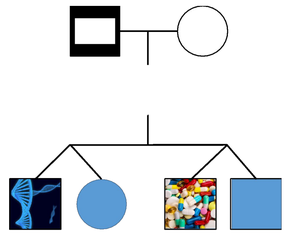
Using Methylation to predict early cancer onset in patients with LFS

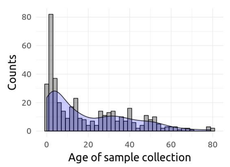
SickKids
THE HOSPITAL FOR SICK CHILDREN
Genetics & Genome Biology Program

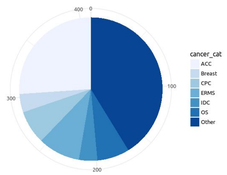
B Brew with L Erdman, T Guha, A Novokmet,
A Doria , J Berman, A Shlien, D Malkin, Anna Goldenberg



Background

Li-Fraumeni syndrome (LFS) is a hereditary cancer predisposition disorder molecularly characterized by a germline *TP53* mutation and a spectrum of early onset cancers. Consistent surveillance is essential for early detection and better survival outcomes.





Research question

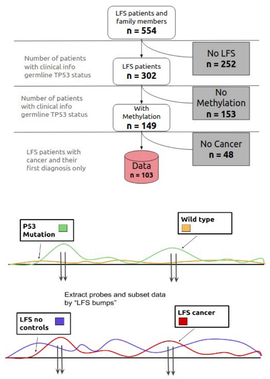
Hypothesis: Machine learning techniques applied to germline methylation profiles can accurately predict cancer onset before the age of 6 in patients with LFS. Further, similar methods detect a cancer signature directly from the methylation profiles.

Methods

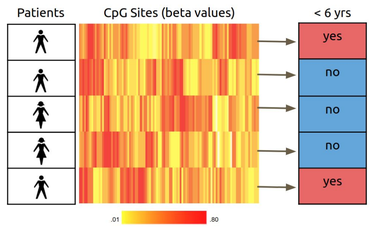
1) **Data:** Clinical and methylation data from The Hospital for Sick Children.

2) **Dimensionality reduction:** Identify "LFS bumps" and remove cancer signature.

3) **Prediction:** Predict age of onset using an elastic net model.

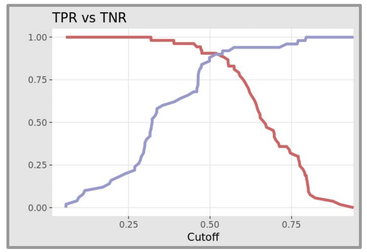


$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|y - X\beta\|^2 + \lambda_2\|\beta\|^2 + \lambda_1\|\beta\|_1).$$



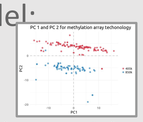
Results

- Our most accurate model has an accuracy of 91%, only missing 7% of the patients that have cancer before the age of 6 (False negatives).



Other variables controlled for in the model:

- Gender
- Technology
- Family members with cancer.



Classification at 6 years		Observed age of onset: Before age 6	Observed age of onset: After age 6
Predictions	Before age 6	.93 (n=50)	.11 (n=5)
	After age 6	.07 (n=4)	.89 (n=44)

Predicted to be earlier

Not detected

Detecting cancer directly		Cancer status: Cancer	Cancer status: No Cancer
Predicted cancer status	Cancer	.90 (n = 93)	.14 (n = 6)
	No Cancer	.10 (n = 10)	.86 (n = 35)

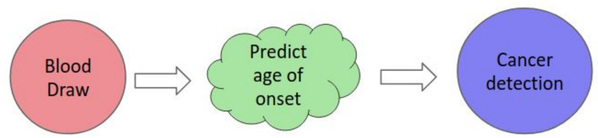
False positive

False negatives

- Our model that detects cancer directly, i.e. applied to the patients that are at risk of getting cancer in the near future (or might already have cancer) - has an accuracy of 88%.

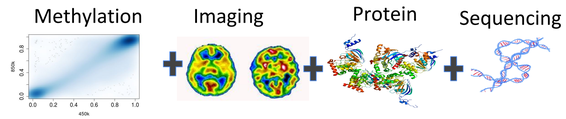
Conclusions

We have devised a two-step strategy which estimates (1) the risk of cancer onset before the age of 6 and (2) a follow up diagnostic tool to detect cancer directly from methylation.



Future Work

Future work will focus on not only leveraging more methylation data, but also incorporating other data types into our analysis. Additional data will allow us to add more sophistication to the model while simultaneously increasing it's statistical power.



Acknowledgements

The Terry Fox Research Institute
L'Institut de recherche Terry Fox

This work is supported by a New Frontiers Program Project Grant from the Terry Fox Research Institute