher risk group, and 302 (57%) into a lower risk group when using BMD measurement in the risk prediction model.

Of patients classed as low risk using only standard fracture risk factors, 37% are classified into high risk using added BMD measurement; and of patients classed as high risk, 23% are reclassified into the high risk category.

Table 8.25 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 3b (standard CVD risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off.

			with BMD			
			<8.5%	≥8.5%	Total	Total No.(%) Reclassified
without BMD	<8.5%	No	391	227	618	227 (36.73%)
		%	63.27%	36.73%		
		No. Events	5	9		
		No. Non events	386	218		
		Observed Event Rate	1.28%	3.96%		
	≥8.5%	No	302	1,040	1,342	302 (22.5%)
		%	22.50%	77.50%		
		No. Events	10	121		
		No. Non events	292	919		
		Observed Event Rate	3.31%	11.63%		
	Total		693	1,267	1,960	529 (26.99%)

Table 8.26 shows a further breakdown of this, which presents the reclassification in patients with and without a fracture, based on the addition of using continuous BMD measurement.

Table 8.26 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 3b (standard fracture risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off.

No Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	386 63.91%	218 36.09%	604
	≥8.5%	292 24.11%	919 75.89%	1211
	Total	678	1137	1815
Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	5 0.357143	9 0.642857	14
	≥8.5%	10 7.63%	121 92.37%	131
	Total	15	130	145

Table 8.26 shows that of the 529 patients reclassified 9 (2%) were reclassified correctly into a higher risk group; and 55% were reclassified correctly into a lower risk category.

Small changes were found in the proportion of patients recommended and not recommend for treatment based on their continuous BMD measurement; 68.47% would be recommend for treatment based standard fracture risk factors only and 64.64% would be recommended using BMD measurement. Model 3b had a slightly smaller proportion of unnecessary treatments (0.55% reduction).

The NRI comparing model 2 with model 3b is 0.03 (Table 8.24), and the integrated discrimination index (IDI) is 0.0214.

8.1.8.4.2.2 Model 2 vs. Model 4b

The reclassification table comparing Model 2 with Model 3b is given in Table 8.27. This shows that when using continuous BMD measurement (t-score) in addition to the data driven standard fracture risk factors, 23.62% (n=463) of patients are reclassified into a lower or higher risk category. Of the 463 patients reclassified, 188 (41%) were reclassified into a higher risk group, and 275 (59%) into a lower risk group when using BMD measurement in the risk prediction model.

Of patients classed as low risk using only standard fracture risk factors, 30% are classified into high risk using added BMD measurement; and of patients classed as high risk, 20% are reclassified into the high risk category.

Table 8.27 Risk Reclassification Table comparing Model 2 (standard fracture risk factors alone) to Model 4b (data driven standard CVD risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off.

			with BMD			
			<8.5%	≥8.5%	Total	Total No.(%) Reclassified
without BMD	<8.5%	No	430	188	618	188 (30.42%)
		%	69.58%	30.42%		
		No. Events	7	7		
		No. Non events	423	181		
		Observed Event Rate	1.63%	3.72%		
	≥8.5%	No	275	1,067	1,342	275 (20.49%)
		%	20.49%	79.51%		
		No. Events	8	123		
		No. Non events	267	944		
		Observed Event Rate	2.91%	11.53%		
	Total		705	1,255	1,960	463 (23.62%)

Table 5.28 shows a further breakdown of this, which presents the reclassification in patients with and without a fracture, based on the addition of using continuous BMD measurement.

Table 8.28 Risk Reclassification Table, by fracture outcome; comparing Model 2 (standard fracture risk factors alone) to Model 4b (data driven standard fracture risk factors with continuous BMD measurement), using a clinical 8.5% risk cut off.

No Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	423 70.03%	181 29.97%	604
	≥8.5%	267 22.05%	944 77.95%	1211
	Total	690	1125	1815
Fracture		**with BMD**		
		<8.5%	≥8.5%	Total
without BMD	<8.5%	7 0.5	7 0.5	14
	≥8.5%	8 6.11%	123 93.89%	131
	Total	15	130	145

Table 8.28 shows that of the 463 patients reclassified 7 (2%) were reclassified correctly into a higher risk group; and 58% were reclassified correctly into a lower risk category.

Small changes were found in the proportion of patients recommended and not recommend for treatment based on their continuous BMD measurement; 68.47% would be recommend for treatment based standard fracture risk factors only and 64.03% would be recommended using BMD measurement. Model 4b had a slightly smaller proportion of unnecessary treatments (0.66% reduction).

The NRI comparing model 2 with model 4b is 0.04, and the integrated discrimination index (IDI) is 0.0181 (Table 8.24).

8.1.8.5 Decision Curve Analysis

A Decision Curve Analysis (DCA) allows new risk factors to be assessed by comparing treatment strategies, which includes treat all; treatment based on each fracture risk model; and treat all. It identifies the strategy which leads to the highest Net Benefit (NB); the biggest increase in the proportion of true positives without a change in the proportion of false positives. This is used to assess improvement by looking at change in net benefit at the 8.5% risk threshold.

Table 8.9 presents the change in net benefit for each comparison considered in the study.

Figure 8.9 Net Benefit for each treatment strategy using an 8.5% risk threshold.

Model	Net Benefit	Change in Net Benefit (compared to Model 2)
Model 2	-0.0512	-
Model 3a	-0.0496	0.0016
Model 3b	-0.0477	0.0035
Model 4a	-0.0498	0.0014
Model 4b	-0.0472	0.0040

Table 8.9 shows very small changes in net benefit based on an 8.5% risk threshold; the biggest increase was found for model including continuous BMD measurement in addition to data driven standard fracture risk (Model 4b) factors; an absolute change of 0.004 was found.

8.1.9 Sensitivity Analysis

Sensitivity analyses were performed by treating missing values as a category for each affected risk factor; no change in results was found.

8.1.10 Section Summary

This section presented the results from the data extraction from AURORA; described the baseline characteristics of the cohort and investigated trends in the incidence of fragility fracture. Cox regression was used to model the data and create seven fracture risk prediction models; including BMD measurement alone in a binary and continuous format (Model 1a and Model 1b, respectively); standard fracture risk factors only (Model 2); standard fracture risk factors with BMD measurement in binary and continuous format (Model 3a and Model 3b, respectively); and data driven standard fracture risk factors with BMD measurement in binary and continuous format (Model 4a and Model 4b, respectively).

The derived models were then validated using current methodology and performance measures describing calibration, discrimination, and reclassification. The section discusses these findings.

8.2 Discussion

Bone mineral density is used to diagnose osteoporosis, and is a recognised risk factor for fragility fractures (153-155); it is embedded in the FRAX risk prediction model; and is treated to slow down bone loss and reduce the risk of fracture (6). However, it is only optionally used in fracture risk assessment in the UK (2). NICE has recommended the added value of BMD to fracture risk prediction be evaluated (2). The aim of this study was to identify and quantify the added value BMD measurement contributes to fracture risk prediction; adding it to standard fracture risk factors.

This was achieved by recalibrating a FRAX based risk prediction model using a bespoke prospective cohort. Models were created including the standard fracture risk factors, with and without BMD measurement (in a binary and continuous format). Models were then validated using current and refined methodology. The previous section presented the results from the current methodology, which is discussed in this section.

8.2.1 Summary of findings

Bone mineral density measurement showed a strong association with fracture throughout the analysis. This was found with both the binary form of BMD (to indicate osteoporosis) and continuous; both showed an increased risk of fracture with more severe bone loss.

The univariate analysis using Cox regression also showed increased hazard of fracture with increasing age, females, decreasing BMI, previous fracture, smokers, patients who consume more than 3 units of alcohol per day, secondary osteoporosis, osteoporosis, and decreasing bone mineral density; this was consistent with the literature (131-134).

Further, previous fracture was also found to be a time varying covariate in the analysis, due the risk of subsequent fracture reducing over time. This is consistent with literature,

as the risk of a subsequent fracture falls after the first year post previous fracture (135-137).

However, the univariate analysis also showed decreased risk of fracture amongst patients with a parental hip fracture, diagnosed with rheumatoid arthritis, and patients taking glucocorticoids for more than 3 months. These associations were not consistent with the literature (13, 164). On further investigation, a decreased age and/or a lower proportion of previous fractures, was also found in these subgroups (Appendix L); patients with a parental hip fracture had a 4 times lower mean age and 32% lower odds of a previous fracture compared to patients without; patients with rheumatoid arthritis had a 43% lower odds of a previous fracture compared to patients without rheumatoid arthritis; and patients taking glucocorticoids for more than 3 months had a 44% lower odds of a previous fracture.

Both age and a previous fracture were strong risk factors for fracture risk and dominated the analysis. Thus, adjusting for age and previous fracture showed an increased risk of fracture was found in patients with a parental hip fracture, diagnosed with rheumatoid arthritis, and patients taking glucocorticoids for more than 3 months.

Further, including BMD measurement in the adjusted risk models also changed the direction of effect for secondary osteoporosis; indicating a confounding effect. This was found when exploring the associations between risk factors, presented in Table 8.10 and Appendix L; and 19% increased odds of osteoporosis was found in patients with secondary osteoporosis, and a -0.09 lower t-score.

Beyond the regression analysis, which showed increased fracture risk with BMD measurements, the validation analysis did not find binary BMD measurement (to indicate osteoporosis) to add any benefit fracture risk prediction; particularly in discriminatory

accuracy and reclassification. A poorer model performance was found compared to using standard fracture risk factors alone.

Added value of BMD measurement to fracture risk prediction was found when included as a continuous risk factor; thus the extent of bone loss at the hip was the primary driver for the added performance between risk models with and without BMD measurement. This was found with increased discriminatory accuracy and good reclassification of patients into their respective risk categories. Although comparison of these results have been difficult as studies have not directly compared models with and without BMD measurement (147, 165-167); consistent findings were found of better model performance using continuous BMD measurement.

8.2.2 Strengths and limitations

This study is the first to explicitly assess the added value of BMD measurement to fracture risk prediction through direct comparison; showing the effects of BMD on the calibration, discrimination, and reclassification of the model.

The data used to evaluate the added value of BMD measurement in UK fracture risk assessment was derived from a Danish population. This limited the analysis with regard to its generalisability to a UK population. However, preliminary investigation into using routinely collected data from UK primary care showed BMD measurement was inadequately recorded. The robustly collected, and validated, AURORA dataset overcame this limitation with minimal missing data and also provided continuous BMD t-score measurement for each patient.

The AURORA data was also linked to the Danish Patient Registry. This provided access to additional risk factors data as well as incident fractures during the follow up period. In particular, additional previous fracture data was also found and used to detail the number

of fractures each patient had. Further, the linked data was also used to retrieve data for risk factors not collected in the AURORA dataset, such as rheumatoid arthritis and secondary osteoporosis.

The AURORA data was collected from patients who presented with to their doctor with at least one fracture risk factor; and were a referred population. This limits and affects the generalisability of the results to a UK population in primary care, which would be a non-referred population. Thus, in hindsight assessing the generalisability of the AURORA dataset to a sample extracted from CPRD is not comparable due to the potential reasoning behind the selection of patients for assessment of fracture risk and bone mineral density.

Further, collecting data from a referred population also led to a biased study sample with a higher risk of a fracture and increased age which may lead to an overestimation of fracture risk amongst patients in the UK.

Further, due to the increased age of the sample, death becomes a competing risk. However, information on death was not collected and could not be retrieved. This limited the analysis of the data as competing risks could not be accounted for which may lead to an overestimation of fracture risk (168). However, as the study compares risk models derived using the same methods, this would not affect the difference in the performance measures described in the analysis; balancing the bias.

The study cohort was followed up for a maximum of 4 years, which was used to create 4 year fracture risk prediction models. This made it less comparable to other risk models that produce 10 year fracture risk estimates. The survival estimate of the study could not be extrapolated as Cox regression was used to derive the risk models; which does not assume any underlying distribution for the baseline hazard function. This meant the 20% (128, 129) pre-defined risk threshold used for 10 year fracture risk estimate was not

applicable to the developed models. In this case, a 4 year risk threshold of 8.5% was used, based on the 10 year risk threshold of 20% using a compounded interest method.

A complete case analysis was performed for the main analysis of the study as a small amount (<5%) of missing data was found for one variable of the analysis. It has been shown using list wise deletion of missing data >20% would produce biased results. As the amount of missing data falls much lower than this threshold, a complete case analysis would not lead to biased results (169, 170). Further, a sensitivity analysis including the missing data as a category did not alter the results.

Another limitation for the study was performing an internal validation, which uses the data from the same population to develop and validate the fracture risk models. This may lead to over optimistic results of the performance of the risk models. To account for this limitation, a commonly practised method which randomly assigns patients to the derivation and validation datasets was used; further, a similar 1:3 ratio was also used to split the data (11, 31, 83).

8.3 Chapter Summary

This chapter presented and discussed the results of the current methodology used to evaluate the added value of BMD measurement to fracture risk prediction. This included a description of the extraction, the characteristics and generalisability of the sample, and the development of the risk models.

The next chapter will present and discuss the results of the refined methodology used to evaluate the added value of BMD measurement to fracture risk prediction.

Chapter 9 - Bone Mineral Density and Fracture Risk: Results and Discussion (Refined Methodology)

The previous chapter presented the extracted cohort from AURORA; it described the baseline characteristics; and also looked at the trends in incidence rates of osteoporotic fractures. Missing data was excluded for the analysis and 7 fracture risk prediction models were created using the derivation dataset (Chapter 8, Section 8.1.7.3). Bone mineral density (BMD) was added to the standard and data driven standard risk factors, in both a binary (to indicate osteoporosis) and continuous format.

These models were then validated using the validation dataset. Current methodology was used to assess each model and evaluate the added value of BMD measurement information, these included measures of discrimination and reclassification.

In this chapter, evaluation of added value of BMD to fracture prediction is continued; using refined methodology; a cost-effectiveness analysis with a decision tree framework.

9.1 Results

9.1.1 Cost-Effectiveness Analysis

A decision tree was used for the cost-effectiveness analysis, which compared risk prediction models with BMD information to models without. Model 2, with standard fracture risk factors only, was compared to Model 3a (with binary BMD), Model 3b (with continuous BMD), Model 4a (data driven with binary BMD), and Model 4b (data driven with continuous BMD). The decision tree used for this approach is presented in Chapter

7, Section 7.2.6.1.5.1, Figure 7.1, and was used for both the deterministic and probabilistic cost effectiveness analyses.

The Youden Index and Harrell's C-Index, measures of discriminatory accuracy, were used as measures for the performance of each model, and the change between these measures were used to describe the added value BMD measurement.

9.1.1.1 Deterministic Cost-Effectiveness Analysis

A deterministic cost-effectiveness analysis was performed for all comparisons described above. This used the sensitivity and specificity for each model using an 8.5% risk cut off. A summary of values used in the decision tree to compare Model 2 with Model 3a, 3b, 4a, and 4b are presented in Table 9.1. Values assume the base case for the analysis, described in Chapter 3, Section 3.3. In their respective comparisons, Model 2 is termed as the original model (without BMD), and Model 3a, 3b, 4a, and 4b are the modified models (with BMD).

Table 9.1 Summary of values inputted into the decision tree to compare Model 2 to Models 3a, 3b, 4a, and 4b.

Inputs	Model 2	Model 3a	Model 3b	Model 4a	Model 4b
P(D+)			0.0745 (0.0042)		
P(D-)			0.9255 (0.0042)		
P(T+)	0.6847 (0.0105)	0.6628 (0.0107)	0.6464 (0.0108)	0.6633 (0.0107)	0.6403 (0.0108)
P(T-)	0.3153 (0.0105)	0.3372 (0.0107)	0.3536 (0.0108)	0.3367 (0.0107)	0.3597 (0.0108)
Sensitivity: P(T+|D+)	0.9034 (0.0246)	0.869 (0.0281)	0.8966 (0.0254)	0.8621 (0.0287)	0.8966 (0.0254)
P(T+|D-)	0.6672 (0.0111)	0.6463 (0.0112)	0.6264 (0.0114)	0.6474 (0.0112)	0.6198 (0.0114)
Specificity: P(T-|D-)	0.3328 (0.0111)	0.3537 (0.0112)	0.3736 (0.0114)	0.3526 (0.0112)	0.3802 (0.0114)
P(T-|D+)	0.0966 (0.0246)	0.131 (0.0281)	0.1034 (0.0254)	0.1379 (0.0287)	0.1034 (0.0254)
Cost of Original Model (without BMD)			£45.63		
Cost of Modified Model (with BMD)			£114.63		
Cost of Low Risk - untreated			£574.87		
Cost of Low Risk - treated			£874.49		
Cost of High Risk - untreated			£1,074.27		
Cost of High Risk - treated			£943.72		

Figure 9.1 shows an example of the decision tree comparing Model 2 to Model 3a.

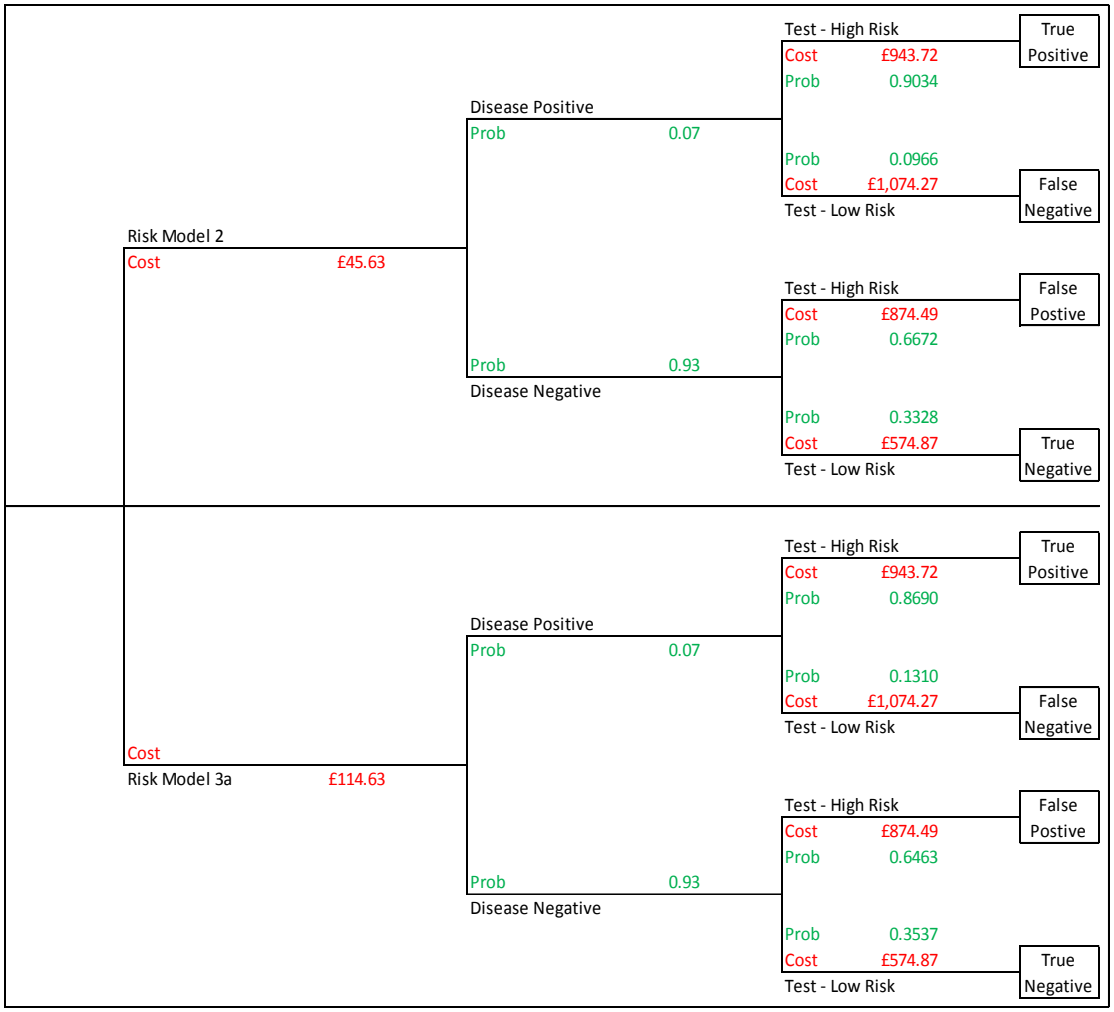


Figure 9.1 Decision tree to compare Model 2 (standard risk factors only) to Model 3a (standard risk factors with binary BMD measurement).

9.1.1.1.1 Model 2 vs. Model 3a (SRF vs. SRF + binary BMD)

Comparing Model 2 and Model 3a, the expected cost of calculating fracture risk based on standard risk factors only is £833.93 per person. Adding binary BMD measurement to this increases the expected cost per person to £897.46. Table 9.2 provides a summary of the output produced by the decision tree.

Table 9.2 Output from each decision tree comparing models with BMD measurement to models without.

		Model 2	Model 3a	Model 3b	Model 4a	Model 4b
True Positive $P(D+ \cap T_n+)$	Probability	0.0673	0.0647	0.0667	0.0642	0.0667
	Proportion, per 1000	67.3	64.7	66.7	64.2	66.7
	Cost	£989.35	£1,058.35	£1,058.35	£1,058.35	£1,058.35
False Positive $P(D- \cap T_n+)$	Probability	0.6175	0.5982	0.5798	0.5992	0.5737
	Proportion, per 1000	617.5	598.2	579.8	599.2	573.7
	Cost	£920.12	£989.12	£989.12	£989.12	£989.12
False Negative $P(D+ \cap T_n-)$	Probability	0.0072	0.0098	0.0077	0.0103	0.0077
	Proportion, per 1000	7.2	9.8	7.7	10.3	7.7
	Cost	£1,119.90	£1,188.90	£1,188.90	£1,188.90	£1,188.90
True Negative $P(D- \cap T_n-)$	Probability	0.3080	0.3274	0.3457	0.3264	0.3519
	Proportion, per 1000	308	327.4	345.7	326.4	351.9
	Cost	£620.50	£689.50	£689.50	£689.50	£689.50
	Expected Cost of Test	£833.93	£897.46	£891.69	£897.83	£889.86
	Expected Youden Index	0.2362	0.2227	0.2701	0.2147	0.2767

Adding binary BMD to standard fracture risk factors lowers the proportion of patients who would be recommended treatment based on their risk and 8.5% treatment threshold. Comparing Model 2 and Model 3a; 68% or 66% of patients would be recommended preventative treatment, respectively. Of these, the same proportion of patients would be treated unnecessarily (90%).

Conversely, adding binary BMD to standard fracture risk factors marginally increases the number of patients who would be considered low risk based on their fracture risk factors; Model 2 would predict 32% to be low risk compared to 34% when using BMD measurement. Of these patients, 2% would be a missed diagnosis using Model 2, compared to 3% using Model 3a.

9.1.1.1.1.1 Incremental Cost-Effectiveness Ratio (ICER)

Based on Table 9.2, the incremental cost of using Model 3a is £63.53; a 6% decrease was also found in the Youden Index. Harrell's C-Index for each model has been calculated in Chapter 8 Section 8.1.8.3.2; a C-Index of 0.7640 (95% CI: 0.7181 to 0.8099) and 0.7592 (95%CI: 0.7192 to 0.8061) was found for Model 2 and Model 3a, respectively, giving a 0.62% decrease in Harrell's C-Index between models. A summary of the incremental costs and effects are presented in Table 9.3.

Table 9.3 Results from the deterministic cost-effectiveness analysis evaluating the added value of BMD information to fracture risk prediction.

Comparison	Change in Cost	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Model 2 vs. Model 3a	£63.53	-5.73%	-£1,107.88	-0.62%	-£10,200.14
Model 2 vs. Model 3b	£57.76	14.34%	£402.80	1.15%	£5,018.85
Model 2 vs. Model 4a	£63.90	-9.12%	-£700.67	-0.69%	-£9,206.31
Model 2 vs. Model 4b	£55.93	17.14%	£326.32	0.88%	£6,359.25

A negative ICER of £1,107.88 per percentage increase in Youden Index was found, which coincided with a negative ICER for the percentage increase in Harrell's C-Index (-£10,200.15). These have been plotted on a cost-effectiveness plane (Figure 9.2). Both measures lie in the north-west quadrant of the plot and show a decreased effect for an increased cost. Model 3a is dominated by Model 2 with regard to both discriminatory accuracy measures.

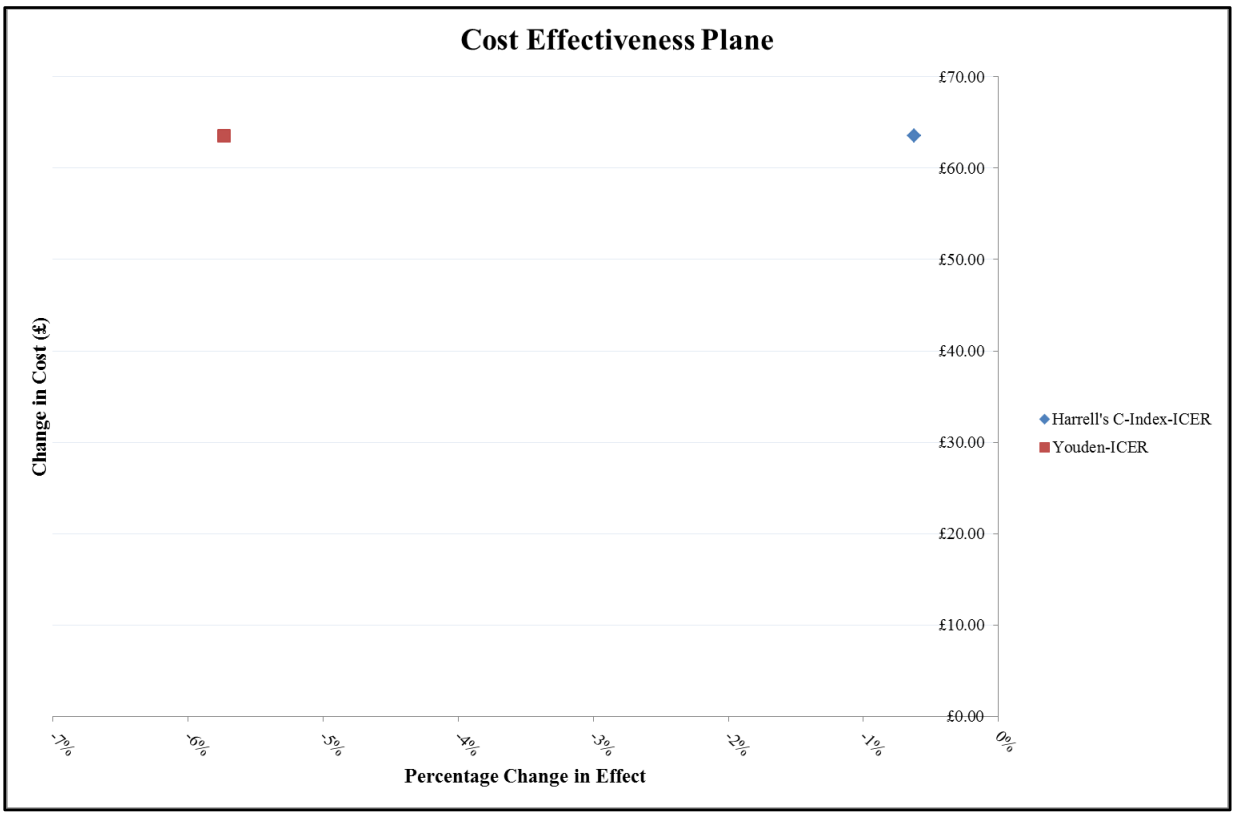


Figure 9.2 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 3a.

9.1.1.1.1.2 Incremental Net Benefit (INB)

The INB was calculated using the ICER calculated above, and varying the willingness to pay for the improvement between risk models using the Youden Index and Harrell's C-Index. Based on a £0 to £10,000 range for the willingness to pay; the INB never rises above 0 for both measures of effect.

Similar results were found when comparing Model 2 (standard fracture risk factors only) with Model 4a (adding binary BMD measurement to data driven standard fracture risk factors).

9.1.1.1.2 Model 2 vs. Model 3b

Adding continuous BMD measurement (Model 3b) to standard fracture risk factors (Model 2) increases the expected cost of calculating fracture risk by £57.76 (Table 9.3). It also lowers the proportion of patients who would be recommended treatment based on their risk and 8.5% treatment threshold. Using Model 2 or Model 3b; 68% or 65% of patients would be

recommended preventative treatment, respectively. Of these, marginally fewer patients would be treated unnecessarily (90% using Model 2 compared to 89% using Model 3b).

Conversely, adding continuous BMD to standard fracture risk factors marginally increases the number of patients who would be considered low risk based on their fracture risk factors; Model 2 would predict 32% to be low risk compared to 35% when using BMD measurement. Of these patients, 2% of patients would be classed as missed diagnoses from using either model.

9.1.1.1.2.1 Incremental Cost-Effectiveness Ratio (ICER)

An increase in discriminatory accuracy was also found when adding continuous BMD measurement to standard fracture risk factors. The Youden Index increased by 14% and Harrell's C-index increased by 1% (Table 9.4).

An ICER of £402.80 per percentage increase in Youden Index was found; and an ICER of £5,018.85 per percentage increase in Harrell's C-Index. These have been plotted on a cost-effectiveness plane (Figure 9.3). Both measures lie in the first quadrant of the graph, indicating non dominance of the Model 3b (with continuous BMD); for both discriminatory accuracy measures, Model 3b is more effective and more costly than Model 2.

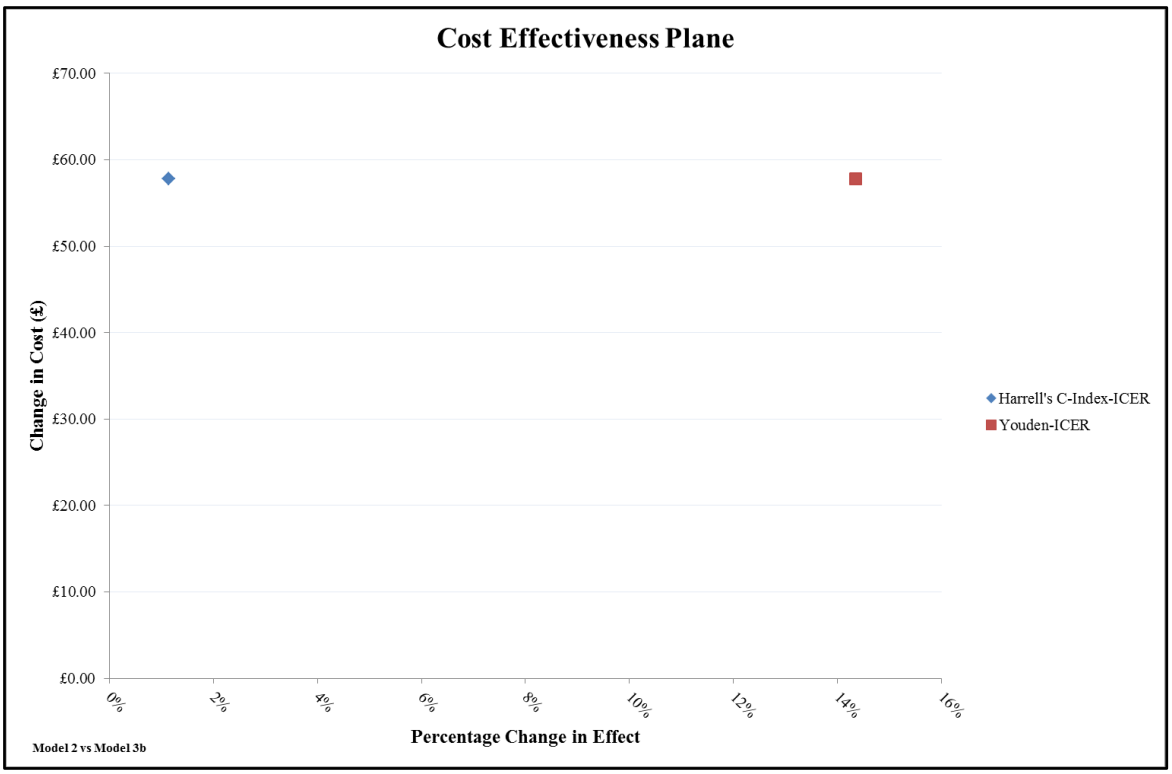


Figure 9.3 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 3b.

9.1.1.1.2.2 Incremental Net Benefit (INB)

The INB was calculated using the ICER calculated above, and varying the willingness to pay for the improvement between risk models using the Youden Index and Harrell's C-Index. Based on a £0 to £10,000 range for the willingness to pay; the INB increases with the willingness to pay for both measure of effect. A higher willingness to pay is required using Harrell's C-Index as the measure of effect (Figure 9.4).

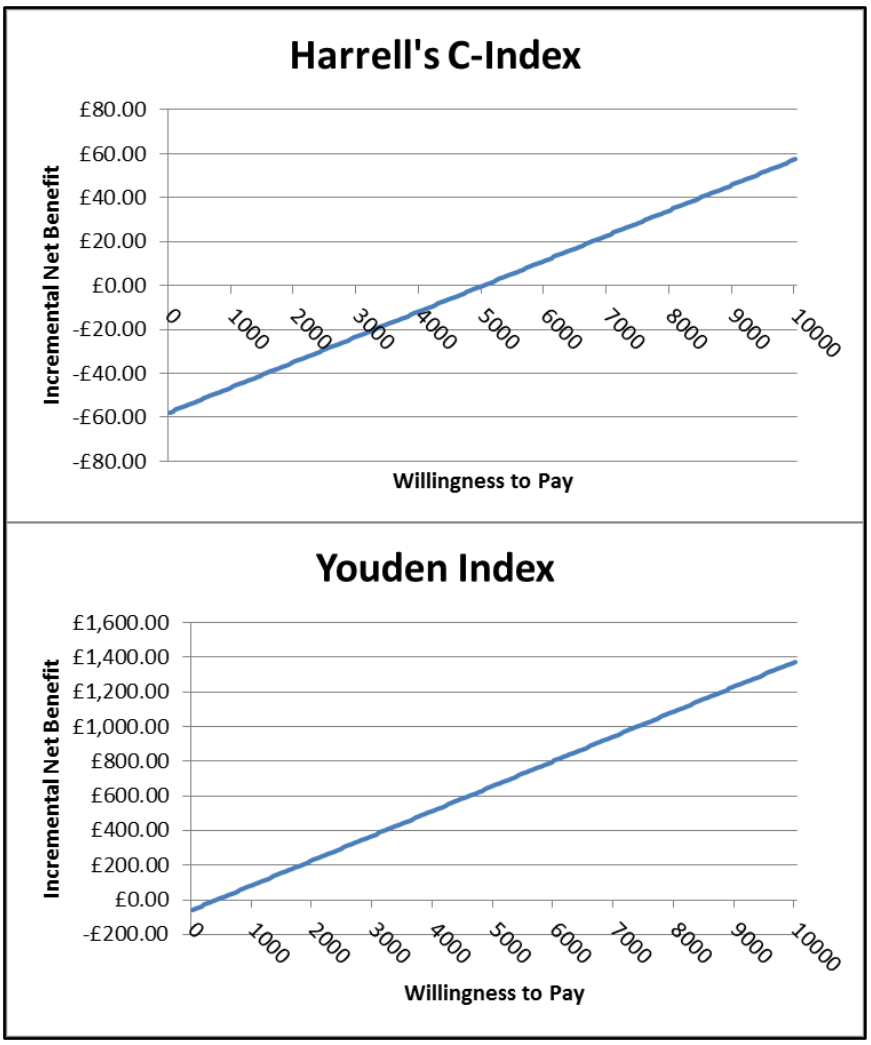


Figure 9.4 Incremental net benefit for Harrell's C-Index and the Youden Index, comparing Model 2 with Model 3b.

9.1.1.1.3 Model 2 vs. Model 4b

Comparing Model 2 with Model 4b (data driven standard fracture risk factors with continuous BMD shows an increase the expected cost of calculating fracture risk by £55.93 (Table 9.3). Further, Model 4b also lowers the proportion of patients who would be recommended treatment based on their risk and 8.5% treatment threshold. Comparing Model 2 or Model 4b; 68% or 64% of patients would be recommended preventative treatment, respectively. Of these, marginally fewer patients would be treated unnecessarily (90% using Model 2 compared to 89% using Model 4b).

Conversely, adding continuous BMD to standard fracture risk factors marginally increases the number of patients who would be considered low risk based on their fracture risk factors; Model 2 would predict 32% to be low risk compared to 36% when using BMD measurement. Of these patients, 2% of patients would be classed as missed diagnoses from using either model.

9.1.1.1.3.1 Incremental Cost-Effectiveness Ratio (ICER)

An increase in discriminatory accuracy was also found when adding continuous BMD measurement to data driven standard fracture risk factors. The Youden Index increased by 17% and Harrell's C-index increased by 0.9% (Table 9.3).

An ICER of £326.32 per percentage increase in Youden Index was found; and an ICER of £6,359.25 per percentage increase in Harrell's C-Index. These have been plotted on a cost-effectiveness plane (Figure 9.5). Both measures lie in the first quadrant of the graph, indicating non dominance of the Model 3b (with continuous BMD); in terms of discriminatory accuracy, Model 3b is more effective and more costly than Model 2.

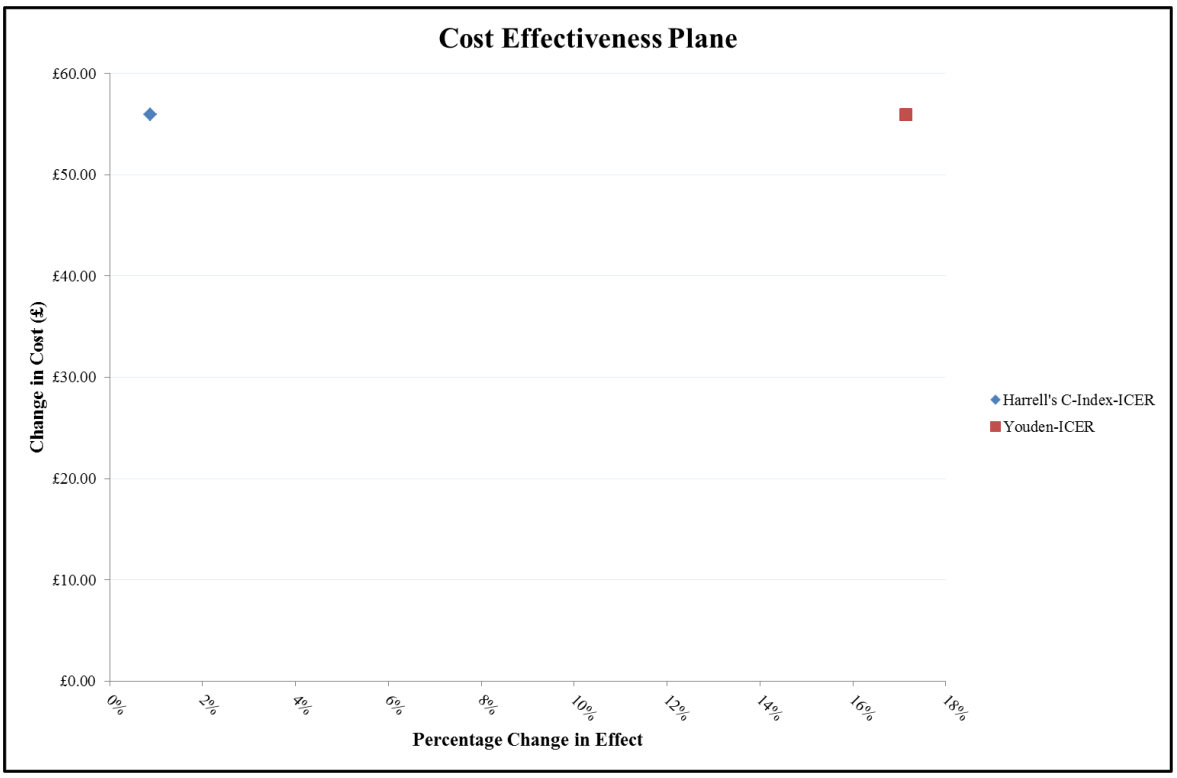


Figure 9.5 Cost-effectiveness Plane for the Youden Index and Harrell's C-Index, comparing Model 2 and Model 4b.

9.1.1.1.3.2 Incremental Net Benefit (INB)

The INB was calculated using the ICER calculated above, and varying the willingness to pay for the improvement between risk models using the Youden Index and Harrell's C-Index. Based on a £0 to £10,000 range for the willingness to pay; the INB increases with the willingness to pay for both measure of effect. A higher willingness to pay is required using Harrell's C-Index as the measure of effect. (Figure 9.6)

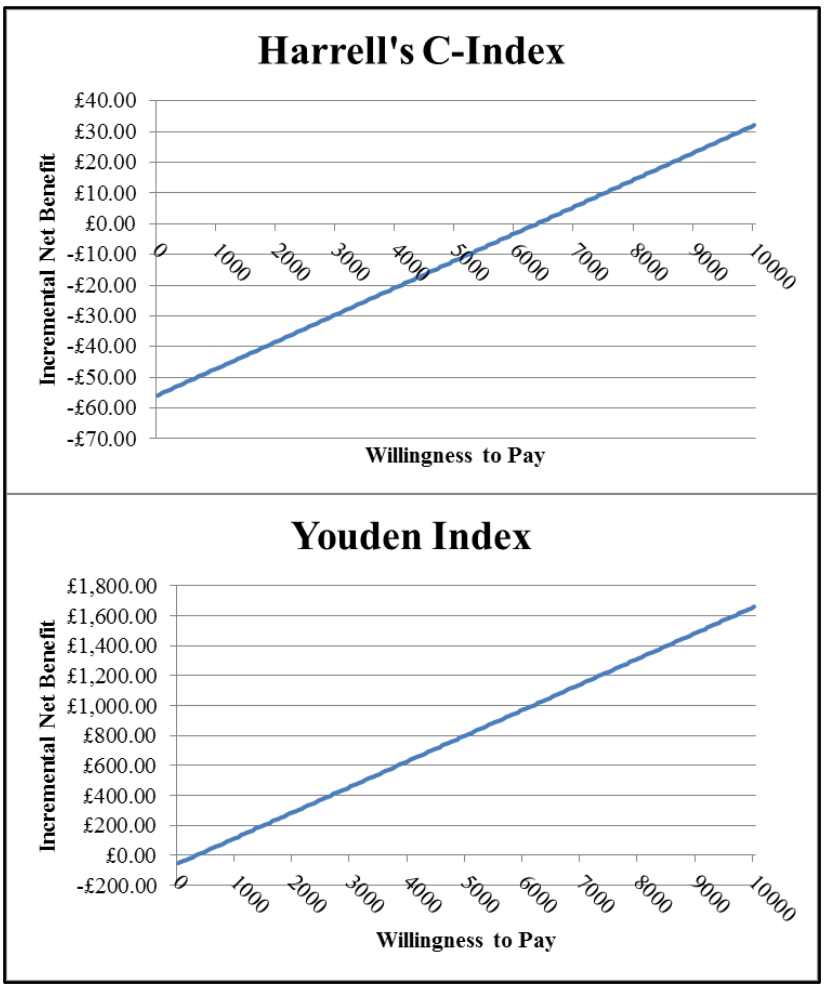


Figure 9.6 Incremental net benefit for Harrell's C-Index and the Youden Index, comparing Model 2 with Model 4b.

9.1.1.2 Probabilistic Sensitivity Analysis

As well as the deterministic cost-effectiveness analysis, a probabilistic analysis was also performed to measure the uncertainty of each ICER. This was performed using the change in the Youden Index and Harrell's C-Index as measures of effect; and were compared for all comparisons described above.

Using the input values for the decision tree presented in Table 9.2, the estimates and their standard errors (if available) were used for the probability distribution. A beta distribution was used for proportion estimates and the gamma distribution was used for costs. The alpha and beta values for costs and their gamma distributions are presented in Chapter 7, Section

7.2.5.8.6.4. Table 9.4 gives the alpha and beta values for probabilities used in the decision tree and their beta distributions.

Table 9.4 Deterministic probabilities used in the decision tree, for each risk model (with and without FH); with standard error, alpha, and beta for estimation of probabilistic probabilities derived from a Beta distribution.

Probability	Value	SE	Alpha	Beta
P(D+)	0.0745	0.0042	287.97	3579.93
P(D-)	0.9255	0.0042	3579.59	287.94
Model 2				
P(T+)	0.6847	0.0105	1340.67	617.39
P(T-)	0.3153	0.0105	617.32	1340.52
Sensitivity: P(T+|D+)	0.9034	0.0246	129.19	13.81
P(T+|D-)	0.6672	0.0111	1209.59	603.3
Specificity: P(T-|D-)	0.3328	0.0111	603.36	1209.72
P(T-|D+)	0.0966	0.0246	13.81	129.19
Model 3a				
P(T+)	0.6628	0.0107	1297.72	660.35
P(T-)	0.3372	0.0107	660.35	1297.72
Sensitivity: P(T+|D+)	0.869	0.0281	124.26	18.74
P(T+|D-)	0.6463	0.0112	1171.65	641.26
Specificity: P(T-|D-)	0.3537	0.0112	641.33	1171.77
P(T-|D+)	0.131	0.0281	18.74	124.26
Model 3b				
P(T+)	0.6464	0.0108	1265.76	692.32
P(T-)	0.3536	0.0108	692.23	1265.6
Sensitivity: P(T+|D+)	0.8966	0.0254	128.21	14.79
P(T+|D-)	0.6264	0.0114	1135.94	677.37
Specificity: P(T-|D-)	0.3736	0.0114	677.21	1135.68
P(T-|D+)	0.1034	0.0254	14.79	128.21

[Table continues on the next page]

Probability cont.	Value	SE	Alpha	Beta
Model 4a				
P(T+)	0.6633	0.0107	1298.43	659.21
P(T-)	0.3367	0.0107	659.35	1298.73
Sensitivity: P(T+|D+)	0.8621	0.0287	123.27	19.72
P(T+|D-)	0.6474	0.0112	1173.63	639.26
Specificity: P(T-|D-)	0.3526	0.0112	639.32	1173.76
P(T-|D+)	0.1379	0.0287	19.72	123.28
Model 4b				
P(T+)	0.6403	0.0108	1253.78	704.31
P(T-)	0.3597	0.0108	704.22	1253.61
Sensitivity: P(T+|D+)	0.8966	0.0254	128.21	14.79
P(T+|D-)	0.6198	0.0114	1123.7	689.2
Specificity: P(T-|D-)	0.3802	0.0114	689.28	1123.82
P(T-|D+)	0.1034	0.0254	14.79	128.21

Ten-thousand iterations were produced and the mean ICER over all iterations was presented with a 95% confidence interval, using the Youden Index and C-Index as measures of effect. A summary of the probabilistic output is given in Table 9.5.

Table 9.5 Results from the probabilistic cost-effectiveness analysis evaluating the added value of BMD measurement to fracture risk prediction.

Comparison	Change in Cost (95% CI)	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Model 2 vs. Model 3a	£63.61 (£37.29 to £89.92)	-4.72 (-39.28 to 29.83)	-£1,347.44 (-£5,945.09 to £5,653.88)	-0.55 (-9.15 to 8.06)	-£11,636.37 (-£22,947.06 to £22,073.67)
Model 2 vs. Model 3b	£57.80 (£30.46 to £85.15)	15.74 (-20.58 to 52.06)	£367.25 (-£4,241.88 to £4,828.50)	1.29 (-6.84 to 9.42)	£4,480.54 (-£22,816.84 to £22,970.55)
Model 2 vs. Model 4a	£63.71 (£37.73 to £89.69)	-7.88 (-42.73 to 26.98)	-£808.97 (-£4,962.67 to £4,393.47)	-0.63 (-9.21 to 7.96)	-£10,140.64 (-£21,872.59 to £20,723.36)
Model 2 vs. Model 4b	£55.93 (£28.40 to £83.46)	18.85 (-18.97 to 56.67)	£296.68 (-£3,208.55 to £3,847.85)	1.04 (-7.11 to 9.18)	£5,381.53 (-£22,104.70 to £20,699.95)

9.1.1.2.1 Model 2 vs. Model 3a

Results from the probabilistic analysis are consistent with the deterministic cost-effectiveness analysis. Based on the change in Youden Index and Harrell's C-Index, Table 9.5 shows, the mean ICER was -£1,347.44 (-£5,945.09 to £5,653.88) and -£11,636.82 (-£22,947.06 to £22,073.67), respectively. A cost-effectiveness plane for both measures are presented in Figure 9.7, and show over half of the iterations from the probabilistic sensitivity analysis fall in the increased cost and decreased effect quadrant (dominance).

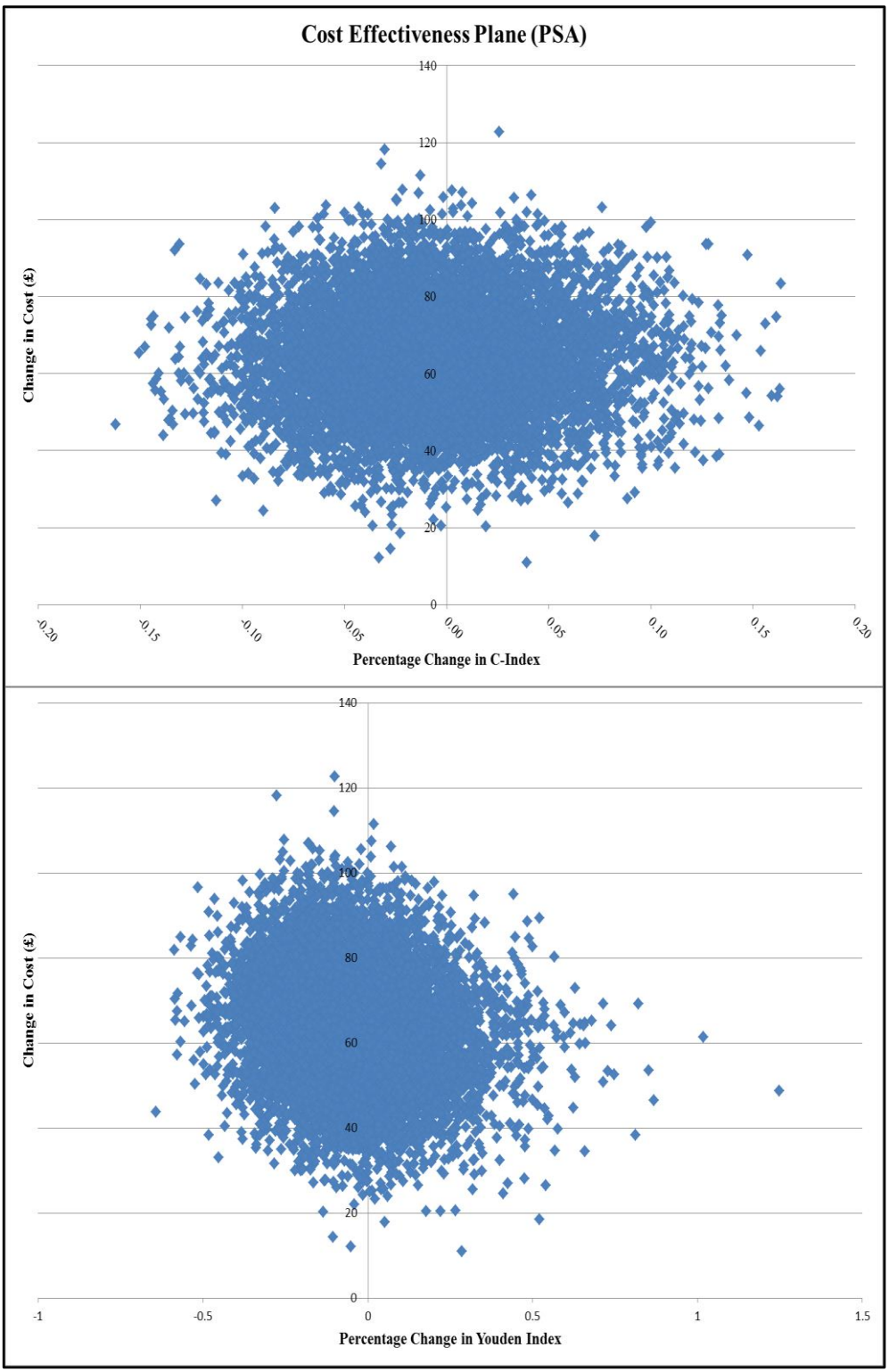


Figure 9.7 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 3a.

9.1.1.2.1.1 Cost Effectiveness Acceptability Curve

A cost effectiveness acceptability curve, varying the willingness to pay for the added effect is presented in Figure 9.8. It shows that given the maximum willingness to pay of

£100,000 per percentage increment of the Youden Index and Harrell's C-Index, the probability Model 3a (with binary BMD measurement) is cost-effective to Model 2 (with only standard fracture risk factors) is 0.36 and 0.43, respectively. Thus, given the data, there is 36% and 43% chance that Model 3a is cost effective compared to Model 2 based on the increment of the Youden and Harrell's C-Index and given a maximum willingness to pay of £100,000 per percentage increment of the Youden and Harrell's C-Index.

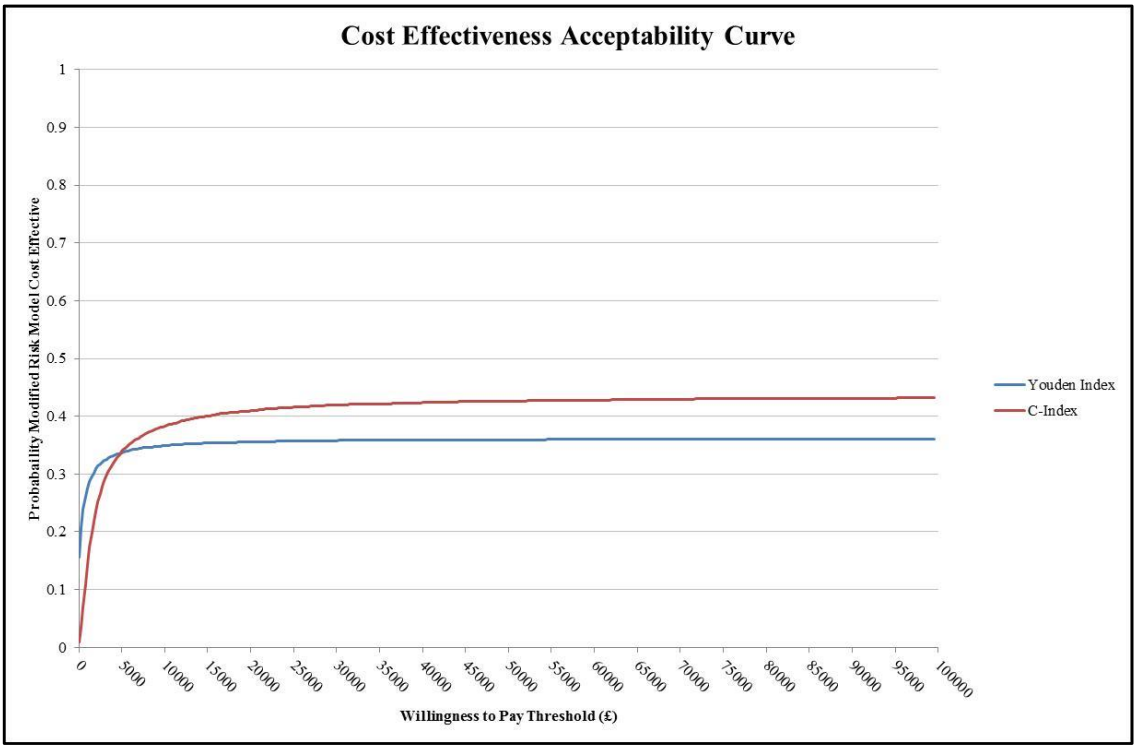


Figure 9.8 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 3a.

Further, Figure 9.8 shows that a minimum willingness to pay needed for a 0.36 probability of adding binary BMD measurement to be cost effective is approximately £19,250 per percentage increment in the Youden Index. However, a minimum willingness to pay for a 0.43 probability of adding BMD measurement to be cost effective is approximately £44,750 per percentage increment in Harrell's C-Index.

Results were the same for the probabilistic analysis comparing Model 2 with a data driven fracture risk model, including binary BMD measurement (Model 4a).

9.1.1.2.2 Model 2 vs. Model 3b

Results from the probabilistic analysis are consistent with the deterministic cost-effectiveness analysis. Table 9.5 shows that based on the change in Youden Index and Harrell's C-Index, the mean ICER was £367.25 (-£4,241.88 to £4,828.50) and £4,480.54 (-£22,816.84 to £22,970.55), respectively. A cost-effectiveness plane for both measures are presented in Figure 9.9, and shows over half of the iterations from the probabilistic sensitivity analysis fall in the increased cost and increased effect quadrant.

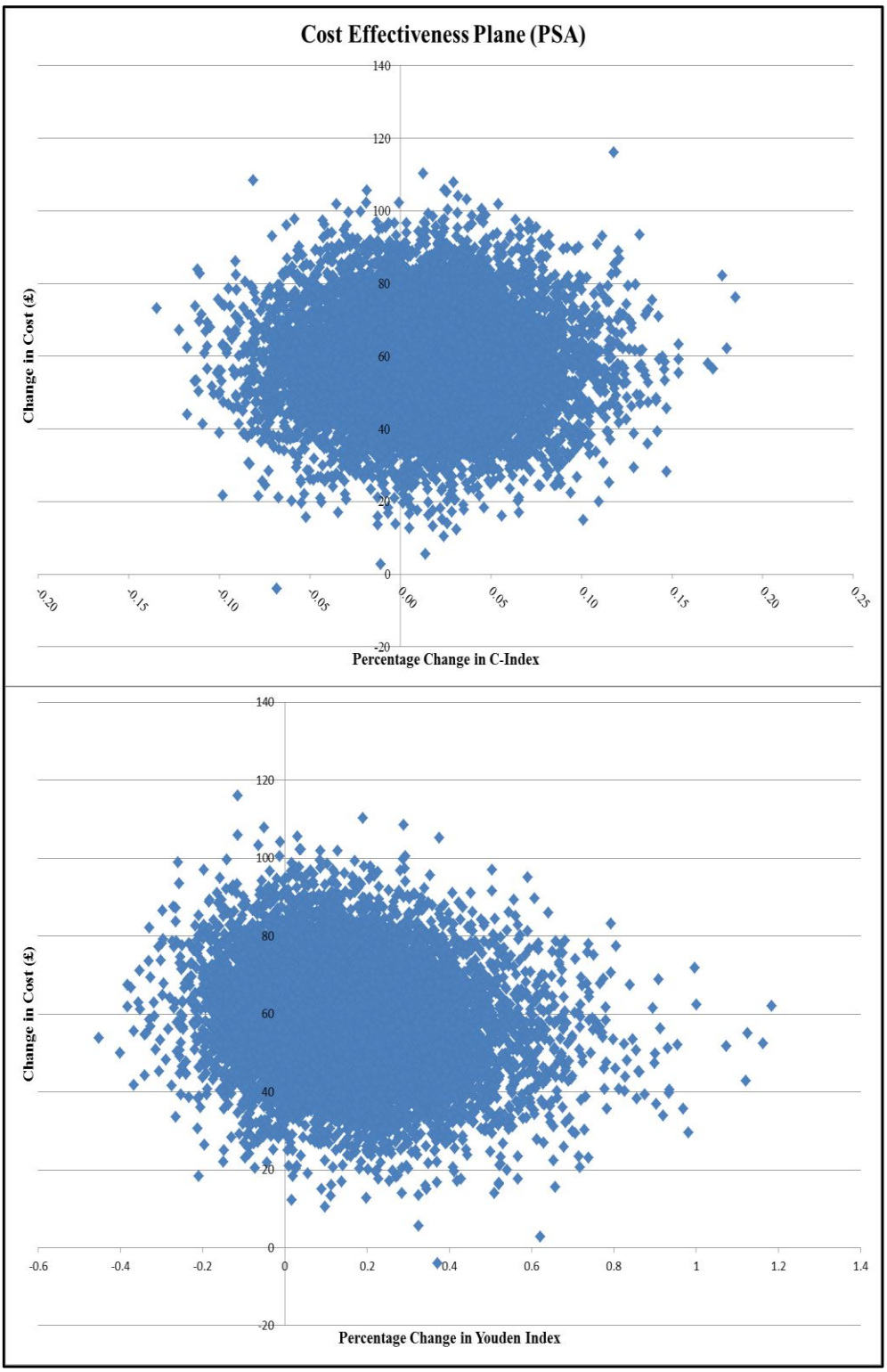


Figure 9.9 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 3b.

9.1.1.2.2.1 Cost Effectiveness Acceptability Curve

A cost effectiveness acceptability curve, varying the willingness to pay for the added effect is presented in Figure 9.10. It shows that given the maximum willingness to pay of

£100,000 per percentage increment of the Youden Index and Harrell's C-Index, the probability Model 3b (with continuous BMD measurement) is cost-effective to Model 2 (with only standard fracture risk factors) is 0.81 and 0.61, respectively. Thus, given the data, there is 81% and 61% chance that Model 3b is cost effective compared to Model 2 based on the increment of the Youden and Harrell's C-Index and given a maximum willingness to pay of £100,000 per percentage increment of the Youden and Harrell's C-Index.

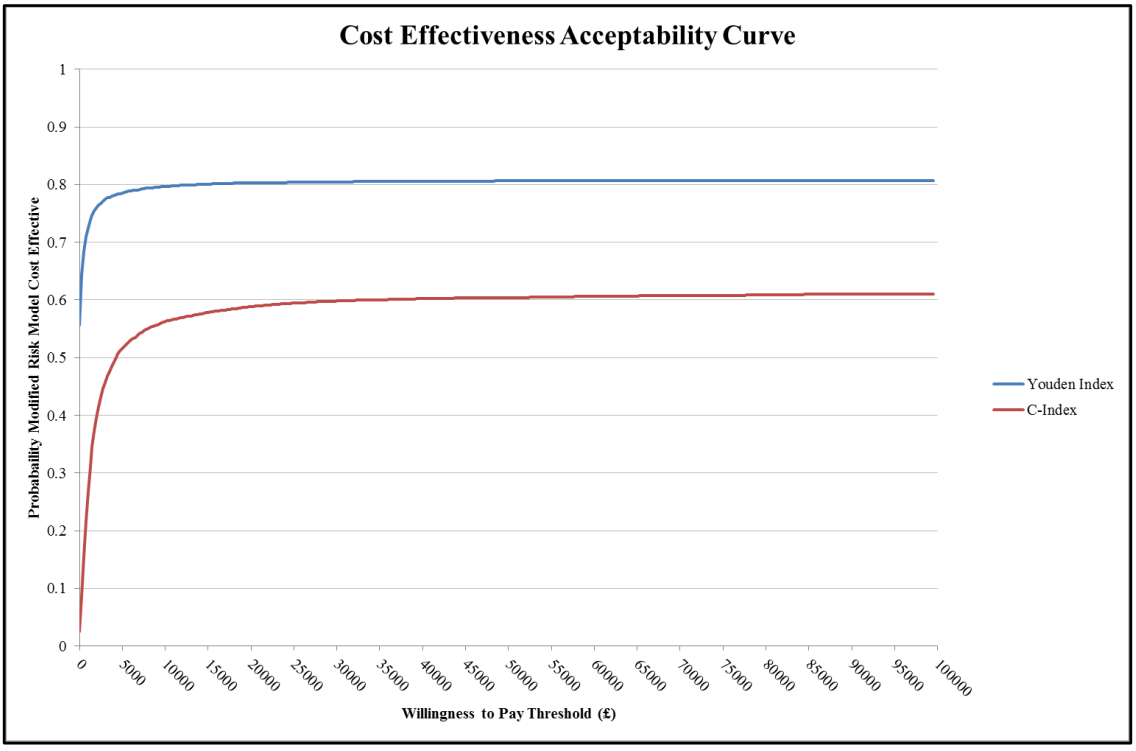


Figure 9.10 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 3b.

Further, Figure 9.10 shows that a minimum willingness to pay needed for a 0.81 probability of adding BMD measurement to be cost effective is approximately £32,500 per percentage increment in the Youden Index. However, a minimum willingness to pay for a 0.61 probability of adding continuous BMD measurement to be cost effective is approximately £56,000 per percentage increment in Harrell's C-Index.

9.1.1.2.3 Model 2 vs. Model 4b

Results from the probabilistic analysis are consistent with the deterministic cost-effectiveness analysis. Table 9.5 shows that based on the change in Youden Index and Harrell's C-Index, the mean ICER was £296.68 (-£3,208.55 to £3,847.85) and £5,381.53 (-£22,104.70 to £20,699.95), respectively. A cost-effectiveness plane for both measures are presented in Figure 9.11, and shows over half of the iterations from the probabilistic sensitivity analysis fall in the increased cost and increased quadrant based on Harrell's C-Index, and over 75% based on the Youden Index.

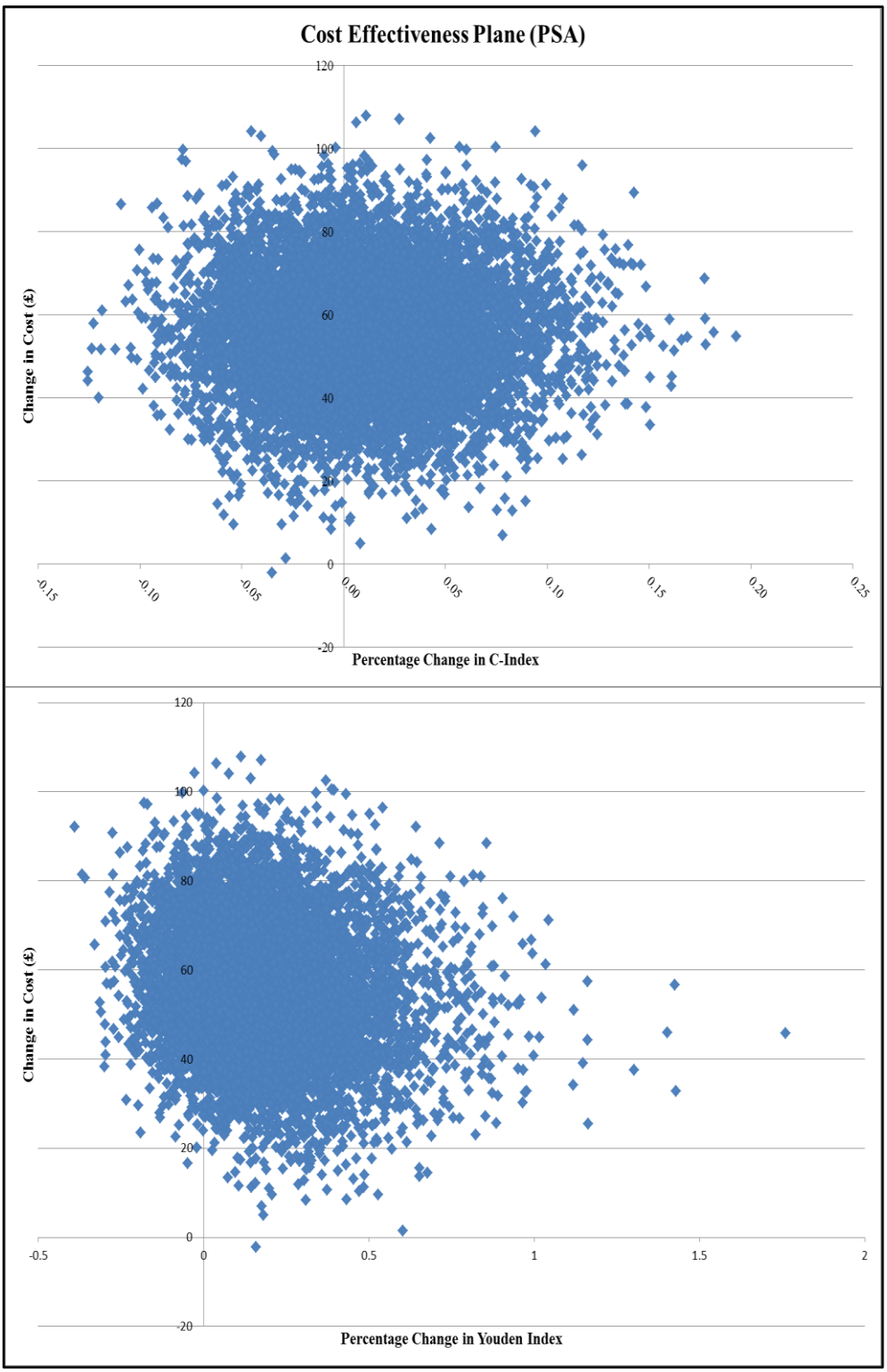


Figure 9.11 Cost Effectiveness Plane for the Youden Index and Harrell's C-Index based on the probabilistic sensitivity analysis, comparing Model 2 and Model 4b.

9.1.1.2.3.1 Cost Effectiveness Acceptability Curve

A cost effectiveness acceptability curve, varying the willingness to pay for the added effect is presented in Figure 9.12. It shows that given the maximum willingness to pay of £100,000 per percentage increment of the Youden Index and Harrell's C-Index, the probability Model 4b (with continuous BMD measurement) is cost-effective to Model 2 (with only standard fracture risk factors) is 0.85 and 0.58, respectively. Thus, given the data, there is 85% and 58% chance that Model 4b is cost effective compared to Model 2 based on the increment of the Youden and Harrell's C-Index and given a maximum willingness to pay of £100,000 per percentage increment of the Youden and Harrell's C-Index.

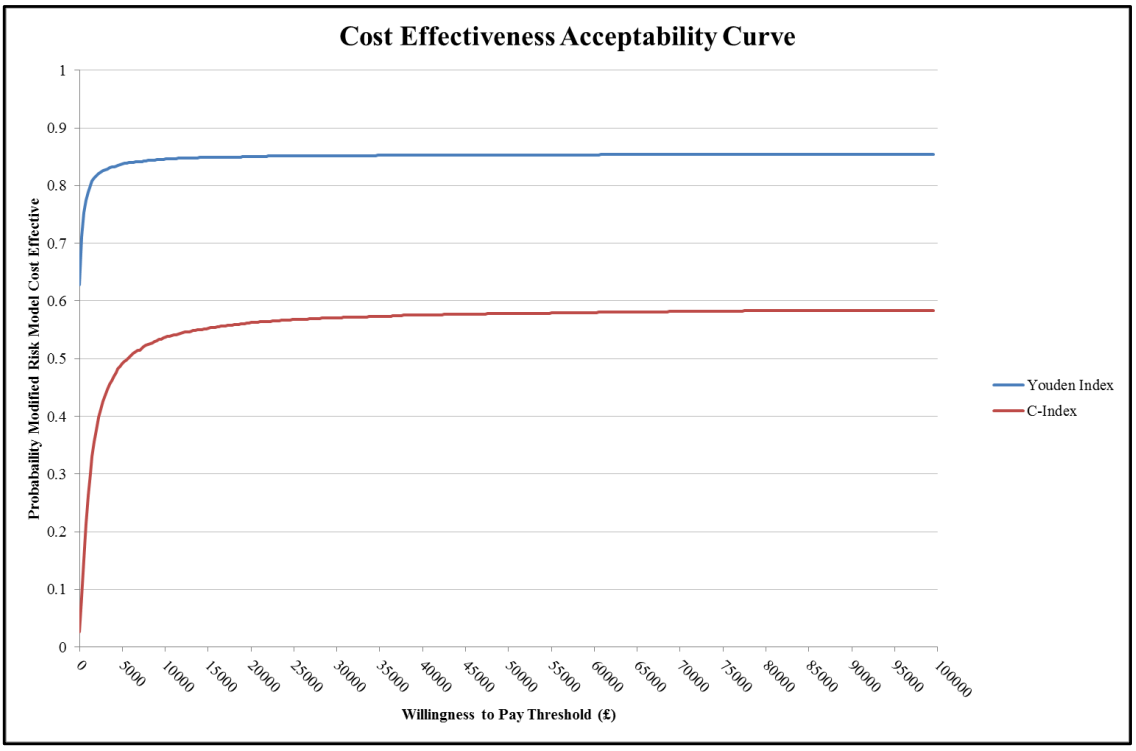


Figure 9.12 Cost effectiveness acceptability curve for the Youden Index and Harrell's C-Index, based on the probabilistic sensitivity analysis; comparing Model 2 and Model 4b.

Further, Figure 9.12 shows that a minimum willingness to pay needed for a 0.85 probability of adding BMD measurement to be cost effective is approximately £9,750 per percentage increment in the Youden Index. However, a minimum willingness to pay for a

0.58 probability of adding BMD measurement to be cost effective is approximately £38,250 per percentage increment in Harrell's C-Index.

9.1.2 Sensitivity Analysis

9.1.2.1 Discount Rate

A sensitivity analysis was performed for each cost-effectiveness analysis, which did not discount the cost in the analysis. This was performed for analyses comparing Model 2 with Model 3a, Model 3b, Model 4a, and Model 4b.

Table 9.6 presents a summary of the output for each comparison; and results from both the deterministic and probabilistic analysis. Small changes were found in the ICERs for each discriminatory accuracy measure in both comparisons. No change was found in the results of the cost effectiveness acceptability curve.

Table 9.6 Results from the deterministic and probabilistic cost-effectiveness analysis, comparing models with and without BMD measurement using a 0% discount rate.

Analysis	Comparison	Change in Cost	% Change in effects (Youden Index)	ICER (Youden Index)	% Change in effects (C-Index)	ICER (C-Index)
Deterministic	Model 2 vs. Model 3a	63.04	-5.73%	-£1,099.39	-0.62%	-£10,121.97
	Model 2 vs. Model 3b	56.76	14.34	£395.83	1.15	£4,931.93
	Model 2 vs. Model 4a	63.45	-9.12	-£695.69	-0.69	£9,140.94
	Model 2 vs. Model 4b	54.76	17.14	£319.54	0.88	£6,226.95
Probabilistic	Model 2 vs. Model 3a	£63.10 (£36.59 to £89.61)	-4.19 (-38.98 to 30.59)	-£1,505.20 (-£5,886.35 to £4,659.08)	-0.56 (-9.12 to 8.01)	-£11,326.49 (-£20,233.38 to £23,761.19)
	Model 2 vs. Model 3b	£56.83 (£28.87 to £84.74)	15.69 (-20.75 to 52.12)	£362.13 (-£4,076.51 to £4,360.35)	1.26 (-6.96 to 9.47)	£4,525.94 (-£23,935.56 to £21,628.27)
	Model 2 vs. Model 4a	£63.38 (£36.93 to £89.83)	-7.93 (-41.77 to 25.91)	-£799.10 (-£5,033.63 to £5,325.42)	-0.62 (-9.20 to 7.96)	-£10,220.96 (-£23,243.01 to £22,918.51)
	Model 2 vs. Model 4b	£54.61 (£26.31 to £82.91)	18.54 (-18.81 to 55.88)	£294.59 (-£3,301.28 to £3,930.26)	0.96 (-7.08 to 9.00)	£5,679.43 (-£18,261.99 to £19,807.38)

9.1.2.2 Including Missing values

The cost effectiveness analyses were re-run using a cohort which included missing values as a category of the affect risk factors. No change was found in the results of the cost-effectiveness analysis; when comparing Model 2 with Model 3a, Model 3b, Model 4a, and Model 4b.

9.2 Discussion

This section discusses the added value of BMD measurement when evaluated using the refined methodology. It combines the findings of the results using current methodology to provide more clinically relevant information with regards to the addition of BMD to fracture risk prediction.

9.2.1 Summary of findings

Adding BMD measurement increased the expected cost of using the fracture risk prediction model, which increased more when adding BMD measurement as a binary risk factor to indicate osteoporosis. Fracture risk prediction with continuous BMD measurement increased the expected cost of the test by £57.76, whereas binary BMD measurement increased it by £63.53.

For this cost, the continuous BMD measurement increased the probability of identifying patients as low risk and without the event (true negative), for a small decrease in the probability of identifying patient as high risk and with the event (true positive). Conversely, adding binary BMD measurement was found to provide a smaller increase in the probability of identifying patients as low risk and without the event (true negative), for a larger decrease in the probability of identifying patient as high risk and with the event (true positive).

The ICER estimates showed increased cost for an increased effect for models with continuous BMD measurement, measured by the discriminatory accuracy of the models. However, increased cost was associated to a decrease in effect when adding binary BMD to the model.

Data driven risk models including BMD measurement were also compared to the original fracture risk model, without BMD. These models showed a further improvement in the

Youden Index, but a smaller improvement was found in Harrell's C-Index when assess the added value of a data driven risk prediction model with continuous BMD measurement. Further, this analysis showed that replacing standard fracture risk factors which are less associated with fracture risk after the addition of BMD measurement is more cost effective than assessing fracture risk using standard risk factors alone (Youden Index ICER=£296.68 (95% CI:-£3,208.55 to £3,847.85)), and more cost effective than including BMD measurement as an additive component to standard risk factors (Youden Index ICER=£367.25 (95% CI:-£4,241.88 to £4,828.50)).

9.2.2 Strengths and Limitations

The effect of BMD information on the expected cost of using fracture risk prediction has not been investigated. This study is the first to provide such information with respect to fracture risk prediction models, and make change in current performance measures more clinically relevant.

Further, the refined method was performed in addition to the current validation methods, and where possible, estimates were derived from the validation study (sensitivity, specificity etc.). Thus, estimates were single study based and more robust.

However, large amounts of uncertainty were found in the ICER estimates for both measures of effect. This can be attributed to the small sample size of the data, as well as large amounts of parameter uncertainty. Although robust estimates were extracted from the validation study with their respective standard errors rather than taken from the literature, the small sample size increased their uncertainty. Similar to the CVD study conducted Chapter 4-6; costs were extracted from current literature, but did not provide corresponding standard errors for the estimates. In line with current practice, the standard error for these estimates was assumed to be 10% of the cost (112-114).

Further, caution is needed as the analysis is dependent on the treatment efficacy; this information is needed for patients considered low and high risk of fracture. In this study, data was taken from a randomised, blinded clinical trial which compared the effects of 5mg/d of Alendronate, versus a placebo. This trial also allowed for the effects of Alendronate to be investigated in low and high fracture risk patients; the fracture event rate was taken from high fracture risk patients in this study (osteoporotic) and low fracture risk patients (not osteoporotic).

Current guidance recommends prescribing 70mg of Alendronate per week or 10mg per day as the primary prevention strategy for high risk fracture patients; hence using clinical trials comparing the effects of a higher dose of Alendronate vs. placebo would be preferable. However, none could be found at the time of this study, which also provided treatment effects on low and high fracture risk patients.

It is also recognised and advised that meta-analyses would be more beneficial to use for data regarding the treatment effects of Alendronate compared to individual clinical trial data. These were searched for, however in this case, none were found presenting information on the primary prevention drug for high fracture risk (Alendronate) versus placebo, and in patients considered at low or high risk of fracture.

The decision model assumes that all patients at high risk are treated with preventative treatment. It does not account for the varying compliance rate patients may have depending on the treatment. However, as the decision tree is designed to look at the difference between models with respect to the number of patients treated or not, the same compliance rate to treatment, if treated, is assumed for both risk models; thus balancing this bias.

Costs data was taken from the validated data sources (25, 70), which have been used to inform national guidance. Information was also extracted in line with current guidance and its future impact. For example, based on current treatment recommendations, Alendronate is the primary preventative treatment measure for fracture risk. Thus, when calculating treatment cost, these were used with the expected prescribing proportions for all drug therapy used to reduce fracture risk.

The study calculates ICERs for each measure of effect and also presents information on the probability adding FH information is cost effective based on willingness to pay. Current NICE guidance uses a £20,000 per quality adjusted life year (QALY) to identify if an intervention is cost-effective. However, QALYs have not been used as the measure of effect, and instead discriminatory performance measures are used. This cannot be translatable to the threshold used by NICE. However, it is proposed to use this information to create a league table for new risk factors, allowing them to be compared and better risk prediction models to be identified.

9.2.3 Clinical implications and Conclusion

Depending on the willingness to pay, the cost of a DXA scan to retrieve BMD measurement in patients has shown to be cost effective. Both, results from the current and refined methods, showed added benefit of including continuous BMD measurement to fracture risk prediction. It is not advised the use BMD measurement in a binary format to indicate presence of osteoporosis.

Bone mineral density is already embedded in the FRAX risk score which is used in UK fracture risk assessment, but is optionally used in routine risk assessment. Further, NICE clinical guidelines have recommended for research to be undertaken, identifying the

added value of BMD measurement. Implications from this study would increase the measurement of BMD in primary care for routine fracture risk assessment.

Increasing the use of BMD in fracture risk prediction would see improved prevention of fractures through better tailored treatment to reduce fracture risk.

9.3 Chapter Summary

This chapter presented and discussed the results of the refined methodology used to evaluate the added value of BMD measurement to fracture risk prediction. This included a cost-effectiveness analysis, adding BMD measurement to standard fracture risk factors.

The next chapter will discuss the findings of the thesis, the strength and limitations of the refined methodology, clinical implications, and will also recommend future research.

Chapter 10 - Discussion

This chapter summarises the findings of this thesis and discusses the results, linking it to current literature. It also describes the strengths and limitations of the thesis, and outlines the implications of the findings. Future work and research recommendations are then discussed, ending with concluding remarks.

10.1 Summary of the findings

10.1.1 Objective 1: Review and Critique of Current Methodology

Collate information of current methods used to assess the incremental value of novel markers. Evaluate these with respect to their application, strengths, and weaknesses.

In Chapter 2, the aim was to review and critically appraise current methods used to evaluate the added value of a new risk factor. The MEDLINE electronic database was searched for papers validating risk prediction models, evaluating new risk factors, and comparing risk prediction models. Citations and references for key papers were also searched.

Reviewing the literature found that current methods used to measure the added value of new risk factors to established risk prediction models are inadequate to provide a full evaluation of its consequences. Evidence suggests current methodology is insensitive to change when adding new risk factors to established risk prediction models (40); and is criticised for lacking clinical interpretability and relevance (19, 21, 30).

Although, the AUC is the most common measure of discriminatory accuracy used to describe the improvement between risk models (39), and is used to inform national guidance recommendations (3); it has shown to give small, clinically irrelevant changes

when describing the added value of a new risk factor to existing risk prediction models. Pepe et al (40) showed that even with large effect sizes describing the association of the new risk factor with the risk of the prediction model outcome, small changes in the discriminatory accuracy were found.

Work by the PHG foundation (30) further indicates that disassociation of current measures makes it difficult to identify the improvement in a risk prediction model with a new risk factor. Calibration and discrimination are two areas in which a risk model is assessed; however, a model may be well calibrated but have poor discriminatory accuracy, and vice versa. Similarly, when assessing discriminatory accuracy of a model using the AUC, this assesses the trade-off between high sensitivity of a model or high specificity. This has shown to produce a trade-off scenario between measures (171), with difficulty to identify which is the better option.

The limitations of current methodology have led to new techniques to evaluate the added value of a new risk factor (23, 24), specifically addressing the clinical added value. Using reclassification tables have become a popular method to evaluate the added value of a new risk factor; which directly compares the shift of patients based on the prediction of two risk models. It presents more clinically relevant results, with regards to the proportion of patients moving between risk categories. Similarly, a decision curve analysis investigates the true positives gained by using a risk prediction model with the new risk factor, without an increment in the false negatives.

However, interpretation remains a limitation (172), as well as a lack of statistical measures to capture uncertainty of estimates. Further, newer methods do not assess the cost implication when describing the clinical impact of adding a new risk factor to an established risk prediction model.

10.1.2 Objective 2: Refine Current Methodology using Health Economic Evaluation Methods

Refine the identified methods to fully assess the clinical consequence when adding a new risk factor to an established risk prediction model; encompassing evaluation of statistical and clinical improvement, with cost implications.

In Chapter 3, health economic evaluation methods were explored, as well and decision analytical frameworks. The aim of this chapter was to describe the proposed analysis to combine current measures of model performance to cost data, to fully evaluate the effectiveness and cost effectiveness of new risk factors.

Health economic evaluation methods were identified to address the limitations of current methodology used to evaluate the added value of a new risk factor; interpretability, clinical relevance, and assessment of costs for clinical impact. Specifically, cost effective and cost utility analyses have been used extensively in medical research to evaluate new interventions compared to 'usual care' (173); they have been used to compare screening methods (174, 175); and are also used to inform national guidance (65). However, they are not used to evaluate the added value of a new risk factor, nor assess cost-effectiveness in relation to model performance measures even though it has been recommended (27, 176).

Within health economic evaluation methods, a cost-effectiveness analysis was found to be most applicable for the evaluating the added value of a new risk factor. This was because it is a recognised approach used in a similar situation to compare screening methods in the economics of screening (174, 177-181). Further, although cost-utility analyses, which use preference based measures (e.g. QALYs) are more commonly used to inform national

guidance (65), a cost-effectiveness analysis for measures effect on their natural scale is more applicable to evaluating the added value of a new risk factor.

The main types of decision modelling used to support cost effectiveness analyses in health research were found to be decision trees and Markov models (174, 182), of which the decision tree is the simplest form of a decision model (183) and was chosen to provide the analytical framework for the cost-effectiveness analysis evaluating the added value of a new risk factor. The decision tree model is another recognised approach used when assessing the economics of screening (174, 175); it provides visual representation and structure to the data needed to evaluate the added value of a new risk factor; and also allows for basic probability relationships to be used. Thus, it explicitly shows every possibly outcome from choosing a risk model with the new risk factor compared to without, based on subsequent events.

In considering this approach to evaluate the added value of the new risk factor; the decision tree would compare two risk prediction models; with and without the new risk factor. This has a similar premise to assessing screening programmes, which would use observed and predicted events to identify improvement between comparators (184), but would differ in relation to the measure of effects.

The measures of effect for the cost-effectiveness analysis were chosen to describe the change in discriminatory accuracy of a model; which is the most popularly reported measure to describe the added value of a new risk factor (39, 74, 149, 185). Specifically, Harrell's C-Index was used as to describe the change in discriminatory accuracy, which accounts for risk prediction models derived using Cox regression (time dependent analysis) (186). The analysis was also performed using the Youden Index, another

measure of discriminatory accuracy, which could be calculated directly from the decision tree.

Thus, the measure of effectiveness would be a model performance based measure rather than the traditional health benefit measures, such as cases detected used in economics of screening (174, 175), or QALYs used to inform national guidance (65).

The cost-effectiveness analysis would then associate the change in effects with the change in the expected cost from the risk prediction model in an ICER calculation; identifying the more cost effective risk model and describing the added value of the new risk factor.

10.1.3 Objective 3: Apply the Refined Methodology to Two Exemplar Datasets

Apply the refined methodology to two common exemplar conditions using generalisable data: assess the added value of family history of cardiovascular disease to cardiovascular disease risk prediction; assess the added value of bone mineral density to fracture risk prediction.

The refined methodology was applied to two exemplar prospective cohort studies. In Chapters 4-6, the added value of FH of CVD to CVD risk prediction was assessed. In Chapters 7-9, the added value of BMD measurement to fracture risk prediction as assessed. The aim of both of these studies was to exemplify the refined methodology described in Chapter 3. Study specific findings can be found in Chapter 5, 6, 8, and 9.

Adding FH to CVD risk prediction gave an increased measure of effect for an increased cost, and given a willingness to pay it was found to be cost effective.

Although costs were incurred when treating patients unlikely to develop CVD in the future, the trade-off of reduced specificity for improved sensitivity is cost-effective. This has also been reflected in current literature and has contributed to the lowering CVD risk

threshold for statin treatment to 10% (3), which would substantially increase the number of patients being correctly and incorrectly treated with statin therapy (187).

Specifically, findings from the cost-effectiveness analyses showed increased cost for an increased measure of effect (discriminatory accuracy); adding family history to standard CVD risk prediction cost £799.91 (-£5,962.15 to £5968.22) per percentage increase in the Youden Index and £7,788.76 (-£42,760.16 to £48,962.39) per percentage increase in Harrell's C-Index. The added value of FH was independent of the risk factors included in the model; results were consistent between adding FH to standard, and standard with additional CVD risk factors.

Adding continuous BMD measurement to fracture risk prediction also gave an increased measure of effect for an increased cost, and given a willingness to pay it was found to be cost effective. However, adding binary BMD measurement to indicate osteoporosis resulted in poorer model performance.

Adding continuous BMD measurement to standard fracture risk prediction cost £367.25 (-£4,241.88 to £4,828.50) per percentage increase in the Youden Index and £4,480.54 (-£22,816.84 to £22,970.55) per percentage increase in Harrell's C-Index. It was also found that although binary BMD measurement was associated with the risk of fracture in the univariate analysis, it showed a decrease in effect for and increased cost; adding binary BMD measurement to standard fracture risk prediction cost -£1,347.44 (-£5,945.09 to £5,653.88) per percentage increase in the Youden Index and -£11,636.82 (-£22,947.06 to £22,073.67) per percentage increase in Harrell's C-Index.

Further, the probabilistic sensitivity analysis showed that adding FH for CVD risk prediction had a maximum cost effectiveness probability of 0.7 for a minimum willingness to pay of £15,000, based on the Youden Index; and adding continuous BMD

to fracture risk prediction had a maximum cost effectiveness probability of 0.8 for a minimum willingness to pay of £32,500, based on the Youden Index. Adding binary BMD measurement to fracture risk prediction had a less than 0.5 probability of being cost effective.

10.2 Strengths and Limitations

The refined methodology is the first method to economically appraise the added value and clinical impact of a new risk factor to established risk prediction models. Current methodology has been criticised to lack interpretability and clinical relevance when evaluating the added value of a new risk factor; this method combines this data with cost information to put changes in model performance into clinical context. As measures of effectiveness and cost effectiveness are increasingly used to inform national guideline recommendations and decisions made by commissioning groups, incorporating cost data to assess the performance of a risk prediction models would enhance the dissemination of the effects of the new risk factor and allow for better decision making.

The refined methodology also combines the different current measures of model performance, which may otherwise be disjointed (e.g. calibration and discrimination (185); and sensitivity and specificity (188)). Poor calibration occurs when the population used to derive the risk prediction models is not generalisable to the population used to validate them; and leads to over or under estimation of risk. This is accounted for in the cost-effectiveness analysis when calculating the output from the decision tree; the product of the disease prevalence in the validation cohort and the sensitivity or specificity components of model performance; which are derived risk models using the disease prevalence in the derivation cohort.

Further, incorporating treatment cost and effect data in the analysis helped the sensitivity/specificity trade off in the risk prediction model. This contextualises the improvement of the risk model and highlights the impact of increased specificity with regards to treatment effect and cost, as well as increased sensitivity. Thus, a more tailored improvement can be sought for the respective disease and treatment options; accounting for the severity, efficacy, and the cost of the treatment. This has not been looked at in relation to costs in current methodology, and would otherwise be overlooked.

The cost-effectiveness analysis also captures uncertainty in the ICER estimates by performing a probabilistic sensitivity analysis. A probabilistic sensitivity analysis was performed in addition to the deterministic cost-effectiveness analysis to explore the uncertainty around the ICER estimates and describe them with 95% confidence intervals. Exploring uncertainty in a cost-effectiveness analysis by a probabilistic sensitivity analysis is a preference stipulated by policy makers (e.g. NICE) (65) and reporting guidelines (66) for health economic evaluations. Capturing the uncertainty around estimates is another limitation of current methodology, which fails to provide this information for measures (e.g. net reclassification index from reclassification tables and net benefit from decision curve analysis).

However, a large amount of uncertainty was depicted by wide confidence intervals for the ICER estimates. This was found to originate from the lack of reporting of standard errors with cost data, as well as using literature based estimates. Three methods were used to counteract this effect. The first was to use primary data from the respective validation study for the cost-effectiveness analysis to populate the probabilities for the decision tree; this would reduce additional bias and uncertainty from using published literature for these estimates. The second method used a standard health economic approach (112-114) and assumed the standard error for costs were 10% of the estimate if the standard error was

missing. The third method used 10,000 iterations of the decision tree to reduce uncertainty.

The cost-effectiveness analysis provides flexibility to use various measures of effect. This thesis has used popular measures of discriminatory accuracy to describe the added value of a new risk factor; which are commonly used to describe risk model performance (8, 9, 39, 74, 149, 185). The expected Youden Index was calculated from the output of the decision tree model and Harrell's C-Index was calculated independently, using current methodology. However, caution is needed with externally calculated performance measures, such as Harrell's C-Index, when quantifying the uncertainty in the ICER estimate as it is not derived from the decision tree. Thus, the probabilistic sensitivity analysis of the decision tree may inaccurately estimate the uncertainty in the Harrell's C-Index ICER value. The effect of this limitation was minimised by also reporting the expected Youden Index, which was calculated from the decision tree output.

As well as capturing uncertainty of the ICER estimates, the probabilistic sensitivity analysis also used incremental net monetary benefit (INB) of the new risk model (with the new risk factor) to illustrate the probability of cost-effectiveness, given the willingness to pay. Again this contextualises the added value of the new risk factor in relation to the willingness to pay for the added effect; something with has not been considered in current methodology and is a strength of the refined methodology.

The willingness to pay threshold for a QALYs is set at £20,000-£30,000 (65) in national guidance. In this case, no threshold has been agreed for the added value of a new risk factor due to the novel nature of this approach. To account for this limitation a range of willingness to pay values were used in the analysis and plotted on the cost effectiveness acceptability curve.

Finally, inconsistency of reporting has made it difficult to directly compare risk prediction models (16), which has led to new models being developed that vary in relation to the included risk factors. The refined methodology is advantageous in this situation which can be extended for such comparisons to be made. Indirect comparisons are accounted for in the analysis by changing the respective costs for each risk models and sensitivity and specificity values.

10.3 Implications

The work of this thesis provides a new way to assess the added value of a new risk factor; one which encompasses statistical, clinical, and cost effective gains. This directly feeds into the development of national guidance, alongside current methodology; adding to the evidence base to allow for better decision making.

Guideline developers and commissioning groups are becoming increasing dependent on effectiveness and cost effectiveness evidence. This methodology will provide enhanced evaluation of both the effectiveness and cost effectiveness of a new risk factor. Policy makers would therefore be more likely to make informed choices regarding new risk factors for clinical practice.

As well as informing national guidance and commissioning groups, the methodology can also be used as part of reporting guidelines and recommendations for validating risk prediction models, to provide a fuller evaluation of the difference between risk prediction models. This would increase the use of the refined methodology producing standardised results and would also allow for better comparability across the diseases.

With increased use of the methodology, comparisons can begin to be made within and between diseases; comparing both new risk factors within disease as well as between. A summary of performance measure from the cost-effectiveness analysis can be used to

identify which risk model is more cost effective within a disease. The ICER and cost effectiveness probability can be used for each new risk factor to produce a league table style approach to evaluate, appraise, and identify the better risk prediction model and better new risk factor to be included in routine risk assessment in primary care.

Using a league table to rank new risk factors by the cost effectiveness (cost per increment of effect) of their added value to risk prediction, would enable commissioners and policy makers to compare ranks of the new risk factors and allow for comparisons of risk prediction models between and within disease areas, similar to QALY league tables. Further, with limited resources if a health service had to choose a better risk prediction model to use, a league table would also provide opportunity costs.

In addition, current policy makers often rely of consensus driven risk inflations factors in the absence of evidence of a new risk factor. Increased use of this methodology would standardise current methodology providing a better and more structured evidence base for policy makers to use when evaluating a new risk factor.

Performing the cost-effectiveness analysis combines probability and costs resulting from the addition of a new risk factor, and looks beyond risk prediction to the preventative treatment for high risk patients. This is another advantage of the refined method as it helps to define what the ideal improvement in the risk model would be in relation to effectiveness and cost effectiveness; we can then start tailoring improvement in risk prediction models for greatest cost effectiveness. For example, if comparing two new risk factors, their affect can be seen on the risk prediction model in terms of sensitivity and specificity; then depending of the cost and efficacy of treatment the more cost effective risk factor can be identified.

To support the clinical impact preliminary focus groups with GP's and commissioners have been positive regarding the refined method and the additional information it will provide. They recognised that current approaches were limited in relation to little clinical value, and supported the incorporation of cost data into future methods.

10.4 Recommendations for Future Research

In this Ph.D. an economic appraisal of the clinical impact of a new risk factor was performed. This consisted of a cost-effectiveness analysis using a decision tree model framework. The next stage of this research would look to use a Markov model to evaluate the probability and cost data for each risk prediction model. This would extend the analysis and allow the decision tree to be run more than once, accounting for recurring events over time in different cycles, and explore additional consequences of the disease.

Further, the proposed analysis was defined for a single, pre-specified risk threshold value to identify patients at high risk for not. This single threshold could be varied, and also increased to see its effects on the risk model performance and also the added value of a new risk factor. It could be varied to identify the point at which the added benefit of the new risk factor would be the greatest.

As previously mentioned, it is difficult to identify a generic willingness to pay threshold for the addition of the new risk factor to be considered cost effective, based on the ICER values. The future research is recommended to identify this threshold to use in conjunction with the proposed league tables to identify cost effective risk models and new risk factors; a similar approach was used to identify QALY thresholds.

It would however, also be useful to make the evaluation of a new risk factor more relevant by producing information which can be used in conjunction with the existing QALY threshold. The methodology can be further developed to produce a cost per QALY to

evaluate the added value of a new risk factor. Data for the QALY can be introduced into the model in a similar manner to the treatment efficacy information; as more or less patients are identified at high risk, this will impact the number of patient being treated. This can then be associated to the expected change in QALYs from taking the preventative medication.

Finally, qualitative research is recommended to work with guideline developers and commissioning groups. Interviews and focus groups with policy makers would further identify limitations in current methodology, which could potentially be accounted for by the refined methodology. Thus, as well as promoting the existing refined method, it would help tailor the refined method to meet their requirements; providing a better evaluation of the added value of the new risk factor.

10.5 Conclusion

Current methodology is inadequate to evaluate the added value of a new risk factor. Limitations were found regarding the sensitivity to change in performance measures, clinical interpretability and clinical relevance. Combining cost data with the current measure of effects was found to remove these limitations.

The proposed cost effectiveness method with a decision tree structure provides a means to assess the performance of a risk model when adding a new risk factor; thus evaluating the added value of a new risk factor. In particular, it combines the statistical improvement of a model, with the clinical improvement and consequences; allowing for better decision making to be made. Estimates from this method quantify uncertainty in the added value of a new risk factor and is recommended to be used in addition to current methodology.

This method contributes an additional measure to current methodology to evaluate the added value of new risk factors. This, used with current measures of evaluation, will help

standardise risk prediction models and access all of their benefits through worthy additional risk factors; improving the prevention of future disease using the most cost effective risk factors, and standardising assessment in future research.

References

1. Brown P, Cassidy B, Cottrell C, Davies M, Greaves C, Hill J, et al. Preventing Type 2 Diabetes: Risk Identification and Interventions for Individuals at High Risk (NICE Public Health Guidance 38). National Institute for Health and Care Excellence, 2012 Contract No.: PH38.

2. Barry P, Aspray T, Briers K, Collins GS, Compston J, Dockery F, et al. Osteoporosis: assessing the risk of fragility fracture (NICE Clinical Guidance 146). National Institute for Health and Care Excellence, 2012 Contract No.: CG146.

3. Wierzbicki A, Ahmad R, Banks L, Clark L, Duerden M, Grey E, et al. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease (NICE Clinical Guidance 181). National Institute for Health and Care Excellence, 2014 Contract No.: CG181.

4. Treasure T, Carter K, Gautam N, Hingorani AD, Hughes R, Hunt B, et al. Venous Thromboembolism: Reducing the Risk (NICE Clinical Guidance 92). National Institute for Health and Care Excellence, 2010 Contract No.: CG92.

5. Stansby G, Borthwick J, Broad N, Bull R, Collier M, McGinns E, et al. Pressure Ulcers: Prevention and Management of Pressure Ulcers (NICE Clinical Guidance 179). National Institute for Health and Care Excellence, 2014 Contract No.: CG179.

6. Adam J, Adler A, Abrams K, J. A, Ades AE, Armstrong R, et al. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women (amended) (Technology Appraisal Guidance 161). National Institute for Health and Care Excellence, 2008 Contract No.: TA161.

7. Ashton P, Briggs A, Capewell S, Cappuccio F, Caraher M, Foster C, et al. Prevention of cardiovascular disease (National Costing Report). National Institute for Health and Care Excellence, 2010 Contract No.: PH25.

8. Woodward M, Brindle P, Tunstall-Pedoe H, Estimation SGR. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart. 2007;93(2):172-6.

9. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-Reactive Protein and Parental History Improve Global Cardiovascular Risk Prediction The Reynolds Risk Score for Men. Circulation. 2008;118(22):2243-4.

10. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743-53.

11. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. British Medical Journal. 2010;341.

12. Hippisley-Cox J, Coupland C. Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: prospective open cohort study. BMJ. 2012;344:e3427.

13. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E. FRAX and the assessment of fracture probability in men and women from the UK. Osteoporosis International. 2008;19(4):385-97.

14. Board JBS. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). Heart. 2014;100 Suppl 2:ii1-ii67.

15. Grant J, Anderson B, Andrew M, Bond C, Brady A, Campbell N, et al. Risk Estimation and the Prevention of Cardiovascular Disease: A National Clinical Guidance. Edinburgh: Scottish Intercollegiate Guidelines Network, 2007 Contract No.: No. 97.

16. Collins GS, Moons KG. Comparing risk prediction models. BMJ. 2012;344:e3186.

17. Cooper A, Nherera L, Calvert N, O'Flynn N, Turnbull N, Robson J. Clinical Guidelines and Evidence Review for Lipid Modification: Cardiovascular Risk Assessment and the Primary and Secondary Prevention of Cardiovascular Disease. National Collaborating Centre for Primary Care and Royal College of General Practitioners. 2008.

18. Wood D, Wray R, Poulter N, Williams B, Kirby M, Patel V, et al. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart. 2005;91.

19. McGeechan K, Macaskill P, Irwig L, Liew G, Wong TY. Assessing New Biomarkers and Predictive Models for Use in Clinical Practice A Clinician's Guide. Archives of Internal Medicine. 2008;168(21):2304-10.

20. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. European Journal of Clinical Investigation. 2012;42(2):216-28.

21. Janes H, Pepe MS, Gu W. Assessing the Value of Risk Predictions by Using Risk Stratification Tables. Annals of Internal Medicine. 2008;149(10):751-W162.

22. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction - Response. Circulation. 2007;116(6):E134-E.

23. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007;115(7):928-35.

24. Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. Medical Decision Making. 2006;26(6):565-74.

25. National Institute for Health and Care Excellence. Osteoporosis: fragility fracture risk. Costing Report: Implementing NICE Guidance. National Institute for Health and Care Excellence, 2012 Contract No.: CG146.

26. National Institute for Health and Care Excellence Guideline Development Group. Costing report: Lipid Modification Implementing the NICE Guideline on Lipid Modification (CG181) National Institute for Health and Care Excellence, 2014 Contract No.: CG181.

27. Hlatky MA, Greenland P, Arnett DK, Ballantyne C, M., Criqui MH, Elkind MS, et al. Criteria for Evaluation of Novel Markers of Cardiovascular Risk: A Scientific Statement From the American Heart Association (vol 119, pg 2408, 2009). Circulation. 2009;119(25):E606-E.

28. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375.

29. Plevritis SK, Kurian AW, Sigal BM, Daniel BL, Ikeda DM, Stockdale FE, et al. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. JAMA. 2006;295(20):2374-84.

30. Dent T, Wright C. Quality standards in risk prediction - summary of an expert meeting. PHG Foundation, 2011.

31. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score.[Erratum appears in JAMA. 2007 Apr 4;297(13):1433]. JAMA. 2007;297(6):611-9.

32. Folsom AR, Chambless LE, Ballantyne CM, Coresh J, Heiss G, Wu KK, et al. An assessment of incremental coronary risk prediction using C-reactive protein and other novel risk markers: the atherosclerosis risk in communities study. Archives of Internal Medicine. 2006;166(13):1368-73.

33. Quante AS, Whittemore AS, Shriver T, Strauch K, Terry MB. Breast cancer risk assessment across the risk continuum: genetic and nongenetic risk factors contributing to differential model performance. Breast Cancer Research. 2012;14(6):R144.

34. Janssens ACJW, Khoury MJ. Assessment of Improved Prediction Beyond Traditional Risk Factors When Does a Difference Make a Difference? Circulation-Cardiovascular Genetics. 2010;3(1):3-5.

35. Greenland S. The need for reorientation toward cost-effective prediction: Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine. Statistics in Medicine. 2008;27(2):199-206.

36. de la Iglesia B, Potter JF, Poulter NR, Robins MM, Skinner J. Performance of the ASSIGN cardiovascular disease risk score on a UK cohort of patients from general practice. Heart. 2011;97(6):491-9.

37. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605.

38. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594.

39. Siontis GCM, Tzoulaki I, Siontis KC, Ioannidis JPA. Comparisions of established risk prediction models for cardiovascular disease: systematic review. BMJ. 2012;344:3318.

40. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. American Journal of Epidemiology. 2004;159(9):882-90.

41. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691-8.

42. Altman DG, Royston P. What do we mean by validating a prognostic model? Statistics in Medicine. 2000;19(4):453-73.

43. Pencina MJ, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics in Medicine. 2008;27(2):157-72.

44. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer Series in Statistics, editor: Springer; 2011.

45. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed: Wiley Interscience Publication; 2000. 392 p.

46. Tzoulaki I, Liberopoulos G, Ioannidis JPA. Use of reclassification for assessment of improved prediction: an empirical evaluation. International Journal of Epidemiology. 2011;40(4):1094-105.

47. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models A Framework for Traditional and Novel Measures. Epidemiology. 2010;21(1):128-38.

48. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis: New York: Springer; 2001.

49. Hanley JA, McNeil BJ. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. Radiology. 1982;143(1):29-36.

50. Vickers AJ, Cronin AM. Everything You Always Wanted to Know About Evaluating Prediction Models (But Were Too Afraid to Ask). Urology. 2010;76(6):1298-301.

51. Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98(9):683-90.

52. Cook NR. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. Clinical Chemistry. 2008;54(1):17-23.

53. Jousilahti P, Vartiainen E, Tuomilehto J, Puska P. Sex, age, cardiovascular risk factors, and coronary heart disease: a prospective follow-up study of 14 786 middle-aged men and women in Finland. Circulation. 1999;99(9):1165-72.

54. Jakobsen A, Laurberg P, Vestergaard P, Andersen S. Clinical risk factors for osteoporosis are common among elderly people in Nuuk, Greenland. International Journal of Circumpolar Health. 2013;72:19596.

55. Pencina MJ, D'Agostino RB, Vasan RS. Statistical methods for assessment of added usefulness of new biomarkers. Clinical Chemistry and Laboratory Medicine. 2010;48(12):1703-11.

56. Greenland P, O'Malley PG. When is a new prediction marker useful? A consideration of lipoprotein-associated phospholipase A(2) and C-reactive protein for stroke risk. Archives of Internal Medicine. 2005;165(21):2454-6.

57. Sever PS, Dahlof B, Poulter NR, Wedel H, Beevers G, Caulfield M, et al. Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial--Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. Lancet. 2003;361(9364):1149-58.

58. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HA, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet. 2004;364(9435):685-96.

59. Cook NR, Ridker PM. Advances in Measuring the Effect of Individual Predictors of Cardiovascular Risk: The Role of Reclassification Measures. Annals of Internal Medicine. 2009;150(11):795-802.

60. Pencina M, D'Agostino R, Steyerberg EW. Authors Reply: Comments on 'Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers' by M. J. Pencina, R. B. D'Agostino, Sr. and E. W. Steyerberg reply. Statistics in Medicine. 2012;31(1):96-7.

61. Cook NR. Clinically relevant measures of fit? A note of caution. American Journal of Epidemiology. 2012;176(6):488-91.

62. Stevenson MD, Davis SE, Kanis JA. The hospitalisation costs and out-patient costs of fragility fractures. Women's Health Medicine. 2006;3(4):149-51.

63. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed: Oxford University Press; 2005.

64. Grossman M. The Human Capital Model. Handbook of Health Economics. 2000;1(Part A):347-408.

65. Boysen M, Robertson J, Woodward L, Longson C, Stevens A, Adler A, et al. Guide to the Methods of Technology Appraisal. National Institute for Health and Care Excellence, 2013 Contract No.: PMG9.

66. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value in Health. 2013;16(2):231-50.

67. HM Treasury. The Green Book: Appraisal and Evaluation in Central Government. In: HM Treasury, editor. 2014. p. 26-8.

68. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32-5.

69. Fleurence RL, Hollenbeak CS. Rates and probabilities in economic modelling: transformation, translation and appropriate application. Pharmacoeconomics. 2007;25(1):3-6.

70. Curtis L. Unit Costs of Health and Social Care 2014. University of Kent: Personal Social Services Research Unit, 2014.

71. Townsend N, Wickramsinghe K, Bhatnagar P, Nichols M, Leal J, Juengo-Fernandez R, et al. Coronary heart disease statistics 2012 edition. London: British Heart Foundation, 2012.

72. Townsend N, Williams J, Bhatnagar P, Wickramsinghe K, Rayner M. Cardiovascular Disease Statistics 2014. London: British Heart Foundation, 2014.

73. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. Plos Medicine. 2006;3(11):e442.

74. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. BMJ. 2012;344:e4181.

75. Mendis S, Puska P, Norrving B, Editors. Global Atlas on Cardiovascular Disease Prevention and Control. Geneva: World Health Organization, 2011.

76. Mosca L, Barrett-Connor E, Wenger NK. Sex/gender differences in cardiovascular disease prevention: what a difference a decade makes. Circulation. 2011;124(19):2145-54.

77. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364(9438):937-52.

78. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ. 2008;336(7659):1475-82.

79. Kane R, Wellings K, Free C, Goodrich J. Uses of routine data sets in the evaluation of health promotion interventions: opportunities and limitations. Health Education. 2000;100(1):33-41.

80. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

81. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I. Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidemiology & Drug Safety. 2010;19(6):618-26.

82. Tunstall-Pedoe H. Deprivation and Health in Scotland. Postgraduate Medical Journal. 1993;69(814).

83. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ. 2007;335(7611):136-41.

84. Cambien F, Tiret L. Genetics of cardiovascular diseases: from single mutations to the whole genome. Circulation. 2007;116(15):1714-24.

85. Sandhu M, Wood A, Young E. Genomic risk prediction. Lancet. 2010;376(9750):1366-7.

86. Scheuner MT, Setodji CM, Pankow JS, Blumenthal RS, Keeler E. General Cardiovascular Risk Profile identifies advanced coronary artery calcium and is improved by family history: the multiethnic study of atherosclerosis. Circulation-Cardiovascular Genetics. 2010;3(1):97-U302.

87. Scheuner MT, Whitworth WC, McGruder H, Yoon PW, Khoury MJ. Familial risk assessment for early-onset coronary heart disease. Genetics in Medicine. 2006;8(8):525-31.

88. Hasanaj Q, Wilson BJ, Little J, Montazeri Z, Carroll JC, Screening ObotCETiG. Family History: Impact on Coronary Heart Disease Risk Assessment beyond Guideline-Defined Factors. Public Health Genomics. 2013;16:208-14.

89. Wilson BJ, Qureshi N, Santaguida P, Little J, Carroll JC, Allanson J, et al. Systematic Review: Family History in Risk Assessment for Common Diseases. Annals of Internal Medicine. 2009;151(12):878-85.

90. Hawe E, Talmud PJ, Miller GJ, Humphries SE. Family history is a coronary heart disease risk factor in the second Northwick Park Heart Study. Annals of Human Genetics. 2003;67:97-106.

91. Butterworth A. Family history as a risk factor for common, complex disease. An independent, epidemiological assessment of the evidence for familial risk of disease. Cambridge: Public Health Genetics Unit., 2007.

92. Valdez R, Yoon PW, Qureshi N, Green RF, Khoury MJ. Family history in public health practice: a genomic tool for disease prevention and health promotion. Annual Review of Public Health. 2010;31:69-87.

93. Dhiman P, Kai J, Horsfall L, Walters K, Qureshi N. Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment. PLoS ONE. 2014;9(1):e81998.

94. Slack J, Evans KA. The increased risk of death from ischaemic heart disease in first-degree relatives of 121 men and 96 women with ischaemic heart disease. Journal of Medical Genetics. 1966;3:239-57.

95. Friedlander Y, Kark JD, Stein Y. Family history of myocardial infarction as an independent risk factor for coronary heart disease. British Heart Journal. 1985;53:382-7.

96. Thomas CB, Cohen BH. The familial occurrence of hypertension and coronary artery disease, with observations concerning obesity and diabetes. Annals of Internal Medicine. 1955;42:90-127.

97. Lloyd-Jones DM, Nam B-H, D'Agostino RB, Sr., Levy D, Murabito JM, Wang TJ, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. JAMA. 2004;291(18):2204-11.

98. Hunt SC, Williams RR, Barlow GK. A comparison of positive family history definitions for defining risk of future disease.[Erratum appears in J Chronic Dis 1987;40(4):369]. Journal of Chronic Diseases. 1986;39(10):809-21.

99. Silberberg JS, Wlodarczyk J, Fryer J, Ray CD, Hensley MJ. Correction for biases in a population-based study of family history and coronary heart disease. The Newcastle Family History Study I. American Journal of Epidemiology. 1998;147(12):1123-32.

100. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiology and Drug Safety. 2009;18(8):704-7.

101. Tzoulaki I, Molokhia M, Curcin V, Little MP, Millett CJ, Ng A, et al. Risk of cardiovascular disease and all cause mortality among patients with type 2 diabetes prescribed oral antidiabetes drugs: retrospective cohort study using UK general practice research database. BMJ. 2009;339:b4731.

102. Welch C, Petersen I, Walters K, Morris RW, Nazareth I, Marston L, et al. Two-stage method to remove population-and individual-level outliers from longitudinal data in a primary care database. Pharmacoepidemiology and Drug Safety. 2012;21:725-32.

103. Boniface S, Bridges S, Craig R, Darton R, Fuller E, Hancock R, et al. Health, social care and lifestyles. Leeds: Health and Social Care Information Centre., 2011.

104. NatCen Social Research and University College London. Department of Epidemiology and Public Health, Health Survey for England, 2011 [computer file]. Colchester, Essex: UK Data Archive [distributor]; 2011.

105. Rubin DB. Multiple imputation for non-response in surveys. 2004 ed. New York: John Wiley & Sons, Inc.; 1987.

106. StataCorp. Stata Multiple-Imputation Reference Manual. College Station, TX: Stata Press; 2013.

107. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine. 1996;15(4):361-87.

108. Cook NR. Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine. Statistics in Medicine. 2008;27(2):191-5.

109. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statistics in Medicine. 2011;30(1):11-21.

110. NHS Business Services Authority NPS. Electronic Drug Tariff. National Health Service: England and Wales; 2014.

111. Ward S, Lloyd Jones M, Pandor A, Holmes M, Ara R, Ryan A, et al. A systematic review and economic evaluation of statins for the prevention of coronary events. Health Technology Assessment (Winchester, England). 2007;11(14):1-160, iii-iv.

112. Anokye NK, Trueman P, Green C, Pavey TG, Hillsdon M, Taylor RS. The cost-effectiveness of exercise referral schemes. BMC Public Health. 2011;11:954.

113. Trapero-Bertran M. Incorporating external effects in economic evaluation: the case of smoking. London: Brunel University; 2011.

114. Giardina S, Pane B, Spinella G, Cafueri G, Corbo M, Brasseur P, et al. An economic evaluation of an abdominal aortic aneurysm screening program in Italy. Journal of Vascular Surgery. 2011;54(4):938-46.

115. StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.

116. Belin TR, Hu M, Young AS, Grusky O. Using Multiple Imputation to Incorporate Cases with Missing Items in a Mental Health Services Study. Health Services and Outcomes Research Methodology. 2000;1(1):7-22.

117. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. Statistical Methods in Medical Research. 2007;16(3):219-42.

118. Pigott TD. A Review of Methods for Missing Data. Educational Research and Evaluation. 2001;7(4):353-83.

119. He Y. Missing data analysis using multiple imputation: getting to the heart of the matter. Circulation Cardiovascular Quality & Outcomes. 2010;3(1):98-105.

120. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. British Journal of Cancer. 2004;91(1):4-8.

121. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. Journal of Clinical Epidemiology. 2010;63(2):205-14.

122. Marsh D, Currie C, Brown P, Cooper A, Elliott J, Griffiths R, et al. The care of patients with fragility fracture. British Orthopaedic Association, 2007.

123. Burge RT, Worley D, Johansen A, Bhattacharyya S, Bose U. The cost of osteoporotic fractures in the UK: projections for 2000-2020. Journal of Medical Economics. 2001;4:51-62.

124. WHO Scientific Group on the Assessment of Osteoporosis at Primary Health Care Level. WHO Press, Switzerland: World Health Organization, 2004.

125. Surnbo I, Johnell O. Consequences of a hip fracture: a prospective study over 1 year. Osteoporosis International. 1993;3:148-53.

126. Prevention and management of osteoporosis. World Health Organization Technical Report Series. 2003;921:1-164. back cover.

127. Kanis JA. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: Synopsis of a WHO report. Osteoporosis International. 1994;4(6):368-81.

128. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, et al. Clinician's Guide to Prevention and Treatment of Osteoporosis. Washington, DC: 2013.

129. Thomsen K. Odense Fallers and Osteoporosis Study: Assessment of osteoporosis among older women presenting with falls. Denmark: Odense University Hospital; 2014.

130. Nguyen ND, Frost SA, Center JR, Eisman JA, Nguyen TV. Development of a nomogram for individualizing hip fracture risk in men and women. Osteoporosis International. 2007;18(8):1109-17.

131. Hui SL, Slemenda CW, Johnston CC. Age and bone mass as predictors of fracture in a prospective study. Journal of Clinical Investigation. 1988;81(6):1804-9.

132. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. Lancet. 2002;359(9319):1761-7.

133. Dennison E, Cole Z, Cooper C. Diagnosis and epidemiology of osteoporosis. Current Opinion in Rheumatology. 2005;17(4):456-61.

134. Center JR, Bliuc D, Nguyen TV, Eisman JA. Risk of subsequent fracture after low-trauma fracture in men and women. JAMA. 2007;297(4):387-94.

135. Giangregorio LM, Leslie WD, Manitoba Bone Density P. Time since prior fracture is a risk modifier for 10-year osteoporotic fractures. Journal of Bone & Mineral Research. 2010;25(6):1400-5.

136. Johnell O, Kanis JA, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Fracture risk following an osteoporotic fracture. Osteoporosis International. 2004;15(3):175-9.

137. Johnell O, Oden A, Caulin F, Kanis JA. Acute and long-term increase in fracture risk after hospitalization for vertebral fracture. Osteoporosis International. 2001;12(3):207-14.

138. De Vries F, Bracke M, Leufkens HG, Lammers JW, Cooper C, Van Staa TP. Fracture risk with intermittent high-dose oral glucocorticoid therapy. Arthritis & Rheumatism. 2007;56(1):208-14.

139. Kanis JA, Johansson H, Oden A, Johnell O, De Laet C, Eisman JA, et al. A family history of fracture and fracture risk: a meta-analysis. Bone. 2004;35(5):1029-37.

140. Kanis JA, Johansson H, Johnell O, Oden A, De Laet C, Eisman JA, et al. Alcohol intake as a risk factor for fracture. Osteoporosis International. 2005;16(7):737-42.

141. Kanis JA, Johansson H, Oden A, De Laet C, Johnell O, Eisman JA, et al. A meta-analysis of milk intake and fracture risk: low utility for case finding. Osteoporosis International. 2005;16(7):799-804.

142. Kanis JA, Johansson H, Oden A, Johnell O, de Laet C, Melton IL, et al. A meta-analysis of prior corticosteroid use and fracture risk. Journal of Bone & Mineral Research. 2004;19(6):893-9.

143. Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. Bone. 2004;35(2):375-82.

144. Kanis JA, Johnell O, Oden A, Johansson H, De Laet C, Eisman JA, et al. Smoking and fracture risk: a meta-analysis. Osteoporosis International. 2005;16(2):155-62.

145. De Laet C, Kanis JA, Oden A. Body mass index as a prediction of fracture risk. A meta-analysis. Osteoporosis International. 2005;16:1330-8.

146. Johansen A. QFracture is better than FRAX tool in assessing risk of hip fracture. BMJ. 2012;345:e4988.

147. Hippisley-Cox J, Coupland C. Validation of QFracture compared with FRAX: Analysis prepared for NICE 2011. Nottingham: University of Nottingham, 2011.

148. Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J, et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. Osteoporosis International. 2007;18(8):1033-46.

149. Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. BMJ. 2011;342:d3651.

150. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart. 2008;94(1):34-9.

151. National Institute for Health and Clinical Excellence. Osteoporosis: fragility fracture risk. Short clinical guideline—evidence and recommendation. London: National Clinical Guidance Centre, 2012 Contract No.: CG146.

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

153. WHO Study Group. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organization Technical Report Series. 1994;843:1-129.

154. McCloskey E, Johansson H, Oden A, Kanis JA. Fracture Risk Assessment. Clinical biochemistry. 2012;45(12):887-93.

155. Briot K, Paternotte S, Kolta S, Eastell R, Felsenberg D, Reid DM, et al. FRAX: Prediction of Major Osteoporotic Fractures in Women from the General Population: The OPUS Study. PLoS ONE. 2013;8(12):e83436.

156. Barry P, Aspray T, Briers K, Collins GS, Compston J, Dockery F, et al. Osteoporosis: assessing the risk of fragility fracture. 2012 Contract No.: CG146.

157. Andersen S, Laurberg P. Age discrimination in osteoporosis screening--data from the Aalborg University Hospital Record for Osteoporosis Risk Assessment (AURORA). Maturitas. 2014;77(4):330-5.

158. Andersen S, Laurberg P. Age impact on clinical risk factors does not justify the age related change in referral pattern for osteoporosis assessment—Data from the Aalborg University Hospital Record for Osteoporosis Risk Assessment (AURORA). Maturitas. 2015;80(3):302-7.

159. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scandinavian Journal of Public Health. 2011;39(7 Suppl):30-3.

160. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiology & Drug Safety. 2009;18(8):704-7.

161. National Institute for Health and Care Excellence. Alendronate, etidronate, risedronate, raloxifene, strontium ranelate and teriparatide for the secondary prevention of osteoporotic fragility fractures in postmenopausal women: Costing Template and report - Implementing NICE Guidance. London: NICE, 2008.

162. Kanis JA, Oden A, Johnell O, Jonsson B, de Laet C, Dawson A. The burden of osteoporotic fractures: a method for setting intervention thresholds. Osteoporosis International. 2001;12(5):417-27.

163. Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA. 1998;280(24):2077-82.

164. Cooper C, Coupland C, Mitchell M. Rheumatoid arthritis, corticosteroid therapy and hip fracture. Annals of the Rheumatic Diseases. 1995;54(1):49-52.

165. Sandhu SK, Nguyen ND, Center JR, Pocock NA, Eisman JA, Nguyen TV. Prognosis of fracture: evaluation of predictive accuracy of the FRAX algorithm and Garvan nomogram. Osteoporosis International. 2010;21(5):863-71.

166. Cummins NM, Poku EK, Towler MR, O'Driscoll OM, Ralston SH. Clinical Risk Factors for Osteoporosis in Ireland and the UK: a Comparison of FRAX and QFracture Scores. Calcified Tissue International. 2011;89(2):172-7.

167. Donaldson MG, Palermo L, Schousboe JT, Ensrud KE, Hochberg MC, Cummings SR. FRAX and risk of vertebral fractures: the fracture intervention trial. Journal of Bone & Mineral Research. 2009;24(11):1793-9.

168. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. British Journal of Cancer. 2004;91(7):1229-35.

169. Demissie S, LaValley MP, Horton NJ, Glynn RJ, Cupples LA. Bias due to missing exposure data using complete-case analysis in the proportional hazards regression model. Statistics in Medicine. 2003;22(4):545-57.

170. Arbukle JL. Full information estimation in the presence of incomplete data. In: Marcoulides, G. A., Schumacker, R. E. Advanced structural equation modeling: Issues and Techniques. Mahwah, NJ: Lawrence Erlbaum Associates.; 1996. p. 243-77.

171. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry. 1993;39(4):561-77.

172. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Annals of Internal Medicine. 2014;160(2):122-31.

173. The CDC Diabetes Cost-Effectiveness Study Group. The Cost-effectiveness of Screening for Type 2 Diabetes. JAMA. 1998;280(20):1757-63.

174. Burke MJ, Shenton RC, Taylor MJ. The economics of screening infants at risk of hearing impairment: an international analysis. International Journal of Pediatric Otorhinolaryngology. 2012;76(2):212-8.

175. Svatek RS, Sagalowsky AI, Lotan Y. Economic impact of screening for bladder cancer using bladder tumor markers: a decision analysis. Urologic Oncology. 2006;24(4):338-43.

176. Greenland S. The need for reorientation toward cost-effective prediction: comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929).[Erratum appears in Stat Med. 2008 Jan 30;27(2):316]. Statistics in Medicine. 2008;27(2):199-206.

177. Slejko JF, Page RL, 2nd, Sullivan PW. Cost-effectiveness of statin therapy for vascular event prevention in adults with elevated C-reactive protein: implications of JUPITER. Current Medical Research & Opinion. 2010;26(10):2485-97.

178. Padula WV, Mishra MK, Makic MB, Sullivan PW. Improving the quality of pressure ulcer care with prevention: a cost-effectiveness analysis. Medical Care. 2011;49(4):385-92.

179. McComb MN, Collins CD. Comparative cost-effectiveness of alternative empiric antimicrobial treatment options for suspected enterococcal bacteremia. Pharmacotherapy:The Journal of Human Pharmacology & Drug Therapy. 2014;34(6):537-44.

180. Lopez-Bastida J, Hart W, Garcia-Perez L, Linertova R. Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease. Journal of Alzheimer's Disease. 2009;16(2):399-407.

181. Jensen CE, Sorensen P, Petersen KD. In Denmark kidney transplantation is more cost-effective than dialysis. Danish Medical Journal. 2014;61(3):A4796.

182. Linssen AM, Anteunis LJC, Joore MA. The Cost-Effectiveness of Different Hearing Screening Strategies for 50- to 70-Year-Old Adults: A Markov Model. Value in Health. 2015;In Press. Available online 2015. doi:10.1016/j.jval.2015.03.1789.

183. Briggs A, Schulfer M, Claxton K. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press Inc.; 2006.

184. Marshall D, Simpson KN, Earle CC, Chu CW. Economic decision analysis model of screening for lung cancer. European Journal of Cancer. 2001;37(14):1759-67.

185. Collins GS, Altman DG. An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. BMJ. 2010;340.

186. Antolini L, Boracchi P, Biganzoli E. A time-dependent discrimination index for survival data. Statistics in Medicine. 2005;24(24):3927-44.

187. McFadden E, Stevens R, Glasziou P, Perera R. Implications of lower risk thresholds for statin treatment in primary prevention: Analysis of CPRD and simulation modelling of annual cholesterol monitoring. Preventive Medicine. 2015;70:14-6.

188. Criqui MH, Fronek A, Klauber MR, Barrett-Connor E, Gabriel S. The sensitivity, specificity, and predictive value of traditional clinical evaluation of peripheral arterial disease: results from noninvasive testing in a defined population. Circulation. 1985;71(3):516-22.

Appendices

Appendix A. Search terms for Literature Review

Risk

Prediction

Risk assessment

Risk models

Risk scores

Risk tools

Risk prediction

Statistical Models

Predictive value of tests

Novel/New Risk factor/marker

Discrimination

Calibration

Reclassification

Model performance

Predictors

Biomarkers

Prognosis/prognostic models

Additional marker/risk factor

Incremental improvement

Screening

Sensitivity/Specificity

ROC/ Receiver Operating Characteristic Curve

AUC/Area under the curve

Risk stratification

Risk classification

Reclassification tables

NRI/Net reclassification improvement/index

Decision/making

Decision trees

Decision curve analysis

Appendix B. Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment (93).

OPEN ACCESS Freely available online

PLOS ONE

Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment

Paula Dhiman[1], Joe Kai[1], Laura Horsfall[2], Kate Walters[2], Nadeem Qureshi[1]*

1 School of Medicine, Division of Primary Care, University of Nottingham, University Park, Nottingham, United Kingdom, 2 The Research Department of Primary Care and Population Health, University College of London (UCL), London, United Kingdom

Abstract

Background: The potential to use data on family history of premature disease to assess disease risk is increasingly recognised, particularly in scoring risk for coronary heart disease (CHD). However the quality of family health information in primary care records is unclear.

Aim: To assess the availability and quality of family history of CHD documented in electronic primary care records

Design: Cross-sectional study

Setting: 537 UK family practices contributing to The Health Improvement Network database.

Method: Data were obtained from patients aged 20 years or more, registered with their current practice between 1st January 1998 and 31st December 2008, for at least one year. The availability and quality of recorded CHD family history was assessed using multilevel logistic and ordinal logistic regression respectively.

Results: In a cross-section of 1,504,535 patients, 19% had a positive or negative family history of CHD recorded. Multilevel logistic regression showed patients aged 50–59 had higher odds of having their family history recorded compared to those aged 20–29 (OR:1.23 (1.21 to 1.25)), however most deprived patients had lower odds compared to those least deprived (OR: 0.86 (0.85 to 0.88)). Of the 140,058 patients with a positive family history recorded (9% of total cohort), age of onset was available in 45%; with data specifying both age of onset and relative affected available in only 11% of records. Multilevel ordinal logistic regression confirmed no statistical association between the quality of family history recording and age, gender, deprivation and year of registration.

Conclusion: Family history of CHD is documented in a small proportion of primary care records; and where positive family history is documented the details are insufficient to assess familial risk or populate cardiovascular risk assessment tools. Data capture needs to be improved particularly for more disadvantaged patients who may be most likely to benefit from CHD risk assessment.

Citation: Dhiman P, Kai J, Horsfall L, Walters K, Qureshi N (2014) Availability and Quality of Coronary Heart Disease Family History in Primary Care Medical Records: Implications for Cardiovascular Risk Assessment. PLoS ONE 9(1): e81998. doi:10.1371/journal.pone.0081998

Editor: Heiner K. Berthold, Charité University Medicine Berlin, Germany

Received May 25, 2013; **Accepted** October 20, 2013; **Published** January 9, 2014

Copyright: © 2014 Dhiman et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The study was supported by pump-primed funding from the Faculty of Medicine, University of Nottingham. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: nadeem.qureshi@nottingham.ac.uk

Introduction

Family history has long been recognised as an important risk factor for a range of common chronic diseases, such as cancers, diabetes, and coronary heart disease (CHD) [1,2]. The recognised predictive accuracy of family history of CHD has opened international interest in integrating family history information into routine clinical practice [3–5]. Currently, several risk prediction models incorporate family history. In the UK, these include cardiovascular risk assessment tools such as ASSIGN [6] and QRISK2 [7]. With the recognition of CHD as the leading cause of death, such assessment has been implemented through national guidance [8] and policy initiatives [9].

In many primary care computer systems, automated risk prediction tools can be populated from coded data in clinical records and provide an ideal environment to collect and assess patient CHD risk. In primary care, the initial consultation and/or health screening questionnaire at new registration is a well-recognised opportunity to collect family history information [10–12]; and the 1990 UK Family Practice contract has encouraged enquiry about inherited predisposition [13].

Data extracted from primary care computer systems have previously been used to identify the extent of missing data for

standard cardiovascular risk factors [14]. A similar approach could be used to explore the nature of the family history information recorded. The potential value of electronic health records to assimilate important information on family history has been recognised [15].

Level of recording of family history has been collated for cancer and CHD in primary care-based cross-sectional surveys and in community-based epidemiological studies [16,17]. Similarly, web based tools have been developed to collate this information [18]. Thus far, quality of family history information has not been evaluated in primary care databases.

The quality of data entry in primary care computer systems has been evaluated for medication recording [19], but is yet to be assessed for family history. Accurate assessment of the relationship of the affected relative to the patient, and the age of onset of the condition will indicate quality of data recording necessary to assess the CHD risk associated with a family history [20,21].

This study aimed to assess the availability and quality of family history of CHD data documented in electronic primary care records at patient registration, and any associations with patients' socio-economic background.

Materials and Methods

Design and Setting

This was a cross-sectional study, using data from The Health Improvement Network (THIN) database, a large anonymised primary care database. At the time the study was conducted, the THIN database contained data from 537 practices, on 10.9 million patients, representing approximately 6% of the UK primary care population. Data in this database has a similar distribution to the UK population. Data is entered using the VISION family practice computer software, by the family physician, nurses and administrative staff during routine clinical practice; and is recorded using a hierarchal coding system (Read Codes), which maps onto the ICD-10 coding system, and free text entries. As well as medical diagnoses and symptoms, the THIN database also includes information on prescriptions, referrals, tests and risk factors relating to lifestyle behaviour, and is linked to UK Census data on local area deprivation. Recording of consultations and prescriptions is comparable to national consultation and prescription statistics [22]. Disease rates have been calculated from the data and compared to externally generated rates in validation studies [23].

Study Population

Eligible patients were aged 20 or greater, and registered with their practice during the decade 1st January 1998 to 31st December 2008 for at least one year. Patients without gender specified or patients registered before the practice began using the VISION medical record software were excluded. Temporary patients are not formally registered with UK Family Practice and thus data input is suboptimal in these patients. In line with established methodology in primary care database research, to avoid the artificial inflation of the denominator with temporary patients, patients who were not registered with the practice for at least one year were removed from the cohort. To identify the level of family history recording, the cohort comprised new patient registrants with any family history of CHD. To assess quality of data, only patients who had presence of a positive FH of CHD recorded were included.

Outcomes

Code lists for family history (FH) were created using recognised primary care database methodology [24] which identified FH recording in the medical records and additional health detail records. Code lists used are available from authors on request. Anonymised and coded free text was also investigated to derive information on specific first and second degree relatives and age of onset.

For recording of CHD the outcome measure was a dichotomised variable indicating whether any FH (positive or negative) of CHD was recorded or not. A positive FH of CHD would encompass any description indicating information on FH of CHD was present, whilst a coded entry 'no family history of CHD' was classified as a negative family history.

The outcome for the quality of data was restricted to those with a positive FH of CHD. It was then sub-classified into four categories: FH with no further details identified; FH with relative identified; FH with age of onset identified; and FH with details on both relative and age of onset.

Covariates

For both outcomes the patient's gender, age (10 year age groups), level of deprivation and year of practice registration (1st Jan 1999–31st Dec 2008) were included as covariates in our study. The level of deprivation was defined using the Townsend Index, derived from the patient's postcode and combining 4 census variables: car ownership, household size, owner occupation, employment status [25]. Each postal enumeration district (around 150 households) is assigned a Townsend Deprivation score. These are divided into national quintiles and in the THIN database patients are assigned a quintile score according to the enumeration district in which they live. This score ranges from 1–5, where a higher score indicates greater deprivation.

Statistical Analyses

Initial descriptive presentation of the covariates included numbers (%) for categorical variables; mean (SD) for normally distributed variables; and median (IQR) for non – normally distributed variables. Associations between groups (categorical data, such as gender) were tested for, using Chi-squared tests, whilst for non-normal ordinal continuous data (such as Townsend Score), Kruskal Wallis tests were used.

Due to the hierarchical structure of the data where patients were clustered by practice, a multilevel approach was adopted for the analyses of both objectives. A two-level model was used, allowing for a random intercept by the cluster variable (practice). An unadjusted and adjusted analysis was performed. The latter adjusted for all the covariates in the study.

A multilevel logistic regression analysis was used to assess recording of FH of CHD data, odds ratios indicating the odds of having a FH recorded at practice registration versus not having FH recorded. Whilst the quality of family history data was assessed by performing a multilevel ordinal logistic regression analysis. For this analysis having family history recordings with both items (age of onset and relatives affected) was given highest quality ranking, whilst having a single item (age of onset or relative) were considered of equal quality. The poorest quality was recording FH of CHD with no further details specified. Results from this analysis gave the odds of having the best FH quality (FH of CHD with relative and age specified) versus the two lower qualities combined (FH of CHD with age or relative given, and FH of CHD alone).

Analysis also estimated practice variation, presented as a value of the variance and standard error at level 2 (practice) and also a

percentage of the remaining unexplained variance of the final model, using the variation partition coefficient (VPC). To calculate the VPC, the variance at level 1 (patient level) is assumed to be 3.29 (the variance of a standard logistic regression). As the models produced are two level random intercept models the VPC can also be interpreted as the Intra Class Correlation (ICC), the amount of variance attributable to practices. Missing values were presented in the descriptive analysis, but were excluded from the logistic regression analysis. Analyses were performed using Stata v11 [26] and MLwiN v2.23 statistical software [27].

Ethics statement

The Health Improvement Network (THIN) scheme of providing anonymised data to researchers was approved by the NHS South-East Multi-centre Research Ethics Committee and the CSD Medical Research Scientific Review Committee approved the present study.

Results

Recording of family history of CHD

Of the total cohort, 283,715 (18.86%) patients had FH of CHD recorded (positive or negative); half (49.37%) of these had a positive FH of CHD recorded, representing 9.31% of the total cohort. Socio-demographic characteristics of the sample are provided in Table 1.

Compared to the total cohort, proportionally fewer men had any (positive or negative) recording of FH of CHD (46.76%, $p = 0.16$), or of a positive FH of CHD (44.87%). Further, patients with a recorded FH were slightly older ($p<0.001$), more recently registered with the practice ($p<0.001$) and had lower deprivation scores ($p = 0.007$).

The proportion of patients with any FH of CHD recorded reduced over time from 20.55% of patients with a record of FH of CHD in 1999 to 17.82% in 2008. The most prevalent recording of any FH was in the age groups 50–59 and 60–69 (Figure 1). This pattern remained the same when exploring FH recording for each age group by year of registration. Patients aged 80 and above had a consistently low recording of FH. On the other hand, no pattern was found between the level of recording and patient deprivation across different years of registration, for example in 2007 the highest level of recording was for those patients in Townsend Score 5, and in 2003 & 1999 the highest score was for patients with Townsend Score 2. Multivariate analysis confirmed that the level of recording fell in patients aged over 70 (OR: 0.66 (0.64 to 0.68)) and with increasing deprivation score (Table 2). For example, patients with a Townsend Score of 5 (most deprived) had OR: 0.86 (95%CI: 0.85 to 0.88) of having their FH recorded compared to patients with a Townsend score of 1 (least deprived). Further, women were slightly more likely to have a FH recording (OR: 1.07 (95%CI: 1.06 to 1.08).

Multivariate analysis also demonstrated marginally better recording of FH in years prior to 2008, but no specific trend with time was found. Analysis indicated that after adjusting for gender; age; deprivation; and year of registration, 45.26% (VPC: 2.72 (SE: 0.13)) of the unexplained variance of having FH of CHD recorded is due to between practice variation.

Quality of family history recording in primary care computer records

Of the 1,504,535 total cohort, 140,058 (9.31%) had presence of a positive FH of CHD recorded. Of these, 33.34% (95% CI: 33.09% to 33.59%) had a FH of CHD recorded with no further information; 10.77% (95% CI: 10.61% to 10.93%) had FH of CHD recorded with information on the affected relative; 45.05% (95% CI: 44.78% to 45.30%) had FH of CHD recorded with the age of onset specified; and 10.85% (95% CI: 10.69% to 11.01%) had FH of CHD recorded with both the relative and age of onset specified.

When analysed in a multilevel adjusted ordinal logistic regression, (collapsing FH categories, age of onset category and relative affected, into a single category) there was no association between quality of FH and the covariates: year of registration, deprivation score and patient's age. Further, as demonstrated in Figure 2, across all of the covariates the highest availability of FH recording was where no further details are specified or age of onset is also identified with the FH; with poor levels of recording in other categories (relative noted, and both relative and age of onset noted).

The unexplained variation at practice level for the final model of the quality of FH recording is 50.67% (VPC: 3.38 (SE: 0.10)), after controlling for gender, age, deprivation and year of registration.

Discussion

Summary of main findings

We have found CHD family history recording in UK primary care is low, and, where a positive history is recorded, it often lacks details to adequately assess familial risk. Only 19% of patients at 537 practices (with in excess of 3 million actively registered patients) had the presence of any family history of CHD recorded, in the decade to 2008. Presence of a positive family history of CHD was noted in 9% of patients, but data combining age of onset and specified relative affected was available in only 11% of these records.

Strengths and limitations

To our knowledge, this is the first study to also explore nature of the family history in primary care computer records, providing a comprehensive cross-section of their status from 1998 to 2008 in the UK.

The THIN database is derived from 537 practices and contains data on 6.2% of the UK population. Data from this database is generalisable to the UK demographic for age and gender distribution. Crude prevalence rates for common chronic diseases in THIN are also similar to national estimates; for example, the crude prevalence of coronary heart disease in THIN was 3.9% compared to 3.7% in the UK national QOF data [28].

However the THIN database contains slightly more affluent patients (23.5% classified least deprived compared to 20% nationally). This may have caused some overestimation of the total proportions we found with FH recorded, but differences are marginal. Further, the THIN database contains fewer patients aged less than 25 years. This does not negate the relationship between age and extent of family history recording. A further limitation of the THIN database is that the Townsend Deprivation Score is not recorded individually, but is based on the patient's residential postcode (the approximately 150 households in the surrounding neighbourhood) and this deprivation status may not reflect that of the individual.

Although having a detailed family history recorded does not imply an accurate one, previous studies have reported the accuracy of a recorded family history of CHD, compared to direct recall by relatives, is as high as 67% [29]. Further, we anticipate that family history recording would be more frequently taken at registration, compared to other clinical encounters [11]. We also attempted to identify free text entries on family history in

Table 1. Demographic characteristics of patients in the complete sample.

Variable		Total (N = 1,504,535)*		Patients with any FH of CHD recorded (n = 283,715)**		Patients with positive FH of CHD recorded (n = 140,058)**	
N (%)							
Gender	Male	703,498	46.8%	132,151	46.6%	62,825	44.9%
Age (years)***	20–29	550,611	36.6%	96,225	33.9%	36,577	26.1%
	30–39	395,708	26.3%	77,306	27.3%	37,204	26.6%
	40–49	205,697	13.7%	44,061	15.5%	26,634	19.0%
	50–59	134,696	9.0%	31,801	11.2%	20,645	14.7%
	60–69	85,901	5.7%	19,726	7.0%	12,062	8.6%
	70–79	55,845	3.7%	9,677	3.4%	5,144	3.7%
	80 and above	61,610	4.1%	4,866	1.7%	1,767	1.3%
	Missing	14,467	1.0%	53	0.0%	25	0.0%
Townsend Score ****	1	303,573	20.2%	55,767	19.7%	29,951	21.4%
	2	279,432	18.3%	54,259	19.1%	28,009	20.0%
	3	302,589	20.1%	56,361	19.9%	27,437	19.6%
	4	299,197	19.9%	56,562	19.9%	25,947	18.2%
	5	222,960	14.8%	40,761	14.4%	17,889	12.8%
	Missing	100,784	6.7%	20,005	7.1%	10,825	7.7%
Median [IQR]							
Year of Registration		2004 [2002–2007]		2004 [2002–2007]		2005 [2002–2007]	
Age (years)		30–39 [20–29, 40–49]		30–39 [20–29, 40–49]		30–39 [20–29, 50–59]	
Townsend Score		3 [2–4]		3 [2–4]		3 [2–4]	

*Percentages are proportions of the total cohort (1,504,535).
**Percentages are proportions of the cohort of patients with any FH recorded (283,715).
***Percentages are proportions of the cohort of patients with positive FH of CHD recorded (140,058).
****Age was missing for 0.96% of the total sample.
*****Townsend Score was missing for 6.70% of the total sample.
doi:10.1371/journal.pone.0081998.t001

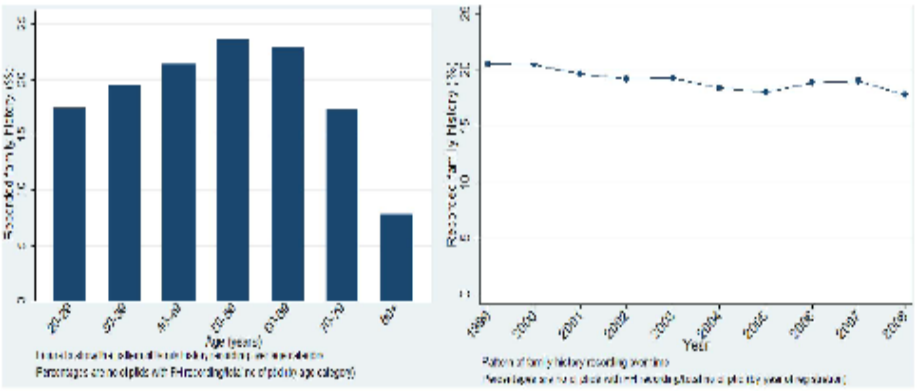


Figure 1. Demographic trends of patients with FH of CHD recorded, by age category and year of registration. Values are out of the total sample 1,389,965 (excl missing values). Reports the proportion of patients with any FH of CHD within each subgroup (age group and year of practice registration).
doi:10.1371/journal.pone.0081998.g001

Table 2. Unadjusted and adjusted odds ratios (95% CI) for recording of FH of CHD, from the multilevel logistic regression model*.

Variable		Unadjusted Analysis	Adjusted Analysis**
Sex	Male	Reference	
	Female	1.02 (1.01 to 1.03)	1.07 (1.06 to 1.08)
Age (years)	20–29	Reference	
	30–39	1.11 (1.10 to 1.12)	1.11 (1.10 to 1.13)
	40–49	1.16 (1.15 to 1.18)	1.17 (1.15 to 1.19)
	50–59	1.23 (1.21 to 1.25)	1.23 (1.21 to 1.25)
	60–69	1.08 (1.05 to 1.10)	1.07 (1.05 to 1.09)
	70–79	0.66 (0.64 to 0.68)	0.66 (0.64 to 0.68)
	80+	0.24 (0.24 to 0.25)	0.24 (0.23 to 0.25)
Townsend Score	1	Reference	
	2	0.96 (0.95 to 0.98)	0.97 (0.95 to 0.98)
	3	0.93 (0.91 to 0.94)	0.94 (0.92 to 0.95)
	4	0.91 (0.90 to 0.93)	0.92 (0.91 to 0.94)
	5	0.86 (0.84 to 0.87)	0.86 (0.85 to 0.88)
Period	1999	1.10 (1.07 to 1.13)	1.09 (1.06 to 1.12)
	2000	1.08 (1.05 to 1.10)	1.07 (1.04 to 1.09)
	2001	1.09 (1.06 to 1.11)	1.08 (1.06 to 1.10)
	2002	1.09 (1.06 to 1.11)	1.08 (1.06 to 1.10)
	2003	1.08 (1.05 to 1.10)	1.07 (1.05 to 1.09)
	2004	1.04 (1.01 to 1.06)	1.03 (1.01 to 1.05)
	2005	0.99 (0.97 to 1.01)	0.99 (0.97 to 1.01)
	2006	1.10 (1.08 to 1.12)	1.09 (1.07 to 1.11)
	2007	1.11 (1.09 to 1.13)	1.11 (1.09 to 1.13)
	2008	Reference	

*Results are derived using 2nd order linearisation and PQL estimation type.
**Adjusted for sex, age category, year of registration and Townsend score.
doi:10.1371/journal.pone.0081998.t002

records, but recognise potentially relevant information to assess its quality may not have been identified, such as, free text on exact relative affected.

Considering the analysis, missing values were handled using a complete case analysis. As a result 8% of the total sample was not included in the analysis; Townsend score was not available for approximately 7% of the sample; and 1% for age. Patients without Townsend score recorded had similar distributions for gender and age; but registered with their practice later. Patients with age recorded had similar distributions for deprivation and year of registration; but had a higher proportion of males compared to the total cohort. The slightly different profile for those with missing values should not impact on the overall study findings.

Existing literature

Secondary findings of similar database studies have shown FH of CHD recording to range between 3.7% [30] and 10.8% [31] of patients, and specifically between 3.7% [30] and 4.4% [32] for studies using the THIN database.

These estimates are significantly lower than our level of recording of 19%, which included patients identified with FH of CHD based on free text extracted from electronic health records. Compared to the low level of recording of family history in primary care records, epidemiological research indicates up to 52% of the population will have a positive family history of CHD, whilst 37% will have a FH of premature CHD [17]. However, similar to this study, research in secondary care indicate poor recording of family history [33].

Case reviews in other countries, looking at small numbers of primary care practices have also reported poor recording of FH in electronic health records. A Swedish study reviewing cardiovascular risk factors in hypertensive patients, showed family history was incompletely recorded (46%) in patient records [34]. Similar studies in New Zealand have indicated varied recording of cardiovascular risk factors. High levels of recording was found of some risk factors such as blood pressure and cholesterol and smoking [35], but levels of FH recording ranged from 19 to 81% in patients aged 60–75 and was dependent on the practice [36].

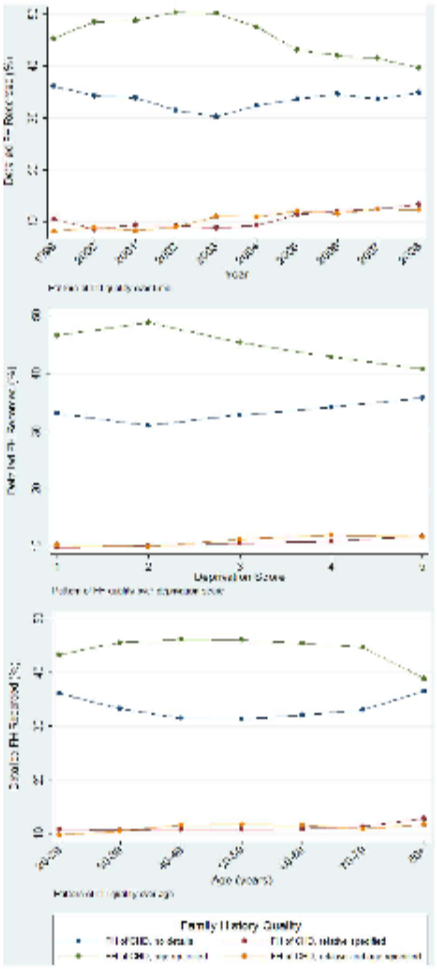


Figure 2. Pattern of FH quality by year of registration, Townsend score and age category (excluding missing values).
doi:10.1371/journal.pone.0081998.g002

In the current study, better recording was anticipated in newly registered patients due to family physicians and practice nurses recognising the value of recording family history on registration [10,11], however the level and quality of recording was still low.

Implications for clinical practice

Availability of family history data was greatest in those aged 50 to 69 years, and was lower in the more socially deprived. Yet recognition of family history of premature CHD in younger age groups and in socially deprived contexts is likely of greater value in targeting interventions to patients with greatest overall risk of premature cardiovascular disease events [6]. This should form a priority for practice.

Optimal interpretation of familial risk requires both age of diagnosis and relative affected [21,37] but this was available in only one in ten patients identified as having a family history of CHD. Poor recording of details on affected relatives may reflect limited choice of history-taking codes in primary care computer systems and the time required in consultations. Currently, electronic health records used in UK primary care lack integrated family history pedigree software, with clinicians relying on unstructured formats to enter FH data [38]. The extent of missing data on affected relatives highlights the possible limitations of risk prediction tools that incorporate family history information from clinical computer medical records. Thus, this study underlines the need to introduce more structured approaches to allow clinicians to improve capture of family history data in clinical practice.

Paucity of family history data may compromise scoring of cardiovascular risk if this is automatically calculated for patients from routinely collected data in primary care records [30,37]. For example, in the Joint British Societies CVD risk assessment tool [37], identifying family history of CHD in a male 1st degree relative under 55 and/or female relative under 65 uplifts the CVD risk score by a factor of 1.5. Applying the current study findings, around 85% of patients in primary care (81% no recording & 4% poor recording) would not have sufficient family history data recorded to benefit from this assessment of their cardiovascular risk. Caution is thus also advised when interpreting studies of risk assessment tools that use family history derived from primary care computer records given the current extent of missing and poor quality data.

A recent randomised trial in primary care suggests the addition of more systematic family history may increase the proportion of individuals identified at high cardiovascular risk for targeting of preventive interventions by up to 40% [39]. We anticipated improved recording of CHD family history over the decade from 1998 given the increased prominence of cardiovascular risk assessment tools in primary care [8], but this was not found. There is a danger that continued incomplete collection of family history data will lead to imprecise cardiovascular risk assessment in those at high familial risk. Implementation of population screening, such as England's Vascular Health Check programme and cardiovascular risk assessment in the UK Primary Care Quality Outcome incentivisation scheme, may help [9].

Future research should explore approaches to improving the more systematic recording of family history of CHD and other common chronic diseases in primary care, together with qualitative exploration of barriers and facilitators to family history collection in practice [40]. Feedback and incentives, for example, may improve data quality [41]. This may enable data on this key risk factor to be used with more confidence and for benefit in patient health outcomes.

Acknowledgments

Acknowledgements: Prof Min Yang (University of Nottingham) for advice on method of analysis.

Author Contributions

Conceived and designed the experiments: NQ PD LH. Performed the experiments: NQ PD LH. Analyzed the data: NQ PD LH JK. Contributed reagents/materials/analysis tools: LH KW. Wrote the paper: PD NQ JK LH KW. Obtained funding: NQ JK. Drafted manuscript: PD NQ JK. Revised the draft manuscript: LH KW.

References

1. Wilson BJ, Qureshi N, Santaguida P, Little J, Carroll JC, et al. (2009) Systematic Review: Family History in Risk Assessment for Common Diseases. Ann Intern Med 151(12): 878–885.
2. Howe E, Talmud PJ, Miller GJ, Humphries SE, Second Northwick Park Heart Study (2003) Family history is a coronary heart disease risk factor in the second Northwick Park Heart Study. Ann Hum Genet 67(Pt 2): 97–106.
3. Berg AO, Baird MA, Botkin JR, Driscoll DA, Fishman PA, et al. (2009) National Institutes of Health State-of-the-Science Conference Statement: Family History and Improving Health. Ann Intern Med 151(12): 872–877.
4. Hall R, Saukko PM, Evans PH, Qureshi N, Humphries SE (2007) Assessing family history of heart disease in primary care consultations: a qualitative study. Fam Pract 24(5): 435–442.
5. Kardia SL, Modell SM, Peyser PA (2003) Family-centered approaches to understanding and preventing coronary heart disease. Am J Prev Med 2003;24(2):143–51.
6. Woodward M, Brindle P, Tunstall-Pedoe H, SIGN group of risk estimation (2007) Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart 93(2): 172–176.
7. Hippisley-Cox J, Coupland C, Robson J, Brindle P (2010) Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. BMJ 341: c6624. DOI: 10.1136/bmj.c6624.
8. Cooper A, Nherera L, Calvert N, O'Flynn N, Turnbull N, et al. (2008) Clinical Guidelines and Evidence Review for Lipid Modification: cardiovascular risk assessment and the primary and secondary prevention of cardiovascular disease London: National Collaborating Centre for Primary Care and Royal College of General Practitioners.
9. Vascular Programme (2009) Putting prevention first – Vascular Checks: Risk assessment and management. London: Department of Health.
10. Bankhead C, Emery J, Qureshi N, Campbell H, Austoker J, et al. (2001) New developments in genetics-knowledge, attitudes and information needs of practice nurses. Fam Pract 18(5): 475–486.
11. Summerton N, Garmod PV (1997) The family history in family practice: a questionnaire study. Fam Pract 14(4): 285–288.
12. Acheson LS, Wiesner GL, Zyzanski SJ, Goodwin MA, Stange KC (2000) Family history-taking in community family practice: Implications for genetic screening. Genet Med 2(3): 180–185.
13. Department of Health (1989) Terms of Service for Doctors in General Practice. London: HMSO.
14. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, et al. (2010) Issues in multiple imputation of missing data for general practice clinical databases. Pharmacoepidemiol Drug Saf 19(6): 618–626.
15. Weiner MG, Lyman JA, Murphy S, Weiner M (2007) Electronic health records: high quality electronic data for higher-quality clinical research. Inform Prim Care 15(2): 121–127.
16. Imperial Cancer Research Fund OXCHECK Study Group (1994) Prevalence of risk factors for heart disease in OXCHECK trial: implications for screening in primary care. BMJ 302(6784): 1057–1060.
17. EUROSPIRE Study Group (1997) A European Society of Cardiology survey of secondary prevention of coronary heart disease: principal results. Eur Heart J 18(10): 1569–1582.
18. O'Neill SM, Rubinstein WS, Wang C, Yoon PW, Acheson LS, et al. (2009) Familial Risk for Common Diseases in Primary Care: The Family Healthware Impact Trial. Am J Prev Med 36(6): 506–514.
19. Pozzi MF, Camporese M, Nobili A, Frau S, Del Zotti F, et al. (2010) Quality assessment of information about medications in primary care electronic patient record (EPR) systems. Inform Prim Care 18(2): 109–116.
20. Hunt SC, Williams RR, Barlow GK (1986) A comparison of positive family history definitions for defining risk of future disease. J Chronic Dis 39(10): 809–821.
21. Bennett RL (2004) The family medical history. Prim Care 31(3): 479–495.
22. Bourke A, Dattani H, Robinson M (2004) Feasibility study and methodology to create a quality-evaluated database of primary care data. Inform Prim Care 12 (3): 171–177.
23. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL (2007) Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiol Drug Saf 16 (4): 393–401.
24. Dave S, Petersen I (2009) Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiol Drug Saf 18(8): 704–707.
25. Townsend P, Phillimore P, Beattie A (1988) Health and Deprivation: Inequality and the North. Routledge, London.
26. StataCorp (2009) Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.
27. Rabash J, Charlton C, Browner WJ, Healy M, Cameron B (2011) MlwiN. 2.23 ed. Centre for Multilevel Modelling, University of Bristol.
28. Blak BT, Thompson M, Dattani H, Bourke A (2011) Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care 19(4): 251–255.
29. Kee F, Tiret L, Robo JY, Nicaud V, McCrum E, et al. (1993) Reliability of reported family history of myocardial infarction. BMJ 307(6918): 1528–1530.
30. de la Iglesia B, Potter JF, Poulter NR, Robins MM, Skinner J (2011) Performance of the ASSIGN cardiovascular disease risk score on a UK cohort of patients from general practice. Heart 97(6): 491–499.
31. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P (2008) Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart 94(1): 34–39.
32. Collins GS, Altman DG (2010) An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. BMJ 340: c2442. DOI: 10.1136/bmj.c2442.
33. Frezzo TM, Rubinstein WS, Dunham D, Ormond KE (2003) The genetic family history as a risk assessment tool in internal medicine. Genet Med 5(2): 84–91.
34. Nielsen I, Kahan T (2010) Cardiovascular risk factors are not treated to target in hypertensive patients in primary care. Blood Press 19(3): 176–181.
35. Rafter N, Wells S, Stewart A, Selak V, Whittaker R, et al. (2008) Gaps in primary care documentation of cardiovascular risk factors. New Z Med J 121(1269): 24–33.
36. Sheerin I, Hamilton G, Humphrey A, Scragg A (2007) Issues in the assessment of cardiovascular risk in selected general practices in Canterbury, New Zealand. New Z Med J 120(1261): U2714.
37. Wood D, Wray R, Poulter N, Williams, Kirby M, et al. (2005) JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart 91 (Suppl 5): v1–52.
38. Scheuner MT, de Vries H, Kim B, Meili RC, Olmsted SH, et al. (2009) Are electronic health records ready for genomic medicine? Genet Med 11(7): 510–517.
39. Qureshi N, Armstrong S, Dhiman P, Saukko P, Middlemass J, et al. (2012) Effect of Adding Systematic Family History Enquiry to Cardiovascular Disease Risk Assessment in Primary Care: A Matched-Pair, Cluster Randomized Trial. Ann Intern Med 156(4): 253–262.
40. Rich EC, Burke W, Heaton CJ, Haga S, Pinsky L, et al. (2004) Reconsidering the family history in primary care. J Gen Intern Med 19(3): 273–280.
41. de Lusignan S, van Weel C (2006) The use of routinely collected computer data for research in primary care: opportunities and challenges. Fam Pract 23(2): 253–263.

Appendix C. Read code and ICD code lists for Cardiovascular Disease

Appendix Table C 1. Read code list for CVD

Read Code	Description
323..00	ECG: myocardial infarction
3232	ECG: old myocardial infarction
3233	ECG: antero-septal infarct.
3234	ECG:posterior/inferior infarct
3235	ECG: subendocardial infarct
3236	ECG: lateral infarction
323Z.00	ECG: myocardial infarct NOS
662K.00	Angina control
662K000	Angina control - good
662K100	Angina control - poor
662K200	Angina control - improving
662Kz00	Angina control NOS
792..00	Coronary artery operations
7920	Saphenous vein graft replacement of coronary artery
7920000	Saphenous vein graft replacement of one coronary artery
7920100	Saphenous vein graft replacement of two coronary arteries
7920200	Saphenous vein graft replacement of three coronary arteries
7920300	Saphenous vein graft replacement of four+ coronary arteries
7920y00	Saphenous vein graft replacement of coronary artery OS
7920z00	Saphenous vein graft replacement coronary artery NOS
7921	Other autograft replacement of coronary artery
7921000	Autograft replacement of one coronary artery NEC
7921100	Autograft replacement of two coronary arteries NEC
7921200	Autograft replacement of three coronary arteries NEC
7921300	Autograft replacement of four of more coronary arteries NEC

[*Table continues on the next page]*

Read Code cont.	Description
7921y00	Other autograft replacement of coronary artery OS
7921z00	Other autograft replacement of coronary artery NOS
7922	Allograft replacement of coronary artery
7922000	Allograft replacement of one coronary artery
7922100	Allograft replacement of two coronary arteries
7922200	Allograft replacement of three coronary arteries
7922300	Allograft replacement of four or more coronary arteries
7922y00	Other specified allograft replacement of coronary artery
7922z00	Allograft replacement of coronary artery NOS
7923	Prosthetic replacement of coronary artery
7923000	Prosthetic replacement of one coronary artery
7923100	Prosthetic replacement of two coronary arteries
7923200	Prosthetic replacement of three coronary arteries
7923300	Prosthetic replacement of four or more coronary arteries
7923z00	Prosthetic replacement of coronary artery NOS
7924100	Revision of bypass for two coronary arteries
7924200	Revision of bypass for three coronary arteries
7925000	Double anastomosis of mammary arteries to coronary arteries
7925100	Double implant of mammary arteries into coronary arteries
7925300	Single anastomosis of mammary artery to coronary artery NEC
7925400	Single implantation of mammary artery into coronary artery
7925y00	Connection of mammary artery to coronary artery OS
7926	Connection of other thoracic artery to coronary artery
7926000	Double anastom thoracic arteries to coronary arteries NEC
7926200	Single anastomosis of thoracic artery to coronary artery NEC
7926300	Single implantation thoracic artery into coronary artery NEC
7926z00	Connection of other thoracic artery to coronary artery NOS

[*Table continues on the next page]*

Read Code cont.	Description
7927	Other open operations on coronary artery
7929	Other therapeutic transluminal operations on coronary artery
7929200	Percut translum inject therap subst to coronary artery NEC
7929400	Insertion of coronary artery stent
7929500	Insertion of drug-eluting coronary artery stent
7929600	Percutaneous transluminal atherectomy of coronary artery
7929y00	Other therapeutic transluminal op on coronary artery OS
792B000	Endarterectomy of coronary artery NEC
792B100	Repair of rupture of coronary artery
792Bz00	Repair of coronary artery NOS
792C.00	Other replacement of coronary artery
792C000	Replacement of coronary arteries using multiple methods
792Cy00	Other specified replacement of coronary artery
792Cz00	Replacement of coronary artery NOS
792y.00	Other specified operations on coronary artery
792z.00	Coronary artery operations NOS
G3...00	Ischaemic heart disease
G3...11	Arteriosclerotic heart disease
G3...12	Atherosclerotic heart disease
G3...13	IHD - Ischaemic heart disease
G30..00	Acute myocardial infarction
G30..11	Attack - heart
G30..12	Coronary thrombosis
G30..13	Cardiac rupture following myocardial infarction (MI)
G30..14	Heart attack
G30..15	MI - acute myocardial infarction
G30..16	Thrombosis - coronary

[*Table continues on the next page]*

Read Code cont.	Description
G30..17	Silent myocardial infarction
G300.00	Acute anterolateral infarction
G301.00	Other specified anterior myocardial infarction
G301000	Acute anteroapical infarction
G301100	Acute anteroseptal infarction
G301z00	Anterior myocardial infarction NOS
G302.00	Acute inferolateral infarction
G303.00	Acute inferoposterior infarction
G304.00	Posterior myocardial infarction NOS
G305.00	Lateral myocardial infarction NOS
G306.00	True posterior myocardial infarction
G307.00	Acute subendocardial infarction
G307000	Acute non-Q wave infarction
G307100	Acute non-ST segment elevation myocardial infarction
G308.00	Inferior myocardial infarction NOS
G309.00	Acute Q-wave infarct
G30A.00	Mural thrombosis
G30B.00	Acute posterolateral myocardial infarction
G30X.00	Acute transmural myocardial infarction of unspecif site
G30X000	Acute ST segment elevation myocardial infarction
G30y.00	Other acute myocardial infarction
G30y000	Acute atrial infarction
G30y100	Acute papillary muscle infarction
G30y200	Acute septal infarction
G30yz00	Other acute myocardial infarction NOS
G30z.00	Acute myocardial infarction NOS
G31..00	Other acute and subacute ischaemic heart disease

[*Table continues on the next page]*

Read Code cont.	Description
G310.00	Postmyocardial infarction syndrome
G310.11	Dressler's syndrome
G311.00	Preinfarction syndrome
G311.11	Crescendo angina
G311.12	Impending infarction
G311.13	Unstable angina
G311.14	Angina at rest
G311000	Myocardial infarction aborted
G311011	MI - myocardial infarction aborted
G311100	Unstable angina
G311200	Angina at rest
G311300	Refractory angina
G311400	Worsening angina
G311500	Acute coronary syndrome
G311z00	Preinfarction syndrome NOS
G312.00	Coronary thrombosis not resulting in myocardial infarction
G31y.00	Other acute and subacute ischaemic heart disease
G31y000	Acute coronary insufficiency
G31y100	Microinfarction of heart
G31y200	Subendocardial ischaemia
G31y300	Transient myocardial ischaemia
G31yz00	Other acute and subacute ischaemic heart disease NOS
G32..00	Old myocardial infarction
G32..11	Healed myocardial infarction
G32..12	Personal history of myocardial infarction
G33..00	Angina pectoris
G330.00	Angina decubitus

[*Table continues on the next page]*

Read Code cont.	Description
G330000	Nocturnal angina
G330z00	Angina decubitus NOS
G331.00	Prinzmetal's angina
G331.11	Variant angina pectoris
G332.00	Coronary artery spasm
G33z.00	Angina pectoris NOS
G33z000	Status anginosus
G33z100	Stenocardia
G33z200	Syncope anginosa
G33z300	Angina on effort
G33z400	Ischaemic chest pain
G33z500	Post infarct angina
G33z600	New onset angina
G33z700	Stable angina
G33zz00	Angina pectoris NOS
G34..00	Other chronic ischaemic heart disease
G340.00	Coronary atherosclerosis
G340.11	Triple vessel disease of the heart
G340.12	Coronary artery disease
G340000	Single coronary vessel disease
G340100	Double coronary vessel disease
G342.00	Atherosclerotic cardiovascular disease
G343.00	Ischaemic cardiomyopathy
G344.00	Silent myocardial ischaemia
G34y.00	Other specified chronic ischaemic heart disease
G34y000	Chronic coronary insufficiency
G34y100	Chronic myocardial ischaemia

[Table continues on the next page]

Read Code cont.	Description
G34yz00	Other specified chronic ischaemic heart disease NOS
G34z.00	Other chronic ischaemic heart disease NOS
G34z000	Asymptomatic coronary heart disease
G35..00	Subsequent myocardial infarction
G350.00	Subsequent myocardial infarction of anterior wall
G351.00	Subsequent myocardial infarction of inferior wall
G353.00	Subsequent myocardial infarction of other sites
G35X.00	Subsequent myocardial infarction of unspecified site
G36..00	Certain current complication follow acute myocardial infarct
G360.00	Haemopericardium/current comp folow acut myocard infarct
G361.00	Atrial septal defect/curr comp folow acut myocardal infarct
G362.00	Ventric septal defect/curr comp fol acut myocardal infarctn
G363.00	Ruptur cardiac wall w'out haemopericard/cur comp fol ac MI
G364.00	Ruptur chordae tendinae/curr comp fol acute myocard infarct
G365.00	Rupture papillary muscle/curr comp fol acute myocard infarct
G366.00	Thrombosis atrium,auric append&vent/curr comp foll acute MI
G38..00	Postoperative myocardial infarction
G380.00	Postoperative transmural myocardial infarction anterior wall
G381.00	Postoperative transmural myocardial infarction inferior wall
G384.00	Postoperative subendocardial myocardial infarction
G38z.00	Postoperative myocardial infarction, unspecified
G3y..00	Other specified ischaemic heart disease
G3z..00	Ischaemic heart disease NOS
G61..00	Intracerebral haemorrhage
G61..11	CVA - cerebrovascular accid due to intracerebral haemorrhage
G61..12	Stroke due to intracerebral haemorrhage
G610.00	Cortical haemorrhage

[Table continues on the next page]

Read Code cont.	Description
G611.00	Internal capsule haemorrhage
G612.00	Basal nucleus haemorrhage
G613.00	Cerebellar haemorrhage
G614.00	Pontine haemorrhage
G615.00	Bulbar haemorrhage
G616.00	External capsule haemorrhage
G618.00	Intracerebral haemorrhage, multiple localized
G61X.00	Intracerebral haemorrhage in hemisphere, unspecified
G61X000	Left sided intracerebral haemorrhage, unspecified
G61X100	Right sided intracerebral haemorrhage, unspecified
G61z.00	Intracerebral haemorrhage NOS
G63y000	Cerebral infarct due to thrombosis of precerebral arteries
G63y100	Cerebral infarction due to embolism of precerebral arteries
G64..00	Cerebral arterial occlusion
G64..11	CVA - cerebral artery occlusion
G64..12	Infarction - cerebral
G64..13	Stroke due to cerebral arterial occlusion
G640.00	Cerebral thrombosis
G640000	Cerebral infarction due to thrombosis of cerebral arteries
G641.00	Cerebral embolism
G641.11	Cerebral embolus
G641000	Cerebral infarction due to embolism of cerebral arteries
G64z.00	Cerebral infarction NOS
G64z.11	Brainstem infarction NOS
G64z.12	Cerebellar infarction
G64z000	Brainstem infarction
G64z100	Wallenberg syndrome

[Table continues on the next page]

Read Code cont.	Description
G64z111	Lateral medullary syndrome
G64z200	Left sided cerebral infarction
G64z300	Right sided cerebral infarction
G64z400	Infarction of basal ganglia
G65..00	Transient cerebral ischaemia
G65..12	Transient ischaemic attack
G65..13	Vertebro-basilar insufficiency
G650.00	Basilar artery syndrome
G650.11	Insufficiency - basilar artery
G651.00	Vertebral artery syndrome
G651000	Vertebro-basilar artery syndrome
G653.00	Carotid artery syndrome hemispheric
G654.00	Multiple and bilateral precerebral artery syndromes
G656.00	Vertebrobasilar insufficiency
G65y.00	Other transient cerebral ischaemia
G65z.00	Transient cerebral ischaemia NOS
G65z000	Impending cerebral ischaemia
G65z100	Intermittent cerebral ischaemia
G65zz00	Transient cerebral ischaemia NOS
G66..00	Stroke and cerebrovascular accident unspecified
G66..11	CVA unspecified
G66..12	Stroke unspecified
G66..13	CVA - Cerebrovascular accident unspecified
G660.00	Middle cerebral artery syndrome
G661.00	Anterior cerebral artery syndrome
G662.00	Posterior cerebral artery syndrome
G663.00	Brain stem stroke syndrome

[Table continues on the next page]

Read Code cont.	Description
G664.00	Cerebellar stroke syndrome
G665.00	Pure motor lacunar syndrome
G666.00	Pure sensory lacunar syndrome
G667.00	Left sided CVA
G668.00	Right sided CVA
G676000	Cereb infarct due cerebral venous thrombosis, nonpyogenic
G677000	Occlusion and stenosis of middle cerebral artery
G677100	Occlusion and stenosis of anterior cerebral artery
G677200	Occlusion and stenosis of posterior cerebral artery
G683.00	Sequelae of cerebral infarction
G68X.00	Sequelae of stroke,not specfd as h'morrhage or infarction
G6W..00	Cereb infarct due unsp occlus/stenos precerebr arteries
G6X..00	Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Gyu3.00	[X]Ischaemic heart diseases
Gyu3000	[X]Other forms of angina pectoris
Gyu3200	[X]Other forms of acute ischaemic heart disease
Gyu3300	[X]Other forms of chronic ischaemic heart disease
Gyu3400	[X]Acute transmural myocardial infarction of unspecif site
Gyu3600	[X]Subsequent myocardial infarction of unspecified site
Gyu6200	[X]Other intracerebral haemorrhage
Gyu6300	[X]Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Gyu6400	[X]Other cerebral infarction
Gyu6500	[X]Occlusion and stenosis of other precerebral arteries
Gyu6600	[X]Occlusion and stenosis of other cerebral arteries
Gyu6F00	[X]Intracerebral haemorrhage in hemisphere, unspecified
Gyu6G00	[X]Cereb infarct due unsp occlus/stenos precerebr arteries
L440.11	CVA - cerebrovascular accident in the puerperium
ZV12D00	[V]Personal history of transient ischaemic attack

Appendix Table C 2. ICD code list for CVD

ICD Code	Description
I20	Angina pectoris
I20.0	Unstable angina
I20.1	Angina pectoris with documented spasm
I20.8	Other forms of angina pectoris
I20.9	Angina pectoris, unspecified
I21	Acute myocardial infarction
I21.0	Acute transmural myocardial infarction of anterior wall
I21.1	Acute transmural myocardial infarction of inferior wall
I21.2	Acute transmural myocardial infarction of other sites
I21.3	Acute transmural myocardial infarction of unspecified site
I21.4	Acute subendocardial myocardial infarction
I21.9	Acute myocardial infarction, unspecified
I22	Subsequent myocardial infarction
I22.0	Subsequent myocardial infarction of anterior wall
I22.1	Subsequent myocardial infarction of inferior wall
I22.8	Subsequent myocardial infarction of other sites
I22.9	Subsequent myocardial infarction of unspecified site
I23	Certain current complications following acute myocardial infarction
I23.0	Haemopericardium as current complication following acute myocardial infarction
I23.1	Atrial septal defect as current complication following acute myocardial infarction
I23.2	Ventricular septal defect as current complication following acute myocardial infarction
I23.3	Rupture of cardiac wall without haemopericardium as current complication following acute myocardial infarction
I23.4	Rupture of chordae tendineae as current complication following acute myocardial infarction
I23.5	Rupture of papillary muscle as current complication following acute myocardial infarction
I23.6	Thrombosis of atrium, auricular appendage, and ventricle as current complications following acute myocardial infarction

[Table continues on the next page]

ICD Code cont.	Description
I23.8	Other current complications following acute myocardial infarction
I24	Other acute ischaemic heart diseases
I24.0	Coronary thrombosis not resulting in myocardial infarction
I24.1	Dressler syndrome
I24.8	Other forms of acute ischaemic heart disease
I24.9	Acute ischaemic heart disease, unspecified
I25	Chronic ischaemic heart disease
I25.0	Atherosclerotic cardiovascular disease, so described
I25.1	Atherosclerotic heart disease
I25.2	Old myocardial infarction
I25.3	Aneurysm of heart
I25.4	Coronary artery aneurysm
I25.5	Ischaemic cardiomyopathy
I25.6	Silent myocardial ischaemia
I25.8	Other forms of chronic ischaemic heart disease
I25.9	Chronic ischaemic heart disease, unspecified
I61	Intracerebral haemorrhage
I61.0	Intracerebral haemorrhage in hemisphere, subcortical
I61.1	Intracerebral haemorrhage in hemisphere, cortical
I61.2	Intracerebral haemorrhage in hemisphere, unspecified
I61.3	Intracerebral haemorrhage in brain stem
I61.4	Intracerebral haemorrhage in cerebellum
I61.5	Intracerebral haemorrhage, intraventricular
I61.6	Intracerebral haemorrhage, multiple localized
I61.8	Other intracerebral haemorrhage
I61.9	Intracerebral haemorrhage, unspecified
I63	Cerebral infarction

[Table continues on the next page]

ICD Code cont.	Description
I63.0	Cerebral infarction due to thrombosis of precerebral arteries
I63.1	Cerebral infarction due to embolism of precerebral arteries
I63.2	Cerebral infarction due to unspecified occlusion or stenosis of precerebral arteries
I63.3	Cerebral infarction due to thrombosis of cerebral arteries
I63.4	Cerebral infarction due to embolism of cerebral arteries
I63.5	Cerebral infarction due to unspecified occlusion or stenosis of cerebral arteries
I63.6	Cerebral infarction due to cerebral venous thrombosis, nonpyogenic
I63.8	Other cerebral infarction
I63.9	Cerebral infarction, unspecified
I64	Stroke, not specified as haemorrhage or infarction
I65	Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
I65.0	Occlusion and stenosis of vertebral artery
I65.1	Occlusion and stenosis of basilar artery
I65.2	Occlusion and stenosis of carotid artery
I65.3	Occlusion and stenosis of multiple and bilateral precerebral arteries
I65.8	Occlusion and stenosis of other precerebral artery
I65.9	Occlusion and stenosis of unspecified precerebral artery
I66	Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction
I66.0	Occlusion and stenosis of middle cerebral artery
I66.1	Occlusion and stenosis of anterior cerebral artery
I66.2	Occlusion and stenosis of posterior cerebral artery
I66.3	Occlusion and stenosis of cerebellar arteries
I66.4	Occlusion and stenosis of multiple and bilateral cerebral arteries
I66.8	Occlusion and stenosis of other cerebral artery
I66.9	Occlusion and stenosis of unspecified cerebral artery
I67.2	Cerebral atherosclerosis
I67.9	Cerebrovascular disease, unspecified

[Table continues on the next page]

ICD Code cont.	Description
I69	Sequelae of cerebrovascular disease
I69.1	Sequelae of intracerebral haemorrhage
I69.2	Sequelae of other nontraumatic intracranial haemorrhage
I69.3	Sequelae of cerebral infarction
I69.4	Sequelae of stroke, not specified as haemorrhage or infarction

Appendix Table C 3 Entity code list for CVD

Entity Code	Description
12	health promotion - fh of cva / ihd
87	family history

Appendix D. Read code list for Family History

Appendix Table D1. Read code list used to search for family history of cardiovascular disease.

Read code	Description	Detail
1154.00	No significant family history	FH only
122..00	No relevant family history	FH only
122..11	No relevant FH: family history	FH only
122Z.00	No FH: NOS	FH only
12V..00	Family history unknown	FH only
1154100	No significant family history of cardiovascular disease	FH, disease
1224.00	No FH: Cardiovascular disease	FH, disease
1226.00	No FH: Ischaemic heart disease	FH, disease
1226.11	No FH: Angina	FH, disease
12C..00	FH: Cardiovascular disease	FH, disease
12C..11	FH: CVS disorder	FH, disease
12C..12	FH: Cardiac disorder	FH, disease
12C..13	FH: Heart disorder	FH, disease
12C..14	FH: Angina	FH, disease
12C5.00	FH: Myocardial infarction	FH, disease
12C5.11	FH: Coronary thrombosis	FH, disease
12C5.12	FH: Ischaemic heart disease	FH, disease
12C6.00	FH: Atherosclerosis	FH, disease
12CI.00	FH: premature coronary heart disease	FH, disease
12CZ.00	FH: CVS disease NOS	FH, disease
ZV17300	[V]Family history of ischaemic heart disease	FH, disease
ZV17311	[V]Family history of ischaemic heart disease (IHD)	FH, disease
ZV17312	[V]Family history of myocardial infarction	FH, disease
ZV17400	[V]Family history of other cardiovascular disease	FH, disease
ZVu6600	[X]Family hist/ischaemic hrt disease+oth dis/circultr system	FH, disease

[Table continues on the next page]

Read code cont.	Description	Detail
12C2.00	FH: Ischaemic heart dis. <60	FH, disease, age
12C2.11	FH: Myocardial infarction < 60	FH, disease, age
12C2.12	FH: MI- Myocardial infarct <60	FH, disease, age
12C2.13	FH: Angina < 60yrs	FH, disease, age
12C3.00	FH: Ischaemic heart dis. >60	FH, disease, age
12C3.11	FH: Myocardial infarction > 60	FH, disease, age
12C3.12	FH: MI- myocardial infarct >60	FH, disease, age
12C3.13	FH: Angina > 60yrs	FH, disease, age
12CL.00	FH: Angina in 1st degree female relative <65 years	FH, disease, age, relative
12CM.00	FH: Angina in 1st degree male relative <55 years	FH, disease, age, relative
12CN.00	FH: Myocardial infarct in 1st degree female relative <65 yrs	FH, disease, age, relative
12CP.00	FH: Myocardial infarct in 1st degree male relative <55 years	FH, disease, age, relative
12CV.00	FH: Cardiovascular disease 1st degree male relative < 55 yrs	FH, disease, age, relative
12CW.00	FH: Cardiovascular disease 1st degree female reltve < 65 yrs	FH, disease, age, relative
12CA.00	FH myocardial infarction male first degree age known	FH, disease, relative
12CB.00	FH myocardial infarction male first degree age unknown	FH, disease, relative
12CC.00	FH myocardial infarction female first degree age known	FH, disease, relative
12CD.00	FH myocardial infarction female first degree age unknown	FH, disease, relative
12CE.00	FH angina male first degree age known	FH, disease, relative
12CF.00	FH angina male first degree age unknown	FH, disease, relative
12CG.00	FH angina female first degree age known	FH, disease, relative
12CH.00	FH angina female first degree age unknown	FH, disease, relative
12K..00	FH: Mother	FH, relative, no disease
12K1.00	FH: Mother alive + well	FH, relative, no disease
12K2.00	FH: Mother alive + problem	FH, relative, no disease
12K3.00	FH: Mother dead	FH, relative, no disease
12K4.00	FH: Mother unwell	FH, relative, no disease

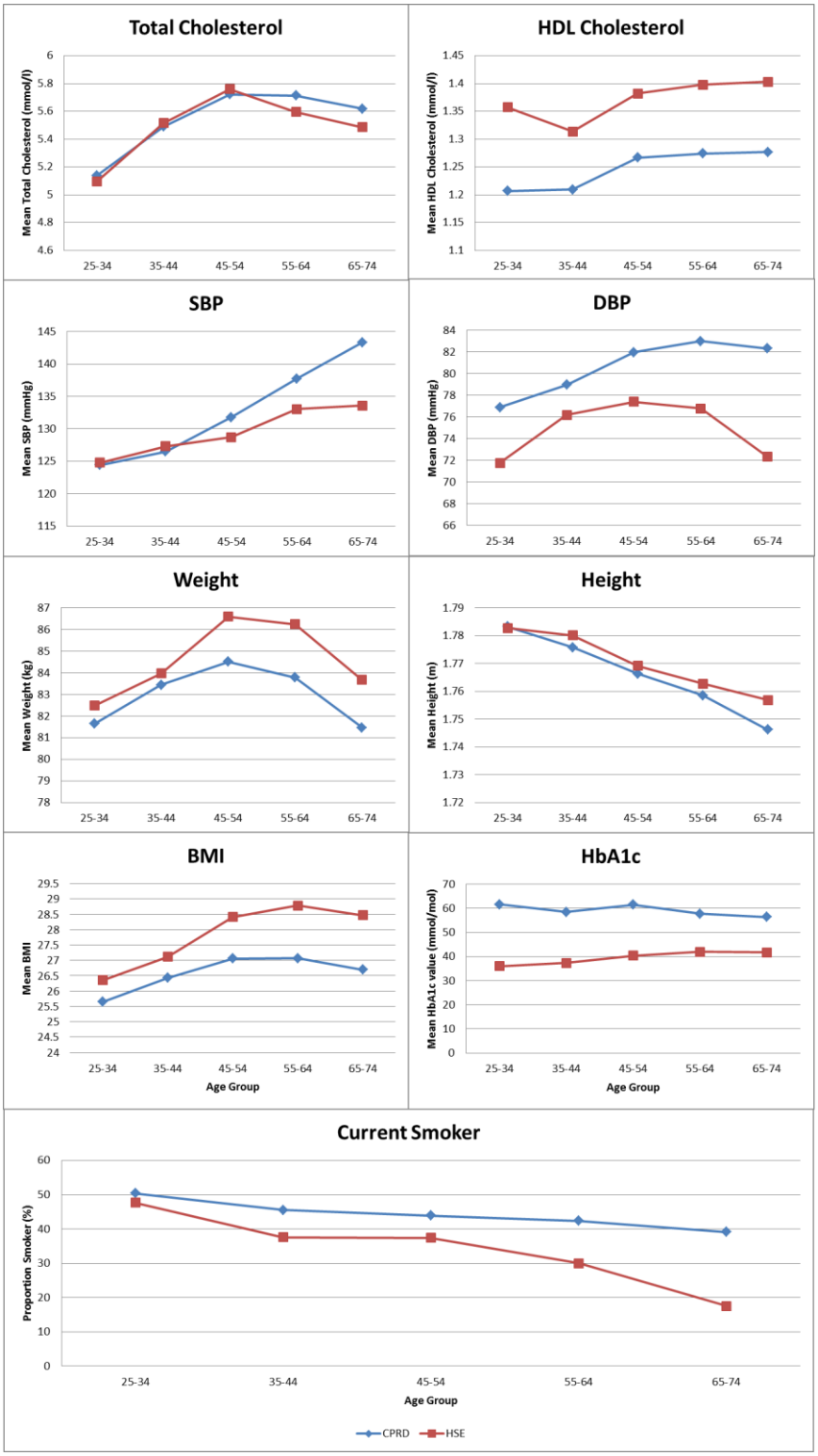
[Table continues on the next page]

Read code cont.	Description	Detail
12KZ.00	FH: Mother NOS	FH, relative, no disease
12L..00	FH: Father	FH, relative, no disease
12L1.00	FH: Father alive + well	FH, relative, no disease
12L2.00	FH: Father alive + problem	FH, relative, no disease
12L3.00	FH: Father dead	FH, relative, no disease
12L4.00	FH: Father deaf	FH, relative, no disease
12LZ.00	FH: Father NOS	FH, relative, no disease
12M..00	FH: Sister	FH, relative, no disease
12M1.00	FH: Sister alive + well	FH, relative, no disease
12M2.00	FH: Sister alive + problem	FH, relative, no disease
12M3.00	FH: Sister dead	FH, relative, no disease
12MZ.00	FH: Sister NOS	FH, relative, no disease
12N..00	FH: Brother	FH, relative, no disease
12N1.00	FH: Brother alive + well	FH, relative, no disease
12N2.00	FH: Brother alive + problem	FH, relative, no disease
12N3.00	FH: Brother dead	FH, relative, no disease
12NZ.00	FH: Brother NOS	FH, relative, no disease
12O..00	FH: Son	FH, relative, no disease
12O1.00	FH: Son alive + well	FH, relative, no disease
12O2.00	FH: Son alive + problem	FH, relative, no disease
12O3.00	FH: Son dead	FH, relative, no disease
12OZ.00	FH: Son NOS	FH, relative, no disease
12P..00	FH: Daughter	FH, relative, no disease
12P1.00	FH: Daughter alive + well	FH, relative, no disease
12P2.00	FH: Daughter alive + problem	FH, relative, no disease
12P3.00	FH: Daughter dead	FH, relative, no disease
12PZ.00	FH: Daughter NOS	FH, relative, no disease

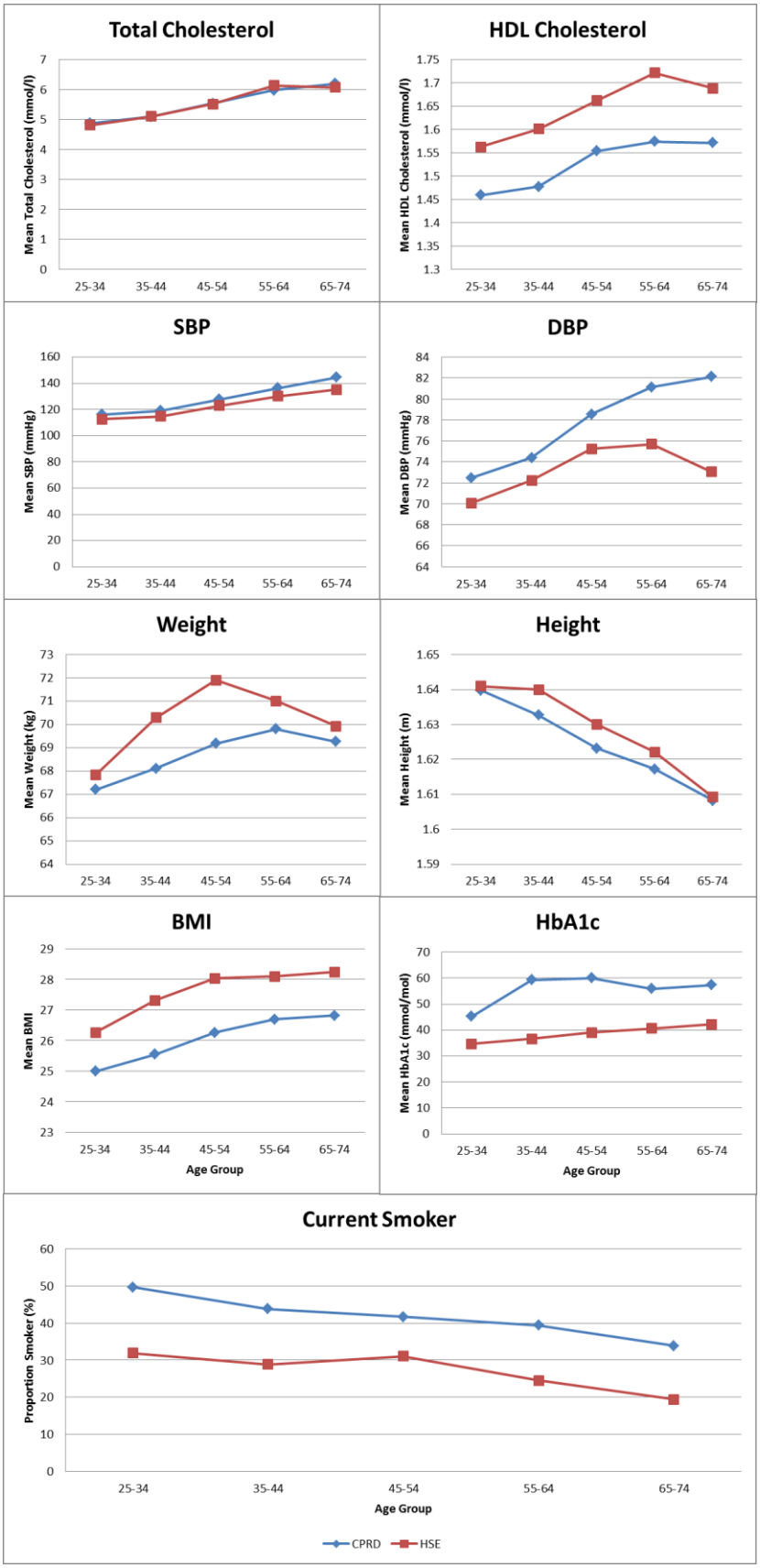
[Table continues on the next page]

Read code cont.	Description	Detail
12Q..00	FH: Relative OS	FH, relative, no disease
12Q1.00	FH: Maternal grandfather	FH, relative, no disease
12Q2.00	FH: Paternal grandfather	FH, relative, no disease
12Q3.00	FH: Maternal grandmother	FH, relative, no disease
12Q4.00	FH: Paternal grandmother	FH, relative, no disease
12Q5.00	FH: Aunt	FH, relative, no disease
12Q6.00	FH: Uncle	FH, relative, no disease
12QZ.00	FH: Relative NOS	FH, relative, no disease
12Z3.00	FH: Death under 60 years	FH, relative, no disease

Appendix E. Generalisability of extracted CPRD data to Health Survey for England 2011.



Appendix Figure E 1. Comparison of extracted CPRD data and HSE data characteristics, by age group (males).



Appendix Figure E 2. Comparison of extracted CPRD data and HSE data characteristics, by age group (females).

Appendix F. Regression output for missing data.

Appendix Table F 1. Unadjusted regression for missing data – Smoking

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		0.002 (0.001)	0.0334	1.002	(1 to 1.005)
Gender	Female	Ref			
	Male	0.415 (0.029)	<0.001	1.514	(1.431 to 1.601)
Total Cholesterol		-0.028 (0.04)	0.4846	0.972	(0.899 to 1.052)
HDL Cholesterol		0.236 (0.173)	0.1737	1.266	(0.901 to 1.779)
SBP		0.008 (0.001)	<0.001	1.008	(1.006 to 1.01)
Family History	Negative	Ref			
	Positive	0.373 (0.029)	<0.001	1.453	(1.373 to 1.537)
CVD	No	Ref			
	Yes	0.079 (0.1)	0.4343	1.082	(0.888 to 1.317)
Townsend Deprivation Score	1	Ref			
	2	-0.063 (0.048)	0.1926	0.939	(0.854 to 1.032)
	3	0.117 (0.048)	0.0156	1.124	(1.022 to 1.236)
	4	-0.228 (0.061)	<0.001	0.796	(0.707 to 0.897)
	5	-0.343 (0.077)	<0.001	0.71	(0.61 to 0.825)
Ethnicity	White	Ref			
	S.Asian	0.093 (0.186)	0.6171	1.097	(0.763 to 1.579)
	Black	-0.774 (0.449)	0.0852	0.461	(0.191 to 1.113)
	Other/Mixed	0.013 (0.051)	0.8043	1.013	(0.917 to 1.118)
BMI		0.019 (0.004)	<0.001	1.019	(1.011 to 1.027)
Consumes Alcohol	No	Ref			
	Yes	0.396 (0.11)	<0.001	1.486	(1.197 to 1.845)
Diabetes	No	Ref			
	Yes	-0.345 (0.128)	0.0068	0.708	(0.551 to 0.909)
Impaired Glucose Regulation	No	Ref			
	Yes	0.021 (0.504)	0.9671	1.021	(0.38 to 2.744)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0.214 (0.583)	0.7143	1.238	(0.395 to 3.884)

Appendix Table F 2. Unadjusted regression for missing data – Total Cholesterol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.035 (0)	<0.001	0.966	(0.965 to 0.967)
Gender	Female	Ref			
	Male	-0.277 (0.011)	<0.001	0.758	(0.743 to 0.774)
Smoking Status	Non/Ex	Ref			
	Current	-0.342 (0.011)	<0.001	0.71	(0.696 to 0.725)
HDL Cholesterol		0.235 (0.103)	0.0226	1.264	(1.033 to 1.547)
SBP		-0.026 (0)	<0.001	0.974	(0.973 to 0.974)
Family History	Negative	Ref			
	Positive	-0.692 (0.011)	<0.001	0.5	(0.49 to 0.511)
CVD	No	Ref			
	Yes	-0.665 (0.031)	<0.001	0.514	(0.484 to 0.547)
Townsend Deprivation Score	1	Ref			
	2	0.144 (0.017)	<0.001	1.155	(1.117 to 1.194)
	3	0.263 (0.018)	<0.001	1.301	(1.255 to 1.348)
	4	0.199 (0.02)	<0.001	1.22	(1.172 to 1.27)
	5	0.207 (0.025)	<0.001	1.23	(1.172 to 1.29)
Ethnicity	White	Ref			
	S.Asian	-0.344 (0.06)	<0.001	0.709	(0.63 to 0.797)
	Black	-0.177 (0.1)	0.0778	0.838	(0.688 to 1.02)
	Other/Mixed	-0.005 (0.018)	0.7875	0.995	(0.962 to 1.03)
BMI		-0.053 (0.001)	<0.001	0.948	(0.946 to 0.95)
Consumes Alcohol	No	Ref			
	Yes	-0.438 (0.026)	<0.001	0.645	(0.614 to 0.679)
Diabetes	No	Ref			
	Yes	-2.853 (0.031)	<0.001	0.058	(0.054 to 0.061)
Impaired Glucose Regulation	No	Ref			
	Yes	-2.514 (0.137)	<0.001	0.081	(0.062 to 0.106)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-1.562 (0.165)	<0.001	0.21	(0.152 to 0.29)

Appendix Table F 3. Unadjusted regression for missing data – HDL Cholesterol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.035 (0.001)	<0.001	0.966	(0.964 to 0.967)
Gender	Female	Ref			
	Male	-0.213 (0.017)	<0.001	0.808	(0.781 to 0.836)
Smoking Status	Non/Ex	Ref			
	Current	-0.35 (0.017)	<0.001	0.705	(0.681 to 0.729)
Total Cholesterol		-0.284 (0.01)	<0.001	0.753	(0.738 to 0.768)
SBP		-0.025 (0)	<0.001	0.975	(0.974 to 0.976)
Family History	Negative	Ref			
	Positive	-0.692 (0.017)	<0.001	0.501	(0.484 to 0.518)
CVD	No	Ref			
	Yes	-0.537 (0.049)	<0.001	0.584	(0.53 to 0.644)
Townsend Deprivation Score	1	Ref			
	2	0.121 (0.029)	<0.001	1.128	(1.066 to 1.195)
	3	0.046 (0.03)	0.1232	1.048	(0.987 to 1.111)
	4	-0.088 (0.033)	0.0066	0.915	(0.859 to 0.976)
	5	0.02 (0.04)	0.6273	1.02	(0.942 to 1.103)
Ethnicity	White	Ref			
	S.Asian	-0.486 (0.09)	<0.001	0.615	(0.516 to 0.734)
	Black	-0.195 (0.162)	0.2281	0.823	(0.6 to 1.13)
	Other/Mixed	-0.012 (0.029)	0.6691	0.988	(0.933 to 1.045)
BMI		-0.048 (0.002)	<0.001	0.954	(0.95 to 0.957)
Consumes Alcohol	No	Ref			
	Yes	-0.185 (0.039)	<0.001	0.831	(0.771 to 0.897)
Diabetes	No	Ref			
	Yes	-2.045 (0.032)	<0.001	0.129	(0.121 to 0.138)
Impaired Glucose Regulation	No	Ref			
	Yes	-2.147 (0.141)	<0.001	0.117	(0.089 to 0.154)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-1.34 (0.223)	<0.001	0.262	(0.169 to 0.405)

Appendix Table F 4. Unadjusted regression for missing data – SBP

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.014 (0)	<0.001	0.986	(0.986 to 0.987)
Gender	Female	Ref			
	Male	0.883 (0.009)	<0.001	2.419	(2.376 to 2.463)
Smoking Status	Non/Ex	Ref			
	Current	-0.429 (0.01)	<0.001	0.651	(0.639 to 0.663)
Total Cholesterol		-0.04 (0.02)	0.0429	0.961	(0.924 to 0.999)
HDL Cholesterol		-0.364 (0.098)	<0.001	0.695	(0.574 to 0.841)
Family History	Negative	Ref			
	Positive	-0.383 (0.01)	<0.001	0.682	(0.669 to 0.695)
CVD	No	Ref			
	Yes	-0.34 (0.036)	<0.001	0.712	(0.663 to 0.764)
Townsend Deprivation Score	1	Ref			
	2	0.049 (0.016)	0.0019	1.05	(1.018 to 1.083)
	3	0.074 (0.017)	<0.001	1.077	(1.043 to 1.113)
	4	-0.028 (0.019)	0.1429	0.973	(0.937 to 1.009)
	5	0.097 (0.022)	<0.001	1.102	(1.055 to 1.15)
Ethnicity	White	Ref			
	S.Asian	-0.352 (0.07)	<0.001	0.703	(0.612 to 0.807)
	Black	-0.335 (0.111)	0.0026	0.715	(0.575 to 0.89)
	Other/Mixed	-0.052 (0.017)	0.0017	0.949	(0.919 to 0.981)
BMI		0.003 (0.002)	0.2501	1.003	(0.998 to 1.008)
Consumes Alcohol	No	Ref			
	Yes	-0.347 (0.037)	<0.001	0.707	(0.657 to 0.76)
Diabetes	No	Ref			
	Yes	-2.199 (0.082)	<0.001	0.111	(0.094 to 0.13)
Impaired Glucose Regulation	No	Ref			
	Yes	-1.782 (0.323)	<0.001	0.168	(0.089 to 0.317)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-3.643 (1.003)	<0.001	0.026	(0.004 to 0.187)

Appendix Table F 5. Unadjusted regression for missing data – Townsend Deprivation Score

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		0.001 (0)	<0.001	1.001	(1.001 to 1.002)
Gender	Female	Ref			
	Male	-0.05 (0.008)	<0.001	0.952	(0.937 to 0.966)
Smoking Status	Non/Ex	Ref			
	Current	0.024 (0.008)	0.0019	1.024	(1.009 to 1.039)
Total Cholesterol		0.033 (0.01)	<0.001	1.034	(1.014 to 1.054)
HDL Cholesterol		-0.018 (0.042)	0.6644	0.982	(0.905 to 1.066)
SBP		0.002 (0)	<0.001	1.002	(1.001 to 1.002)
Family History	Negative	Ref			
	Positive	0.008 (0.008)	0.2759	1.008	(0.993 to 1.024)
CVD	No	Ref			
	Yes	-2.568 (0.063)	<0.001	0.077	(0.068 to 0.087)
Ethnicity	White	Ref			
	S.Asian	1.039 (0.053)	<0.001	2.826	(2.546 to 3.135)
	Black	-0.19 (0.122)	0.1196	0.827	(0.651 to 1.05)
	Other/Mixed	-1.444 (0.032)	<0.001	0.236	(0.222 to 0.251)
BMI		0.005 (0.001)	<0.001	1.005	(1.004 to 1.007)
Consumes Alcohol	No	Ref			
	Yes	-0.413 (0.017)	<0.001	0.662	(0.64 to 0.684)
Diabetes	No	Ref			
	Yes	0.089 (0.028)	0.0015	1.094	(1.035 to 1.156)
Impaired Glucose Regulation	No	Ref			
	Yes	0.175 (0.131)	0.1809	1.192	(0.922 to 1.54)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0.552 (0.164)	<0.001	1.737	(1.26 to 2.395)

Appendix Table F 6. Unadjusted regression for missing data – Ethnicity

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.012 (0)	<0.001	0.988	(0.987 to 0.989)
Gender	Female	Ref			
	Male	0.255 (0.007)	<0.001	1.29	(1.271 to 1.309)
Smoking Status	Non/Ex	Ref			
	Current	-0.007 (0.007)	0.369	0.993	(0.979 to 1.008)
Total Cholesterol		-0.001 (0.009)	0.877	0.999	(0.981 to 1.017)
HDL Cholesterol		0.004 (0.04)	0.9283	1.004	(0.928 to 1.086)
SBP		-0.003 (0)	<0.001	0.997	(0.997 to 0.997)
Family History	Negative	Ref			
	Positive	-0.051 (0.007)	<0.001	0.95	(0.937 to 0.964)
CVD	No	Ref			
	Yes	-2.374 (0.047)	<0.001	0.093	(0.085 to 0.102)
Townsend Deprivation Score	1	Ref			
	2	-0.227 (0.014)	<0.001	0.797	(0.776 to 0.819)
	3	-0.315 (0.015)	<0.001	0.73	(0.71 to 0.751)
	4	-0.414 (0.017)	<0.001	0.661	(0.64 to 0.683)
	5	-0.545 (0.021)	<0.001	0.58	(0.557 to 0.604)
BMI		-0.013 (0.001)	<0.001	0.987	(0.985 to 0.988)
Consumes Alcohol	No	Ref			
	Yes	-0.058 (0.017)	<0.001	0.944	(0.913 to 0.976)
Diabetes	No	Ref			
	Yes	-0.368 (0.028)	<0.001	0.692	(0.655 to 0.731)
Impaired Glucose Regulation	No	Ref			
	Yes	-0.345 (0.132)	0.0087	0.708	(0.547 to 0.916)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-0.024 (0.164)	0.8828	0.976	(0.708 to 1.346)

Appendix Table F 7. Unadjusted regression for missing data – BMI

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		0.005 (0)	<0.001	1.005	(1.004 to 1.005)
Gender	Female	Ref			
	Male	0.441 (0.008)	<0.001	1.554	(1.531 to 1.577)
Smoking Status	Non/Ex	Ref			
	Current	-0.593 (0.008)	<0.001	0.553	(0.544 to 0.561)
Total Cholesterol		0.022 (0.011)	0.0456	1.022	(1 to 1.045)
HDL Cholesterol		0.126 (0.048)	0.0092	1.134	(1.032 to 1.247)
SBP		0.008 (0)	<0.001	1.008	(1.007 to 1.008)
Family History	Negative	Ref			
	Positive	-0.344 (0.008)	<0.001	0.709	(0.698 to 0.72)
CVD	No	Ref			
	Yes	-0.053 (0.027)	0.051	0.948	(0.898 to 1)
Townsend Deprivation Score	1	Ref			
	2	-0.005 (0.013)	0.697	0.995	(0.97 to 1.02)
	3	0.057 (0.014)	<0.001	1.058	(1.031 to 1.087)
	4	-0.085 (0.015)	<0.001	0.919	(0.892 to 0.947)
	5	-0.015 (0.018)	0.4165	0.985	(0.951 to 1.021)
Ethnicity	White	Ref			
	S.Asian	-0.337 (0.053)	<0.001	0.714	(0.643 to 0.792)
	Black	-0.238 (0.083)	0.004	0.788	(0.67 to 0.927)
	Other/Mixed	-0.072 (0.013)	<0.001	0.931	(0.907 to 0.955)
Consumes Alcohol	No	Ref			
	Yes	-0.172 (0.027)	<0.001	0.842	(0.799 to 0.888)
Diabetes	No	Ref			
	Yes	-1.714 (0.045)	<0.001	0.18	(0.165 to 0.197)
Impaired Glucose Regulation	No	Ref			
	Yes	-1.412 (0.193)	<0.001	0.244	(0.167 to 0.355)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-0.472 (0.183)	0.0101	0.624	(0.435 to 0.894)

Appendix Table F 8. Unadjusted regression for missing data – Consumes Alcohol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		0.01 (0)	<0.001	1.01	(1.01 to 1.011)
Gender	Female	Ref			
	Male	0.139 (0.007)	<0.001	1.149	(1.132 to 1.166)
Smoking Status	Non/Ex	Ref			
	Current	-0.892 (0.008)	<0.001	0.41	(0.403 to 0.416)
Total Cholesterol		0.037 (0.01)	<0.001	1.038	(1.018 to 1.058)
HDL Cholesterol		0.024 (0.043)	0.5781	1.024	(0.941 to 1.115)
SBP		0.01 (0)	<0.001	1.01	(1.009 to 1.01)
Family History	Negative	Ref			
	Positive	-0.351 (0.008)	<0.001	0.704	(0.693 to 0.714)
CVD	No	Ref			
	Yes	0.07 (0.027)	0.0087	1.072	(1.018 to 1.129)
Townsend Deprivation Score	1	Ref			
	2	0.002 (0.013)	0.8804	1.002	(0.977 to 1.027)
	3	0.079 (0.013)	<0.001	1.082	(1.054 to 1.11)
	4	0.063 (0.015)	<0.001	1.065	(1.034 to 1.096)
	5	0.065 (0.018)	<0.001	1.068	(1.031 to 1.105)
Ethnicity	White	Ref			
	S.Asian	-0.101 (0.049)	0.0395	0.904	(0.821 to 0.995)
	Black	-0.001 (0.078)	0.9891	0.999	(0.858 to 1.163)
	Other/Mixed	-0.09 (0.013)	<0.001	0.914	(0.891 to 0.938)
BMI		0.041 (0.001)	<0.001	1.041	(1.039 to 1.044)
Diabetes	No	Ref			
	Yes	-0.03 (0.028)	0.2876	0.971	(0.92 to 1.025)
Impaired Glucose Regulation	No	Ref			
	Yes	-0.479 (0.138)	<0.001	0.62	(0.473 to 0.812)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-0.65 (0.181)	<0.001	0.522	(0.366 to 0.744)

Appendix Table F 9. Adjusted regression for missing data. – Smoking

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		0.013 (0.021)	0.5439	1.013	(0.972 to 1.055)
Gender	Female	Ref			
	Male	0.882 (0.454)	0.0519	2.417	(0.993 to 5.882)
Total Cholesterol		0.005 (0.196)	0.9816	1.005	(0.684 to 1.476)
HDL Cholesterol		1.001 (0.438)	0.0222	2.722	(1.154 to 6.422)
SBP		0.024 (0.012)	0.0499	1.024	(1 to 1.048)
Family History	Negative	Ref			
	Positive	0.443 (0.426)	0.2986	1.557	(0.676 to 3.586)
CVD	No	Ref			
	Yes	0.598 (0.643)	0.3527	1.818	(0.515 to 6.417)
Townsend Deprivation Score	1	Ref			
	2	-0.118 (0.499)	0.8126	0.888	(0.334 to 2.364)
	3	-0.13 (0.523)	0.804	0.878	(0.315 to 2.45)
	4	-1.608 (1.054)	0.1272	0.2	(0.025 to 1.581)
	5	-1.009 (1.057)	0.3397	0.365	(0.046 to 2.894)
Ethnicity	White	Ref			
	S.Asian	0 (0)	<0.001	1	(1 to 1)
	Black	0 (0)	<0.001	1	(1 to 1)
	Other/Mixed	-0.753 (0.622)	0.226	0.471	(0.139 to 1.594)
BMI		-0.059 (0.053)	0.2625	0.943	(0.85 to 1.045)
Consumes Alcohol	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)
Diabetes	No	0 (0)	<0.001	1	(1 to 1)
	Yes	-0.527 (1.045)	0.6145	0.591	(0.076 to 4.584)
Impaired Glucose Regulation	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)

Appendix Table F 10. Adjusted regression for missing data. – Total Cholesterol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.027 (0.008)	<0.001	0.974	(0.959 to 0.989)
Gender	Female	Ref			
	Male	0.167 (0.179)	0.3523	1.182	(0.831 to 1.679)
Smoking Status	Non/Ex	Ref			
	Current	-0.123 (0.17)	0.4664	0.884	(0.634 to 1.232)
HDL Cholesterol		0.495 (0.195)	0.0113	1.641	(1.119 to 2.406)
SBP		-0.014 (0.006)	0.0122	0.986	(0.976 to 0.997)
Family History	Negative	Ref			
	Positive	-0.187 (0.166)	0.2601	0.83	(0.6 to 1.148)
CVD	No	Ref			
	Yes	-0.118 (0.429)	0.783	0.889	(0.383 to 2.059)
Townsend Deprivation Score	1	Ref			
	2	0.386 (0.245)	0.1148	1.471	(0.911 to 2.378)
	3	0.536 (0.242)	0.0271	1.709	(1.063 to 2.749)
	4	0.461 (0.259)	0.0755	1.586	(0.954 to 2.637)
	5	0.645 (0.3)	0.0317	1.905	(1.058 to 3.43)
Ethnicity	White	Ref			
	S.Asian	0 (0)	<0.001	1	(1 to 1)
	Black	0 (0)	<0.001	1	(1 to 1)
	Other/Mixed	-0.033 (0.195)	0.8678	0.968	(0.66 to 1.419)
BMI		0.019 (0.018)	0.2864	1.019	(0.984 to 1.056)
Consumes Alcohol	No	Ref			
	Yes	-0.346 (0.32)	0.2799	0.708	(0.378 to 1.325)
Diabetes	No	Ref			
	Yes	-0.776 (0.4)	0.0523	0.46	(0.21 to 1.008)
Impaired Glucose Regulation	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)

Appendix Table F 11. Adjusted regression for missing data. – HDL Cholesterol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.005 (0.002)	0.0087	0.995	(0.991 to 0.999)
Gender	Female	Ref			
	Male	-0.02 (0.038)	0.5954	0.98	(0.909 to 1.056)
Smoking Status	Non/Ex	Ref			
	Current	-0.097 (0.039)	0.0126	0.907	(0.84 to 0.979)
Total Cholesterol		-0.335 (0.019)	<0.001	0.715	(0.689 to 0.742)
SBP		-0.002 (0.001)	0.0494	0.998	(0.995 to 1)
Family History	Negative	Ref			
	Positive	-0.008 (0.038)	0.8372	0.992	(0.92 to 1.07)
CVD	No	Ref			
	Yes	0.127 (0.084)	0.1283	1.136	(0.964 to 1.339)
Townsend Deprivation Score	1	Ref			
	2	0.089 (0.051)	0.0822	1.093	(0.989 to 1.208)
	3	-0.245 (0.054)	<0.001	0.783	(0.704 to 0.87)
	4	-0.318 (0.058)	<0.001	0.728	(0.649 to 0.816)
	5	-0.145 (0.07)	0.0393	0.865	(0.753 to 0.993)
Ethnicity	White	Ref			
	S.Asian	-0.473 (0.165)	0.0041	0.623	(0.451 to 0.861)
	Black	-0.296 (0.27)	0.2733	0.744	(0.438 to 1.263)
	Other/Mixed	-0.055 (0.046)	0.2292	0.947	(0.865 to 1.035)
BMI		0.002 (0.004)	0.6613	1.002	(0.994 to 1.009)
Consumes Alcohol	No	Ref			
	Yes	0.106 (0.08)	0.1824	1.112	(0.951 to 1.301)
Diabetes	No	Ref			
	Yes	-0.234 (0.067)	<0.001	0.791	(0.694 to 0.902)
Impaired Glucose Regulation	No	Ref			
	Yes	-0.452 (0.287)	0.1151	0.637	(0.363 to 1.116)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-0.75 (0.454)	0.0991	0.473	(0.194 to 1.152)

Appendix Table F 12. Adjusted regression for missing data. – SBP

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.043 (0.016)	0.0052	0.958	(0.929 to 0.987)
Gender	Female	Ref			
	Male	0.233 (0.376)	0.5351	1.262	(0.605 to 2.635)
Smoking Status	Non/Ex	Ref			
	Current	-0.302 (0.354)	0.3936	0.74	(0.37 to 1.479)
Total Cholesterol		-0.286 (0.17)	0.0925	0.751	(0.538 to 1.048)
HDL Cholesterol		-0.439 (0.503)	0.3822	0.644	(0.241 to 1.726)
Family History	Negative	Ref			
	Positive	0.002 (0.351)	0.9954	1.002	(0.504 to 1.992)
CVD	No	Ref			
	Yes	0.664 (0.753)	0.378	1.943	(0.444 to 8.504)
Townsend Deprivation Score	1	Ref			
	2	0.124 (0.455)	0.7846	1.132	(0.464 to 2.761)
	3	0.198 (0.456)	0.6645	1.219	(0.499 to 2.978)
	4	-0.053 (0.516)	0.9177	0.948	(0.345 to 2.606)
	5	0 (0)	<0.001	1	(1 to 1)
Ethnicity	White	Ref			
	S.Asian	0.303 (1.073)	0.778	1.353	(0.165 to 11.096)
	Black	Ref			
	Other/Mixed	-0.621 (0.49)	0.2053	0.538	(0.206 to 1.405)
BMI		-0.079 (0.043)	0.0644	0.924	(0.85 to 1.005)
Consumes Alcohol	No	Ref			
	Yes	-0.601 (0.649)	0.3545	0.548	(0.154 to 1.957)
Diabetes	No	Ref			
	Yes	-1.121 (1.036)	0.279	0.326	(0.043 to 2.481)
Impaired Glucose Regulation	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)

Appendix Table F 13. Adjusted regression for missing data. – Townsend Deprivation Score

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.01 (0.006)	0.1014	0.99	(0.979 to 1.002)
Gender	Female	Ref			
	Male	-0.129 (0.133)	0.3304	0.879	(0.677 to 1.14)
Smoking Status	Non/Ex	Ref			
	Current	0.303 (0.131)	0.0209	1.354	(1.047 to 1.75)
Total Cholesterol		0 (0.057)	0.994	1	(0.894 to 1.119)
HDL Cholesterol		0.173 (0.158)	0.2762	1.188	(0.871 to 1.621)
SBP		0.006 (0.004)	0.1451	1.006	(0.998 to 1.013)
Family History	Negative	Ref			
	Positive	0.222 (0.126)	0.0785	1.248	(0.975 to 1.598)
CVD	No	Ref			
	Yes	-2.262 (0.715)	0.0016	0.104	(0.026 to 0.423)
Ethnicity	White	Ref			
	S.Asian	1.032 (0.297)	<0.001	2.806	(1.568 to 5.02)
	Black	0.561 (0.631)	0.3745	1.752	(0.508 to 6.039)
	Other/Mixed	-1.182 (0.218)	<0.001	0.307	(0.2 to 0.47)
BMI		0.017 (0.013)	0.1738	1.017	(0.993 to 1.042)
Consumes Alcohol	No	Ref			
	Yes	-0.427 (0.223)	0.0552	0.653	(0.422 to 1.01)
Diabetes	No	Ref			
	Yes	0.27 (0.191)	0.1587	1.31	(0.9 to 1.906)
Impaired Glucose Regulation	No	Ref			
	Yes	0.582 (0.638)	0.3615	1.79	(0.513 to 6.251)
Left Ventricular Hypertrophy	No	Ref			
	Yes	1.246 (0.802)	0.1204	3.475	(0.722 to 16.735)

Appendix Table F 14. Adjusted regression for missing data. – Ethnicity

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.031 (0.003)	<0.001	0.97	(0.963 to 0.976)
Gender	Female	Ref			
	Male	0.427 (0.073)	<0.001	1.533	(1.33 to 1.768)
Smoking Status	Non/Ex	Ref			
	Current	0.049 (0.07)	0.484	1.05	(0.916 to 1.203)
Total Cholesterol		0.093 (0.032)	0.0033	1.097	(1.031 to 1.167)
HDL Cholesterol		0.221 (0.089)	0.0135	1.247	(1.047 to 1.486)
SBP		0.001 (0.002)	0.6707	1.001	(0.997 to 1.005)
Family History	Negative	Ref			
	Positive	0.356 (0.07)	<0.001	1.428	(1.246 to 1.636)
CVD	No	Ref			
	Yes	-1.565 (0.301)	<0.001	0.209	(0.116 to 0.377)
Townsend Deprivation Score	1	Ref			
	2	-0.085 (0.087)	0.3334	0.919	(0.774 to 1.091)
	3	-0.221 (0.094)	0.0187	0.802	(0.667 to 0.964)
	4	-0.453 (0.106)	<0.001	0.636	(0.516 to 0.783)
	5	-0.346 (0.132)	0.0089	0.707	(0.545 to 0.917)
BMI		-0.025 (0.008)	0.0018	0.975	(0.96 to 0.991)
Consumes Alcohol	No	Ref			
	Yes	0.386 (0.174)	0.0263	1.471	(1.046 to 2.067)
Diabetes	No	Ref			
	Yes	-0.233 (0.14)	0.0954	0.792	(0.603 to 1.042)
Impaired Glucose Regulation	No	Ref			
	Yes	-0.41 (0.625)	0.5121	0.664	(0.195 to 2.26)
Left Ventricular Hypertrophy	No	Ref			
	Yes	-0.438 (0.787)	0.5779	0.645	(0.138 to 3.019)

Appendix Table F 15. Adjusted regression for missing data. – BMI

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.008 (0.005)	0.1121	0.992	(0.981 to 1.002)
Gender	Female	Ref			
	Male	0.309 (0.119)	0.0094	1.362	(1.079 to 1.719)
Smoking Status	Non/Ex	Ref			
	Current	0.019 (0.114)	0.8664	1.019	(0.815 to 1.275)
Total Cholesterol		-0.025 (0.052)	0.6209	0.975	(0.881 to 1.078)
HDL Cholesterol		0.146 (0.138)	0.2894	1.157	(0.883 to 1.516)
SBP		0.007 (0.003)	0.0401	1.007	(1 to 1.014)
Family History	Negative	Ref			
	Positive	-0.085 (0.111)	0.445	0.919	(0.739 to 1.142)
CVD	No	Ref			
	Yes	0.432 (0.211)	0.0403	1.541	(1.019 to 2.33)
Townsend Deprivation Score	1	Ref			
	2	-0.083 (0.157)	0.6	0.921	(0.676 to 1.253)
	3	0.003 (0.159)	0.9854	1.003	(0.734 to 1.371)
	4	0.195 (0.161)	0.2249	1.215	(0.887 to 1.666)
	5	0.319 (0.191)	0.0949	1.375	(0.946 to 1.999)
Ethnicity	White	Ref			
	S.Asian	-0.757 (0.609)	0.2141	0.469	(0.142 to 1.548)
	Black	0.555 (0.558)	0.3192	1.743	(0.584 to 5.199)
	Other/Mixed	-0.129 (0.137)	0.3454	0.879	(0.672 to 1.149)
Consumes Alcohol	No	Ref			
	Yes	-0.421 (0.215)	0.0507	0.657	(0.431 to 1.001)
Diabetes	No	Ref			
	Yes	-1.867 (0.39)	<0.001	0.155	(0.072 to 0.332)
Impaired Glucose Regulation	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)

Appendix Table F 16. Adjusted regression for missing data. – Consumes Alcohol

Risk factor		Beta (SE)	p-value	OR	95% CI
Age		-0.001 (0.004)	0.7349	0.999	(0.992 to 1.006)
Gender	Female	Ref			
	Male	-0.189 (0.078)	0.0154	0.828	(0.71 to 0.964)
Smoking Status	Non/Ex	Ref			
	Current	-1.176 (0.074)	<0.001	0.308	(0.267 to 0.356)
Total Cholesterol		-0.018 (0.034)	0.5986	0.982	(0.919 to 1.05)
HDL Cholesterol		-0.096 (0.098)	0.3265	0.908	(0.749 to 1.101)
SBP		0.006 (0.002)	0.0069	1.006	(1.002 to 1.01)
Family History	Negative	Ref			
	Positive	-0.123 (0.072)	0.0884	0.885	(0.768 to 1.019)
CVD	No	Ref			
	Yes	0.168 (0.146)	0.2489	1.183	(0.889 to 1.576)
Townsend Deprivation Score	1	Ref			
	2	0.143 (0.099)	0.1518	1.153	(0.949 to 1.401)
	3	0.066 (0.103)	0.5218	1.068	(0.873 to 1.306)
	4	0.112 (0.108)	0.2983	1.119	(0.905 to 1.383)
	5	-0.084 (0.135)	0.5344	0.92	(0.706 to 1.198)
Ethnicity	White	Ref			
	S.Asian	-0.227 (0.296)	0.4422	0.797	(0.446 to 1.422)
	Black	-0.272 (0.481)	0.5711	0.761	(0.297 to 1.955)
	Other/Mixed	-0.132 (0.09)	0.1439	0.876	(0.734 to 1.046)
BMI		0.016 (0.007)	0.0328	1.016	(1.001 to 1.031)
Diabetes	No	Ref			
	Yes	0.922 (0.099)	<0.001	2.515	(2.073 to 3.051)
Impaired Glucose Regulation	No	Ref			
	Yes	0.385 (0.369)	0.2961	1.47	(0.714 to 3.029)
Left Ventricular Hypertrophy	No	Ref			
	Yes	0 (0)	<0.001	1	(1 to 1)

Appendix Table F 17. Unadjusted regression for missing data.

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
Smoking	Missing Total Cholesterol	No			Ref	
		Yes	0.227 (0.045)	<0.001	1.255	(1.149 to 1.371)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.302 (0.077)	<0.001	1.352	(1.163 to 1.572)
	Missing SBP	No			Ref	
		Yes	0.843 (0.03)	<0.001	2.324	(2.192 to 2.463)
	Missing Townsend Score	No			Ref	
		Yes	-0.025 (0.03)	0.4008	0.975	(0.92 to 1.034)
	Missing Ethnicity	No			Ref	
		Yes	0.065 (0.029)	0.0233	1.067	(1.009 to 1.129)
	Missing BMI	No			Ref	
		Yes	0.872 (0.029)	<0.001	2.391	(2.258 to 2.532)
	Missing Alcohol Consumption	No			Ref	
		Yes	1.197 (0.032)	<0.001	3.309	(3.109 to 3.521)
Total Cholesterol	Missing Smoking Status	No			Ref	
		Yes	0.227 (0.045)	<0.001	1.255	(1.149 to 1.371)
	Missing HDL Cholesterol	No			Ref	
		Yes	5.504 (0.045)	<0.001	245.683	(224.935 to 268.344)
	Missing SBP	No			Ref	
		Yes	1.578 (0.021)	<0.001	4.848	(4.649 to 5.055)
	Missing Townsend Score	No			Ref	
		Yes	0.268 (0.011)	<0.001	1.308	(1.279 to 1.337)
	Missing Ethnicity	No			Ref	
		Yes	0.272 (0.011)	<0.001	1.313	(1.286 to 1.34)
	Missing BMI	No			Ref	
		Yes	0.851 (0.012)	<0.001	2.342	(2.286 to 2.4)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.56 (0.011)	<0.001	1.751	(1.713 to 1.79)
HDL Cholesterol	Missing Smoking Status	No			Ref	
		Yes	0.302 (0.077)	<0.001	1.352	(1.163 to 1.572)
	Missing Total Cholesterol	No			Ref	
		Yes	5.504 (0.045)	<0.001	245.682	(224.935 to 268.344)
	Missing SBP	No			Ref	
		Yes	1.627 (0.038)	<0.001	5.087	(4.72 to 5.482)
	Missing Townsend Score	No			Ref	
		Yes	0.077 (0.018)	<0.001	1.08	(1.043 to 1.118)
	Missing Ethnicity	No			Ref	
		Yes	0.123 (0.017)	<0.001	1.131	(1.094 to 1.169)
	Missing BMI	No			Ref	
		Yes	0.847 (0.021)	<0.001	2.332	(2.24 to 2.428)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.557 (0.018)	<0.001	1.745	(1.684 to 1.809)

[Table continues on the next page]

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
SBP	Missing Smoking Status	No			Ref	
		Yes	0.843 (0.03)	<0.001	2.324	(2.192 to 2.463)
	Missing Total Cholesterol	No			Ref	
		Yes	1.578 (0.021)	<0.001	4.848	(4.649 to 5.055)
	Missing HDL Cholesterol	No			Ref	
		Yes	1.627 (0.038)	<0.001	5.087	(4.72 to 5.482)
	Missing Townsend Score	No			Ref	
		Yes	0.211 (0.009)	<0.001	1.235	(1.213 to 1.258)
	Missing Ethnicity	No			Ref	
		Yes	0.305 (0.009)	<0.001	1.356	(1.332 to 1.38)
	Missing BMI	No			Ref	
		Yes	3.174 (0.014)	<0.001	23.914	(23.288 to 24.556)
	Missing Alcohol Consumption	No			Ref	
		Yes	2.758 (0.013)	<0.001	15.771	(15.368 to 16.185)
Townsend Score	Missing Smoking Status	No			Ref	
		Yes	-0.025 (0.03)	0.4008	0.975	(0.92 to 1.034)
	Missing Total Cholesterol	No			Ref	
		Yes	0.268 (0.011)	<0.001	1.308	(1.279 to 1.337)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.077 (0.018)	<0.001	1.08	(1.043 to 1.118)
	Missing SBP	No			Ref	
		Yes	0.211 (0.009)	<0.001	1.235	(1.213 to 1.258)
	Missing Ethnicity	No			Ref	
		Yes	2.654 (0.01)	<0.001	14.211	(13.944 to 14.483)
	Missing BMI	No			Ref	
		Yes	0.18 (0.008)	<0.001	1.197	(1.18 to 1.216)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.19 (0.008)	<0.001	1.21	(1.192 to 1.228)
Ethnicity	Missing Smoking Status	No			Ref	
		Yes	0.065 (0.029)	0.0233	1.067	(1.009 to 1.129)
	Missing Total Cholesterol	No			Ref	
		Yes	0.272 (0.011)	<0.001	1.313	(1.286 to 1.34)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.123 (0.017)	<0.001	1.131	(1.094 to 1.169)
	Missing SBP	No			Ref	
		Yes	0.305 (0.009)	<0.001	1.356	(1.332 to 1.38)
	Missing Townsend Score	No			Ref	
		Yes	2.654 (0.01)	<0.001	14.211	(13.944 to 14.483)
	Missing BMI	No			Ref	
		Yes	0.126 (0.008)	<0.001	1.135	(1.118 to 1.151)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.053 (0.007)	<0.001	1.054	(1.039 to 1.07)

[Table continues on the next page]

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
BMI	Missing Smoking Status	No			Ref	
		Yes	0.872 (0.029)	<0.001	2.391	(2.258 to 2.532)
	Missing Total Cholesterol	No			Ref	
		Yes	0.851 (0.012)	<0.001	2.342	(2.286 to 2.4)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.847 (0.021)	<0.001	2.332	(2.24 to 2.428)
	Missing SBP	No			Ref	
		Yes	3.174 (0.014)	<0.001	23.914	(23.288 to 24.556)
	Missing Townsend Score	No			Ref	
		Yes	0.18 (0.008)	<0.001	1.197	(1.18 to 1.216)
	Missing Ethnicity	No			Ref	
		Yes	0.126 (0.008)	<0.001	1.135	(1.118 to 1.151)
	Missing Alcohol Consumption	No			Ref	
		Yes	3.232 (0.01)	<0.001	25.343	(24.84 to 25.855)
Alcohol Consumption	Missing Smoking Status	No			Ref	
		Yes	1.197 (0.032)	<0.001	3.309	(3.109 to 3.521)
	Missing Total Cholesterol	No			Ref	
		Yes	0.56 (0.011)	<0.001	1.751	(1.713 to 1.79)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.557 (0.018)	<0.001	1.745	(1.684 to 1.809)
	Missing SBP	No			Ref	
		Yes	2.758 (0.013)	<0.001	15.771	(15.368 to 16.185)
	Missing Townsend Score	No			Ref	
		Yes	0.19 (0.008)	<0.001	1.21	(1.192 to 1.228)
	Missing Ethnicity	No			Ref	
		Yes	0.053 (0.007)	<0.001	1.054	(1.039 to 1.07)
	Missing BMI	No			Ref	
		Yes	3.232 (0.01)	<0.001	25.343	(24.84 to 25.855)

Appendix Table F 18. Adjusted regression for missing data.

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
Smoking	Missing Total Cholesterol	No			Ref	
		Yes	-0.042 (0.053)	0.4292	0.959	(0.864 to 1.064)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.105 (0.09)	0.242	1.111	(0.932 to 1.324)
	Missing SBP	No			Ref	
		Yes	0.301 (0.036)	<0.001	1.351	(1.259 to 1.449)
	Missing Townsend Score	No			Ref	
		Yes	-0.144 (0.035)	<0.001	0.866	(0.808 to 0.928)
	Missing Ethnicity	No			Ref	
		Yes	0.104 (0.034)	0.0023	1.11	(1.038 to 1.187)
	Missing BMI	No			Ref	
		Yes	0.084 (0.04)	0.0356	1.088	(1.006 to 1.177)
	Missing Alcohol Consumption	No			Ref	
		Yes	1.026 (0.04)	<0.001	2.789	(2.576 to 3.019)
Total Cholesterol	Missing Smoking Status	No			Ref	
		Yes	-0.038 (0.053)	0.4767	0.963	(0.868 to 1.068)
	Missing HDL Cholesterol	No			Ref	
		Yes	5.464 (0.045)	<0.001	236.021	(215.922 to 257.992)
	Missing SBP	No			Ref	
		Yes	1.212 (0.027)	<0.001	3.36	(3.189 to 3.539)
	Missing Townsend Score	No			Ref	
		Yes	0.204 (0.016)	<0.001	1.227	(1.189 to 1.266)
	Missing Ethnicity	No			Ref	
		Yes	0.174 (0.015)	<0.001	1.19	(1.155 to 1.225)
	Missing BMI	No			Ref	
		Yes	0.416 (0.018)	<0.001	1.516	(1.462 to 1.572)
	Missing Alcohol Consumption	No			Ref	
		Yes	-0.042 (0.016)	0.0106	0.959	(0.929 to 0.99)
HDL Cholesterol	Missing Smoking Status	No			Ref	
		Yes	0.07 (0.09)	0.436	1.073	(0.899 to 1.28)
	Missing Total Cholesterol	No			Ref	
		Yes	5.464 (0.045)	<0.001	236.002	(215.91 to 257.963)
	Missing SBP	No			Ref	
		Yes	0.305 (0.047)	<0.001	1.356	(1.237 to 1.487)
	Missing Townsend Score	No			Ref	
		Yes	-0.203 (0.026)	<0.001	0.816	(0.775 to 0.859)
	Missing Ethnicity	No			Ref	
		Yes	-0.055 (0.025)	0.0263	0.946	(0.901 to 0.993)
	Missing BMI	No			Ref	
		Yes	0.08 (0.029)	0.0061	1.084	(1.023 to 1.148)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.035 (0.025)	0.1691	1.036	(0.985 to 1.088)

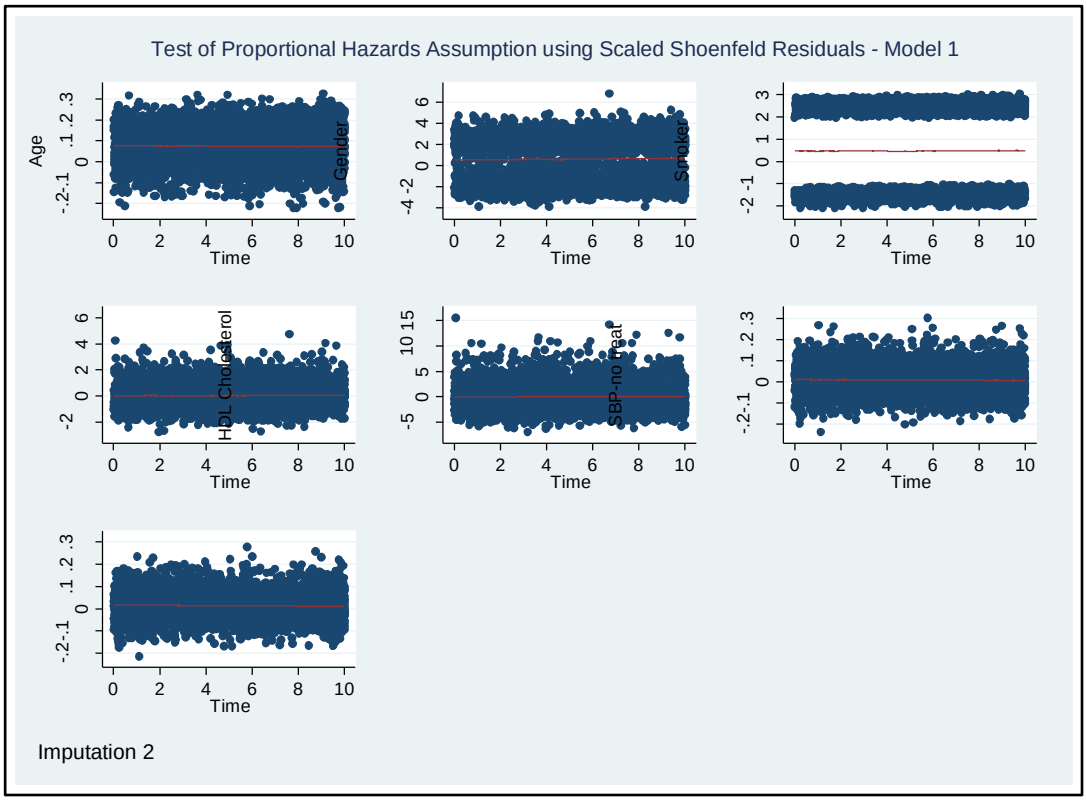
[Table continues on the next page]

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
SBP	Missing Smoking Status	No			Ref	
		Yes	0.334 (0.036)	<0.001	1.397	(1.302 to 1.499)
	Missing Total Cholesterol	No			Ref	
		Yes	1.184 (0.027)	<0.001	3.266	(3.098 to 3.443)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.274 (0.048)	<0.001	1.315	(1.197 to 1.445)
	Missing Townsend Score	No			Ref	
		Yes	-0.089 (0.013)	<0.001	0.914	(0.891 to 0.939)
	Missing Ethnicity	No			Ref	
		Yes	0.352 (0.013)	<0.001	1.422	(1.386 to 1.459)
	Missing BMI	No			Ref	
		Yes	2.321 (0.016)	<0.001	10.181	(9.876 to 10.496)
	Missing Alcohol Consumption	No			Ref	
		Yes	1.426 (0.016)	<0.001	4.16	(4.034 to 4.29)
Townsend Score	Missing Smoking Status	No			Ref	
		Yes	-0.151 (0.036)	<0.001	0.86	(0.802 to 0.922)
	Missing Total Cholesterol	No			Ref	
		Yes	0.21 (0.016)	<0.001	1.234	(1.195 to 1.273)
	Missing HDL Cholesterol	No			Ref	
		Yes	-0.193 (0.025)	<0.001	0.824	(0.784 to 0.867)
	Missing SBP	No			Ref	
		Yes	-0.116 (0.013)	<0.001	0.89	(0.868 to 0.914)
	Missing Ethnicity	No			Ref	
		Yes	2.66 (0.01)	<0.001	14.297	(14.028 to 14.573)
	Missing BMI	No			Ref	
		Yes	0.032 (0.013)	0.0156	1.032	(1.006 to 1.059)
	Missing Alcohol Consumption	No			Ref	
		Yes	0.249 (0.012)	<0.001	1.283	(1.253 to 1.314)
Ethnicity	Missing Smoking Status	No			Ref	
		Yes	0.107 (0.034)	0.0018	1.113	(1.04 to 1.19)
	Missing Total Cholesterol	No			Ref	
		Yes	0.173 (0.015)	<0.001	1.189	(1.154 to 1.224)
	Missing HDL Cholesterol	No			Ref	
		Yes	-0.065 (0.024)	0.007	0.937	(0.893 to 0.982)
	Missing SBP	No			Ref	
		Yes	0.375 (0.013)	<0.001	1.455	(1.418 to 1.493)
	Missing Townsend Score	No			Ref	
		Yes	2.66 (0.01)	<0.001	14.301	(14.031 to 14.576)
	Missing BMI	No			Ref	
		Yes	0.016 (0.013)	0.2171	1.016	(0.991 to 1.042)
	Missing Alcohol Consumption	No			Ref	
		Yes	-0.233 (0.012)	<0.001	0.792	(0.774 to 0.811)

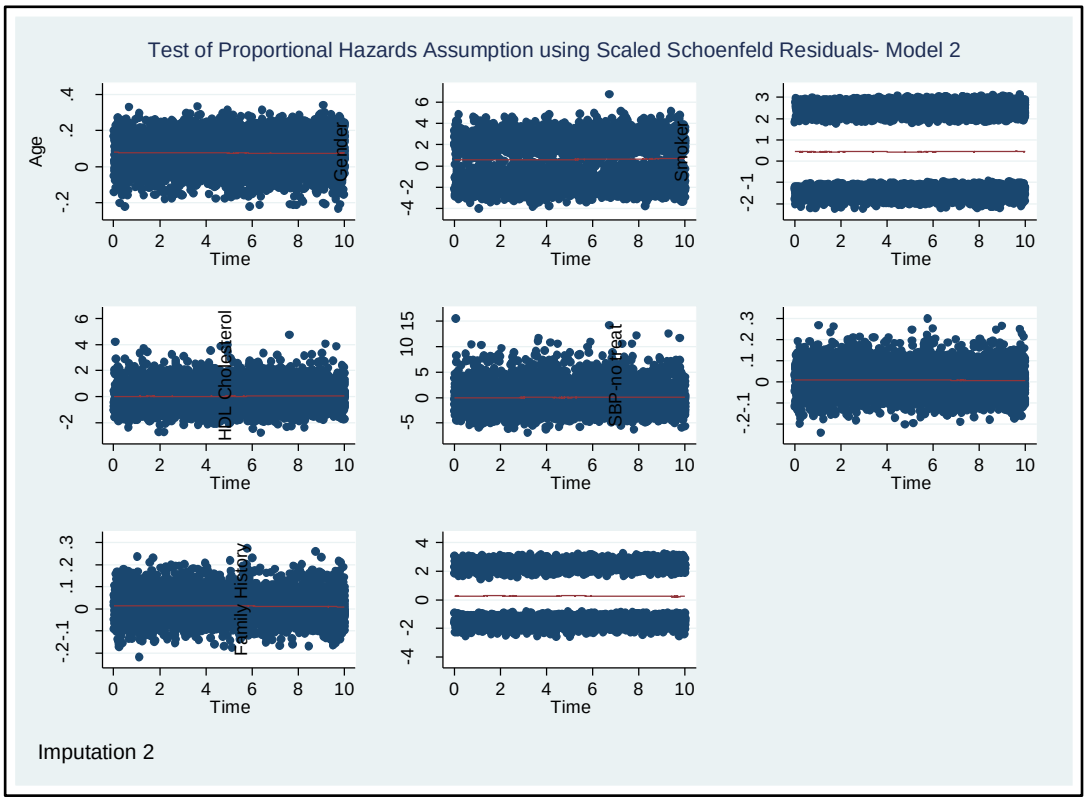
[Table continues on the next page]

Missingness Risk Factor	Risk factor		Beta (SE)	p-value	OR	95% CI
BMI	Missing Smoking Status	No			Ref	
		Yes	0.085 (0.039)	0.0311	1.089	(1.008 to 1.176)
	Missing Total Cholesterol	No			Ref	
		Yes	0.417 (0.018)	<0.001	1.517	(1.463 to 1.573)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.076 (0.03)	0.0113	1.079	(1.017 to 1.145)
	Missing SBP	No			Ref	
		Yes	2.322 (0.015)	<0.001	10.2	(9.896 to 10.514)
	Missing Townsend Score	No			Ref	
		Yes	0.03 (0.013)	0.0219	1.031	(1.004 to 1.058)
	Missing Ethnicity	No			Ref	
		Yes	0.018 (0.013)	0.1548	1.018	(0.993 to 1.045)
	Missing Alcohol Consumption	No			Ref	
		Yes	2.793 (0.011)	<0.001	16.331	(15.987 to 16.682)
Alcohol Consumption	Missing Smoking Status	No			Ref	
		Yes	1.001 (0.041)	<0.001	2.721	(2.512 to 2.946)
	Missing Total Cholesterol	No			Ref	
		Yes	-0.043 (0.016)	0.0081	0.958	(0.927 to 0.989)
	Missing HDL Cholesterol	No			Ref	
		Yes	0.021 (0.026)	0.4192	1.021	(0.97 to 1.075)
	Missing SBP	No			Ref	
		Yes	1.439 (0.016)	<0.001	4.217	(4.09 to 4.348)
	Missing Townsend Score	No			Ref	
		Yes	0.248 (0.012)	<0.001	1.281	(1.251 to 1.313)
	Missing Ethnicity	No			Ref	
		Yes	-0.232 (0.012)	<0.001	0.793	(0.775 to 0.812)
	Missing BMI	No			Ref	
		Yes	2.793 (0.011)	<0.001	16.333	(15.988 to 16.684)

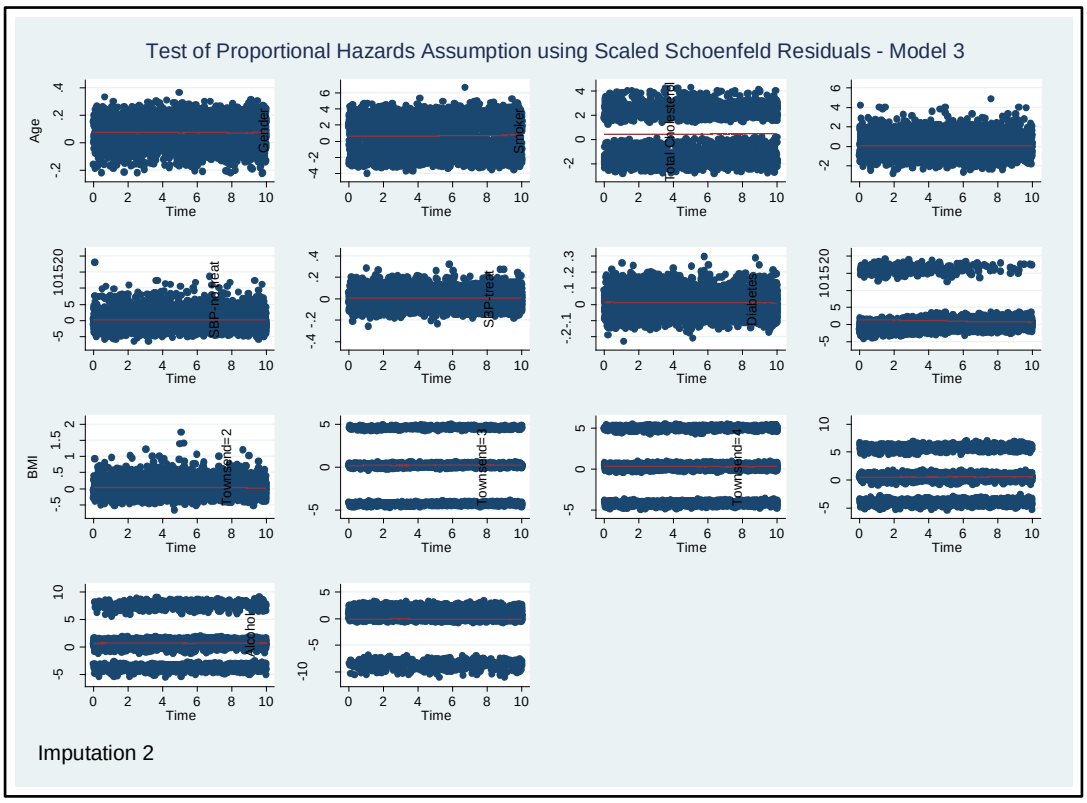
Appendix G. Schoenfeld residual plots for Proportional Hazards.



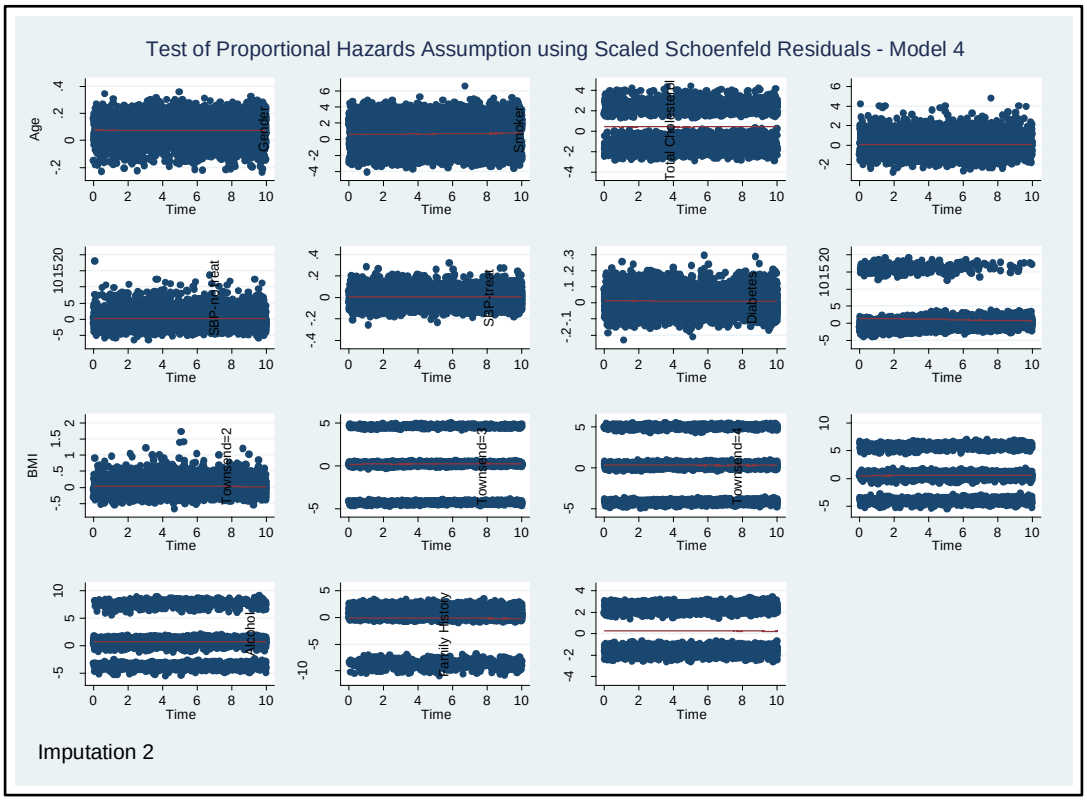
Appendix Figure G 1. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 2.



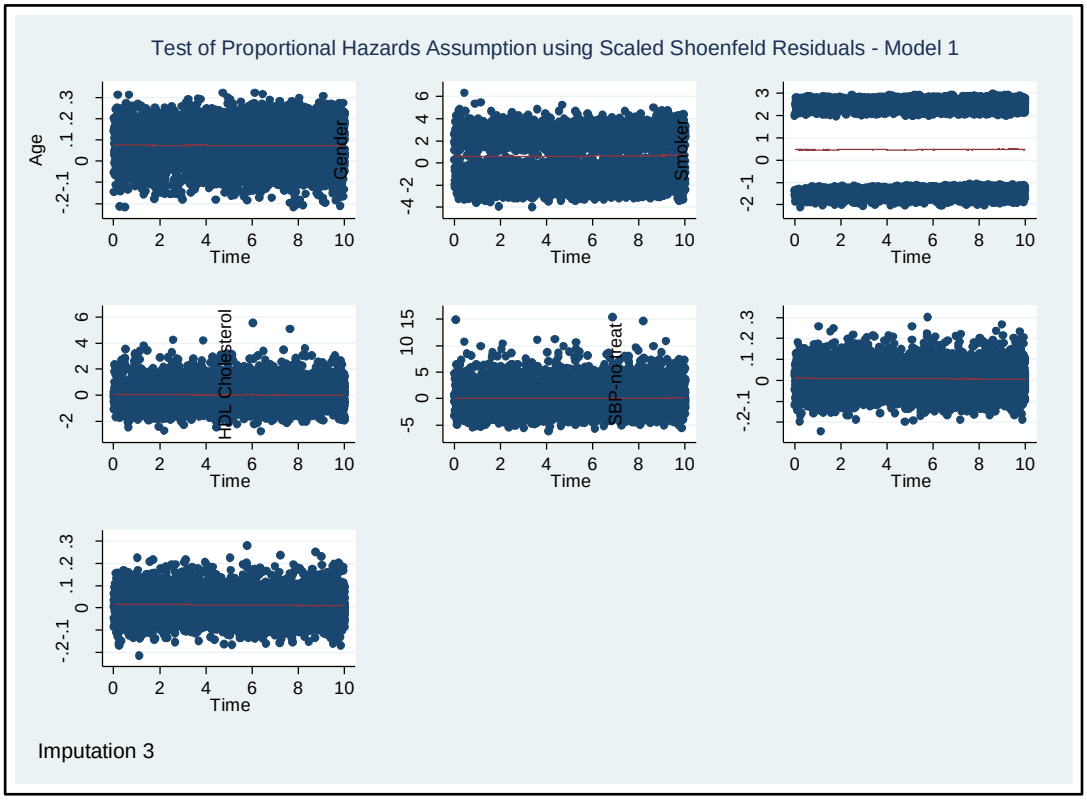
Appendix Figure G 2. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 2.



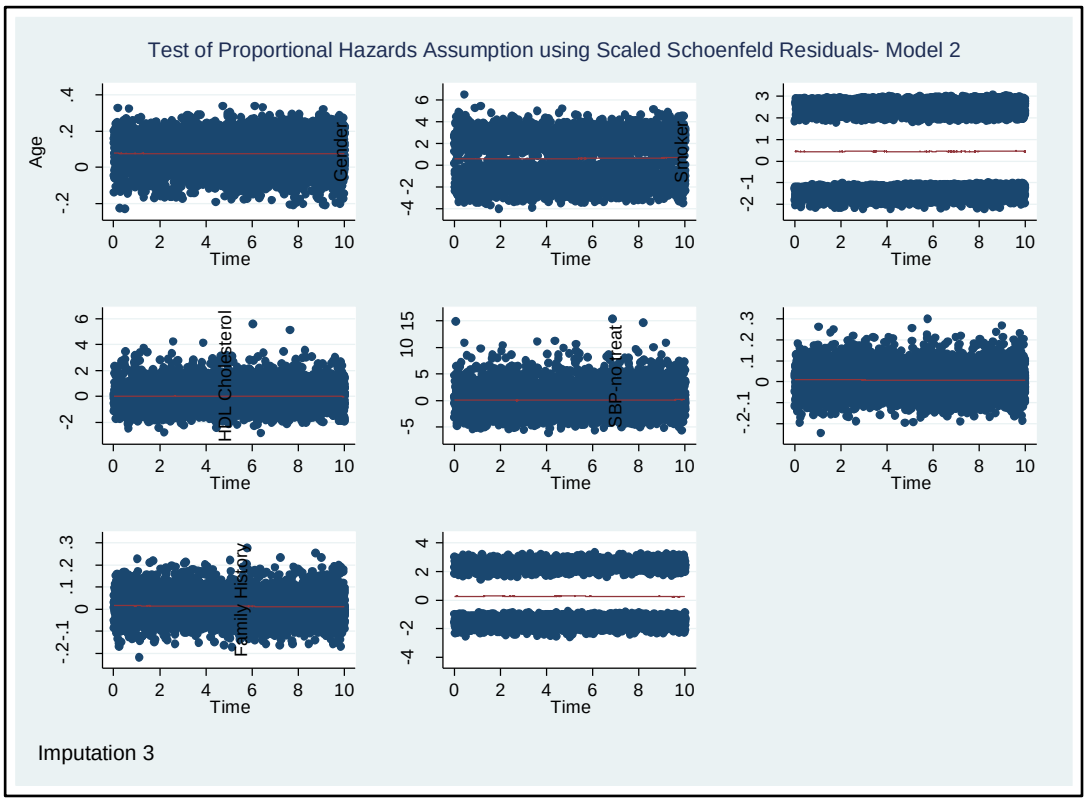
Appendix Figure G 3. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 2.



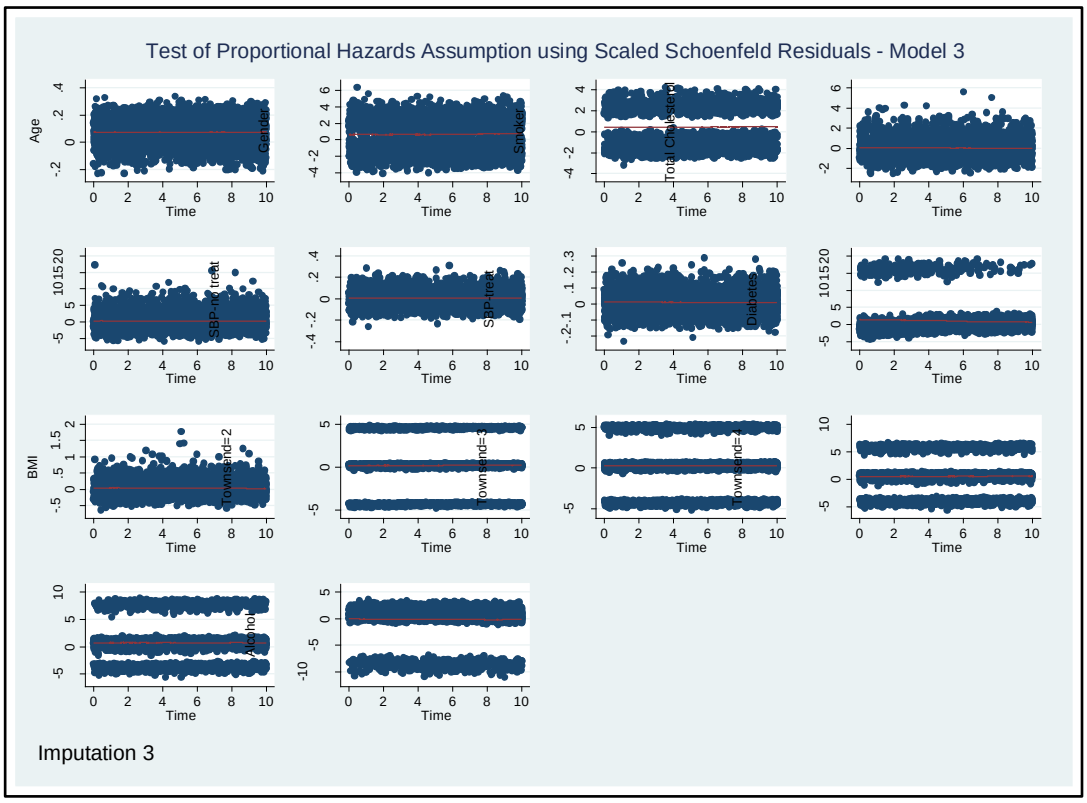
Appendix Figure G 4. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 2.



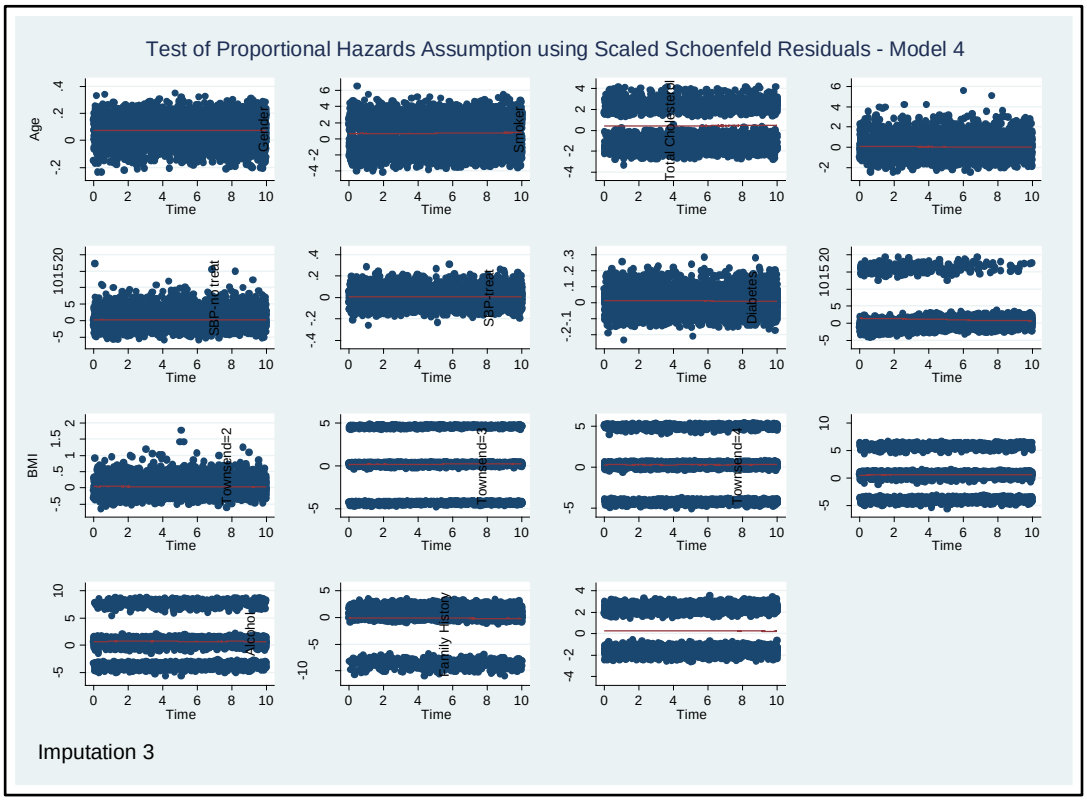
Appendix Figure G 5. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 3.



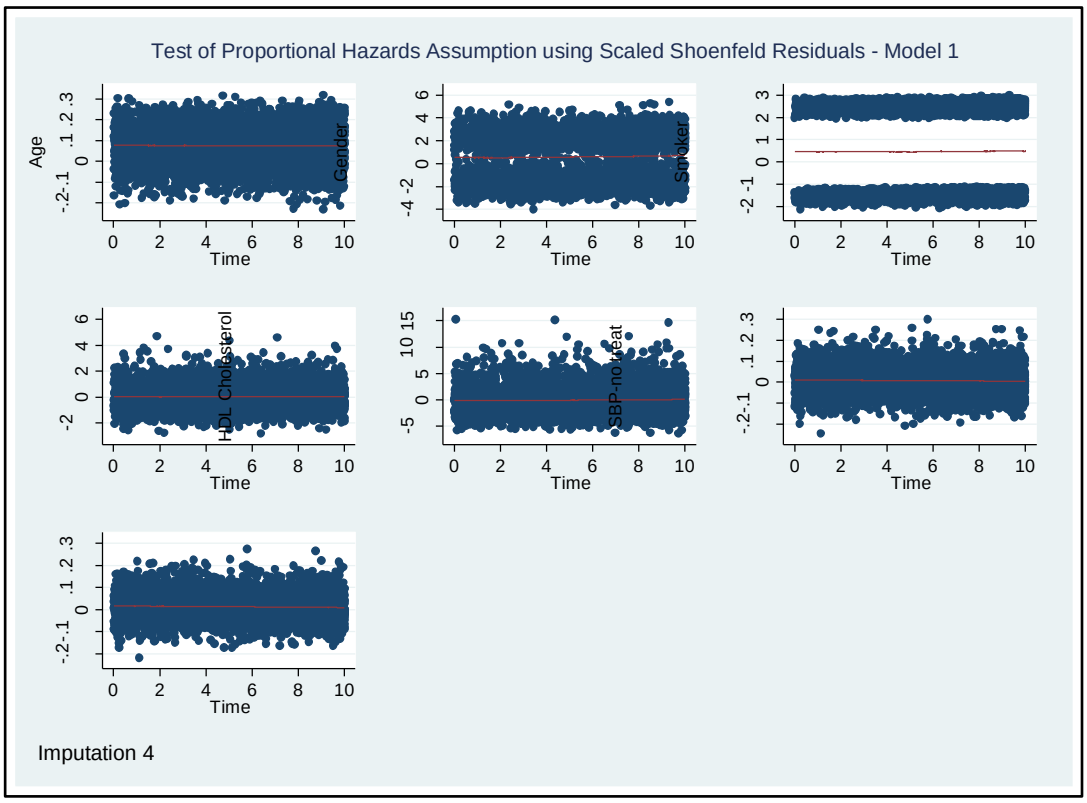
Appendix Figure G 6. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 3.



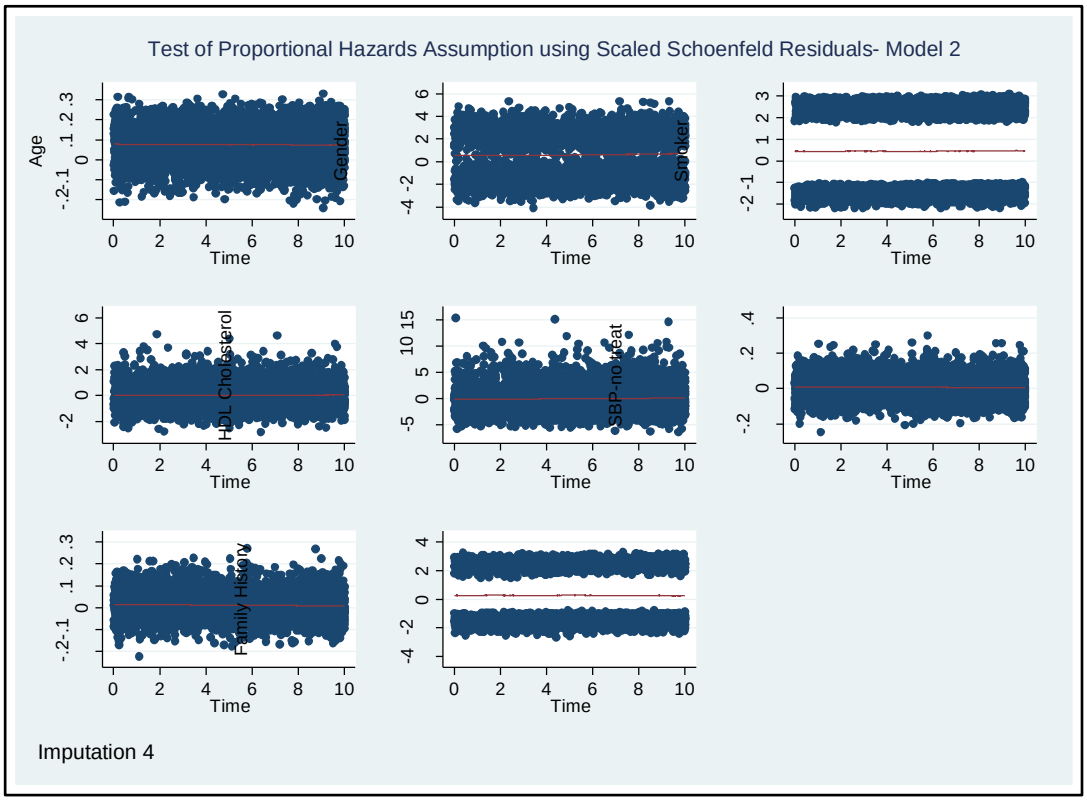
Appendix Figure G 7. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 3.



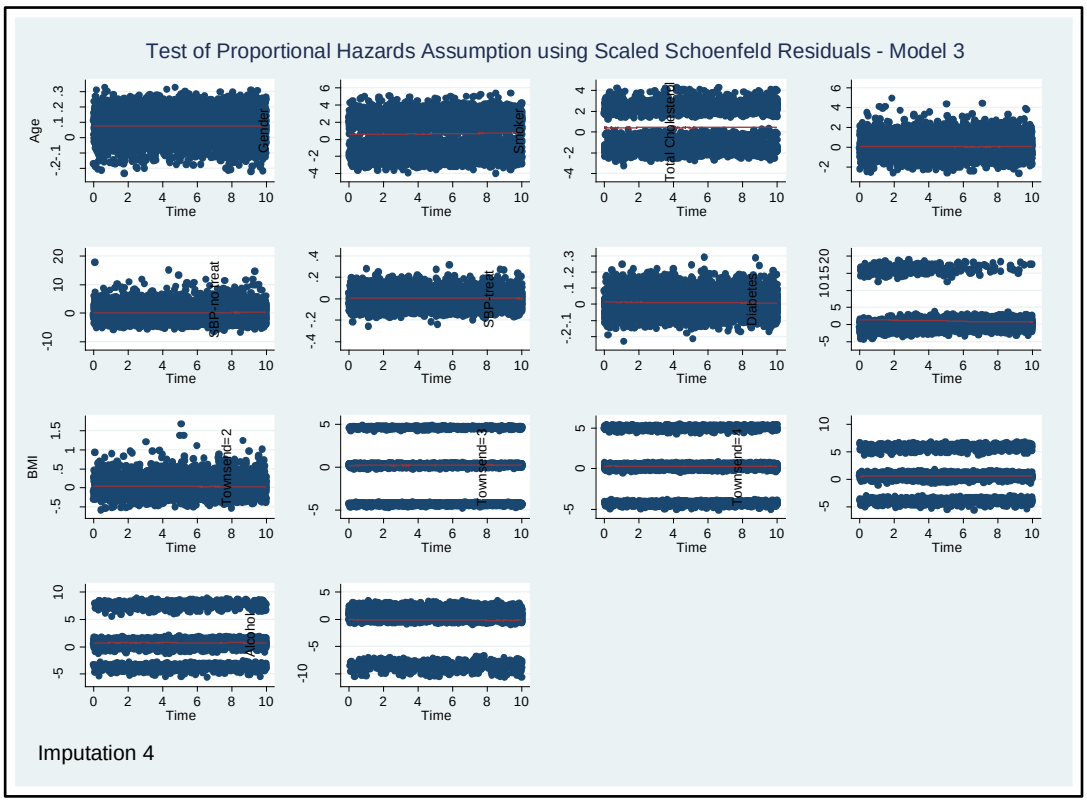
Appendix Figure G 8. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 3.



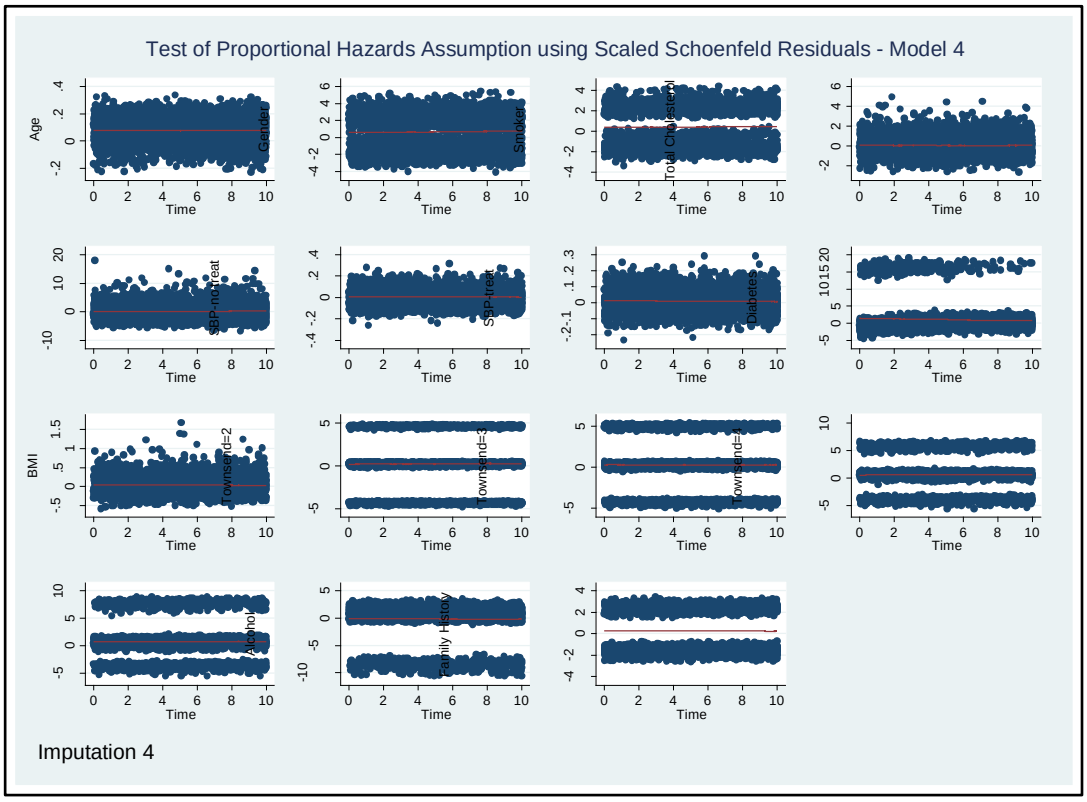
Appendix Figure G 9. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 4.



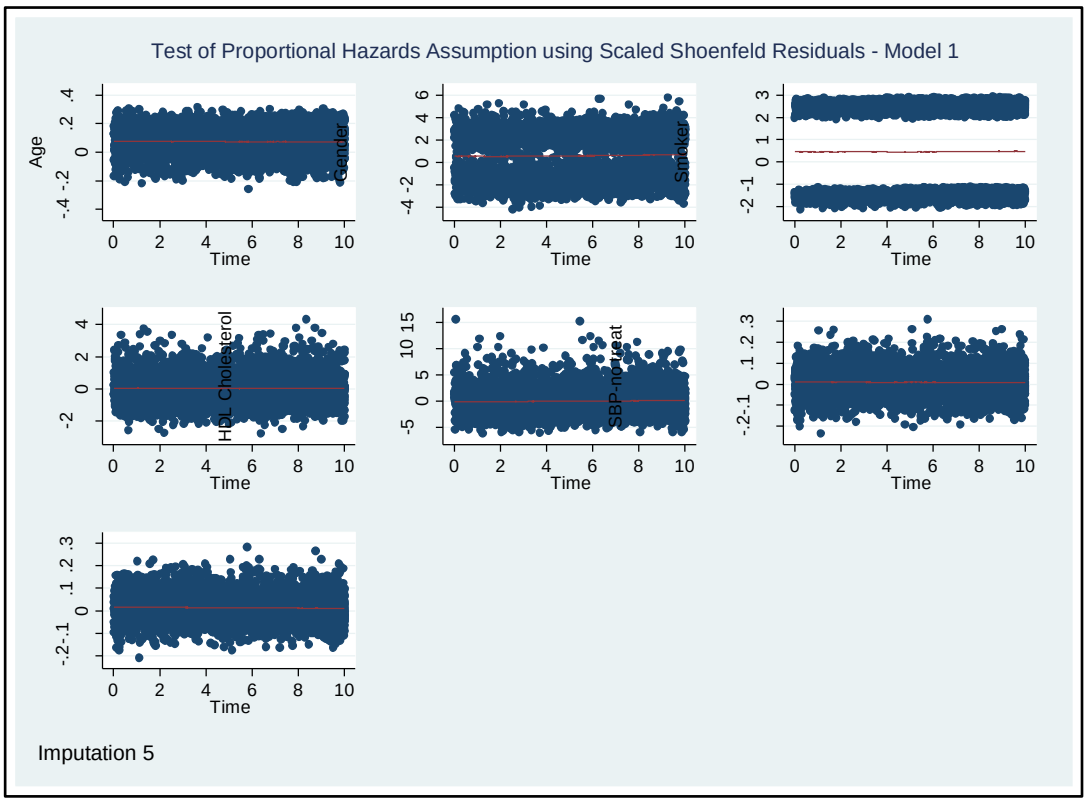
Appendix Figure G 10. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 4.



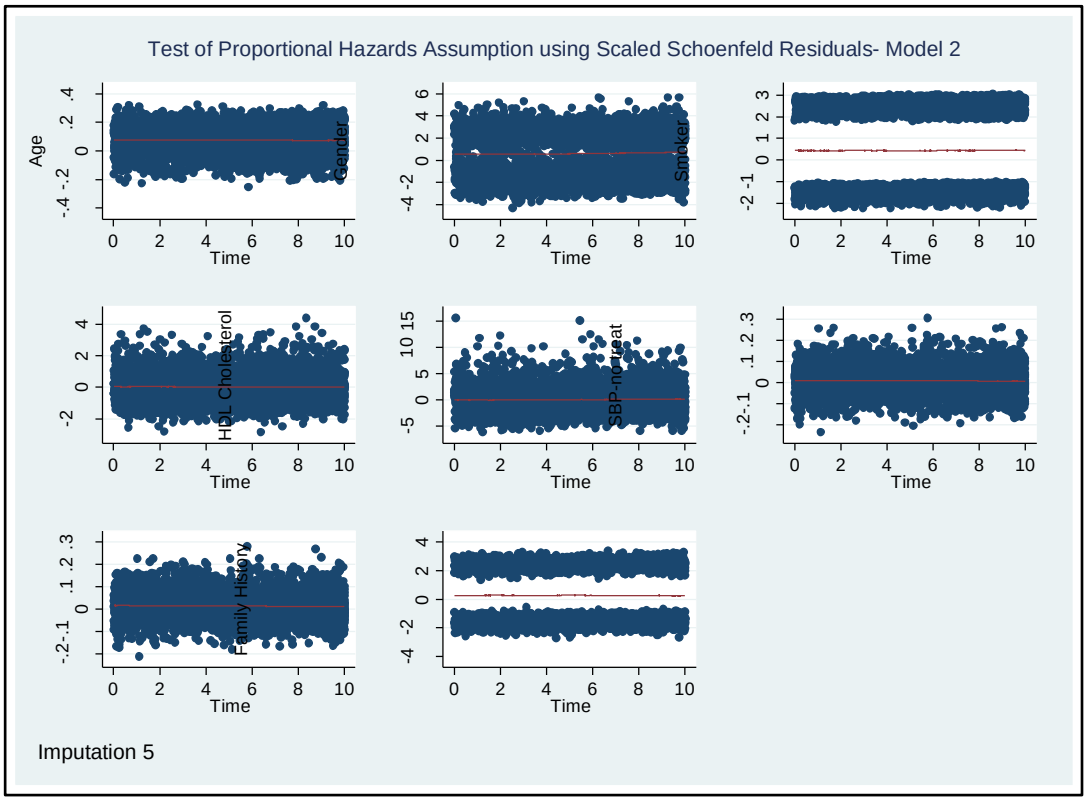
Appendix Figure G 11. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 4.



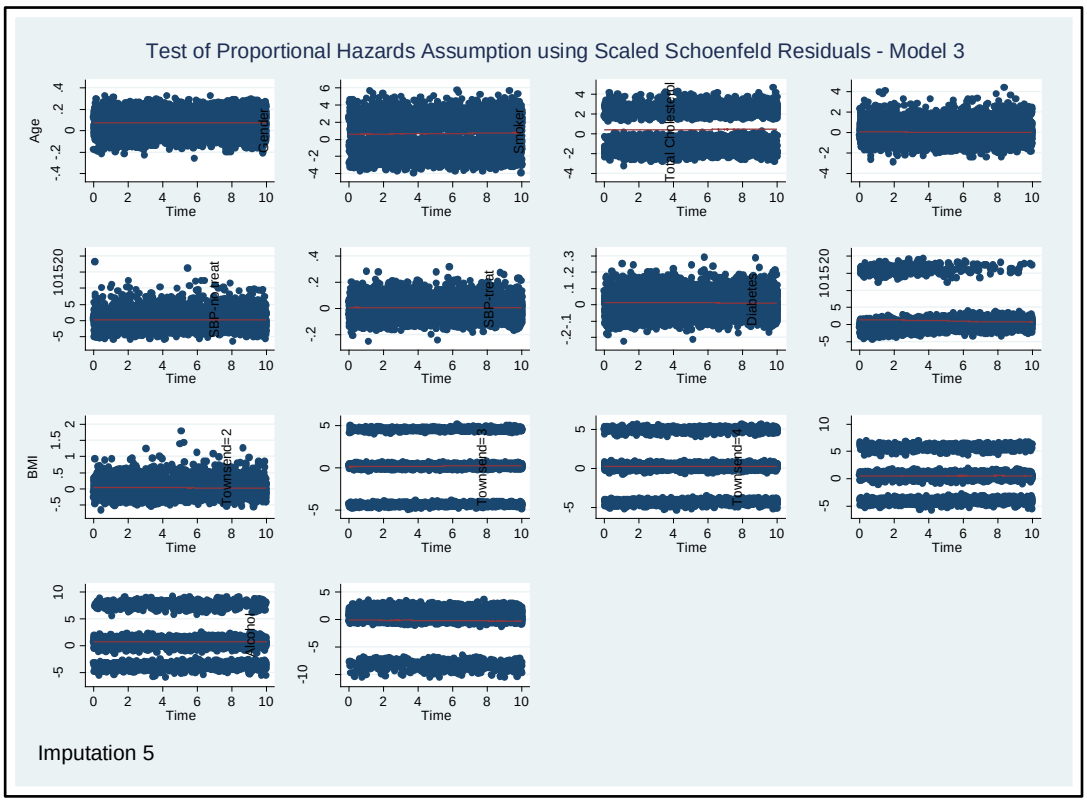
Appendix Figure G 12. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 4.



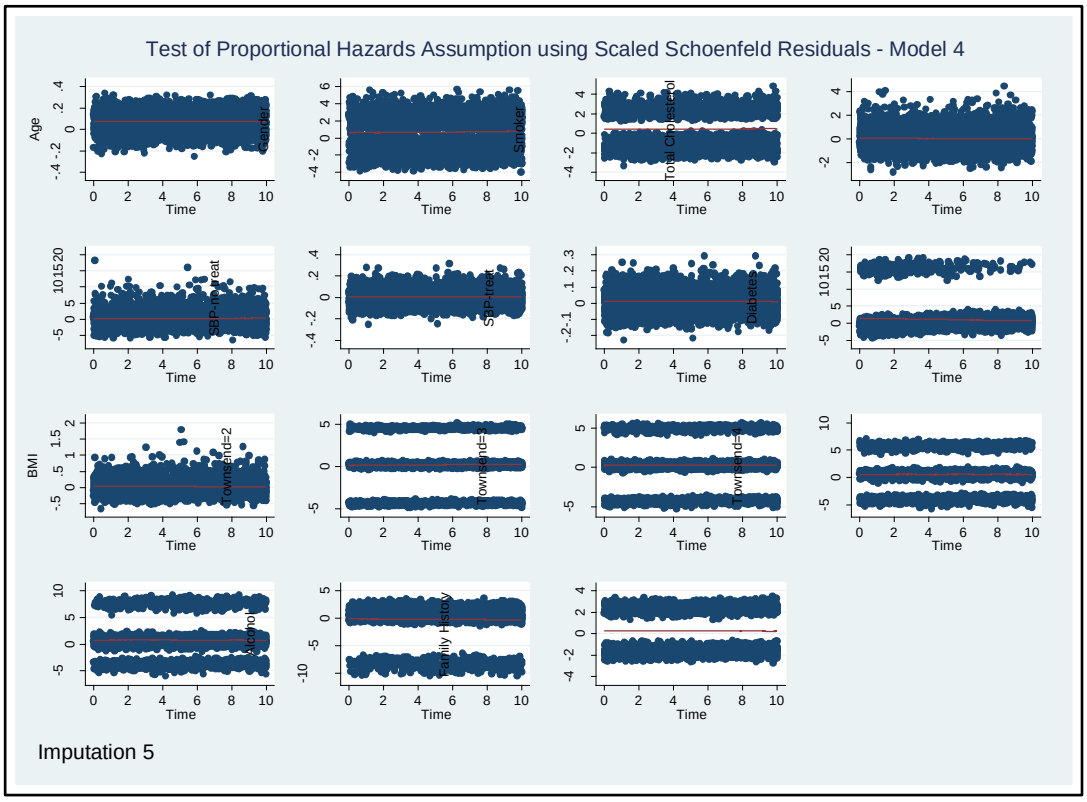
Appendix Figure G 13. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 5.



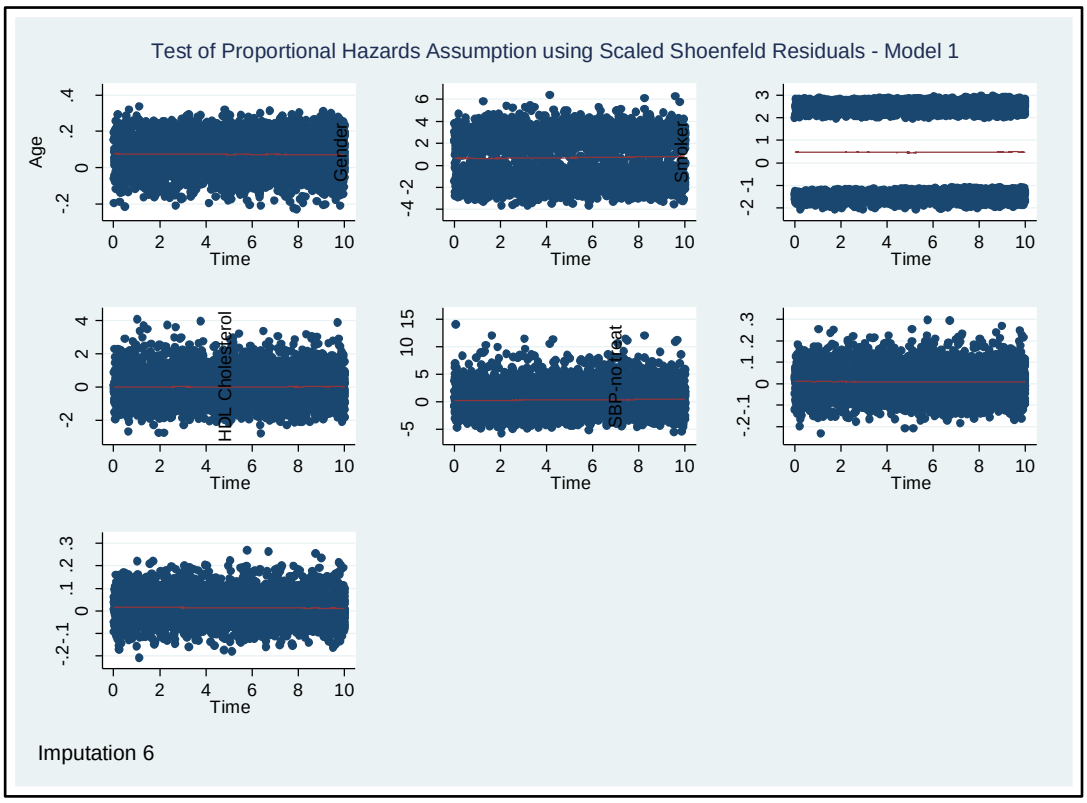
Appendix Figure G 14. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 5.



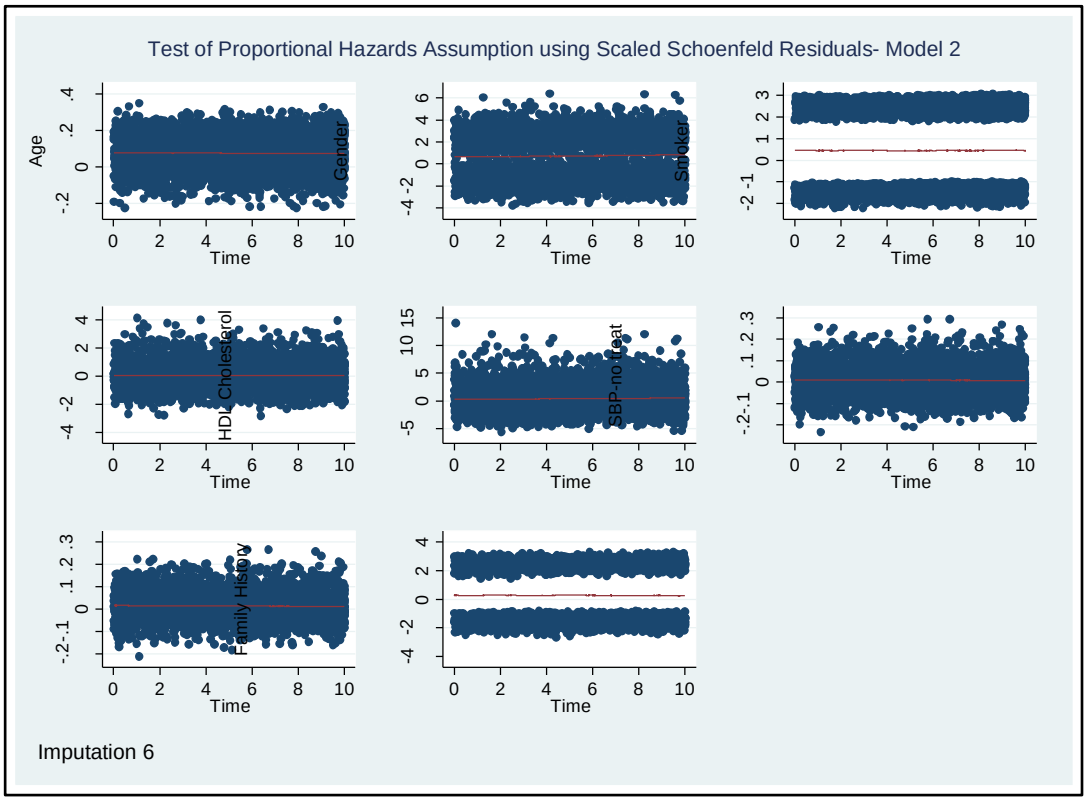
Appendix Figure G 15. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 5.



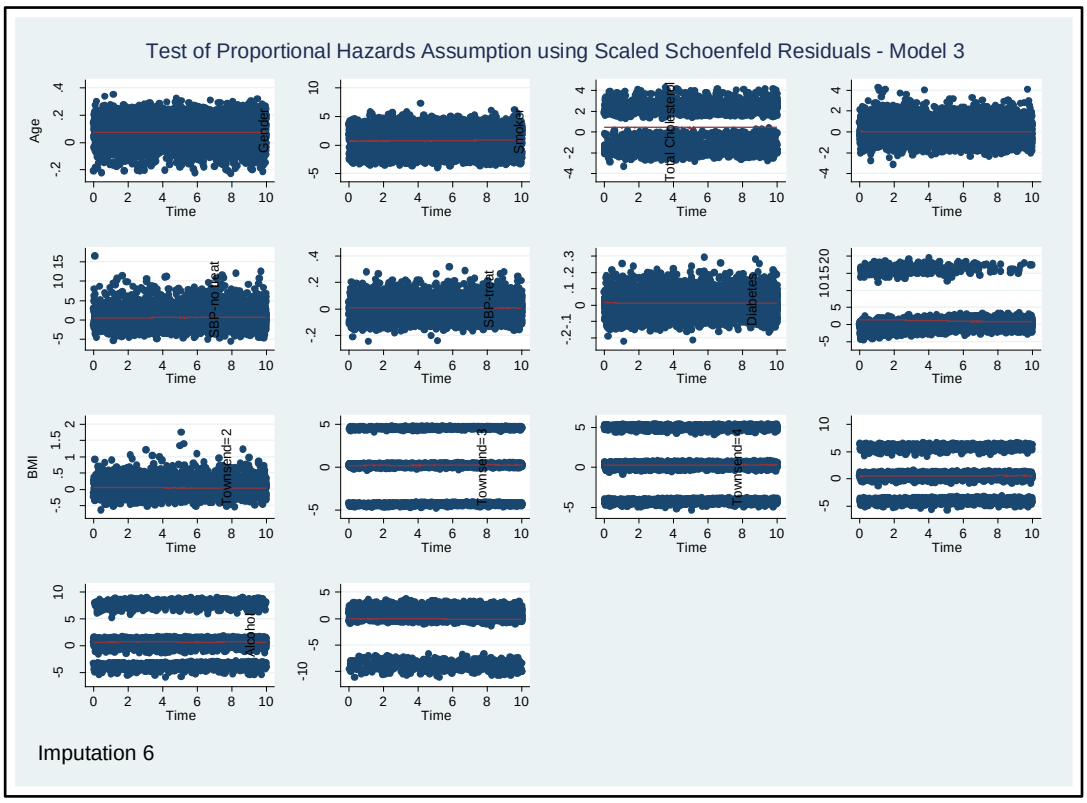
Appendix Figure G 16. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 5.



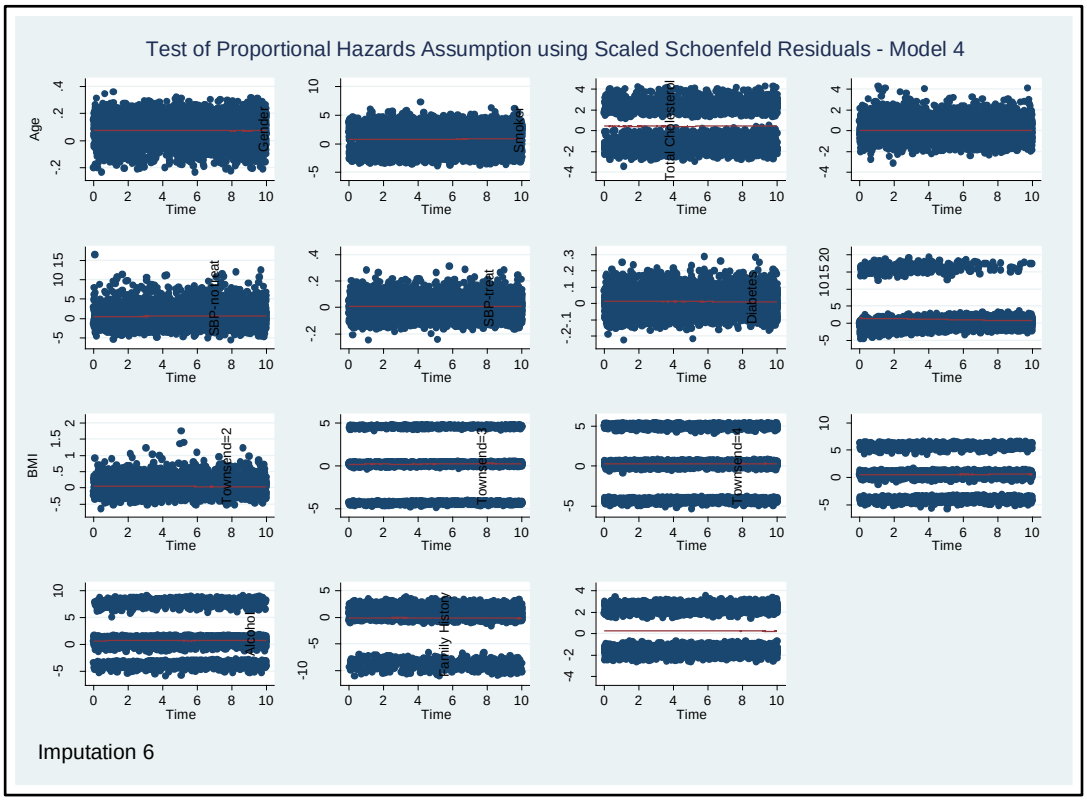
Appendix Figure G 17. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 6.



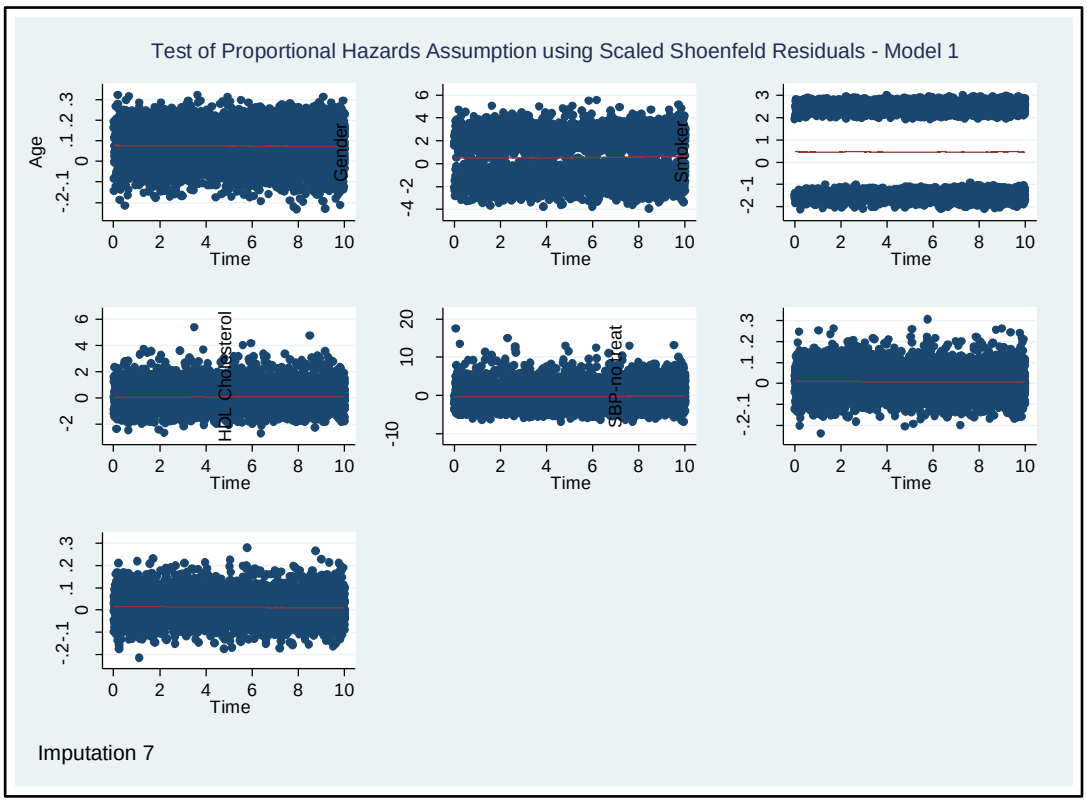
Appendix Figure G 18. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 6.



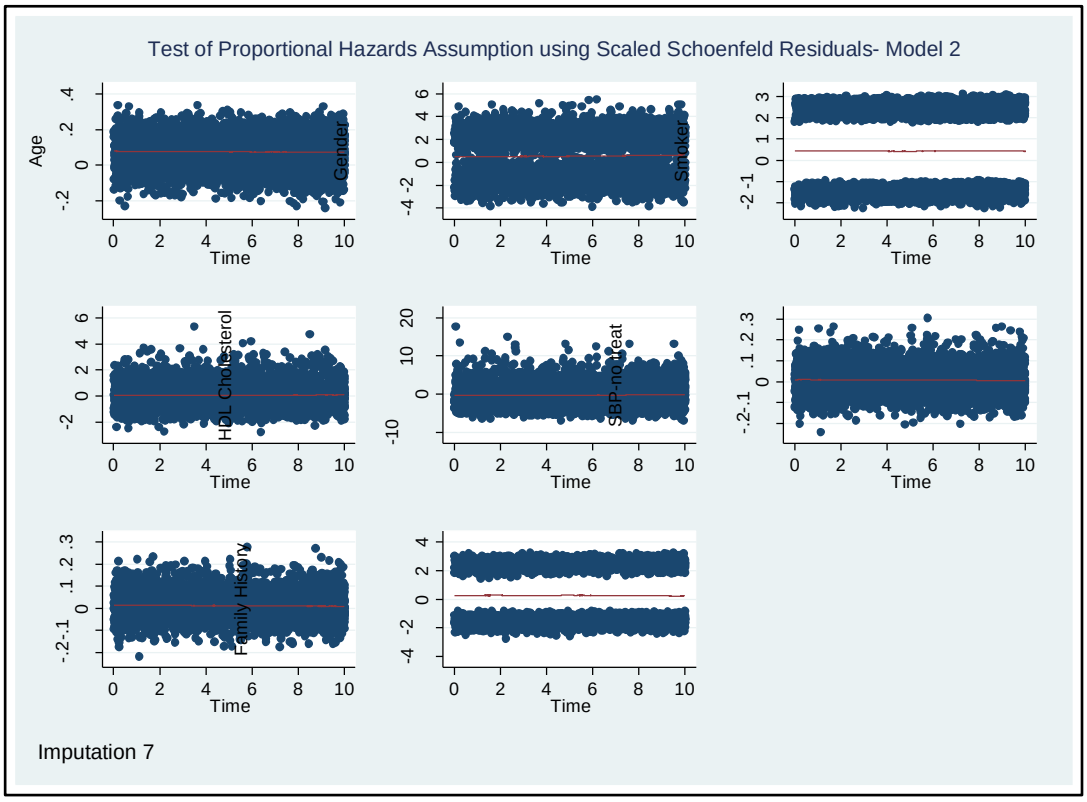
Appendix Figure G 19. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 6.



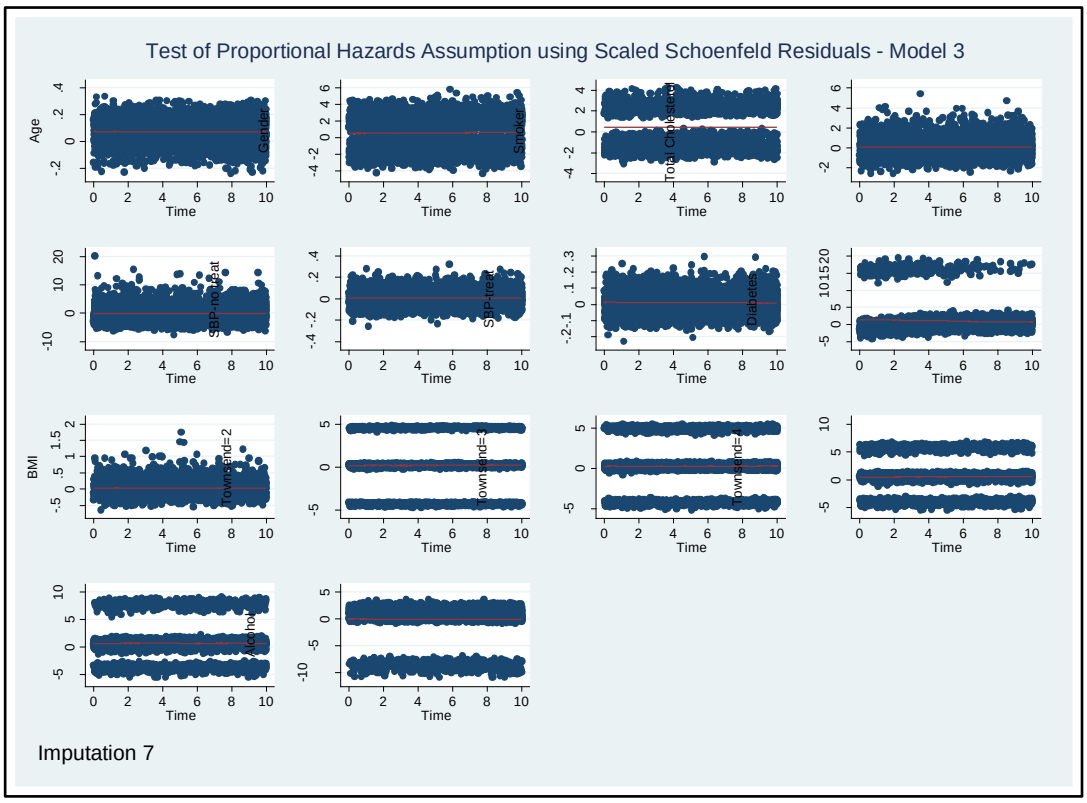
Appendix Figure G 20. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 6.



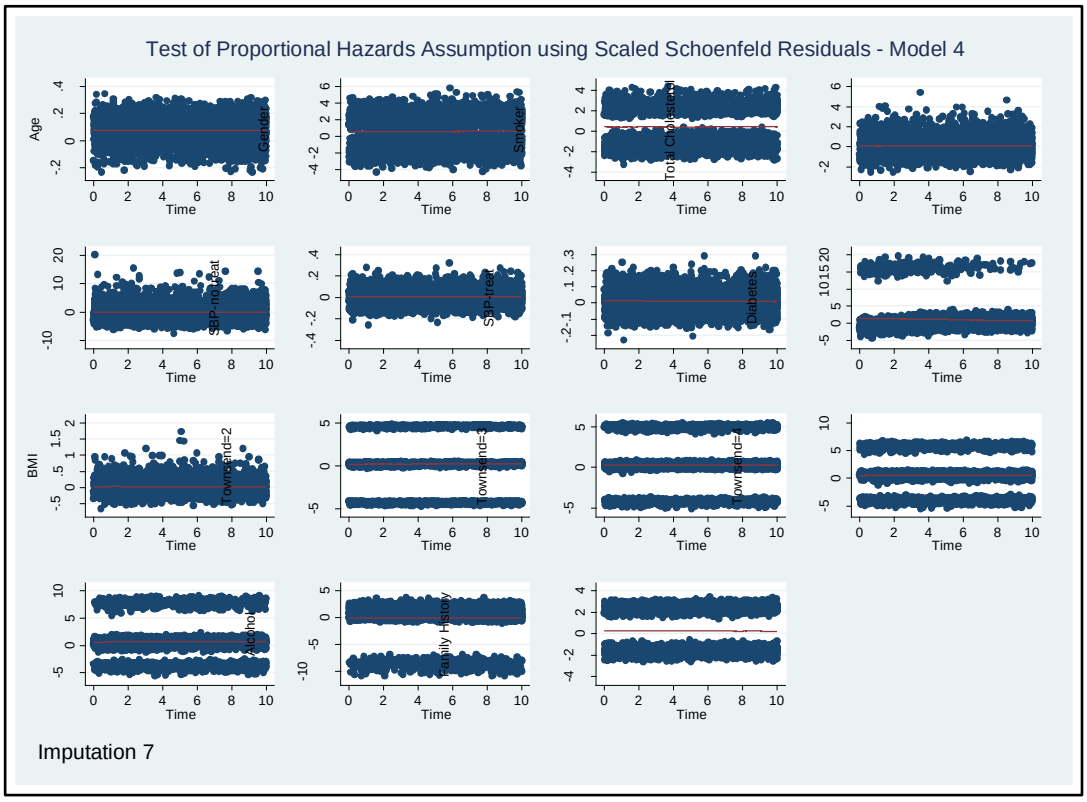
Appendix Figure G 21. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 7.



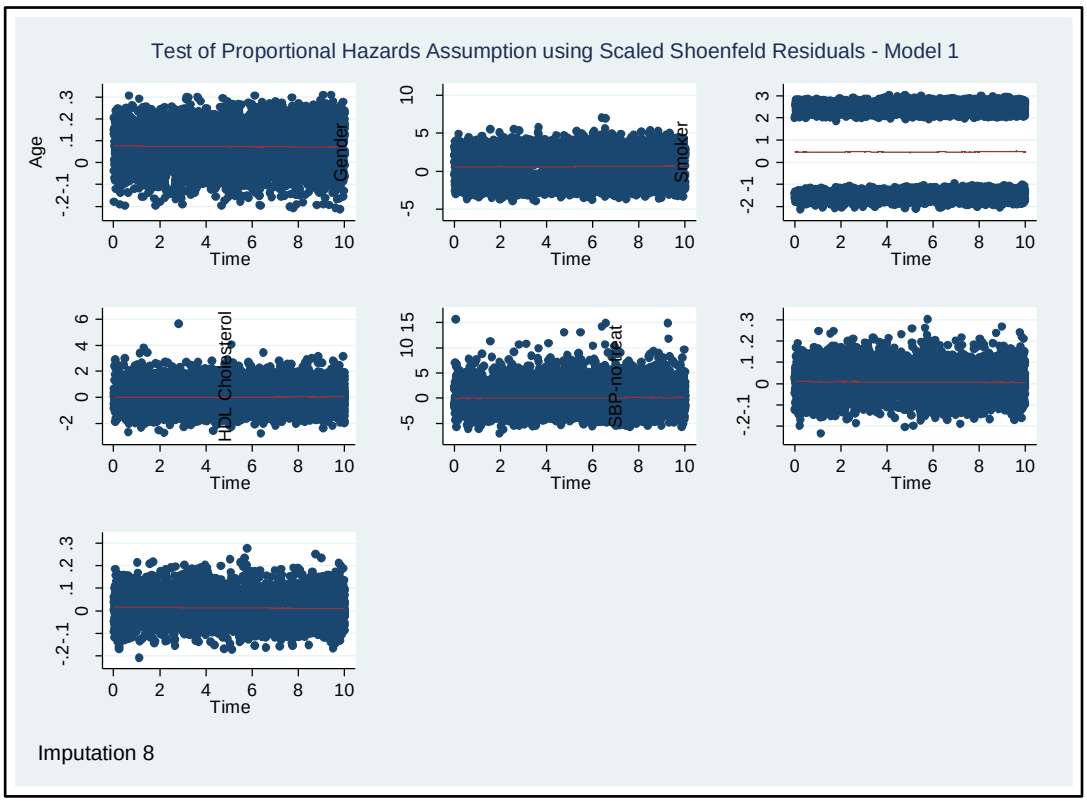
Appendix Figure G 22. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 7.



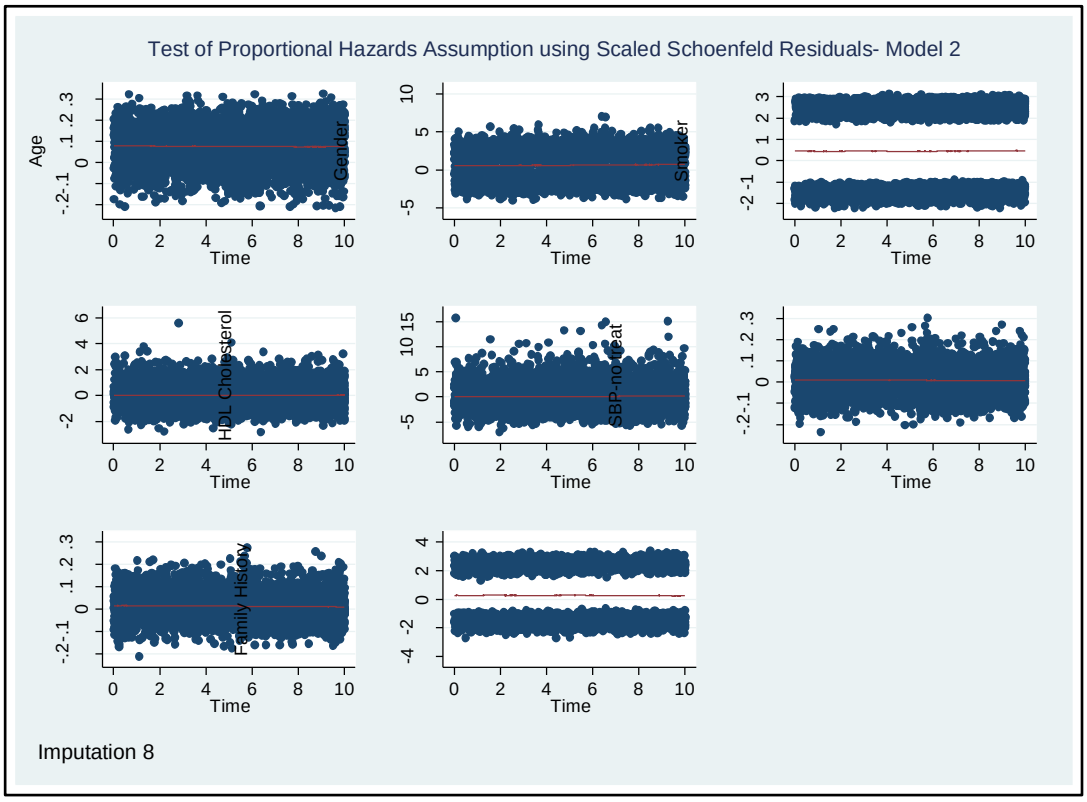
Appendix Figure G 23. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 7.



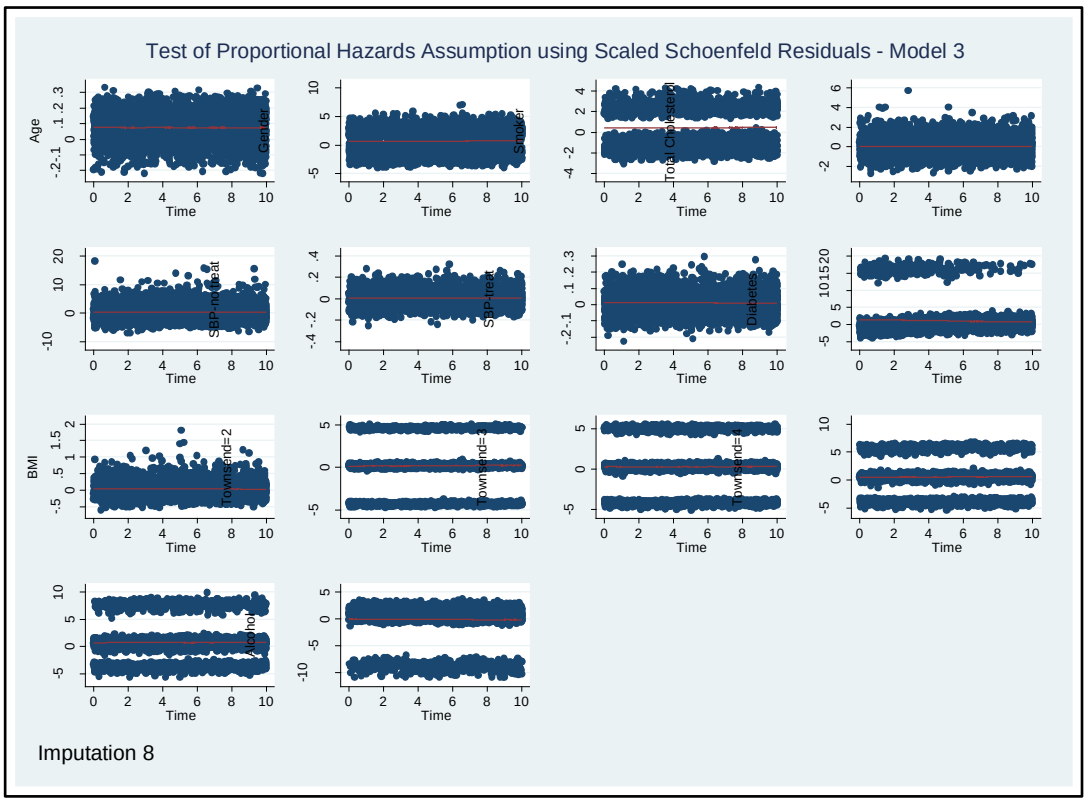
Appendix Figure G 24. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 7.



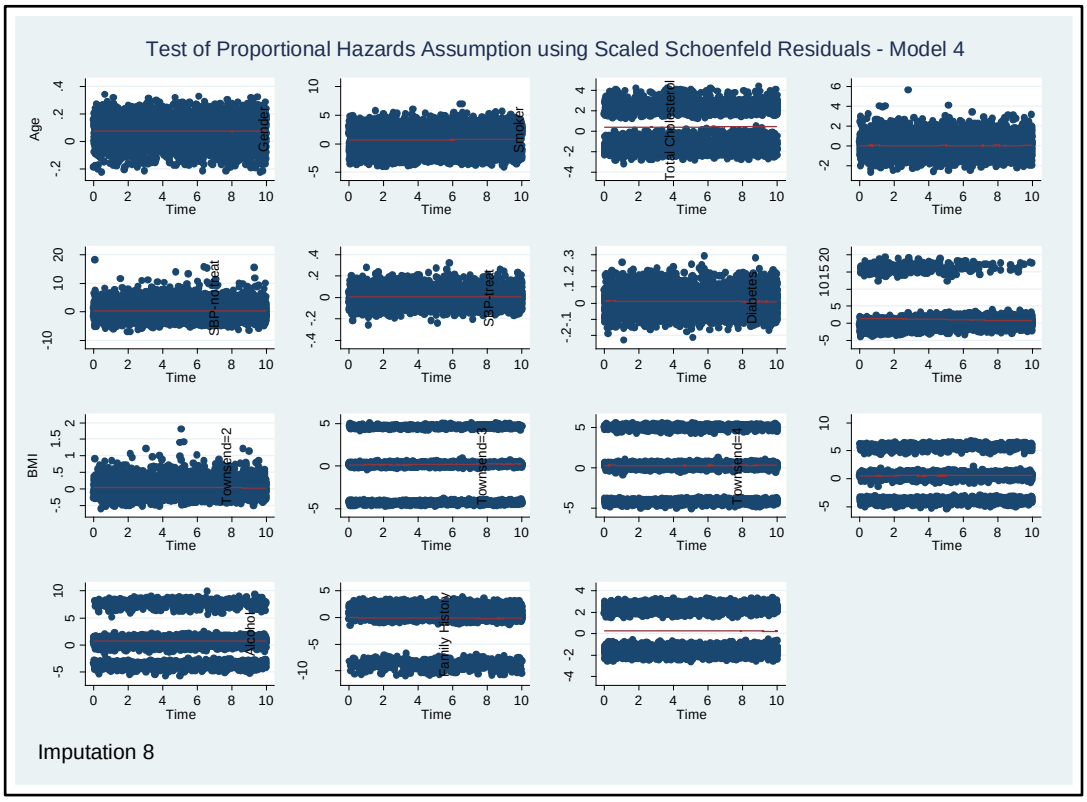
Appendix Figure G 25. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 8.



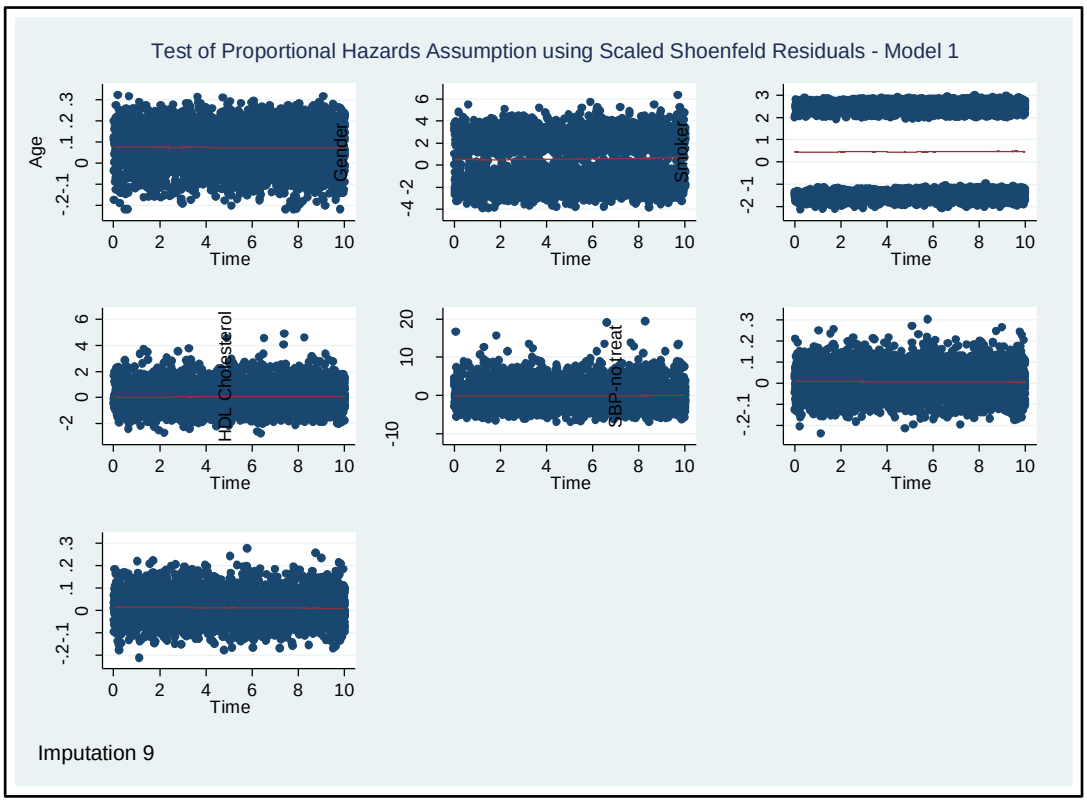
Appendix Figure G 26. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 8.



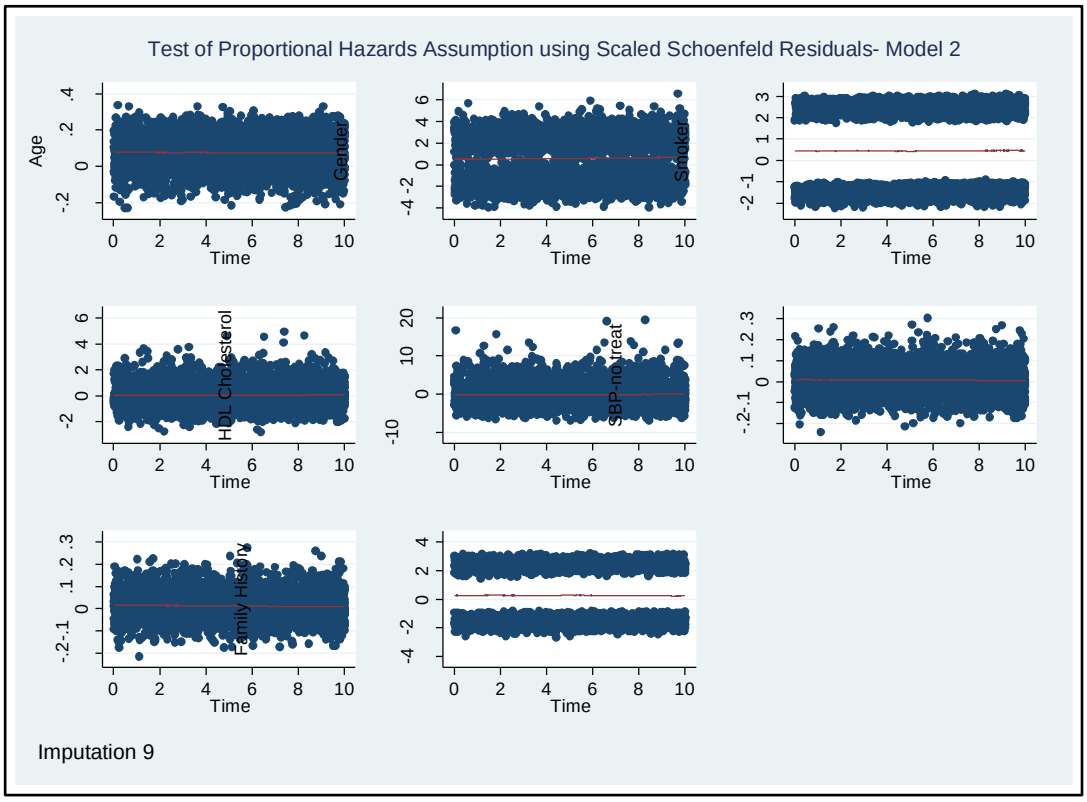
Appendix Figure G 27. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 8.



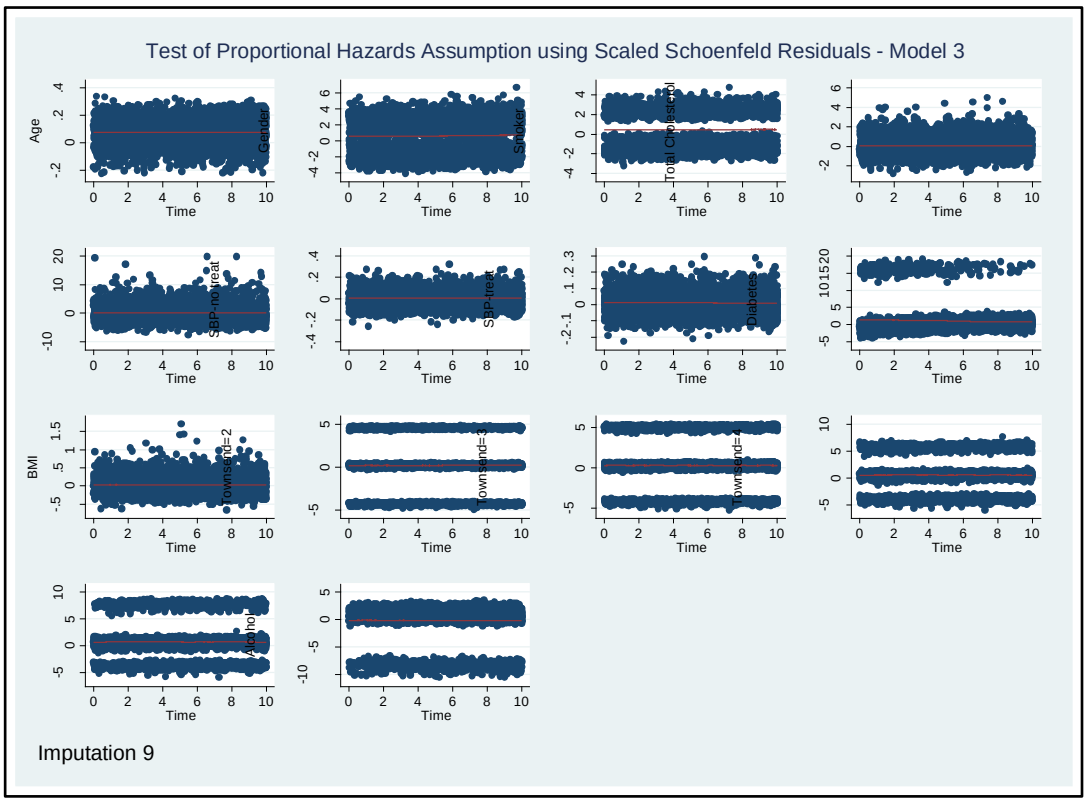
Appendix Figure G 28. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 8.



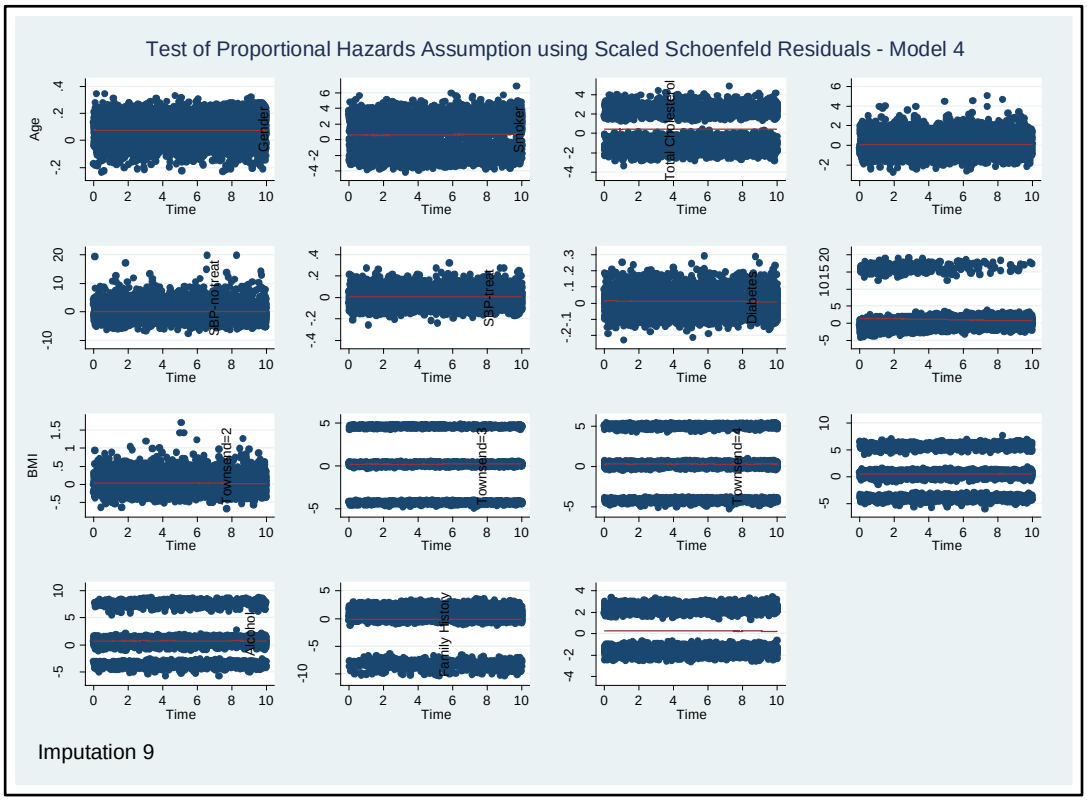
Appendix Figure G 29. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 9.



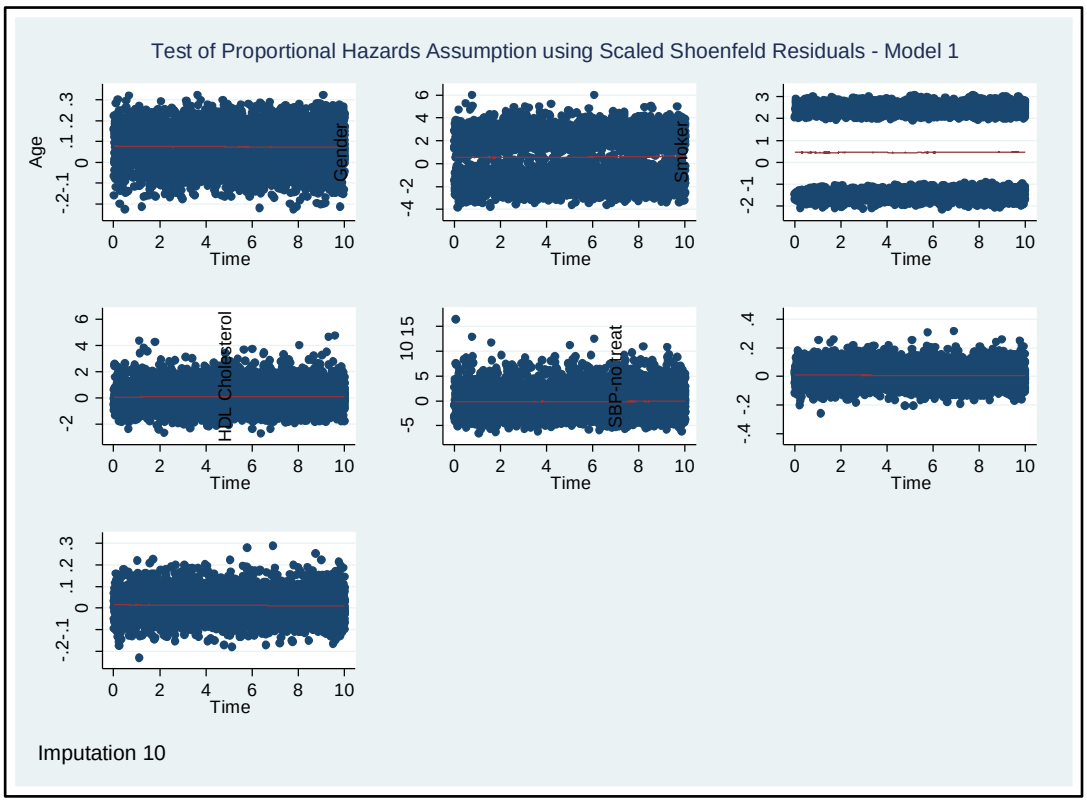
Appendix Figure G 30. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 9.



Appendix Figure G 31. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 9.



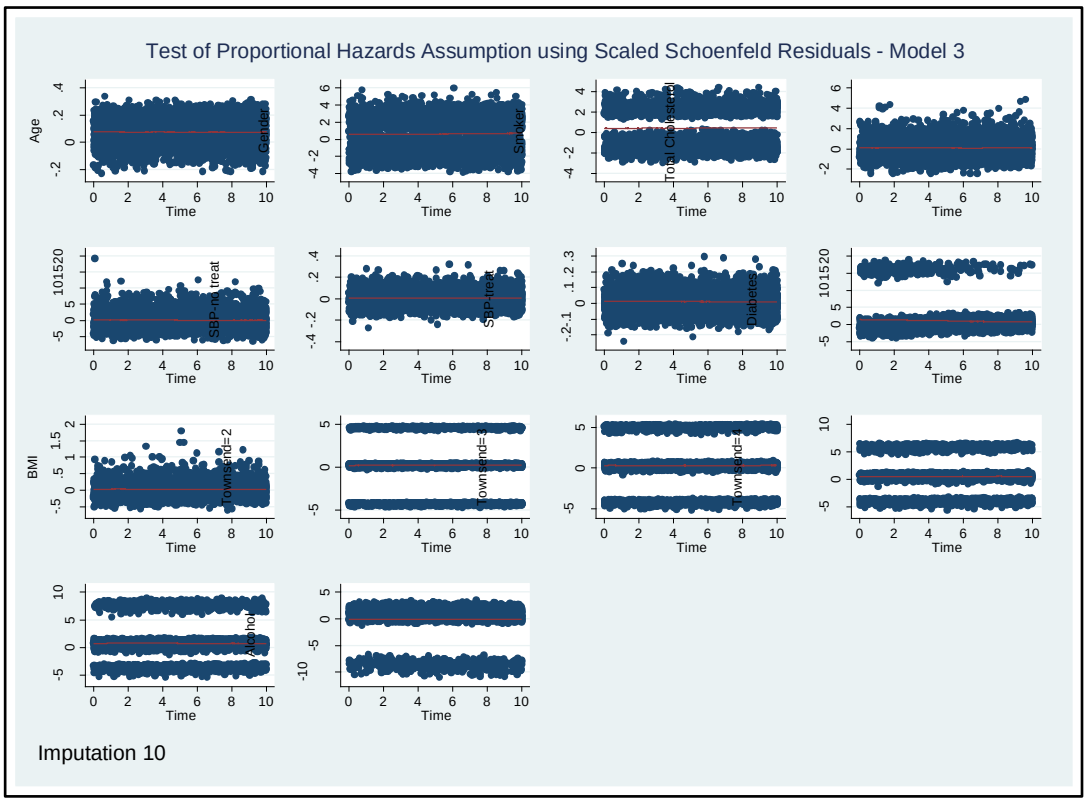
Appendix Figure G 32. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 9.



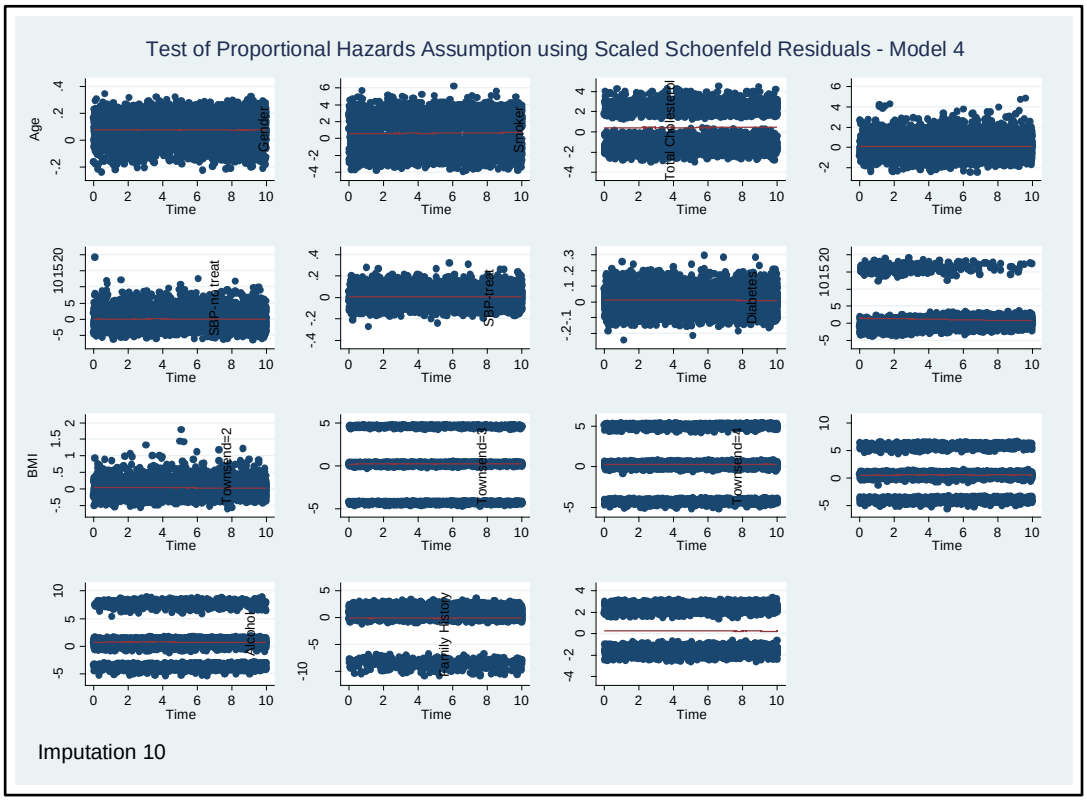
Appendix Figure G 33. Schoenfeld residual plots for Proportional Hazards, Model 1, Imputation 10.



Appendix Figure G 34. Schoenfeld residual plots for Proportional Hazards, Model 2, Imputation 10.



Appendix Figure G 35. Schoenfeld residual plots for Proportional Hazards, Model 3, Imputation 10.



Appendix Figure G 36. Schoenfeld residual plots for Proportional Hazards, Model 4, Imputation 10.

Appendix H. AIC, BIC, R^2, Likelihood Ratio values for each model and imputed dataset.

Appendix Table H 1. AIC, BIC, R^2, Likelihood Ratio values for each model and imputed dataset.

	Model	AIC	BIC	R^2	LR	df	ΔLR	p (df)
Imputation 1	Model 1	88601.98	88673.51	66.75%	5679.2	7		
							62.53	<0.0001 (1)
	Model 2	88541.44	88623.2	67.21%	5741.73	8		
	Model 3	88113.48	88256.55	70.31%	6181.69	14		
							60.56	<0.0001 (1)
	Model 4	88054.92	88208.21	70.72%	6242.25	15		
Imputation 2	Model 1	88600.45	88671.98	66.76%	5680.73	7		
							62.09	<0.0001 (1)
	Model 2	88540.36	88622.11	67.22%	5742.82	8		
	Model 3	88090.89	88233.96	70.46%	6204.28	14		
							59.74	<0.0001 (1)
	Model 4	88033.15	88186.45	70.86%	6264.02	15		
Imputation 3	Model 1	88600.78	88672.32	66.76%	5680.39	7		
							63.12	<0.0001 (1)
	Model 2	88539.66	88621.42	67.23%	5743.51	8		
	Model 3	88103.14	88246.21	70.38%	6192.03	14		
							59.44	<0.0001 (1)
	Model 4	88045.7	88199	70.78%	6251.47	15		
Imputation 4	Model 1	88606.97	88678.51	66.71%	5674.2	7		
							62.45	<0.0001 (1)
	Model 2	88546.52	88628.28	67.18%	5736.65	8		
	Model 3	88078.37	88221.44	70.55%	6216.8	14		
							61.38	<0.0001 (1)
	Model 4	88018.99	88172.28	70.95%	6278.18	15		
Imputation 5	Model 1	88592.2	88663.74	66.82%	5688.97	7		
							61.82	<0.0001 (1)
	Model 2	88532.38	88614.14	67.28%	5750.79	8		
	Model 3	88080.86	88223.93	70.53%	6214.31	14		
							59.59	<0.0001 (1)
	Model 4	88023.27	88176.57	70.93%	6273.9	15		
Imputation 6	Model 1	88507.87	88579.41	67.44%	5773.3	7		
							64.93	<0.0001 (1)
	Model 2	88444.94	88526.7	67.92%	5838.23	8		
	Model 3	87945.32	88088.4	71.42%	6349.85	14		
							60.35	<0.0001 (1)
	Model 4	87886.97	88040.26	71.82%	6410.2	15		

[Table continues on the next page]

	Model cont	AIC	BIC	R^2	LR	df	ΔLR	p (df)
Imputation 7	Model 1	88559.99	88631.53	67.06%	5721.18	7	59.84	<0.0001 (1)
	Model 2	88502.15	88583.91	67.50%	5781.02	8		
	Model 3	88117.24	88260.31	70.28%	6177.93	14	56.74	<0.0001 (1)
	Model 4	88062.5	88215.79	70.67%	6234.67	15		
Imputation 8	Model 1	88602.15	88673.69	66.75%	5679.02	7	62.92	<0.0001 (1)
	Model 2	88541.23	88622.98	67.22%	5741.94	8		
	Model 3	88093.1	88236.17	70.45%	6202.07	14	60.05	<0.0001 (1)
	Model 4	88035.05	88188.34	70.85%	6262.12	15		
Imputation 9	Model 1	88592.91	88664.45	66.82%	5688.26	7	61.14	<0.0001 (1)
	Model 2	88533.77	88615.52	67.27%	5749.4	8		
	Model 3	88116.09	88259.17	70.29%	6179.08	14	58.84	<0.0001 (1)
	Model 4	88059.25	88212.54	70.69%	6237.92	15		
Imputation 10	Model 1	88580.59	88652.13	66.91%	5700.58	7	60.76	<0.0001 (1)
	Model 2	88521.83	88603.59	67.36%	5761.34	8		
	Model 3	88114.47	88257.54	70.30%	6180.7	14	57.6	<0.0001 (1)
	Model 4	88058.87	88212.16	70.69%	6238.3	15		
Mean	Model 1	88584.59	88656.13	66.88%	5696.583	7	62.16	<0.0001 (1)
	Model 2	88524.43	88606.19	67.34%	5758.743	8		
	Model 3	88085.3	88228.37	70.50%	6209.874	14	59.429	<0.0001 (1)
	Model 4	88027.87	88181.16	70.89%	6269.303	15		

Appendix I. FH and CVD Study: Complete Case Analysis

Appendix Table I 1. Unadjusted Complete Case Cox Regression

Risk Factor		Unadjusted Analysis					
		Number in analysis	Beta	SE	p value	HR	95% CI
Age		202714	0.08	0	<0.001	1.09	(1.09 to 1.09)
Gender	Female	202714	Ref				
	Male		0.49	0.03	<0.001	1.32	(1.53 to 1.73)
Smoker	No	199391	Ref				
	Yes		0.28	0.03	<0.001	1.32	(1.24 to 1.41)
Total Cholesterol		28159	0.25	0.03	<0.001	1.28	(1.21 to 1.36)
HDL Cholesterol		9678	-0.17	0.14	0.24	0.84	(0.64 to 1.12)
SBP	Not treated	161024	0.03	0	<0.001	1.03	(1.03 to 1.03)
SBP	Treated	161024	0.04	0	<0.001	1.04	(1.03 to 1.04)
Family History	Negative	202714	Ref				
	Positive		0.3	0.03	<0.001	1.34	(1.26 to 1.43)
Ethnicity	White	103405	Ref				
	S. Asian		-0.39	0.19	0.04	0.68	(0.47 to 0.98)
	Black		-0.21	0.27	0.43	0.81	(0.48 to 1.37)
	Other/Mixed		-0.43	0.05	<0.001	0.65	(0.59 to 0.72)
Exercise	Inactive	84105	Ref				
	Active		-0.36	0.07	<0.001	0.7	(0.61 to 0.81)
Exercise, detailed	Inactive	80996	Ref				
	Gentle		-0.07	0.08	0.41	0.94	(0.8 to 1.09)
	Moderate		-0.5	0.08	<0.001	0.61	(0.52 to 0.71)
	Vigorous		-0.95	0.13	<0.001	0.39	(0.3 to 0.5)

[Table continues on the next page]

Risk Factor cont		Unadjusted Analysis					
		Number in analysis	Beta	SE	p value	HR	95% CI
Townsend Deprivation Score	1	126800	Ref				
	2		0.23	0.05	<0.001	1.26	(1.14 to 1.38)
	3		0.34	0.05	<0.001	1.4	(1.27 to 1.54)
	4		0.61	0.05	<0.001	1.84	(1.66 to 2.03)
	5		0.83	0.06	<0.001	2.28	(2.05 to 2.55)
Consumes Alcohol	No		Ref				
	Yes		-0.21	0.07	0	0.81	(0.7 to 0.93)
Alcohol Consumption, units per week		79820	0.01	0	<0.001	1.01	(1.01 to 1.01)
Diabetes	No	202714	Ref				
	Yes		2.11	0.06	<0.001	8.28	(7.3 to 9.39)
Impaired Glucose Intolerance	No	202714	Ref				
	Yes		1.19	0.41	0	3.28	(1.47 to 7.31)
Left Ventricular Hypertrophy	No	202714	Ref				
	Yes		1.01	0.5	0.04	2.74	(1.03 to 7.32)
Random Blood Glucose		1300	0.15	0.03	<0.001	1.16	(1.1 to 1.23)
Fasting Blood Glucose		1428	0.19	0.04	<0.001	1.2	(1.11 to 1.3)
Triglyceride Value		13056	0.15	0.04	<0.001	1.16	(1.07 to 1.27)
C-Reactive Protein		1971	0.01	0	0.05	1.01	(1 to 1.02)
BMI		126048	0.06	0	<0.001	1.06	(1.05 to 1.06)
ALT		11027	0	0	0.71	1	(0.99 to 1)
AST		9479	0.01	0	<0.001	1.01	(1.01 to 1.02)
GGT		8469	0	0	<0.001	1	(1 to 1)

Appendix Table I 2. Adjusted Complete Case Cox Regression

Model	Risk Factor		Adjusted Analysis					
			Number in analysis	Beta	SE	p value	HR	95% CI
Model 1	Age (years)		8775	0.06	0.01	<0.001	1.06	(1.04 to 1.07)
	Gender	Female				Ref		
		Male	8775	0.63	0.13	<0.001	1.88	(1.45 to 2.45)
	Smoker	No				Ref		
		Yes	8775	0.26	0.12	0.0333	1.3	(1.02 to 1.66)
	Total Cholesterol		8775	0.2	0.06	0.0012	1.22	(1.08 to 1.37)
	HDL Cholesterol		8775	-0.2	0.17	0.2294	0.82	(0.59 to 1.14)
	SBP	Not treated	8775	0.01	0	0.2185	1.01	(1 to 1.01)
		Treated	8775	0.01	0	0.0063	1.01	(1 to 1.02)
Model 2	Age (years)		8775	0.05	0.01	<0.001	1.06	(1.04 to 1.07)
	Gender	Female				Ref		
		Male	8775	0.62	0.13	<0.001	1.86	(1.43 to 2.42)
	Smoker	No				Ref		
		Yes	8775	0.27	0.12	0.0287	1.31	(1.03 to 1.67)
	Total Cholesterol		8775	0.19	0.06	0.0012	1.22	(1.08 to 1.37)
	HDL Cholesterol		8775	-0.2	0.17	0.2283	0.82	(0.59 to 1.13)
	SBP	Not treated	8775	0.01	0	0.2197	1.01	(1 to 1.01)
		Treated	8775	0.01	0	0.0062	1.01	(1 to 1.02)
	Family History	Negative				Ref		
		Positive	8775	-0.13	0.12	0.3044	0.88	(0.69 to 1.12)

[Table continues on the next page]

Model cont.	Risk Factor		Adjusted Analysis					
			Number in analysis	Beta	SE	p value	HR	95% CI
Model 3	Age (years)		4621	0.05	0.01	<0.001	1.05	(1.04 to 1.07)
	Gender	Female				Ref		
		Male	4621	0.7	0.15	<0.001	2.02	(1.5 to 2.72)
	Smoker	No				Ref		
		Yes	4621	0.16	0.14	0.2573	1.17	(0.89 to 1.54)
	Total Cholesterol		4621	0.2	0.07	0.0025	1.23	(1.07 to 1.4)
	HDL Cholesterol		4621	0.25	0.19	0.1748	1.29	(0.89 to 1.86)
	SBP	Not treated	4621	0	0	0.759	1	(0.99 to 1.01)
		Treated	4621	0.01	0	0.0795	1.01	(1 to 1.02)
	Diabetes	No				Ref		
		Yes	4621	1.23	0.17	<0.001	3.41	(2.43 to 4.77)
	BMI		4621	0.06	0.01	<0.001	1.06	(1.03 to 1.09)
	Townsend Deprivation Score	1				Ref		
		2	4621	0.05	0.21	0.8217	1.05	(0.69 to 1.59)
		3	4621	0.34	0.2	0.0943	1.4	(0.94 to 2.08)
		4	4621	0.4	0.21	0.0499	1.5	(1 to 2.24)
		5	4621	0.53	0.24	0.0297	1.69	(1.05 to 2.72)

[Table continues on the next page]

Model cont.	Risk Factor		Adjusted Analysis					
			Number in analysis	Beta	SE	p value	HR	95% CI
Model 4	Age (years)		4621	0.05	0.01	<0.001	1.05	(1.04 to 1.07)
	Gender	Female				Ref		
		Male	4621	0.69	0.15	<0.001	2	(1.49 to 2.69)
	Smoker	No				Ref		
		Yes	4621	0.17	0.14	0.2315	1.18	(0.9 to 1.56)
	Total Cholesterol		4621	0.2	0.07	0.0027	1.22	(1.07 to 1.4)
	HDL Cholesterol		4621	0.26	0.19	0.17	1.29	(0.9 to 1.87)
	SBP	Not treated	4621	0	0	0.7601	1	(0.99 to 1.01)
		Treated	4621	0.01	0	0.079	1.01	(1 to 1.02)
	Diabetes	No				Ref		
		Yes	4621	1.21	0.17	<0.001	3.36	(2.39 to 4.71)
	BMI		4621	0.06	0.01	<0.001	1.06	(1.03 to 1.09)
	Townsend Deprivation Score	1				Ref		
		2	4621	0.05	0.21	0.8206	1.05	(0.69 to 1.59)
		3	4621	0.33	0.2	0.0991	1.4	(0.94 to 2.07)
		4	4621	0.4	0.21	0.0518	1.49	(1 to 2.23)
		5	4621	0.52	0.24	0.0312	1.68	(1.05 to 2.7)
	Family History	No				Ref		
		Yes	4621	-0.12	0.14	0.4119	0.89	(0.68 to 1.17)

Appendix Table I 3. AIC, BIC R^2, and Likelihood Ratio tests, comparing models with FH (Model2 2 and 4) to models without FH (Models 1 and 3).

Comparison	Models	AIC	BIC	R^2 (%)	LR	df	ΔLR	p-value (df)
Model 1 vs Model 2	Model 1	4499.903	4549.46	44.89	226.97	7	1.06	0.3045 (1)
	Model 2	4500.848	4557.486	45.06	228.03	8		
Model 3 vs Model 4	Model 3	3236.762	3320.461	62.32	279.87	13	0.67	0.4118 (1)
	Model 4	3238.088	3328.225	62.42	280.54	14		

Appendix Table I 4. Reclassification Table for Model 1 and Model 2.

			with FH			
			<10%	≥10%	Total	Total No.(%) Reclassified
without FH	<10%	No	3,585	22	3,607	22 (0.61%)
		%	99.39%	0.61%		
		No. Events	89	0		
		No. Non events	3,496	22		
		Observed Event Rate	2.48%	0.00%		
	≥10%	No	66	696	762	66 (8.66%)
		%	8.66%	91.34%		
		No. Events	3	57		
		No. Non events	63	639		
		Observed Event Rate	4.55%	8.19%		
	Total		3,651	718	4,369	88 (2.01%)

Appendix Table I 5. Risk Reclassification for Model 1 and Model 2, by CVD outcome.

No CVD		**with FH**		
without FH		<10%	≥10%	Total
	<10%	3,496 99.37%	22 0.63%	3,518
	≥10%	63 8.97%	639 91.03%	702
	Total	3,559	661	4,220
CVD		**with FH**		
without FH		<10%	≥10%	Total
	<10%	89 100.00%	0 0.00%	89
	≥10%	3 5.00%	57 95.00%	60
	Total	92	57	149

Appendix Table I 6. Reclassification Table for Model 3 and Model 4.

			with FH			
			<10%	≥10%	Total	Total No.(%) Reclassified
without FH	<10%	No	1,746	7	1,753	7 (0.4%)
		%	99.60%	0.40%		
		No. Events	61	0		
		No. Non events	1,685	7		
		Observed Event Rate	3.49%	0.00%		
	≥10%	No	29	535	564	29 (5.14%)
		%	5.14%	94.86%		
		No. Events	3	53		
		No. Non events	26	482		
		Observed Event Rate	10.34%	9.91%		
	Total		1,775	542	2,317	36 (1.55%)

Appendix Table I 7. Risk Reclassification for Model 3 and Model 4, by CVD outcome.

No CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	1,685 99.59%	7 0.41%	1,692
	≥10%	26 5.12%	482 94.88%	508
	Total	1,711	489	2,200
CVD		**with FH**		
		<10%	≥10%	Total
without FH	<10%	61 100.00%	0 0.00%	61
	≥10%	3 5.36%	53 94.64%	56
	Total	64	53	117

Appendix J. Literature based Survival

Appendix Table J 1. Reclassification Table for Model 1 and Model 2.

			with FH			
			<10%	≥10%	Total	Total No.(%) Reclassified
without FH	<10%	No	73,647	1,533	75,180	1533 (2.04%)
		%	97.96%	2.04%		
		No. Events	588	48		
		No. Non events	73,059	1,485		
		Observed Event Rate	0.80%	3.13%		
	≥10%	No	1215	23,426	24,641	1215 (4.93%)
		%	4.93%	95.07%		
		No. Events	30	1,287		
		No. Non events	1185	22,139		
		Observed Event Rate	2.47%	5.49%		
	Total		74,862	24,959	99,821	2748 (2.75%)

Appendix Table J 2. Risk Reclassification for Model 1 and Model 2, by CVD outcome.

no CVD		**with FH**		
without FH		<10%	≥10%	Total
	<10%	73,059 98.01%	1,485 1.99%	74,544
	≥10%	1185 5.08%	22,139 94.92%	23,324
	Total	74,244	23,624	97,868
CVD		**with FH**		
without FH		<10%	≥10%	Total
	<10%	588 92.45%	48 7.55%	636
	≥10%	30 2.28%	1,287 97.72%	1,317
	Total	618	1,335	1,953

Appendix Table J 3. Reclassification Table for Model 3 and Model 4.

			with FH			
			<10%	≥10%	Total	Total No.(%) Reclassified
without FH	<10%	No	73,387	1,521	74,908	1521 (2.03%)
		%	97.97%	2.03%		
		No. Events	570	33		
		No. Non events	72,817	1,488		
		Observed Event Rate	0.78%	2.17%		
	≥10%	No	1149	23,764	24,913	1149 (4.61%)
		%	4.61%	95.39%		
		No. Events	28	1,322		
		No. Non events	1121	22,442		
		Observed Event Rate	2.44%	5.56%		
	Total		74,536	25,285	99,821	2670 (2.67%)

Appendix Table J 4. Risk Reclassification for Model 3 and Model 4, by CVD outcome.

no CVD		**with FH**		
without FH		<10%	≥10%	Total
	<10%	72,817 98.00%	1,488 2.00%	74,305
	≥10%	1121 4.76%	22,442 95.24%	23,563
	Total	73,938	23,930	97,868
CVD		with FH		
without FH		<10%	≥10%	Total
	<10%	570 94.53%	33 5.47%	603
	≥10%	28 2.07%	1,322 97.93%	1,350
	Total	598	1,355	1,953

Appendix K. Availability of bone mineral density in routinely collected data

Recording of Bone Mineral Density in General Practice Databases

Study Design:

Prospective Cohort Study

Data Source:

Data for the study was extracted from the Clinical Practice Research Datalink (CPRD), formerly known as the General Practice Research Database (GPRD). This database is a large, validated primary care research database with information on over 12 million patients and 650 practices. It collects information of patient demographics, clinical events, referrals, tests, prescriptions and deaths. It also includes lifestyle data such as BMI and also self-reported lifestyle information such as smoking status and alcohol consumption.

Study Population:

The baseline date for the study is 1st Jan 2004, which will allow a maximum 10 year follow up period to 1st Dec 2014. In line with the age range given with the FRAX risk score, patients must be aged between 40-90 years and will be excluded if they are on any osteoporotic treatment or taking any bone sparing drugs prior to baseline. Further, patients must be registered with their practice for at least a year, removing any temporary patients, and also have complete data on the 'core' risk factors (age, gender, height, weight, current smoking status, and alcohol consumption) within 12 months prior to baseline.

Patients were also excluded from the study if they exited from their practice the same time as or before the baseline study date. Patients were followed up for 10 years.

Outcome:

The outcome for this study is a fragility fracture or fracture at an osteoporotic site (clinical spine, hip, forearm, pelvis and proximal humerus) within the period of interest, 1st Jan 2004 and 31st Dec 2013. A comprehensive Read Code list was developed, based on QOF business rules for fragility fractures. This was added to, by also searching the Read Codes using the methodology described above. Keywords such as "fracture", "hip", "femur", "spine" and "forearm" were used to identify the relevant fractures. The developed Read Code list was then mapped to related ICD – 10 codes, which were used to identify the outcome in the linked HES and ONS data.

Using the code lists, patients were identified to be outcome positive if they had a fracture code in their primary care medical records (clinical file), linked secondary care HES records, or in their ONS records. This will be during the 10 year follow up period (1st January 2004 to 31st December 2013).

To identify a fragility fracture, ideally both the fracture and mechanism need to be knows, as they are specifically non traumatic events. As both pieces of information are not explicitly found in CPRD records, it was assumed that events osteoporotic sites within the age range for the study were fragility fractures.

For the statistical analysis, outcome data from each of these sources (clinical, HES and ONS) were aggregated into a single binary variable indicating the presence or absence of a fracture at an osteoporotic site, using the date of the first record within the follow up period.

Core Risk Factors:

Age

Using the patient date of birth, their age was calculated at the baseline date of the study, given in years. Age was treated as a continuous variable, in the analysis.

Gender

Gender was specified for each patient; they were classed as either male or female.

Body Mass Index (BMI)

Height, weight and BMI were derived using the additional clinical file using entity codes. If any one of these measurements were missing, the other two would be used to calculate the third measurement.

Smoking status

Smoking data was used to create three smoking related variables; smoking status, smoking amount, and level of smoking. Smoking status was given as an ordinal categorical variable, identifying patients to be a current smoker, ex-smoker or non-smoker. Smoking amount was presented as a continuous variable, giving the number of cigarettes smoked per day. Both variables were then combined to give the level of smoking; heavy smokers (>20/day); moderate smokers (10-19/day); light smokers (<10/day); ex-smokers; and non-smokers.

Smoking information was found in both the clinical and additional clinical details files. Smoking data in each of these files were then combined to create a single smoking status and smoking amount variable.

Discrepancies in the data between files were first dealt with by taking data from the file with the most information. For example, if a patient classed a smoker with information regarding the amount smoked in the clinical file, was classed as an ex-smoker in the additional clinical file; they would be classed as a smoker for the study.

The worst case scenario was then assumed for remaining patients with inconsistent smoking data. For example, if a patient was classed as a smoker in the clinical file and ex-smoker in the additional clinical file; they would be classed a smoker in the study

Finally, for all other discrepancies, data from the additional clinical details file was given precedence over data from the clinical file.

If smoking status was not found in either file, but smoking amount was given and was more than 0, the patient was assumed to be a smoker.

Further, codes indicating health promotion of smoking or if smoking advice was given, assumed the patient was a smoker. Passive smokers with no indication of smoking were assumed to be non-smokers. Data was also checked for ex-smokers and if they had not smoked for 10 or more years, they were classed as non-smokers for the risk model.

Alcohol Consumption

Alcohol consumption was derived using similar methodology when deriving smoking status for each patient. Each patient was classed as a non/past drinker or drinker based on their recorded drinking status and the number of units per week.

Additional Risk Factor

Bone Mineral Density

Bone Mineral Density was identified using a Read code list. Codes were identified using recognised primary care database methodology (160). This code list included codes identifying osteoporosis through bone mineral density score; however, results from the Dual energy X-ray –absorptiometry (DXA) scan were not provided.

Codes were included if the DXA scan was taken at the hip/femur, spine, heel, and forearm. Patients were identified is being normal, osteopenic or osteoporotic based on a DXA scan. No details were provided of the criteria used to diagnose patients based on their DXA scan.

Statistical Analysis:

A descriptive analysis was performed, reporting number (%), mean (SD) and median (IQR) for categorical, normal continuous and non-normal continuous variables, respectively. Missing values will also be presented. This analysis will also investigate the statistical significance of the association of additional factors with the outcome. Appropriate statistics tests such as $\chi 2$ tests for binary variables, t-tests to compare means of two normal continuous variables, and analysis of variance tests (ANOVA) to compare means of three or more groups, will be used to assess the differences between the groups of interest.

Results:

A total of 60,658 patients were extracted from the CPRD database. Table 1 provides a description of the baseline characteristics of the data. Patients had a mean age of 56 years (SD=13.08) and consisted of more females (52.17%) than males (47.83%).

Approximately 7% (n=4,142) of patients had a fracture during the 10 year follow up period.

Of the 60,658 extracted patients, BMD was recorded in 3 patients (<0.01%). Thus no further analysis could be performed.

Appendix Table K 1. Baseline Characteristics

Characteristics		CPRD (n=60,658)	
		Mean	SD
Age (years)		56.09	13.08
Height (m)		1.68	0.10
Weight (kg)		74.89	16.32
BMI		26.32	4.96
		No	%
Gender	Female	31,646	52.17
	Male	29,012	47.83
Previous Fracture	No	54,504	89.85
	Yes	6,154	10.15
Parental History of Hip Fracture	No	60,658	100.00
	Yes	0	0.00
Current Smoking Status	No	20,490	33.78
	Yes	31,842	52.49
Alcohols Consumption greater than 3 units per day	No	44,851	73.94
	Yes	15,807	26.06
Glucocorticoid Use	No	59,019	97.30
	Yes	1,639	2.70
Rheumatoid Arthritis	No	60,151	99.16
	Yes	507	0.84
Secondary Osteoporosis	No	56,678	93.44
	Yes	3,980	6.56

Appendix L. Associations between fracture risk factors

Appendix Table L 1. Mean Age, by risk factor

Mean Age	Group 1	Group 2	p value
Gender (females, males)	62.3 (10.8)	63.95 (10.86)	<0.001
Previous Fracture (no, yes)	61.28 (10.66)	66.18 (10.45)	<0.001
Parental History of Hip Fracture (no, yes)	64.23 (11.22)	59.37 (9.15)	<0.001
Current Smoker (no, yes)	63.19 (11.01)	60.65 (9.91)	<0.001
Glucocorticoid Use (no, yes)	62.27 (10.7)	65.09 (11.38)	<0.001
Alcohol Consumption (>3 units/day) (no, yes)	62.57 (10.89)	63.53 (9.49)	0.215
Rheumatoid Arthritis (no, yes)	62.73 (10.83)	61.71 (10.79)	0.077
Secondary Osteoporosis (no, yes)	62.95 (10.84)	60.92 (10.62)	<0.001
Osteoporotic (no, yes)	61.83 (10.54)	70.3 (10.53)	<0.001

Appendix Table L 2. Mean BMI, by risk factor

Mean BMI	Group 1	Group 2	p value
Gender (females, males)	26.27 (5.24)	26.91 (4.11)	0.002
Previous Fracture (no, yes)	26.4 (5.06)	26.37 (5.01)	0.839
Parental History of Hip Fracture (no, yes)	26.5 (5.08)	26.17 (4.97)	0.052
Current Smoker (no, yes)	26.74 (5.02)	25.18 (4.95)	<0.001
Glucocorticoid Use (no, yes)	26.31 (5.05)	26.96 (4.97)	0.008
Alcohol Consumption (>3 units/day) (no, yes)	26.44 (5.07)	25.48 (4.4)	0.008
Rheumatoid Arthritis (no, yes)	26.32 (5.04)	27.02 (5.07)	0.009
Secondary Osteoporosis (no, yes)	26.41 (4.97)	26.32 (5.4)	0.684
Osteoporotic (no, yes)	26.76 (5.03)	22.89 (3.71)	<0.001

Appendix Table L 3 Proportion female, by risk factor

Proportion female	Group 1	Group 2	p value
Previous Fracture (no, yes)	79.44	83.2	0.008
Parental History of Hip Fracture (no, yes)	76.86	87.78	<0.001
Current Smoker (no, yes)	80.36	80.84	0.754
Glucocorticoid Use (no, yes)	83.24	60.98	<0.001
Alcohol Consumption (>3 units/day) (no, yes)	81.8	56.31	<0.001
Rheumatoid Arthritis (no, yes)	81.22	73.74	<0.001
Secondary Osteoporosis (no, yes)	77.97	93.55	<0.001
Osteoporotic (no, yes)	79.48	90.05	<0.001

Appendix Table L 4. Proportion with previous fracture, by risk factor

Proportion with Previous Fracture	Group 1	Group 2	p value
Gender (females, males)	21.16	14.1	0.008
Parental History of Hip Fracture (no, yes)	21.48	16.35	<0.001
Current Smoker (no, yes)	19.96	19.16	0.588
Glucocorticoid Use (no, yes)	20.6	14.02	<0.001
Alcohol Consumption (>3 units/day) (no, yes)	19.8	19.42	0.009
Rheumatoid Arthritis (no, yes)	20.41	14.14	<0.001
Secondary Osteoporosis (no, yes)	19.21	22.8	0.208
Osteoporotic (no, yes)	17.99	37.1	<0.001

Appendix Table L 5. Proportion with parental hip fracture, by risk factor

Proportion with Parental History of Hip Fracture	Group 1	Group 2	p value
Gender (females, males)	36.08	20.7	<0.001
Previous Fracture (no, yes)	34.49	27.33	<0.001
Current Smoker (no, yes)	33.22	32.54	0.703
Glucocorticoid Use (no, yes)	35.86	13.41	<0.001
Alcohol Consumption (>3 units/day) (no, yes)	33.32	28.64	0.166
Rheumatoid Arthritis (no, yes)	34.9	16.67	<0.001
Secondary Osteoporosis (no, yes)	33.23	32.23	0.623
Osteoporotic (no, yes)	34.41	20.16	<0.001

Appendix Table L 6. Proportion smoker, by risk factor

Proportion Smoker	Group 1	Group 2	p value
Gender (females, males)	22.39	21.86	0.754
Previous Fracture (no, yes)	22.46	21.58	0.588
Parental History of Hip Fracture (no, yes)	22.46	21.93	0.703
Glucocorticoid Use (no, yes)	22.62	19.92	0.178
Alcohol Consumption (>3 units/day) (no, yes)	20.98	46.12	<0.001
Rheumatoid Arthritis (no, yes)	22.32	21.97	0.874
Secondary Osteoporosis (no, yes)	20.56	31.29	<0.001
Osteoporotic (no, yes)	21.47	30.11	<0.001

Appendix Table L 7. Proportion taking glucocorticoids, by risk factor

Proportion taking glucocorticoids	Group 1	Group 2	p value
Gender (females, males)	9.42	24.84	<0.001
Previous Fracture (no, yes)	13.32	8.81	<0.001
Parental History of Hip Fracture (no, yes)	16.08	5.04	<0.001
Current Smoker (no, yes)	12.81	11.11	0.178
Alcohol Consumption (>3 units/day) (no, yes)	12.23	16.02	0.110
Rheumatoid Arthritis (no, yes)	11.9	17.17	0.003
Secondary Osteoporosis (no, yes)	12.94	9.75	0.026
Osteoporotic (no, yes)	12.24	14.25	0.265

Appendix Table L 8 Proportion consuming more than 3 units of alcohol, by risk factor

Proportion consuming >3units alcohol per day	Group 1	Group 2	p value
Gender (females, males)	3.64	11.64	<0.001
Previous Fracture (no, yes)	5.23	5.11	0.009
Parental History of Hip Fracture (no, yes)	5.55	4.51	0.166
Current Smoker (no, yes)	3.61	10.77	<0.001
Glucocorticoid Use (no, yes)	4.99	6.71	0.110
Rheumatoid Arthritis (no, yes)	5.5	2.53	0.014
Secondary Osteoporosis (no, yes)	5.09	5.82	0.448
Osteoporotic (no, yes)	5.08	6.45	0.257

Appendix Table L 9. Proportion diagnosed with RA, by risk factor

Proportion diagnosed with RA	Group 1	Group 2	p value
Gender (females, males)	9.17	13.45	<0.001
Previous Fracture (no, yes)	10.71	7.15	<0.001
Parental History of Hip Fracture (no, yes)	12.46	5.04	<0.001
Current Smoker (no, yes)	10.05	9.86	0.874
Glucocorticoid Use (no, yes)	9.46	13.82	0.003
Alcohol Consumption (>3 units/day) (no, yes)	10.29	4.85	0.014
Secondary Osteoporosis (no, yes)	10.39	8.02	0.069
Osteoporotic (no, yes)	10.04	9.68	0.825

Appendix Table L 10. Proportion diagnosed with secondary osteoporosis, by risk factor

Proportion diagnosed with Secondary Osteoporosis	Group 1	Group 2	p value
Gender (females, males)	18.68	5.3	<0.001
Previous Fracture (no, yes)	15.46	18.52	0.208
Parental History of Hip Fracture (no, yes)	16.27	15.66	0.623
Current Smoker (no, yes)	14.21	22.56	<0.001
Glucocorticoid Use (no, yes)	16.56	12.6	0.026
Alcohol Consumption (>3 units/day) (no, yes)	15.96	17.96	0.448
Rheumatoid Arthritis (no, yes)	16.42	12.88	0.069
Osteoporotic (no, yes)	15.84	18.28	0.223

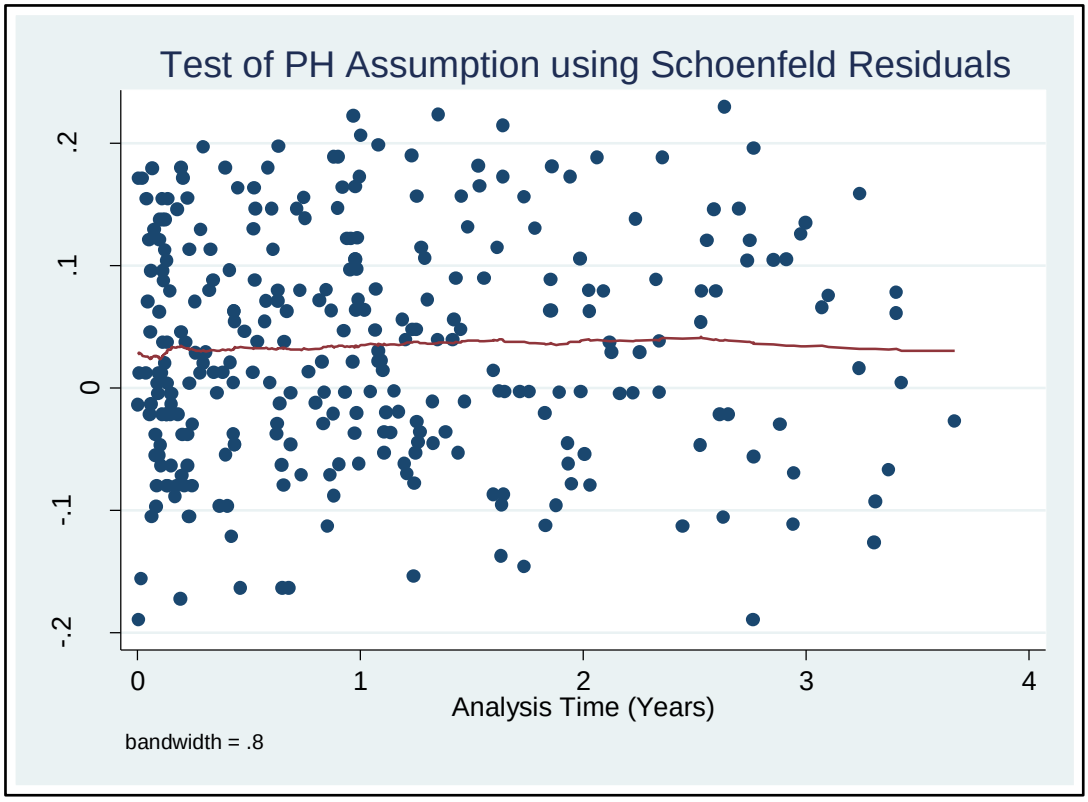
Appendix Table L 11. Proportion diagnosed with osteoporosis, by risk factor

Proportion Osteoporotic	Group 1	Group 2	p value
Gender (females, males)	10.52	4.79	<0.001
Previous Fracture (no, yes)	7.37	17.62	<0.001
Parental History of Hip Fracture (no, yes)	11.21	5.73	<0.001
Current Smoker (no, yes)	8.45	12.7	<0.001
Glucocorticoid Use (no, yes)	9.2	10.77	0.265
Alcohol Consumption (>3 units/day) (no, yes)	9.28	11.65	0.257
Rheumatoid Arthritis (no, yes)	9.43	9.09	0.825
Secondary Osteoporosis (no, yes)	9.15	10.69	0.223

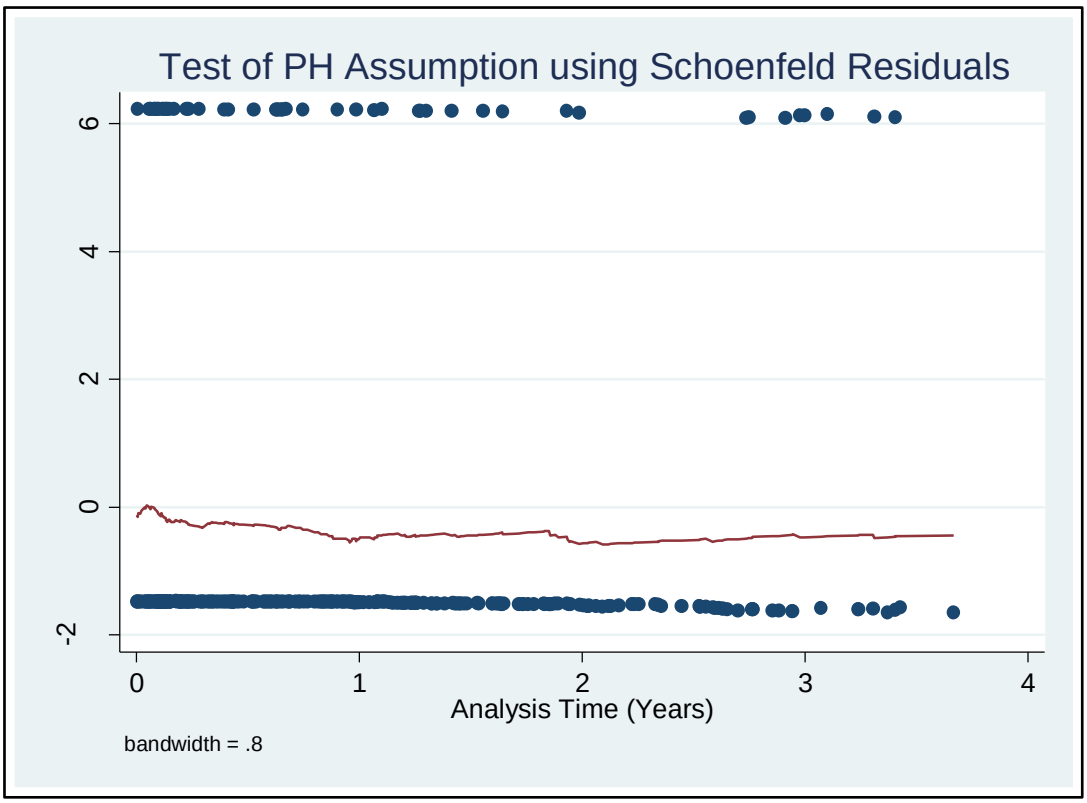
Appendix Table L 12. Mean BMD t-score, by risk factor

Mean Hip DXA t-score	Group 1	Group 2	p value
Gender (females, males)	-1.16 (1.1)	-0.95 (0.93)	<0.001
Previous Fracture (no, yes)	-0.96 (1.04)	-1.53 (1.05)	<0.001
Parental History of Hip Fracture (no, yes)	-1.16 (1.11)	-1.01 (1)	<0.001
Current Smoker (no, yes)	-1.05 (1.08)	-1.32 (1.04)	<0.001
Glucocorticoid Use (no, yes)	-1.11 (1.07)	-1.12 (1.09)	0.953
Alcohol Consumption (>3 units/day) (no, yes)	-1.11 (1.07)	-1.2 (1.09)	0.241
Rheumatoid Arthritis (no, yes)	-1.13 (1.07)	-0.96 (1.11)	0.003
Secondary Osteoporosis (no, yes)	-1.1 (1.07)	-1.19 (1.1)	0.044
Osteoporotic (no, yes)	-0.92 (0.92)	-3.01 (0.49)	<0.001

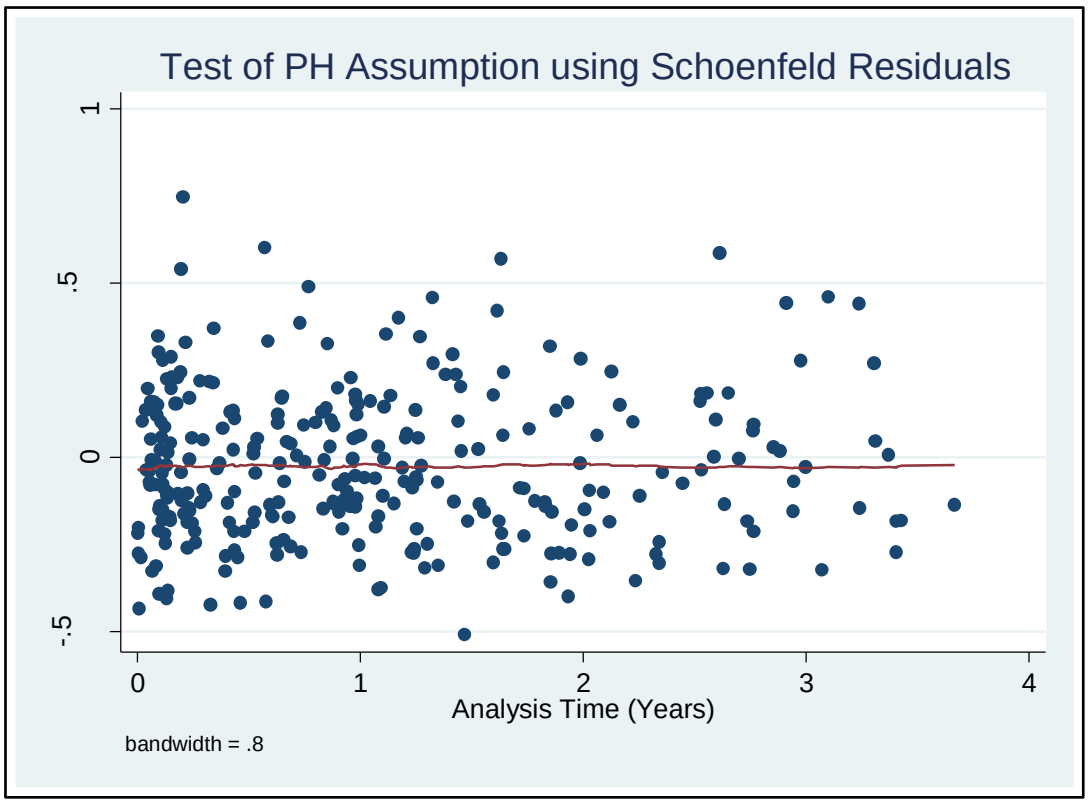
Appendix M. Schoenfeld Residual Plots to test for proportional hazards for each risk factor.



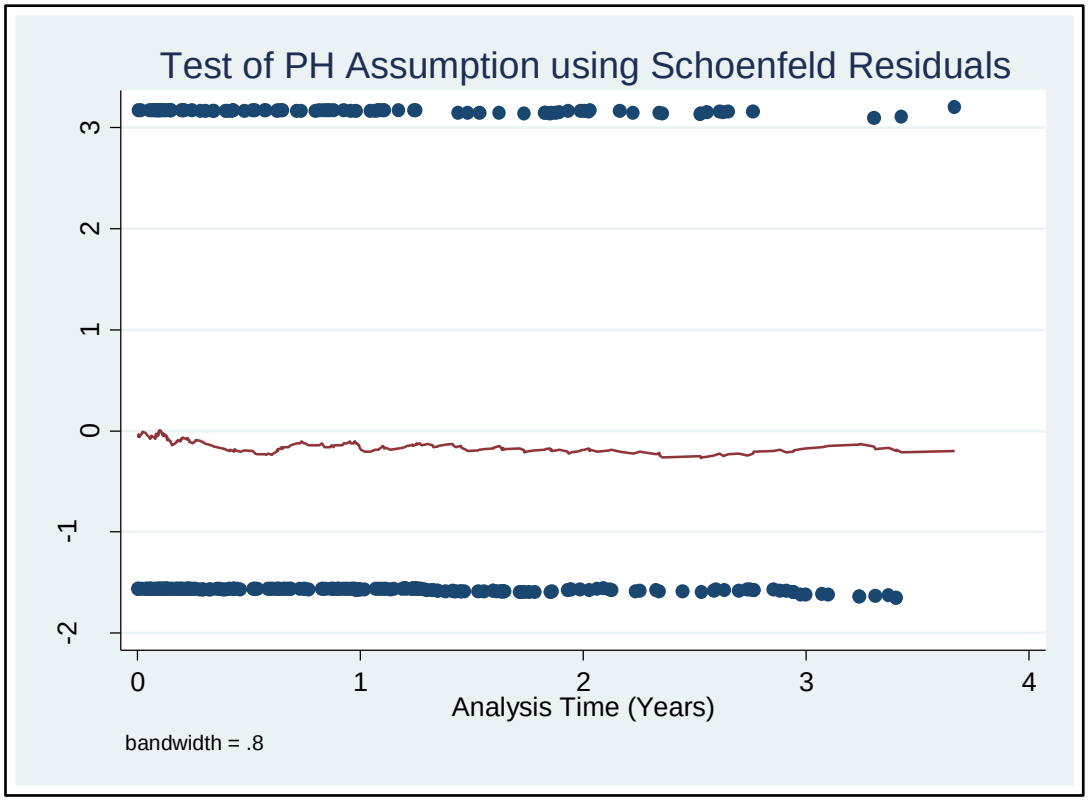
Appendix Figure M 1. Age



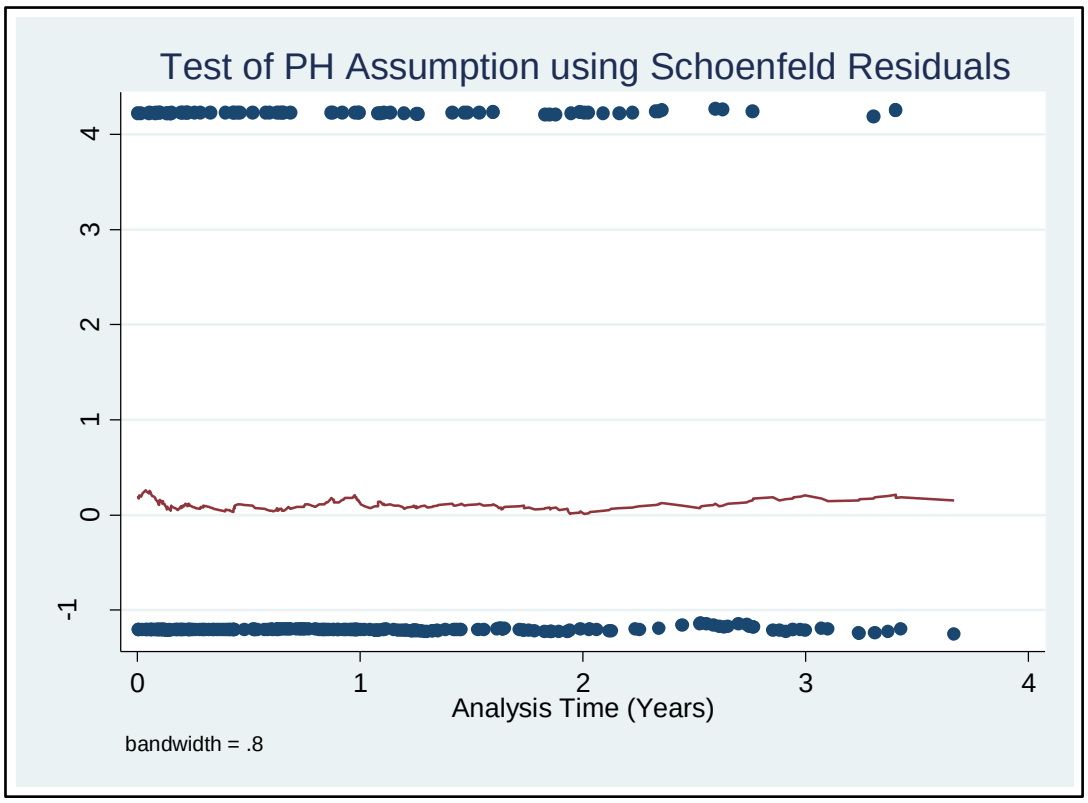
Appendix Figure M 2. Gender



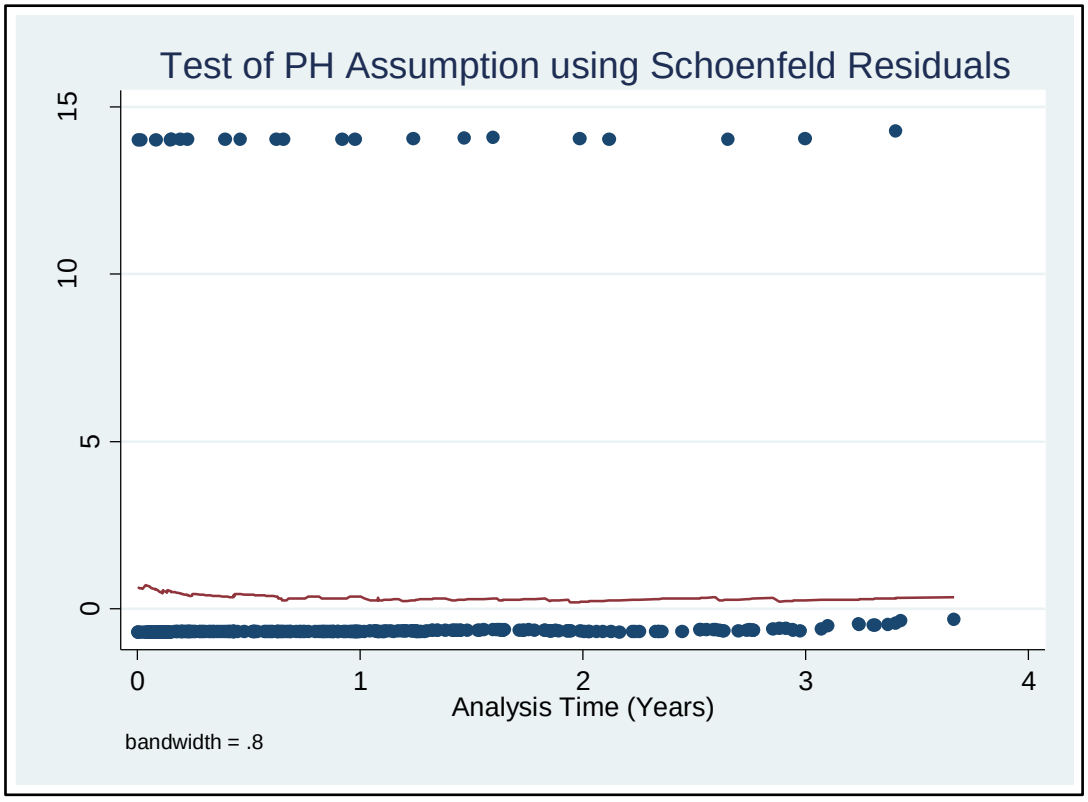
Appendix Figure M 3. BMI



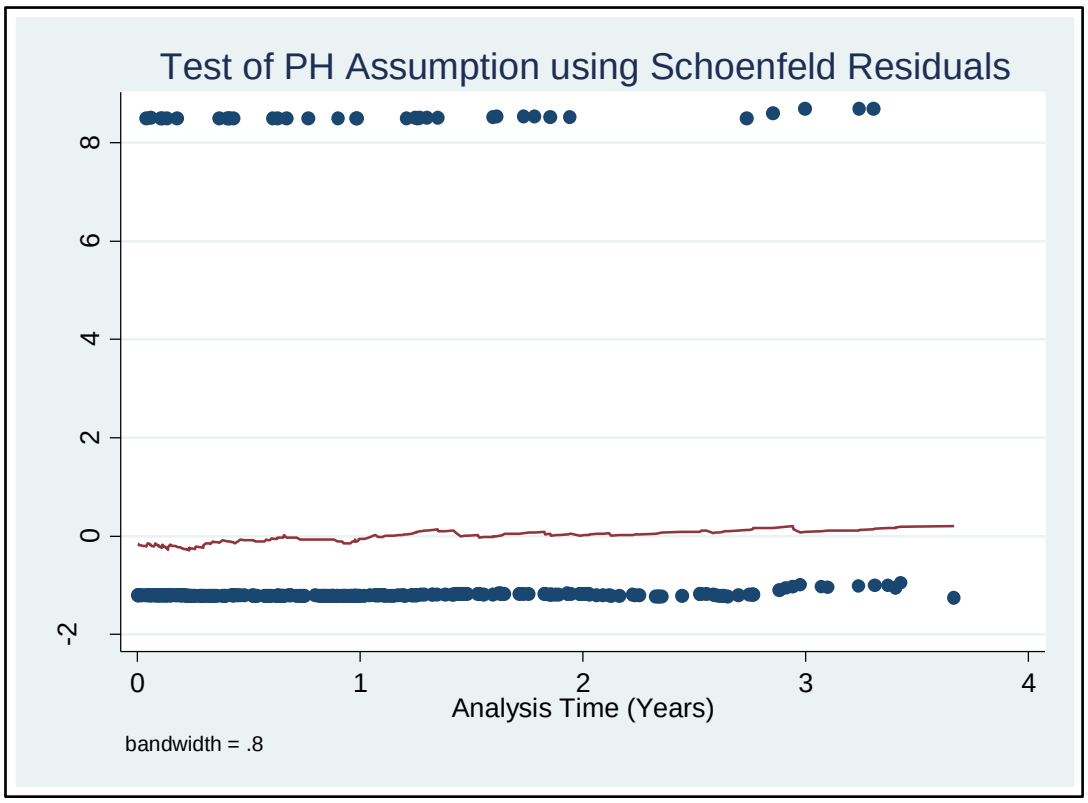
Appendix Figure M 4. Parental History of Hip Fracture



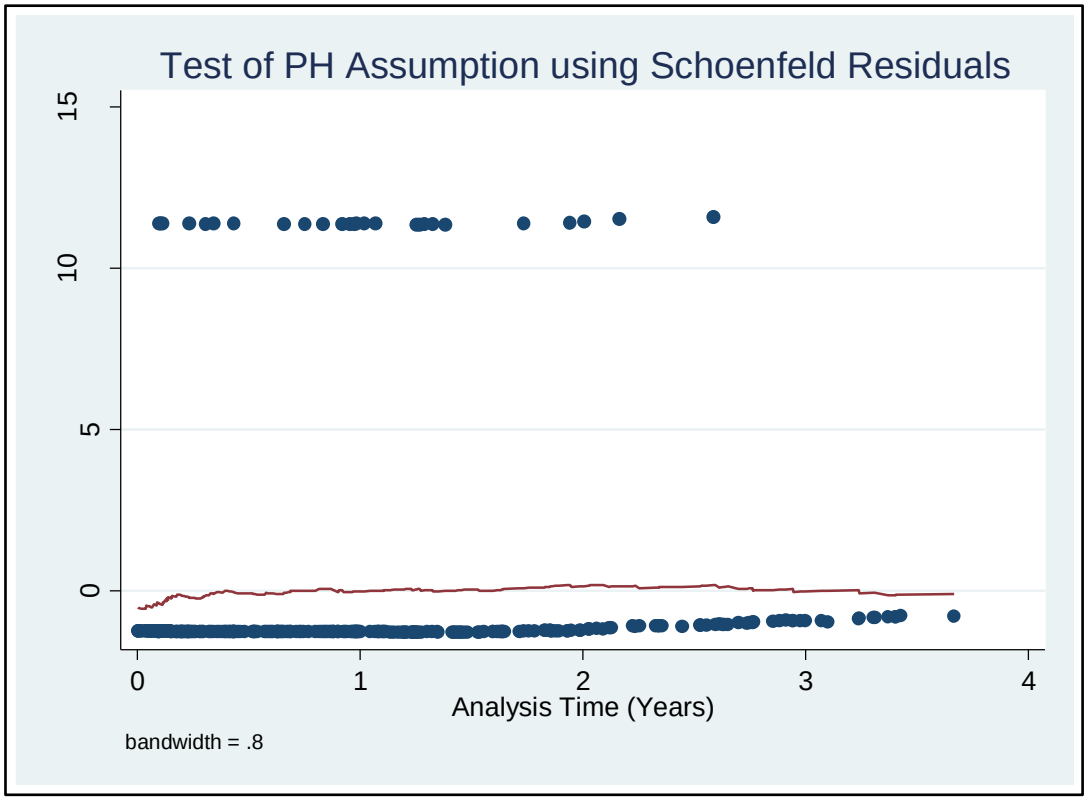
Appendix Figure M 5. Smoking Status



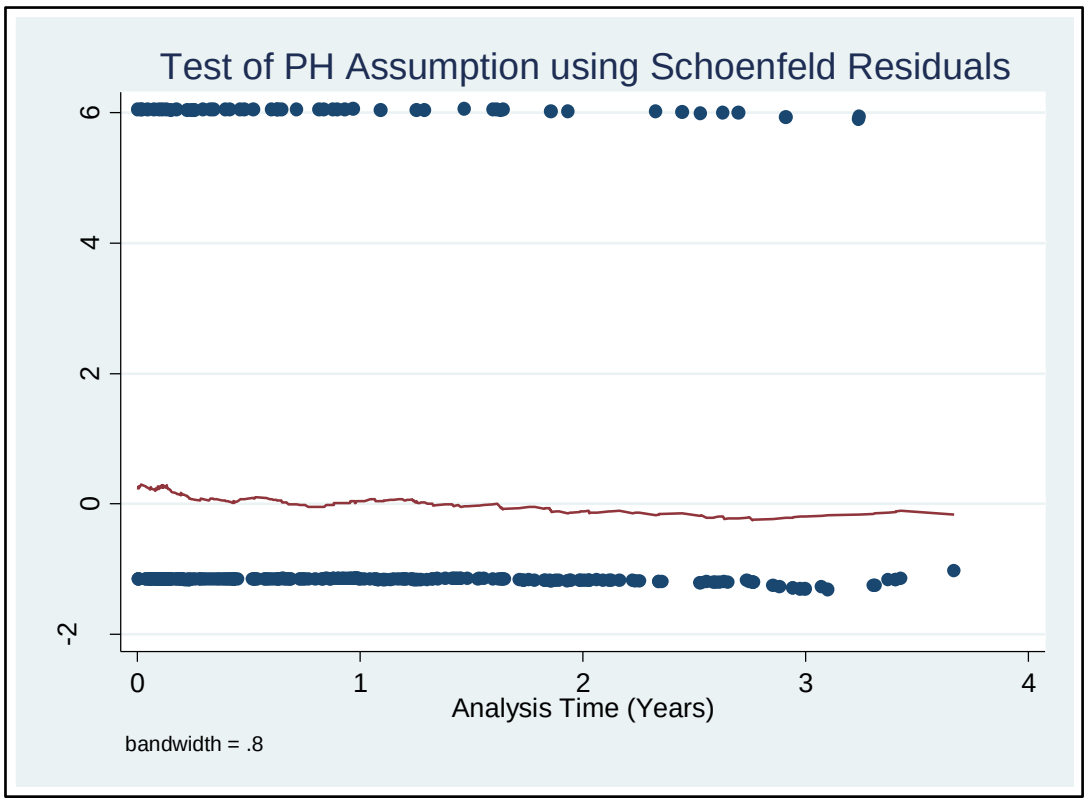
Appendix Figure M 6. Alcohol Consumption ≥3 units per day.



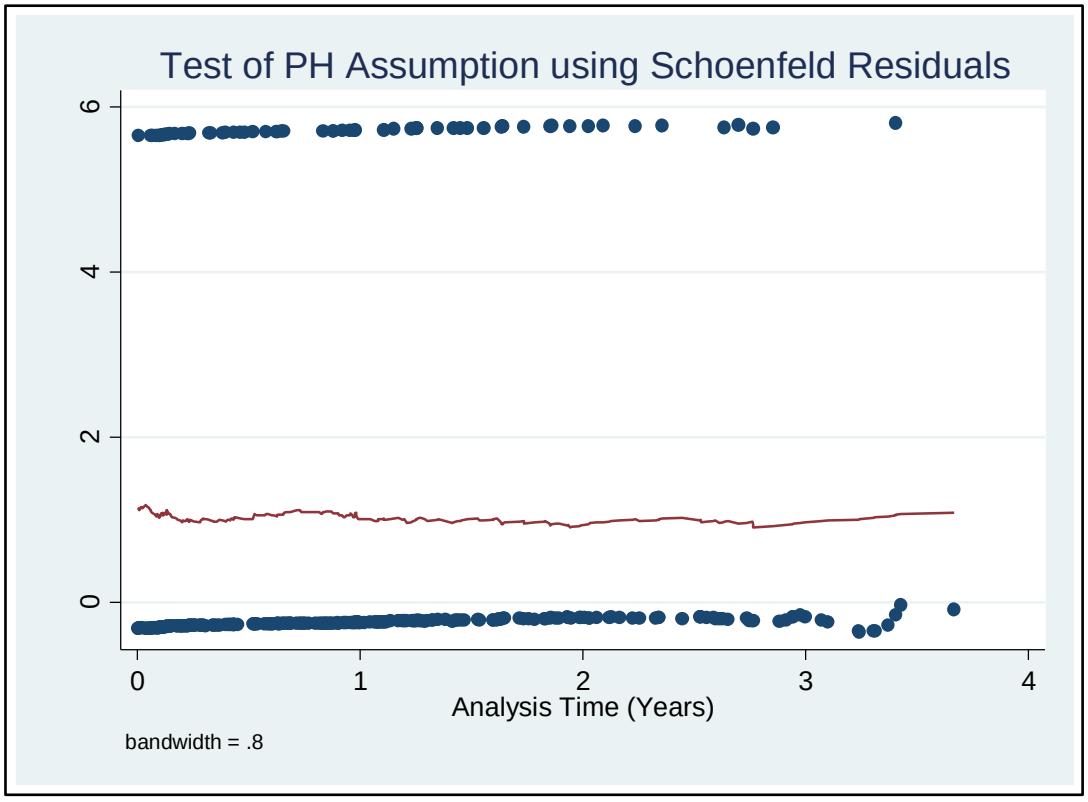
Appendix Figure M 7. Glucocorticoid Use ≥3 months



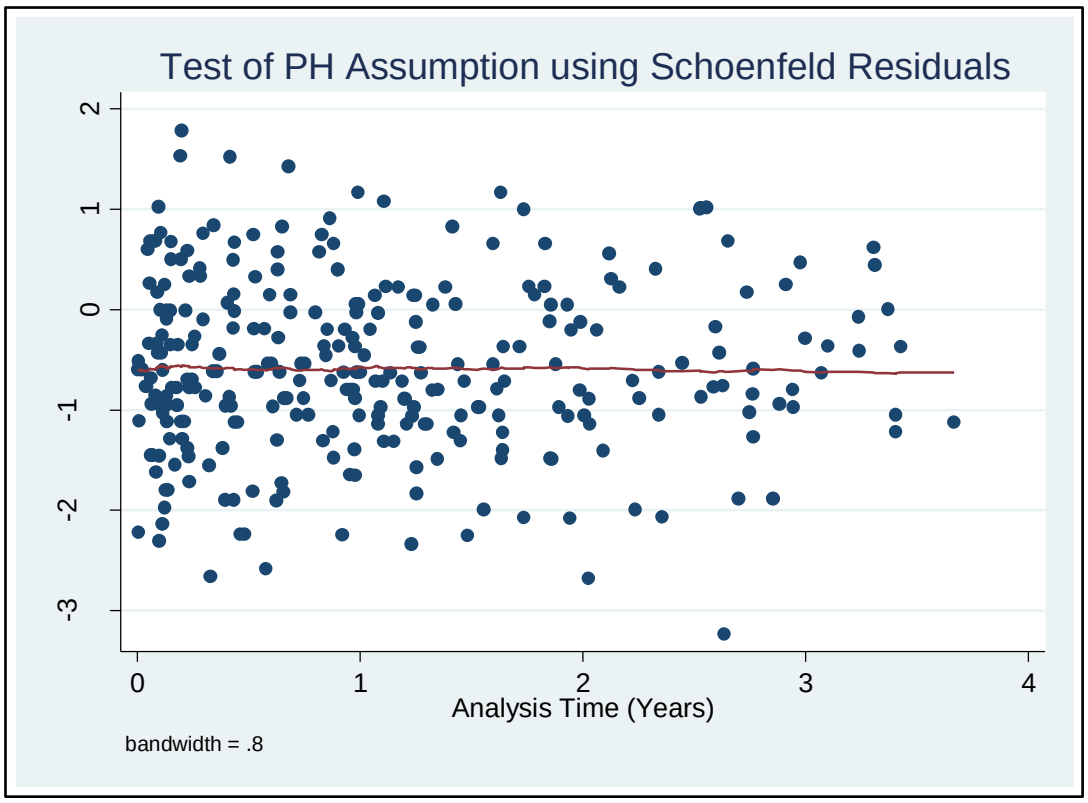
Appendix Figure M 8. Rheumatoid Arthritis



Appendix Figure M 9. Secondary Osteoporosis

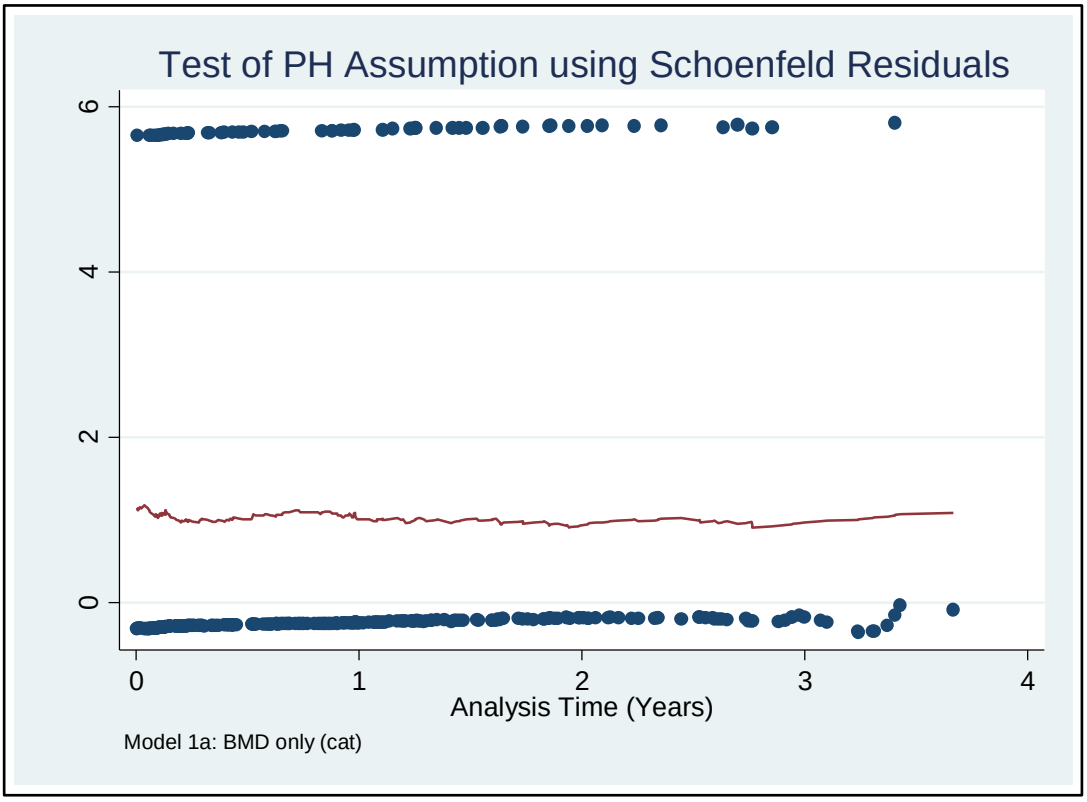


Appendix Figure M 10. Osteoporosis (t-score≤-2.5)

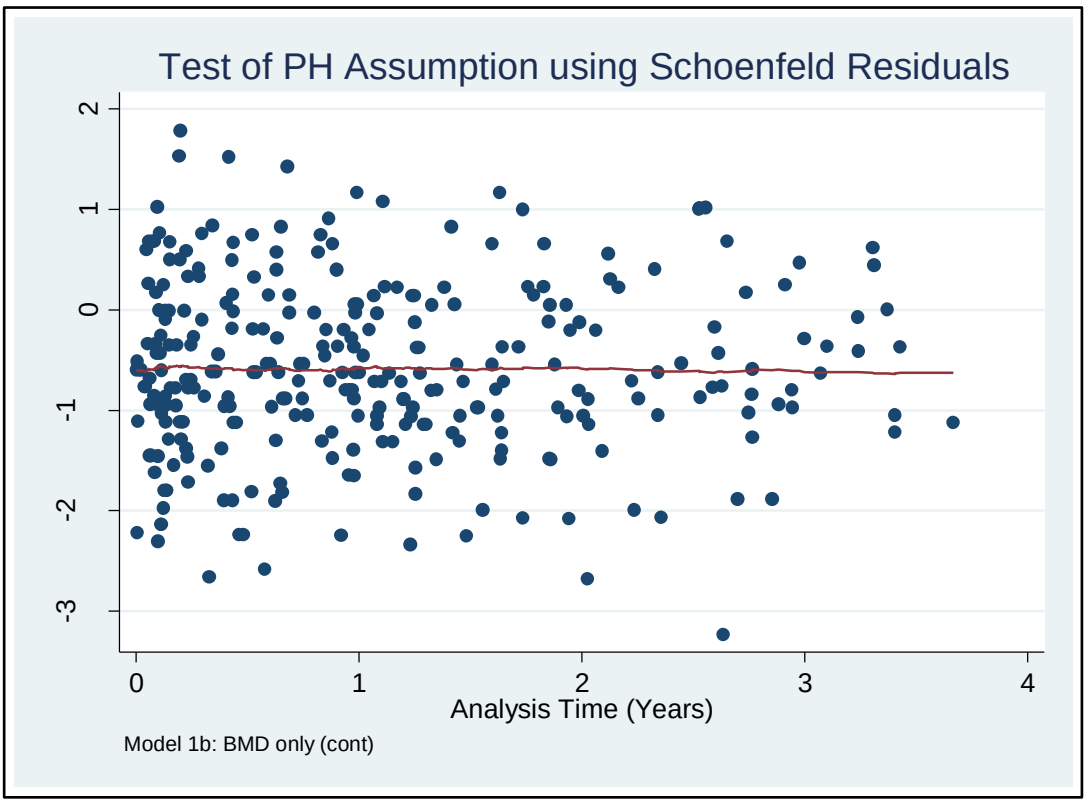


Appendix Figure M 11. BMD measurement T-score.

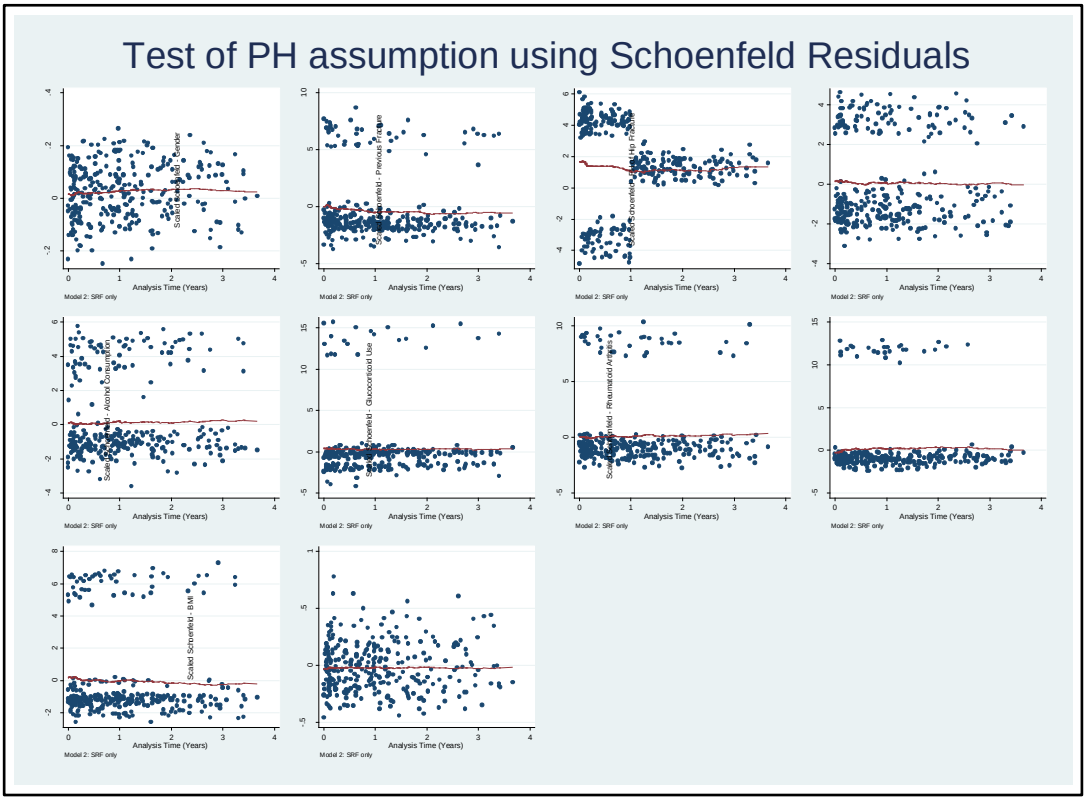
Appendix N. Schoenfeld Residual Plots to test for proportional hazards for each risk model.



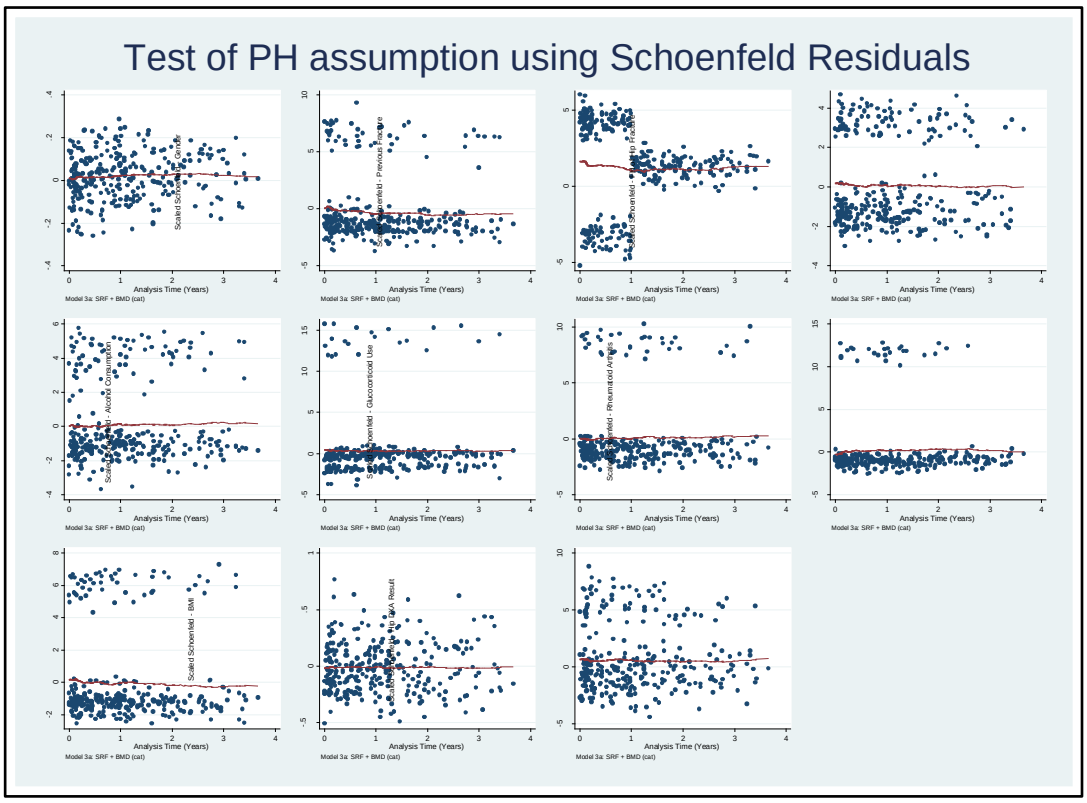
Appendix Figure N 1. Model 1a



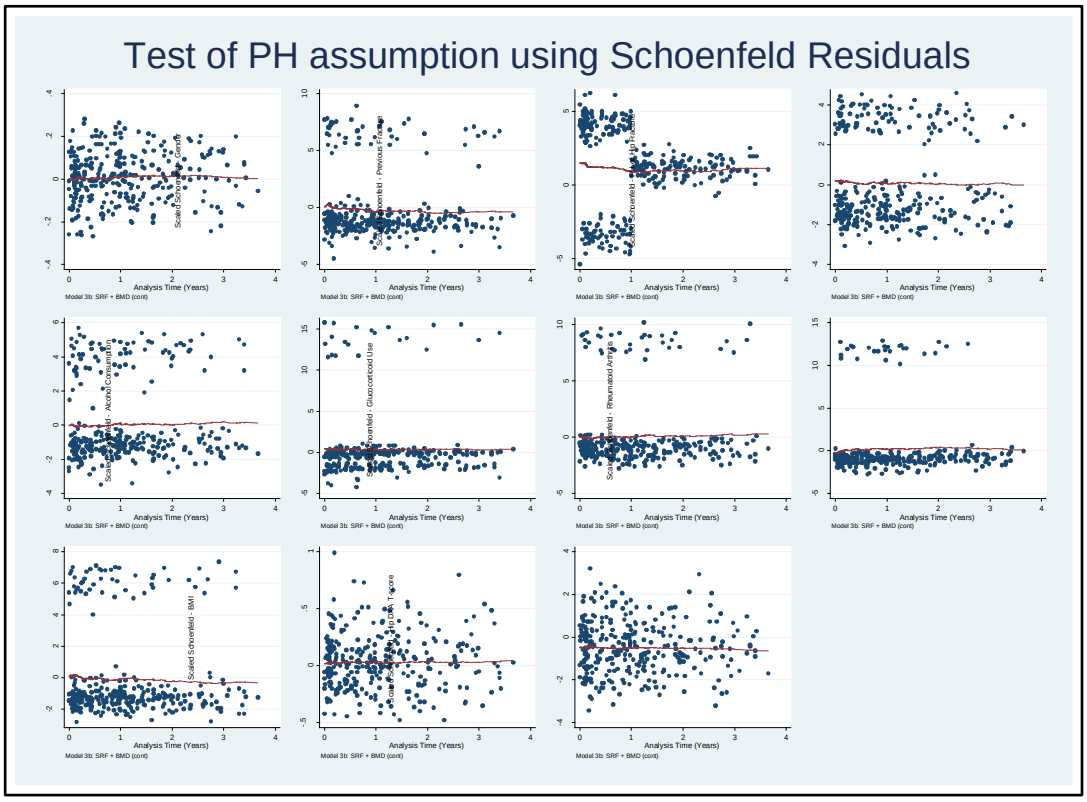
Appendix Figure N 2. Model 1b



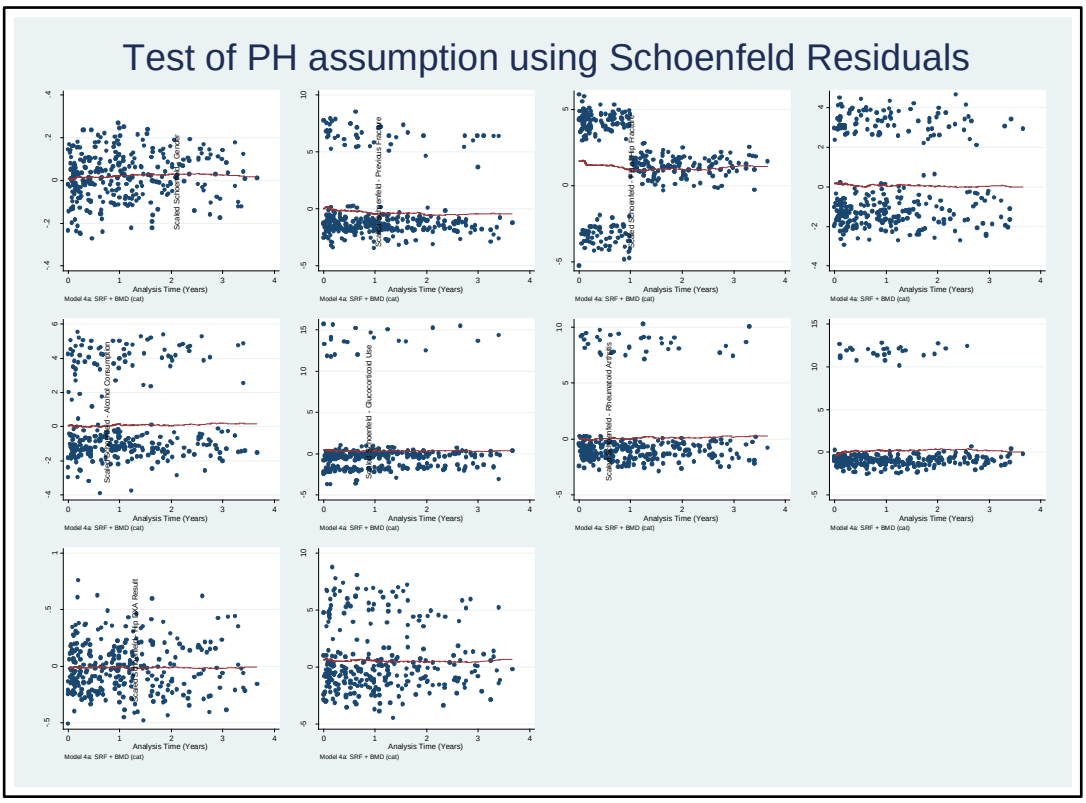
Appendix Figure N 3. Model 2



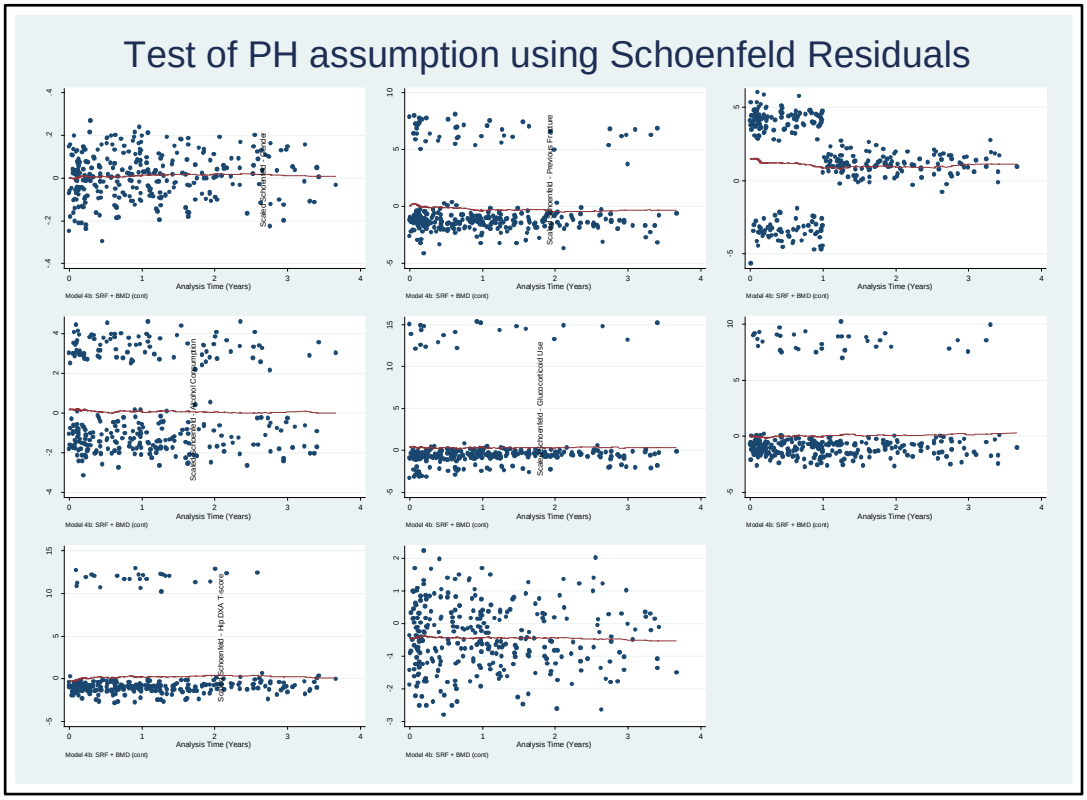
Appendix Figure N 4. Model 3a



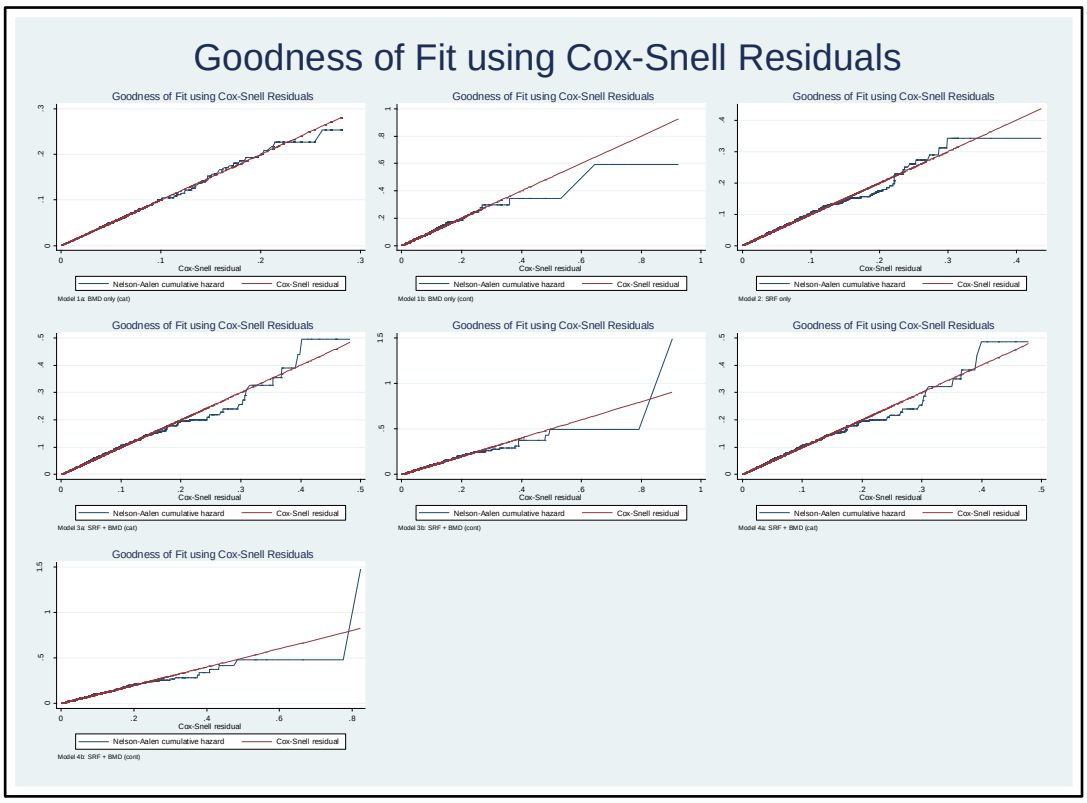
Appendix Figure N 5. Model 3b



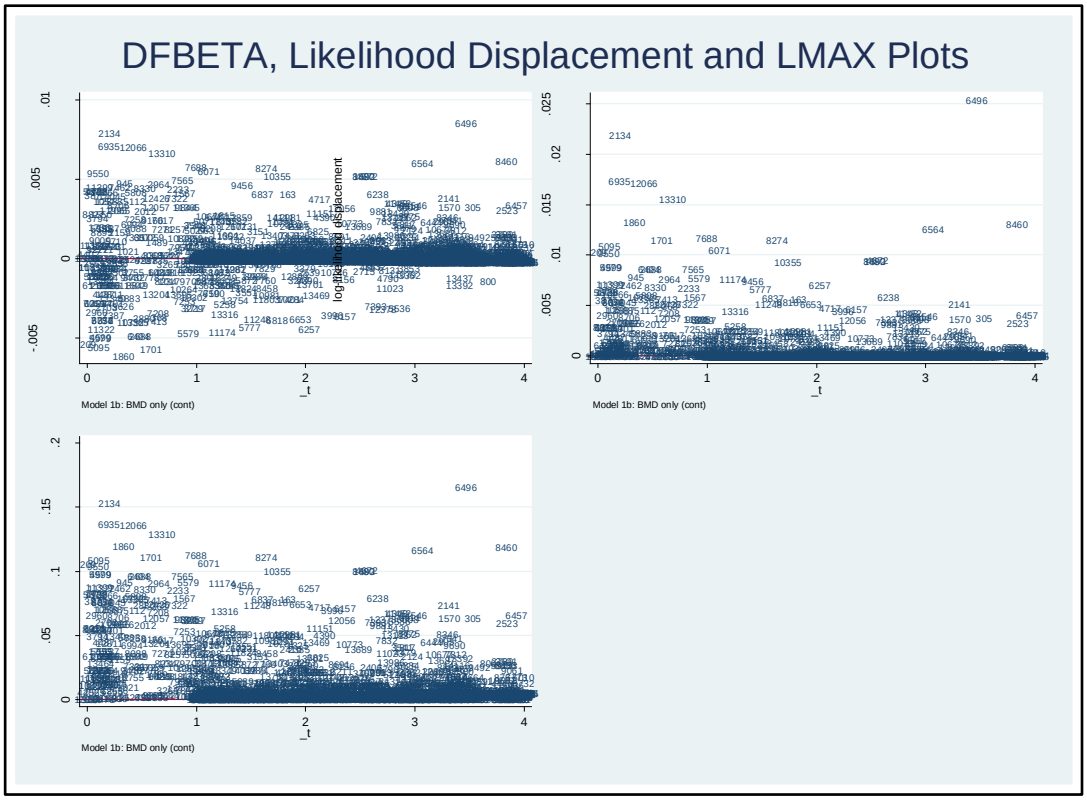
Appendix Figure N 6. Model 4a



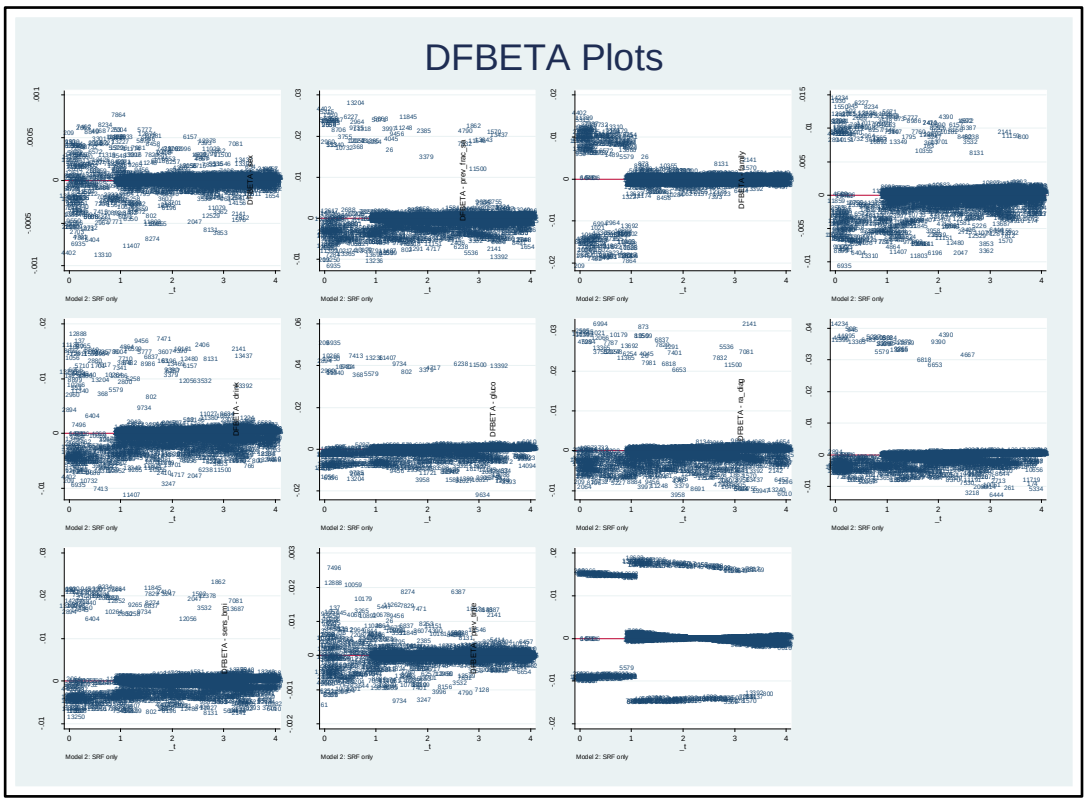
Appendix Figure N 7. Model 4b



Appendix Figure N 8. Cox-Snell Plots for Goodness of Fit, for all models

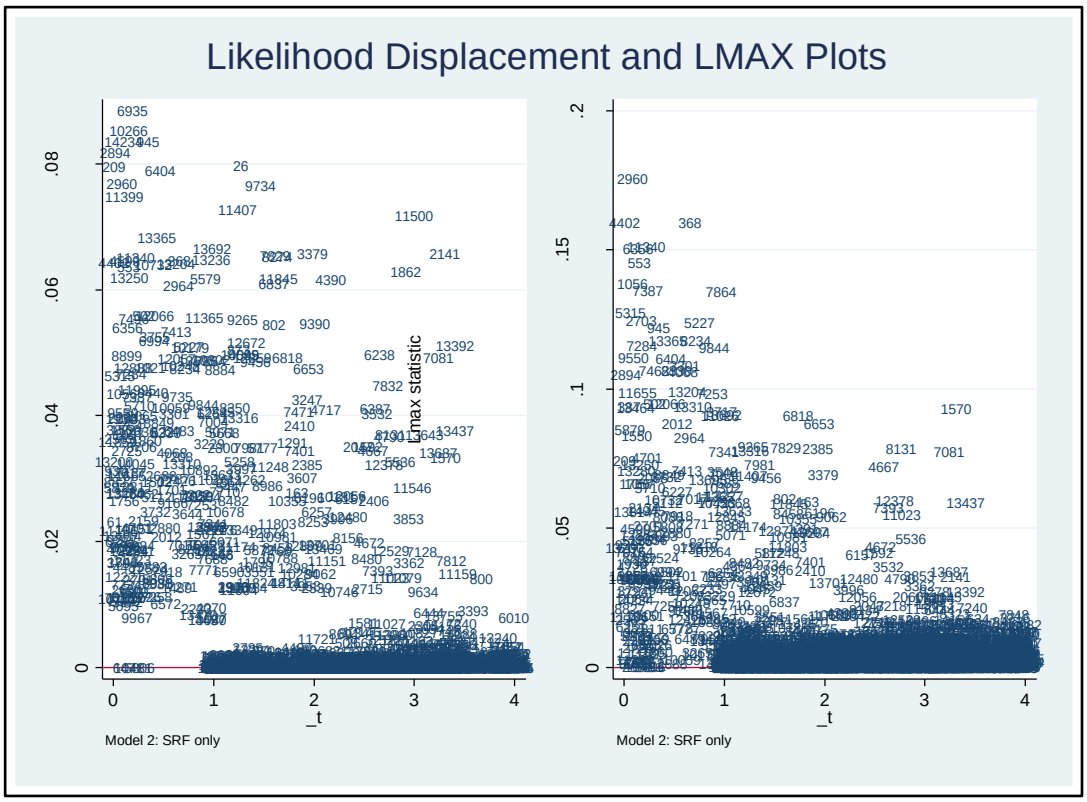


Appendix Figure N 9. DFBETA, L-displacement and LMAX Plots, Model 1b

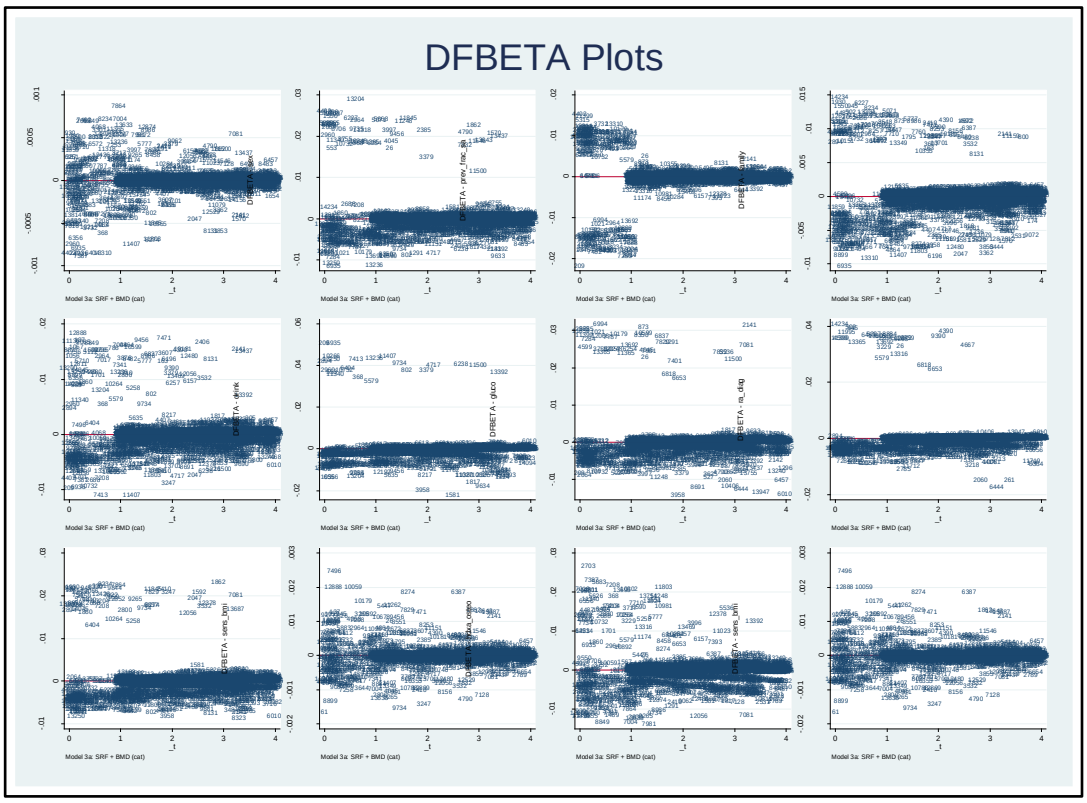


Appendix Figure N 10. DFBETA, Model 2

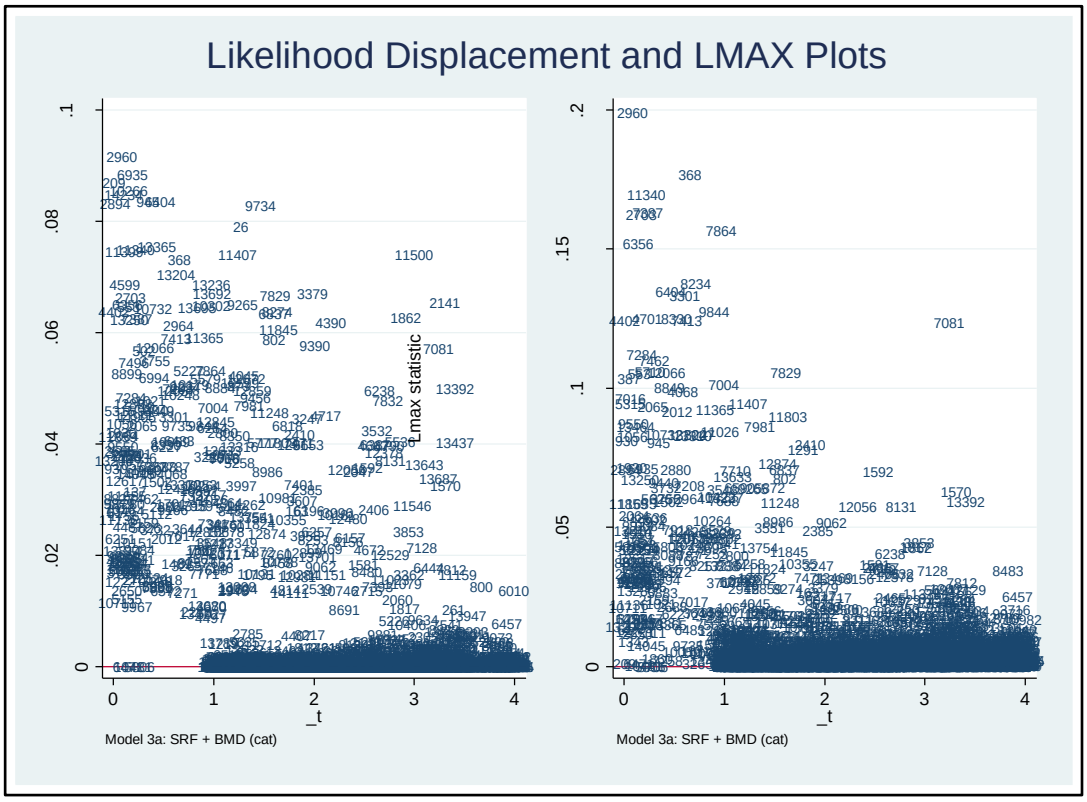
Appendix Figure N.11 L-displacement and LMAX Plots, Model 2



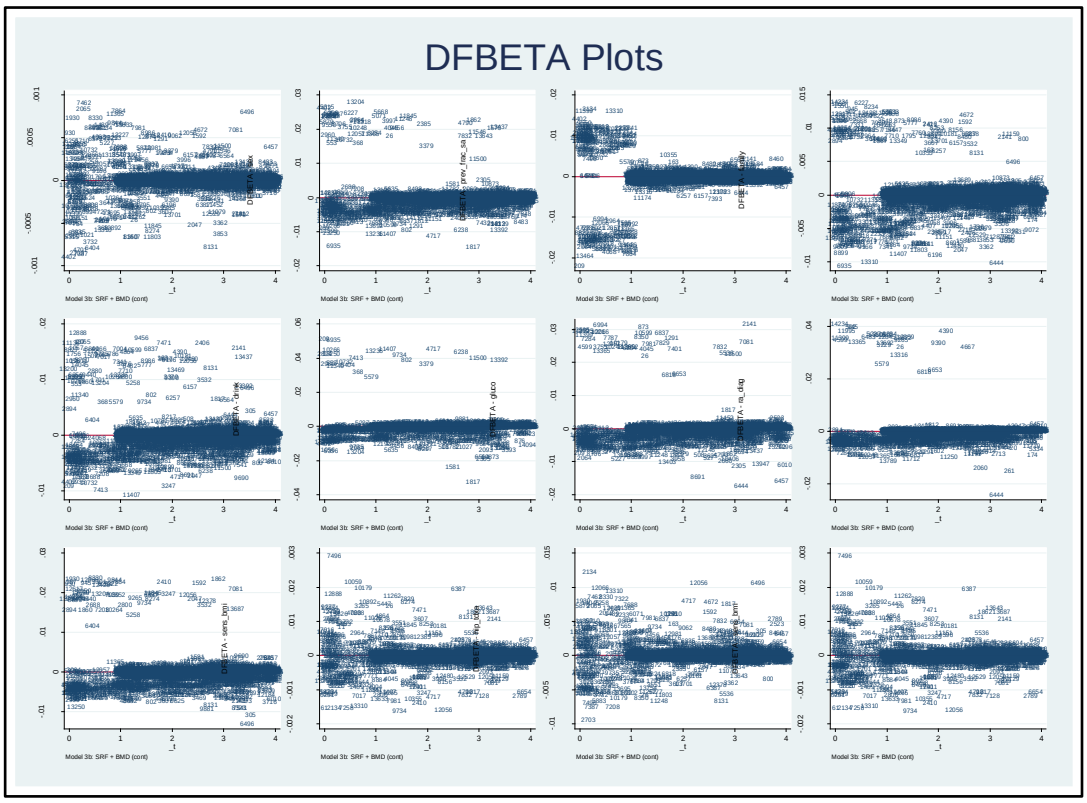
Appendix Figure N 11. L-displacement and LMAX Plots, Model 2



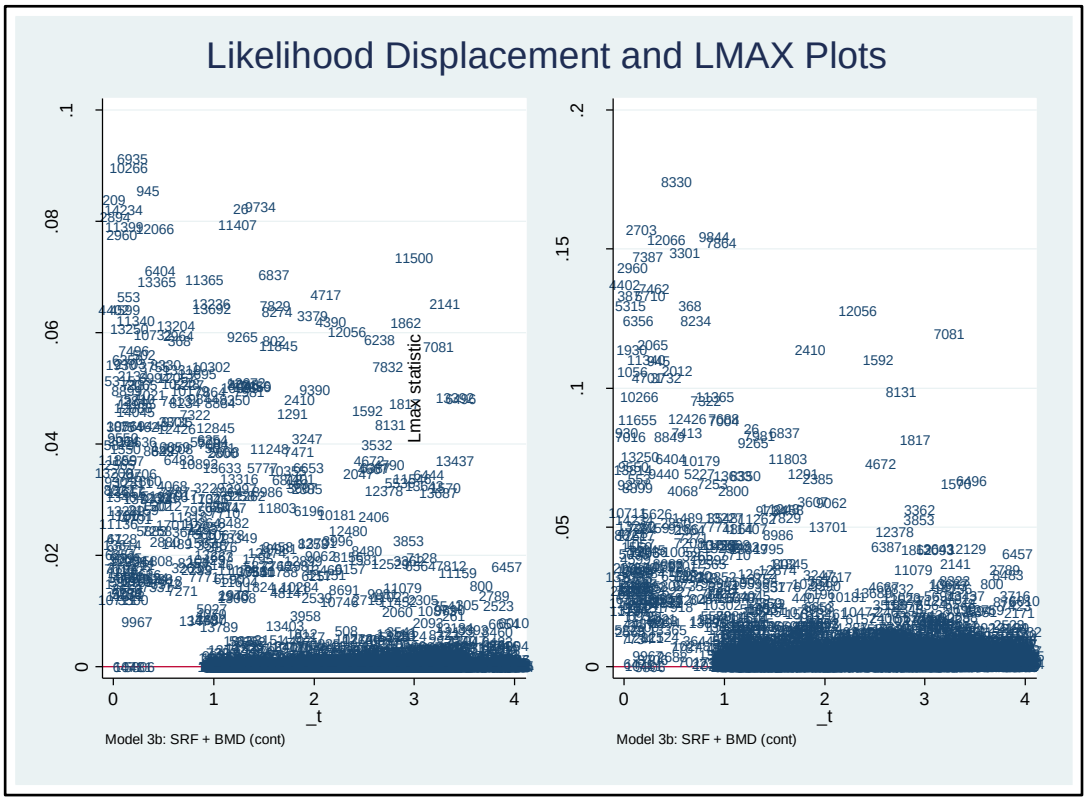
Appendix Figure N 12. DFBETA, Model 3a



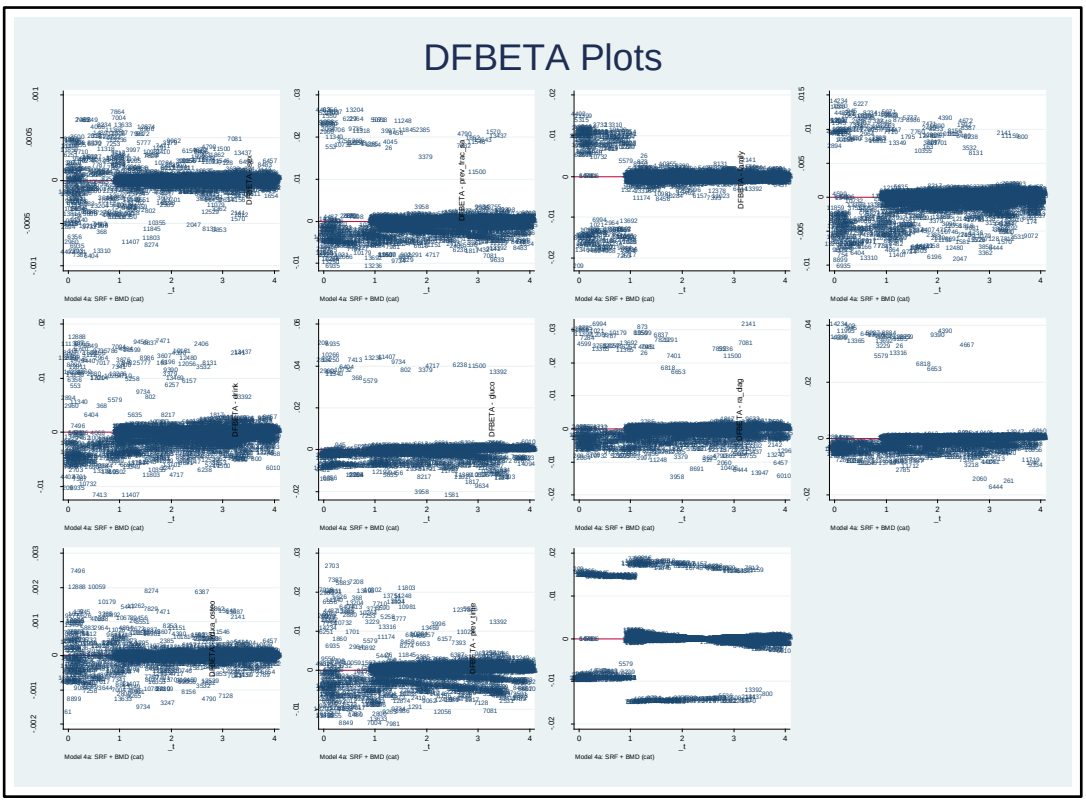
Appendix Figure N 13. L-displacement and LMAX Plots, Model 3a



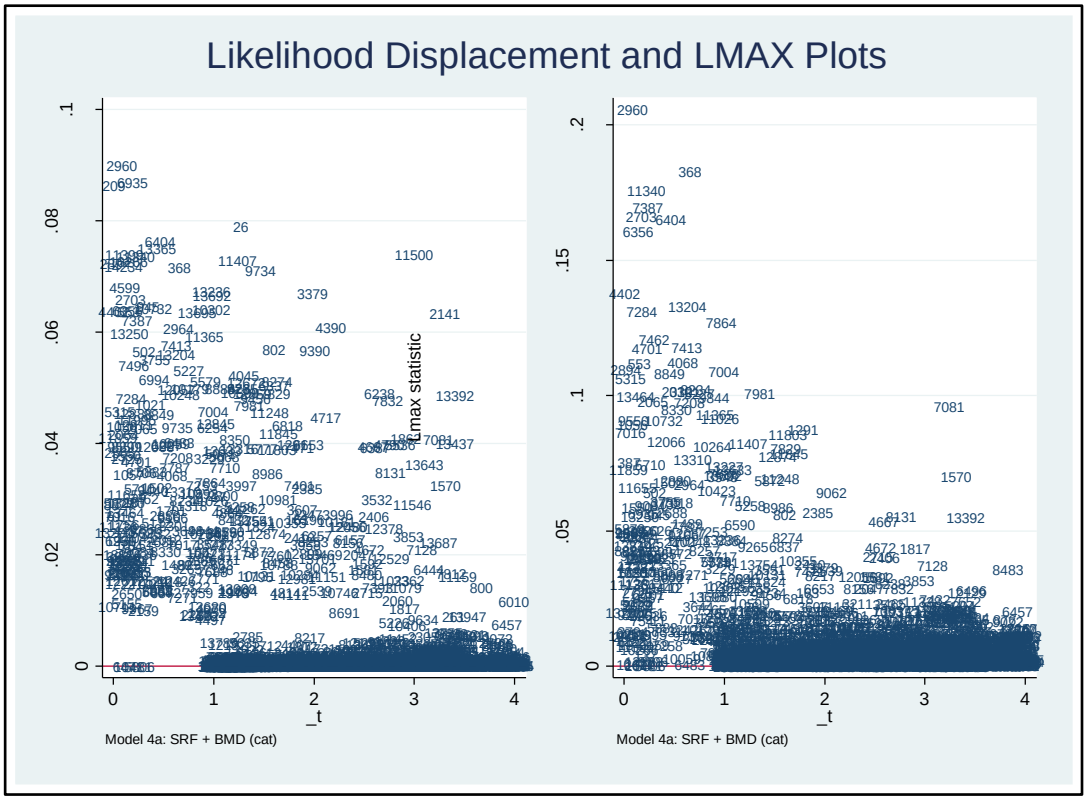
Appendix Figure N 14. DFBETA, Model 3b



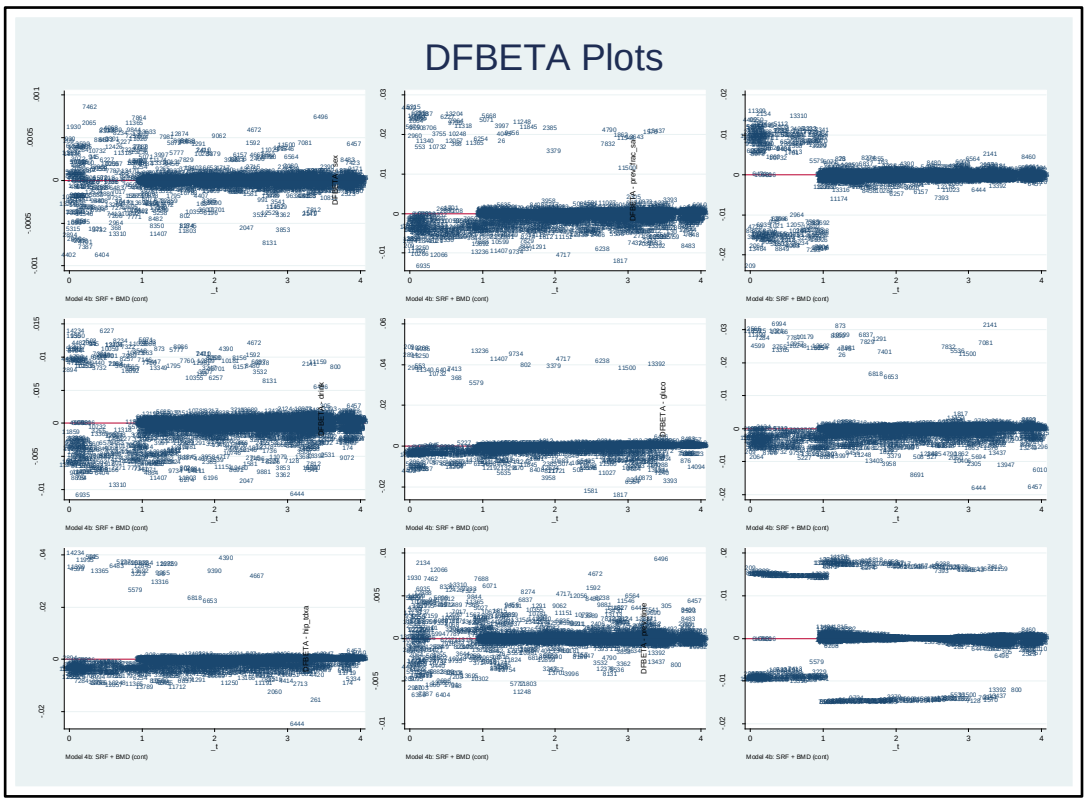
Appendix Figure N 15. L-displacement and LMAX Plots, Model 3b



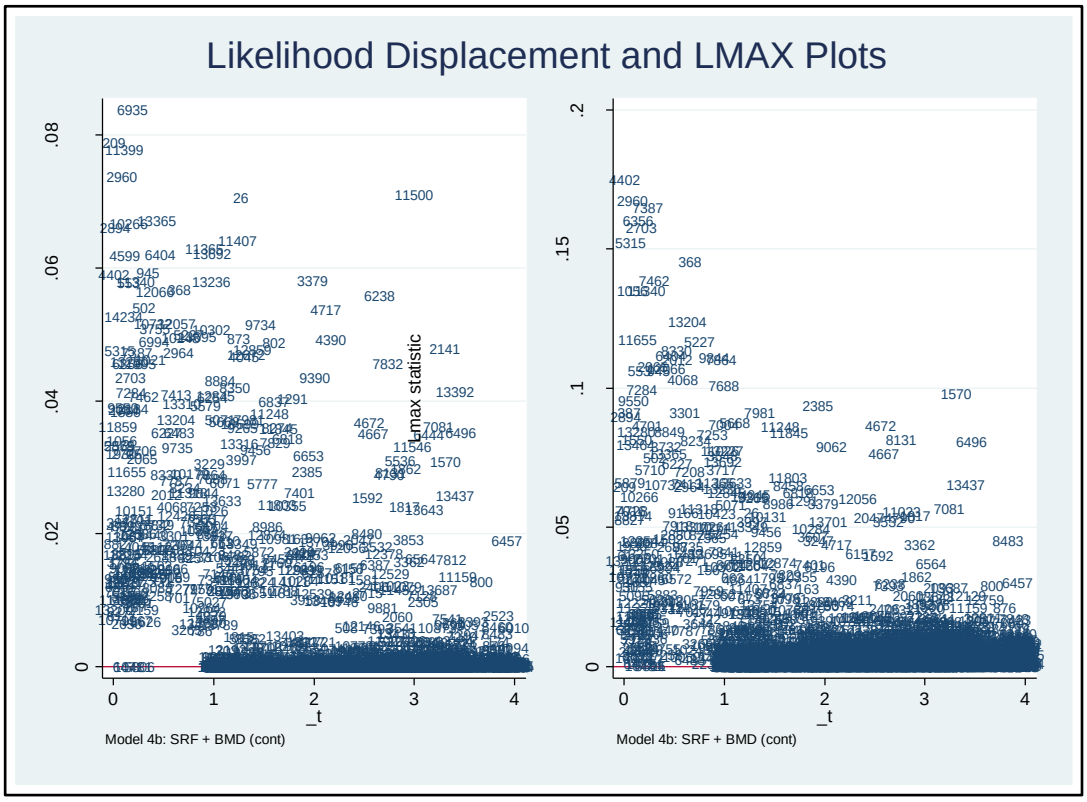
Appendix Figure N 16. DFBETA, Model 4a



Appendix Figure N 17. L-displacement and LMAX Plots, Model 4a



Appendix Figure N 18. DFBETA, Model 4b



Appendix Figure N 19. L-displacement and LMAX Plots, Model 4b