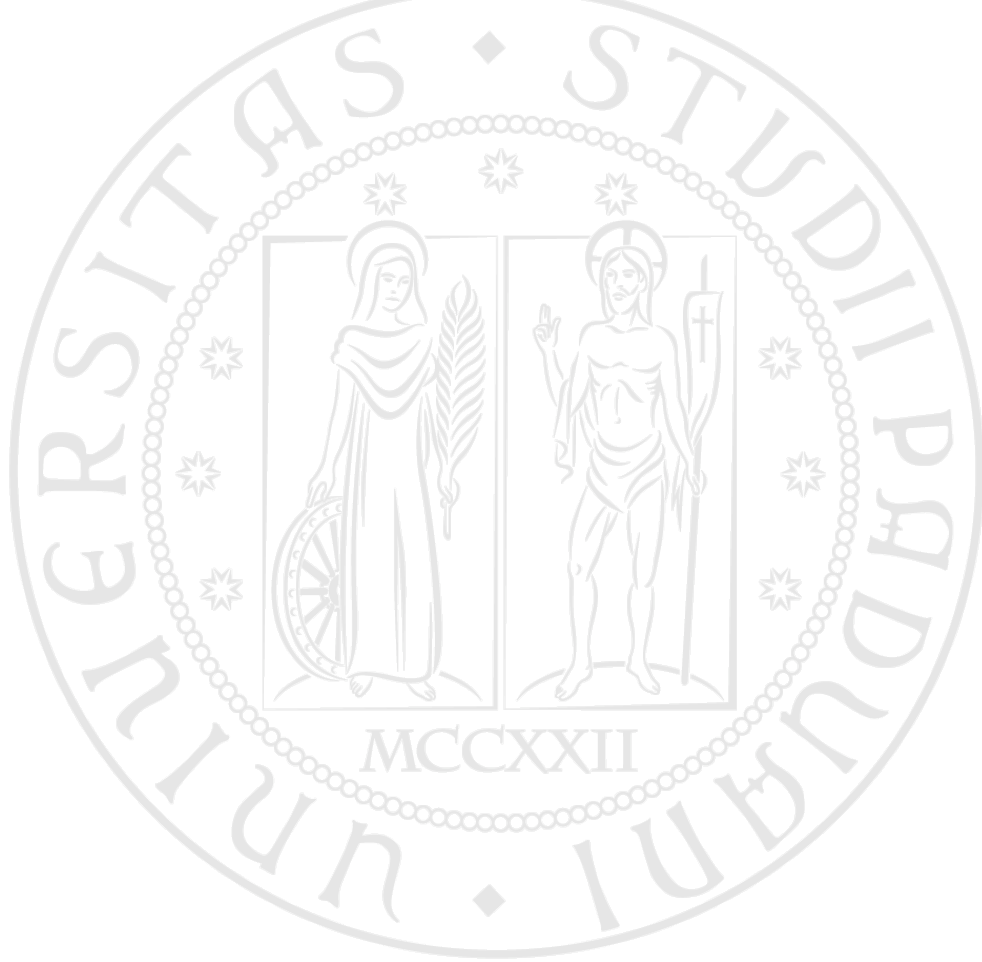


CAGI slide pack guidelines

I agree to my slides being used according to the following terms: [Marco Carraro](#)

- Authors; please include THIS slide at the beginning of your presentation to indicate that you agree with reuse of your slides in CAGI presentations, according to the terms below, and digitally sign by writing your name above.
- To disseminate the CAGI results, I encourage all CAGI participants to remix these CAGI slides for inclusion in their oral and poster presentations
- CAGI participants may use or publish any content of these slides only in compliance with the CAGI Data Use Agreement
- CAGI participants may not cite or publish any content of these slides except with the written permission of the originating author or in the primary CAGI publications
- CAGI participants must acknowledge the original authors of these slides and CAGI 2013. Include the acknowledgement banner across the bottom
- CAGI participants must include the CAGI credits slide in their presentations
- Slides with a red slash (no remix) may not be included in a presentation unless all attendees are registered CAGI participants who signed the data use agreement
- Slides with a red X over the entire slide may not be used in any circumstances
- Slides with a blue slash (embargoed) may not be used unless the embargo is explicitly lifted
- Slides without slash or X may be tweeted unless flagged with the “No tweet” icon.
- CAGI participants using these slides must explain the “No tweet” icon, which means that the slide content should not be disseminated outside the presenting conference hall. If the presentation venue permits pervasive tweeting, you may not include these slides in your talk
- Note for slide authors: we expect all CAGI participants to abide by these restrictions, and will make best effort to ensure they are followed. However, bear in mind that we have limited means of enforcing them, and therefore the restrictions cannot be guaranteed.





MARCO CARRARO

*BioComputing UP, Dept. of Biomedical Sciences,
University of Padova Italy*

Intellectual Disability panel challenge

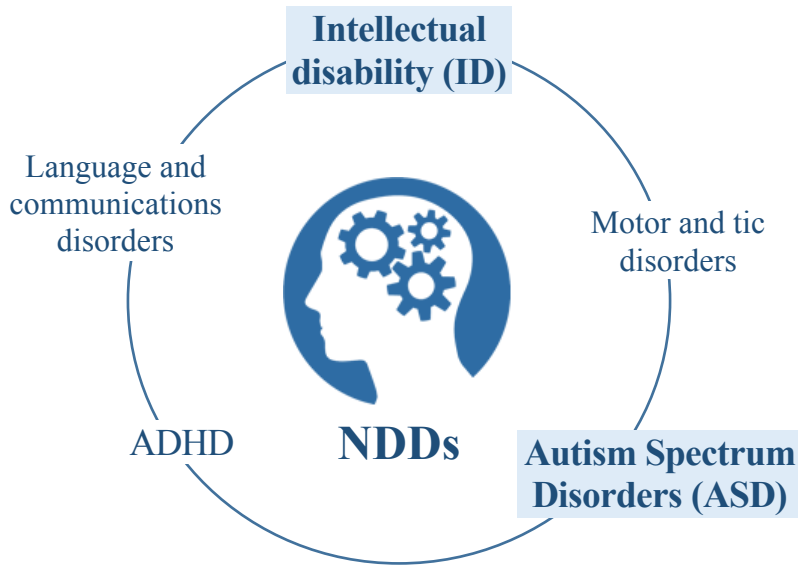
CAGI

07/07/2018

Overview

- The challenge
- Dataset
- Assessment metrics
- Variants assessment
- Results

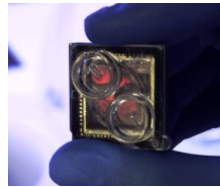
Experimental data



Neurodevelopmental disorders (NDDs):

- Heritable diseases
- Clinically and genetically heterogeneous
- **Comorbidity** = overlapping impaired molecular pathways

74 genes panel to investigate ID-ASD comorbidity



190 screened patients:

- In **33 cases**, we were able to assign a **molecular diagnosis**;
- For additional 29 probands, at least one likely pathogenic variant has been detected;
- **Putative diagnostic yield: ~ 30%**

**Predict patients' clinical descriptions and pathogenic variants
from gene panel sequences**

The Challenge

Paz. ID	ID	Variant	ASD	Variant	Epilepsy	...
MR2201_01	1	11:706655:C:T	1	7:2586:T:C	NA	...
MR2202_01	1	1:6455:A:T	0	NA	NA	...
MR2214_01	1	NA	NA	NA	0	...
MR2227_01	0	7:25896:T:C	1	NA	0	...
MR2275_01	1	NA	1	1:6455:A:T	0	...
...	...	...	...	...	...	...
MR2290_01	1	11:704655:C:T	0	11:706455:C:T	1	...

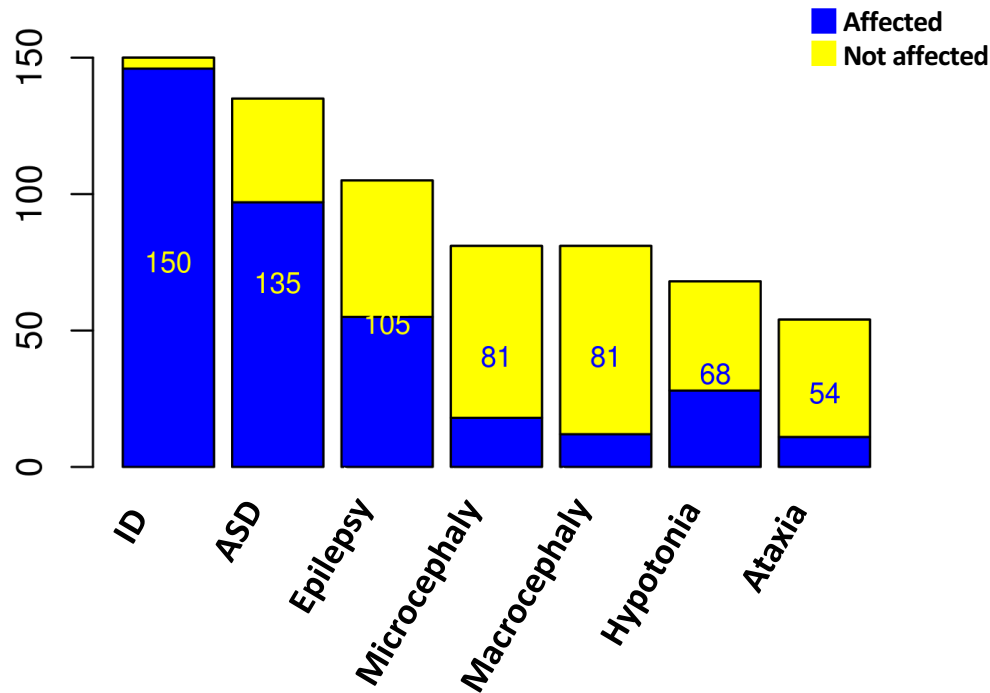
The challenge:

- predict **p [0,1]** to be affected (+ SD)
- match the putative **causative variant(s)** (+ confidence)

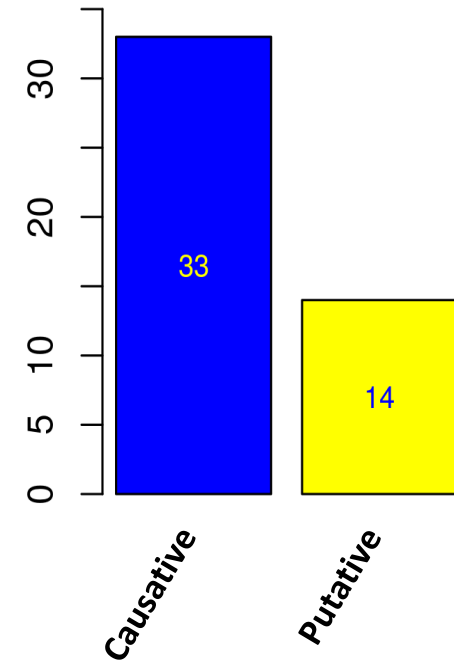
For each of the 7 disease phenotypes

Dataset

Patients with known features



Known variants



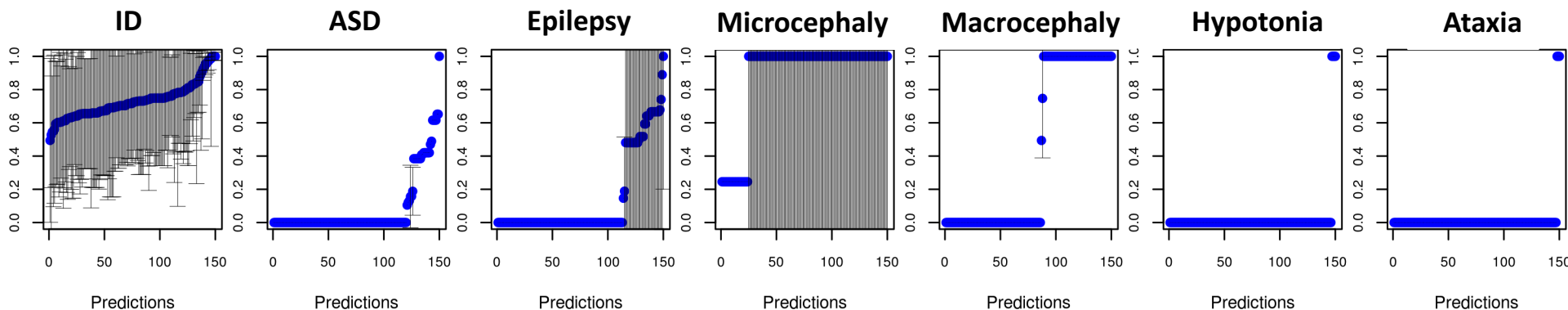
Submissions

- **5 participating groups**
- **14 submissions** (6 submissions from group 2)
- **1 late submission** (only p=1)

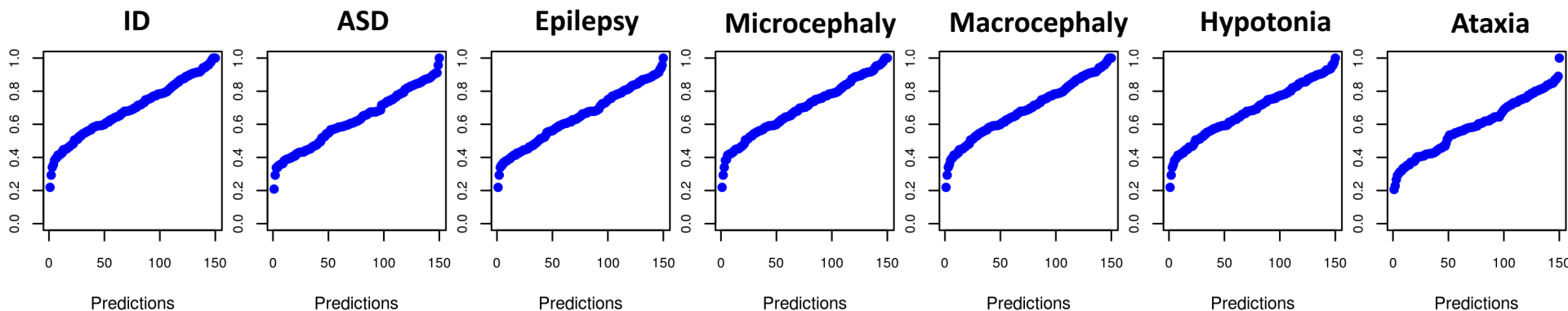
Group ID	Group Name
Group 1	Mooney-Radivojac
Group 2	Moult Lab.
Group 3	Lichtarge Lab.
Group 4	Brenner Lab.
Group 5	enGenome

Submissions

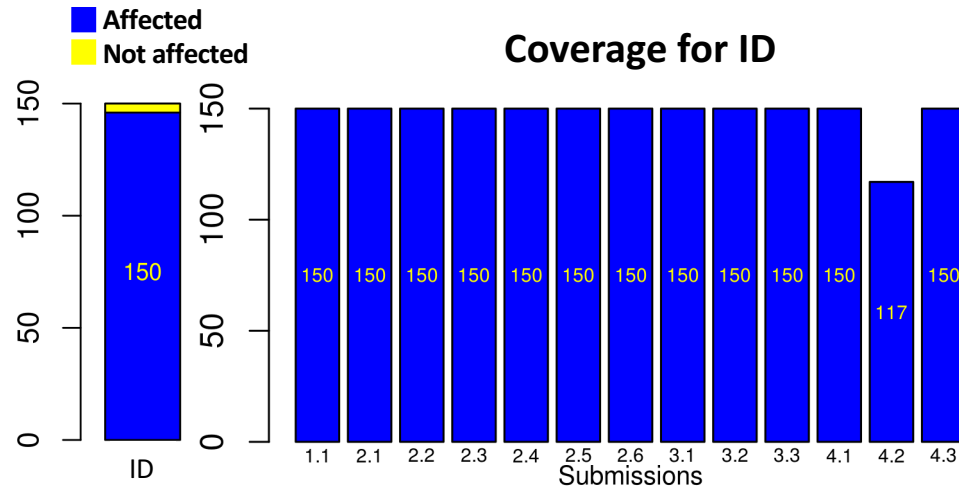
Submission 3.3



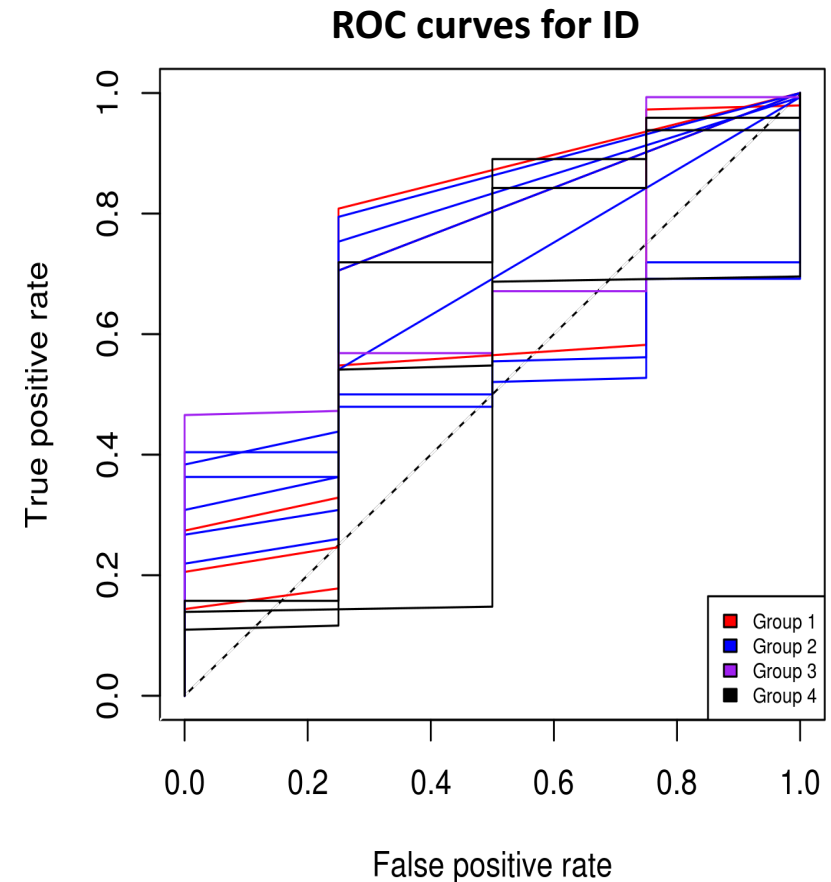
Submission 4.3



Performance assessment ID



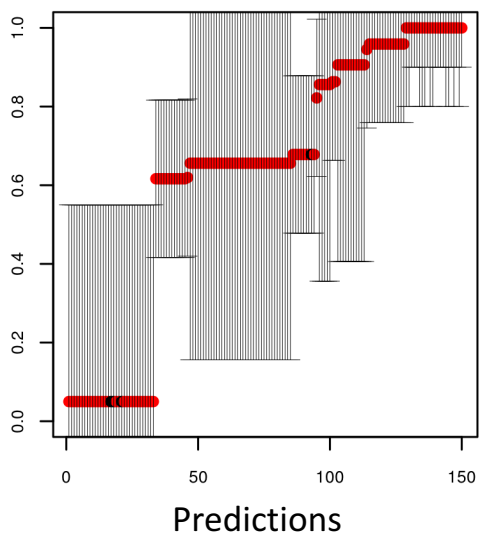
Submission	AUC	MCC	TNR	NPV
Sub. 2.4	0.78	0.21	0.75	0.10
Sub. 2.2	0.75	0.22	0.75	0.10
Sub. 2.6	0.74	0.19	0.75	0.08



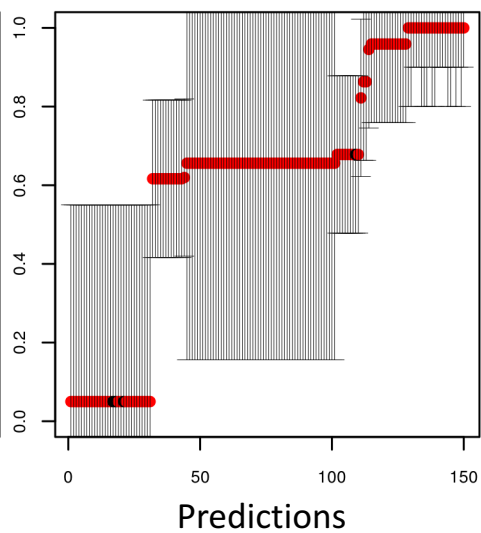
ID best performers

● Affected
● Not affected

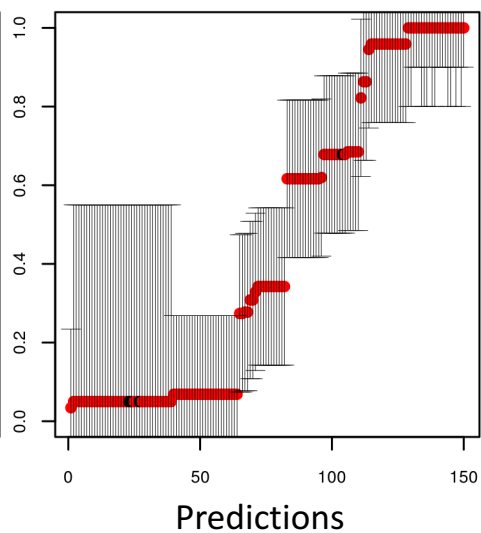
Submission 2.4



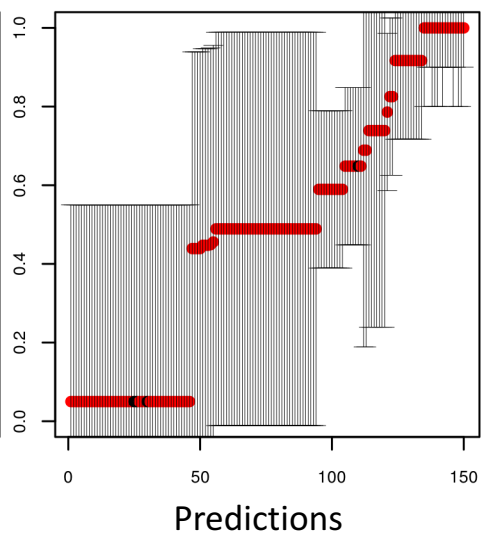
Submission 2.4



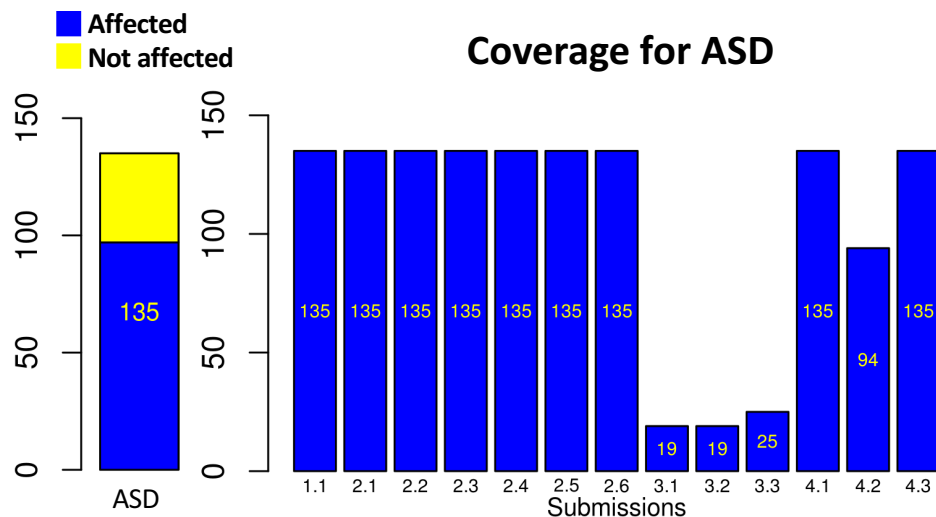
Submission 2.6



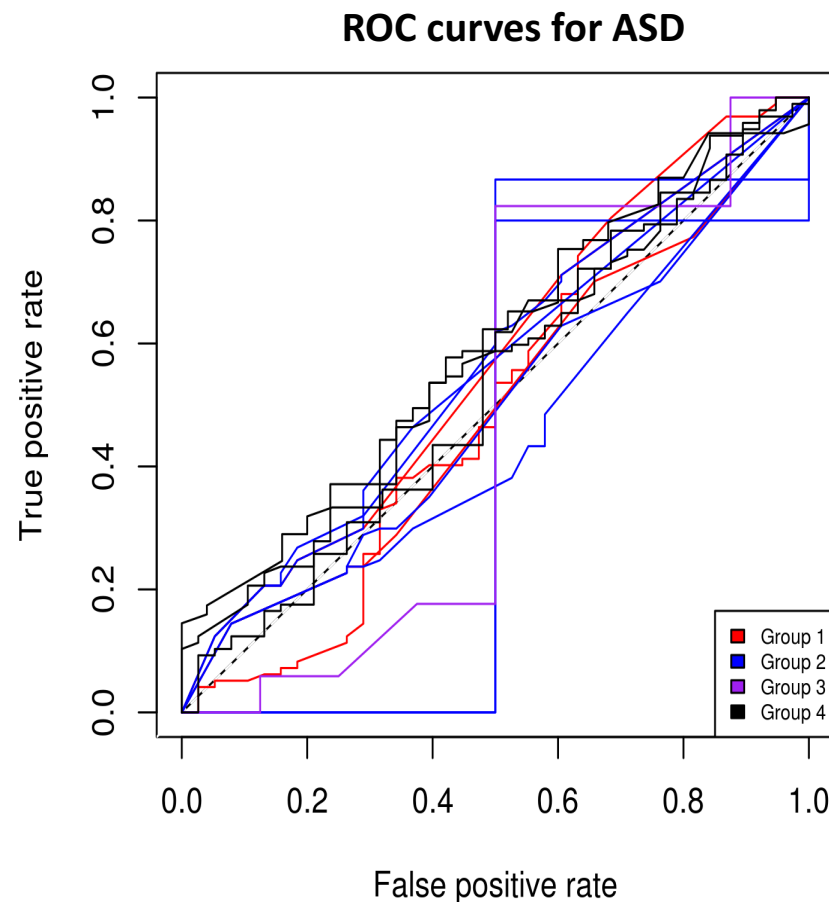
Submission 2.3



Performance assessment ASD



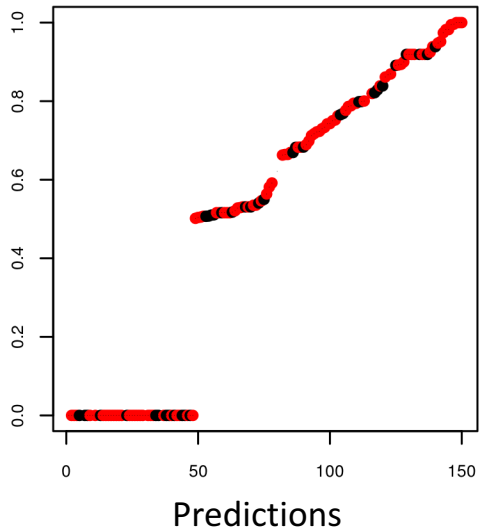
Submission	AUC	MCC	FNR	TNR	NPV
Sub. 4.2	0.57	0.21	0.86	1	0.30
Sub. 4.1	0.56	0.18	0.90	1	0.30
Sub. 2.3	0.56	0.11	0.38	0.50	0.34



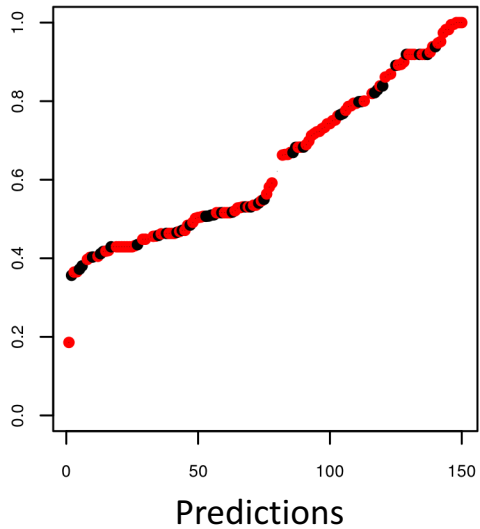
ASD best performers

● Affected
● Not affected

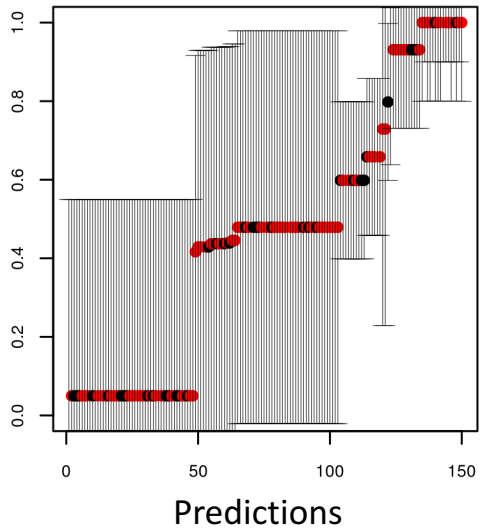
Submission 4.2



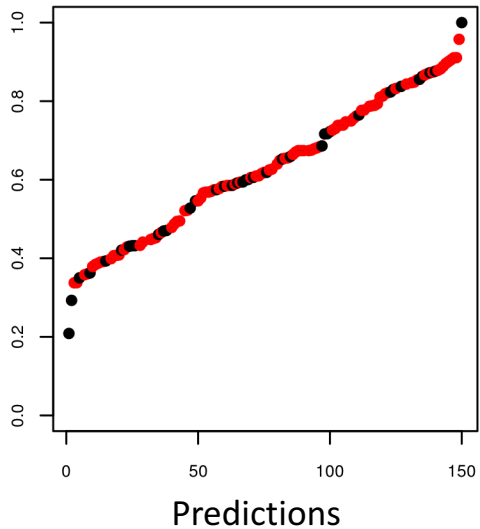
Submission 4.1



Submission 2.3



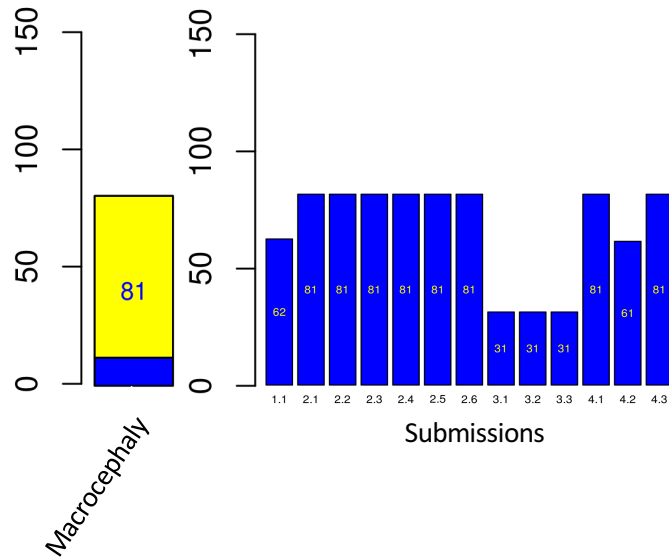
Submission 4.3



Performance assessment Macrocephaly

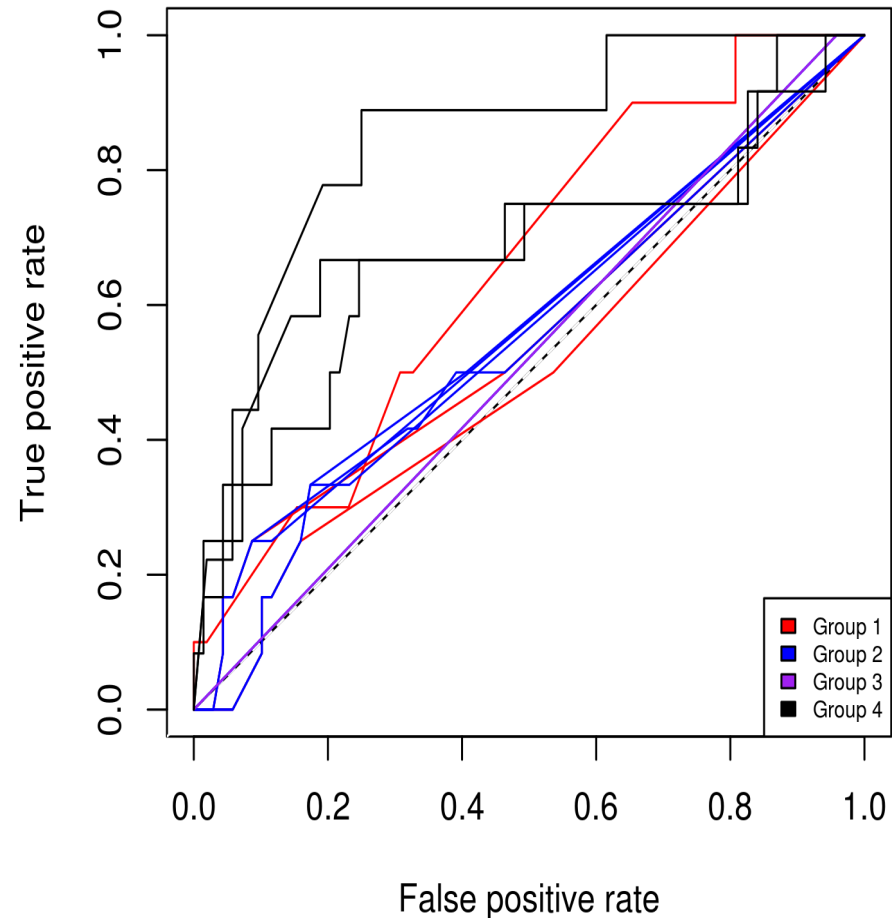
■ Affected
■ Not affected

Coverage for Macrocephaly



Submission	AUC	MCC	F1
Sub. 4.2	0.85	0.48	0.53
Sub. 4.1	0.70	0.39	0.49
Sub. 4.4	0.67	0.39	0.35

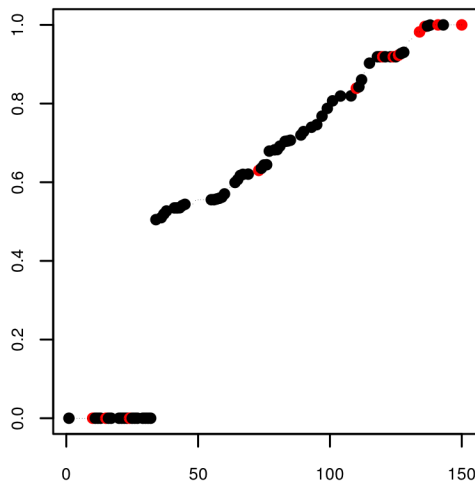
ROC curves for Macrocephaly



Macrocephaly best performers

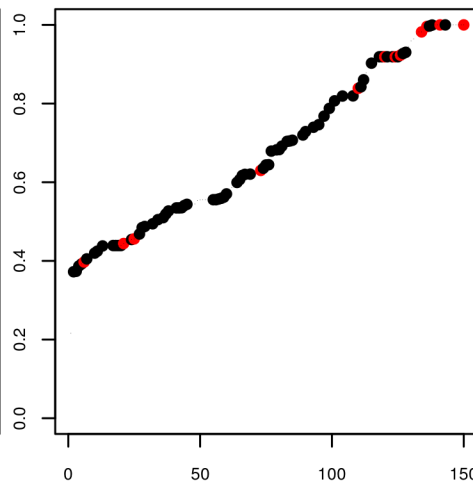
● Affected
● Not affected

Submission 4.2



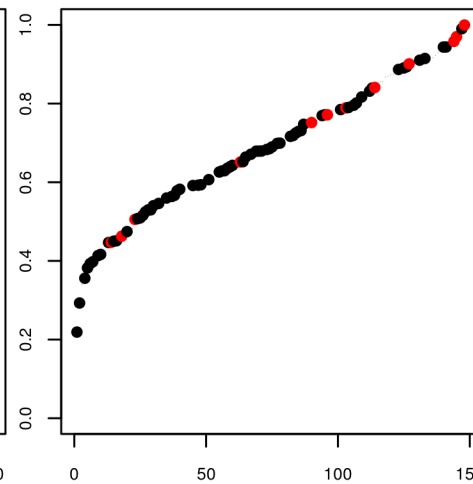
Predictions

Submission 4.1



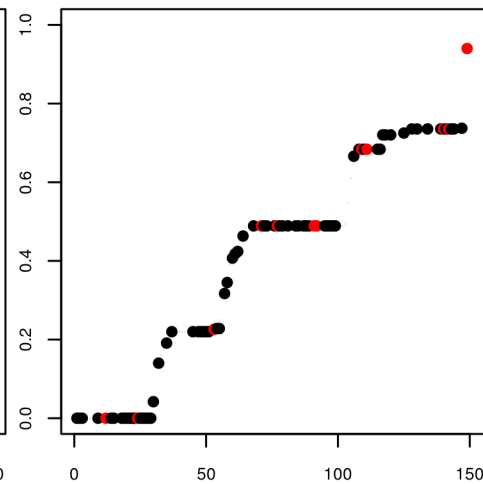
Predictions

Submission 4.3



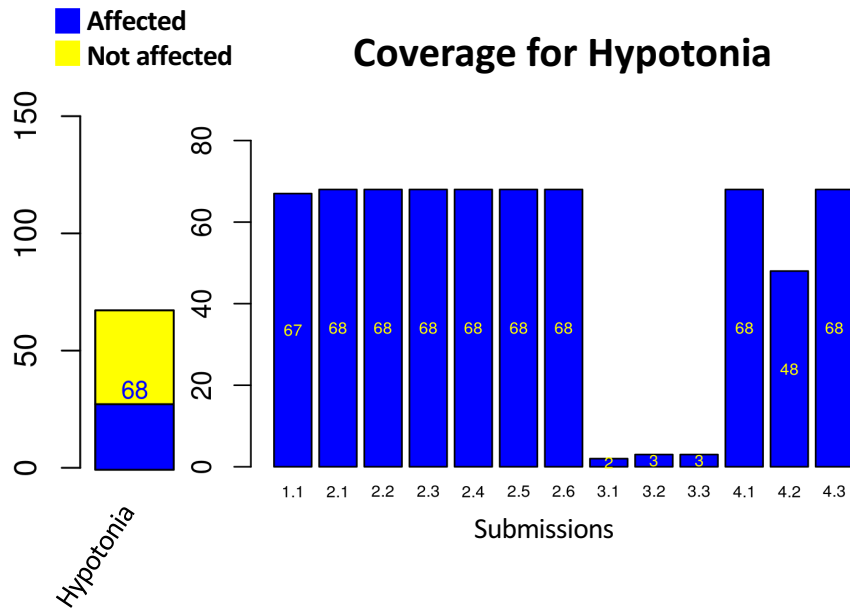
Predictions

Submission 1.1

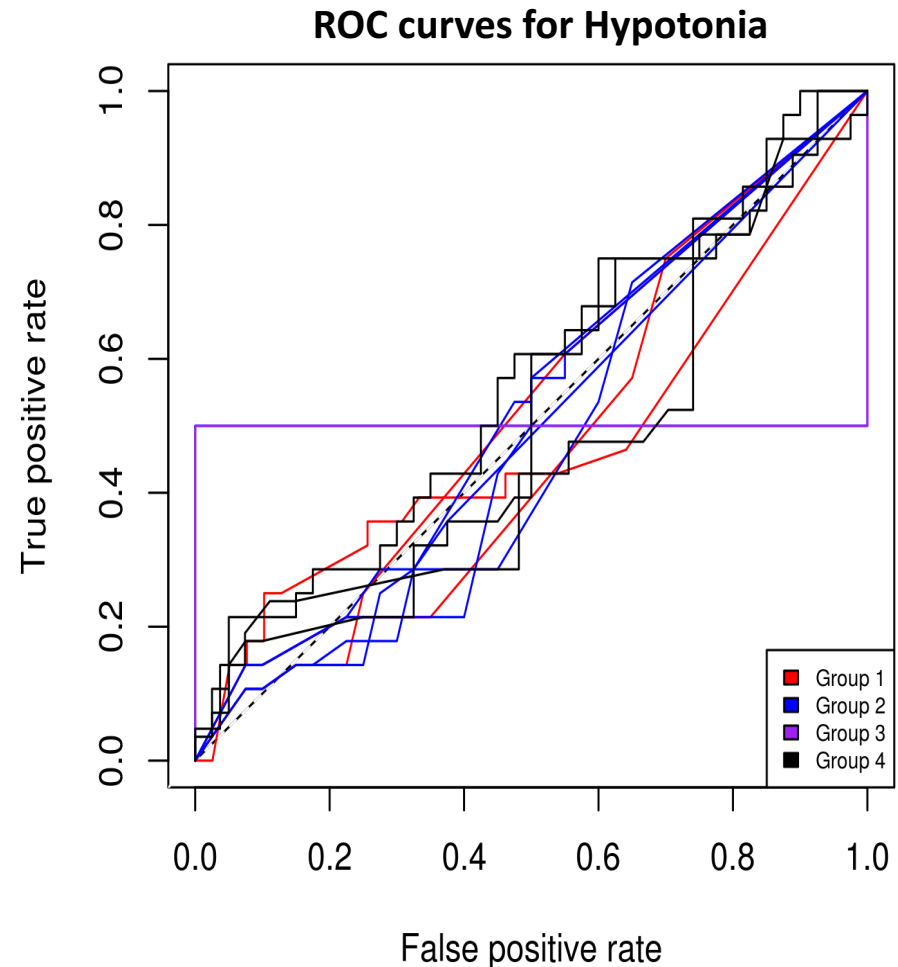


Predictions

Performance assessment Hypotonia



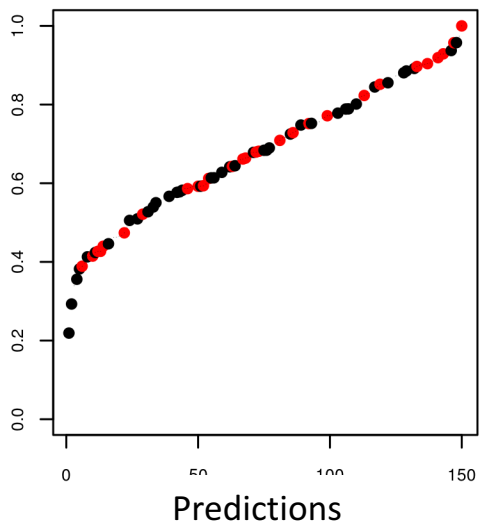
Submission	AUC	MCC	F1
Sub. 4.3	0.56	0.25	0.33
Sub. 3.1	0.52	0.10	0.38
Sub. 2.1	0.52	0.07	0.51



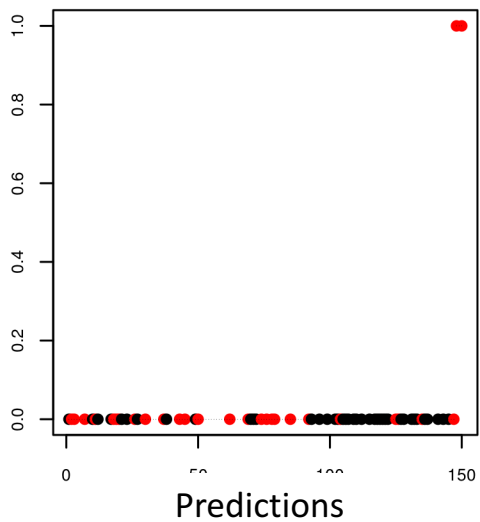
Hypotonia best performers

● Affected
● Not affected

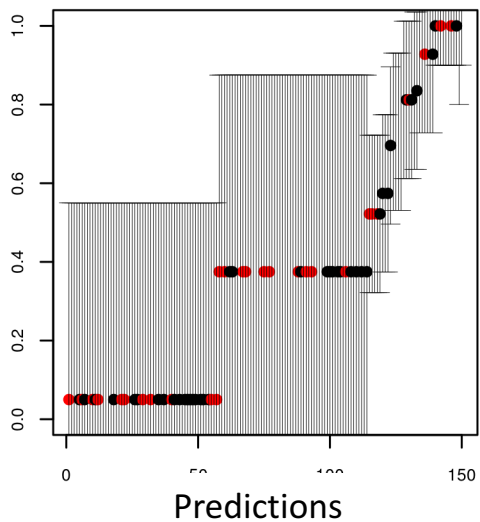
Submission 4.3



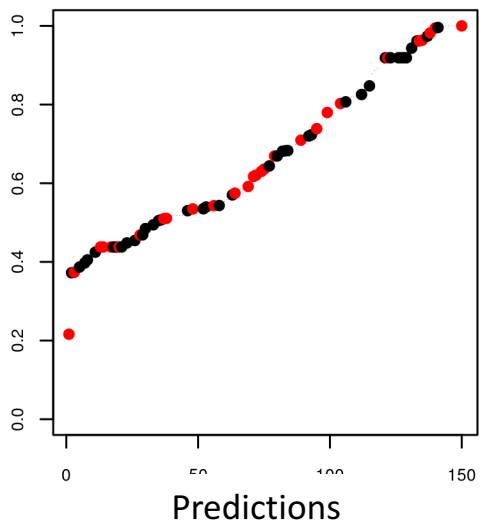
Submission 3.1



Submission 2.1



Submission 4.1

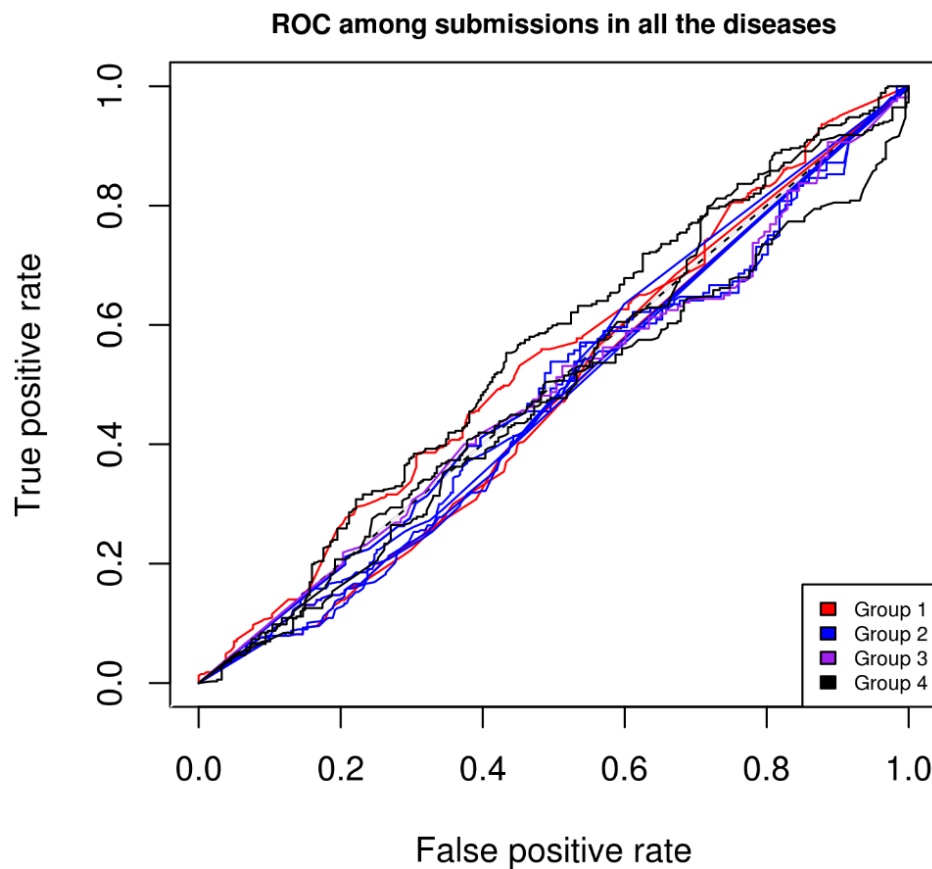


Overall performance assessment

	ID	ASD	Epilepsy	Microcephaly	Macrocephaly	Hypotonia	Ataxia
Sub. 4.3	0.68	0.56	0.56	0.63	0.67	0.56	0.46
Sub. 4.1	0.61	0.56	0.54	0.57	0.70	0.51	0.45
Sub. 4.2	0.42	0.57	0.54	0.51	0.85	0.47	0.49
Sub. 2.3	0.71	0.56	0.39	0.49	0.56	0.51	0.66
Sub. 2.1	0.70	0.55	0.38	0.51	0.55	0.52	0.61
Sub. 2.4	0.78	0.49	0.41	0.48	0.56	0.49	0.72
Sub. 1.1	0.57	0.51	0.52	0.57	0.65	0.47	0.32
Sub. 2.5	0.64	0.55	0.40	0.53	0.57	0.49	0.56
Sub. 2.2	0.75	0.50	0.41	0.55	0.50	0.47	0.66
Sub. 2.6	0.74	0.46	0.41	0.49	0.55	0.50	0.66
Sub. 3.3	0.68	0.48	0.45	0.50	0.52	0.50	0.00
Sub. 3.2	0.55	0.43	0.44	0.50	0.52	0.50	0.00
Sub. 3.1	0.51	0.40	0.42	0.50	0.52	0.00	0.00

Statistical significance

- AUC overall
- 10000 random predictions
- same 1/0 ratio of real data (for every phenotype)



AUC Overall

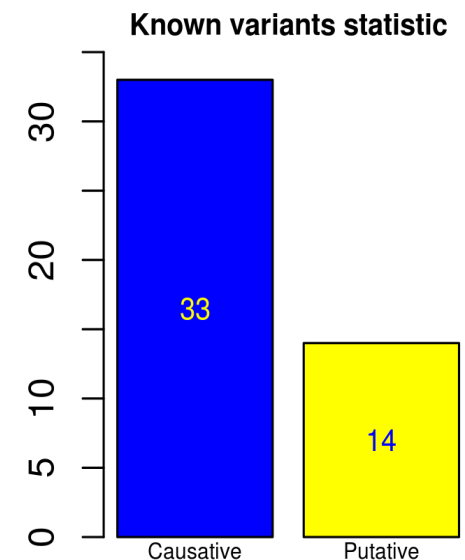
0.55	Sub. 4.3
0.54	Sub. 1.1
0.51	Sub. 4.1
0.49	Sub. 2.4
0.49	Sub. 3.3
0.49	Sub. 3.2
0.48	Sub. 2.2
0.48	Sub. 2.6
0.48	Sub. 3.1
0.47	Sub. 2.3
0.47	Sub. 2.1
0.47	Sub. 2.5
0.46	Sub. 4.2

Causative variants

- Predicted variants have been associated to a patient and not to each phenotype

Submission ID	# of causative variant found
Sub. 1.1	2
Sub. 2.1	10
Sub. 2.2	15
Sub. 2.3	10
Sub. 2.4	15
Sub. 2.5	10
Sub. 2.6	15
Sub. 3.1	4
Sub. 3.2	8
Sub. 3.3	8
Sub. 4.1	4
Sub. 4.2	3
Sub. 4.3	7

Submission ID	# of putative variant found
Sub. 1.1	1
Sub. 2.1	2
Sub. 2.2	1
Sub. 2.3	2
Sub. 2.4	1
Sub. 2.5	2
Sub. 2.6	4
Sub. 3.1	4
Sub. 3.2	4
Sub. 3.3	1
Sub. 4.1	1
Sub. 4.2	3
Sub. 4.3	0



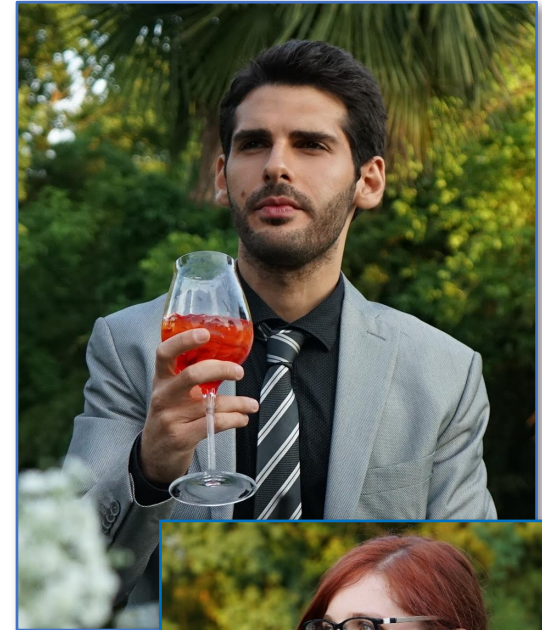
Acknowledgments

Prof. Silvio CE Tosatto



Luigi Chiricosta

MSc



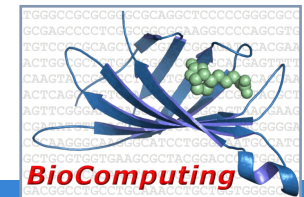
**CAGI Organizers
Data providers
Predictors**

Alessandra Gasparini

PhD

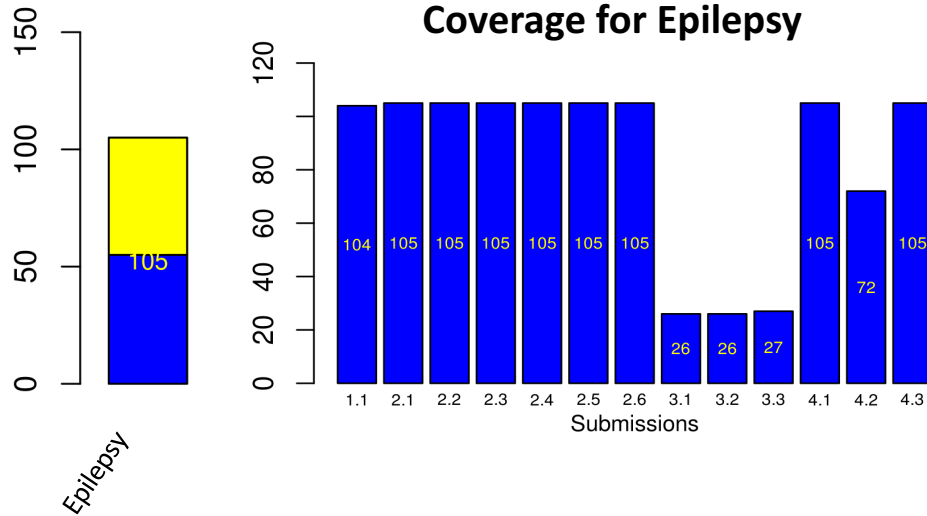


CAGI

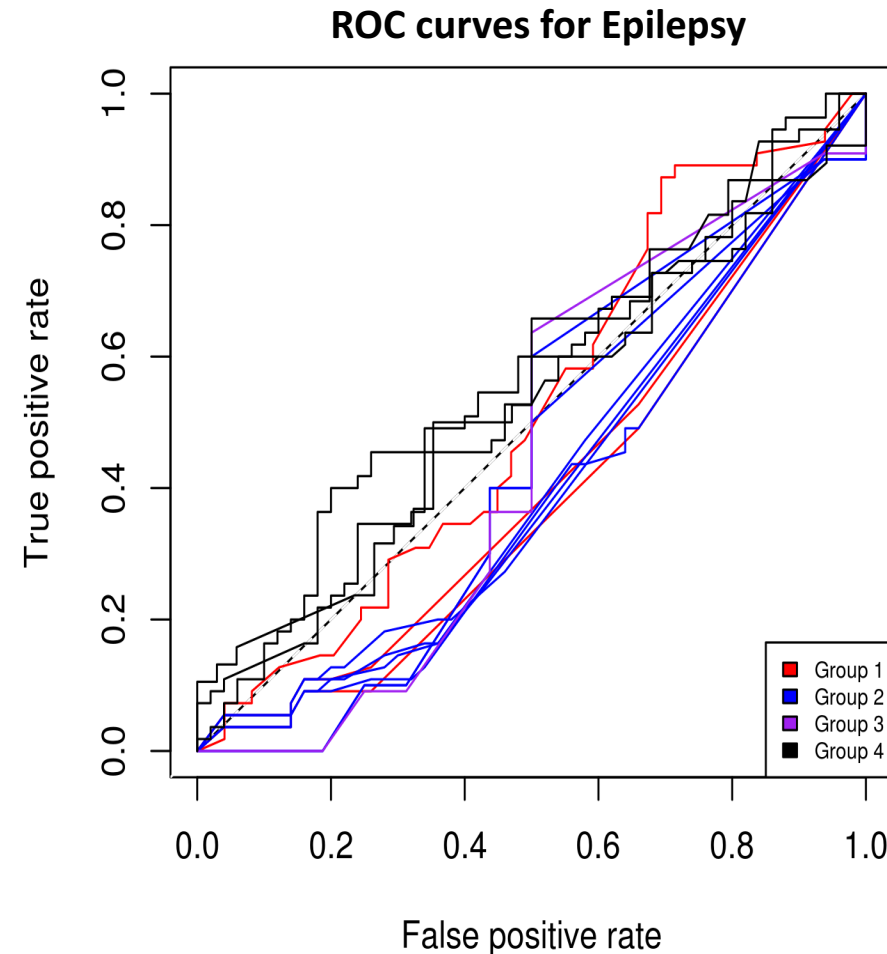


Performance assessment Epilepsy

■ Affected
■ Not affected



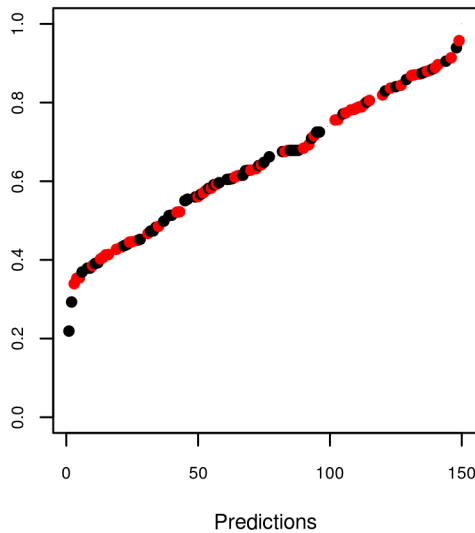
Submission	AUC	MCC
4.3	0.56	0.22
4.2	0.55	0.23
4.1	0.54	0.19



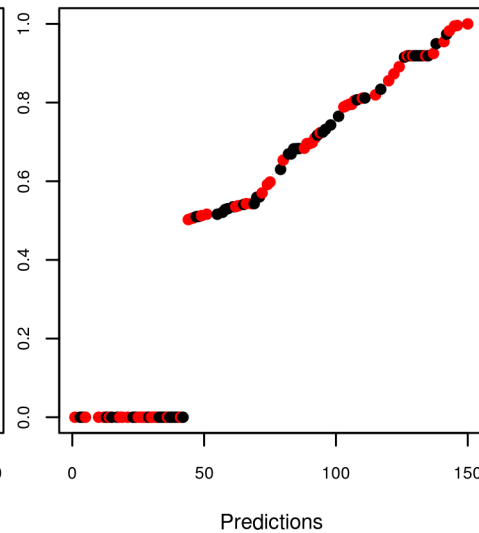
Epilepsy best performers

● Affected
● Not affected

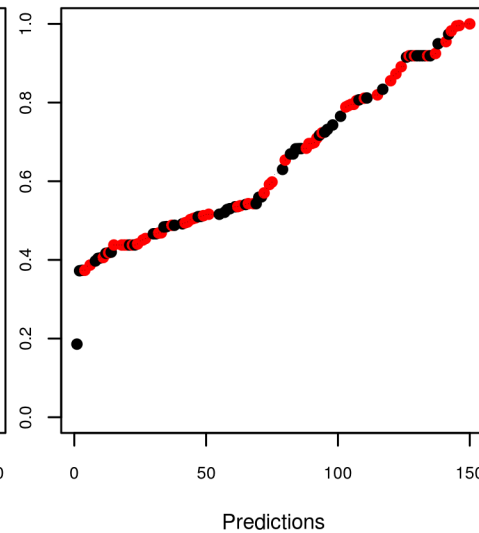
Submission 4.3



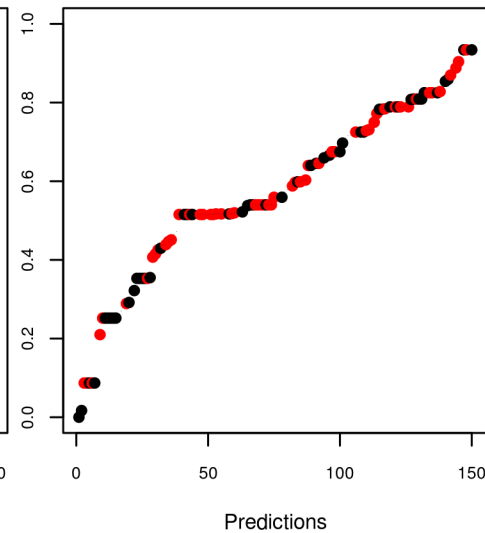
Submission 4.2



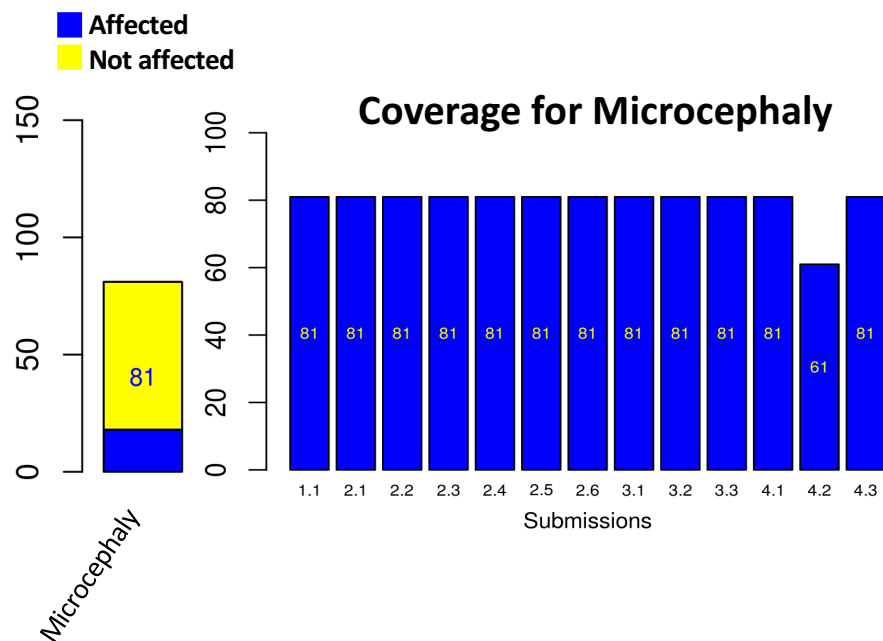
Submission 4.1



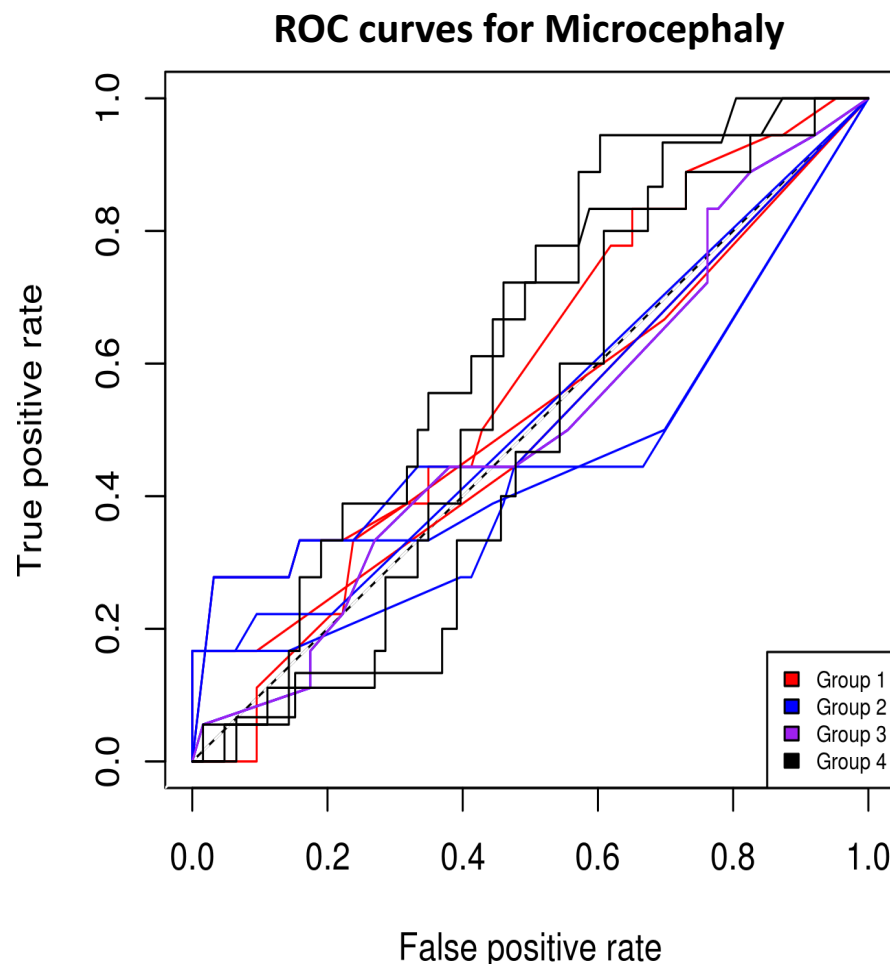
Submission 1.1



Performance assessment Microcephaly



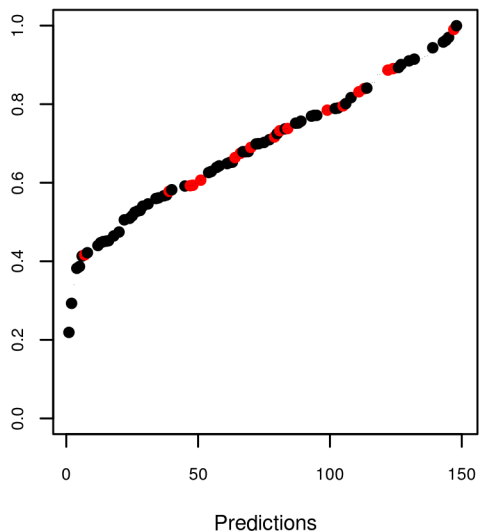
Submission	AUC	MCC	F1
4.3	0.63	0.30	0.47
1.1	0.57	0.16	0.41
4.1	0.57	0.23	0.44



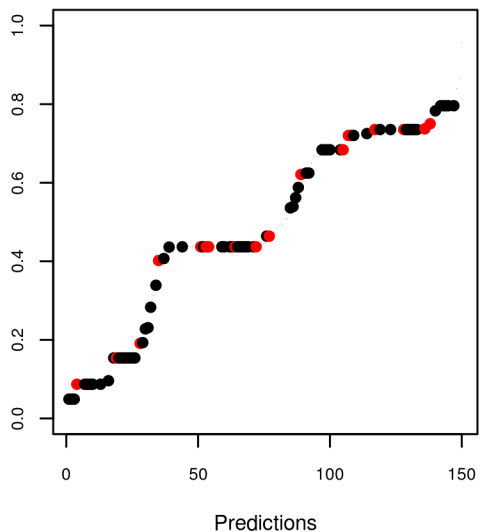
Microcephaly best performers

● Affected
● Not affected

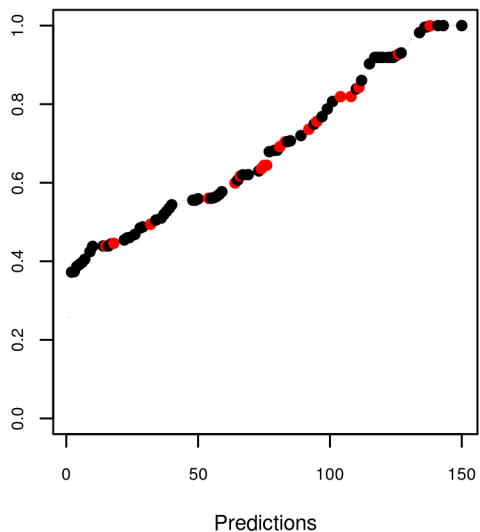
Submission 4.3



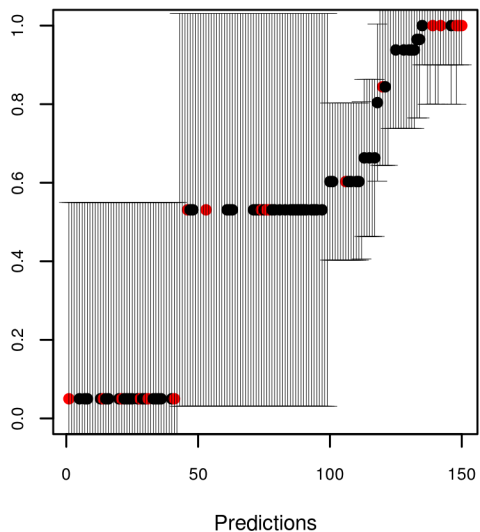
Submission 1.1



Submission 4.1

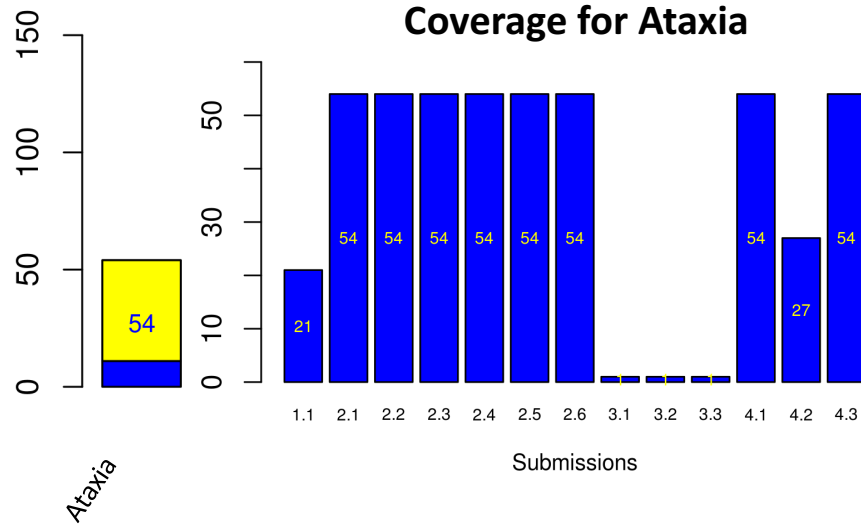


Submission 2.2

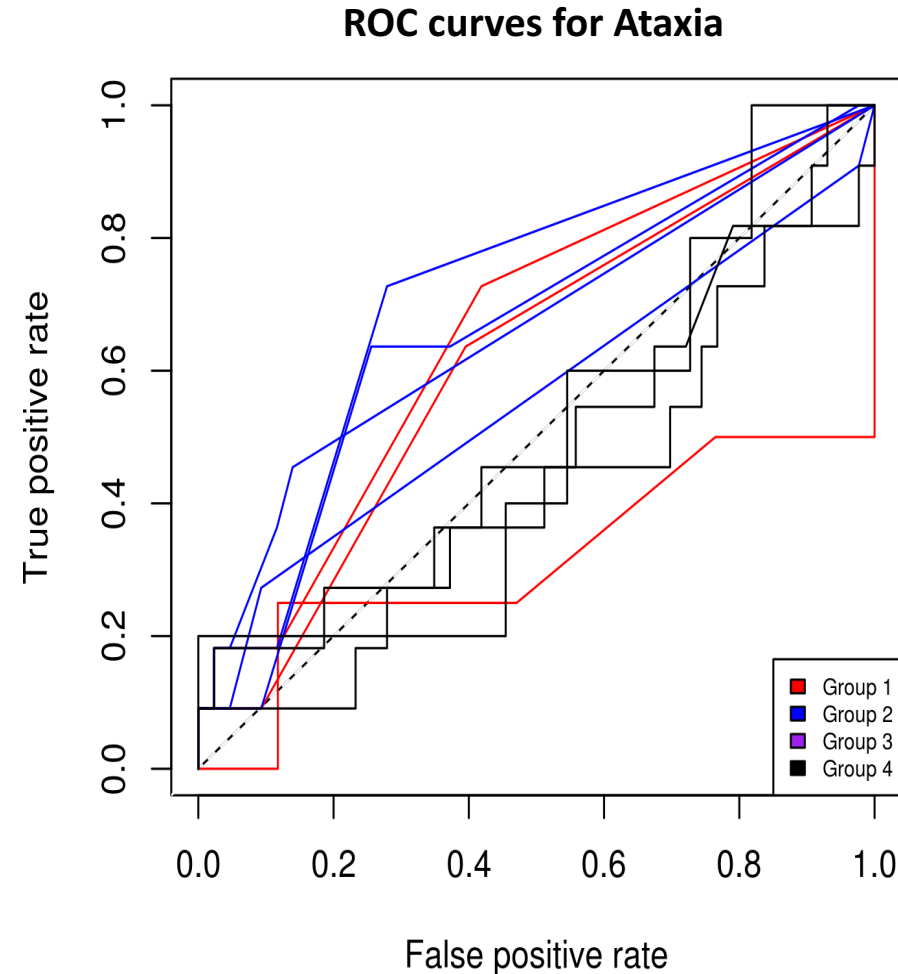


Performance assessment Ataxia

■ Affected
■ Not affected



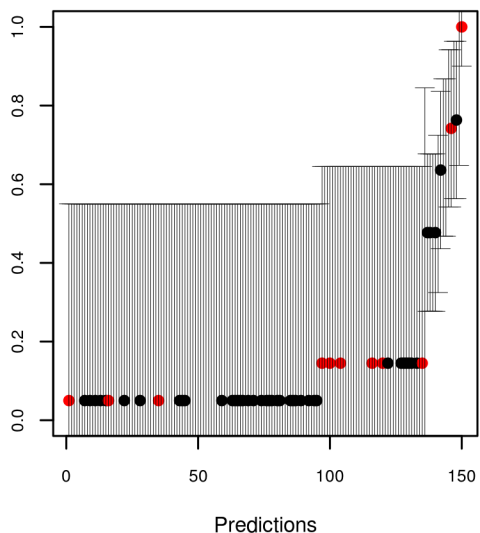
Submission	AUC	MCC	F1
2.4	0.72	0.37	0.52
2.6	0.66	0.32	0.46
2.3	0.66	0.33	0.48



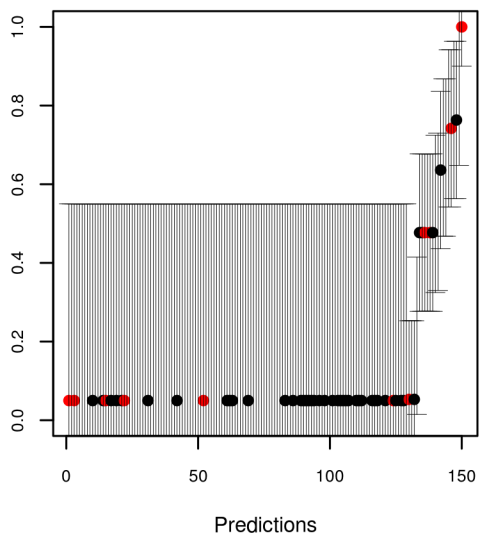
Ataxia best performers

● Affected
● Not affected

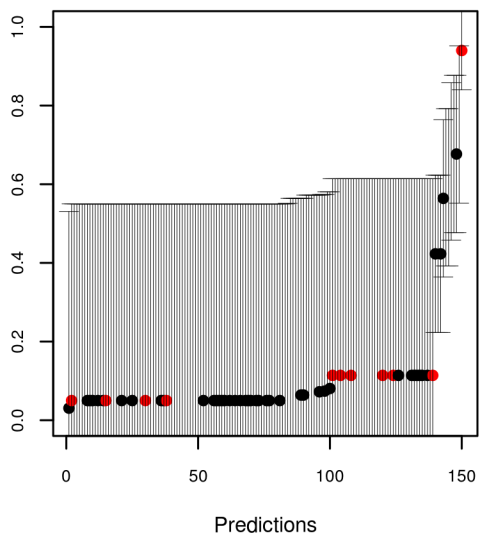
Submission 2.3



Submission 2.6



Submission 2.3



Submission 2.2

