WP5: Exploiting Large Datasets to Understand the Genetics and Clinical Heterogeneity of PID

Chris Wallace, Ken Smith
Tom Willis

Aims of WP5

Task 5.1 Coheritability of IMD and other molecular phenotypes with PID

Task 5.2: Identify common PID-associated variants

Task 5.3: Leverage common disease genetics to aid rare variant analysis

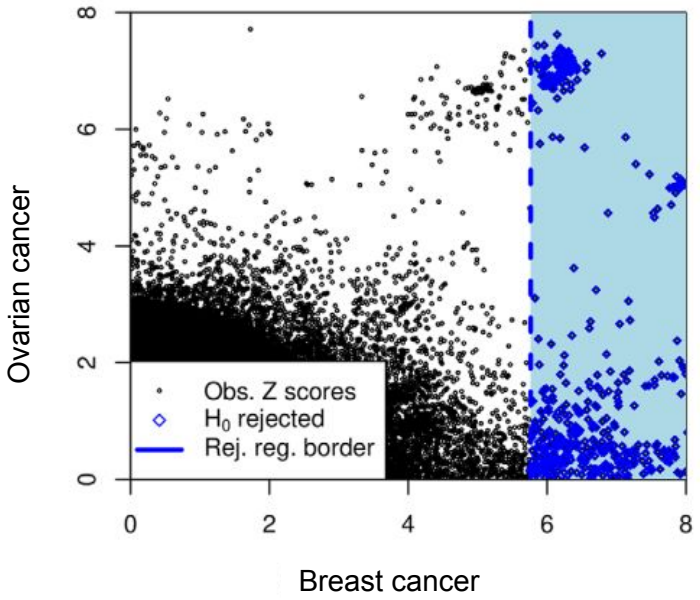
Task 5.4: Genetic regulation of immunophenotypes generated in WP2, WP3

Task 5.5: Reverse genetics: phenotype of carriers of PID variants in EPIC, UK Biobank

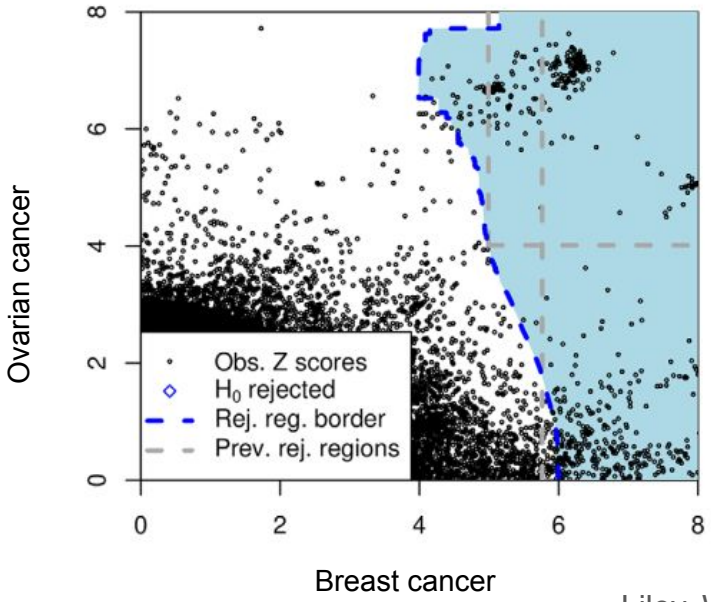
Integration of PID GWAS with other IMD (Task 5.2)

With conditional FDR we can adapt our rejection region to reject slightly larger p values when a related trait has smaller p values

Breast cancer only

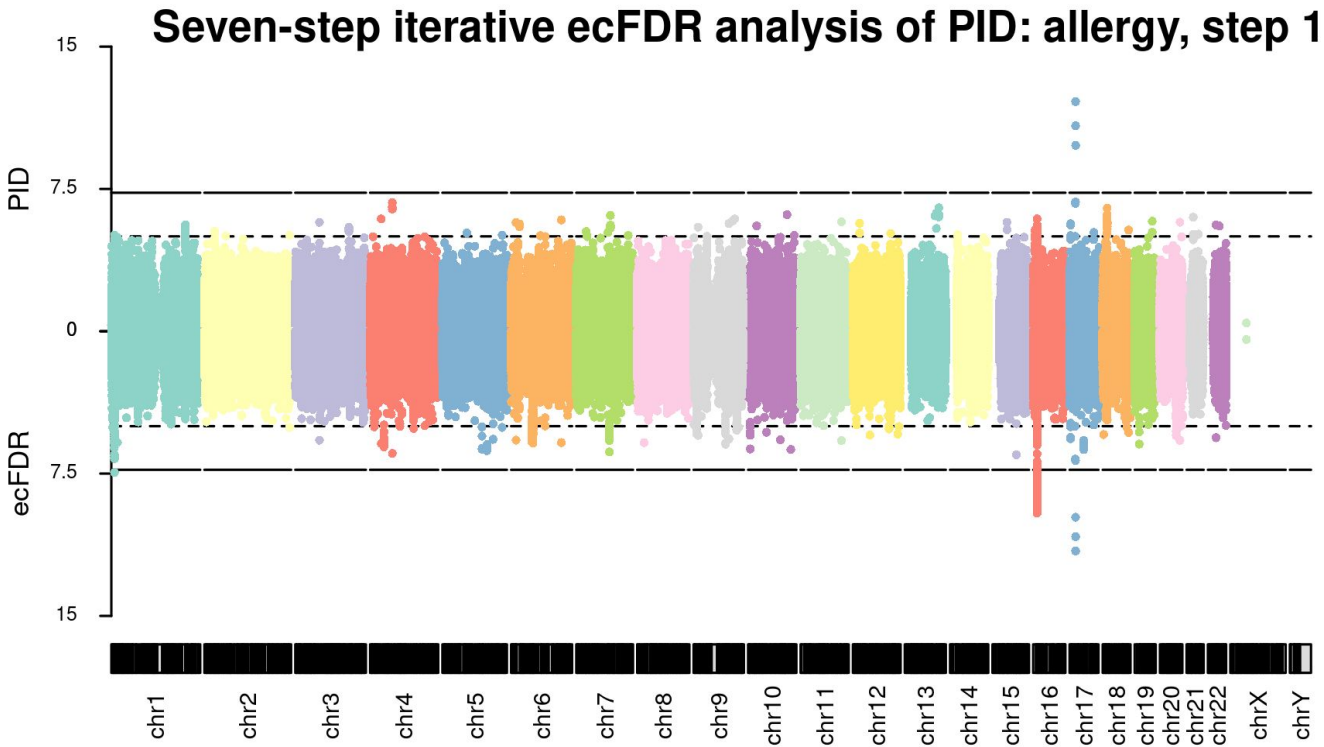


Breast cancer | ovarian cancer

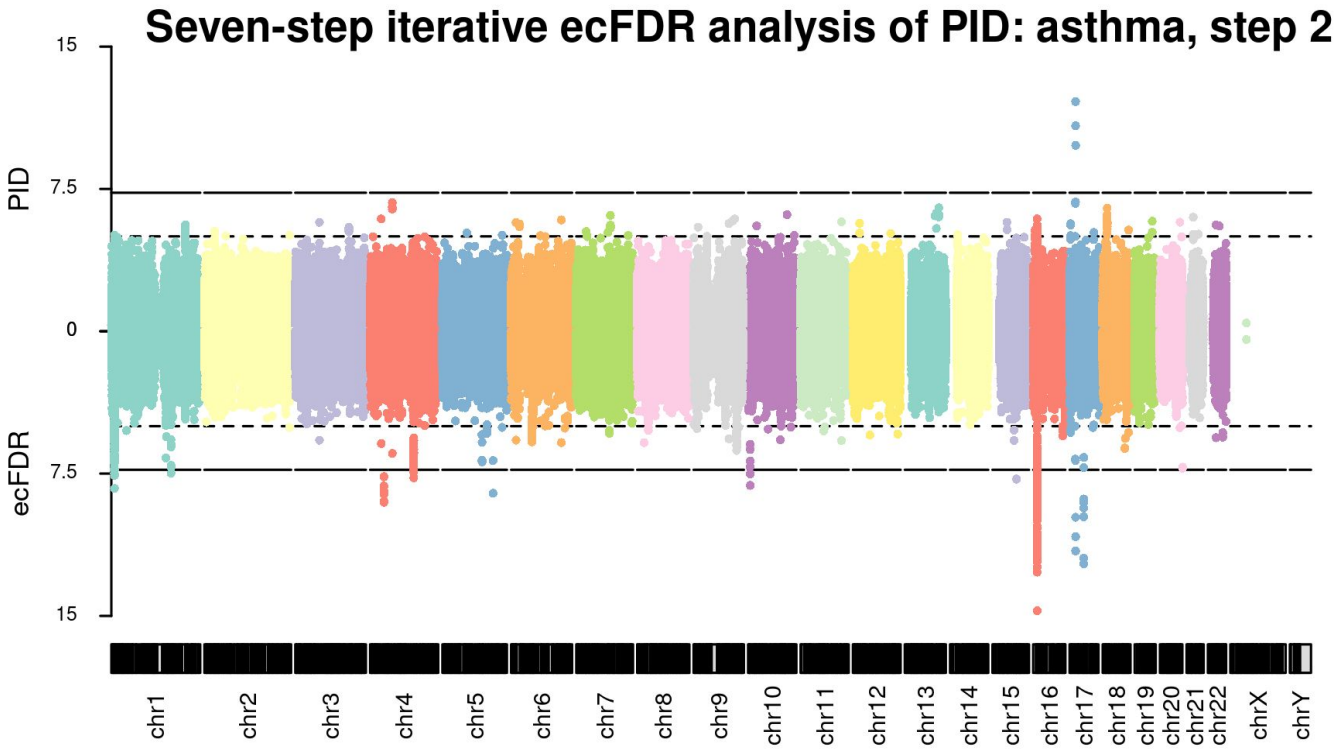


Liley, Wallace 2021

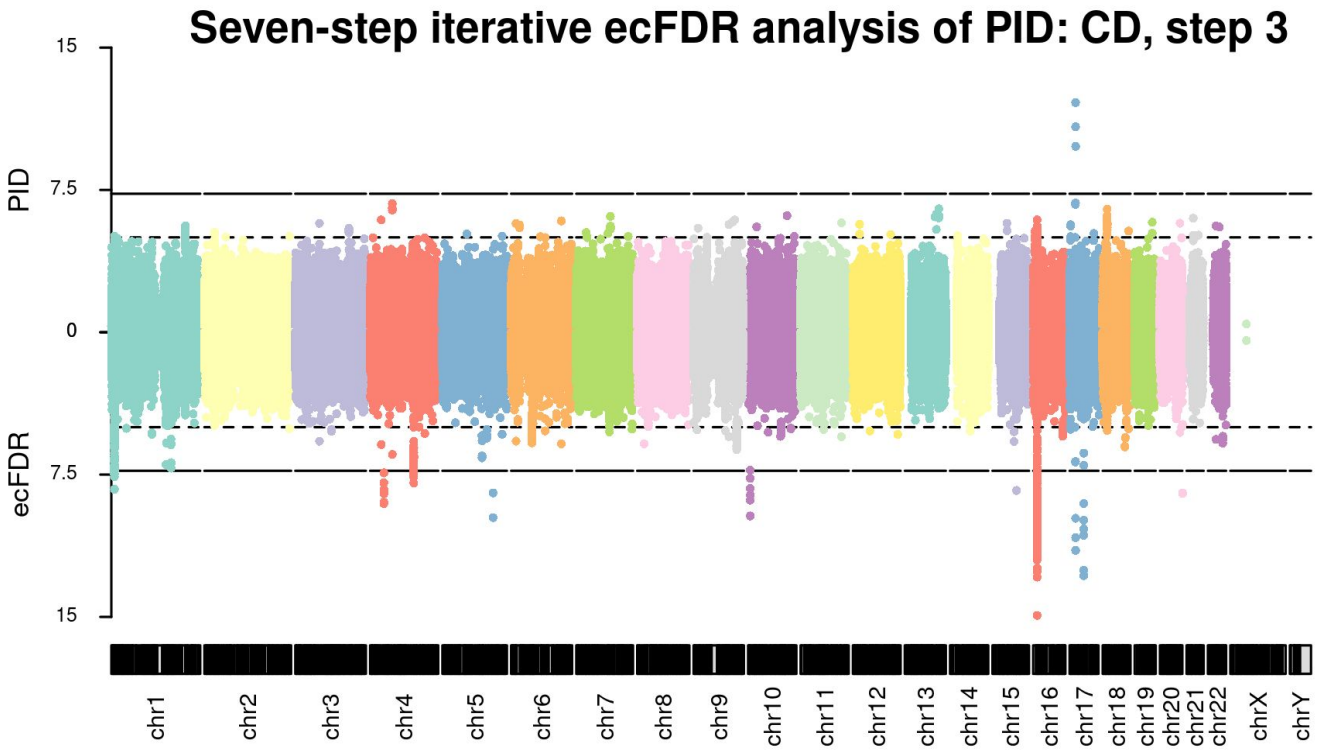
Preliminary results of leveraging seven IMD with PID



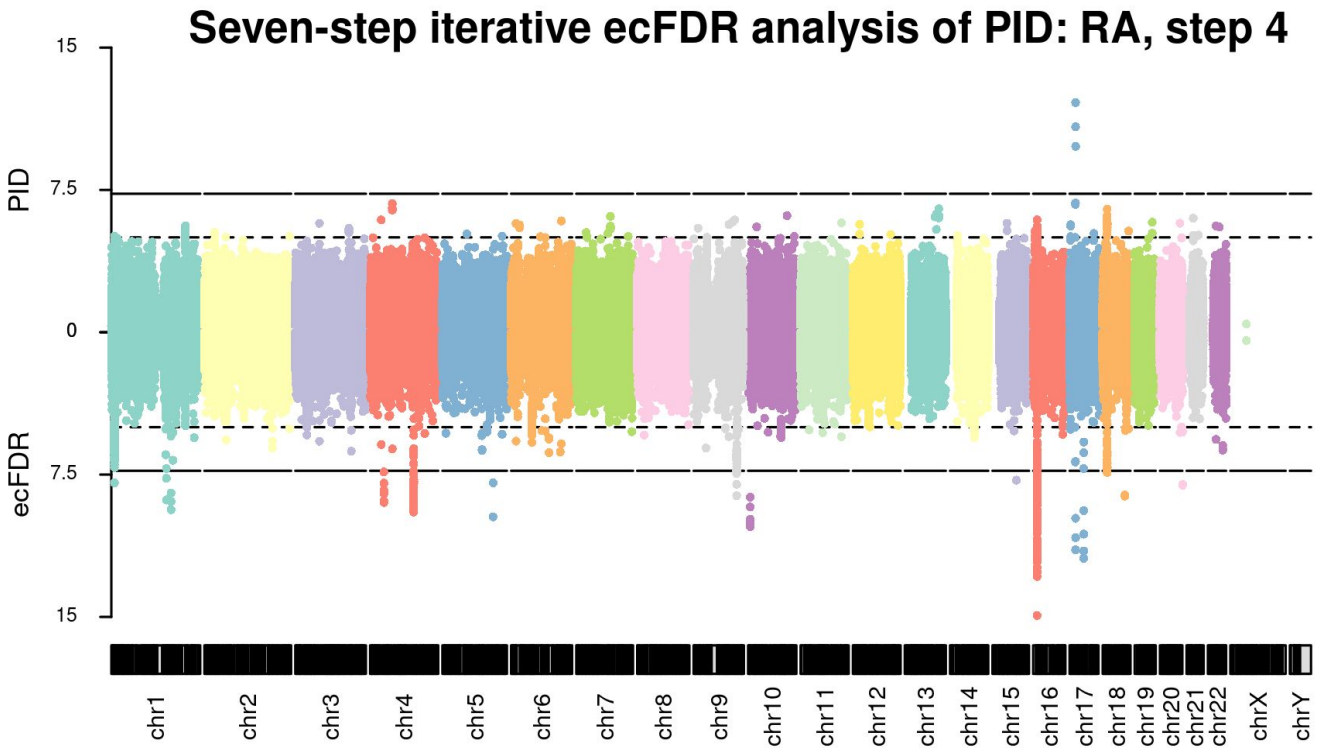
Preliminary results of leveraging seven IMD with PID



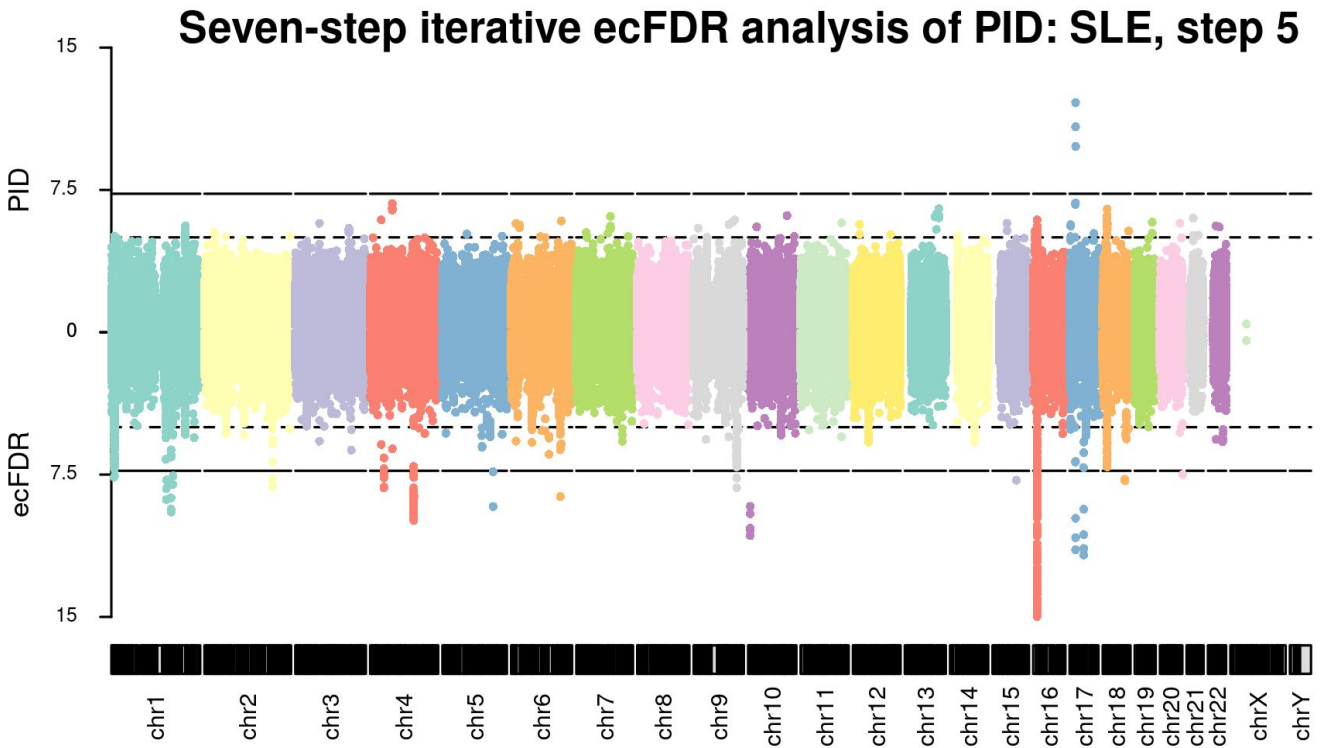
Preliminary results of leveraging seven IMD with PID



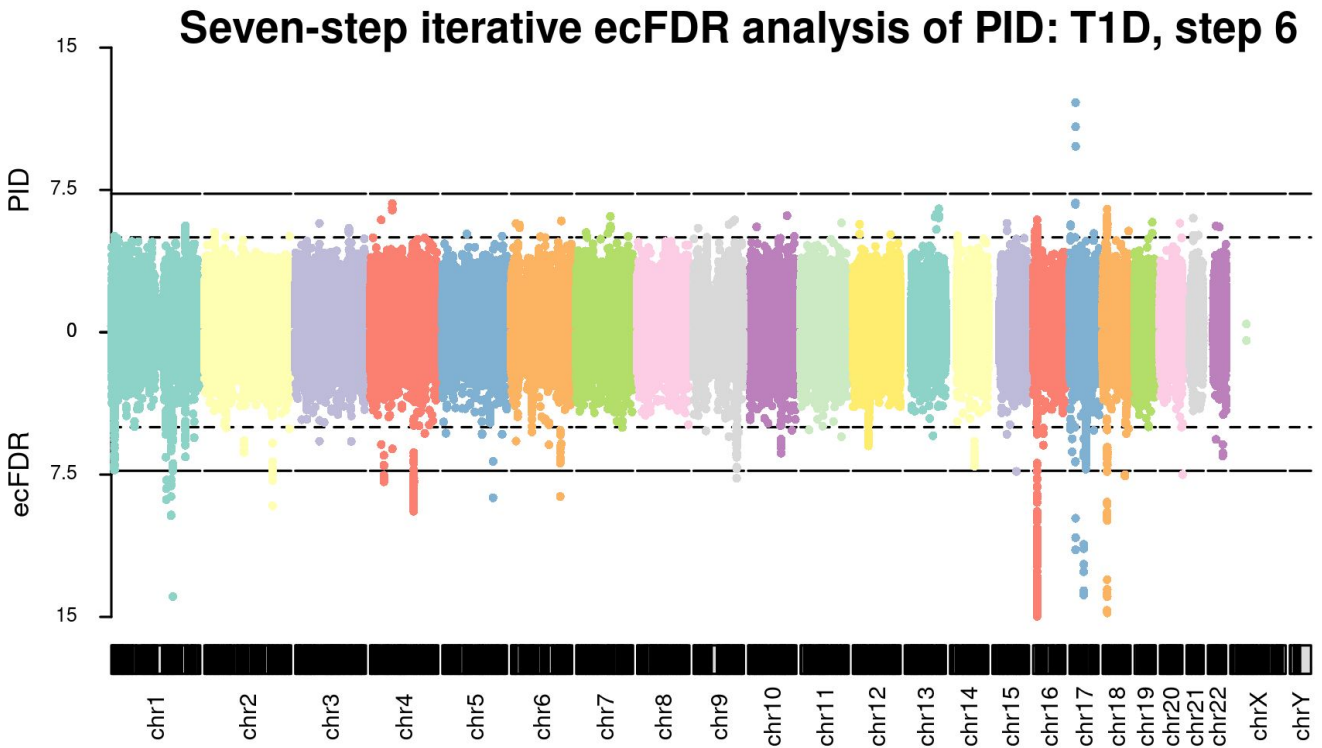
Preliminary results of leveraging seven IMD with PID



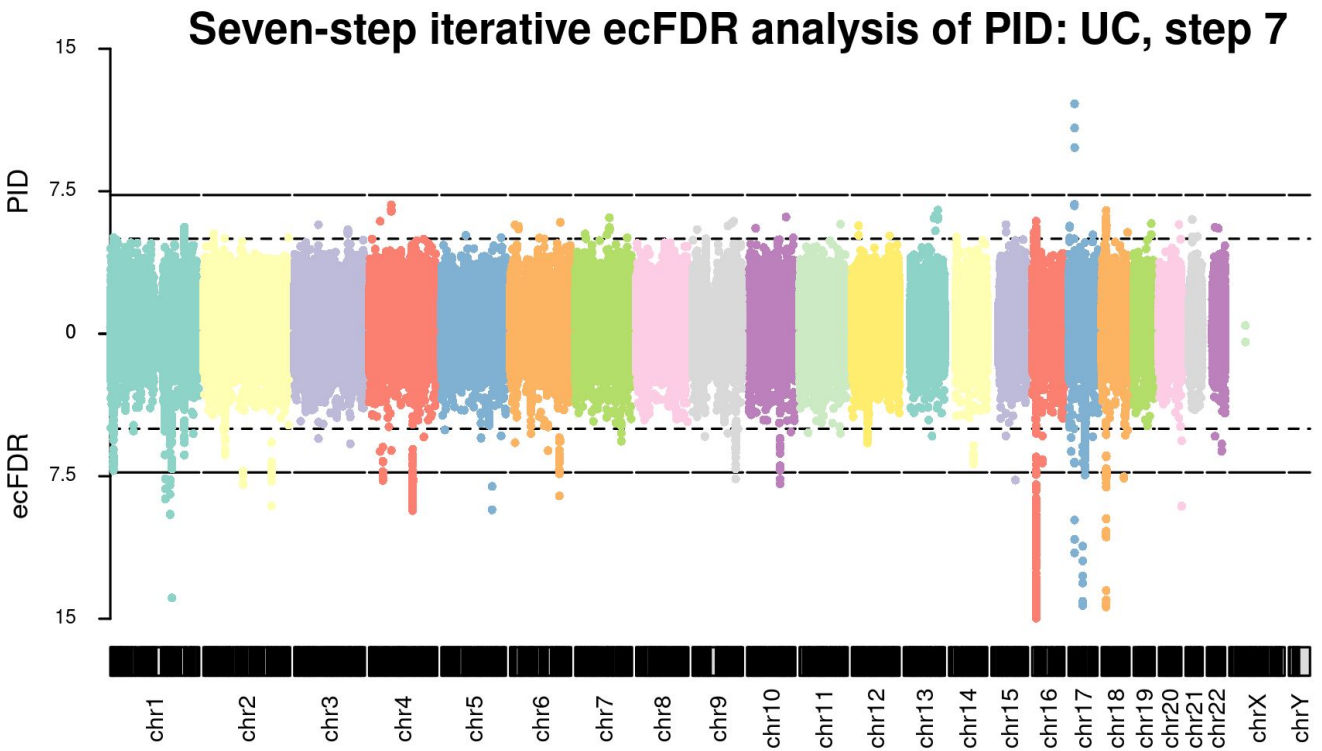
Preliminary results of leveraging seven IMD with PID



Preliminary results of leveraging seven IMD with PID



Preliminary results of leveraging seven IMD with PID



rsID	Chr.	v-value	Prin. p-value	MAF	Gene	Local genes
rs6679677	1	1.54e-23	9.58e-03	0.09		*PHTF1, RSBN1*
rs4845623	1	2.00e-09	7.97e-03		*IL6R*	*RP11-350G8.5, PSMD8P1, SHE, RP11-350G8.9*
rs840016	1	3.05e-10	9.89e-03	0.38	*CD247*	*POU2F1, RP11-104L21.2, RP11-104L21.3*
rs78037977	1	1.26e-14	3.79e-03	0.12		*SLC25A38P1, RP3-332O11.2, RP1-15D23.2*
rs72837847	2	1.17e-08	3.18e-03	0.13	*MIR4435-2HG*	*AC108463.1, AC108463.2, RP11-181E10.3, AC068491.2, RP11-803D5.4*
rs13390894	2	8.77e-10	8.91e-03	0.26	*STAT1*	*GLS, AC067945.3, AC067945.2, AC067945.4, STAT4*
rs62323881	4	1.20e-09	7.24e-04	0.08		*AC097533.1, RN7SL335P, KIAA1109*
rs76487164	4	4.96e-10	7.94e-04	0.08	*IL21-AS1*	*IL21*
rs55649498	4	1.90e-08	6.51e-03	0.25		*RNA5SP158, TLR10, TLR1*
rs10068466	5	5.42e-10	9.70e-03	0.30	*NDFIP1*	*CTC-463A16.1*

rsID	Chr.	v-value	Prin. p-value	MAF	Gene	Local genes
rs67297943	6	2.89e-09	5.88e-03	0.19		*TNFAIP3, RP11-10J5.1, RP11-240M16.1*
rs10760122	9	2.29e-08	6.66e-03	0.36	*PHF19*	*PSMD5, PSMD5-AS1, TRAF1*
rs61839660	10	1.65e-22	5.63e-03	0.07	*IL2RA*	*SNORA14, RP11-536K7.5, RP11-414H17.2, RBM17*
rs9419741	10	1.25e-08	3.33e-03	0.48		*KIF11, RN7SL644P, EIF2S2P3, HHEX, Y_RNA*
rs147793459	15	1.97e-08	2.09e-05	0.04	*SMAD3*	*RP11-342M21.2*
rs12927355	16	3.09e-21	2.53e-04	0.33	*CLEC16A*	*RP11-66H6.4, RP11-66H6.3*
rs10445308	17	5.02e-15	7.16e-03	0.48	*IKZF3*	*GRB7*
rs77301847	17	3.65e-08	6.65e-03	0.24	*CRHR1*	*RP11-293E1.2, RP11-293E1.1*
rs1893217	18	9.97e-18	4.67e-04	0.14	*PTPN2*	*Y_RNA, RP11-973H7.1*
rs4369774	18	2.54e-08	7.19e-03	0.45	*TNFRSF11A*	*KIAA1468, RP11-173A16.1, Y_RNA, RP11-640A1.3*
rs909334	20	8.28e-10	5.33e-03			*GMEB2, CTD-3184A7.4, STMN3, RTEL1, RTEL1-TNFRSF6B*

Choice of diseases to condition on (Task 5.1)

Can’t iterate forever: these diseases are related and we would lose control of type 1 error

Must choose a small number that are

a. Informative for PID
b. Relatively independent of each other

Genetic correlation is not useful here because PID sample size too small.

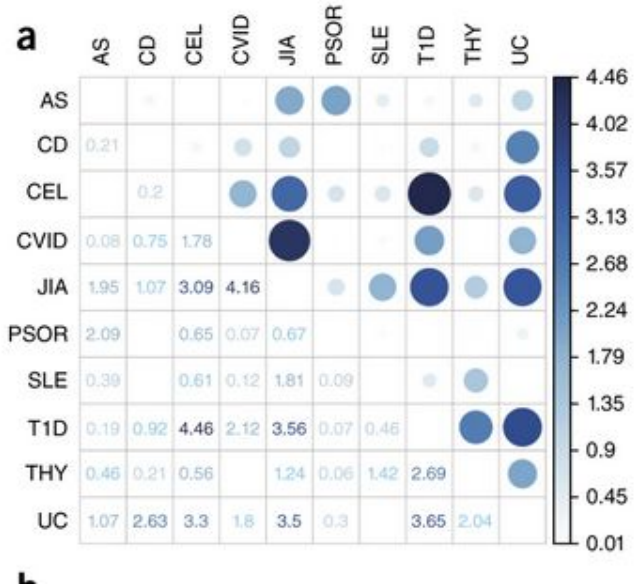
GPS statistic from Li et al (2015)

Published: 24 August 2015

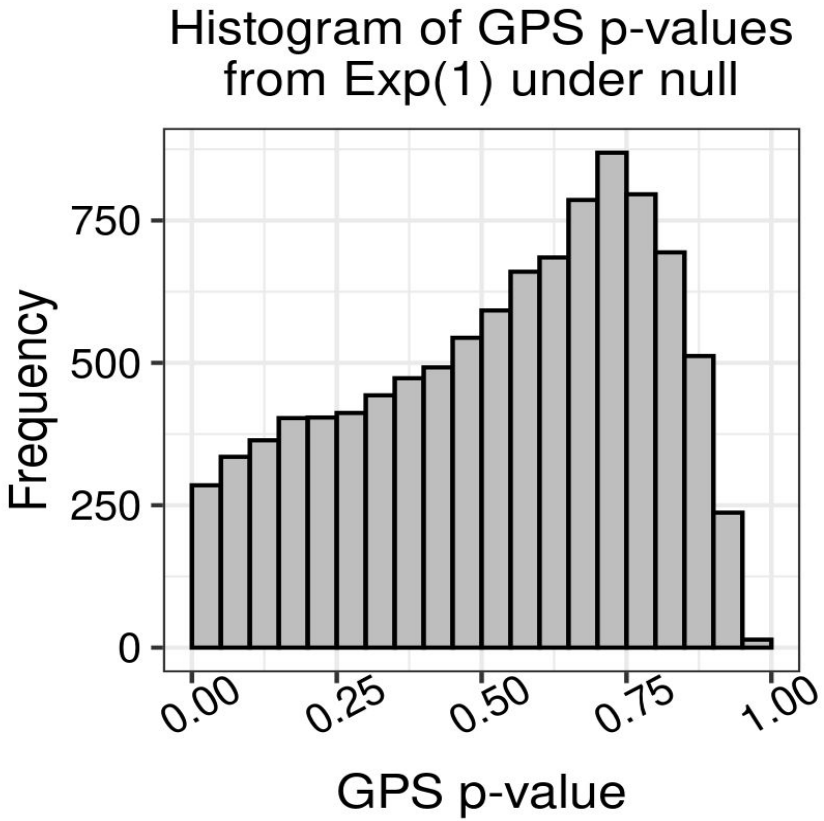
Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases

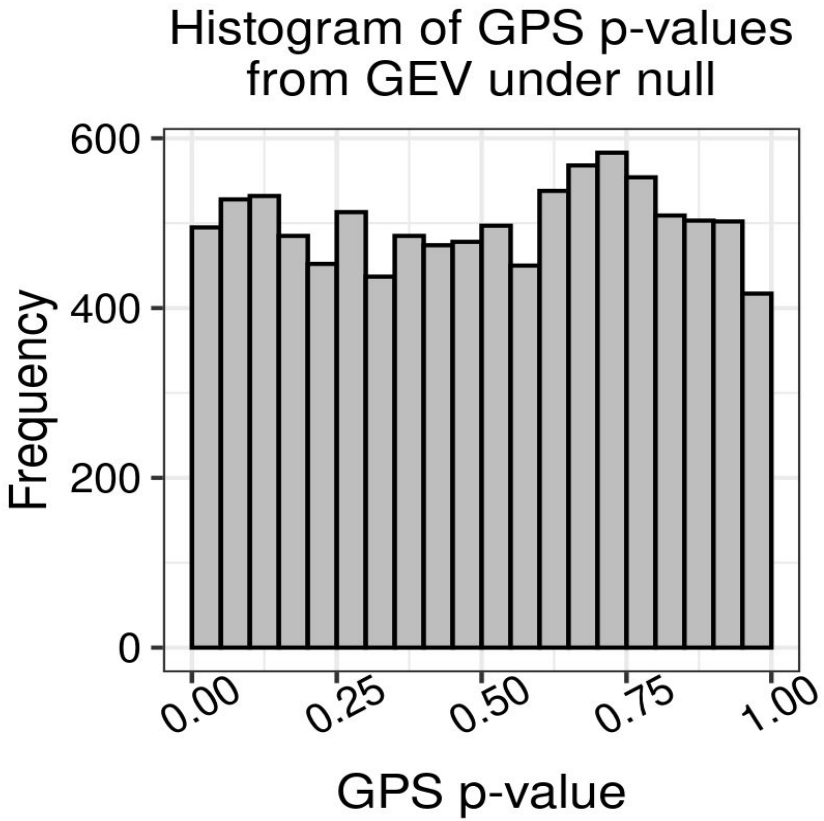
Yun R Li, Jin Li, [...] Hakon Hakonarson

Nature Medicine **21**, 1018–1027 (2015) | Cite this article



GPS statistic from Li et al (2015)

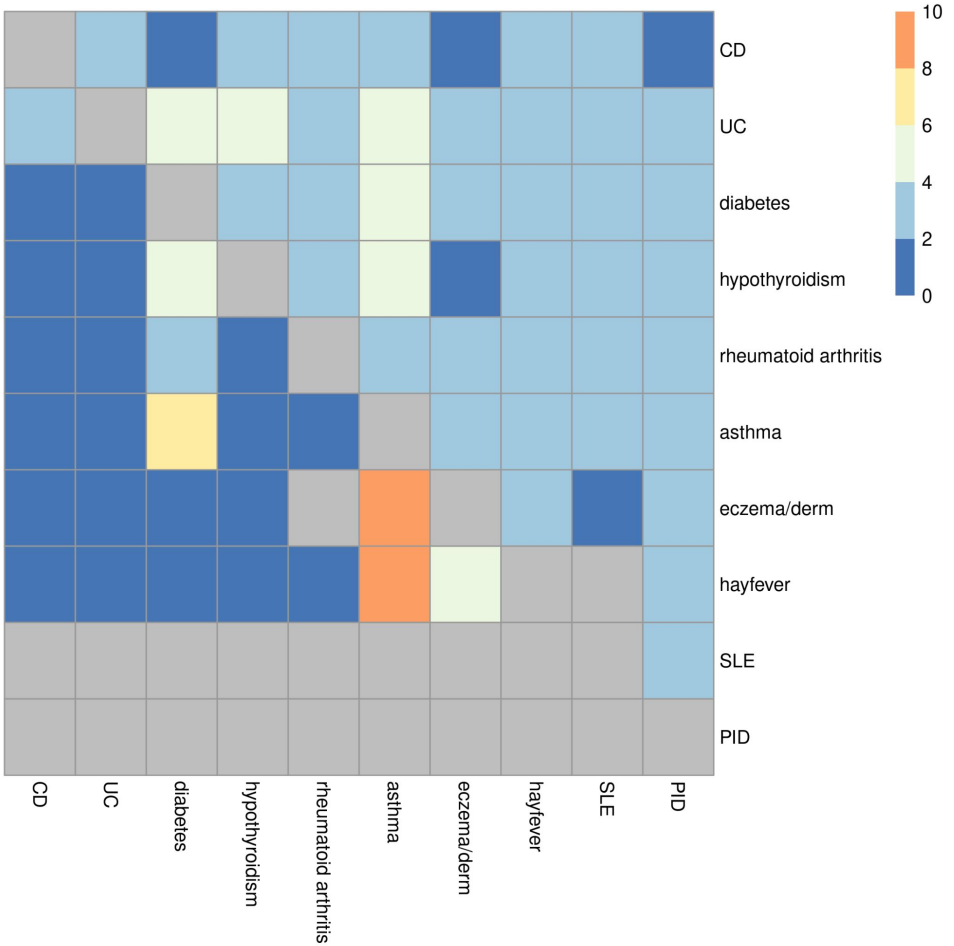




GPS statistic from Li et al (2015)

GPS p value gives results across range of sample sizes

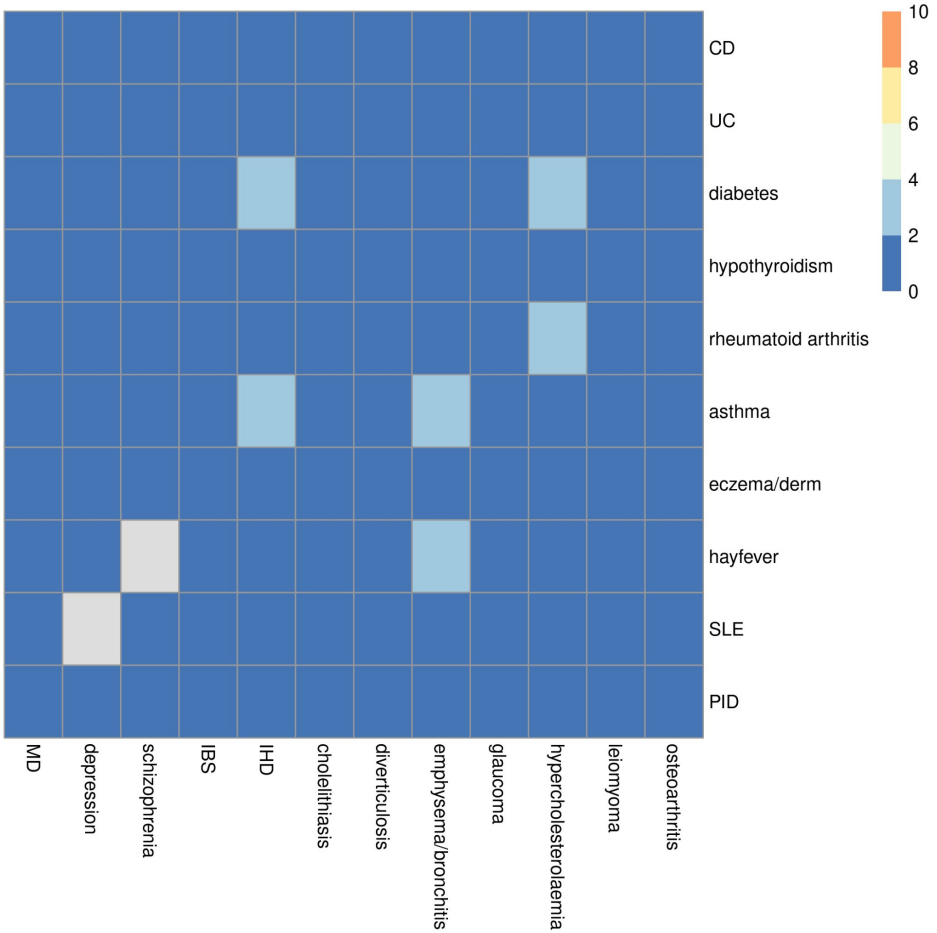
r_g p value dependent on sample size



GPS statistic from Li et al (2015)

GPS p value

looks flat when comparing immune & non-immune traits



Next steps

Full coheritability analysis across broad range of diseases, IgG/M/A levels in 8,000 samples from EPIC-Norfolk

Select optimal subset and run cFDR using existing PID GWAS data

Update PID GWAS when additional 800 case data are available

Update coheritability and cFDR analyses

Extend method to work with rare variants, leveraging “linked” common variants

Acknowledgements

Tom Willis, PhD student has led the analyses presented today

UNIVERSITY OF CAMBRIDGE

MRC Biostatistics Unit

wellcome

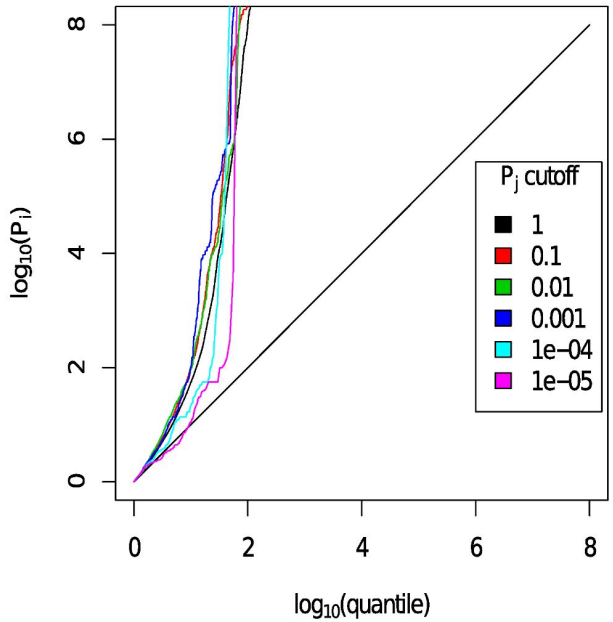
gsk

MERCK

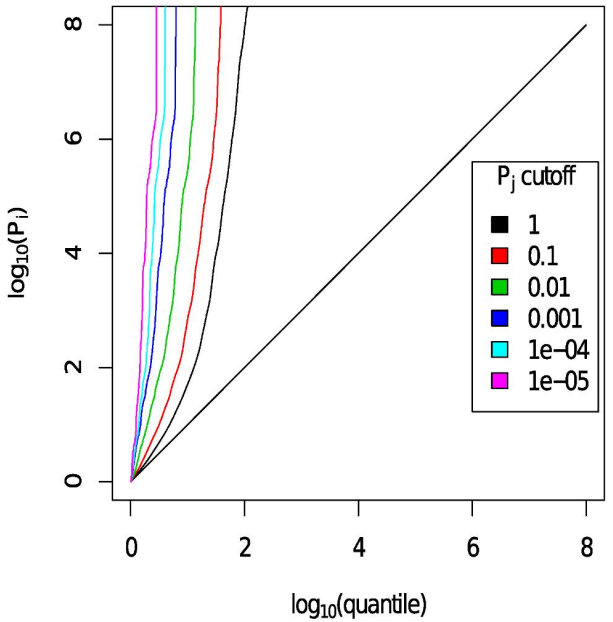
Integration of PID GWAS with other IMD (Task 5.2)

Conditional FDR is one way to leverage information from a larger GWAS

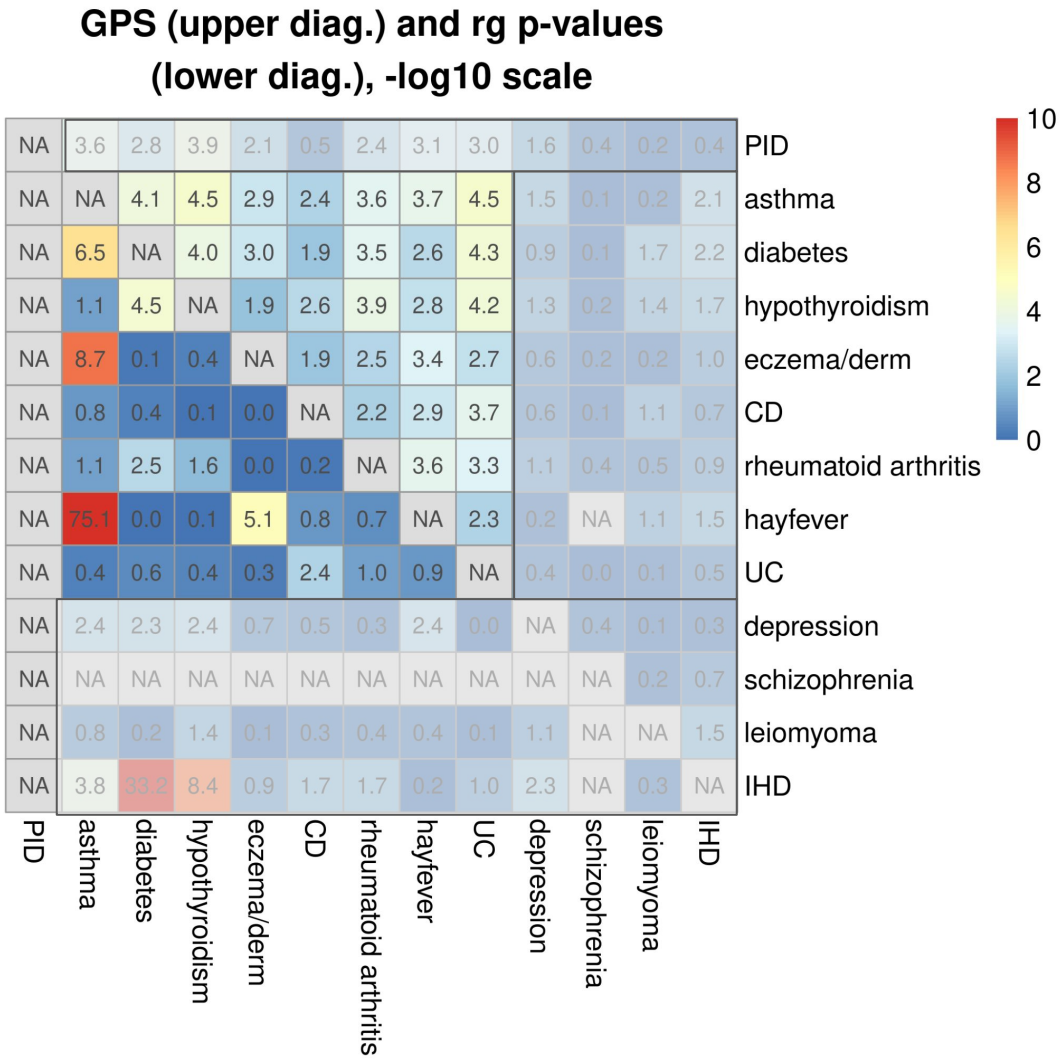
Type 1 diabetes | psoriasis



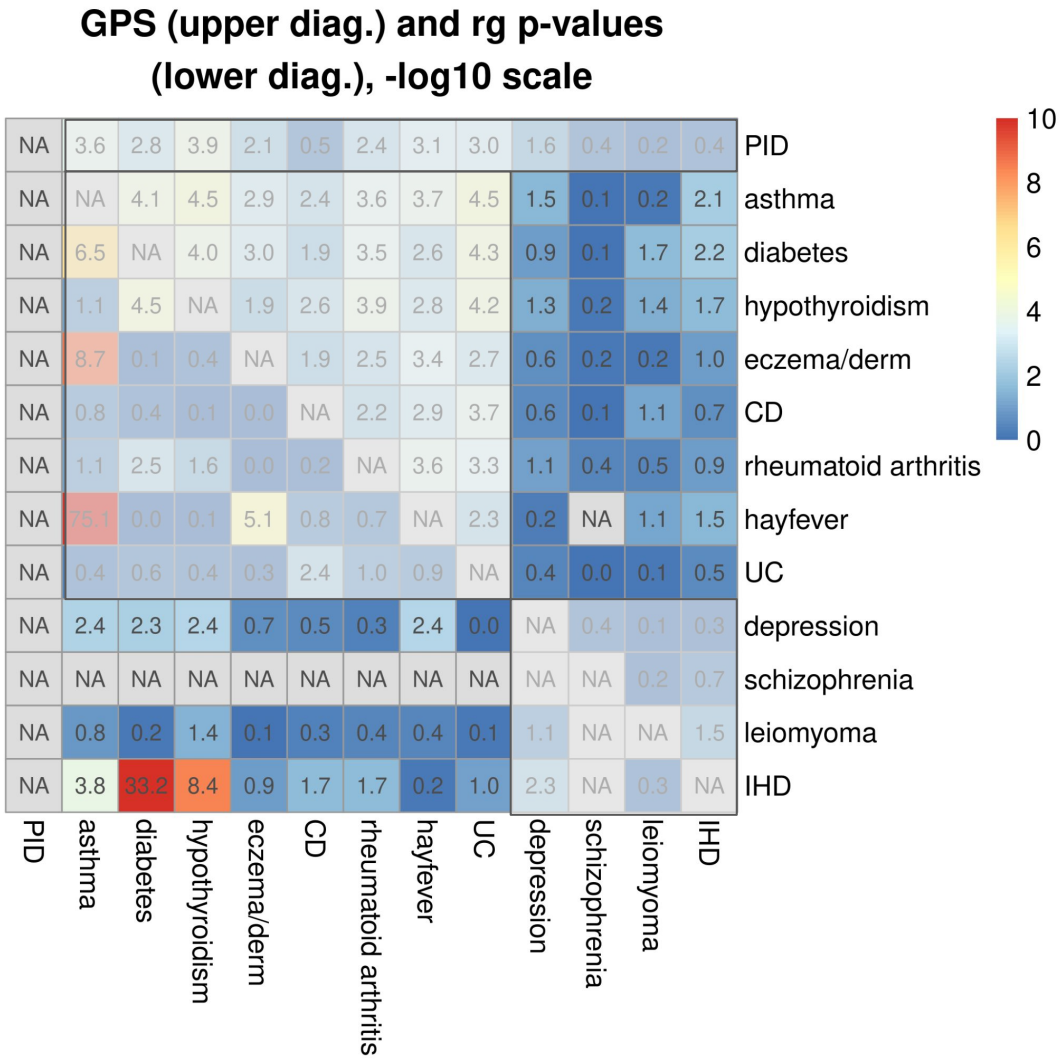
Type 1 diabetes | rheumatoid arthritis



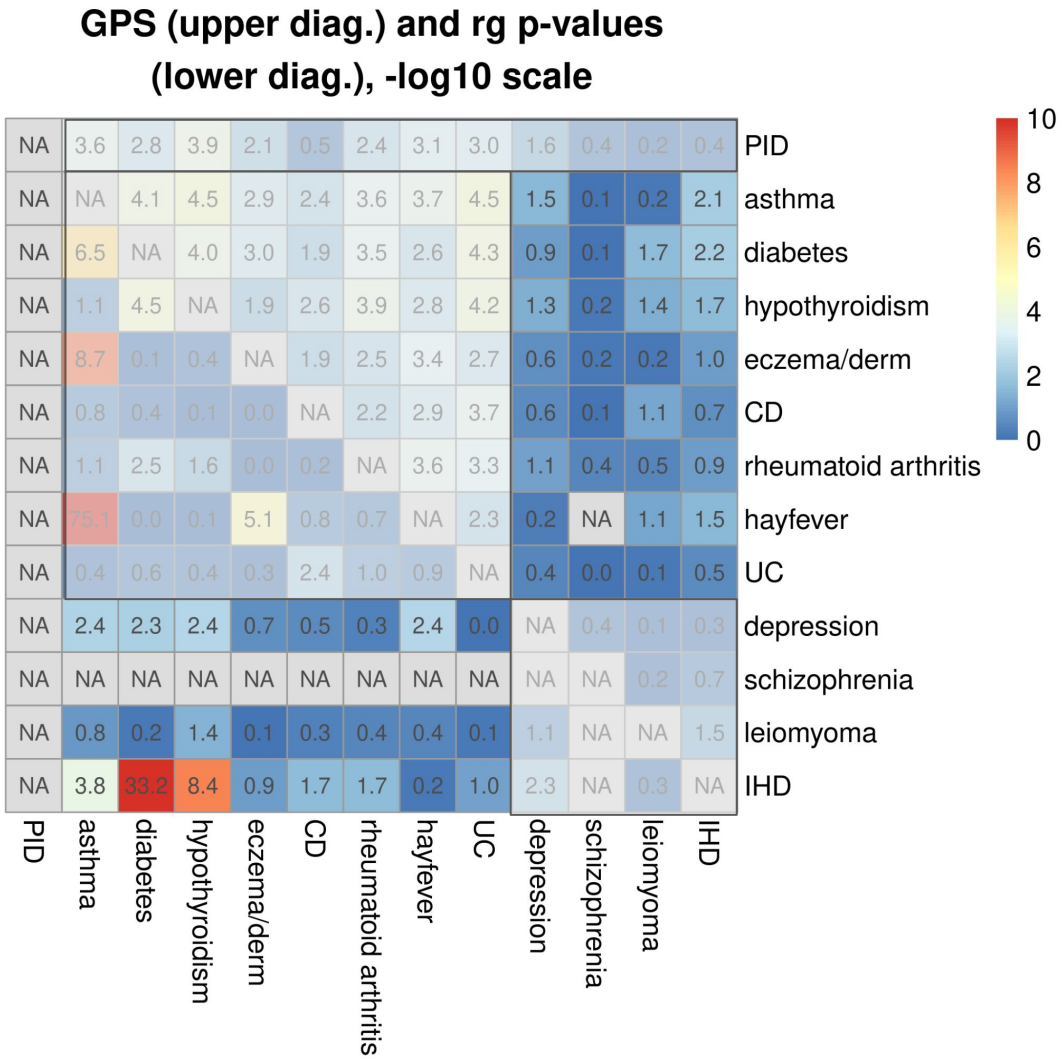
Solution: GPS statistic from Li et al (2015)



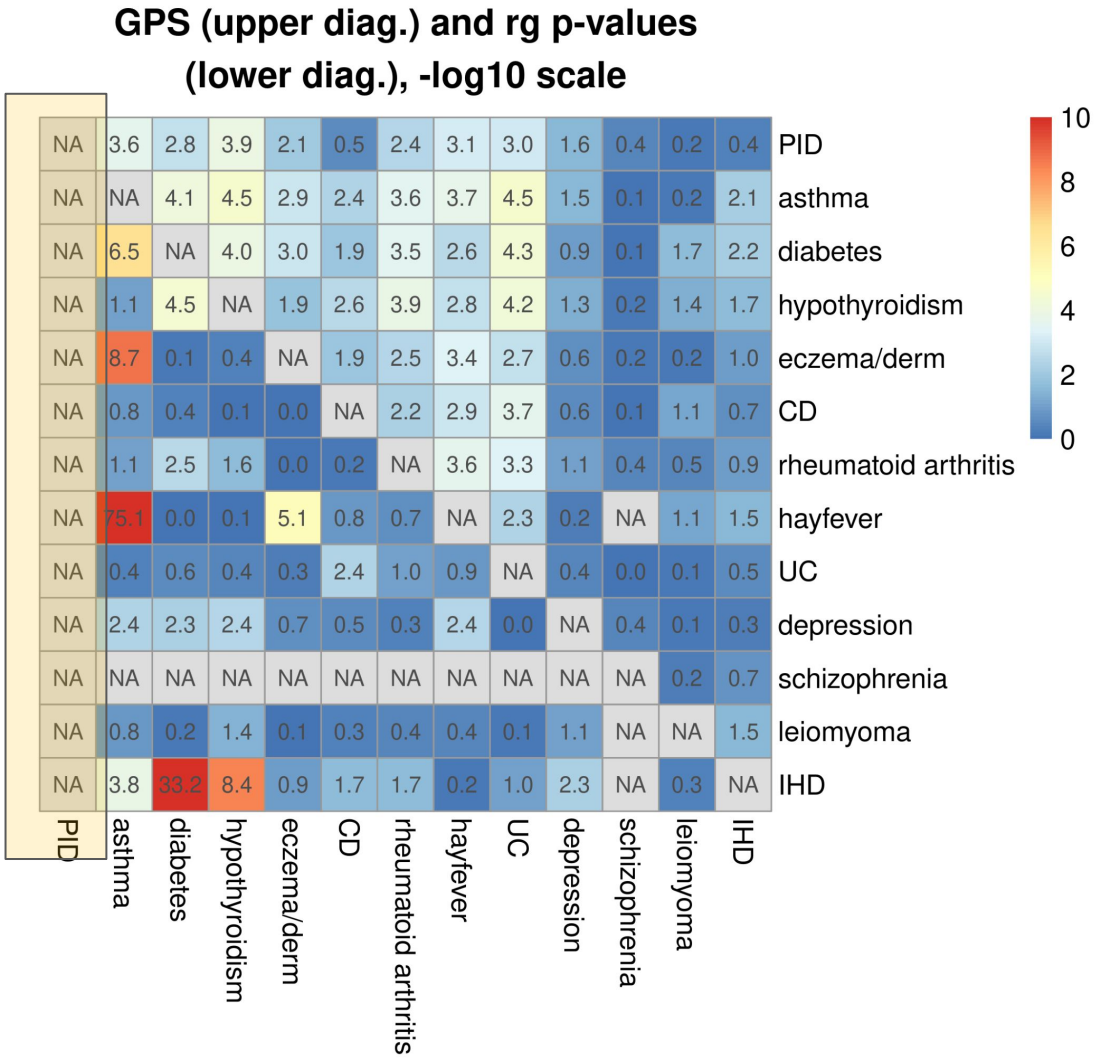
Solution: GPS statistic from Li et al (2015)



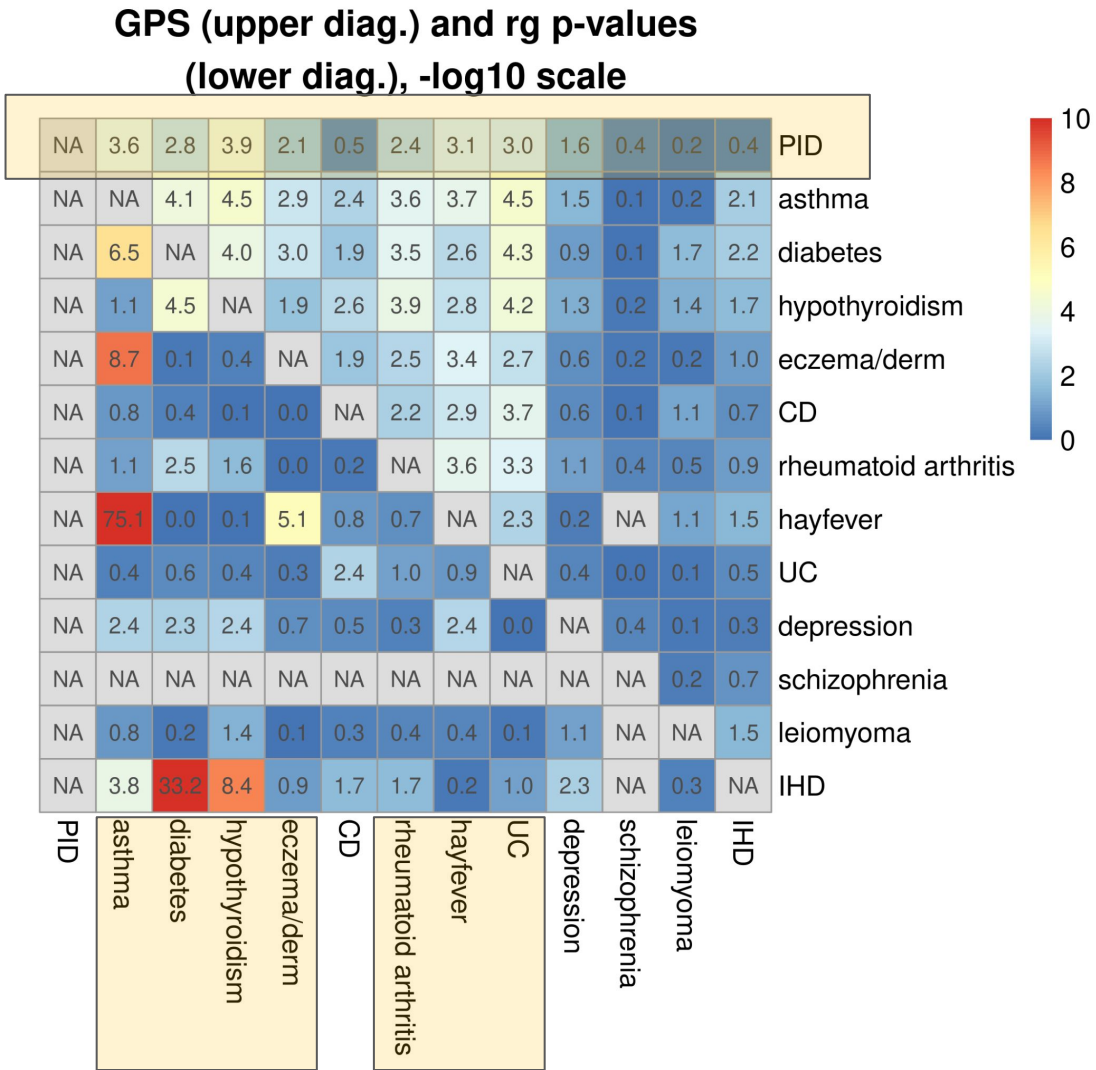
Solution: GPS statistic from Li et al (2015)



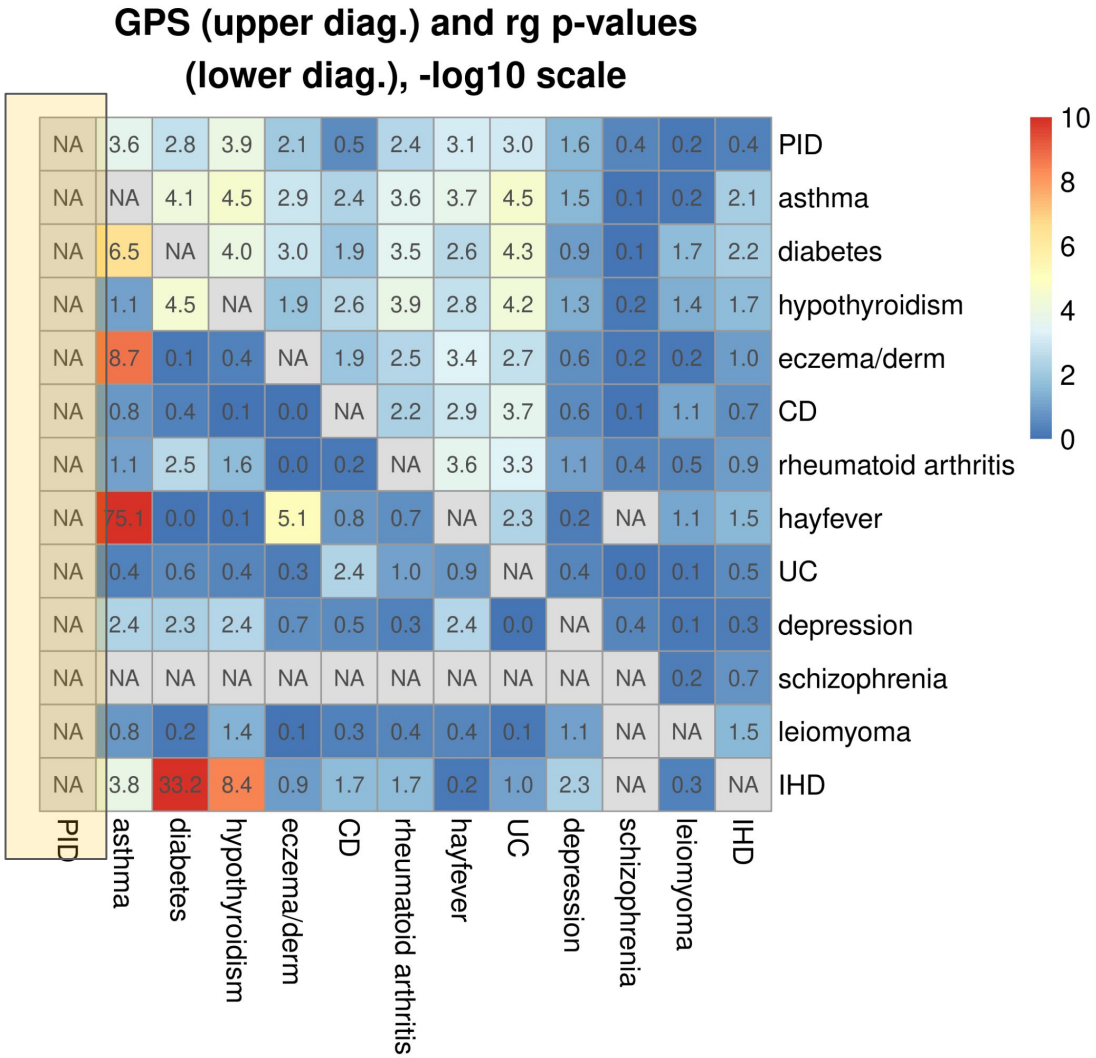
Solution: GPS statistic from Li et al (2015)



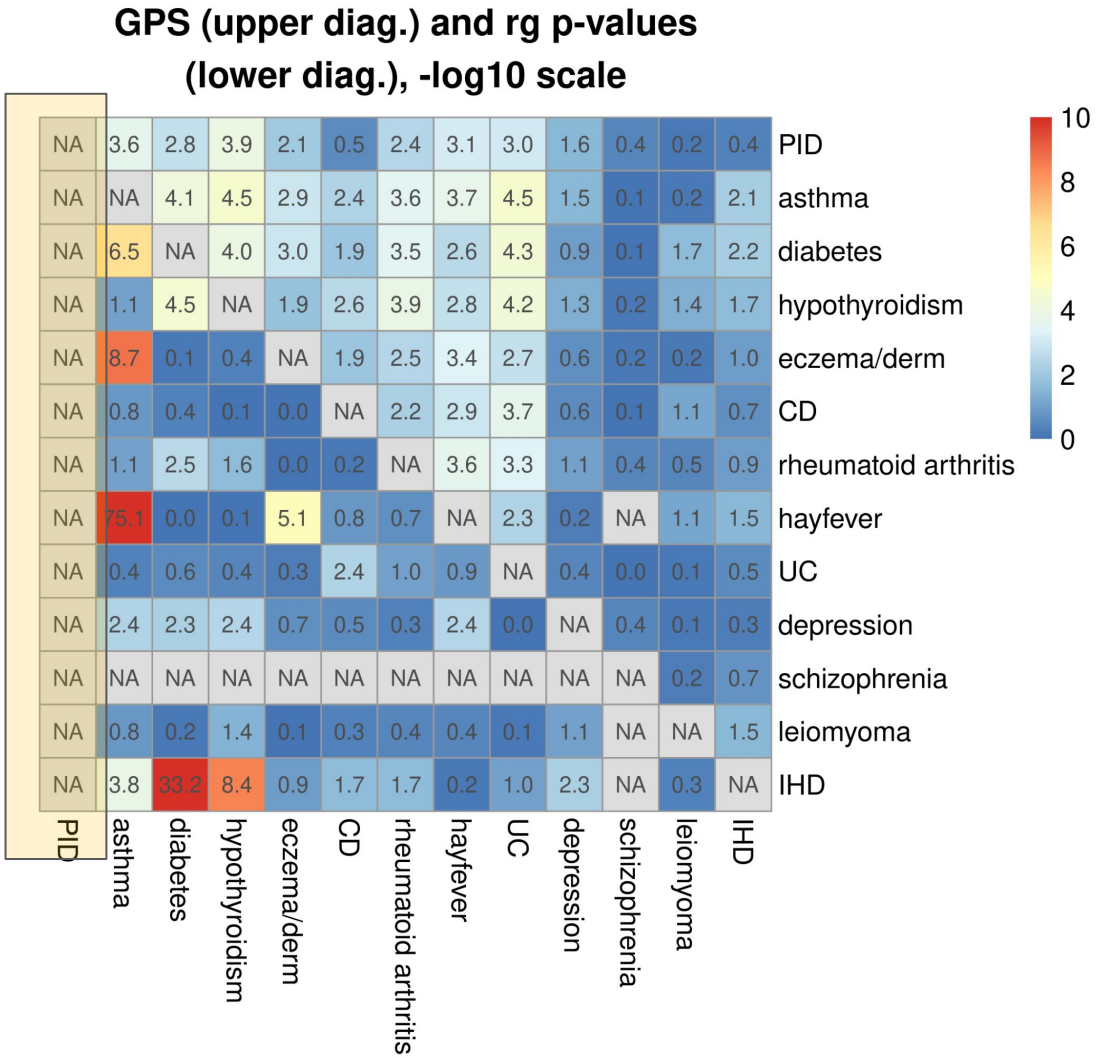
Solution: GPS statistic from Li et al (2015)



Solution: GPS statistic from Li et al (2015)



Solution: GPS statistic from Li et al (2015)



Choice of diseases to condition on (Task 5.1)

Can’t iterate forever: these diseases are related and we would lose control of type 1 error

Choice of diseases to condition on (Task 5.1)