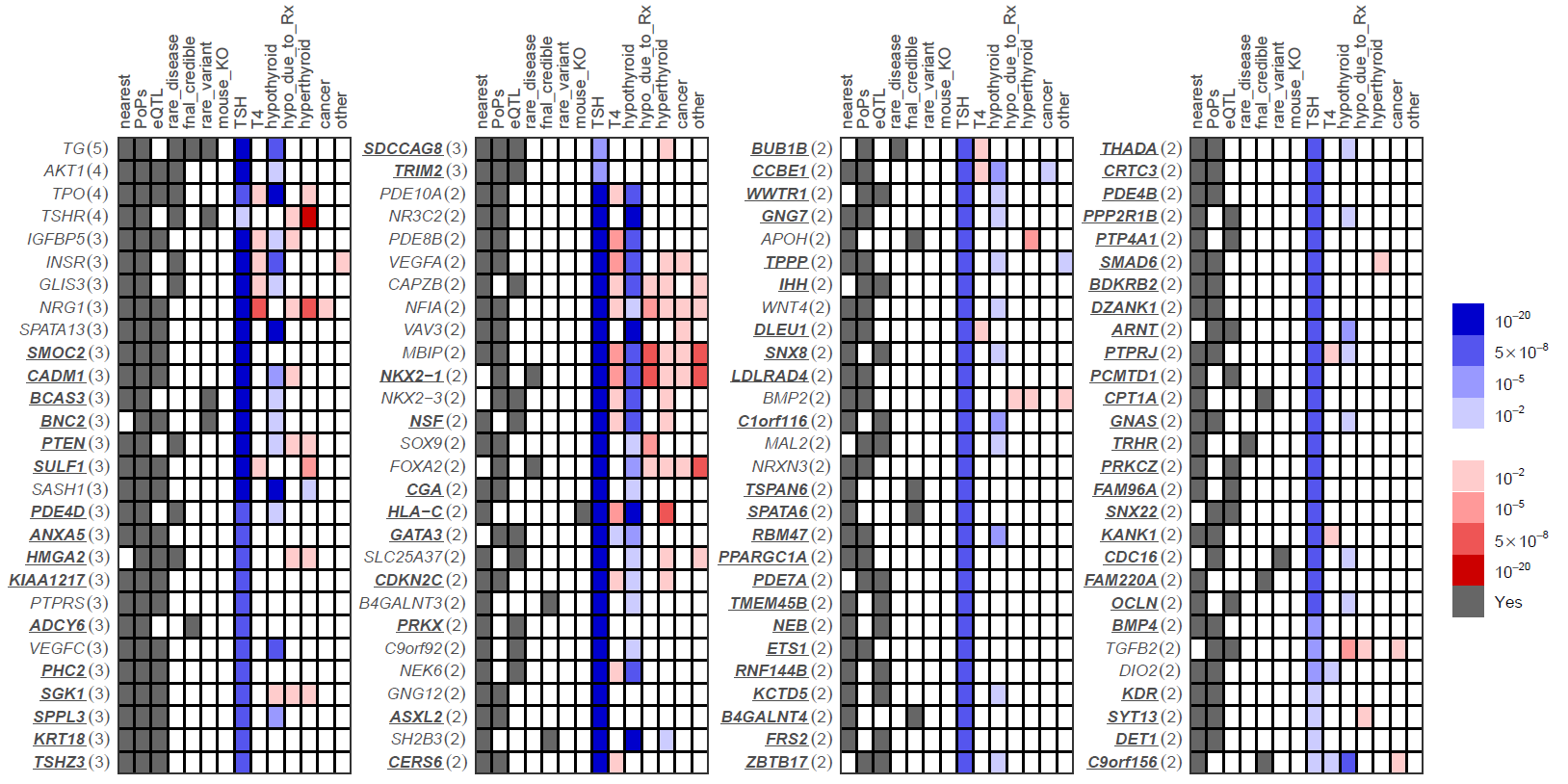
Supplementary Tables and Figures for the manuscript entitled "Genome-wide association study of thyroid-stimulating hormone highlights new genes, pathways and association with thyroid disease susceptibility and age-of-onset".

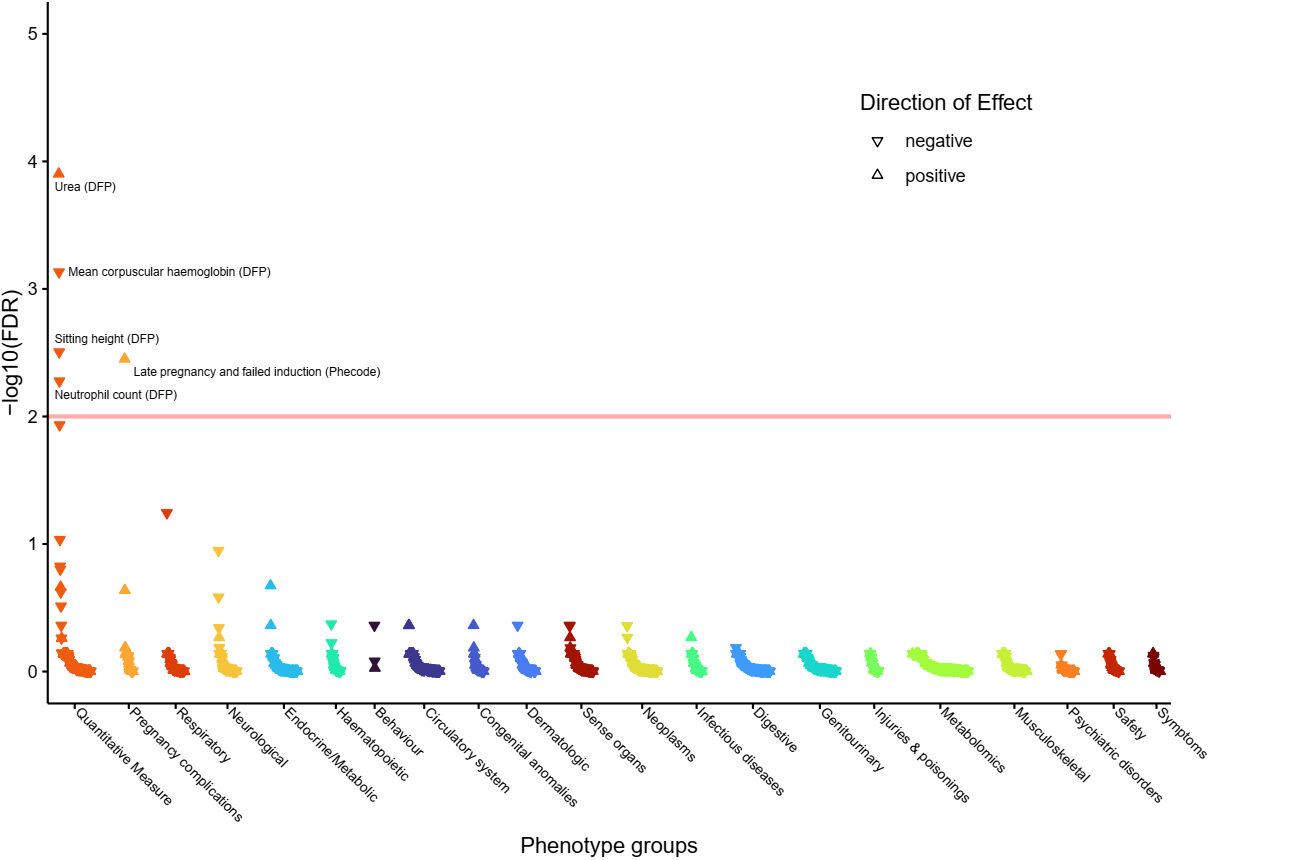
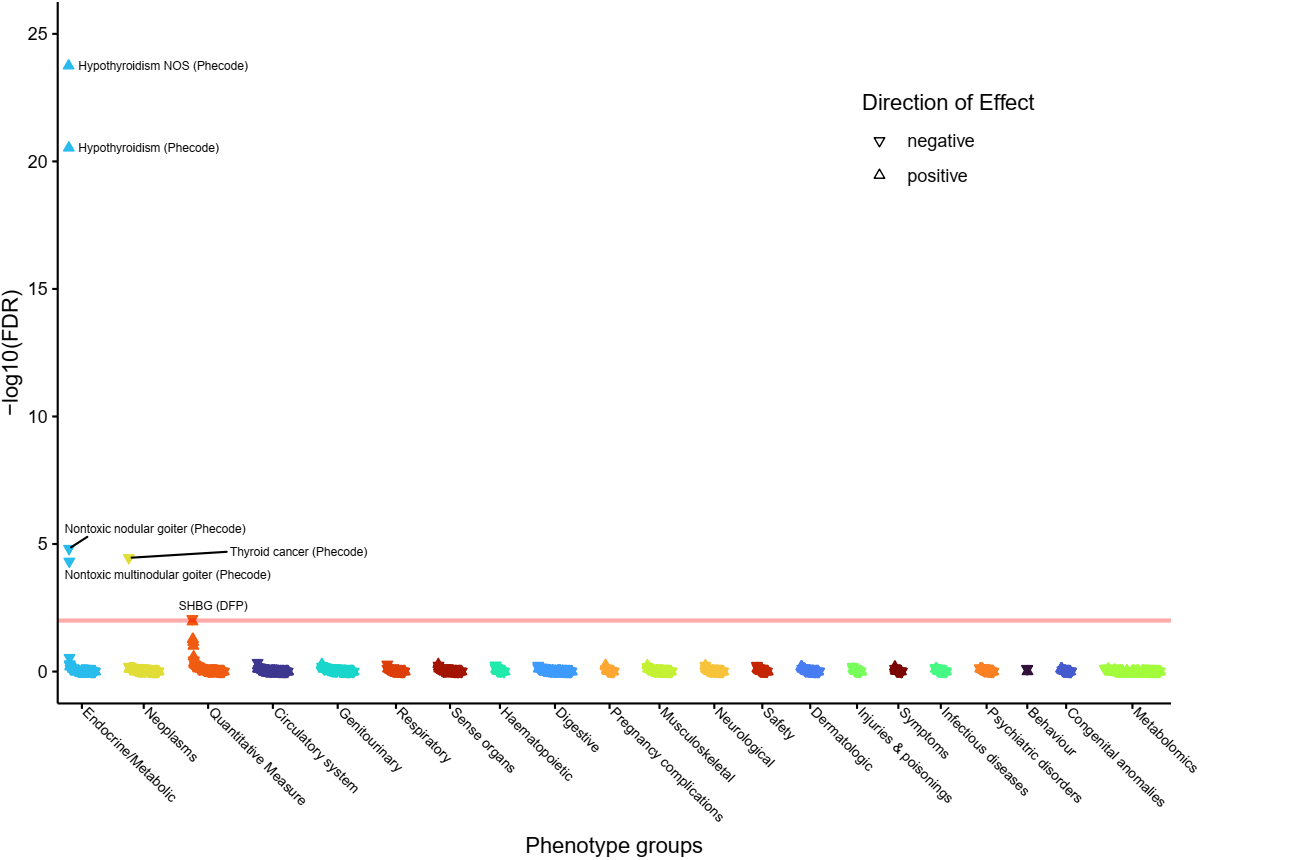
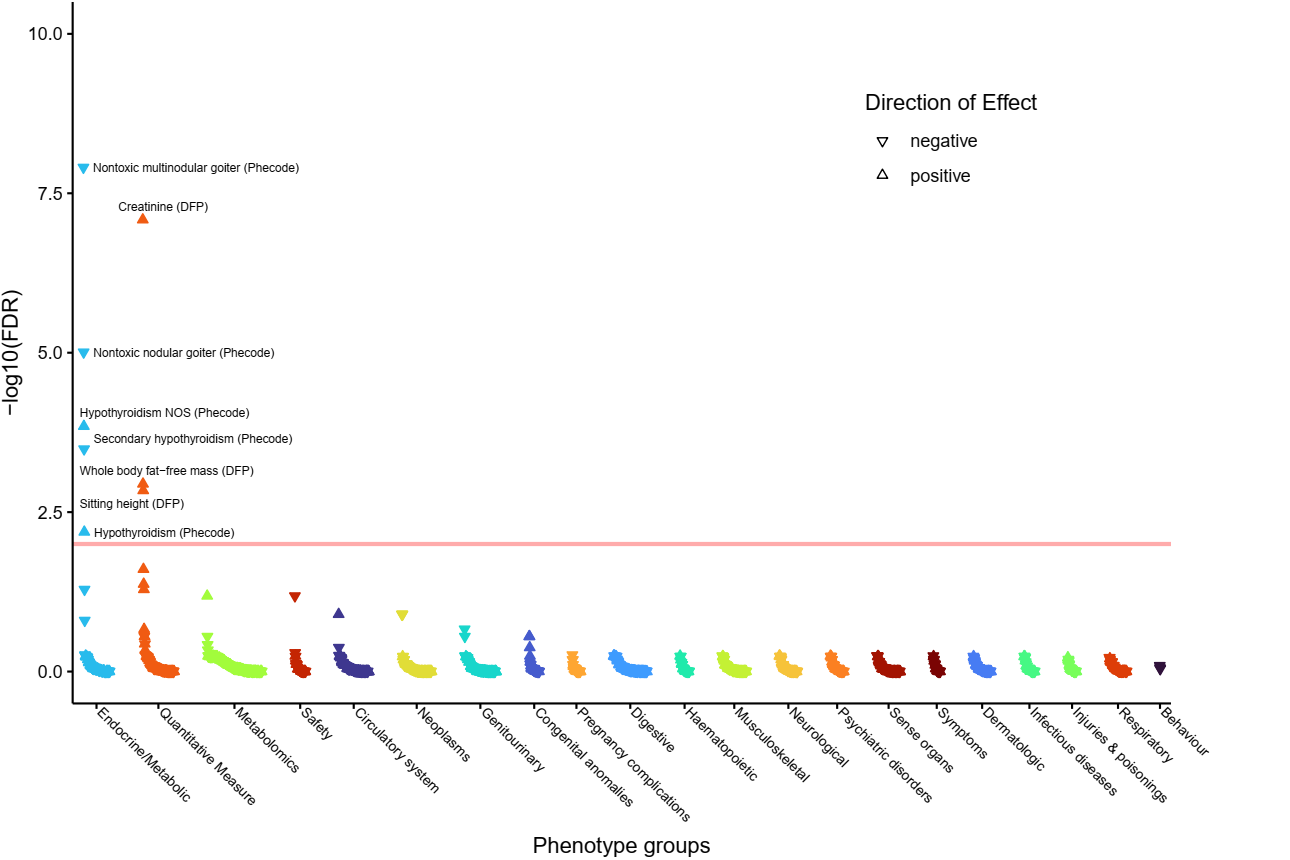
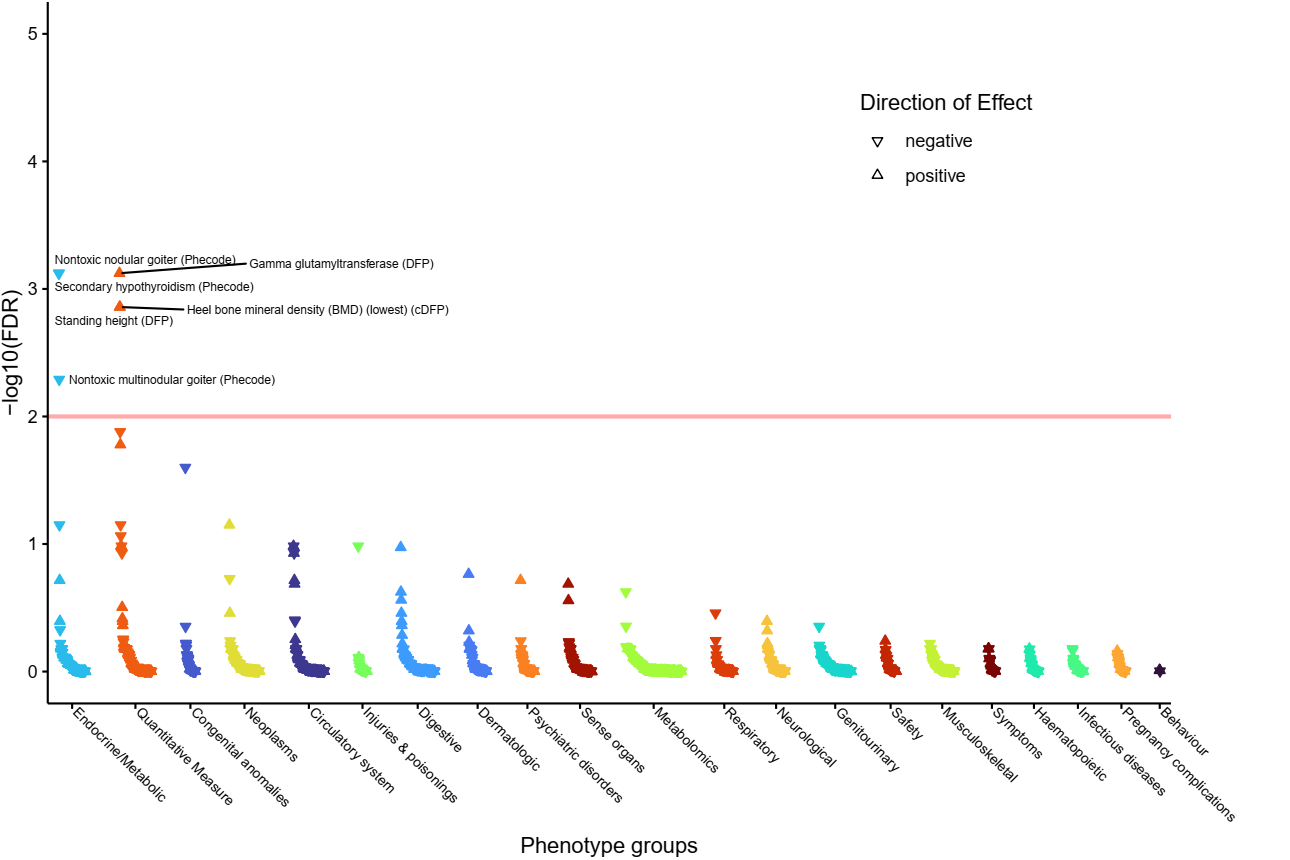
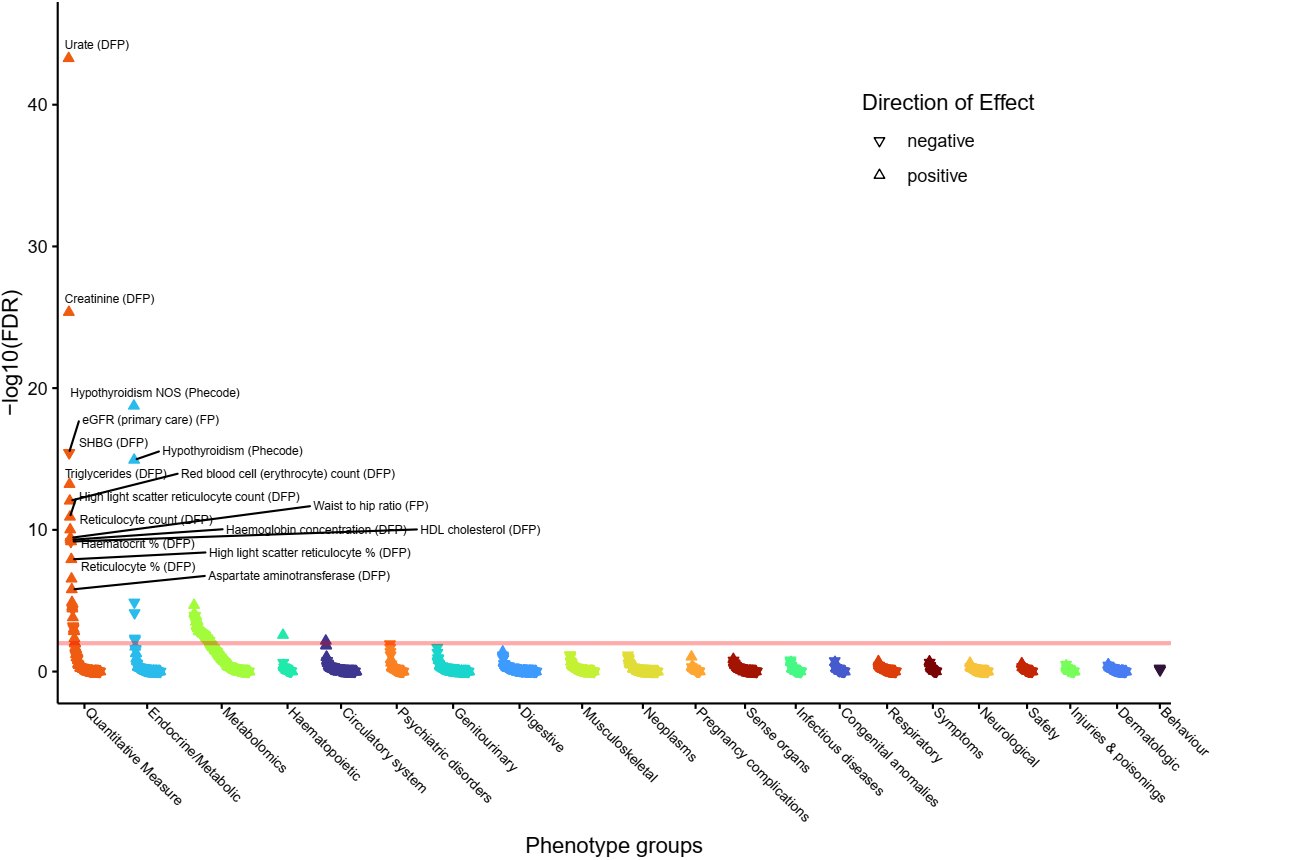
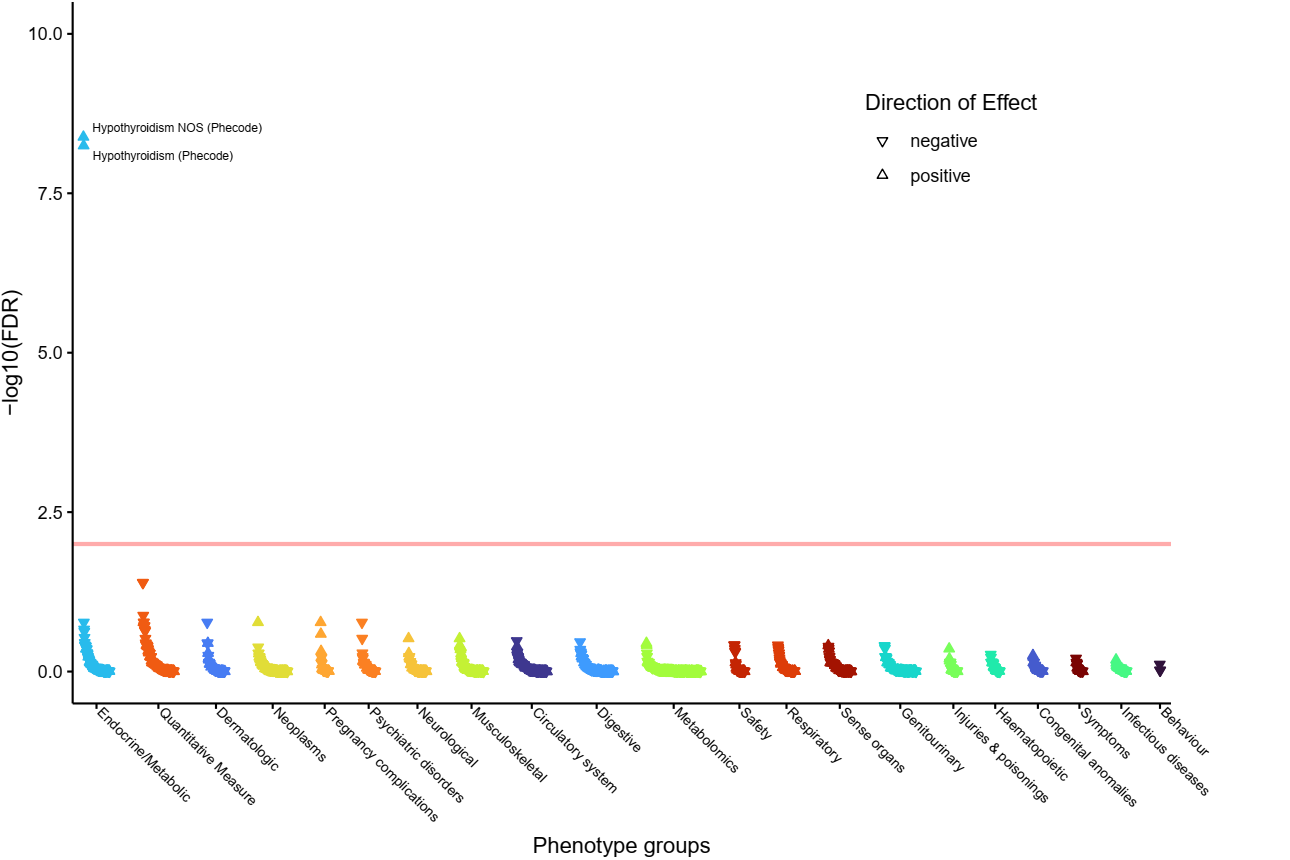
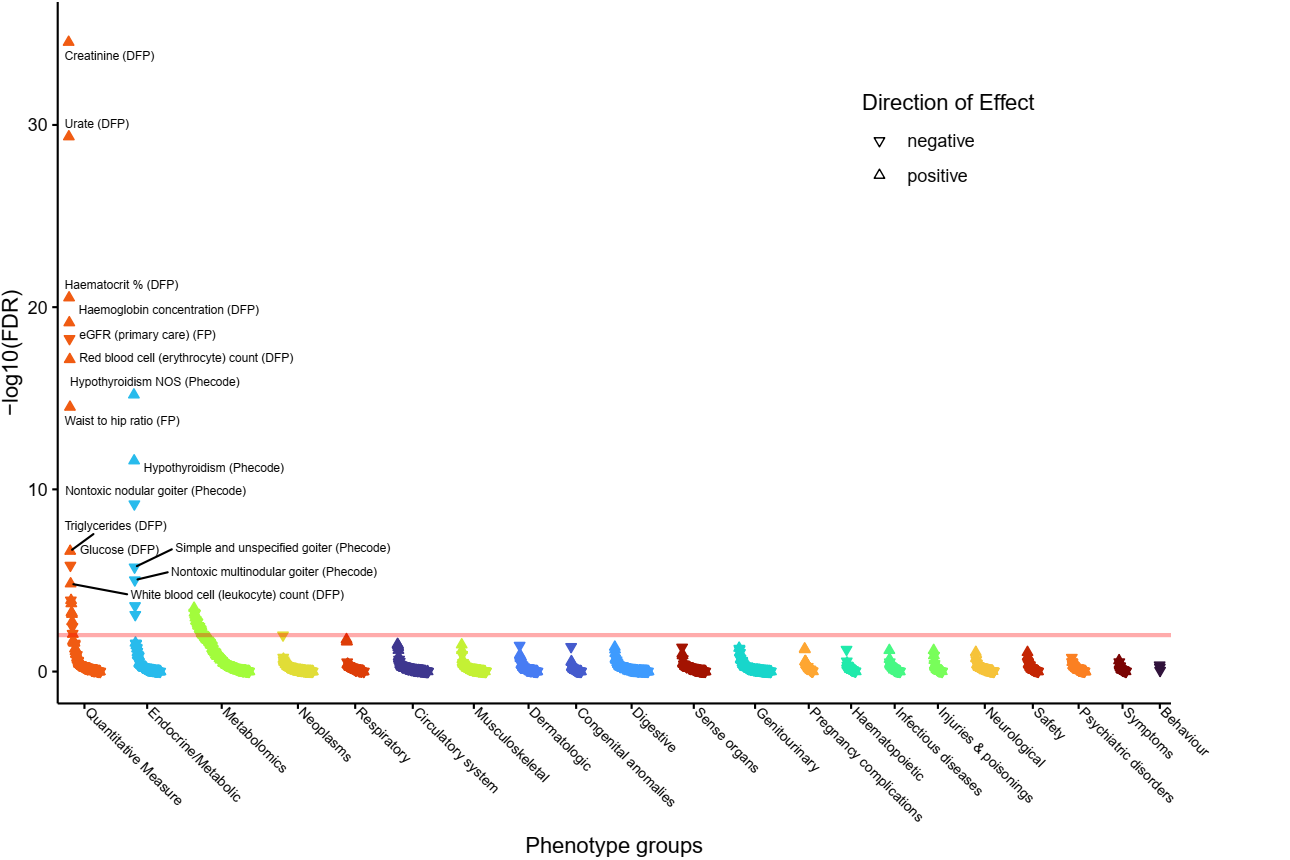
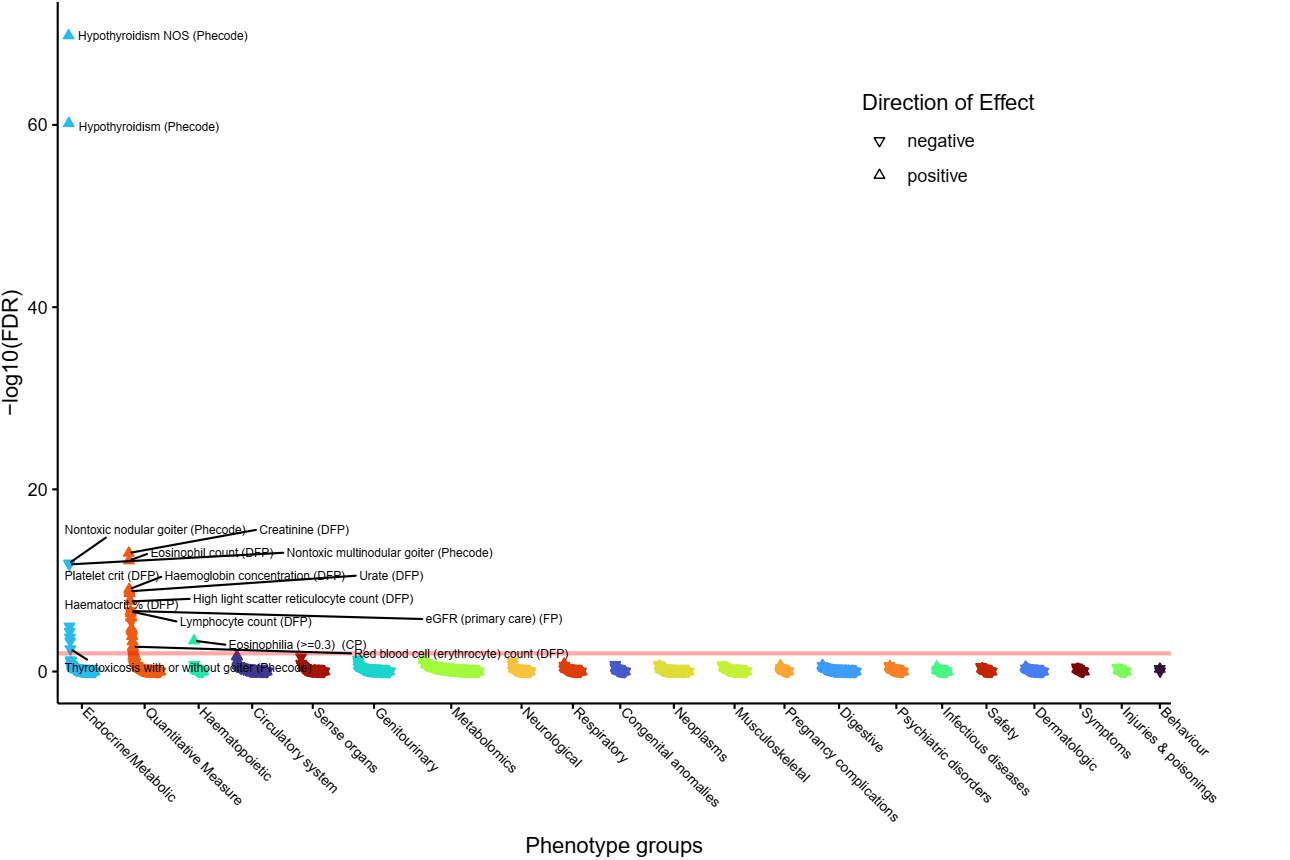
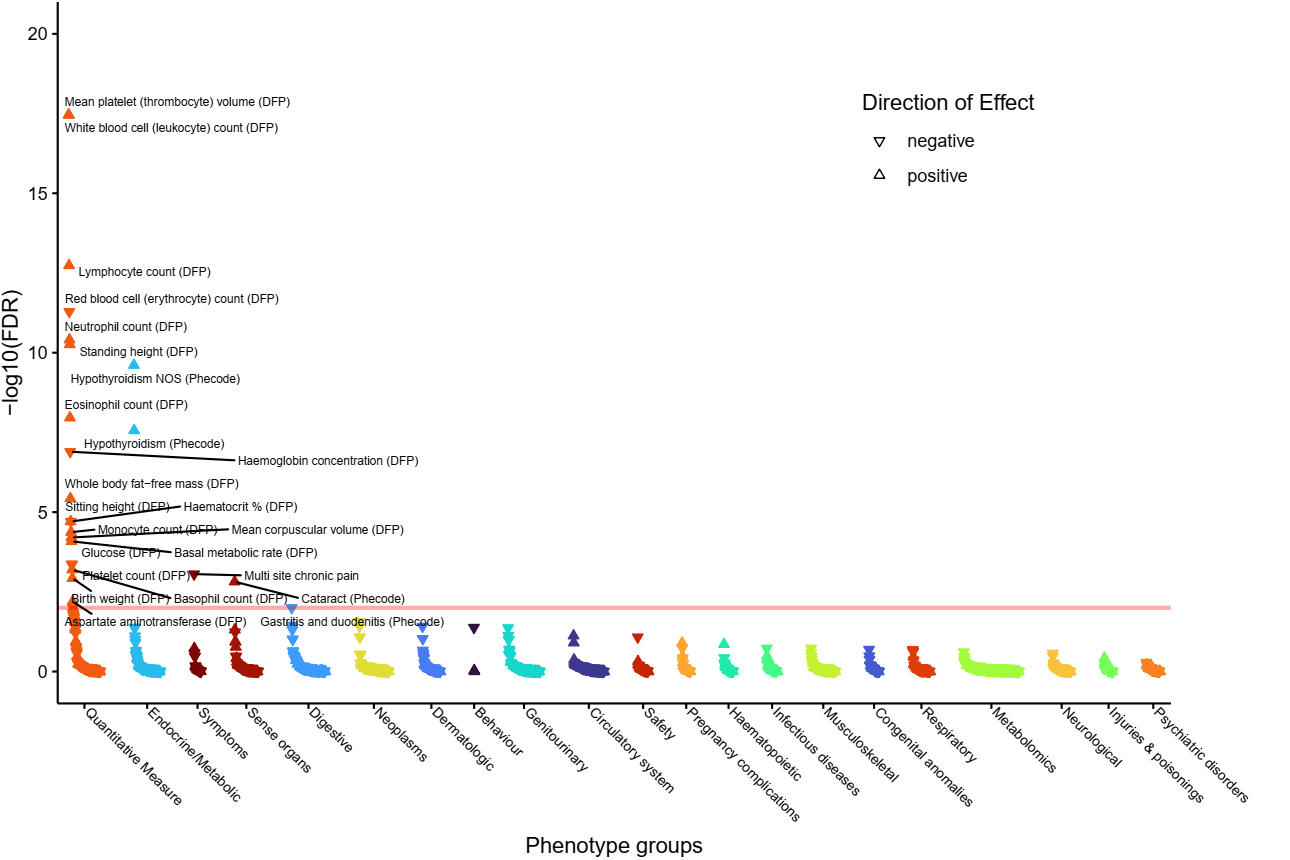
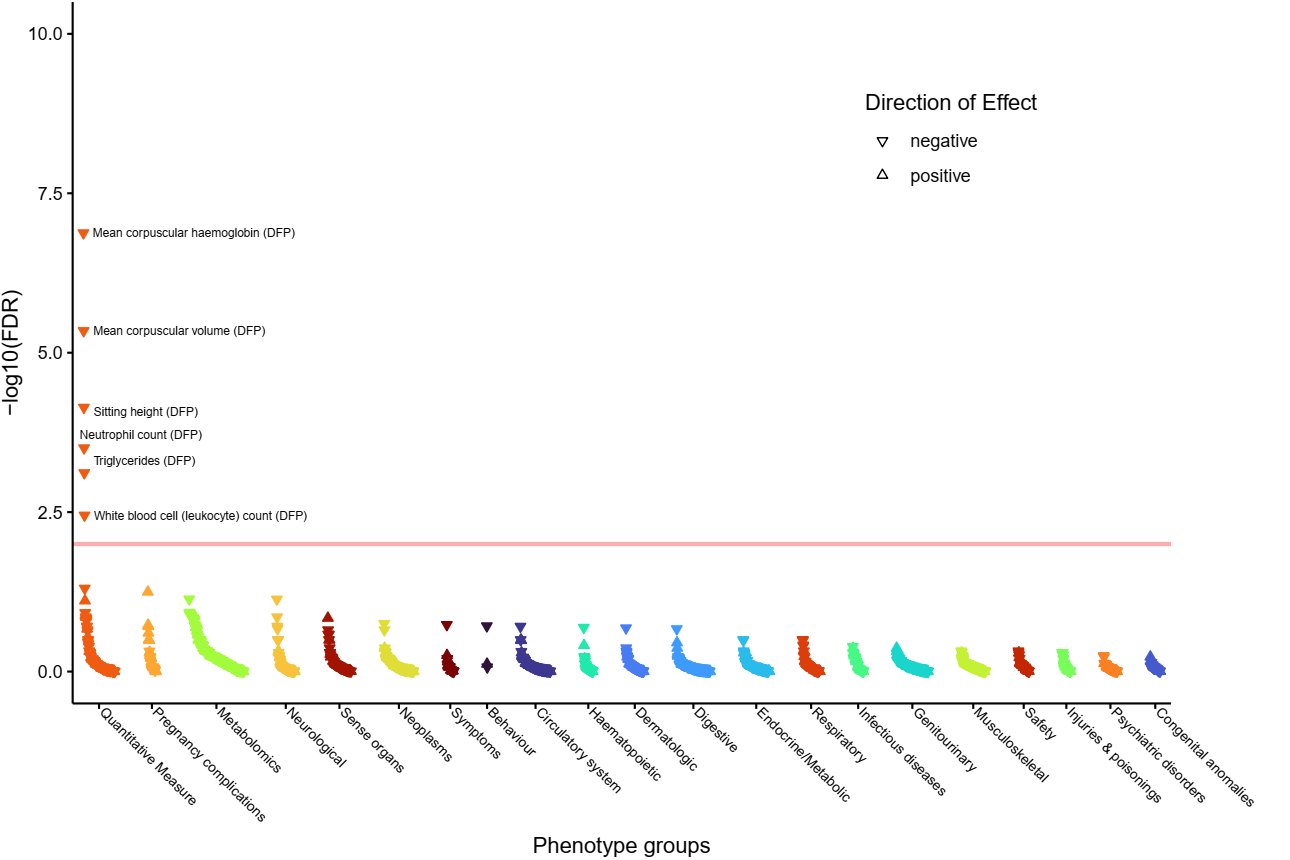
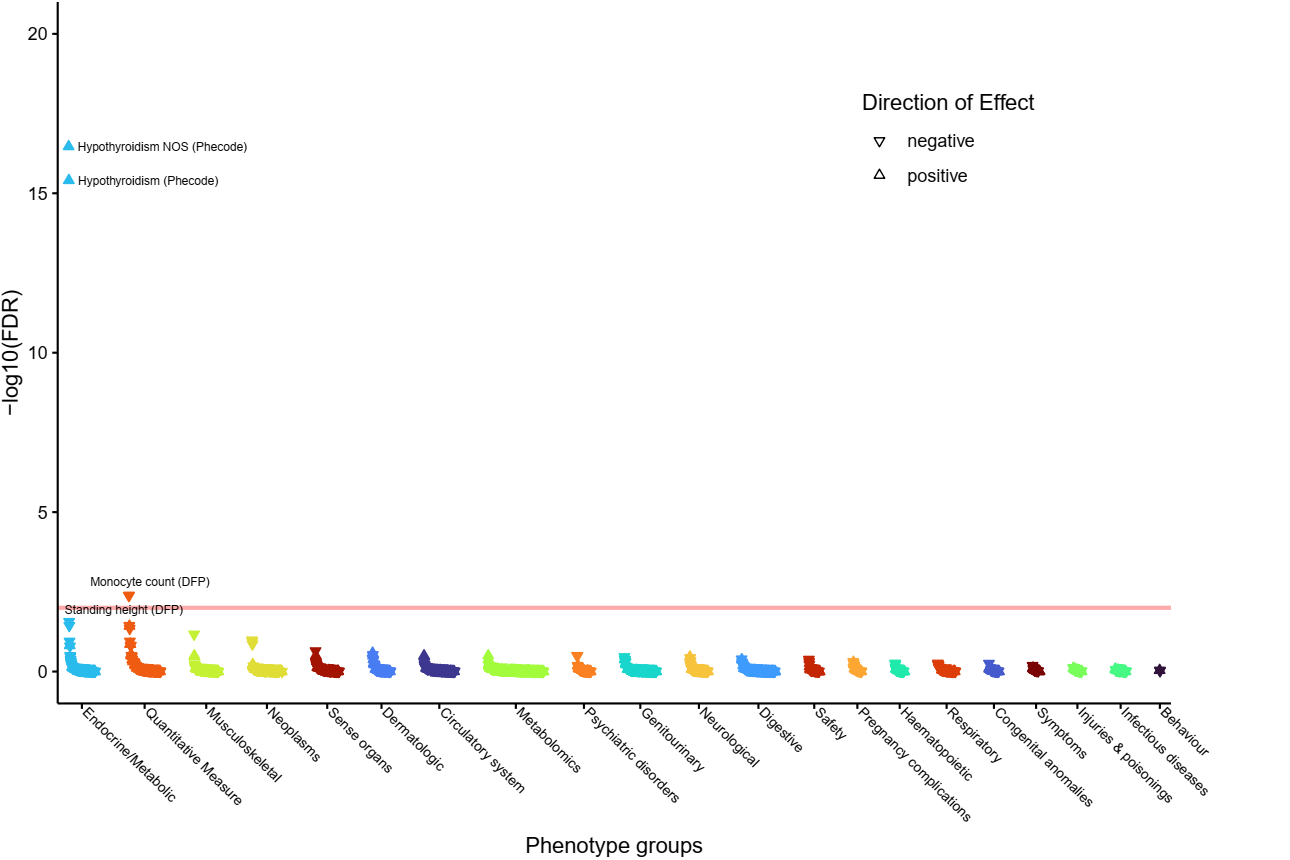
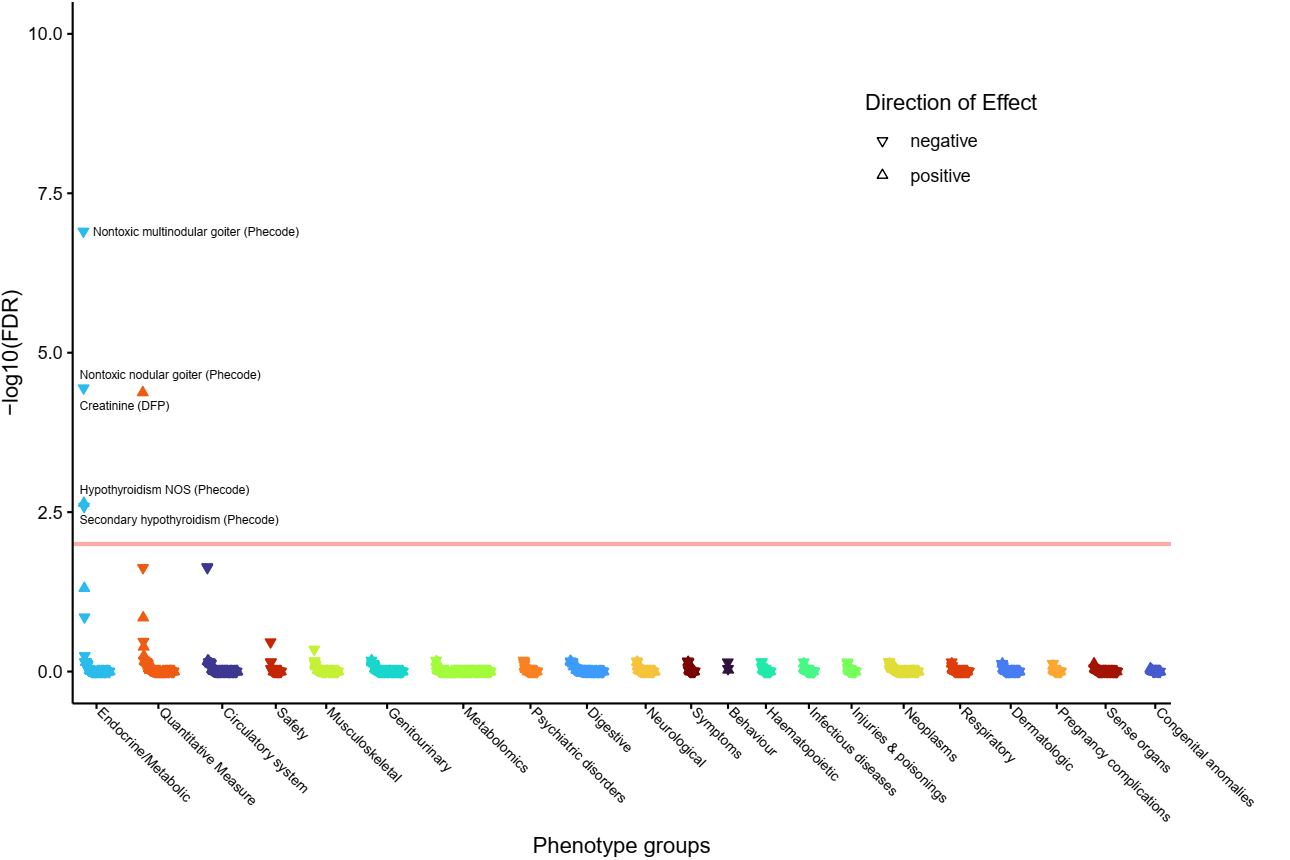
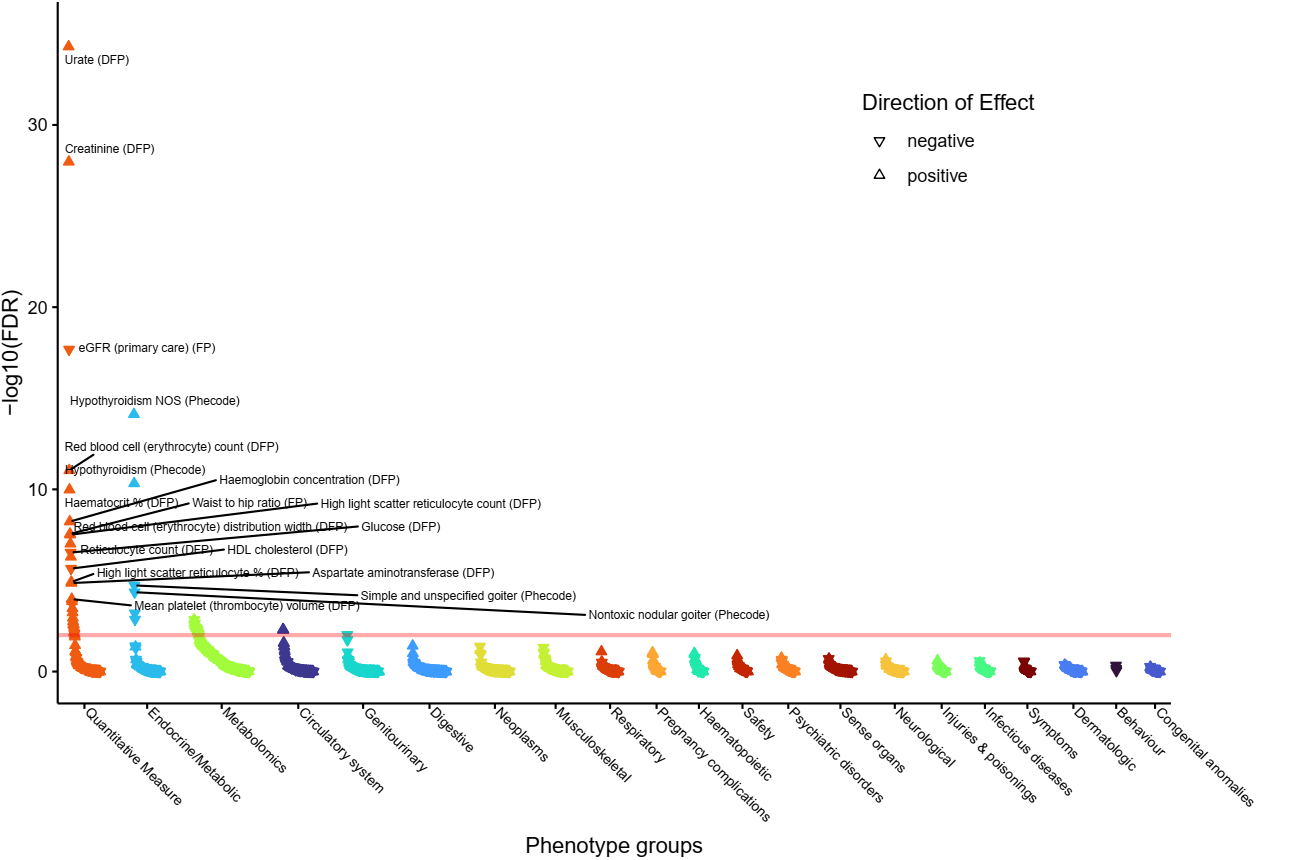
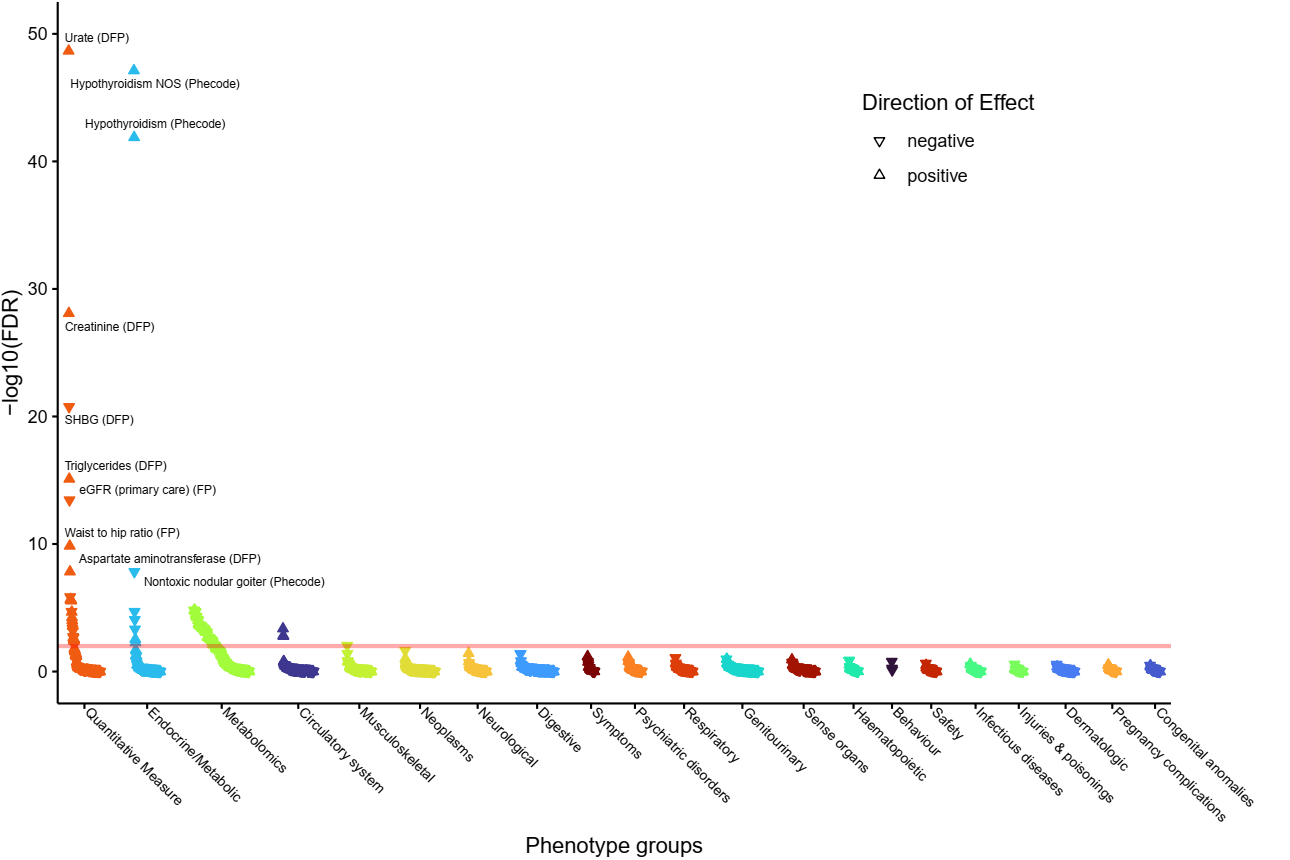
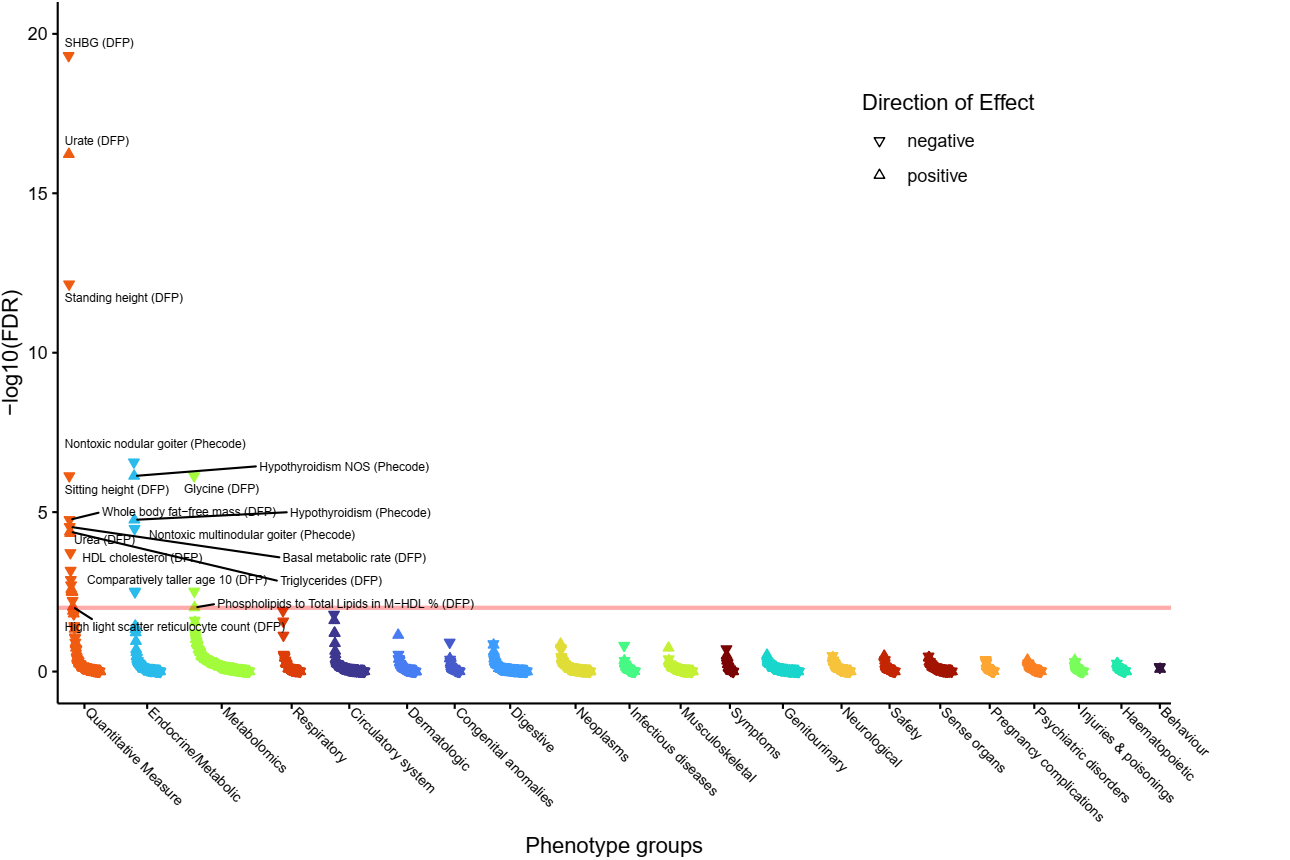
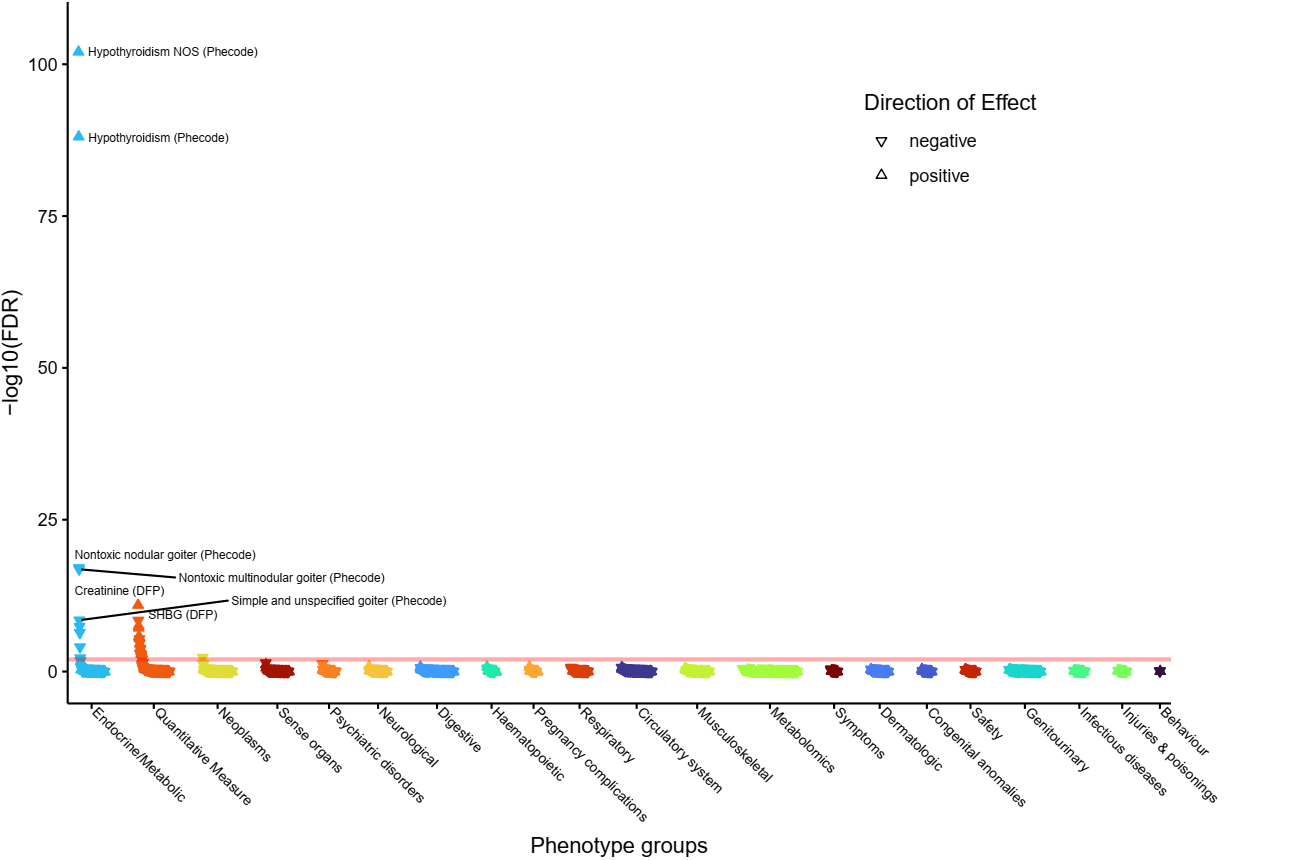
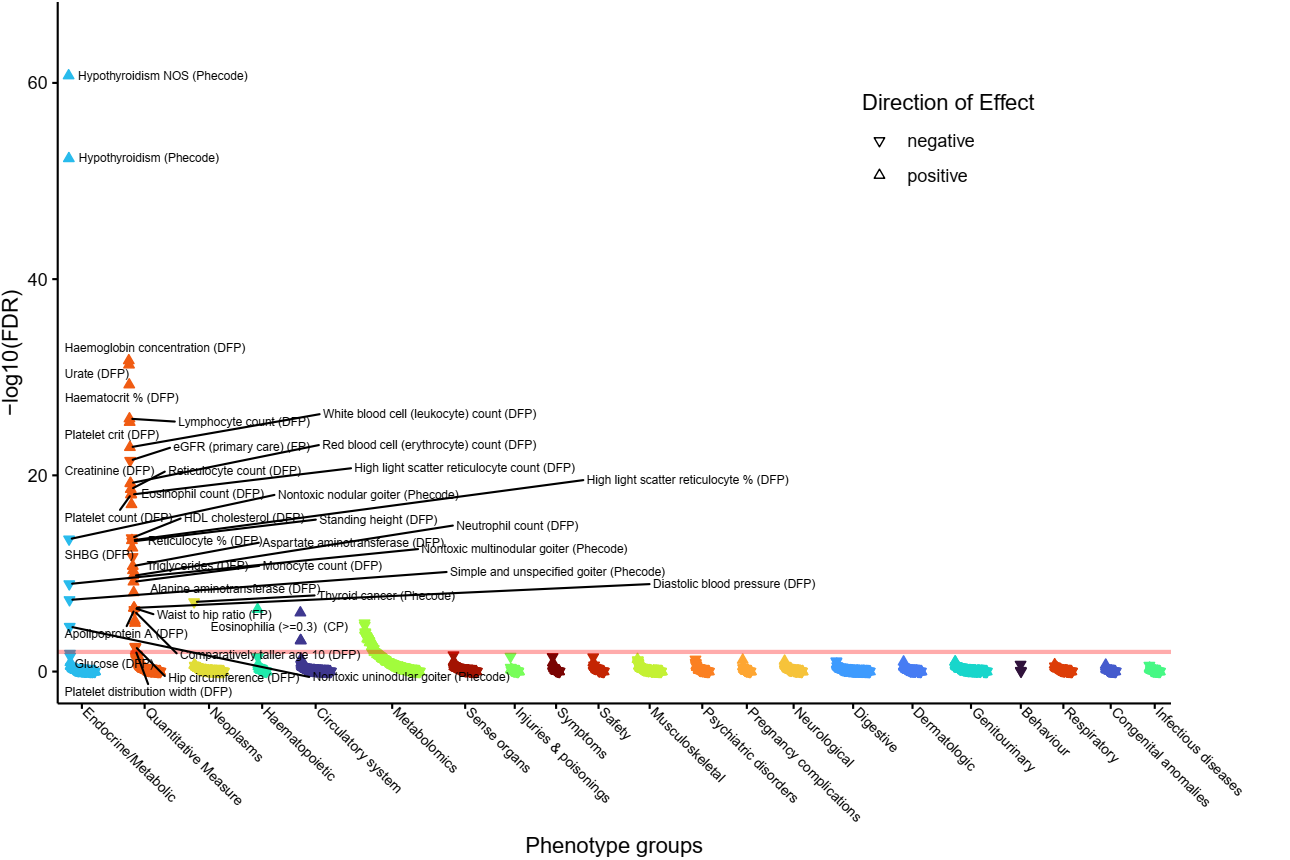
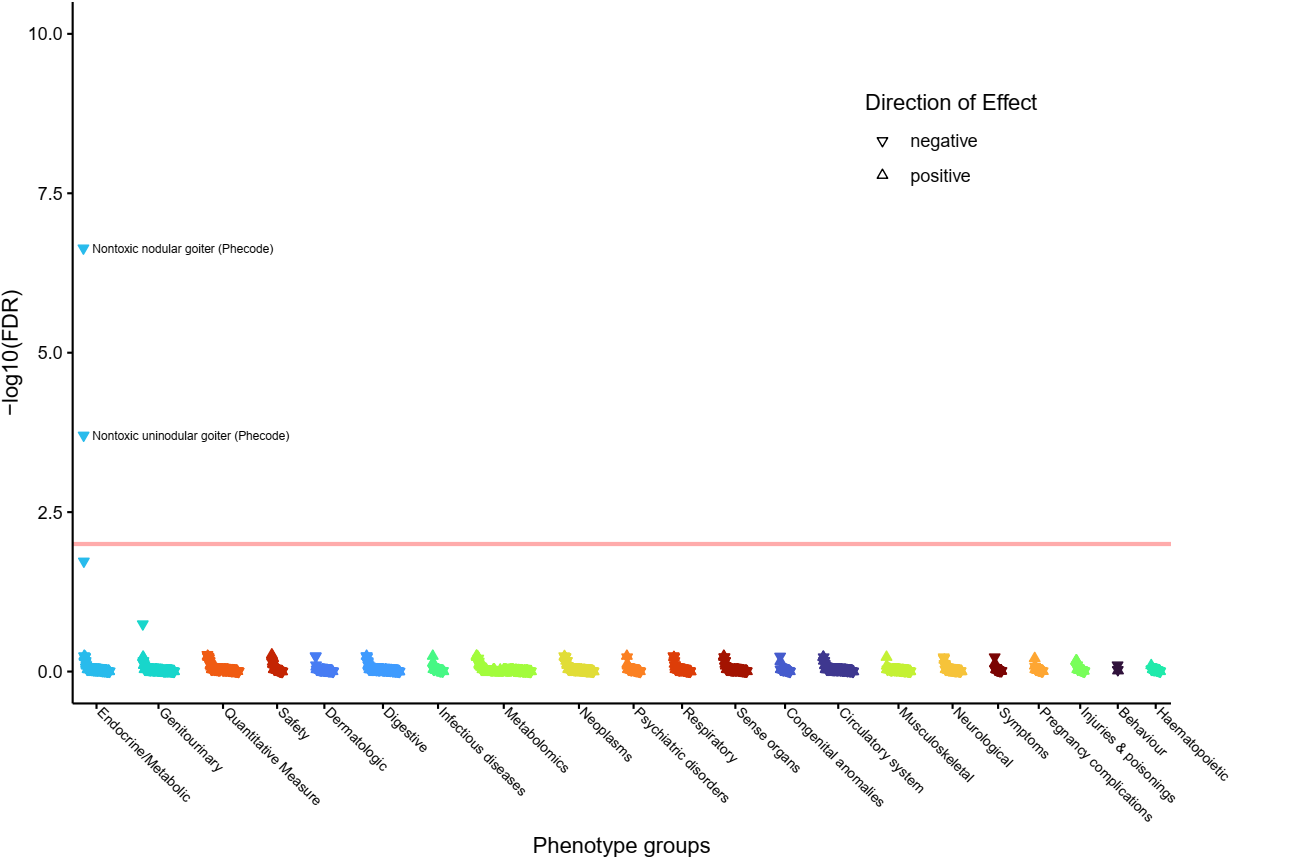
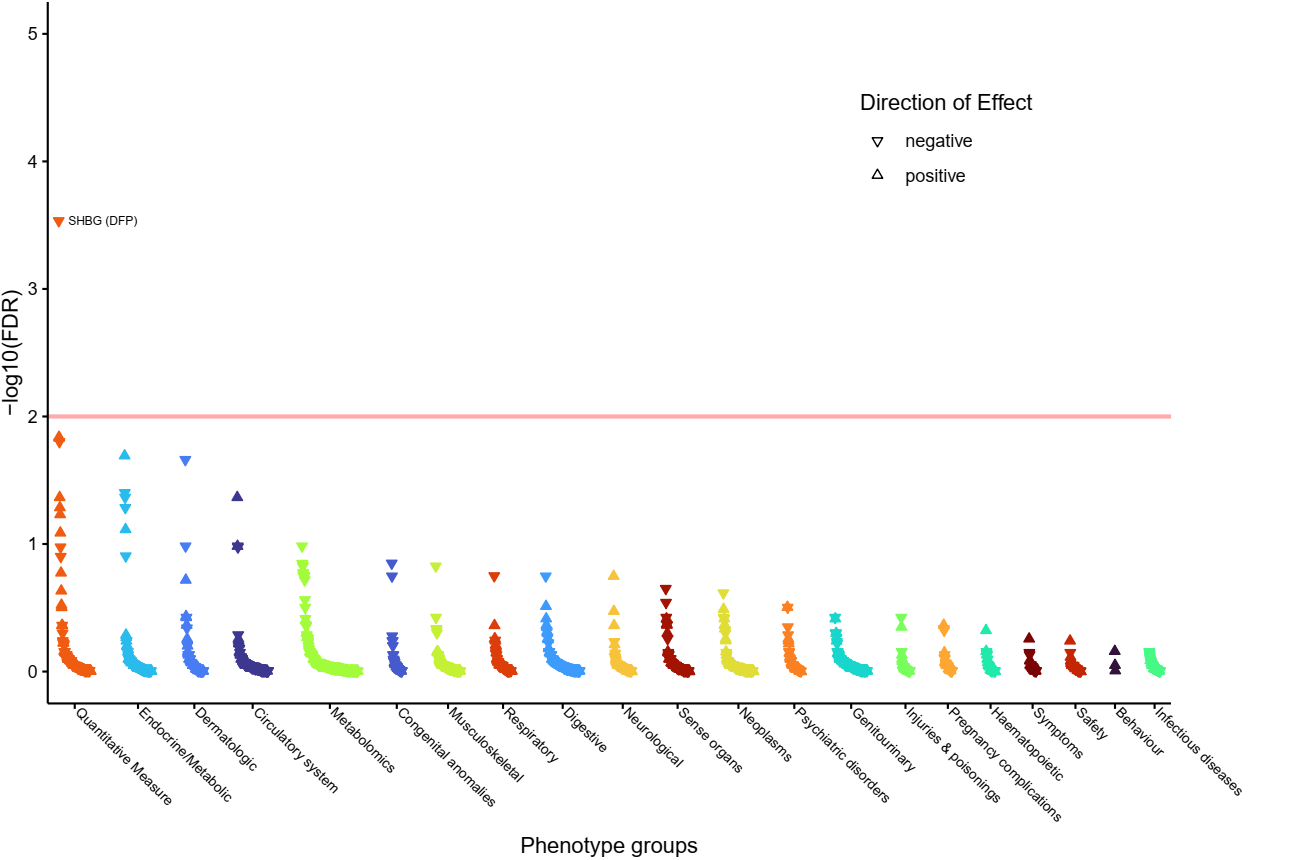
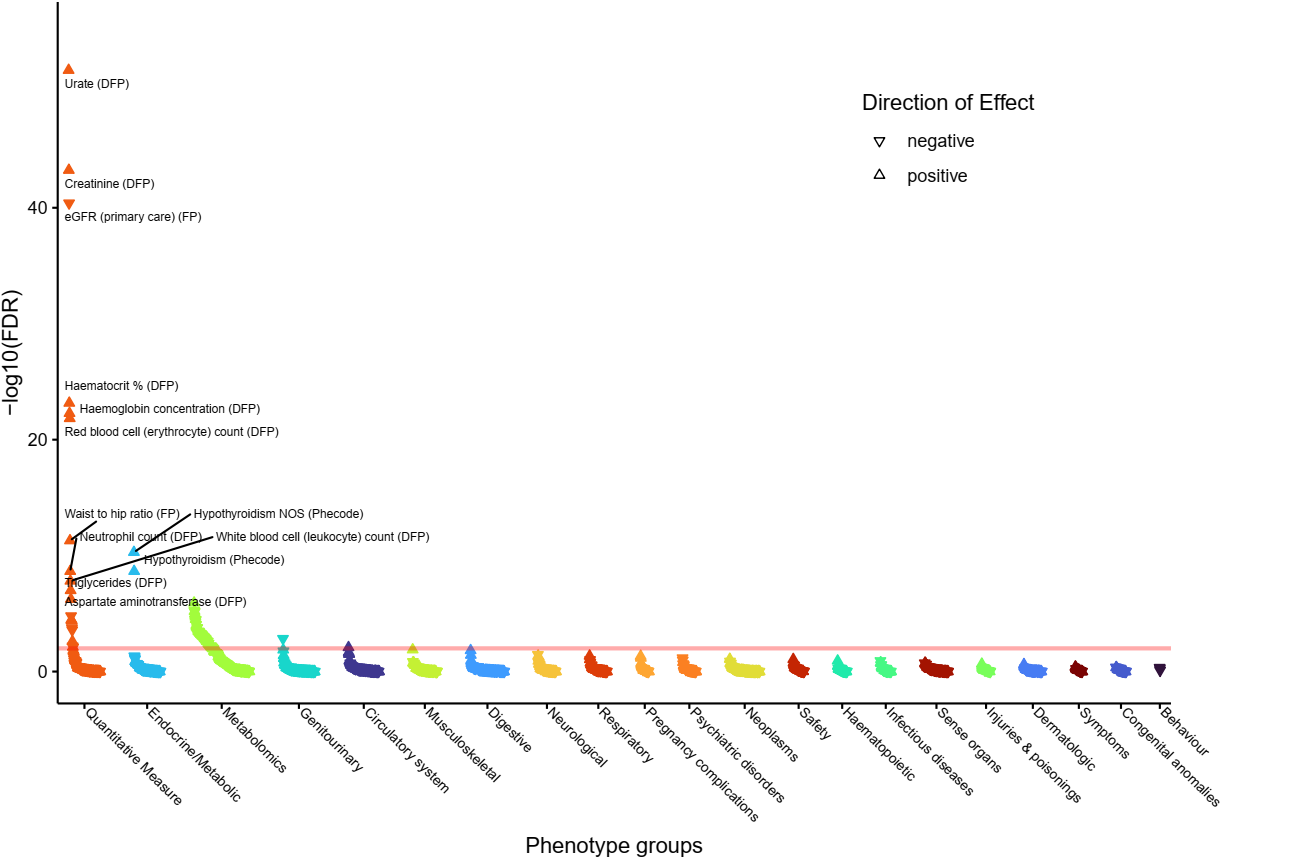
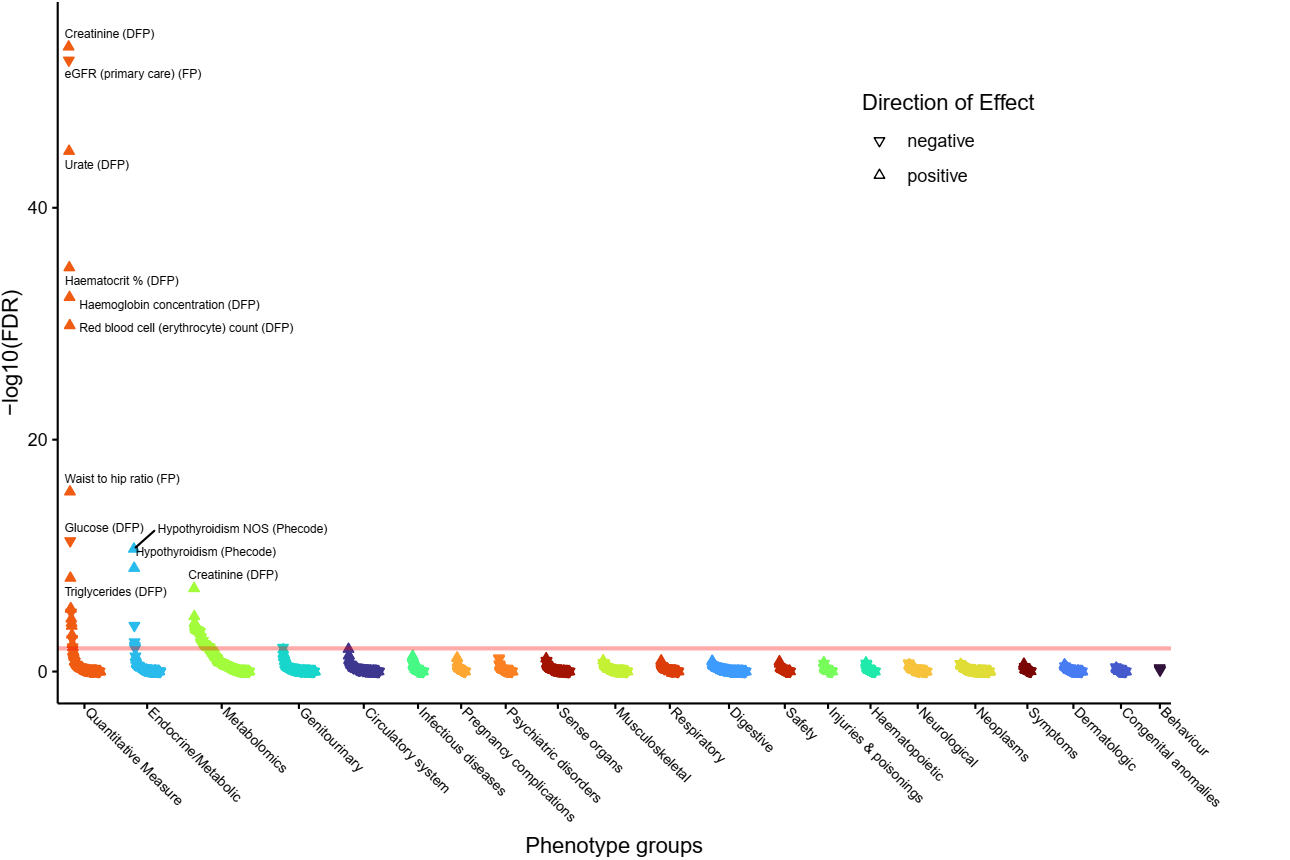
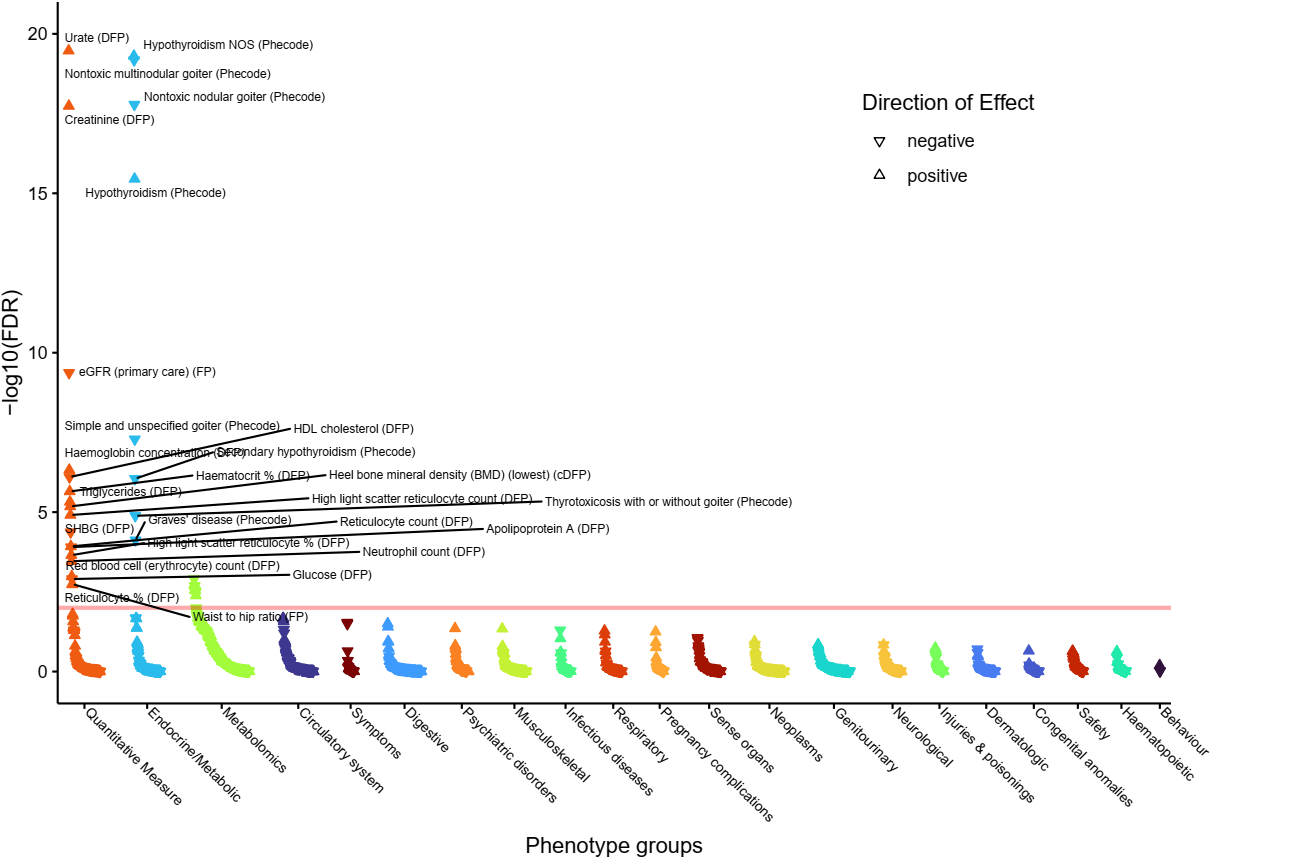
# Supplementary Figures

Supplementary Figure 1: 112 genes prioritised by two or more variant-to-gene criteria. The first seven columns indicate that at least one variant implicates the corresponding gene via the evidence for that column. The remaining seven columns indicate the strength of association of the most significant variant implicating the corresponding gene as causal with respect to the TSH increasing allele, such that shades of blue represent associations with the other thyroid phenotypes that have the same direction of effect as the TSH association and shades of red represent an opposite direction of effect to the TSH association.



**Supplementary Figure 2**: Pathway-specific polygenic score phenome-wide association studies (PheWAS)

1. **cAMP (KEGG) signalling pathway**  
   Text, application

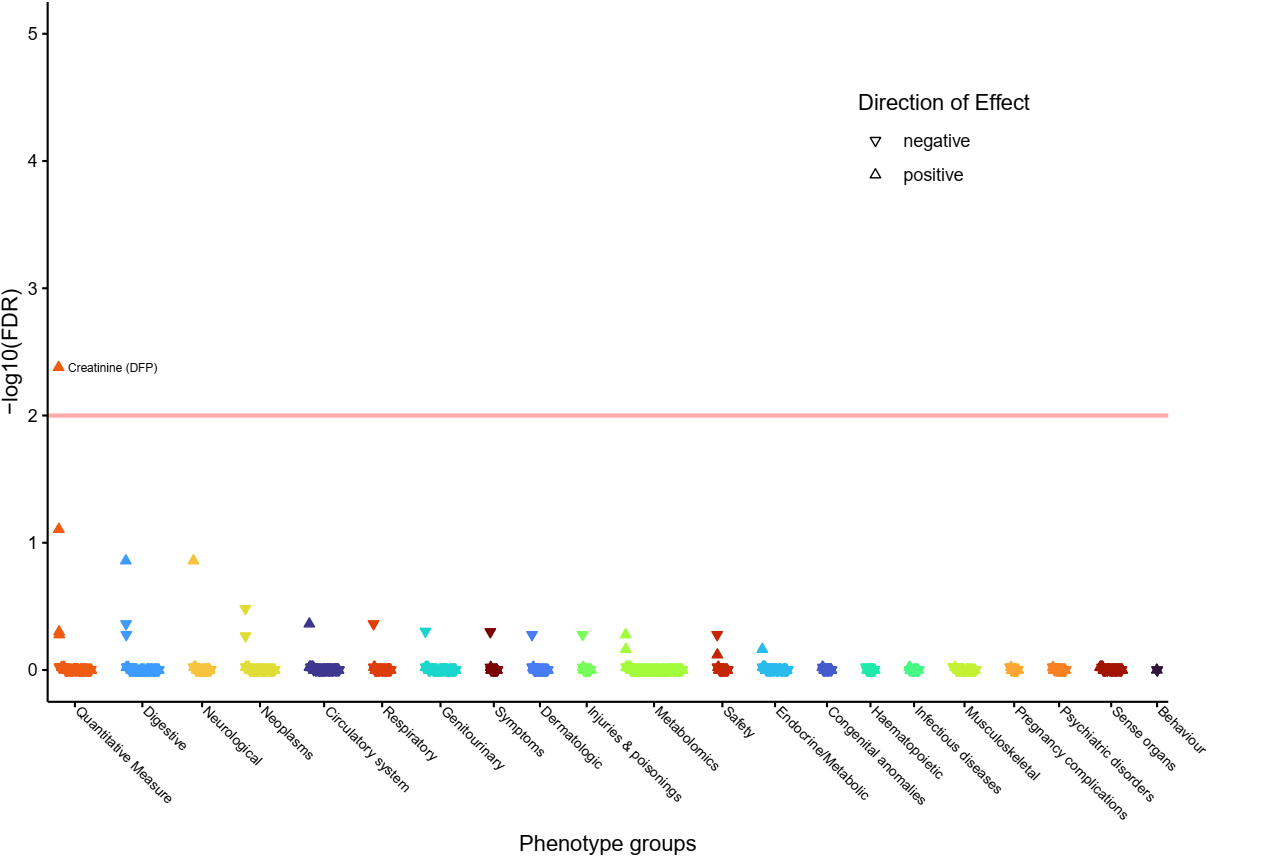
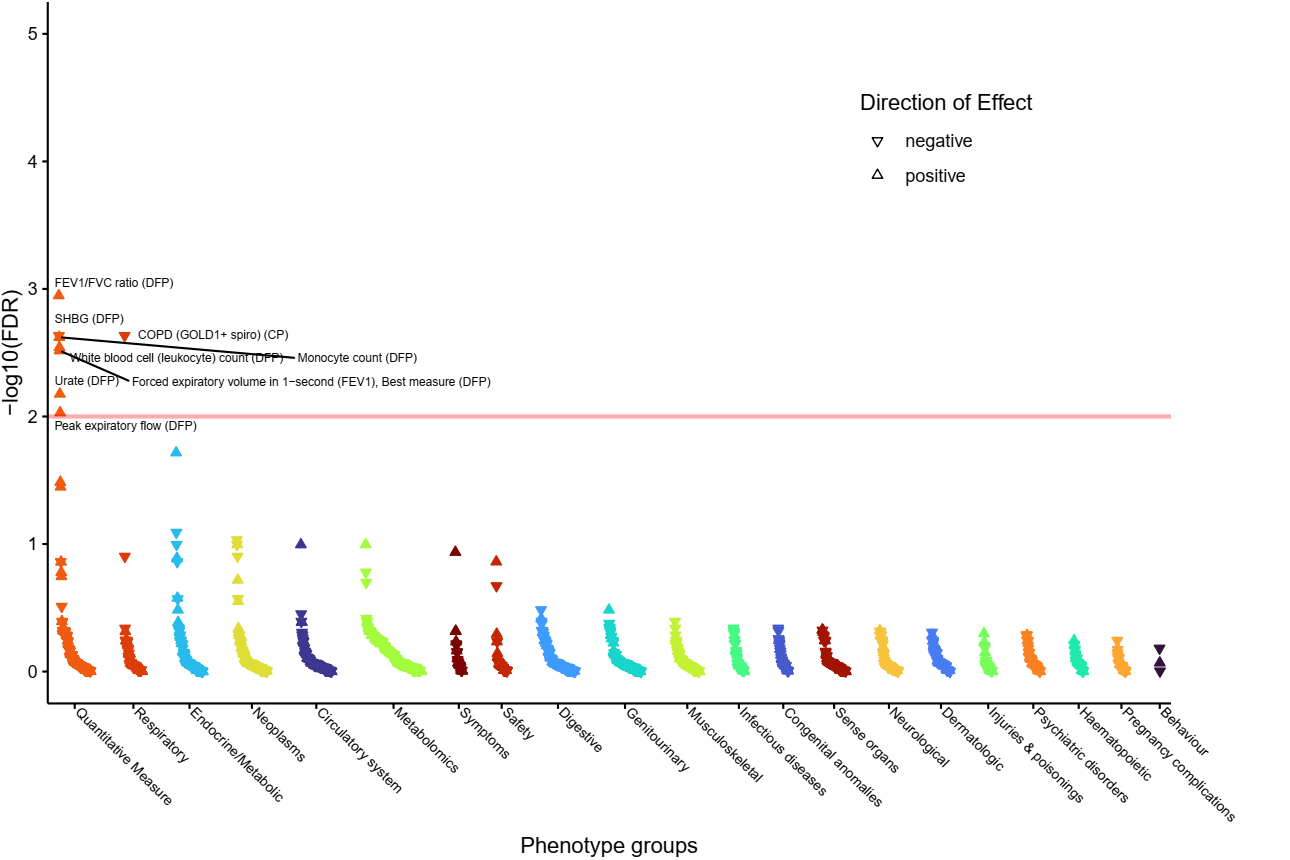
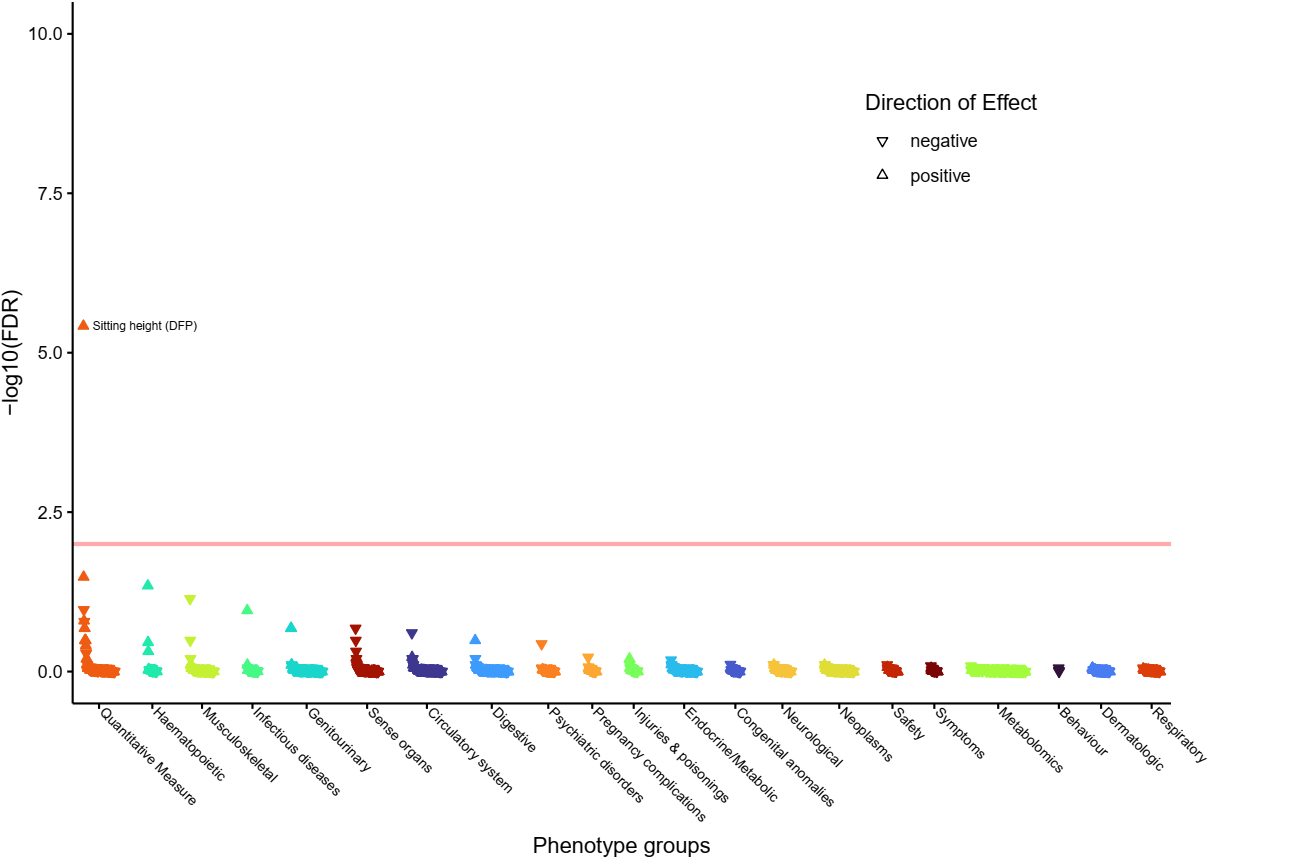
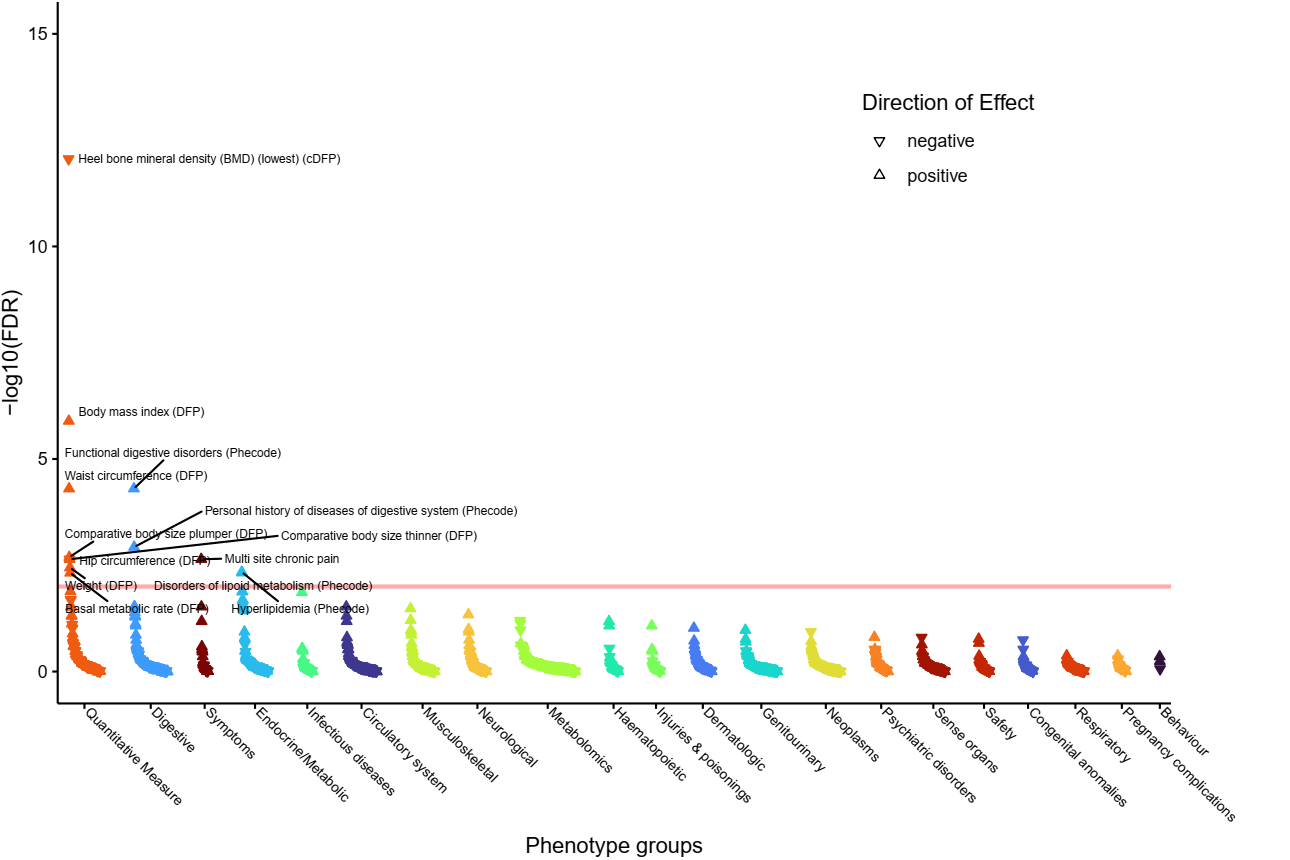
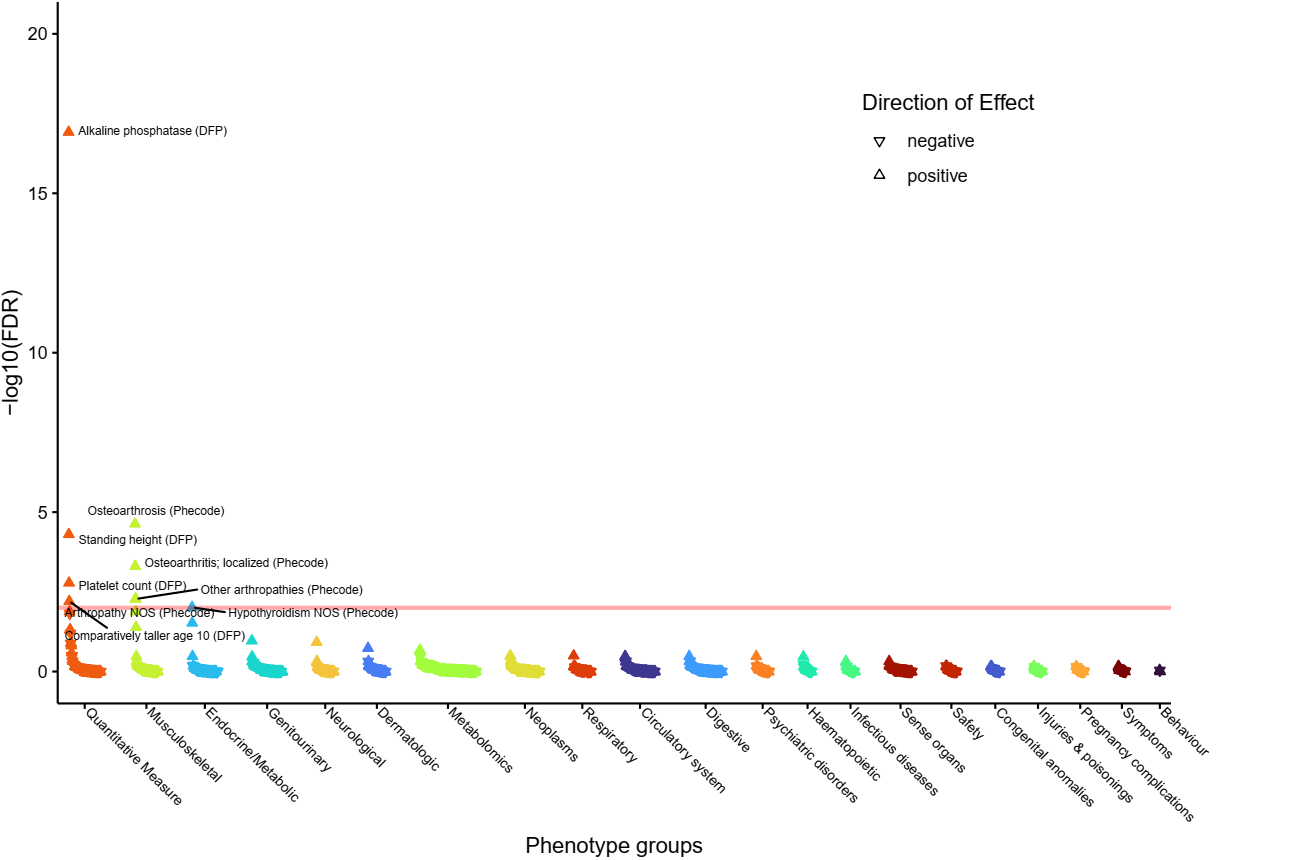
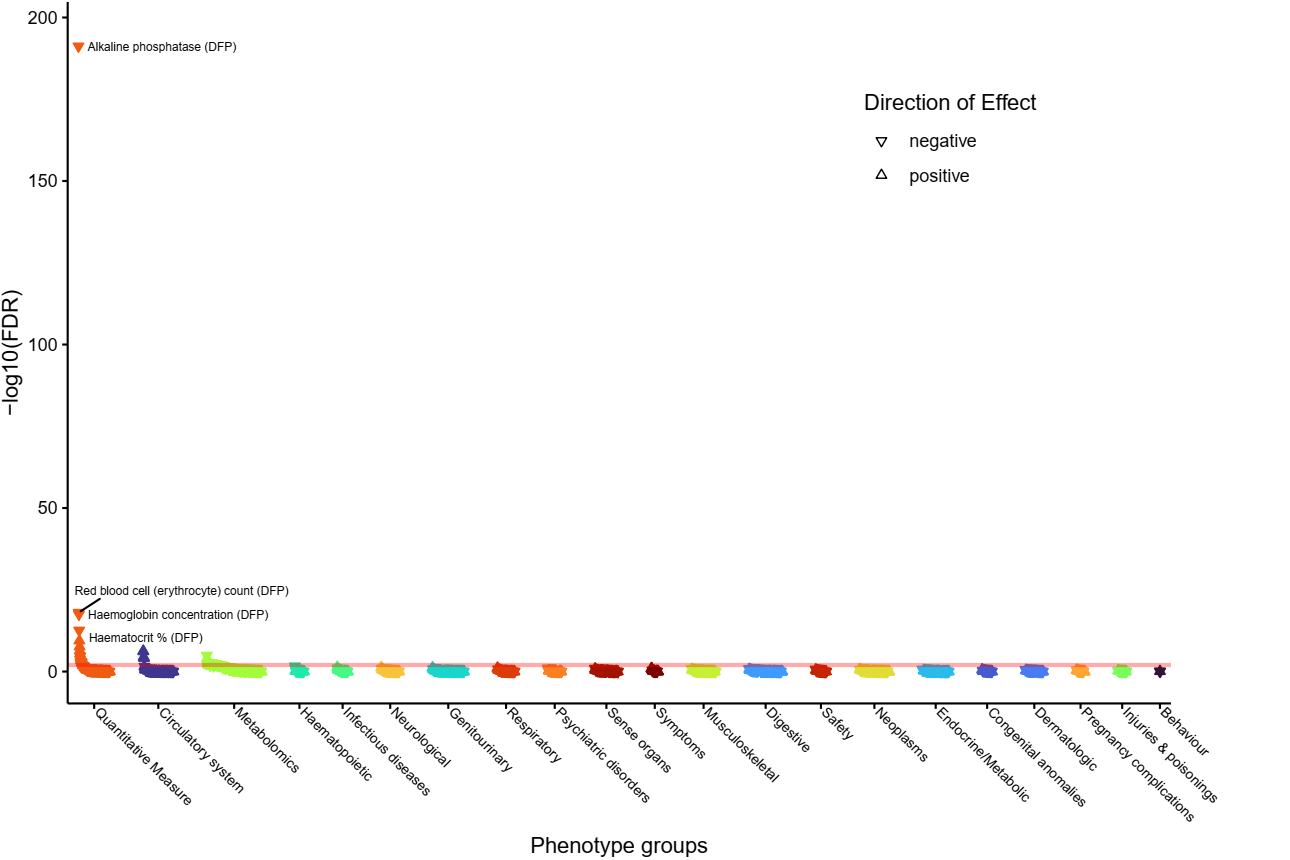
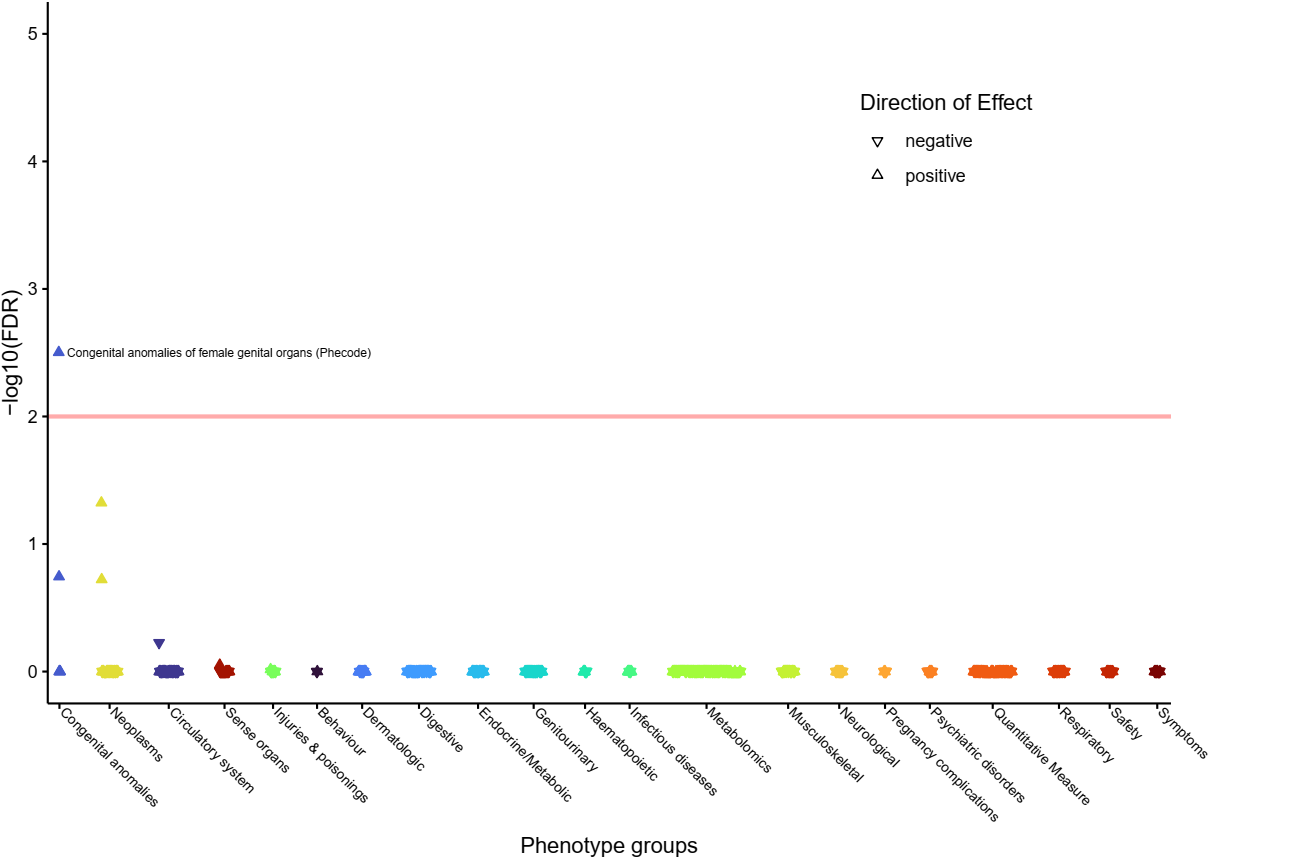
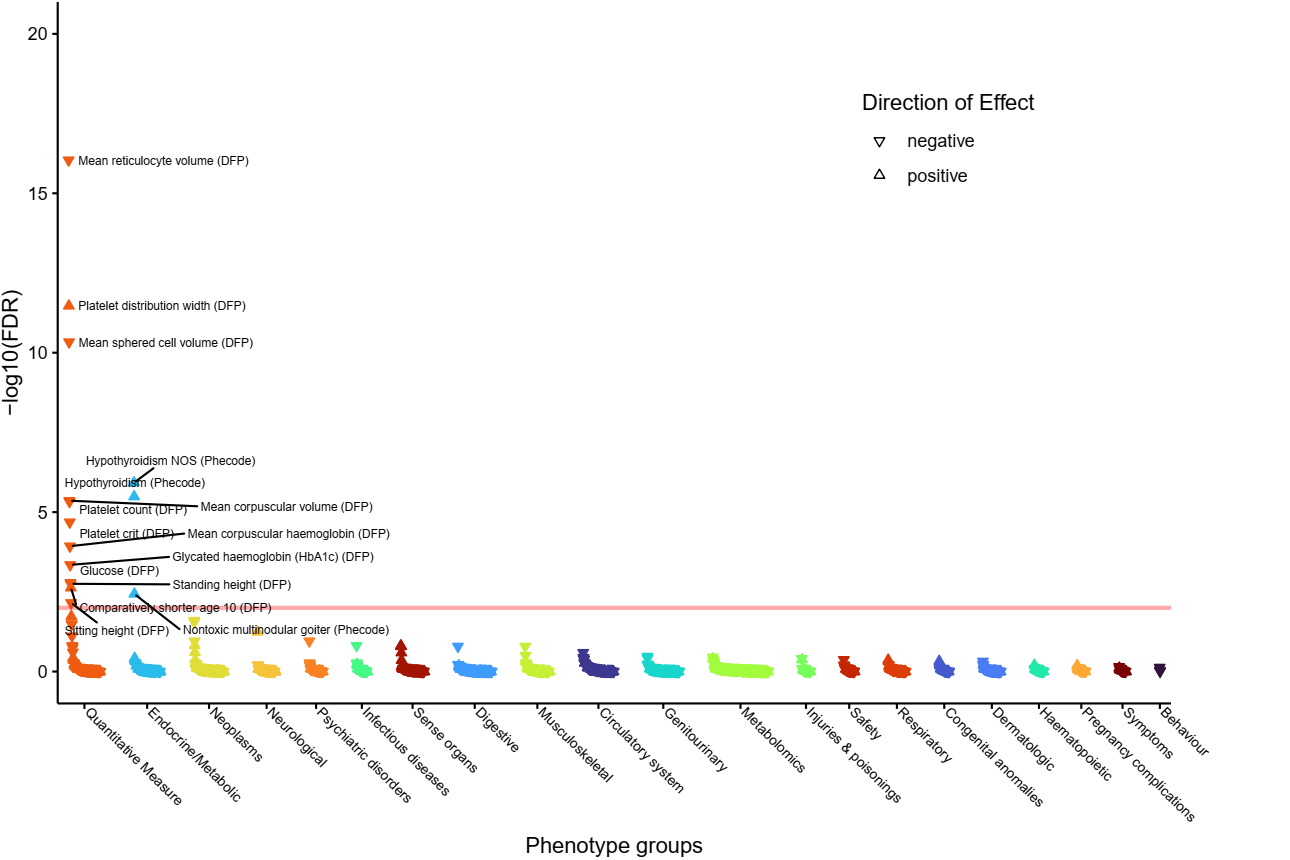
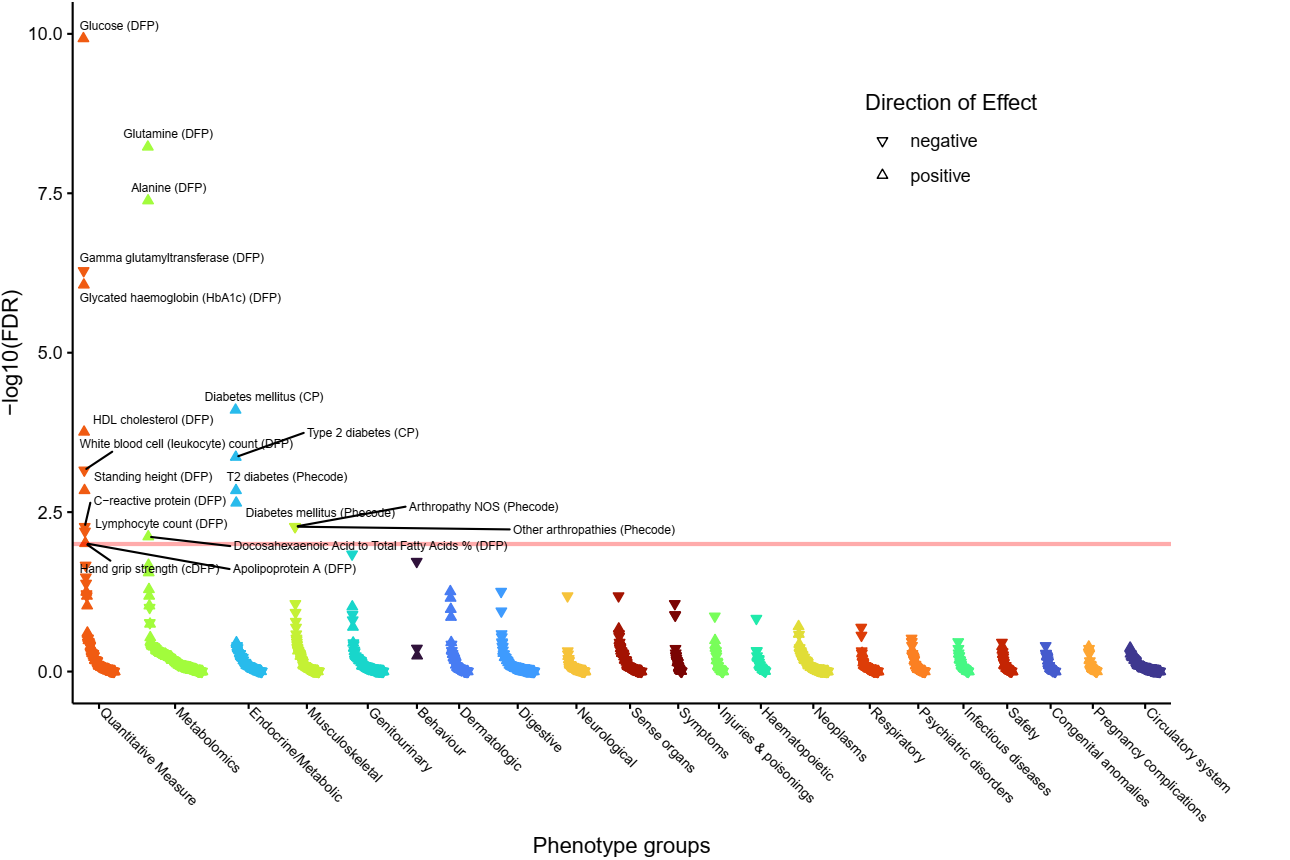
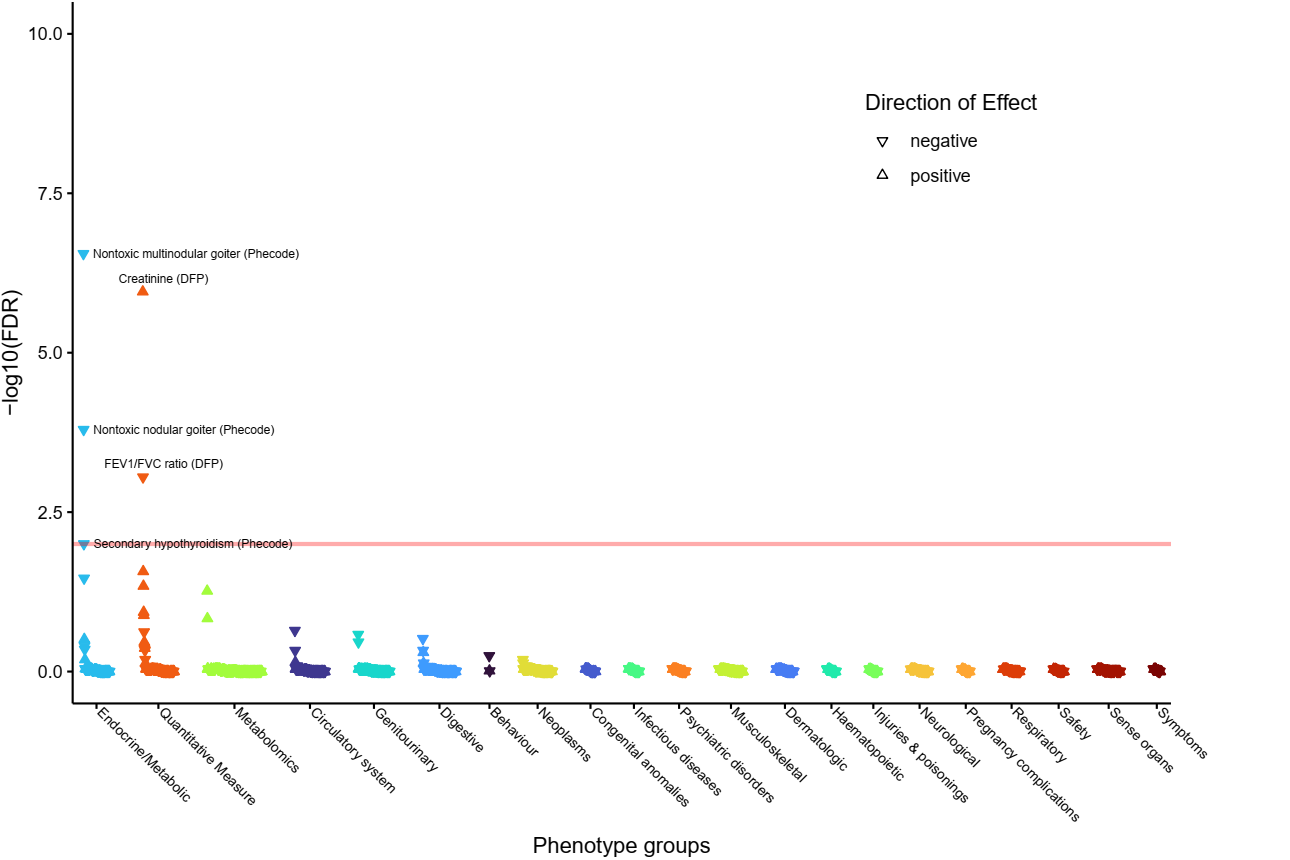
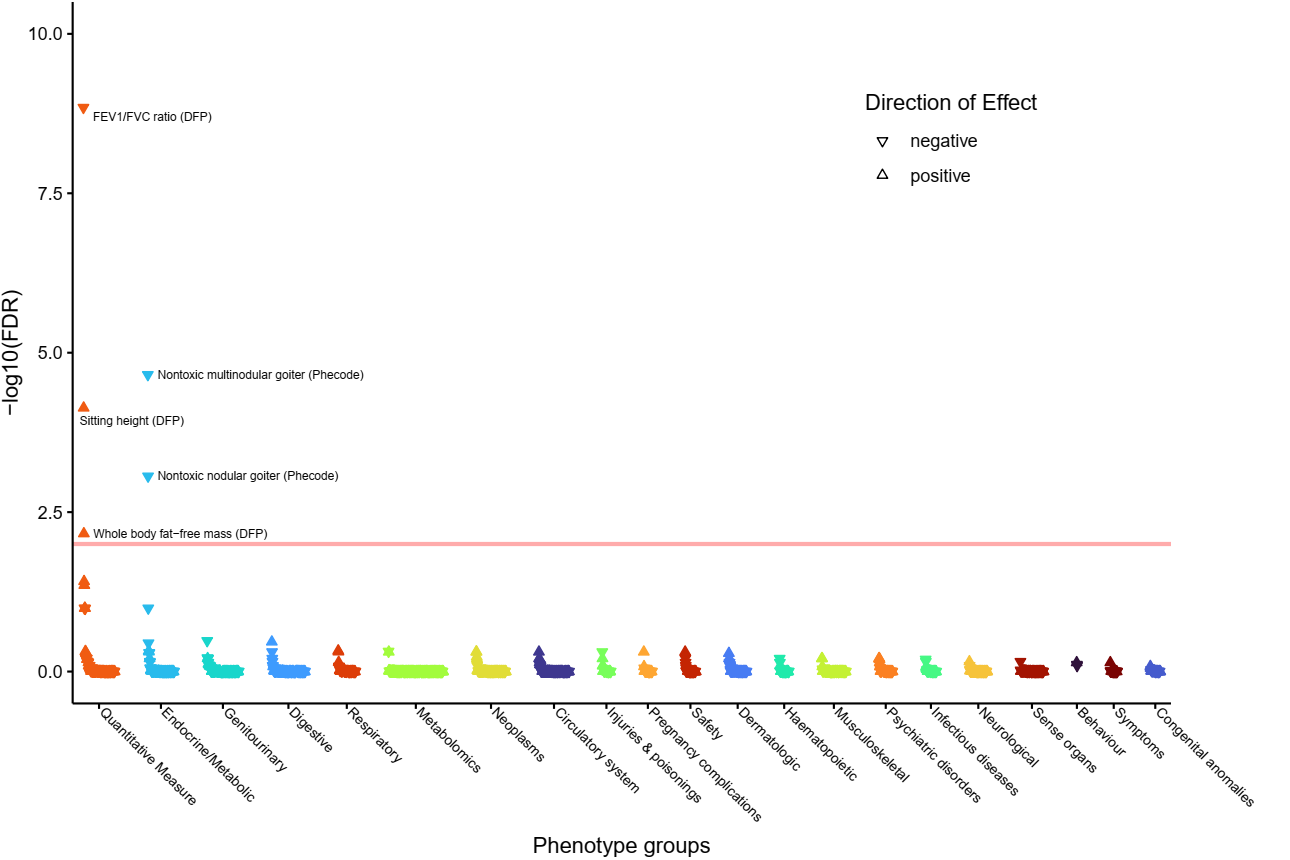
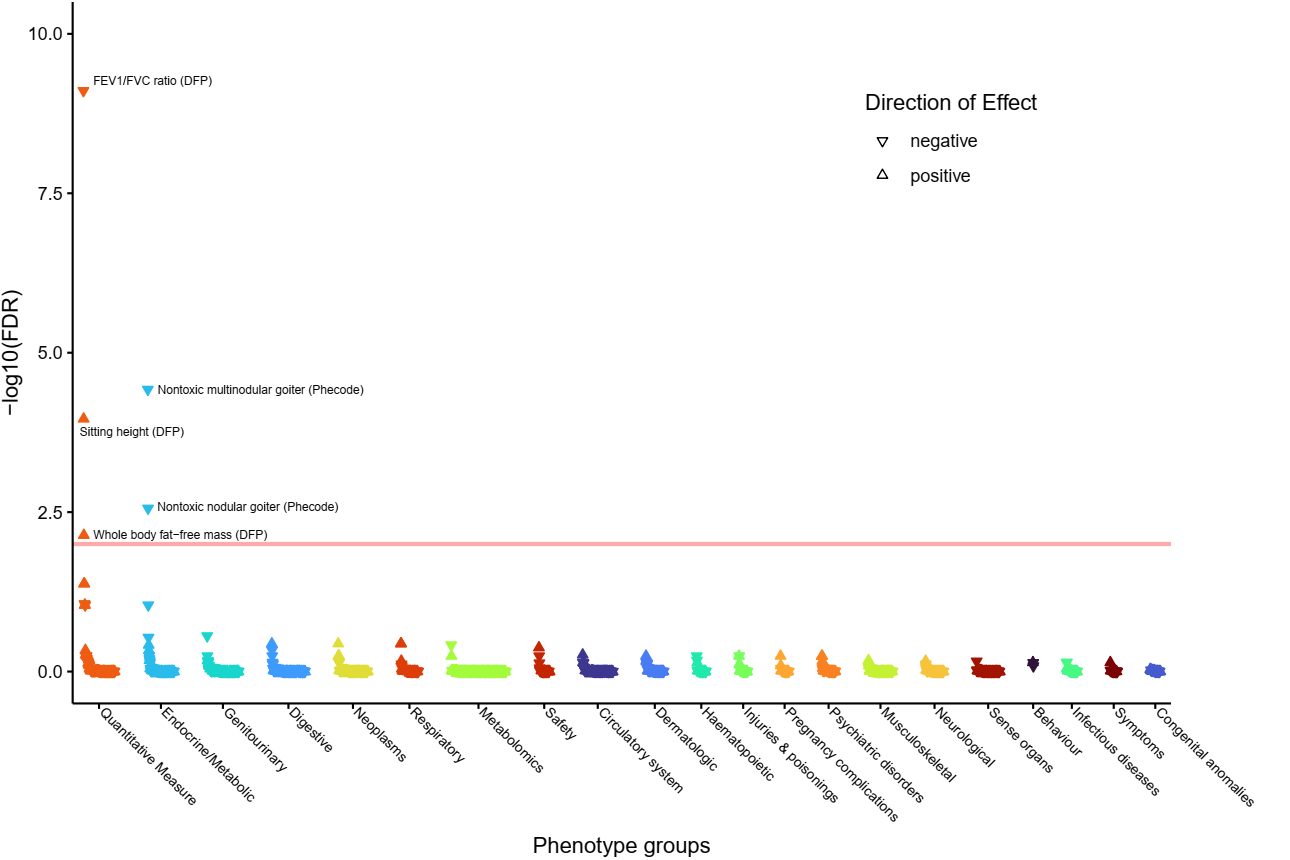
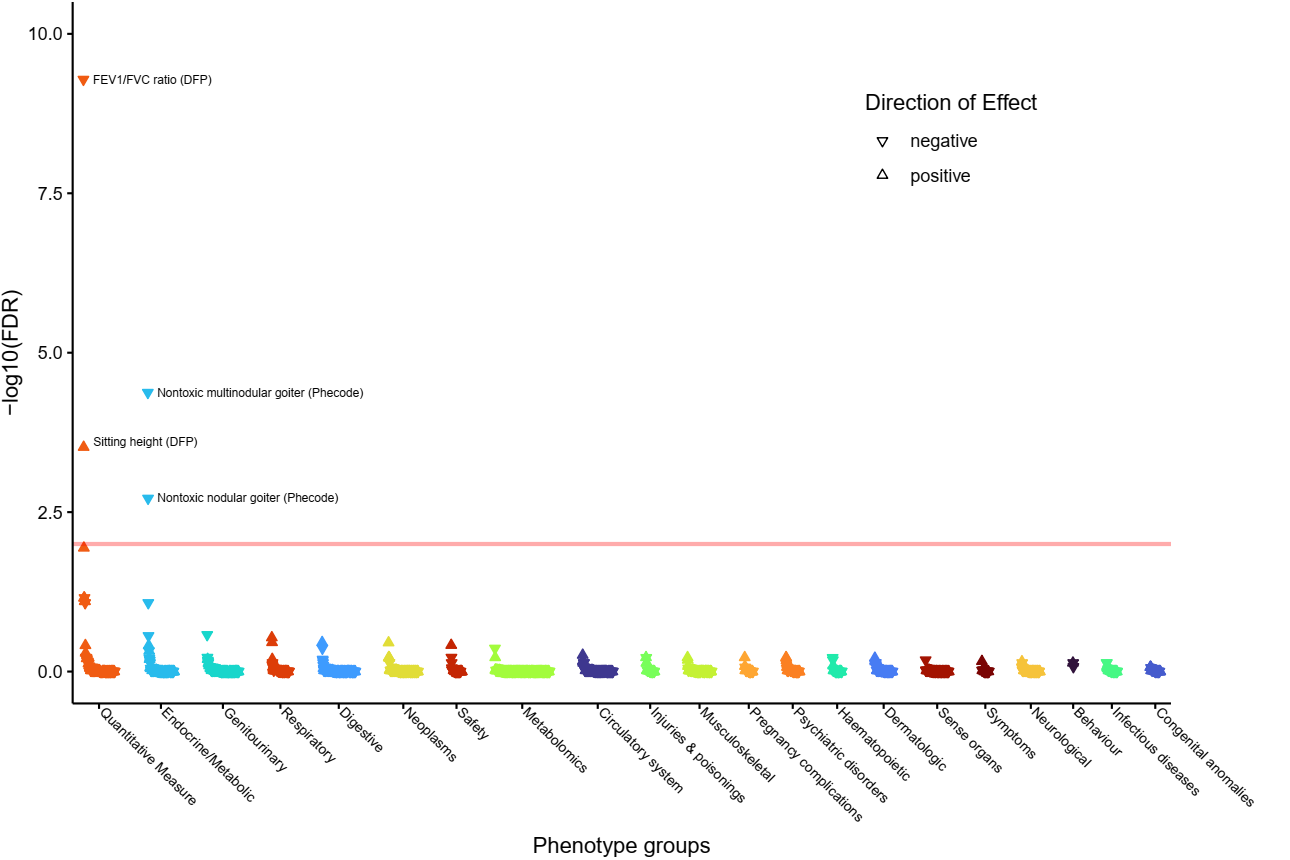
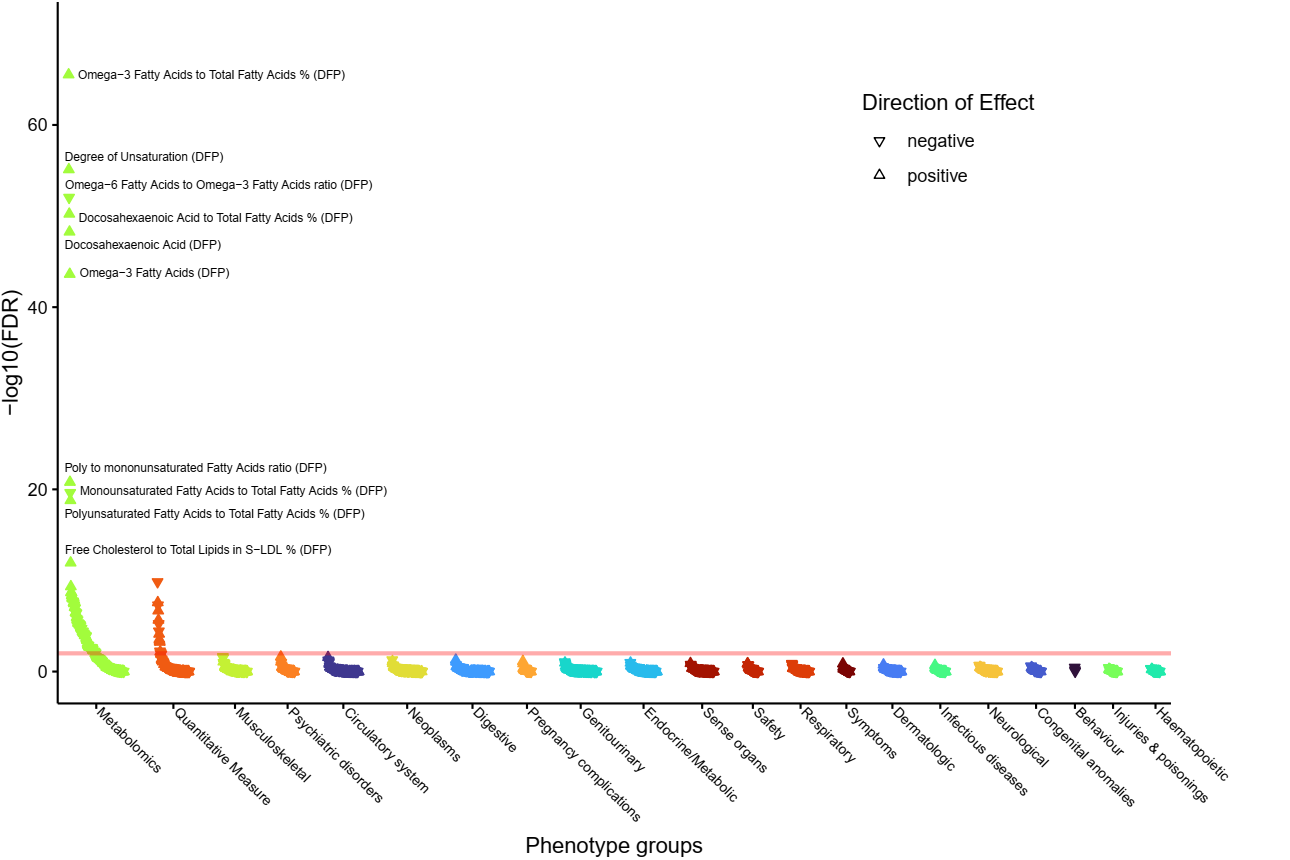
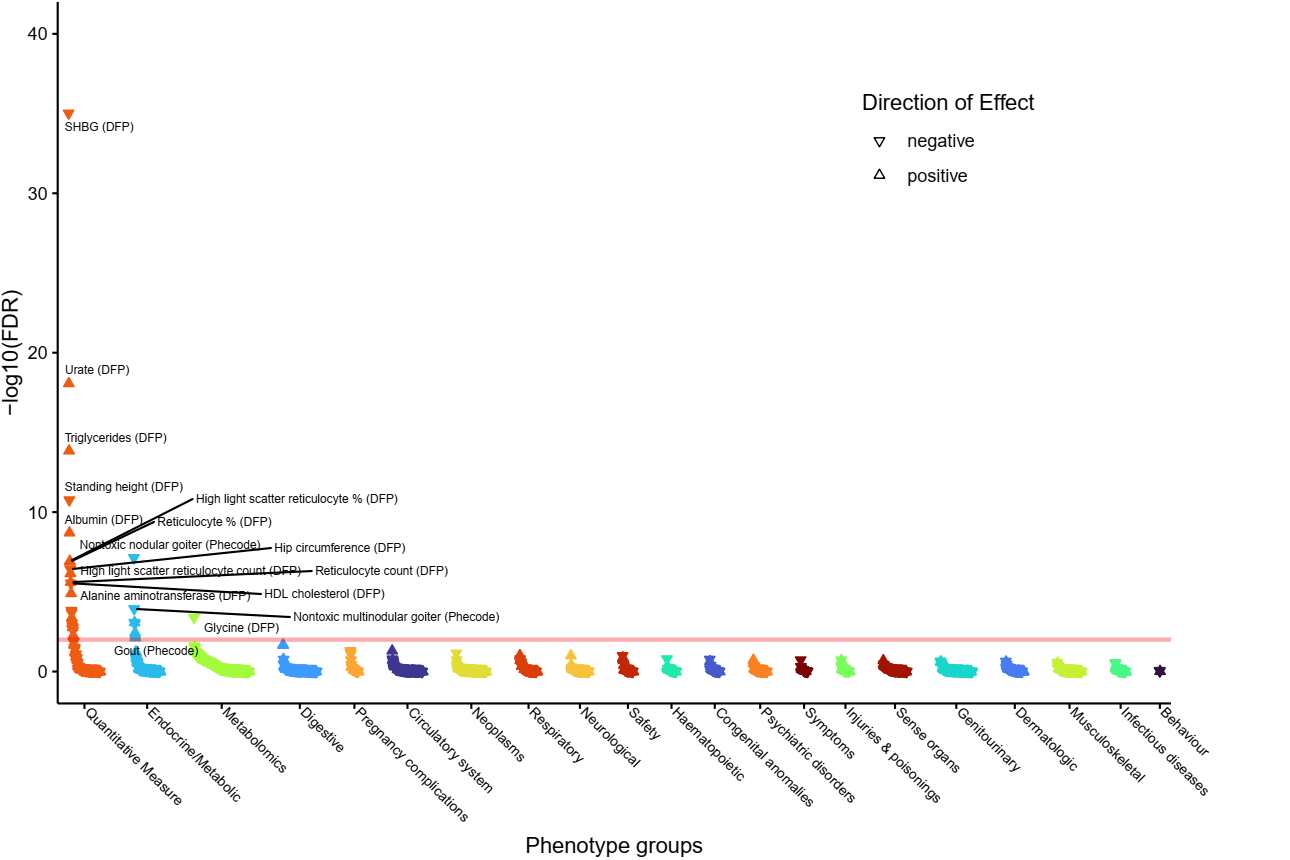
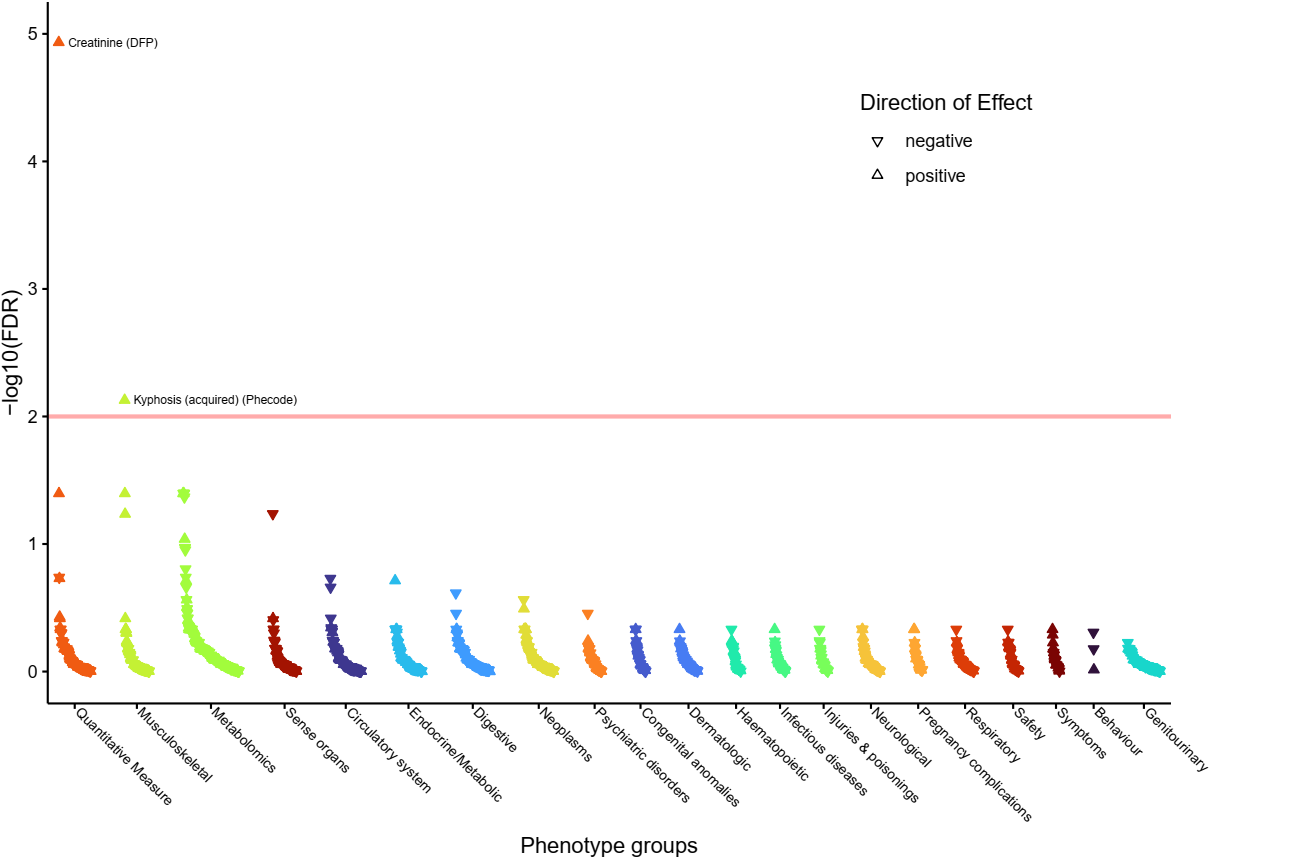
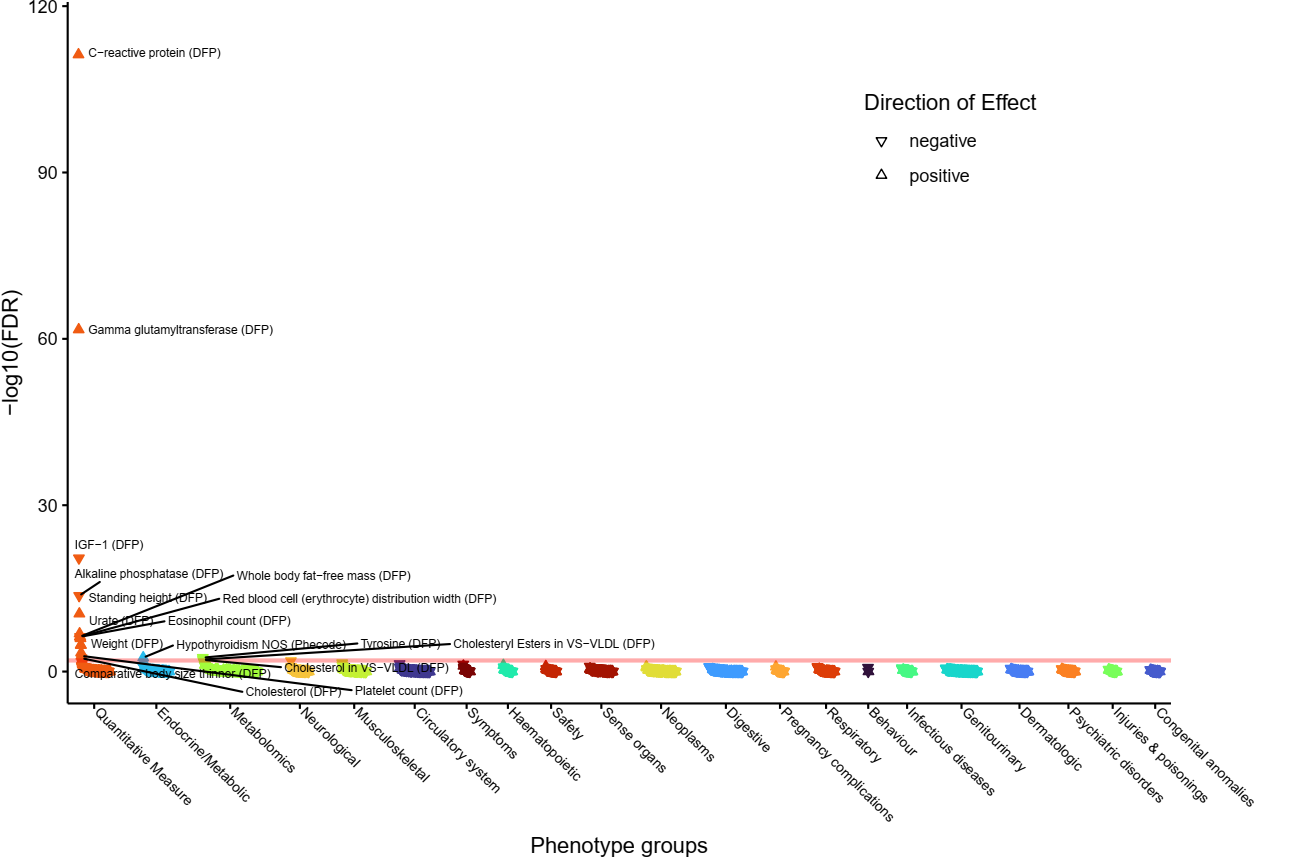
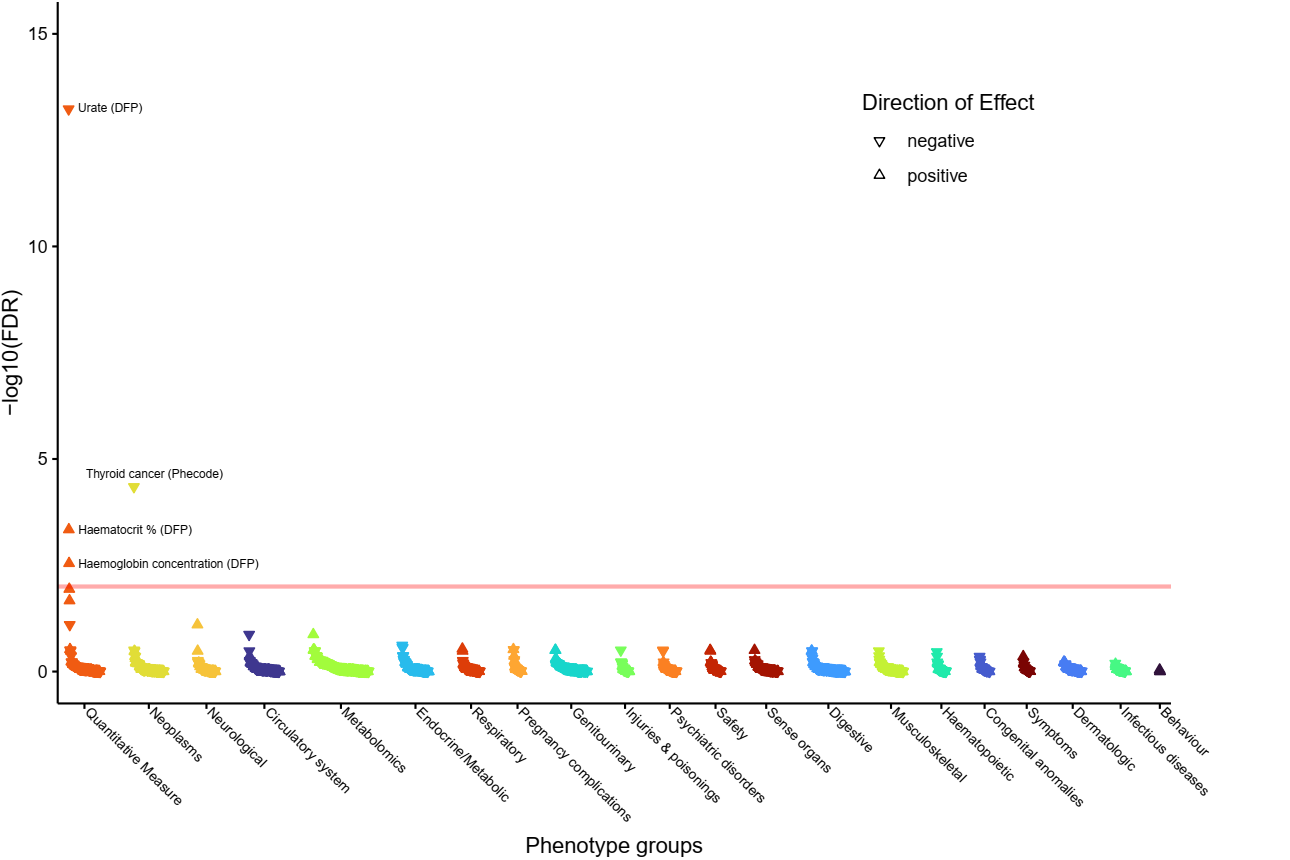
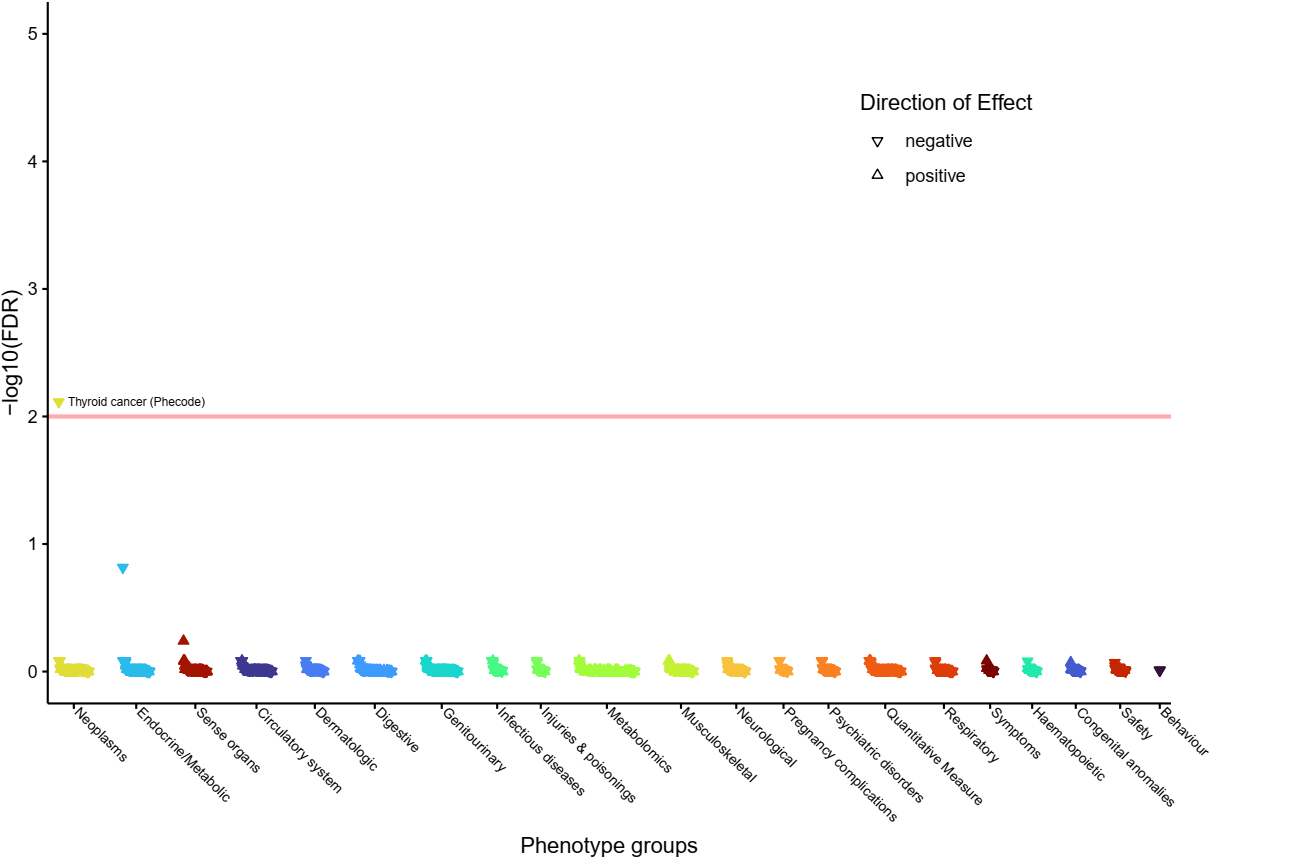
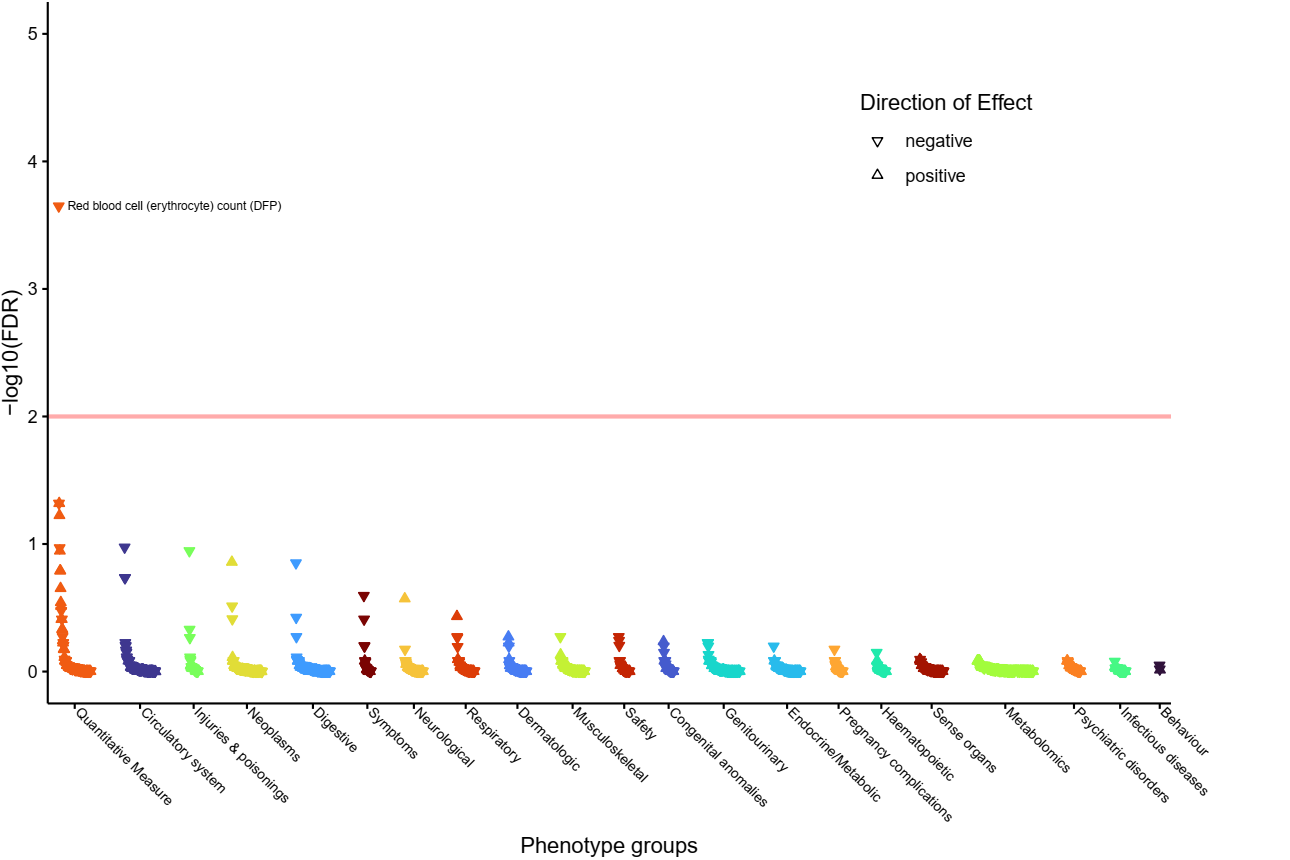
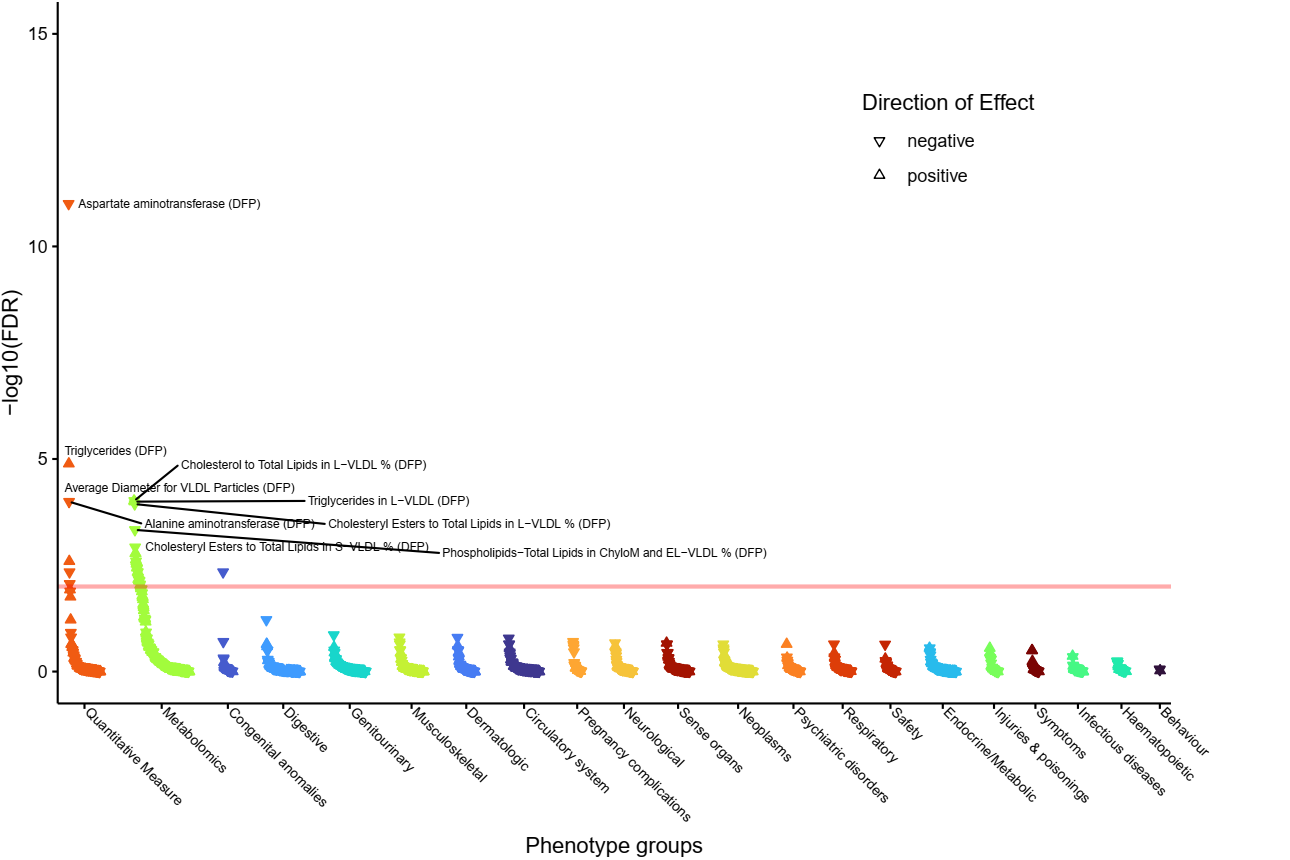
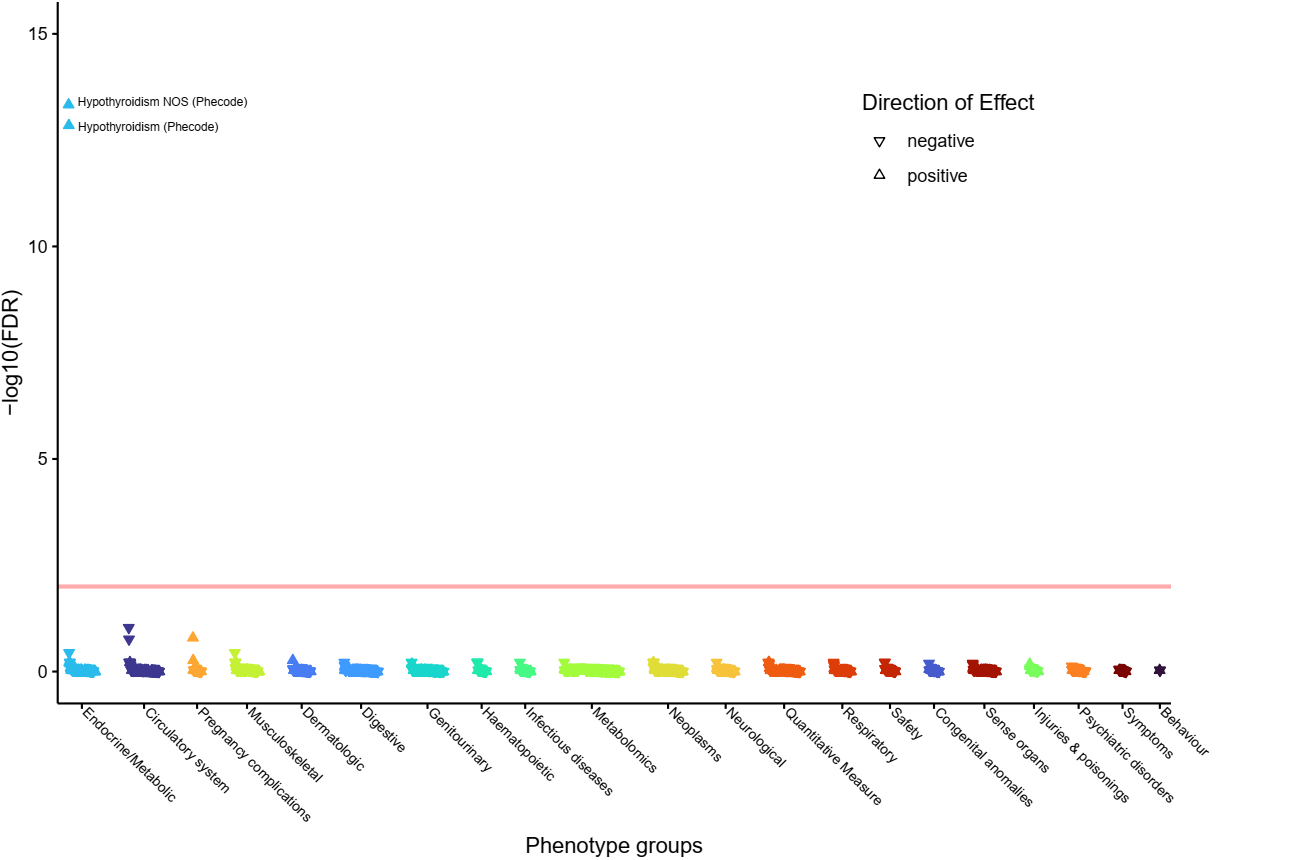
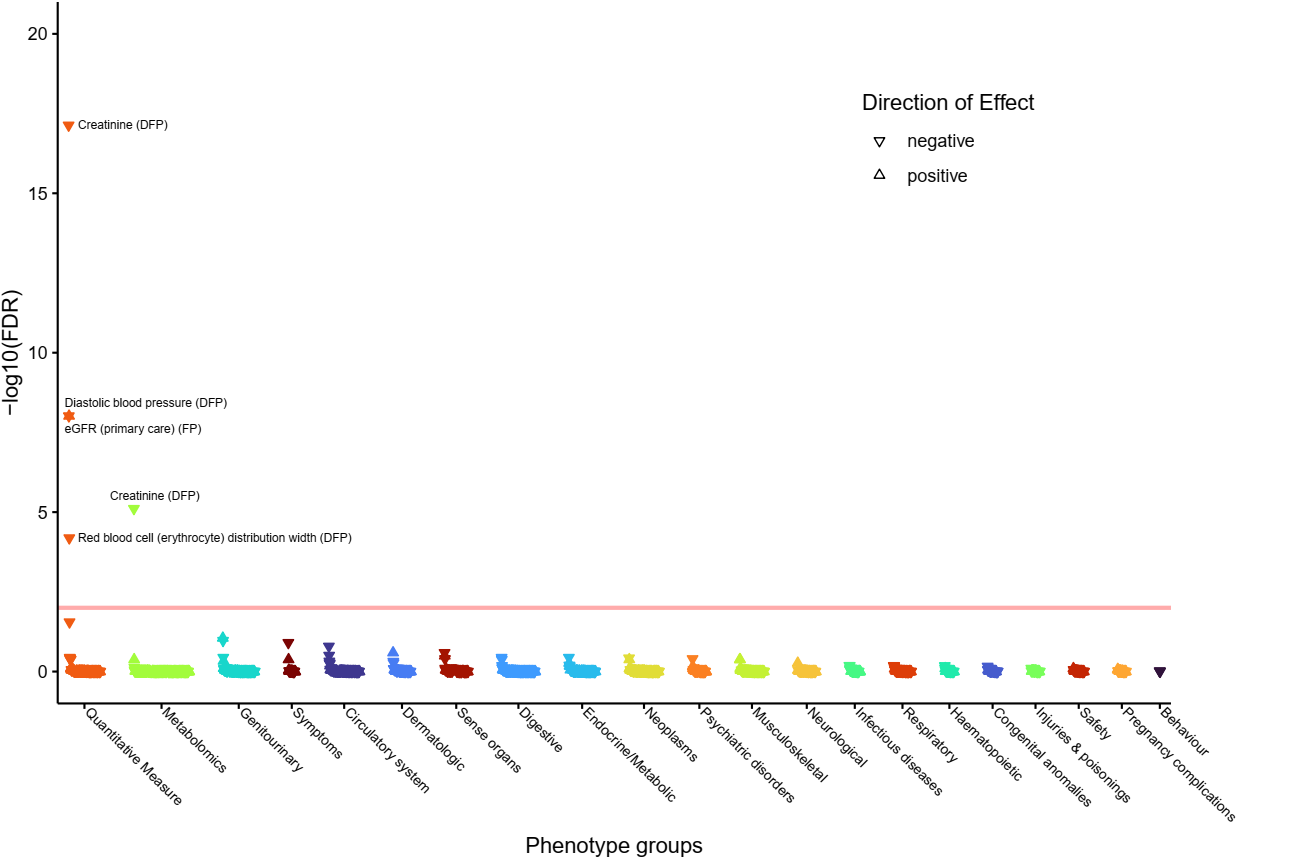
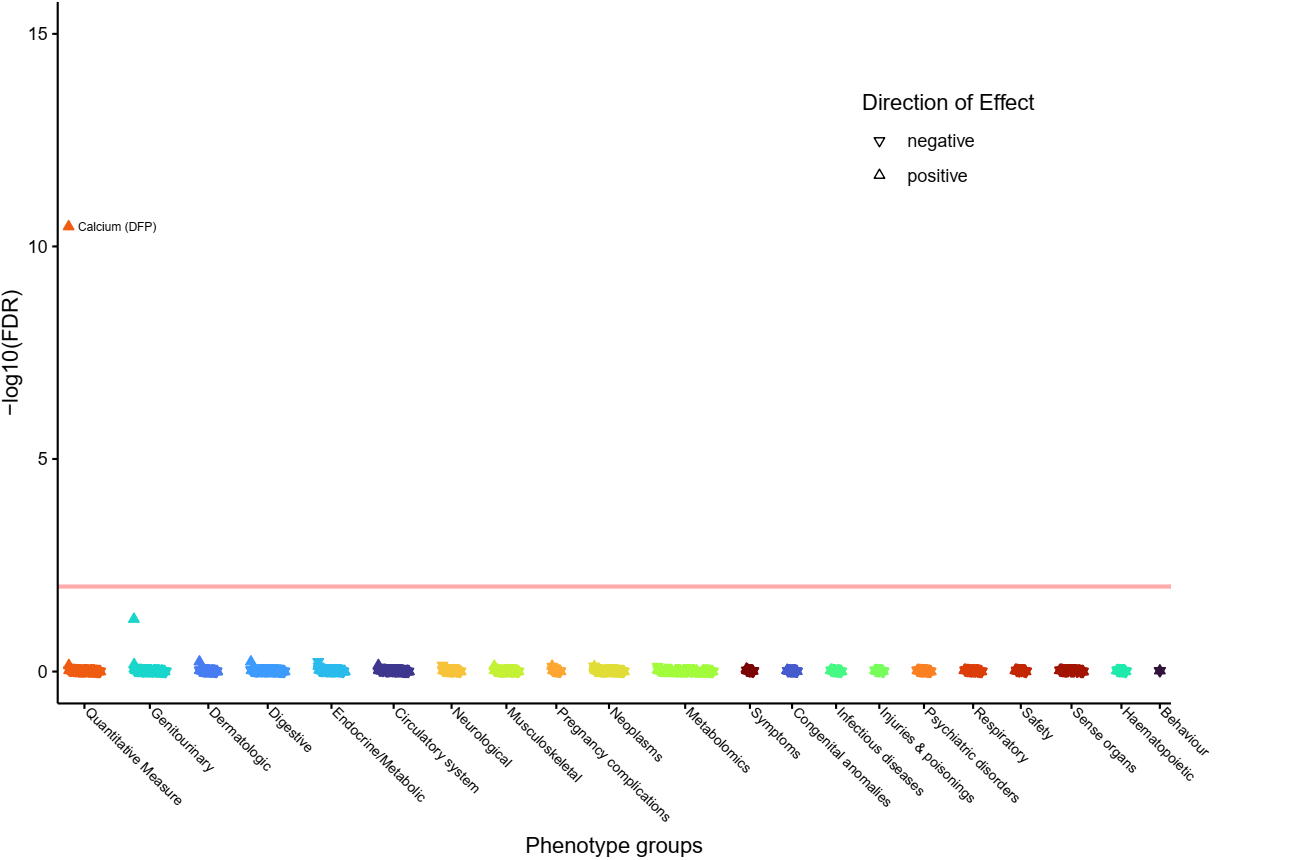
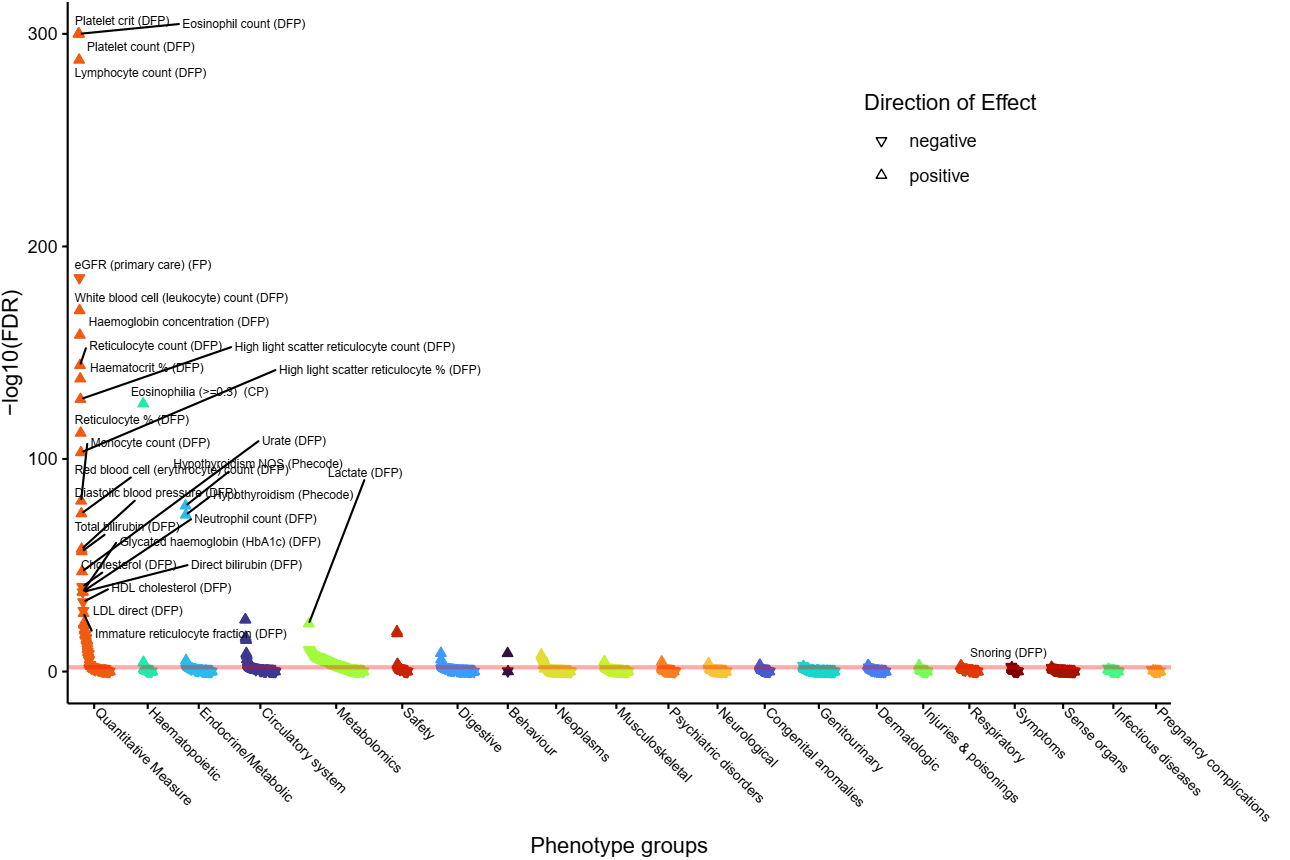
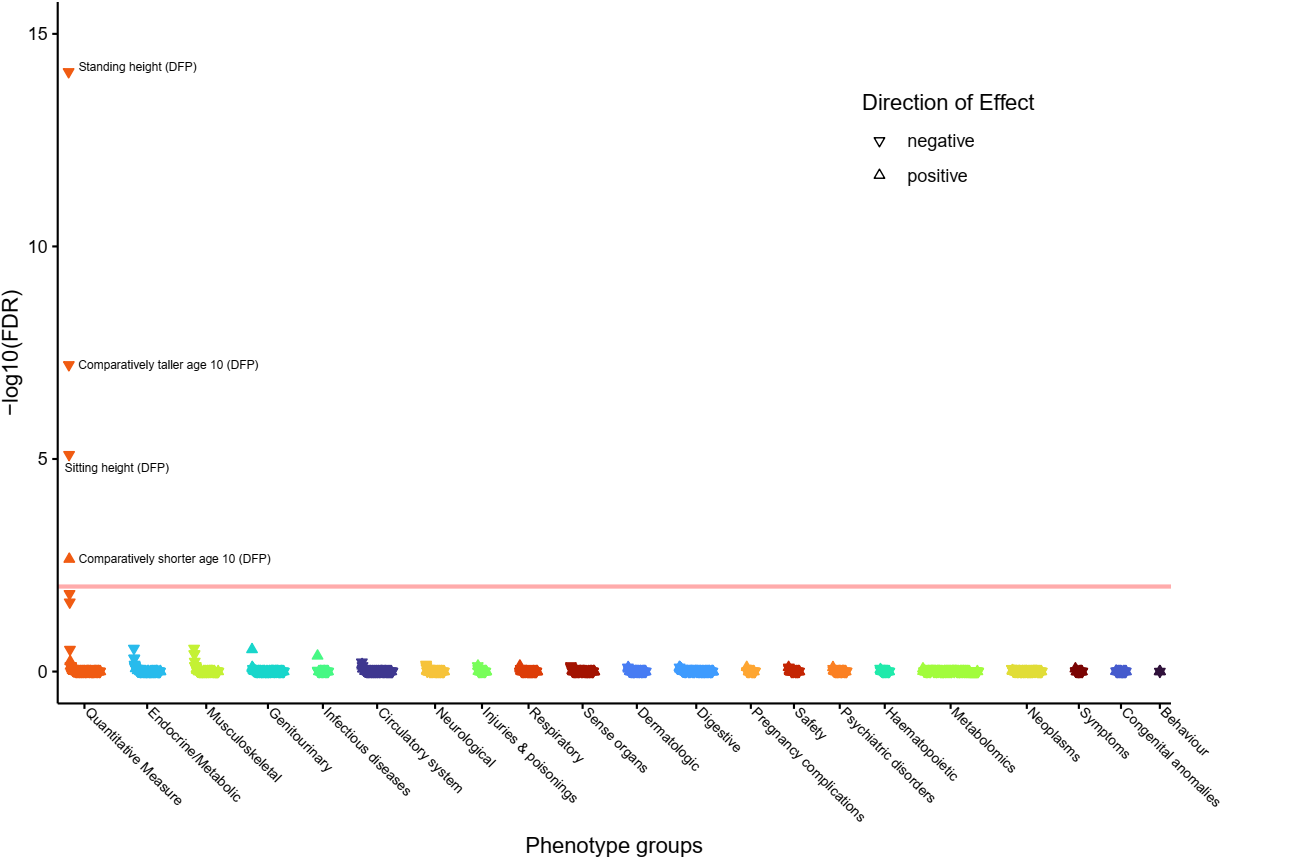
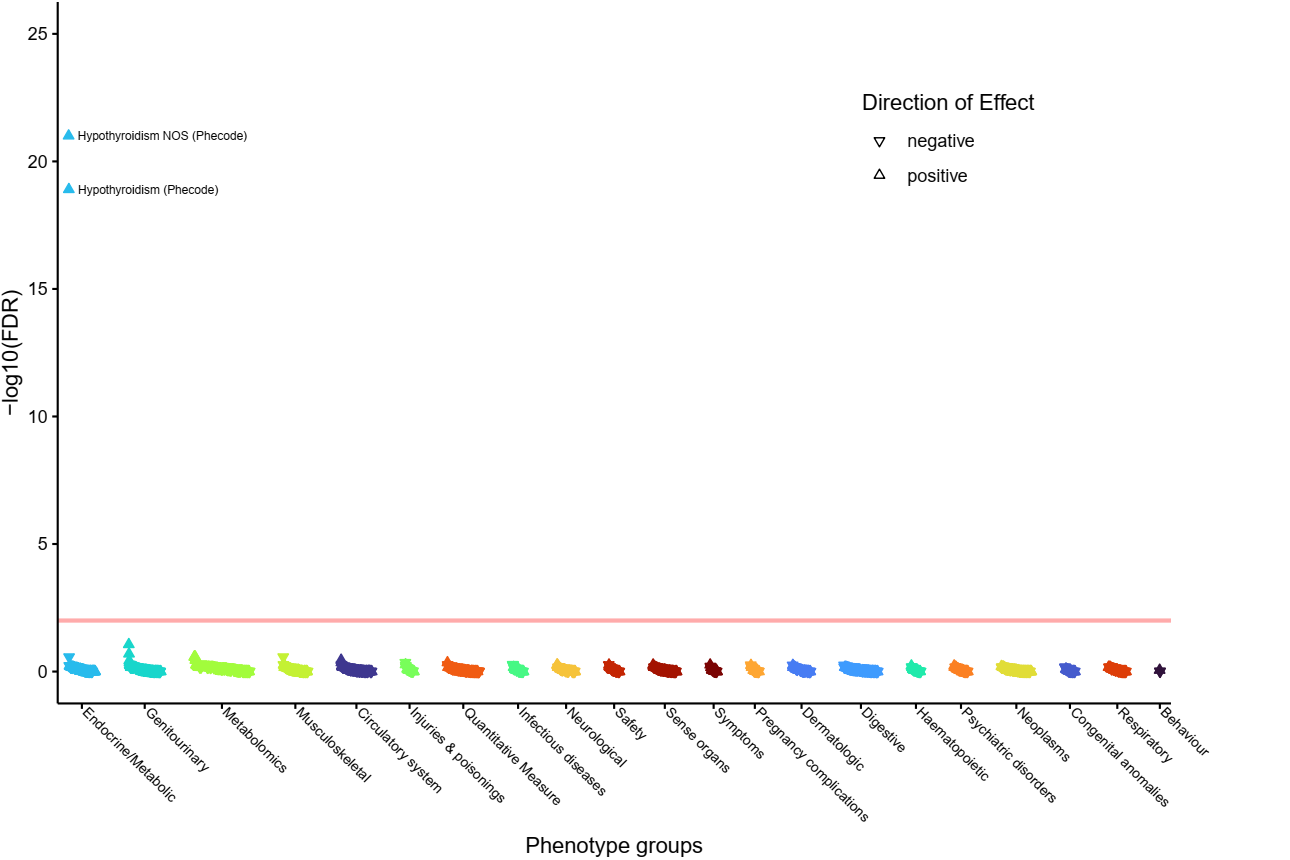
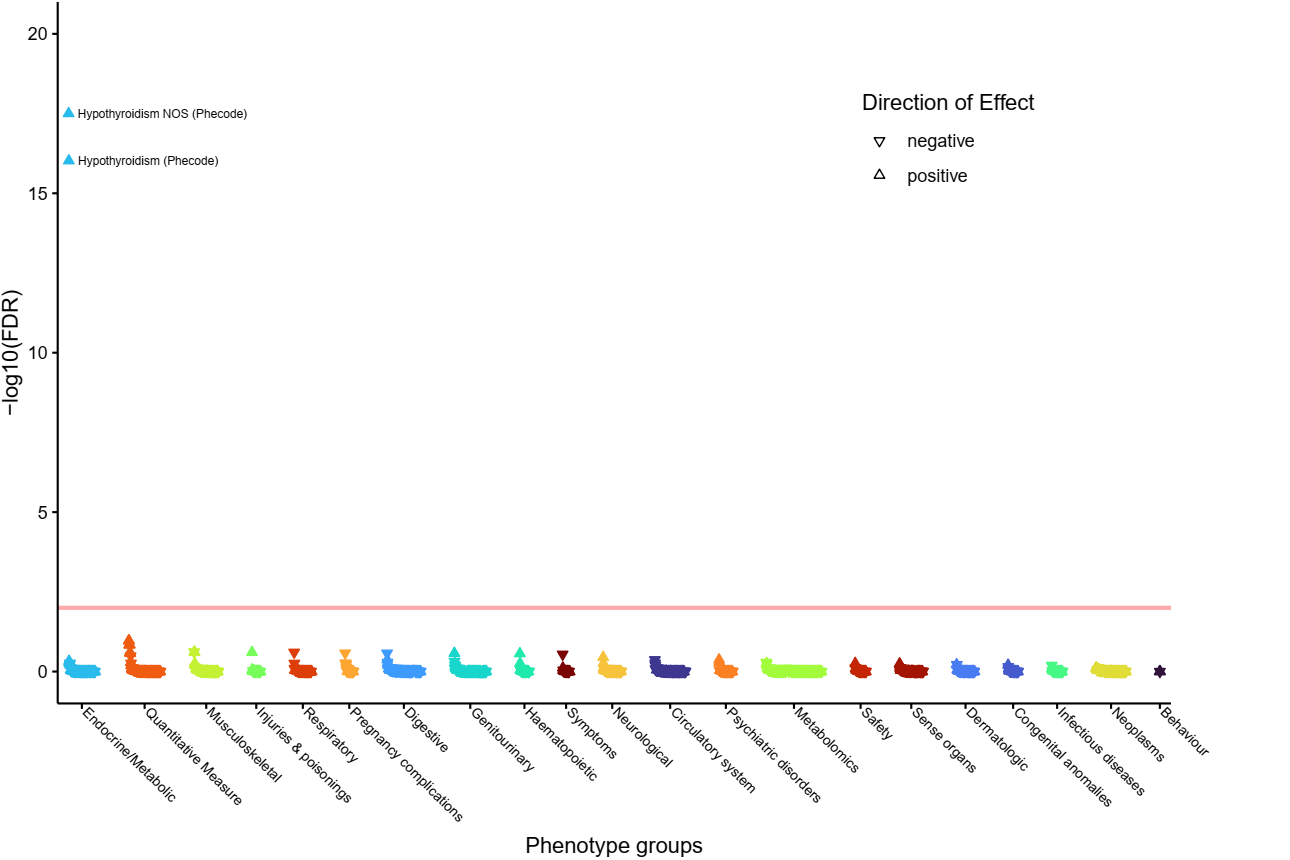
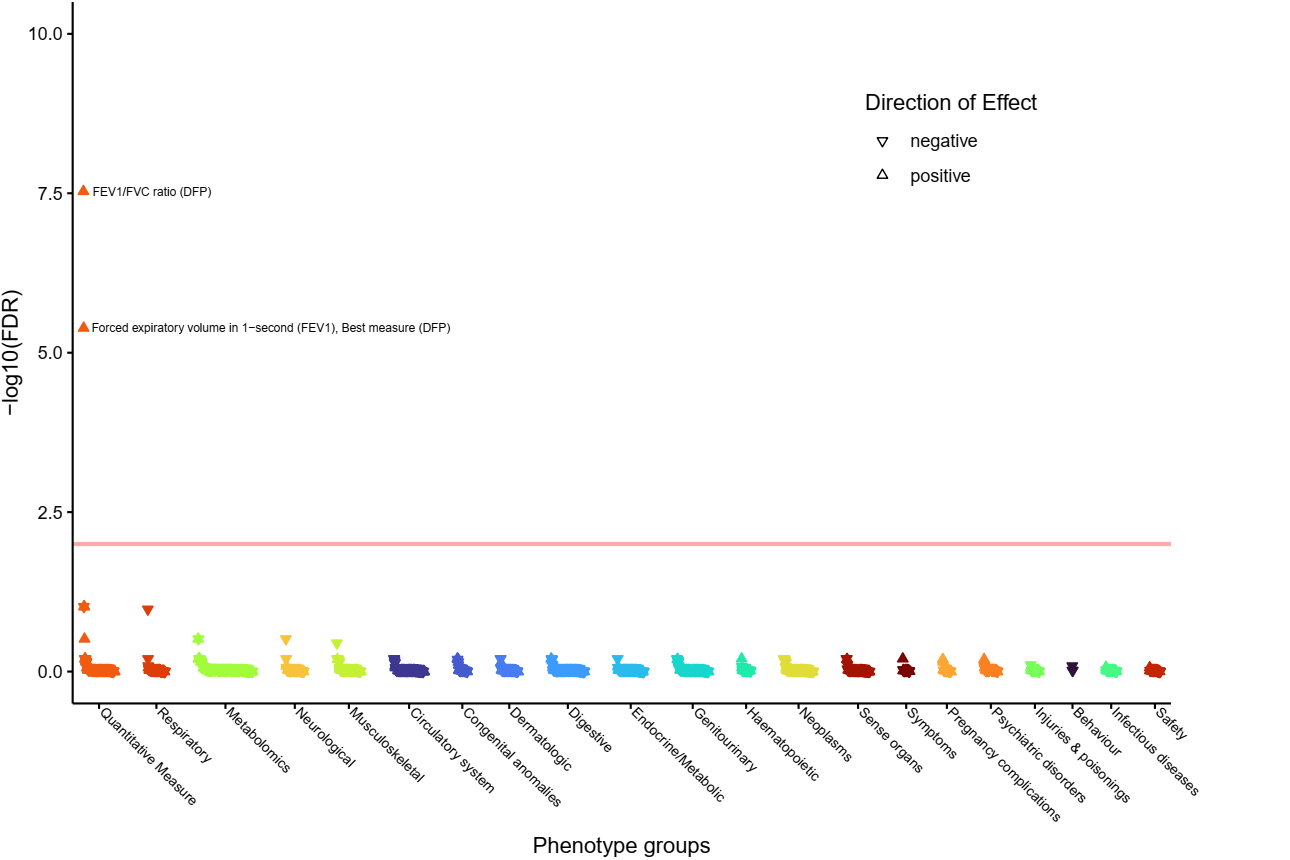
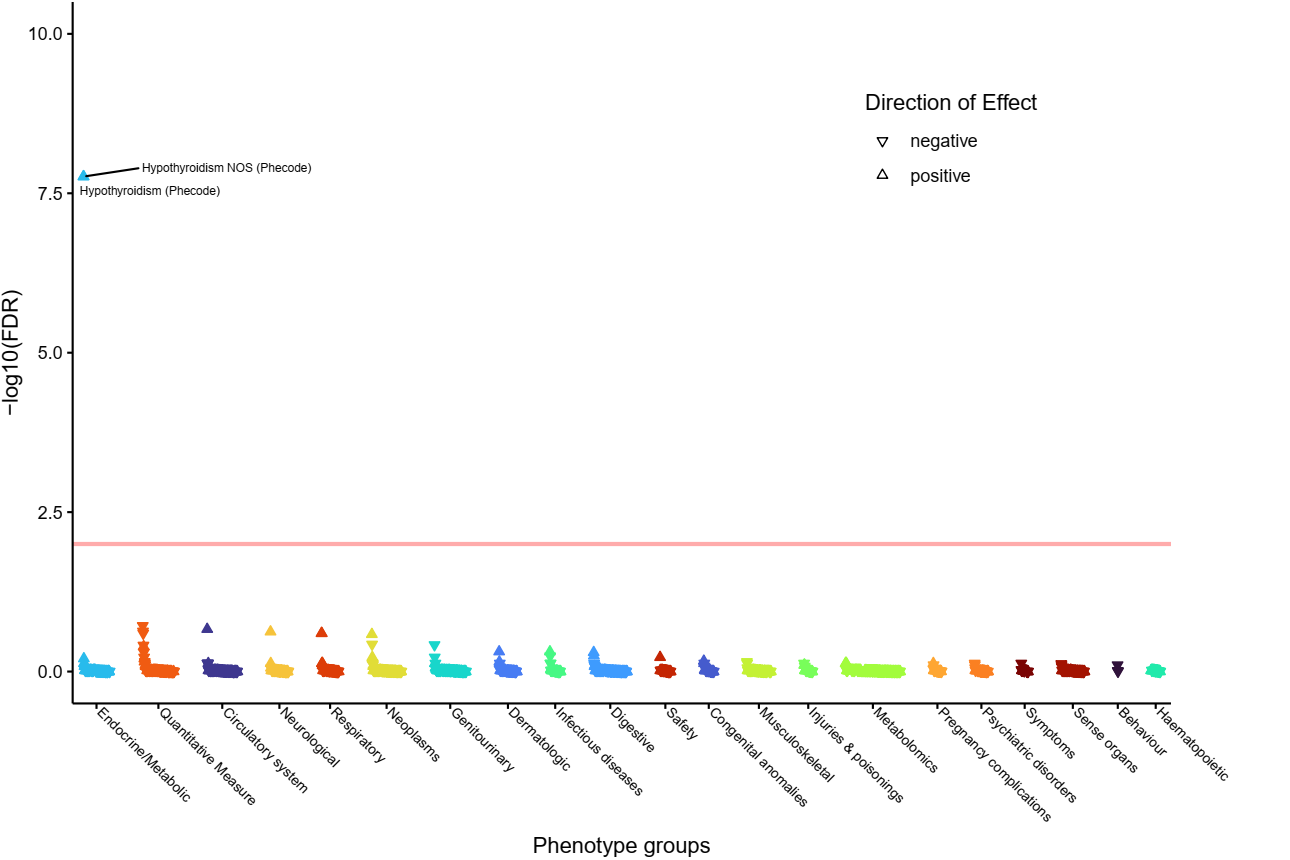
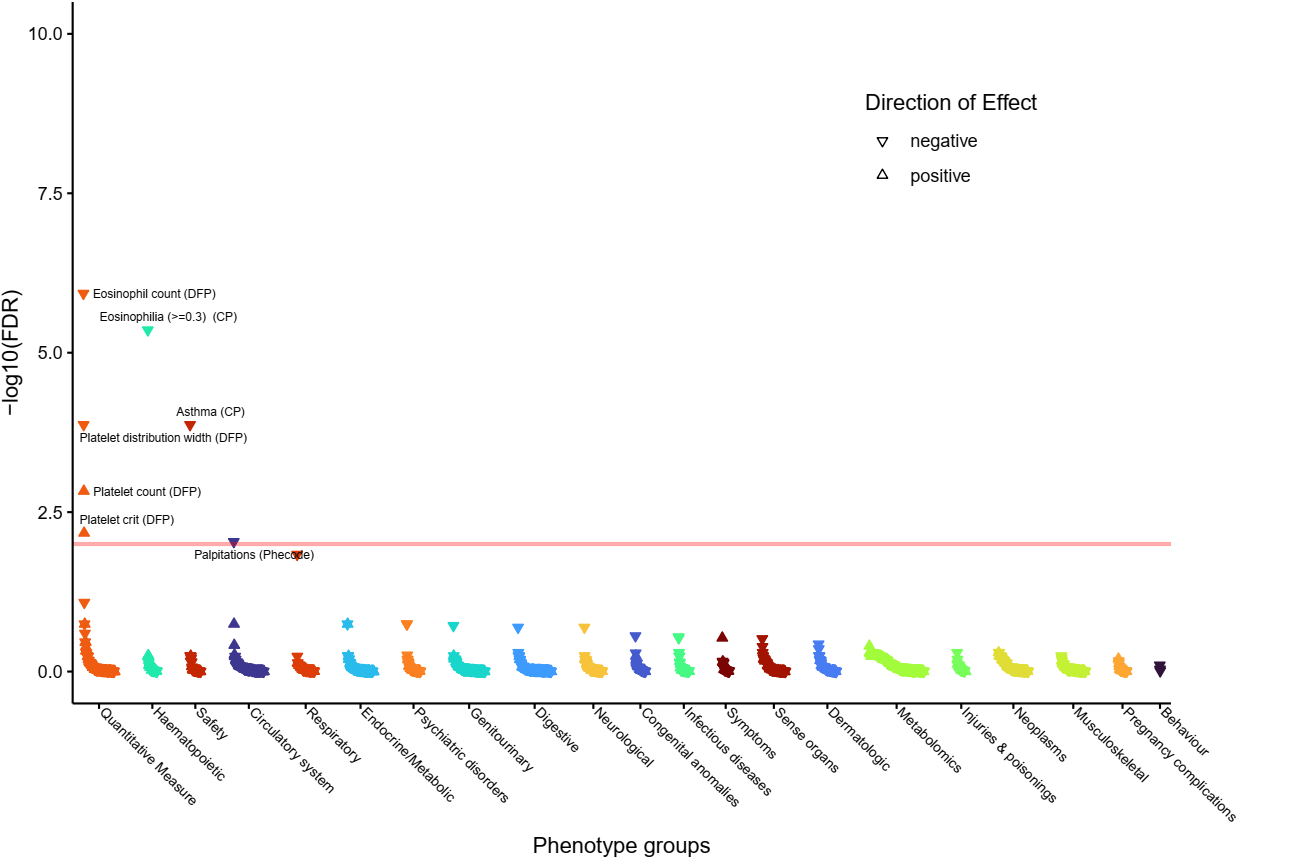
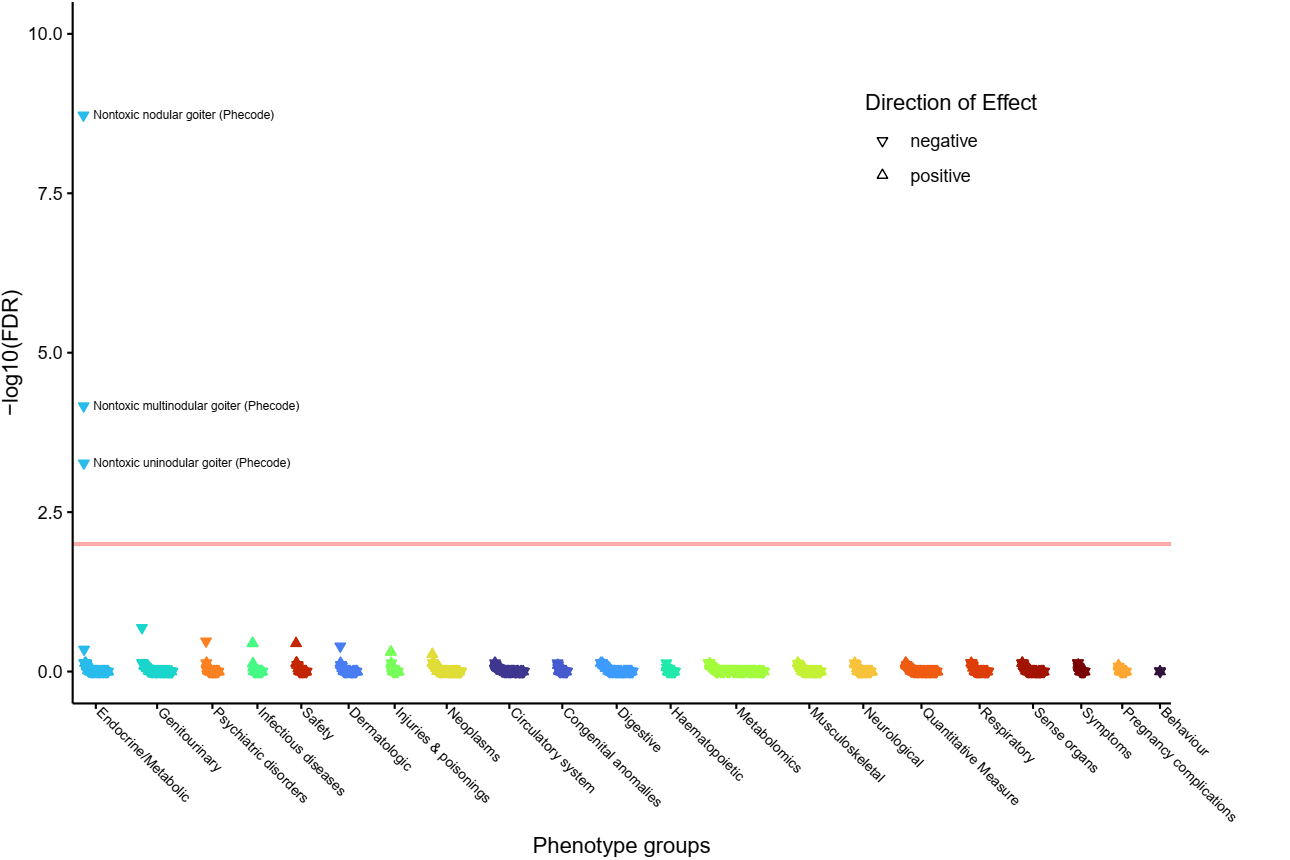
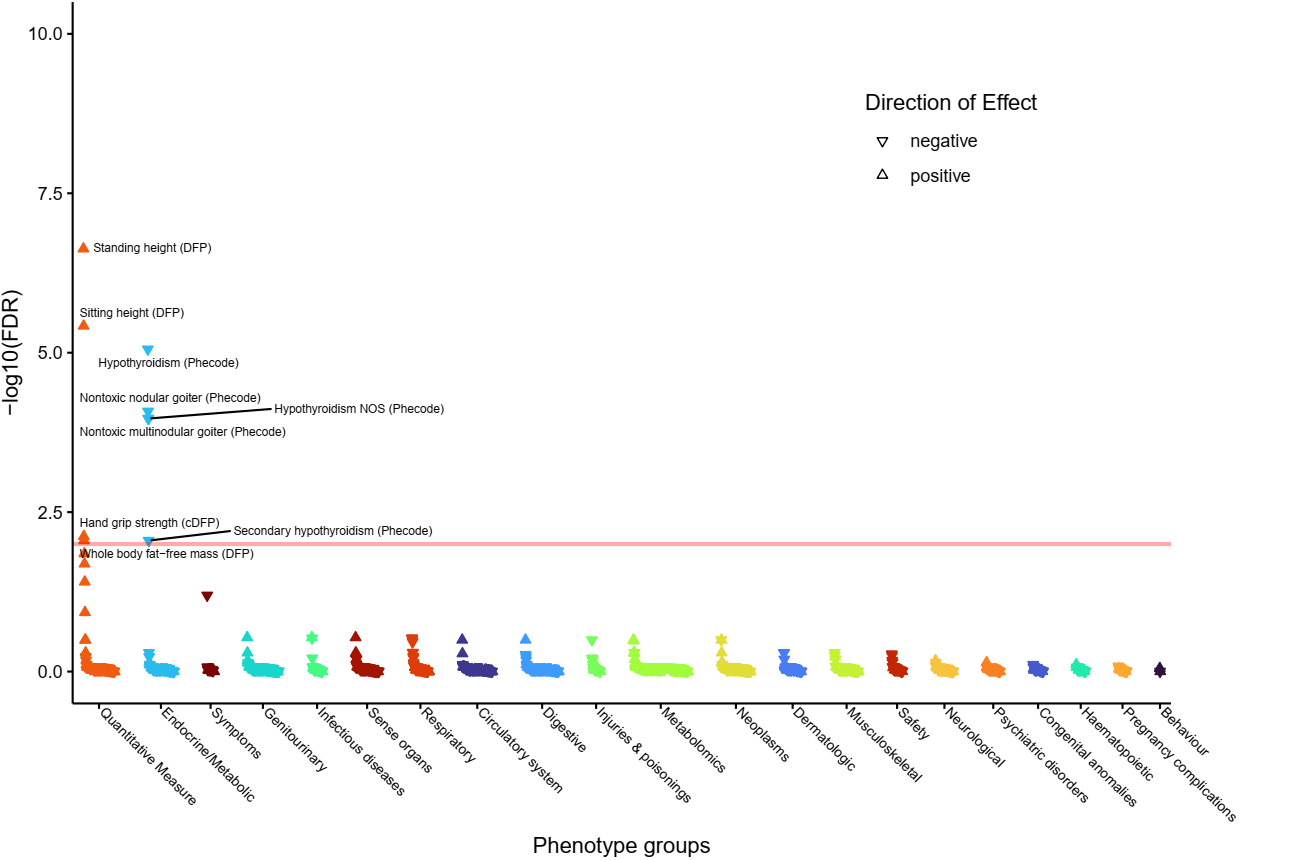
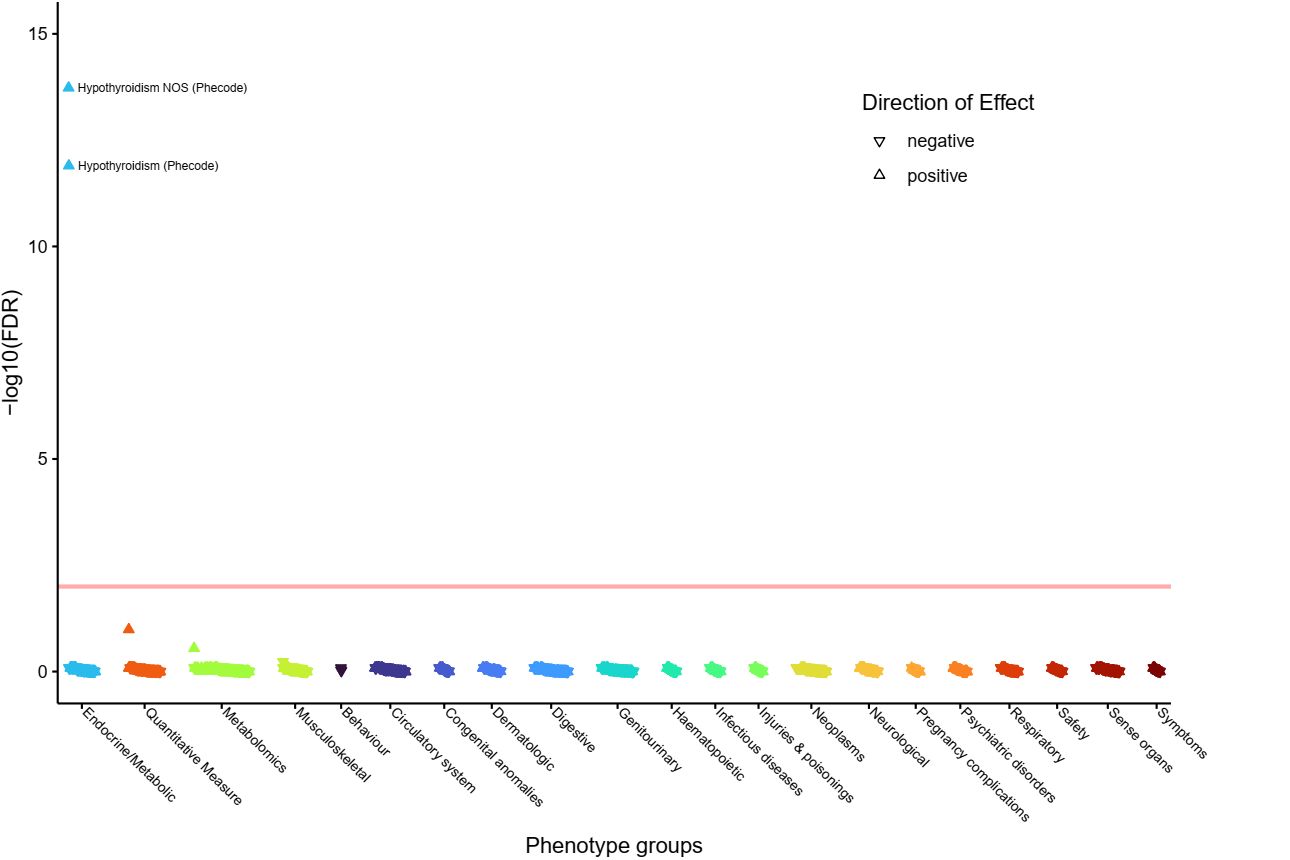
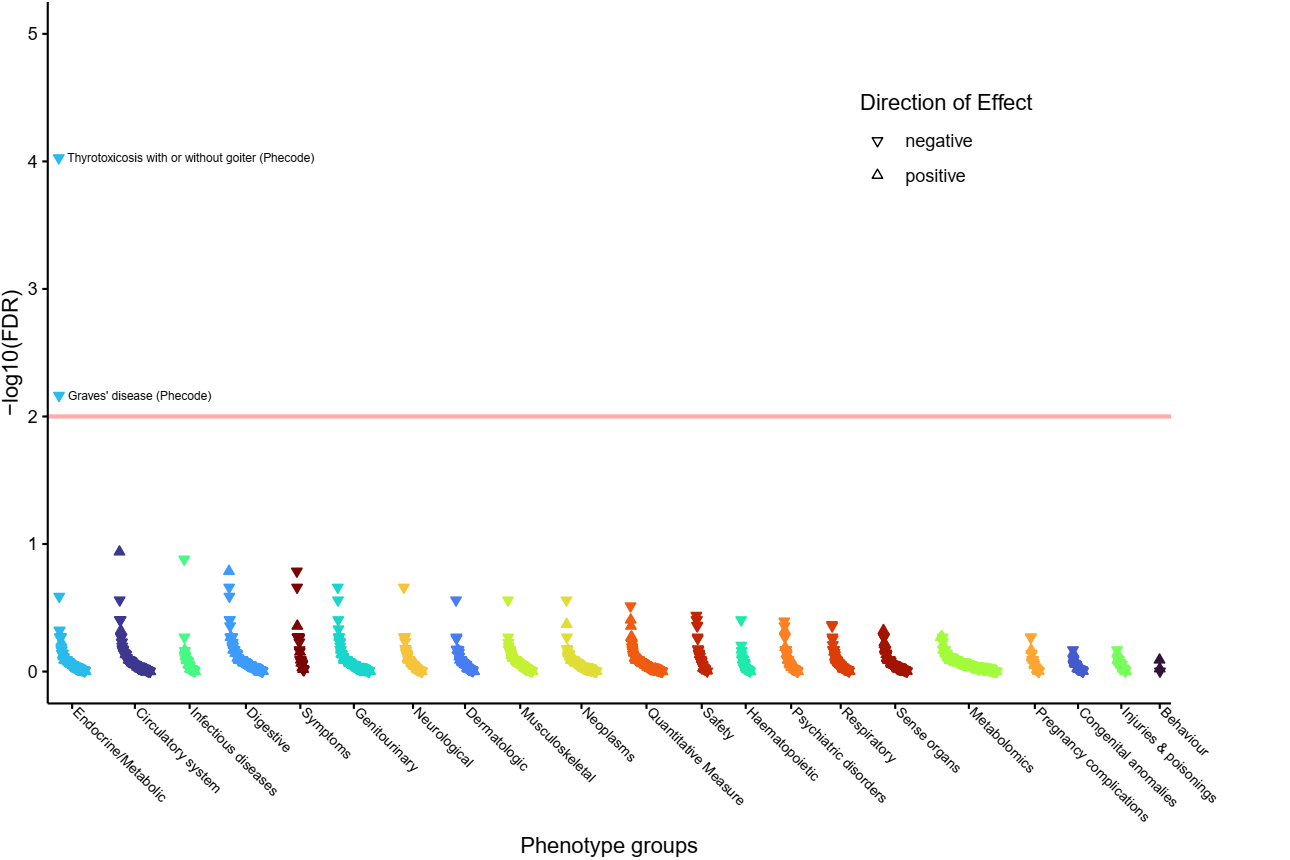
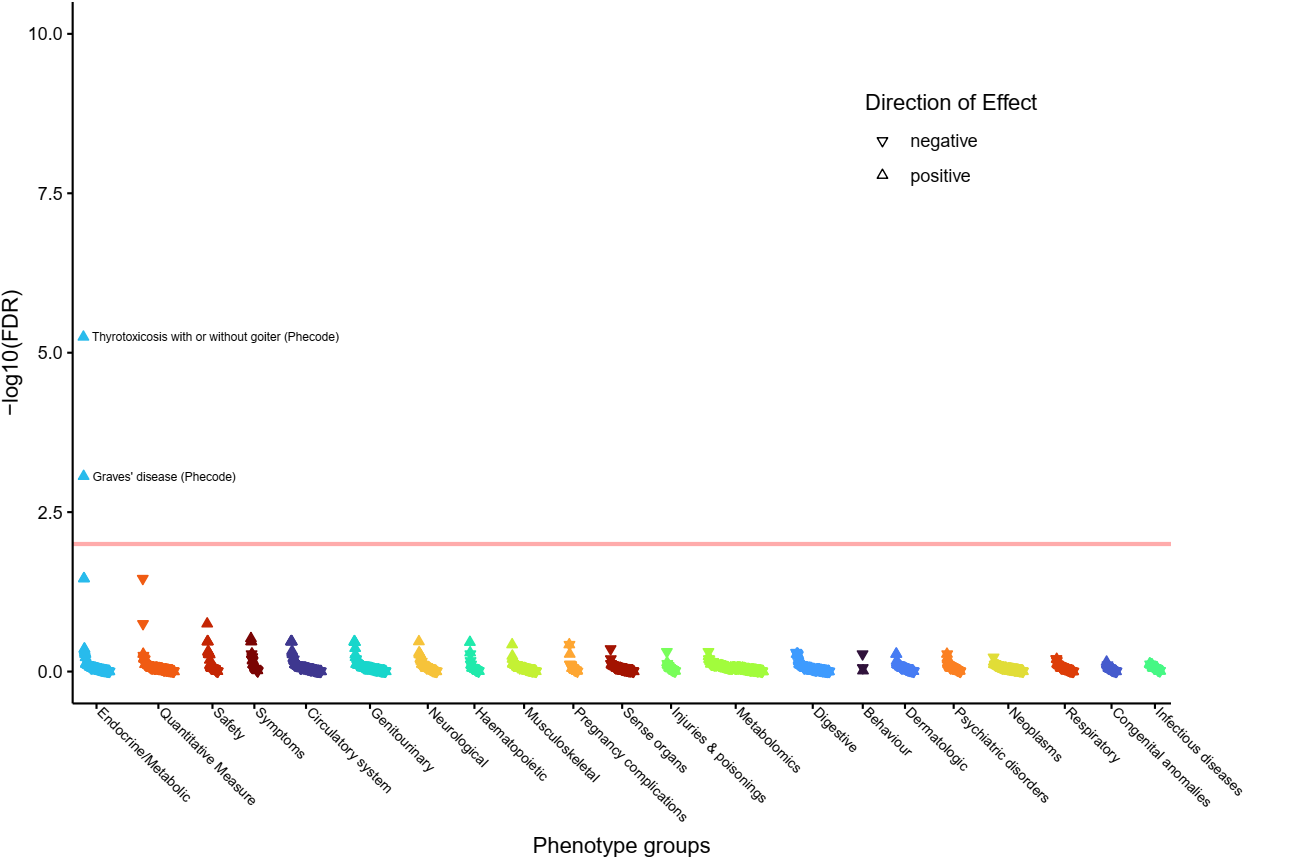
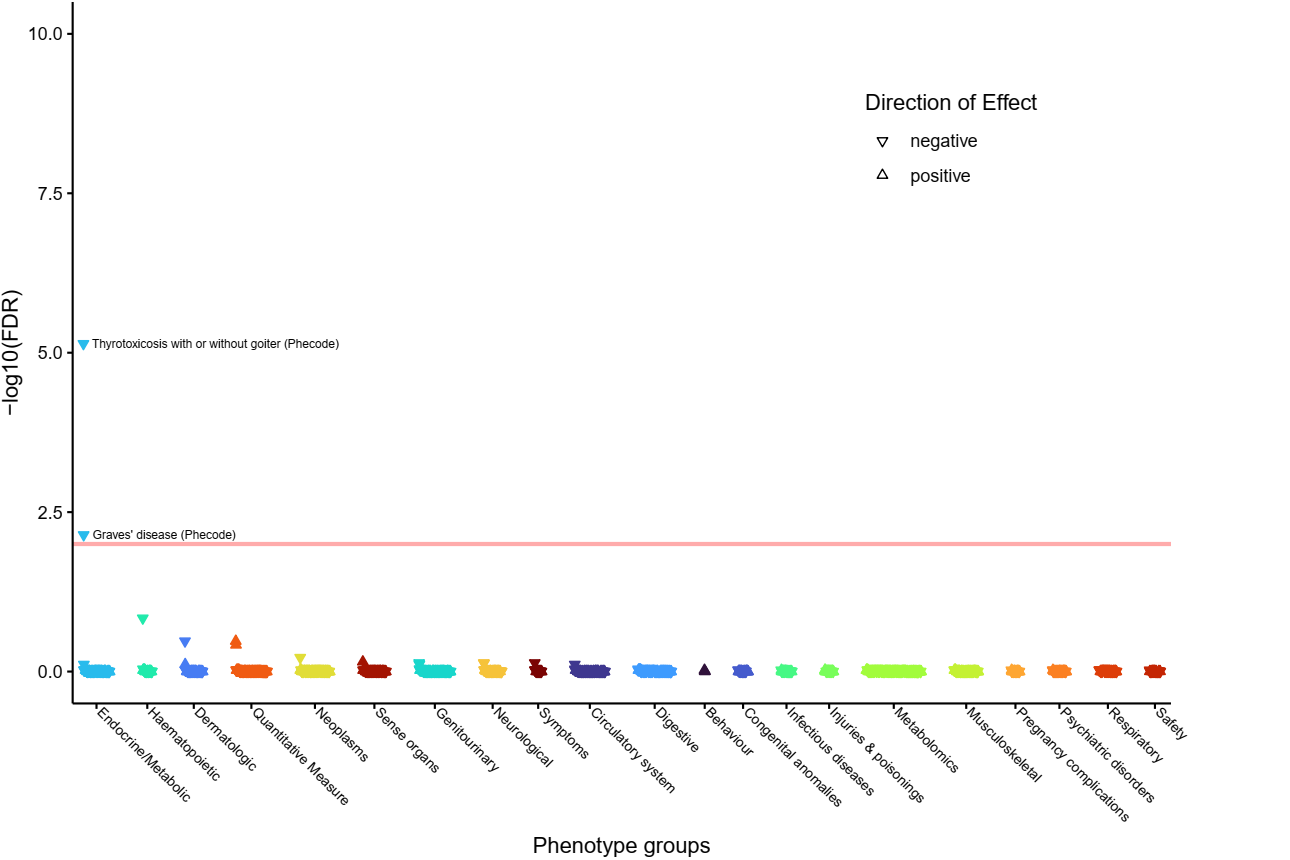
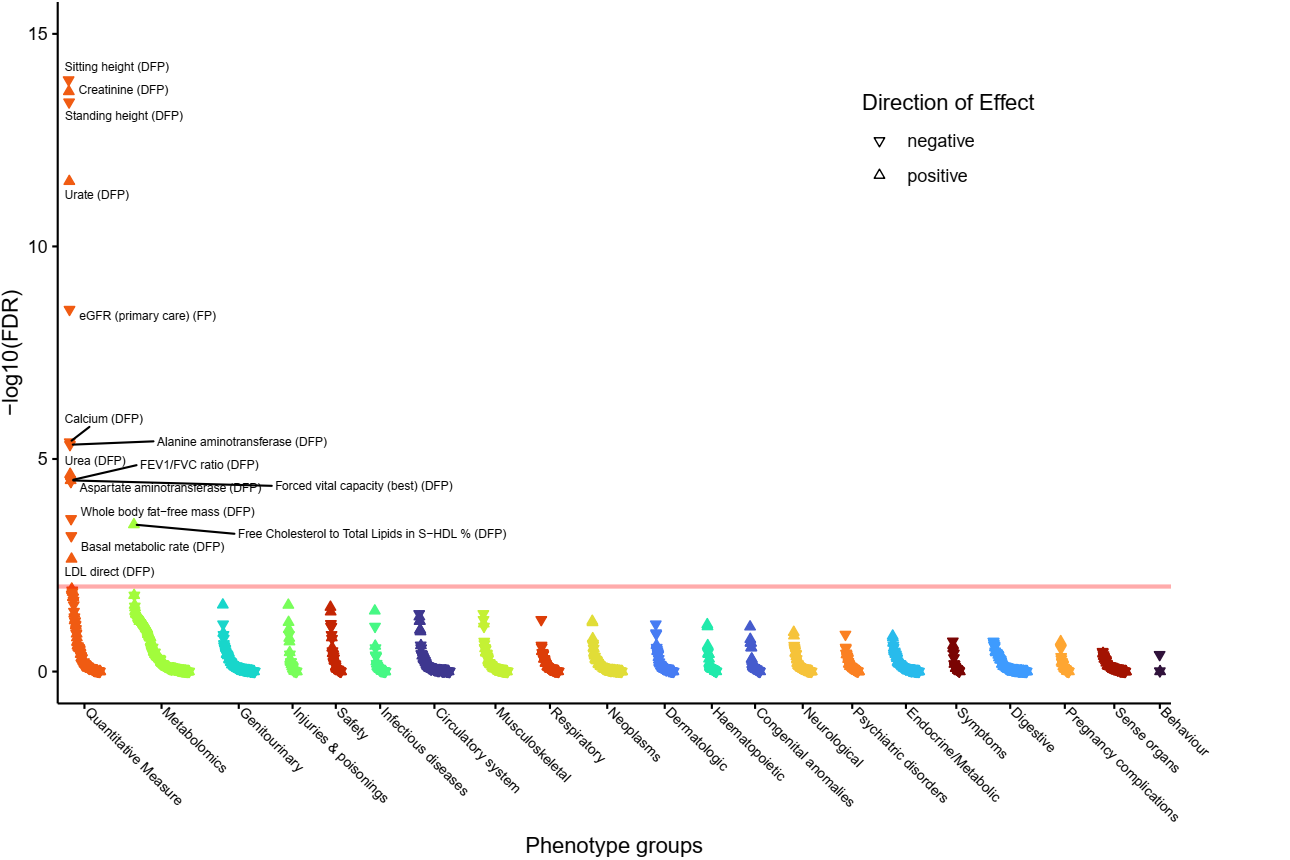
   Description automatically generated
2. **ADORA2B mediated anti-inflammatory cytokines production pathway.**  
   ****
3. **Apoptosis-related network due to altered Notch3 in ovarian cancer pathway.  
   **
4. **Factors and pathways affecting insulin-like growth factor (IGF1)-Akt signaling pathway.  
   **
5. **FGFR3 signaling in chondrocyte proliferation and terminal differentiation pathway.  
   **
6. **Focal Adhesion-PI3K-Akt-mTOR-signaling pathway.  
   **
7. **Supplementary Figure 7: PheWAS for the G Protein Signaling Pathways pathway.  
   **
8. **Heart Development pathway.  
   **
9. **Hemostasis pathway.  
   **
10. **Human T-cell leukemia virus 1 infection - Homo sapiens (human) pathway.  
    **
11. **Intracellular Signalling Through FSH Receptor and Follicle Stimulating Hormone pathway.  
    **
12. **Morphine addiction - Homo sapiens (human) pathway.  
    **
13. **Myometrial relaxation and contraction pathways.  
    **
14. **Pathways in cancer - Homo sapiens (human) pathway.  
    **
15. **Rap1 signaling pathway - Homo sapiens (human) pathway.  
    **
16. **Regulation of lipolysis in adipocytes - Homo sapiens (human) pathway.  
    **
17. **Signal Transduction pathway.**
18. **Signalling by Receptor Tyrosine Kinases pathway.  
    **
19. **Thyroxine (Thyroid Hormone) Production pathway.  
    **
20. **Transcriptional regulation by RUNX2 pathway.  
    **
21. **Vegf hypoxia and angiogenesis pathway.  
    **
22. **VEGF ligand-receptor interactions pathway.  
    **
23. **VEGFA-VEGFR2 Signaling Pathway.  
    **

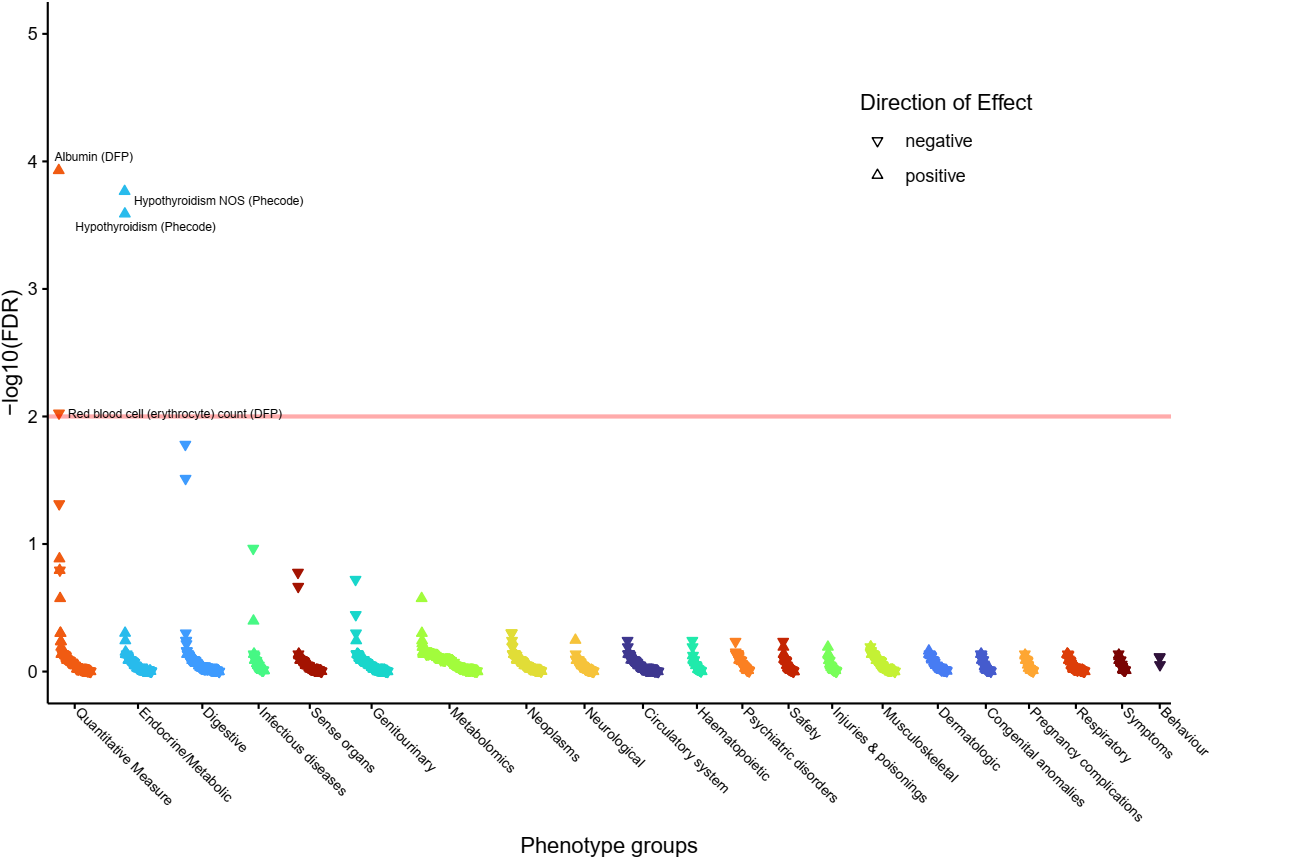
Supplementary Figure 3: Thyroid-stimulating hormone polygenic score PheWAS

**Timeline

Description automatically generated**

Supplementary Figure 4: Single variant PheWAS

1. **rs115315671 (implicating *ADCY6*).  
   **
2. **rs2494734 (implicating *AKT1* and *ZBTB42*).  
   **
3. **rs17450274 (implicating *ANXA5*).  
   **
4. **rs11605461 (implicating *CADM1*).  
   **
5. **rs149363012 (implicating *CCDC77*, *WNK1* and *B4GALNT3*).  
   **
6. **rs3124747 (implicating *DBH*, *C9orf96* and *STKLD1*).  
   **
7. **rs139411099 (implicating *DIO2*).  
   **
8. **rs2282192 (implicating *FOXE1*, *ANP32B* and *C9orf156*).  
   **
9. **rs10814915 (implicating *GLIS3*).  
   **
10. **rs1861628 (implicating *IGFBP5*).  
    **
11. **rs1861630 (implicating *IGFBP5*).  
    **
12. **Supplementary Figure 36: PheWAS for rs6711608 (implicating *IGFBP5*).  
    **
13. **Supplementary Figure 37: PheWAS for rs57866041 (implicating *IGFBP5*).  
    **
14. **rs2229738 (implicating *IGHMBP2*, *CPT1A* and *SCD*).  
    **
15. **rs4804416 (implicating *INSR* and *ACTL9*).  
    **
16. **rs7306508 (implicating *KRT18*).  
    **
17. **rs2393717 (implicating *MLEC* and *SPPL3*).  
    **
18. **rs36213229 (implicating *NRG1*).  
    **
19. **PheWAS for rs68030583 (implicating *NRG1*).  
    **
20. **rs11554674 (implicating *PHC2*, *ZNF362* and *A3GALT2*).  
    **
21. **rs1801690 (implicating *PRKCA* and *APOH*).  
    **
22. **rs9497965 (implicating *SASH1*).  
    **
23. **rs10926981 (implicating *SDCCAG8*).  
    **
24. **Supplementary Figure 48: PheWAS for rs1743963 (implicating *SGK1*).  
    **
25. **rs3184504 (implicating *SH2B3*).  
    **
26. **rs751171 (implicating *SMOC2*).  
    **
27. **rs9507279 (implicating *SPATA13*).  
    **
28. **rs17364832 (implicating *SPATA13*).  
    **
29. **rs10113355 (implicating *SULF1* and *LOC100505739*).  
    **
30. **rs2069556 (implicating *TG* and *CCN4*).  
    **
31. **rs2069568 (implicating *TG* and *CCN4*).  
    **
32. **rs114322847 (implicating *TG* and *CCN4*).  
    **
33. **rs519709 (implicating *TGFB2* and *MIR548F3*).  
    **
34. **rs11694732 (implicating *TPO*).  
    **
35. **Supplementary Figure 59: PheWAS for rs3742721 (implicating *TSHR* and *CEP128*).  
    **
36. **rs6574611 (implicating *TSHR* and *CEP128*).  
    **
37. **rs2110696 (implicating *TSHR*).  
    **
38. **rs1802288 (implicating *TSPAN6*).  
    **
39. **rs3755972 (implicating *VEGFC*).**

****

# Supplementary Tables

Supplementary Table 1: (Excel sheet) Signal selection results.  
Each row represents the sentinel variant for an independent signal of association, the sentinel is the variant with the highest posterior probability (PIP column) returned from Polyfun Susie or Wakefield (method column). Allele1 is the coded/effect allele. The novel and novel locus columns indicated a previously unreported signal or locus respectively and the SNP reported, R2 and PMID columns give the previously reported SNP at the locus, it’s r2 with the signal reported here and the publication of the previously reported signal. The n\_genes, genes and evidence columns give the genes implicated by the sentinel and the variant-to-gene evidence. Columns AA-AF give the VEP annotation for the sentinel and columns AG-AI give the most highly associated SNP in the locus. The study-level statistics from the 3 studies in the meta-analysis are in columns AJ-AR. The “consistent” column indicates signals that showed consistent direction of effect and P<0.01 in at least two of the three contributing datasets in addition to reaching P<5×10-8 in the meta-analysis.

Supplementary Table 2: (Excel sheet) Prioritised genes  
List of prioritised genes according to the number of lines of variant-to-gene evidence implicates the gene (n\_evidence column). The “Novel” column indicates whether the gene was previously reported. The n\_novel\_signals lists the number of novel signals implicating the gene.

Supplementary Table 3: (Excel sheet) Previously reported genes  
Genes typically implicated by a single criterion and the reference for the previous publication reporting the gene.

Supplementary Table 4: (Excel sheet) Epidemiological associations  
Epidemiological associations between the 257 top PIP variants available in UK Biobank and thyroxine (T4) levels or clinical thyroid diseases: hypothyroidism, hyperthyroidism, thyroid cancer, and other non-cancer thyroid diseases. Also shown are epidemiological associations with hypothyroidism due to medication defined using phecode 244.1 - these associations were only assessed in genetic variants implicating genes that met two or more variant-to-gene criteria.

Supplementary Table 5: (Excel sheet) Single-variant PheWAS associations  
Phenome-wide association study (PheWAS) results for select TSH sentinel variants for associations at FDR <1% (European ancestry). N\_ID = Number of participants in the analysis, FDR = false discovery rate, P = p value, OR = odds ratio, L95 = lower boundary of the 95% confidence interval, L95 = upper boundary of the 95% confidence interval, MAF = minor allele frequency, MAC = minor allele count, MAC\_cases = minor allele count in cases, MAC\_controls = minor allelecount in controls. Z\_T\_STAT = the Z or T STAT output from Plink2, SE = standard error. The P value is the P value from Plink2 for association for a linear model or logistic regression for the phenotype after adjusting for age, sex and first 10 principal components.

Supplementary Table 6: (Excel sheet) Variants selected at each gene for look up of epidemiological associations  
For each gene the variant with the most significant association across either 6 or 7 binary traits is shown.

**Supplementary Table 7: (Excel sheet) Druggability**

Drug Gene Interaction Database (DGIDB) results. Columns:

* Drug; drug/compound name,
* ChEMBL\_ID; drug/compound identification number from ChEMBL,
* Gene; mapped gene(s),
* Gene\_source; line(s) of evidence for each mapped gene, and signal implicated (including the associated lung function trait),
* Indication(Phase); drug indication phase. Phase 1: Testing of drug on healthy volunteers for dose-ranging; Phase 2: Testing of drug on patients to assess efficacy and safety; Phase 3: Testing of drug on patients to assess efficacy, effectiveness and safety; and Phase 4: Approval of drug and post-marketing surveillance,
* Thyroid\_related: whether the drug is used to for the treatment of thyroid conditions.
* Cancer\_related: whether the drug is used to for the treatment of some form of cancer.

**Supplementary Table 8: (Excel sheet) ConsensuspathDB results**ConsensuspathDB pathways enriched (FDR <5%) for genes from our 112 with at least 2 lines of evidence for being causal. The P value and FDR for enrichment of genes in pathways are those returned by ConsensuspathDB{Herwig, 2016 #78}.

**Supplementary Table 9: (Excel sheet) Pathway-based PheWAS results**Results of 29 pathway-specific PheWAS for FDR <1%: FDR = false discovery rate, P = p value, OR = odds ratio, L95 = lower boundary of the 95% confidence interval, L95 = upper boundary of the 95% confidence interval.

**Supplementary Table 10: (Excel sheet) TSH polygenic risk score PheWAS results**Results of TSH PGS PheWAS for FDR <1%: FDR = false discovery rate, P = p value, OR = odds ratio, L95 = lower boundary of the 95% confidence interval, L95 = upper boundary of the 95% confidence interval.

**Supplementary Table 11: (Excel sheet) Association results of TSH and free T4 with TSH PGS in ancestry groups in UK Biobank**

**Supplementary Table 12: (Excel sheet) Association results of diseases with TSH PGS in ancestry groups in UK Biobank**

**Supplementary Table 13: (Excel sheet) 95% credible sets**All variants in the 95% credible sets are shown. The posterior probability (PIP column) returned from Polyfun Susie or Wakefield (method column). Allele1 is the coded/effect allele**.**

Supplementary Table : Clinical codes used to define thyroid-stimulating hormone (TSH), free T4, hypo- and hyperthyroidism, other thyroid diseases, and thyroid cancer.

| **TSH** | **T4** | **Hypothyroidism** | **Hyperthyroidism** | **Other thyroid diseases** | **Thyroid cancer** |
| --- | --- | --- | --- | --- | --- |
| 442..  442A.  442A0  442A1  442e.  442K.  442L.  442M.  442N.  442O.  442P.  442Q.  442R.  442S.  442T.  442W.  442X.  X80Gc  XaBj2  XaELV  XaELW  XaESa  XaESb  XaESc  XaESX  XaESY  XaESZ  XaET1  XaET2  XaET7  XaETX  XaIzE  XE27u  XE2wy | 4427.  442V.  442c.  4426.  XaERr  XaERs | C04..  C0430  C0431  C0432  C043z  C044.  C046.  C047.  C04y.  C04z.  C0A5.  Cyu11  PK251  X40HD  X40HE  X40HF  X40HG  X40HH  X40HI  X40HL  X40HM  X40HN  X40HO  X40HP  X40Hv  X40Hw  X40Hx  X40Hy  X40Hz  X40IB  X40IQ  X40OP  X50Fa  XE108  XE10A  XE124  Xa08g  Xa3ec  Xa3ed  XaJ9F | C02..  C020.  C0200  C0201  C020z  C021.  C0210  C0211  C021z  C022.  C0220  C0221  C022z  C023.  C0230  C0231  C023z  C024.  C0240  C0241  C024z  C025.  C02y.  C02y0  C02y1  C02y2  C02yz  C02z.  C02z0  C02z1  C02zz  C05y4  Cyu13  X40Gj  X40Gk  X40Gl  X40Gn  X40Go  X40Gt  X40Gv  X40Gw  X40Gx  X40Gy  X40Gz  X40H0  X40H1  X40H2  X40H4  X40H5  X40H7  X40Hd  X40I4  X40OI  X40OK  X40OO  XE104  XE105  XE106  XE122  XE1g7  Xa09v  Xa3eb  XaZtG | C0...  C00..  C000.  C00z.  C01..  C010.  C011.  C01z.  C05..  C050.  C0500  C0501  C050z  C051.  C052.  C053.  C054.  C05y.  C05z.  C061.  C0A3.  C0A4.  C0AX.  Cyu12  Cyu14  Cyu16  L1810  L1811  L1812  L1813  L1814  L1815  L181z  X40Dj  X40Ge  X40Gf  X40Gi  X40HR  X40HS  X40HT  X40HU  X40HV  X40HW  X40HY  X40HZ  X40Hc  X40He  X40Hf  X40Hg  X40Hh  X40Hu  X40IE  X40IG  X40II  X40IJ  X40OR  X76FB  XE101  XE102  XE103  XE10B  XE10C  XE10D  XE120  XE126  XM0Ap  Xa08f  Xa3ea | C73  193  V1087 |

**Supplementary Table 15: (Excel sheet) FUMA eQTL look up results**Look up of eQTL results from the GTEx v8 (thyroid, hypothalamus and pituitary tissues) and eQTLGen (blood, cis- and trans-eQTLs) datasets for variants in our 95% credible sets.

**Supplementary Table 16: (Excel sheet) Polygenic Priority Score (PoPS) results - 500KB window**Polygenic Priority Score (PoPS) results using a +/-250KB window

**Supplementary Table 17: (Excel sheet) Genes near our signals associated with rare Mendelian respiratory diseases**34 genes associated with a rare Mendelian respiratory disease were implicated.

**Supplementary Table 18: (Excel sheet) Mouse ortholog genes near our signals associated with a respiratory disease**  
6 genes were implicated

**Supplementary Table 19: (Excel sheet) Look up of UK Biobank WES variants**Look up of UK Biobank WES variants within +/-500kb of TSH sentinels.

**Supplementary Table 20: (Excel sheet) Putative causal variants**Variants from credible sets annotated as missense/damaging/deleterious and with a PIP>50%