ELMED219

Computational medicine: Introduction and motivation

Arvid Lundervold (UiB) and Alexander S. Lundervold (HVL)

Learning outcomes

Get to know **vocabulary** and **central concepts** within **computational medicine**
incl. machine learning

with a focus on multiscale, model organisms, molecules / cells / tissues / organs, “biomarkers”, “deep phenotyping”, “convergence”,

predictive models $y \approx f(X; \theta)$

TOPICS

- Model organisms: fish / flies / worms /… / man
- *In vivo* - *in vitro* - *in silico*
- Wet-lab / dry-lab / moist-lab
- On measurement and math (Lord Kelvin, Richard Feynman)
- What is computational medicine?
- Personalized medicine / precision medicine
- Computational imaging - generic methods
- Kidney function & brain (tumor) segmentation
- Predictive models: $y \approx f(X; \theta)$ y: outcome, f: model, X: data, θ: model parameters
- P4-medicine (predictive, preventive, personalized, participatory)

Computational approaches to organisms $\mathcal{O}_i$

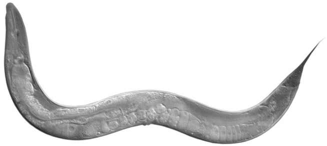
$\mathcal{O}_1$ **fish** *Zebrafish*

http://thezebrafishlab.com/zebrafish-and-human1

$\mathcal{O}_2$ **flies** *Drosophila*

http://www.cam.ac.uk/research/features/how-close-are-you-to-a-fruit-fly

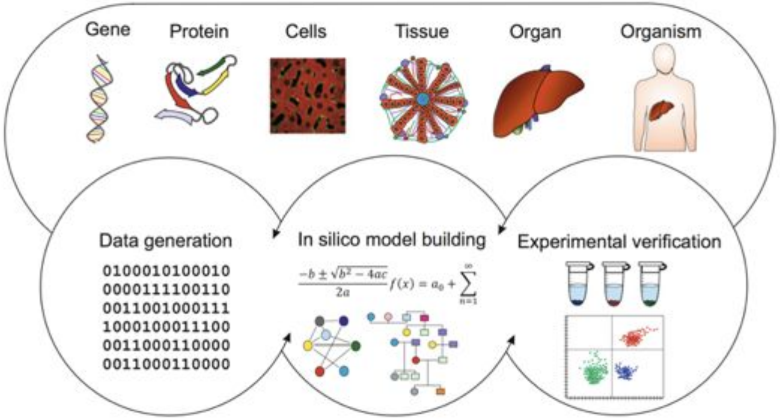
$\mathcal{O}_3$ **worms** *C. elegans*



https://macaulay.cuny.edu/eportfolios/wormsandlearning/what-is-caenorhabditis-elegans

. . . . .

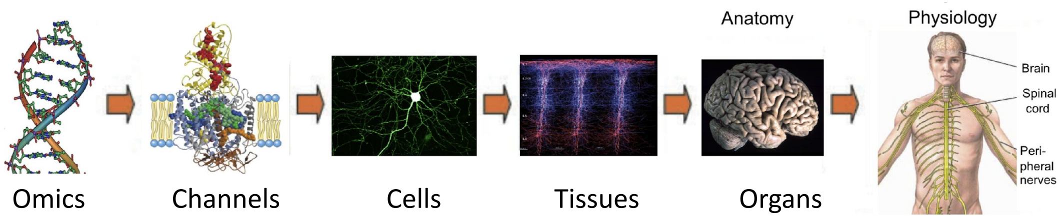
$\mathcal{O}_n$ **man** (in health & disease)



Wet-lab **Dry-lab** **Moist-lab**

(trans-disciplinary teams)

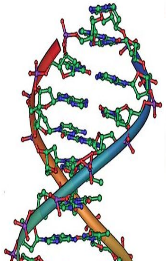
Python
Jupyter notebooks
GitHub



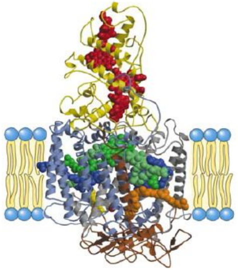
Computational medicine – from molecules to man

Space:

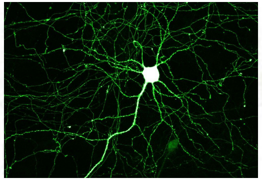
Atom
10^{-12} m



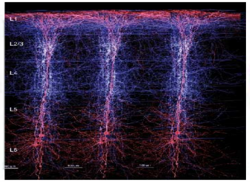
Protein
10^{-9} m



Cell
10^{-6} m

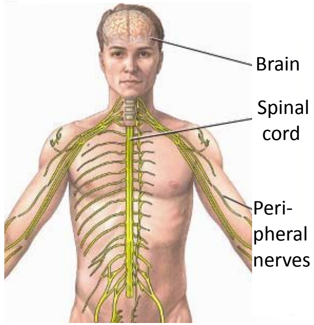


Tissue
10^{-3} m



Organ
10^{0} m
Anatomy

Organ system & organism
Physiology



Modelling :

Gene networks

Pathway models

Stochastic models

Ordinary Differential Equations

Continuum models (Partial Differential Equations)

Systems models

Time:

10^{-6} s
molecular events
(ion channel gating)

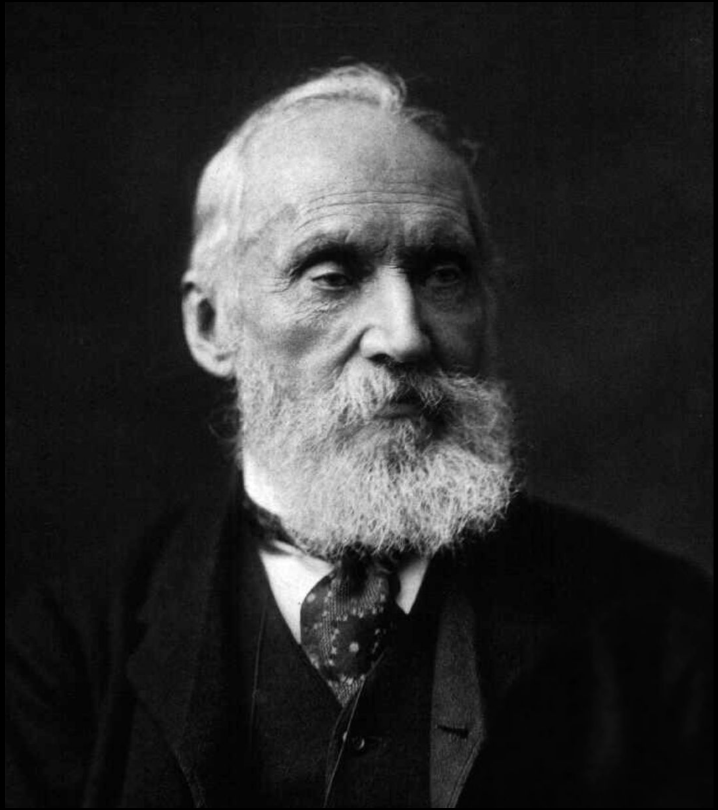
10^{-3} s
diffusion
cell signalling

10^{0} s
motility

10^{3} s
mitosis

10^{6} s
protein
turnover

10^{9} s
human
lifetime



"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be."

- Lord Kelvin, 1883

“ People who wish to analyze nature without using mathematics must settle for a reduced understanding.”

- Richard Feynman (1918-1988)

clinicians

MEDICAL PRACTICE and the «new»

COMPUTATIONAL DISCIPLINES (computational X)

«*body engineers*»

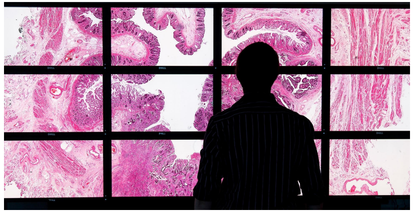
Biomedicine: *wet-lab* ➡ *dry-lab*

What is "**computational medicine**"?

"Application of methods from engineering, mathematics and computational sciences aiming for deeper understanding and better treatment of diseases and disease processes"

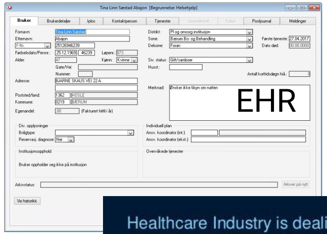
– Rai Winslow, Director, Institute for Computational Medicine, Johns Hopkins University

... by acquisition, organization, analysis, modelling, and visualization of large, rich, and heterogeneous data:

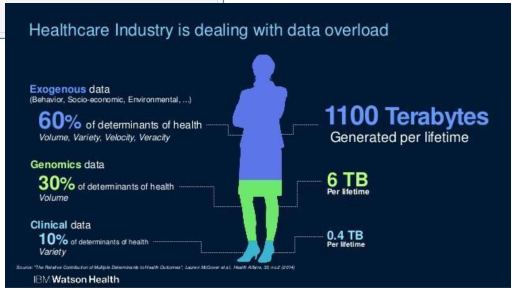


Digital pathology

CT/MRI/PET scanners

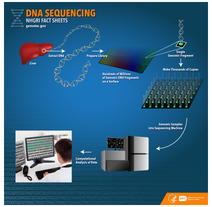


Electronic health records



Clinical chemistry

Sequencing / OMICS

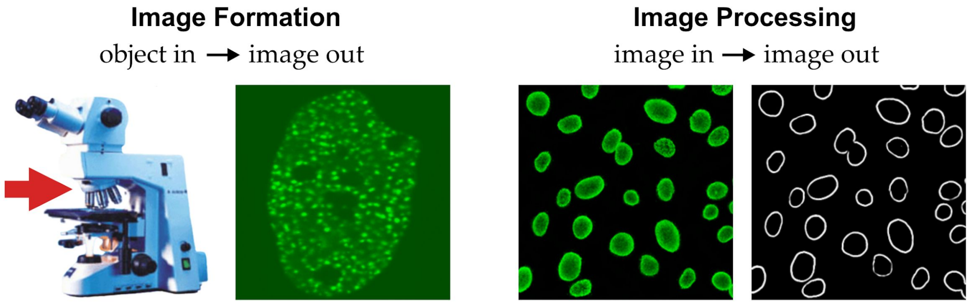


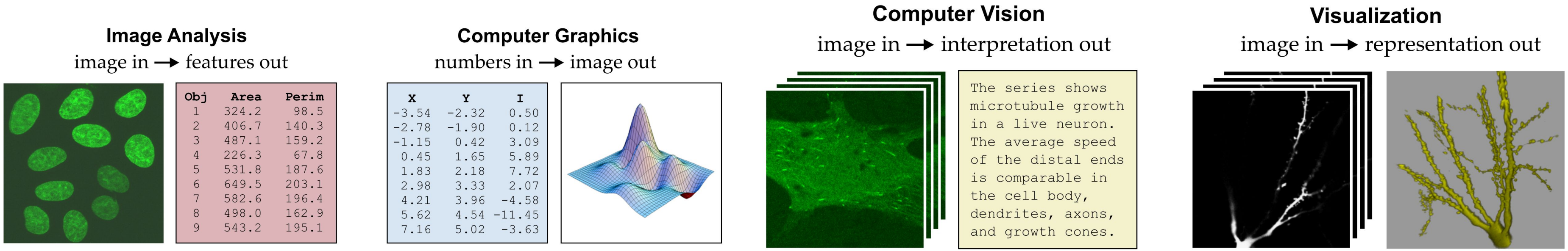
Smartphones / wearables

Computational imaging

- Machine learning

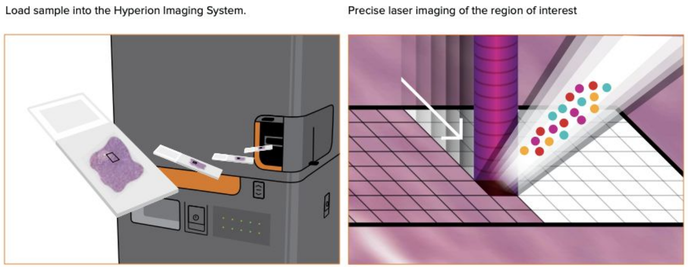
- Visualization

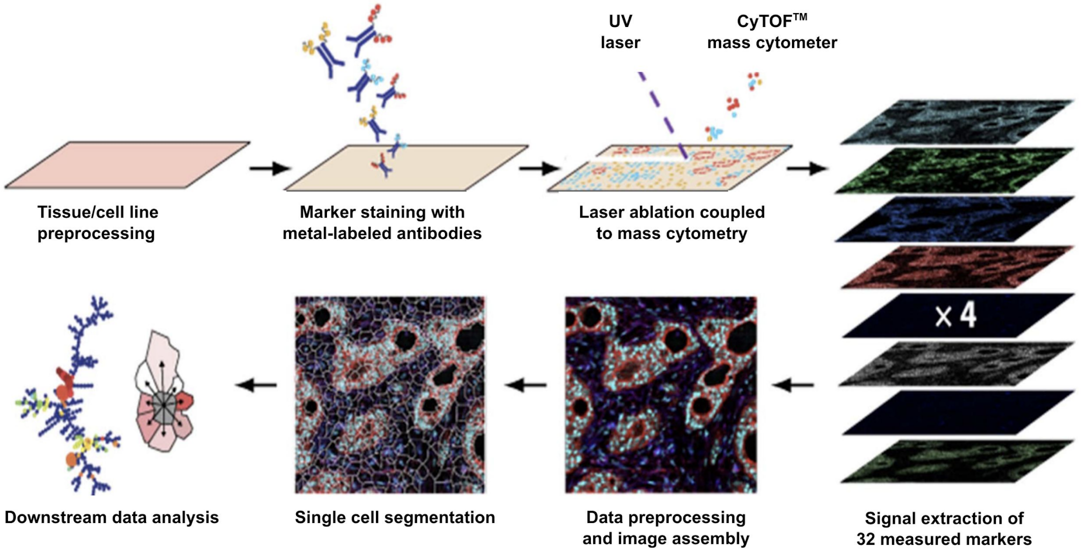


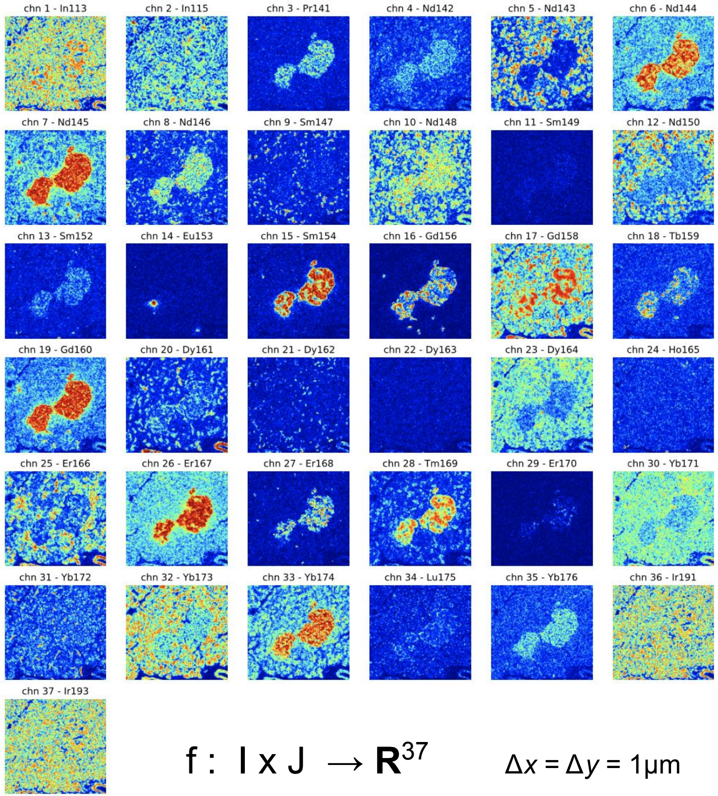


http://www.imagescience.org/meijering/publications/download/biap2006.pdf

Imaging Mass Cytometry (combining molecular biology and imaging)





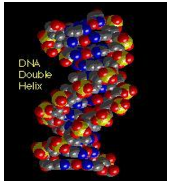


$f: I \times J \rightarrow \mathbf{R}^{37}$ $\Delta x = \Delta y = 1\mu m$

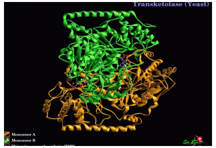
Figure 3: Mosaic of color-coded channel images in the E08 IMC data set.

Biomedical imaging

DNA

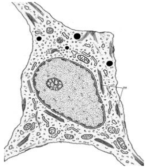


Protein



Organelle

Cell

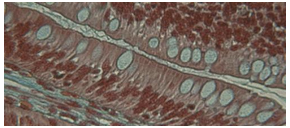


Clinical MRI

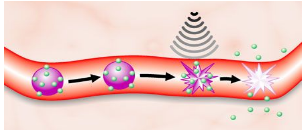
Clinical US

Clinical PET/CT

Tissue



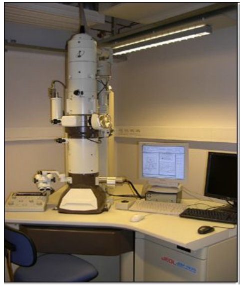
Organ



Organism

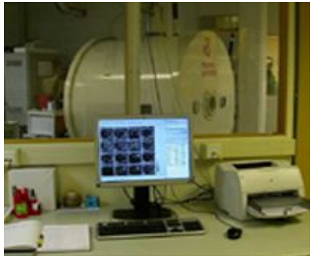
Microscopy

EM



Confocal

Animal MRI

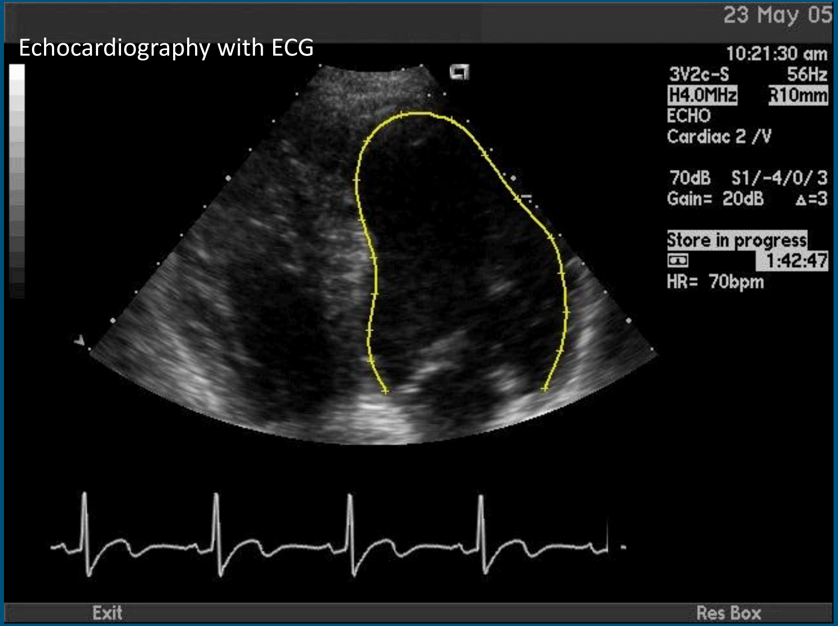


$$\frac{\partial M_x(t)}{\partial t} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_x - \frac{M_x(t)}{T_2}$$
$$\frac{\partial M_y(t)}{\partial t} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_y - \frac{M_y(t)}{T_2}$$
$$\frac{\partial M_z(t)}{\partial t} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t))_z - \frac{M_z(t) - M_0}{T_1}$$

SHORT BREAK

Computational imaging & machine learning

… generic technologies



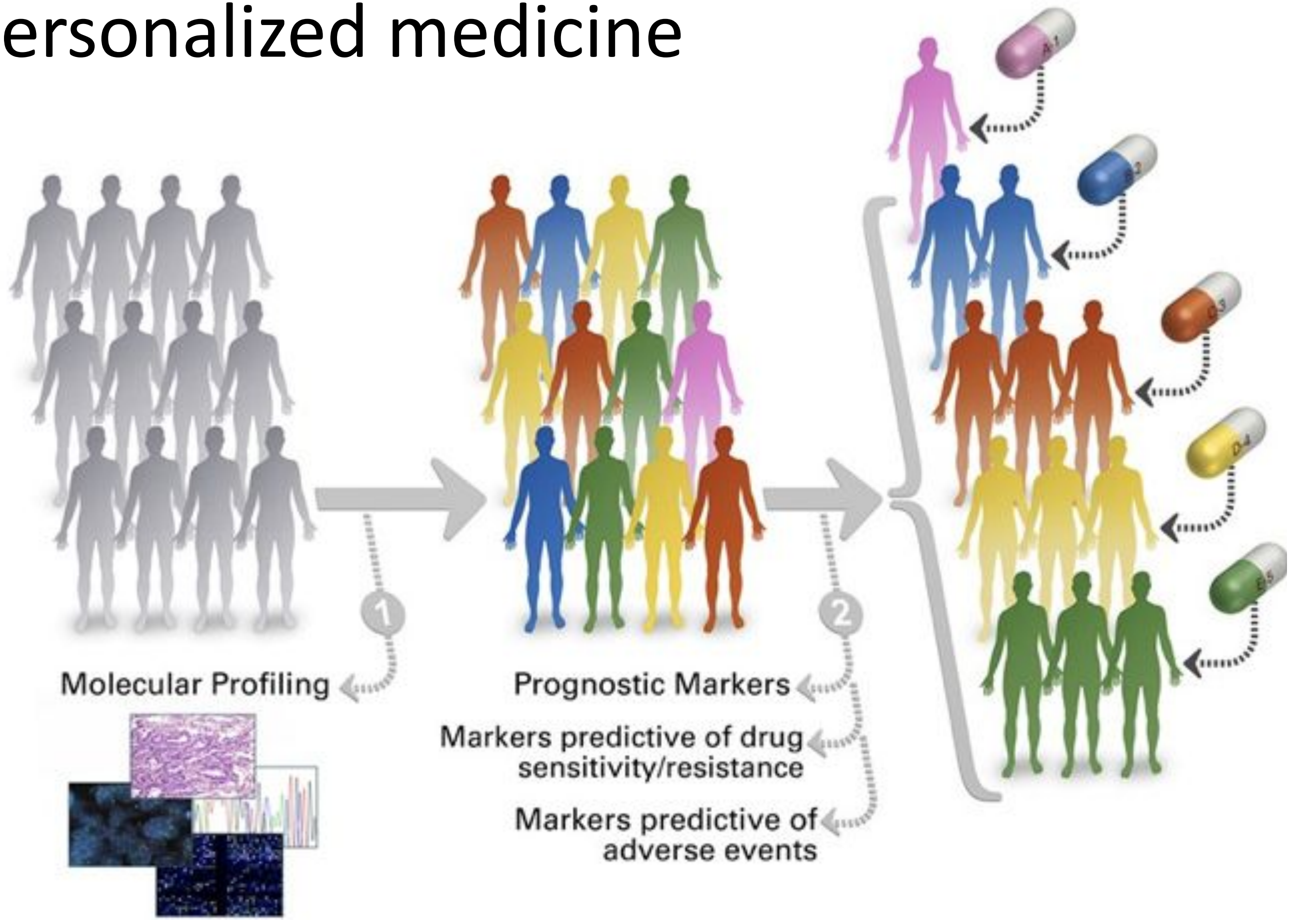
Contour detection + tracking + volume estimation
=> cardiac output [ml/min]

Peng Wang, S. Zhou, M. Szucs, Endocardium tracking by fusing optical flows in straightened images with learning based detections, IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), 2011

Object detection + tracking + number-plate recognition
=> electronic toll collection / vehicle speed [km/h]

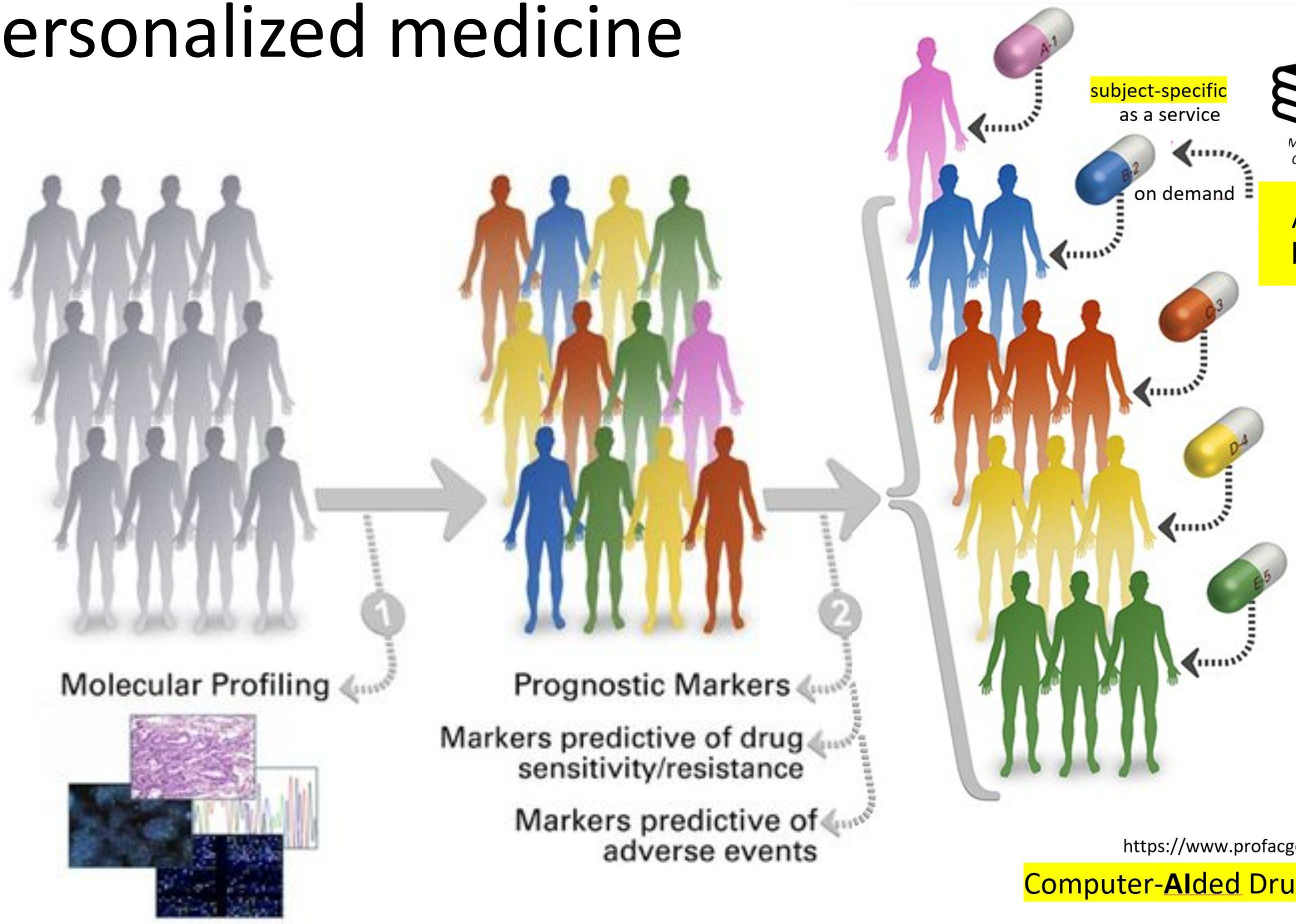
https://developer.ibm.com/code/2018/05/11/using-computer-vision-to-detect-and-track-moving-objects-in-video/

Personalized medicine



https://pct.mdanderson.org

Personalized medicine



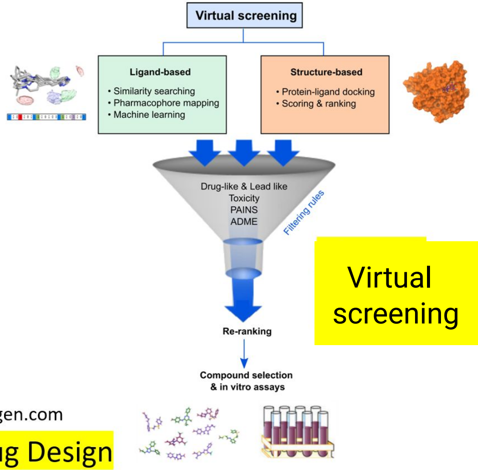
subject-specific as a service

on demand

Artificial intelligence driven biopharmaceutical enterprises

e.g. twoXAR

http://www.twoxar.com



Virtual screening

https://www.profacgen.com

Computer-**AI**ded Drug Design

https://pct.mdanderson.org

The **Cancer Genome Atlas Glioblastoma Multiforme** (TCGA-GBM) [**open**] data collection
https://wiki.cancerimagingarchive.net/display/Public/TCGA-GBM

Team-based project ...

https://github.com/MMIV-ML/ELMED219-2022/tree/main/project

K-means clustering of MRI data set from TCGA-GBM

We will be using a four-channel multispectral image (an axial slice from a multispectral 3D recording is shown below), downloded from the TCGA-GBM data collection - i.e. study TCGA-06-1802. The DICOM images were converted to NIFTI using the dcm2niix software.

```
from IPython.display import Image
Image(filename='./assets/TCGA-GBM-dataset.png', width=900)
```

Data Access Detailed Description Citations & Data Usage Policy Versions

Version 4 (Current): Updated 2020/05/29

Data Type	Download all or Query/Filter
Images (DICOM, 73.5GB) *	Download Search
Tissue Slide Images (web)	Search
Clinical Data (TXT)	Download
Biomedical Data (TXT)	Download
Genomics (web)	Search

Updated clinical data link with latest spreadsheets from GDC. Added new biomedical

Data Access Detailed Description Citations

Detailed Description

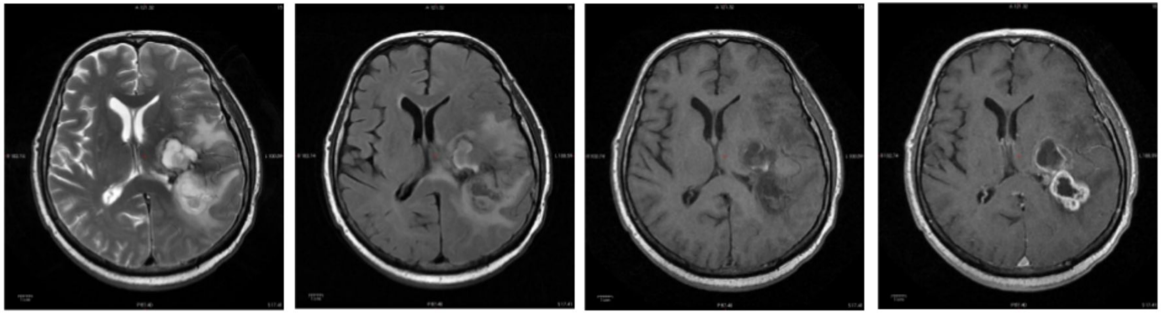
Image Statistics	
Modalities	MR
Number of Participants	262
Number of Studies	575
Number of Series	5,412
Number of Images	481,158
Images Size (GB)	73.5

*) TCIA_TCGA-GBM_09-16-2015.tcia -> TCGA-06-1802_gdc_download_20200630_211154.634082.tar.gz, TCGA-06-1802.zip

The TCGA-06-1802 data set including metadata

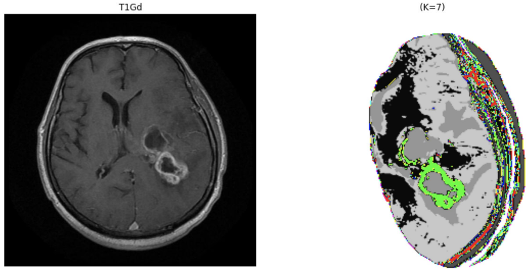
(DICOM images converted to NIFTI format using dcm2niix)

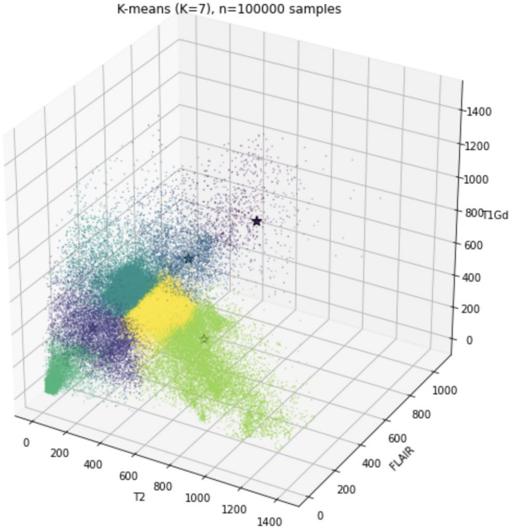
```
TCGA-06-1802_5_AX_T2_FR-FSE.json
TCGA-06-1802_5_AX_T2_FR-FSE.nii.gz
TCGA-06-1802_6_Ax_T2_FLAIR.json
TCGA-06-1802_6_Ax_T2_FLAIR.nii.gz
TCGA-06-1802_7_AX_T1_pre_GD_FLAIR.json
TCGA-06-1802_7_AX_T1_pre_GD_FLAIR.nii.gz
TCGA-06-1802_8_AX_T1_POST_GD_FLAIR.json
TCGA-06-1802_8_AX_T1_POST_GD_FLAIR.nii.gz
TCGA-06-1802_clinical.tsv
TCGA-06-1802_exposure.tsv
TCGA-06-1802_family_history.tsv
```



T2w FLAIR T1w T1wGd

The multispectral MRI slice and K-means clustering

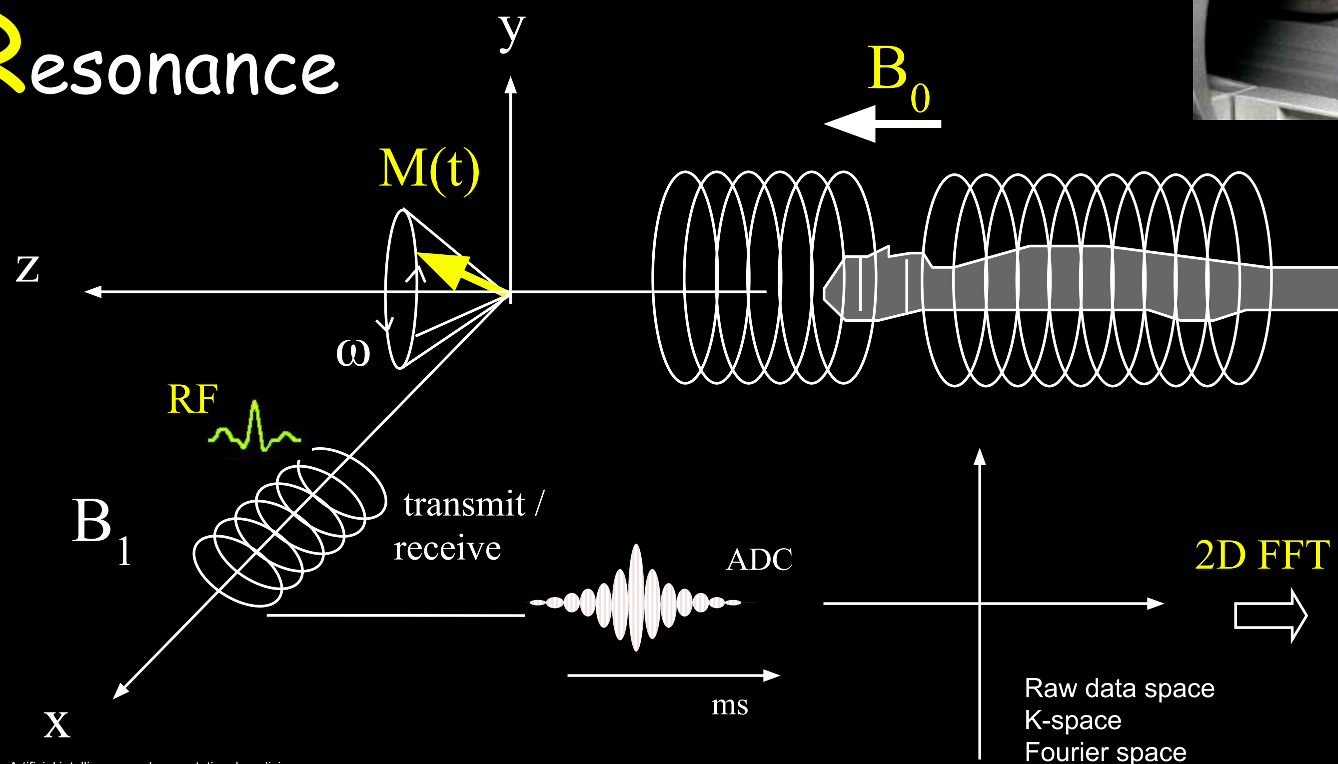




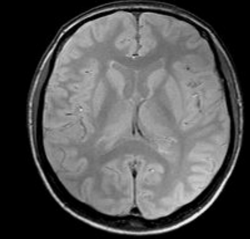
MRI principles ...

Magnetic

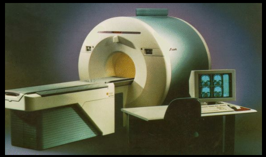
Resonance



Imaging



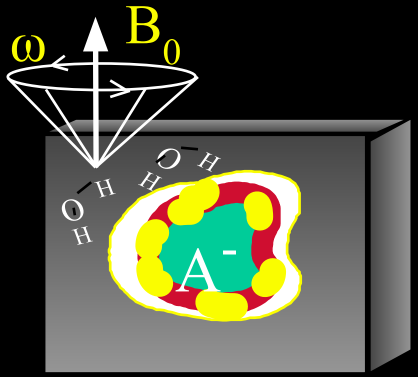
MRI ...



We are imaging:

water protons
in cellular environment

Contrast-mechanisms



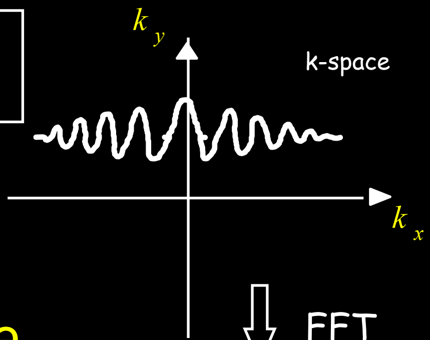
by:

experimental manipulation
of proton spin populations

Pulse sequences

using:

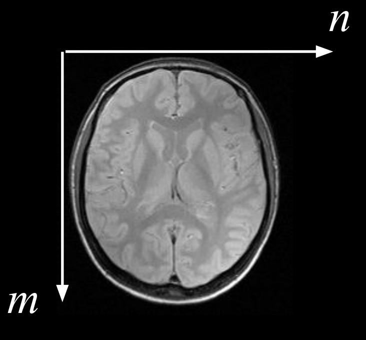
Fourier transform
(2D FFT)



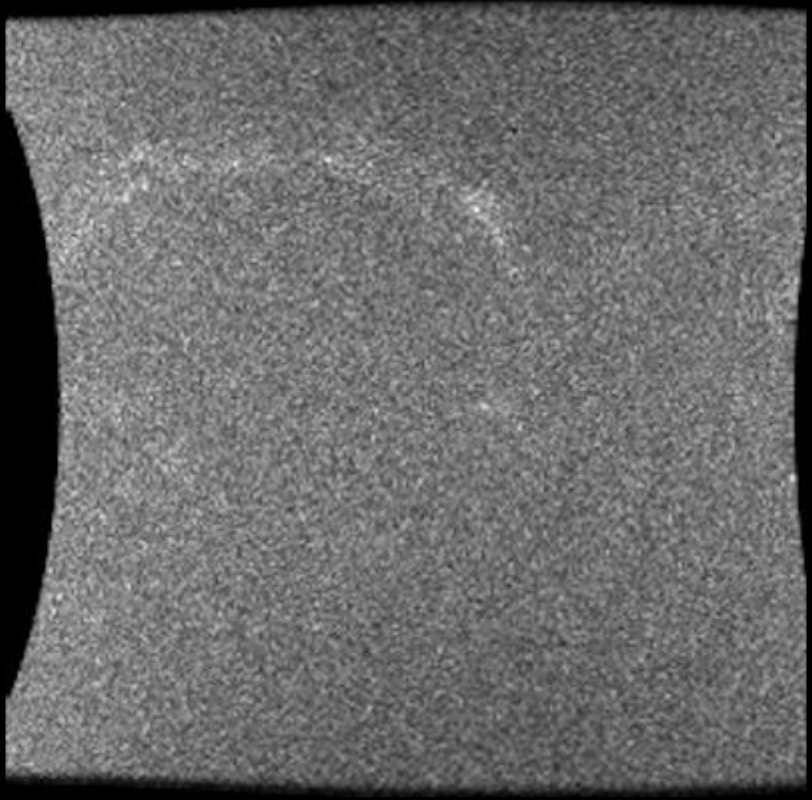
FFT

for:

Image reconstruction

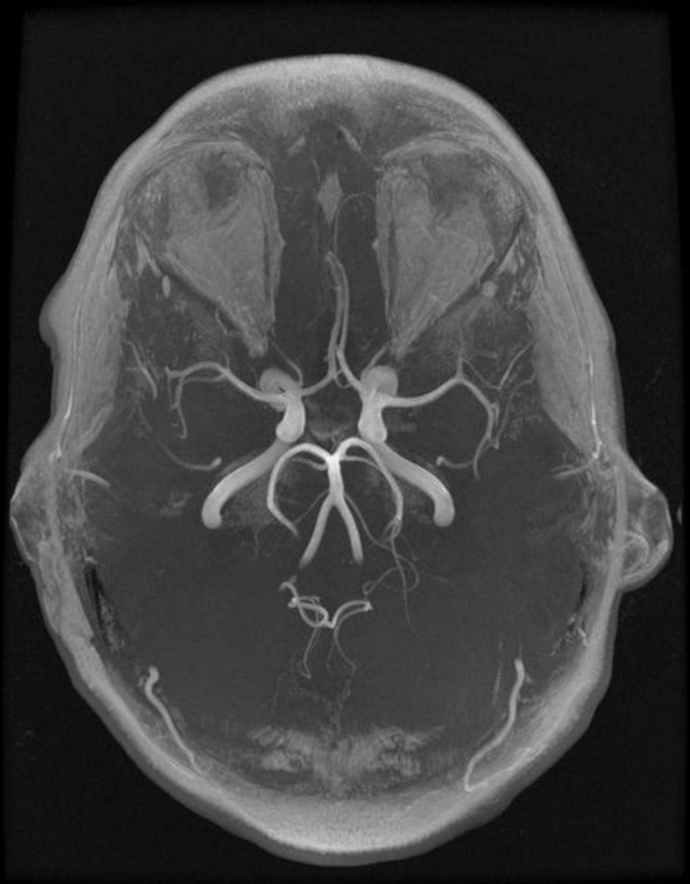


3D T1-weighted MRI



3DT1_spgr_out_of_phase_2_2_mov.gif

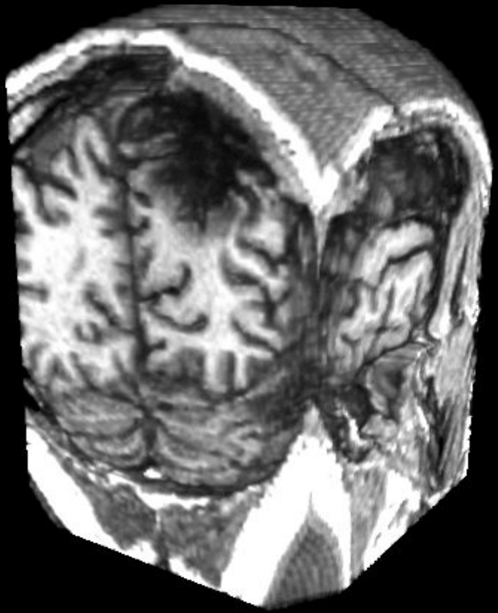
Time-Of-Flight MRA



3D_TOF_1024_3_3_mov.gif

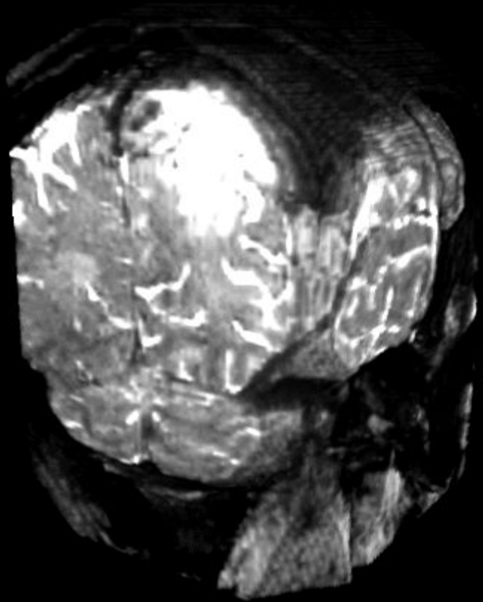
Multispectral 3D MRI

FLASH



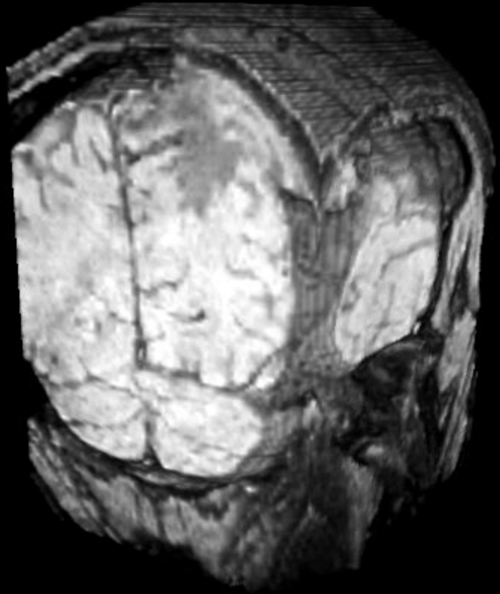
T1-w

DESS



T2-w

FISP



ρ-w

OpenViz

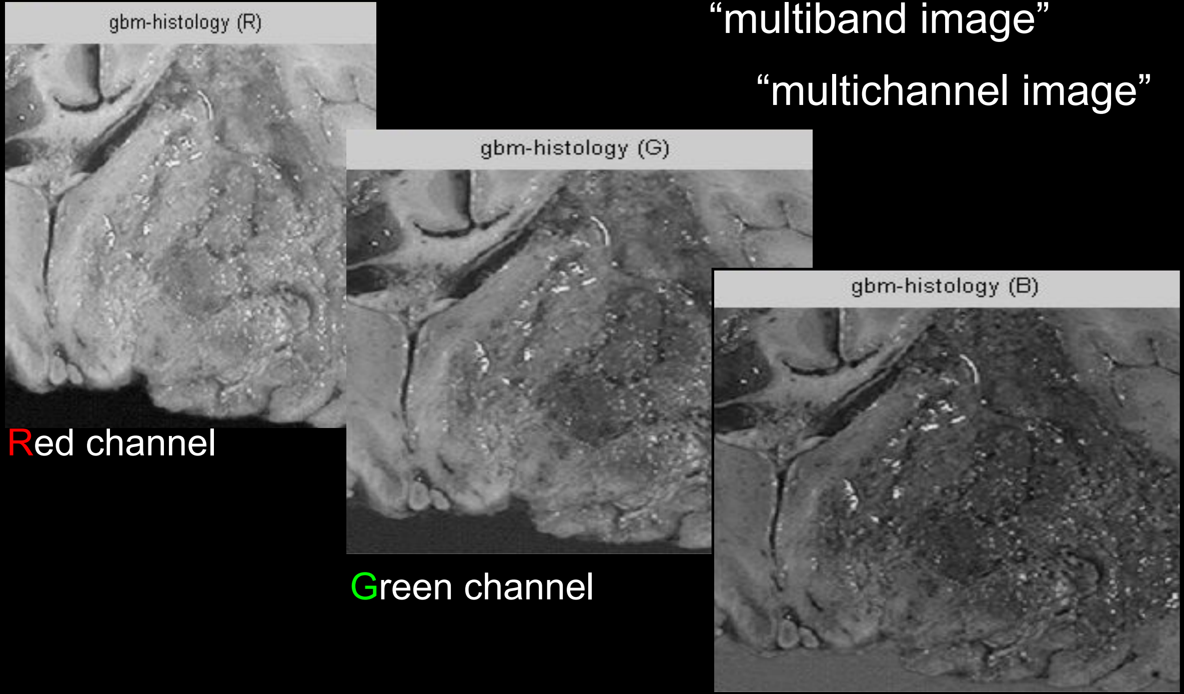
Multispectral image

gbm_histology.bmp

"multispectral image"

"multiband image"

"multichannel image"



Red channel

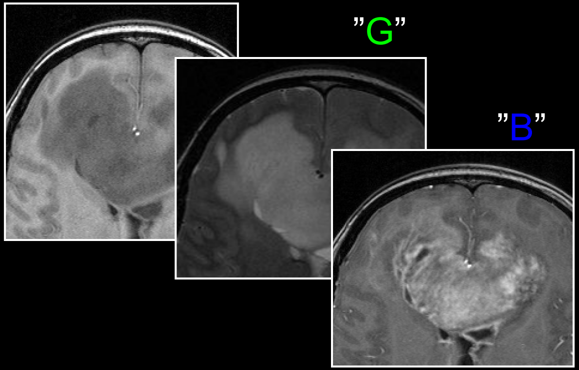
Green channel

Blue channel

"R"

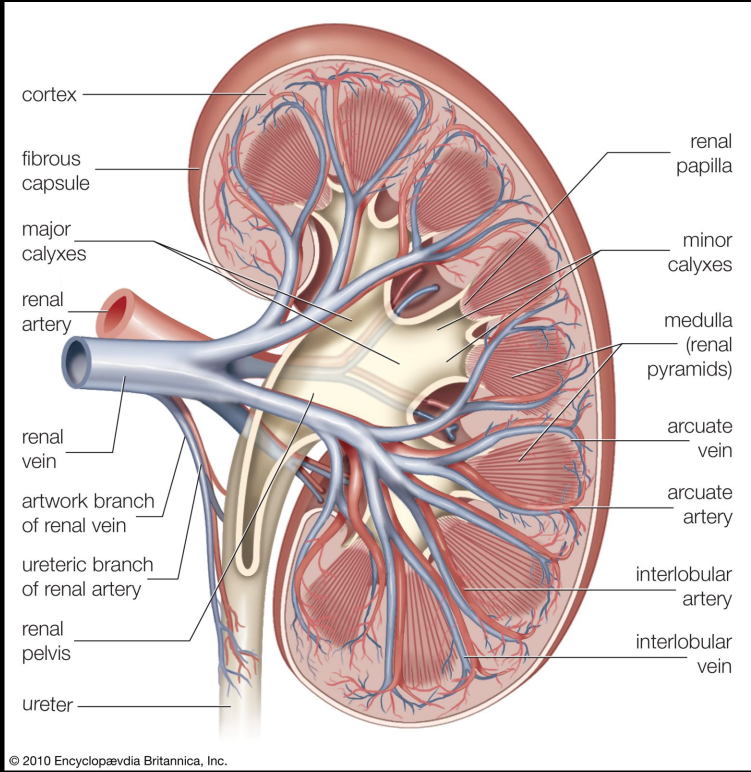
"G"

"B"

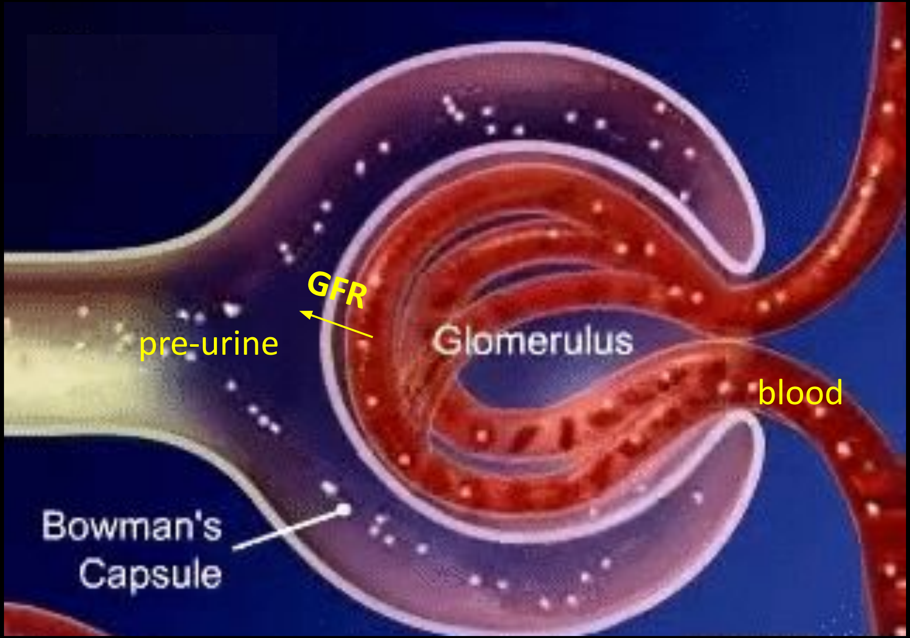


different pulse sequences in MRI

A little (computational) kidney physiology



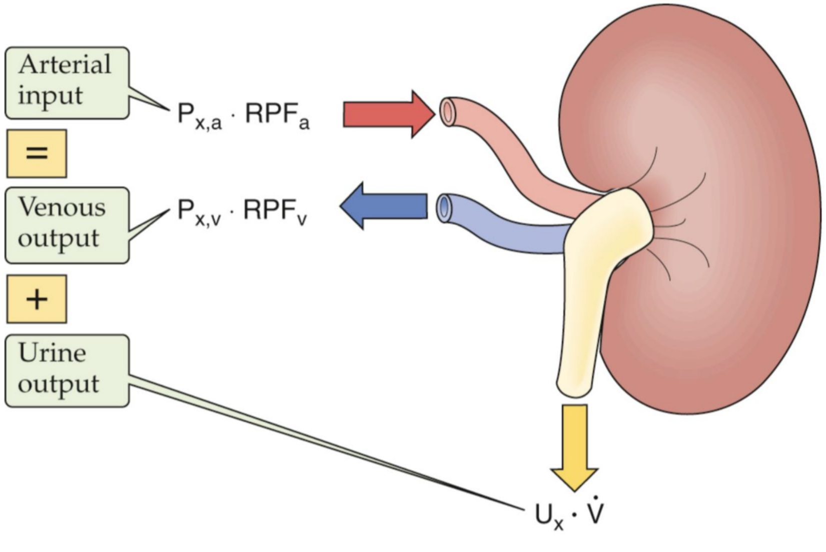
Glomerular filtration



https://www.youtube.com/watch?v=oCQ-5iwTQvM

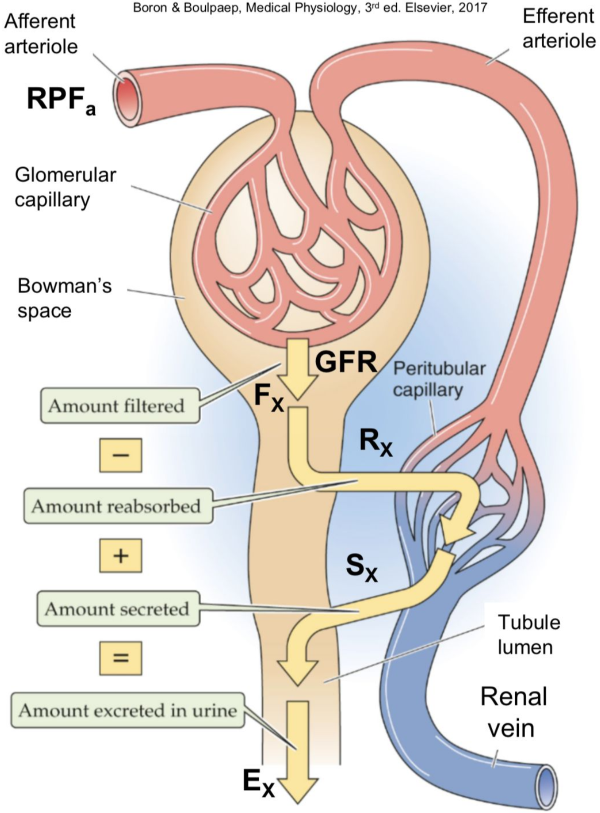
A little (computational) kidney physiology

Mass balance

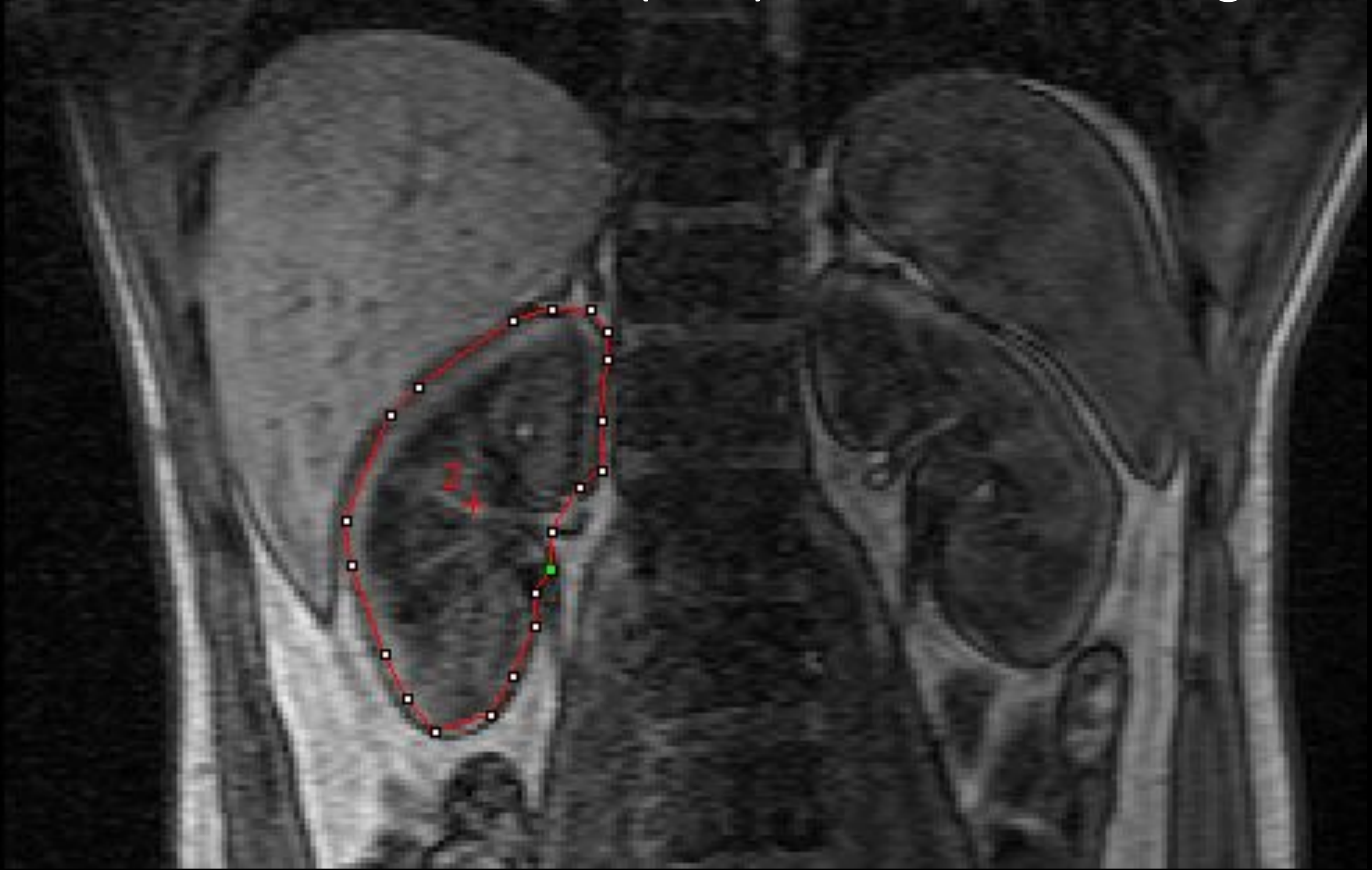


$$\overbrace{\underbrace{P_{X,a}}_{\frac{\text{mmole}}{\text{mL}}} \cdot \underbrace{RPF_a}_{\frac{\text{mL}}{\text{min}}}}^{\text{Arterial input of X}} = \overbrace{\left(\underbrace{P_{X,v}}_{\frac{\text{mmole}}{\text{mL}}} \cdot \underbrace{RPF_v}_{\frac{\text{mL}}{\text{min}}}\right)}^{\text{Venous output of X}} + \overbrace{\left(\underbrace{U_X}_{\frac{\text{mmole}}{\text{mL}}} \cdot \underbrace{\dot{V}}_{\frac{\text{mL}}{\text{min}}}\right)}^{\text{Urine output of X}}$$

Boron & Boulpaep, Medical Physiology, 3rd ed. Elsevier, 2017

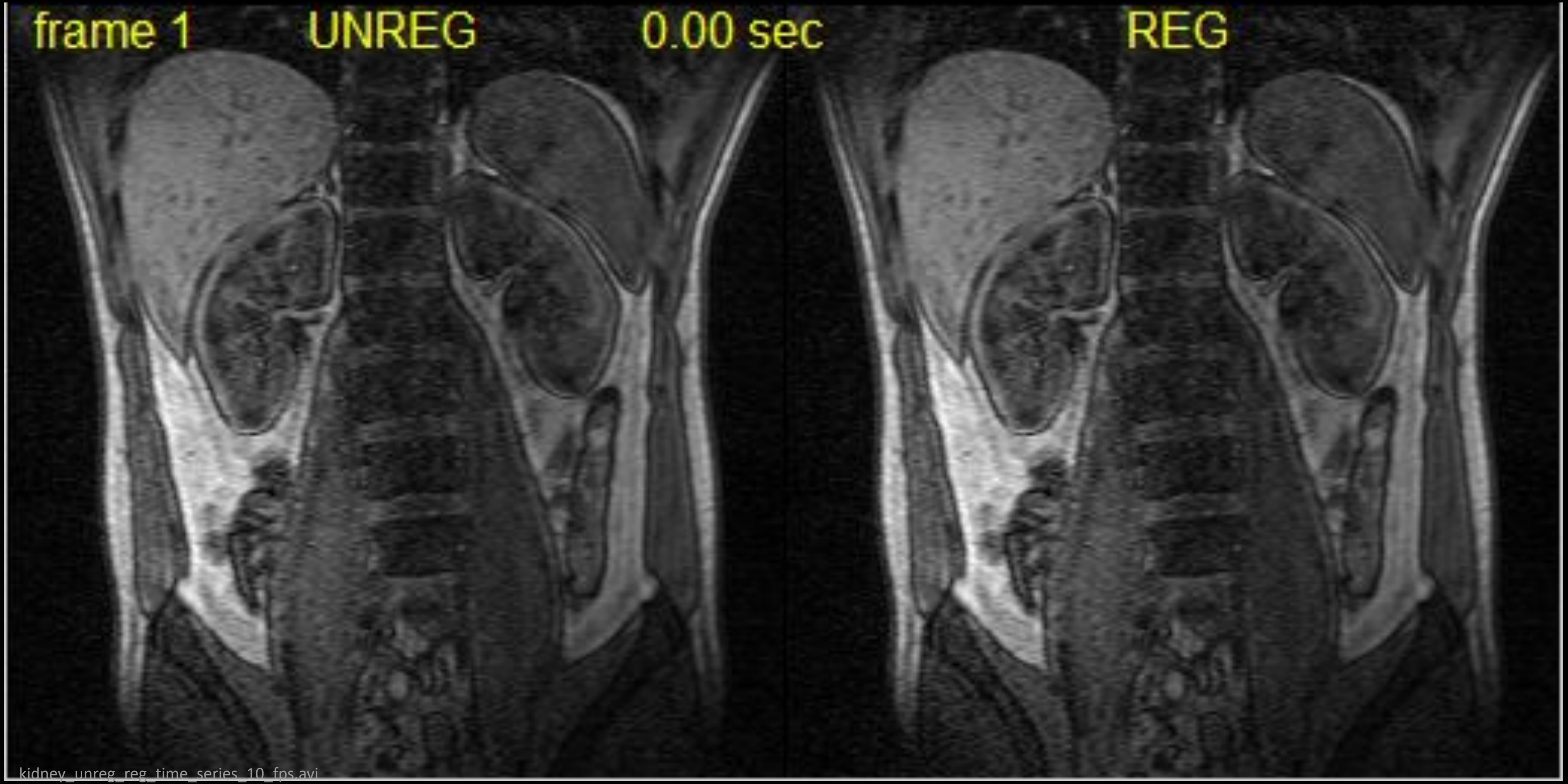


Dynamic contrast enhanced (DCE) MRI of the moving kidney



kidney_slice10_frame10_roi.avi

DCE-MRI of the moving kidney with **motion correction** (image registration)

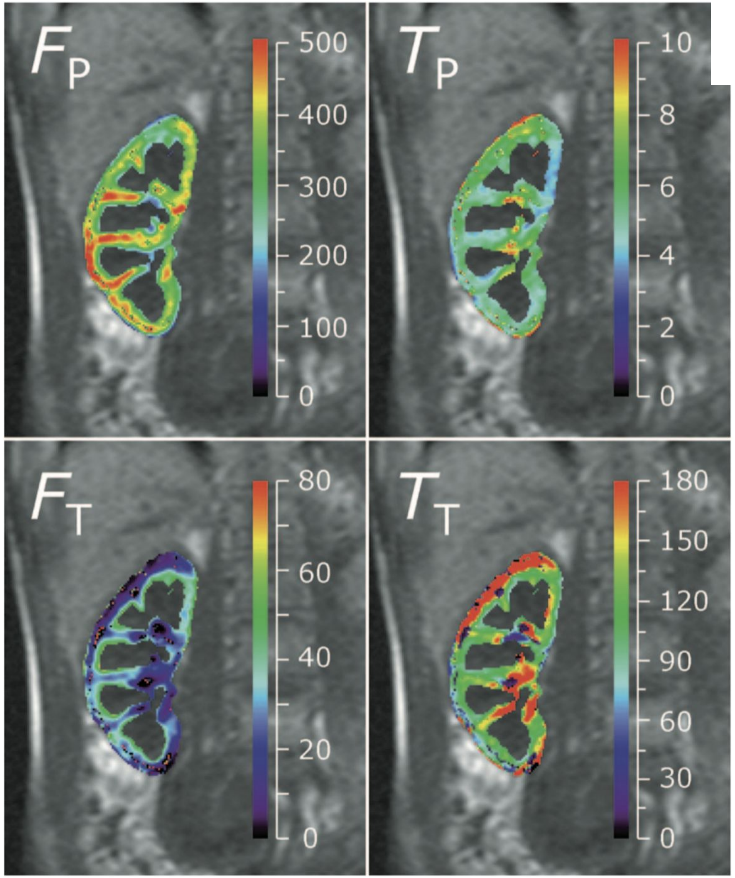


Voxel-wise estimation of the four model parameters

$\theta = (F_P, T_P, F_T, T_T)$

F: flow (perfusion)
MTT: mean transit time

P: plasma compartment
T: tubular compartment

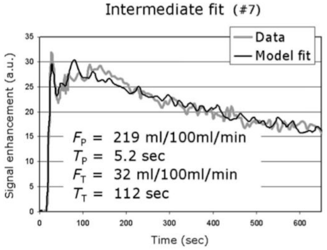


F_P (mL/100 mL/Min)

T_P = MTT_P (sec)

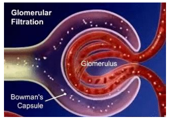
F_T (mL/100 mL/ Min)

T_T = MTT_T (sec)



for the data with the intermediate fit accuracy

F_P = GFR (glomerular filtration rate)

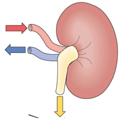


θ: image-based biomarkers

The parametric maps (colored) are superposed on a precontrast image (gray) for anatomic reference.

Sourbron et al. MRI-Measurement of Perfusion and Glomerular Filtration in the Human Kidney With a Separable Compartment Model. Investigative Radiology 2008;43:40-48.

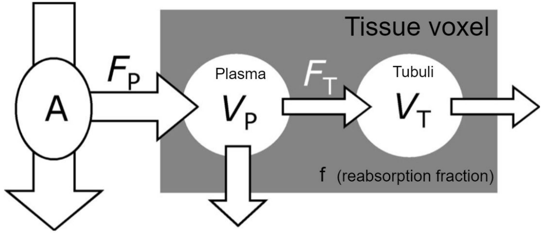
Renal perfusion and filtration:



Kidney GFR

(Sourbron et al., 2008)

Feeding artery



AV_pV_T compartments

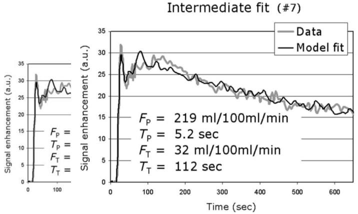
The four independent model parameters

F_P (mL/100 mL/Min)

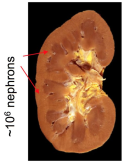
T_P = MTT_P (sec)

F_T (mL/100 mL/ Min)

T_T = MTT_T (sec)



for the data with the intermediate fit accuracy

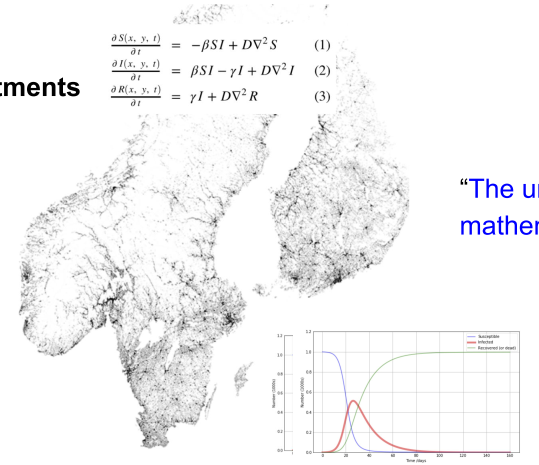


Outbreak science:

COVID-19

SIR compartments

$$\frac{\partial S(x,y,t)}{\partial t} = -\beta SI + D\nabla^2 S \quad (1)$$
$$\frac{\partial I(x,y,t)}{\partial t} = \beta SI - \gamma I + D\nabla^2 I \quad (2)$$
$$\frac{\partial R(x,y,t)}{\partial t} = \gamma I + D\nabla^2 R \quad (3)$$



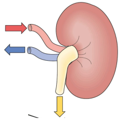
"The unreasonable effectiveness of mathematics in the natural sciences."

(Wigner*, 1960)

*) … *The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.*

We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure, even though perhaps also to our bafflement, to wide branches of learning.

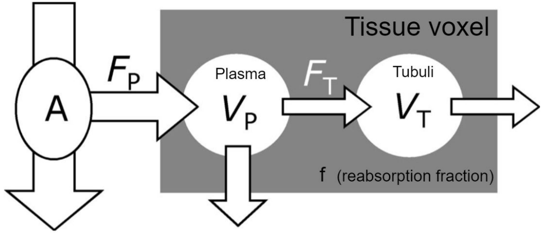
Renal perfusion and filtration:



Kidney GFR

(Sourbron et al., 2008)

Feeding artery



AV_pV_T compartments

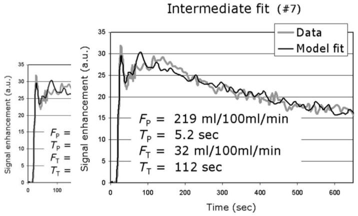
The four independent model parameters

F_P (mL/100 mL/Min)

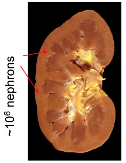
T_P = MTT_P (sec)

F_T (mL/100 mL/ Min)

T_T = MTT_T (sec)

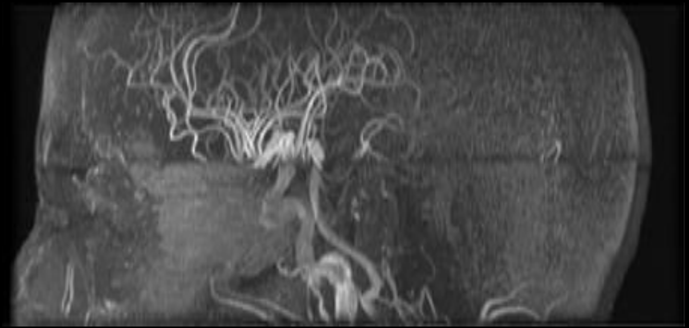


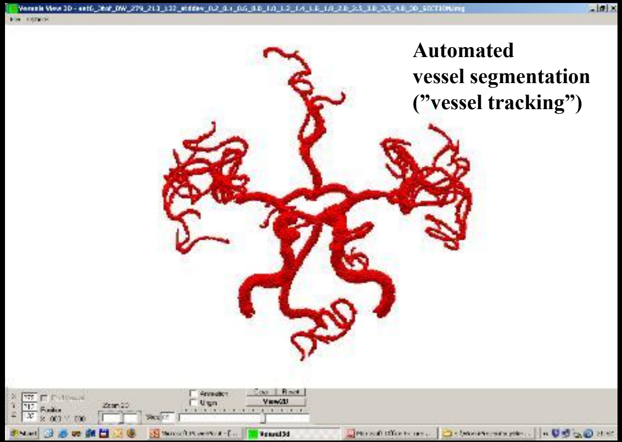
for the data with the intermediate fit accuracy



Cerebral circulation: Vessels – Flow – Perfusion - Permeability

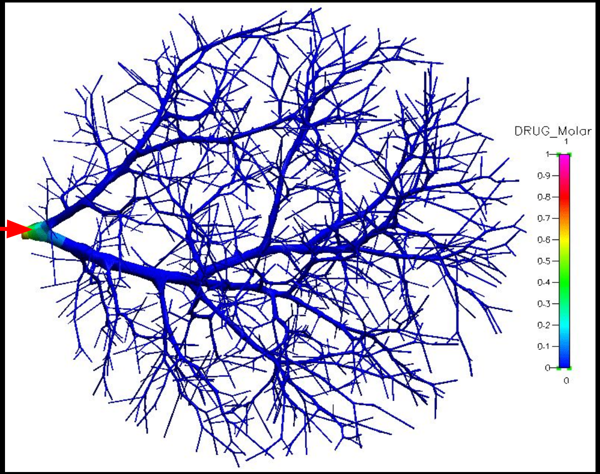
MRA



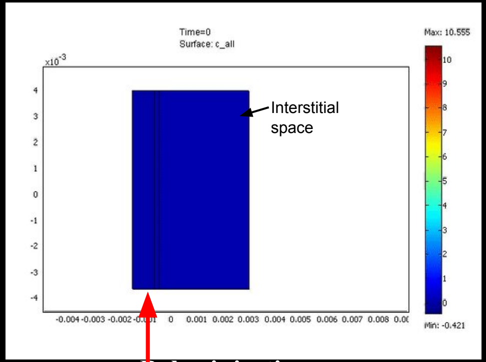


Modelling and simulation of ”drug delivery” in a vessel tree (Kocinski, TUL)

Bolus injection



Modelling of capillary wall leakage, assuming a closed extravascular compartment (interstitial space) and a combined diffusive and convective transcapillary transport



Bolus injection

COMSOL Multiphysics

Biomedical & technological revolutions → "Convergence"

1. Omics, multiscale & multimodal imaging, EHR, ... ***data***: x

"big data" & "rich data"

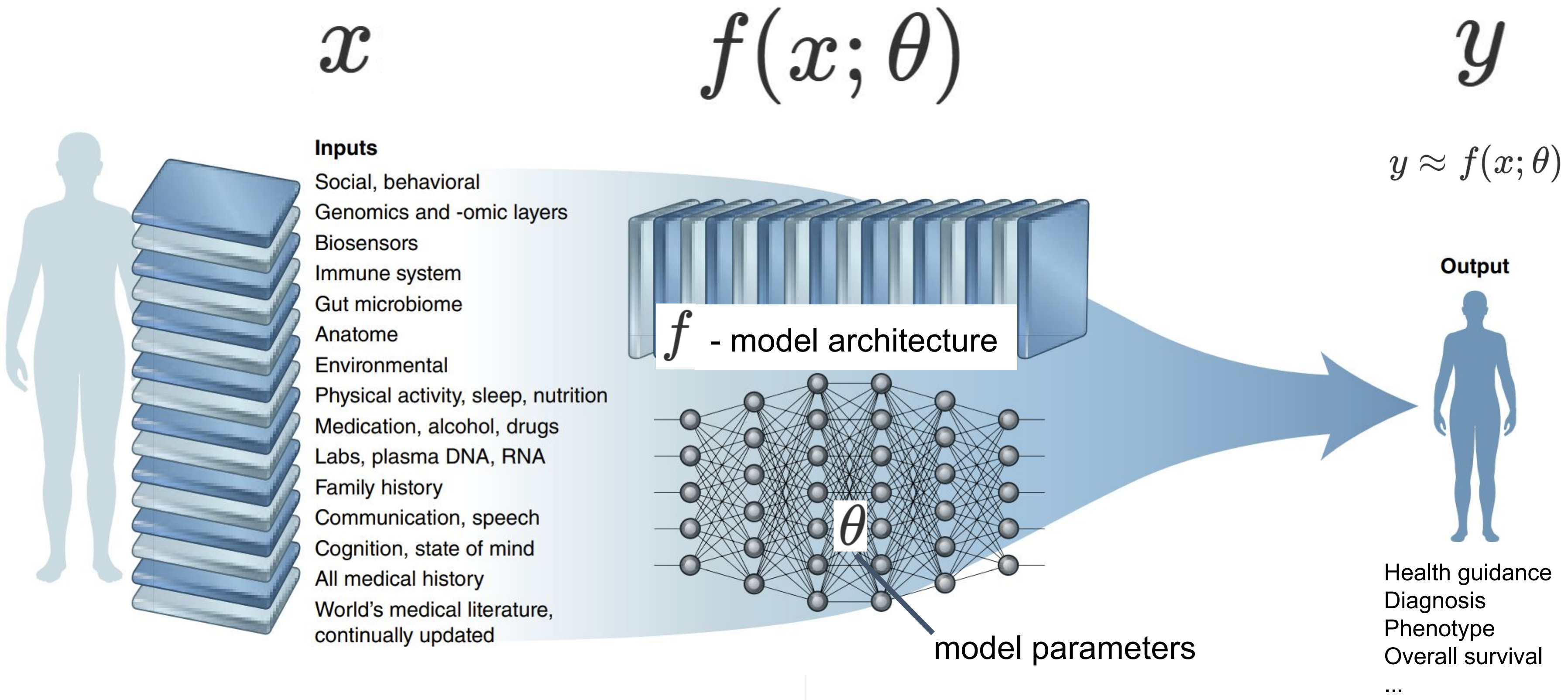
2. Computational sciences ***models***: f

Artificial intelligence, ML/DL ***learning***: θ parameters

3. Diagnosis, phenotype, overall survival ***outcome***: y

prediction: $y \approx f(x;\theta)$

" personalized medicine "

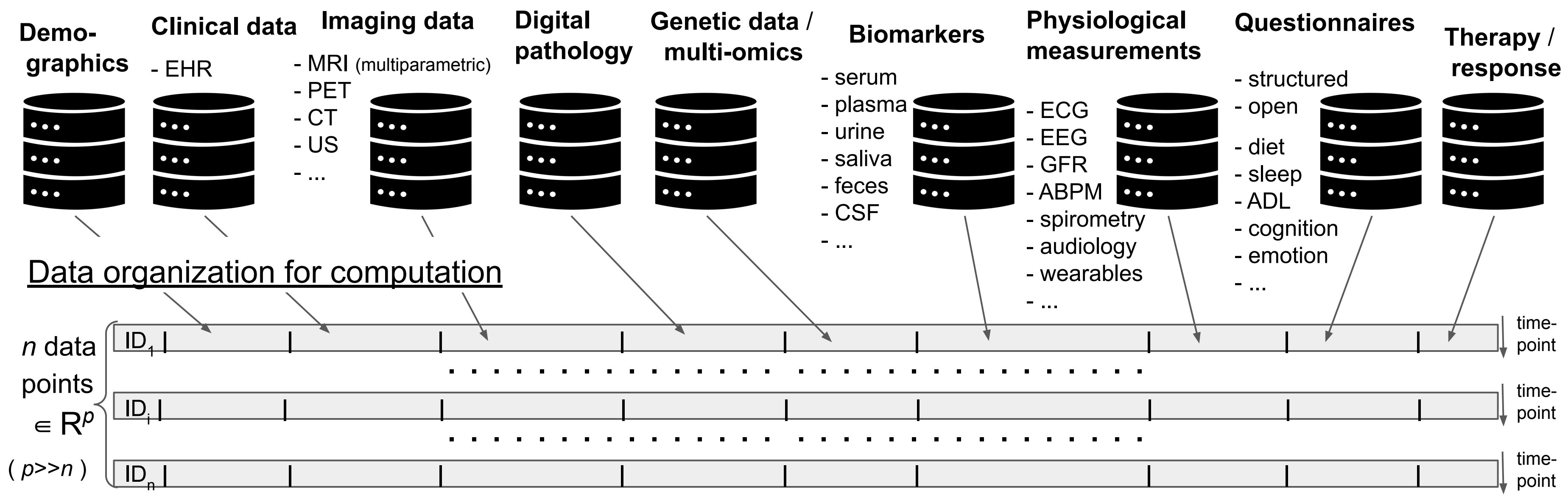


The virtual medical coach model with multi-modal data inputs and algorithms to provide individualised guidance

Modified from: E. J. Topol, 'High-performance medicine: the convergence of human and artificial intelligence', *Nature Medicine* 2019;25:44−56. https://www.nature.com/articles/s41591-018-0300-7

A patient-oriented, spatio-temporal digital biobank for clinical data science and medical AI

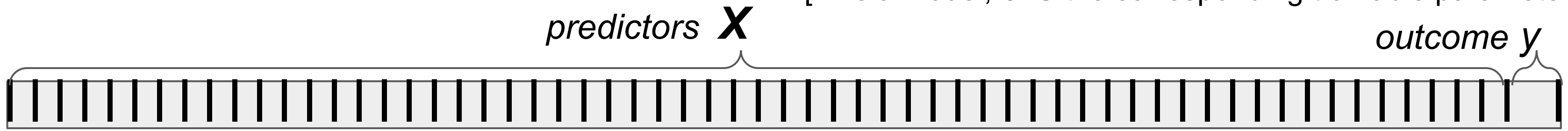
MACHINE LEARNING & ARTIFICIAL INTELLIGENCE | ICT INFRASTRUCTURE | (BIO)MEDICAL TECHNOLOGIES



<u>Explorative data analysis</u> (visualization, normalization, correlations, feature importance, dimensionality reduction, TDA, …)

<u>Machine learning</u> Splitting the data: (X_{train}, y_{train}) & (X_{test}, y_{test}) ML & DL models: $y \approx f(\boldsymbol{X}, \theta)$

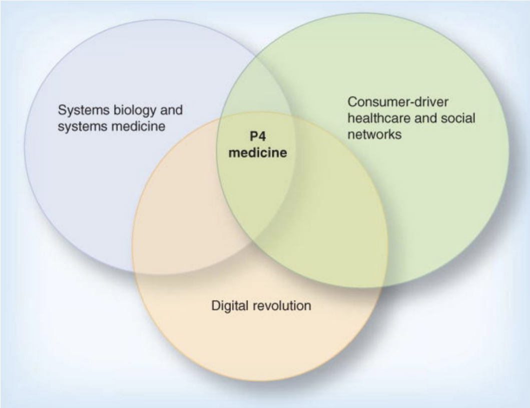
[f is a model; θ is the corresponding trainable parameters]



http://www.esprevmed.org/p4-medicine

“P4 medicine”

incl. systems biology and systems medicine



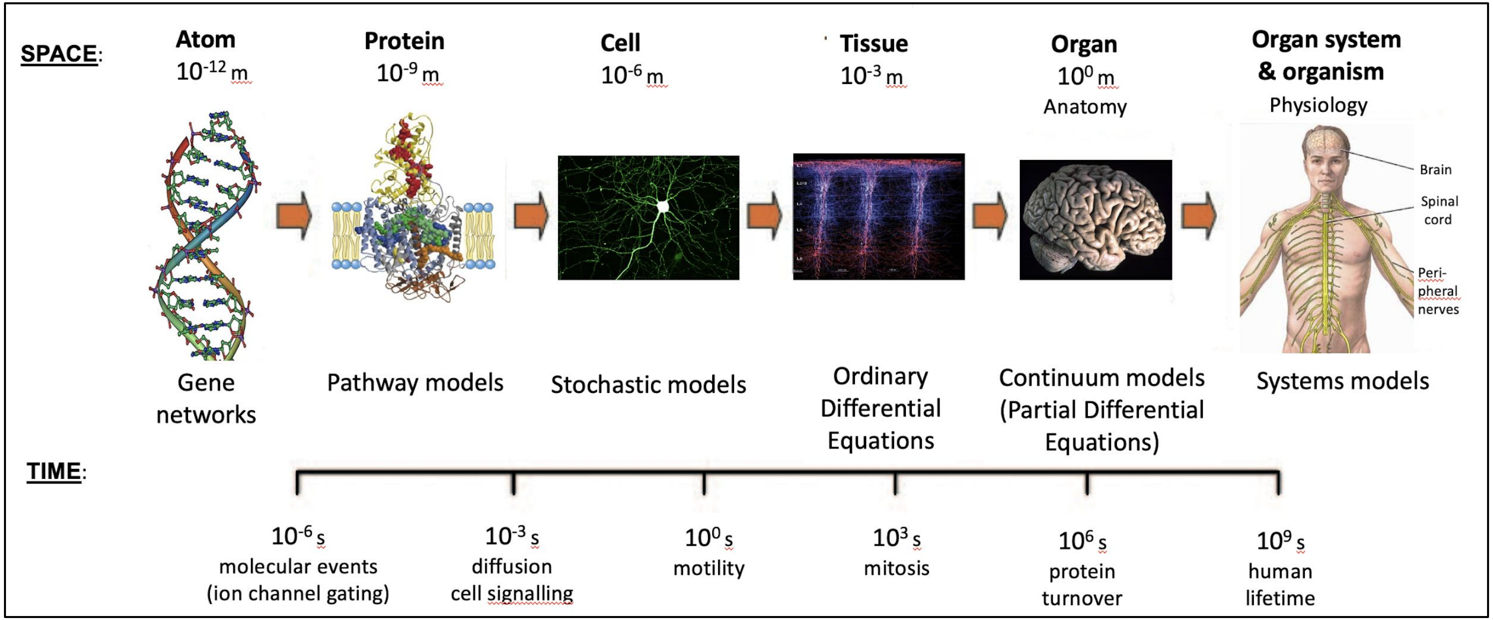
P4: Predictive Preventive Personalised Participatory

Based on

- The increasing ability of systems biology and systems medicine to decipher the biological complexity of disease.
- The digital revolution´s radically enhanced capabilities for collecting, integrating, storing, analyzing and communicating data and information (conventional medical histories, clinical tests and the results of the tools of systems medicine)
- Consumer access to information and consequent interest in managing their own health.
  Consumers are driving the transformation of healthcare by these megatrends.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC420440

The future of computational medicine, modeling and machine learning…



Challenges:

Δ mindset

Δ skillset

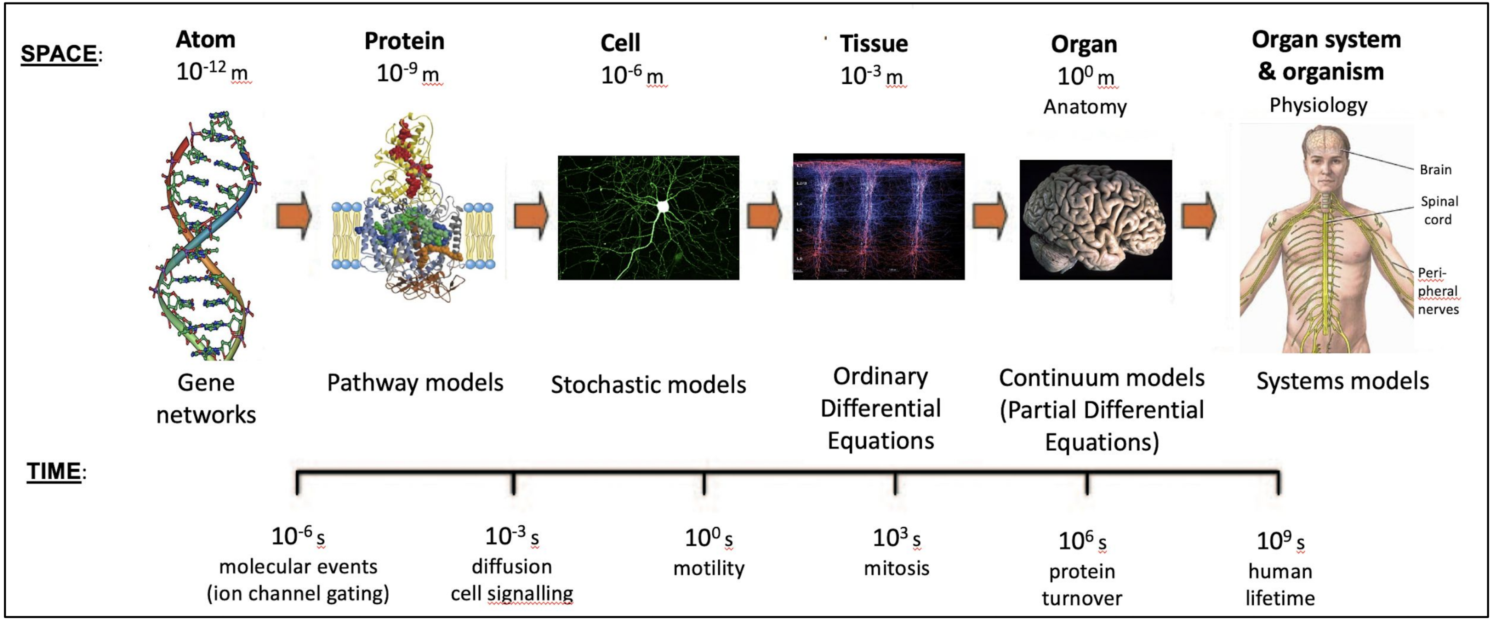
Δ toolset

- open science

- reproducible research

- education
- training

The future of computational medicine, modeling and machine learning…



Challenges:

Δ mindset

Δ skillset

Δ toolset

- open science

- reproducible research

- education
- training

- Why Python ?

- Why Jupyter notebooks ?

- Why GitHub ?

Challenges:

Δ mindset

Δ skillset

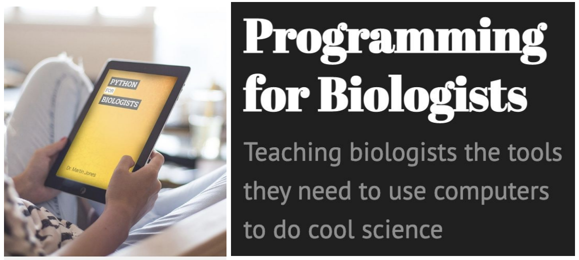
Δ toolset

- open science

- reproducible research

- education
- training

- # Why Python ?



http://www.programmingforbiologists.org/about/why-python

1) Easy to Learn and Use
2) Mature and Supportive Python Community
3) Support from Renowned Corporate Sponsors
4) Hundreds of Python Libraries and Frameworks
5) Versatility, Efficiency, Reliability, and Speed
6) Big data, Machine Learning and Cloud Computing
7) First-choice Language
8) The Flexibility of Python Language
9) Use of python in academics
10) Automation

https://www.upgrad.com/blog/reasons-why-python-popular-with-developers

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6920002

Journal List › J Med Libr Assoc › v.108(1); 2020 Jan › PMC6920002

JMLA Journal of the Medical Library Association
See also Bulletin of the Medical Library Association, v. 1–89
MLANET SUBSCRIBE SUBMISSIONS

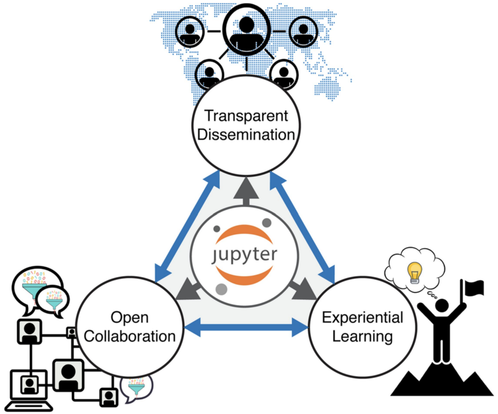
J Med Libr Assoc. 2020 Jan; 108(1): 29–35. PMCID: PMC6920002
Published online 2020 Jan 1. doi: 10.5195/jmla.2020.819 PMID: 31897049

J Med

Why do biomedical researchers learn to program? An exploratory investigation

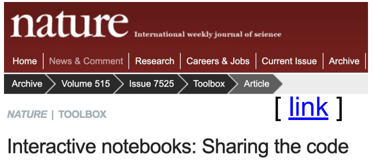
Ariel Deardorff

We use the Python language because it now pervades virtually every domain of the biosciences, from sequence-based bioinformatics and molecular evolution to phylogenomics, systems biology, structural biology, and beyond. [link]



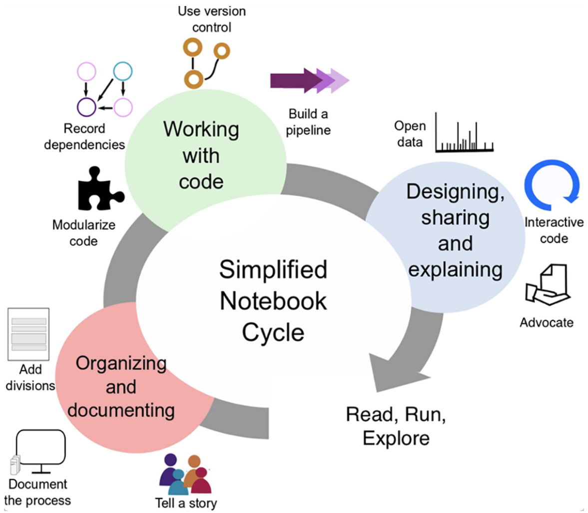
… their interactive and easily deployable framework can drive experiential learning opportunities for computational novices to develop their own skills and better understand metabolomics data analysis [link]

- Why [Jupyter](link) notebooks ?



[link]

Jupyter notebooks provide an environment where you can freely combine human-readable narrative with computer-readable code.



Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks [link]

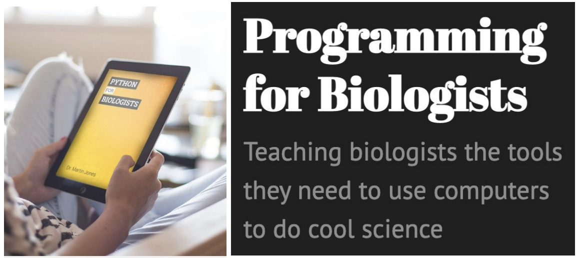
- Why GitHub ?

Github is like facebook for programmers. Everyone's on there. You can look at what they're working on and easily peruse their code and make suggestions or changes.

It's really open source. "Open source" is not so open if you can't easily study it. With github, all of the code is easily inspected, as is its entire history.

Github lowers the barriers to collaboration. [link]

- # Why [Python](#) ?



http://www.programmingforbiologists.org/about/why-python

1) Easy to Learn and Use
2) Mature and Supportive Python Community
3) Support from Renowned Corporate Sponsors
4) Hundreds of Python Libraries and Frameworks
5) Versatility, Efficiency, Reliability, and Speed
6) Big data, Machine Learning and Cloud Computing
7) First-choice Language
8) The Flexibility of Python Language
9) Use of python in academics
10) Automation

https://www.upgrad.com/blog/reasons-why-python-popular-with-developers

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6920002

Journal List › J Med Libr Assoc › v.108(1); 2020 Jan › PMC6920002

JMLA Journal of the Medical Library Association — See also Bulletin of the Medical Library Association, v. 1–89 — MLANET | SUBSCRIBE | SUBMISSIONS

J Med Libr Assoc. 2020 Jan; 108(1): 29–35. PMCID: PMC6920002
Published online 2020 Jan 1. doi: 10.5195/jmla.2020.819 PMID: 31897049

Why do biomedical researchers learn to program? An exploratory investigation

Ariel Deardorff

We use the Python language because it now pervades virtually every domain of the biosciences, from sequence-based bioinformatics and molecular evolution to phylogenomics, systems biology, structural biology, and beyond. [link]

- # Why [Jupyter](#) notebooks ?

nature International weekly journal of science

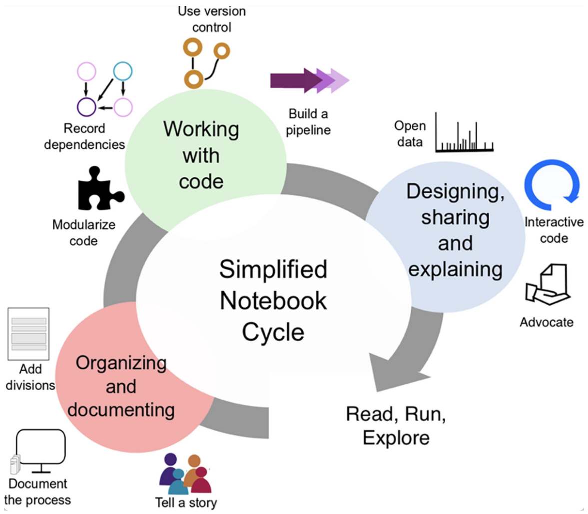
Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive

Archive › Volume 515 › Issue 7525 › Toolbox › Article

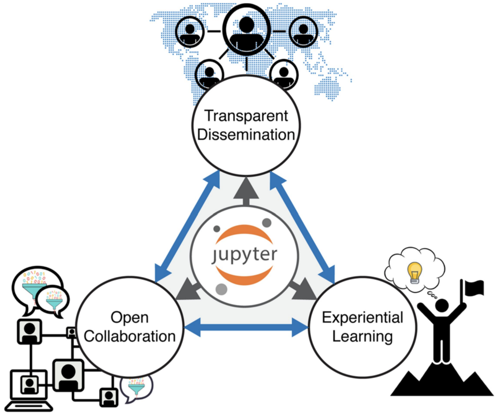
NATURE | TOOLBOX [link]

Interactive notebooks: Sharing the code

Jupyter notebooks provide an environment where you can freely combine human-readable narrative with computer-readable code.



Ten simple rules for writing and sharing computational analyses in Jupyter Notebooks [link]



… their interactive and easily deployable framework can drive experiential learning opportunities for computational novices to develop their own skills and better understand metabolomics data analysis [link]

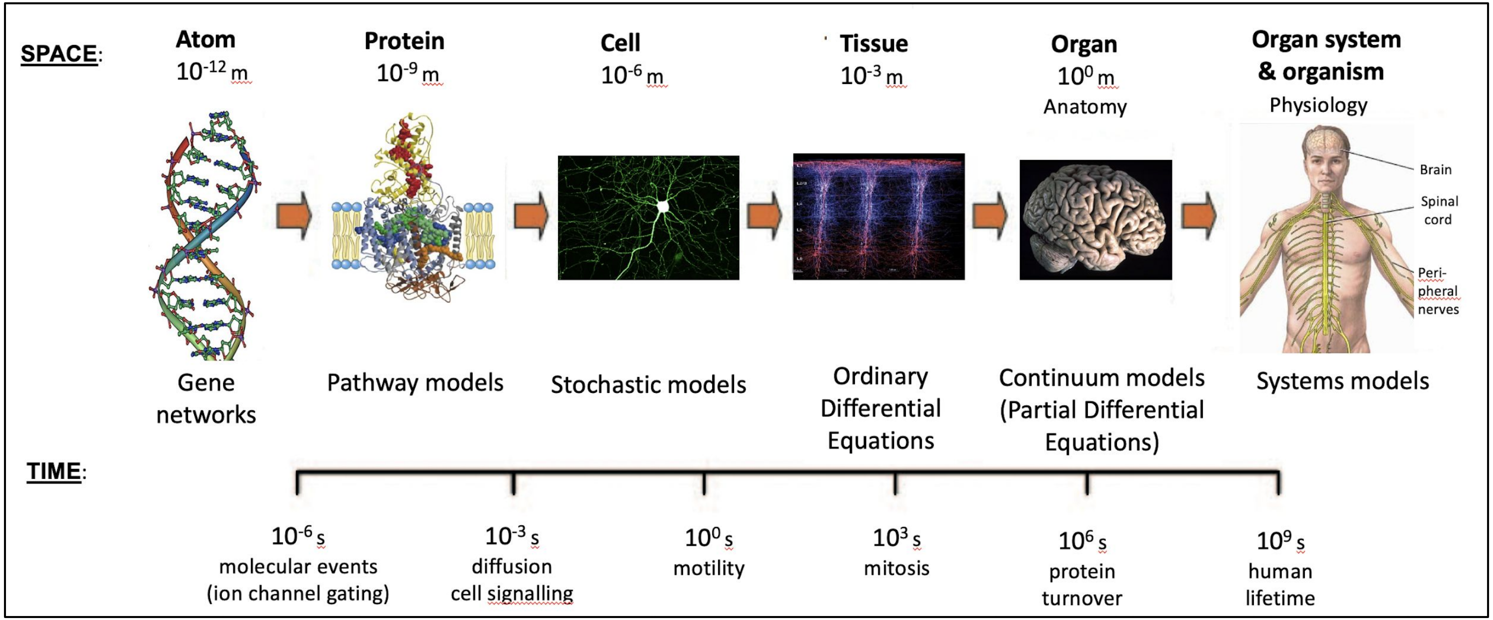
- # Why [GitHub](#) ?

Github is like facebook for programmers. Everyone's on there. You can look at what they're working on and easily peruse their code and make suggestions or changes.

It's really open source. "Open source" is not so open if you can't easily study it. With github, all of the code is easily inspected, as is its entire history.

Github lowers the barriers to collaboration. [link]

The future of computational medicine, modeling and machine learning…



Challenges:

Δ mindset

Δ skillset

Δ toolset

- **open science**
- **reproducible research**
- **education**
- **training**

The future of computational medicine, modeling and machine learning…

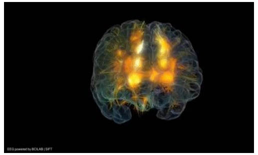
SPA

TIM

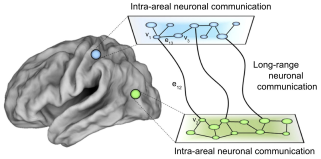
- **Interdisciplinarity**

  (MED, TECH, ENG)

https://www.abdn.ac.uk/strategy-development/key-themes/interdisciplinary-104.php



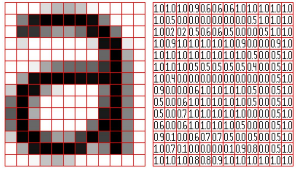
$5 \cdot 10^9$ neurons firing

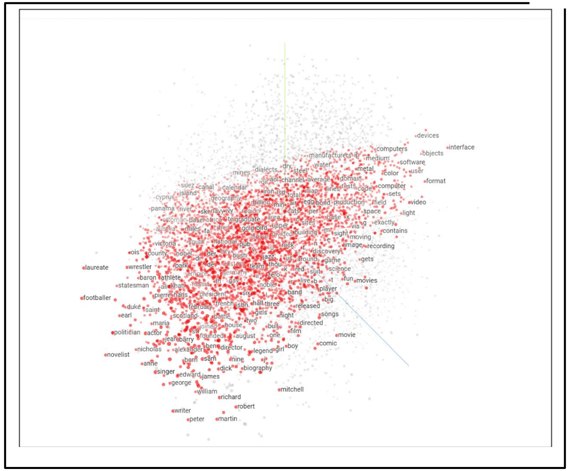


- **Representations**

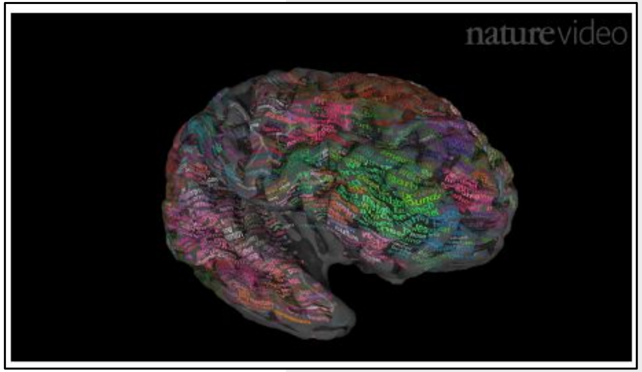
  (images, text, knowledge, …)

$y \approx f(X; \theta)$





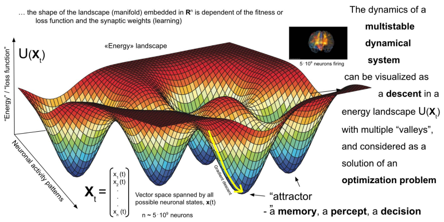
http://projector.tensorflow.org



Natural speech reveals the semantic maps that tile human cerebral cortex, Nature 2016
https://gallantlab.org/huth2016

- **Computations**

Neural networks & neurodynamics in brain and machine



Challenges:

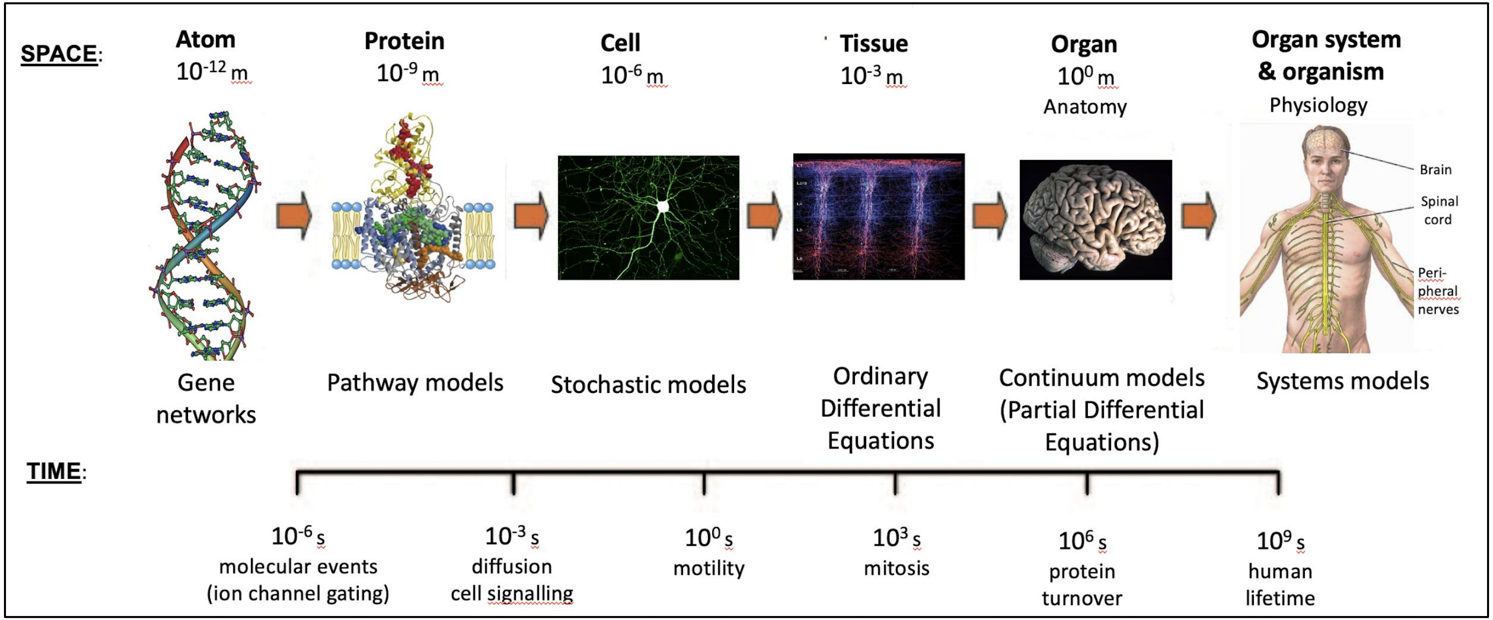
Δ mindset

Δ skillset

Δ toolset

A **graph** G is an ordered pair $G=(V, E)$ where V denote a set of **vertices** (nodes) and $E \subseteq V \times V$ are the **edges** (arcs, connections) of the graph

The future of computational medicine, modeling and machine learning…



Challenges:

Δ mindset

Δ skillset

Δ toolset

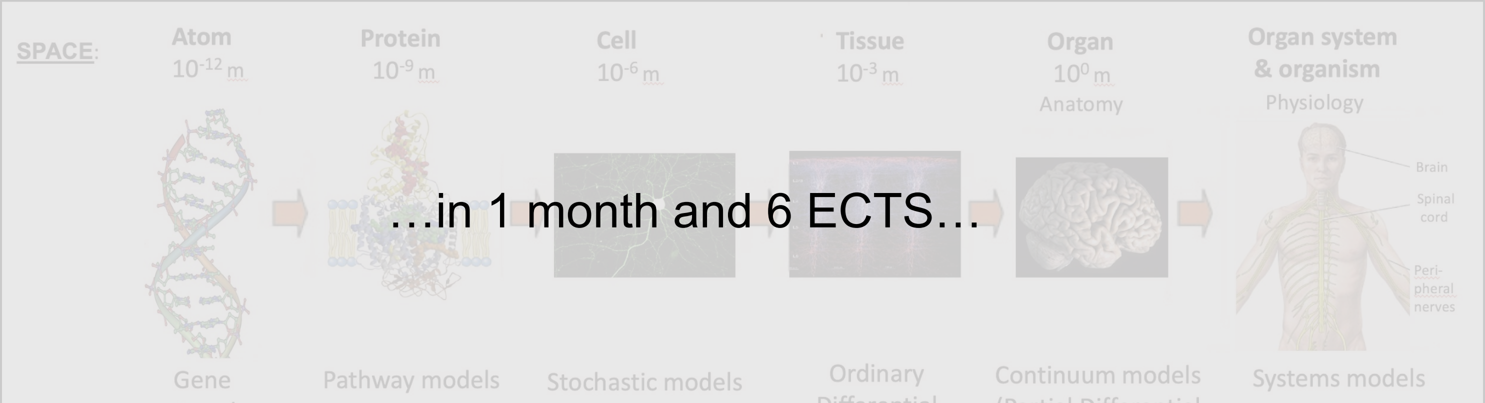
- open science

- reproducible research

- education
- training

The future of computational medicine, modeling and machine learning…

Predictive models: $y \approx f(X; \theta)$ y: outcome, f: model, X: data, θ: model parameters



…in 1 month and 6 ECTS…

Challenges:

Δ mindset

Δ skillset

Δ toolset

- open science

- reproducible research

- education
- training


ELMED219
helping to see the landscape


… in a new and
enlightening perspective