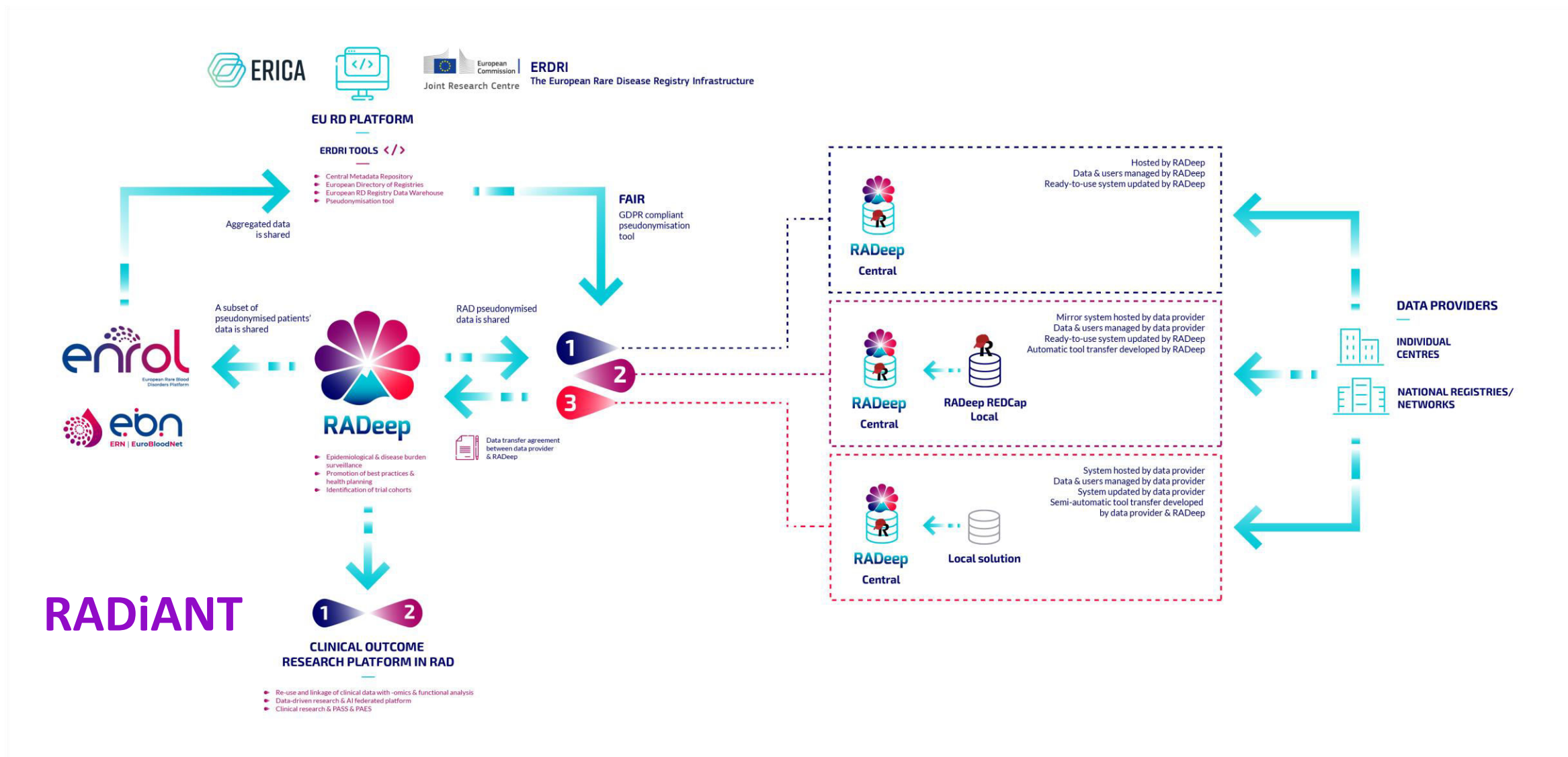# AI Federated platform in anemia
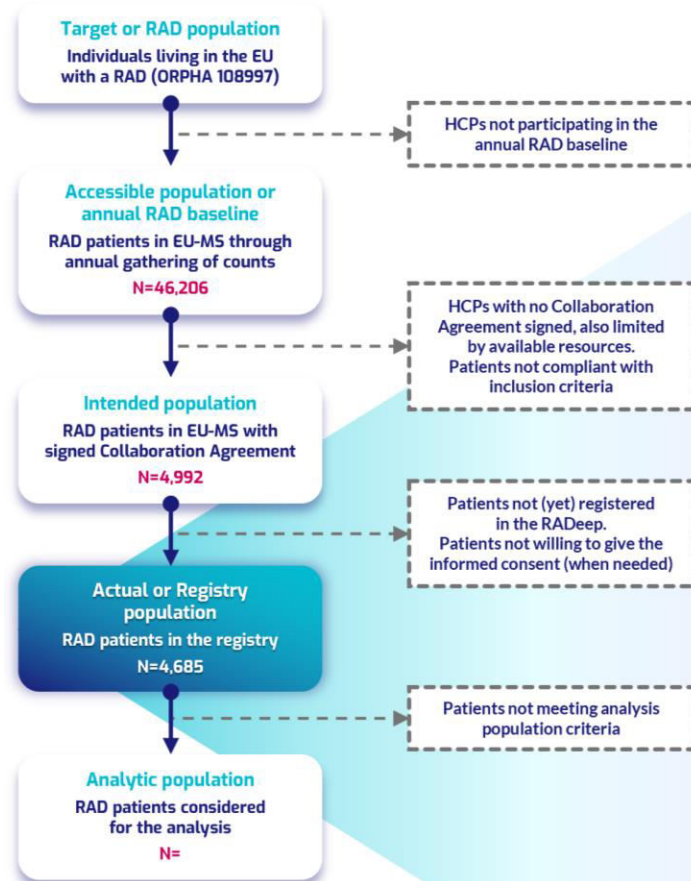
RADeep

RADeep: Clinical data standardization in RADs

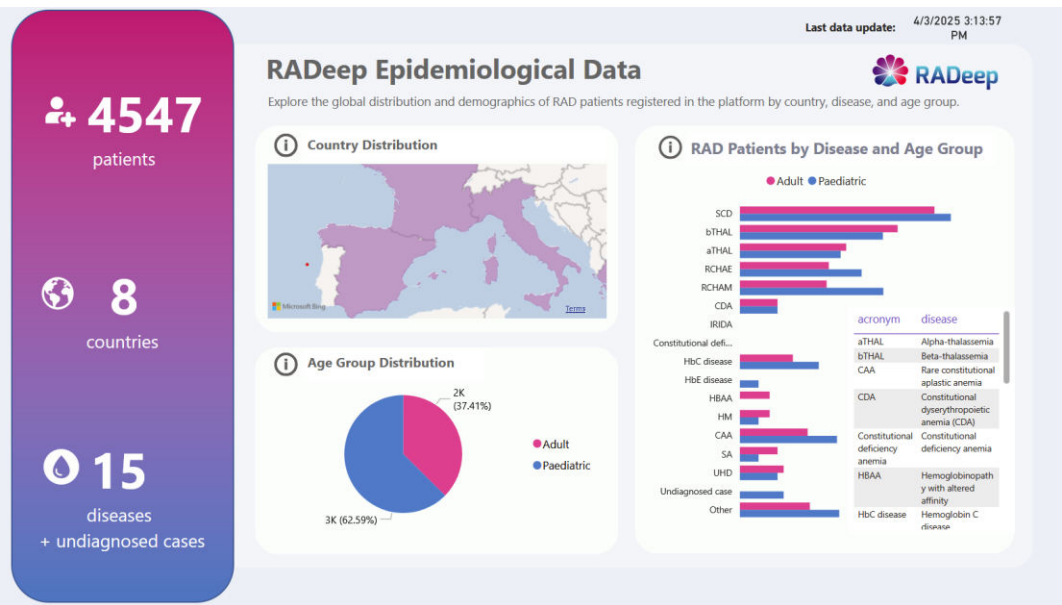# RADeep study populations and CRF summary

Target or RAD population
Individuals living in the EU with a RAD (ORPHA 108997)

HCPs not participating in the annual RAD baseline

Accessible population or annual RAD baseline
RAD patients in EU-MS through annual gathering of counts
N=46,206

HCPs with no Collaboration Agreement signed, also limited by available resources. Patients not compliant with inclusion criteria

Intended population
RAD patients in EU-MS with signed Collaboration Agreement
N=4,992

Patients not (yet) registered in the RADeep. Patients not willing to give the informed consent (when needed)

Actual or Registry population
RAD patients in the registry
N=4,685

Patients not meeting analysis population criteria

Analytic population
RAD patients considered for the analysis
N=

Group		Subgroup		Parameters	Mandatory	Longitudinal
1	Patient pseudonym, permissions and biobanking information	1.1.	Pseudonym	1	1	0
		1.2.	Permissions	5	5	0
		1.3.	Biobanking	3	3	0
2	Demographics	2.1.	Population distribution	12	8	4
		2.2.	Mortality and comorbidities	7	7	6
3	Diagnosis	3.1.	Diagnosis	6	6	0
		3.2.	Genotype	5	5	0
		3.3.	Undiagnosis	2	2	0
		3.4.	Disease Onset	4	2	0
		3.5.	Neonatal Manifestations	5	0	0
4	Physical examination	4.1.	Development	0	0	0
5	Organ Damage	5.1.	Organ damage assessment	14	3	14
		5.2.	Chronic complications of bones and extremities	8	4	8
		5.3.	Chronic cardiac and pulmonary disease	6	5	6
		5.4.	Chronic neurological disease	7	6	7
		5.5.	Chronic endocrinologic disease	8	7	8
		5.6	Chronic liver and renal disease	8	7	8
		5.7.	Visual and hearing disease	5	4	5
6	Acute Complications	6.1.	Acute complications in RADs (Except SCD) requiring hospitalization or emergency admission for more than 24 hours	4	3	4
		6.2.	Acute complications in SCD requiring hospitalization or emergency admission for more than 24 hours	4	3	4
		6.3.	Intesive Care Unit Admission in the last 12 months	1	1	1
7	Clinical manifestations and surgery	7.1.	Spleen	4	4	4
		7.2.	Gallbladder	3	0	3
8	Treatments	8.1.	Blood transfusion	14	7	13
		8.2.	Chelation	5	5	5
		8.3.	Hydroxyurea	4	1	4
		8.4.	Specific treatment(s)	2	2	2
		8.5.	Haematopoietic stem cell transplantation (HSCT) / gene therapy	5	5	5
		8.6.	Inclusion in clinical trial protocol	2	2	2
9	Fertility	9.1.	Fertility and Offspring	2	2	2
10	Disability	10.1.	Disability	3	1	3
11	Laboratory tests	11.1	Complete blood count	17	17	17
		11.2	Biochemical tests	14	14	14
		11.3	Hemoglobin tests	7	4	1
		11.4	Enzyme tests	17	17	0
		11.5	Membrane tests	3	3	0
		11.6	Other Laboratory tests	4	4	0
		Total		221	170	150

Gliklich RE, Dreyer NA, Leavy MB, editors. Registries for Evaluating Patient Outcomes: A User's Guide [Internet]. 3rd ed. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014 Apr. Report No.: 13(14)-EHC111. PMID: 24945055.

Public dashboard: https://www.radeepnetwork.eu/epidemiological-data/

- **Real time data quality rules**
  - Alerts when data is being entered

- **Asynchronous data quality rules:**
  - Data quality module accesible for data entrists

Private dashboards: https://radeep.vhir.org/

VHIR Shiny Applications

Interactive data visualization tools

RADeep

RADeep Descriptive Tables
RADeep Descriptive tables with granular information per country, disease, sub-disease, age-group and gender.

Launch Application

Last updated: March 2025

RADeep Data Completeness Report
RADeep Completeness reports by variable with filters for: mandatory and optional parameters, active patients, country, hcpid and diagnosis.

Launch Application

Last updated: March 2025

RADeep Data Coverage Report
RADeep coverage reports by data parameter with filters for: country, RADs and mapping status

Launch Application

Last updated:May 2025

Quality management
Data Quality flow

- Central Solution

1. **RADeep Data Monitor** runs the data quality rules.

2. **RADeep Data Monitor** open and assigns the query.

3. **Data Entries** resolves the query assigned to them.

4. **RADeep Data Monitor** closes the query.

5. **RADeep Data Monitor** generates periodically a report for each HCP with open queries.

- Local Solution

1. **Data Entries** generates their own internal data quality flow.

2. **Data Entries** resolves the query before sending the data for batch transfer.

3. **RADeep Data Monitor** runs the data quality rules.

4. **RADeep Data Monitor** generates a report for each HCP.

The query in RADeep REDCap central is closed when a new data batch transfer is performed.

RADeep ⟷ **RADiANT**

14 Ongoing collaboration agreements involving 202 HCPs in 13 EU countries:
12 Member States

- Belgium
- Cyprus
- Denmark
- France (2)
- Germany
- Greece
- Ireland
- Italy
- Portugal
- Spain
- Sweden
- The Netherlands
- + Norway

Countries onboard in the research platform in RADs: 4 Members States

- France (2/2)
- Italy (1/1)
- Spain (1/10)
- The Netherlands (1/4)

- Belgium (2/2)
- Portugal (3/3)
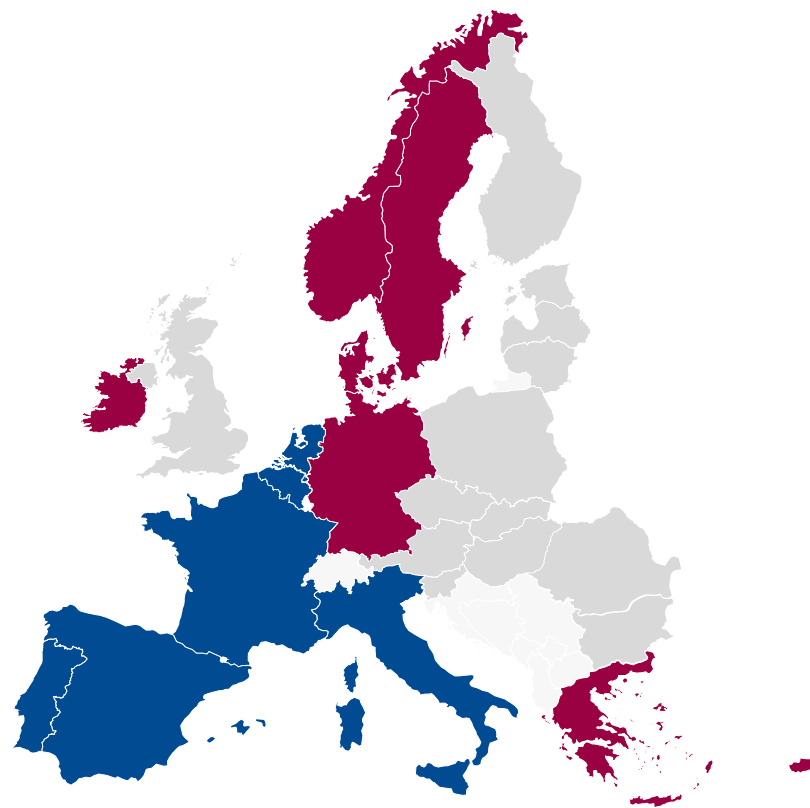
AI aims in SCD

Development of AI Solutions to Improve SCD Clinical Management through a Personalized Precision Medicine Approach:

❑ Omics-based Classification and Prognosis of SCD and Omics-based Clinical Decision Making in SCD

- Aim 1/2: Cluster SCD patients according to their genomic /metabolomic profile
- Aim 3: To develop a probability score using AI-based brain MRI image analysis (radiomics) to detect existing lesions for early diagnosis of cerebral silent infarcts
- Aim 4: Based on results from objectives 1 to 3, to develop predictive risk scores for recurrent vaso-occlusive crisis, acute chest syndrome, cerebral silent infarct, kidney disease

❑ Synthetic data generation in Kidney disease
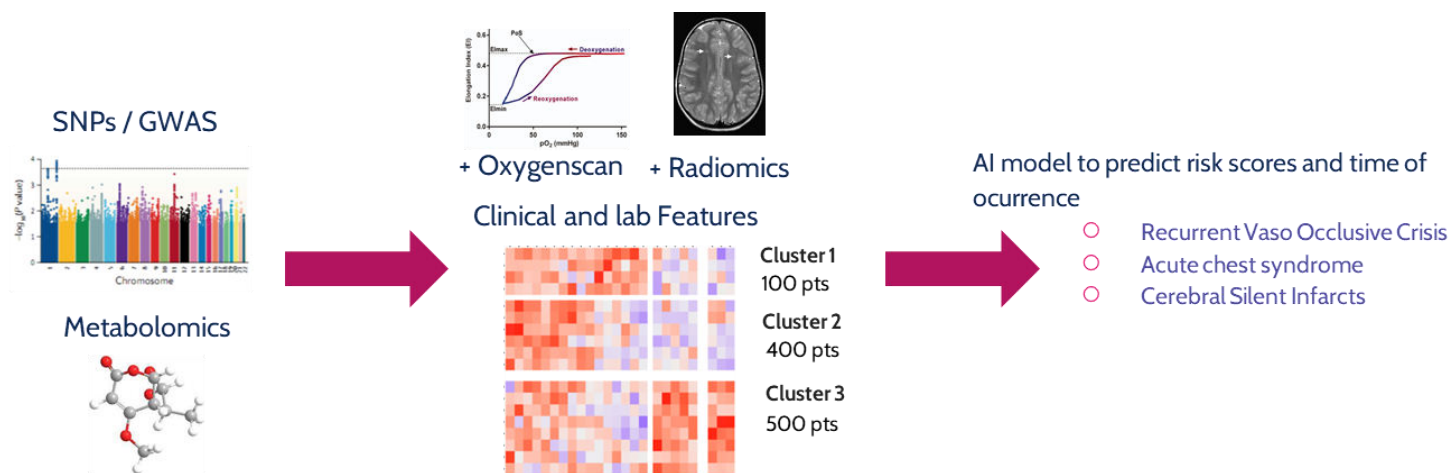
SCD Use Case

Target 1,000 SCD Patients

Genetic modifiers 1000+ GWAS

Metabolomics 1900+ metabolites
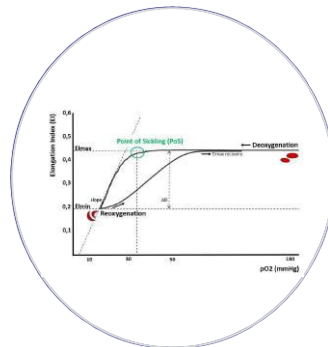
Functional test Rheology LoRRca - Oxygenscan
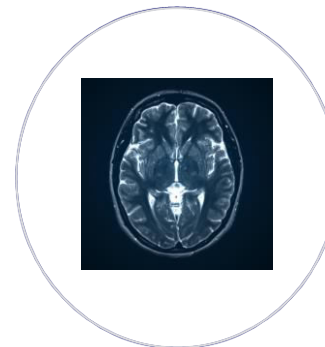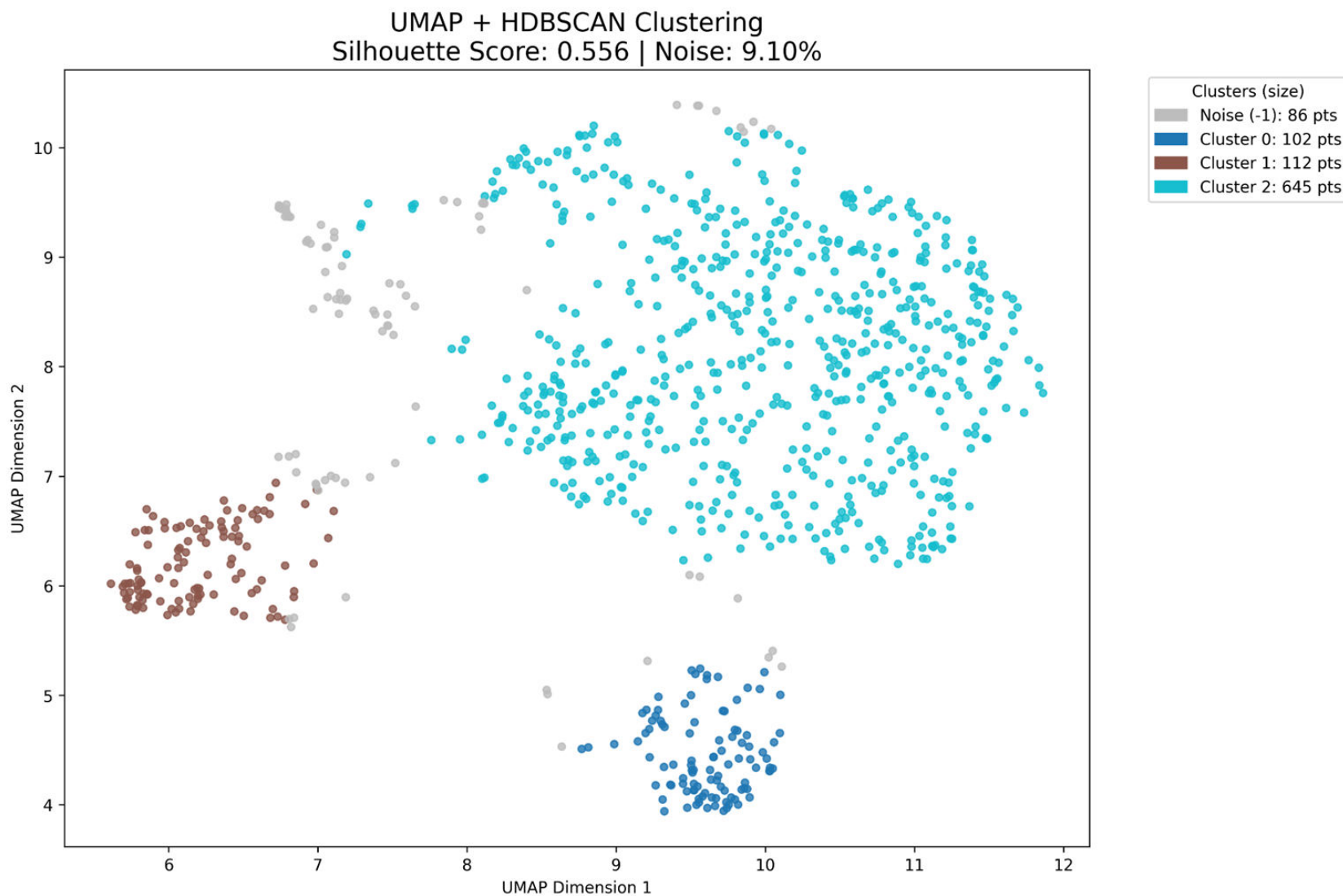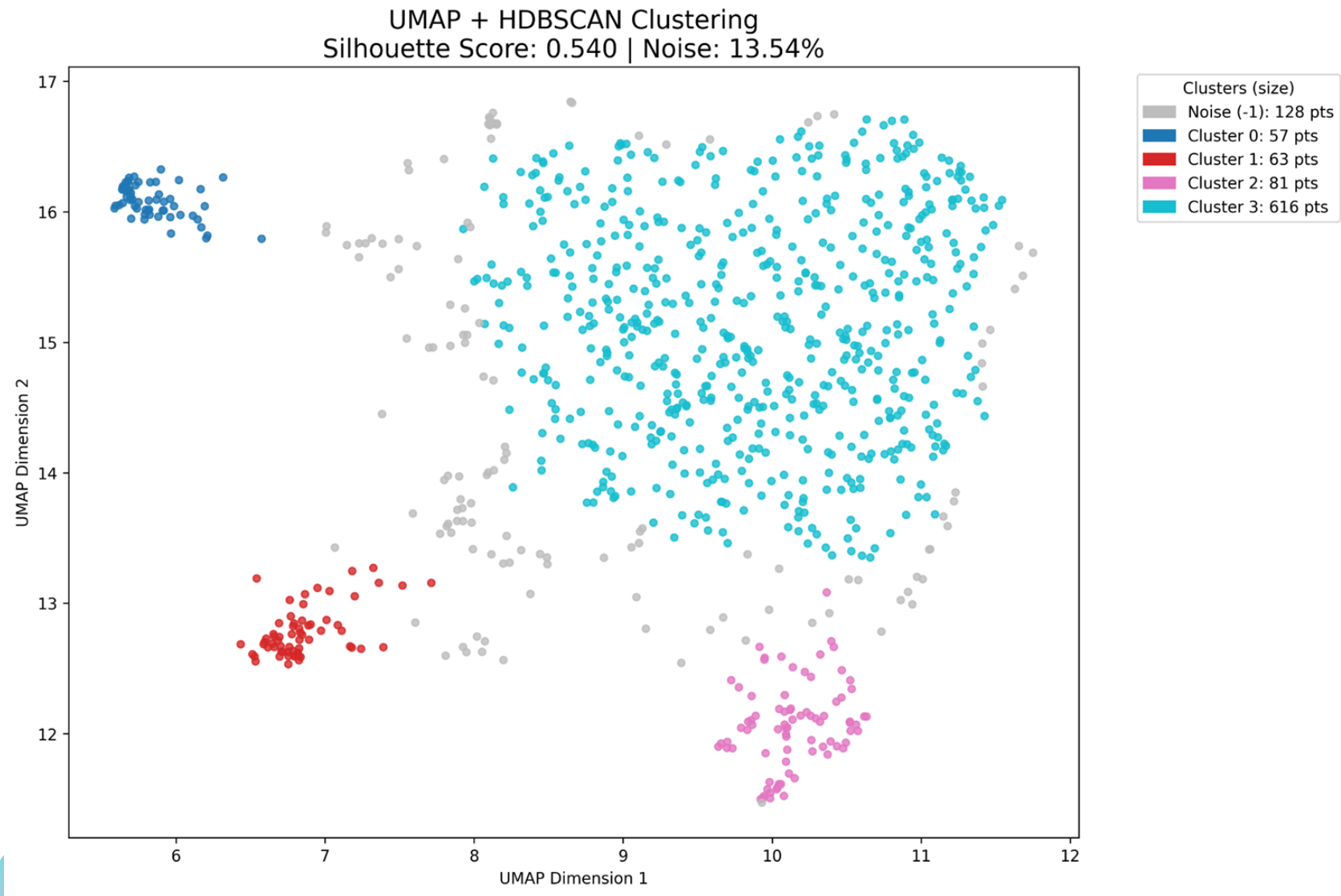
MRIs

Clinical and lab data

RADeep

Spain (INTEGRA 10 sites)
334 (76% PED)

France (2 site)
494 (37.6% PED)

Netherlands (SCORE 4 sites)
487 (45% PED)

Italy (1 site)
130 (75% PED)
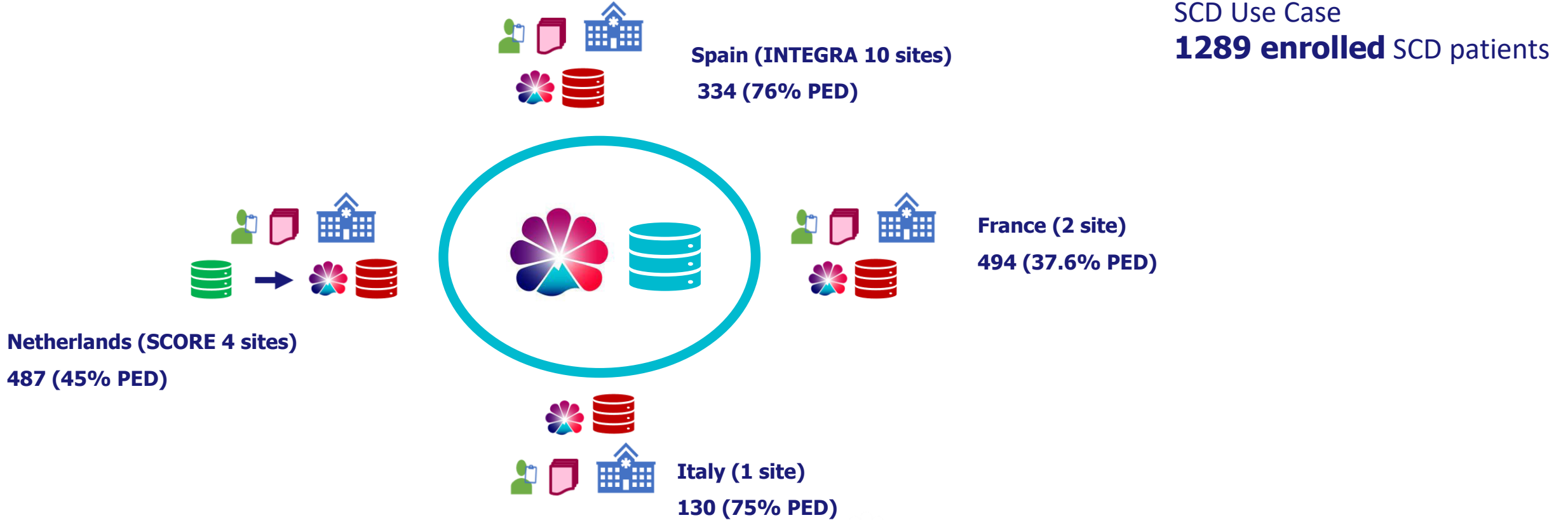
AI Methods

- Clinical dataset: 1274 patients
- GWAS dataset: 991 patients

$\longrightarrow$

- Overlap: 959 patients

- Clinical dataset: 1274 patients
- Metabolomics dataset: 979 patients

$\longrightarrow$

- Overlap: 924 patients

GWAS Clustering (945 samples): Binarized data

GWAS Clustering (945 samples): {0,0.5,1} mapping

UMAP + HDBSCAN Clustering
Silhouette Score: 0.540 | Noise: 13.54%

Clusters (size)
- Noise (-1): 128 pts
- Cluster 0: 57 pts
- Cluster 1: 63 pts
- Cluster 2: 81 pts
- Cluster 3: 616 pts

GWAS Clustering (945 samples): One-Hot Encoding
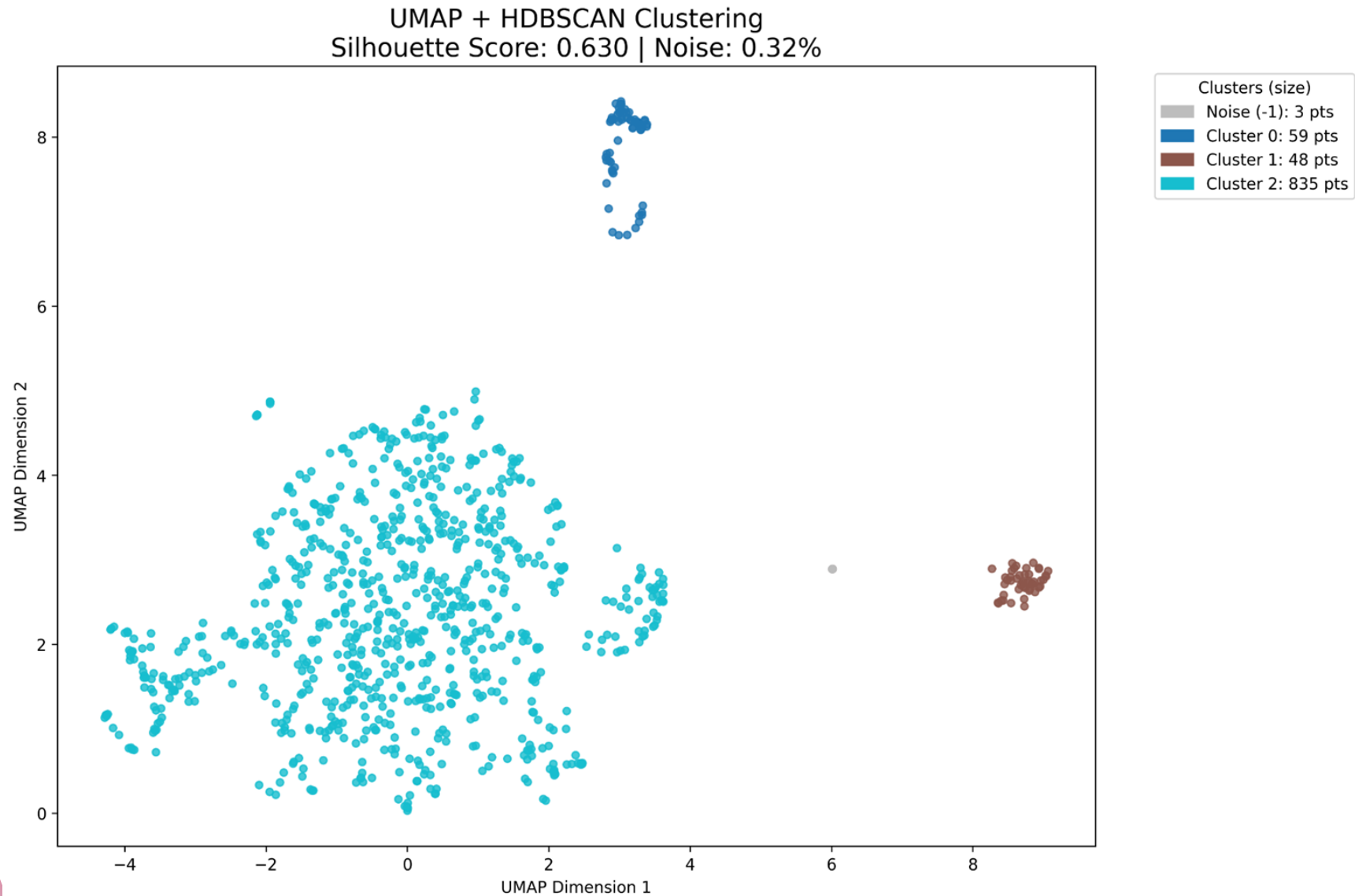
GWAS Clustering (945 samples): {0,0.5,1} mapping

HBA:
5, αα/ααα |
4, αα/αα |
3, αα/α- |
2, α-/α- |
1, α-/-- |
8, Inconclusive
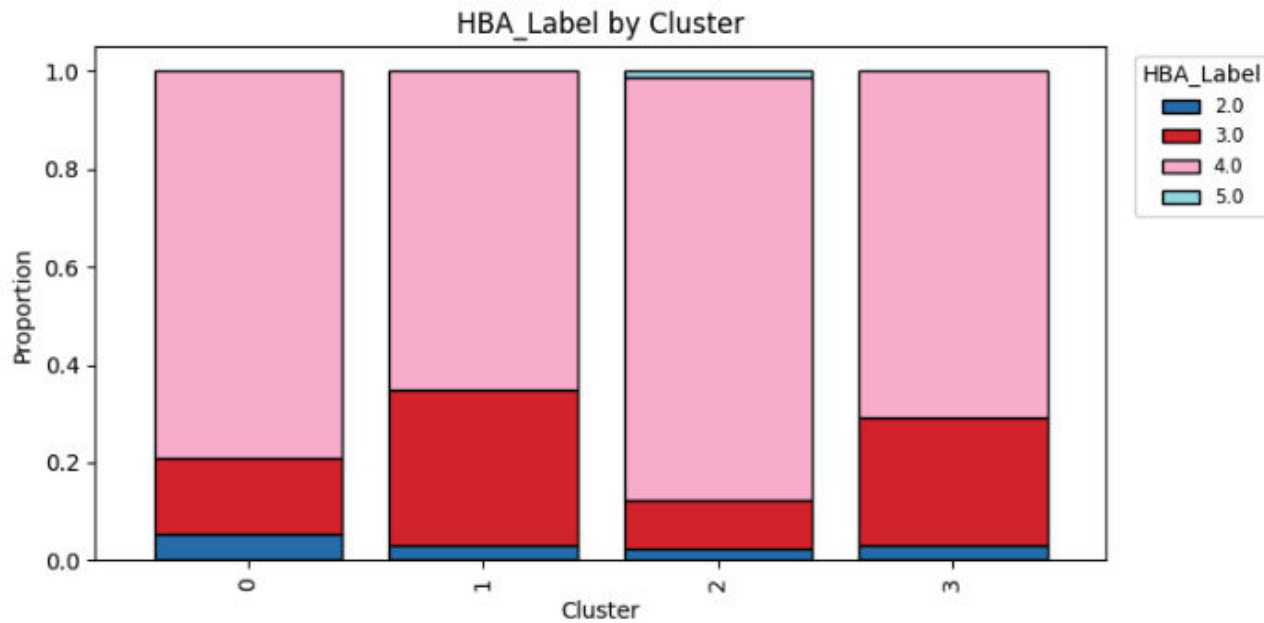
Signif. pvalues:
(2, 3) **
(2, 1) *
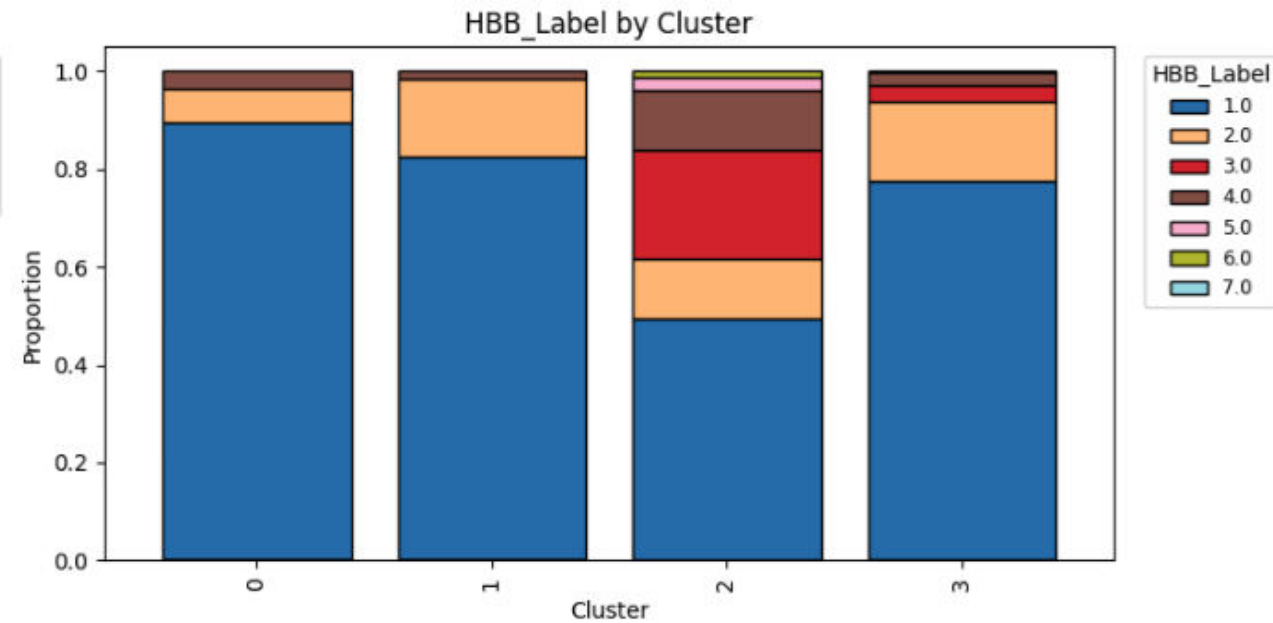
HBB:
1, HbSS |
2, HbSC |
3, HbSbeta0 |
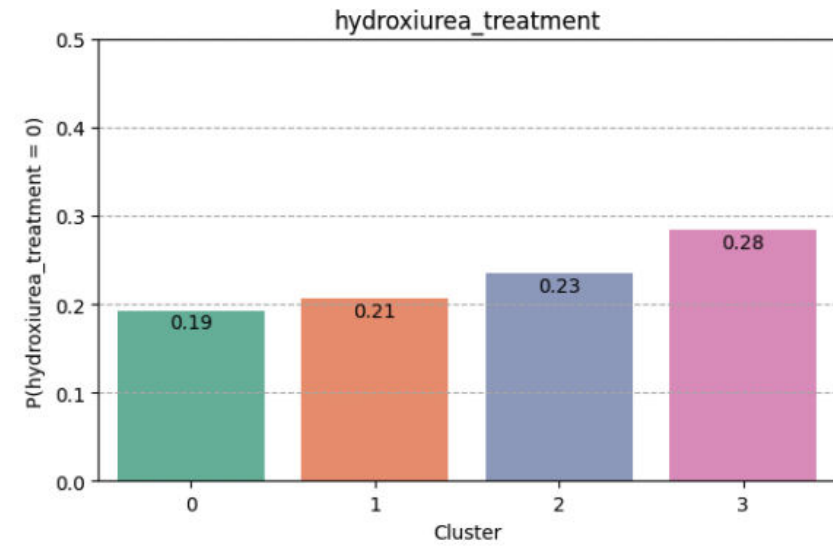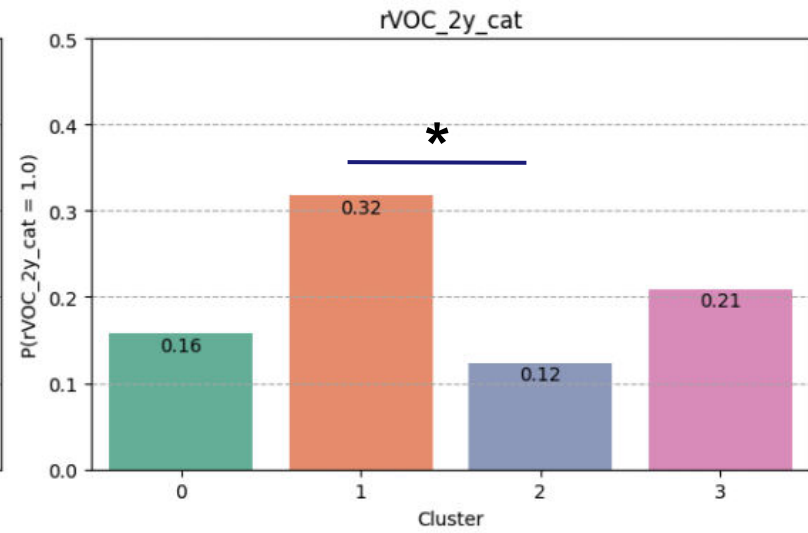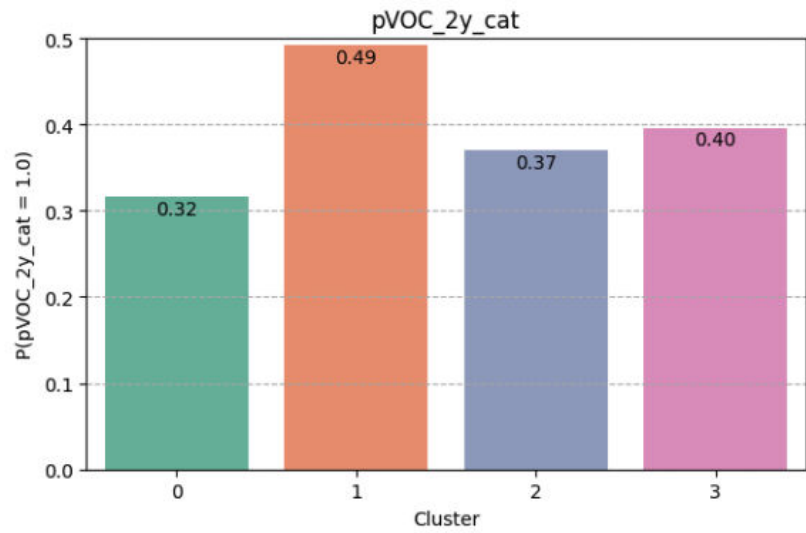4, HbSbeta+ |
5, HbSD |
6, Other |
7, HbS HPHF | 8,
Inconclusive

Signif. pvalues:
(2, 3) ***
(2, 1) ***
(2, 0) ***

GWAS Clustering (945 samples): {0,0.5,1} mapping
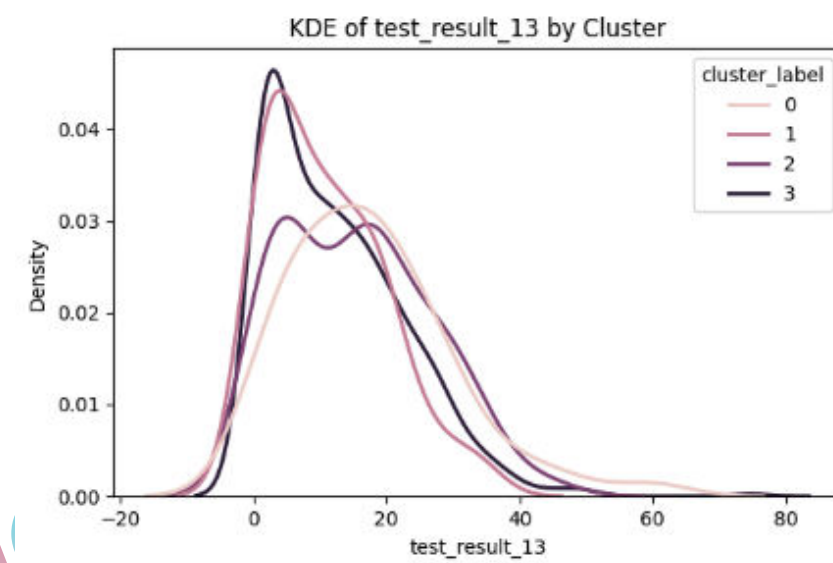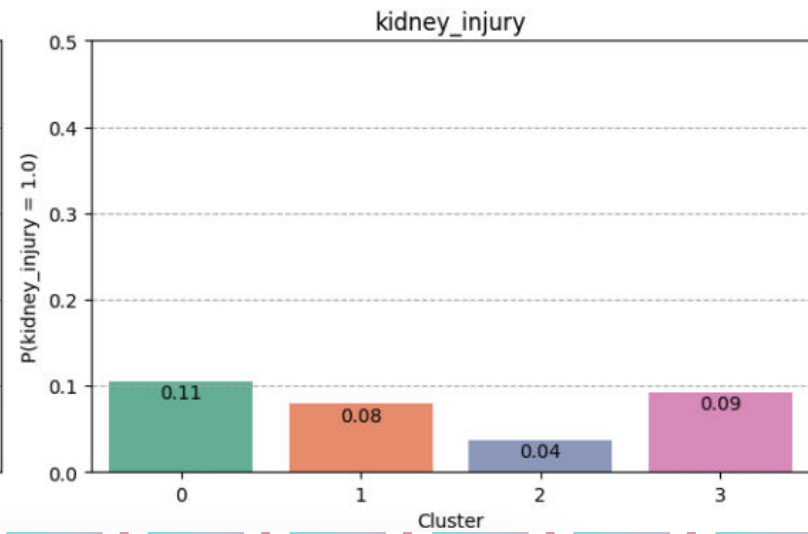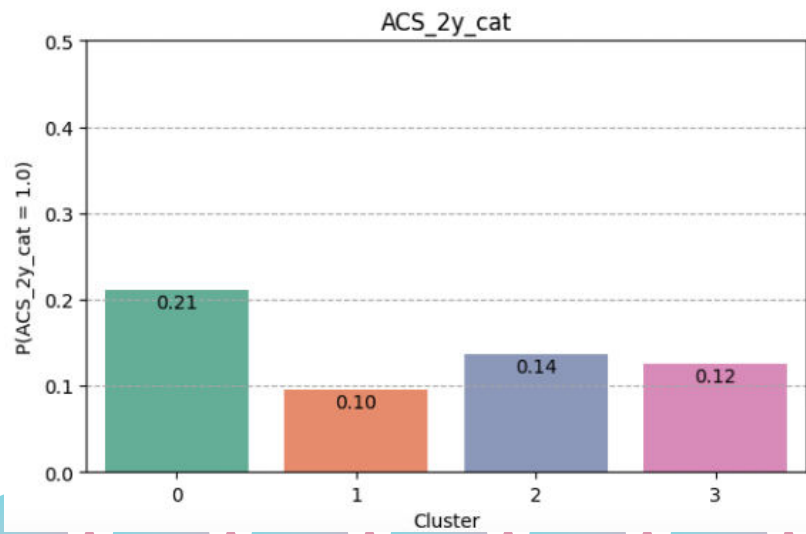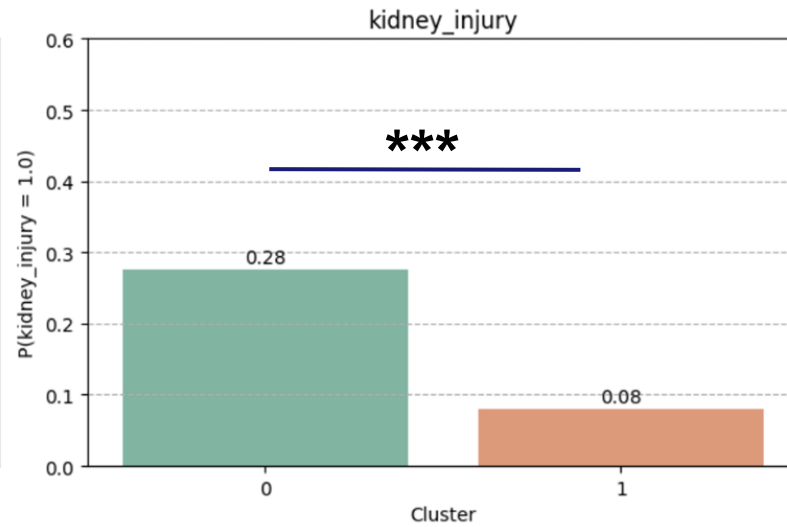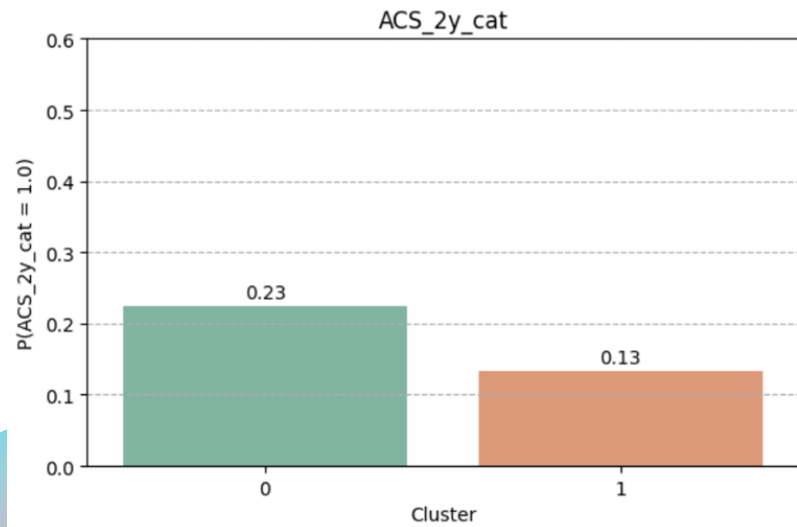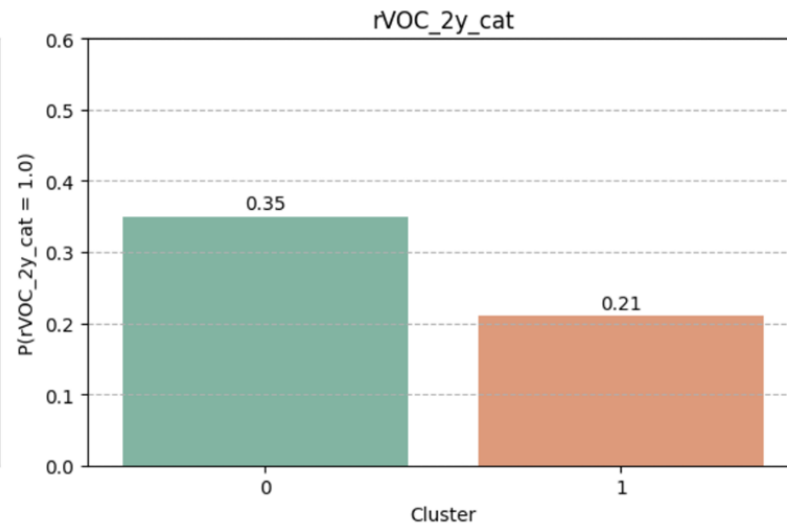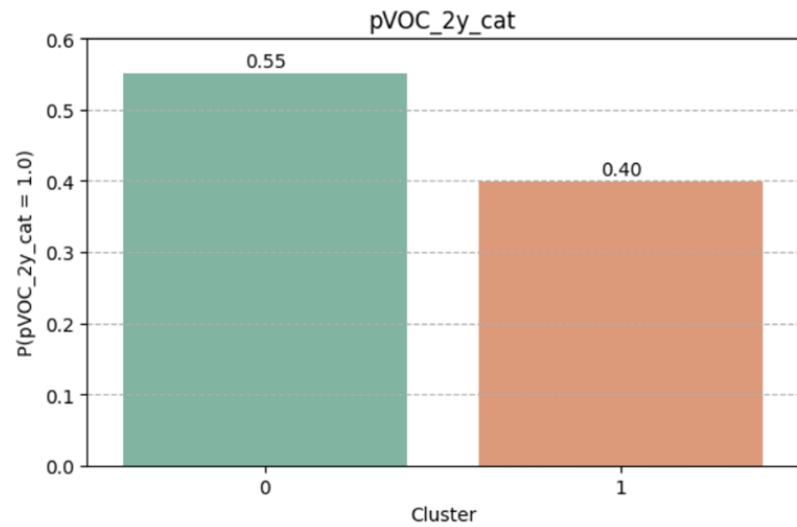
Clustering on Metabolomics (on 961 samples)

Clustering on Metabolomics (on 961 samples)
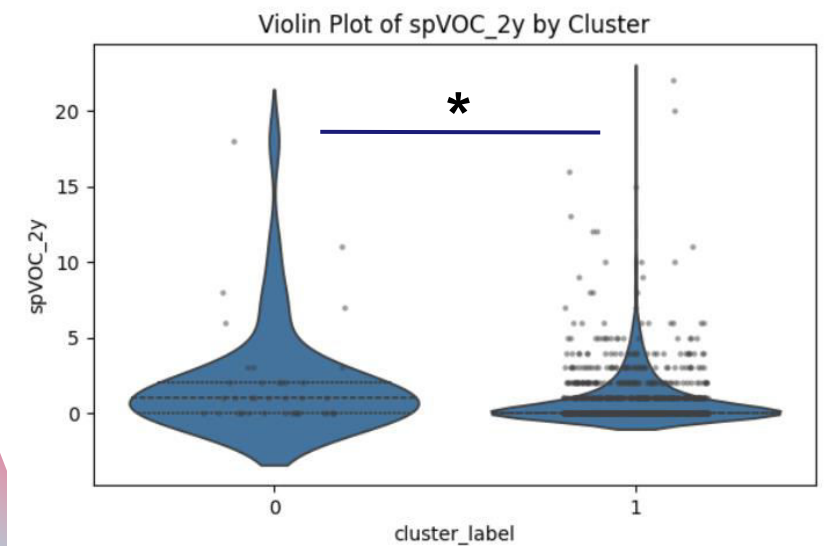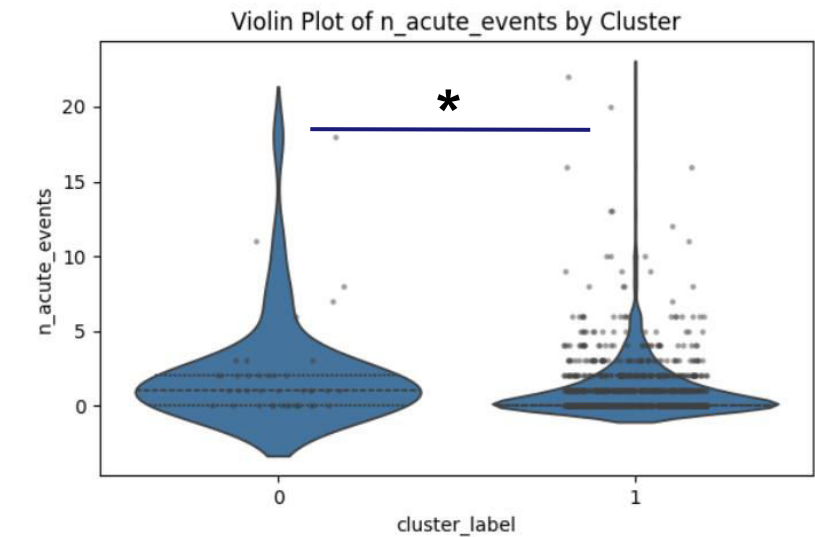
Supervised model

Outcomes to predict:

- ❖ ACS (0/1)
- ❖ pVOC (0/1)
- ❖ spVOC (cont.)
- ❖ ACS (0/1)
- ❖ Kidney Injury (0/1)

How:

Train a unique Neural Network model that exploits at the same time:

- ❖ Clinical and laboratory test data
- ❖ Metabolomics (1800 metabolites)
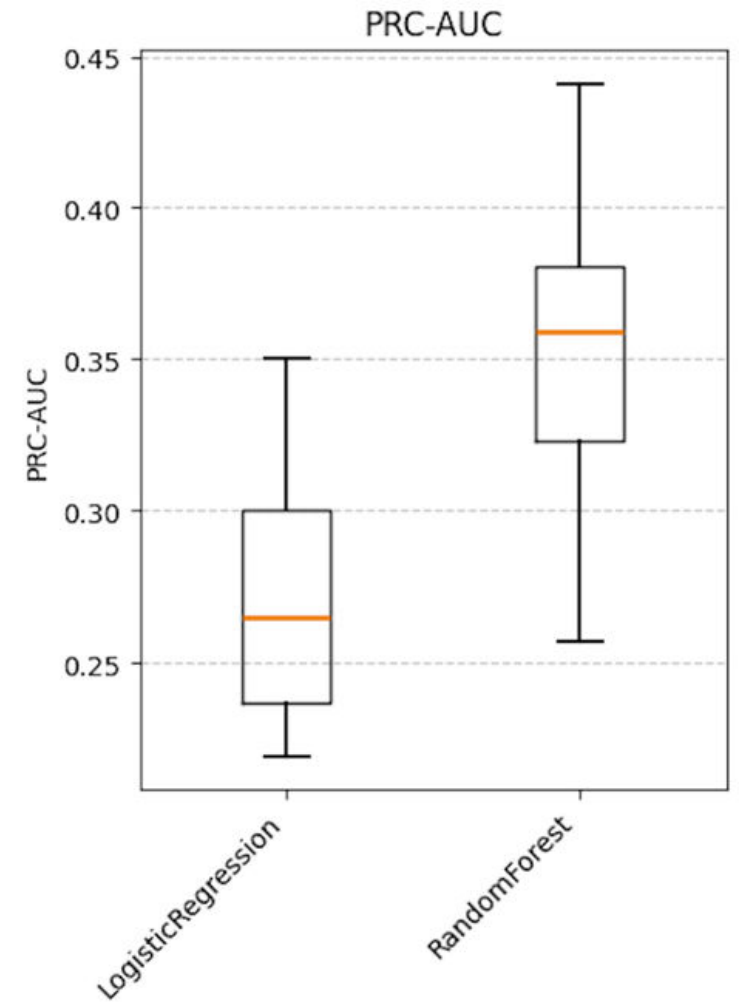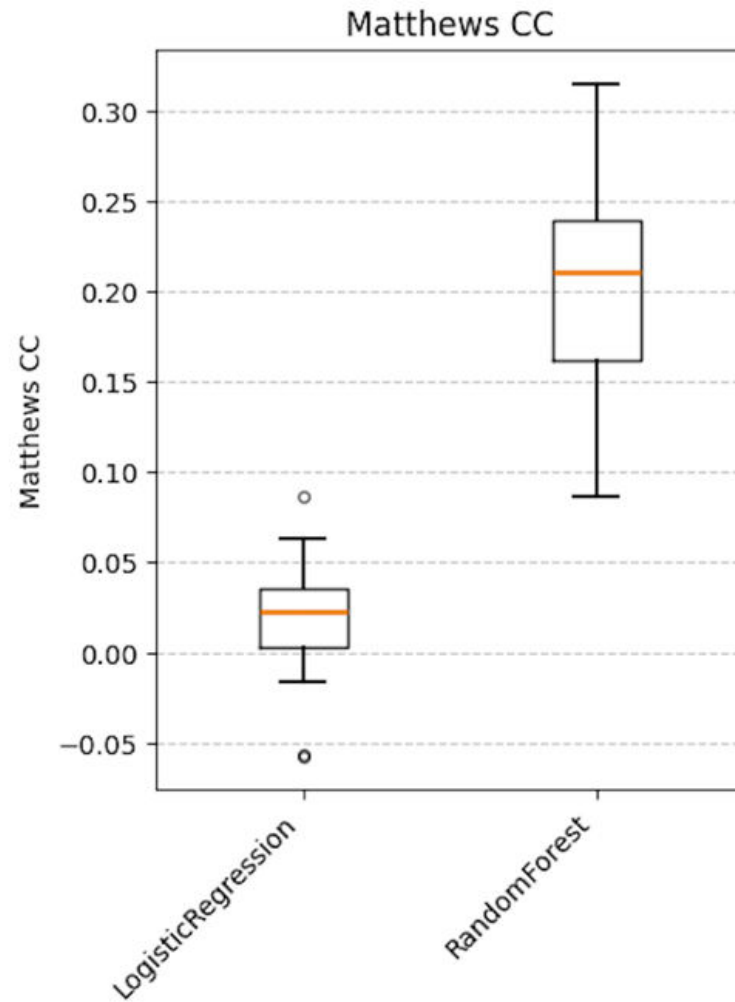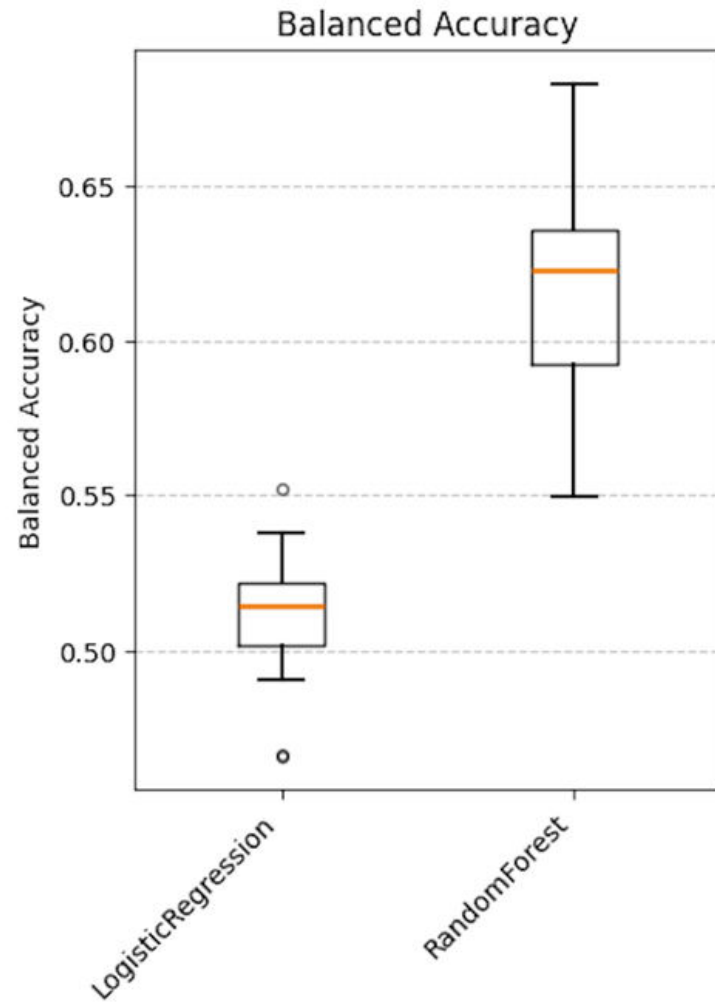- ❖ GWAS (800 SNPs selected)
- ❖ OxygenScan data

Supervised model: Clinical variables

gender	test_result_1	test_result_81	neuro_answer_3
first_hcp_contact_cum	test_result_2	test_result_82	neuro_answer_4
clin_diag_cum	test_result_3	test_result_23	priapism_2y_cat
age	test_result_4	test_result_83	splenic_seq_2y_cat
bmi	test_result_5	test_result_10	bones_complication_answer_1
oxygen_saturation_rate	test_result_6	test_result_62	bones_complication_answer_4
heart_rate	test_result_7	test_result_63	bones_complication_answer_5
diastolic_blood_pressue	test_result_8	test_result_64	bones_complication_answer_3
systolic_blood_pressure	test_result_9	test_result_65	endo_answer_2
hydroxiurea_treatment	test_result_11	test_result_66	lkd_answer_2
cumulative_years	test_result_13	test_result_84	lkd_answer_4
cumulative_years_age	test_result_69	test_result_86	lkd_answer_1
hydroxiurea_dose	test_result_70	test_result_29	lkd_answer_5
chelation_treat	test_result_71	test_result_88	lkd_answer_6
splenomegaly	test_result_72	test_result_87	vhd_answer_3
cholelithiasis	test_result_73	test_result_91	cpd_value_5
cholecystectomy	test_result_74	test_result_92_recalc	splenectomy
ab_test_id___1	test_result_24	test_result_89	Stroke
treatid___4	test_result_25	test_result_90_recalc	CerebralOth
treatid___5	test_result_26	comorbiditiesid___1	
treatid___11	test_result_27	comorbiditiesid___2	sub_orpha_id
treatid___20	test_result_18	comorbiditiesid___4	HBA_Label
treatid___12	test_result_19	comorbiditiesid___14	HBB_Label
treatid___3	test_result_20	comorbiditiesid___22	
treatid___19	test_result_21	comorbiditiesid___19	
treatid___21	test_result_75	comorbiditiesid___29	
treatid___18			

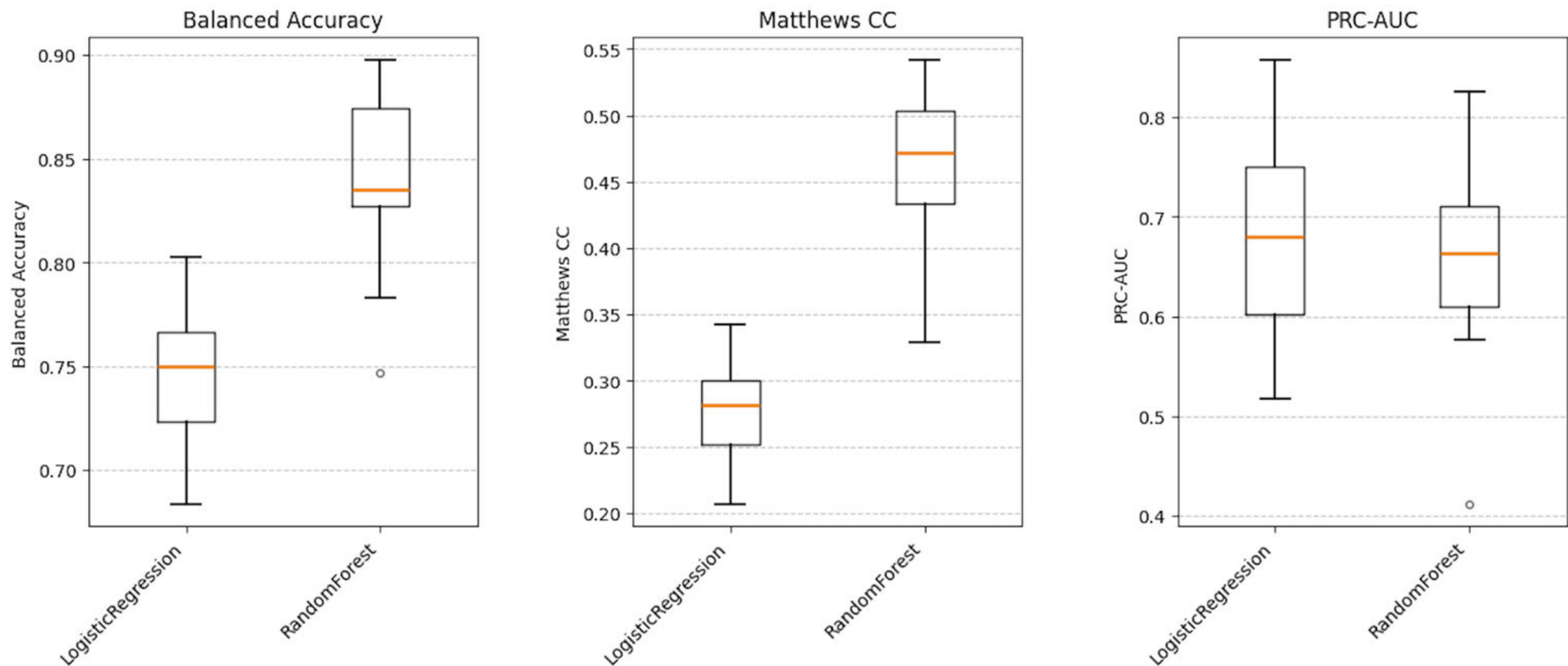# rVOC outcome - Baseline (only clinical) models
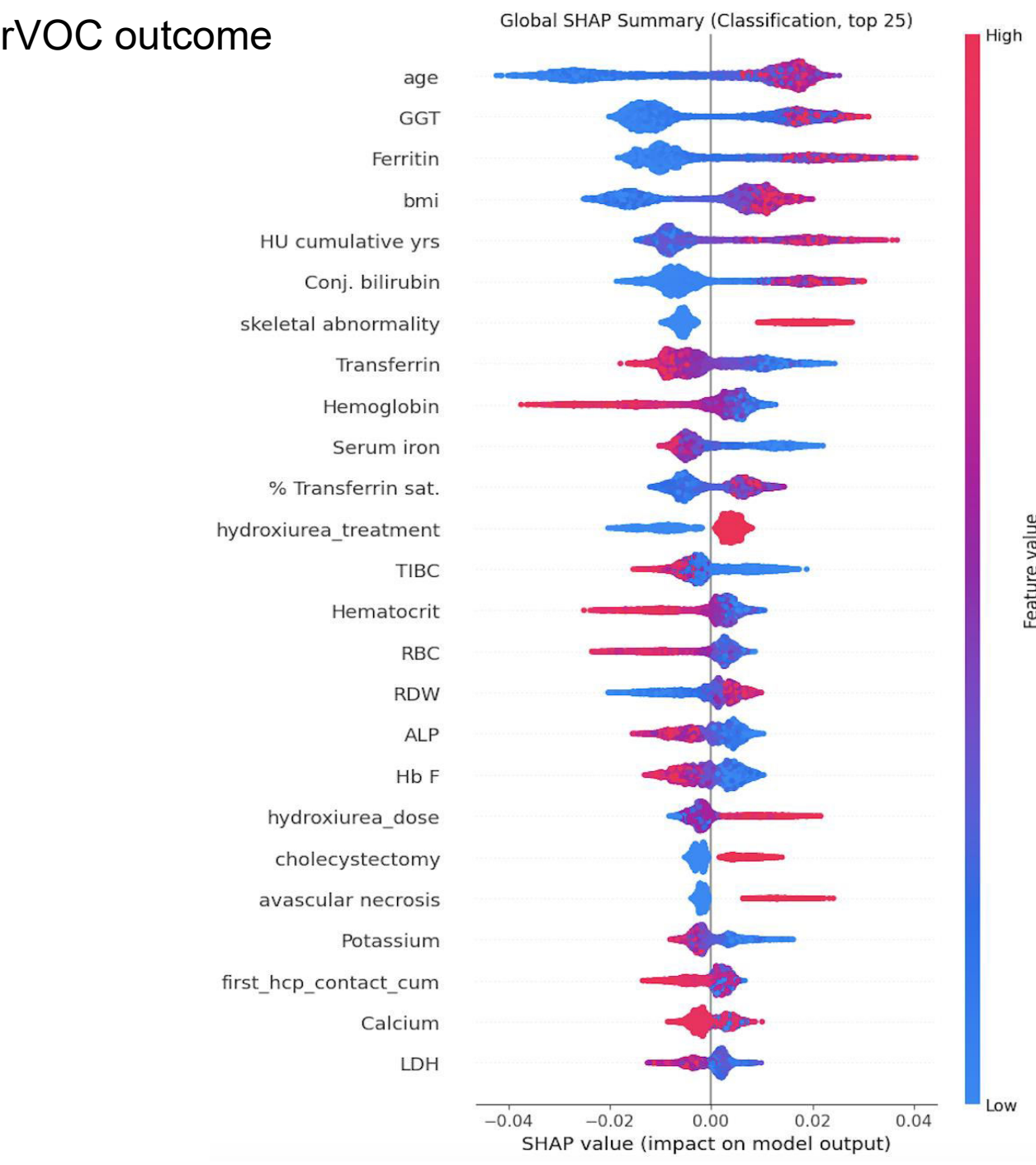

rVOC_2y_cat — Test-Set Distributions
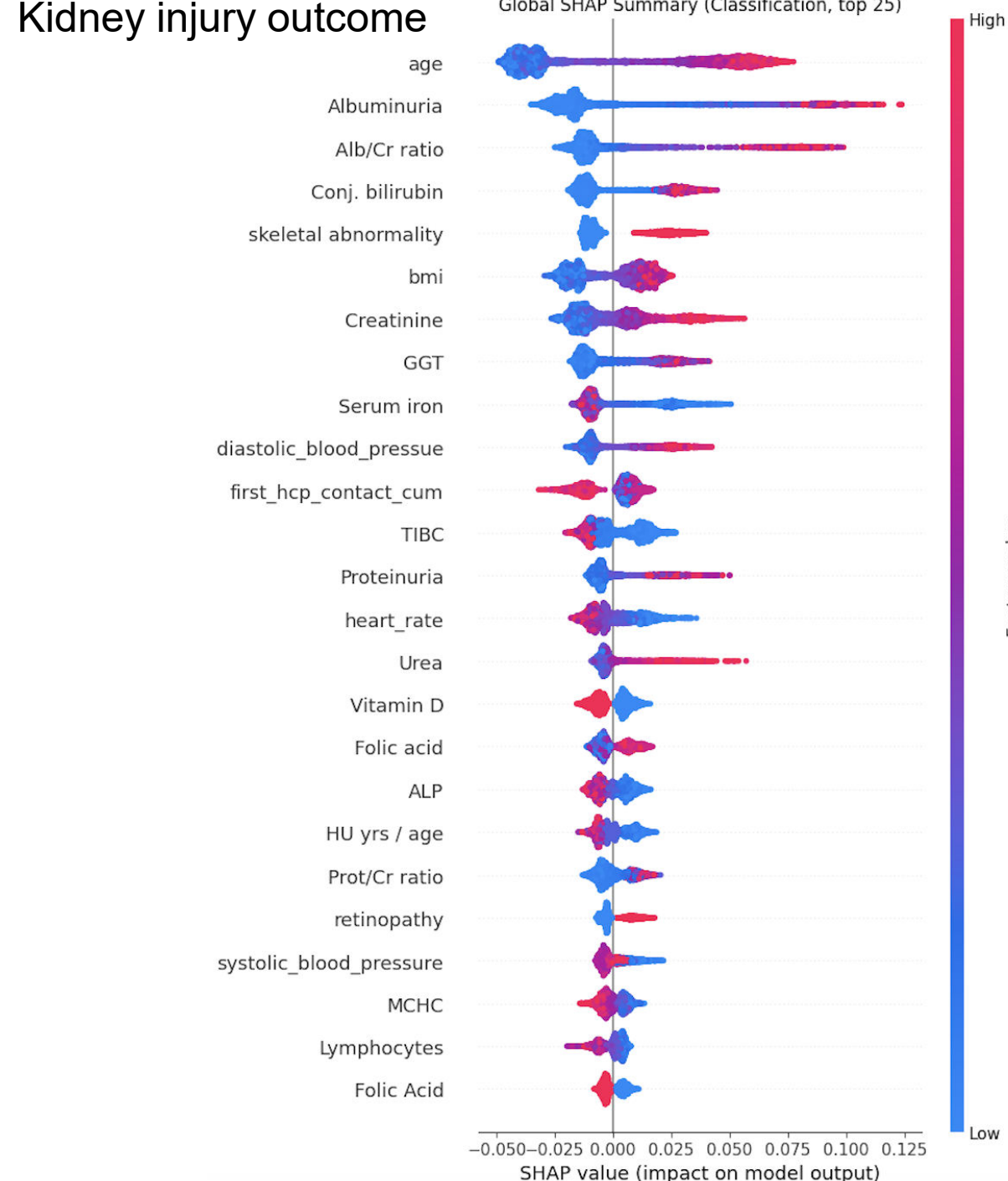
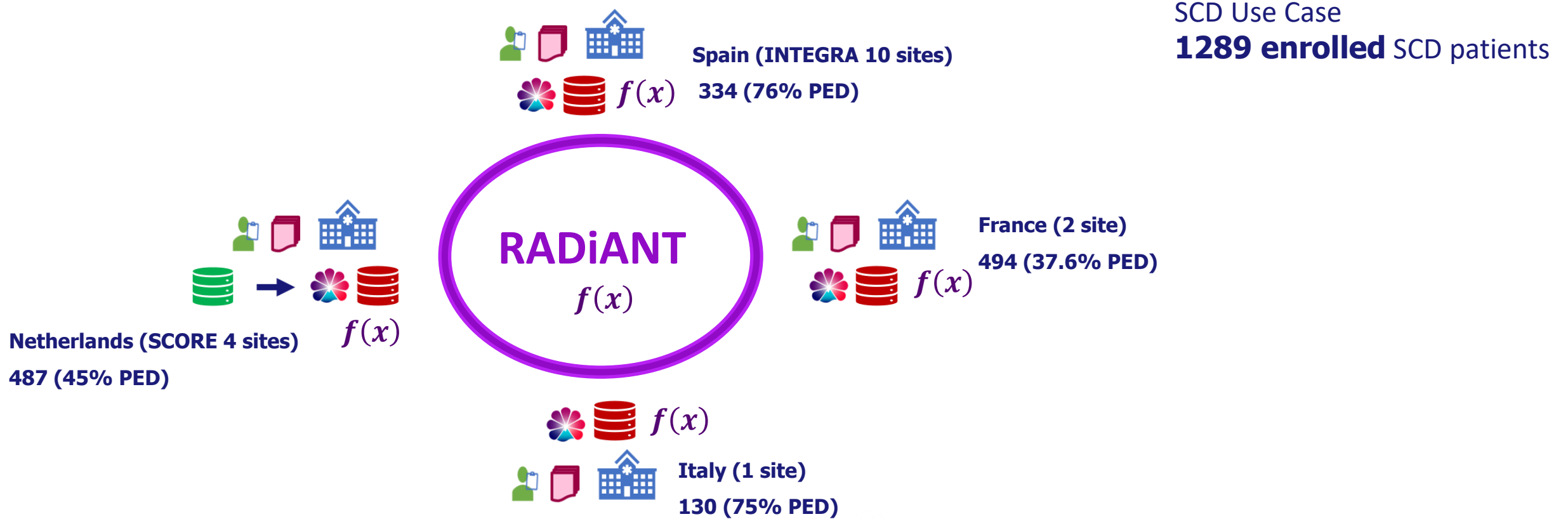# Kidney injury outcome - Baseline (only clinical) models

rVOC outcome — Global SHAP Summary (Classification, top 25)

Kidney injury outcome — Global SHAP Summary (Classification, top 25)

RADiANT

Spain (INTEGRA 10 sites)
334 (76% PED)

Netherlands (SCORE 4 sites)
487 (45% PED)

France (2 site)
494 (37.6% PED)

Italy (1 site)
130 (75% PED)

RADiANT
$f(x)$

SCD Use Case
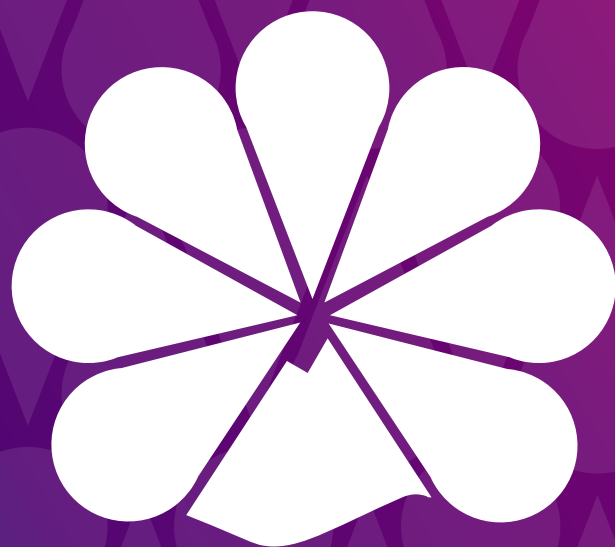**1289 enrolled** SCD patients

RADeep

RADiANT

www.radeepnetwork.eu