

PVP Evaluation

- Synthetic data
 - 8,746 ClinVar pathogenic variants appended to WEX from 1000 Genome Project
 - 11,251 ClinVar pathogenic variants appended to WGS from 1000 Genome Project
- Real patient data
 - **19 exomes** from **UK10K_RARE_THYROID** study sample dataset.
 - 43 exomes from Addenbrooke's Hospital at Cambridge, UK with diseases of diverse phenotypes (e.g Bardett-Biedl Syndrome)
 - 1 patient's exome (Glutaric aciduria)
 - 5 genomes from Personal Genome Project

Synthetic WEX patients

	Top Hit	Top 10	ROC AUC
GWAVA	1.41%	6.32%	0.711
DANN	6.06%	26.69%	0.928
eXtasy	14.85%	42.99%	0.782
CADD	15.15%	32.05%	0.946
Exomiser	24.65%	58.56%	0.89
Phevor	28.25%	64.70%	0.912
PVP	78.80%	89.50%	0.978

Synthetic WGS patients

	Top Hit	Top 10	ROC AUC
GWAVA	0.46%	0.70%	0.949
DANN	4.95%	18.99%	0.99
CADD	5.76%	24.30%	0.962
Genomiser	31.35%	77.98%	0.948
PVP	76.47%	88.61%	0.995

UK10K Data

Sample ID	Mutation	PVP	Phevor	eXtasy	Exomiser
UK10K_THY5329055	TG c.1583C>A (hom)	1	2	209	5
UK10K_THY5370898	TG c.638+5G>A (hom)	3		92	13851
UK10K_THY5329054	TG c.4478G>A (hom)	3	11	127	29
UK10K_THY5329061	TG c.8054G>T (hom)	2	8	74	40
UK10K_THY5236180	TG c.5071T>C (het)	3	5	67	2189
UK10K_THY5236181	TG c.7640T>A (het)	1	4	306	11078
UK10K_THY5329045	TG c.2311C>T (het)	1	8	91	110
UK10K_THY5329046	TSHR c.1637G>A (het)	1	130	9	14
UK10K_THY5068932	TG c.3433+3_3433+6delGAGT (het)	8		246	21654
UK10K_THY5329060	DUOX2 C.2056C>T (het)	1	43	268	4327