BOSTON UNIVERSITY
Virology, Immunology & Microbiology | Boston University Chobanian & Avedisian School of Medicine
Biomedical Engineering | College of Engineering
aTOMIClab
NEIDL

Unboxing PANDORA: Predicting immunological memory to newly emerging viruses

From Theory to Practice: AI in Immunological Cases

Module 4 – Day 2

Adriana Tomic

adrianatomic
@TomicAdriana
atomic@bu.edu
atomiclaboratory

atomic-lab.org

Module 4 – Day 2 - overview

Part I – Lecture plus Q&A

- Lecture "*Unboxing PANDORA:Predicting immunological memory to newly emerging viruses*"

Lecture (3:30-4:30pm) ~1h

- Discussion

Discussion (4:30-4:45pm) ~15min

Part II – Demonstration of the PANDORA software

- Discussion about team-task and practical course

Hands-on (4:45-5:15pm) ~30min

BOSTON UNIVERSITY
Virology, Immunology & Microbiology | Boston University Chobanian & Avedisian School of Medicine
Biomedical Engineering | College of Engineering
aTOMIClab
NEIDL

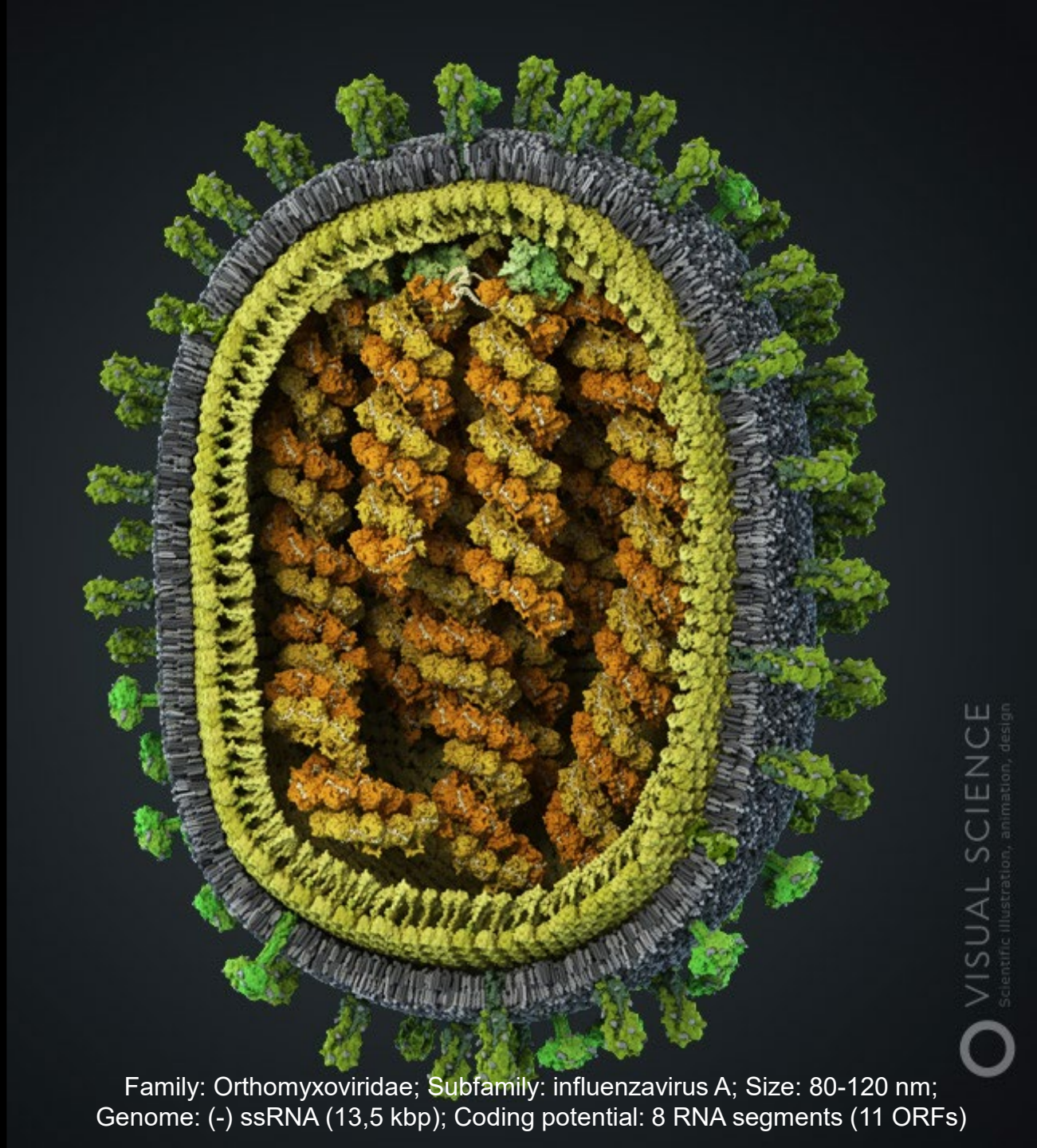
Unboxing PANDORA: Predicting immunological memory to newly emerging viruses

Adriana Tomic

adrianatomic
@TomicAdriana
atomic@bu.edu
atomiclaboratory

atomic-lab.org

Viruses are the simplest life forms



Family: Orthomyxoviridae; Subfamily: influenzavirus A; Size: 80-120 nm;
Genome: (-) ssRNA (13,5 kbp); Coding potential: 8 RNA segments (11 ORFs)

Viruses are the simplest life forms

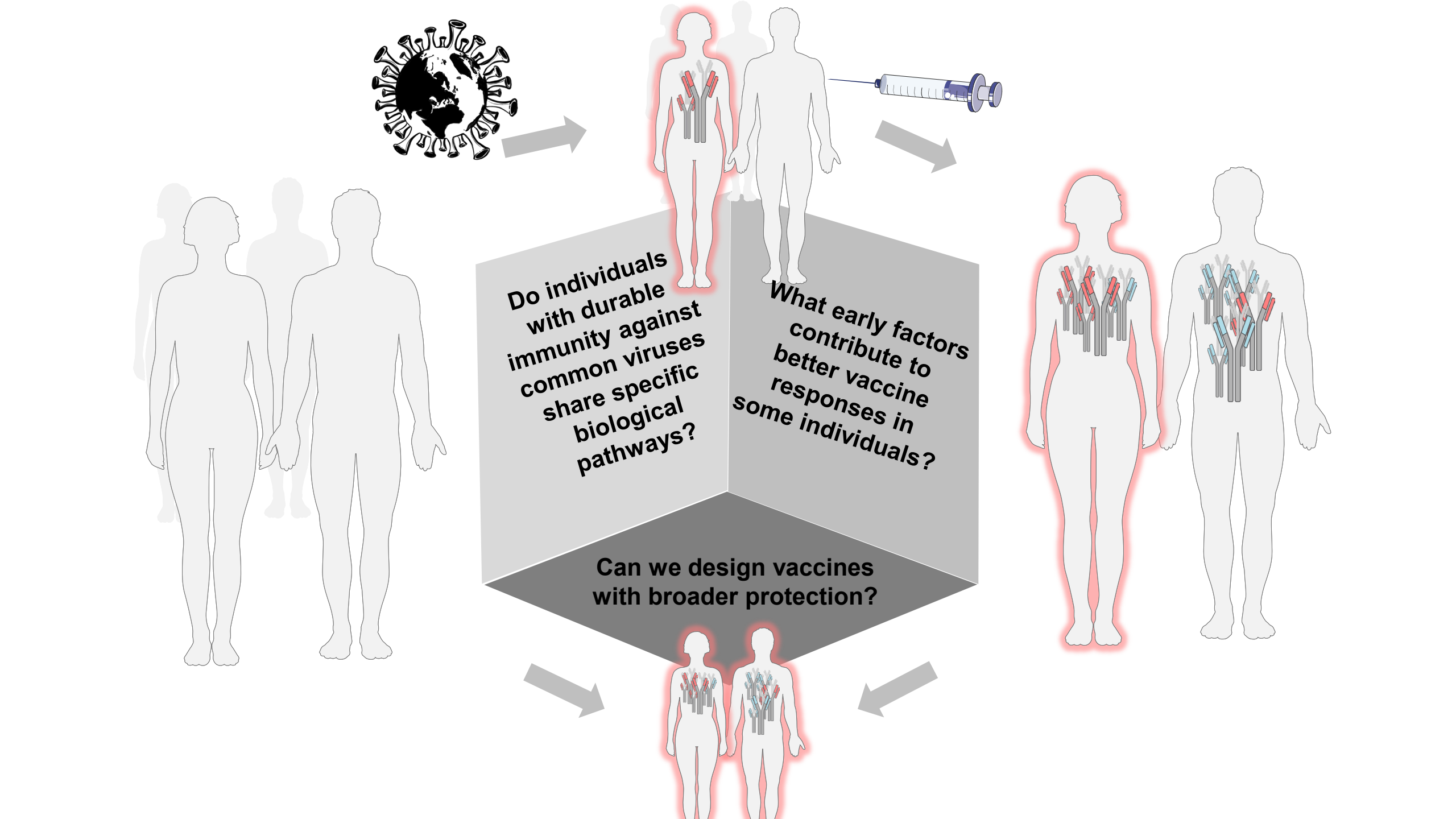
…that represent the greatest naturally occurring threat to human health

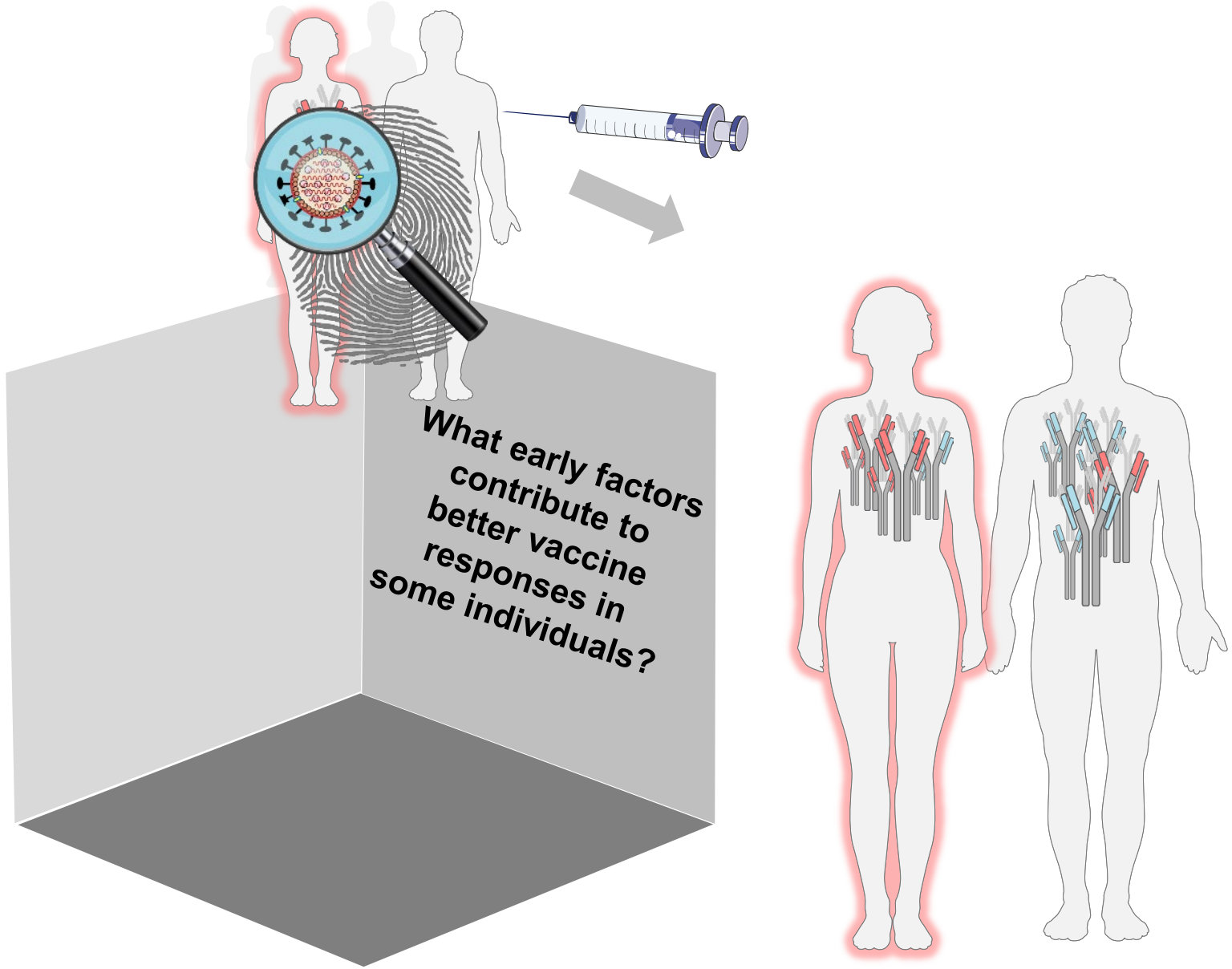
1918 Spanish Flu
50-100 million deaths

1968 Hong Kong Flu
1-4 million deaths

2019 Coronavirus pandemic
7.1 million deaths

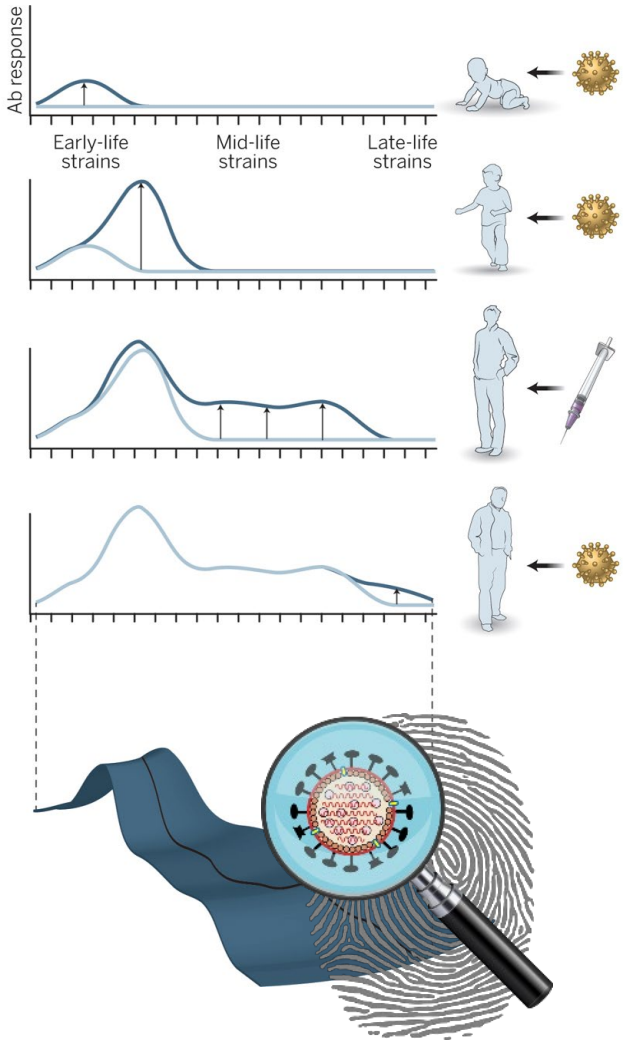
To **STOP** a pandemic, a **GLOBAL** solution is imperative.


Do individuals with durable immunity against common viruses share specific biological pathways?
What early factors contribute to better vaccine responses in some individuals?
Can we design vaccines with broader protection?


What early factors contribute to better vaccine responses in some individuals?

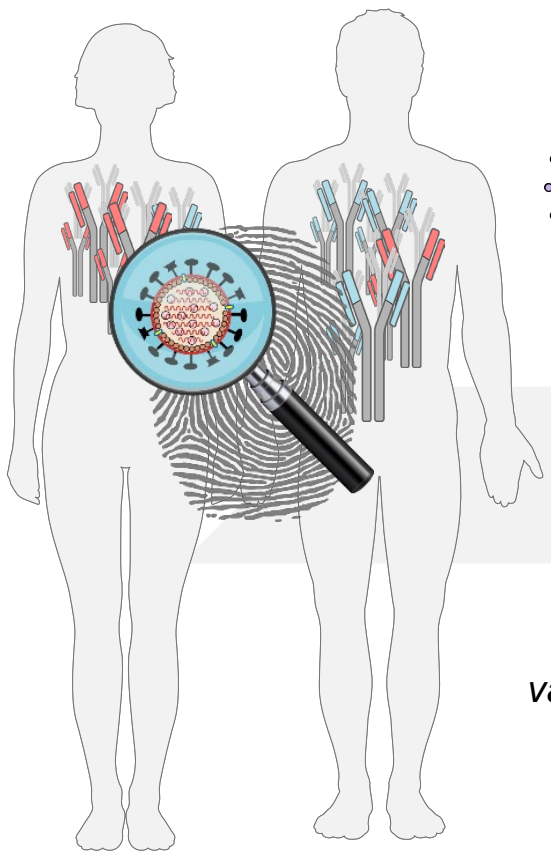
Part 1: Deciphering FluPRINT

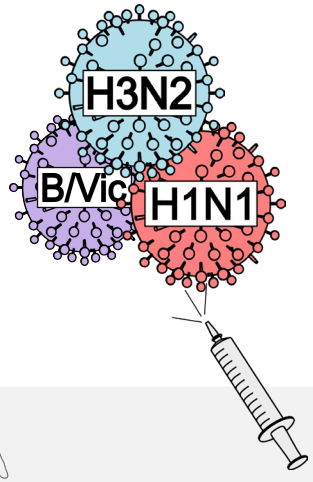
FLUPRINT: lifetime of exposure to influenza



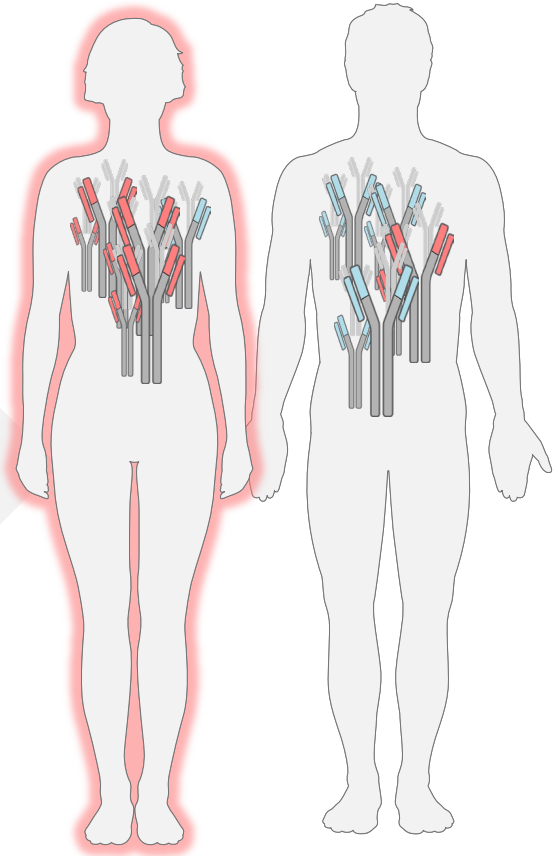
Justin Lessler, Science, 2013

INDUCTION OF HAI ANTIBODIES (HAI titer >=40) associated with reduction in influenza disease





Seasonal inactivated flu vaccine (Fluzone)



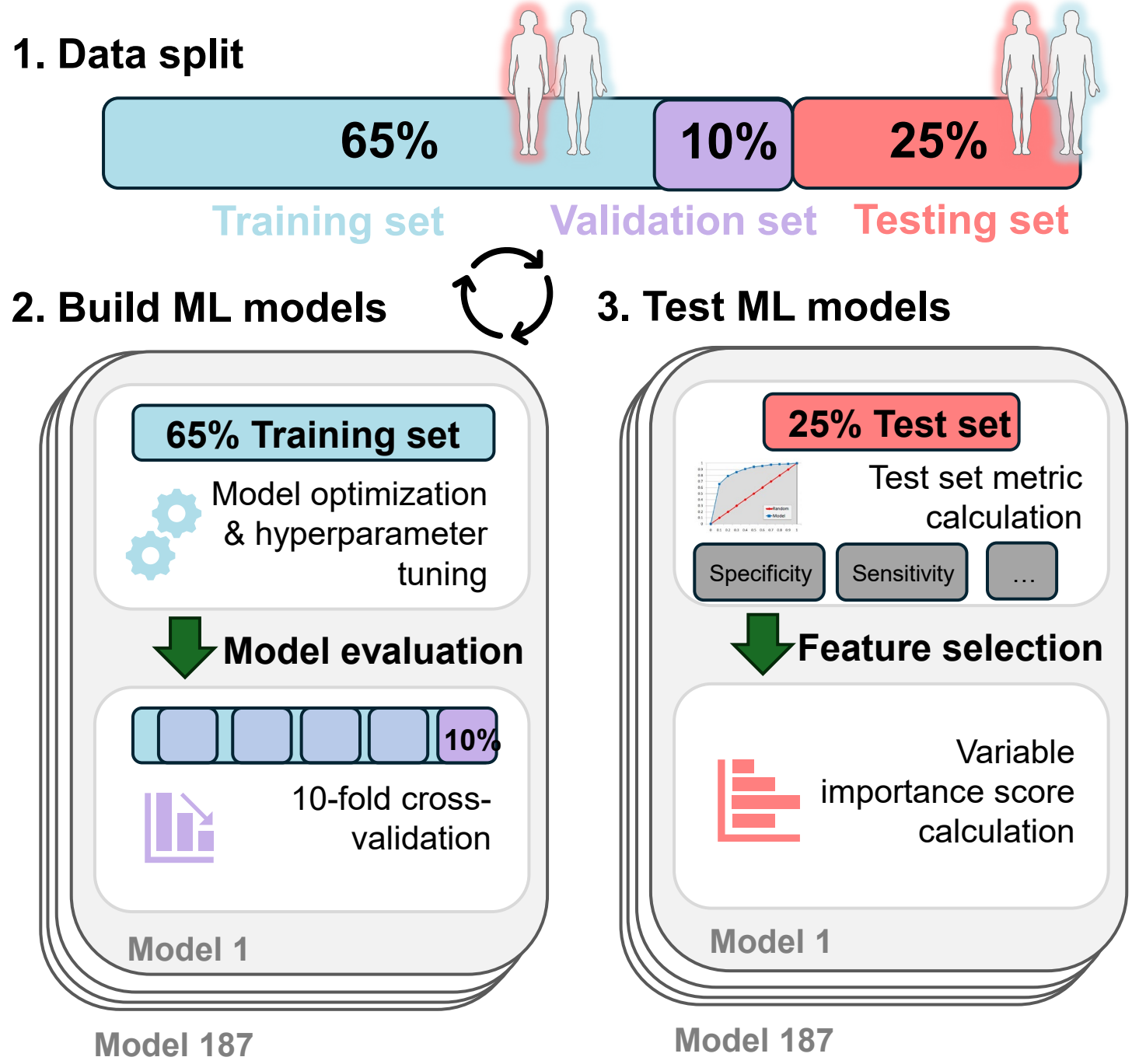
'HIGH RESPONDER'

Tomic et al, JI, 2019

SIMON

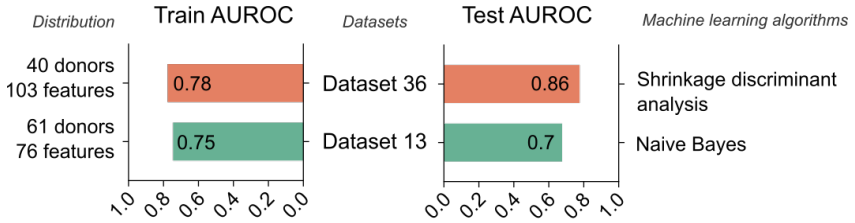
Sequential
Iterative
Modelling
Over-
Night

Tomic et al, JI, 2019;
Tomic et al, Patterns, 2021

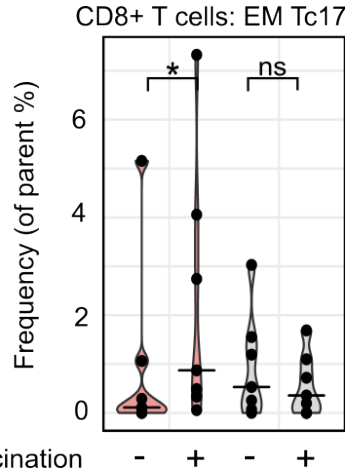


Immune status defines vaccine responses

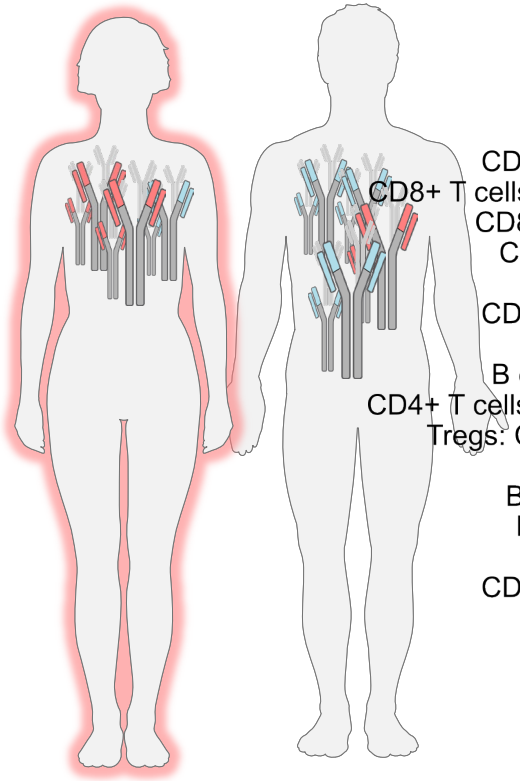
SIMON forecasts who will be **high responder** based on pre-existing immune status

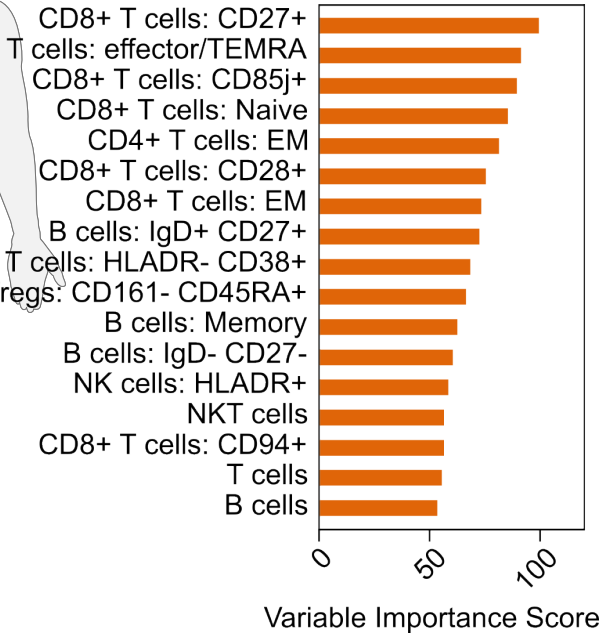


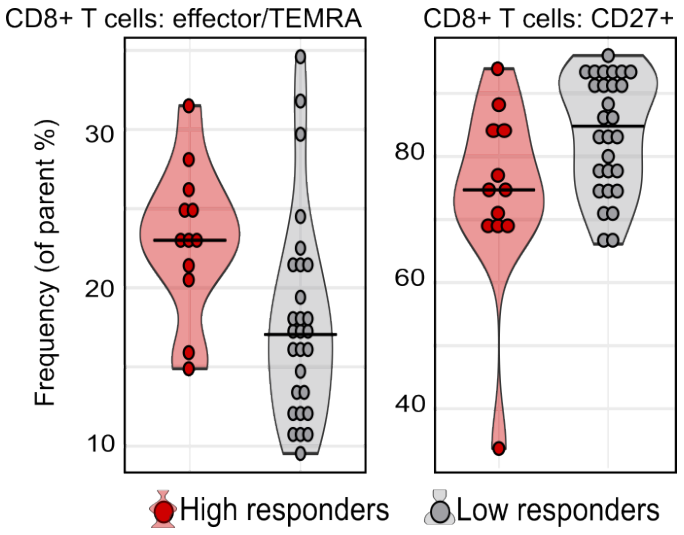
High responders harbor flu-specific memory in EM CD8 T cell pool

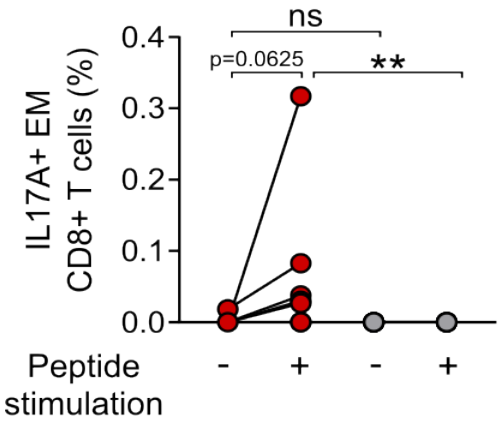


High responders have increased frequency of immune memory cells before vaccination

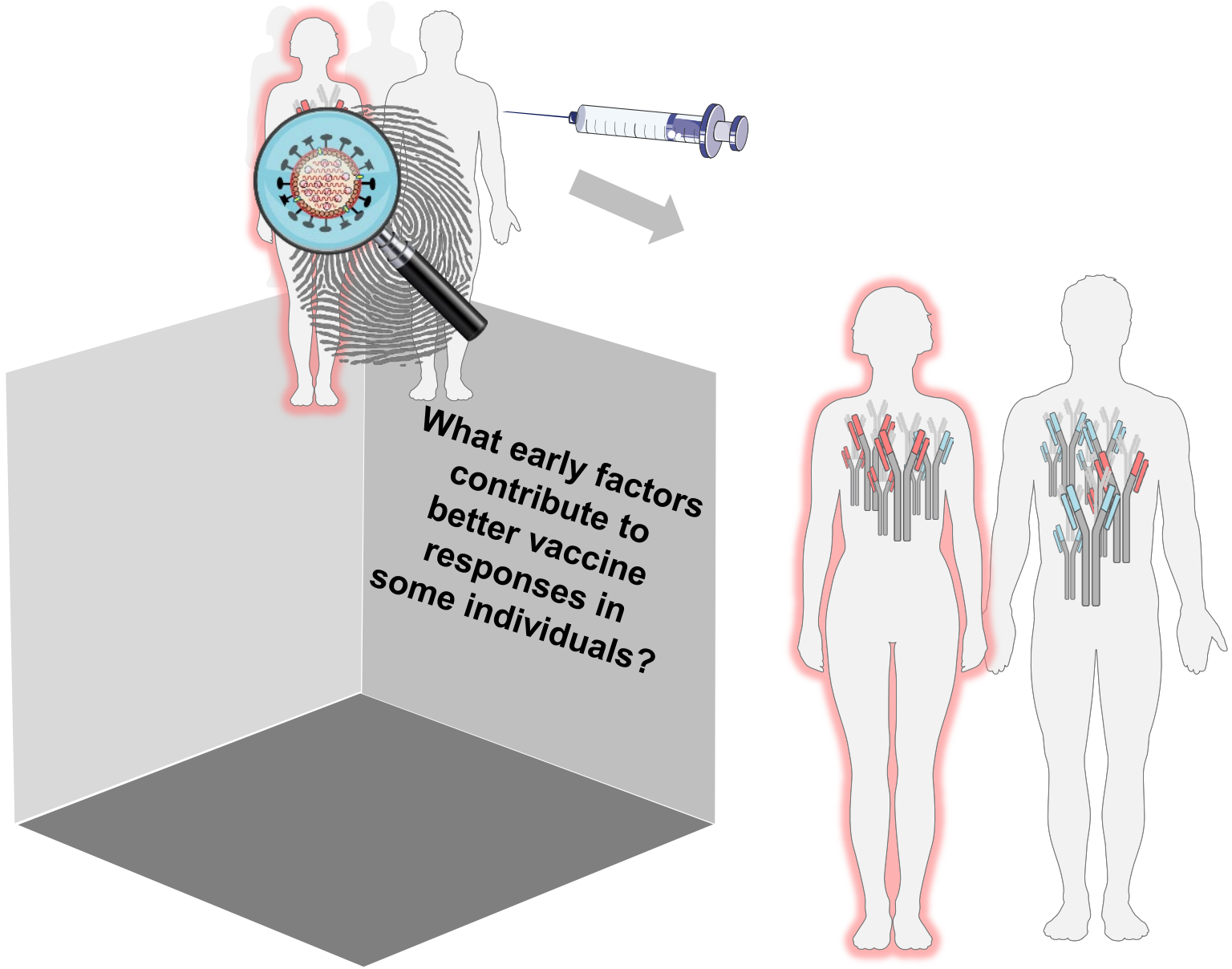


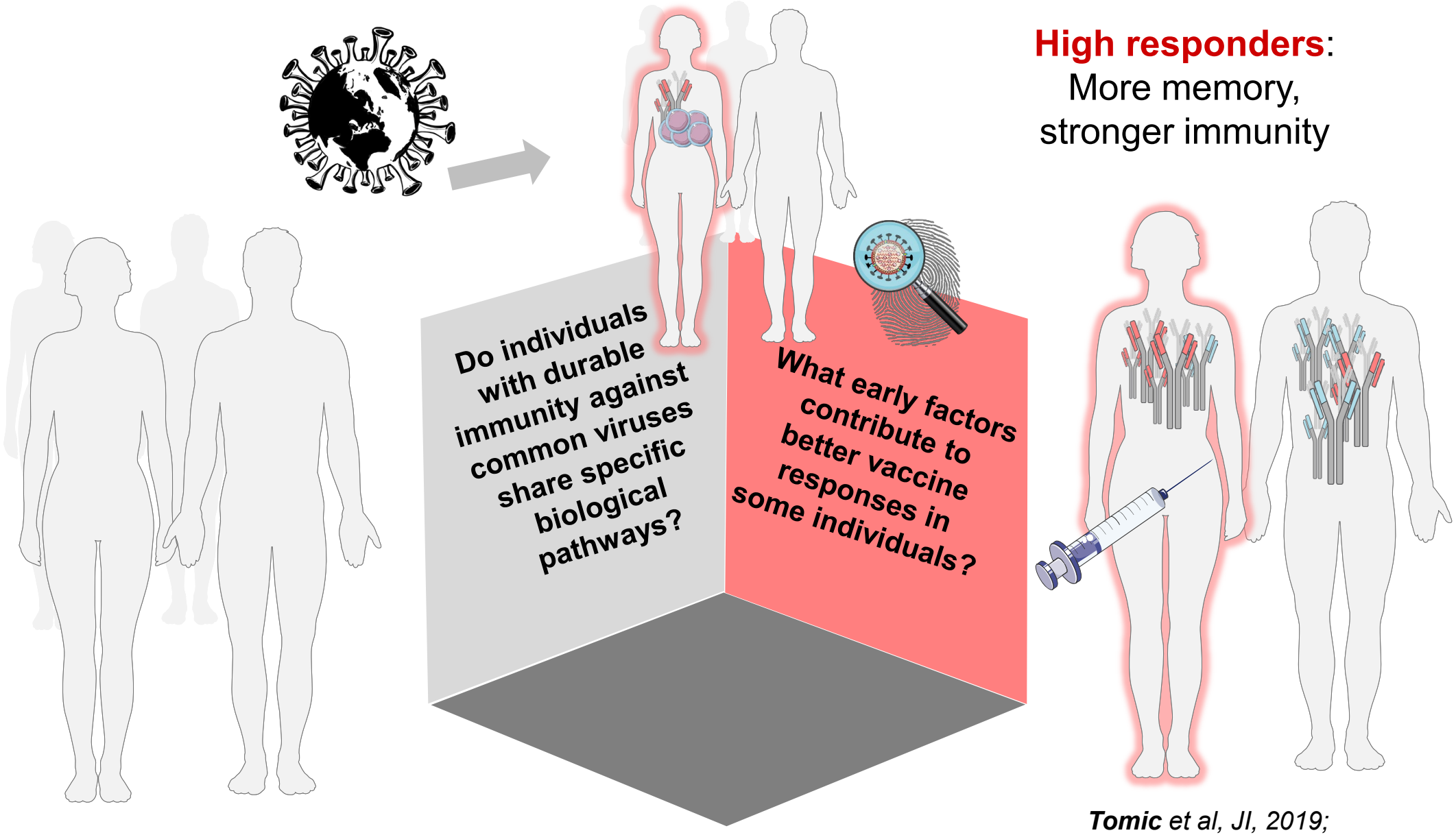






Tomic et al, JI, 2019

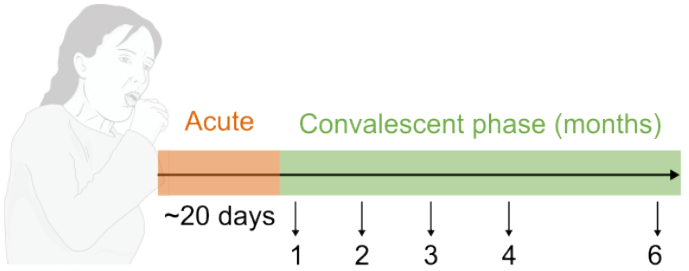

What early factors contribute to better vaccine responses in some individuals?

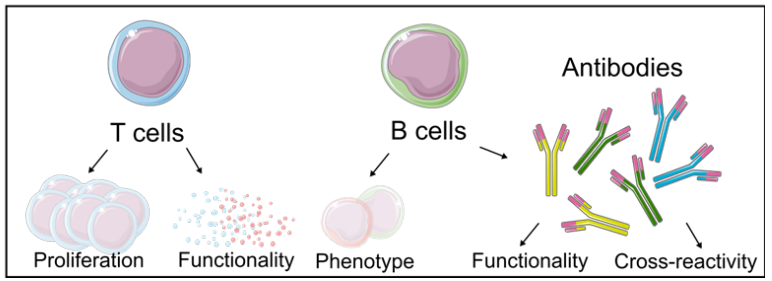


High responders:
More memory,
stronger immunity

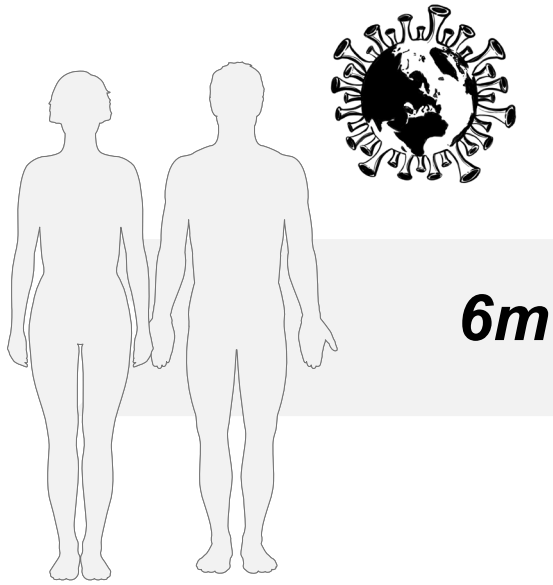
Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

Part 2: COVID-19 memory

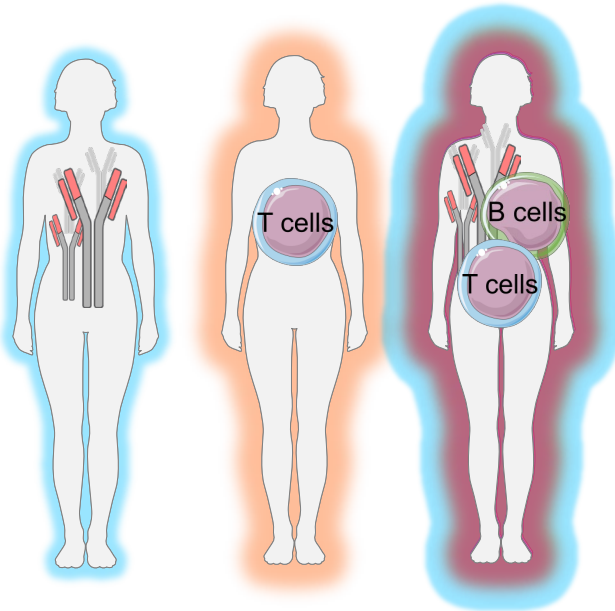




T cells						
IFNγ ELISpot		●		●		●
Proliferation		●				●
ICS		●				●
Antibodies						
Nucleocapsid IgG	●	●	●	●	●	●
Spike IgG	●	●	●	●	●	●
PseudoNA	●	●	●	●	●	●
MSD(9 strains)	●	●	●	●	●	●
Spike isotypes	●	●	●	●	●	●
ADCD	●	●	●	●	●	●
ADMP	●	●	●	●	●	●
ADNP	●	●	●	●	●	●
ADNKA	●	●	●	●	●	●
B cells						
Memory ELISpot	●	●	●	●	●	●
ELISpot	●	●	●	●	●	●



6mo pi



PITCH

PROTECTIVE IMMUNITY from T CELLS in HEALTHCARE WORKERS

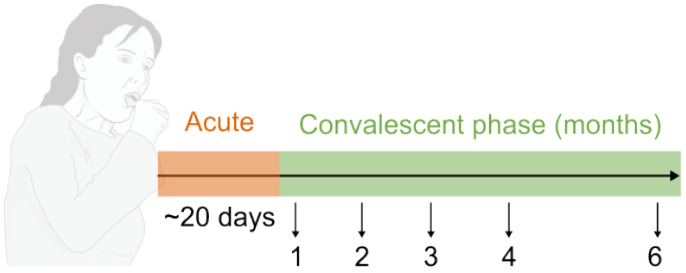
PITCH consortium - 5 clinical sites in the UK

78 individuals recruited

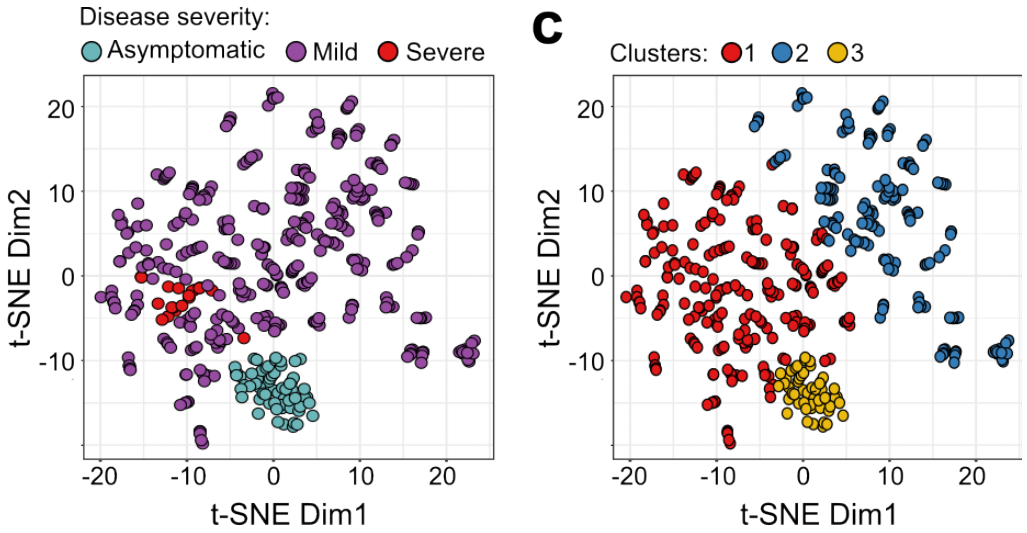
433 samples total analysed

Tomic et al, Nat Comms, 2022

Unveiling the spectrum of COVID-19 immunity

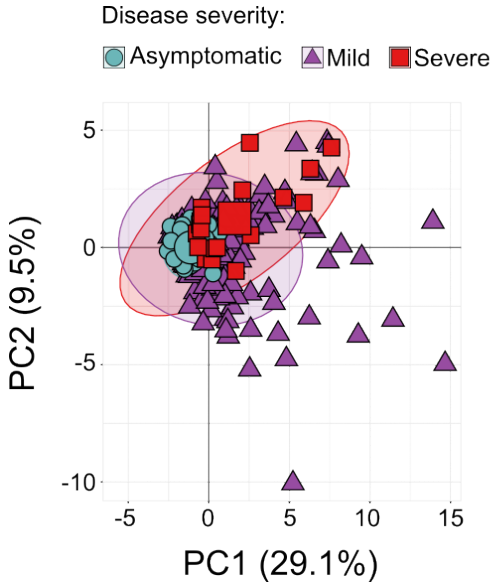


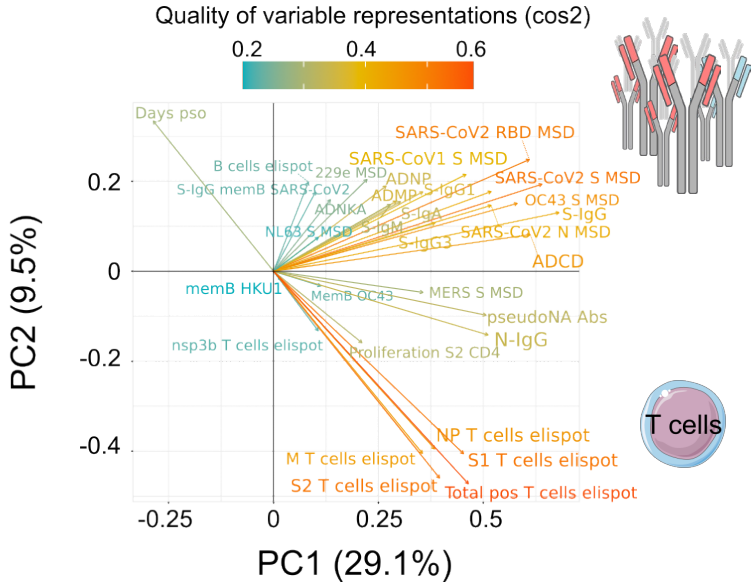
Following infection **3 distinct COVID immune profiles** are observed

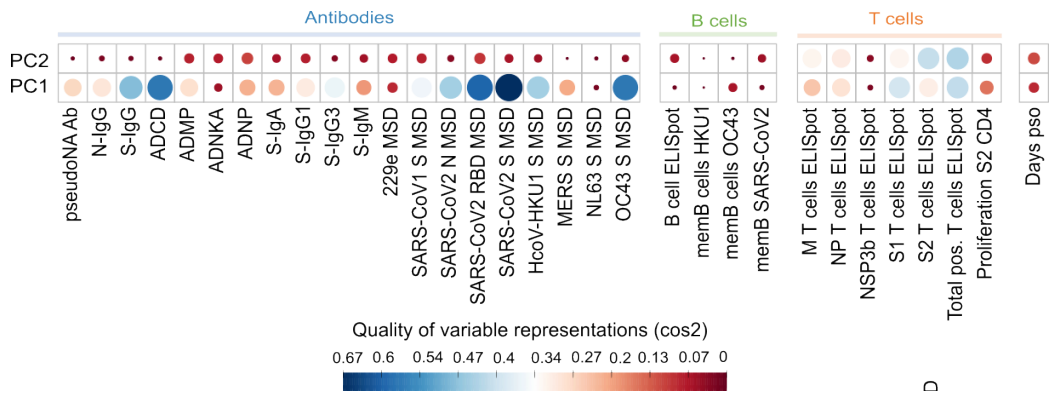


tSNE analysis of integrated datasets:
Clinical data: 20 parameters + Immunological data: 51 parameters; 433 samples total analysed

COVID Immunity Spectrum: T Cells vs. Antibodies

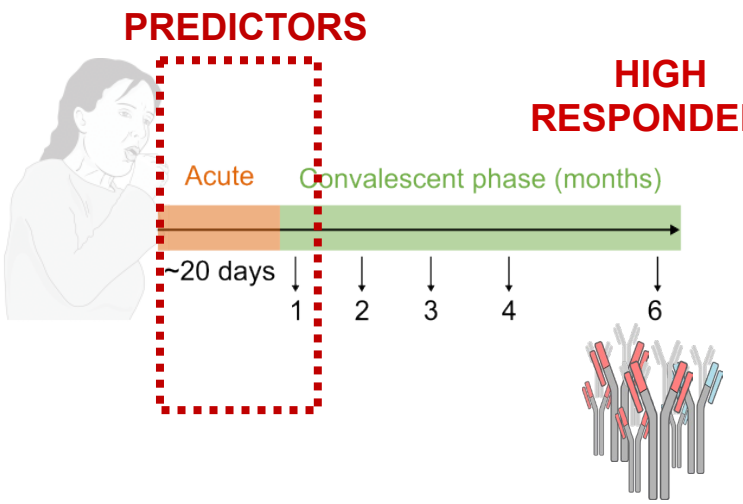






Tomic et al, Nat Comms, 2022

Early immune activation: Key to better COVID defense

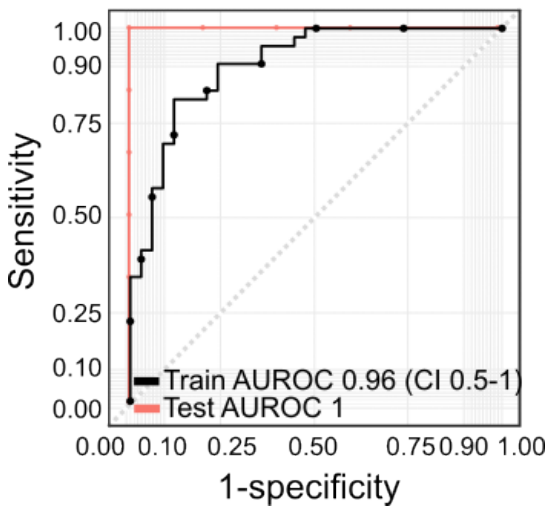


Lumley et al, NEJM, 2021

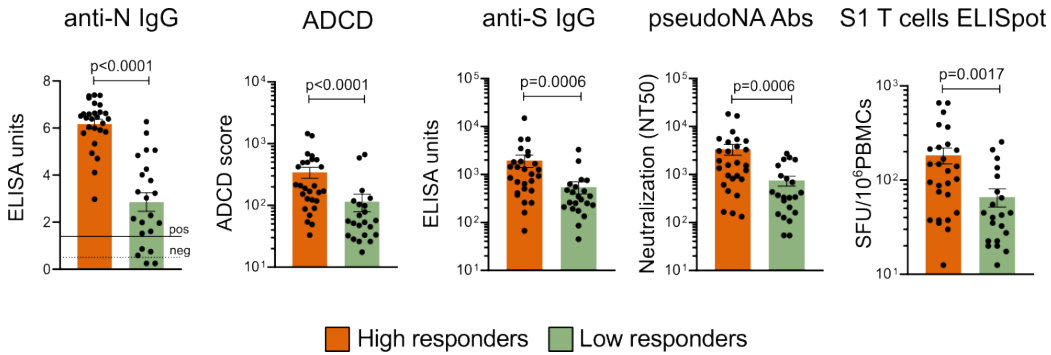
Integrative immunological dataset on day 28 pso (n=3,626 datapoints)

SIMON analysis
Prediction of antibody responder status 6 months pso
(172 ML algorithms, 3,565 models tested)

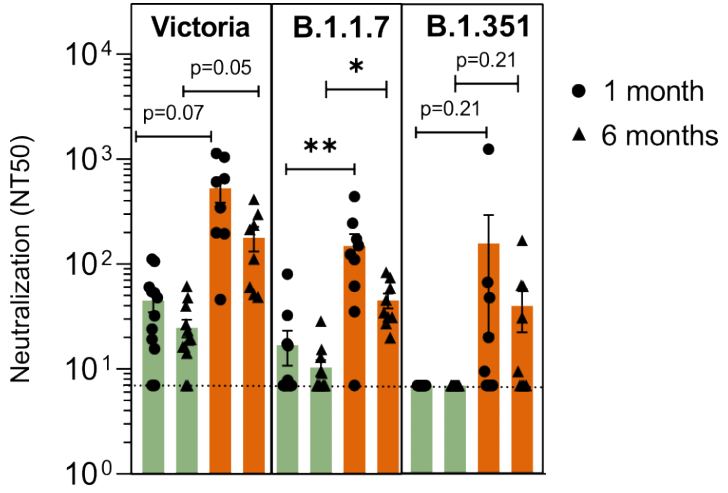
SIMON predicts who will be **high responder** based on early infection immune status



High responders mount stronger antibody and T cell immunity early after infection

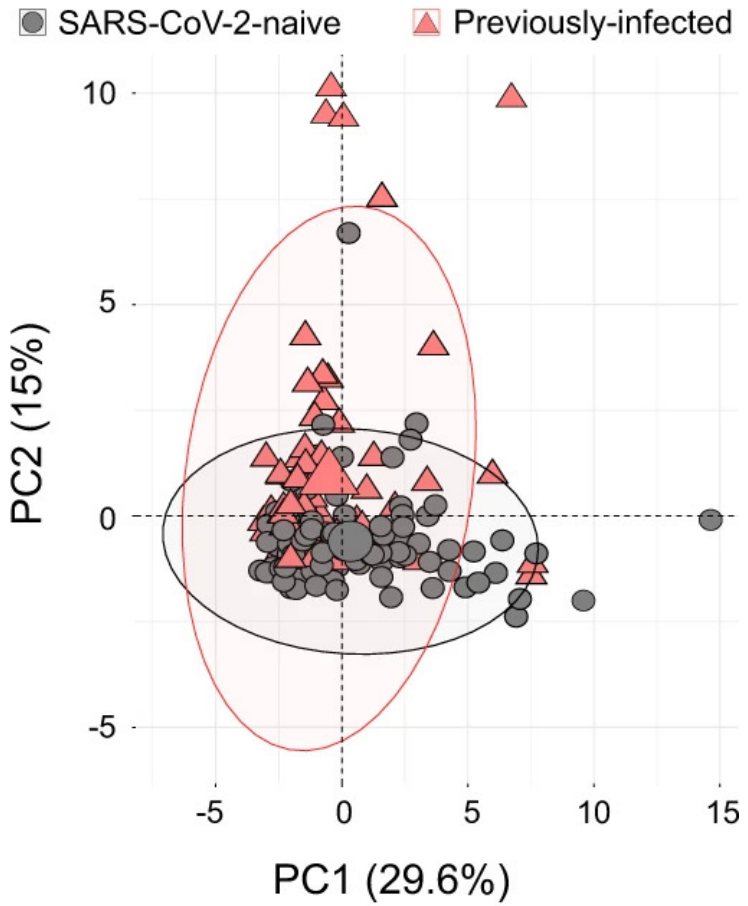


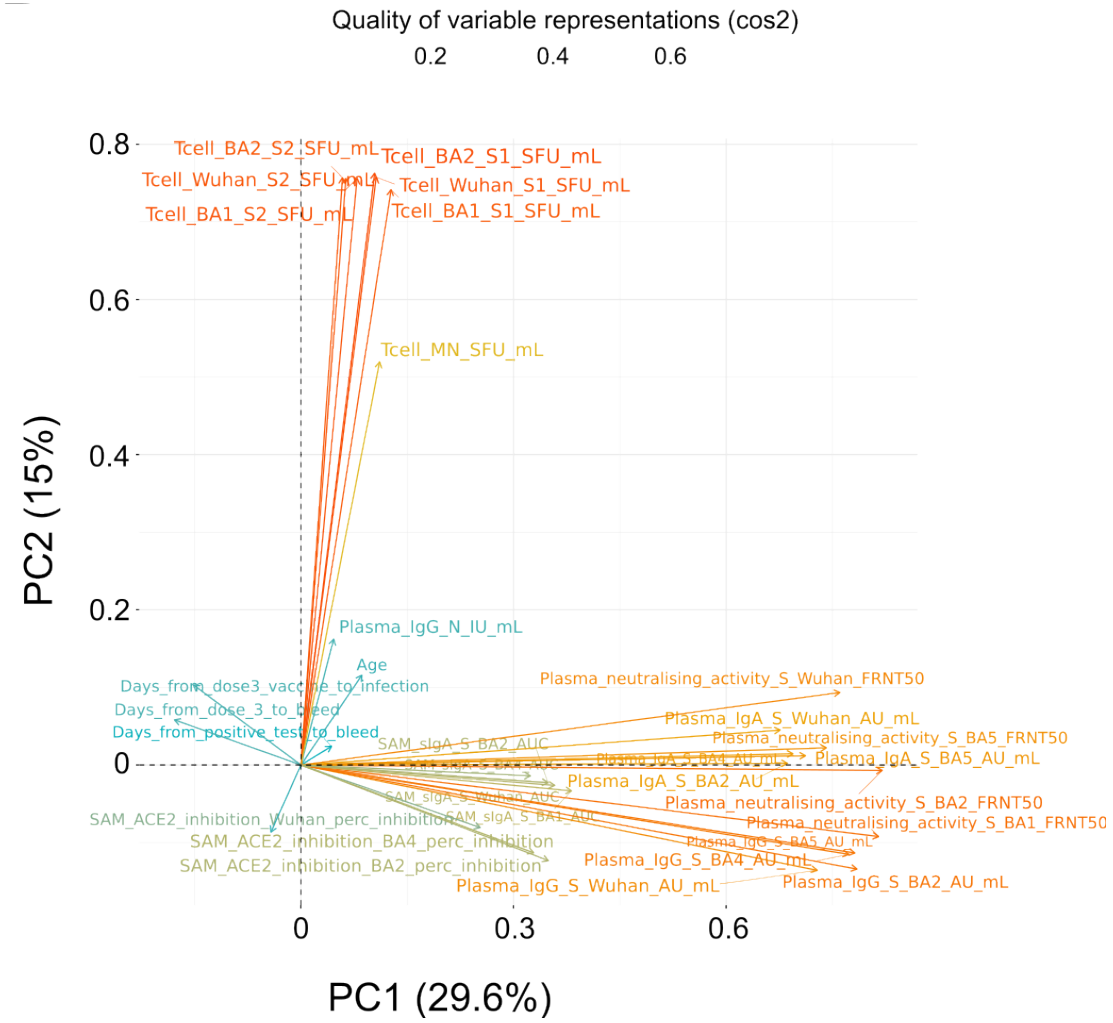
High responders can neutralize VOCs



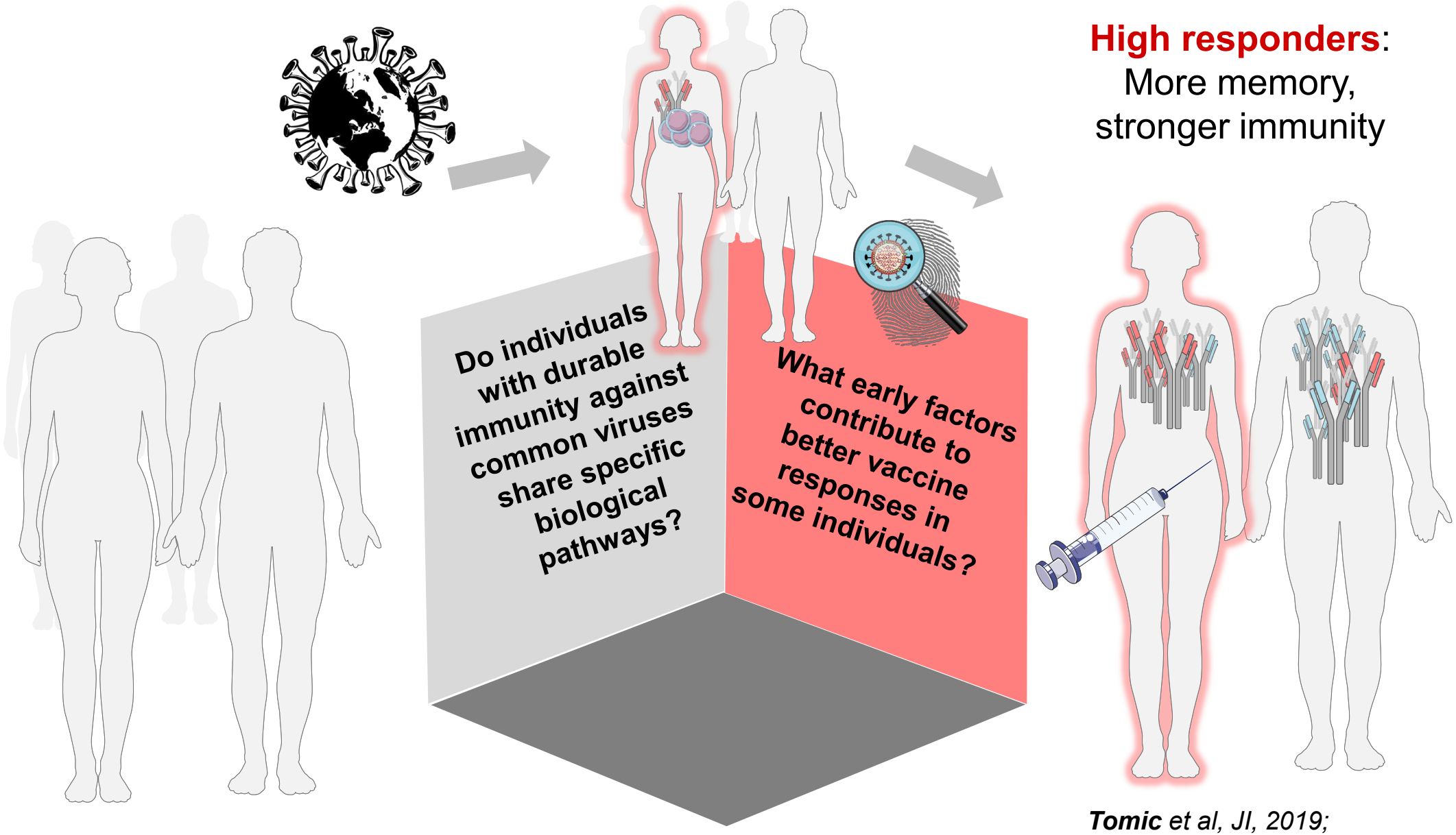
Tomic et al, Nat Comms, 2022

Pre-existing immunity after infection and vaccination





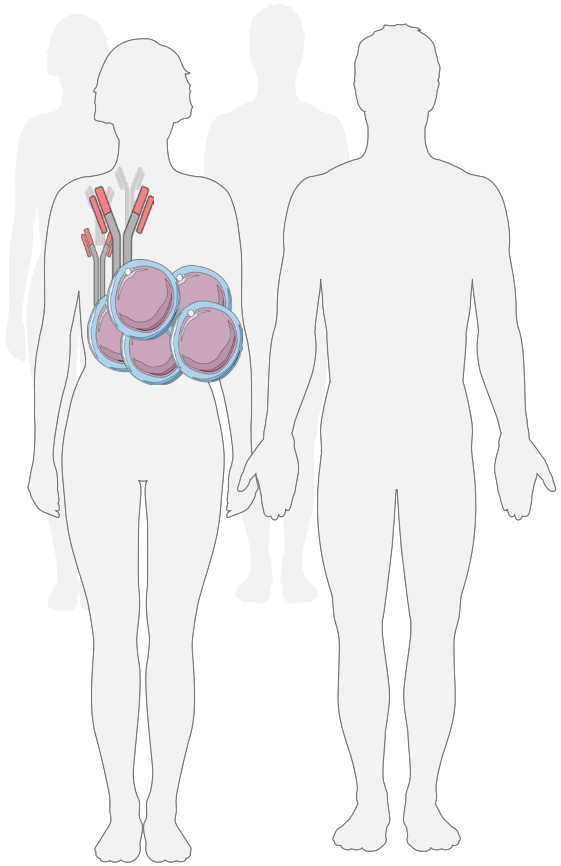
Hornsby et al, Nat Comms, 2023



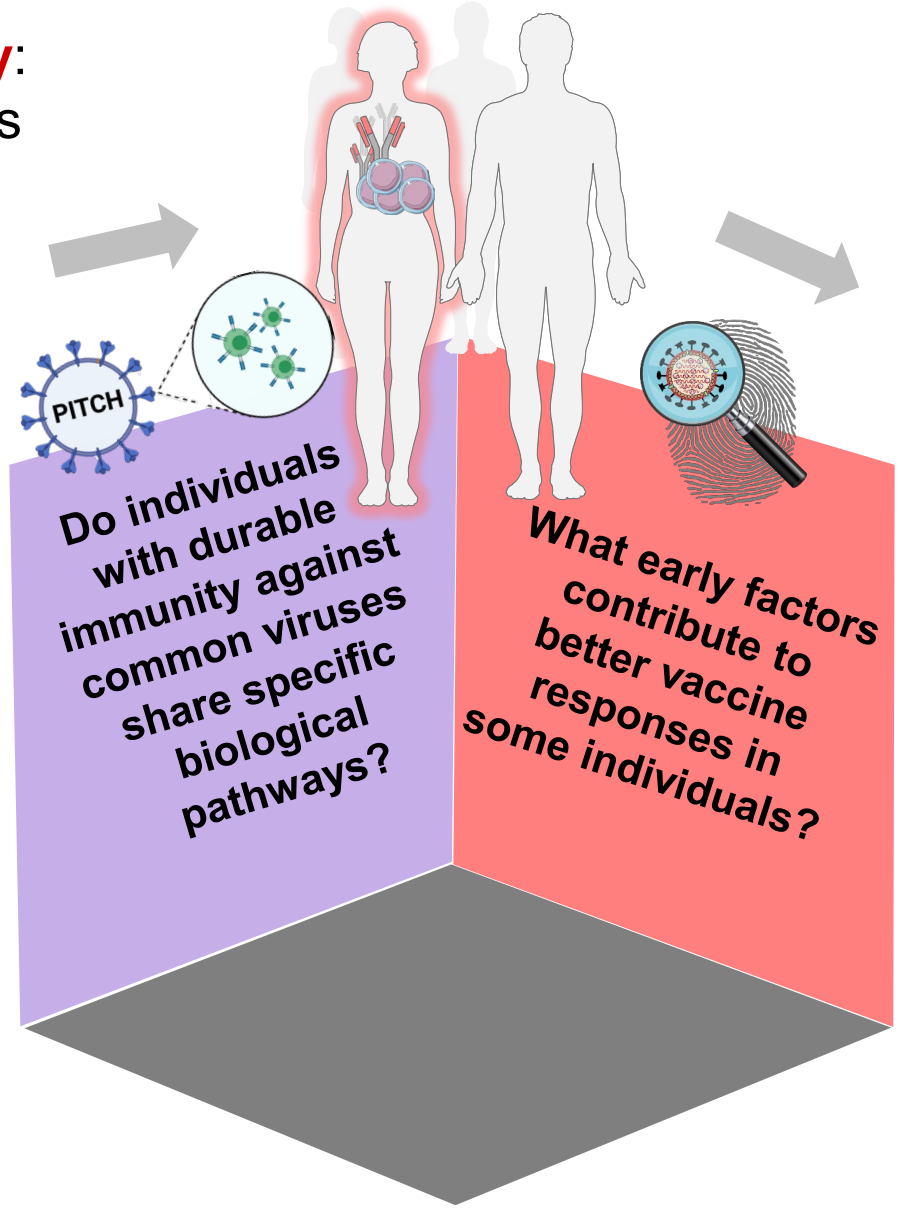
High responders:
More memory,
stronger immunity

Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

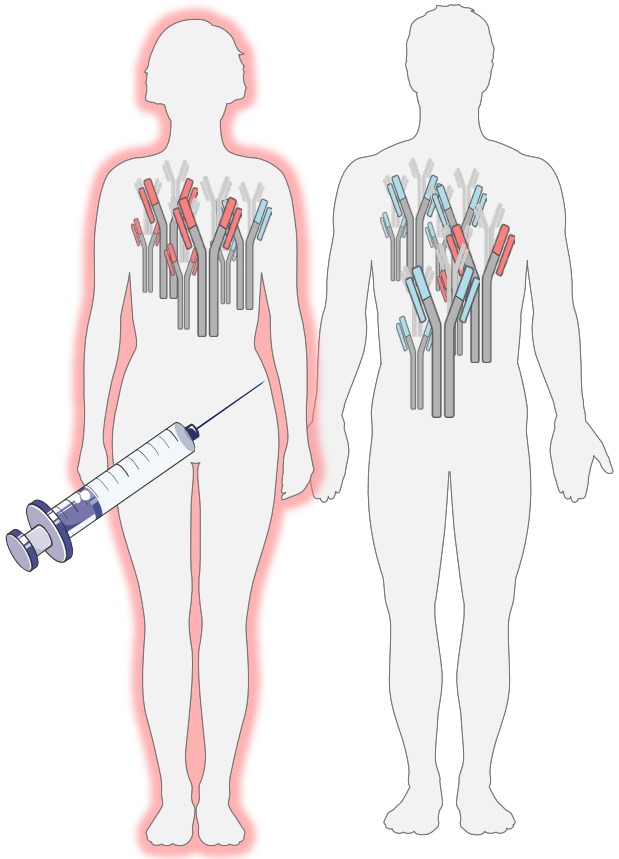
Durable COVID-19 immunity:
Early immune teamwork leads to longer memory



Tomic *et al, Nat Comms, 2022;*
COMBAT consortium, Cell, 2022;
Hornsby et al, Nature Comms, 2023

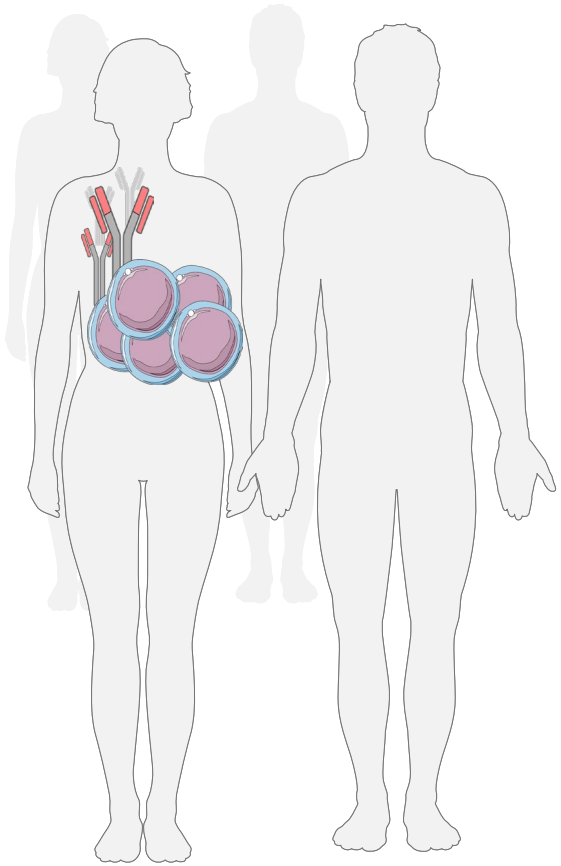


High responders:
More memory, stronger immunity

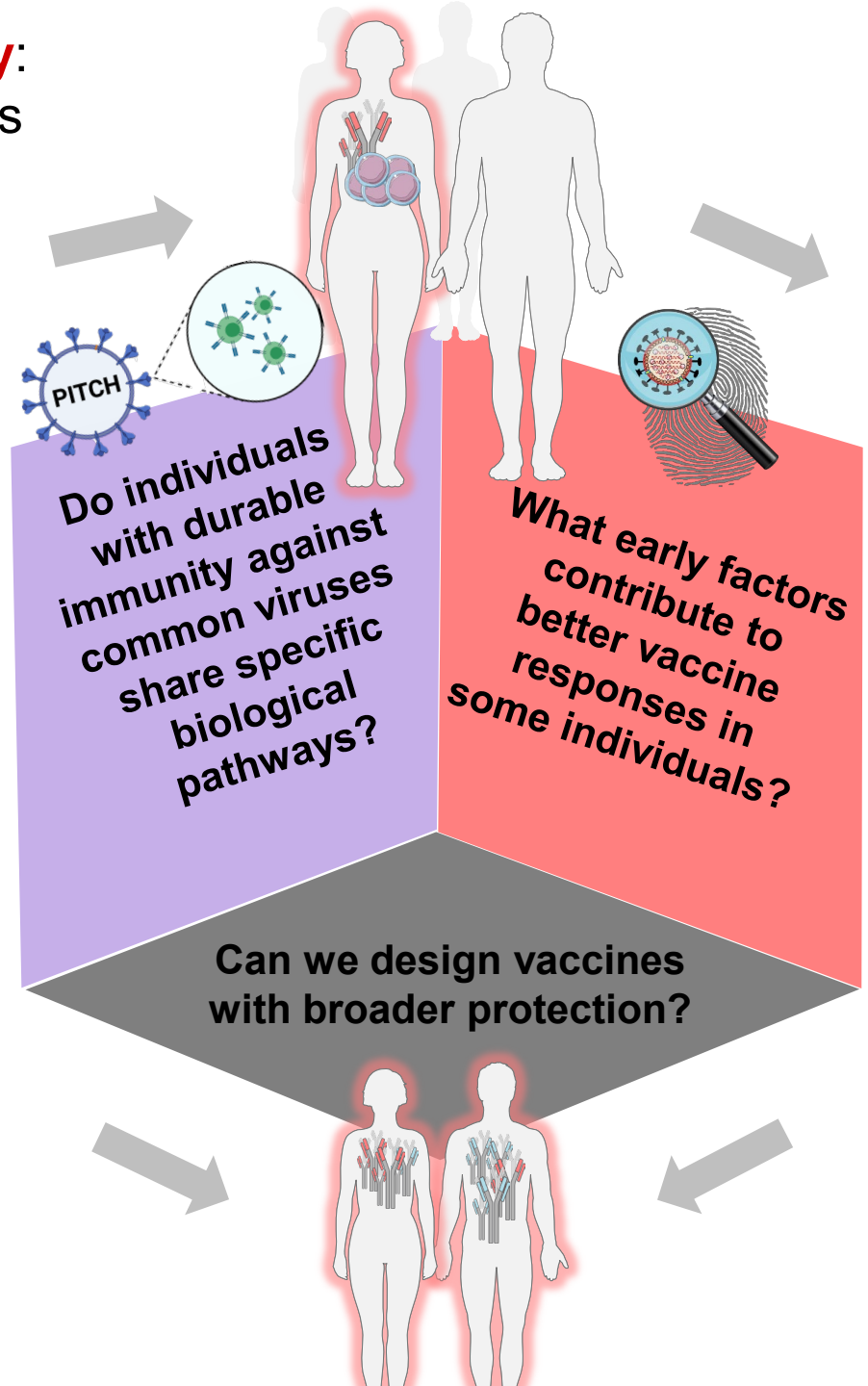


Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

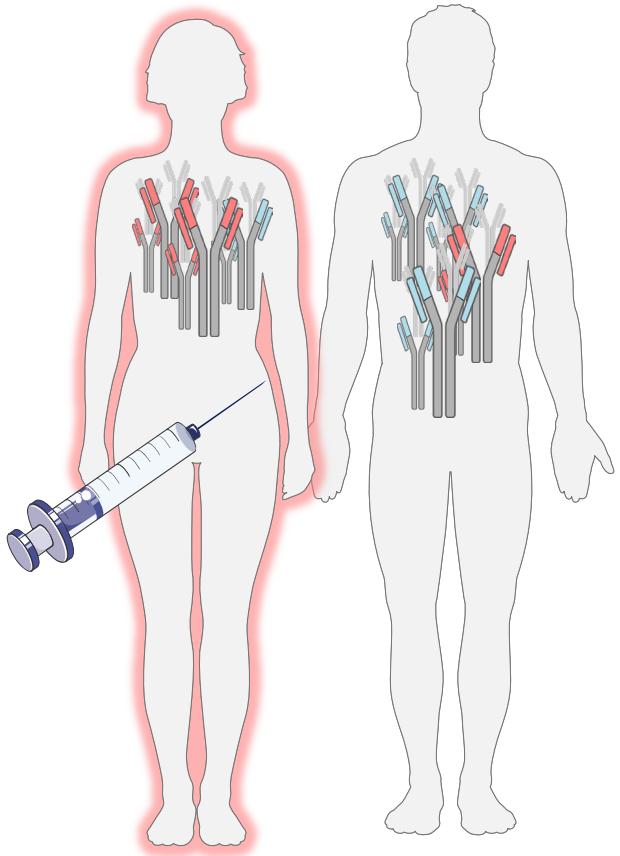
Durable COVID-19 immunity:
Early immune teamwork leads to longer memory



Tomic *et al, Nat Comms, 2022;*
COMBAT consortium, Cell, 2022;
Hornsby et al, Nature Comms, 2023



High responders:
More memory, stronger immunity



Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

Part 3: Towards precision vaccinology

Influenza A virus

- **Cross-species infectivity**
- **High mutation rates** – antigenic drift
- **Antigenic shift** (reassortment) – mixing of genetic materials leading to new subtypes

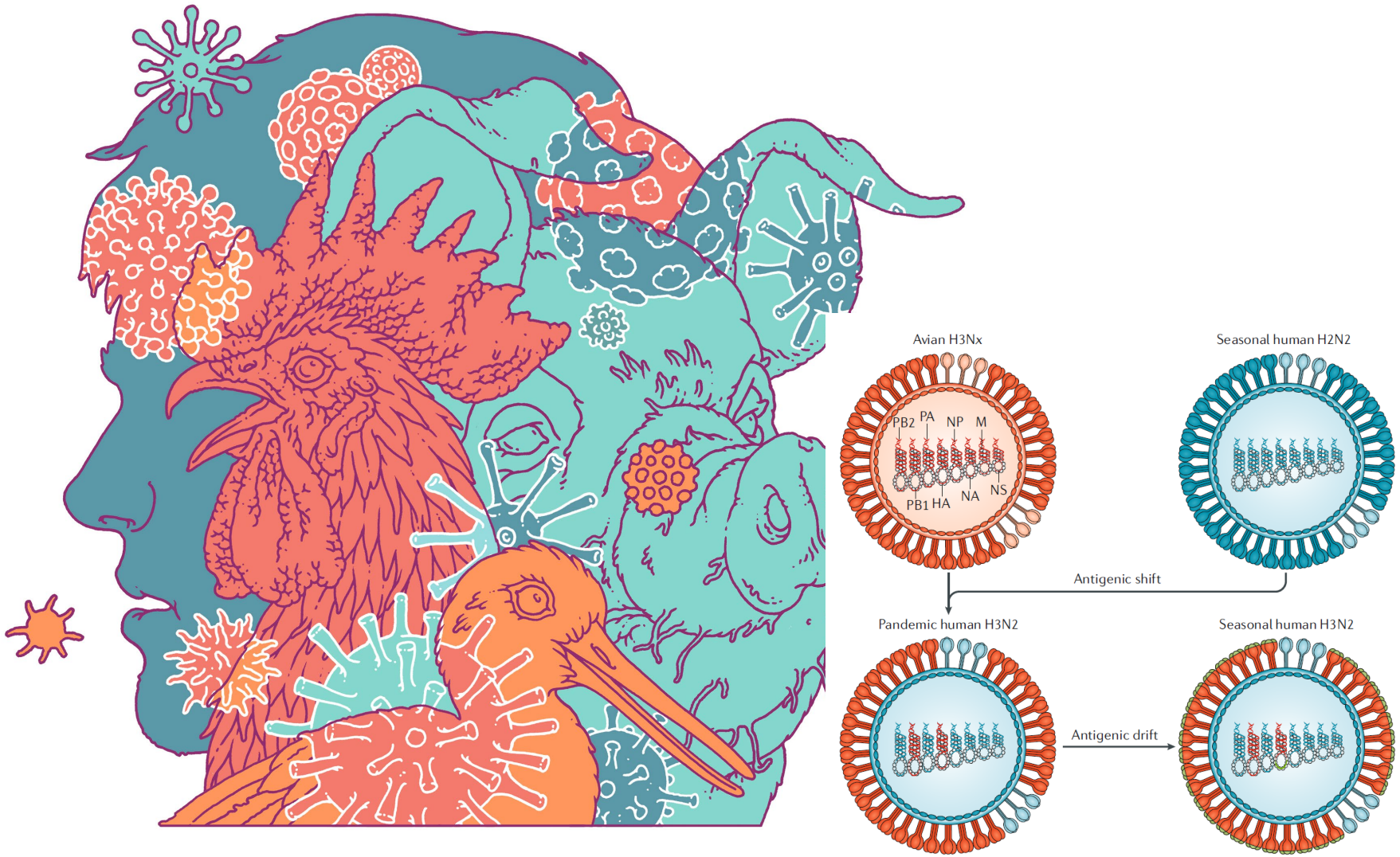
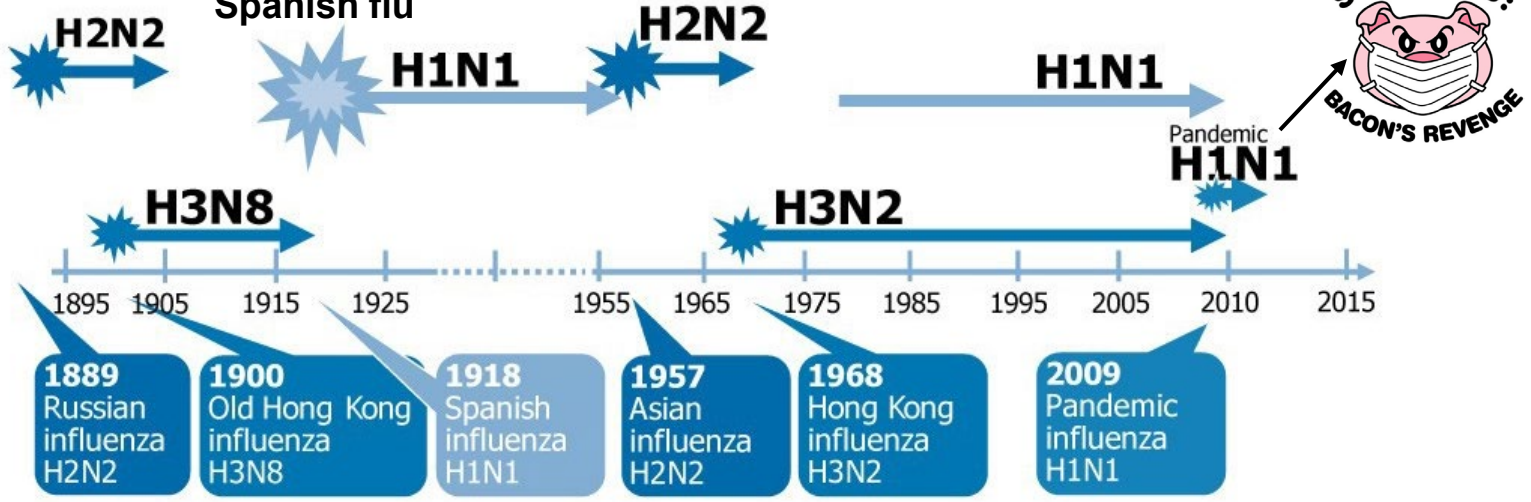


Illustration: Eric Nyquist
https://www.audubon.org/magazine/fall-2020/how-migrating-birds-could-warn-us-next-pandemic

Krammer et al, Nat Reviews, 2018

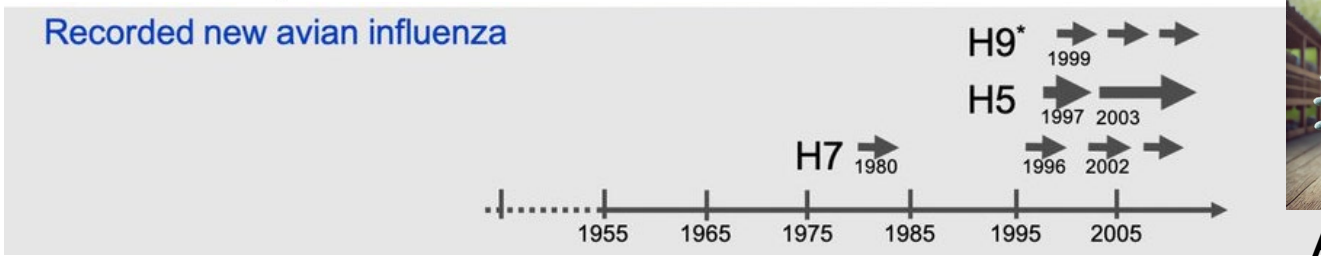
Part 3: Towards precision vaccinology

Spanish flu



Source: European Centre for Disease Prevention and Control (ECDC) 2009

Reproduced and adapted (2009) with permission of Dr Masato Tashiro, Director, Center for Influenza Virus Research, National Institute of Infectious Diseases (NIID), Japan.



Avian flu

H5N1 & H7N9 – newly emerging strains
HIGH MORTALITY RATES

30-60%

20x higher

than Spanish flu

Transmissibility and control of pandemic Influenza

1 child infects 2.4 children

70% vaccinated children could prevent pandemics

Yang et al, Science, 2009

Children vaccination prevents spread of influenza

LAIV vaccine given only to primary school age children

Public Health England

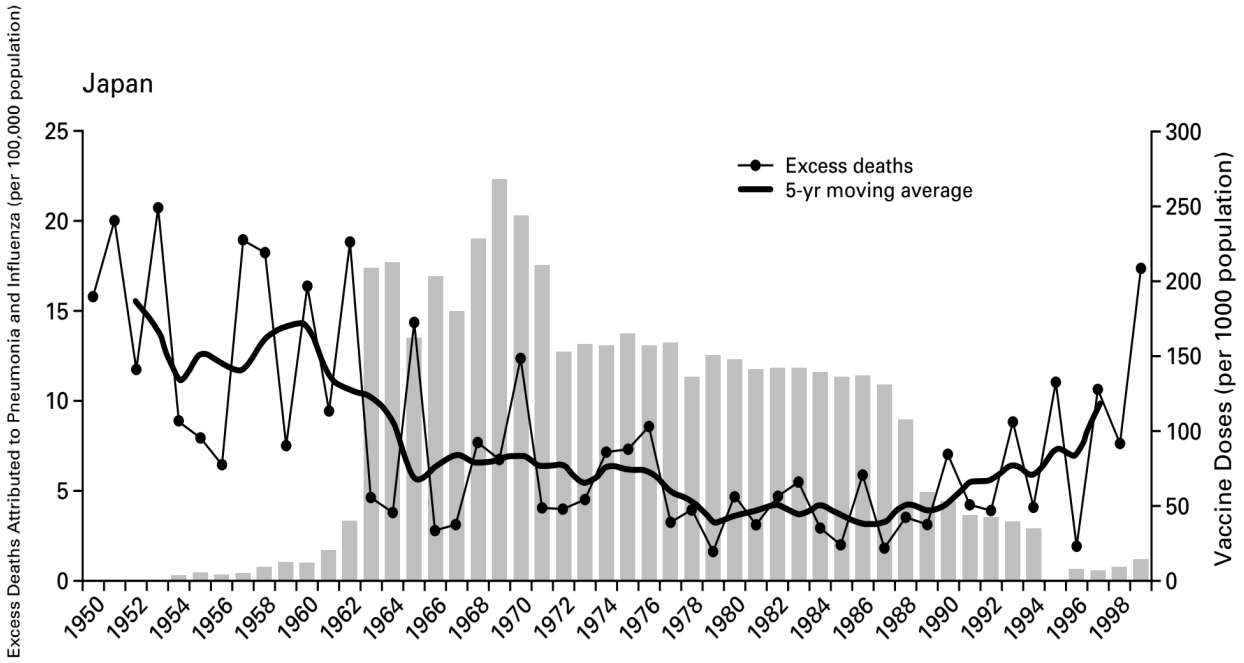
93% less

Children hospital admissions due to confirmed influenza

59% less

Consultation with GP about influenza like illness in adults

Period of mandatory school children vaccination lowered mortality of elderly people in Japan



Uptake and impact of vaccinating school age children against influenza during a season with circulation of drifted Influenza A and B strains. England, 2014/2015

Reichert et al, NEJM, 2009

Universal annual influenza vaccination

CDC ACIP recommendation 2010.

Influenza vaccines

INACTIVATED INFLUENZA VACCINE (IIV)

40-60%
*matched strain

- "detergent"- split vaccine
- excellent safety profile
- low immunogenicity (adjuvants)
- Ab wane over time, no CD8+ T cells

CHILDREN FROM 6 MONTHS OF AGE

?

85%
*mismatched strain

Rhorer et al, Vaccine, 2009

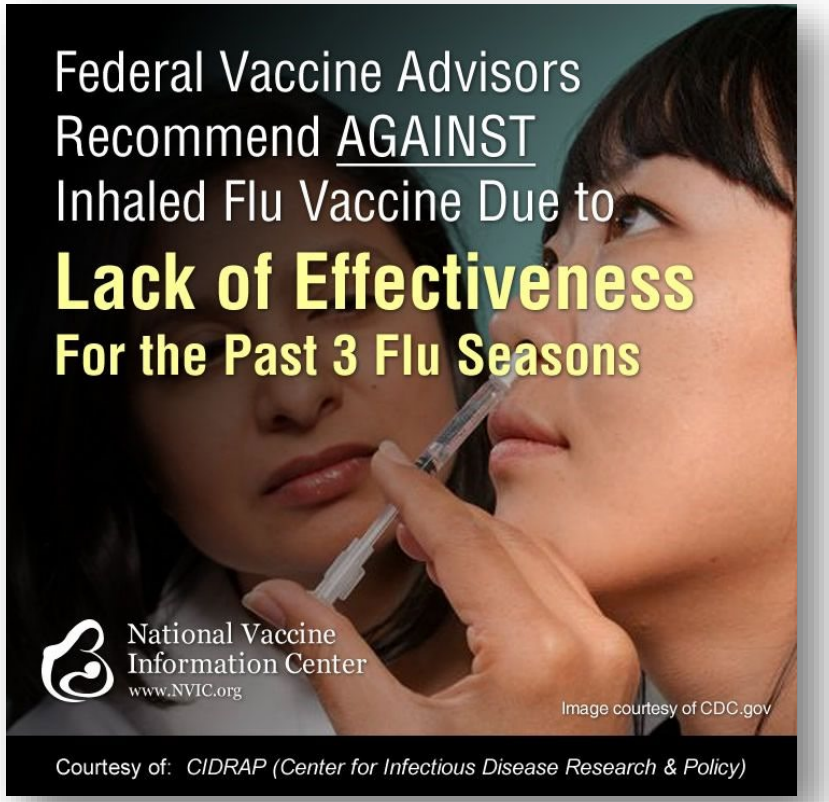
LIVE ATTENUATED INFLUENZA VACCINE (LAIV)

- temperature-sensitive phenotype – grows at 25°C (nasal passage)
- heterosubtypic immunity
- Ab and T cell responses maintained (IFNγ responses, γδT cells)

CHILDREN FROM 2 YEARS OF AGE

LAIV effectiveness?

“ACIP votes down use of LAIV for 2016-2017 flu season”



CDC Advisory Committee on Immunization Practices, June 2016

LAIV has high effectiveness in 2015/2016

Public Health England

Population	N	Cases: unvac; vaccinated	Controls: unvac; vaccinated	Adjusted VE by scheme, age, month, gender (CI**)
All ages	3841	990; 165	1959; 727	52.4% (41, 61.6)
2-17 years*	729	212; 26	402; 89	57.6% (25.1, 76)
18-44 years	1551	486;43	862;160	55.3% (34.2, 69.6)
45-64 years	908	223;49	432;204	55.4% (34.6, 69.5)
65+ years	409	24;39	105;241	29.1% (-31.4, 61.8)

* LAIV Live attenuated influenza vaccine only ** CI 95% Confidence Interval

LAIV is as effective as TIV in 2015/2016

Laboratory-confirmed influenza	Person-years			Adjusted effectiveness (95% confidence intervals)	
Type	Not vaccinated	LAIV	TIV	LAIV	TIV
A and B	29,984	3,965	1,954	50.7% (28.4%–66.1%)	61.2% (30.7%–78.3%)

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Nohynek et al. Euro Surveill. 2016

LAIV effectiveness?

“ACIP votes down use of LAIV for 2016-2017 flu season”

48-53%
2 – 17 year old

CDC Advisory Committee on Immunization Practices, June 2016

“ACIP votes up the use of LAIV for 2018-2019 flu season”

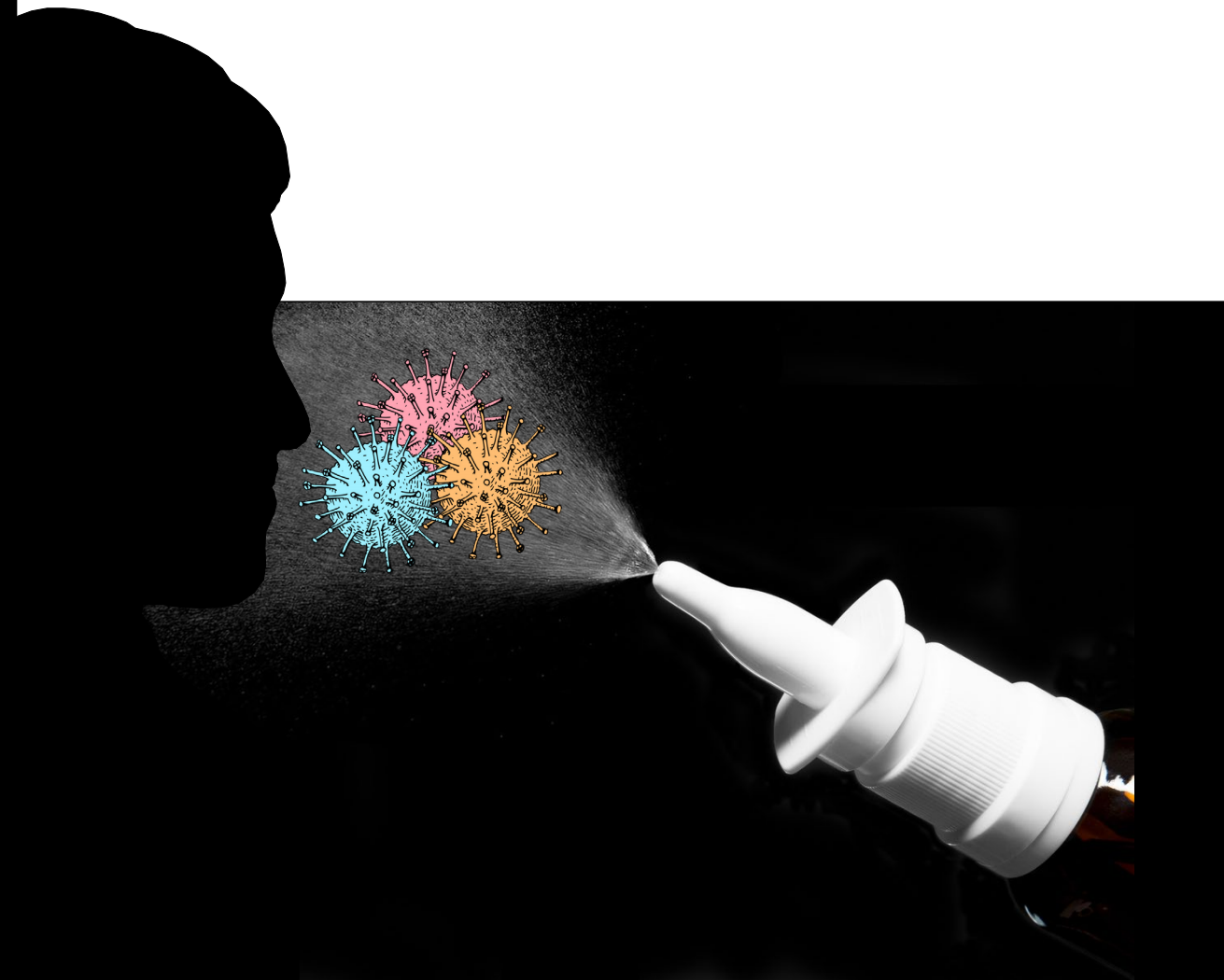
CDC Morbidity and Mortality Weekly Report (*MMWR*) Search

Update: ACIP Recommendations for the Use of Quadrivalent Live Attenuated Influenza Vaccine (LAIV4) — United States, 2018–19 Influenza Season

Weekly / June 8, 2018 / 67(22);643–645

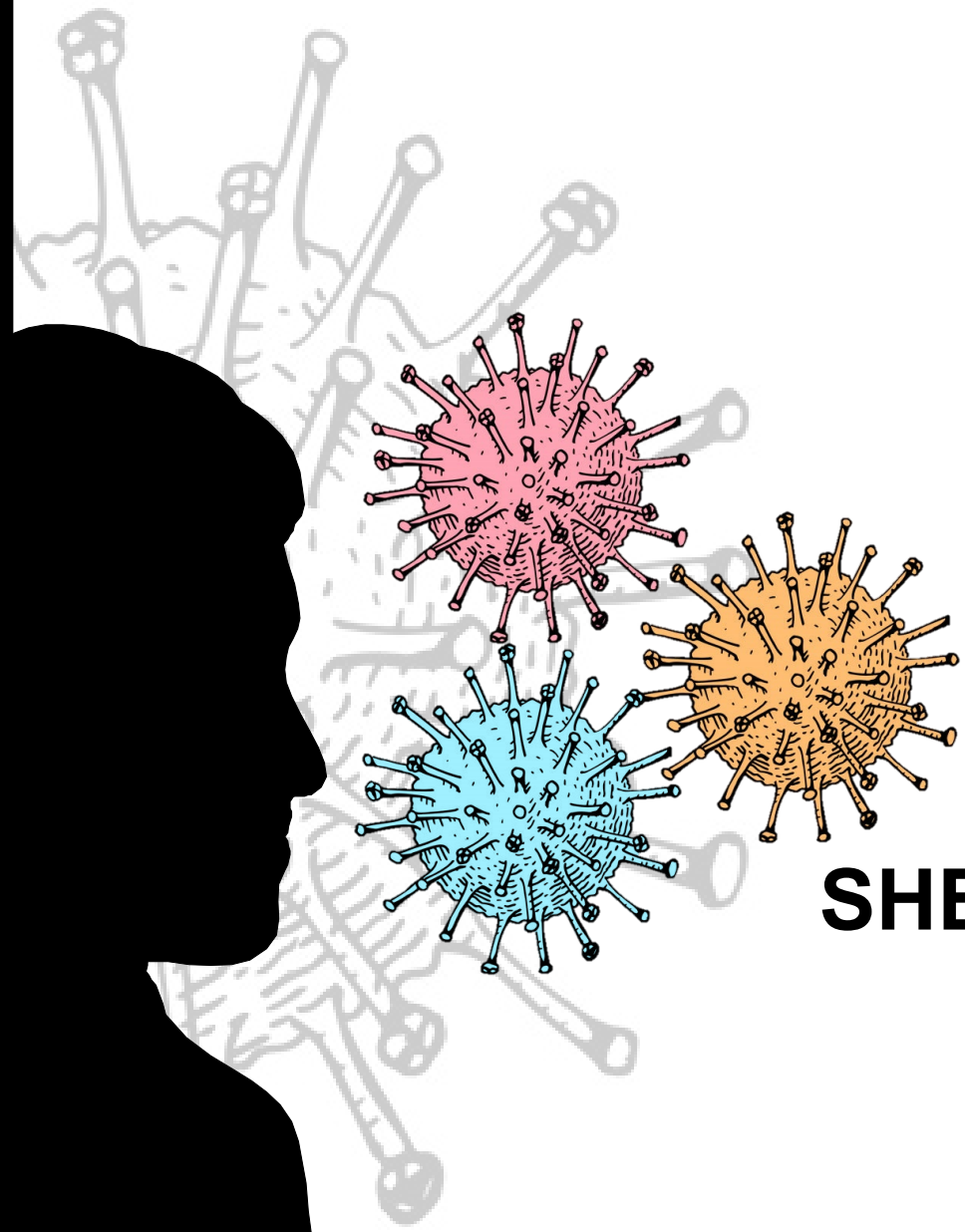
LAIV vaccination

- Intranasal route of infection



LAIV vaccination

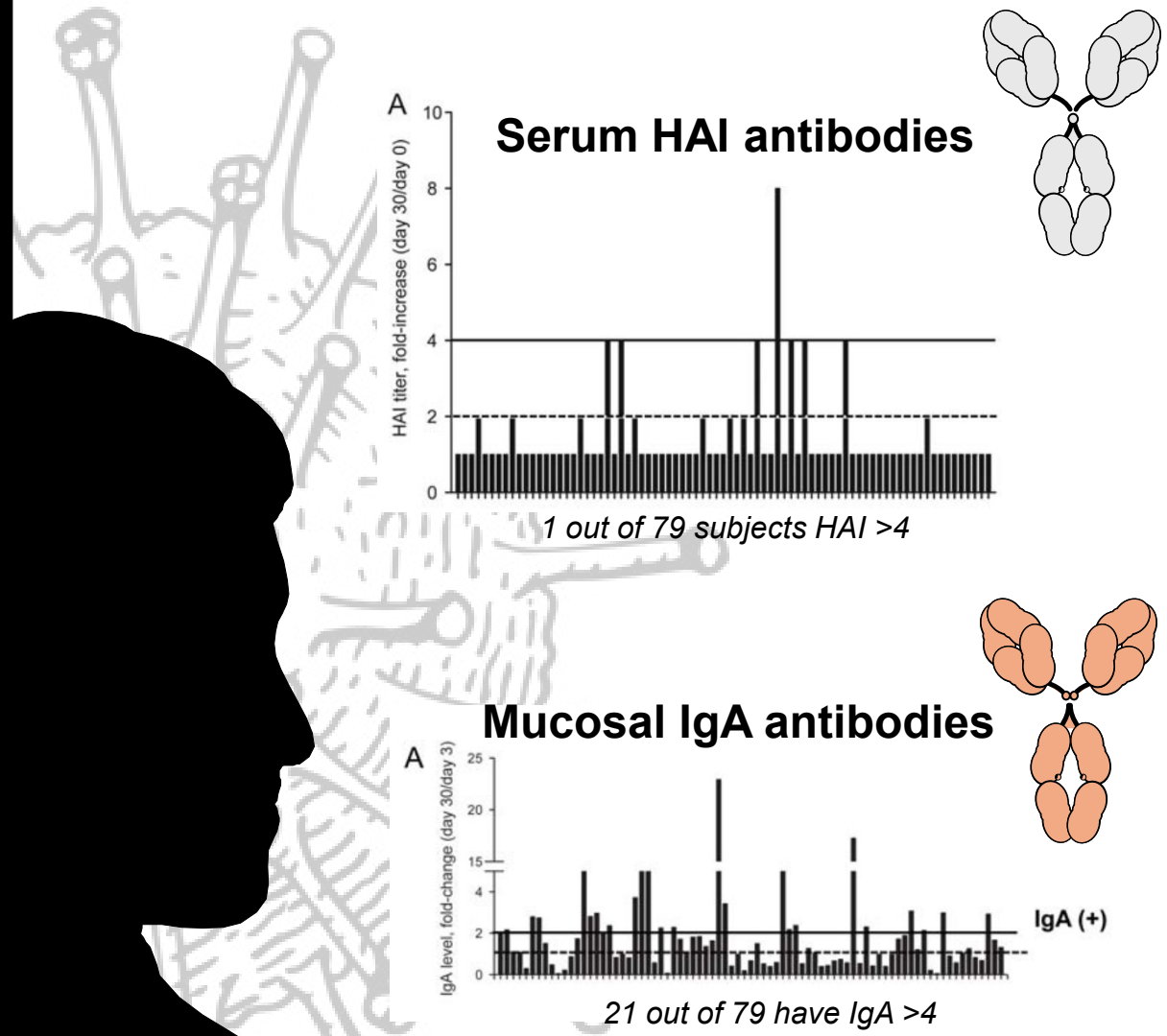
- Intranasal route of infection
- Replication necessary for the immunogenicity



SHEDDING

LAIV vaccination

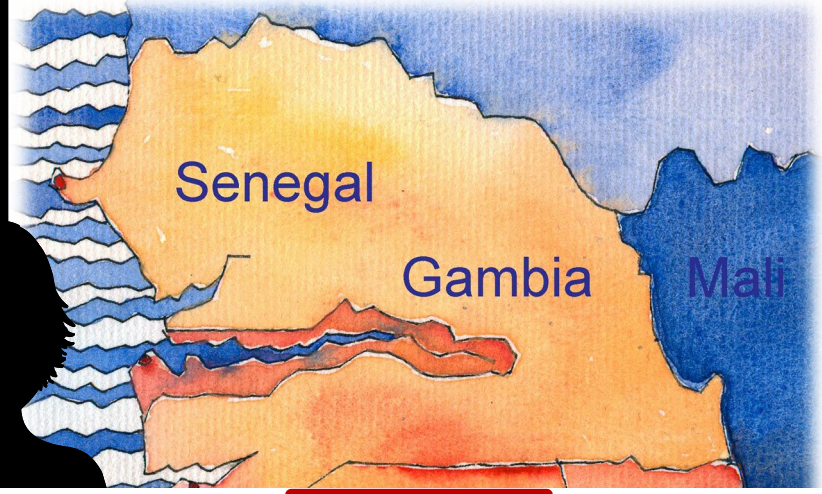
- Intranasal route of infection
- Replication necessary for the immunogenicity
- Lower HAI induction but increased mucosal immunity

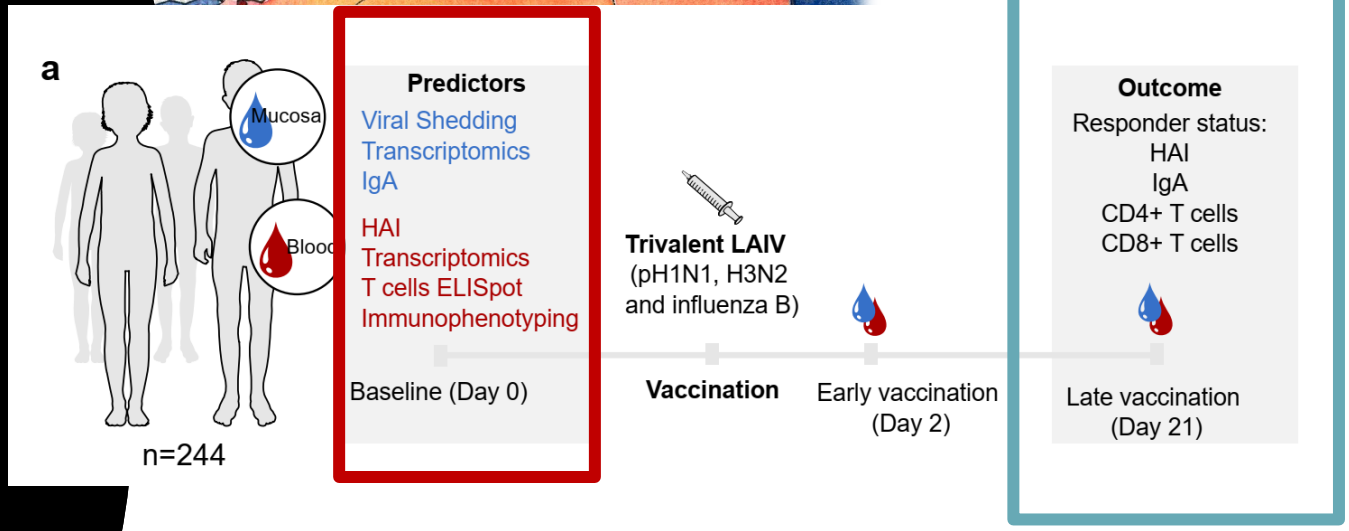


Barrie et al, JID, 2013

LAIV model

- 2 clinical studies in the Gambia
- LAIV studies (Nasovac-S) in 2016/2017 & 2017/2018
- 244 children, 2-5yo



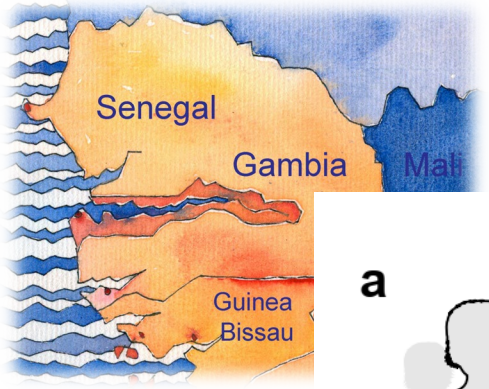


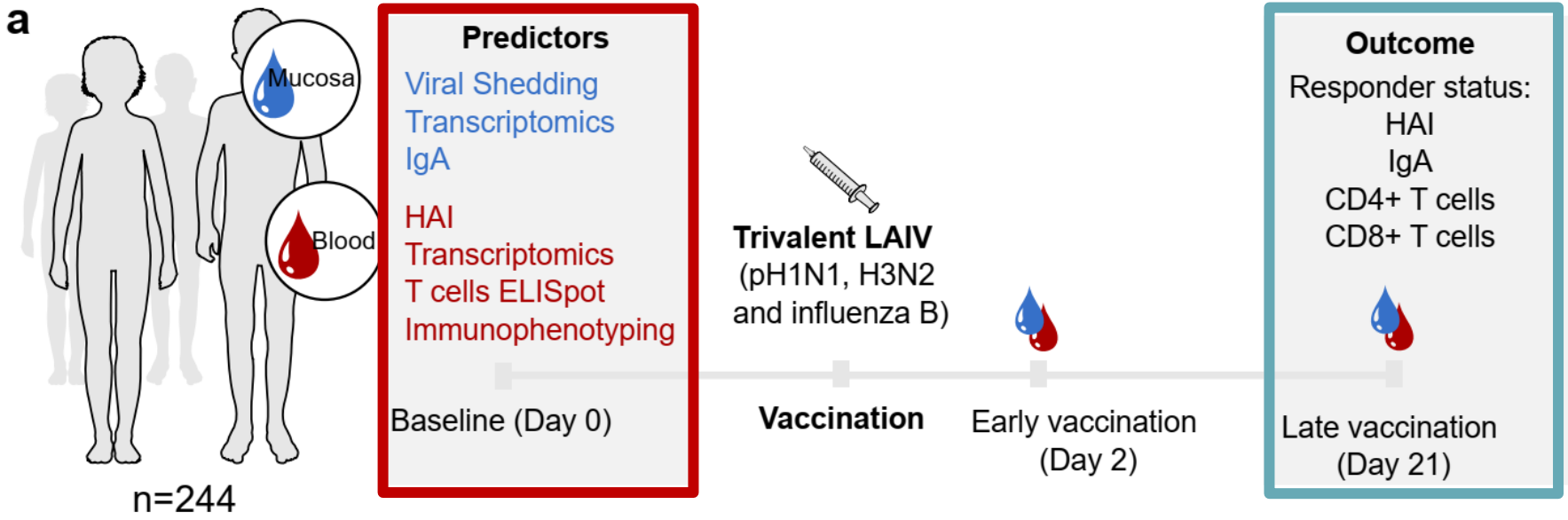
Question 2.
What early factors determine responder status?

Question 1.
How to define responders ?

Collaboration with Thushan de Silva, MRC The Gambia

Part 3: Towards precision vaccinology



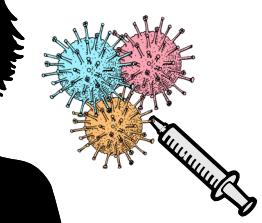


Question 2.
What early factors determine responder status?

Question 1.
How to define responders?

Collaboration with Thushan de Silva, MRC The Gambia; 2 clinical studies in the Gambia; LAIV (Nasovac-S) in 2016/2017 & 2017/2018; 244 children, 2-5yo

How to define responders based on multiple immunological measures?

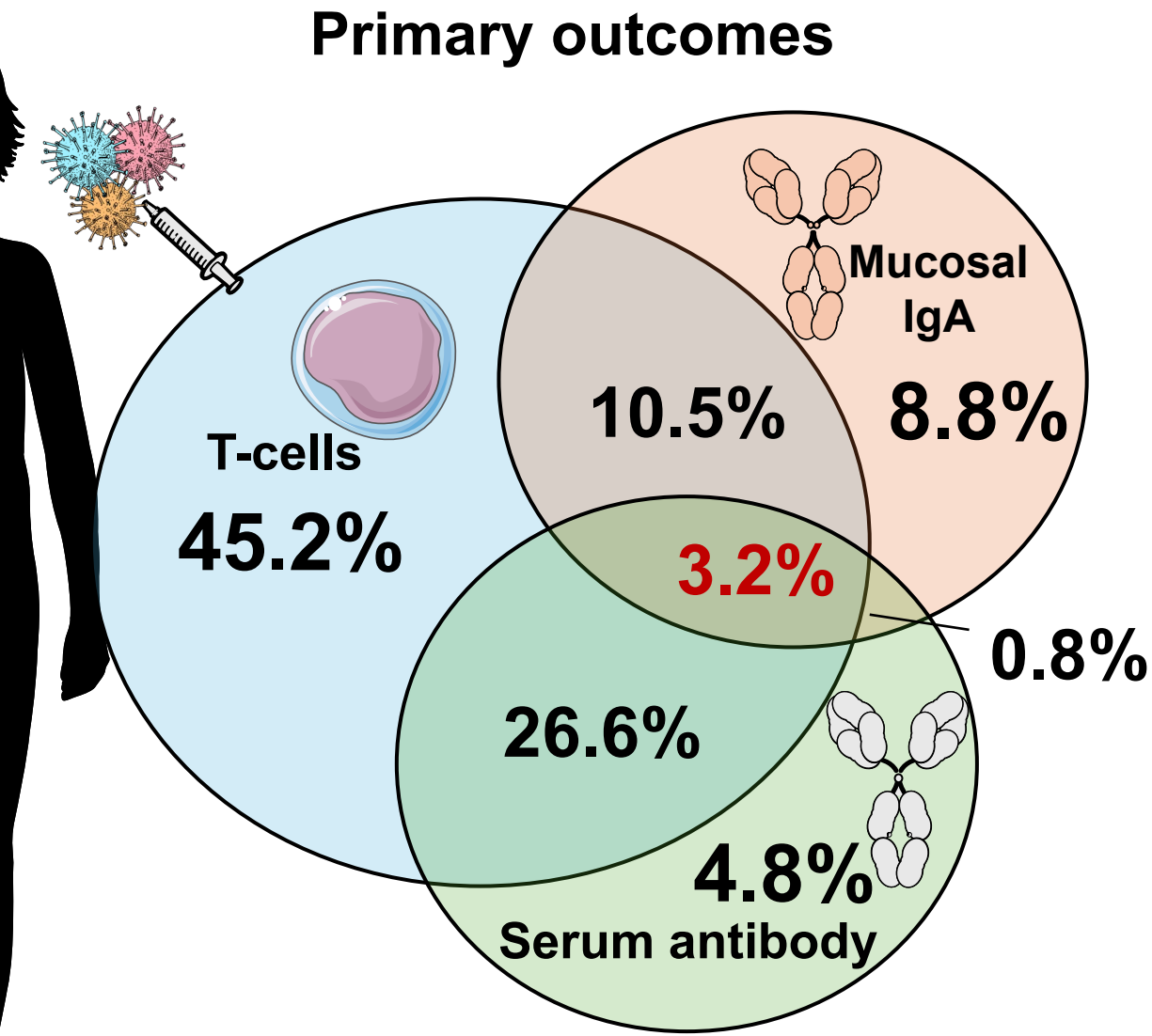


Primary outcomes (21d post-LAIV)

- Serum HAI antibodies
- Mucosal HA-specific IgA
- HA-specific T-cells

How to define responders based on multiple immunological measures?

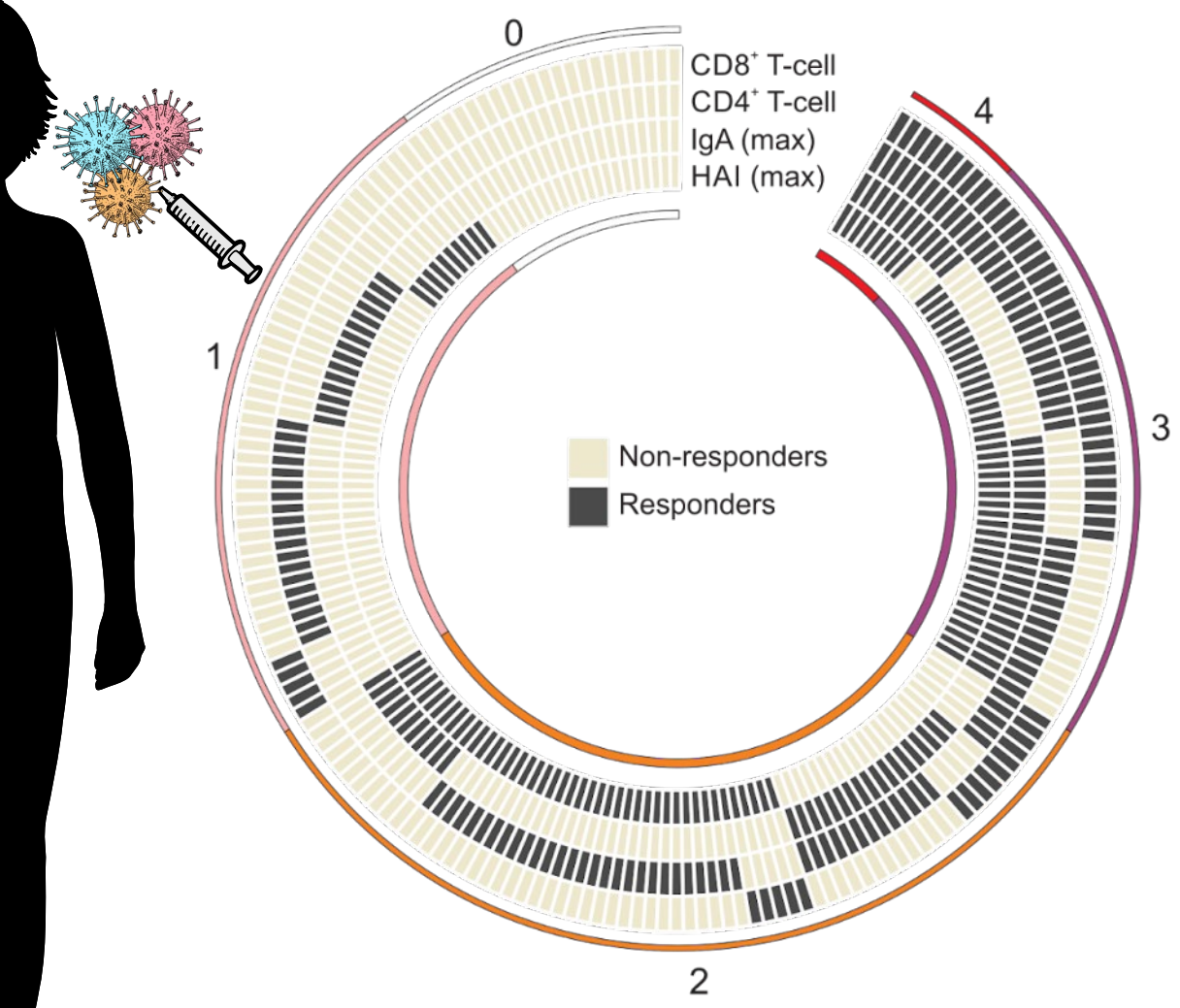
- Responses to LAIV are highly divergent



Responses to H3 HA

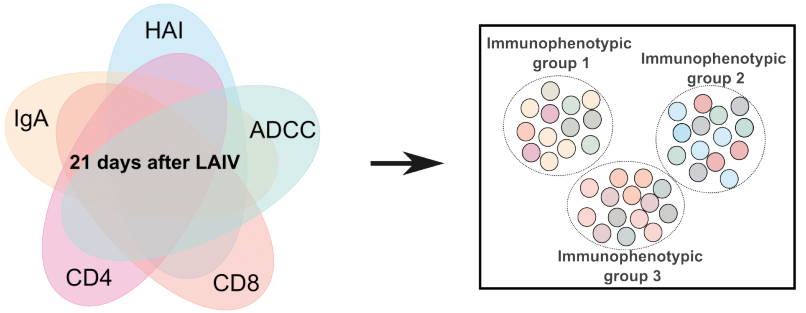
How to define responders based on multiple immunological measures?

- Responses to LAIV are highly divergent
- **Multi-functional responders to LAIV identified**

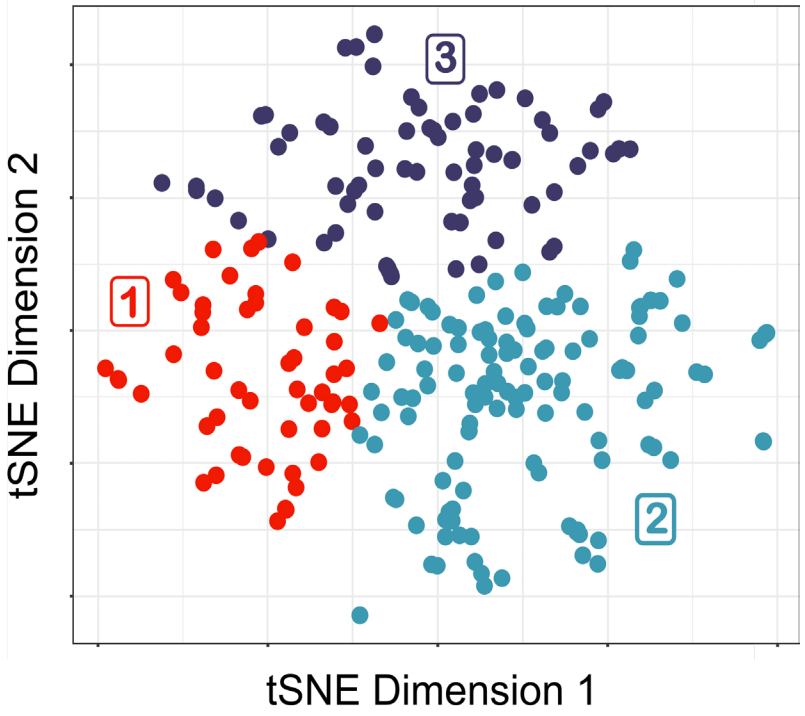


How to define responders based on multiple immunological measures?

- Responses to LAIV are highly divergent
- Multi-functional responders to LAIV identified
- **Unsupervised ML approach: t-SNE**



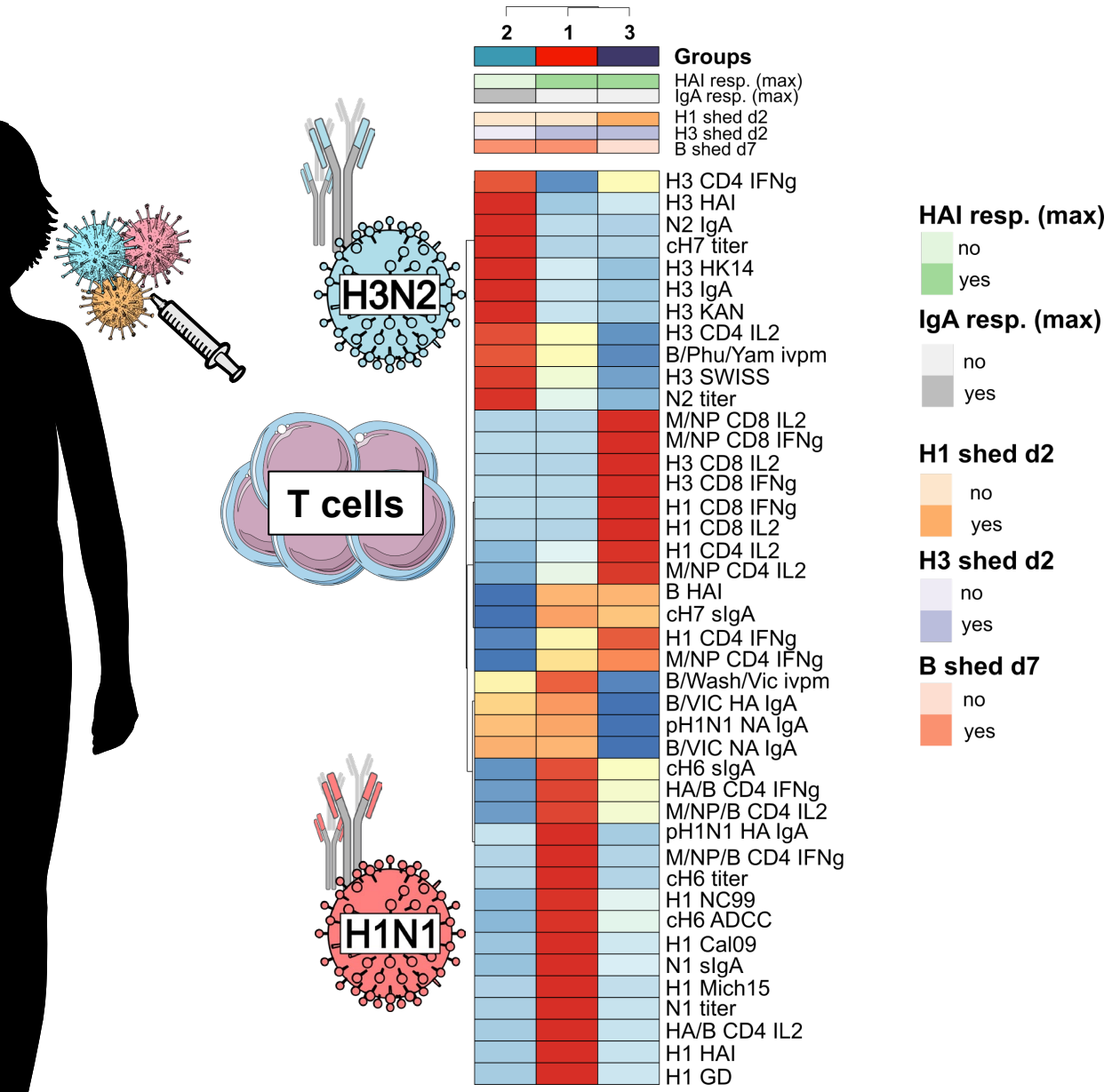
Individuals are separated into 3 groups based on the responses to LAIV



Mucosal/serum Ab and T cells responses to H1N1, H3N2 and B influenza types analyzed (21d post-LAIV)

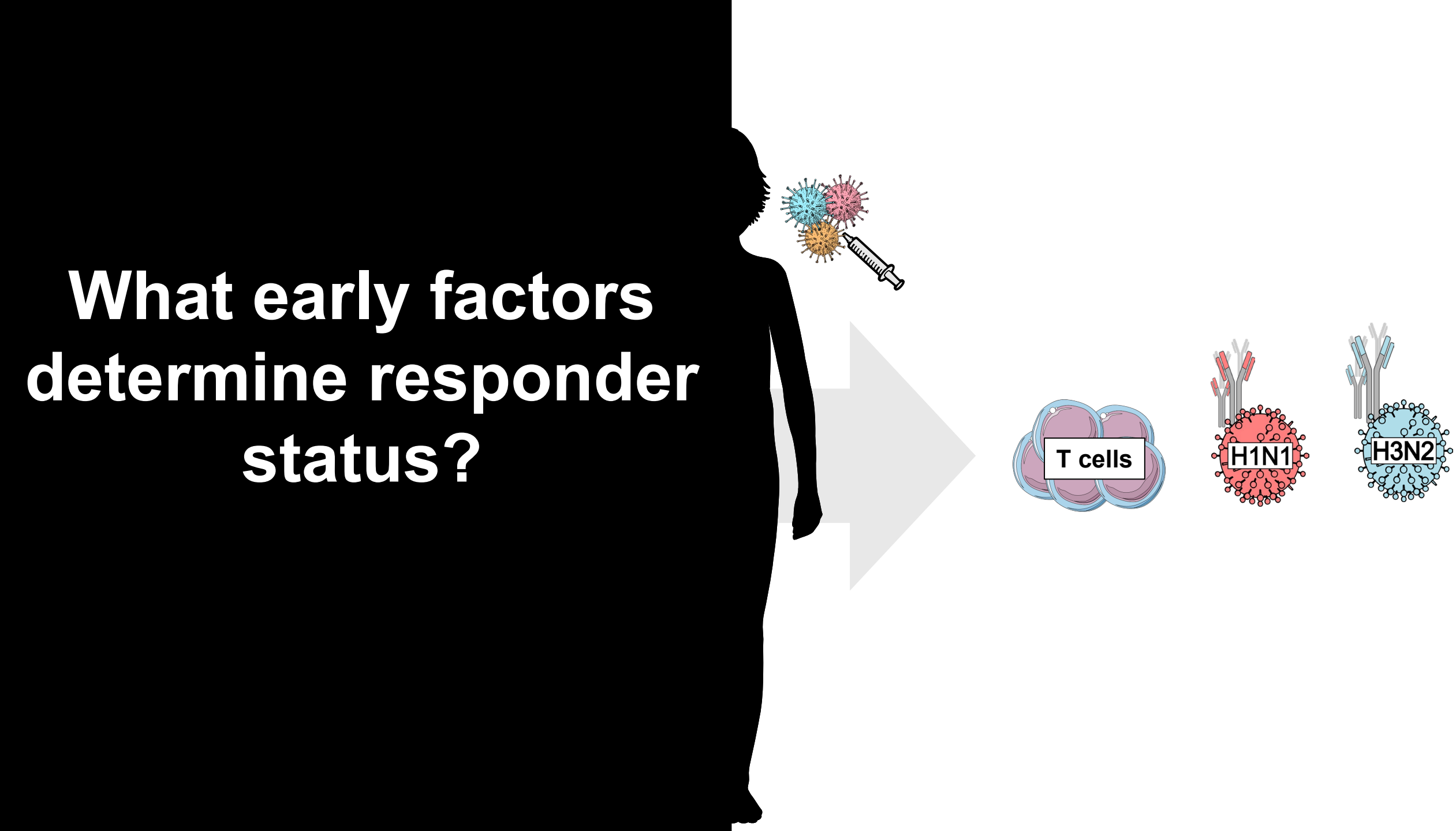
How to define responders based on multiple immunological measures?

- Responses to LAIV are highly divergent
- Multi-functional responders to LAIV identified
- Unsupervised ML approach: t-SNE
- **3 group of responders identified**



IVPM* - *Signal by microarray;* ***N1/N2 titer* *– N1 and N1 NA Ab in serum;* ***cH6/cH7*** *– Groups 1/2 stalk specific serum antibody;* ***cH6/H7 sIgA*** *- Groups 1/2 stalk specific secretory antibody;* ***cH6 ADCC AUC*** *- Group 1 stalk specific ADCC activity*

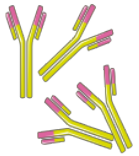
What early factors determine responder status?



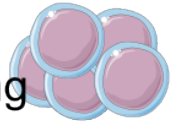
What early factors determine responder status?

Baseline samples

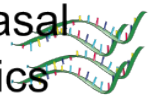
Serum and mucosal antibodies

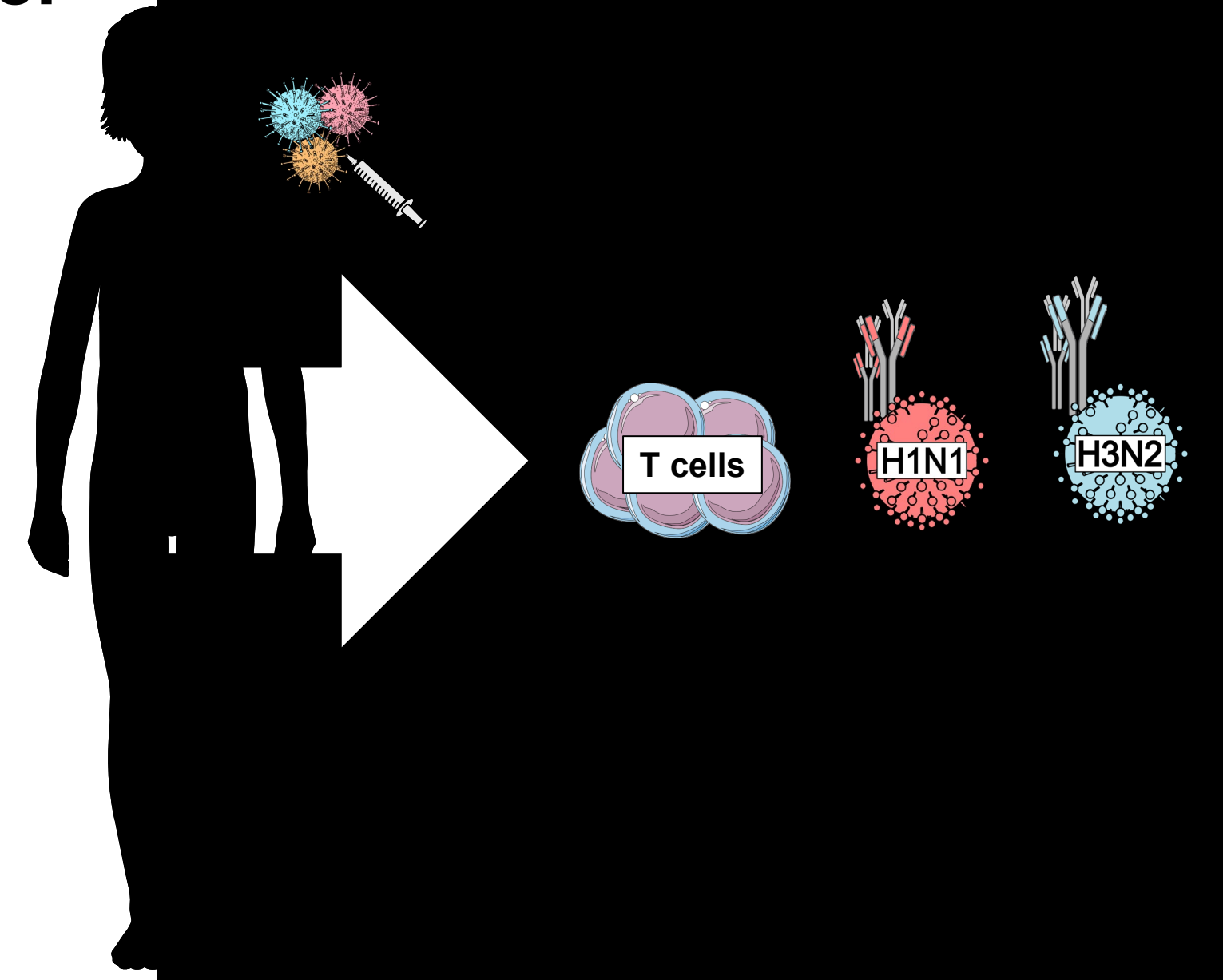


Immuno -phenotyping

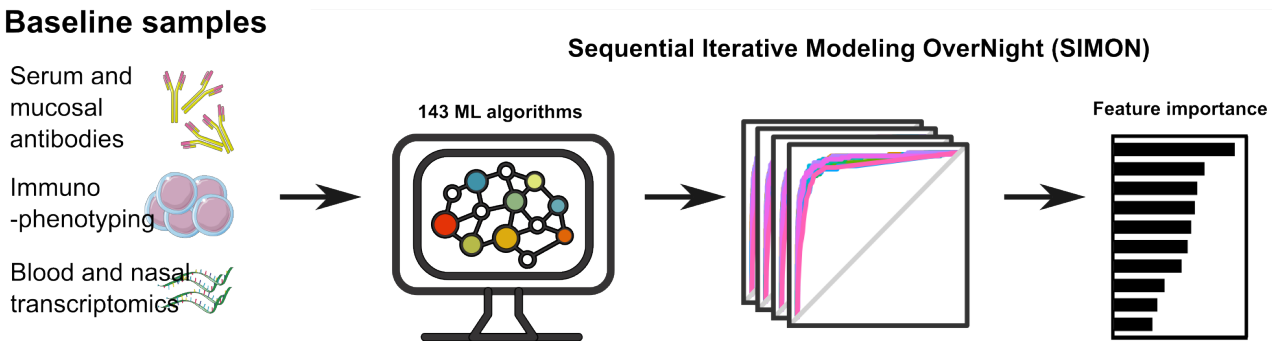


Blood and nasal transcriptomics





What early factors determine responder status?

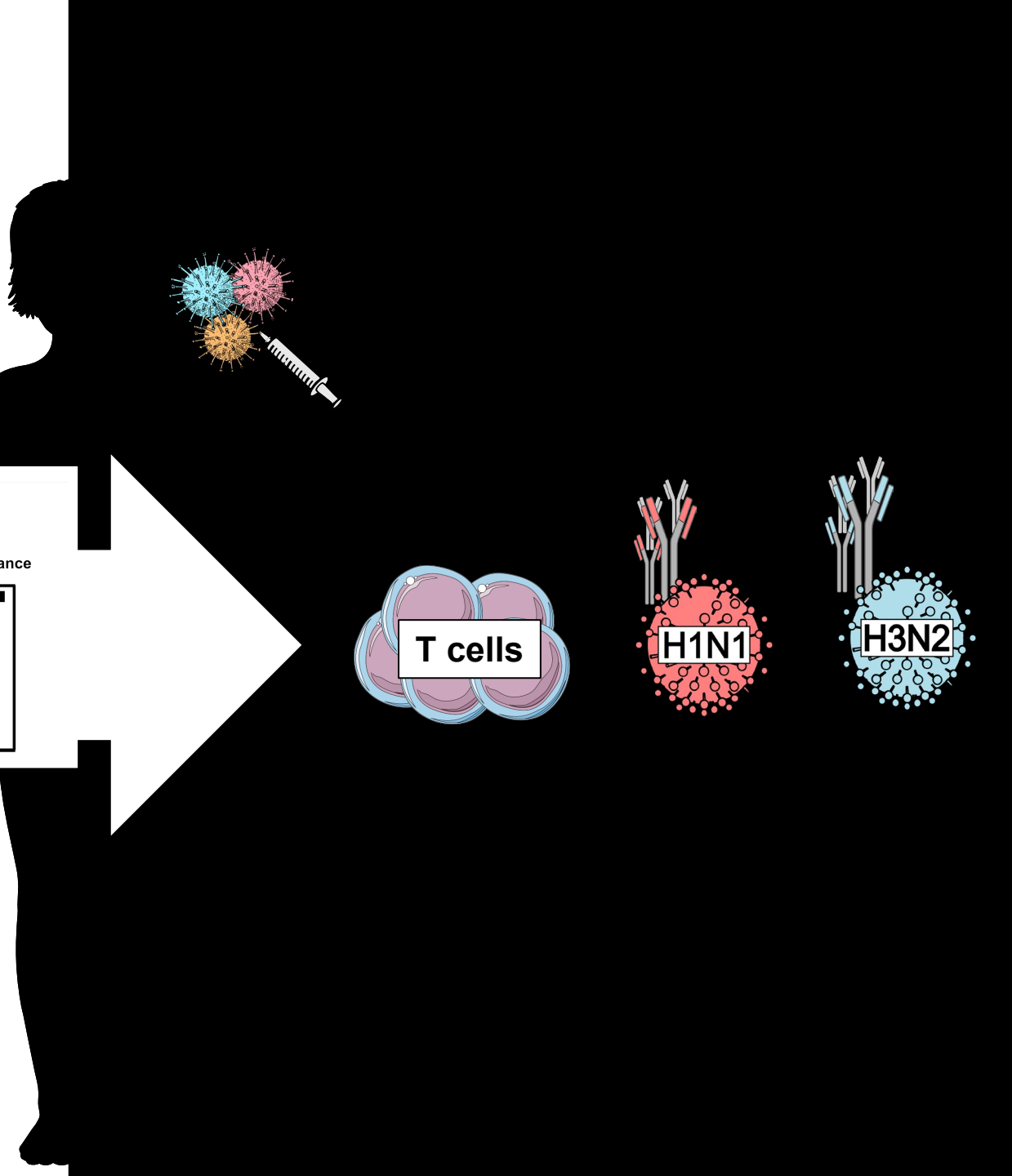


→ 143 ML algorithms tested

→ 69 models successfully built

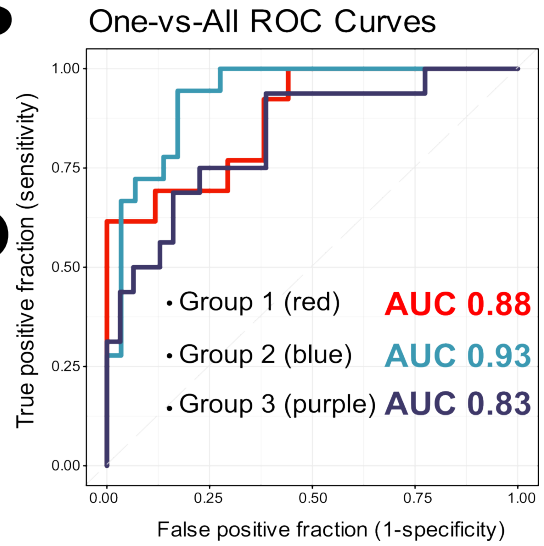
→ 52 models with AUC >0.7

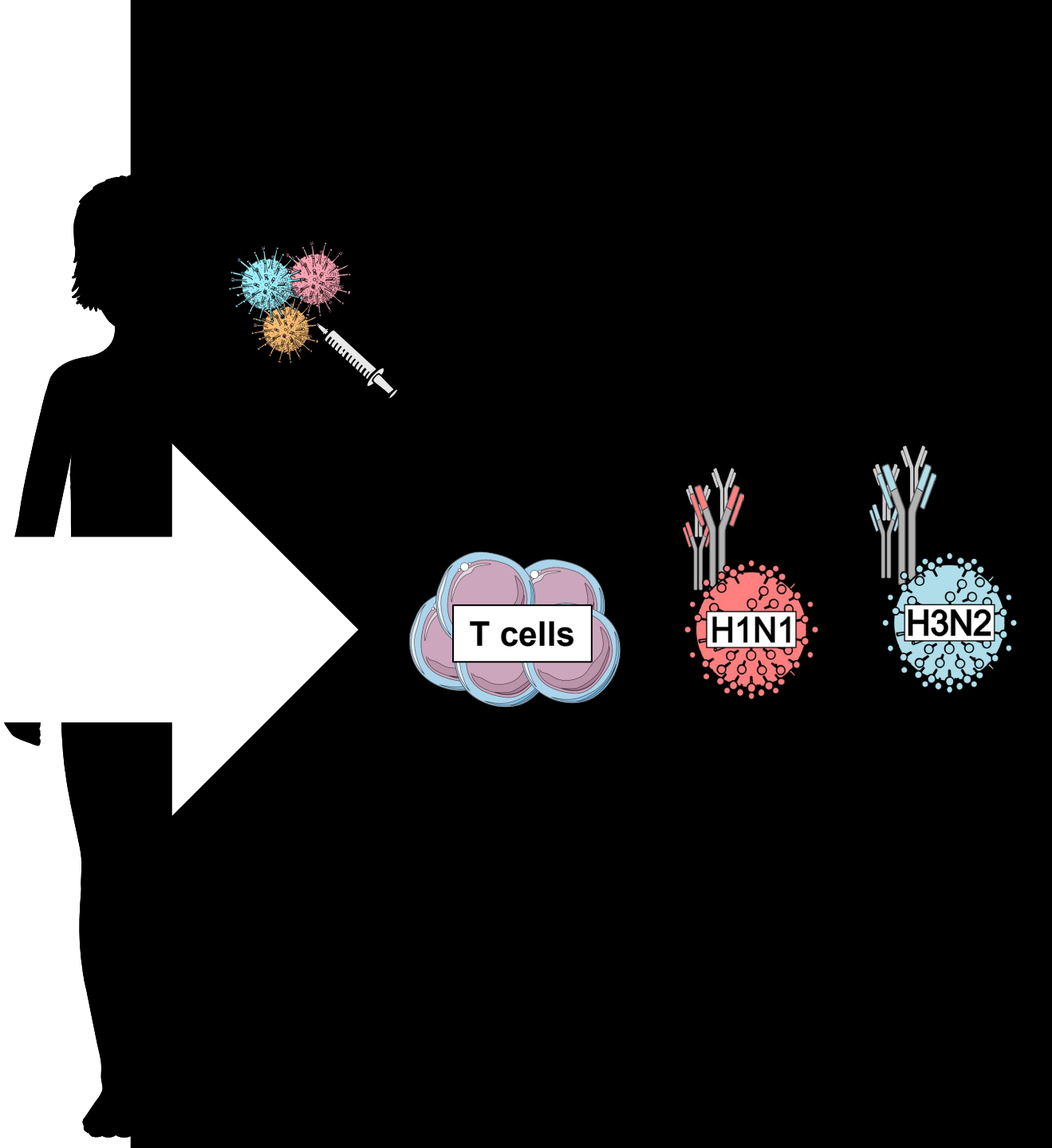
→ 9 ML models not overfitted



What early factors determine responder status?

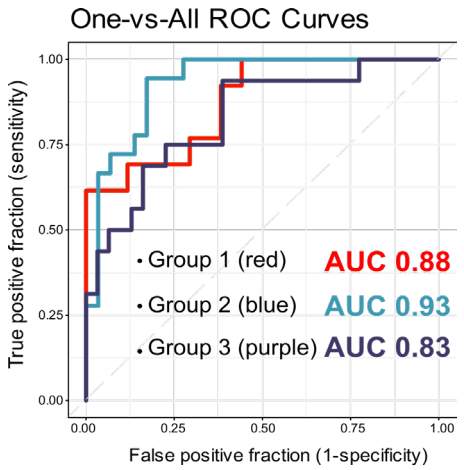
- **SIMON builds predictive model able to discriminate between 3 groups of responders**

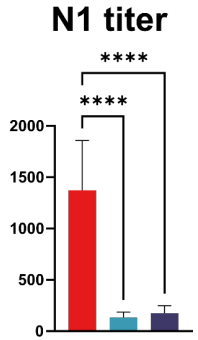


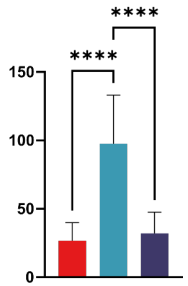


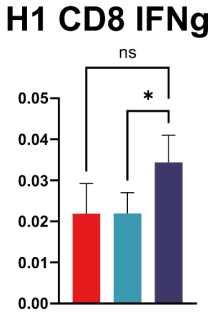
Method. *SIMON; predictors: baseline measurements, outcome: identified clusters; preprocessing (center, scale, medianImpute, corr, zv, nzv); data split: train 80%, test 20%; training performance: 10-fold cross-validation repeated 3 times; test performance: confusion matrix and AUC on test set*

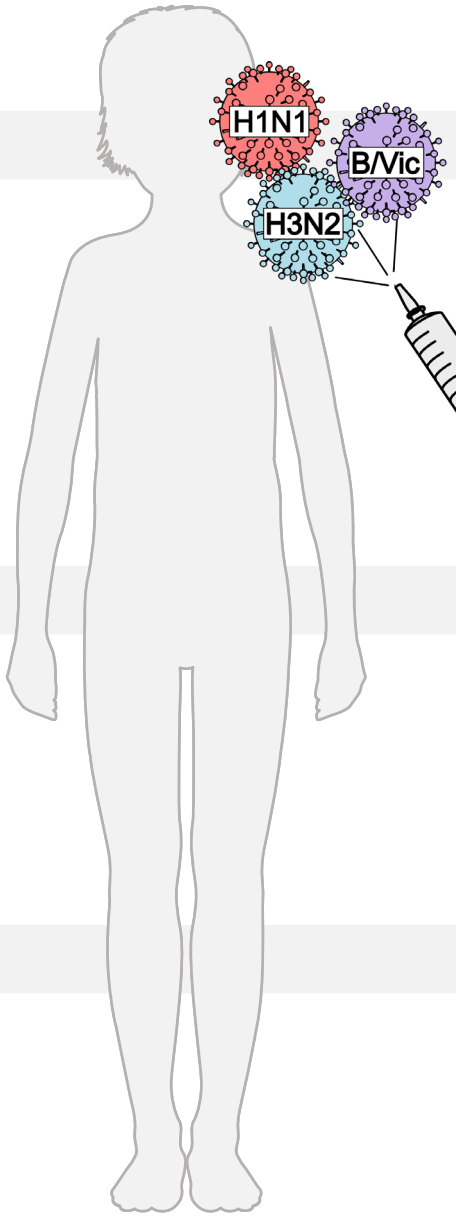
Strain-specific immunity dictates LAIV responses

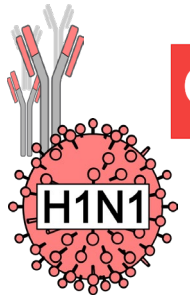




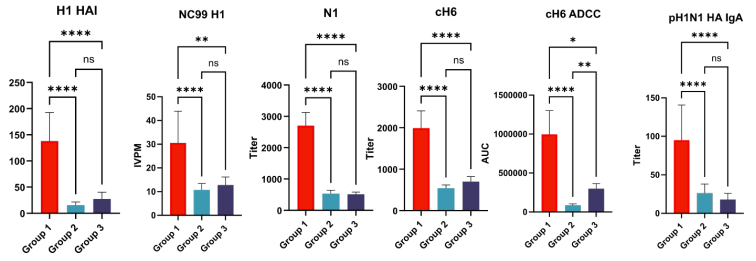




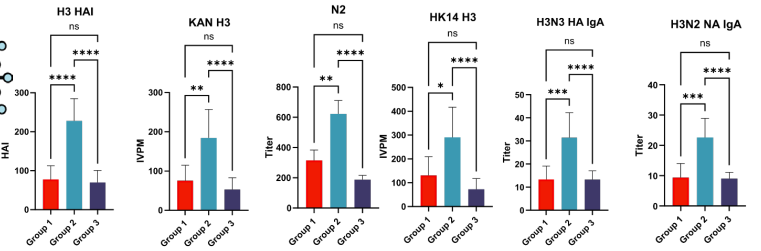


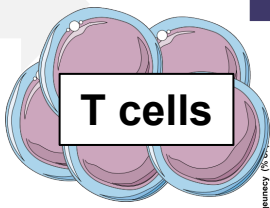


Group 1 - H1N1 responders

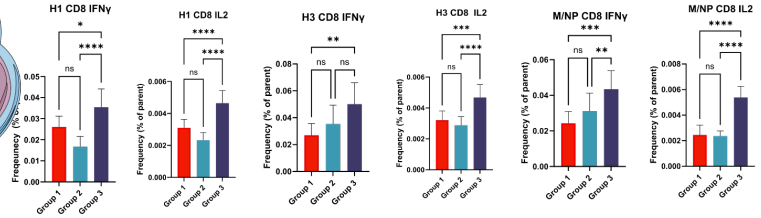


Group 2 - H3N2 responders



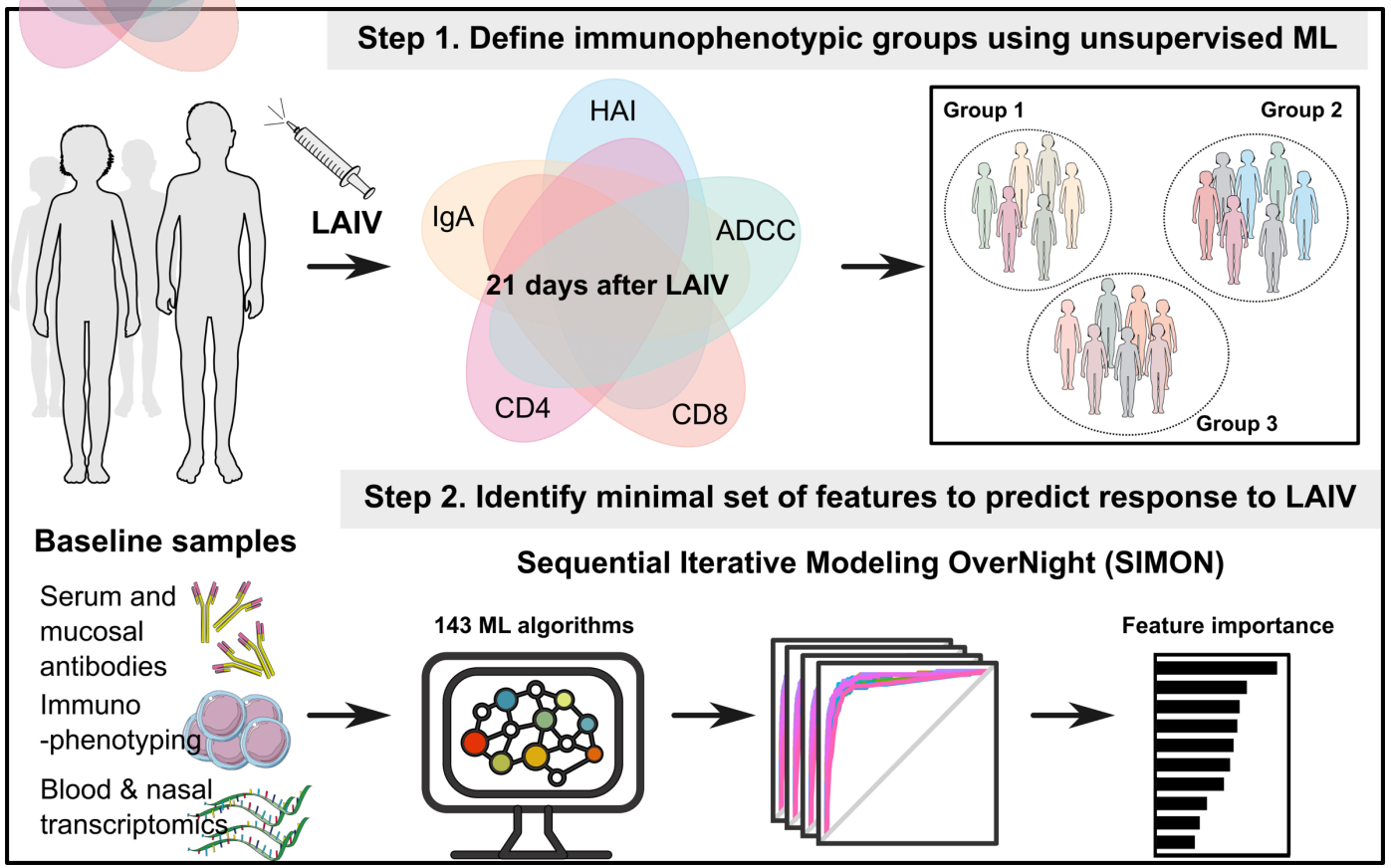


Group 3 - T cell responders



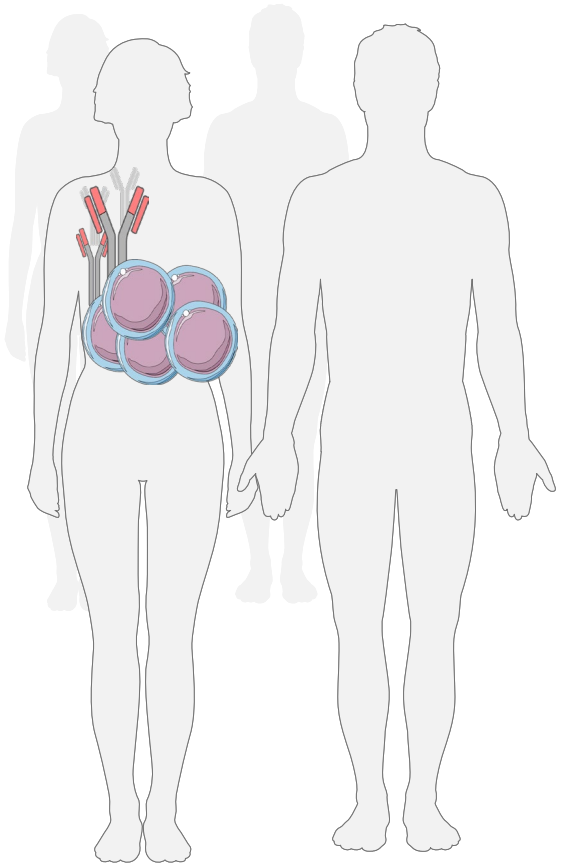
Part 3: Towards precision vaccinology

Immunaut: Unsupervised approach for understanding vaccine responses

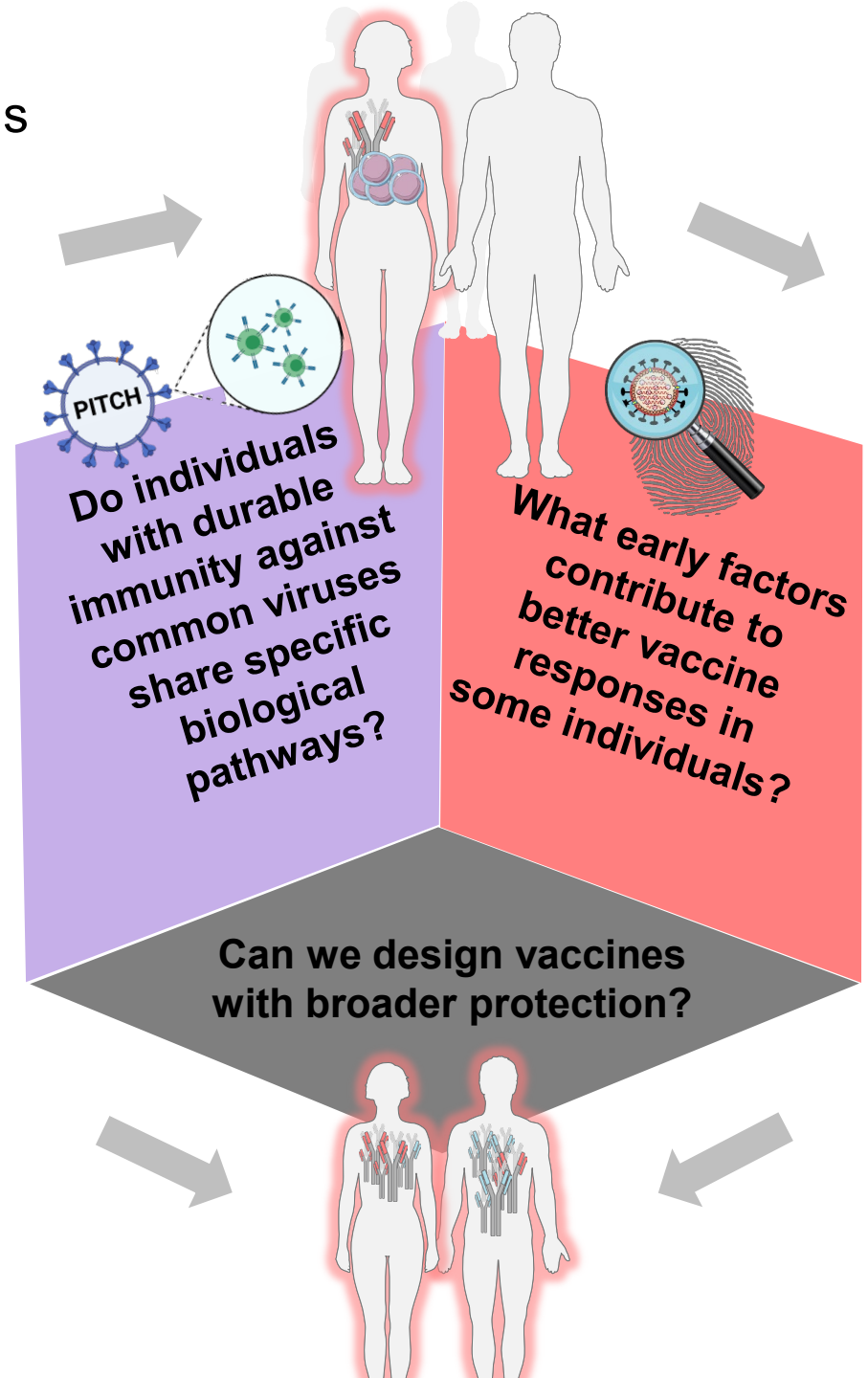


Tomic et al, unpublished

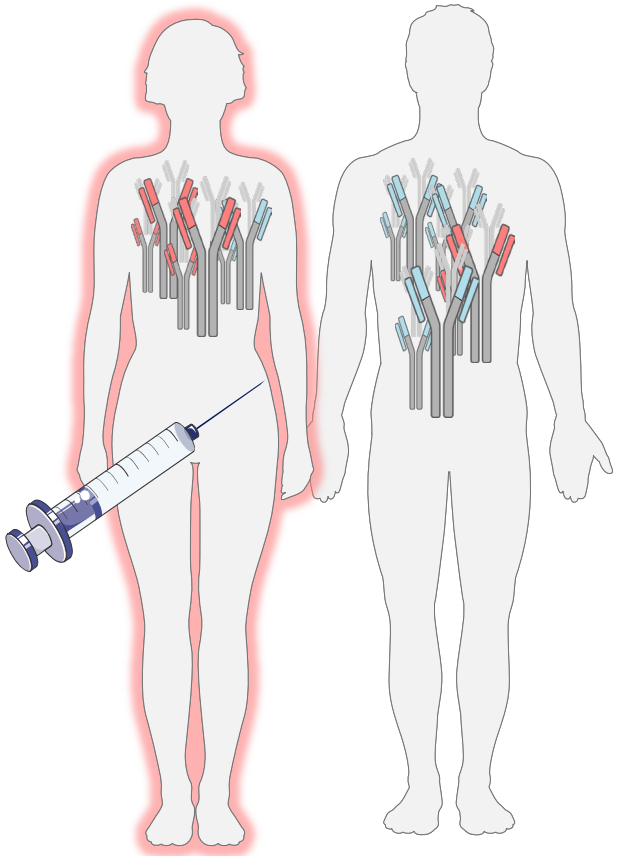
Durable immunity:
Early immune teamwork leads to longer memory



Tomic *et al, Nat Comms, 2022;*
COMBAT consortium, Cell, 2022;
Hornsby et al, Nature Comms, 2023

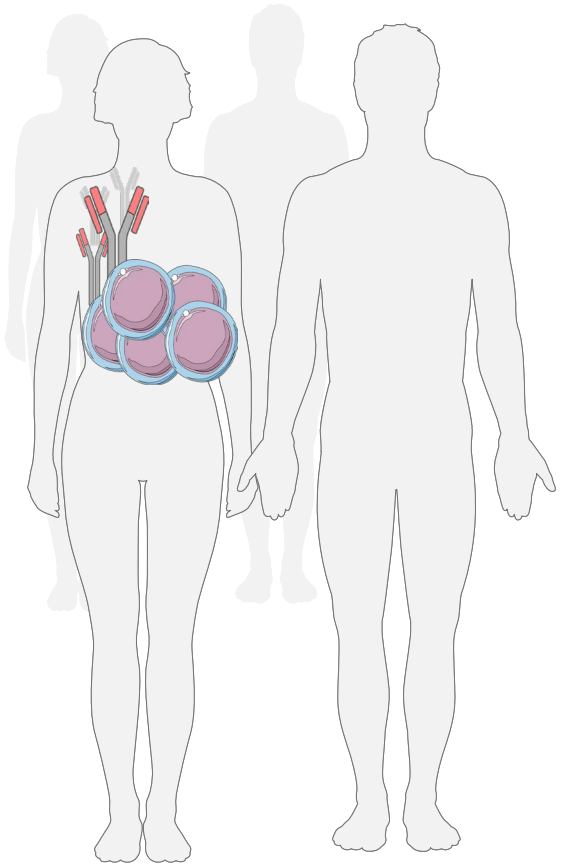


High responders:
More memory, stronger immunity

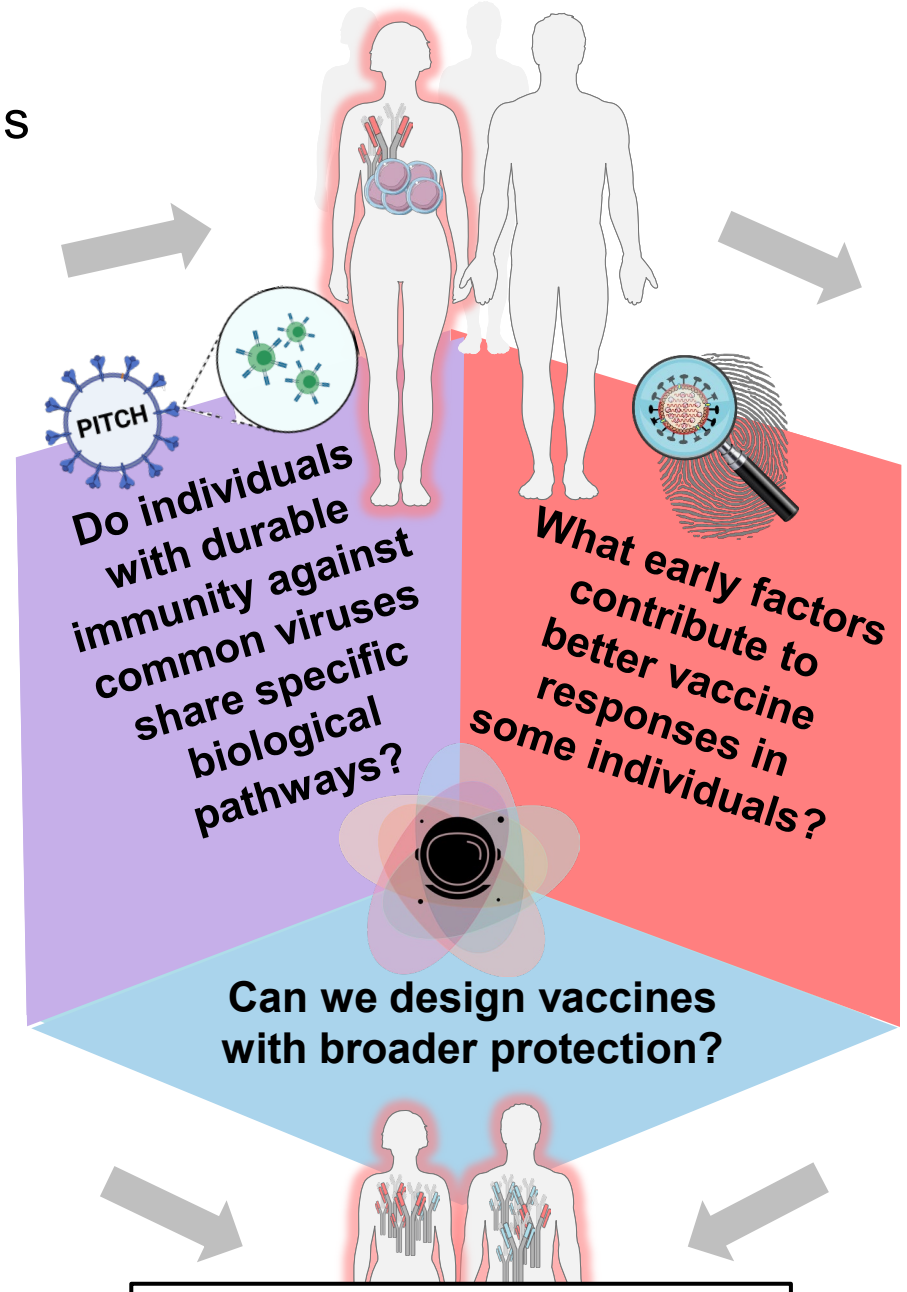


Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

Durable immunity:
Early immune teamwork leads to longer memory

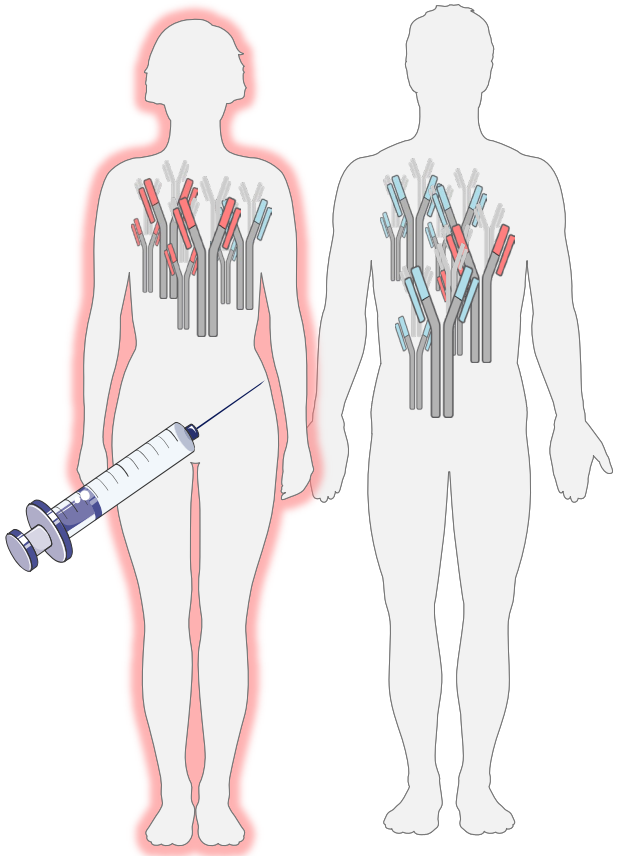


Tomic *et al, Nat Comms, 2022;*
COMBAT consortium, Cell, 2022;
Hornsby et al, Nature Comms, 2023



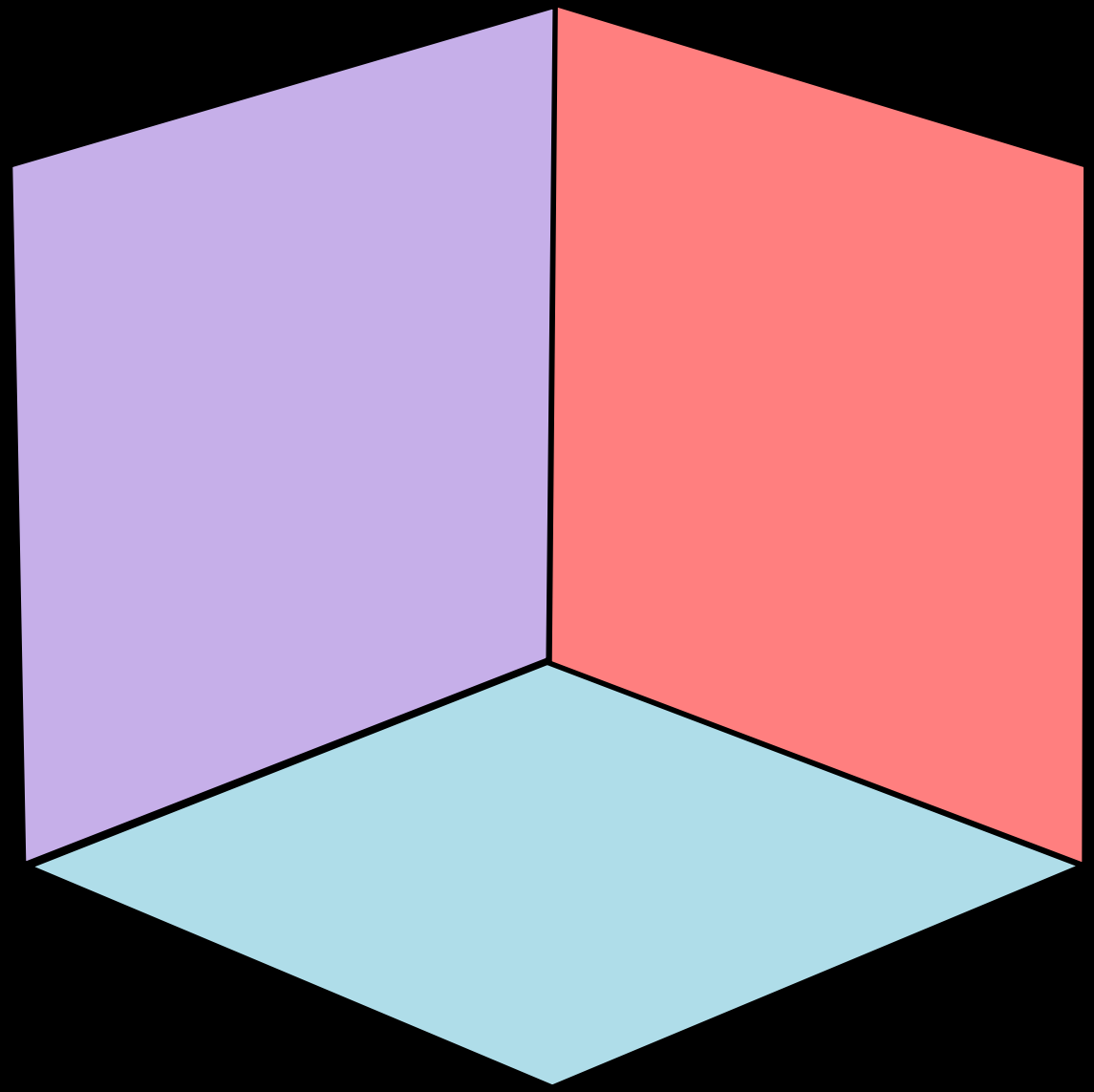
Superior immunity:
Tailored vaccination

High responders:
More memory, stronger immunity



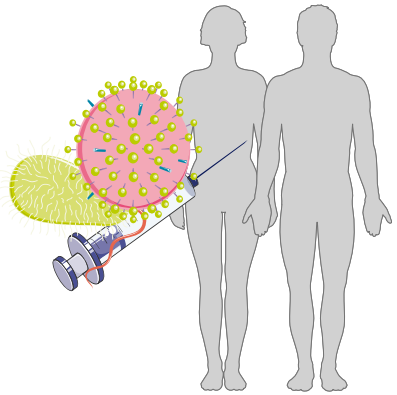
Tomic *et al, JI, 2019;*
O'Connor et al, Mol Syst Biol, 2020
Tomic *et al, Cell Patterns, 2021;*
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clin & Exp Immunology, 2024

PANDORA

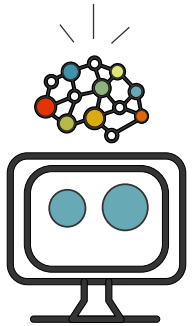


PANDORA: Free & Open-source ML Platform

Predicting immune responses

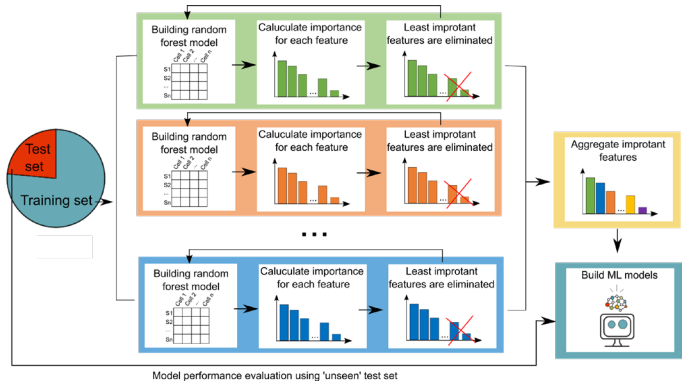


Why do some people develop more robust responses to vaccines and pathogens?



Tomic et al, JI, 2019; Tomic Sci Data 2019;
Tomic et al, Patterns, 2021;
O'Connor et al, Mol Syst Biol, 2020;
Stockdale et al, Frontiers Ana Sci, 2022;
Ali et al, Clinical & Exp Immunology, 2024

Predicting disease severity

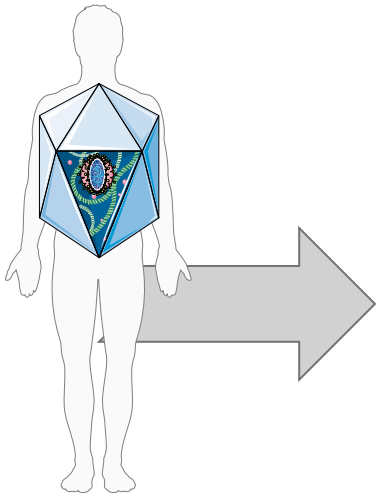


COMBAT
Oxford COVID-19 Multi-omic Blood ATlas

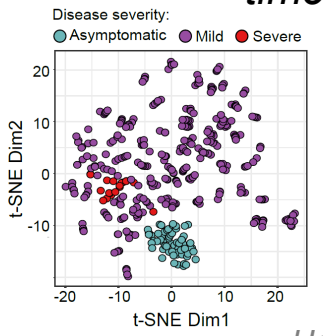
Why do people develop more severe diseases?

COMBAT consortium, Cell, 2022;
Kotanidis et al, Lancet Digital Health, 2022

Predicting duration of protection

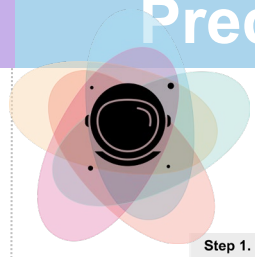


Why are some people protected for more extended time after infection?

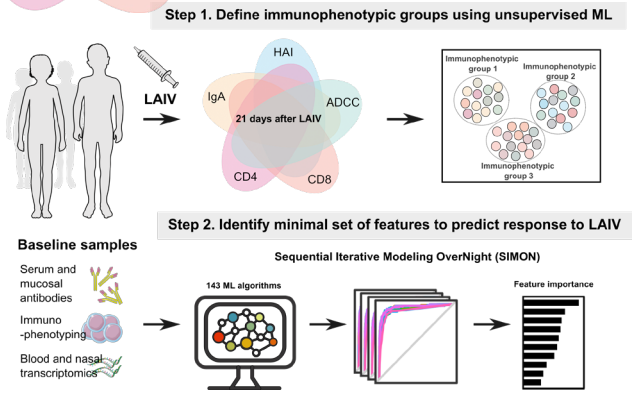


Tomic et al, Nature Comms, 2022;
Hornsby et al, Nature Comms, 2023;
Neale et al, mBio, 2023

Predicting correlates of protection



Can we capture individual variability in their immune responses to vaccines?



Tomic et al, unpublished, 2024

Leading statistical programming languages in data science – available ML tools

R-project (https://www.r-project.org/):

- Mlr3 (https://mlr3.mlr-org.com)
- Classification and regression training (**CARET**) (https://rdrr.io/cran/caret)

Python (https://www.python.org/):

- Scikit-learn (https://scikit-learn.org)
- mlPy (https://mlpy.fbk.eu)
- SciPy (https://www.scipy.org/)

Extensive programming experience and general knowledge of R or Python **essential**, making them inaccessible for many life science researchers

Deep learning libraries:

https://www.tensorflow.org/

https://keras.io/

Available ML software

Commercial software

- Google's cloud-based AutoML (https://cloud.google.com/automl)
- DataRobot (https://www.datarobot.com/)
- BigML (https://bigml.com/)
- MLjar (https://mljar.com)
- RapidMiner (https://rapidminer.com/)

Features

- Closed source – unknown/hidden ML methods and algorithms
- No specific algorithms to deal with biomedical datasets (missingness, heterogenous data types, etc)
- High price (DataRobot$50k/licence!)

Academia-released software

- Waikato Environment for Knowledge Analysis (WEKA) (https://www.cs.waikato.ac.nz/~ml/weka/),
- Orange (https://orange.biolab.si/)
- Konstanz Information Miner (KNIME) https://www.knime.com/)
- ELKI (https://elki-project.github.io/)

Features

- Free and open source – explained/published ML methods and algorithms
- Requires knowledge of ML process
- Lack some of the advance features of commercial software (autoML)

PANDORA