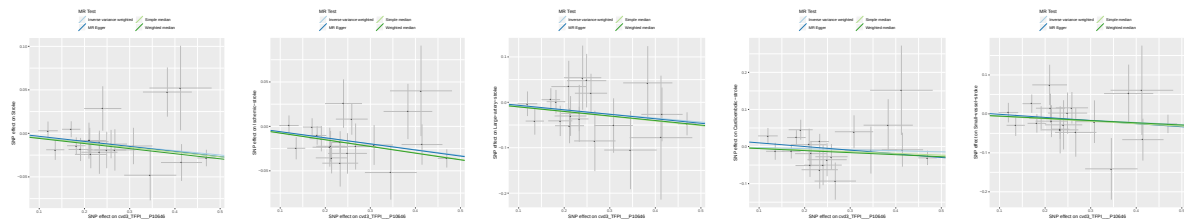


Discovery of potential therapeutic targets for stroke using human plasma proteome through systematic Mendelian randomization

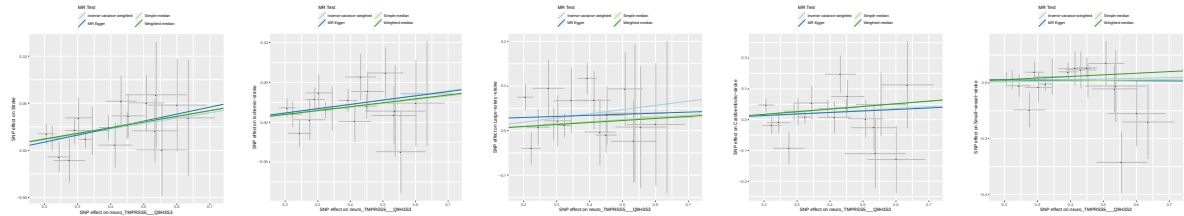
Supplementary Figures

Supplementary Figure 1. Standard MR plots for proteins and risk of stroke.	2
Supplementary Figure 2. Colocalization plots of pQTLs and genetic associations of stroke outcomes and risk factor(s).	10
Supplementary Figure 3. Standard MR plots for risk factors and stroke.	11
Supplementary Figure 4. Standard MR plots for proteins and stroke risk factors.....	12
Supplementary Figure 5. Manhattan plots of the potential on-target side-effects from Phe-MR analysis.	14
Supplementary Figure 6. Forest plots illustrating the potential on-target side-effects associated with causal proteins revealed by Phe-MR analysis.	16
Supplementary Figure 7. Correlations of causal effect sizes of six proteins on stroke outcomes with IVs derived from primary method and others.....	17

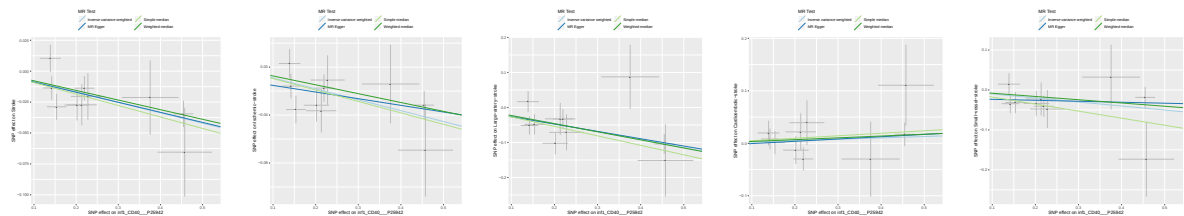
TFPI



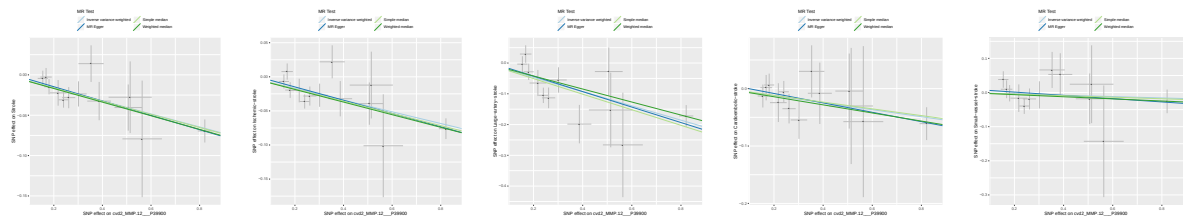
TMPRSS5



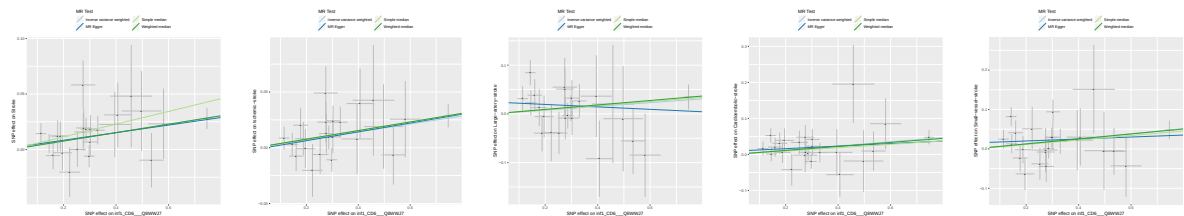
CD40



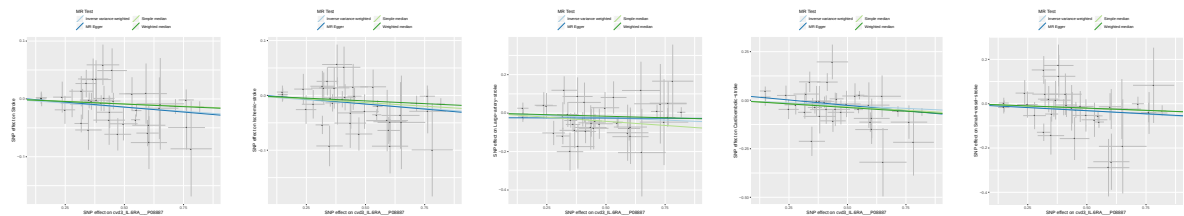
MMP12



CD6



IL6RA



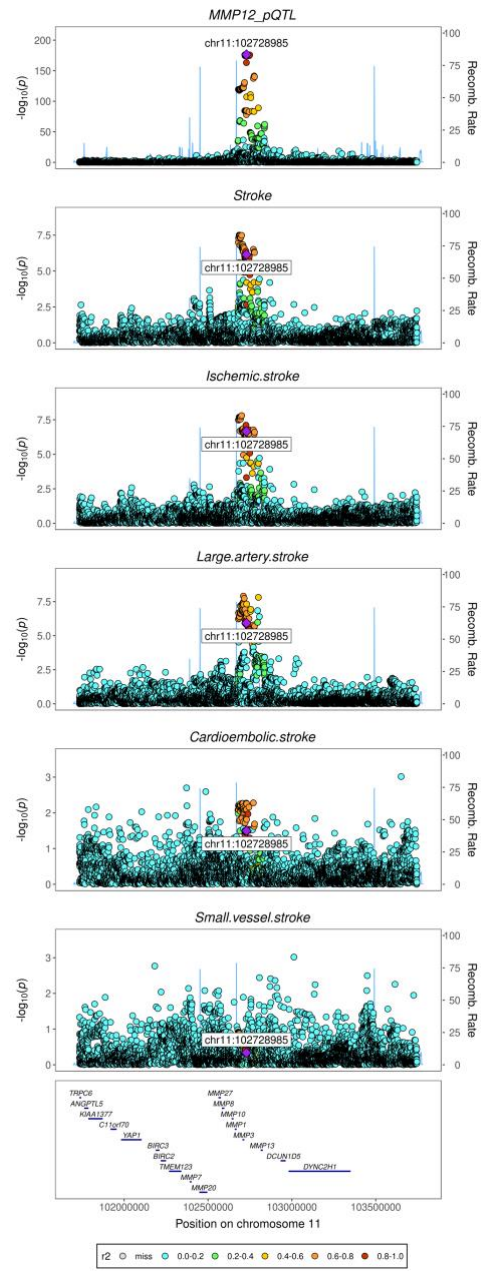
Supplementary Figure 1. Standard MR plots for proteins and risk of stroke.

Scatter plot showing the estimated effect sizes and 95% confidence intervals on plasma proteins (from linear regression; horizontal lines) and risk of five stroke outcomes (from logistic regression; vertical lines) for each variant used in the MR analyses. The different regression lines indicate the effect sizes as calculated by different MR tests (Methods). Abbreviations: CD40 = B Cell Surface Antigen CD40;

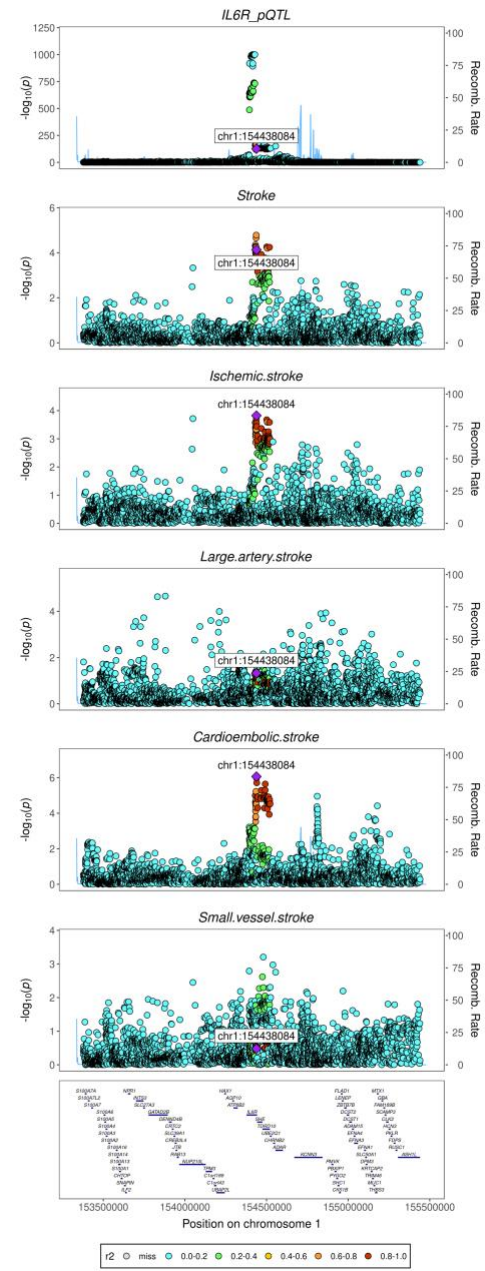
TFPI = Tissue Factor Pathway Inhibitor; MMP12 = Matrix Metalloproteinase 12; IL6RA = Interleukin 6 Receptor Subunit Alpha; TMPRSS5= Transmembrane Serine Protease 5; CD6 = T-Cell Differentiation Antigen CD6.



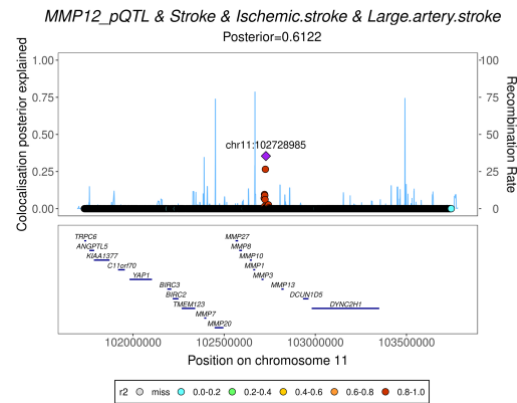
i



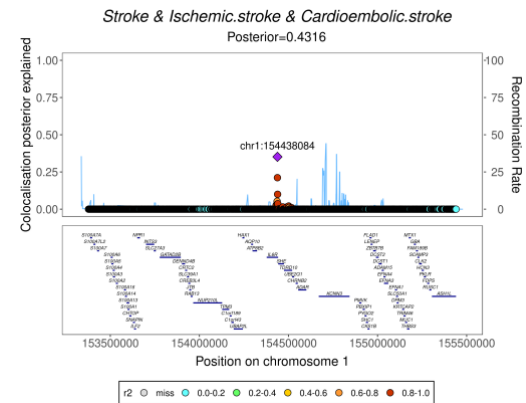
k



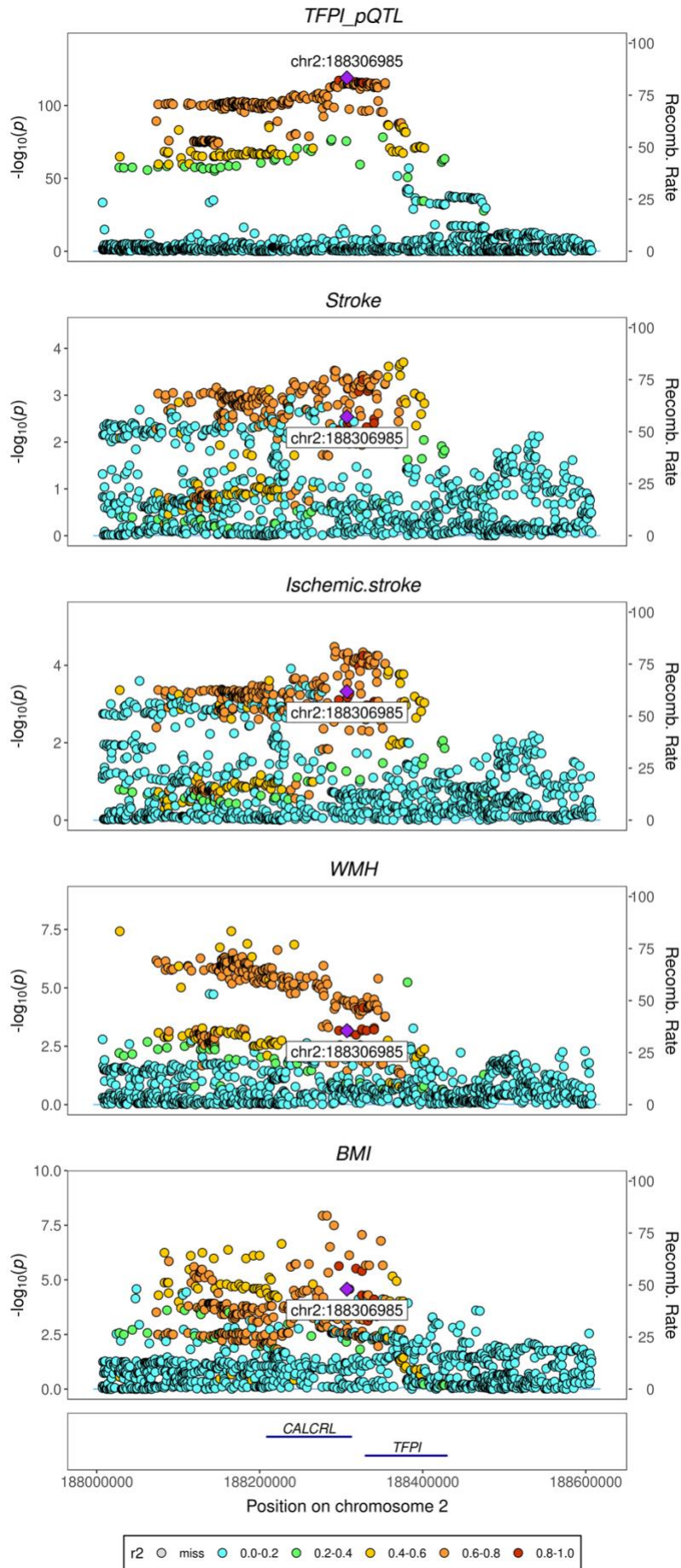
j



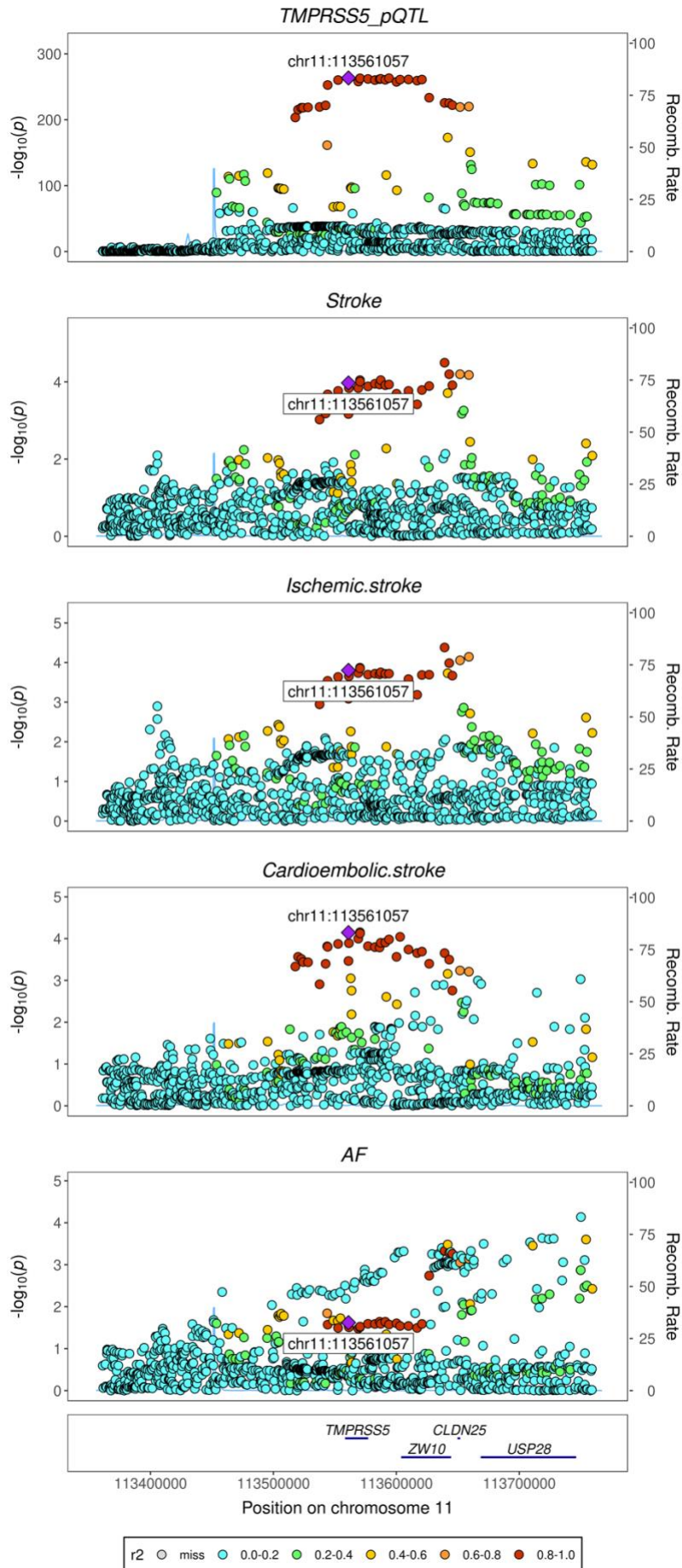
l



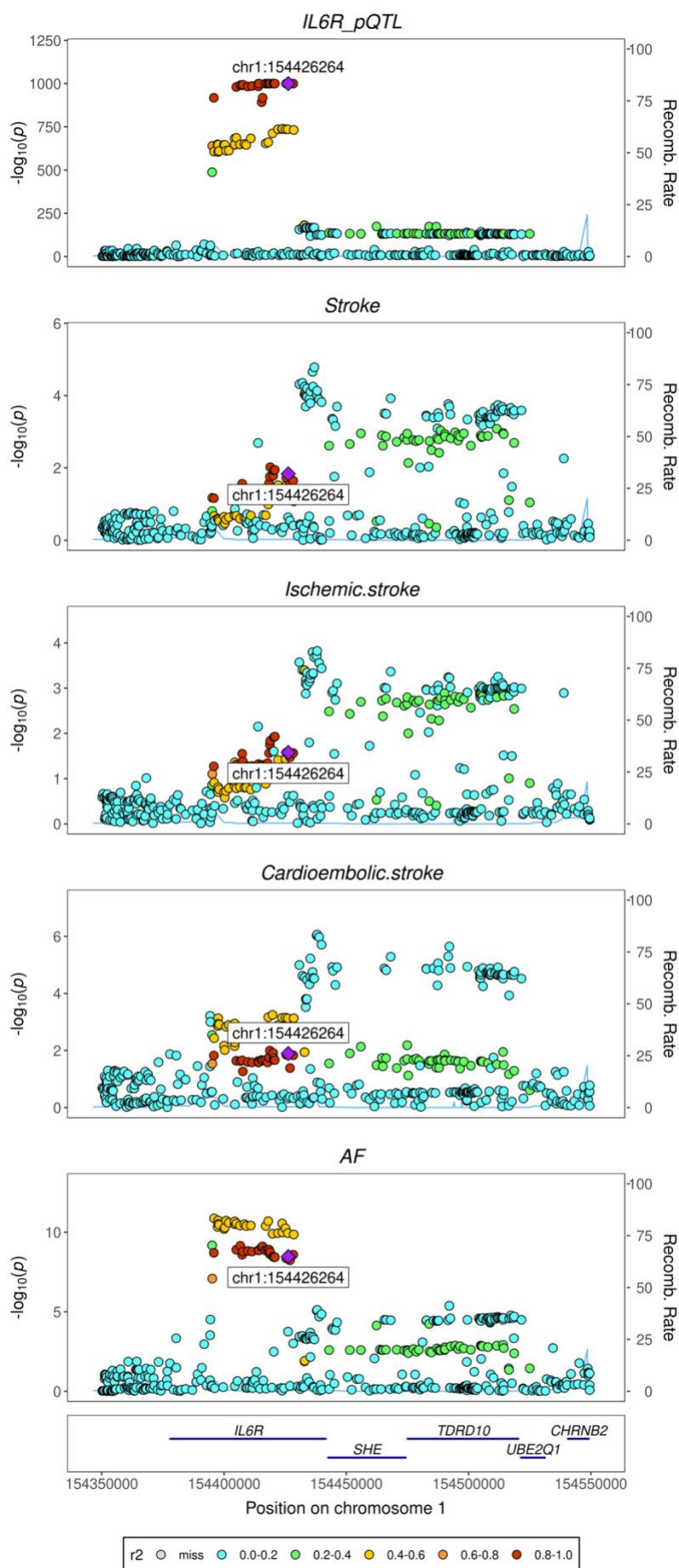
m



n



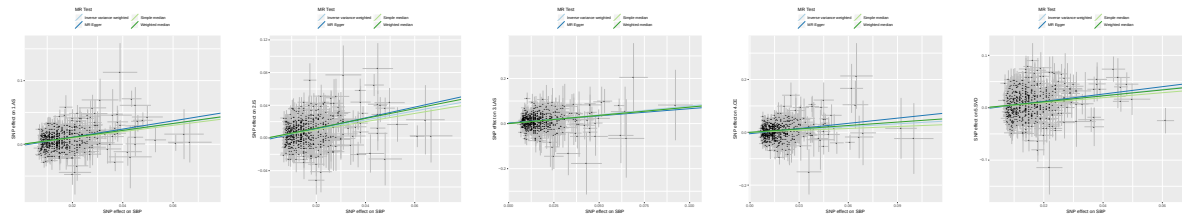
0



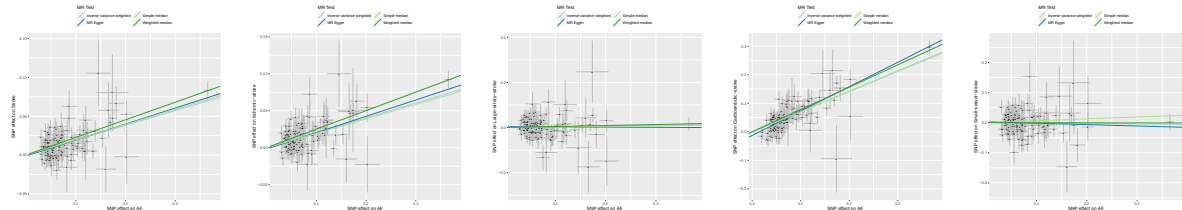
Supplementary Figure 2. Colocalization plots of pQTLs and genetic associations of stroke outcomes and risk factor(s).

a & b: colocalization of TFPI pQTLs and stroke association signals; c & d: colocalization of CD6 pQTLs and stroke association signals; e & f: colocalization of TMPRSS5 pQTLs and stroke association signals; g & h: colocalization of CD40 pQTLs and stroke association signals; i & j: colocalization of MMP12 pQTLs and stroke association signals; k & l: colocalization of IL6RA pQTLs and stroke association signals; m: stacked genetic association plots of TFPI pQTLs, stroke, ischemic-stroke, WHM and BMI associations; n: stacked genetic association plots of TMPRSS5 pQTLs, stroke, ischemic-stroke, cardioembolic-stroke and AF associations; o: stacked genetic association plots of IL6RA pQTLs, stroke, ischemic-stroke, cardioembolic-stroke and AF associations. a/c/e/g/i/k/m/n/o: regional association plot of pQTLs and stroke/risk factor associations. SNPs are plotted by their positions on the chromosome against association with the protein level and stroke/ risk factor ($-\log_{10} P$ value) on the left y axis. Points are coloured by their local linkage disequilibrium (LD) pattern with their “candidate causal SNP” (purple diamond). Below the main plot is the track showing the position of local genes. b/d/f/h/j/l: plot of percentage of Posterior probability (PP) of colocalization explained by each of the SNP within the corresponding region. The traits that colocalized are presented in the title of the plot and the PP of colocalization is presented in the subtitle of the plot. SNPs are plotted by their positions on the chromosome against the percentage explained on the left y axis. The best candidate causal SNP is presented in purple diamond. Abbreviations: CD40 = B Cell Surface Antigen CD40; TFPI = Tissue Factor Pathway Inhibitor; MMP12 = Matrix Metalloproteinase 12; IL6RA = Interleukin 6 Receptor Subunit Alpha; TMPRSS5 = Transmembrane Serine Protease 5; CD6 = T-Cell Differentiation Antigen CD6; AF = Atrial Fibrillation; WMH = White Matter Hyperintensity; BMI = Body Mass Index.

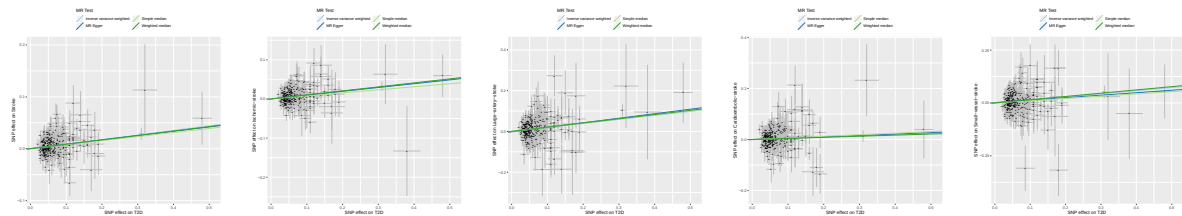
SBP



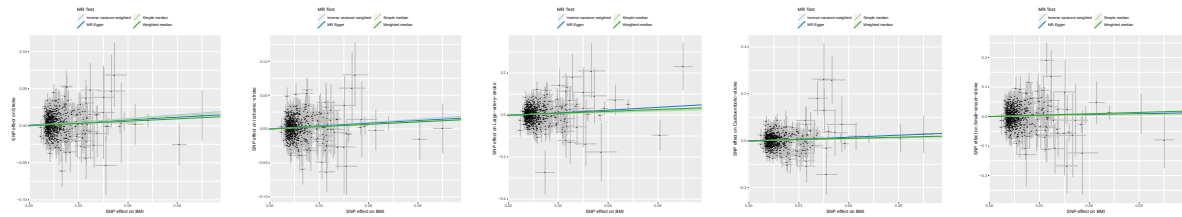
AF



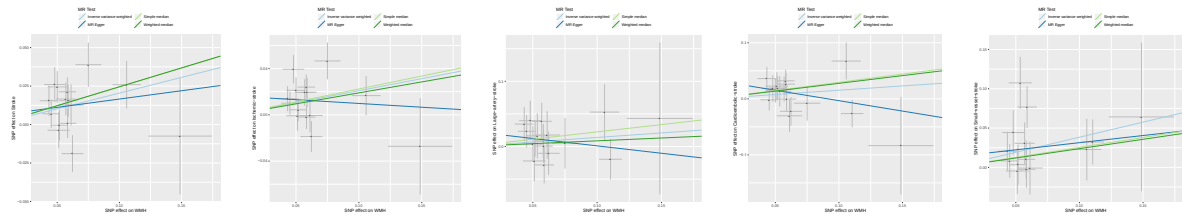
T2D



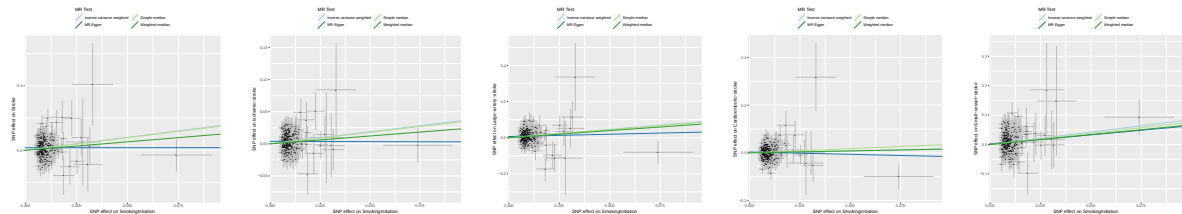
BMI



WMH



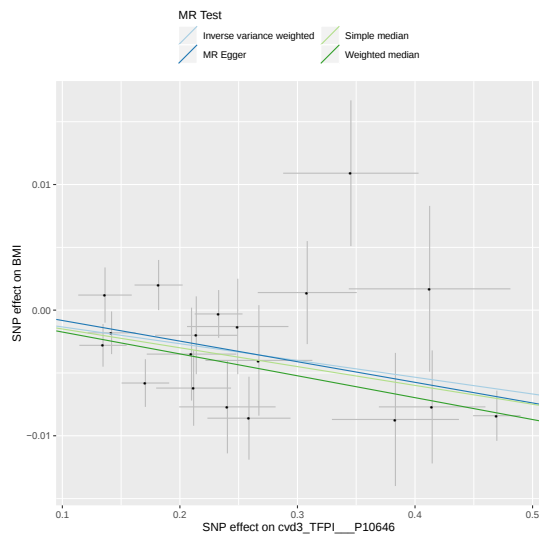
Smoking



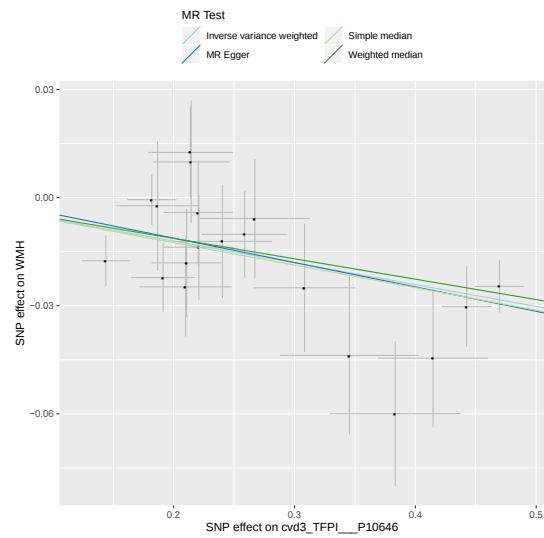
Supplementary Figure 3. Standard MR plots for risk factors and stroke.

Scatter plot showing the estimated effect sizes and 95% confidence intervals on risk factors and risk of five stroke outcomes for each variant used in the MR analyses. The different regression lines indicate the effect sizes as calculated by different MR tests (Methods). Abbreviations: SBP = Systolic Blood Pressure; AF = Atrial Fibrillation; WMH = White Matter Hyperintensity; T2D = Type 2 Diabetes; BMI = Body Mass Index; Smoking = Smoking Initiation.

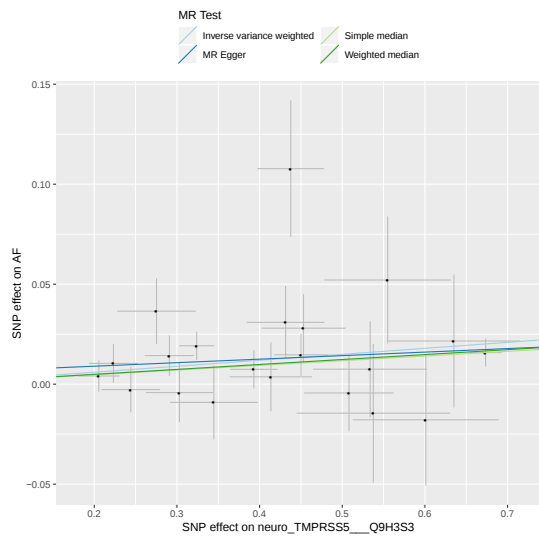
TFPI and BMI



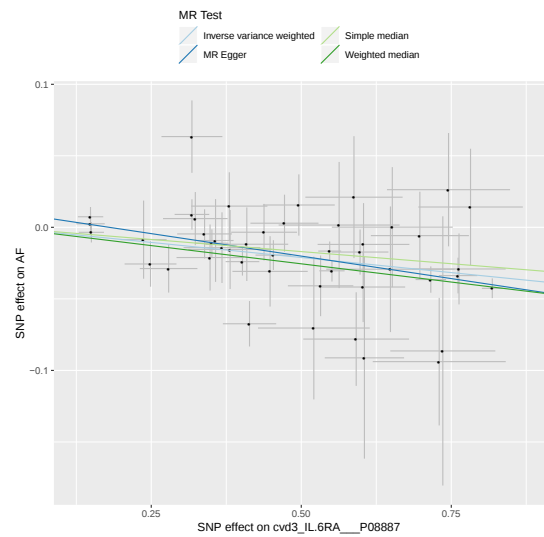
TFPI and WMH



TMPRSS5 and AF



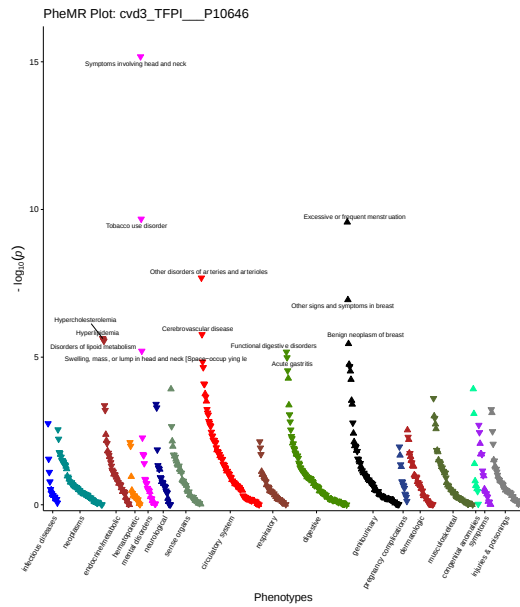
IL6RA and AF



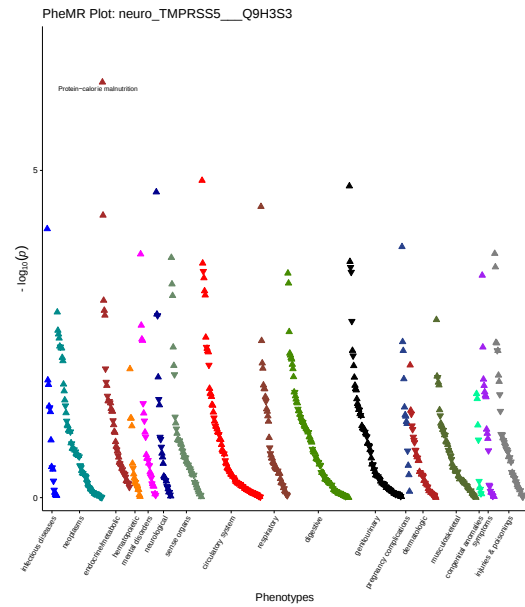
Supplementary Figure 4. Standard MR plots for proteins and stroke risk factors.

Scatter plot showing the estimated effect sizes and 95% confidence intervals on proteins and risk of stroke risk factors for each variant used in the MR analyses. The different regression lines indicate the effect sizes as calculated by different MR tests (Methods). Abbreviations: TFPI = Tissue Factor Pathway Inhibitor; IL6RA = Interleukin 6 Receptor Subunit Alpha; TMPRSS5= Transmembrane Serine Protease 5; AF = Atrial Fibrillation; WMH = White Matter Hyperintensity; BMI = Body Mass Index.

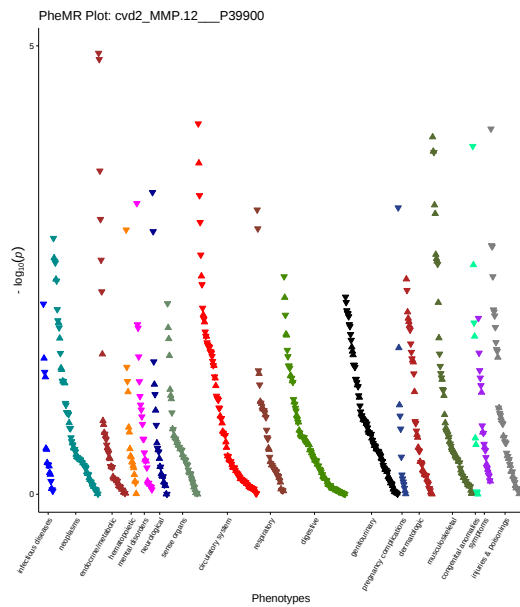
TFPI



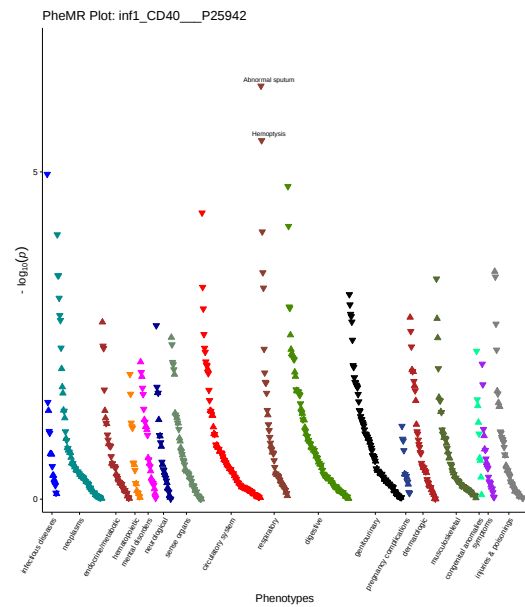
TMPRSS5

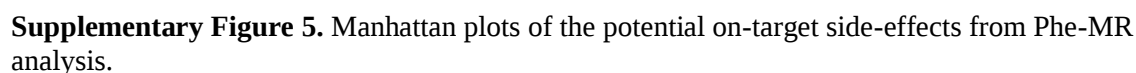


MMP12



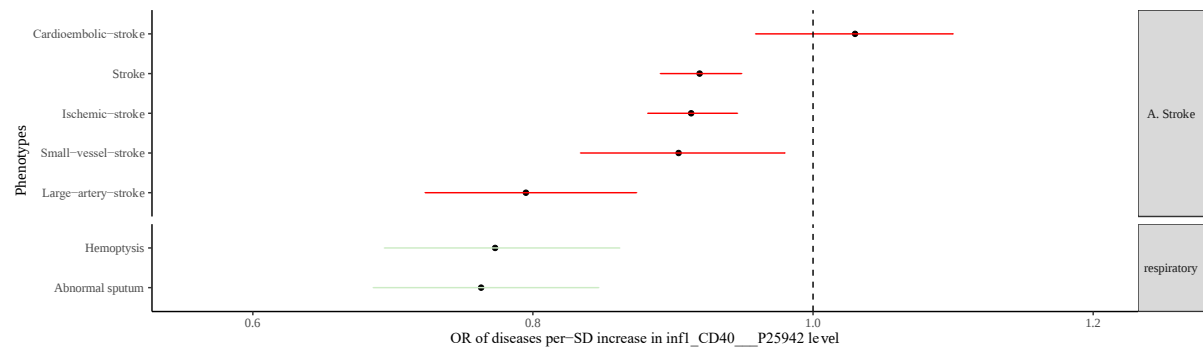
CD40



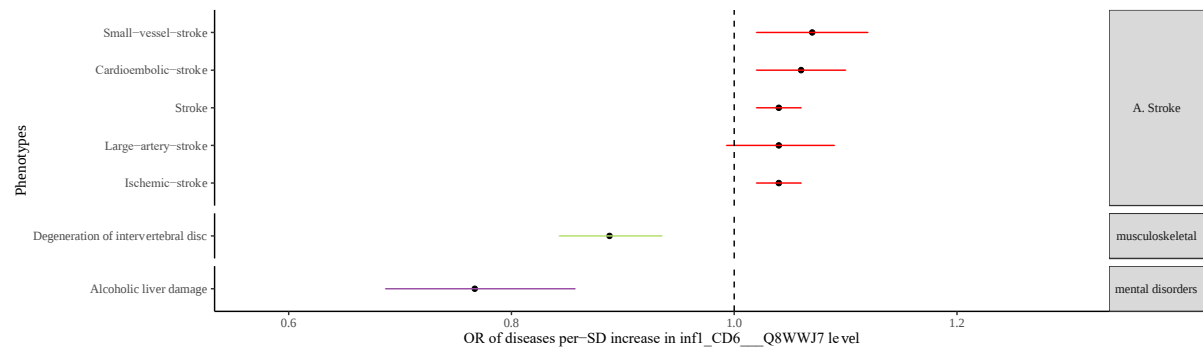


Results can be perceived as the effects of per SD higher circulating protein level on each phenotype - triangle up represent risk-conferring effects and triangle down represents protective effects. If the effect direction of the target protein on the phenotype is consistent with that on stroke outcomes, it represents “beneficial” additional indications through intervention of circulating protein level. Conversely, opposing effect directions of the target protein on the phenotype and stroke represents “deleterious” side-effects. For example, a higher level of TFPI is associated with lower risk of ischemic stroke and so downface triangle represents “beneficial effects”, while the upward triangle represents “deleterious effects” when the hypothetical intervention increases TFPI levels. Associations above the pink line passed Bonferroni significance $P \leq 0.05/6784 = 1.06 \times 10^{-5}$).

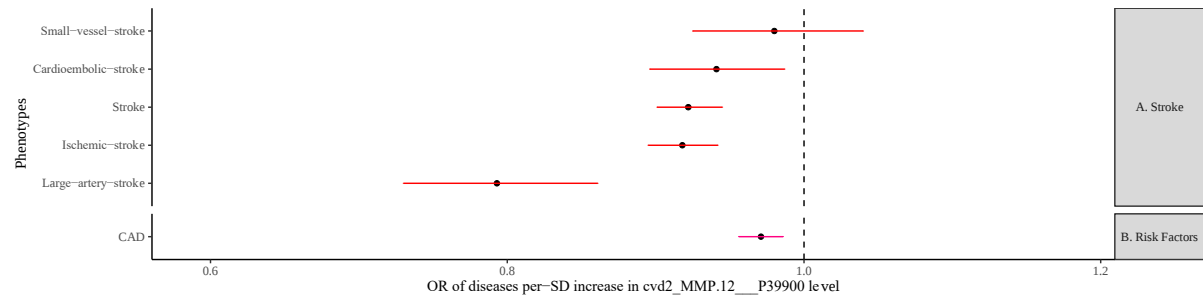
CD40



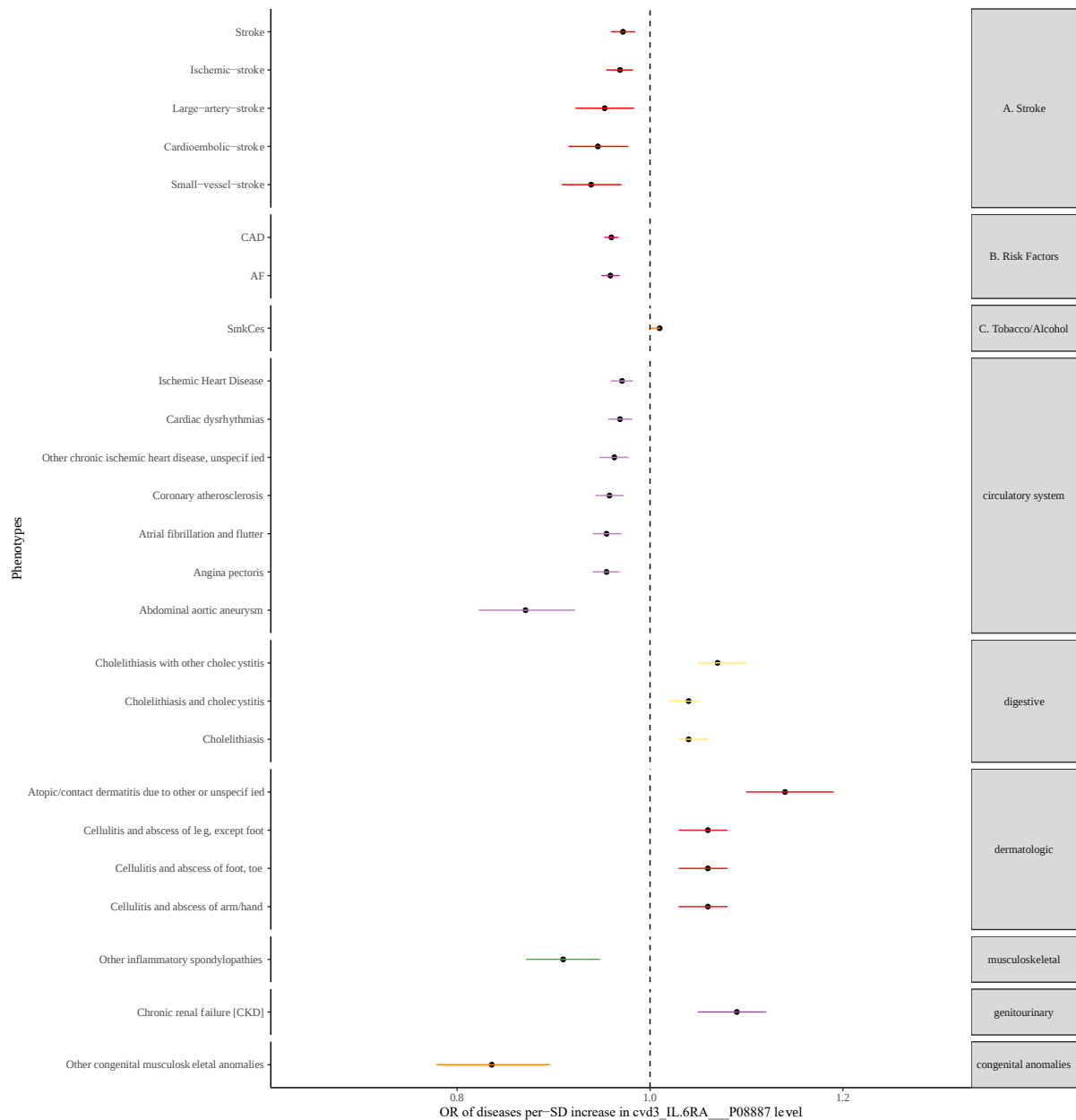
CD6



MMP12

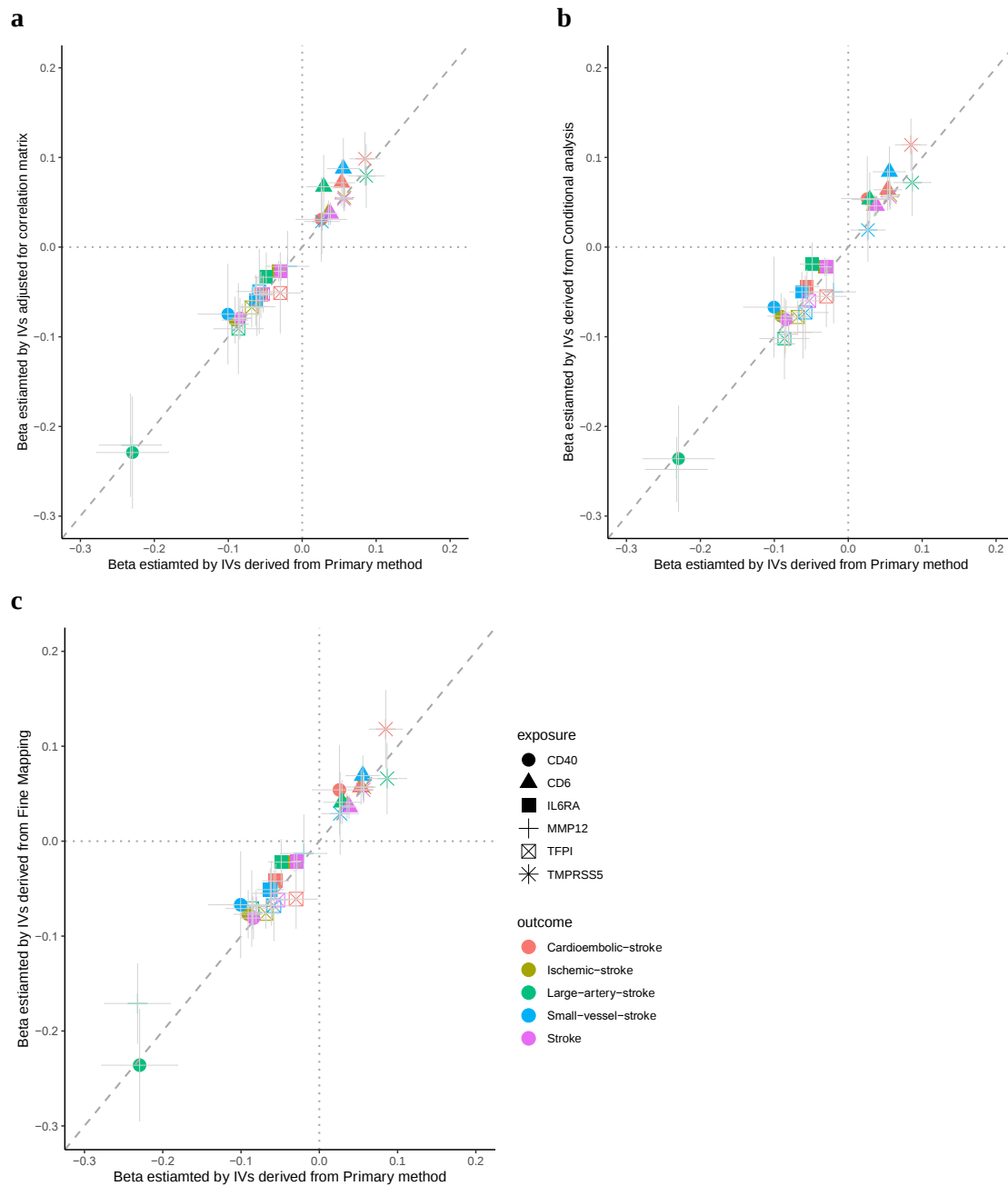


IL6RA



Supplementary Figure 6. Forest plots illustrating the potential on-target side-effects associated with causal proteins revealed by Phe-MR analysis.

Results can be perceived as the effects of per SD higher circulating protein level on each phenotype. If the effect direction of the target protein on the phenotype is consistent with that on stroke outcomes, it represents “beneficial” additional indications through intervention of circulating protein level. Conversely, opposing effect directions of the target protein on the phenotype and stroke represents “deleterious” side-effects. For example, a higher level of TFPI is associated with lower risk of ischemic stroke and so phenotypes with $OR < 1$ represents “beneficial effects”, $OR > 1$ represents “deleterious effects” when the hypothetical intervention increases TFPI levels. Only significant associations that passed Bonferroni correction ($P \leq 0.05/6/784 = 1.06 \times 10^{-5}$) were plotted. The dots are the causal estimates on the OR scale, and whiskers represent the 95% confidence intervals for these ORs.



Supplementary Figure 7. Correlations of causal effect sizes of six proteins on stroke outcomes with IVs derived from primary method and others.

Primary method (IVs derived from LD clumping at $P_{pQTLs} < 5 \times 10^{-8}$ at $R^2 < 0.1$) versus IVs adjusted for correlation matrix (a); IVs derived from conditional analysis (b); and IVs derived from fine mapping (c). Scatter plots show the estimated effect sizes and 95% confidence intervals on proteins and risk of stroke outcomes, with each shape represents each protein and each colour represents each stroke outcome. The dashed grey line is a reference line representing $y=x$. Abbreviations: CD40 = B Cell Surface Antigen CD40; TFPI = Tissue Factor Pathway Inhibitor; MMP12 = Matrix Metalloproteinase 12; IL6RA = Interleukin 6 Receptor Subunit Alpha; TMPRSS5 = Transmembrane Serine Protease 5; CD6 = T-Cell Differentiation Antigen CD6.