## Supplementary Online Content

The Emerging Risk Factors Collaboration. Lipid-Related Markers and Cardiovascular Disease Prediction. *JAMA*. 2012;307(23):DOI:10.1001/jama.2012.6571.



This supplementary material has been provided by the authors to give readers additional information about their work.

**eTable 1.** Characteristics of 26 prospective studies contributing data to analyses of apoB and apoAI (references are listed in eAppendix 1)

| Study abbreviation, year of publication | Year(s) of baseline survey | Total participants | Age at survey (yrs) mean (sd) | Men n (%) | Current smoking n (%) | History of Diabetes n (%) | ApoB (g/L) mean(sd) | ApoAI (g/L) mean(sd) | Duration of follow-up (yrs) median (5th, 95th percentiles) | No. of outcomes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | CVD | CHD | Stroke |
| ARIC[2], 2001 | 1986-99 | 14467 | 54.48 (5.74) | 6250 (43.2%) | 3706 (25.6%) | 1652 (11.4%) | 0.93 (0.29) | 1.34 (0.31) | 14.0 (5.0, 15.7) | 1424 | 868 | 556 |
| BRUN[4], 2006 | 1990-95 | 797 | 57.48 (11.21) | 388 (48.7%) | 193 (24.2%) | 25 (3.1%) | 1.20 (0.32) | 1.64 (0.29) | 20.2 (4.2, 20.5) | 123 | 63 | 60 |
| CAPS[6], 2000 | 1984-88 | 1162 | 57.12 (4.54) | 1162 (100.0%) | 500 (43.0%) | 27 (2.3%) | 0.95 (0.20) | 1.26 (0.21) | 8.0 (4.4, 8.3) | 72 | 70 | 2 |
| CASTEL[7],* | 1983-90 | 1550 | 73.22 (5.20) | 570 (36.8%) | 208 (13.4%) | 167 (10.8%) | 1.33 (0.30) | 1.58 (0.31) | 11.2 (2.1, 13.0) | 116 | 58 | 58 |
| CHARL[8], * | 1987-89 | 199 | 70.29 (6.77) | 102 (51.3%) | 30 (15.1%) | 52 (26.1%) | 1.38 (0.37) | 1.46 (0.30) | 12.6 (1.4, 13.2) | 39 | 24 | 15 |
| COPEN[11], 1998 | 1991-03 | 8258 | 58.17 (14.68) | 3516 (42.6%) | 4039 (48.9%) | 241 (2.9%) | 0.87 (0.23) | 1.43 (0.29) | 13.2 (2.5, 14.9) | 1137 | 521 | 616 |
| DRECE[13], * | 1991-06 | 2642 | 40.51 (10.65) | 1290 (48.8%) | 1057 (40.0%) | 148 (5.6%) | 1.21 (0.36) | 1.55 (0.26) | 16.4 (15.5, 16.6) | 21 | 16 | 5 |
| DUBBO[14], 2001 | 1988-89 | 2045 | 68.39 (6.70) | 860 (42.1%) | 321 (15.7%) | 135 (6.6%) | 1.07 (0.28) | 1.52 (0.34) | 14.1 (1.8, 14.9) | 467 | 282 | 185 |
| EPICNOR[16], 2007 | 1993-98 | 16050 | 58.86 (9.08) | 6921 (43.1%) | 1823 (11.4%) | 286 (1.8%) | 0.97 (0.24) | 1.57 (0.32) | 10.1 (6.1, 12.0) | 707 | 352 | 355 |
| GRIPS[19], 1997 | 1982 | 5611 | 47.74 (5.10) | 5611 (100.0%) | 2124 (37.9%) | 112 (2.0%) | 1.20 (0.24) | 1.26 (0.28) | 9.8 (4.8, 10.0) | 389 | 287 | 102 |
| KIHD[20], 2000 | 1984-93 | 1992 | 52.52 (5.28) | 1992 (100.0%) | 598 (30.0%) | 85 (4.3%) | 1.02 (0.24) | 1.34 (0.25) | 20.9 (3, 25.1) | 538 | 389 | 149 |
| MOGERAUG[21], 2005 | 1989-95 | 849 | 53.91 (5.83) | 849 (100.0%) | 284 (33.5%) | 37 (4.4%) | 0.91 (0.21) | 1.37 (0.21) | 13.0 (3.4, 13.4) | 83 | 78 | 5 |
| MOSWEGOT[23], * | 1990-96 | 1046 | 48.11 (9.29) | 466 (44.6%) | 268 (25.6%) | 17 (1.6%) | 1.07 (0.29) | 1.40 (0.21) | 9.6 (8.9, 10.2) | 33 | 12 | 21 |
| NHANESIII[24], 1994 | 1990-93 | 3557 | 50.05 (17.97) | 1750 (49.2%) | 1044 (29.4%) | 417 (11.7%) | 1.07 (0.27) | 1.44 (0.26) | 16.2 (4.0, 17.9) | 280 | 208 | 72 |
| PREVEND[26], 2011 | 1997-01 | 6425 | 50.57 (11.44) | 3113 (48.5%) | 2181 (33.9%) | 238 (3.7%) | 1.05 (0.31) | 1.39 (0.30) | 7.6 (4.7, 8.2) | 153 | 125 | 28 |
| PRIME[27], 2002 | 1991-94 | 9523 | 54.84 (2.91) | 9523 (100.0%) | 2526 (26.5%) | 312 (3.3%) | 1.27 (0.33) | 1.49 (0.25) | 5.2 (5.0, 7.3) | 187 | 145 | 42 |
| PROCAM[28], 2001 | 1981-95 | 1802 | 46.61 (8.22) | 1355 (75.2%) | 577 (32.0%) | 44 (2.4%) | 0.92 (0.22) | 1.45 (0.23) | 7.9 (3.2, 17.3) | 33 | 26 | 7 |
| PROSPER[31], 2005 | 1998-99 | 3212 | 75.06 (3.28) | 1337 (41.6%) | 1069 (33.3%) | 396 (12.3%) | 1.14 (0.23) | 1.35 (0.25) | 3.2 (1.1, 3.8) | 375 | 262 | 113 |
| QUEBEC[33], 1996 | 1985-86 | 1293 | 56.05 (6.82) | 1293 (100.0%) | 482 (37.3%) | 83 (6.4%) | 1.17 (0.30) | 1.65 (0.26) | 15.1 (4.2, 15.7) | 144 | 105 | 39 |
| REYK[35], * | 1968-95 | 14956 | 52.81 (8.28) | 6979 (46.7%) | 6895 (46.1%) | 313 (2.1%) | 1.32 (0.35) | 1.74 (0.52) | 24.6 (6.3, 36.5) | 3559 | 2870 | 689 |
| SHS[38], 1997 | 1989-95 | 3979 | 56.12 (8.00) | 1566 (39.4%) | 1326 (33.3%) | 1951 (49.0%) | 1.09 (0.34) | 1.49 (0.31) | 12.5 (2.1, 14.3) | 639 | 433 | 206 |
| TARFS[39], 2004 | 1998-07 | 1769 | 50.75 (10.60) | 857 (48.4%) | 393 (22.2%) | 145 (8.2%) | 1.11 (0.35) | 1.34 (0.31) | 6.3 (0.3, 10.0) | 35 | 28 | 7 |
| ULSAM[40], 2004 | 1970-01 | 1471 | 52.70 (7.67) | 1471 (100.0%) | 884 (60.1%) | 98 (6.7%) | 1.21 (0.28) | 1.41 (0.26) | 24.0 (4.8, 37.3) | 588 | 387 | 201 |
| WHITEI[41], 2000 | 1997-98 | 3812 | 76.42 (4.60) | 3812 (100.0%) | 519 (13.6%) | 173 (4.5%) | 0.86 (0.23) | 0.96 (0.14) | 8.2 (2.0, 8.4) | 333 | 203 | 130 |
| WHITEII[42], 2003 | 1987-93 | 7989 | 49.06 (5.70) | 5487 (68.7%) | 1621 (20.3%) | 93 (1.2%) | 1.25 (0.30) | 2.06 (0.39) | 12.3 (5.1, 16.9) | 298 | 292 | 6 |
| WHS[44], 2005 | 1993-04 | 23125 | 55.12 (7.22) | 0 (0.0%) | 2652 (11.5%) | 724 (3.1%) | 1.05 (0.28) | 1.51 (0.26) | 10.2 (8.4, 10.8) | 461 | 217 | 244 |
| TOTAL (26 studies) | | 139581 | 56.42 (8.41) | 68520 (49.1%) | 37320 (26.7%) | 7971 (5.7%) | 1.10 (0.29) | 1.46 (0.32) | 10.5 (3.4, 26.7) | 12234 | 8321 | 3913 |

*studies that had not previously published their findings on apolipoprotein B or AI and vascular risk.

**eTable 2.** Characteristics of 24 prospective studies contributing data to analyses of lipoprotein(a) (references are listed in eAppendix 1)

| Study abbreviation, year of publication | Year(s) of baseline survey | Total participants | Age at survey (yrs) mean (sd) | Men n (%) | Current smoking n (%) | History of Diabetes n (%) | Lp(a) (mg/dl) median (IQR) | Duration of follow-up (yrs) median (5th, 95th percentiles) | No. of outcomes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | CVD | CHD | Stroke |
| AFTCAPS[1], * | 1990-93 | 896 | 58.84 (7.12) | 739 (82.5%) | 115 (12.8%) | 16 (1.8%) | 7.6 (3.3, 17.9) | 5.7 (4.5, 6.8) | 32 | 21 | 11 |
| ARIC[2], 2001 | 1986-99 | 14001 | 54.45 (5.74) | 6069 (43.3%) | 3574 (25.5%) | 1557 (11.1%) | 18.3 (6.9, 43.8) | 14.1 (5.0, 15.7) | 1378 | 846 | 532 |
| BRUN[5], 2007 | 1990-95 | 779 | 57.52 (11.25) | 376 (48.3%) | 189 (24.3%) | 25 (3.2%) | 8.9 (4.3, 21.7) | 20.2 (4.2, 20.5) | 121 | 62 | 59 |
| CHARL[8], 1993 | 1987-89 | 159 | 69.63 (7.71) | 159 (100.0%) | 29 (18.2%) | 35 (22.0%) | 9.1 (3.4, 19.7) | 11.9 (1.6, 12.9) | 32 | 21 | 11 |
| CHS1[9], 2003 | 1989-93 | 3753 | 72.23 (5.18) | 1427 (38.0%) | 448 (11.9%) | 489 (13.0%) | 12.6 (4.8, 22.2) | 12.1 (2.0, 12.9) | 1008 | 568 | 440 |
| COPEN[12], 2008 | 1991-03 | 7999 | 58.25 (14.64) | 3385 (42.3%) | 3872 (48.4%) | 234 (2.9%) | 18.5 (6.8, 41.5) | 13.1 (2.5, 14.9) | 1090 | 497 | 593 |
| DUBBO[14], 2001 | 1988-89 | 1998 | 68.41 (6.74) | 841 (42.1%) | 312 (15.6%) | 132 (6.6%) | 11 (5, 27.5) | 14.1 (1.8, 14.9) | 452 | 273 | 179 |
| EAS[15], 2001 | 1987-88 | 619 | 64.35 (5.58) | 314 (50.7%) | 130 (21.0%) | 29 (4.7%) | 9.2 (3.7, 25.8) | 19.6 (2.2, 21.0) | 115 | 62 | 53 |
| EPICNOR[48], 2010 | 1993-98 | 16444 | 58.82 (9.10) | 7044 (42.8%) | 1858 (11.3%) | 288 (1.8%) | 11.6 (6.2, 27.4) | 10.1 (6.1, 12.1) | 720 | 357 | 363 |
| FINRISK92[17], 2005 | 1992 | 2194 | 53.52 (6.29) | 1017 (46.4%) | 551 (25.1%) | 55 (2.5%) | 12.2 (4.5, 31.6) | 16.8 (4.4, 16.9) | 170 | 100 | 70 |
| GRIPS[19], 1997 | 1982 | 5784 | 47.74 (5.11) | 5784 (100.0%) | 2178 (37.7%) | 113 (2.0%) | 9 (4, 25) | 9.8 (4.8, 10.0) | 402 | 299 | 103 |
| KIHD[20], * | 1984-93 | 1980 | 52.52 (5.28) | 1980 (100.0%) | 601 (30.4%) | 86 (4.3%) | 9.6 (3.9, 22.1) | 20.9 (3.0, 25.1) | 538 | 389 | 149 |
| NHANESIII[24], * | 1990-93 | 2932 | 50.50 (17.10) | 1215 (41.4%) | 723 (24.7%) | 389 (13.3%) | 22 (9, 45) | 13.3 (4.3, 15.1) | 174 | 127 | 47 |
| PREVEND[26], * | 1997-01 | 6425 | 50.57 (11.44) | 3113 (48.5%) | 2181 (33.9%) | 238 (3.7%) | 5 (2, 14) | 7.6 (4.7, 8.2) | 152 | 124 | 28 |
| PRIME[27], 2002 | 1991-94 | 7433 | 54.84 (2.90) | 7433 (100.0%) | 1972 (26.5%) | 229 (3.1%) | 10 (5, 30) | 5.2 (5.0, 7.3) | 144 | 114 | 30 |
| PROCAM[29], 1996 | 1981-95 | 3192 | 42.95 (10.38) | 2250 (70.5%) | 1141 (35.7%) | 54 (1.7%) | 4 (2, 13) | 17.4 (5.3, 18.5) | 112 | 94 | 18 |
| QUEBEC[32], 1998 | 1985-86 | 1277 | 56.08 (6.81) | 1277 (100.0%) | 476 (37.3%) | 83 (6.5%) | 18.8 (7.6, 47) | 15.1 (4.2, 15.7) | 143 | 104 | 39 |
| REYK[35], * | 1968-95 | 14685 | 52.83 (8.30) | 6782 (46.2%) | 6767 (46.1%) | 306 (2.1%) | 5.2 (2.4, 11.8) | 24.6 (6.3, 36.5) | 3498 | 2822 | 676 |
| SHS[37], 2002 | 1989-95 | 3740 | 56.15 (8.00) | 1484 (39.7%) | 1240 (33.2%) | 1832 (49.0%) | 3 (1.1, 6.7) | 12.5 (2.2, 14.3) | 603 | 408 | 195 |
| ULSAM[40], * | 1970-01 | 1461 | 52.65 (7.57) | 1461 (100.0%) | 882 (60.4%) | 93 (6.4%) | 8.3 (3.5, 22.5) | 24 (5.1, 37.3) | 583 | 385 | 198 |
| WHITEII[43], 2011 | 1987-93 | 7706 | 49.50 (6.04) | 5331 (69.2%) | 1478 (19.2%) | 89 (1.2%) | 20 (12, 46) | 12.3 (5.0, 13.0) | 280 | 274 | 6 |
| WHS[45], 2006 | 1993-04 | 23124 | 55.12 (7.22) | 0 (0.0%) | 2653 (11.5%) | 724 (3.1%) | 10.8 (4.4, 33) | 10.2 (8.4, 10.8) | 461 | 217 | 244 |
| WOSCOPS[46], 2000 | 1989-91 | 4617 | 55.20 (5.57) | 4617 (100.0%) | 2080 (45.1%) | 53 (1.1%) | 17 (7, 50) | 5.0 (2.8, 6.0) | 360 | 299 | 61 |
| ZUTE[47], * | 1990 | 304 | 75.46 (4.41) | 304 (100.0%) | 103 (33.9%) | 25 (8.2%) | 12.1 (5.8, 28.7) | 9.1 (1.1, 10.1) | 71 | 42 | 29 |
| TOTAL (24 studies) | | 133502 | 57.00 (8.36) | 64402 (48.2%) | 35553 (26.6%) | 7174 (5.4%) | 10.9 (4.4, 28) | 10.6 (4.4, 26.9) | 12639 | 8505 | 4134 |

*studies that had not previously published their findings on lipoprotein (a) and vascular risk.

**eTable 3.** Characteristics of 11 prospective studies contributing data to analyses of $LpPLA_2$ mass or activity (references are listed in eAppendix 1)

| Study abbreviation, year of publication | Year(s) of baseline survey | Total participants | Age at survey (yrs) mean (sd) | Men n (%) | Current smoking n (%) | History of Diabetes n (%) | Lp-$PLA_2$ activity nmol/min/ml mean(sd) | Lp-$PLA_2$ mass µg/l mean(sd) | Duration of follow-up (yrs) median (5th, 95th percentiles) | No. of outcomes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | CVD | CHD | Stroke |
| ARIC[3], 2004 | 1986-99 | 1314 | 58.73 (5.59) | 758 (57.7%) | 334 (25.4%) | 303 (23.1%) | - | 402.54 (140.09) | 10.6 (1.6, 12.7) | 416 | 250 | 166 |
| BRUN[5], 2007 | 1990-95 | 665 | 61.54 (10.93) | 319 (48.0%) | 134 (20.2%) | 35 (5.3%) | 772.63 (196.24) | - | 10.4 (3.1, 10.4) | 53 | 33 | 20 |
| CHS1[10], 2010 | 1989-93 | 3641 | 72.21 (5.17) | 1352 (37.1%) | 426 (11.7%) | 465 (12.8%) | 39.67 (12.77) | 347.27 (117.77) | 12.1 (2.6, 12.9) | 911 | 484 | 427 |
| CHS2[10], 2010 | 1989-93 | 451 | 72.28 (5.21) | 169 (37.5%) | 72 (16.0%) | 95 (21.1%) | 32.16 (10.13) | 288.18 (95.25) | 9.1 (1.7, 9.5) | 93 | 48 | 45 |
| FRAMOFF[18], 2009 | 1998-01 | 2928 | 60.31 (9.30) | 1288 (44.0%) | 0 (0.0%) | 336 (11.5%) | 143.09 (35.62) | 300.33 (93.24) | 6.2 (2.9, 7.5) | 102 | 66 | 36 |
| MOGERAUG[22], 2004 | 1984-95 | 873 | 53.94 (5.80) | 873 (100.0%) | 295 (33.8%) | 37 (4.2%) | - | 267.34 (84.18) | 13 (3.6, 13.4) | 84 | 79 | 5 |
| NPHSII[25], 2009 | 1989-94 | 1648 | 56.07 (3.49) | 1648 (100.0%) | 443 (26.9%) | 27 (1.6%) | 50.13 (16.17) | - | 13.9 (7.5, 15.4) | 83 | 59 | 24 |
| PROSPER[30], 2010 | 1989-99 | 3174 | 75.06 (3.27) | 1315 (41.4%) | 1057 (33.3%) | 386 (12.2%) | 146.84 (29.26) | 294.89 (67.46) | 3.2 (1.1, 3.8) | 375 | 262 | 113 |
| RANCHO[34], 2008 | 1984-87 | 1190 | 70.33 (10.75) | 656 (55.1%) | 155 (13.0%) | 45 (3.8%) | - | 539.09 (218.27) | 13.3 (1.9, 18.8) | 320 | 177 | 143 |
| REYK[35], * | 1968-95 | 14962 | 52.82 (8.29) | 6984 (46.7%) | 6896 (46.1%) | 314 (2.1%) | 119.96 (28.02) | 926.32 (295.42) | 24.6 (6.3, 36.5) | 3565 | 2876 | 689 |
| ROTT[36], 2005 | 1989-99 | 1229 | 68.28 (8.20) | 452 (36.8%) | 291 (23.7%) | 128 (10.4%) | 44.47 (11.18) | - | 6.2 (1.9, 8.3) | 148 | 89 | 59 |
| TOTAL (11 studies) | | 32075 | 63.78 (7.52) | 15814 (49.3%) | 10103 (31.5%) | 2171 (6.8%) | 166.83 (39.59) | 420.74 (228.45) | 12.7 (2.9, 34.6) | 6150 | 4423 | 1727 |

*studies that had not previously published their findings on $LpPLA_2$ mass and/or activity and vascular risk.

**eTable 4.** Assays characteristics of studies contributing data

| Study abbreviation | Year of publication | Apolipoprotein B | Apolipoprotein AI | Lp(a) | Lp-PLA2 mass | Lp-PLA2 activity |
|---|---|---|---|---|---|---|
| AFTCAPS[1] | * | - | - | NS | - | - |
| ARIC[2,3] | 2001; 2004 | RIA (In-house) | RIA (In-house) | ELISA (In-house) | PLAC I | - |
| BRUN[4,5] | 2006; 2007 | INA (Behring) | INA (Behring) | ELISA (Immuno) | - | Colorimetric (Azwell) |
| CAPS[6] | 2000 | INA (In-house) | INA (In-house) | - | - | - |
| CASTEL[7] | * | RIA (In-house) | RIA (In-house) | - | - | - |
| CHARL[8] | 1993 | NS | NS | ELISA (Terumo) | - | - |
| CHS[9,10] | 2003; 2010 | - | - | ELISA (Genetech) | PLAC II | Radiometric |
| COPEN[11,12] | 1998; 2008 | ITA (Boehringer) | ITA (Boehringer) | ITA (Dako) | - | - |
| DRECE[13] | * | NS | NS | - | - | - |
| DUBBO[14] | 2001 | INA (Bechkman) | INA (Bechkman) | ELISA (Biopool) | - | - |
| EAS[15] | 2001 | - | - | ELISA (Biopool) | - | - |
| EPICNOR[16,48] | 2007; 2010 | ITA (Boehringer) | ITA (Boehringer) | ITA (Denka Seiken) | - | - |
| FINRISK92[17] | 2005 | - | - | IRMA (Pharmacia) | - | - |
| FRAMOFF[18] | 2009 | - | - | - | PLAC II | Colorimetric (CAM) |
| GRIPS[19] | 1997 | INA (Behring) | INA (Behring) | ELISA (Immuno) | - | - |
| KIHD[20] | 2000 | ITA (Kone) | ITA (Kone) | IRMA (Pharmacia) | - | - |
| MOGERAUG[21, 22] | 2005; 2004 | ITA (In-house) | ITA (In-house) | - | PLAC I | - |
| MOSWEGOT[23] | * | NS | NS | - | - | - |
| NHANESIII[24] | 1994 | INA (Bechkman) | RID (Tago) | ELISA (Strategic Diagnostic) | - | - |
| NPHSII[25] | 2009 | - | - | - | - | Radiometric |
| PREVEND[26] | 2011 | INA (Behring) | INA (Behring) | INA (Behring) | - | - |
| PRIME[27] | 2002 | INA (Behring) | INA (Behring) | ELISA (In-house) | - | - |
| PROCAM[28,29] | 2001; 1996 | ITA (Boehringer) | ITA (Boehringer) | EID (Behringwerke) | - | - |
| PROSPER[30,31] | 2005; 2010 | ITA (Roche) | ITA (Roche) | - | PLAC II | Colorimetric (CAM) |
| QUEBEC[32,33] | 1996; 1998 | IE (In-house) | IE (In-house) | ELISA (Biopool) | - | - |
| RANCHO[34] | 2008 | - | - | - | PLAC II | - |
| REYK[35] | * | ITA (Randox) | ITA (Randox) | ITA (Randox) | autoPLAC | Colorimetric (CAM) |
| ROTT[36] | 2005 | - | - | - | - | Radiometric |
| SHS[37,38] | 1997; 2002 | RIA (NS) | RIA (NS) | ELISA (Terumo) | - | - |
| TARFS[39] | 2004 | ITA (Behring) | ITA (Behring) | - | - | - |
| ULSAM[40] | 2004 | RIA (Pharmacia) | RIA (Pharmacia) | IRMA (Pharmacia) | - | - |
| WHITEI[41] | 2000 | ITA (Immuno) | ITA (Immuno) | - | - | - |
| WHITEII[42,43] | 2003; 2011 | ITA (NS) | ITA (NS) | ITA (NS) | - | - |
| WHS[44,45] | 2005; 2006 | ITA (NS) | ITA (NS) | ITA (Denka Seiken) | - | - |
| WOSCOPS[46] | 2000 | - | - | ELISA (Innogenetics) | - | - |
| ZUTE[47] | * | - | - | ELISA (In-house) | - | - |

*studies that had not previously published their findings
ELISA, enzyme linked immunosorbent assay; IC, immunochemical; IE, immunoelectrophoresis; INA, immunonephelometric assay; ITA, immunoturbidimetric assay; IRMA, immunoradiometric assay; NS, not specified; RIA, radioimmuno assay.

**eTable 5.** Baseline characteristics of participants included in the current analysis and in the AMORIS study

| | Studies excluding AMORIS 26 studies, 139581 participants, 12234 CVD cases | | AMORIS study 151667 participants, 3183 CVD cases | |
|---|---|---|---|---|
| **Conventional risk factors** | Mean (SD) or No. (%) | Hazard ratio * (95% CI) | Mean (SD) or No. (%) | Hazard ratio * (95% CI) |
| Ages at survey (years) | 56.42 (8.41) | 1.98 (1.85, 2.13) | 51.03 (12.05) | 4.51 (4.34, 4.69) |
| Male sex | 68520 (49.1%) | NA‡ | 84580 (55.8%) | NA‡ |
| Total-cholesterol (mmol/L) | 5.86 (1.10) | 1.25 (1.19, 1.30) | 6.01 (1.17) | 1.13 (1.09, 1.17) |
| HDL-C (mmol/L) | 1.33 (0.38) | 0.79 (0.75, 0.84) | 1.54 (0.44) | 0.74 (0.71, 0.76) |
| Non-HDL-C (mmol/L) | 4.52 (1.13) | 1.31 (1.25, 1.37) | 4.47 (1.22) | 1.25 (1.21, 1.29) |
| Triglycerides (mmol/L) † | 1.3 (0.9, 1.9) § | 1.28 (1.23, 1.33) | 1.2 (0.8, 1.8) | 1.34 (1.29, 1.39) |
| **Lipids combinations and lipid-related markers** | | | | |
| Total-cholesterol / HDL-C ratio † | 4.4 (3.5, 5.4) § | 1.38 (1.30, 1.46) | 3.9 (3.1, 4.9) | 1.37 (1.33, 1.42) |
| Apolipoprotein AI (g/L) | 1.46 (0.32) | 0.85 (0.82, 0.89) | 1.44 (0.23) | 0.79 (0.76, 0.82) |
| Apolipoprotein B (g/L) | 1.10 (0.29) | 1.28 (1.22, 1.35) | 1.29 (0.34) | 1.23 (1.19, 1.27) |
| Apo-B / Apo-AI ratio † | 0.7 (0.6, 0.9) § | 1.36 (1.29, 1.43) | 0.9 (0.7, 1.1) | 1.39 (1.34, 1.44) |

* Hazard ratio (95% confidence interval) per 1-SD higher age and measured biomarker level. Analyses were adjusted for age and stratified by sex. † Variables were log transformed. § Median and interquartile range. ‡ Models were stratified by sex.

**eTable 6.** Changes in CVD risk discrimination and reclassification when replacing total and HDL cholesterol with different combinations of lipids and apolipoproteins in a model including conventional risk factors

| | Change in C-index (95% CI) | P-value for change in C-index | $I^2$ (95% CI) for Change in C-index | NRI (95% CI) | P-value for NRI | IDI (95% CI) |
|---|---|---|---|---|---|---|
| | **26 studies 139581 participants 12234 CVD cases** | | | **17 studies 77886 participants 5436 CVD cases** | | |
| Total cholesterol / HDL-C | -0.0098 (-0.0114, -0.0082) | <0.0001 | 84 (77, 88) | -2.74% (-3.77%, -1.71%) | <0.0001 | -0.0048 (-0.0058, -0.0037) |
| $Log_e$ (Total cholesterol / HDL-C) | -0.0002 (-0.0007, 0.0003) | 0.491 | 21 (0, 52) | -1.04% (-1.64%, -0.45%) | 0.001 | 0.0000 (-0.0004, 0.0005) |
| HDL-C / Total cholesterol | -0.0017 (-0.0022, -0.0011) | <0.0001 | 28 (0, 56) | -1.43% (-2.04%, -0.82%) | <0.0001 | -0.0008 (-0.0012, -0.0003) |
| $Log_e$ (HDL-C / Total cholesterol) | -0.0002 (-0.0007, 0.0003) | 0.491 | 21 (0, 52) | -1.04% (-1.64%, -0.45%) | 0.001 | 0.0000 (-0.0004, 0.0005) |
| Non-HDL-C | -0.0024 (-0.0032, -0.0016) | <0.0001 | 43 (10, 64) | -0.48% (-1.16%, 0.21%) | 0.171 | -0.0010 (-0.0015, -0.0005) |
| Apo-B & Apo-AI | -0.0028 (-0.0039, -0.0017) | <0.0001 | 72 (58, 81) | -1.08% (-1.94%, -0.22%) | 0.014 | -0.0016 (-0.0024, -0.0008) |
| Apo-B / Apo-AI | -0.0056 (-0.0070, -0.0043) | <0.0001 | 80 (71, 86) | -2.03% (-2.98%, -1.09%) | <0.0001 | -0.0034 (-0.0043, -0.0024) |
| $Log_e$ (Apo-B / Apo-AI) | -0.0036 (-0.0048, -0.0024) | <0.0001 | 79 (69, 85) | -1.75% (-2.63%, -0.87%) | <0.0001 | -0.0017 (-0.0025, -0.0009) |
| Apo-AI / Apo-B | -0.0051 (-0.0064, -0.0038) | <0.0001 | 79 (70, 85) | -1.81% (-2.69%, -0.93%) | <0.0001 | -0.0021 (-0.0029, -0.0013) |
| $Log_e$ (Apo-AI / Apo-B) | -0.0036 (-0.0048, -0.0024) | <0.0001 | 79 (69, 85) | -1.75% (-2.63%, -0.87%) | <0.0001 | -0.0017 (-0.0025, -0.0009) |
| Total cholesterol & Apo-AI | -0.0025 (-0.0035, -0.0014) | <0.0001 | 37 (0, 61) | -0.58% (-1.36%, 0.20%) | 0.146 | -0.0013 (-0.0019, -0.0007) |
| Apo-B & HDL-C | -0.0020 (-0.0029, -0.0011) | <0.0001 | 69 (54, 80) | -1.33% (-2.07%, -0.59%) | <0.001 | -0.0008 (-0.0015, -0.0002) |
| | **21 studies 83573 participants 9116 CVD cases** | | | **13 studies 44396 participants 3233 CVD cases** | | |
| Calculated LDL cholesterol | -0.0049 (-0.0063, -0.0036) | <0.0001 | 54 (24, 72) | -1.73% (-2.86%, -0.59%) | 0.003 | -0.0025 (-0.0034, -0.0015) |

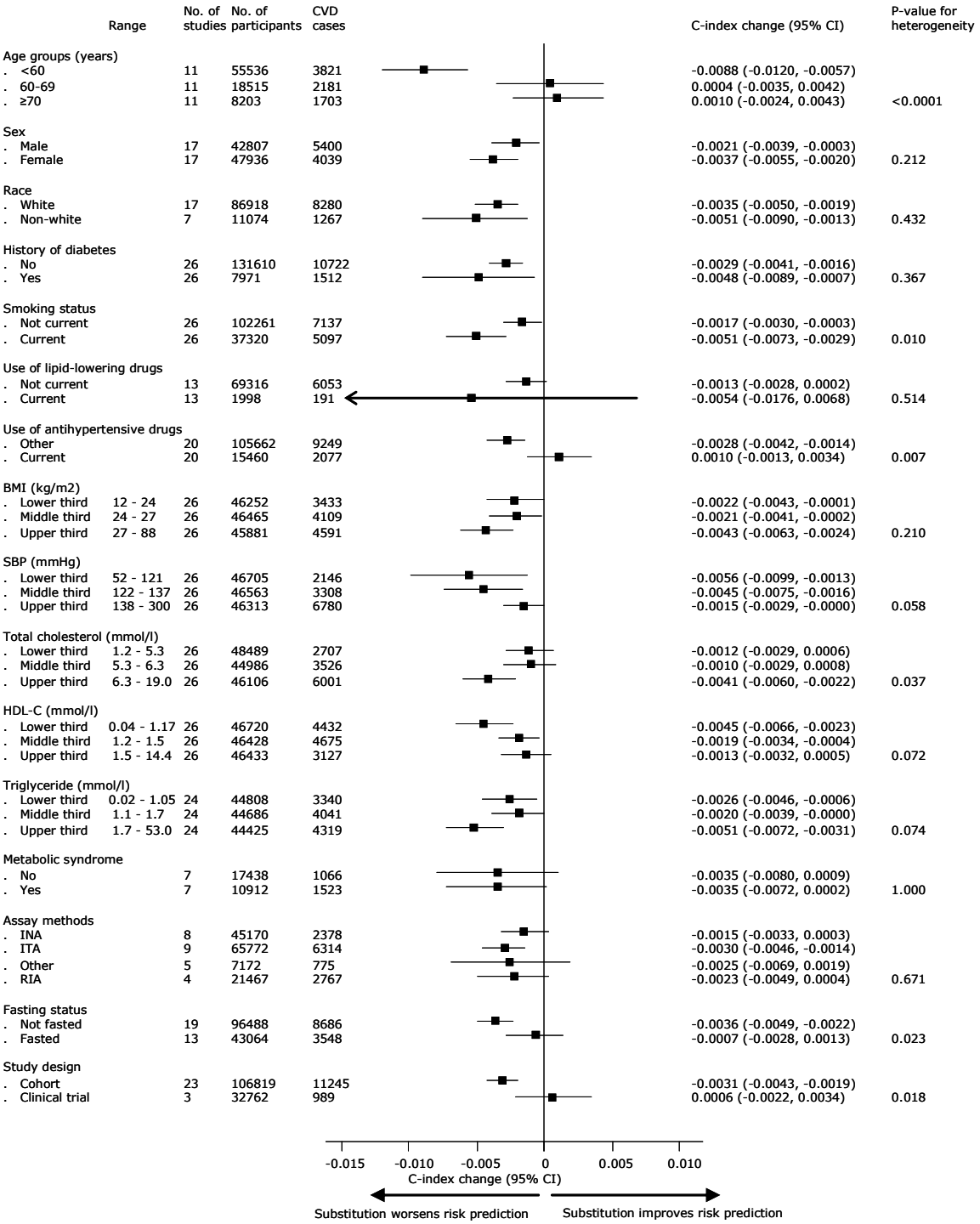
Conventional risk factors were age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol, each included as individual linear term. Models were stratified by sex. NRI, net reclassification improvement. IDI, integrated discrimination index. For changes in the C-index, NRI and IDI values less than zero indicate worsened predictive ability and values greater than zero improved ability. C-index for a model including containing conventional risk factors (ie, age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol, each included as individual linear terms) was 0.7244 (95% CI 0.7200, 0.7289).

**eTable 7.** Changes in CVD risk discrimination and classification after addition of lipid-related markers to a model including conventional risk factors

| | C-index* (95% CI) | Change in C-index (95% CI) | p-value for change in C-index | I2 for change in C-index (95% CI) | NRI (95% CI) | P-value for NRI | IDI (95% CI) |
|---|---|---|---|---|---|---|---|
| **Adding triglycerides** | | **24 studies 133919 participants 11700 CVD cases** | | | **16 studies 76575 participants 5324 CVD cases** | | |
| + Triglycerides | 0.7268 (0.7223, 0.7313) | 0.0000 (-0.0001, 0.0002) | 0.824 | 23 (0, 54) | -0.49% (-0.76%, -0.21%) | 0.001 | -0.0000 (-0.0001, 0.0001) |
| **Adding apolipoproteins** | | **26 studies 139581 participants 12234 CVD cases** | | | **17 studies 77886 participants 5436 cases** | | |
| + Apo-B | 0.7244 (0.7200, 0.7289) | 0.0001 (-0.0002, 0.0003) | 0.639 | 0 (0, 43) | -0.17% (-0.52%, 0.18%) | 0.337 | 0.0002 (0.0000, 0.0004) |
| + Apo-AI | | 0.0005 (0.0002, 0.0007) | <0.0001 | 67 (50, 78) | -0.20% (-0.49%, 0.08%) | 0.166 | 0.0001 (0.0000, 0.0002) |
| + Apo-B & Apo-AI | | 0.0006 (0.0002, 0.0009) | 0.001 | 39 (2, 62) | -0.10% (-0.49%, 0.29%) | 0.621 | 0.0003 (0.0001, 0.0005) |
| + Apo-B / Apo-AI | | 0.0004 (0.0002, 0.0006) | 0.001 | 24 (0, 53) | 0.14% (-0.18%, 0.45%) | 0.400 | 0.0002 (0.0000, 0.0003) |
| **Adding lipoprotein(a)** | | **24 studies 133502 participants 12639 CVD cases** | | | **18 studies 82968 participants 6308 cases** | | |
| + Lp(a) | 0.7229 (0.7185, 0.7272) | 0.0016 (0.0009, 0.0023) | <0.0001 | 42 (5, 64) | 0.05% (-0.59%, 0.70%) | 0.871 | 0.0009 (0.0004, 0.0014) |
| **Adding Lp-PLA2** | | **8 studies 28567 participants 5299 CVD cases** | | | **4 studies 19289 participants 1534 cases** | | |
| + Lp-PLA2 activity | 0.7068 (0.6997, 0.7138) | 0.0001 (-0.0003, 0.0006) | 0.630 | 65 (26, 84) | 0.21% (-0.45%, 0.86%) | 0.535 | -0.0002 (-0.0006, 0.0002) |
| | | **8 studies 28494 participants 5859 CVD cases** | | | **5 studies 20282 participants 2107 cases** | | |
| + Lp-PLA2 mass | 0.7025 (0.6958, 0.7092) | 0.0018 (0.0010, 0.0026) | <0.0001 | 40 (0, 74) | 0.81% (-0.15%, 1.77%) | 0.099 | 0.0016 (0.0008, 0.0024) |

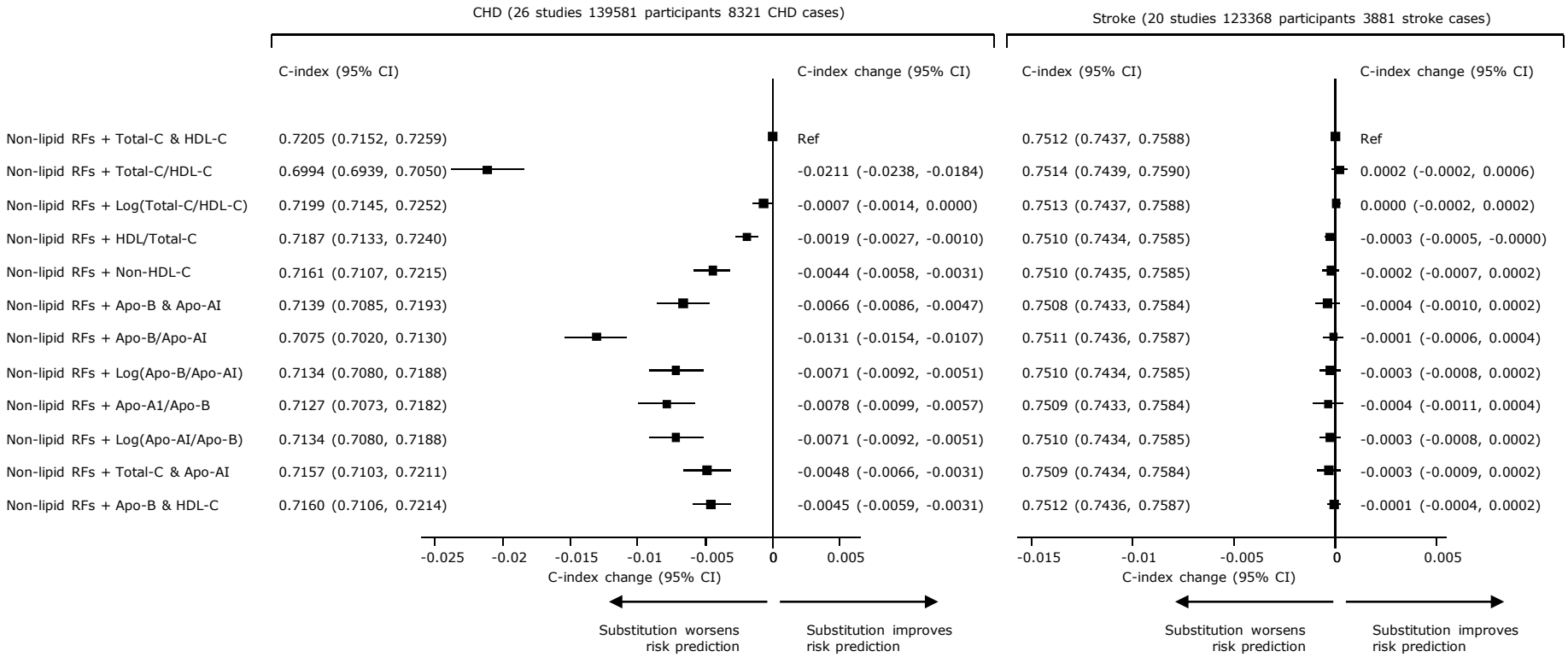
*C-index for a model including containing conventional risk factors. Conventional risk factors were age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol, each included as individual linear term. Models were stratified by sex. NRI, net reclassification improvement. IDI, integrated discrimination index. For changes in the C-index, NRI and IDI values less than zero indicate worsened predictive ability and values greater than zero improved ability.

**eFigure 1.** Changes in C-index for CVD when replacing total and HDL cholesterol with apoB and apoAI according to different subgroups



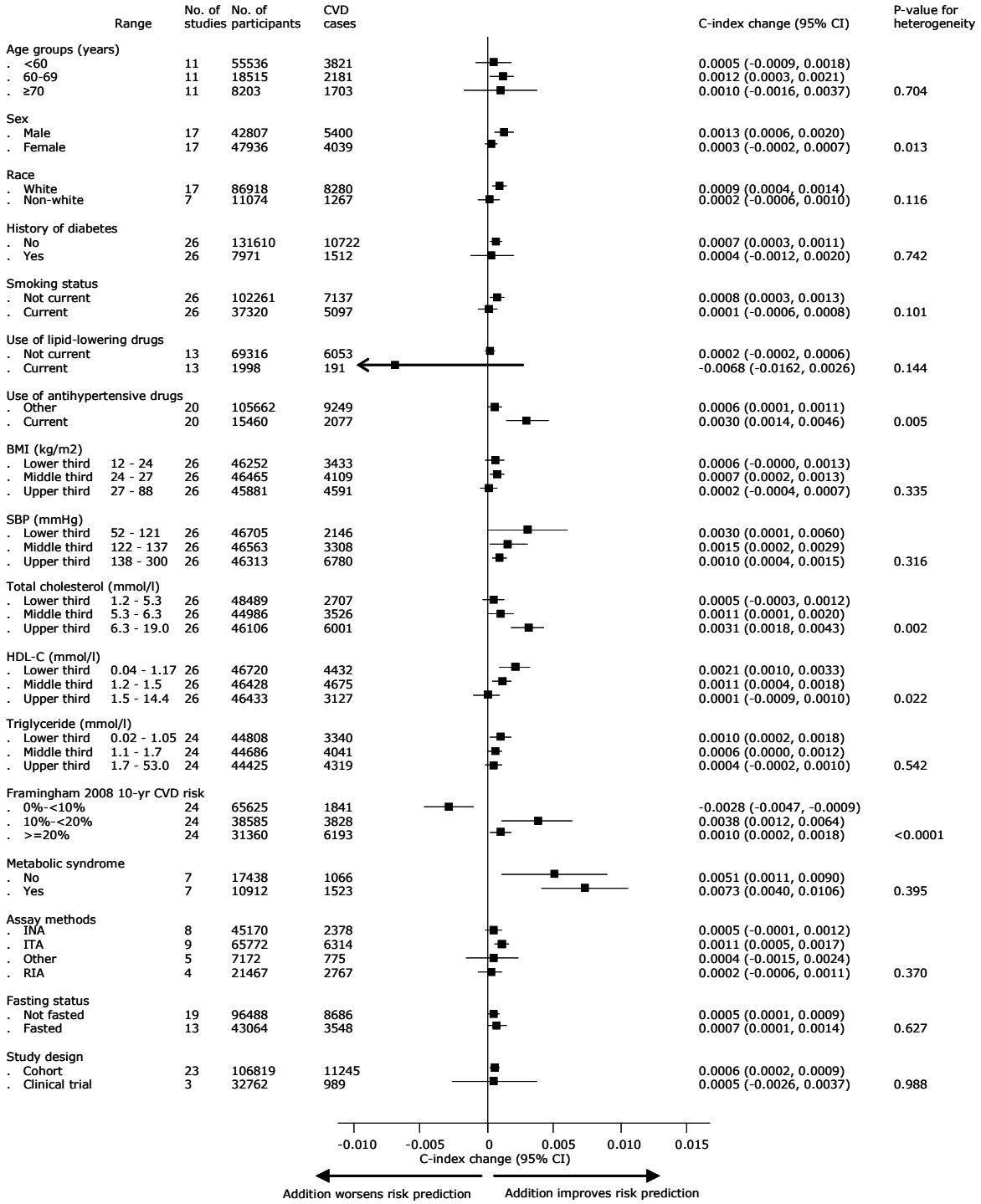

Models contain conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol). Error bars indicate 95% CIs. In each category only studies with information on all subgroup levels are used. Not all studies used had full information across all subgroup levels so comparisons across subgroups (e.g. men vs smokers) may not be reliable due to between study differences. Assay methods abbreviation as in eTable 4.

**eFigure 2.** Changes in risk discrimination for CHD and stroke outcomes when replacing total and HDL cholesterol with different combinations of lipids and apolipoproteins in a model including conventional risk factors



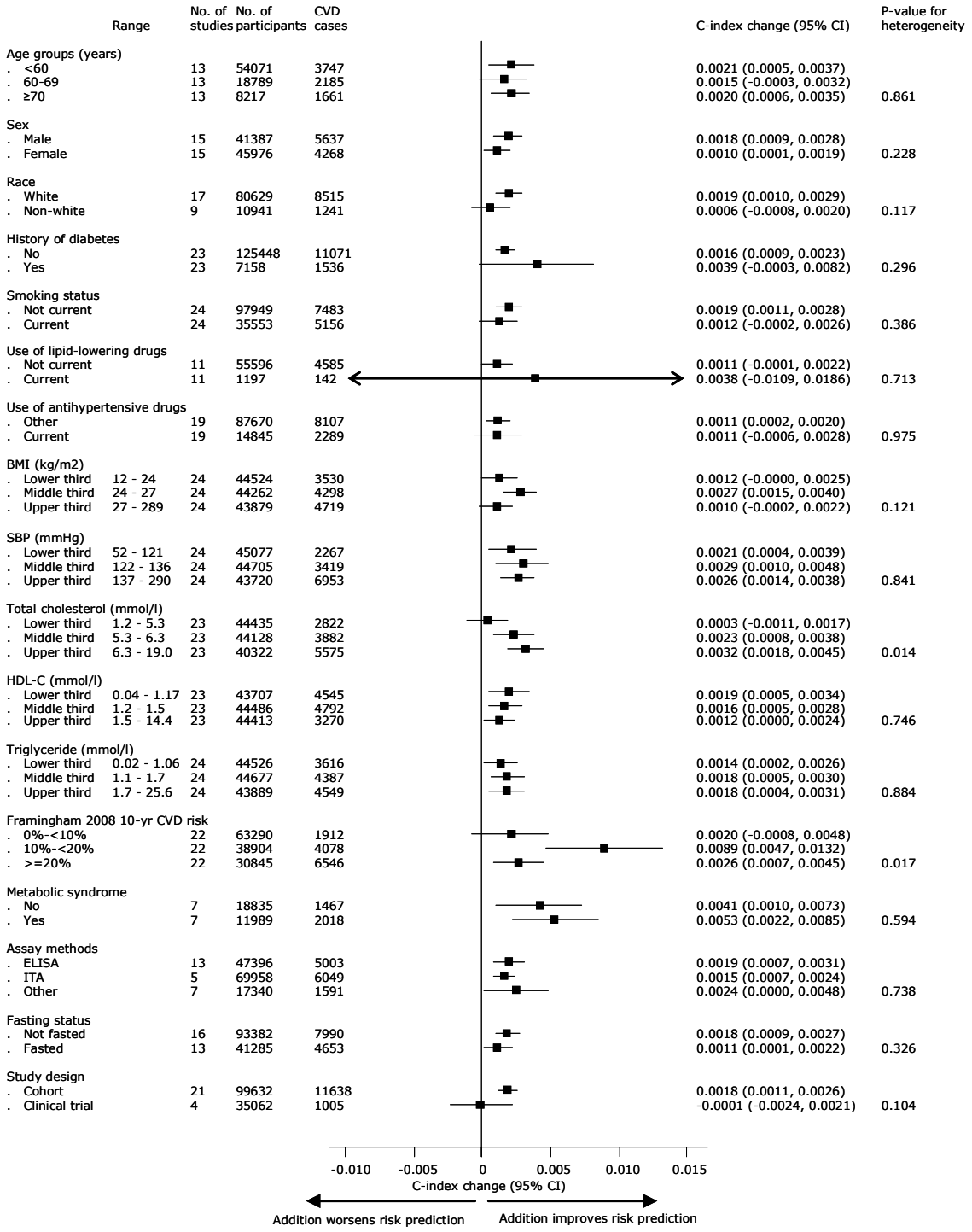

Non-lipid risk factors (RFs) were age, systolic blood pressure, smoking status, history of diabetes, each included as individual linear term. Models were stratified by sex. For changes in the C-index values less than zero indicate worsened predictive ability and values greater than zero improved ability. Error bars indicate 95% CIs.

**eFigure 3.** Changes in C-index for CVD upon addition of apoB and apoAI to conventional risk factors, in different subgroups



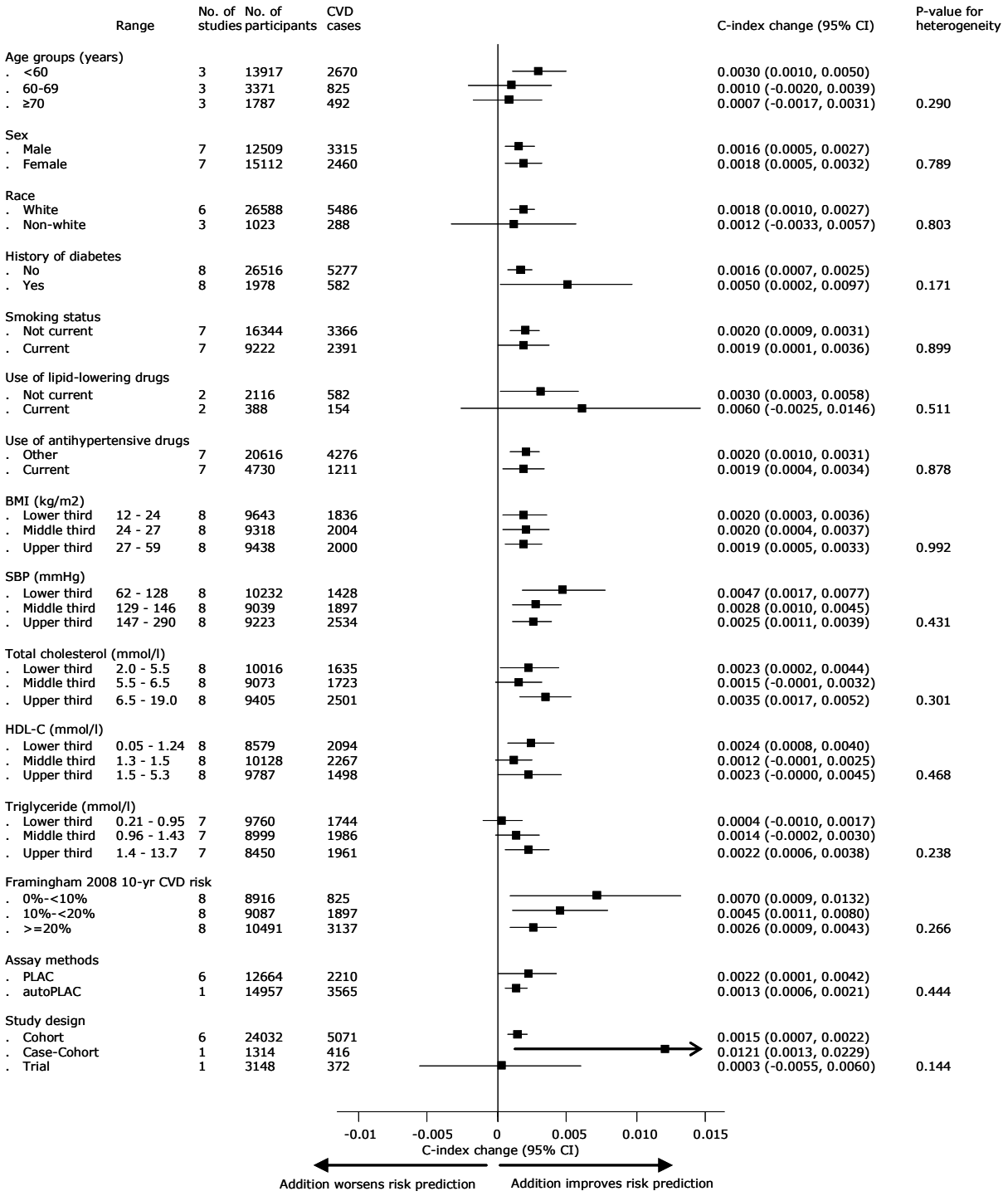

Models contain conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total cholesterol, and HDL cholesterol) with and without inclusion of apoB and apoAI. Error bars indicate 95% CI. In each category only studies with information on both subgroup levels are used. Not all studies used had full information across all subgroup levels so comparisons across subgroups (e.g. smoking status use vs lipid lowering drug use) are not reliable due to inclusion of between study differences. Assay methods abbreviation as in eTable 4.

**eFigure 4.** Changes in C-index for CVD upon addition of Lp(a) to conventional risk factors, in different subgroups



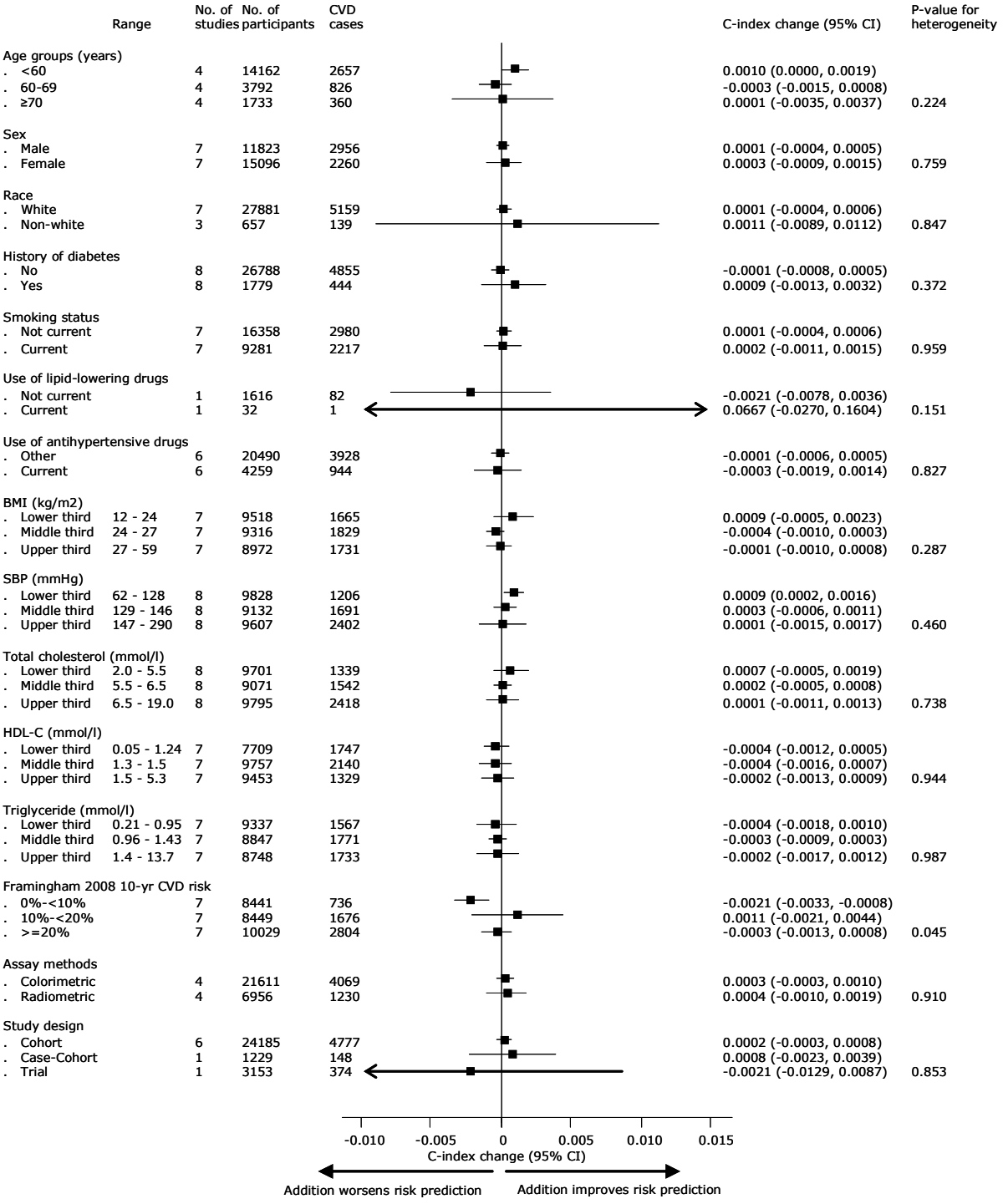

Models contain conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total cholesterol, and HDL cholesterol) with and without inclusion of apoB and apoAI. Error bars indicate 95% CI. In each category only studies with information on both subgroup levels are used. Not all studies used had full information across all subgroup levels so comparisons across subgroups (e.g. smoking status use vs lipid lowering drug use) are not reliable due to inclusion of between study differences. Assay methods abbreviation as in eTable 4.

**eFigure 5.** Change in C-index for CVD upon addition of Lp-$PLA_2$ mass to conventional risk factors, in different subgroups



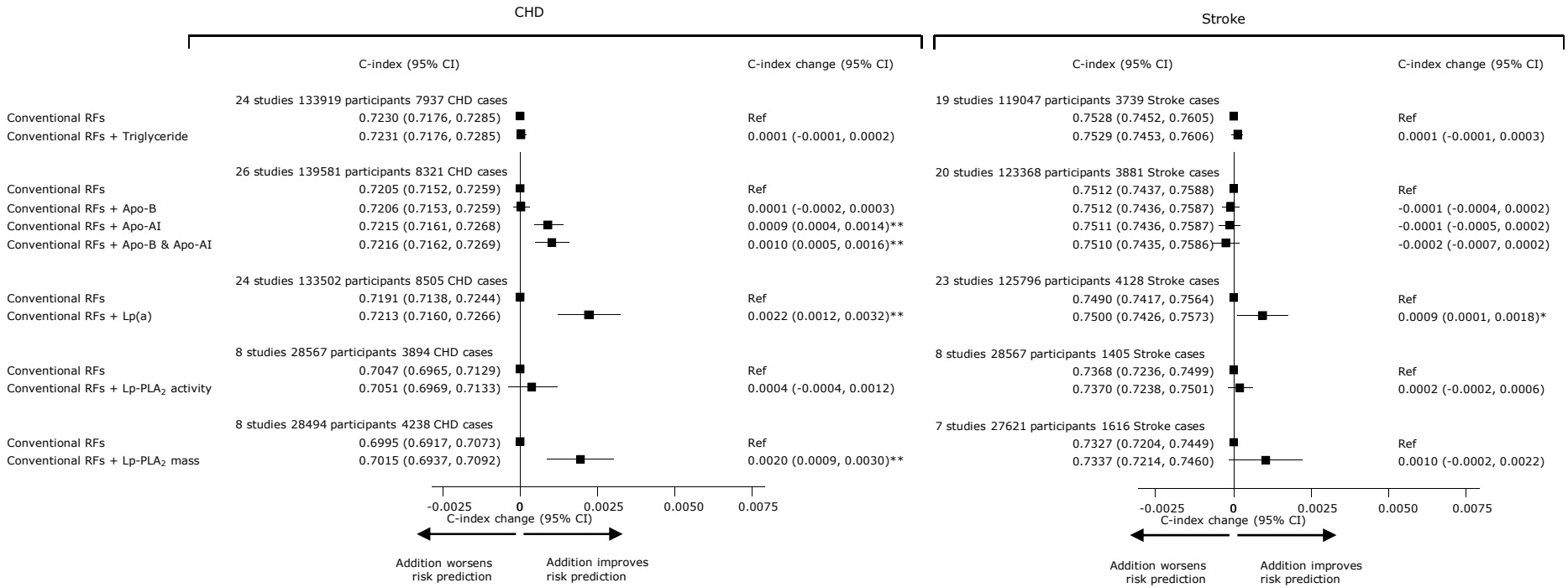

Models contain conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total cholesterol, and HDL cholesterol) with and without inclusion of apoB and apoAI. Error bars indicate 95% CI. In each category only studies with information on both subgroup levels are used. Not all studies used had full information across all subgroup levels so comparisons across subgroups (e.g. smoking status use vs lipid lowering drug use) are not reliable due to inclusion of between study differences. Assay methods abbreviation as in eTable 4.

**eFigure 6.** Change in C-index for CVD upon addition of Lp-$PLA_2$ activity to conventional risk factors, in different subgroups



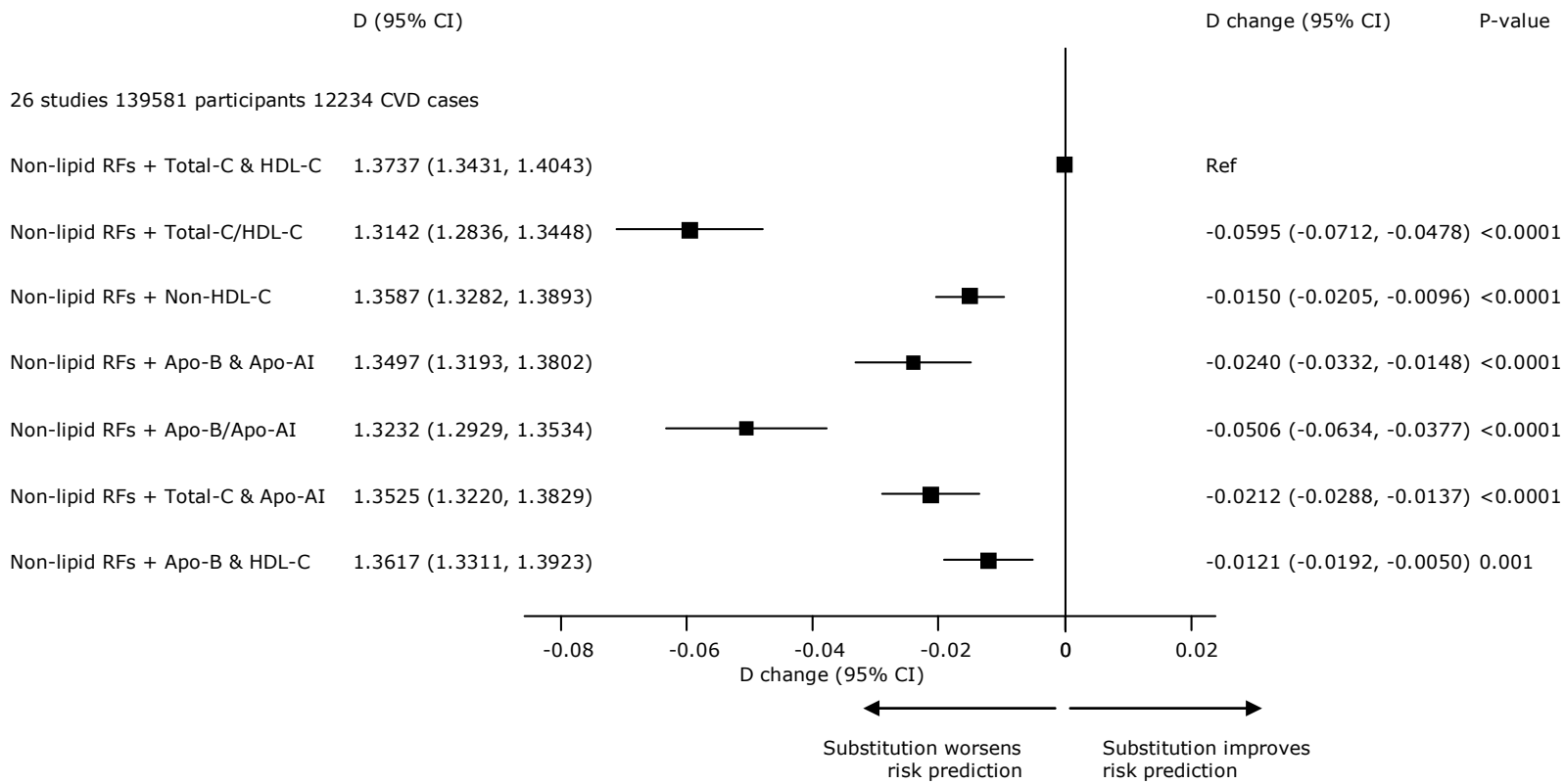

Models contain conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total cholesterol, and HDL cholesterol) with and without inclusion of apoB and apoAI. Error bars indicate 95% CI. In each category only studies with information on both subgroup levels are used. Not all studies used had full information across all subgroup levels so comparisons across subgroups (e.g. smoking status use vs lipid lowering drug use) are not reliable due to inclusion of between study differences. Assay methods abbreviation as in eTable 4.

**eFigure 7.** Changes in risk discrimination for CHD and stroke outcomes after addition of non-traditional lipids parameters to a model including conventional risk factors*



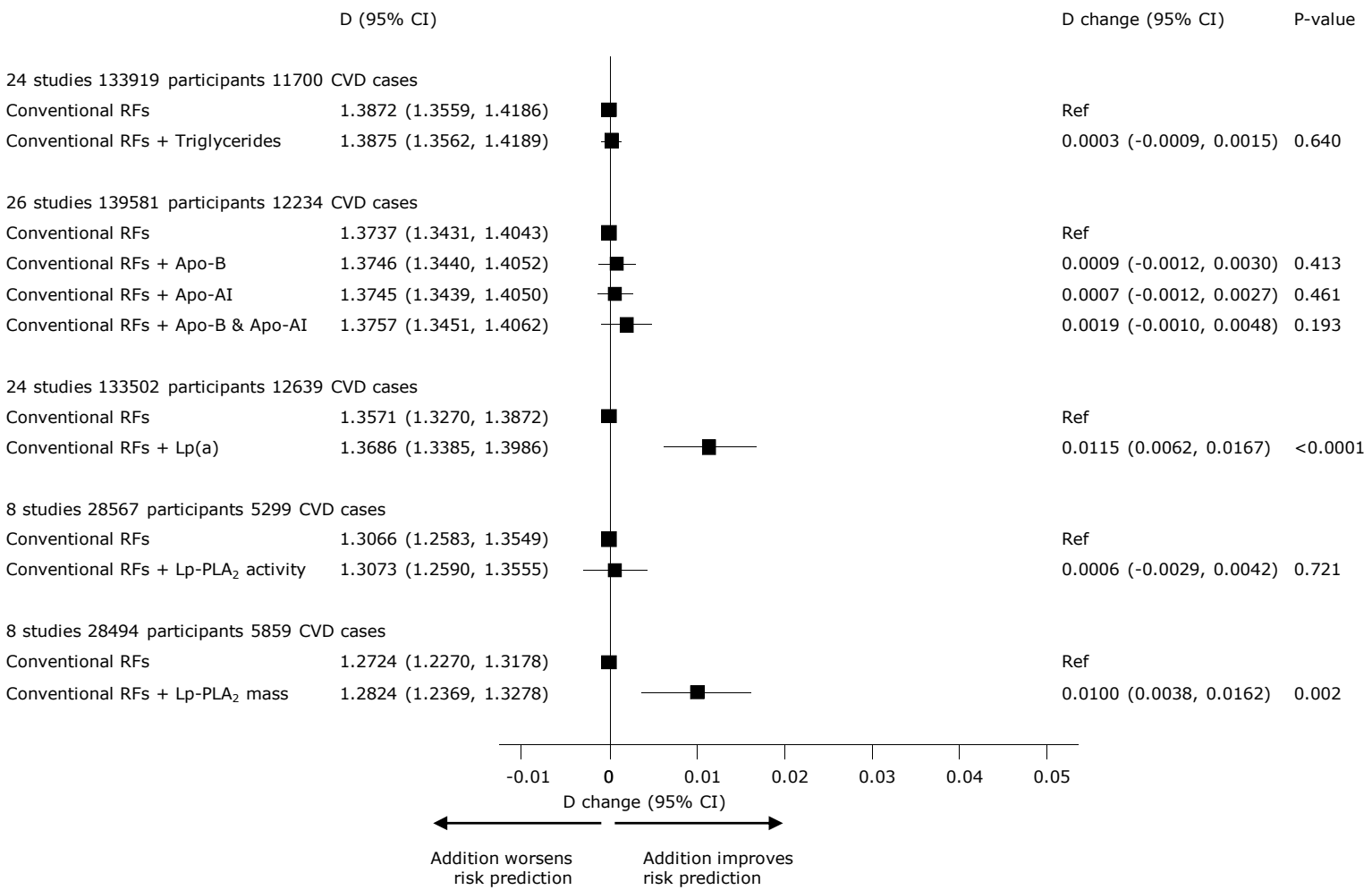

** p<0.001 * p<0.05 for comparison against model containing conventional risk factors (age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol, each included as a linear term). Models were stratified by sex. Triglyceride values were log-transformed. Lp(a) was modelled nonlinearly by including linear and quadratic terms of log-transformed Lp(a). RFs, risk factors.

**eFigure 8.** Changes in D measure for CVD outcomes when replacing total and HDL cholesterol with different combinations of lipids and apolipoproteins in a model including conventional non-lipid risk factors*



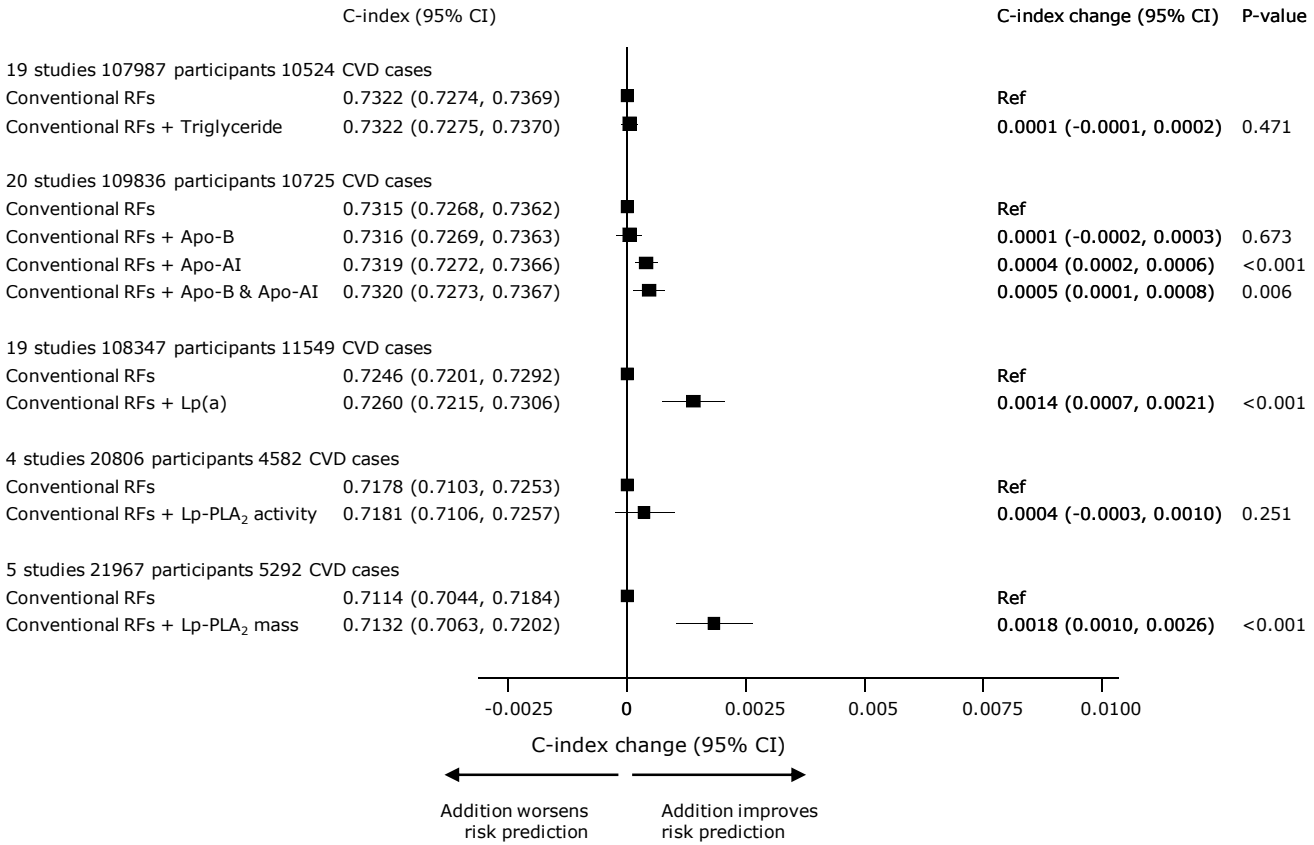

*Non-lipid risk factors (RFs) in each model included age, systolic blood pressure, smoking status, and history of diabetes. Models were stratified by sex.

**eFigure 9.** Changes in D measure for CVD outcomes after addition of lipid-related markers to a model including conventional risk factors*



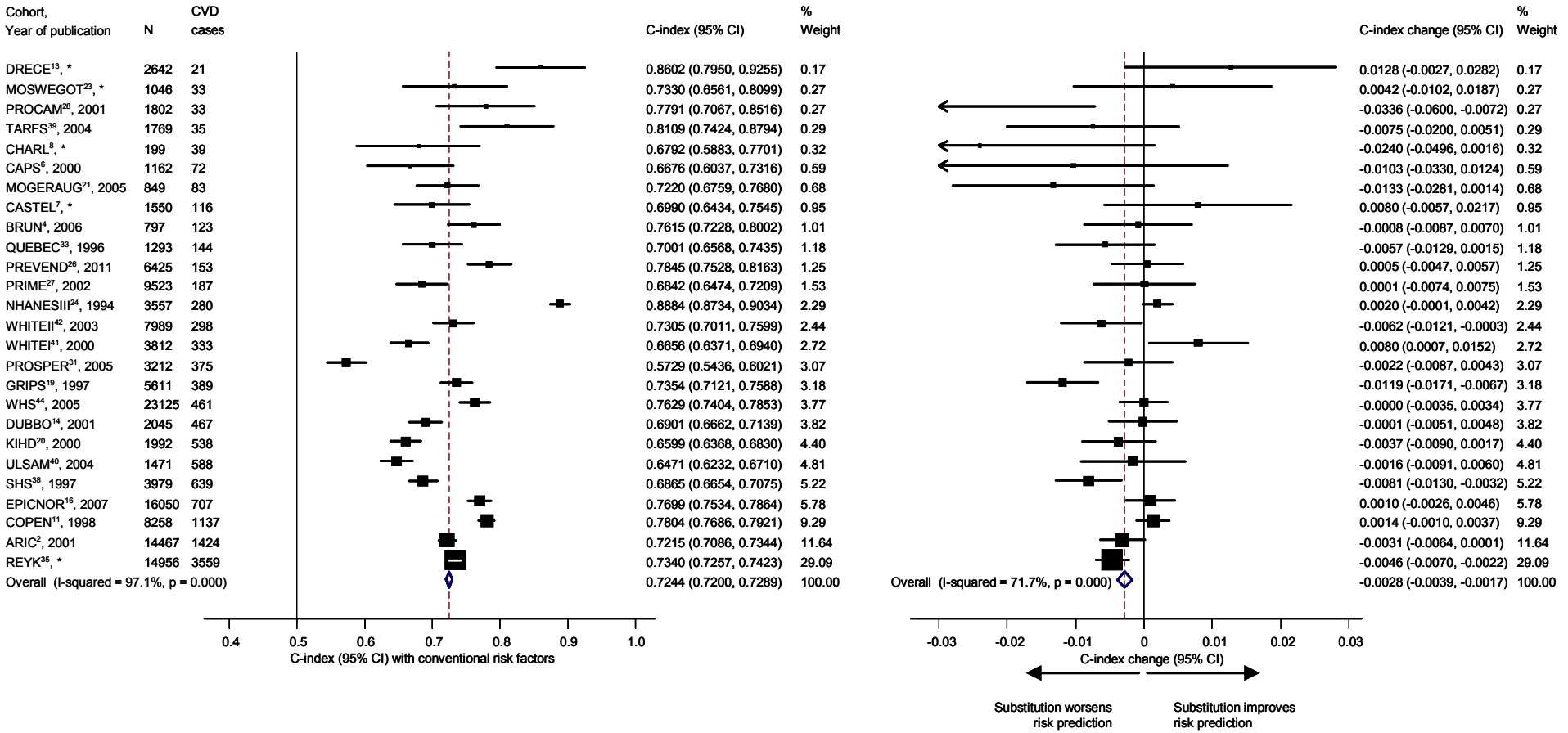

*Conventional risk factors (RFs) in each model included age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol. Models were stratified by sex.

**eFigure 10.** Changes in CVD risk discrimination after addition of lipid-related markers to a model including conventional risk factors*, restricted to studies with at least 10 years of follow-up



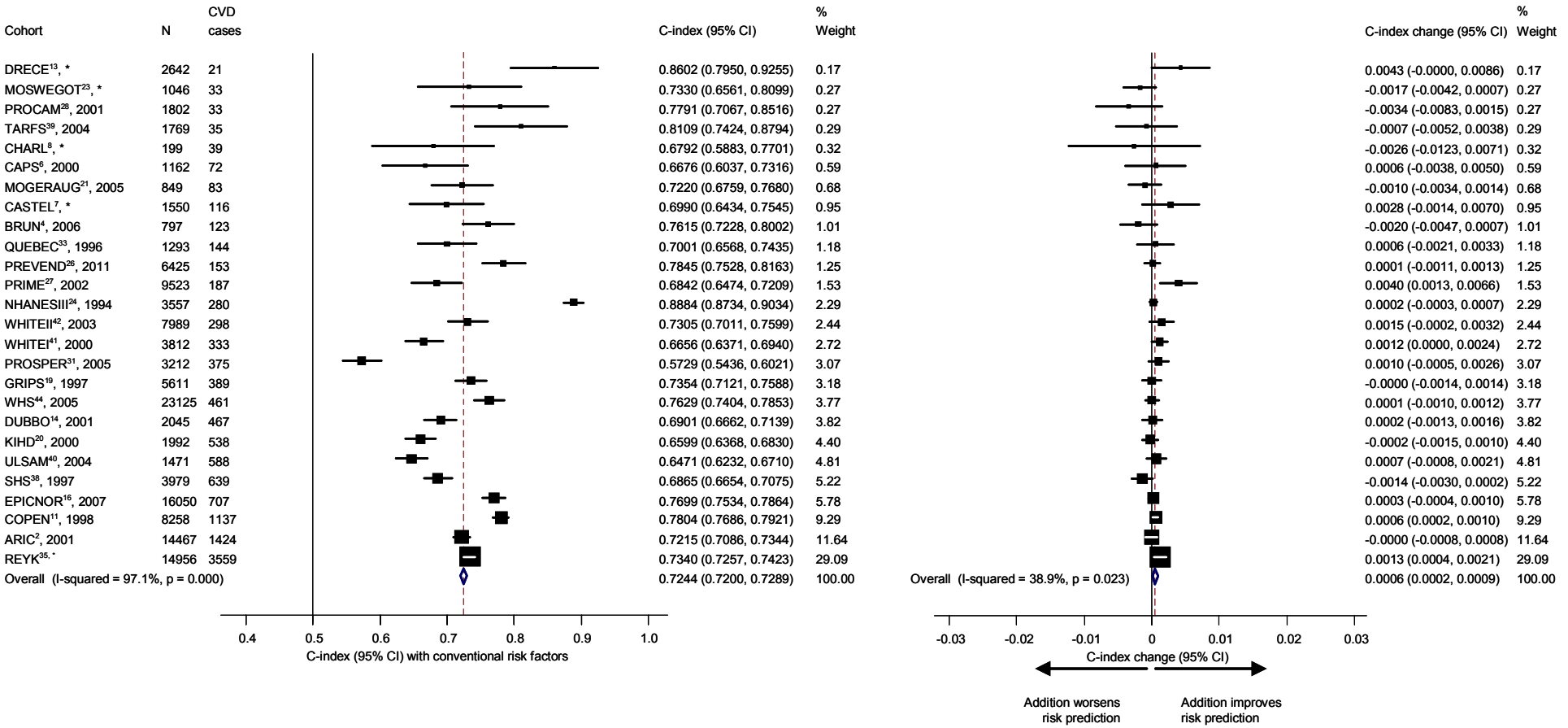

*Conventional risk factors (RFs) in each model included age, systolic blood pressure, smoking status, history of diabetes, total and HDL-cholesterol. Models were stratified by sex.

**eFigure 11.** Study-specific C-index and change in C-index when replacing total and HDL cholesterol with apolipoproteins



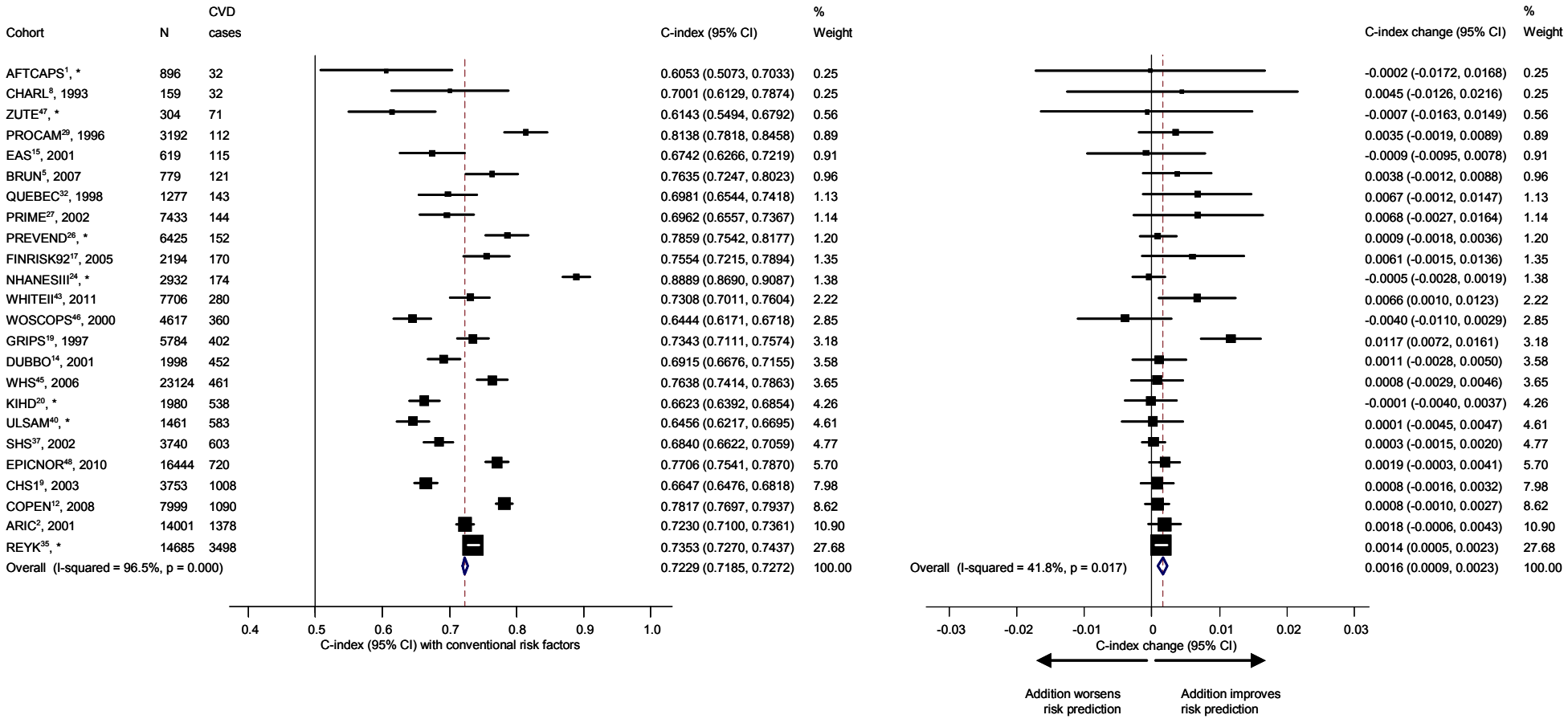

$I^2$ were 97% (95% CI 96%-98%) for the overall C-index and 72% (95% CI 58%-81%) for the C-index change. Sizes of the data markers are proportional to the number of cases in each study.

**eFigure 12.** Study-specific C-index and change in C-index upon addition of apoB and apoAI to conventional risk factors



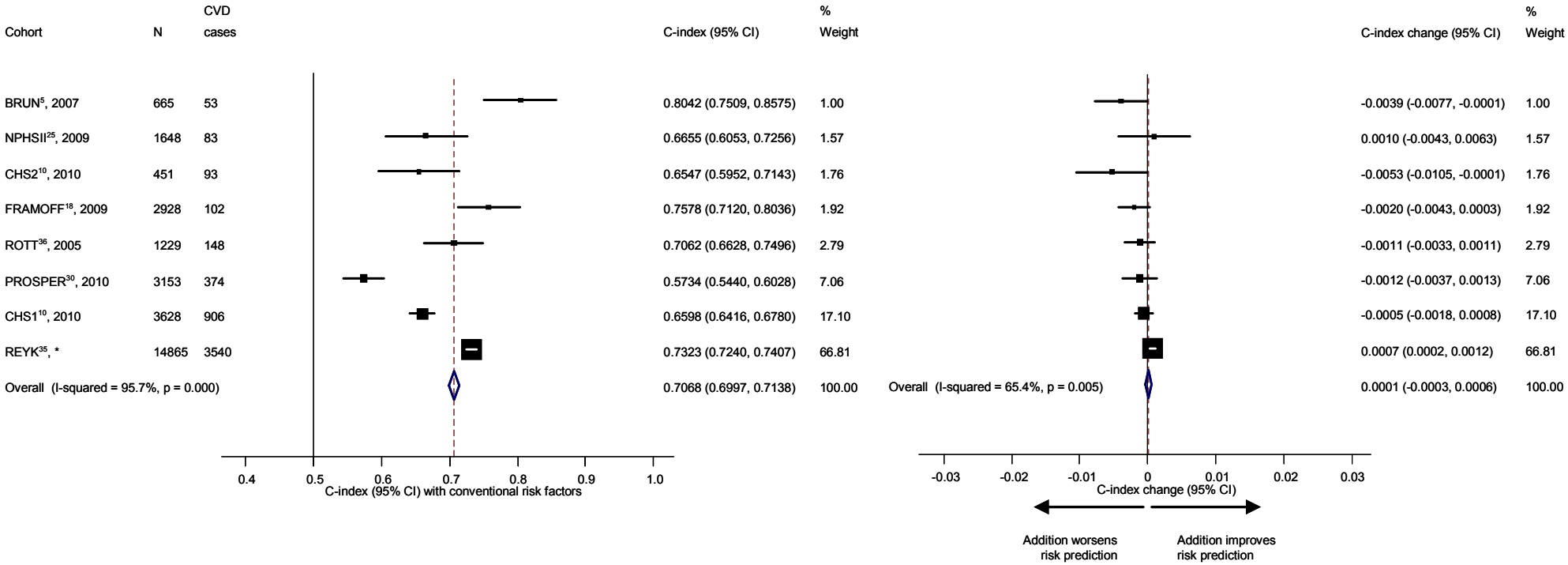

$I^2$ were 97% (95% CI 96%-98%) for the overall C-index and 39% (95% CI 2%-62%) for the C-index change. Sizes of the data markers are proportional to the number of cases in each study.

**eFigure 13.** Study-specific C-index and change in C-index upon addition of Lp(a) to conventional risk factors



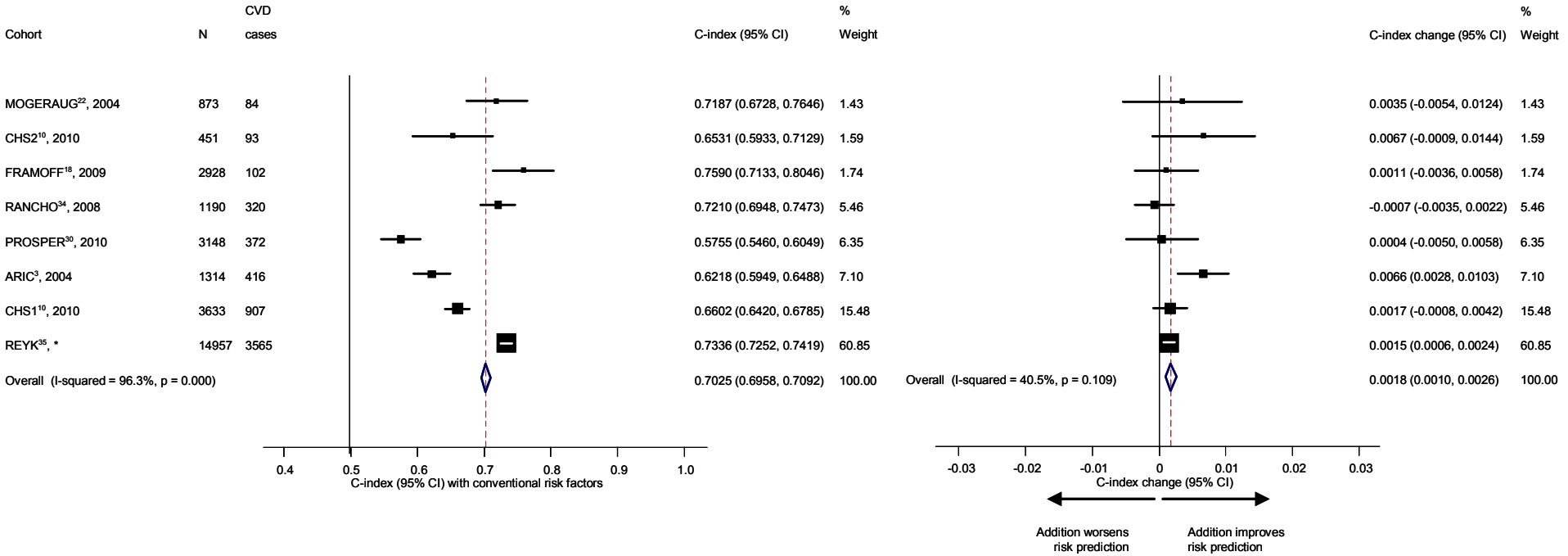

$I^2$ were 96% (95% CI 95%-97%) for the overall C-index and 42% (95% CI 5%-64%) for the C-index change. Sizes of the data markers are proportional to the number of cases in each study.

**eFigure 14.** Study-specific C-index and change in C-index upon addition of Lp-PLA$_2$ activity to conventional risk factors



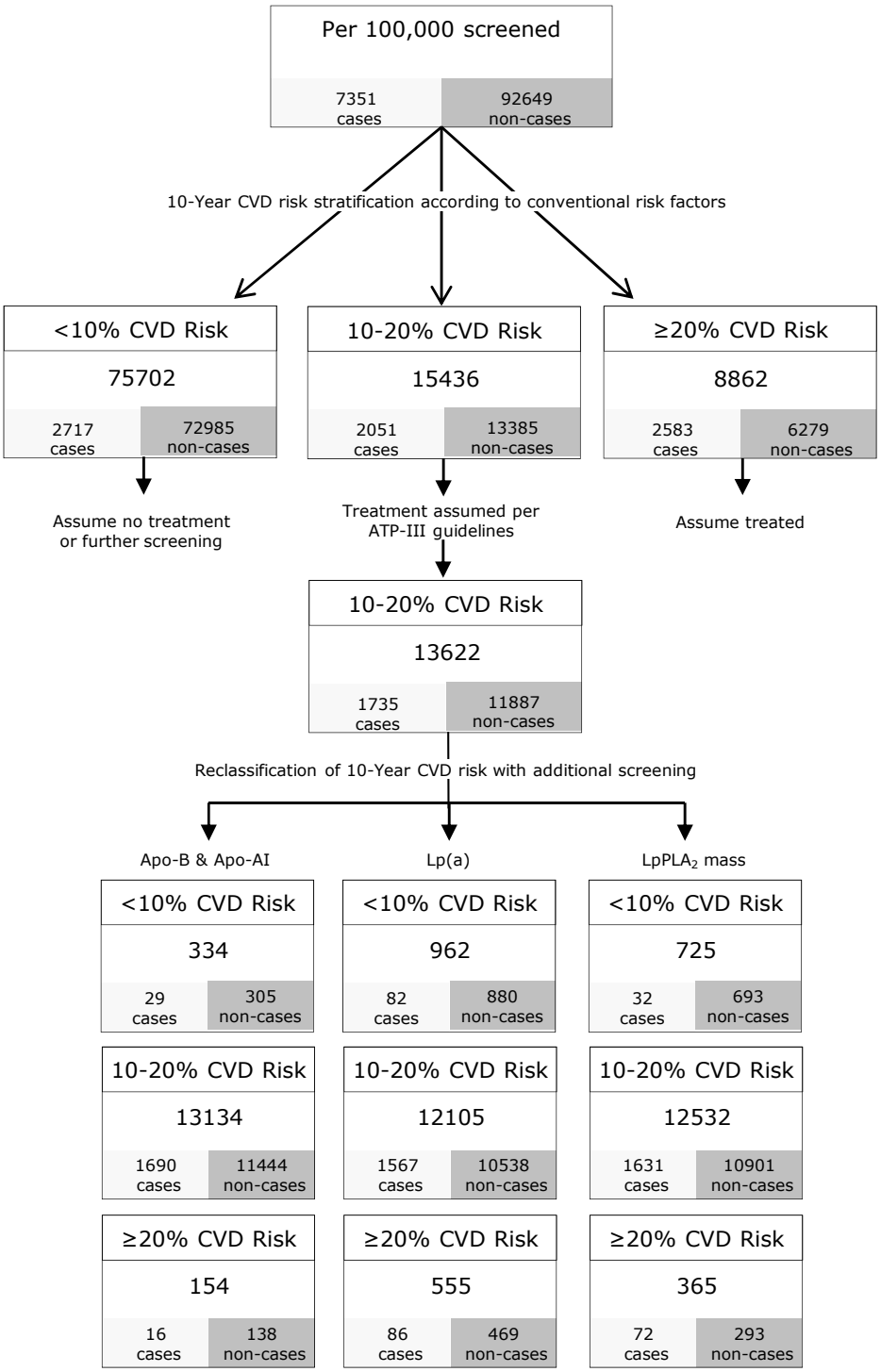

$I^2$ were 96% (95% CI 94%-97%) for the overall C-index and 65% (95% CI 26%-84%) for the C-index change. Sizes of the data markers are proportional to the number of cases in each study.

**eFigure 15.** Study-specific C-index and change in C-index upon addition of Lp-PLA$_2$ mass to conventional risk factors




$I^2$ were 96% (95% CI 94%-98%) for the overall C-index and 41% (95% CI 0%-74%) for the C-index change. Sizes of the data markers are proportional to the number of cases in each study.

**eFigure 16.** Modelling of reclassification per 100,000 people initially screened with conventional risk factors and then additional targeted apolipoproteins, Lp(a) or Lp-$PLA_2$ mass assessment




† Conventional risk factors include: age, smoking status, systolic blood pressure, history of diabetes, total and HDL cholesterol (stratified by sex).
* Following ATP-III guidelines, this model assumes that people who should receive statins are: (1) those at ≥20% predicted 10-year CVD risk and (2) other people (eg, those with diabetes) who merit statins irrespective of predicted 10-year CVD risk. People reporting statin use at "baseline" were also assumed to merit statin allocation.

## eAppendix 1: List of studies' acronyms

**AFTCAPS**, Air Force/Texas Coronary Atherosclerosis Prevention Study[1]
**ARIC**, Atherosclerosis Risk in Communities Study[2;3]
**BRUN**, Bruneck Study[4;5]
**CaPS**, Caerphilly Prospective Study[6]
**CASTEL**, Cardiovascular Study in the Elderly[7]
**CHARL**, Charleston Heart Study[8]
**CHS**, Cardiovascular Health Study[9;10]
**COPEN**, Copenhagen City Heart Study[11;12]
**DRECE**, Diet and Risk of Cardiovascular Disease in Spain[13]
**DUBBO**, Dubbo Study of the Elderly[14]
**EAS**, Edinburgh Artery Study[15]
**EPICNOR,** EPIC Norfolk Study[16, 48]
**FINRISK-92**, Finrisk Cohort 1992[17]
**FRAMOFF**, Framingham Offspring Study[18]
**GRIPS**, Göttingen Risk Incidence and Prevalence Study[19]
**KIHD**, Kuopio Ischaemic Heart Disease Study[20]
**MOGERAUG**, MONICA/KORA Augsburg[21;22]
**MOSWEGOT**, MONICA Göteborg Study[23]
**NHANES III**, Third National Health and Nutrition Examination Survey[24]
**NPHSII**, Northwick Park Heart Study II[25]
**PREVEND**, Prevention of Renal and Vascular End Stage Disease Study[26]
**PRIME**, Prospective Epidemiological Study of Myocardial Infarction[27]
**PROCAM**, Prospective Cardiovascular Münster Study[28;29]
**PROSPER**, Prospective Study of Pravastatin in the Elderly at Risk[30;31]
**QUEBEC**, Quebec Cardiovascular Study[32;33]
**RANCHO**, Rancho Bernardo Study[34]
**REYK**, Reykjavik Study[35]
**ROTT**, The Rotterdam Study[36]
**SHS**, Strong Heart Study[37;38]
**TARFS**, Turkish Adult Risk Factor Study[39]
**ULSAM**, Uppsala Longitudinal Study of Adult Men[40]
**WHITE I**, Whitehall I Study[41]
**WHITE II**, Whitehall II Study[42;43]
**WHS**, Womens Health Study[44;45]
**WOSCOPS**, West of Scotland Coronary Prevention Study[46]
**ZUTE**, Zutphen Elderly Study[47]

## eReferences eTables 1-3 and eAppendix 1


(1) Gotto AM, Jr., Whitney E, Stein EA et al. Relation between baseline and on-treatment lipid parameters and first acute major coronary events in the Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS/TexCAPS). *Circulation* 2000;101:477-484.

(2) Sharrett AR, Ballantyne CM, Coady SA et al. Coronary heart disease prediction from lipoprotein cholesterol levels, triglycerides, lipoprotein(a), apolipoproteins A-I and B, and HDL density subfractions: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation* 2001;104:1108-1113.

(3) Ballantyne CM, Hoogeveen RC, Bang H et al. Lipoprotein-associated phospholipase A2, high-sensitivity C-reactive protein, and risk for incident coronary heart disease in middle-aged men and women in the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 2004;109:837-842.

(4) Mayr M, Kiechl S, Tsimikas S et al. Oxidized Low-Density Lipoprotein Autoantibodies, Chronic Infections, and Carotid Atherosclerosis in a Population-Based Study. *Journal of the American College of Cardiology* 2006;47:2436-2443.

(5) Kiechl S, Willeit J, Mayr M et al. Oxidized Phospholipids, Lipoprotein(a), Lipoprotein-Associated Phospholipase A2 Activity, and 10-Year Cardiovascular Outcomes. *Arterioscler Thromb Vasc Biol* 2007;27:1788-1795.

(6) Sweetnam PM, Bolton CH, Downs LG et al. Apolipoproteins ApoI, ApoII and B, lipoprotein(a) and the risk of ischaemic heart disease: the Caerphilly Study. *European Journal of Clinical Investigation* 2000;30:947-956.

(7) Casiglia E, Palatini P. Cardiovascular risk factors in the elderly. *J Hum Hypertens* 1998;12:575-581.

(8) Knapp RG, Schreiner PJ, Sutherland SE, Keil JE, Gilbert GE, Klein RL, Hames C, Tyroler HA. Serum lipoprotein(a) levels in elderly black and white men in the Charleston Heart Study. *Clin Genet* 1993;44(5):225-31.

(9) Ariyo AA, Thach C, Tracy R. Lp(a) Lipoprotein, Vascular Disease, and Mortality in the Elderly. *New England Journal of Medicine* 2003;349:2108-2115.

(10) Jenny NS, Solomon C, Cushman M et al. Lipoprotein-associated phospholipase A2 (Lp-PLA2) and risk of cardiovascular disease in older adults: Results from the Cardiovascular Health Study. *Atherosclerosis* 2010;209:528-532.

(11) Tybjaerg-Hansen A, Steffensen R, Meinertz H, Schnohr P, Nordestgaard BG. Association of mutations in the apolipoprotein B gene with hypercholesterolemia and the risk of ischemic heart disease. *New England Journal of Medicine* 1998;338:1577-84.

(12) Kamstrup PR, Benn M, Tybjaerg-Hansen A, Nordestgaard BG. Extreme Lipoprotein(a) Levels and Risk of Myocardial Infarction in the General Population. *Circulation* 2008;117:176-184.

(13) Ballesteros-Pomar MD, Rubio-Herrera MA, Gutierrez-Fuentes JA et al. Dietary Habits and Cardiovascular Risk in the Spanish Population: The DRECE Study (I). *Annals of Nutrition and Metabolism* 2000;44:108-114.

(14) Simons LA, Simons J, Friedlander Y, McCallum J. Cholesterol and other lipids predict coronary heart disease and ischaemic stroke in the elderly, but only in those below 70 years. *Atherosclerosis* 2001;159:201-208.

(15) Price JF, Lee AJ, Rumley A, Lowe GDO, Fowkes FG. Lipoprotein (a) and development of intermittent claudication and major cardiovascular events in men and women: The Edinburgh Artery Study. *Atherosclerosis* 2001;157:241-249.

(16) van der Steeg WA, Boekholdt SM, Stein EA, El-Harchaoui K, Stroes ES, Sandhu MS, Wareham NJ, Jukema JW, Luben R, Zwinderman AH, Kastelein JJ, Khaw KT. Role of the apolipoprotein B-apolipoprotein A-I ratio in cardiovascular risk assessment: a case-control analysis in EPIC-Norfolk. *Ann Intern Med* 2007;146:640-8.

(17) Rajecki M, Pajunen P, Jousilahti P, Rasi V, Vahtera E, Salomaa V. Hemostatic factors as predictors of stroke and cardiovascular diseases: the FINRISK '92 Hemostasis Study. *Blood Coagulation & Fibrinolysis* 2005;16.

(18) Schnabel R, Dupuis J, Larson MG et al. Clinical and genetic factors associated with lipoprotein-associated phospholipase A2 in the Framingham Heart Study. *Atherosclerosis* 2009;204:601-607.

(19) Cremer P, Nagel D, Mann H et al. Ten-year follow-up results from the Goettingen Risk, Incidence and Prevalence Study (GRIPS). I. Risk factors for myocardial infarction in a cohort of 5790 men. *Atherosclerosis* 1997;129:221-230.

(20) Lakka HM, Lakka TA, Tuomilehto J, Sivenius J, Salonen JT. Hyperinsulinemia and the Risk of Cardiovascular Death and Acute Coronary and Cerebrovascular Events in Men: The Kuopio Ischaemic Heart Disease Risk Factor Study. *Arch Intern Med* 2000;160:1160-1168.

(21) Meisinger C, Loewel H, Mraz W, Koenig W. Prognostic value of apolipoprotein B and A-I in the prediction of myocardial infarction in middle-aged men and women: results from the MONICA/KORA Augsburg cohort study. *European Heart Journal* 2005;26:271-278.

(22) Koenig W, Khuseyinova N, Löwel H, Trischler G, Meisinger C. Lipoprotein-Associated Phospholipase A2 Adds to Risk Prediction of Incident Coronary Events by C-Reactive Protein in Apparently Healthy Middle-Aged Men From the General Population. *Circulation* 2004;110:1903-1908.

(23) Wilhelmsen L, Johansson S, Rosengren A, Wallin I, Dotevall A, Lappas G. Risk factors for cardiovascular disease during the period 1985-1995 in Goteborg, Sweden. The GOT-MONICA Project. *J Intern Med* 1997;242:199-211.

(24) Bachorik PS, Lovejoy KL, Carroll MD, Johnson CL, Albers JJ, Marcovina SM. Measurement of apolipoproteins A-I and B during the National Health and Nutrition Examination Survey (NHANES) III. . *Clin Chem* 1994;40:1915-20.

(25) Drenos F, Talmud PJ, Casas JP et al. Integrated associations of genotypes with multiple blood biomarkers linked to coronary heart disease risk. *Human Molecular Genetics* 2009;18:2305-2316.

(26) Kappelle PJ, Gansevoort RT, Hillege JL, Wolffenbuttel BH, Dullaart RP. Apolipoprotein B/A-I and total cholesterol/high-density lipoprotein cholesterol ratios both predict cardiovascular events in the general population independently of nonlipid risk factors, albuminuria and C-reactive protein. *J Intern Med* 2011;269:232-242.

(27) Luc G, Bard JM, Ferrieres J et al. Value of HDL Cholesterol, Apolipoprotein A-I, Lipoprotein A-I, and Lipoprotein A-I/A-II in Prediction of Coronary Heart Disease: The PRIME Study. *Arterioscler Thromb Vasc Biol* 2002;22:1155-1161.

(28) von Eckardstein A, Fischer F, Schulte H, Tataru M, Köhler E, Assmann G. Association of serum apolipoprotein A-I (but not high-density lipoprotein cholesterol) with healed myocardial infarction in men independent of serum insulin and C-peptide. *Am J Cardiol* 2001;88(7):723-6.

(29) Assmann G, Schulte H, von Eckardstein A, Huang Y. High-density lipoprotein cholesterol as a predictor of coronary heart disease risk. The PROCAM experience and pathophysiological implications for reverse cholesterol transport. *Atherosclerosis* 1996;124:S11-S20.

(30) Caslake MJ, Packard CJ, Robertson M et al. Lipoprotein-associated phospholipase A2, inflammatory biomarkers, and risk of cardiovascular disease in the Prospective Study of Pravastatin in the Elderly at Risk (PROSPER). *Atherosclerosis* 2010;210:28-34.

(31) Packard CJ, Ford I, Robertson M et al. Plasma lipoproteins and apolipoproteins as predictors of cardiovascular risk and treatment benefit in the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER). *Circulation* 2005;112:3058-3065.

(32) Cantin B, Gagnon F, Moorjani S et al. Is Lipoprotein(a) an Independent Risk Factor for Ischemic Heart Disease in Men? The Quebec Cardiovascular Study. *Journal of the American College of Cardiology* 1998;31:519-525.

(33) Lamarche B, Moorjani S, Lupien PJ et al. Apolipoprotein A-I and B Levels and the Risk of Ischemic Heart Disease During a Five-Year Follow-up of Men in the Quebec Cardiovascular Study. *Circulation* 1996;94:273-278.

(34) Daniels LB, Laughlin GA, Sarno MJ, Bettencourt R, Wolfert RL, Barrett-Connor E. Lipoprotein-Associated Phospholipase A2 Is an Independent Predictor of Incident Coronary Heart Disease in an Apparently Healthy Older Population: The Rancho Bernardo Study. *Journal of the American College of Cardiology* 2008;51:913-919.

(35) Jonsdottir LS, Sigfusson N, Gudnason V, Sigvaldason H, Thorgeirsson G. Do lipids, blood pressure, diabetes, and smoking confer equal risk of myocardial infarction in women as in men? The Reykjavik Study. *J Cardiovasc Risk* 2002;9:67-76.

(36) Oei HH, van der Meer IM, Hofman A et al. Lipoprotein-Associated Phospholipase A2 Activity Is Associated With Risk of Coronary Heart Disease and Ischemic Stroke. *Circulation* 2005;111:570-575.

(37) Wang W, Hu D, Lee ET et al. Lipoprotein(a) in American Indians is Low and Not Independently Associated with Cardiovascular Disease: The Strong Heart Study. *Annals of Epidemiology* 2002;12:107-114.

(38) Gray RS, Robbins DC, Wang W, Yeh JL, Fabsitz RR, Cowan LD, Welty TK, Lee ET,Krauss RM, Howard BV. Relation of LDL size to the insulin resistance syndrome and coronary heart disease in American Indians. The Strong Heart Study. *Arterioscler Thromb Vasc Biol* 1997;17(11):2713-20.

(39) Onat A, Yazici M, Can G, Sniderman A. Evidence for a complex risk profile in obese postmenopausal Turkish women with hypertriglyceridaemia and elevated apolipoprotein B. *Clin Sci (Lond)* 2004;107:97-104.

(40) Dunder K, Lind L, Zethelius B, Berglund L, Lithell H. Evaluation of a scoring scheme, including proinsulin and the apolipoprotein B/apolipoprotein A1 ratio, for the risk of acute coronary events in middle-aged men: Uppsala Longitudinal Study of Adult Men (ULSAM). *Am Heart J* 2004;148:596-601.

(41) Clarke R, Breeze E, Youngman L et al. Re-survey of the Whitehall study of London civil servants: changes in risk factors for cardiovascular disease during 29 years of follow-up. *J Cardiovasc Risk* 2000;7:251-257.

(42) Brunner EJ, Marmot MG, White IR et al. Gender and employment grade differences in blood cholesterol, apolipoproteins and haemostatic factors in the Whitehall II study. *Atherosclerosis* 1993;102:195-207.

(43) Zabaneh D, Kumari M, Sandhu M et al. Meta analysis of candidate gene variants outside the LPA locus with Lp(a) plasma levels in 14,500 participants of six White European cohorts. *Atherosclerosis* 2011;217(2)447-51.

(44) Ridker PM, Rifai N, Cook NR, Bradwin G, Buring JE. Non-HDL Cholesterol, Apolipoproteins A-I and B100, Standard Lipid Measures, Lipid Ratios, and CRP as Risk Factors for Cardiovascular Disease in Women. *JAMA* 2005;294:326-333.

(45) Suk Danik J, Rifai N, Buring JE, Ridker PM. Lipoprotein(a), Measured With an Assay Independent of Apolipoprotein(a) Isoform Size, and Risk of Future Cardiovascular Events Among Initially Healthy Women. *JAMA* 2006;296:1363-1370.

(46) Gaw A, Brown EA, Docherty G, Ford I. Is lipoprotein(a)-cholesterol a better predictor of vascular disease events than total lipoprotein(a) mass? A nested case control study from the West of Scotland Coronary Prevention Study. *Atherosclerosis* 2000;148:95-100.

(47) Weijenberg MP, Feskens EJM, Kromhout D. Total and High Density Lipoprotein Cholesterol as Risk Factors for Coronary Heart Disease in Elderly Men during 5 Years of Follow-up: The Zutphen Elderly Study. *Am J Epidemiol* 1996;143:151-158.

(48) Tsimikas S, Mallat Z, Talmud PJ, Kastelein JJ, Wareham NJ, Sandhu MS, Miller ER, Benessiano J, Tedgui A, Witztum JL, Khaw KT, Boekholdt SM. Oxidation-specific biomarkers, lipoprotein(a), and risk of fatal and nonfatal coronary events. *J Am Coll Cardiol* 2010;56(12):946-55..

**eAppendix 2:** Statistical Methods

**Formulation of the risk prediction models**

The risk prediction models were based on a Cox proportional hazards (PH) model,[1] stratified by study, sex, and (where applicable) randomised trial arm. So, for each stratum $k = 1, 2, \cdots, K$ (i.e. $K$ distinct combinations of study, sex, and trial arm), with $i = 1, 2, \cdots, n_k$ individuals in stratum $k$ and the $i^{th}$ individual having baseline covariate values $\mathbf{X}_i = (x_{i1}, x_{i2}, \cdots x_{ip})$, the the probability of surviving beyond $t$ years after baseline without a cardiovascular disease (CVD) event was modelled as

$$S(t \mid \mathbf{X}_i, k) = S_{0k}(t)^{\exp(\boldsymbol{\beta}\mathbf{X}_i')} \qquad (1)$$

and the probability of a CVD event within $t$ years after baseline was simply one minus the survival probability, i.e.

$$\Pr(T \le t \mid \mathbf{X}_i, k) = 1 - S(t \mid \mathbf{X}_i, k) = 1 - S_{0k}(t)^{\exp(\boldsymbol{\beta}\mathbf{X}_i')} \qquad (2)$$

Deaths from non-CVD causes were considered censored observations, so the above probability refers to the risk of a CVD event in the absence of deaths from other causes.

In (1) and (2) the evolution of risk over time was modelled independently for each stratum (i.e. study, sex, and trial arm), as represented by the non-parametric baseline survival $S_{0k}(t)$. This avoids having to make the potentially inappropriate assumption of proportional hazards across studies and sexes, and also prevents these characteristics from contributing to the measures of predictive ability (see below). For sex this is desirable in the multi-study setting, since by design studies are likely to differ in the proportion of males and females included, and could lead to heterogeneous results in the predictive ability of sex across studies due to design features but otherwise not real.

Parallel analyses involved a multivariate random-effects meta-analysis approach,[2, 3] that allowed for between study heterogeneity in estimating the overall log hazard ratios $\boldsymbol{\beta}$ in (1), thereby relaxing the common effects assumption. The random effects model yielded similar point estimates for $\boldsymbol{\beta}$ but with wider confidence intervals. Since only the point estimates were necessary for making the absolute risk predictions and calculating measures of discrimination, the simpler stratified Cox PH model assuming common effects was used for derivation of the risk prediction models.

**Covariates assessed**

The core group of risk factors $\mathbf{X}_i$ were age, systolic blood pressure, smoking status, and history of diabetes (i.e. non-blood based covariates), which was extended to include either the separate or combined effects of blood-based covariates, including: total cholesterol, HDL cholesterol, and/or other lipid-related markers and C-reactive protein. All continuous covariates, except Lp(a), were modelled as linear terms in the risk prediction model and dummy variable indicators were used for categorical variables. Lp(a) was modelled by including linear and quadratic terms of log-tranformed Lp(a).

**Evaluating predictive ability**

The Cox PH models were fitted to data from all participants and then predictive ability was assessed using measures of risk discrimination and reclassification as described below. We did not employ cross-validation approaches to these analyses (e.g. performing model derivation and validation in separate subsets of participants) because the dataset used was of substantial size. The latter minimises the chance of overfitting, a phenomenon which can occur when sample sizes are small and leads to optimism in estimates of predictive ability. Phillips et al[4] suggest that in a validation sample, the regression coefficients $\hat{\boldsymbol{\beta}}$ are approximately attenuated to $\kappa\hat{\boldsymbol{\beta}}$ where $\kappa = 1-(q-2)/\chi^2$, $q$ is the number of parameters in the model, and $\chi^2$ is the likelihood ratio Chi-square statistic for the model. In our modelling the $\kappa$ values ranged from 0.999 to 1.000, implying overall negligible bias in regression coefficients from overfitting.

**Assessment of discrimination**

Discrimination was assessed using the C-index[5] and the D measure.[6] The C-index is the probability that, for a randomly selected pair of participants, the individual with the shorter survival time has the higher value of the linear predictor $\hat{\boldsymbol{\beta}}\mathbf{X}_i'$ (corresponding to the worse prognosis).[5] It is estimated by examining all possible pairs of participants for which the participant who has the shorter participation time fails. It classifies each pair as concordant (matching in rank according to the magnitude of the linear predictor and the order of failure), discordant (opposite in rank) or undecided (tied in either category). The overall measure is calculated as

$$C = \frac{n_c + 0.5n_u}{n_c + n_d + n_u} \qquad (3)$$

where $n_c$, $n_d$, and $n_u$ are the number of concordant, discordant, and undecided pairs respectively. In our case, the pair-wise comparisons were constrained to allow only pairing of

participants within the same strata, meaning the concordance/discordance counts did not include comparison of males to females nor participants from different studies.

The measure D is computed by first transforming each participant's linear predictor $\hat{\boldsymbol{\beta}}\mathbf{X}_i'$ from the fitted Cox PH model to give standard normal order rank statistics (rankits - formed using Blom's approximation).[6] In our analyses the standard normal rank order statistics were formed within studies to avoid potential influences from between study differences in covariate distributions. The rank statistics are divided by a factor of $\sqrt{\pi/8}$ to give $z_i$, and a second stratified Cox PH model is then fitted to these values; D is the regression coefficient of $z$ from this second model. The distribution of the transformed variable $z$ is approximately $N(0,\pi/8)$, which has the property that the mean of the negative and positive $z$ values is -0.5 and 0.5 respectively (i.e. 1 unit apart). Hence, the regression coefficient of the continuous $z$ in the Cox PH model can be interpreted as the log hazard ratio for "high" predicted risk versus "low" predicted risk groups.

*Combining measures of discrimination across studies*

A two-stage procedure was used to calculate overall discrimination measures across studies. In an initial step the C-index and D measure were calculated within each study, denoted by, say, $\hat{\theta}_s$ with variance $\hat{\sigma}_s^2$, $s=1,2,\cdots,S$. The variance of the C-index was estimated using an efficient jackknife approach for rank statistics,[7] which is computationally faster than bootstrapping[8] especially for large datasets.[7, 9-11] The variance of the D-measure was the standard maximum likelihood estimate of the variance of a log hazard ratio in Cox PH regression. In the second step, the estimated C-indices and D-measures and corresponding variances were each combined across studies using a weighted average:

$$\hat{\theta}=\frac{\sum w_s\hat{\theta}_s}{\sum w_s} \quad \text{and} \quad \hat{\sigma}_{\hat{\theta}}=\sqrt{\frac{\sum w_s^2\hat{\sigma}_s^2}{\left(\sum w_s\right)^2}} \qquad (4,5)$$

where $s=1,2,\cdots,S$ studies and $w_s$ is the weight applied to each study's estimate, which in our case was chosen as the number of CVD events observed in each study. Alternative weights were also considered, including standard inverse-variance weights in fixed and random effects meta-analysis models.[3] However, weighting by the number of events in a study was considered the most appropriate, as it best matches the weighting applied across studies in derivation of the original stratified Cox PH prediction model (1). This scheme also ensured consistency between the weights given to the C-index and D for each study, and between the difference of the pooled model specific $\hat{\theta}$ estimates (for example $\hat{\theta}_{\text{model2}}-\hat{\theta}_{\text{model1}}$) and the result obtained by pooling the within study differences $\hat{\Delta}_s$, as described in the following section. Between study heterogeneity in risk discrimination measures was estimated by calculating the

$Q$ statistic for testing heterogeneity (by comparison to $\chi^2_{S-1}$ distribution) and its corresponding transformation to the $I^2$ statistic for quantifying the extent of heterogeneity

$$Q = \sum \frac{1}{\hat{\sigma}_s^2}(\hat{\theta}_s - \hat{\theta})^2 \quad \text{and} \quad I^2 = \frac{Q-(S-1)}{Q} \times 100\% \qquad (6,7)$$

where $s = 1, 2, \cdots, S$ studies. Confidence intervals for the $I^2$ statistic were calculated as recommended by Higgins and Thompson.[12]

*The change in measures of discrimination on addition of a new marker*

To investigate the change in measures of discrimination on addition of a new risk factor, two risk prediction models were fitted, one model with the core risk factors only (e.g. non-blood based risk factors and conventional lipids), and the second model with the core risk factors plus the new risk factor (e.g. other lipid-related markers). The within-study $\hat{\theta}_s$ were estimated for both models, and the within-study difference simply calculated and denoted as $\hat{\Delta}_s$. The standard error for the within-study change in C-index was directly estimable using the efficient jackknife procedure for rank statistics.[7, 9-11] Standard error estimates for the change in D were obtained within each study using non-parametric bootstrapping,[13] repeated 200 times in random samples drawn within strata as previously described.[8] These within-study differences were then combined using the meta-analysis procedure above, weighting by the number of CVD events with its variance and between study heterogeneity similarly derived as (4,5) and (6,7) above, respectively:

$$\hat{\Delta} = \frac{\sum w_s \hat{\Delta}_s}{\sum w_s} \quad \text{and} \quad \hat{\sigma}_{\hat{\Delta}} = \sqrt{\frac{\sum w_s^2 \hat{\sigma}^2_{\hat{\Delta}s}}{(\sum w_s)^2}} \qquad (8,9)$$

$$Q = \sum \frac{1}{\hat{\sigma}^2_{\hat{\Delta}s}}(\hat{\Delta}_s - \hat{\Delta})^2 \quad \text{and} \quad I^2 = \frac{Q-(S-1)}{Q} \times 100\% \qquad (10,11)$$

where $s = 1, 2, \cdots, S$ studies.

**Assessment of reclassification**

For reclassification, 10-year CVD risk predictions were necessary, which were calculated only in studies that had had some participants followed up for >10-years (since otherwise the baseline survival $S_{0k}(10)$ was not estimable). Furthermore, among the studies with ≥10 years follow-up, risk classification was investigated only among individuals with known CVD status at 10 years after baseline (i.e. ignoring those who were censored before 10 years),[14] since the methods for calculating reclassification measures while allowing for inclusion of censored observations have only been recently proposed in a single study setting.[15] We are in the process of adapting such methods for use in a multi-study setting, which requires consideration of several as yet unclear issues, including appropriate weighting of the measures while combining across studies. The reference risk prediction model included conventional risk

factors (age, systolic blood pressure, smoking status, history of diabetes, total cholesterol, and HDL cholesterol) which was then extended to additionally include novel risk markers as the alternative model. Using the estimated parameters from each model, the 10-year risk predictions were made according to equation (2). The individuals were then classified into three standard 10-year risk categories: 0% to <10%, 10% to <20% risk and ≥20% risk according to each risk prediction model. Cross-tabulation of the risk categories was stratified by CVD status at 10 years and the reclassification of individuals between risk categories was examined, and deemed "appropriate" for CVD cases moving up the risk categories and for individuals without a CVD event moving down the risk categories on addition of other markers. Reclassification was summarised using the Net Reclassification Improvement (NRI)[14] and the Integrated Discrimination Improvement (IDI).[14]

*Estimating the impact of sequential screening with other markers on a hypothetical population*

To express our findings in a more clinically accessible manner, we used the information observed in the reclassification tables to generalize our findings to the context of population screening, thereby quantifying the potential impact of addition of other markers into current risk prediction algorithms. We specifically modelled the case of sequential screening, in which screening and interventions are targeted to clinically relevant subgroups, such as first screening all participants with conventional risk factors and then targeting individuals at intermediate risk for additional evaluation with novel risk markers. In an ideal situation, such head-to-head comparisons would be best done based on analysis of data from a common set of participants with complete information on the conventional risk factors and all the additional markers compared. But, as this was not possible in our analysis, our projections of the impact of screening with various other markers in a standard population of 100,000 participants (eg European standard population) involved the additional modelling assumptions explained below.

We use $H_0$ to denote the set of participants with complete information on conventional risk factors, and $r_0$ to denote the category of predicted 10-year risk from the model with conventional risk factors. Similarly, $H_m$ refers to the subset of participants with additional information on other markers and $r_m$ is the category of predicted 10-year risk from the corresponding models. In our case, $m$ = 1, 2, 3, 4 represented the models with additional information on apolipoproteins, Lp(a), and Lp-$PLA_2$ mass, respectively. We consider the general case with three 10-year CVD risk categories, coded 1 for <10% ("low risk"), 2 for ≥10% to <20% ("intermediate risk"), and 3 for ≥20% ("high risk") and indexed by $i$ or $j$ in the equations below. We also index the CVD status ($c$) at 10-years by $c$=0 for non-cases and $c$=1 for cases. **Table A1** summarises the notation in the more familiar reclassification table format, also showing the implied nested structure of $H_m$ within $H_0$.

To target individuals at intermediate risk, the survival probability for each individual is first calculated using (2) based on the conventional risk factors and categorised as above. The probability of being allocated to risk category $i$ based on this conventional model, can then be simply estimated by the proportion of people within each group, i.e. $P(r_0 = i \mid H_0) = \frac{n_{i\forall}^{(0)}}{n^{(0)}}$ (**Table A1**). To estimate the benefit of additional screening with other markers, the joint probability distribution i.e. $P(r_0 = i, r_m = j \mid H_0)$ upon re-estimation of 10-year risk under the alternative risk prediction model, was required in order to estimate the probability of being reclassified from a risk category of interest (e.g. intermediate risk) to another category. This joint probability distribution could have been straight forwardly estimated as the proportion of participants in each cell had all the participants had complete information on the additional other markers (i.e. no missing values in the $r_m$ missing column in **Table A1**). Moreover, because the characteristics of the subset of participants with the additional other markers may in some instances have differed from one marker to the other (e.g. due to study-wise missing information for some markers) it was not possible to make valid comparisons based on analyses restricted separately to only those with each novel marker measured. In addition, the observed reclassification tables stratified by CVD status at 10-years ignored information from participants censored before 10-years (since their CVD status was otherwise unclear), and hence a direct extrapolation based on the observed counts would necessarily have underestimated the number of incident CVD cases in the population. To circumvent these limitations, we first applied the Bayes rule to express the joint probability of interest, stratified by CVD status, in terms of conditional probabilities that can be estimated from the data, i.e.:

$$P(r_0 = i, r_m = j \mid C = c, H_0) = P(r_m = j \mid r_0 = i, C = c, H_0) P(r_0 = i \mid C = c, H_0) \quad (12)$$

Assuming that the participants in $H_m$ were a random sample of those in $H_0$, we approximated the first right hand side term as a proportion of the observed row margins in $H_m$, ie:,

$$P(r_m = j \mid r_0 = i, C = c, H_0) \approx P(r_m = j \mid r_0 = i, C = c, H_m) = \frac{n_{ij,c}^{(m)}}{n_{i\forall,c}^{(m)}} \quad (13)$$

To generalise the results to the European standard population, we estimated the above probabilities from reclassification tables stratified by sex and age group to accommodate the different age and sex distributions between ERFC cohorts and the European standard population. Thus for stratification by sex and 4 age groups (40-50, 50-60, 60-70 and 70+) ($G$) collectively indexed by $g = 1,2,....8$ respectively, (13) can then be re-written as:

$$P(r_m = j \mid r_0 = i, C = c, G = g, H_m) = \frac{n_{ij,c}^{(m,g)}}{n_{i\forall,c}^{(m,g)}} \quad (14)$$

Therefore, the expected joint probability distribution with additional screening with novel risk markers in the European standard population is given by

$$
\begin{aligned}
&P(r_0 = i, r_m = j \mid H_{eu}) \\
&= \sum_{c,g} P(r_0 = i, r_m = j \mid C = c, G = g, H_{eu}) P(C = c, G = g \mid H_{eu}) \\
&= \sum_{c,g} P(r_m = j \mid r_0 = i, C = c, G = g, H_{eu}) P(r_0 = i \mid C = c, G = g, H_{eu}) P(C = c, G = g \mid H_{eu})
\end{aligned}
\qquad (15)
$$

where $H_{eu}$ represents the hypothetical standard European population aged ≥ 40 years (structured as shown in **Table A2)**[16]. $P(r_m = j \mid r_0 = i, C = c, G = g, H_{eu})$ is approximated by (14) assuming that the probability distribution observed in the ERFC data would be applicable to the European standard population. $P(r_0 = i \mid C = c, G = g, H_{eu})$ is also estimated from the ERFC data using (16), and represents the distribution of predicted risk according to the conventional risk factors model in each stratum of the population.

$$
P(r_0 = i \mid C = c, G = g, H_0) = \frac{n_{i\forall,c}^{(0,g)}}{n_c^{(0,g)}} \qquad i = 1,2,3 \qquad (16)
$$

The last term in (15) $P(C = c, G = g \mid H_{eu})$ defines the expected distribution of the standard population by sex, age group, and CVD status at 10-years, assuming that the sex- and age group-specific incidence rates observed in ERFC were applicable. It is calculated as:

$$
P(C = c, G = g \mid H_{eu}) = P(C = c \mid G = g, H_{eu}) P(G = g \mid H_{eu}) \qquad (17)
$$

where $P(G = g \mid H_{eu})$ is given *a priori* by the corresponding proportion in EU standard population (**Table A2**) and $P(C = c \mid G = g, H_{eu})$ is estimated using ERFC incidence rates assuming the exponential survival function.

$$
P(C = 1 \mid G = g, H_{eu}) = 1 - \exp(-I_g T) \qquad (18)
$$

where $I_g$ denotes the incidence rate in each stratum $g = 1,2,....8$ and $T = 10$ for the calculation of average 10-year CVD risk

The expected number of people in cell $ij$ for European standard population is extracted by simply multiplying $P(r_0 = i, r_m = j \mid H_{eu})$ by 100,000 participants. To estimate the clinical benefit of treatment among those initially classified as being at intermediate risk according to the conventional risk factors model, we assumed that treatment with statins would reduce the risk of CVD by 20% and that treatment would only be allocated to those estimated to be at high risk according to model $m$. Such a sequential strategy would therefore lead to additional treatment of a total of $P(r_0 = 2, r_m = 3 \mid H_{eu}) \times 10^5$ participants, of whom approximately

$$
N_{23,1} = P(r_0 = 2, r_m = 3, C = 1 \mid H_{eu}) \times 10^5 \qquad (19)
$$

are expected to be CVD cases. Hence assuming a 20% CVD risk reduction with statin treatment, it would be expected that an extra $0.20 \times N_{23,1}$ CVD events would be prevented among those initially at intermediate risk according to the conventional model.

*Examination of the simplifying assumption made in (13)*

The method described above involved different subsets for several novel markers and relied on the assumption of the approximation of the conditional probability in equation (13) above being satisfied. This would be valid if the participants with additional information on other markers were a random sample of those with complete information on the conventional risk factors in each risk category, ie the participants in cell $n_{i.,c}^{(m)}$ are randomly dropped from the maximum of $n_{i\forall,c}^{(0)}$ (**Table A1**). To examine whether this assumption was reasonable, we calculated and compared the predicted risks using the conventional risk factors in both $H_0$ and $H_m$ ,.ie $\Pr(T \le t \mid X_i, c, H_0)$ and $\Pr(T \le t \mid X_i, c, H_m)$. Optimally, $\Pr(T \le t \mid X_i, c, H_0) = \Pr(T \le t \mid X_i, c, H_m)$ if the assumption of randomness is fulfilled, otherwise, systematic bias will be incurred. We computed the absolute mean difference as well as the mean ratio between $\Pr(T \le t \mid X_i, c, H_0)$ and $\Pr(T \le t \mid X_i, c, H_m)$ in each strata. As a sensitivity analysis, we excluded strata for which the absolute mean difference in predicted risk was greater than 0.01 in the estimation of the conditional probability in (13), which made no difference to the results. The latter was also the same in sensitivity analyses excluding strata with mean ratio outside the range of [0.95, 1.05].

**Table A1.** Reclassification table showing observed counts stratified by case-control status.

| | $r_m = 1$ | $r_m = 2$ | $r_m = 3$ | Total in $H_m$ | $r_m$ is missing | Total in $H_0$ |
|---|---|---|---|---|---|---|
| Cases | | | | | | |
| $r_0 = 1$ | $n_{11,1}^{(m)}$ | $n_{12,1}^{(m)}$ | $n_{13,1}^{(m)}$ | $n_{1\forall,1}^{(m)}$ | $n_{1.,1}^{(m)}$ | $n_{1\forall,1}^{(0)}$ |
| $r_0 = 2$ | $n_{21,1}^{(m)}$ | $n_{22,1}^{(m)}$ | $n_{23,1}^{(m)}$ | $n_{2\forall,1}^{(m)}$ | $n_{2.,1}^{(m)}$ | $n_{2\forall,1}^{(0)}$ |
| $r_0 = 3$ | $n_{31,1}^{(m)}$ | $n_{32,1}^{(m)}$ | $n_{33,1}^{(m)}$ | $n_{3\forall,1}^{(m)}$ | $n_{3.,1}^{(m)}$ | $n_{3\forall,1}^{(0)}$ |
| | $n_{\forall 1,1}^{(m)}$ | $n_{\forall 2,1}^{(m)}$ | $n_{\forall 3,1}^{(m)}$ | $n_1^{(m)}$ | $n_{\forall.,1}^{(m)}$ | $n_1^{(0)}$ |
| Controls | | | | | | |
| $r_0 = 1$ | $n_{11,0}^{(m)}$ | $n_{12,0}^{(m)}$ | $n_{13,0}^{(m)}$ | $n_{1\forall,0}^{(m)}$ | $n_{1.,0}^{(m)}$ | $n_{1\forall,0}^{(0)}$ |
| $r_0 = 2$ | $n_{21,0}^{(m)}$ | $n_{22,0}^{(m)}$ | $n_{23,0}^{(m)}$ | $n_{2\forall,0}^{(m)}$ | $n_{2.,0}^{(m)}$ | $n_{2\forall,0}^{(0)}$ |
| $r_0 = 3$ | $n_{31,0}^{(m)}$ | $n_{32,0}^{(m)}$ | $n_{33,0}^{(m)}$ | $n_{3\forall,0}^{(m)}$ | $n_{3.,0}^{(m)}$ | $n_{3\forall,0}^{(0)}$ |
| | $n_{\forall 1,0}^{(m)}$ | $n_{\forall 2,0}^{(m)}$ | $n_{\forall 3,0}^{(m)}$ | $n_0^{(m)}$ | $n_{\forall.,0}^{(m)}$ | $n_0^{(0)}$ |
| Total | | | | | | |
| $r_0 = 1$ | $n_{11}^{(m)}$ | $n_{12}^{(m)}$ | $n_{13}^{(m)}$ | $n_{1\forall}^{(m)}$ | $n_{1.}^{(m)}$ | $n_{1\forall}^{(0)}$ |
| $r_0 = 2$ | $n_{21}^{(m)}$ | $n_{22}^{(m)}$ | $n_{23}^{(m)}$ | $n_{2\forall}^{(m)}$ | $n_{2.}^{(m)}$ | $n_{2\forall}^{(0)}$ |
| $r_0 = 3$ | $n_{31}^{(m)}$ | $n_{32}^{(m)}$ | $n_{33}^{(m)}$ | $n_{3\forall}^{(m)}$ | $n_{3.}^{(m)}$ | $n_{3\forall}^{(0)}$ |
| | $n_{\forall 1}^{(m)}$ | $n_{\forall 2}^{(m)}$ | $n_{\forall 3}^{(m)}$ | $n^{(m)}$ | $n_{\forall.}^{(m)}$ | $n^{(0)}$ |

$n_{ij,1}^{(m)}$ = No of cases in cell $ij$

$n_{ij,0}^{(m)}$ = No of controls in cell $ij$

$n_{ij}^{(m)}$ = Total No of participants in cell $ij$ = $n_{ij,1}^{(m)} + n_{ij,0}^{(m)}$

$n_{i\forall,c}^{(m)}$ = Total No of participants in row $i$ in $H_m$ = $n_{i1,c}^{(m)} + n_{i2,c}^{(m)} + n_{i3,c}^{(m)}$

$n_{i\forall,c}^{(0)}$ = Total No of participants in row $i$ in $H_0$ = $n_{i1,c}^{(m)} + n_{i2,c}^{(m)} + n_{i3,c}^{(m)} + n_{i.,c}^{(m)}$

**Table A2.** European standard population of 100,000 participants with age ≥ 40.

| Age group | No. of people (%) | |
|---|---|---|
| | Men | Women |
| 40-50 | 16,300 (16.3%) | 16,300 (16.3%) |
| 50-60 | 15,100 (15.1%) | 15,100 (15.1%) |
| 50-70 | 10,500 (10.5%) | 10,500 (10.5%) |
| ≥ 70 | 8,100 (8.1%) | 8,100 (8.1%) |

**References** eAppendix 2


1. Cox DR. Regression models and life tables. J Roy Stat Soc B 1972; 74:187-220.

2. White IR. Multivariate random-effects meta-analysis. The Stata Journal 2009; 9(1):40-56.

3. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. Int J Epidemiol 2010; 39(5)1345-1359.

4. Phillips AN, Thompson SG, Pocock SJ. Prognostic scores for detecting a high risk group: estimating the sensitivity when applied to new data. Stat Med 1990; 9(10):1189-1198.

5. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982; 247(18):2543-2546.

6. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med 2004; 23(5):723-748.

7. Newson R. Efficient calculation of jackknife confidence intervals for rank statistics. J Stat Softw 2005; 1(15):1-10.

8. Fibrinogen Studies Collaboration, Pennells L, White IR, Wood AM, Kaptoge S, Sarwar N. Measures to assess the prognostic ability of the stratified Cox proportional hazards model. Stat Med 2009; 28(3):389-411.

9. Newson R. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differrences. The Stata Journal 2002; 2(1):45-64.

10. Newson R. Confidence intervals for rank statistics: Somers' D and extensions. The Stata Journal 2006; 6(3):309-334.

11. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. The Stata Journal 2010; 10(3):339-358.

12. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21(11):1539-1558.

13. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000; 19(9):1141-1164.

14. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008; 27(2):157-172.

15. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011; 30(1):11-21.

16. European Cancer Observatory, http://eu-cancer.iarc.fr/5-glossary.html,en.