

St Paul Shenanigans

Or a deep look into cytokines



[BigData2018ubc-stPaul](https://github.com/BigData2018ubc-stPaul)

In Canada

1 of 18

Deaths are due to sepsis

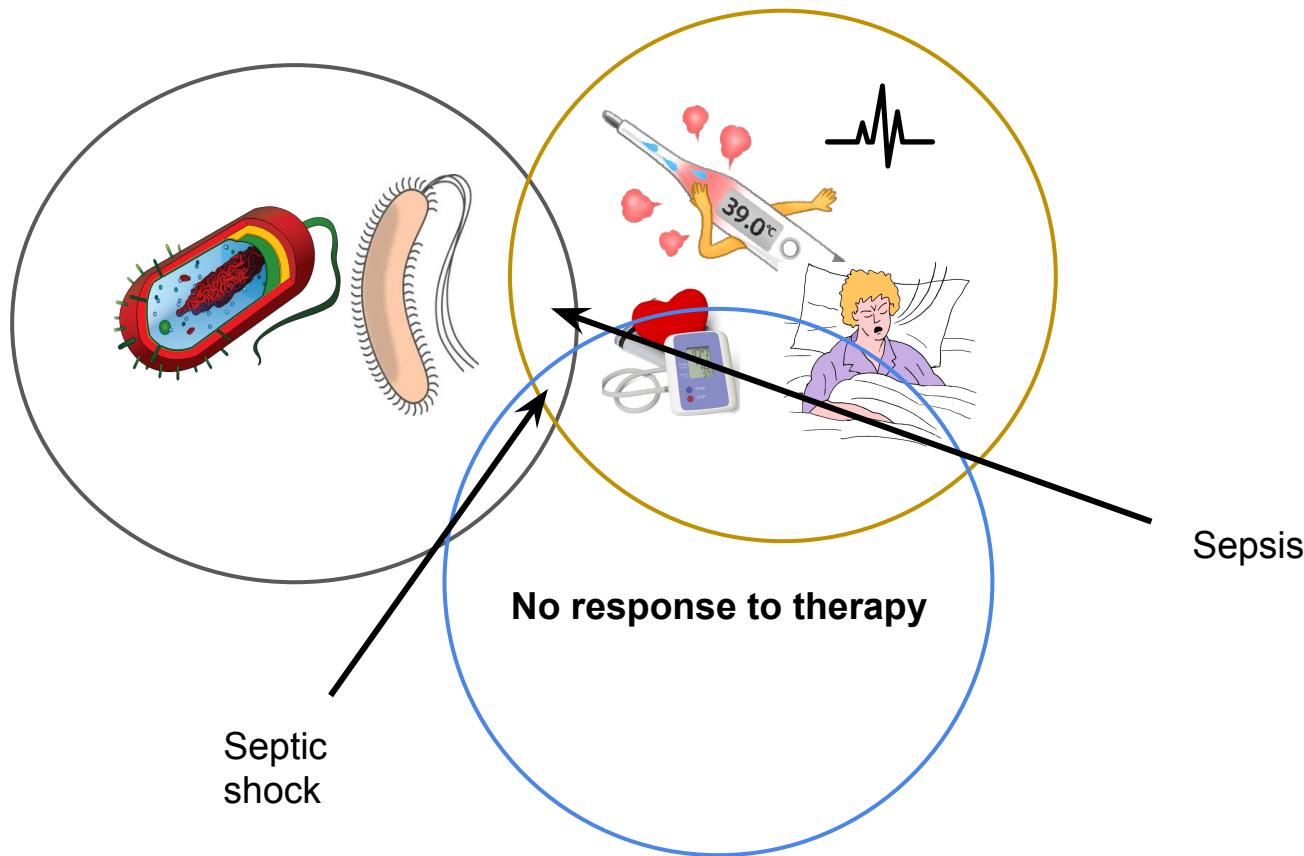
33 %

Deaths within 1st month

[Navaneelan et al., Health at a Glance, Statistics Canada, 2016]



Septic shock: any cytokines involved?



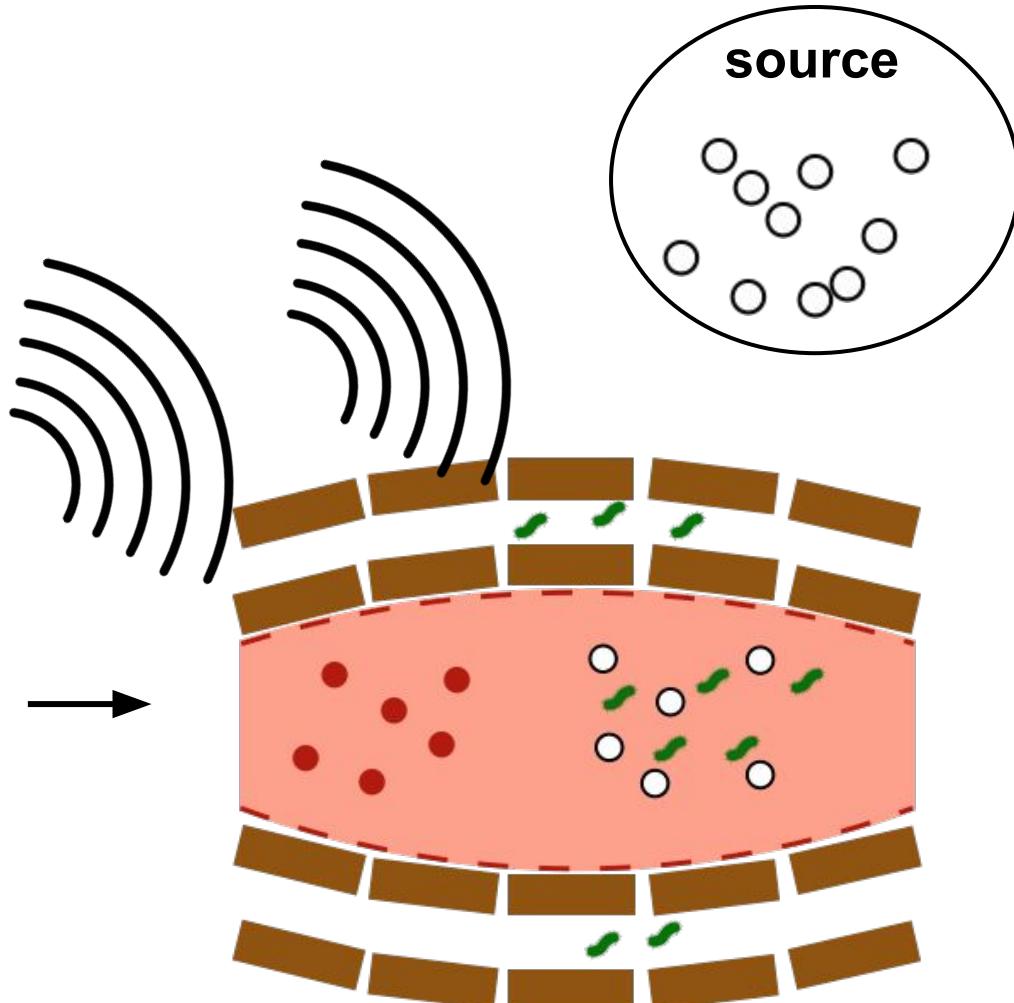
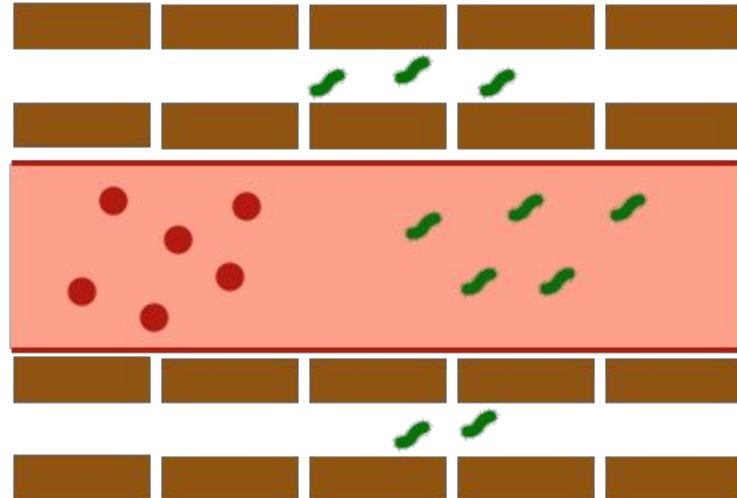
What is Septic shock?

● Red blood cell

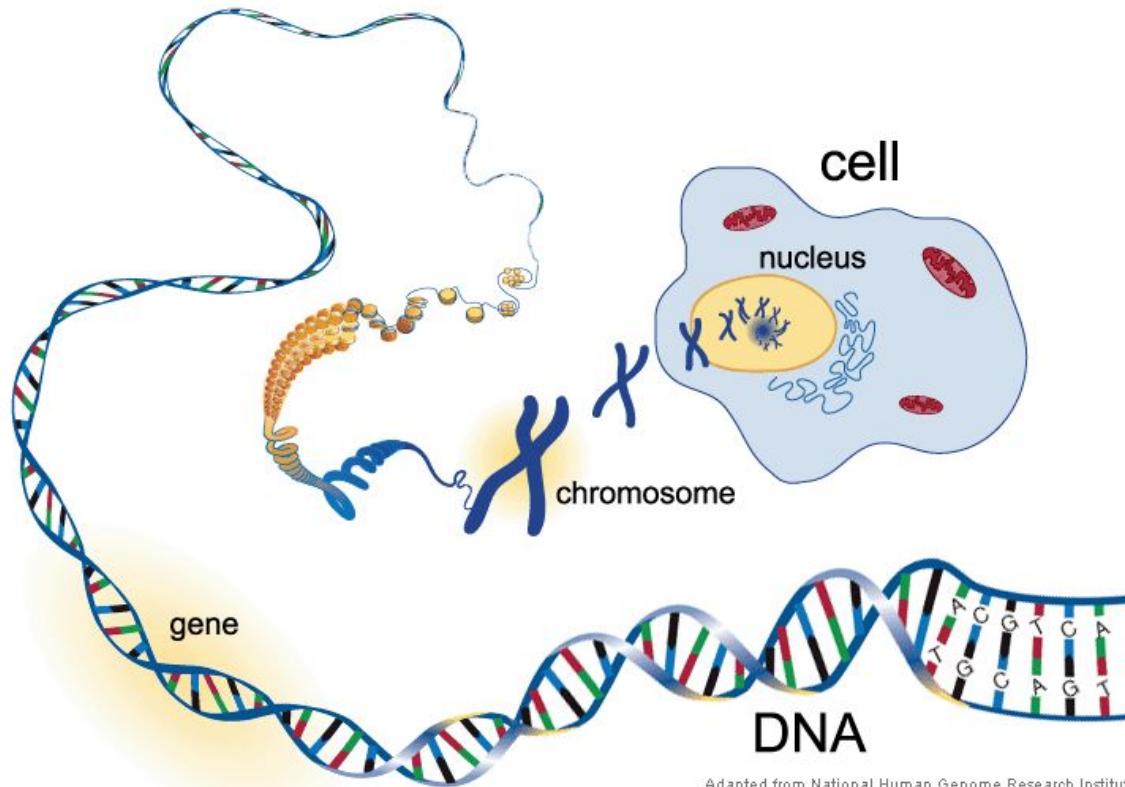
■ Tissue cell

● Infectious cell

○ White blood cell

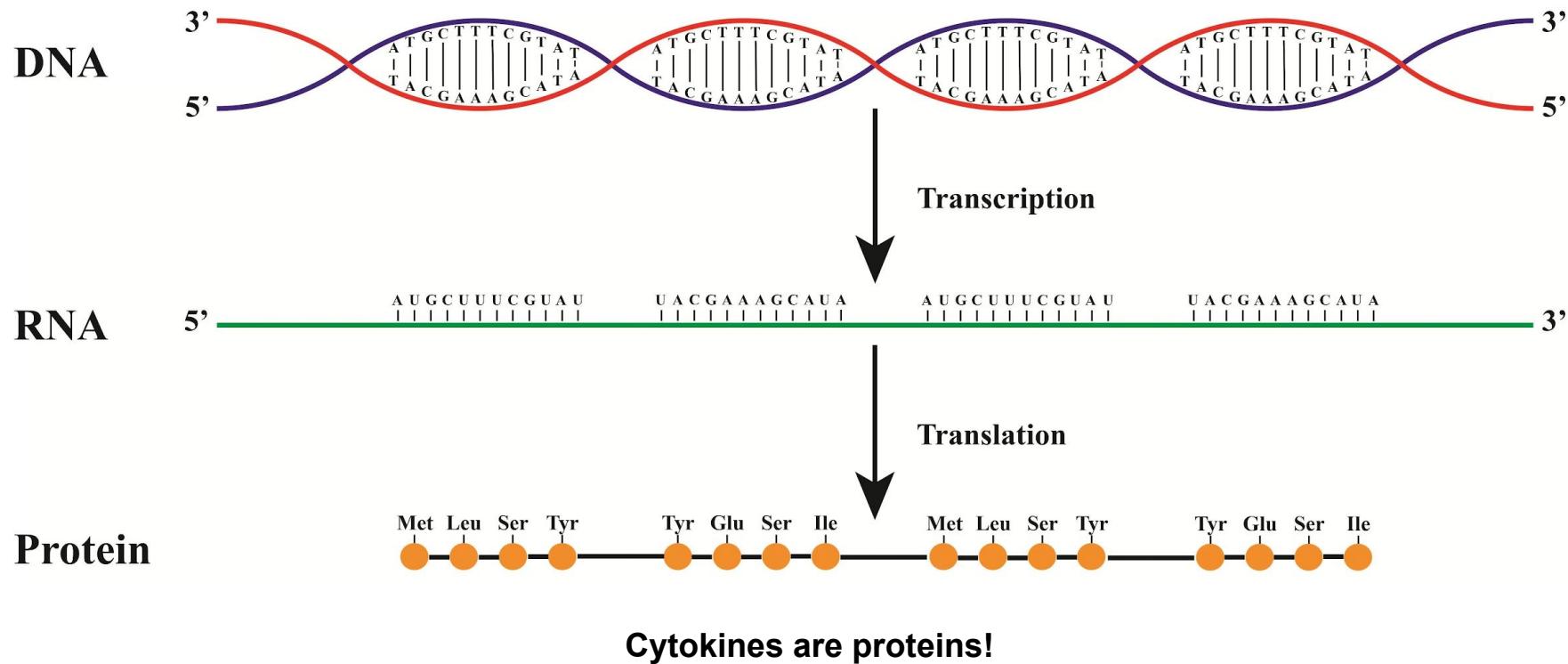


All the information is in the DNA

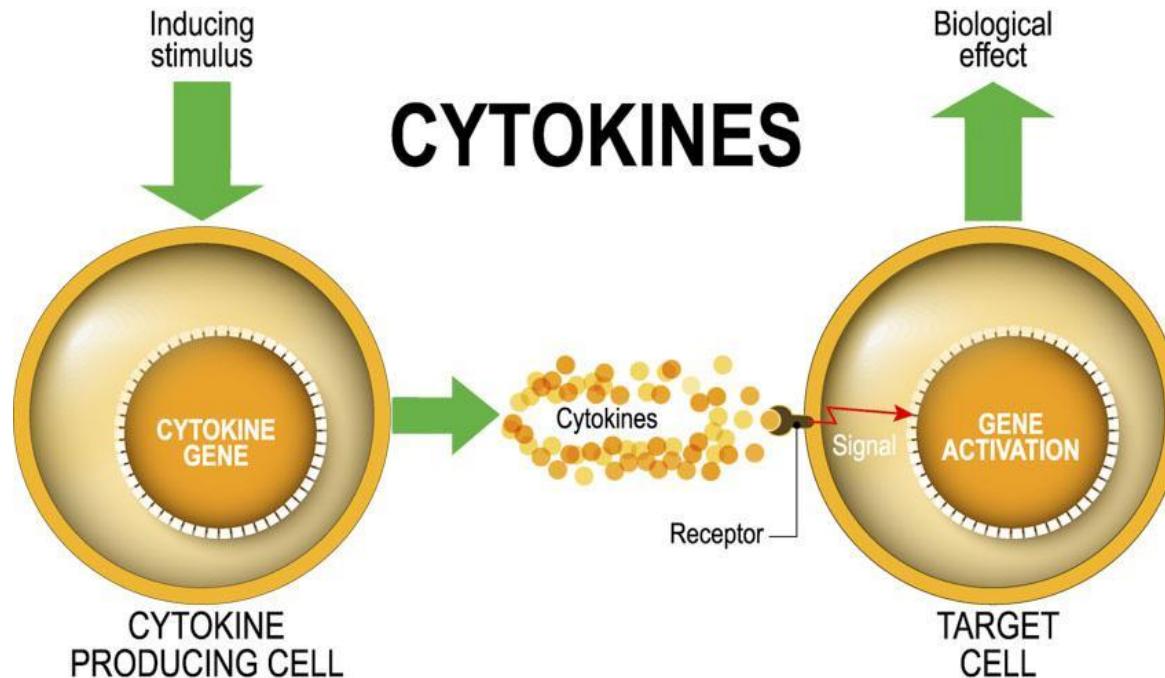


Adapted from National Human Genome Research Institute

Central Dogma (DNA → RNA → Protein)

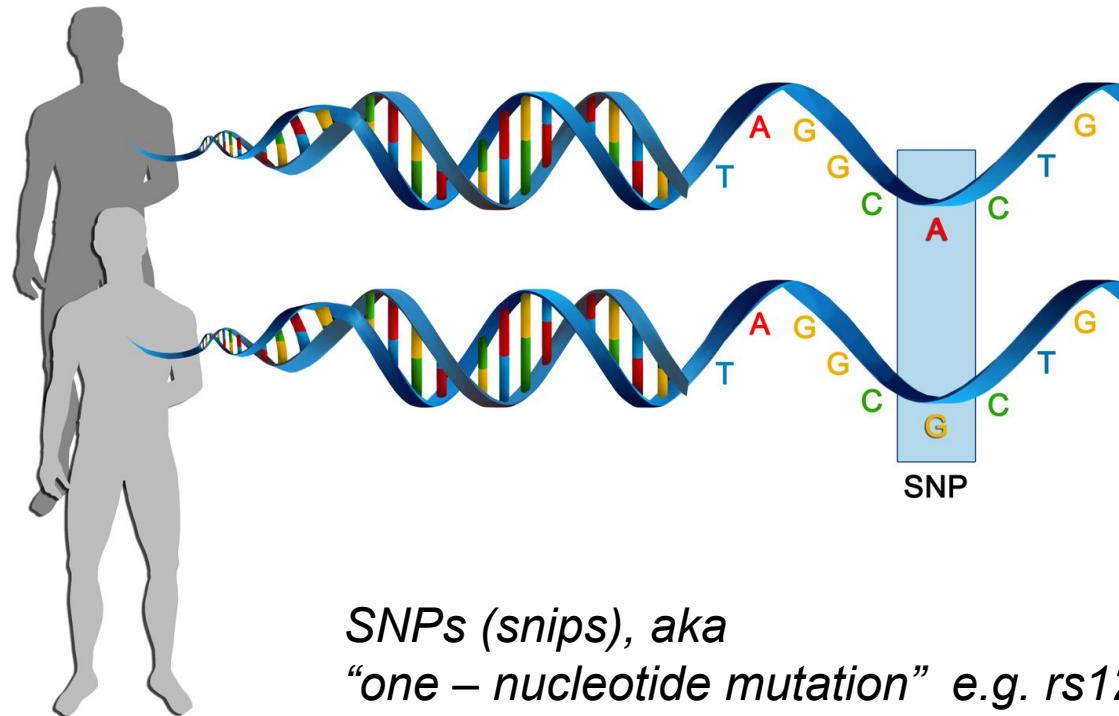


What are cytokines?



Cytokines - are small proteins that mediate cell signalling e.g IL1, MIP1A

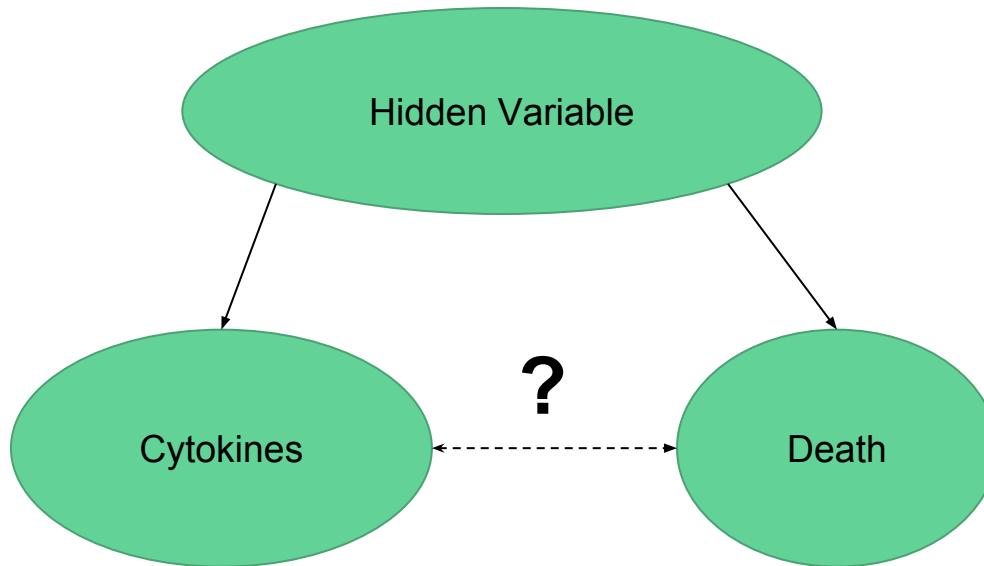
Single Nucleotide Polymorphisms (SNPs)



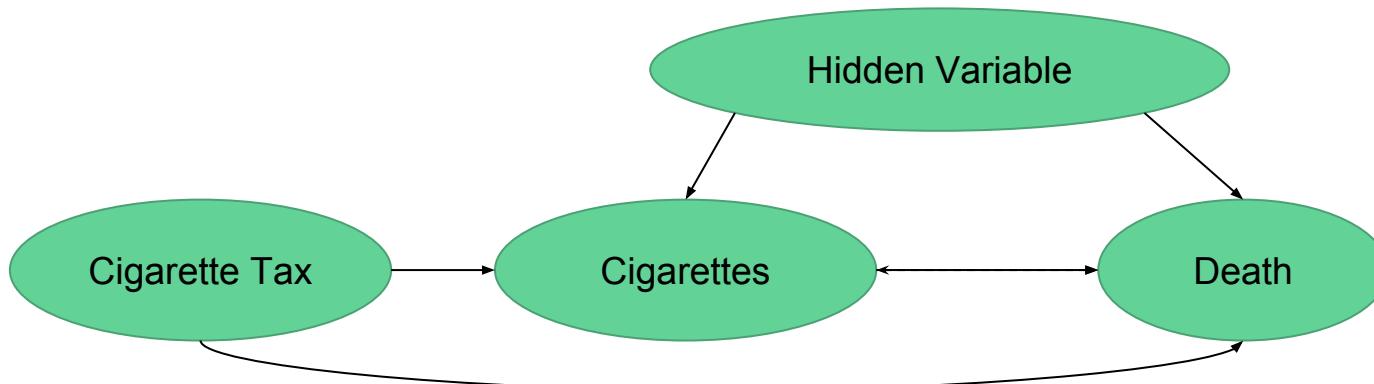
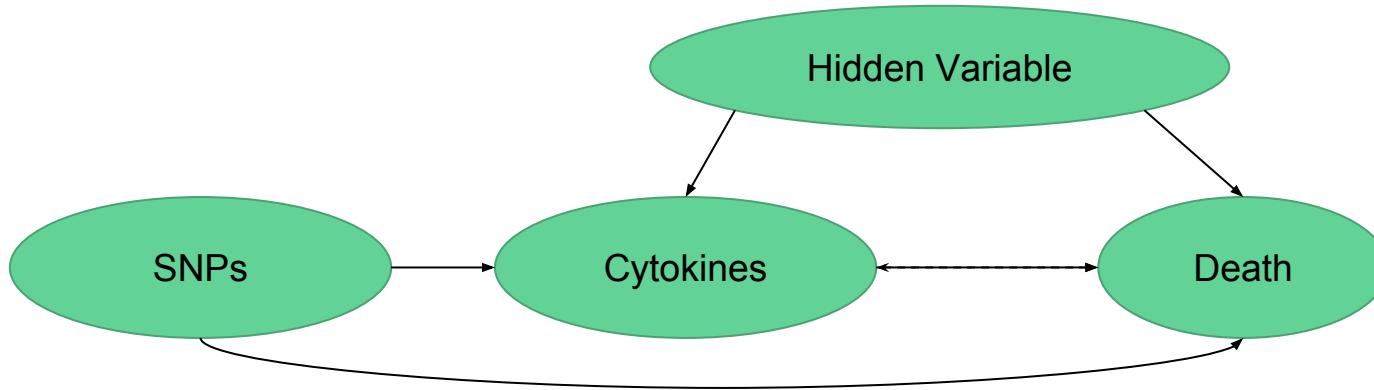
*SNPs (snips), aka
“one – nucleotide mutation” e.g. rs123456*

Certain SNPs can modify cytokine levels.

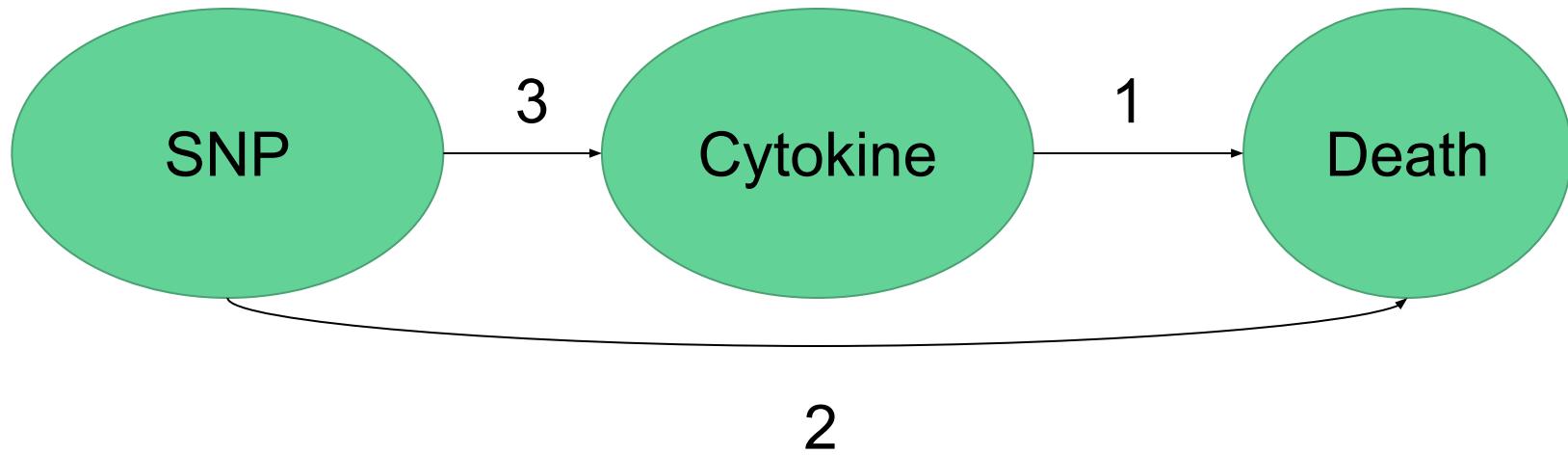
Correlation does not imply causation



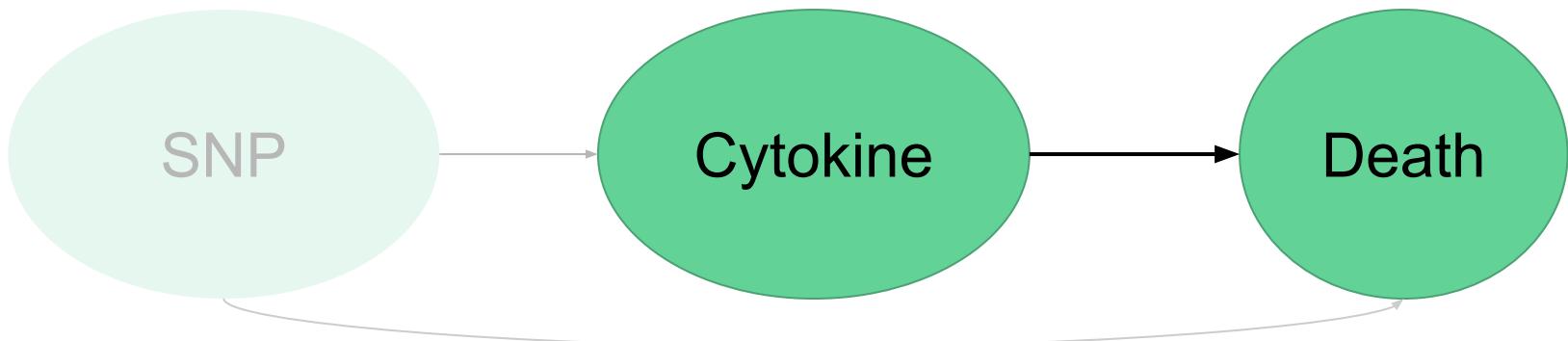
SNPs can help reveal causal relationships



Three points of investigation



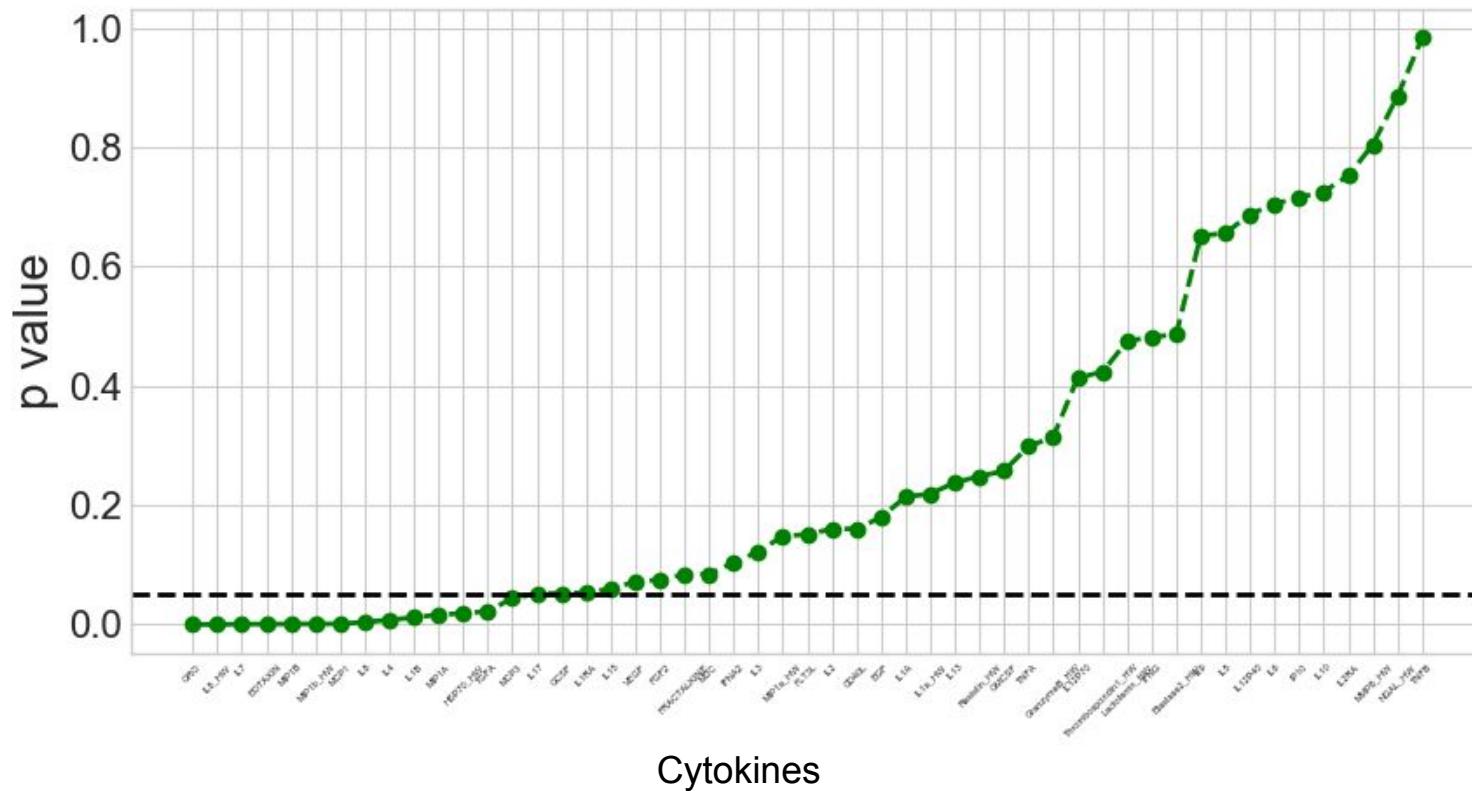
1) Cytokine level → Death



Our Dataset

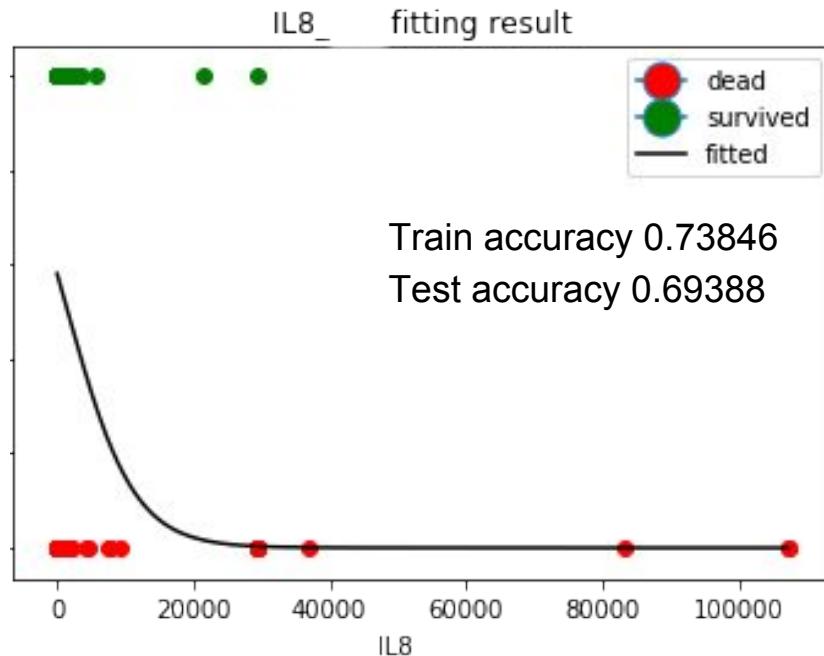
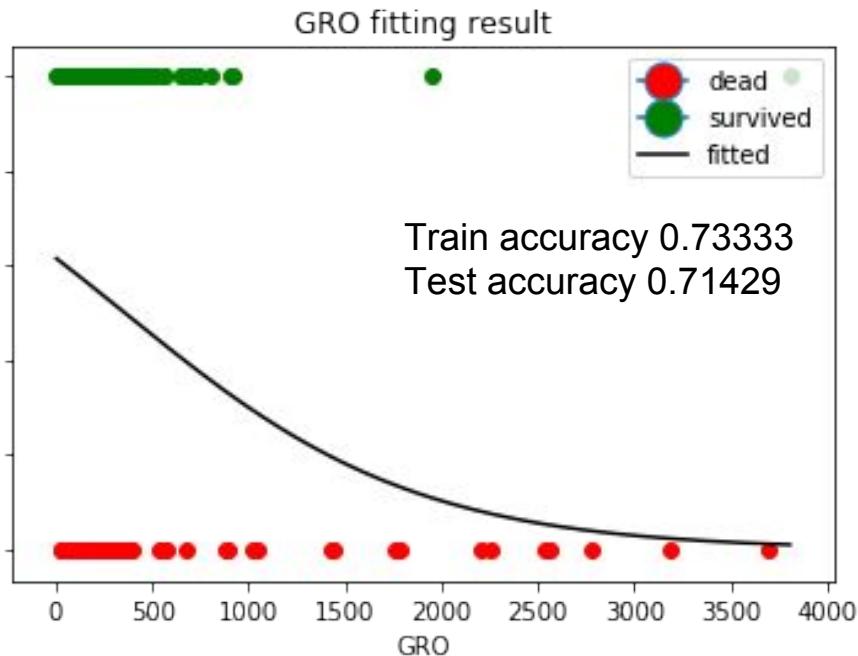
- Data originated from VASST study to determine effects of a certain vasopressin treatment on health outcomes of septic patients
- Type of data arising from 330 septic patients:
 - Blood serum concentration measurements of 52 cytokines per patient
 - Approx. 1.2 million SNP measurements for each patient

Finding “culprit” cytokines



Logistic regression

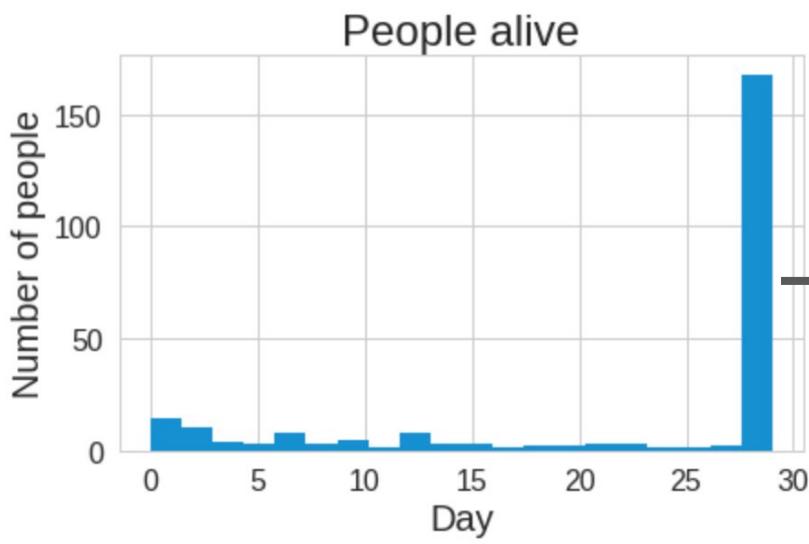
- Single variable
 - L1 loss, higher weight for death cases, inv regularizer=0.5
 - Error usually one-sided:



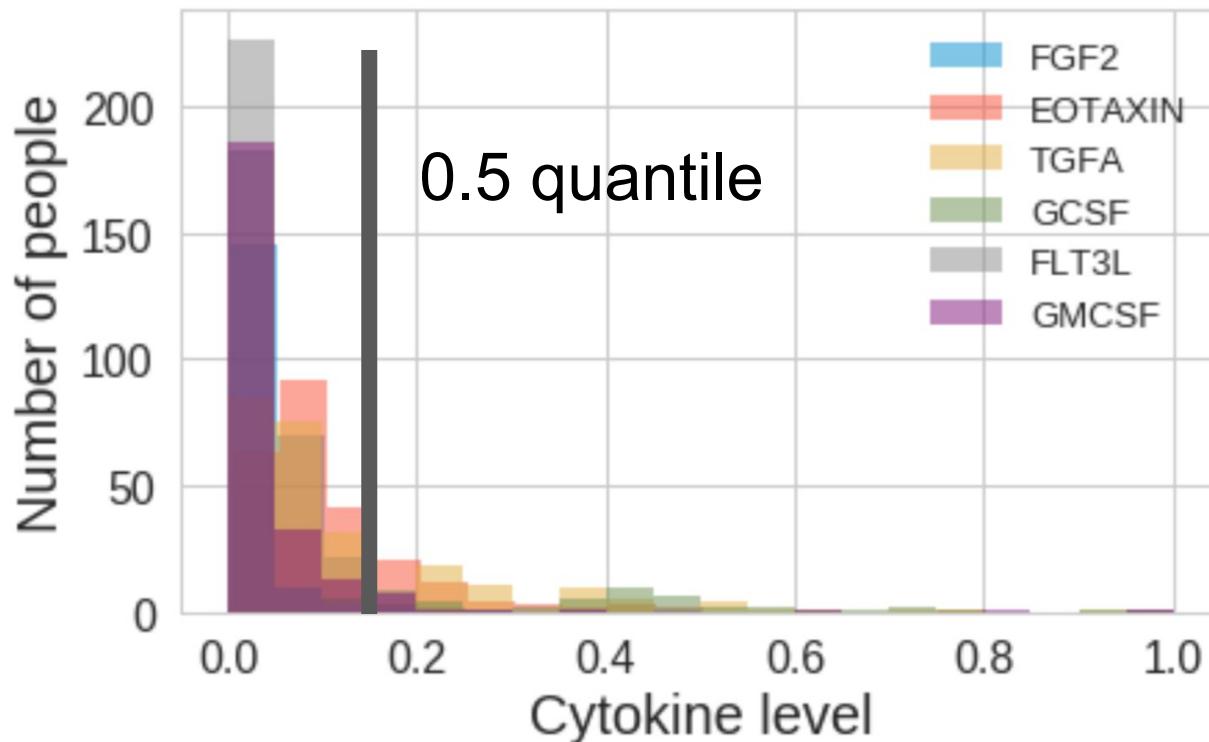
Other models

- Multivariable logistic regression:
 - Using only “useful cytokines”: Train accuracy 0.70588, Test accuracy 0.67568
 - Using all cytokines: Train accuracy 0.75000, Test accuracy 0.64000
- Decision tree (depth 1 to 15):
 - Best: depth=5, Train accuracy is 0.78636, Test accuracy is 0.68000
- Neural network:
 - Various configurations, train accuracy usually high, test accuracy <0.7
- Challenge: overfitting due to small dataset, noise

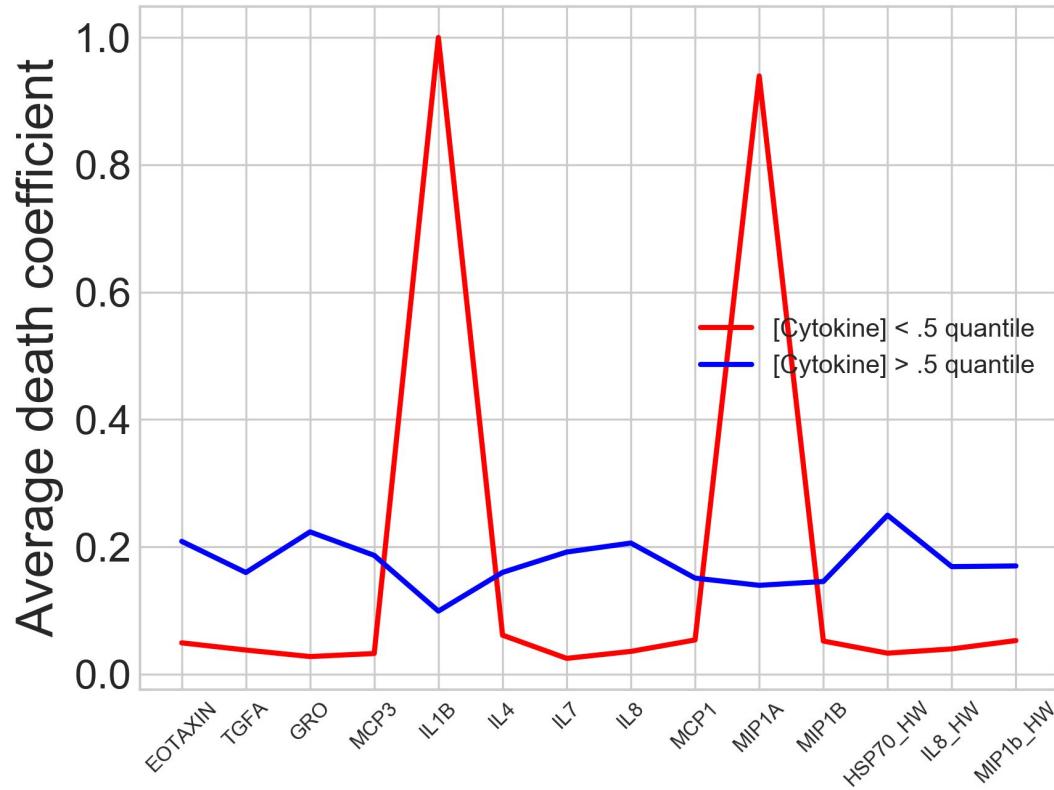
People dying (people that live this many days)



Cytokine levels among patients



Cytokine - Death correlation



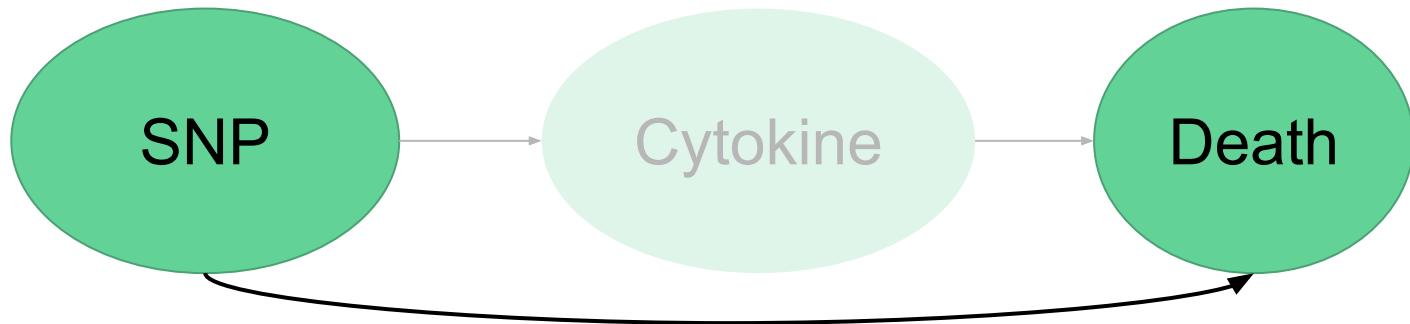
Culprits:



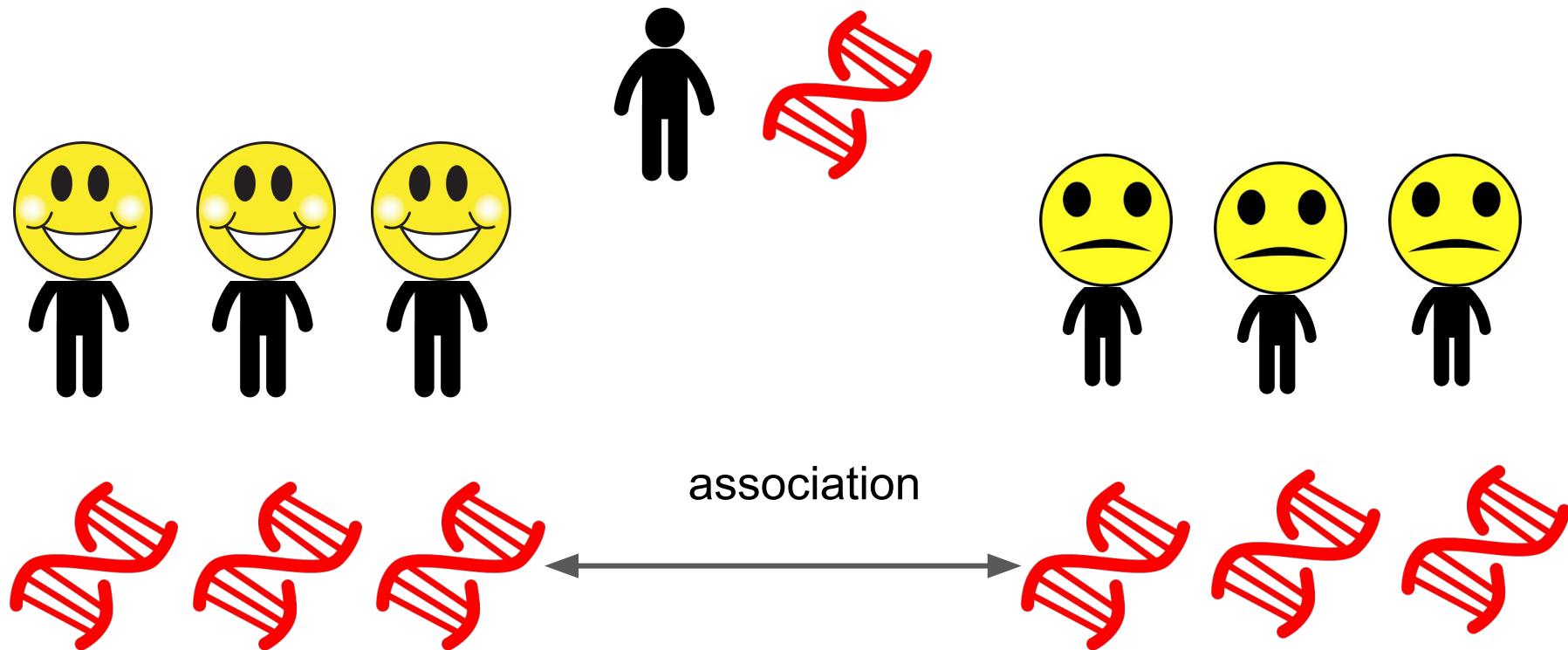
Special culprits:



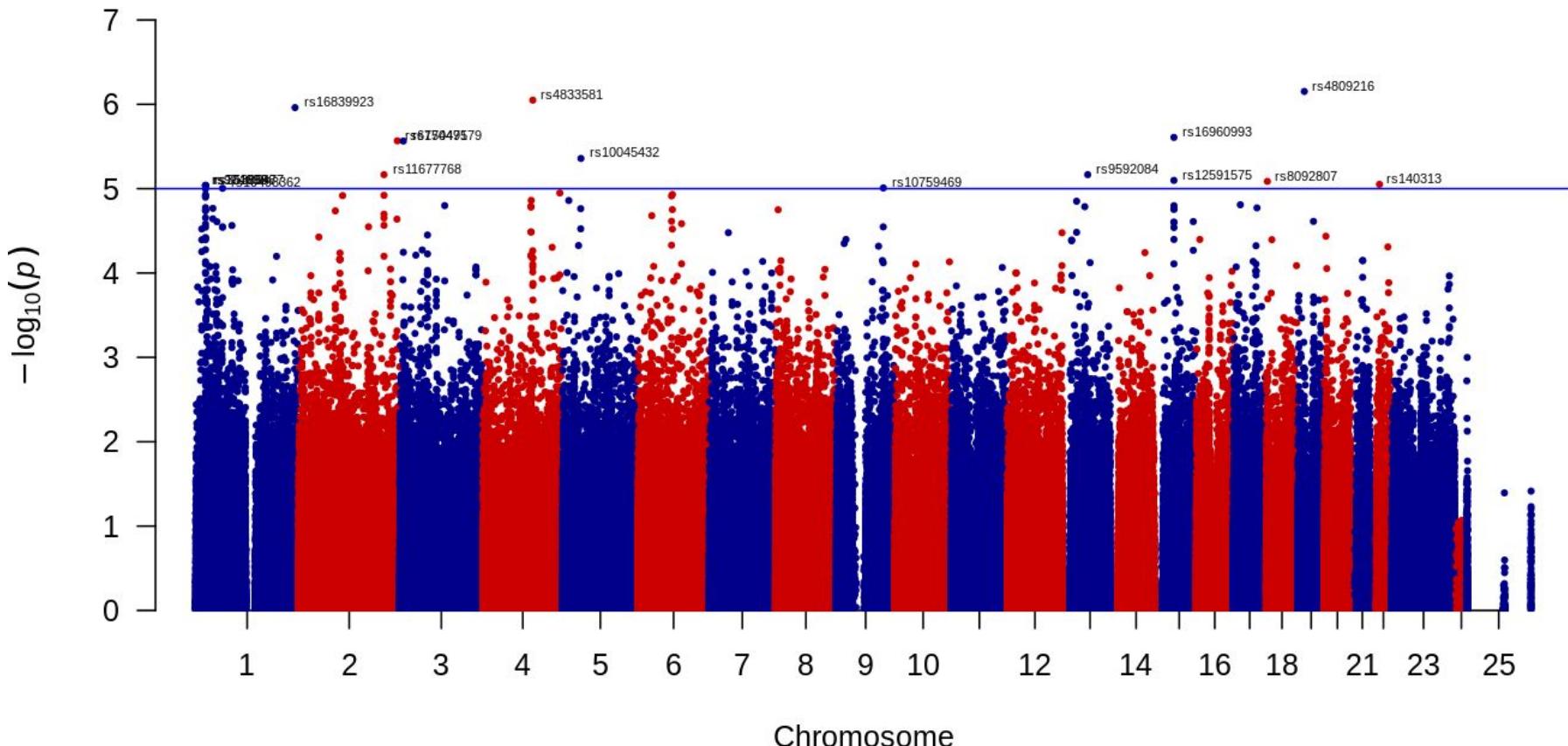
2) SNPs → Death



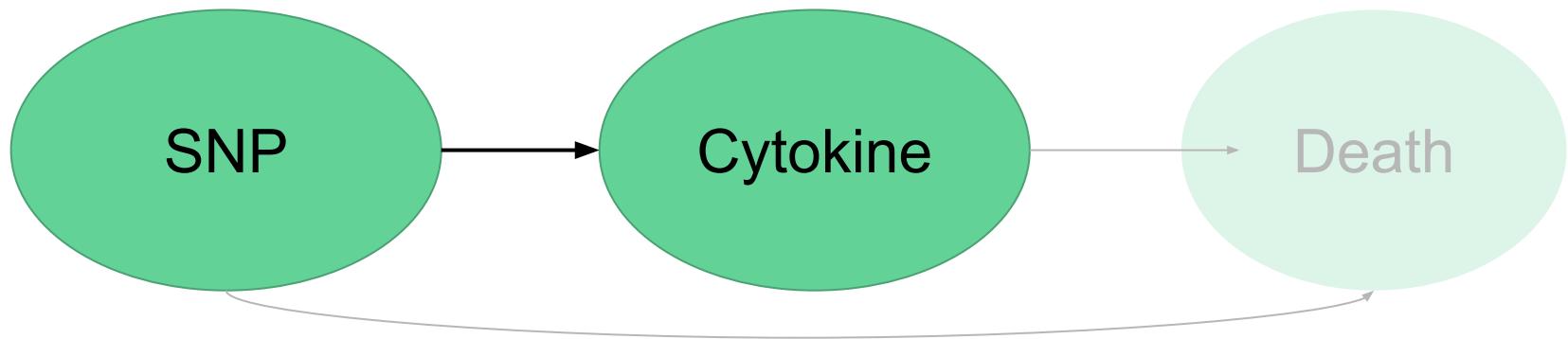
Genome Wide Association Study (GWAS)



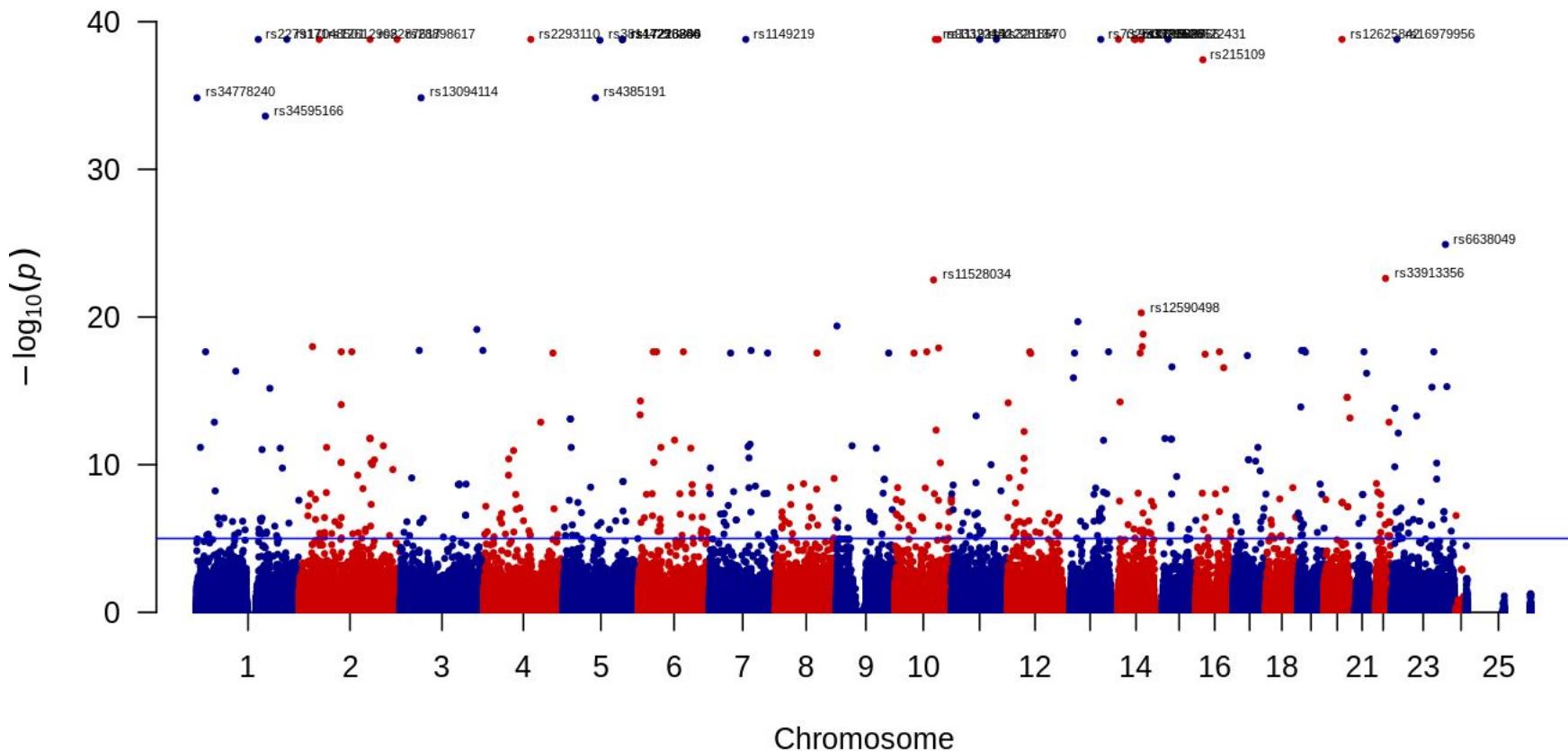
GWAS: SNP association with survival



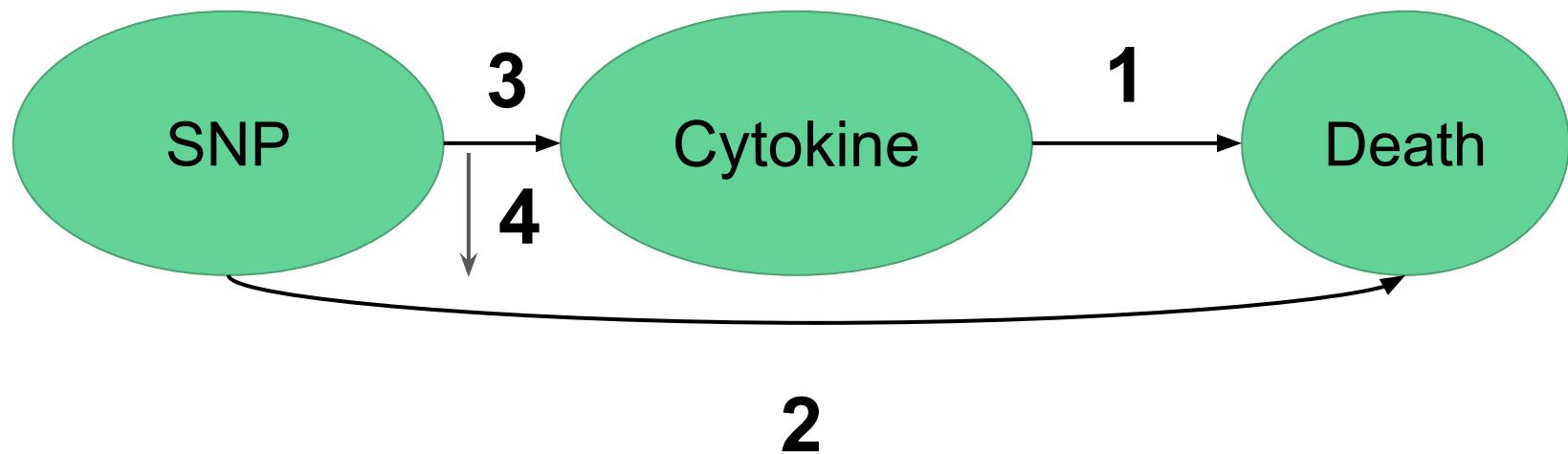
3) SNPs → Cytokine level



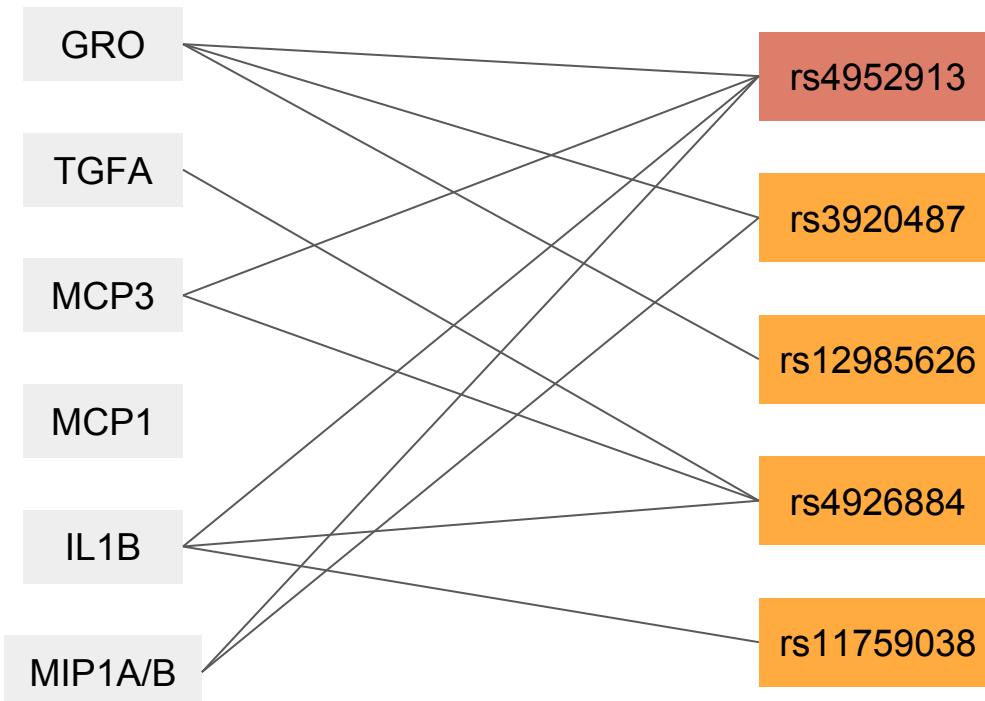
GWAS: SNP association with cytokine MCP3



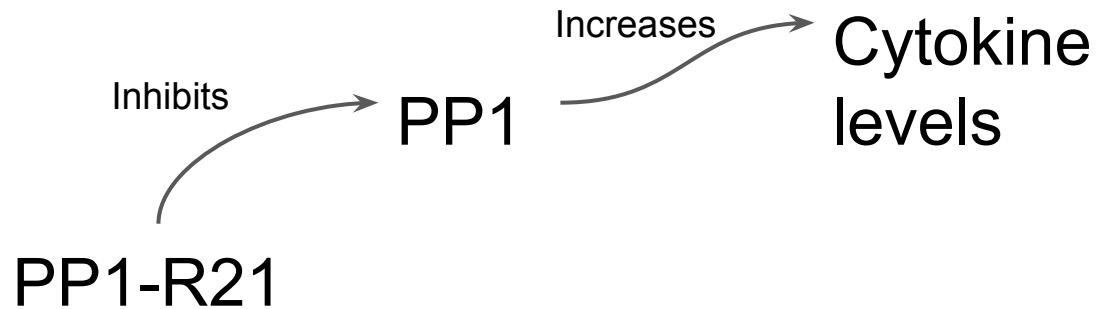
4) Assembling results



Correlating significant Cytokines with SNPs



Protein phosphatase 1 regulatory subunit 21 (PP1R21)

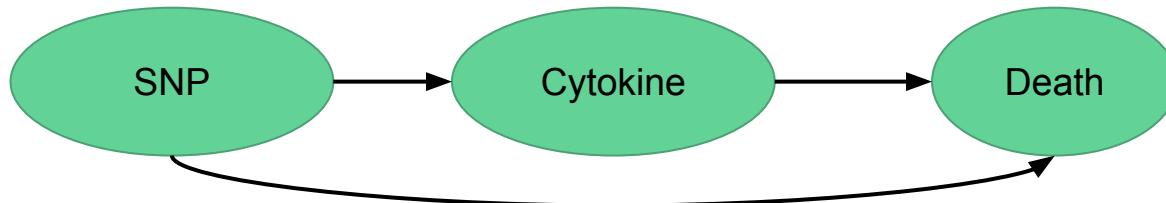


- SNP Rs4952913
- SNP present in:
 - GRO, MCP3, IL1B, MIP1A, MIP1B
 - All are proinflammatory



[Yamaoka *et al.*, Protein phosphatase 1 is involved in IL-2-induced IL-5 and IL-13 expression in human T cells, *Genes to Cells*, 2012]

Concluding Remarks



Evidence supporting: several cytokines cause death/survival through PP1 PP1R21 mechanism

Next steps:

- More data for better training the machine learning
- Deep dive into investigation of PP1 and PP1R21
- Connection with other inflammatory diseases

Acknowledgement

Keith Walley (St. Paul's Hospital)

Brian Wetton

Aaron Berk

India Heisz

Matteo Lepur

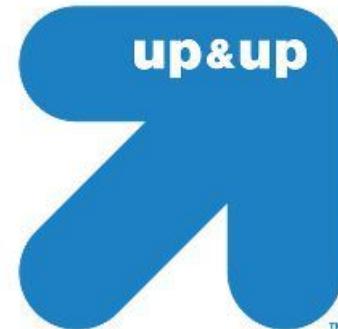


The shenanigans:



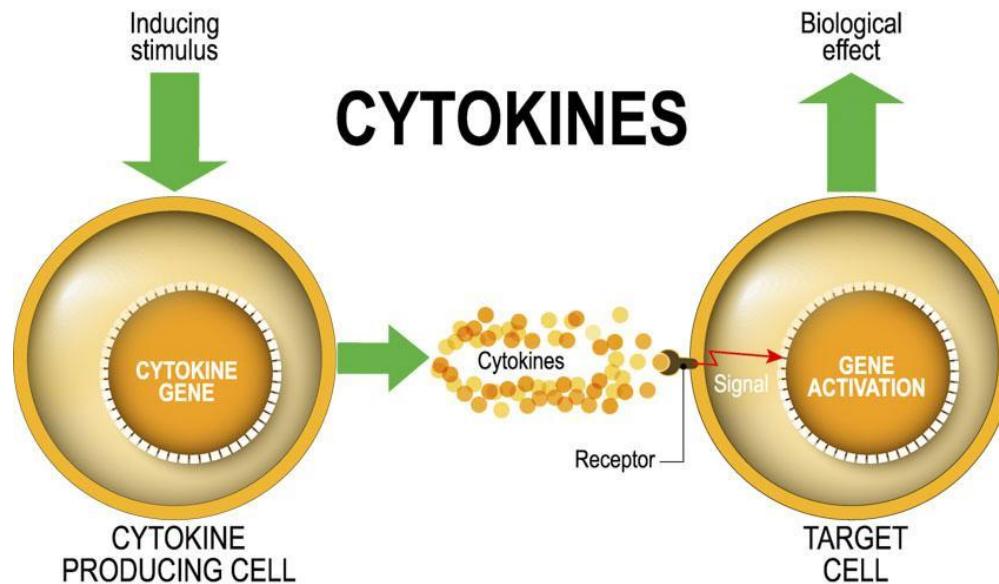
Left to right: Mingfeng Qiu, Matthew Nguyen, Bowen Chen, Sean La, Yue Liu, Babak Nasouri, Ka Mun Nip, and Vasilii Triandafilidi

BACKUP



What is cytokine?

- Small molecular weight proteins (~5-20 kDa) that mediates cell signaling
- Secreted by cells in response to stimuli (eg. inflammation and infection)
- Bind to receptors of target cells and alter their gene expression



Inflammatory Mechanisms



Infection: Viral, Fungal, Bacterial, Parasitic

Tissue Injury



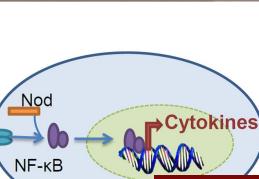
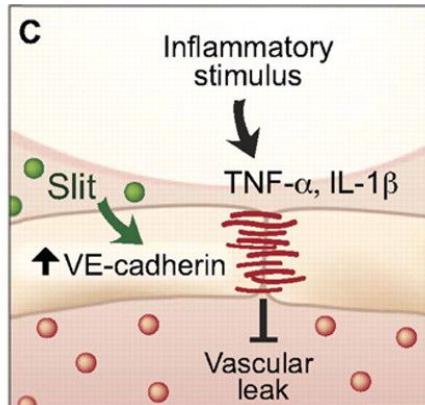
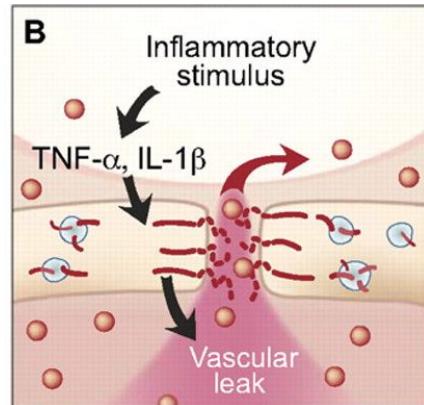
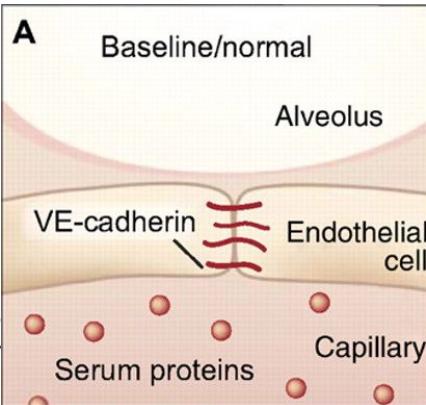
Trauma, Burn, Pancreatitis, Surgery

Pathogen-associated Molecular Pattern (PAMPs)

PAMPs

Damage-associated Molecular Pattern (DAMPs)

DAMPs



Pattern Recognition Receptors (PRRs)

Cytokines

SEPSIS SYMPTOMS

S	E	P	S	I	S
SHIVERING, FEVER, OR VERY COLD	EXTREME PAIN OR DISCOMFORT	PALE OR DISCOLORED SKIN	SLEEPY, DIFFICULT TO ROUSE, CONFUSED	"FEEL LIKE I MIGHT DIE"	SHORT OF BREATH

Inflammatory Cascade
"Cytokine Storm"
Tissue/Organ Damage

BACTERIA OR VIRUS LEAKS FROM BLOOD VESSELS INTO SURROUNDING TISSUE

RED BLOOD CELLS

BLOOD VESSEL

GOTO MAJOR ORGANS AND PERIPHERAL SYSTEMS GO INTO OVERDRIVE
BODY INFLAMMATION, SEPTIC SHOCK, ORGAN FAILURE

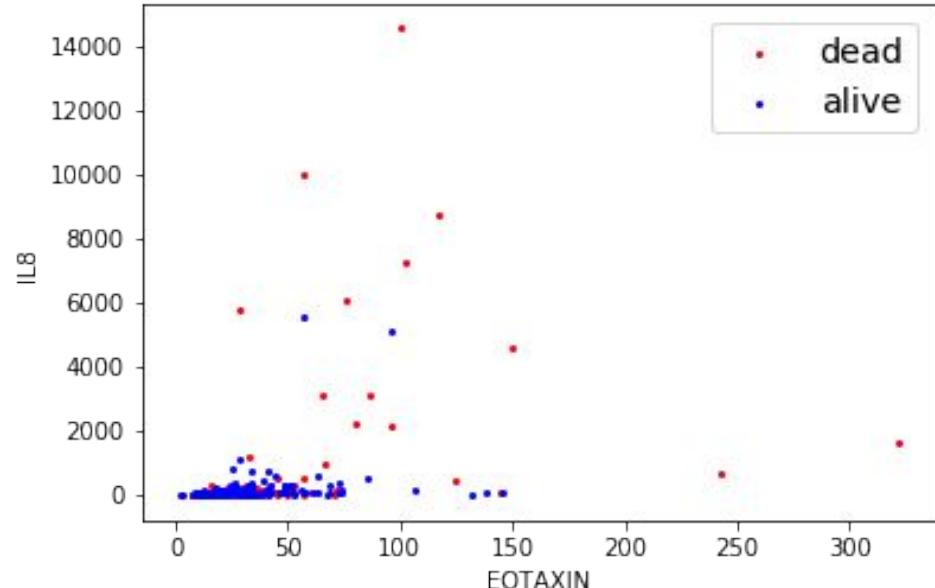
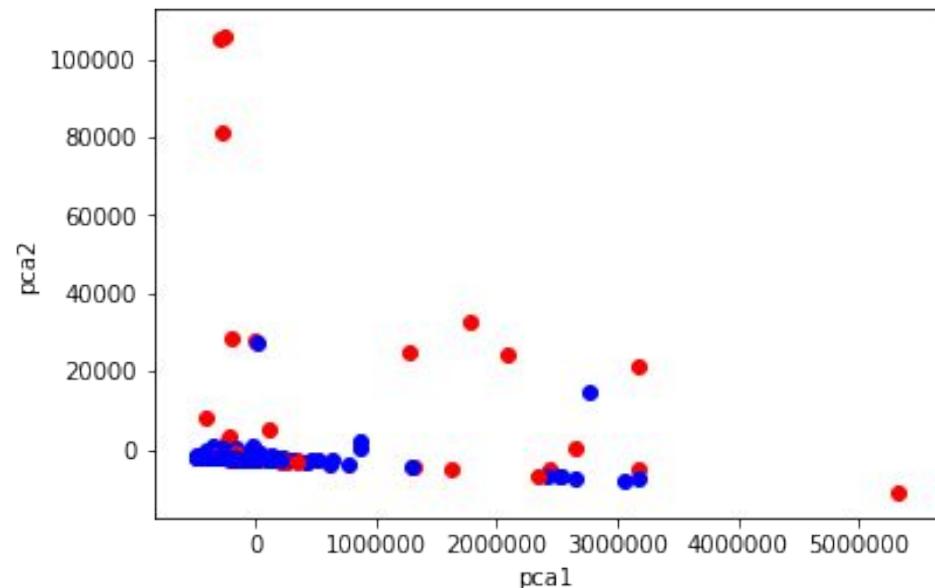


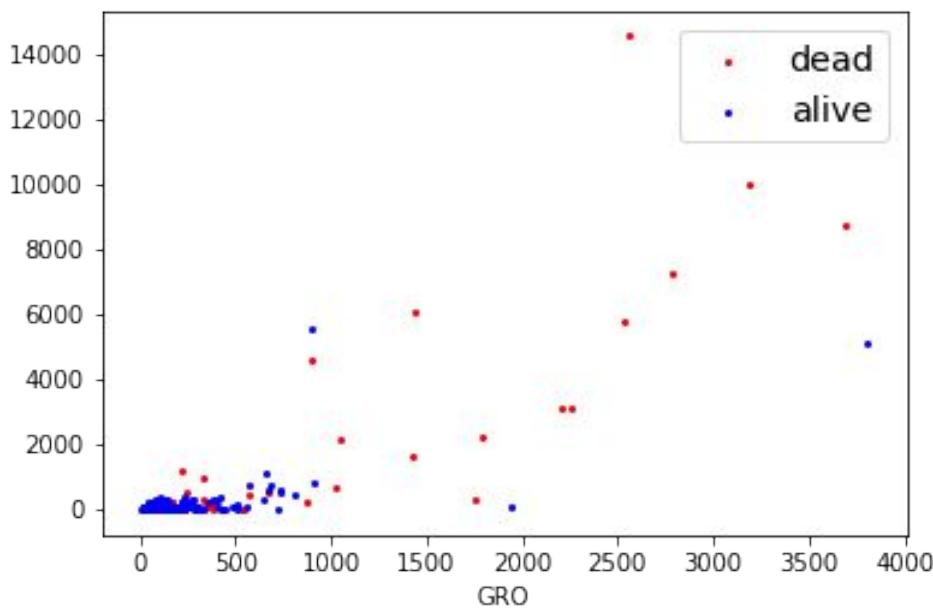
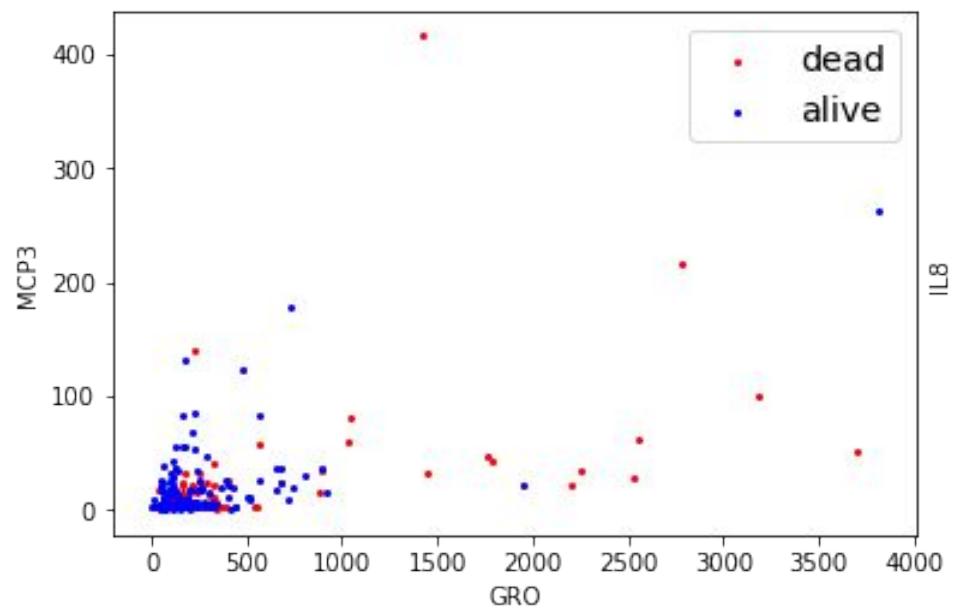
S	E	P	S	I	S
Shivering, fever, or very cold	Extreme pain or general discomfort	Pale or discolored skin	Sleepy, difficult to wake	"I feel like I might die"	Short of breath

2-dimensional visualization

Red = dead, blue = survived

Clustering did not yield good results



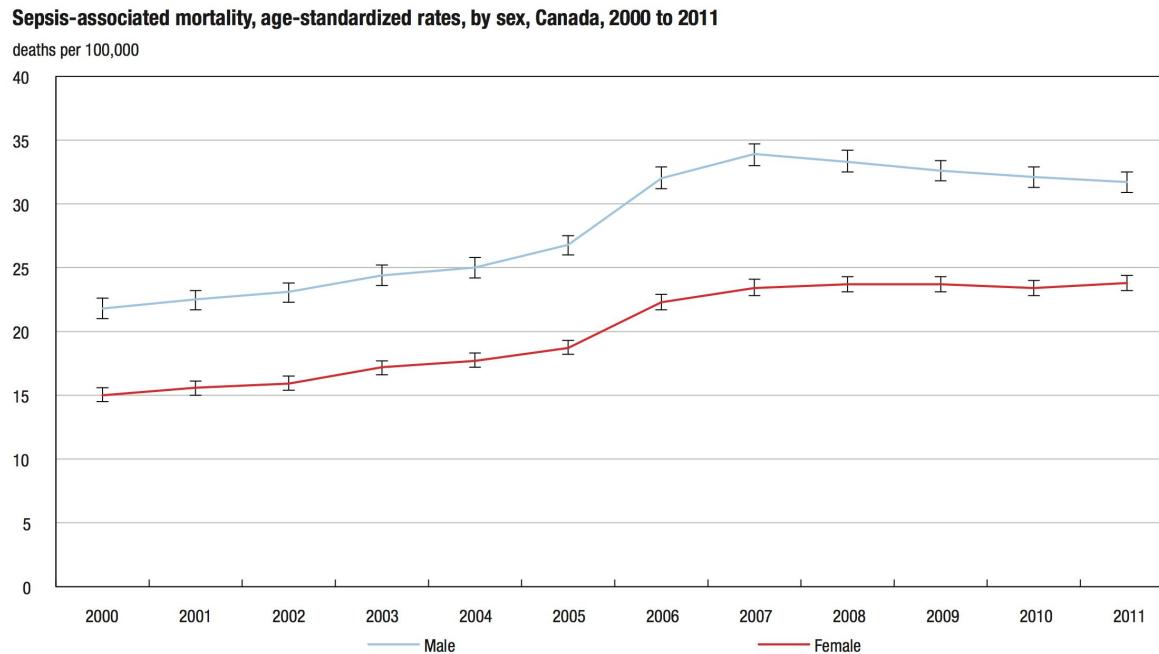


What do cytokines do?

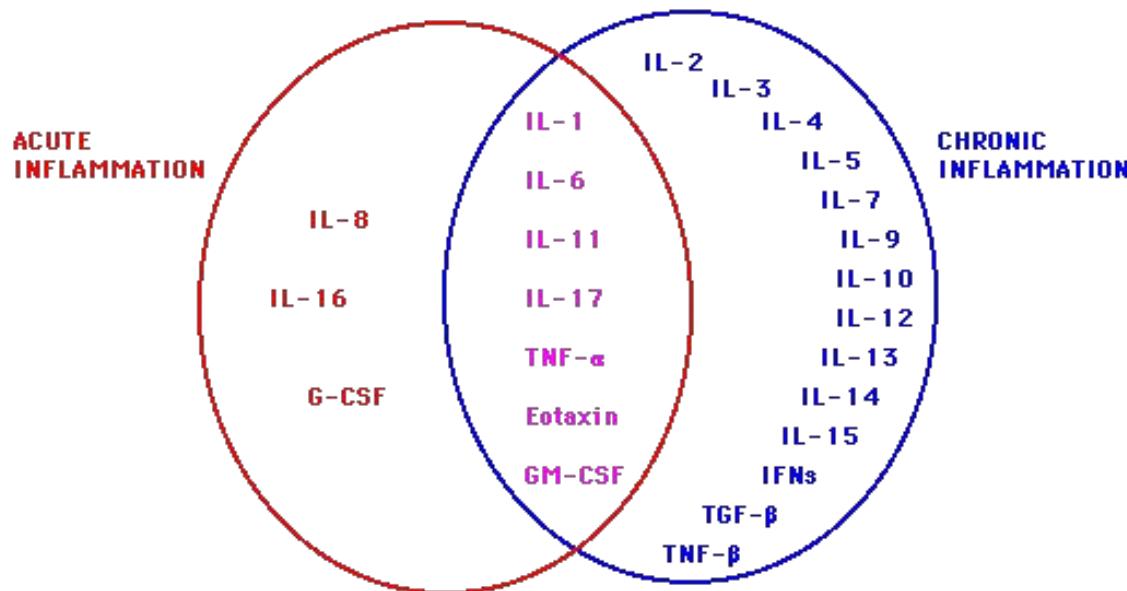
Regulation and mediation of
processes in our bodies

Septic shock: How big of a problem is it?

1 in every 18 deaths in Canada is sepsis related
Around 33% mortality rate within the first month

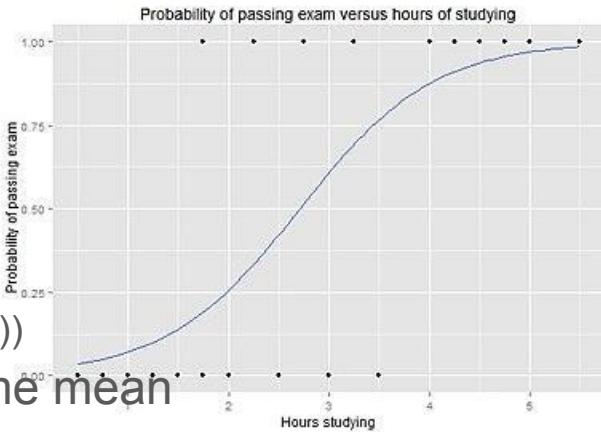


Cytokines and their role

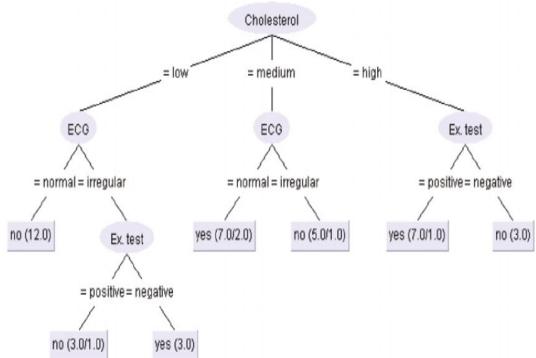


Statistical techniques

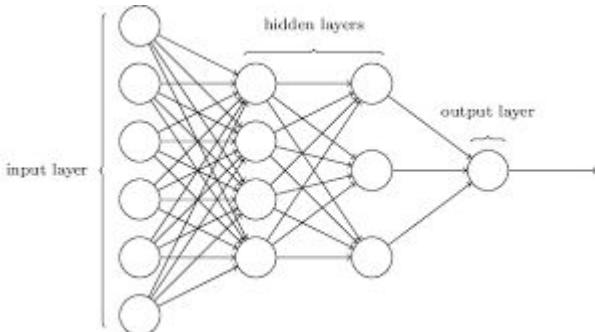
- Logistical regression: binary classifier
 - Fit $y=f(x)$, where y is binary and f is of the form $1/(1+\exp(a+bx))$
- T-test: test whether two sets of numbers has the same mean



Decision tree



Neural network



GWAS: what are Manhattan plots?

