



Machine & deep learning for personalized oncology

Loic Verlingue

MD medical oncology

PhDc data science

IFSBM UE Big Data
24/01/19



This is real: Go to doc



Chinese robot becomes world's first machine to pass medical exam

By Ma Si and Cheng Yu | chinadaily.com.cn | Updated: 2017-11-10 15:32

f t in +



En médecine, les impacts réels de l'intelligence artificielle

- Aller plus vite vs **aller bien**
- Renforcer, et non remplacer, le **lien patients-médecins**
- Des applications qui assistent les **personnels de santé**
- Accélérer et démocratiser la **recherche**
- De nouvelles pistes de réflexions **éthiques et juridiques**

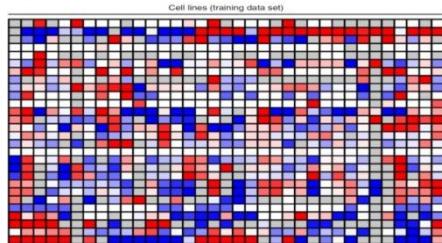
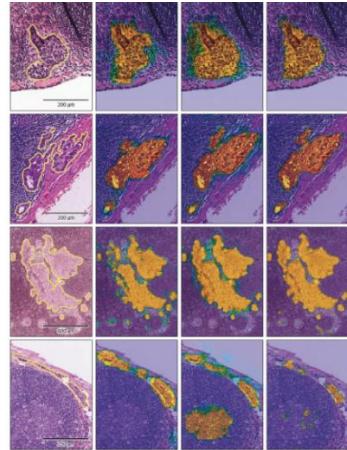
Plan

ML / DL is impacting

- Diagnosis procedures
- Drug development
- Patients' monitoring

Because of many cool things in deep learning

ML/DL in oncology



"Here's my sequence"
The New Yorker

Diagnostics

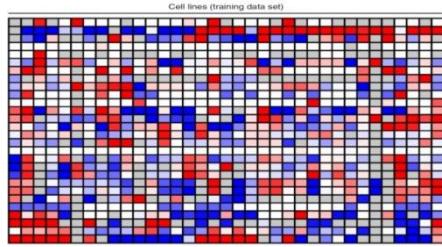
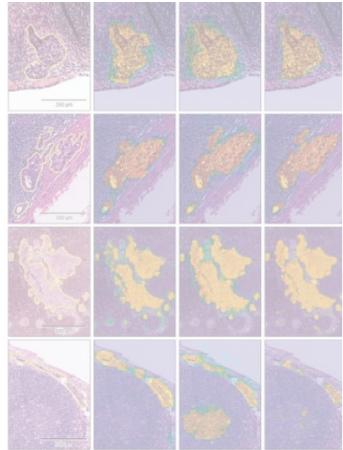
Molecular biology
& prognosis

Drug development
& prediction

Monitoring patients



ML/DL in oncology



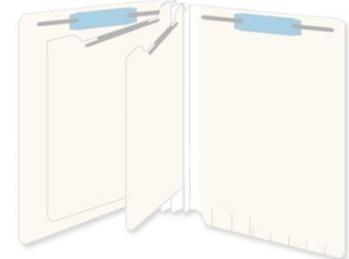
Diagnostics

Molecular biology
& pronostic



"Here's my sequence"
The New Yorker

Drug development
& prediction



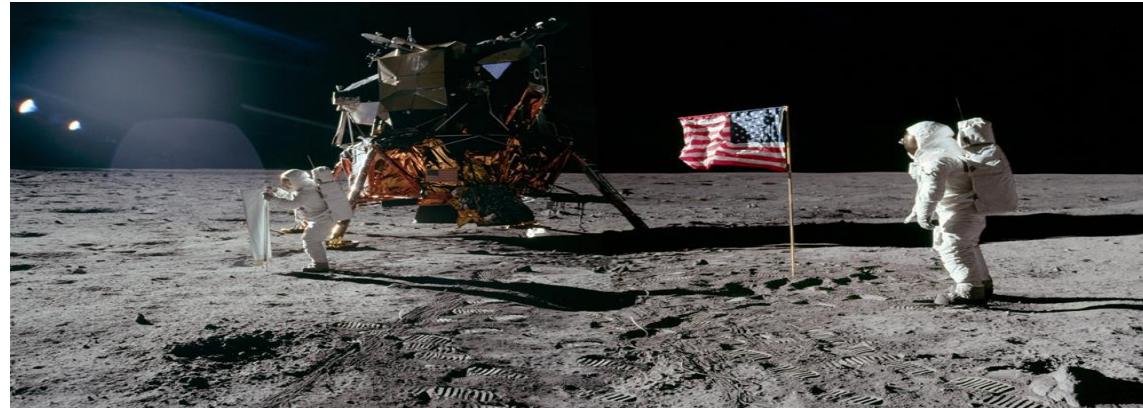
Monitoring patients

Molecular biology & genetics

Table 2. Comparison of prices of large government projects circa 1990 with their projected useful life-span.

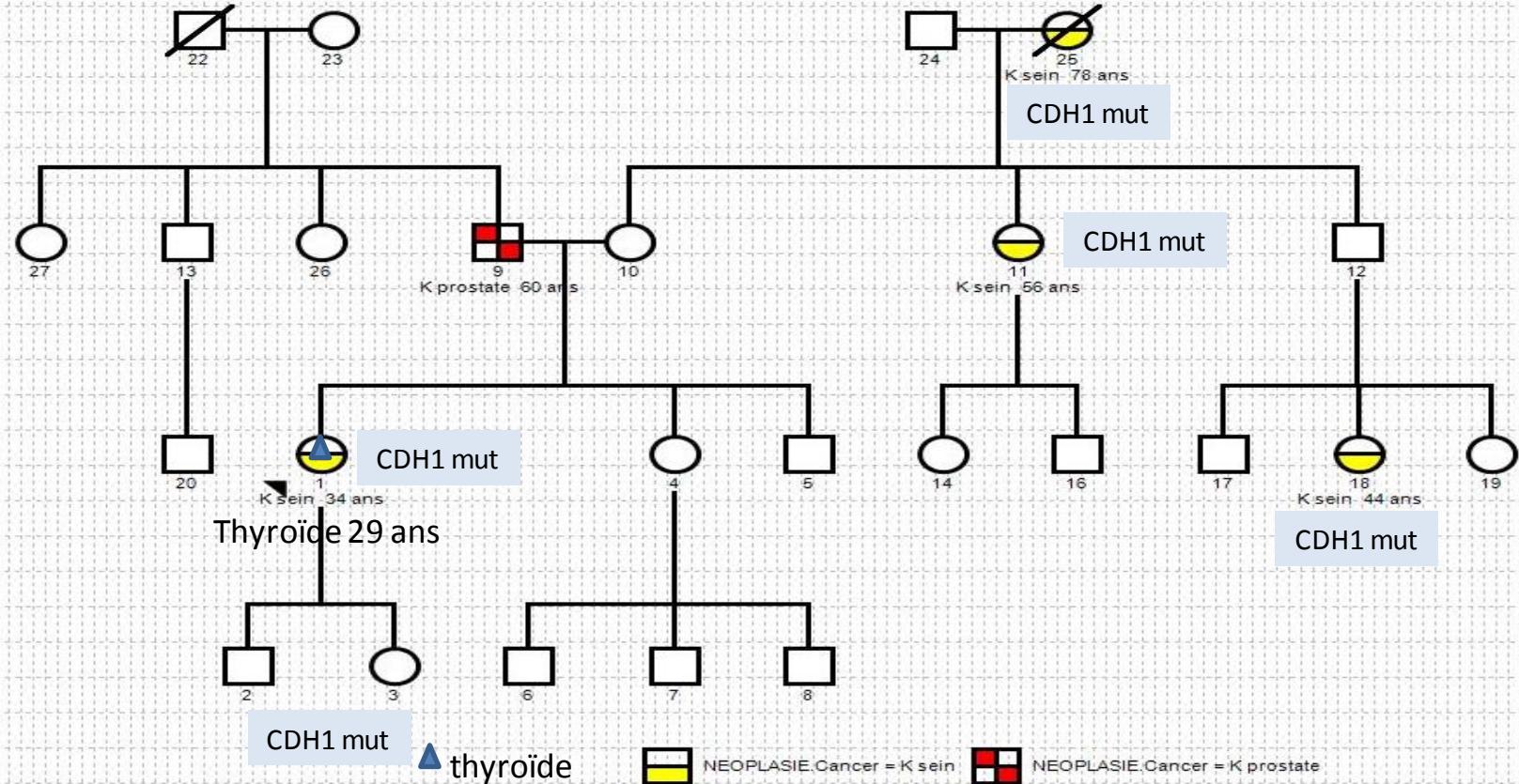
| Proposed project | Projected cost (\$ billion) | Target completion date | Estimated life-span (years) |
|--------------------------------|--|-----------------------------------|--|
| Space Station Freedom | 30.0 | 1999 | 30 |
| Earth Observing System | 17.0 | 2000 | 15 |
| Superconducting Super Collider | 11.0 | 1999 | 30 |
| Human Genome Project | 3.0 | 2005 | Perpetual |
| Hubble Space Telescope | 1.5 | 1990 | 15 to 20 |

Collins et al, Science 2003



Apollo 11, Sea of Tranquility, 20 July 1969, frames A11-40-5872 and A11-40-5874 : Buzz Aldrin is shown deploying the Solar Wind Collector and saluting the U.S. Flag

Difference between
somatic and germline
alteration?

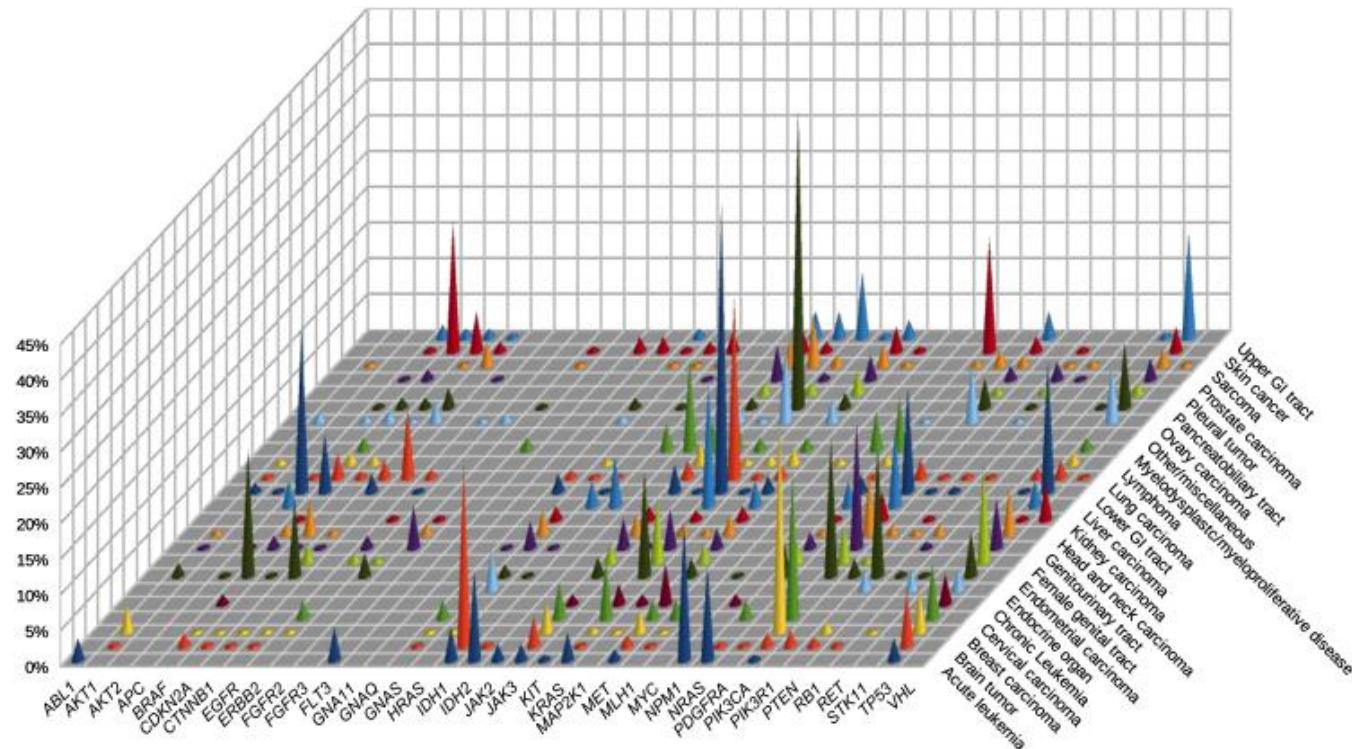


Germline

Somatic alterations

Which are the most frequent **somatic** alterations across tumors?

Heterogeneous somatic alterations

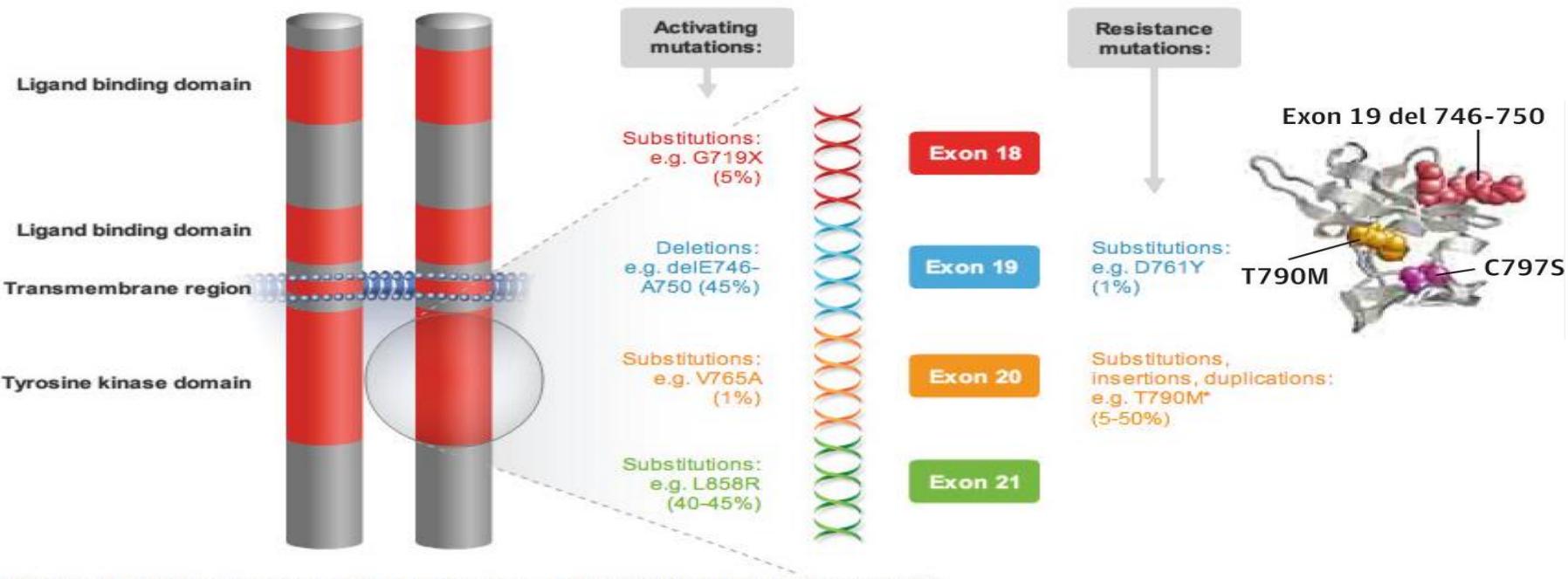




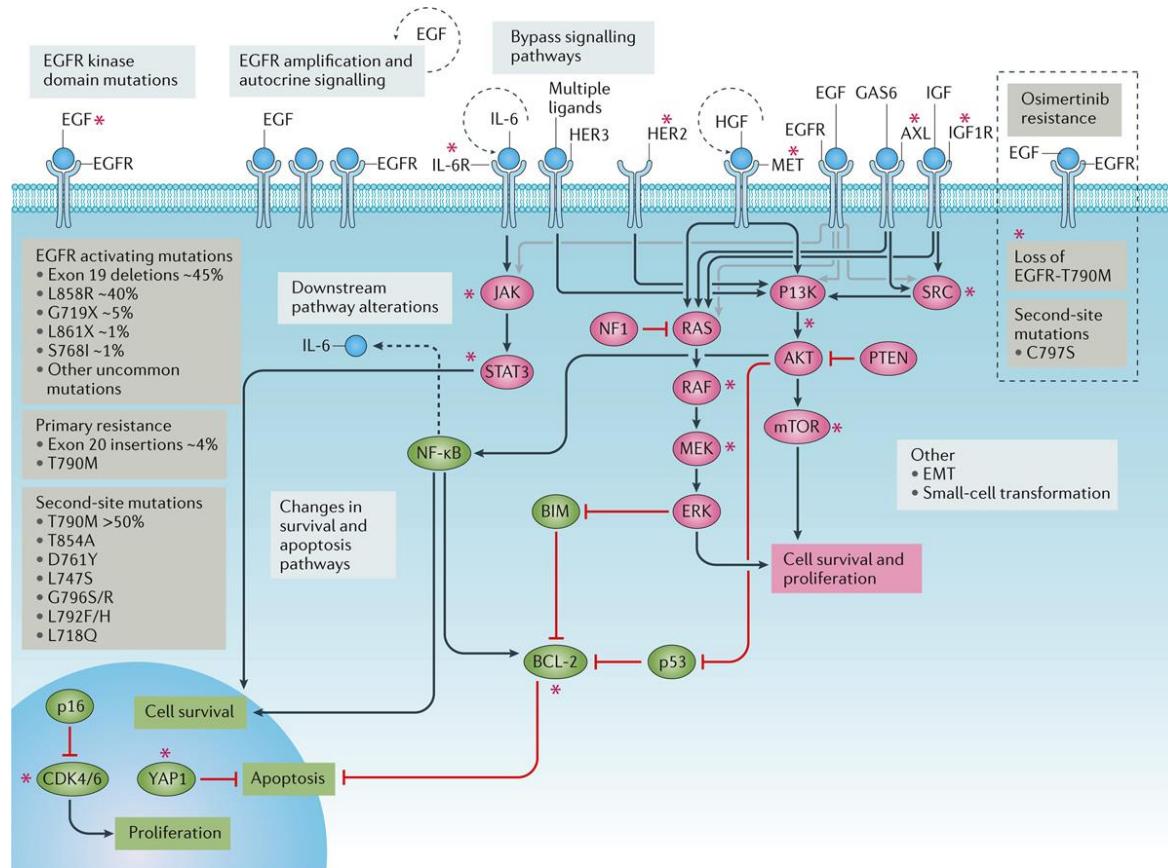
John Godfrey Saxe: "Each was partly in the right, and all were in the wrong"
Opinion by Jacques S. Beckmann and Daniel Lew, tribute to Ronan Flippot

Example of lung adk

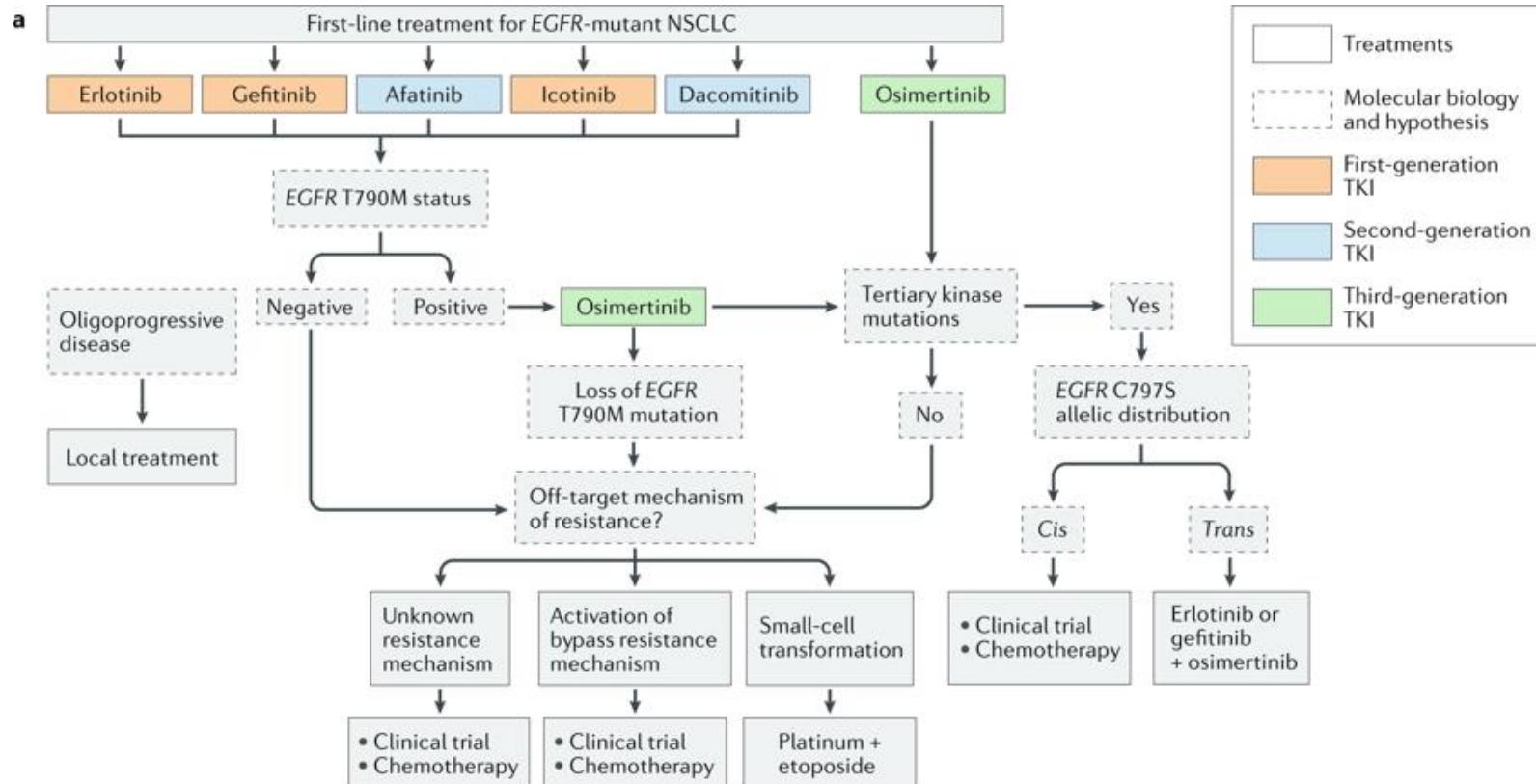
EGFR treatment and resistance mechanisms



EGFR treatment and resistance mechanisms



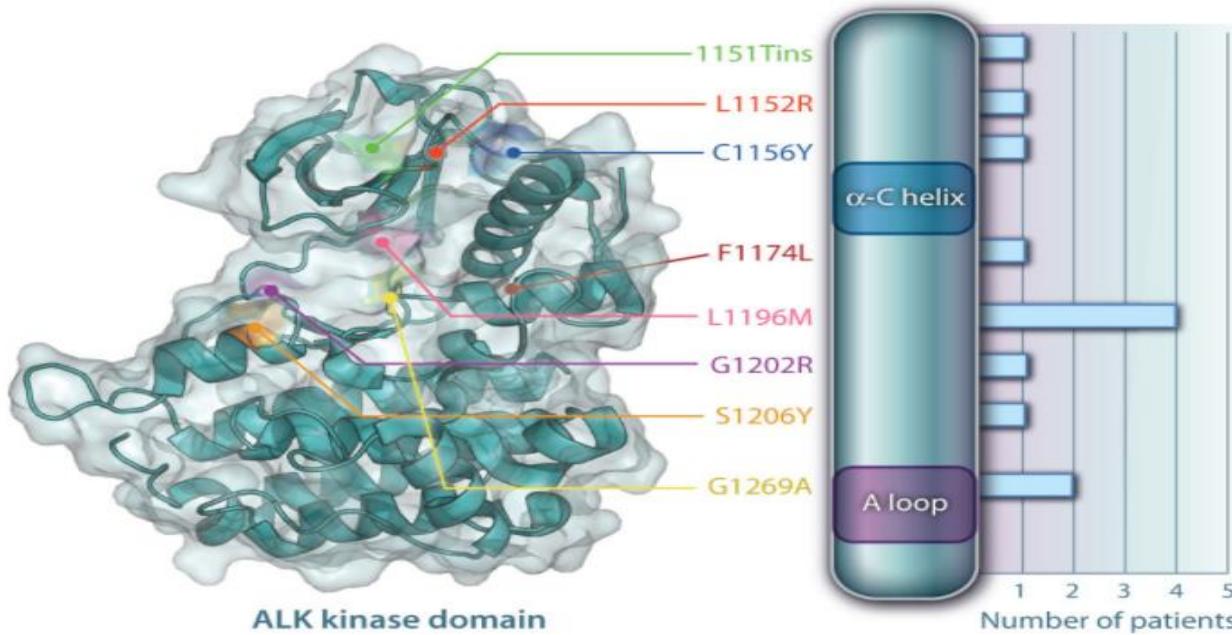
Impact on treatments' strategies



Recondo G, Facchinetto F, Olaussen KA, Besse B, Friboulet L. Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI? Nat Rev Clin Oncol. 2018 Nov;

Example of lung adk

ALK mutations associated to fusion



Targeting mutations

| Mutation status | Crizotinib | Ceritinib | Alectinib | Brigatinib | Lorlatinib |
|-------------------------------|------------|-----------|-----------|------------|------------|
| Parental Ba/F3 | | | | | |
| <i>EML4-ALK</i> V1 | | | | | |
| <i>EML4-ALK</i> C1156Y | | | | | |
| <i>EML4-ALK</i> I1171N | | | | | |
| <i>EML4-ALK</i> I1171S | | | | | |
| <i>EML4-ALK</i> I1171T | | | | | |
| <i>EML4-ALK</i> F1174C | | | | | |
| <i>EML4-ALK</i> L1196M | | | | | |
| <i>EML4-ALK</i> L1198F | | | | | |
| <i>EML4-ALK</i> G1202R | | | | | |
| <i>EML4-ALK</i> G1202del | | | | | |
| <i>EML4-ALK</i> D1203N | | | | | |
| <i>EML4-ALK</i> E1210K | | | | | |
| <i>EML4-ALK</i> G1269A | | | | | |
| <i>EML4-ALK</i> D1203N+F1174C | | | | | |
| <i>EML4-ALK</i> D1203N+E1210K | | | | | |

Eg for ALKi

Targeting mutations

Cellular ALK phosphorylation mean IC₅₀ (nmol/L)

| Mutation status | Crizotinib | Ceritinib | Alectinib | Brigatinib | Lorlatinib |
|-------------------------------|------------|-------------------|-------------------|------------|------------|
| Parental Ba/F3 | 763.9 | 885.7 | 890.1 | 2774.0 | 11293.8 |
| <i>EML4-ALK</i> V1 | 38.6 | 4.9 | 11.4 | 10.7 | 2.3 |
| <i>EML4-ALK</i> C1156Y | 61.9 | 5.3 | 11.6 | 4.5 | 4.6 |
| <i>EML4-ALK</i> I1171N | 130.1 | 8.2 | 397.7 | 26.1 | 49.0 |
| <i>EML4-ALK</i> I1171S | 94.1 | 3.8 | 177.0 | 17.8 | 30.4 |
| <i>EML4-ALK</i> I1171T | 51.4 | 1.7 | 33.6 ^a | 6.1 | 11.5 |
| <i>EML4-ALK</i> F1174C | 115.0 | 38.0 ^a | 27.0 | 18.0 | 8.0 |
| <i>EML4-ALK</i> L1196M | 339.0 | 9.3 | 117.6 | 26.5 | 34.0 |
| <i>EML4-ALK</i> L1198F | 0.4 | 196.2 | 42.3 | 13.9 | 14.8 |
| <i>EML4-ALK</i> G1202R | 381.6 | 124.4 | 706.6 | 129.5 | 49.9 |
| <i>EML4-ALK</i> G1202del | 58.4 | 50.1 | 58.8 | 95.8 | 5.2 |
| <i>EML4-ALK</i> D1203N | 116.3 | 35.3 | 27.9 | 34.6 | 11.1 |
| <i>EML4-ALK</i> E1210K | 42.8 | 5.8 | 31.6 | 24.0 | 1.7 |
| <i>EML4-ALK</i> G1269A | 117.0 | 0.4 | 25.0 | ND | 10.0 |
| <i>EML4-ALK</i> D1203N+F1174C | 338.8 | 237.8 | 75.1 | 123.4 | 69.8 |
| <i>EML4-ALK</i> D1203N+E1210K | 153.0 | 97.8 | 82.8 | 136.0 | 26.6 |

IC₅₀ ≤ 50 nmol/L

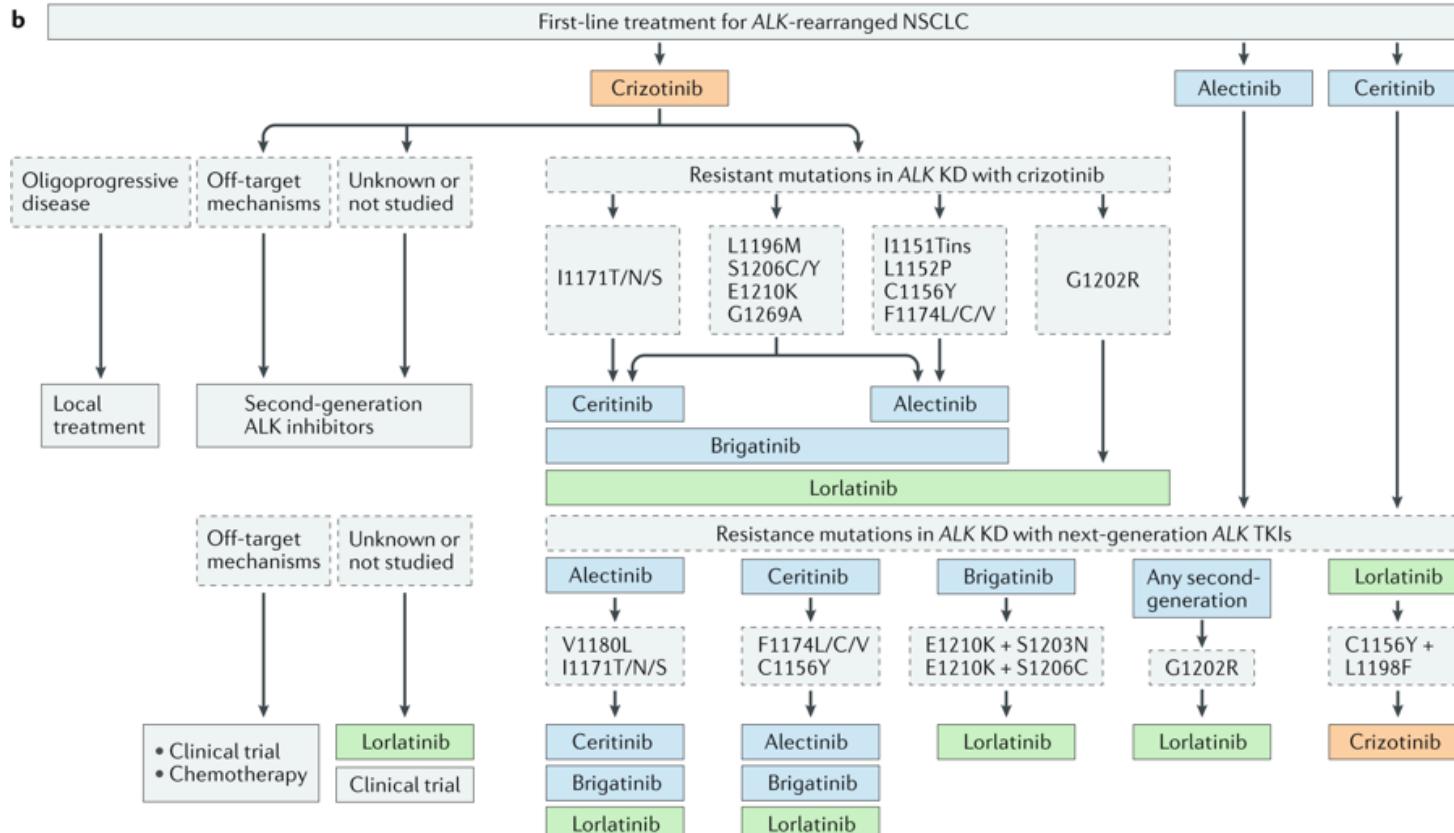
IC₅₀ > 50 < 200 nmol/L

IC₅₀ ≥ 200 nmol/L

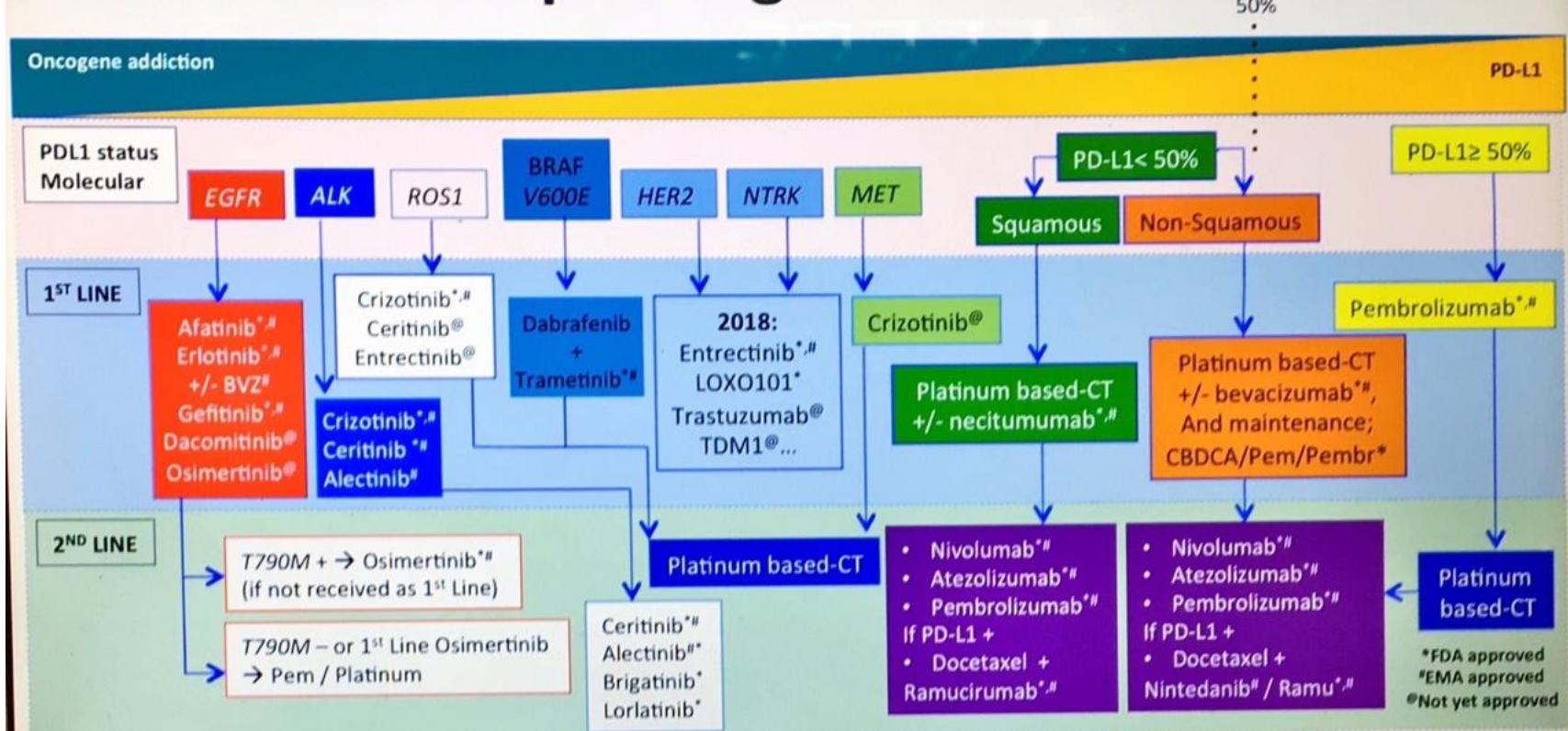
Eg for ALKi

Impact on treatments' strategies

b

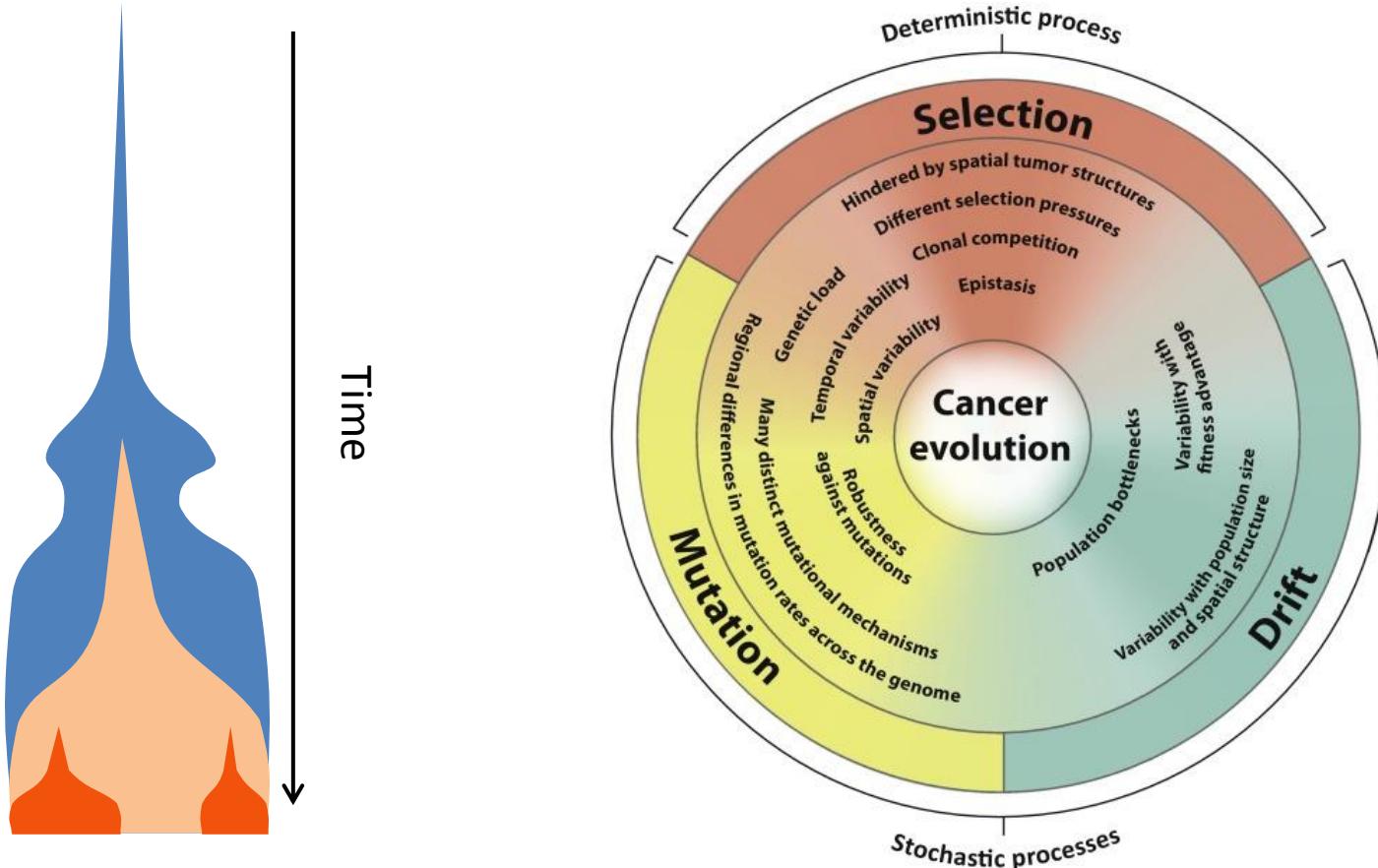


New treatment paradigm in NSCLC



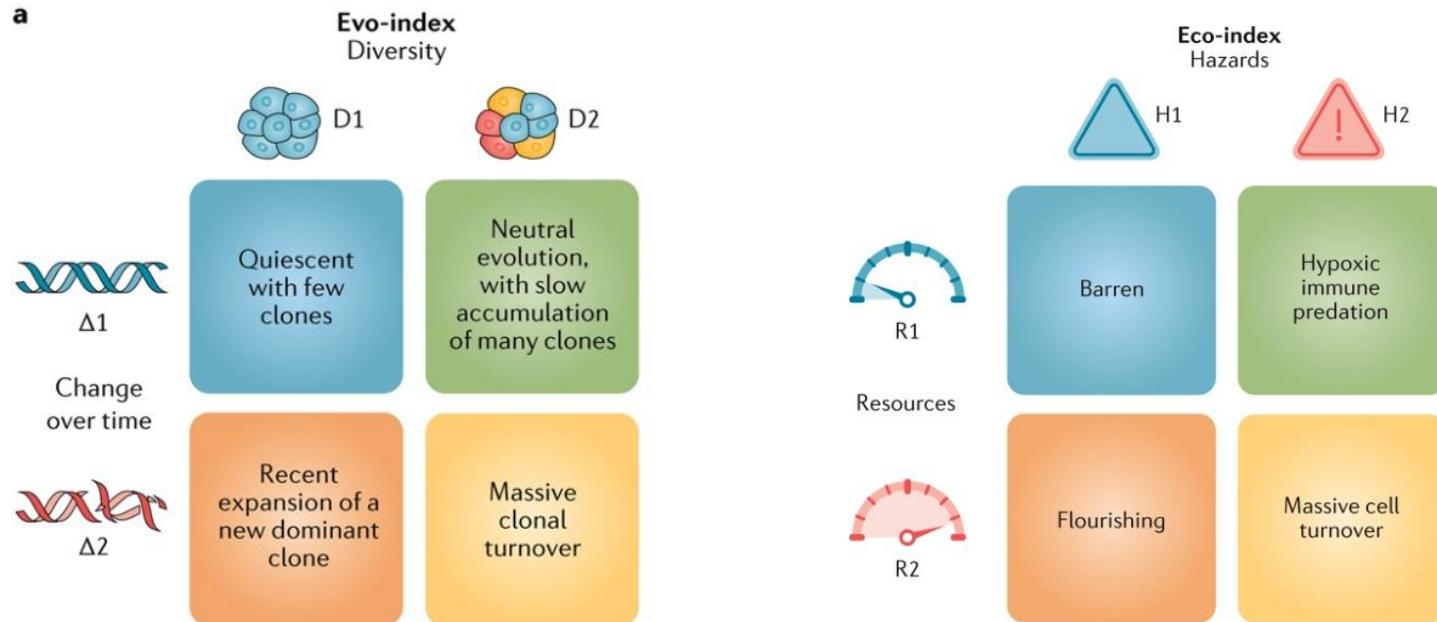
Courtesy Dr Jordi Remon

Cancer Evolution



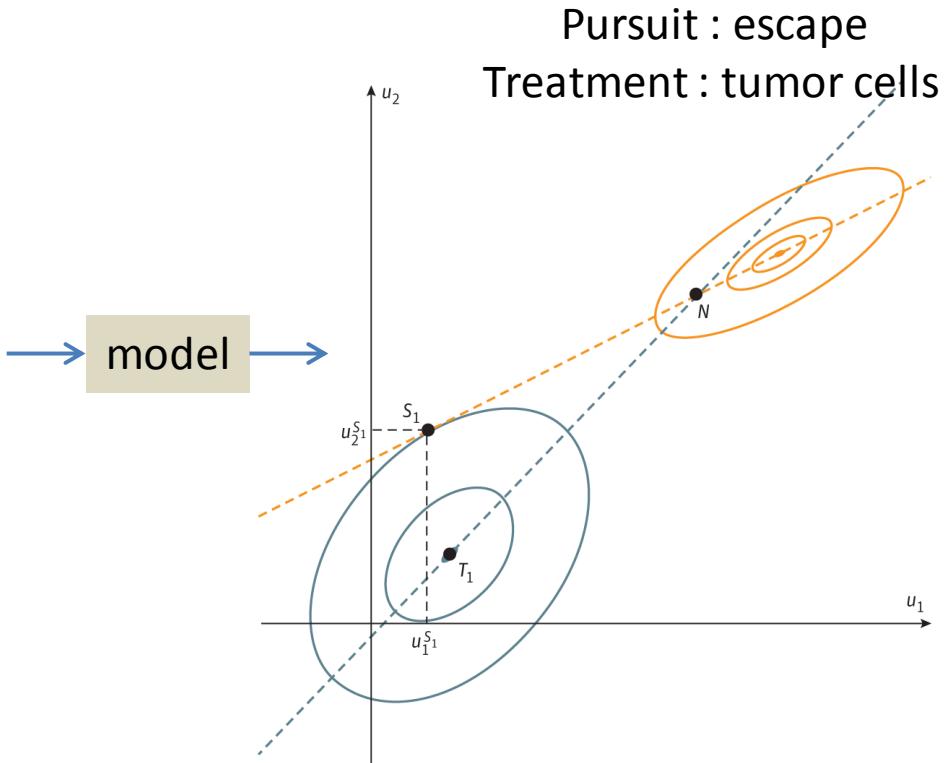
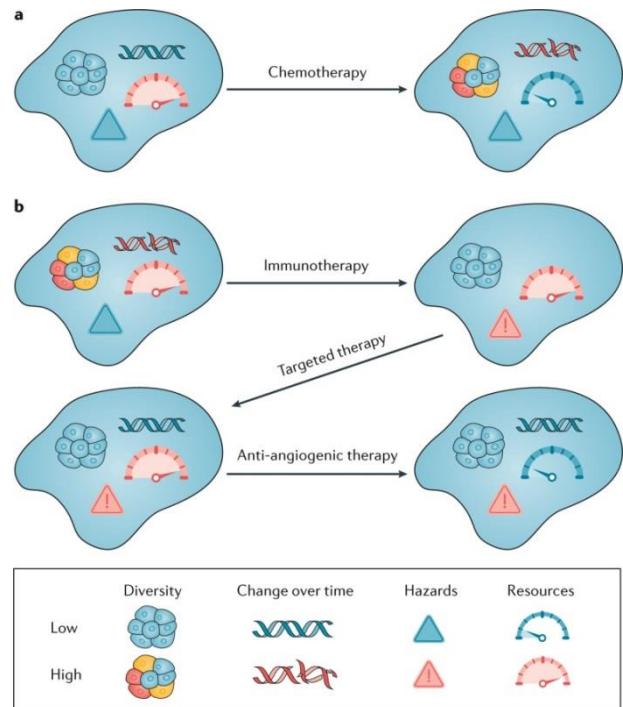
Trends in Cancer

Evolutionary and ecological features of neoplasms



Nature Reviews | Cancer

Cancer evolution -> game theory



Maley CC, et al. Classifying the evolutionary and ecological features of neoplasms. Nat Rev Cancer. 2017 Oct

Stanková K et al. Optimizing Cancer Treatment Using Game Theory: A Review. JAMA Oncol. 2018

The complex molecular landscape

&

The dynamic evolution of tumor biology

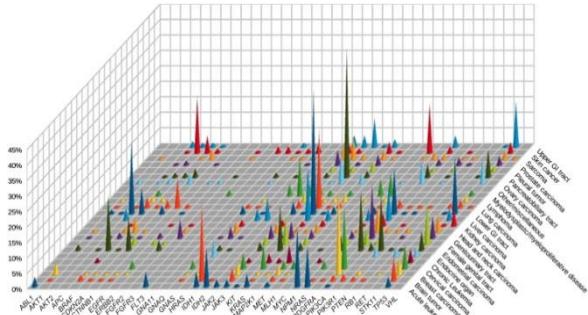
Are new challenges of precision oncology

at least in lung adenocarcinomas...

How to manage it in the clinic?

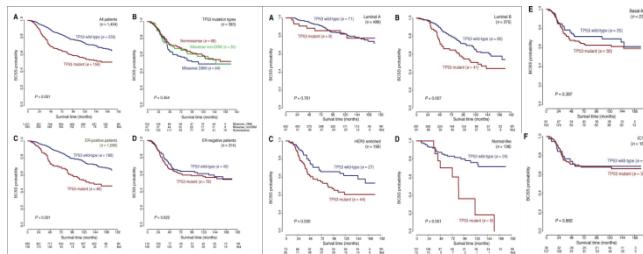
Multi-tumor omics... Problems...

Heterogeneous somatic alterations



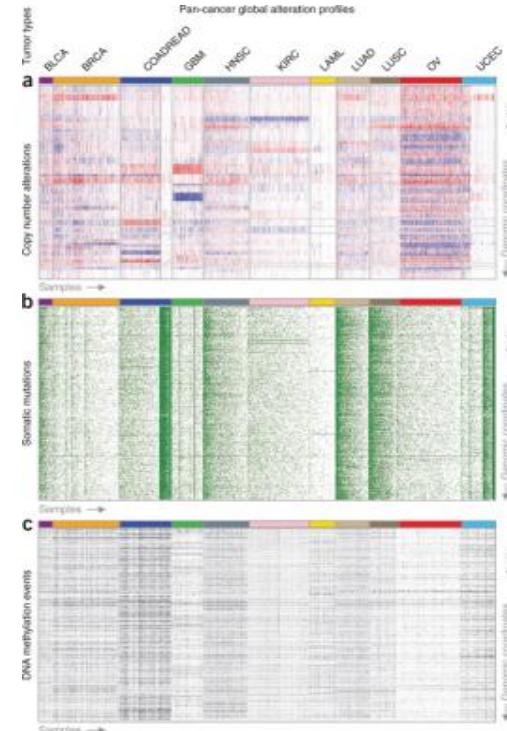
MacConaill et al. Prospective enterprise-level molecular genotyping of a cohort of cancer patients. *J Mol Diagn.* 2014

Effect of mutations are context dependant



Silwal-Pandit L, et al A. TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance. *Clin Cancer Res.* 2014

Tumors are dependant of different layers of alterations



Somatic mutations

Methylation

Ciriello G, Miller ML, Aksoy BA, Senbabaooglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. *Nat Genet.* 2013 Oct

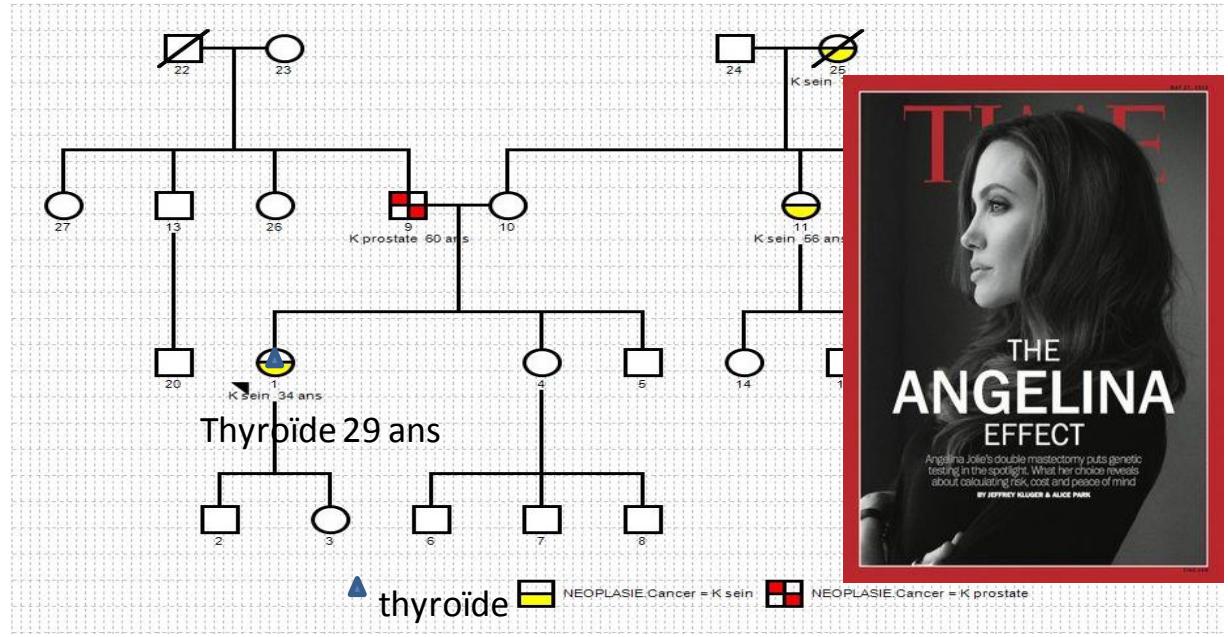


+ clinical variables

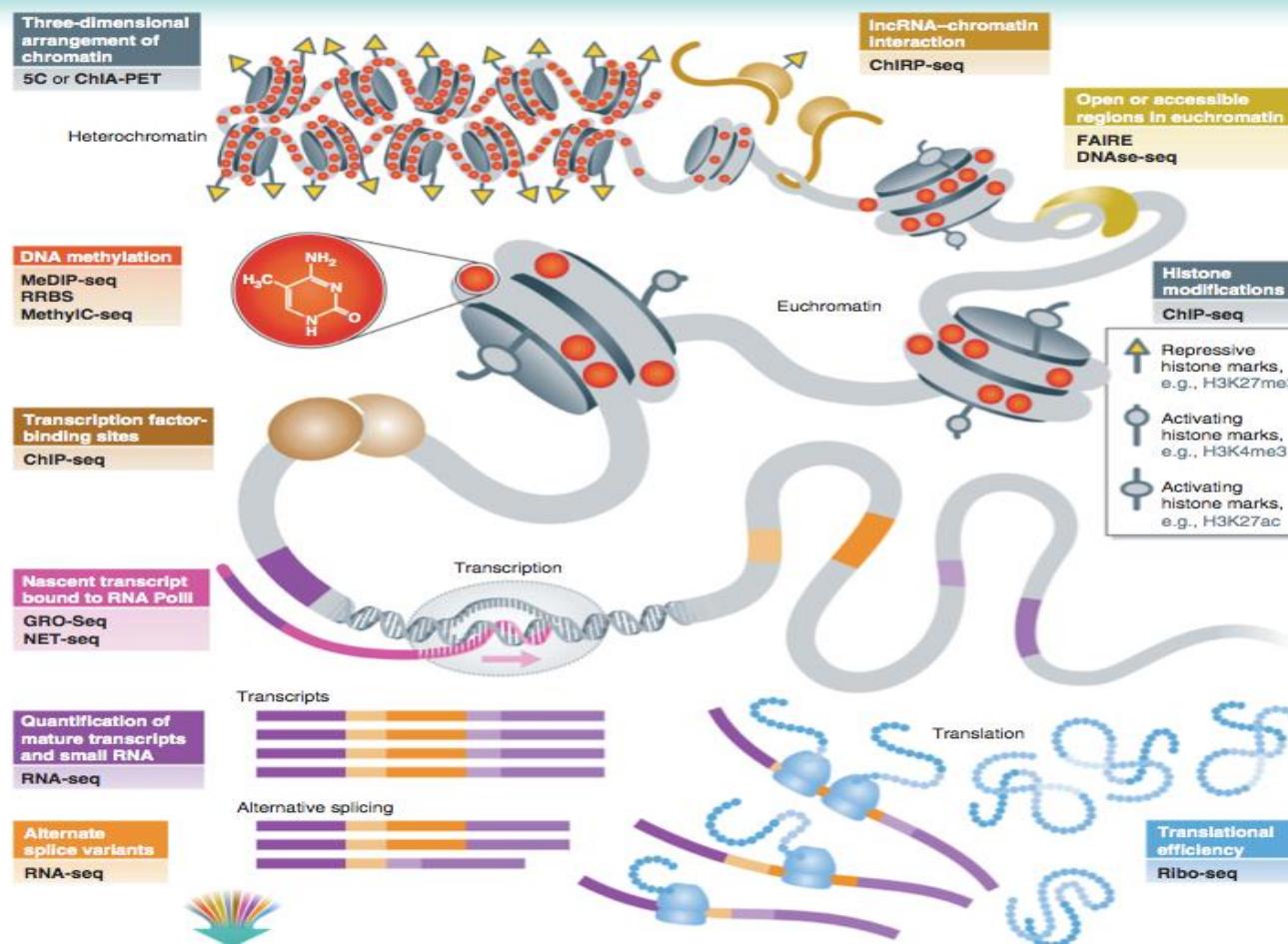
Localisation matters



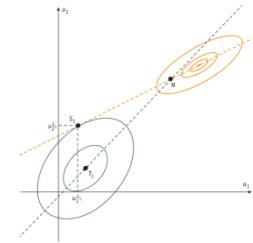
Host matters



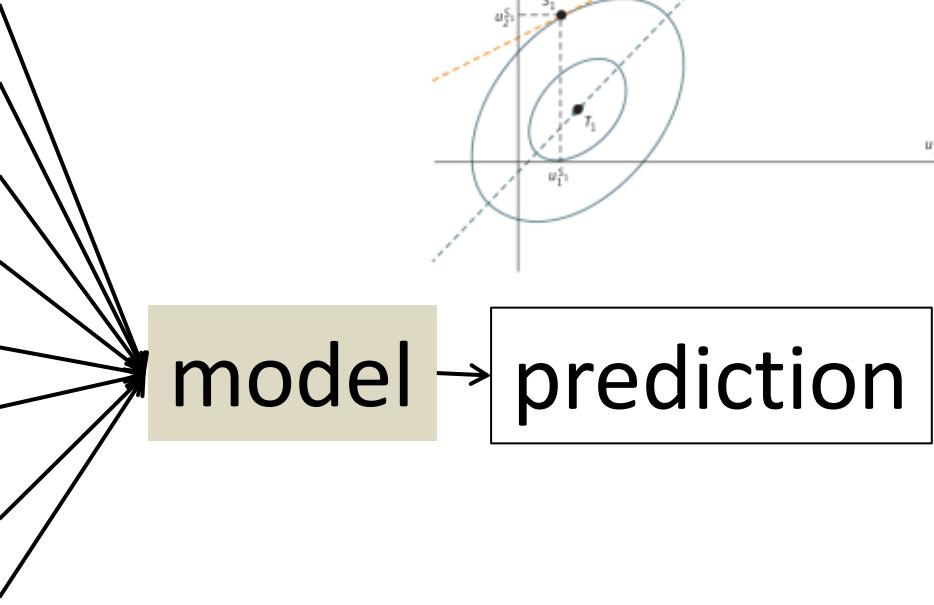
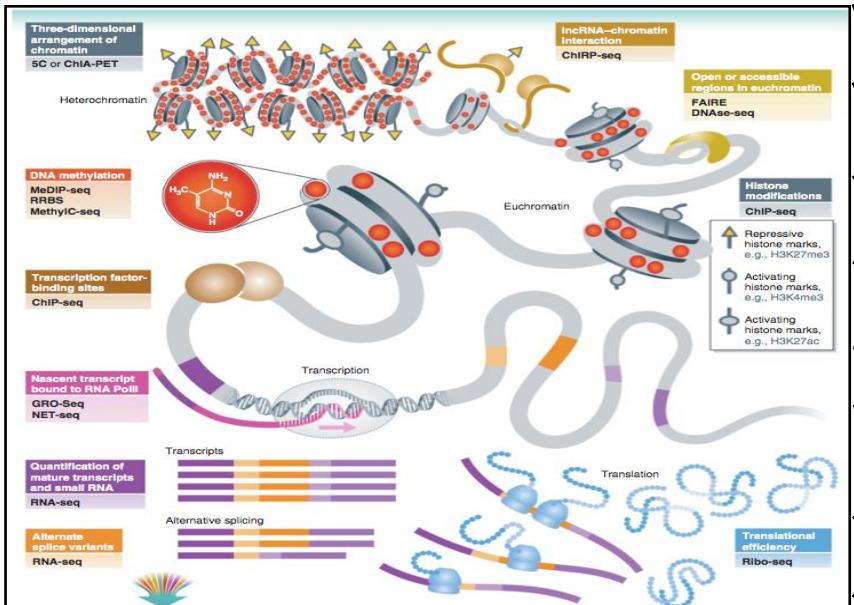
Complexity



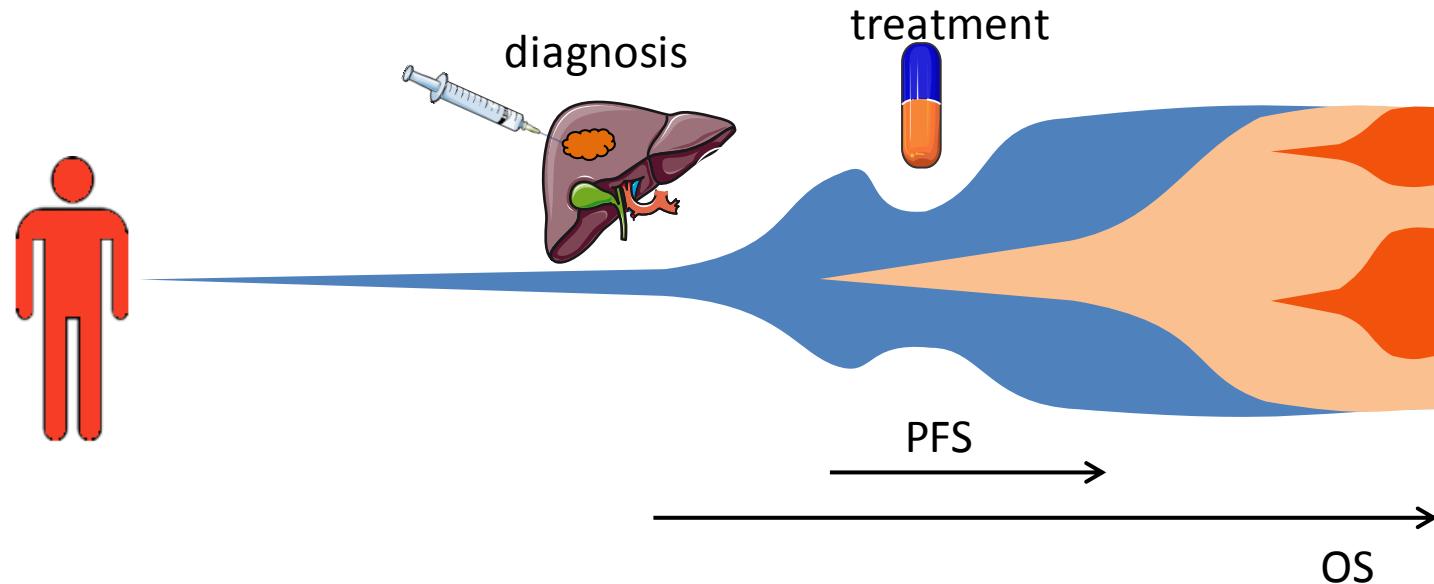
model → prediction



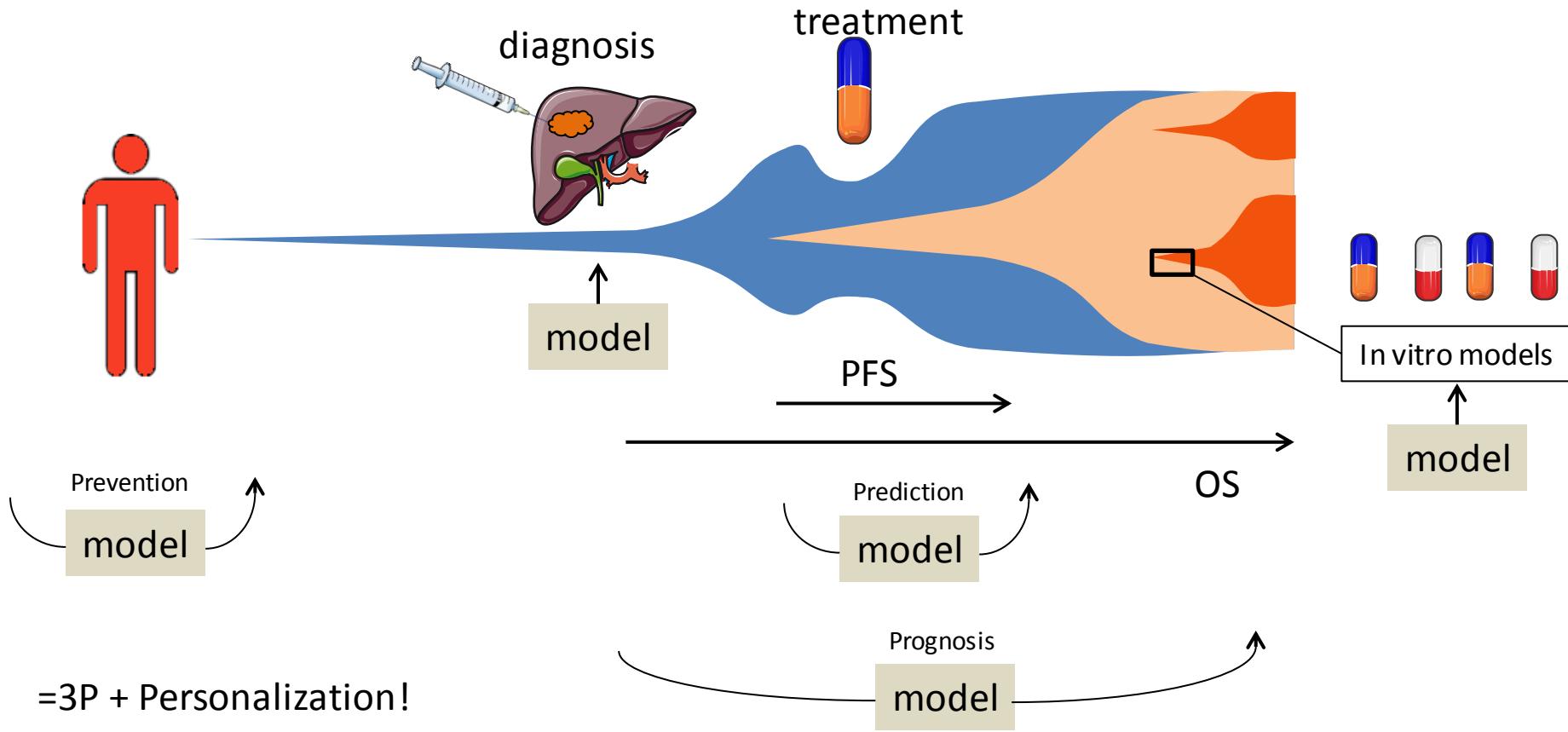
Models



The patient & his tumor's biology

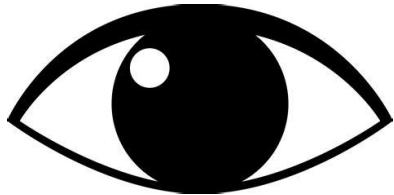


What do we expect from models?



The models you have

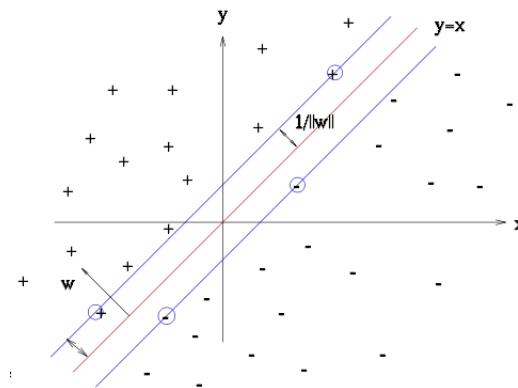
Human



Use your knowledge
and reasoning

Examples: Christophe Massard

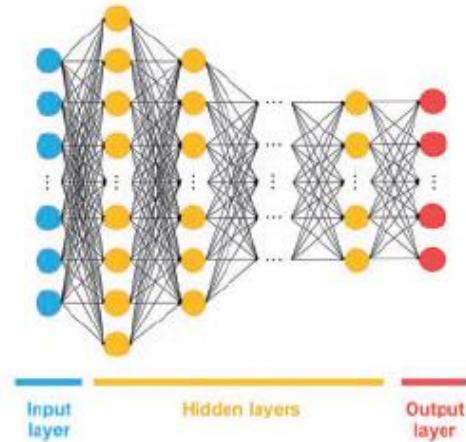
Statistics



Linear models easy to
explain

Cox model
logReg eg: RMH

Deep learning



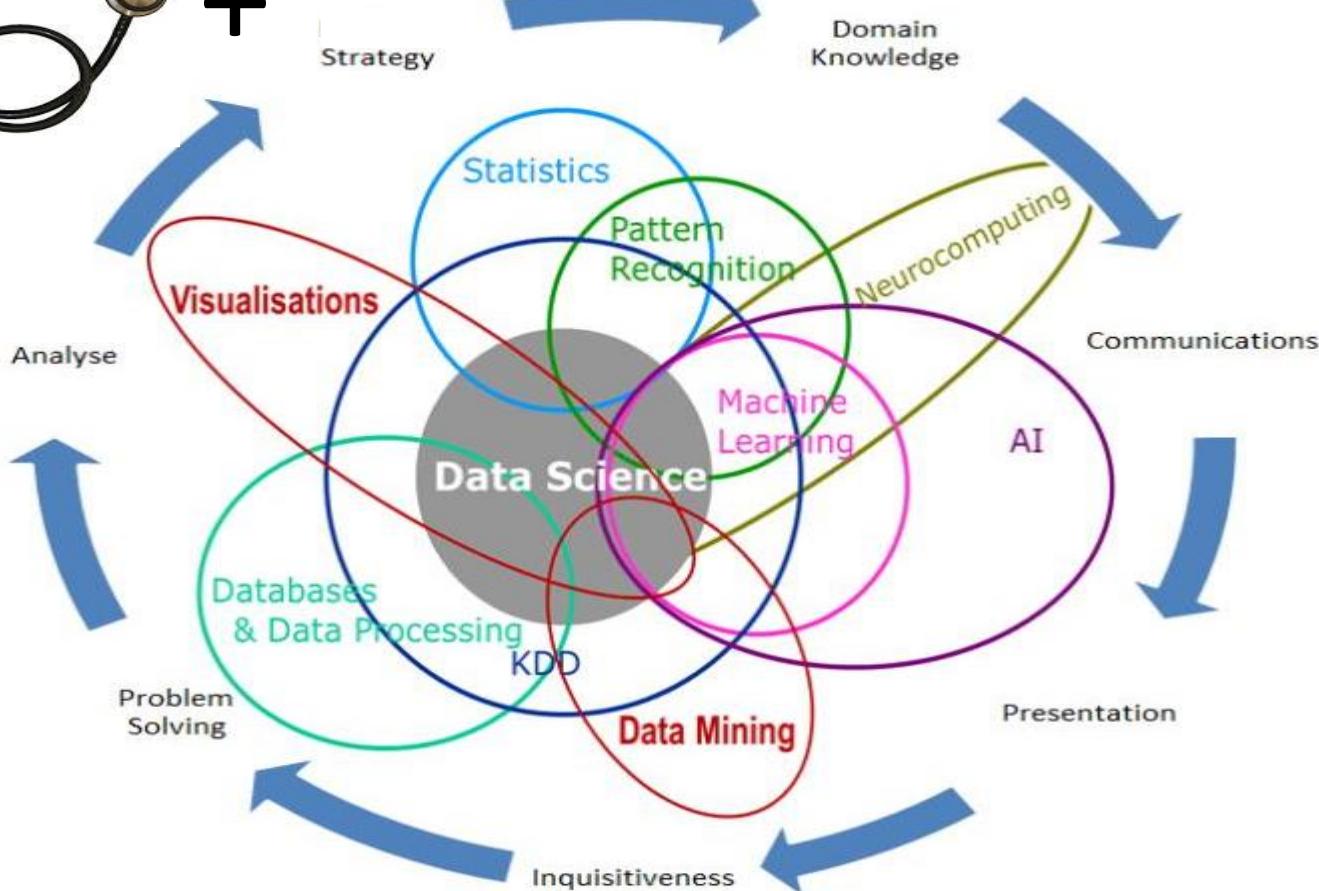
Non-linear models
data-expensive

Deep neural network

Cancer/patient selection

The tools you need

By Brendan Tierney, 2012



Where are the data?



A screenshot of the European Medicines Agency (EMA) website. The header includes the EMA logo and the text "EUROPEAN MEDICINES AGENCY SCIENCE MEDICINES HEALTH". The navigation menu at the top includes links for Home, Find medicine, Human regulatory, Veterinary regulatory, Committees, News & events, and Partners. A sidebar on the left shows categories like Overview, Research and development, Marketing authorisation (which is currently selected), and Advanced therapies. The main content area displays a breadcrumb trail: Home > Human regulatory > Marketing authorisation > Clinical data publication. Below this, there is a section titled "Clinical data publication" with a sub-section about publishing clinical data from October 2016.



A screenshot of the "Study sponsors" section on the ClinicalStudy DataRequest.com website. At the top, there is a purple "View" button and a message encouraging users to select a sponsor logo to view studies from that company. Below this is a grid of pharmaceutical company logos, including Astellas, Boehringer Ingelheim, GSK, Eli Lilly, Novartis, Roche, Sanofi, Takeda, UCB, and ViiV Healthcare.

A screenshot of an article titled "SOUNDING BOARD" from The New England Journal of Medicine. The title is in red capital letters. The article is titled "Data Sharing from Clinical Trials — A Research Funder's Perspective" by Robert Kiley, Tony Peatfield, Jennifer Hansen, and Fiona Reddington.

Data Sharing from Clinical Trials — A Research Funder's Perspective

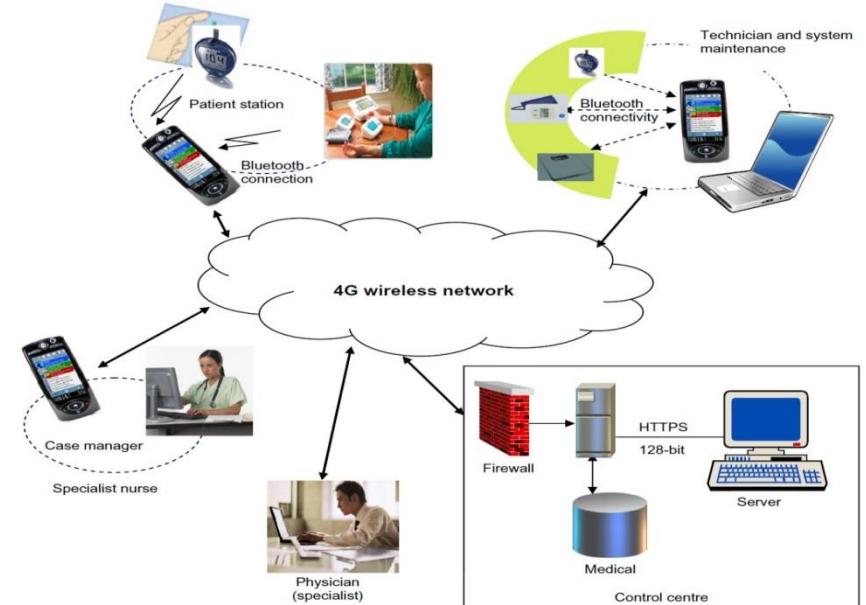
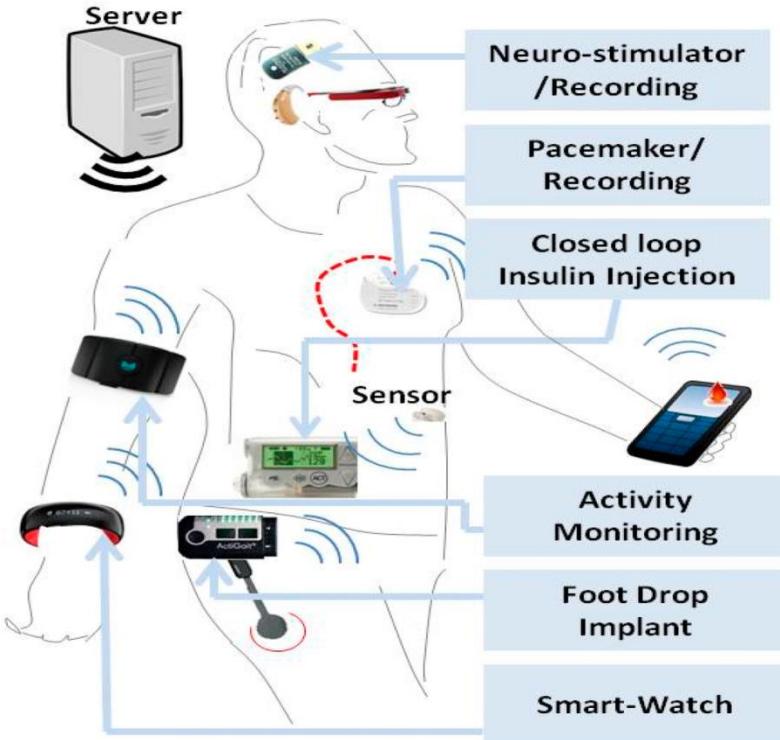
Robert Kiley, Tony Peatfield, Jennifer Hansen, and Fiona Reddington

A screenshot of a "SPECIAL ARTICLE" box from The New England Journal of Medicine. The title of the article is "Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing" by Michelle M. Mello, J.D., Ph.D., Van Lieou, B.S., and Steven N. Goodman, M.D., Ph.D.

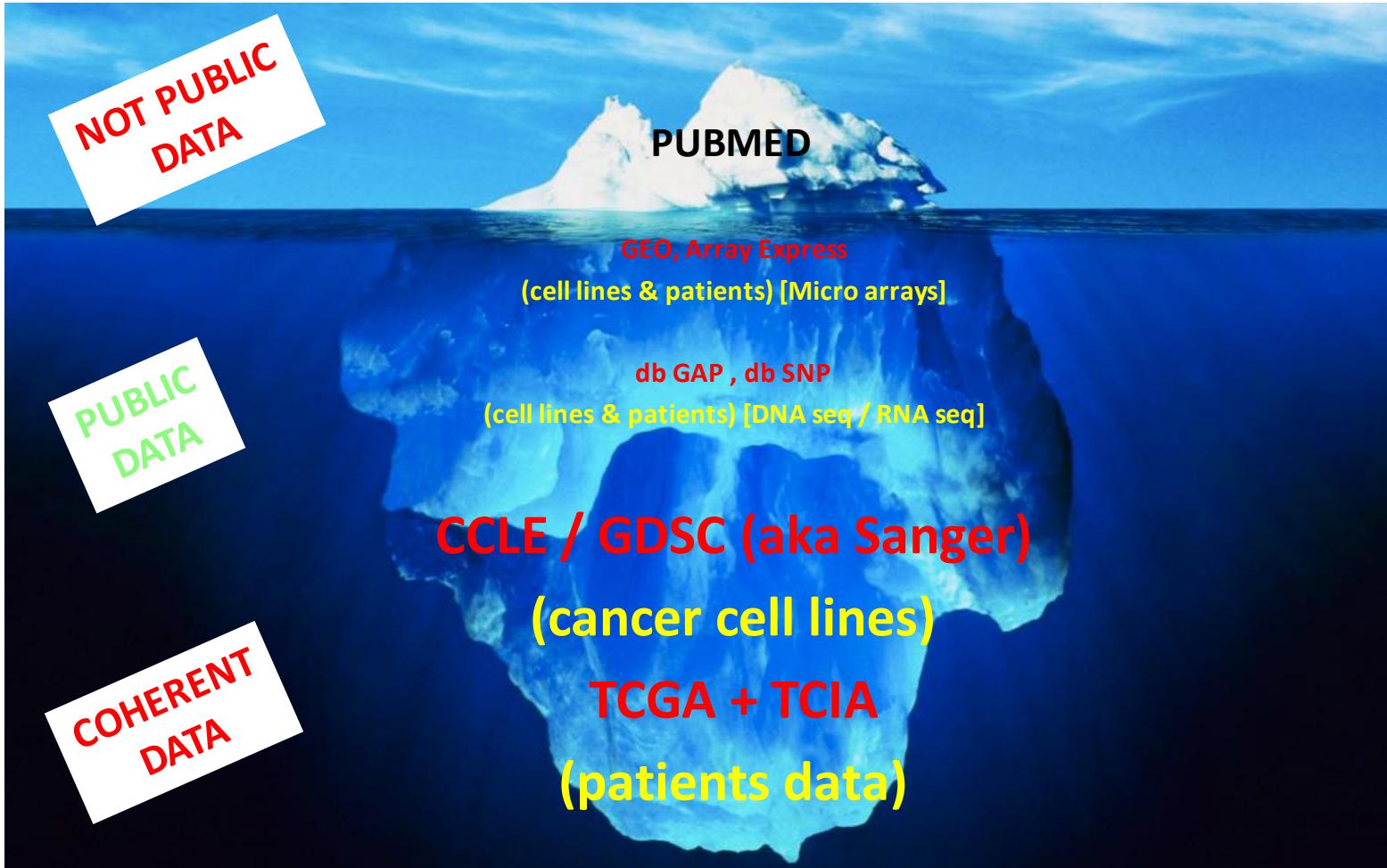
Clinical Trial Participants' Views
of the Risks and Benefits of Data Sharing

Michelle M. Mello, J.D., Ph.D., Van Lieou, B.S.,
and Steven N. Goodman, M.D., Ph.D.

New flow of clinical data



Where are the data?



Search is made easy

Traditional databases



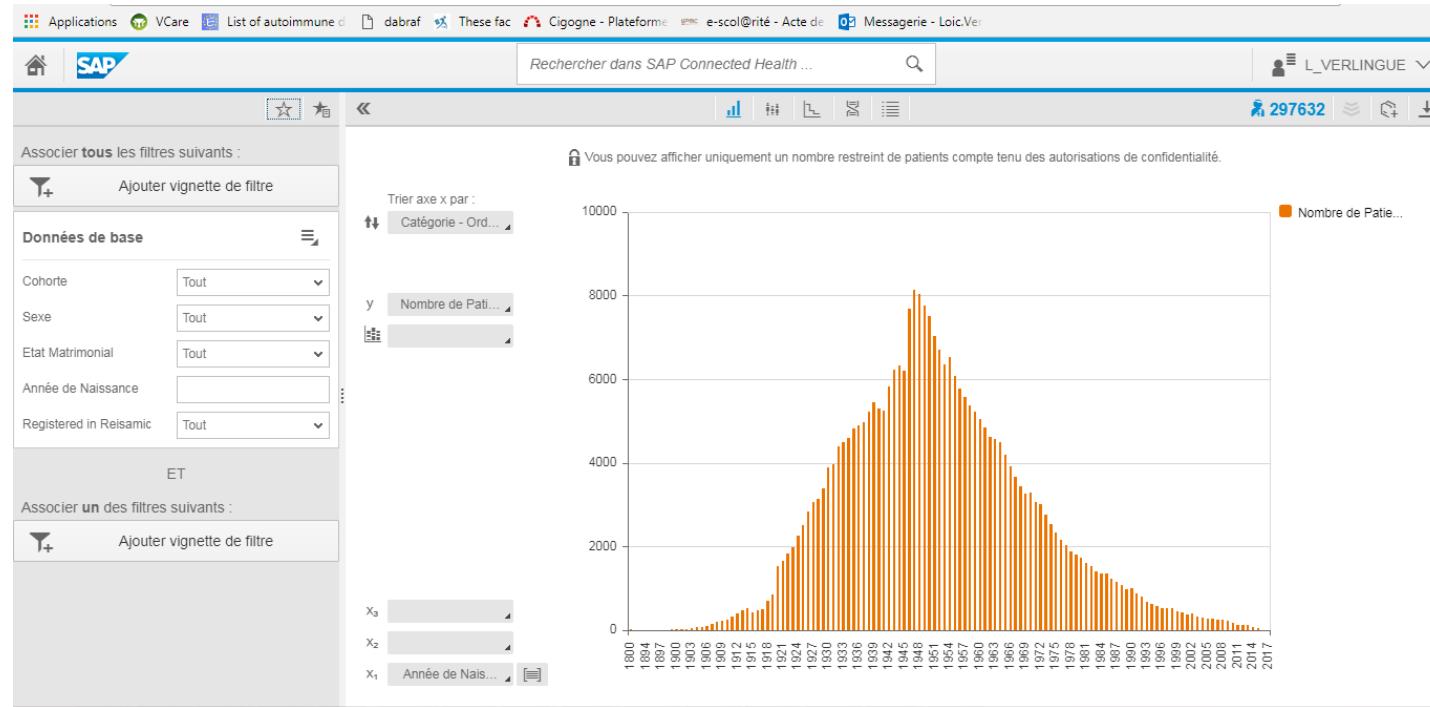
Specialized search engine

The image displays three screenshots illustrating specialized search engines:

- Harmonizome:** A screenshot of the Harmonizome website (amp.pharm.mssm.edu/Harmonizome/). The page features a logo consisting of four colored triangles (orange, yellow, green, blue) forming a larger triangle. Below the logo is the text "Harmonizome" and a search bar with the placeholder "Search for genes or proteins and their functional terms extracted and organized from over a hundred publicly available resources".
- OmicsDI:** A screenshot of the OmicsDI website (OmicsDI.Home). The page has a search bar with dropdown options like "All" and "Example searches" showing results for "achilles", "STAT3", and "breast cancer". Below the search bar is a large word cloud visualization where words related to biological concepts like sequencing, transcriptome, methylation, and expression are prominently displayed.
- Human Protein Atlas:** A screenshot of the Human Protein Atlas website (www.proteinatlas.org). The page includes a search bar with results for "P041", "RBM3", "insulin", and "CD36". It also features three image thumbnails labeled "TISSUE ATLAS", "CELL ATLAS", and "PATHOLOGY ATLAS".

At Gustave Roussy

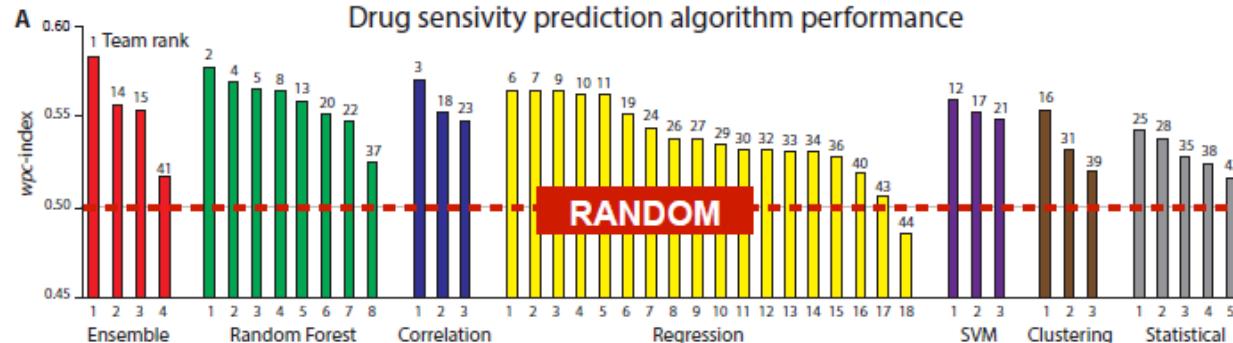
<\\nas-01\SAP-MRI-doc>



Does it work?

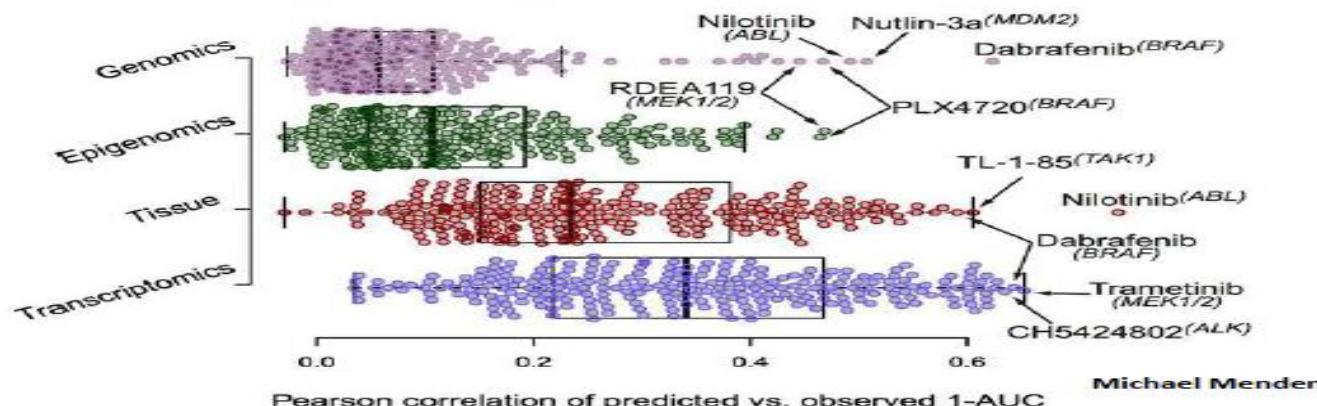
In cell lines

Saez-Rodriguez « predictability is very low »



DREAM CHALLENGES
powered by Sage Bionetworks

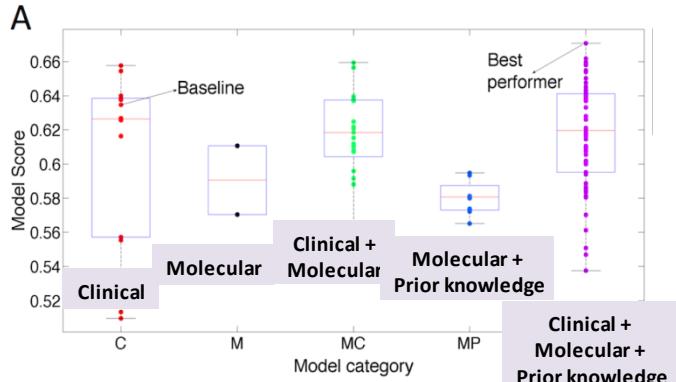
Costello JC, et al. A community effort to assess and improve drug sensitivity prediction algorithms. Nat Biotechnol. 2014



Francesco Iorio et al, Cell, 2016

In the clinic

In 2013



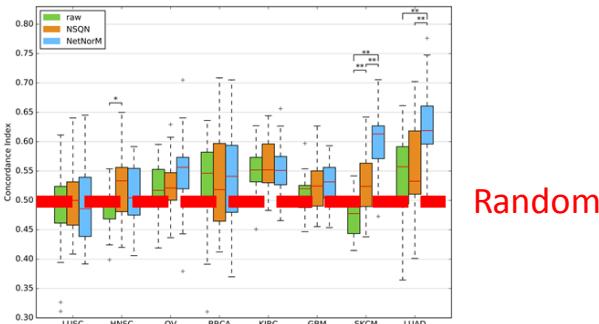
Bilal E, et al. Improving breast cancer survival analysis through competition-based multidimensional modeling. PLoS Comput Biol. 2013

In 2014

- molecular data + clinical variables improved predictions for 3 / 4 cancers
- but gains were limited (2.2-23.9%).

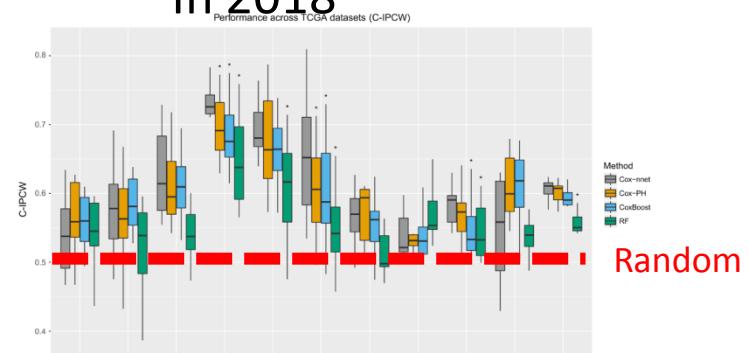
Yuan Y, et al. Assessing the clinical utility of cancer genomic and proteomic data across tumor types. Nat Biotechnol. 2014

In 2017



Adding gene networks information

Using neural networks



Le Morvan M, Zinov'yev A, Vert JP. NetNorM: Capturing cancer-relevant information in somatic exome mutation data with gene networks for cancer stratification and prognosis. PLoS Comput Biol. 2017 Jun 26

Ching T, Zhu X, Garmire LX. Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data. PLoS Comput Biol. 2018 Apr

How to improve models' efficacy?

- Need more samples (cell lines, organoids)
- Need more molecular layers ?
- Need better readouts
- Need more functional (perturbation) data
- Adapt your models to the tasks you want

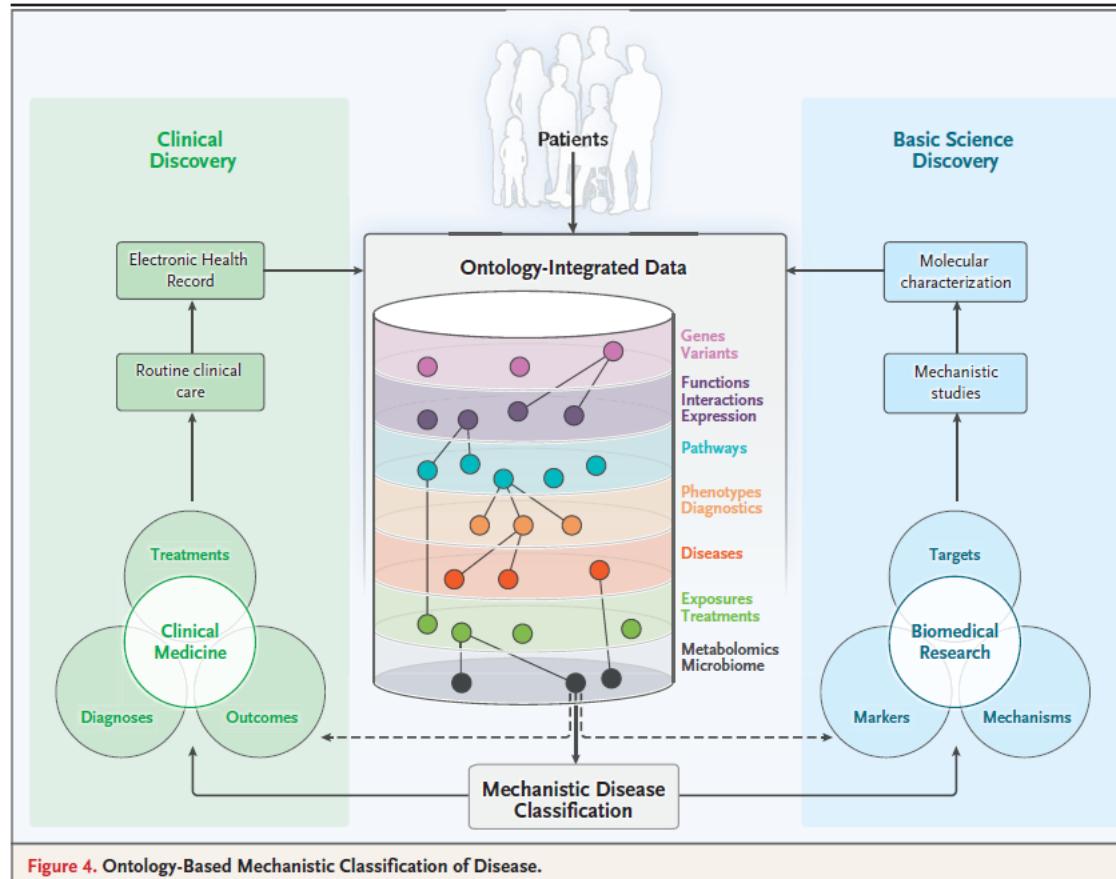
→ Saez-Rodriguez's take

How to improve models' efficacy?

- Focus on our data-analytical approaches
- Use mechanistic prior knowledge
- Adapt your models to the tasks you want

→ Camacho, Costello, Collins takes
→ & my take!

“A goal of precision medicine is to stratify patients in order to improve diagnosis and medical treatment”



Haendel MA, Chute CG,
Robinson PN.
Classification, Ontology,
and Precision Medicine. N
Engl J Med. 2018 Oct

What is cool about deep learning?

You can play with formalisms & architectures

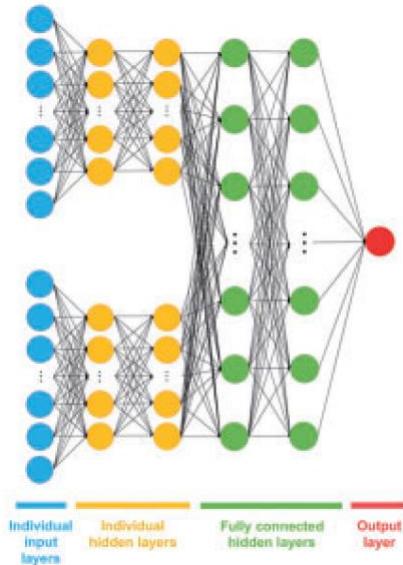
For what ?

To better fit your data and your questions!

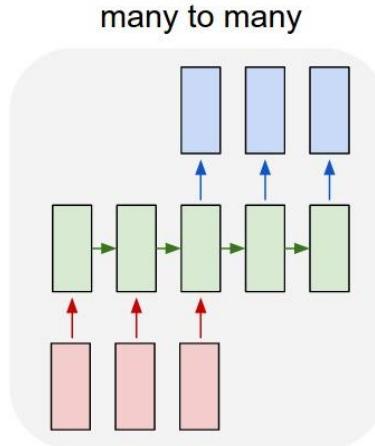
So be innovative!

Find your architecture

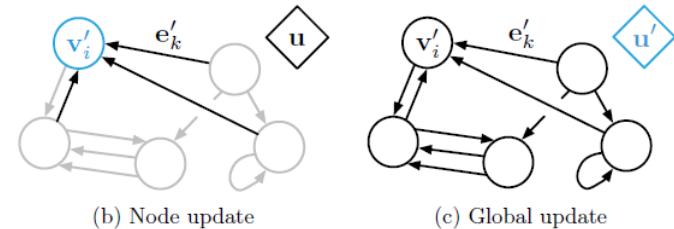
Drugs and proteins
to predict bioactivity
→ **Multitask NN**



Patients' folder
to predict patients' outcome
→ **Recurrent neural nets**

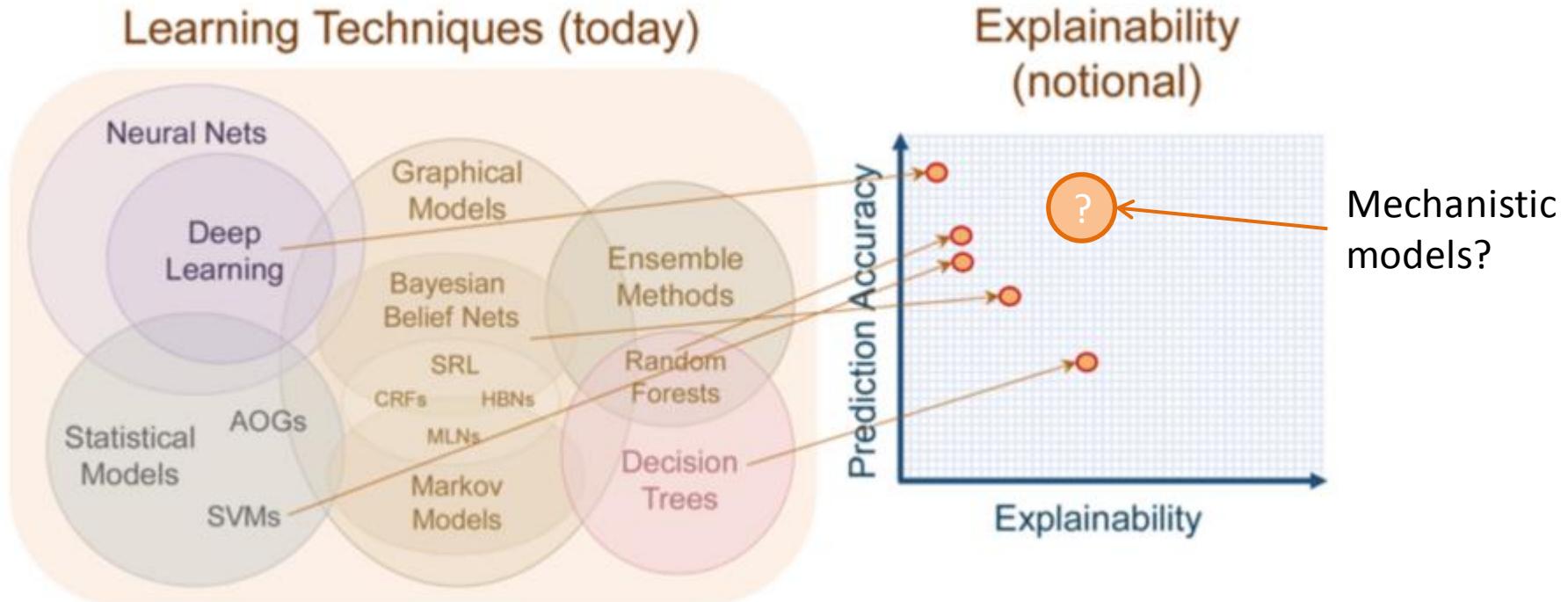


Molecular data
to predict biological behaviors
→ **Graphical networks**



Data science team at

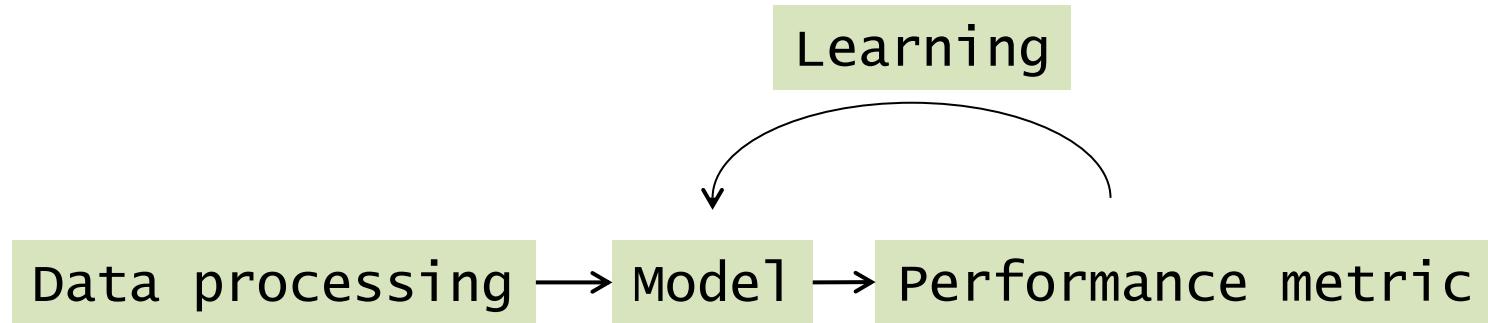
Intepretability of ML/DL



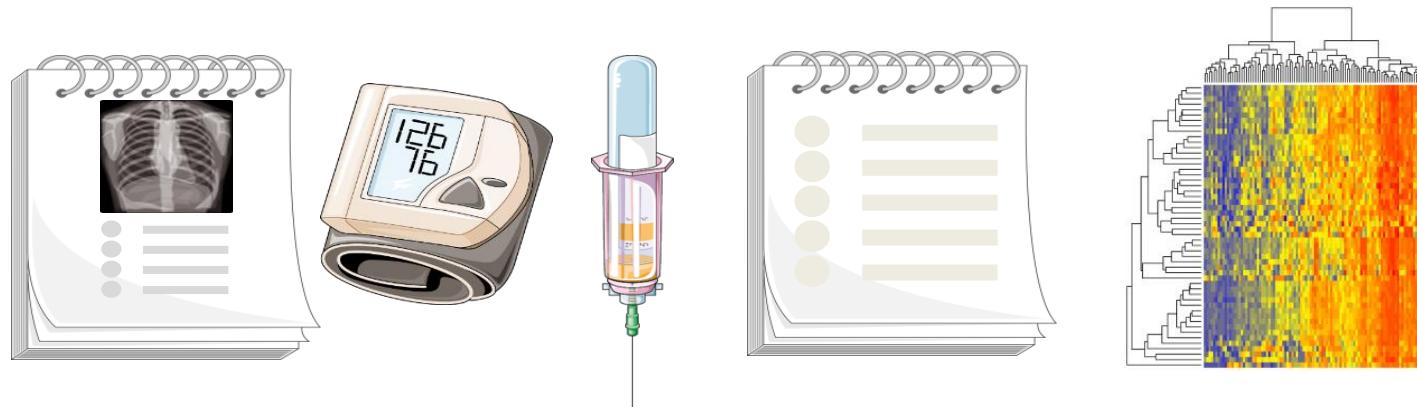
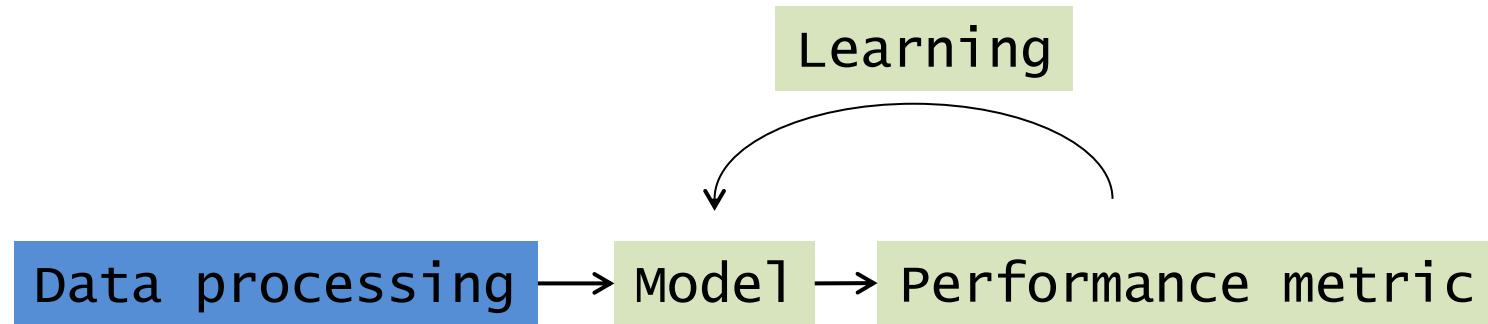
For patients' prognostic estimation

Models Should be improved...

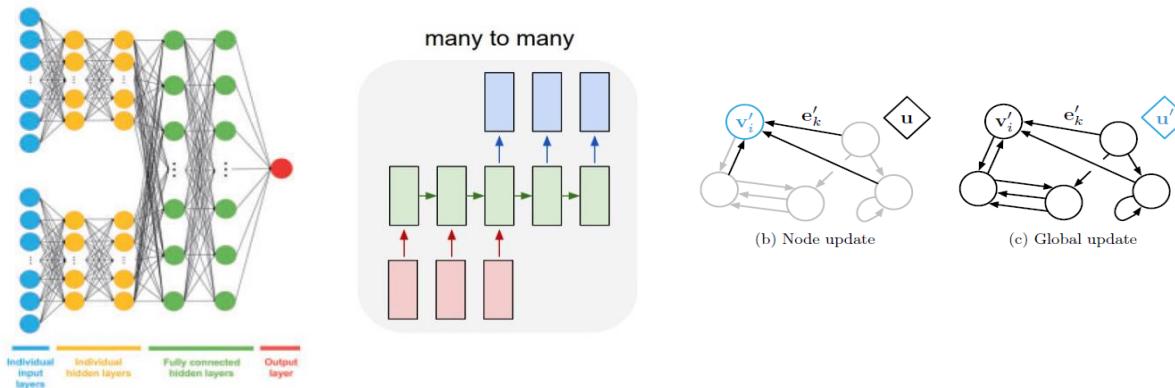
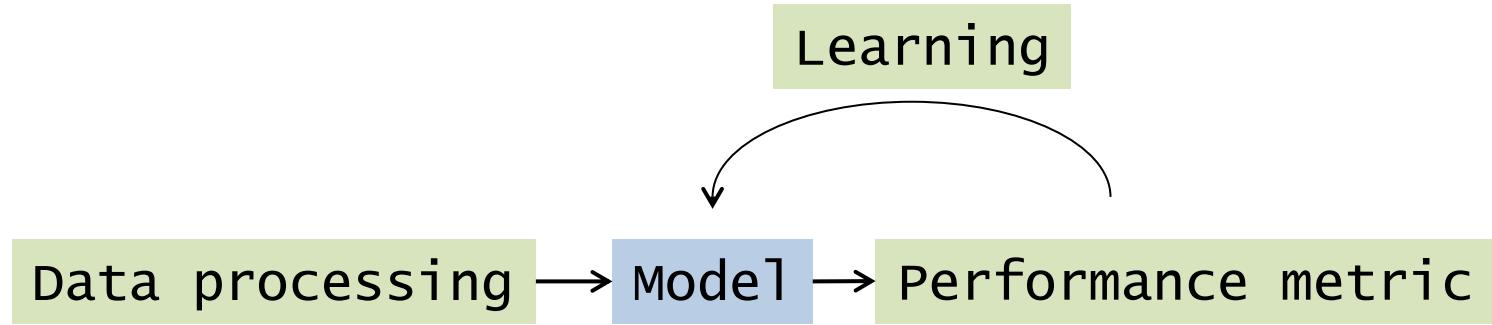
Good practice: design your workflow



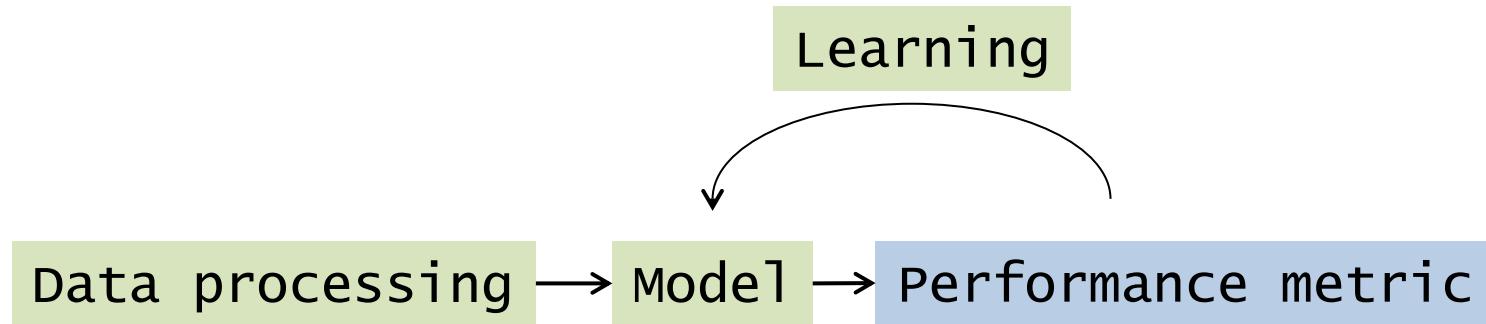
Good practice: design your workflow



Good practice: design your workflow



Good practice: design your workflow



Classification: matrice de confusion

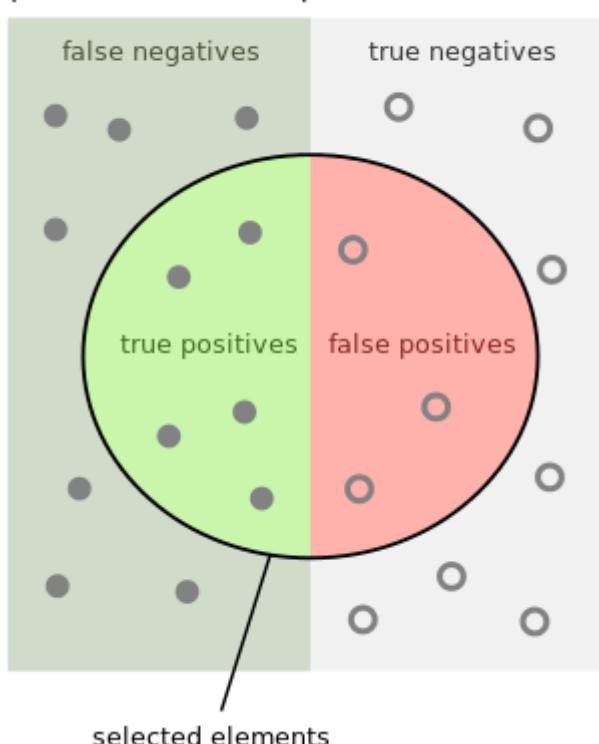
| | | Classe réelle | |
|-----------------|---|---------------|----|
| | | 0 | 1 |
| Classe prédicté | 0 | TN | FN |
| | 1 | FP | TP |

Regression: mean squared error,
corelation, C-index...



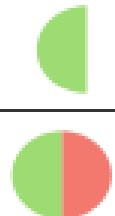
Métriques pour classification binaire

Positives Negatives



How many selected items are relevant?

Precision =



How many relevant items are selected?

Recall =

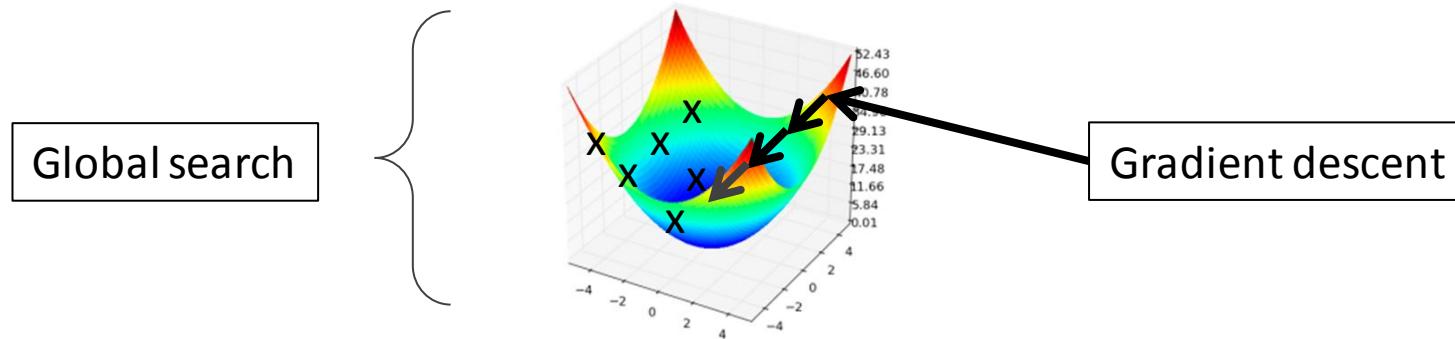
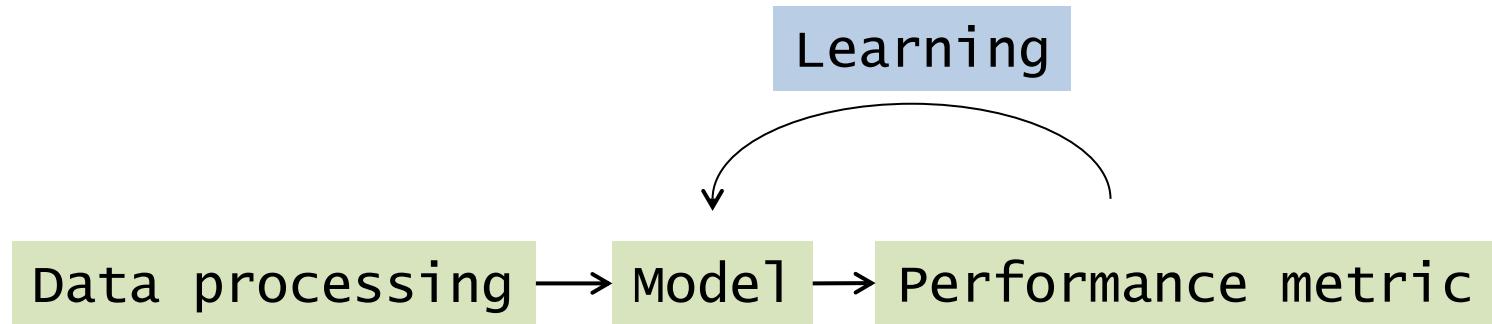


| | | Classe réelle | |
|-------------------|---|---------------|----|
| | | 0 | 1 |
| Classe prédictive | 0 | TN | FN |
| | 1 | FP | TP |

Métriques pour classification binaire

| | Total population | True condition | | Prevalence = $\frac{\sum \text{Condition positive}}{\sum \text{Total population}}$ | Accuracy (ACC) = $\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$ |
|---------------------|---|--|--|--|---|
| Predicted condition | Predicted condition positive | Condition positive True positive, Power | Condition negative False positive, Type I error | Positive predictive value (PPV), Precision = $\frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}}$ | False discovery rate (FDR) = $\frac{\sum \text{False positive}}{\sum \text{Predicted condition positive}}$ |
| | Predicted condition negative | False negative, Type II error | True negative | False omission rate (FOR) = $\frac{\sum \text{False negative}}{\sum \text{Predicted condition negative}}$ | Negative predictive value (NPV) = $\frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}}$ |
| | True positive rate (TPR), Recall, Sensitivity, probability of detection = $\frac{\sum \text{True positive}}{\sum \text{Condition positive}}$ | False positive rate (FPR), Fall-out, probability of false alarm = $\frac{\sum \text{False positive}}{\sum \text{Condition negative}}$ | Positive likelihood ratio (LR+) = $\frac{\text{TPR}}{\text{FPR}}$ | Diagnostic odds ratio (DOR) = $\frac{\text{LR+}}{\text{LR-}}$ | $F_1 \text{ score} = \frac{2}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}}$ |
| | False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$ | Specificity (SPC), Selectivity, True negative rate (TNR) = $\frac{\sum \text{True negative}}{\sum \text{Condition negative}}$ | Negative likelihood ratio (LR-) = $\frac{\text{FNR}}{\text{TNR}}$ | | |

Good practice: design your workflow

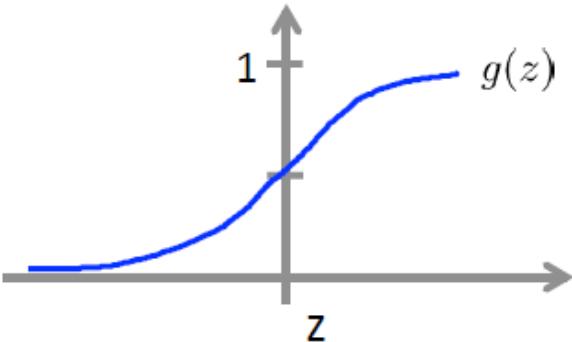


Probability to belong to a class

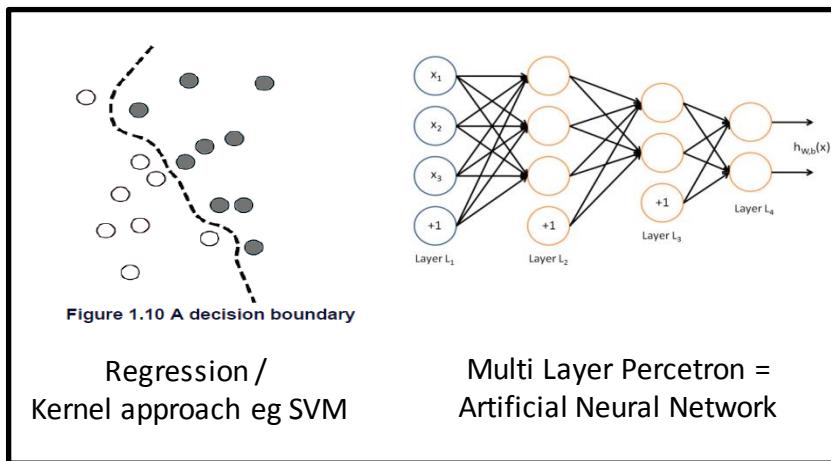
Logistic regression with a sigmoid function

$$h_{\theta}(x) = g(\theta^T x)$$

$$g(z) = \frac{1}{1+e^{-z}}$$



Ok for:



Calculate the cost / error

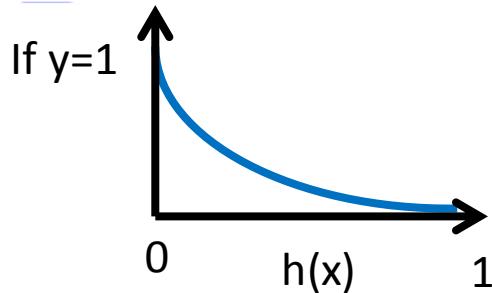
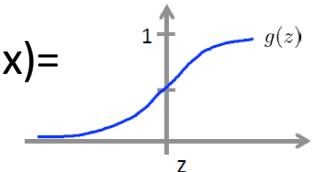
Logistic regression cost function

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

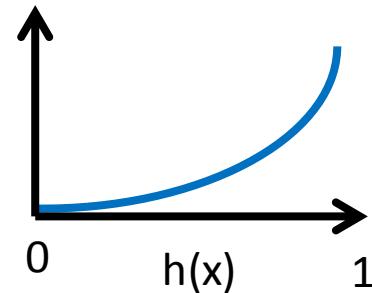
$$= -\frac{1}{m} \left[\sum_{i=1}^m y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)})) \right]$$

prediction

truth



If $y=0$



Parameters' updates with gradient descent

Compute the cost

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

Gradient descent:

Repeat {

Parameters' updates

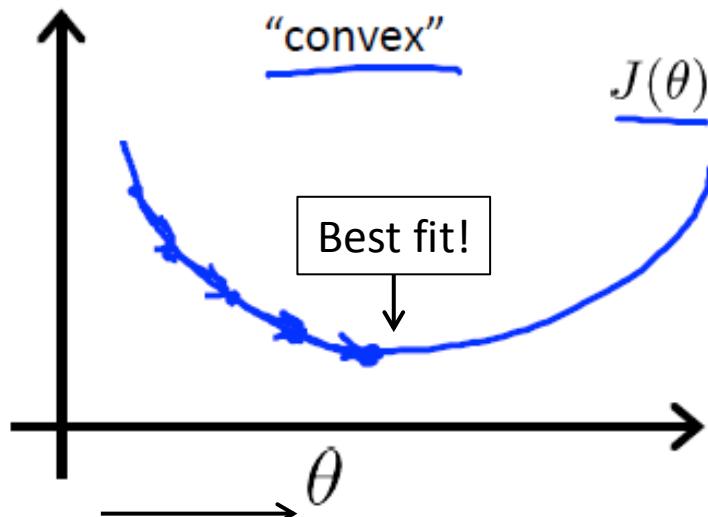
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

Derivative of the cost

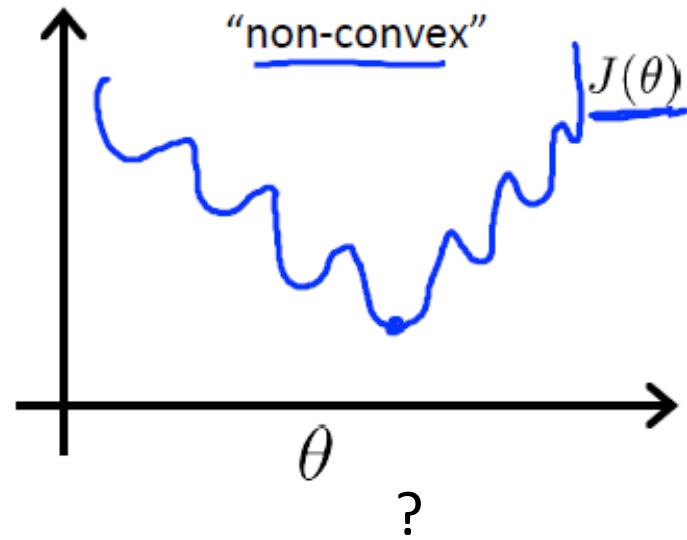
}

Learning rate

Optimisation of the cost



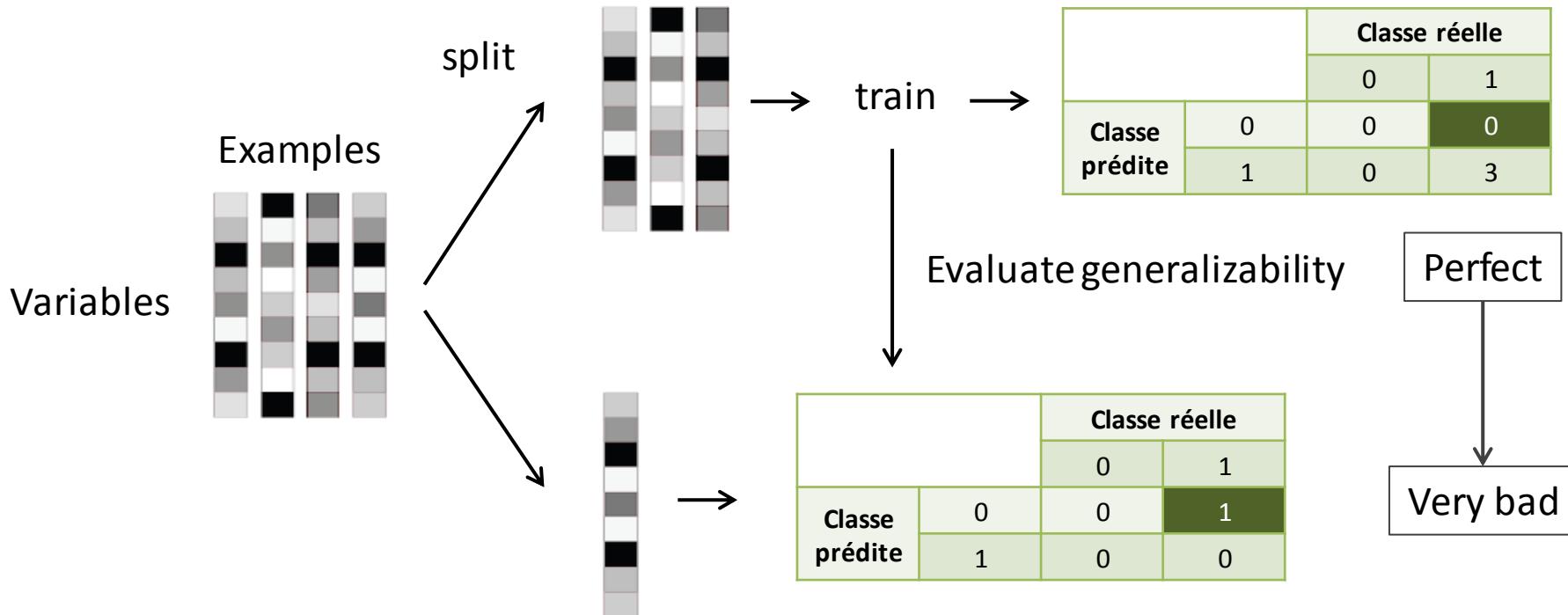
Updates with
gradient descent



Figures from Andrew Ng Coursera

Good practice: test generalization!

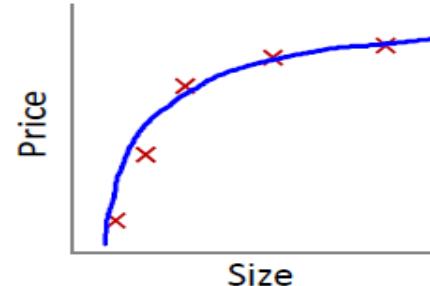
Common to statistics, machine learning, deep learning and sysbio learning



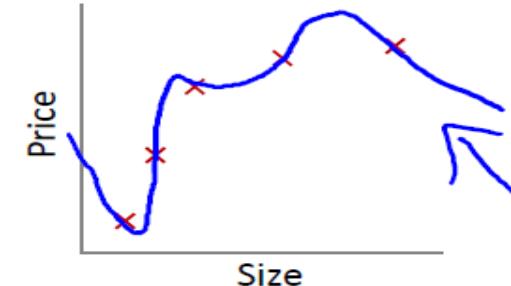
The problem of overfitting



$\rightarrow \theta_0 + \theta_1 x$
 "Underfit" "High bias"



$\rightarrow \theta_0 + \theta_1 x + \theta_2 x^2$
 "Just right"



$\rightarrow \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$
 "Overfit" "High variance"

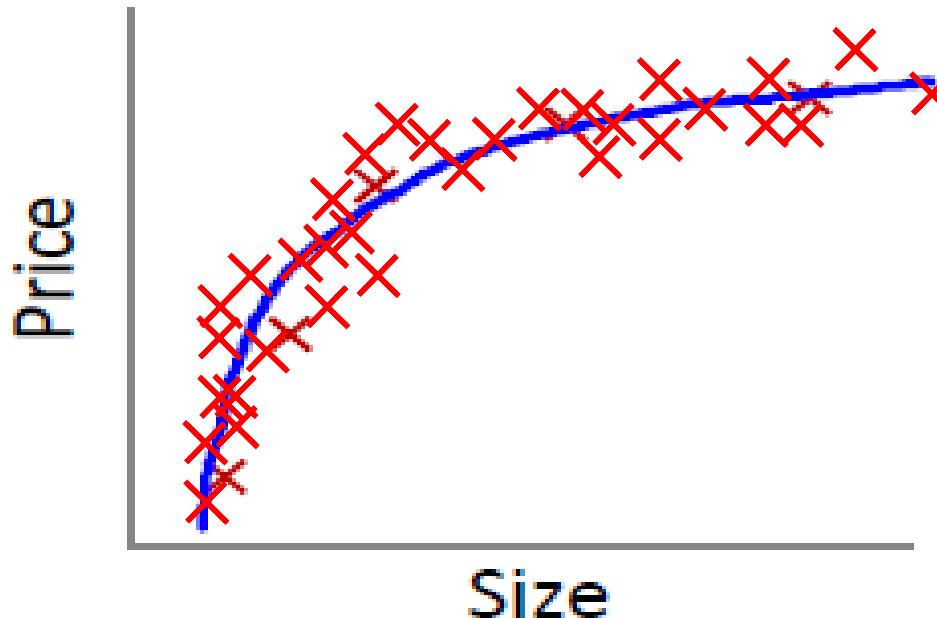
Overfitting: If we have too many features, the learned hypothesis may fit the training set very well ($J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 \approx 0$), but fail to generalize to new examples (predict prices on new examples).

Strategies to address overfitting

- Penalize models
- Add ontologies / knowledge (difficult)
- Increase data for better learning generalization:
 - Very similar data to yours
 - Close data → transfer learning

The problem of overfitting

One solution: increase observations



Other solutions: penalization strategies, simple model....

Price matters



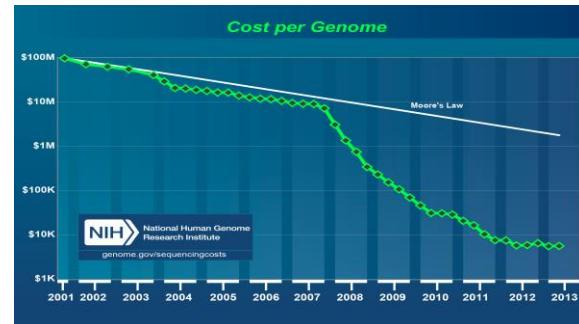
Illumina Unveils New High-Throughput Sequencing Instrument at JP Morgan

Jan 10, 2017 | Monica Heger

This story has been updated to clarify that the cost of sequencing on the NovaSeq will be less expensive than the HiSeq X by 2018, not at launch.

SAN FRANCISCO (GenomeWeb) – Illumina launched a new high-throughput sequencing instrument, NovaSeq, at the JP Morgan Healthcare Conference here this week, which it will begin shipping in limited quantities by the end of the first quarter.

The platform makes use of nanofabricated flow cells and specially designed optics that CEO Francis deSouza said would in future iterations enable a human genome to be sequenced for \$100, although he did not provide a timeline.

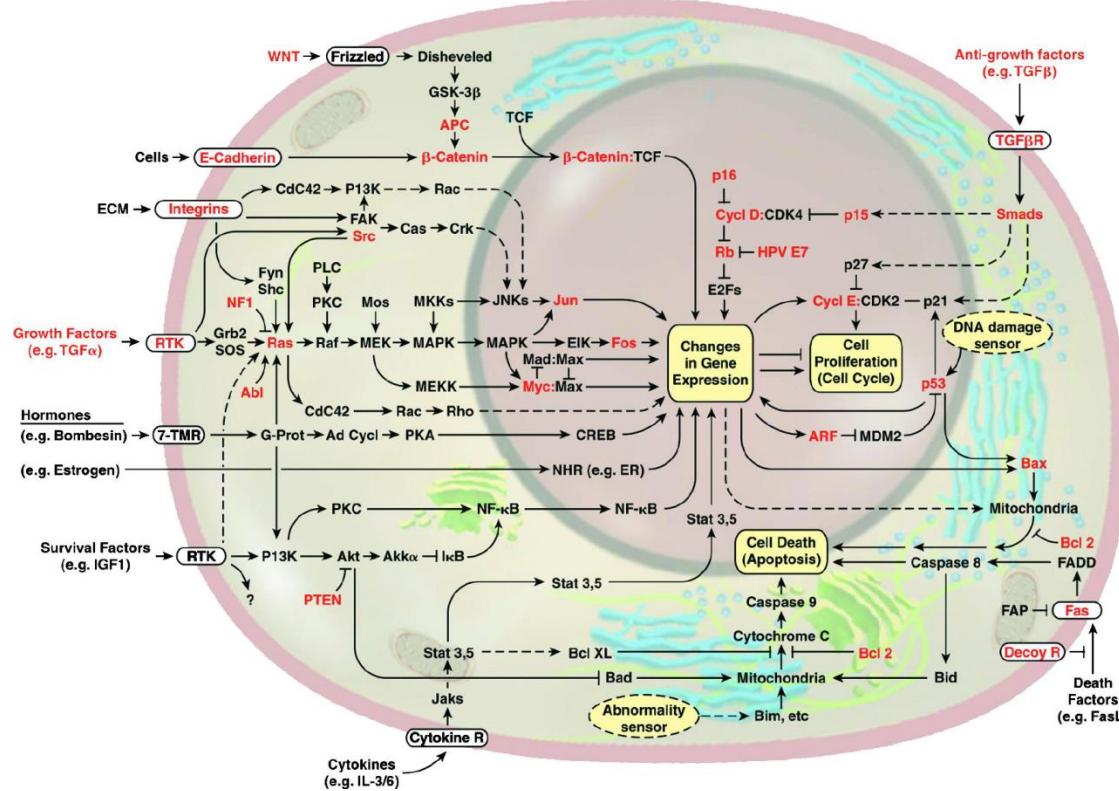


Towards genome seq ~ \$50-100 !!!

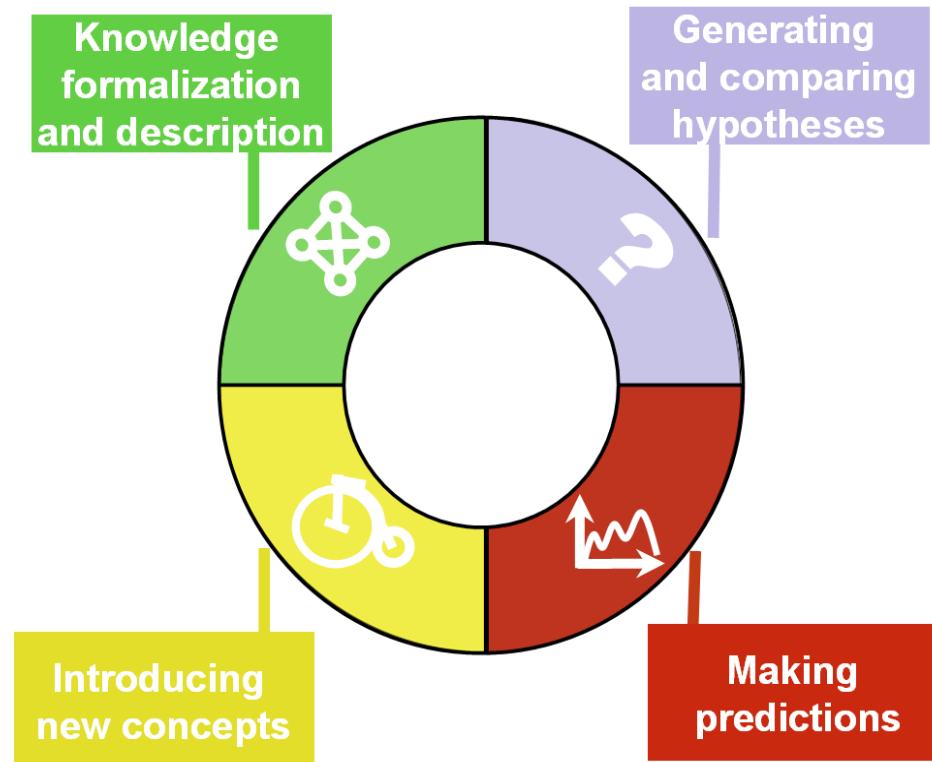
Model generalization with limited observations

Add ontologies / knowledge

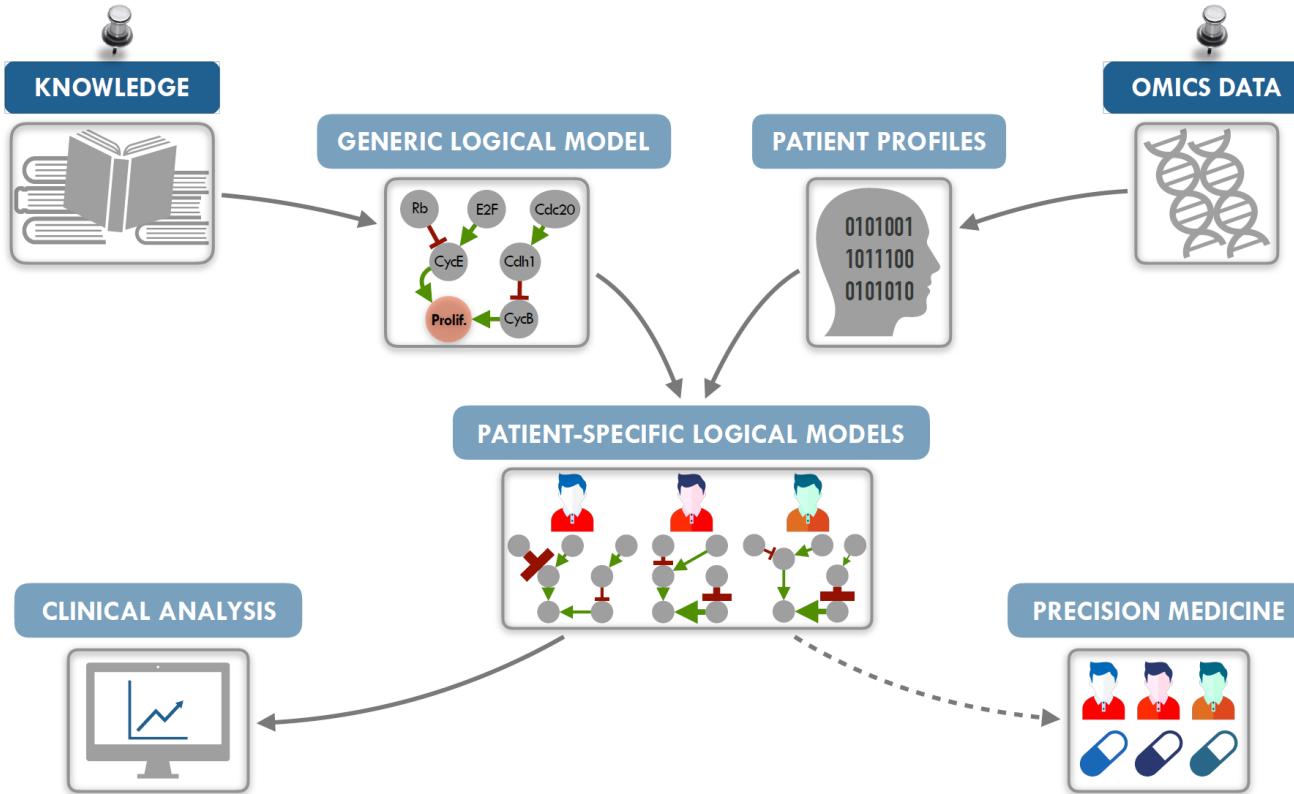
The great thing about modeling biology!



Why would you use it?



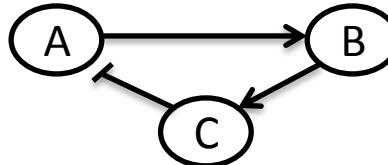
KNOWLEDGE + DATA = ACTIONNABLE MODELS



Jonas Beal

Why it is relevant to biology?

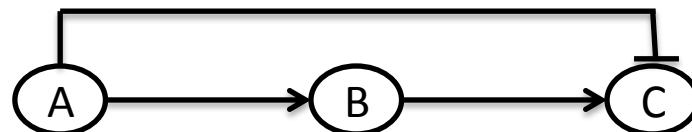
- ⇒ **Negative** feedback loops can give rise to oscillations, e.g. to control protein levels, in cell cycle and circadian cycle...



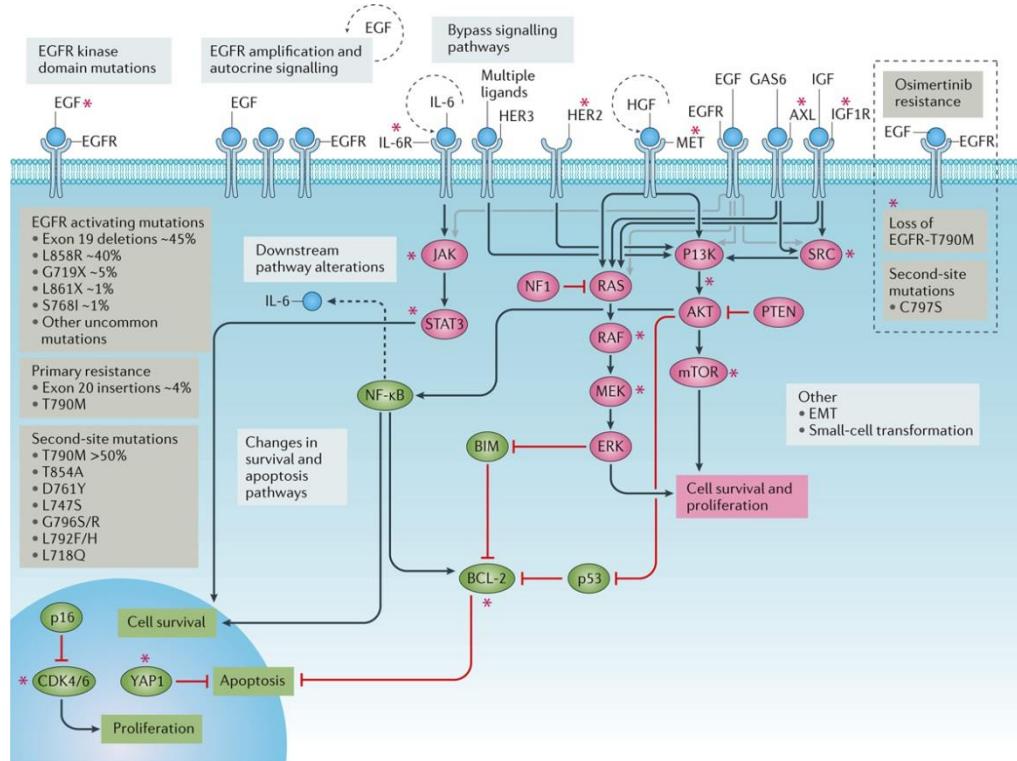
- ⇒ **Positive** feedback loops can give rise to ultrasensitive (all or none) responses, filter or amplify signals, e.g. MAPK signalling pathway



- ⇒ **Feed-forward** loops can generate delays, etc.



Remember that most resistance mechanisms are pathways driven

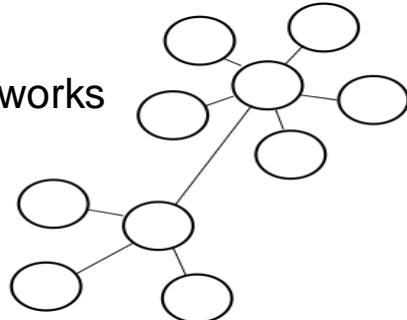


How would you use it?

→ with mathematical modeling!

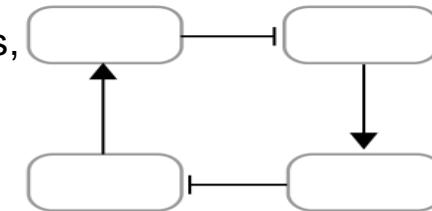
Statistical modelling

PPI networks

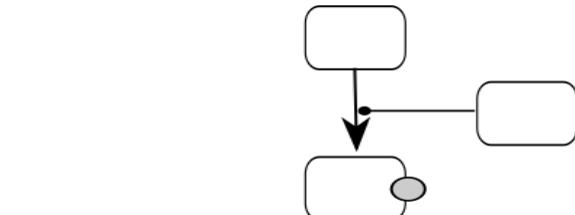


Logical modelling

Signalling pathways,
gene regulatory
networks

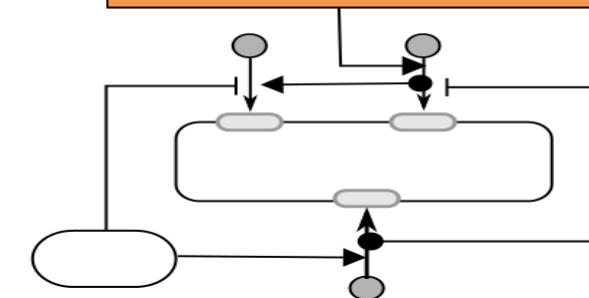


Process modelling (FBA, ODEs...)



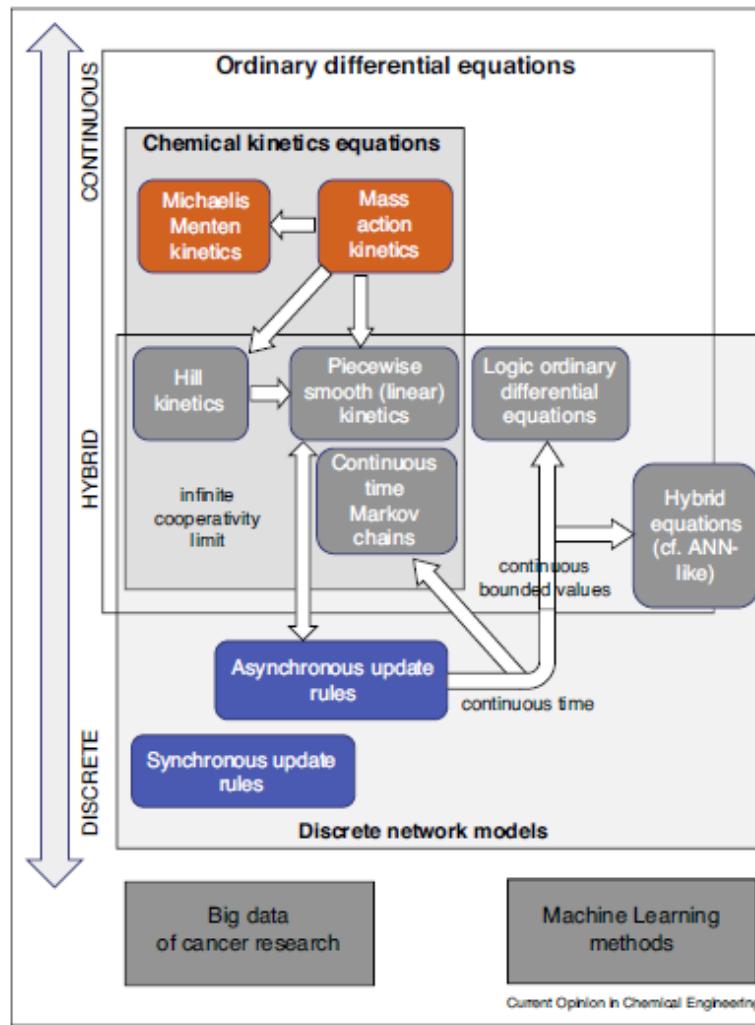
Metabolic networks, reaction networks

Rule-based modelling



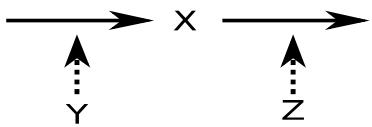
Molecular interactions, reaction networks

Ok but how do I choose my maths?



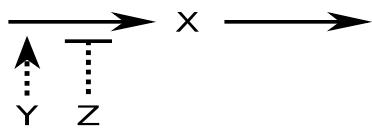
Logical versus kinetic modeling of
biological networks: applications in
cancer research Laurence Calzone,
Emmanuel Barillot and Andrei Zinovyev

Translation of a reaction network into a set of ODEs

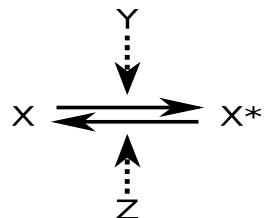


$$\frac{dX}{dt} = \dot{X} = k_1 \cdot Y - k_2 \cdot Z \cdot X$$

$$\frac{dX}{dt} = \dot{X} = \frac{k_1 \cdot Y^n}{K^n + Y^n} - k_2 \cdot Z \cdot X$$



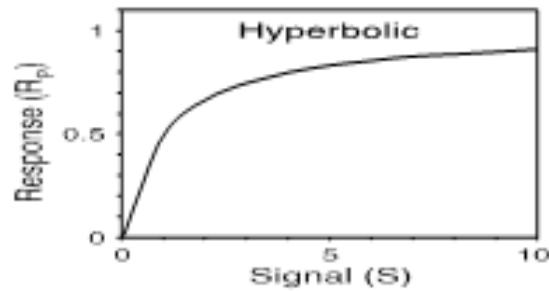
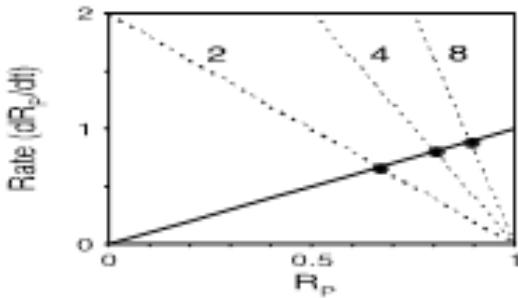
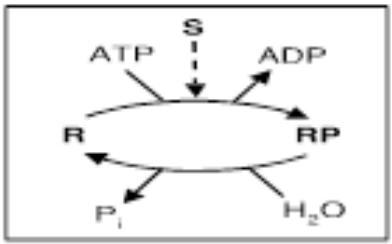
$$\frac{dX}{dt} = \dot{X} = \frac{k_1 \cdot Y}{k_{1'} + k_{1''} \cdot Z} - k_2 \cdot X$$



$$\frac{dX}{dt} = \dot{X} = \frac{k_1 \cdot Y \cdot (X_{tot} - X)}{K_1 + (X_{tot} - X)} - \frac{k_2 \cdot Z \cdot X}{K_2 + X}$$

Goldbeter-Koshland switch

Simulating ODEs

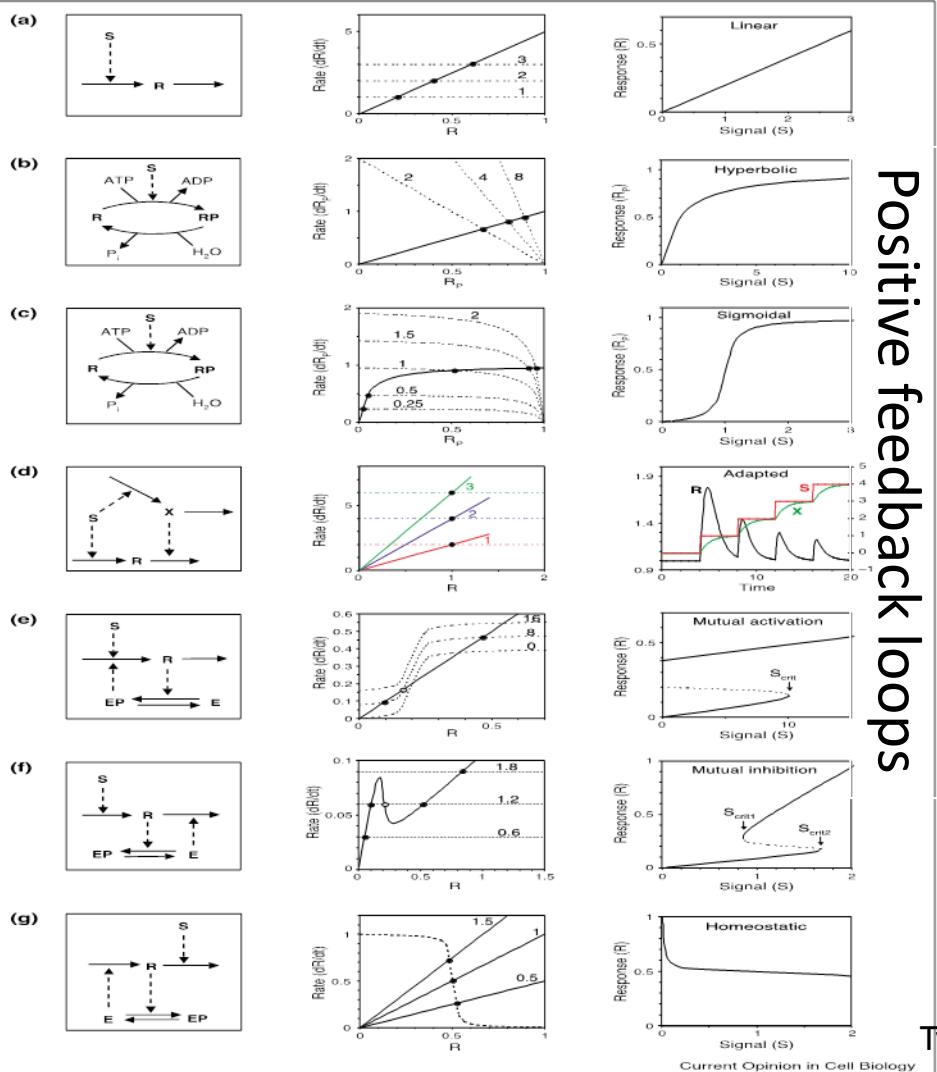


Simulating ODEs

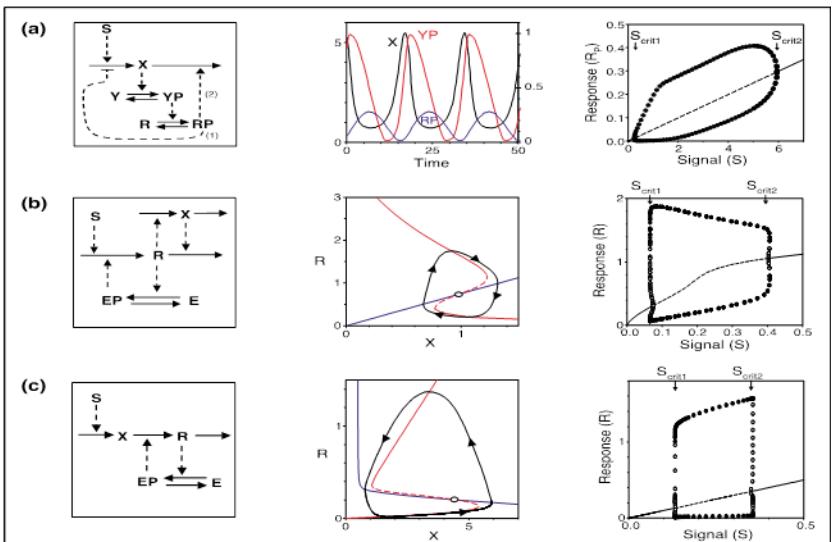
Motifs in biochemical reaction networks

Responses to positive and negative feedback loops

Negative feedback loops

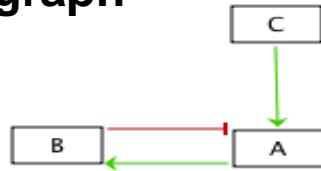


Tyson, Chen and Novak, 2003, Current Opinion in Cell Biology



Boolean logic

Regulatory graph



Each variable can take two states: 0 or 1

Logical rules

$$A = !B \& C$$

$$B = A$$

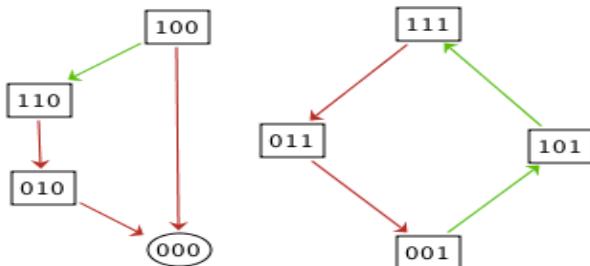
$$C = \text{input}$$

Boolean logic:

Connectors: AND (&), OR (|), NOT (!), XOR (/)

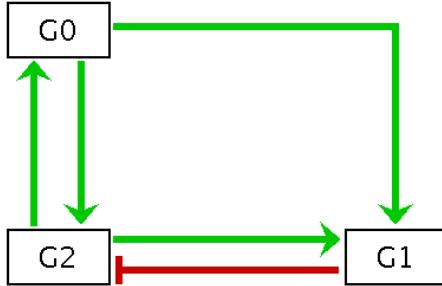
Logic depends on incoming arrows

Solutions



Attractors are subgraphs of the **state transition graph** with no outgoing arrows.
Ex: stable steady states & cyclic attractors

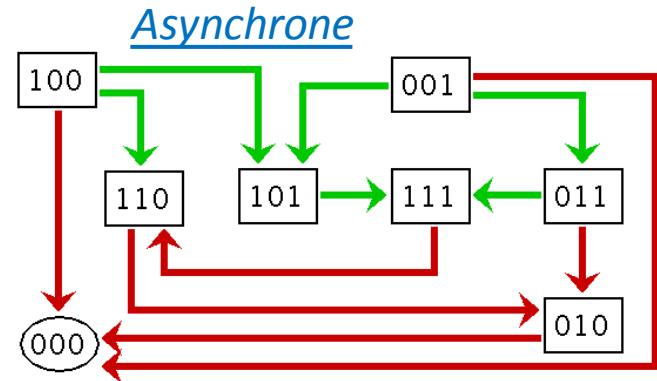
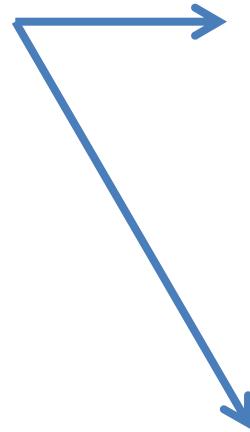
Toy model's simulation



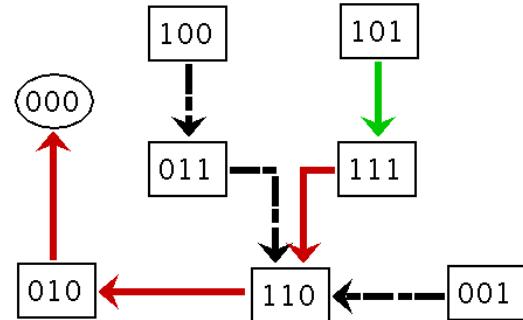
$G_0 = G_2$

$G_1 = G_0 \mid G_2$

$G_2 = G_0 \& \neg G_1$



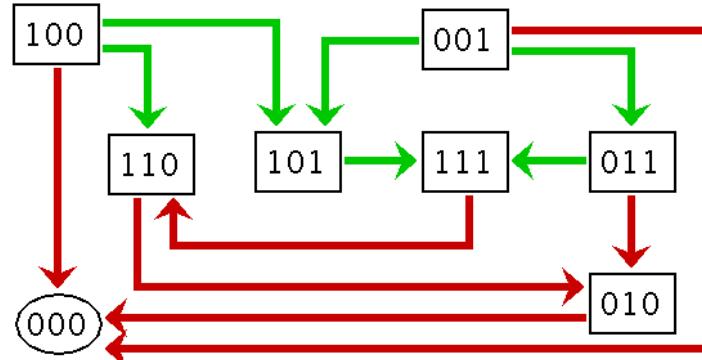
Synchrone



Simulation BKMC with MABOSS*

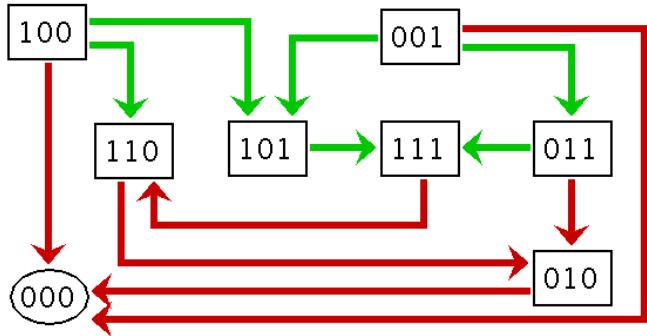
(Boolean Kinetic Monte Carlo)

- 1) Chaîne de Markov → la probabilité d'atteindre un état ne dépend que de l'état où on se trouve actuellement
- 2) A temps continu → Le prochain état visité est indépendant du temps passé dans l'état actuel
→ Le temps moyen passé dans un état ne dépend que des taux de transitions

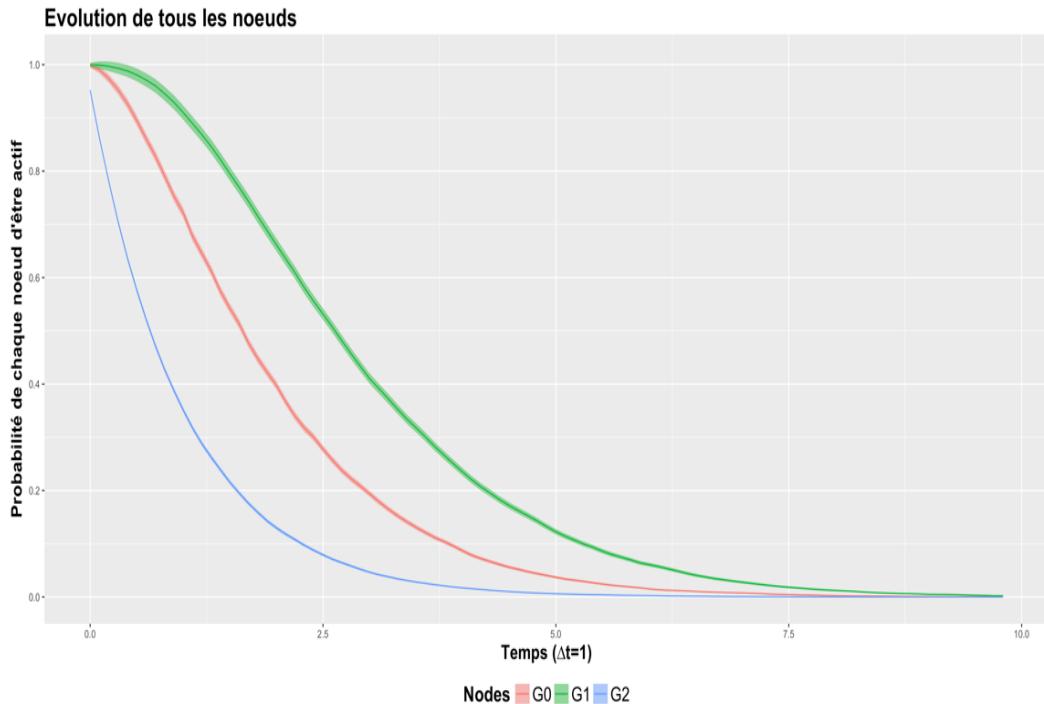


Simulation BKMC

Trajectoire des états

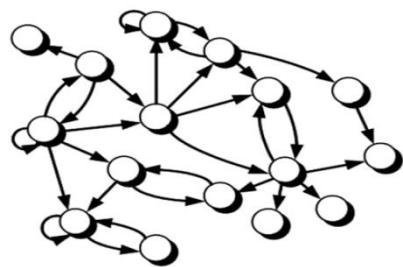


Trajectoire des nœuds

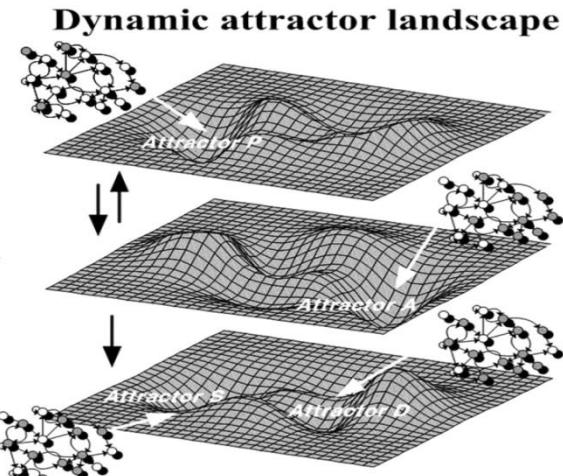


3D simulation of boolean model

Complex regulatory network
(p53 regulatory network)



No DNA damage
Reparable DNA damage
Irreparable DNA damage



Cellular state transition

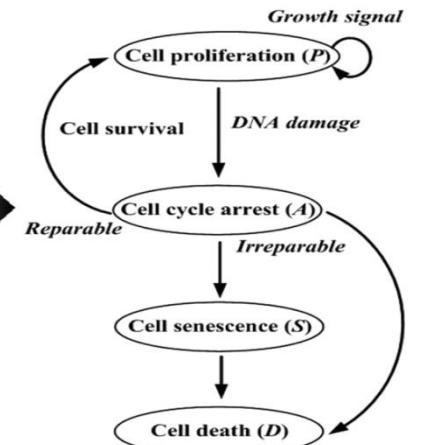
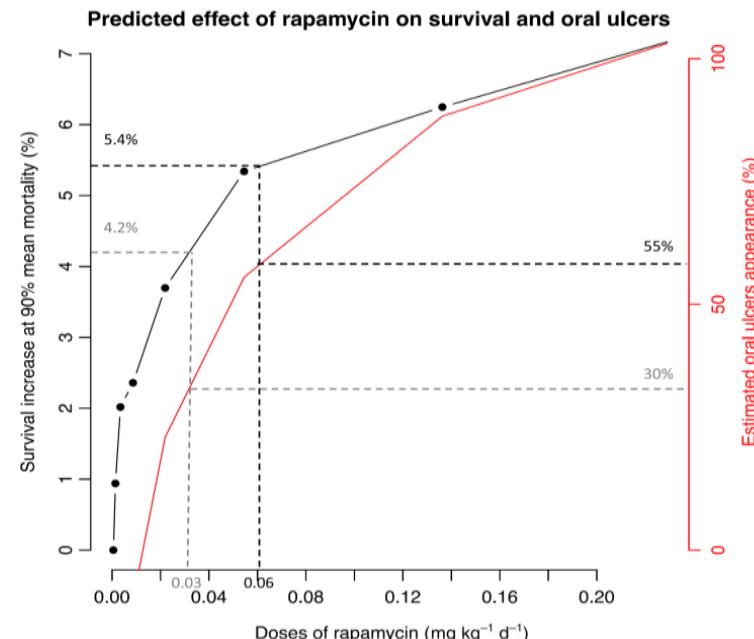
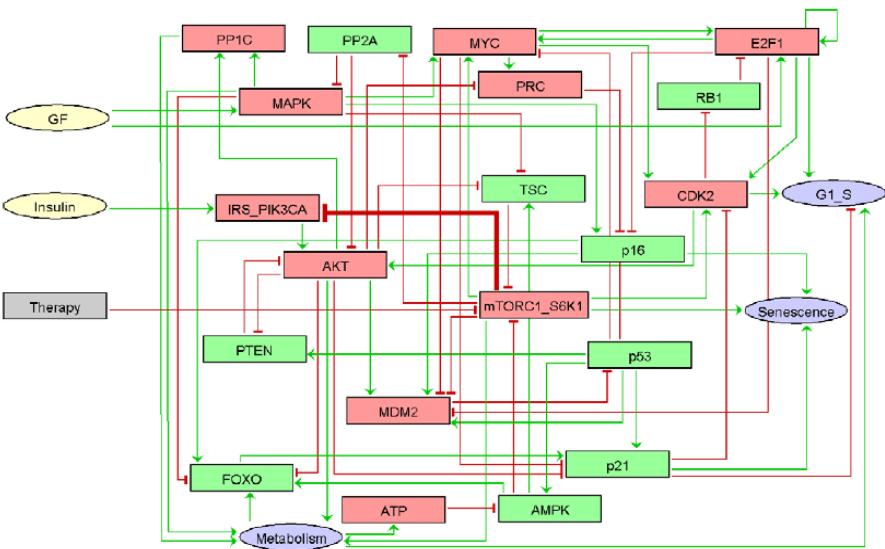


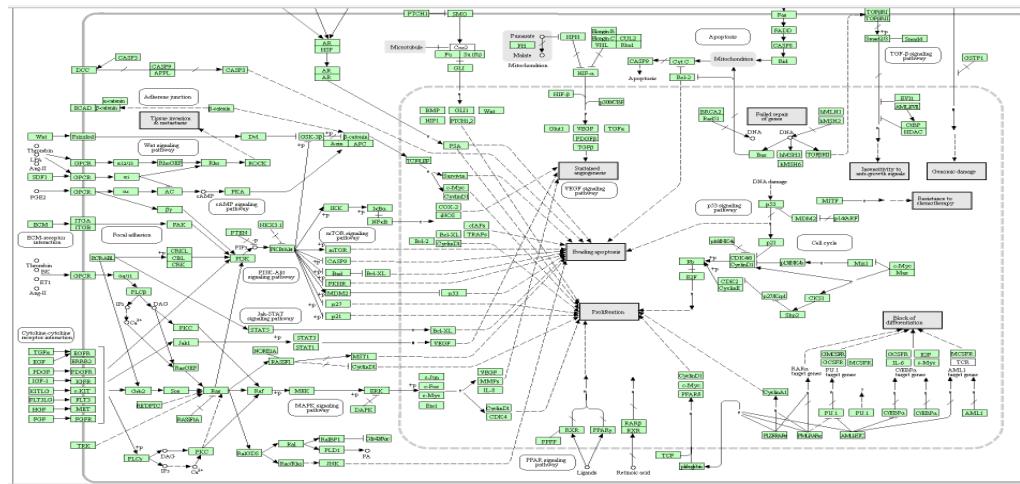
Fig. 6. A simplified view of cellular state transitions identified from the effects of inhibitory treatments on the dynamic landscape in normal cells. Complex regulatory networks are represented by a dynamic landscape that

An application: predict MTORi effect



Logical modeling can be used to simulate molecular alterations, pharmacodynamics, ..., to estimate phenotypes (ORR, PFS, OS)

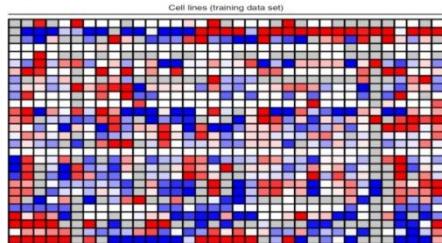
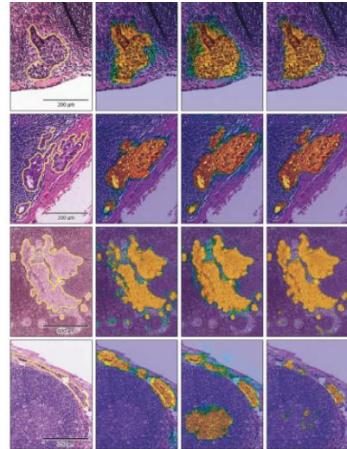
Dynamic regulation networks



Cancer cells from KEGG pathway database

Are possible to turn into models
Based on prior knowledge
Simulated with genomics and treatments

ML/DL in oncology



"Here's my sequence"

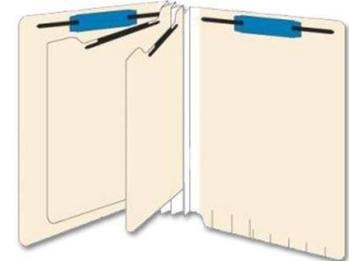
The New Yorker

Diagnostics

Molecular biology
& pronostic

Drug development
& prediction

Monitoring patients



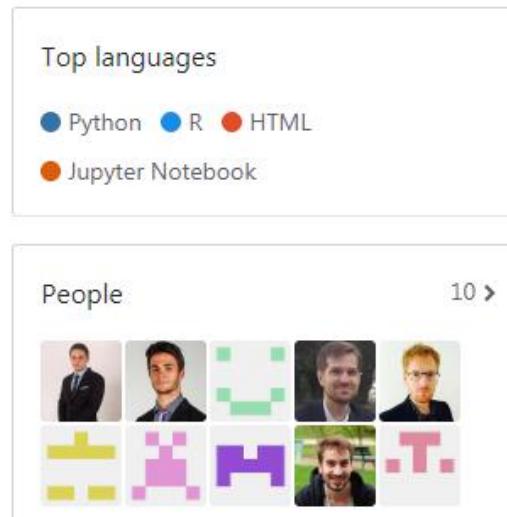
Data Science at DITEP

Info <https://github.com/DITEP>

Medical team



Data Science team



- 1 Senior Bioinformatician
Leo Colmet Daage
- Students from CentraleSuppelec
- 1 Student from Telecom Paristech
- For 2019: MSc, PhDc

loicverlingue@yahoo.fr