MOLECULAR BIOMARKERS FOR PREDICTION, DIAGNOSIS AND MONITORING OF ACUTE MULTIPLE SCLEROSIS RELAPSE

NEUROGENOMICS LABORATORY
MULTIPLE SCLEROSIS CENTER
SHEBA MEDICAL CENTER, ISRAEL

GUREVICH M. KHAVKIN Y.

DEFINITION OF MEDICAL PREDICTIONS

- **Diagnosis :** Predicts if the patient currently has the
- disease
- **Risk:** if subject will exhibit incident disease over a specified time interval.
- **Prediction:** Predicts a future disease-related outcome in a patient with the disease, over a specified time interval

- **Natural history:** Probability of a disease-related outcome if the patient does not receive any therapy/intervention, over a specified time interval

PREDICTIVE FACTORS AND BIOMARKERS

Predictive factors and biomarkers:
any measureable attribute of an individual that can be used to infer a health-related outcome.

The most important how strongly biomarker connected to the disease
Direct - biomarker is a part of the causal disease process
Indirect - involved but it is not necessary
Weakest – not involved (byproduct of the disease process)

Causal biomarkers most probably are predictive, but not all predictive factors are causal

Disease related biomarkers more likely predict disease related outcome and they are the best targets for intervention

CLINICAL USES OF PROGNOSTIC BIOMARKERS.

A natural history prognostic biomarker
The goal is to determine if a therapy is necessary, (does the patient have a sufficiently high likelihood of a good/poor outcome that a therapy should be considered).

A therapy-specific prognostic biomarker
The goal of a therapy-specific prognostic biomarker is to determine the optimal therapy.

A post-therapy prognostic biomarker (surrogate biomarkers)
The goal of a post-therapy prognostic biomarker is to determine whether the therapy was effective without waiting for the occurrence of a clinical outcome.

PREDICTING IS NOTHING ELSE, BUT TELLING THE FUTURE – BASED ON YOUR HISTORICAL DATA.

METHODOLOGY

MACHINE LEARNING ALGORITHM VERSUS TRADITIONAL PROGRAM

Prediction performed by applying Machine Learning Algorithm to large historical data (BIG DATA)

Traditional programs take data as input and produces data as output.

Machine learning algorithm takes data as input but produces a program as an output.
This machine generated program can now take new unknown data, process it and produce output data (prediction results) .

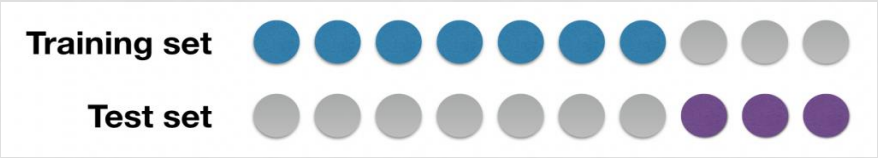
MOST IMPORTANT STEPS

Step 1 – Select outcome variable
Numerical or categorical (this will define predictive algorithm)

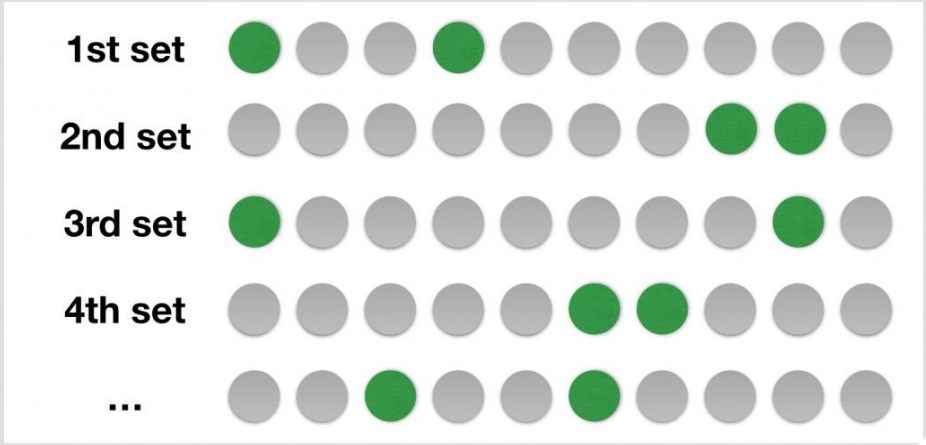
Step 2 – Collect historical data set
Step 3 – Split your data

Training set: the dataset that we use to teach model.

Independent Test set: the dataset that we use to validate model before using it on real life future data.

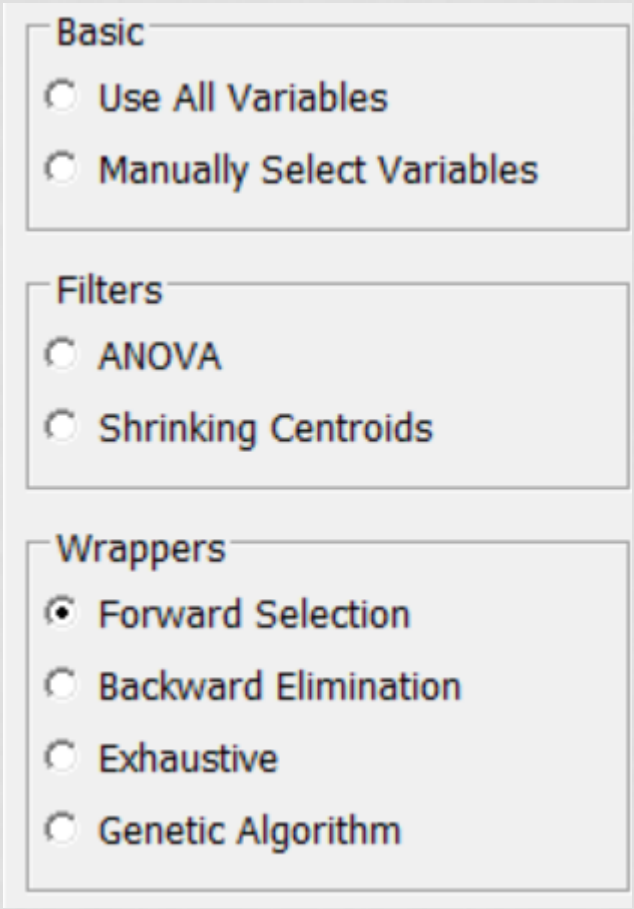


Monte Carlo cross-validation

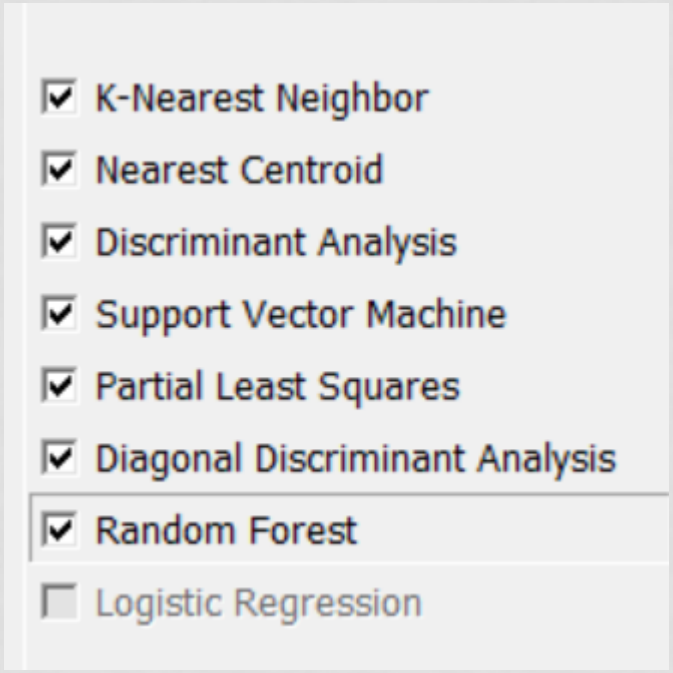


SAMPLES OF CLASSIFICATION ALGORITHMS

Variable Selection

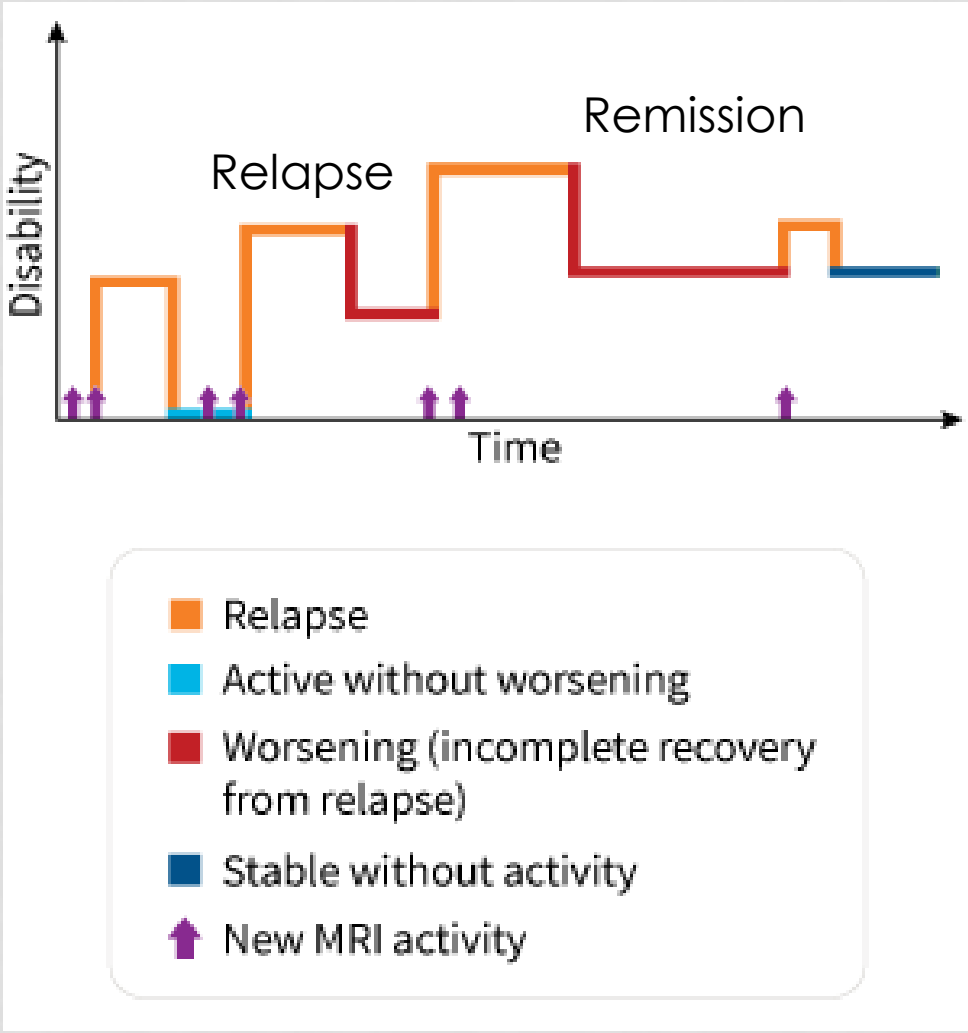


Classification



Predictive models itself is not ready for use It need a choices on when, to whom and how to apply predictions

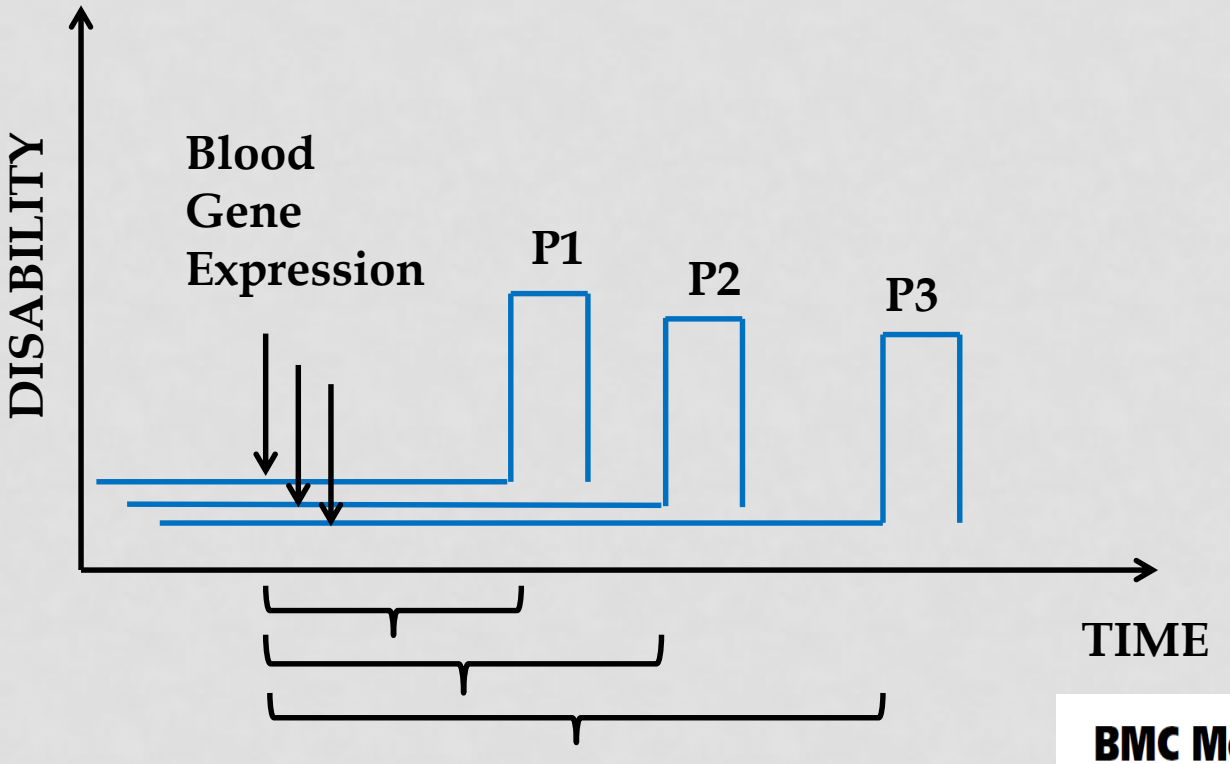
PREDICTION OF ACUTE MULTIPLE SCLEROSIS RELAPSE



Clinical outcome:

- **Experience of acute MS attack**
- **Prediction parameters:**
  Combination of clinical,
  Demographical and
  Blood gene expression
  (most assessable tissue)

EXPERIMENTAL MODEL



Time to next relapse ?
Probability of next relapse during defined period of time?

MS Relapse Biomarkers

Many biomarkers are suggested to be involved in relapse pathogenesis.
Cytokines, chemokines, adhesion molecules, etc.
However results were not reproducible.
Underestimating!?

BMC Medical Genomics
BioMed Central

Research article
Open Access

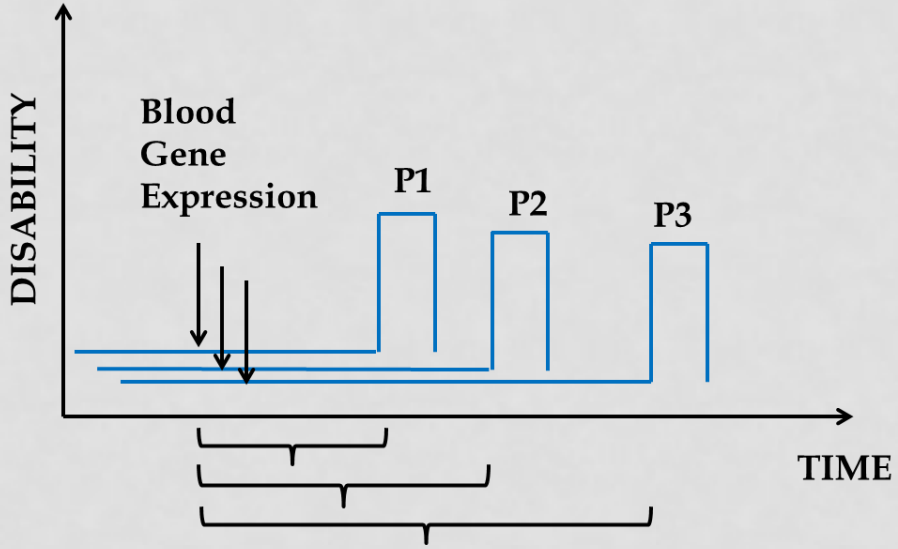
Prediction of acute multiple sclerosis relapses by transcription levels of peripheral blood cells

Michael Gurevich[†1], Tamir Tuller[*†1], Udi Rubinstein[2], Rotem Or-Bach[1] and Anat Achiron[*1]

FACTORS AFFECTING PROBABILITY OF NEXT RELAPSE (NON-MODIFIABLE /MODIFIABLE)

- Demographical:
- Age (decline with age)
- Gender (females are more likely to experience relapses)

- Clinical:
- Disease duration (decline with DD)
- Relapse Rate before
- EDSS rate before
- Previous Treatments
- Current Treatment
- Next Treatment



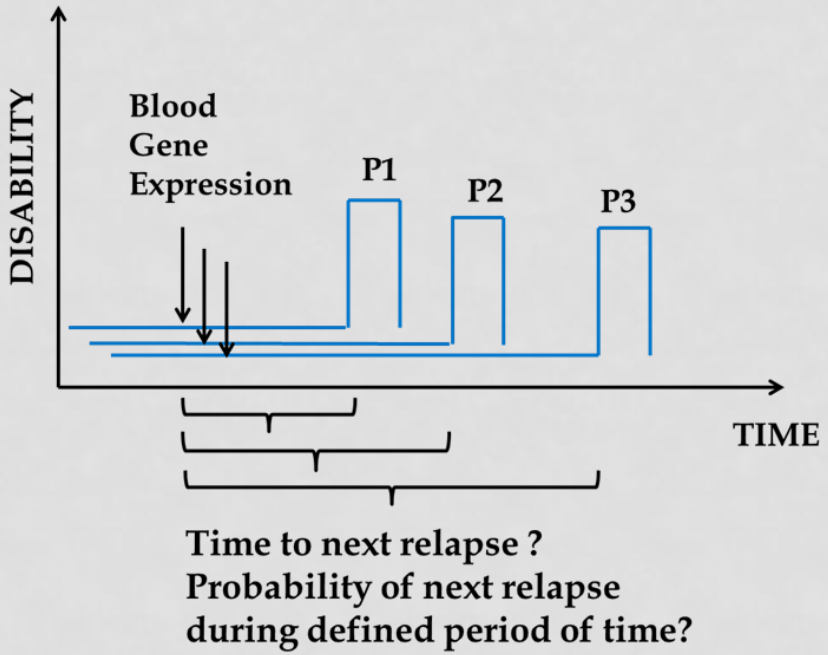
Time to next relapse ?
Probability of next relapse during defined period of time?

Placebo arms from 32 randomized clinical trials: annualized relapse rates ranged from 0.27 to 1.66

FACTORS AFFECTING PROBABILITY OF NEXT RELAPSE

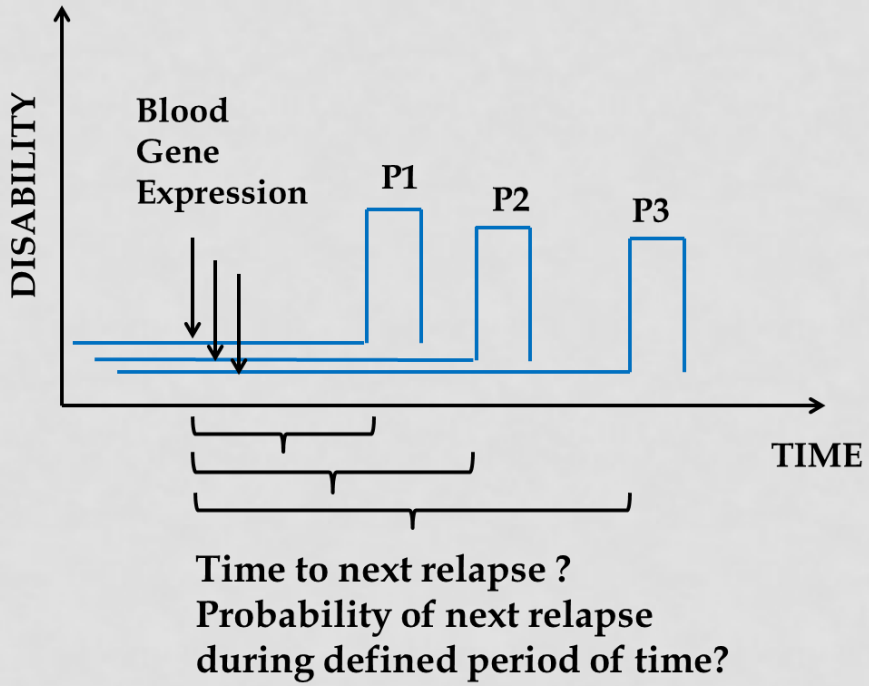
- **Environmental Factors:**
- **Vit D level (increment in vitamin D concentration by 10 nmol/l was associated with reduced hazard of relapses by 9% in adults and 34% in pediatric-onset MS)**
- **Smoking**
- **Heating**
- **Viruses**

Identical twins necessarily generate the same predictions but different real outcomes.



FACTORS THAT COULD AFFECT GENE TRANSCRIPTIONAL LEVELS AT TIME OF ANALYSIS

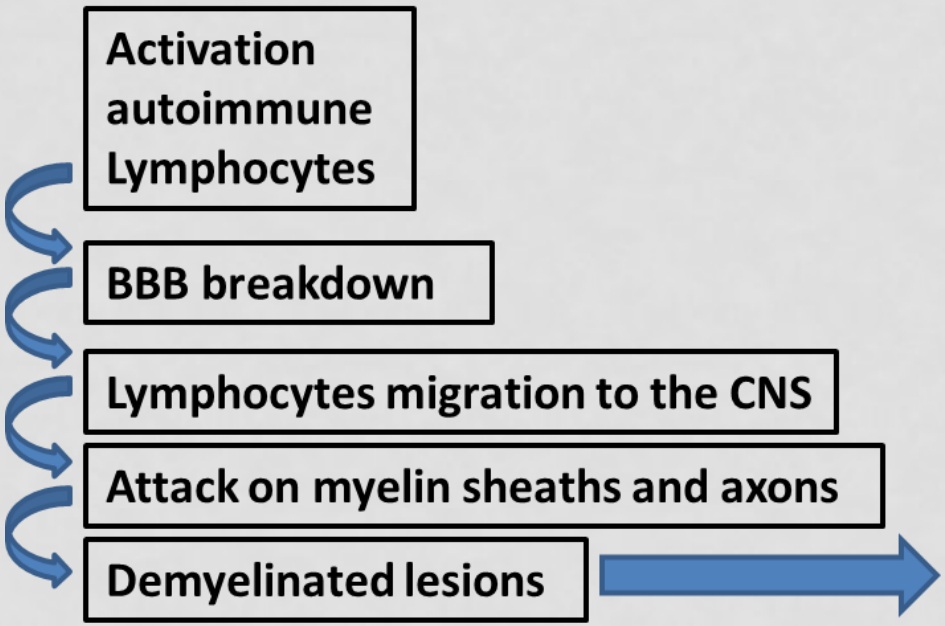
- Disability (EDSS)
- Current treatment,
- Pervious treatment
- Previous acute relapse
- Previous Steroid treatment
- **Next relapse !!!**

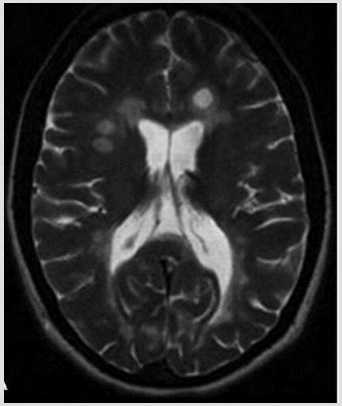


All above factors should be taken to consideration or implemented in statistical model when looking for potential predictors

ACUTE MS RELAPSE AS A CLINICAL OUTCOME

Pathophysiology of acute MS relapse



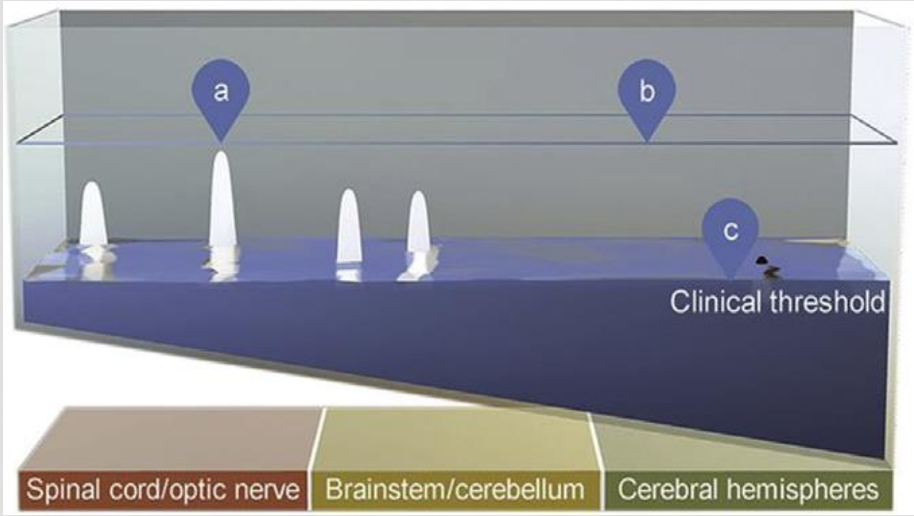


Relapse it is clinical presentation of underlying CNS damage

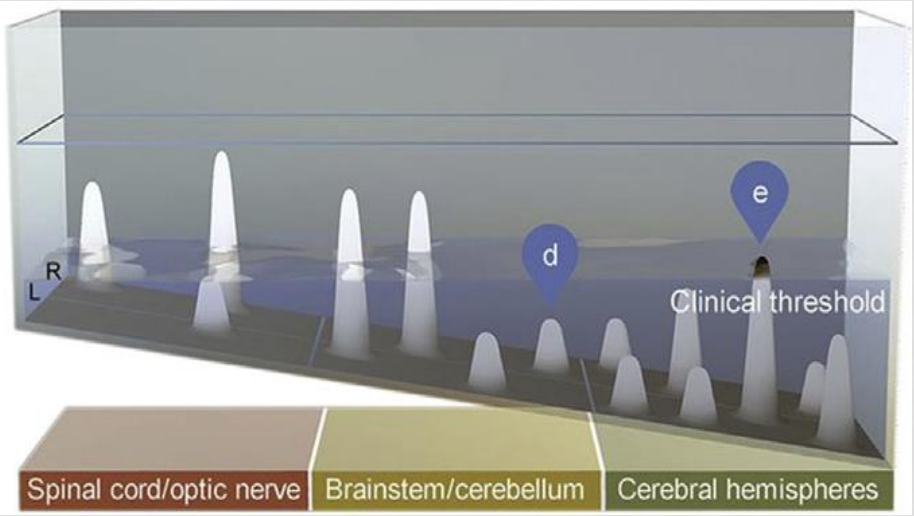
Low correlation between relapse clinical presentation and MRI data

ACUTE MS RELAPSE AS A CLINICAL OUTCOME

Topographic model of clinical and sub-clinical relapses



Clinical view



Sub-clinical view

- New lesions are formed during remission.
- The lesions accumulates.

(About 10 new lesions is required for a clinical symptoms)

- Does the inflammation accumulate?

RADIOLOGICAL PROGRESSION (MRI LESION LOAD) SHOULD BE DEFINED AS CLINICAL OUTCOME

Workflow:

- Find biomarkers of radiological MS progression
- This biomarkers will work on clinical relapse
- Find biomarkers associated with “natural history”
- progression
- Find specific biomarkers for patients on DMD treatment

RESTRICTIONS, TYPES OF ERRORS

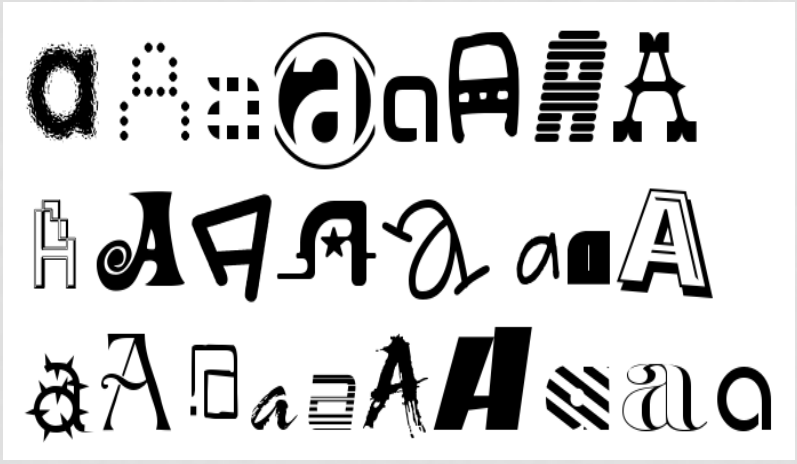
Underestimation

Over fitting

Multiple classifiers

No computing power can squeeze out information that is not present in training set

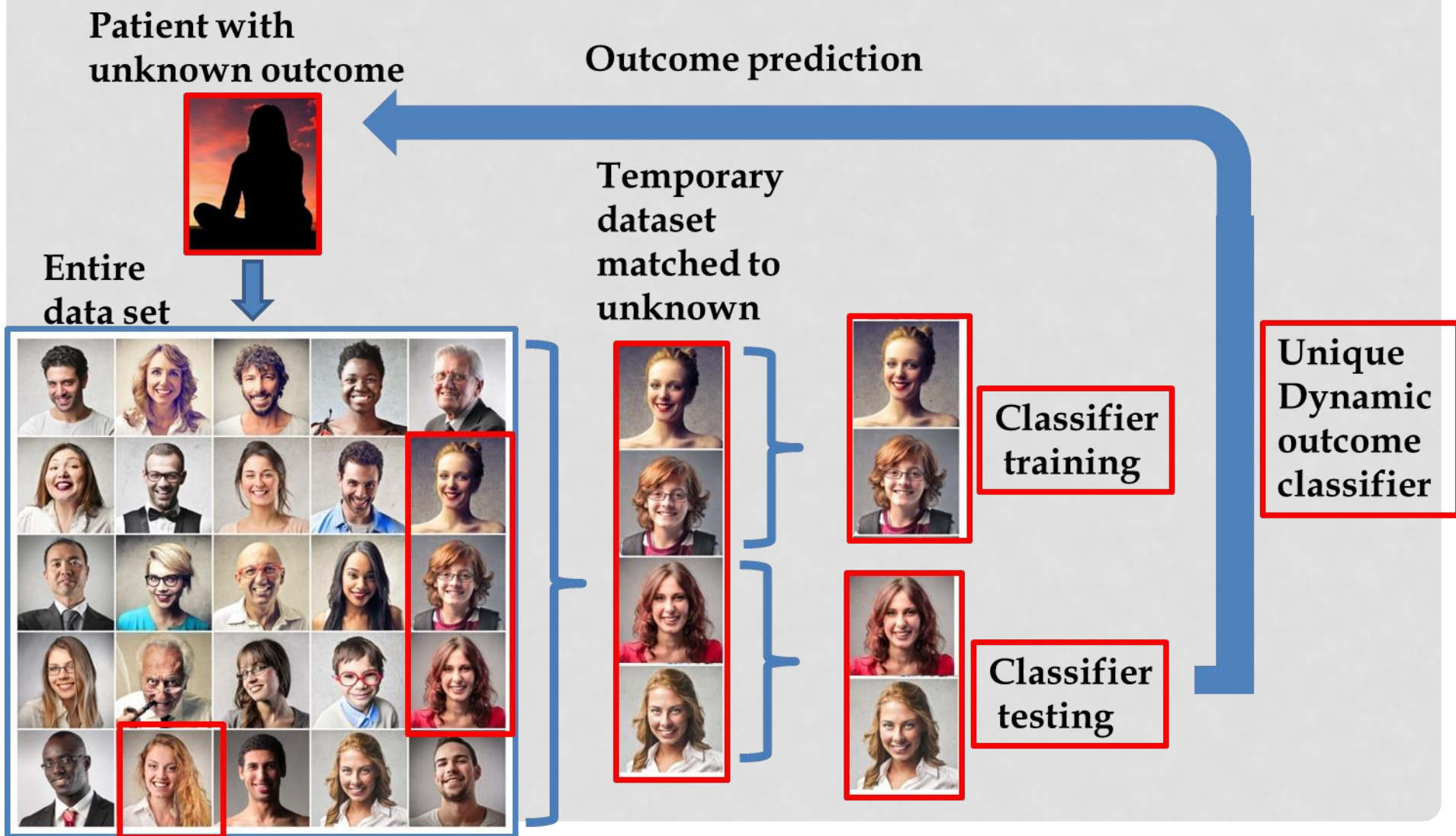
PBMC vs single cell gene expression



THE BAYESIAN WORLD VIEW

Everything has a probability distribution attached to it!

ALTERNATIVE SOLUTION.
DYNAMIC PREDICTORS



DAILY PREDICTIONS

You are in the grocery store.
You are ready to pay.
But which line you choose?
Your brain starts to run a built-in "*predictive algorithm*"

Expected output: the fastest line

Target variable: time spent in the line

Historical data: your experience from previous shopping

Predictors : length of line, number of items in the baskets, average age of customers in the line, etc…

The predictive model: the process your brain goes through while make a decision

SAMPLE OF LEARNING ALGORITHM