

Towards implementation of precision medicine across countries

Eivind Hovig

Precision medicine – Precision diagnostics

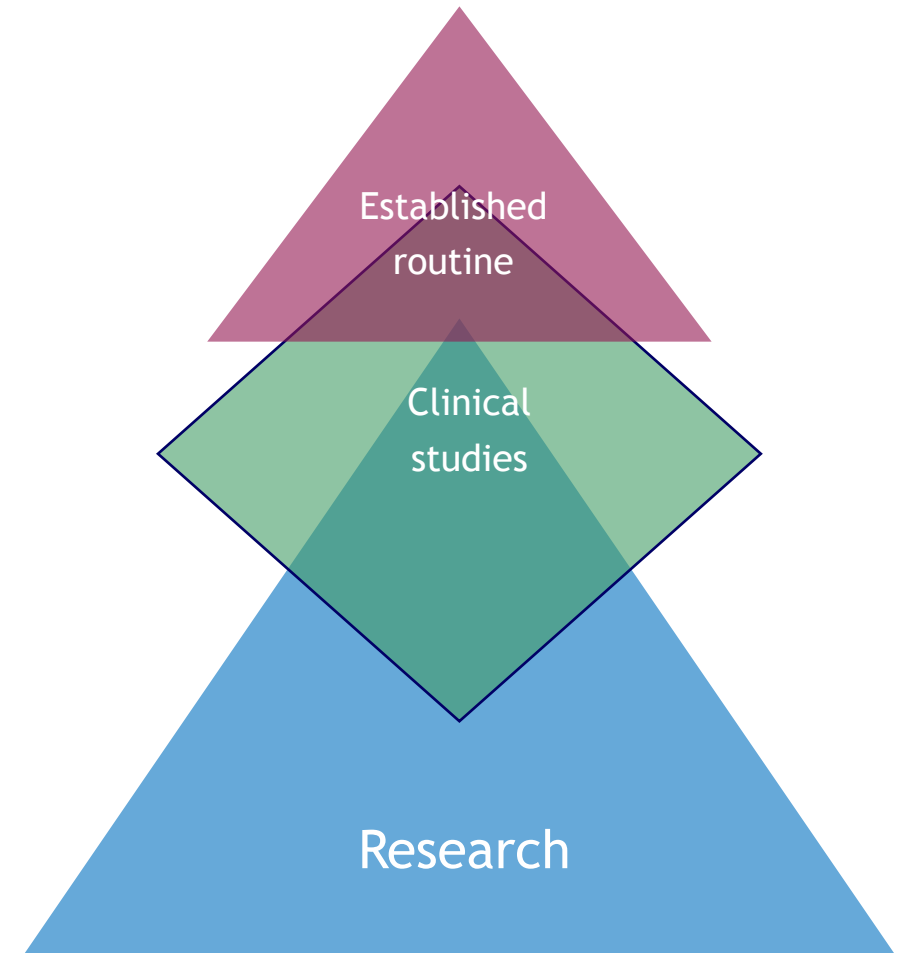
Routine molecular diagnostics

Molecular diagnostics for selection to clinical trials

Molecular diagnostics in clinical trials

Molecular diagnostics for PM targeted treatment

Next generation molecular diagnostics



Genome wide association studies

- Goal: find connections between:
 - A phenotype: height, type-I diabetes, etc., known to be heritable
 - Whole-genome genotype
 - Usually SNP arrays: Sampled points along the genome known to vary between individuals
 - Can be whole genome sequencing (mostly rare due to cost)
 - Mostly case-control comparing variant frequencies at each position to identify deviating frequencies indicating a hit along the genome
 - A hit is an association, not a direct hit
 - Need to huge numbers of individuals for statistical significance

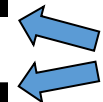
Find the associated SNP

Cases:

AGAGC**A**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**T**G**T**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**A**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****T**C
AGAGC**A**GT**C**GAC**A**GGT**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**A**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C****A**ACAT**G**AG**A**T**C**G**C**T**A**G**A**GC**A**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**T**G**T**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**C**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**T**G**T**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**A**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**A**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****T**C
AGAGC**A**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C

Controls:

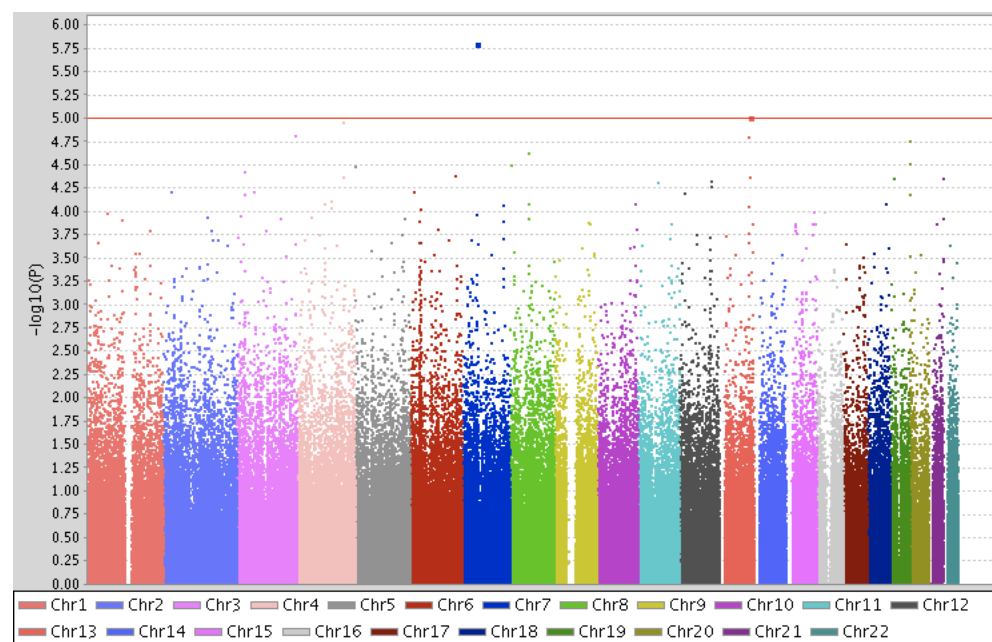
AGAGC**A**GT**C**GAC**A**T**G**T**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C**G**C**T**A**G**A**GC**A**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****C**C
AGAGC**A**GT**C**GAC**A**T**G**T**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C****A**ACAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**A**GT**C**GAC**A**T**G**T**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****T**C
AGAGC**C**GT**C**GAC**A**GGT**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C****A**ACAT**G**A**T**A**G****C**C
AGAGC**A**GT**C**GAC**A**GGT**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**A**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**GGT**A**T**A**GC**C**CT**A**CAT**G**AG**A**T**C**G**A**CAT**G**AG**A**T**C****T**CT**A**G**A**GC**C**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****C**C
AGAGC**C**GT**C**GAC**A**GGT**A**T**A**GC**T**CT**A**CAT**G**AG**A**T**C****A**ACAT**G**AG**A**T**C****T**CT**A**G**A**GC**A**GT**G**AG**A**T**C**G**A**CAT**G**A**T**A**G****T**C



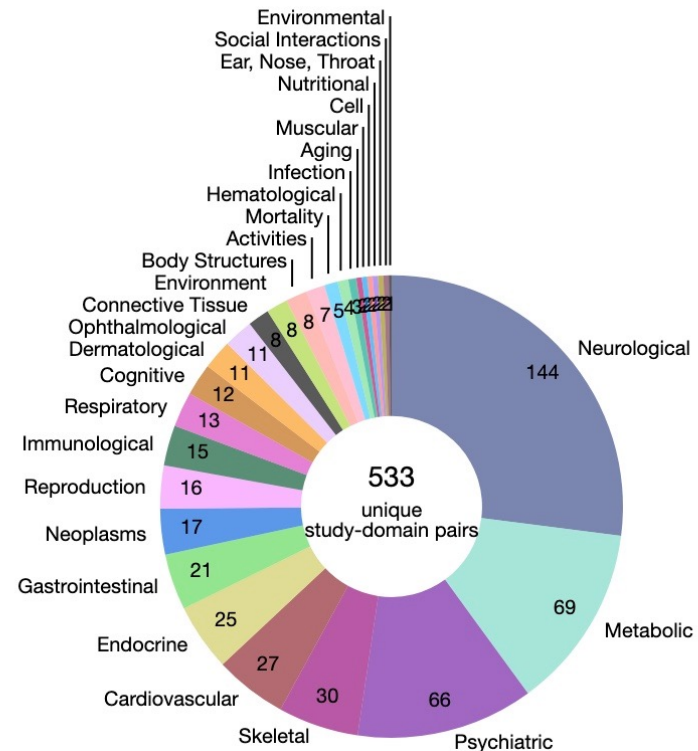
Associated SNP



“Manhattan plot” of GWAS results



GWAS: A large undertaking



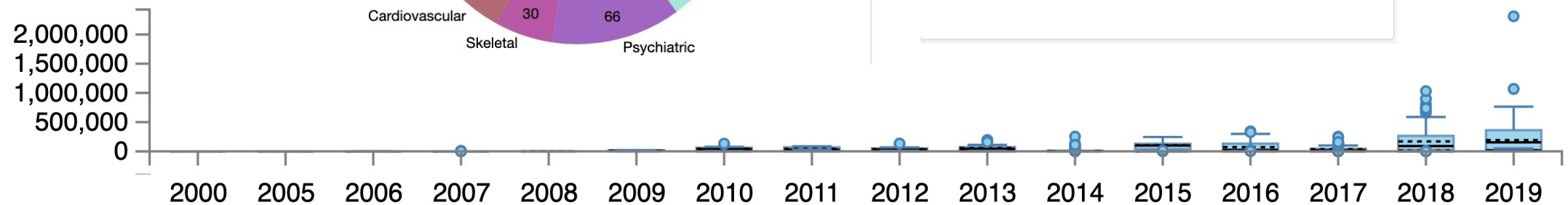
Database Summary

4,756
GWAS

3,302
Unique Traits

473
Unique Studies

28
Domains

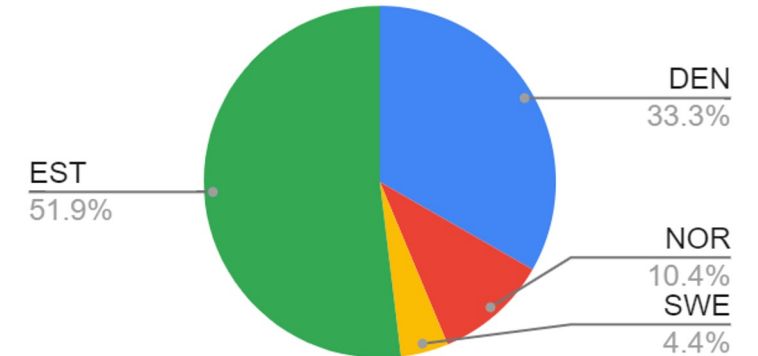


<https://atlas.ctglab.nl/>

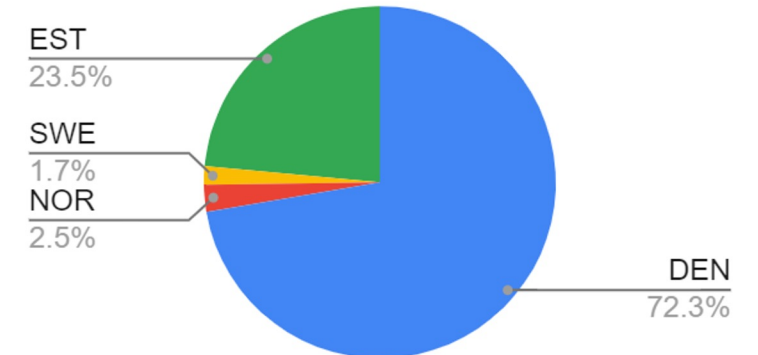
Sample size for cohorts used in an ongoing study on Major Depression Disease

Name	N MDD	N eoMDD	N control	Covariates
iPSYCH2012	20,804	18,429	23,854	Sex, birthyear, 25PCs
iPSYCH2015	10,487	8,105	15,772	Sex, birthyear, 25PCs
EstBB	48,804	8,768	127,395	Sex, birthyear, 10PCs
MoBa	8,824	908	~128,000	Sex, birthyear, 20PCs, batch
PREFECT	1,796	470	6,613	4PCs
UKB	76,828	21,499	418,765	Sex, 20PCs

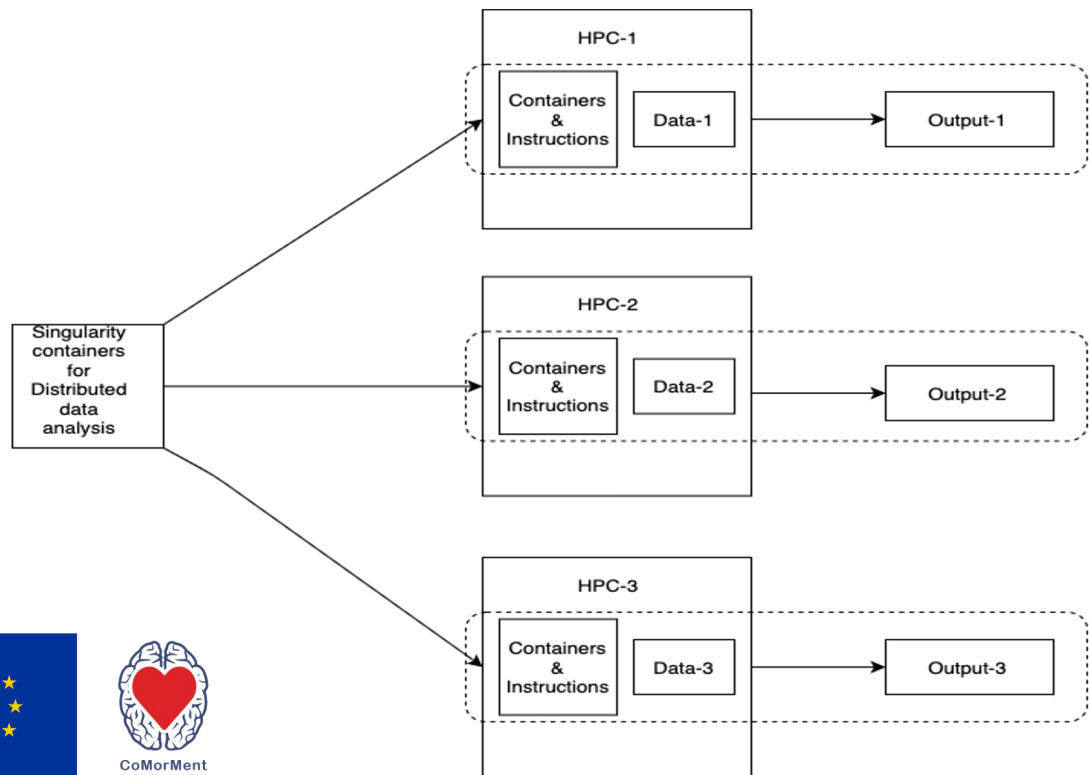
Major Depression disease



early onset Major Depression disease



Distributed & standardized procedure for uniform handling



System requirements have been defined and approved by sites

Available for download :
<https://github.com/comorment/containers/tree/main/singularity>

Inclusion of tools for GWAS and post-GWAS analysis and visualization and available sample datasets and reference data

instructions for use cases :
<https://github.com/comorment/containers/tree/main/usecases>

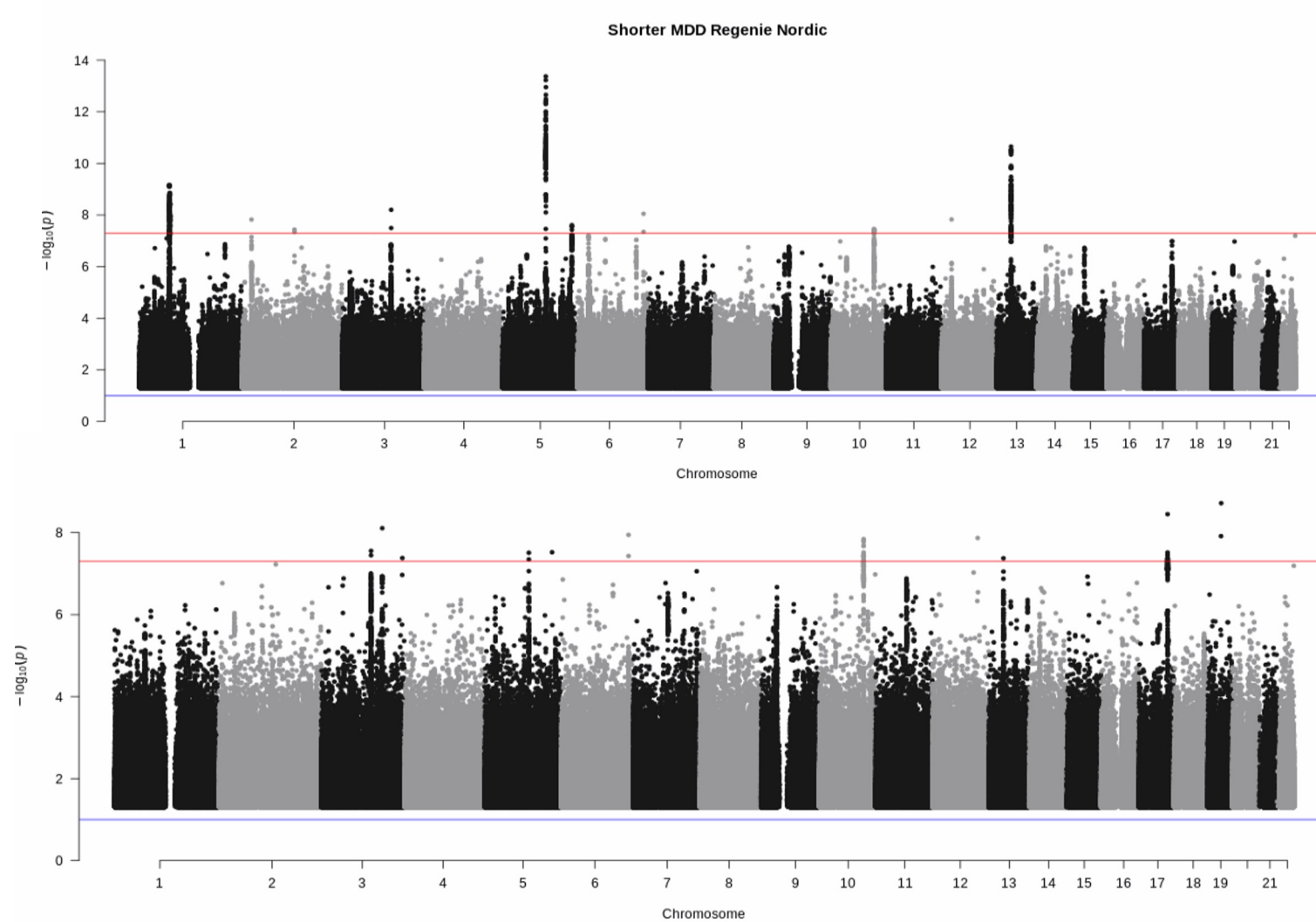


Container content

container	tool				
hello.sif	demo example	gwas.sif	king	python3.sif	python3
gwas.sif	plink	gwas.sif	metal	python3.sif	ldpred
gwas.sif	plink2	gwas.sif	vcftools	python3.sif	mixer
gwas.sif	plink2_avx2	gwas.sif	bcftools	python3.sif	python_convert
gwas.sif	PRSice_linux	gwas.sif	flashpca_x86-64	r.sif	R
gwas.sif	simu_linux	gwas.sif	regenie	r.sif	seqminer
gwas.sif	bolt	gwas.sif	GWAMA	r.sif	rareGWAMA
gwas.sif	gcta64	gwas.sif	magma	r.sif	GenomicSEM
gwas.sif	gctb	gwas.sif	shapeit2	r.sif	TwoSampleMR
		gwas.sif	impute4	r.sif	GSMR
		gwas.sif	minimac4	r.sif	LAVA
		gwas.sif	bgenix	r.sif	LAVA partitioning
		gwas.sif	cat-bgen	saige.sif	SAIGE
		gwas.sif	edit-bgen	enigma-cnv.sif	PennCNV



Nordic MDD and eoMDD show significant loci



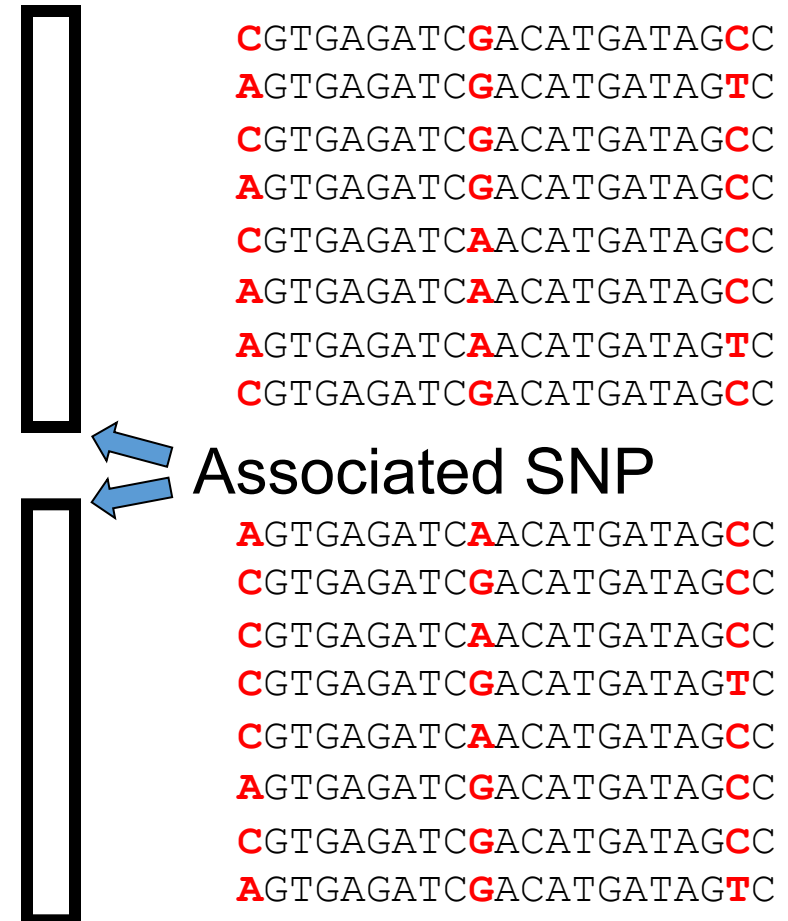
Find the associated SNP

Cases:

AGAGC**A**GTCGACAG
AGAGC**C**GTCGACAT
AGAGC**A**GTCGACAG
AGAGC**A**GTCGACAG
AGAGC**C**GTCGACAT
AGAGC**C**GTCGACAT
AGAGC**C**GTCGACAG
AGAGC**A**GTCGACAG

Controls:

AGAGC**A**GTCGACAT
AGAGC**A**GTCGACAT
AGAGC**A**GTCGACAT
AGAGC**C**GTCGACAG
AGAGC**C**GTCGACAG
AGAGC**A**GTCGACAG
AGAGC**C**GTCGACAG
AGAGC**C**GTCGACAG



Imputation...

Genotype data with missing data at untyped SNPs (grey question marks)

1	?	?	?	1	?	1	?	0	2	2	?	?	2	?	0
0	?	?	?	2	?	2	?	0	2	2	?	?	2	?	0
1	?	?	?	2	?	2	?	0	2	1	?	?	2	?	0
1	?	?	?	2	?	1	?	1	2	2	?	?	2	?	0
2	?	?	?	2	?	2	?	1	2	1	?	?	2	?	0
1	?	?	?	1	?	1	?	1	2	2	?	?	2	?	0
1	?	?	?	2	?	2	?	0	2	1	?	?	2	?	1
2	?	?	?	1	?	1	?	1	2	1	?	?	2	?	1
1	?	?	?	0	?	0	?	2	2	2	?	?	2	?	0

Reference set of haplotypes, for example, HapMap

0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	0
1	1	1	1	0	1	0	0	1	0	0	0	1	0	1	
0	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	0
0	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	1	1	0	1	0	0	1	0	0	0	1	0	1
1	1	1	0	0	1	0	0	1	1	1	0	1	1	1	0
0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	0	0	1	0	0	1	1	1	0	1	1	1	0

The reference haplotypes are used to impute alleles into the samples to create imputed genotypes (orange)

1	1	1	1	1	2	1	0	0	2	2	0	2	2	2	0
0	0	1	0	2	2	2	0	0	2	2	2	2	2	2	0
1	1	1	1	2	2	2	0	0	2	1	1	2	2	2	0
1	1	2	0	2	2	1	0	1	2	2	1	2	2	2	0
2	2	2	2	2	1	2	0	1	2	1	1	2	2	2	0
1	1	1	0	1	2	1	0	1	2	2	1	2	2	2	0
1	1	2	1	2	1	2	0	0	2	1	1	1	2	1	1
2	2	2	1	1	1	1	0	1	2	1	0	1	2	1	1
1	2	2	0	0	2	0	0	2	2	2	1	2	2	2	0

Each sample is phased and the haplotypes are modelled as a mosaic of those in the haplotype reference panel

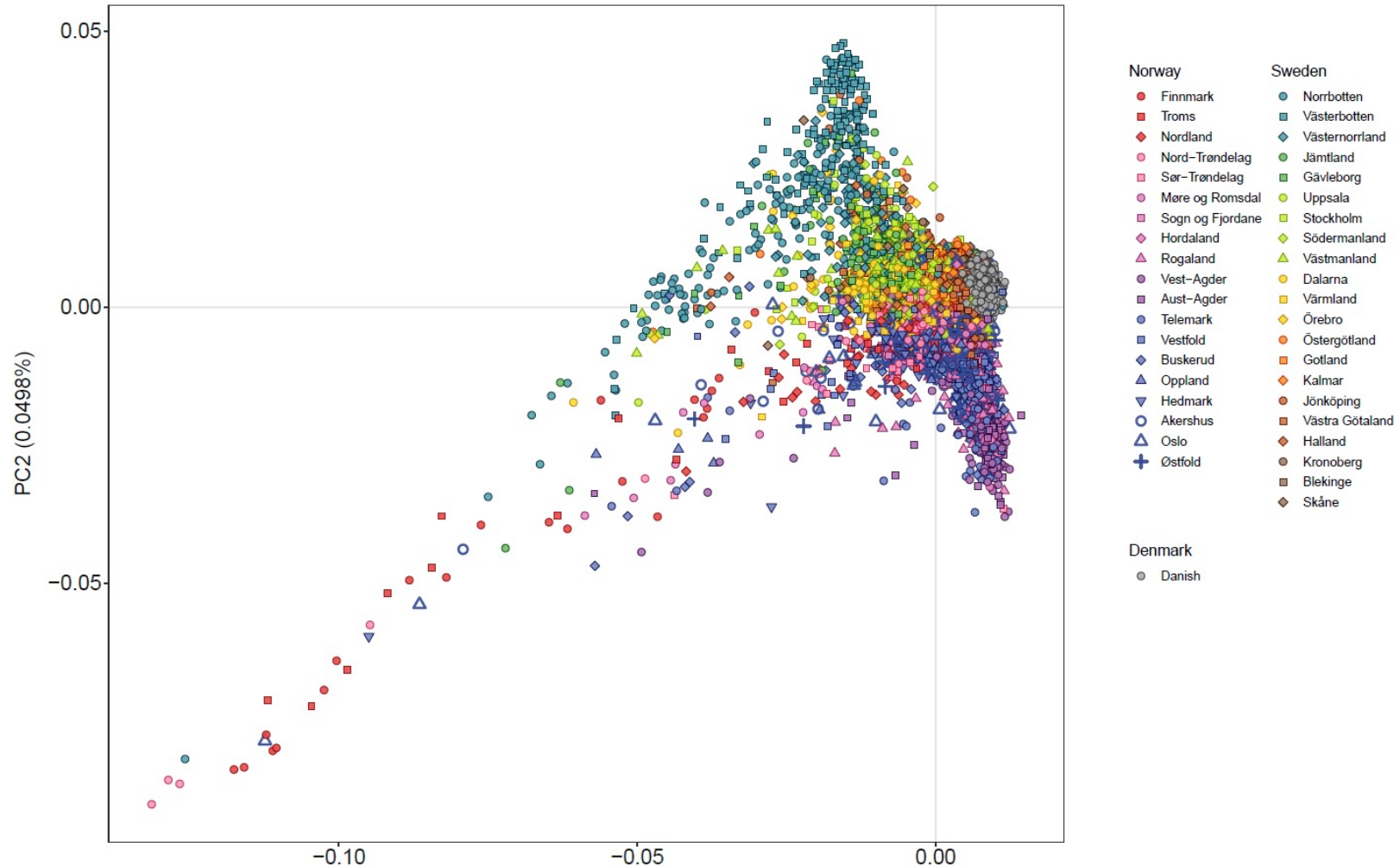
0	?	?	?	1	?	1	?	0	1	1	?	?	1	?	0
1	?	?	?	1	?	1	?	0	1	0	?	?	1	?	0
1	?	?	?	1	?	1	?	1	1	1	?	?	1	?	0
1	?	?	?	0	?	0	?	1	1	1	?	?	1	?	0
0	?	?	?	0	?	0	?	1	1	1	?	?	1	?	0



Genetic structure in the Nordics

2985 Norwegians
3519 Swedes
1606 Danes

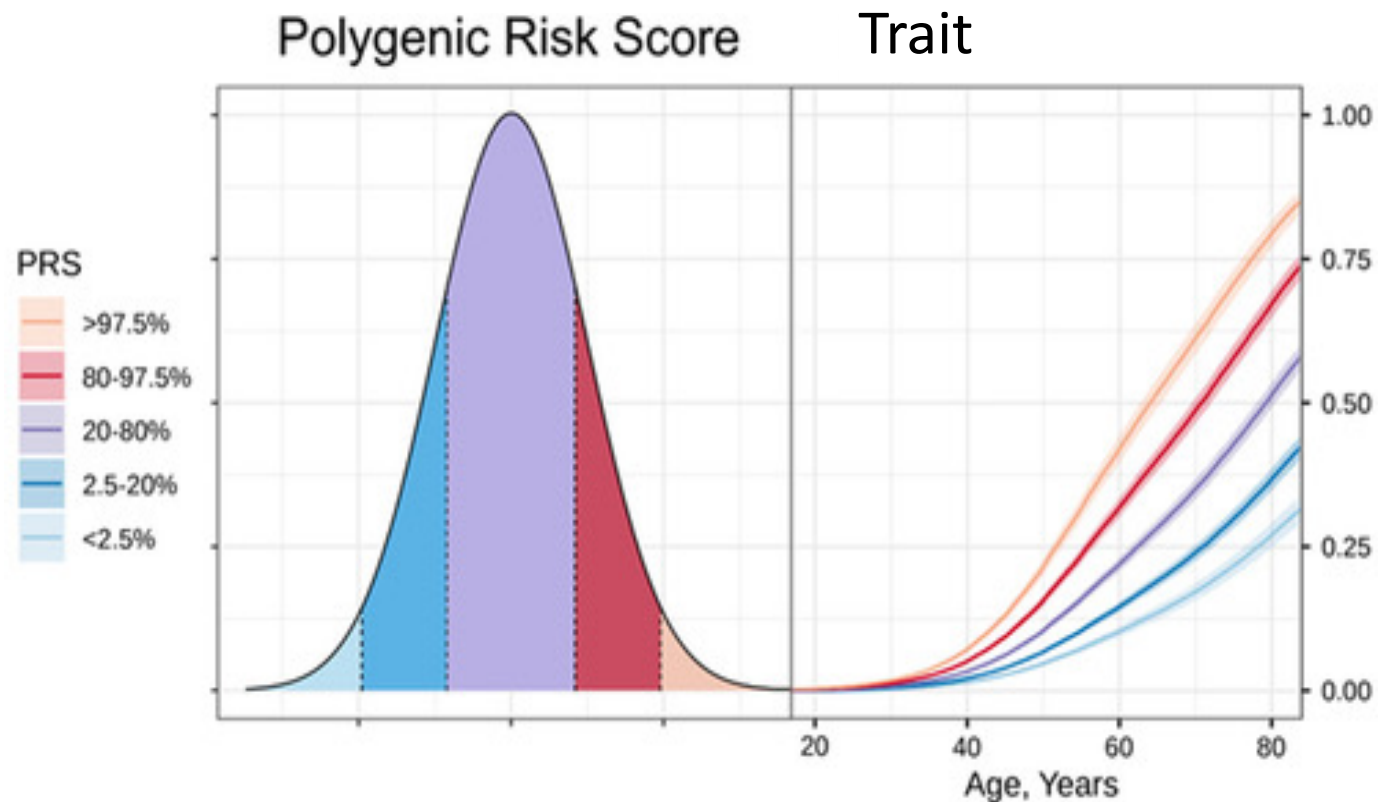
PCA plot



Example imputation

chr1	Estonian	HRC	NORGENE
N imputed SNPs	2,203,149	3,066,134	1,768,482
N imputed SNPs (INFO > 0.8)	863,824	1,498,665	1,084,341
N imputed SNPs (MAF 1%)	672,704	610,806	854,52
Median INFO	0.56	0.73	0.89
N imputed SNPs (INFO > 0.8 & MAF > 1%)	601,772	598,272	768,197
MEDIAN INFO (INFO > 0.8 & MAF > 1%)	0.9900	0.9982	0.9960
Overlap (position and change, no imp qc)	Estonian	NORGENE	874,791
	HRC	NORGENE	1,111,409
	HRC	Estonian	1,936,512

Polygenic risk score



Polygenic risk score:

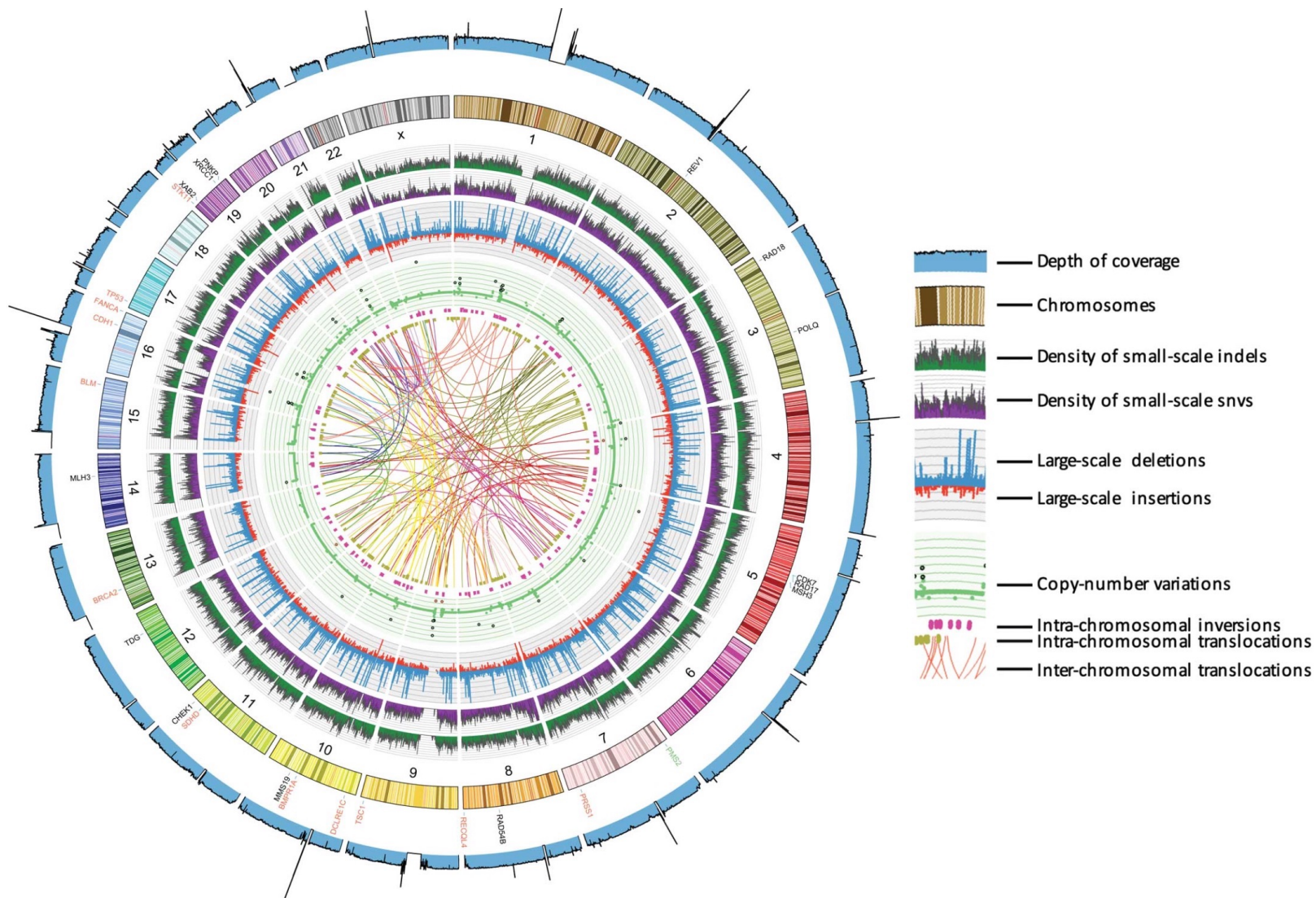
A set of SNPs that collectively may predict a risk, based on GWAS studies

e.g.: Does an individual have a high risk that warrants monitoring or treatment

May depend on the population

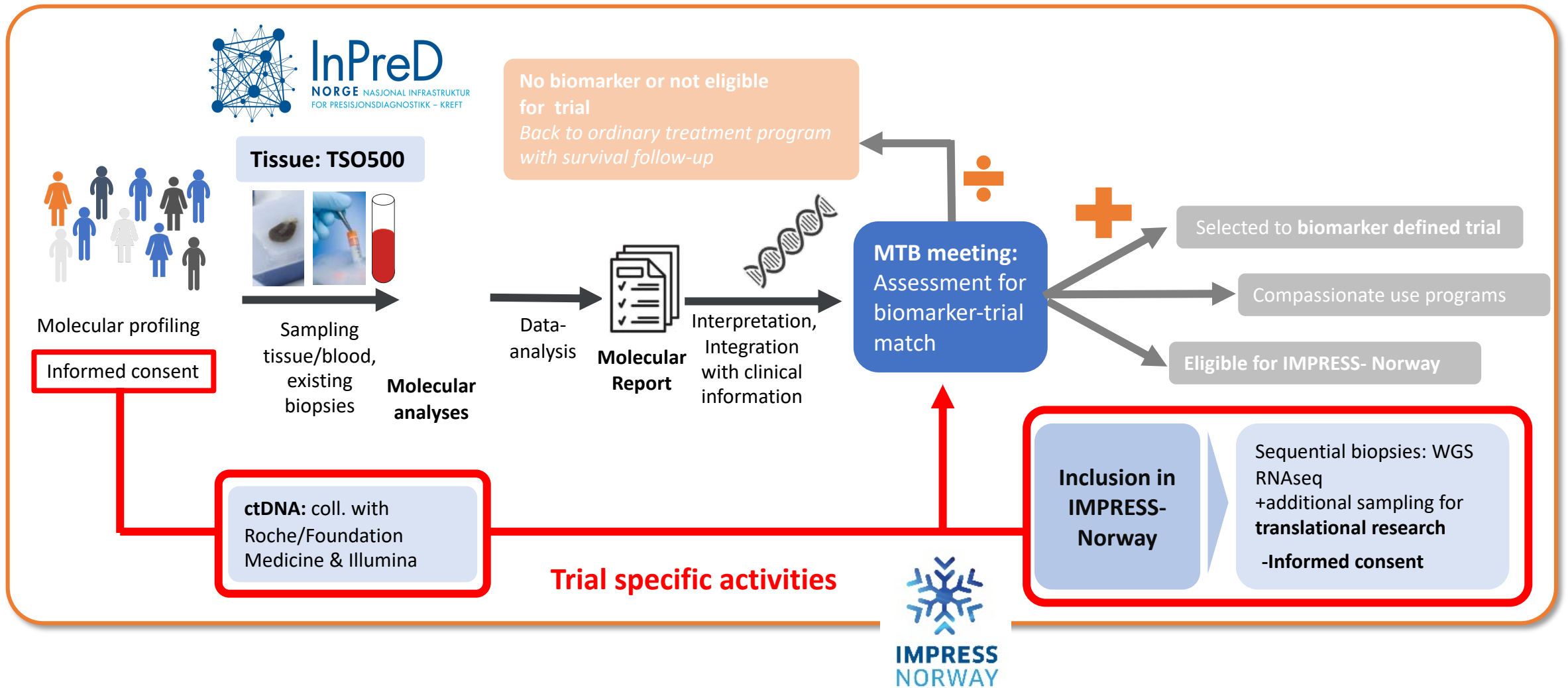
Cancer risks, diabetes, etc

Cancer: Many layers of perturbation



Genetics
Epigenetics

The interaction between InPreD and IMPRESS – Norway



SUMMARY OF KEY FINDINGS

SNVs/ indels	Total nr of SNVs/indels in protein coding sequence: alter protein coding sequence	15 12 9	Copy number variants	None > 6 copies Loss of CDKN2A/2B	Gene fusions RNA	CCDC6-RET
	TMB	Mut/ Mb				

Biomarkers and variant/gene-drug association

Biomarkers potentially relevant for immune therapy	None reported
---	---------------

Gene	Variant	Type	VAF/ CN	Variant GoF/ LoF	Pathway/ function	Therapeutic context		Level of Evidence	
						Sensitive	Resistant	Patient's tumor type	Other tumor type
RET	CCDC6 -RET	fusion		GoF	MAPK Pi3K/AKT	TKI		1	
PIK3CA (F1 Liquid)	E545K E542K	missense	1,2% 1.2%	GoF	Pi3K/AKT	PI3Ki		3	1

Variant classification (RET)
The fusion CCDC6 (exon 2) and RET (exon 12) was assessed to be in frame and kinase domain intact. The fusion is a known oncogenic variant with constitutive kinase activation. RET fusion is a biomarker for RET targeted therapy in subtypes of thyroid and lung cancers. Clinical data is emerging from trials with RET-selective inhibitors*.

Variant classification (PIK3CA)
The PIK3CA hotspot variants lead to constitutively activated PI3K/AKT pathway. The low variant allele frequency (F1 Liquid) suggest subclones, possibly as a result of treatment resistance. Targeting PIK3CA is app. for a subgroup of breast cancer but the variants have no clear treatment implication in thyroid cancer.

Additional results: Variants of unknown significance in treatment relevant genes

Gene/Variant	Comments
RAD51C R370Q	RAD51C is involved in HR DNA repair. The variant is in the C terminal of the gene but of unknown biological significance.

Variants that alter protein coding sequence (N=10)

Gene Symbol	Protein change	Coding status	VAF tumor
DICER1	E1705Q	missense	0.024
IRF4	G202D	missense	0.323
SOX17	D131N	missense	0.358
MST1R	X1401=	stop_retained	0.062
PPP6C	E237K	missense	0.357
C11ORF30	R840H	missense	0.485
IGF1R	S752R	missense	0.473
NSD1	Q784E	missense	0.437
RAD51C	R370Q	missense	0.336
TERT	c.-124 C>T	Non-coding	0.471

According to ESMO guidelines, recommendation for genetic counseling does not apply for the variants reported by TSO500 for this sample.

 Mutation hotspot

*; [Registrational results of LOXO-292 in patients with RET-altered thyroid cancers - ScienceDirect](#); and PMID: 32846061.

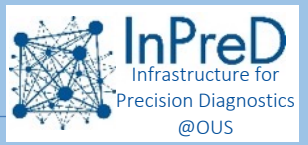
Patient
Radium F/XXy

Tumor
type
Papillary thyroid
cancer
Y of D: 20xx

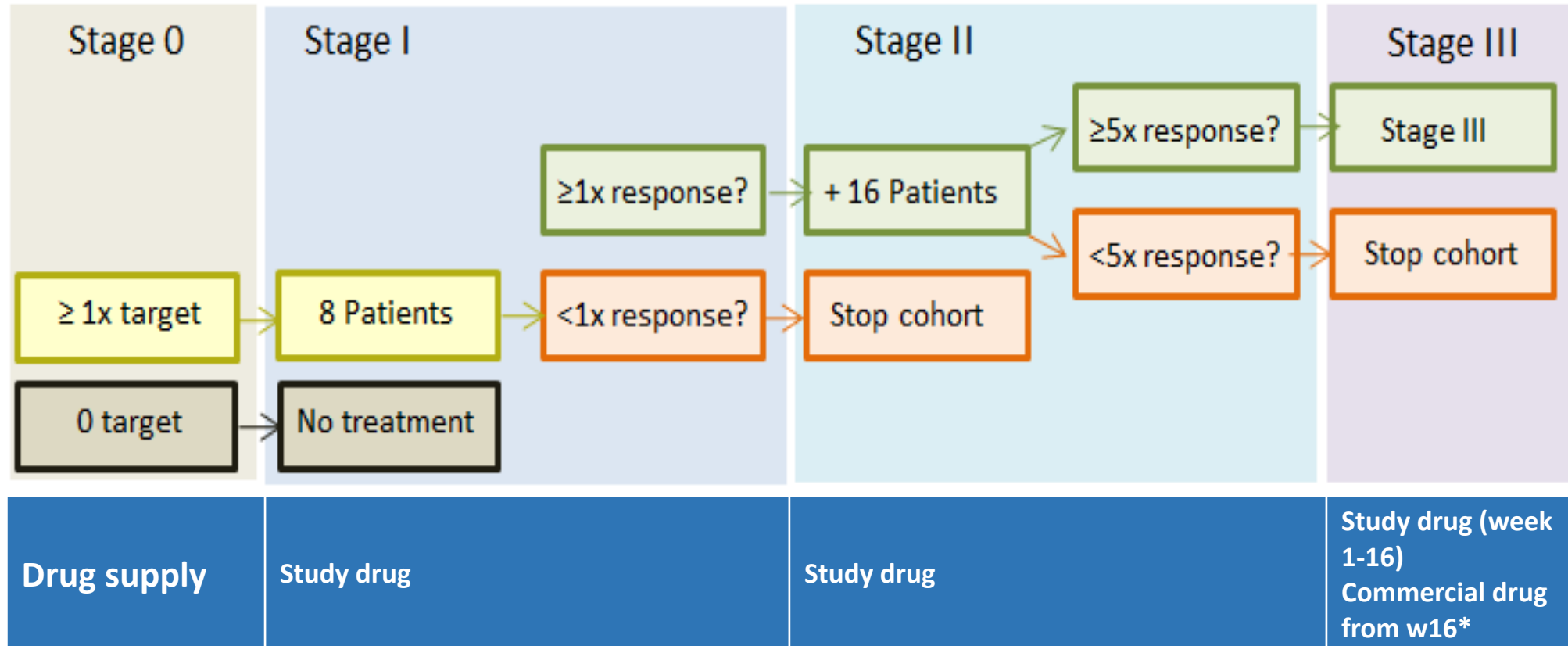
Bio-
material
Distal met
post-treatment
FFPE
Xx-xxxx

Tumor
content
40%

F1
Liquid
CDx
Gene Alterations
Overlap TSO500: YES
Add. biomarkers: YES
PIK3CA E545K, E542K



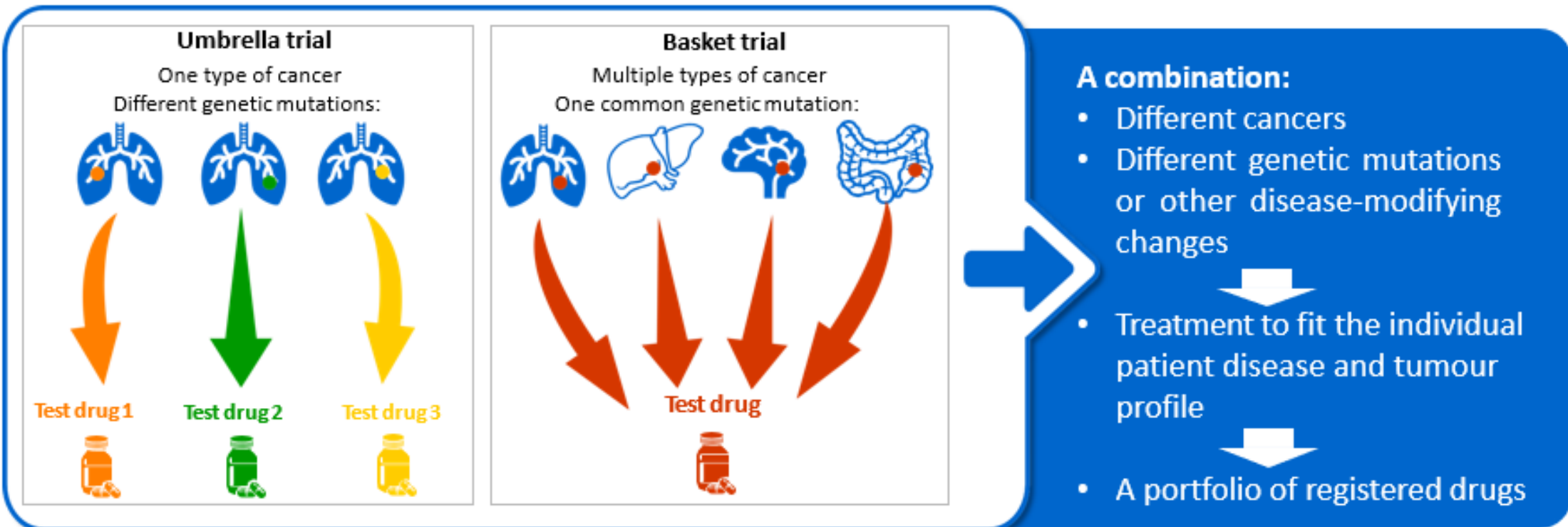
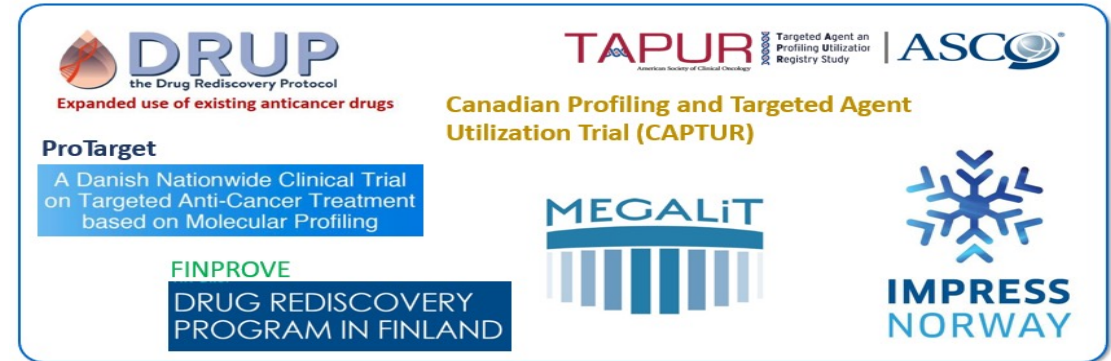
Novel reimbursement models for PCM - from a trial setting to implementation



*The design of stage III is a phase 1 (week 1 to 16) where patients will be treated with the investigational products (IP), provided by the manufacturer. If clinical benefit (defined as confirmed objective response or stable disease at 16 weeks and measured more than once, at least 28 days apart) is observed for an individual patient at 16 weeks, treatment for that individual patient with the IP will continue in phase 2

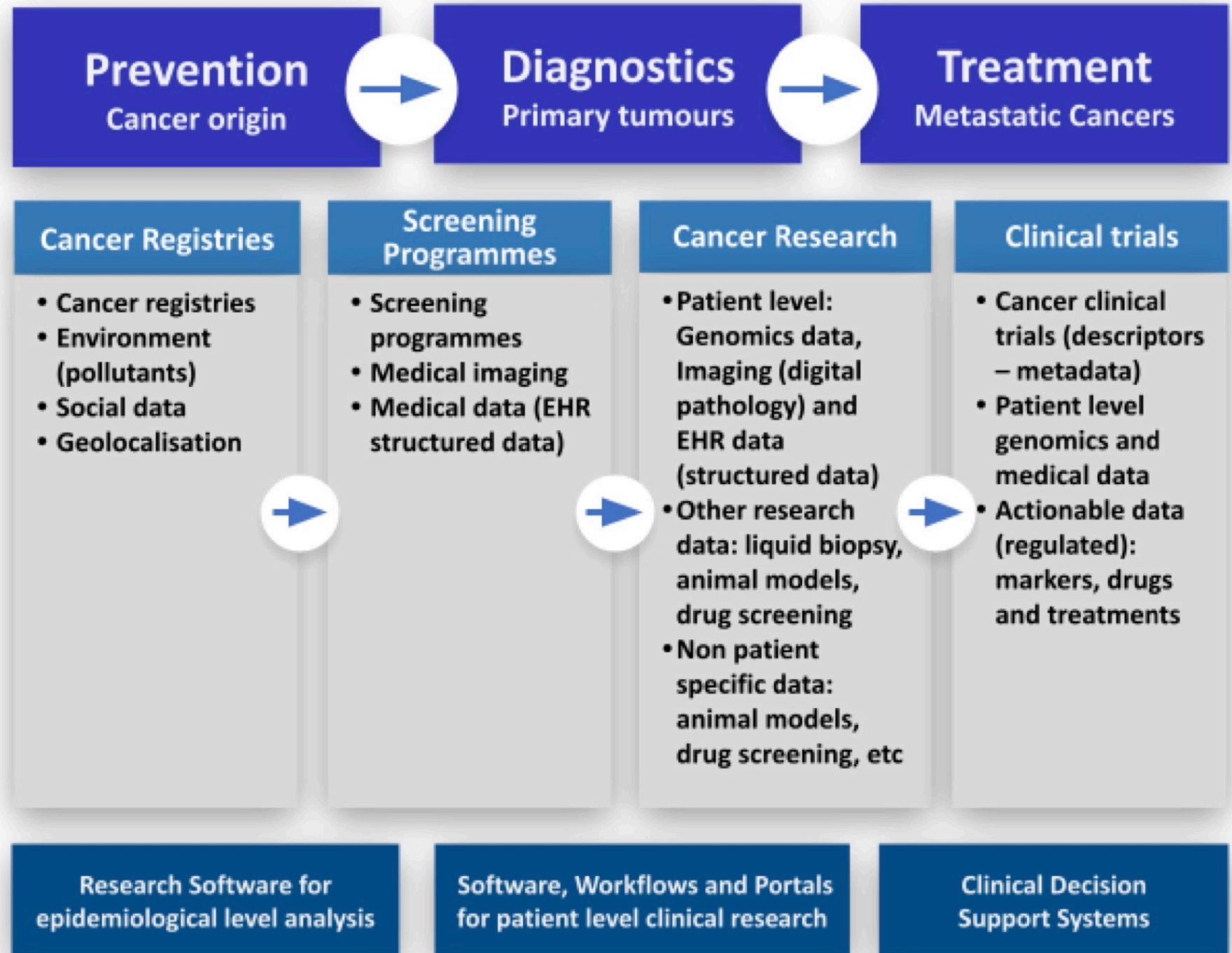
Shared study-design across Nordic – DRUP collaboration based on tumor-agnostic approach

- Individual treatment approach
- Targeted treatment



EOSC4Cancer

Starting EU project



<http://ega.elixir.no>

Federated EGA Norway node

 LS LOGIN

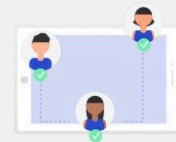
The Federated EGA (European Genome-phenome Archive) is a distributed solution for sharing and exchange of human -omics data across national borders. Federated instance collects metadata of -omics data collections stored in national or regional archives and make them available for search through the main EGA portal.



Searchable



Secure



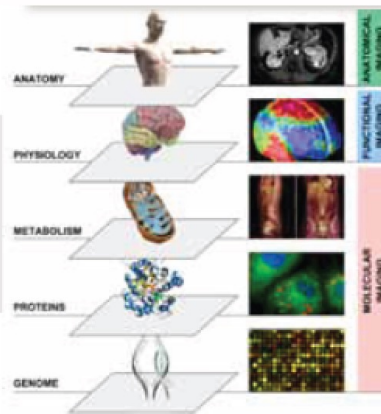
Shareable

Big Data

Individual



Biomarkers
Treatment
Reponse eval.
Side effects



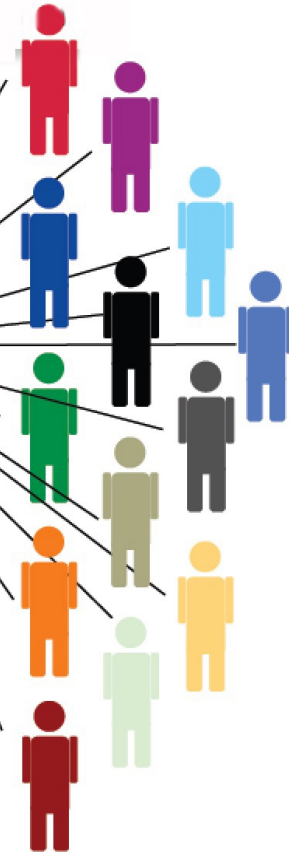
Database

Biomarkers
Treatment
Reponse
Side effects
Survival
Historical controls



Statistics

Frequencies
Machine learning



BACKGROUND DOCUMENT

Proposal for a regulation - The European Health Data Space

COM(2022) 197/2

PAGE CONTENTS

[Details](#)

[Description](#)

[Files](#)

Details

Publication date

3 May 2022

Author

Directorate-General for Health and Food
Safety

Acknowledgments:

- CoMorMent
 - Bayram Akdeniz
 - Oleksandr Frei
- IMPRESS
 - Åslaug Helland
 - Hege Russnes
 - Leonardo Meza-Zepeda
- ELIXIR Norway
 - Kjell Petersen
- USIT
 - Gard Thomassen