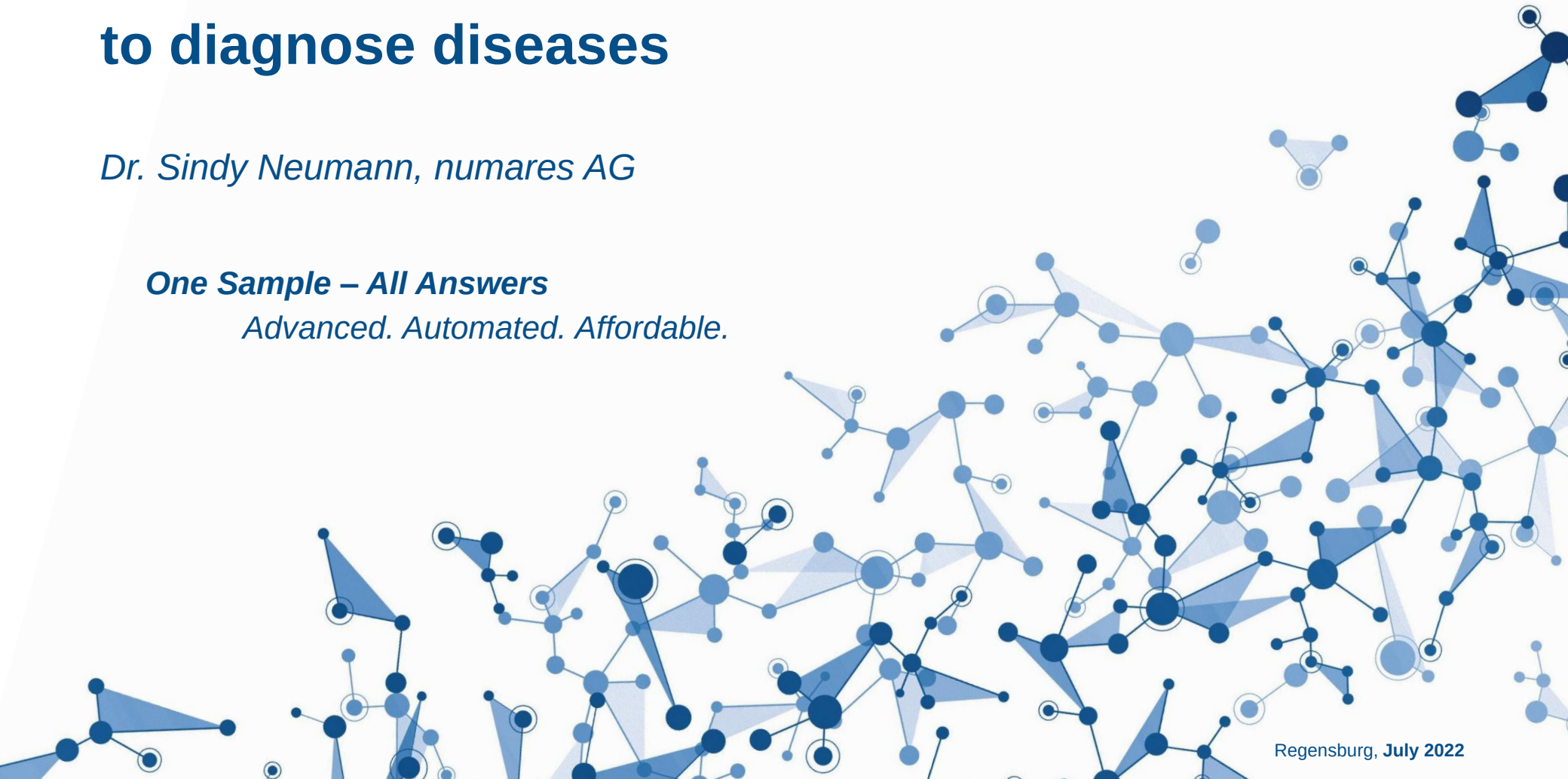
Machine learning in healthcare: How metabolic biomarkers can be identified to diagnose diseases

Dr. Sindy Neumann, numares AG

One Sample – All Answers

Advanced. Automated. Affordable.



Regensburg, **July 2022**

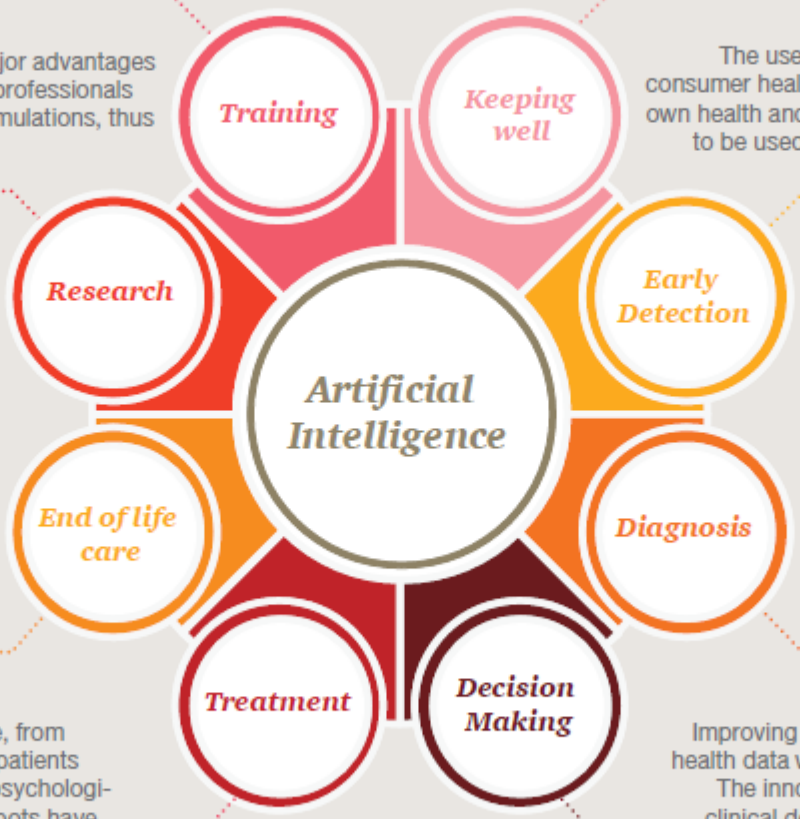
numares at a glance

numares is a NMR diagnostics company shaping future diagnostics using Metabolomics and Machine Learning for software-controlled multimarker diagnostics

2004	60+	Patents	2
Funded: 2004 as a spin-off of Institute of Biophysics & Physical Biochemistry of Regensburg University	Employees, 12 nationalities, with backgrounds in medicine, biochemistry, physics, computational science	15 patent families, 30+ peer-reviewed publications, several ongoing clinical studies	Locations worldwide: Boston, USA Regensburg, Germany

Applications of AI in healthcare

AI applications in healthcare



Training

The use of VR in healthcare has shown major advantages including improving training of healthcare professionals through providing realistic and accurate simulations, thus improving comprehension and retention.

Keeping well

The use of AI and the internet of medical things in consumer health applications helps people manage their own health and keeps them well. VR has a high potential to be used as a tool to improve public health issues.

Research

Research has shown that digitization can and will help healthcare professionals in making decisions and discoveries much faster than they previously could have. It will also pave the way to new and emerging jobs.

Early Detection

AI, the use of wearables and other devices can be applied to detect diseases such as Cancer, or to monitor cardiac diseases at a very early stage, allowing doctors to closely monitor those patients and helping them save lives.

End of life care

Robotics is being used to care for the elderly, helping them remain independent, reminding them of daily tasks and tracking their progress. Advancements might even go further to include having "conversations" and socialization with the elderly.

Diagnosis

AI is being implemented across different hospitals worldwide to solve issues of misdiagnoses. AI has the ability to process information much faster than any human can, thus increasing efficiencies and accuracy.

Treatment

Robotics is being widely used in healthcare, from surgery to supporting self management of patients with long term conditions and for treating psychological conditions. Compared with humans, robots have the ability to achieve better health outcomes.

Decision Making

Improving patient care requires the alignment of big health data with the appropriate and timely decisions. The innovation of predictive analytics will support clinical decision-making and deliver administration priorities and actions.

[Source: Sherlock in Health, *How Artificial Intelligence may improve quality and efficiency whilst reducing healthcare costs in Europe*; pwc, June 2017]

Introduction to Metabolomics

[Gonzalez-Covarrubias et al. 2022; Zhang et al. 2015; Zhang et al. 2020]

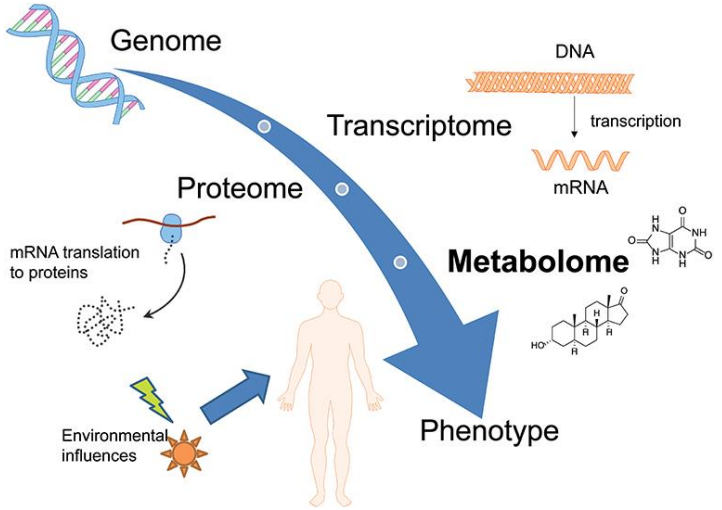
- **Definition**: study of small molecules, commonly known as metabolites, within an organism
- **Matrices**: cells, tissues, body fluids (urine, serum, plasma, CSF etc.)
- **Applications**:

  Nutrition Science, Drug Development,

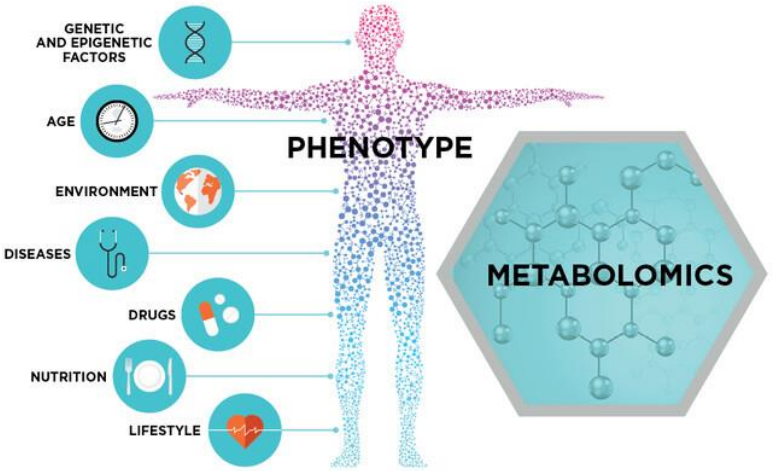
  Diagnostics: Alzheimer‘s disease,

  Diabetes, Prostate cancer, CVD etc.

  → **Clinical metabolomics**
- **Technology**: Mass spectrometry (MS), NMR (nuclear magnetic resonance)

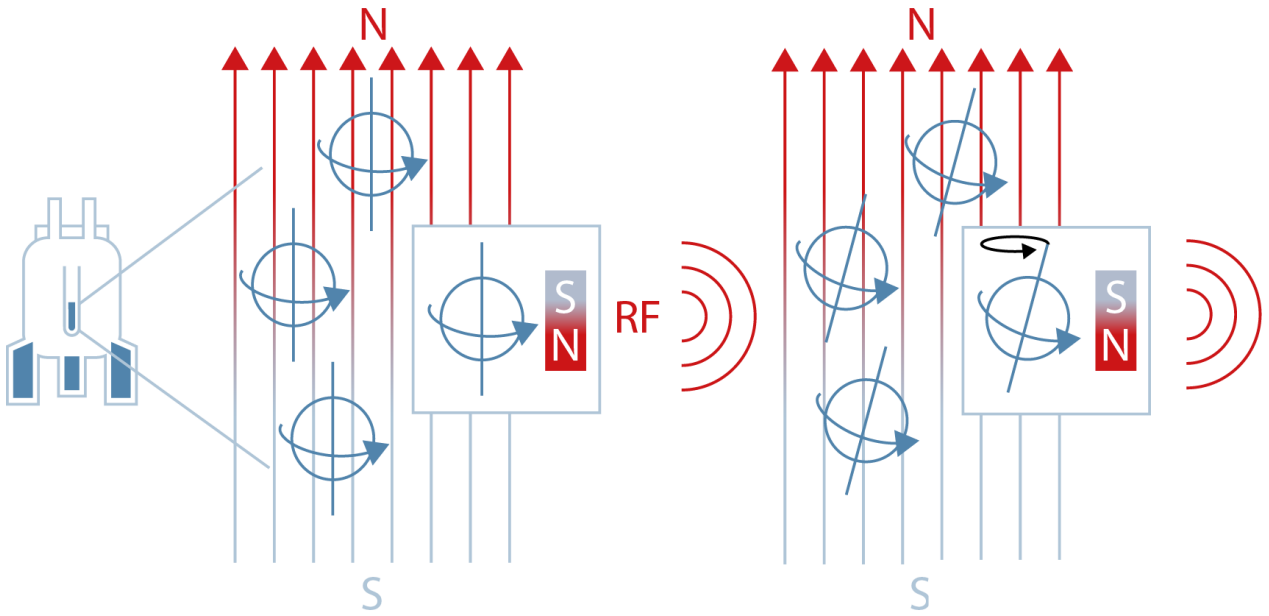


Source: https://www.frontiersin.org/articles/10.3389/fchem.2019.00319/full



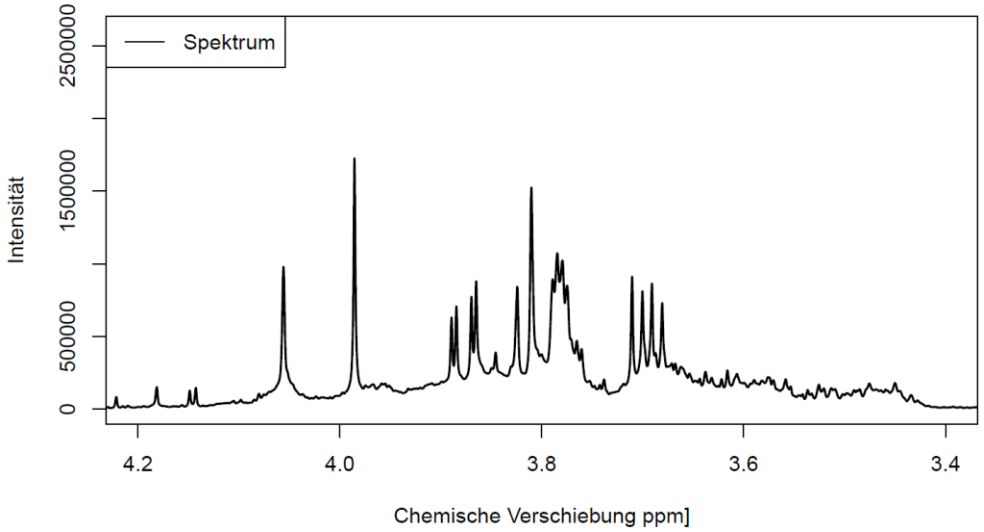
Source: https://www.mtidx.com/our-technology/metabolomics

NMR – the technology in a nutshell

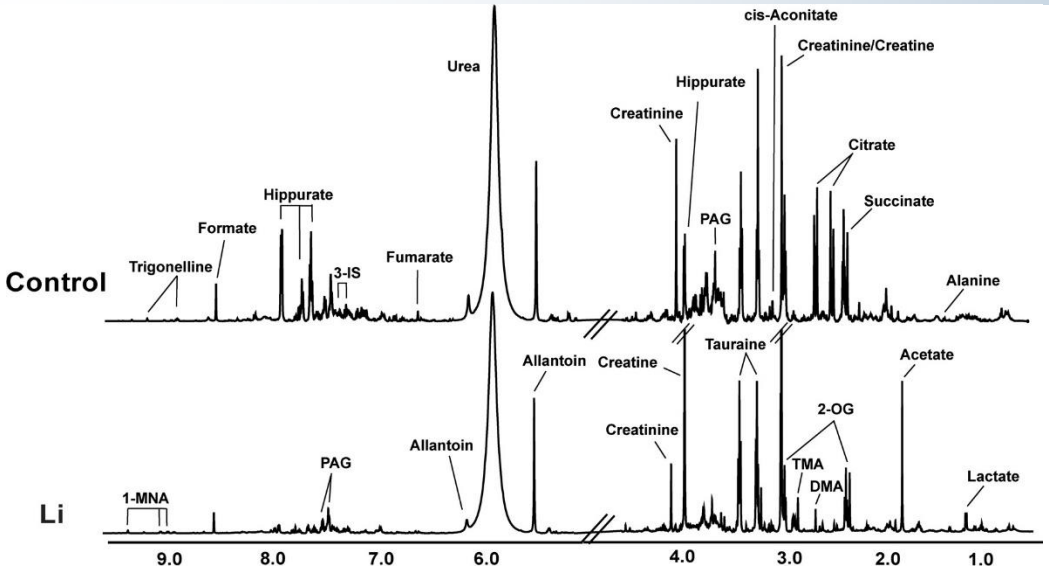


Different chemical environments cause the NMR signal to be slightly “out of tune” with nominal resonance frequency (chemical shift)
→ NMR spectrum

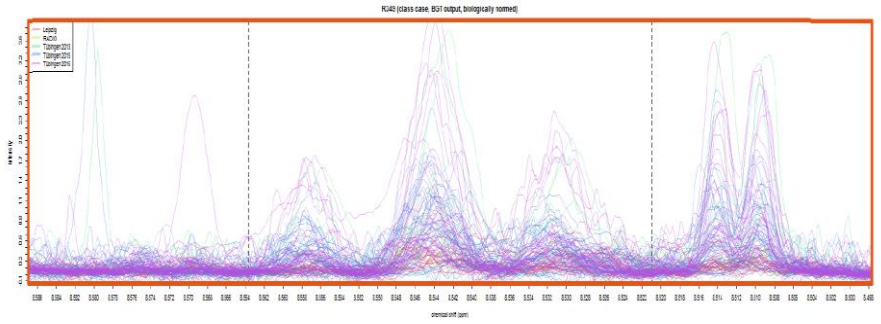
- One substance can be associated with one or multiple signals
- Integrals of signals are proportional to concentrations of corresponding substances



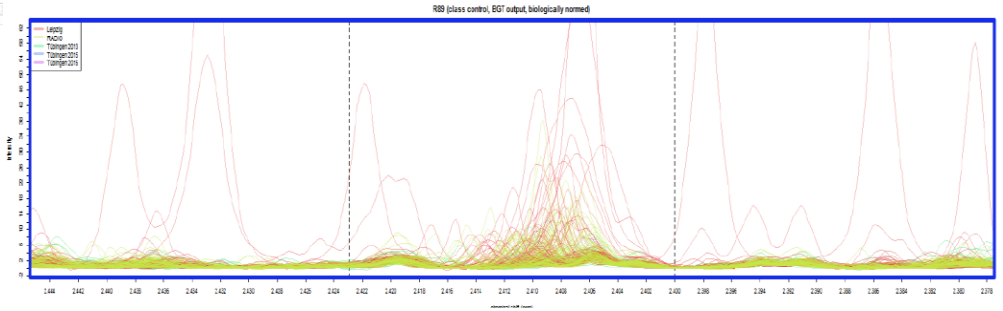
NMR data can be quite challenging



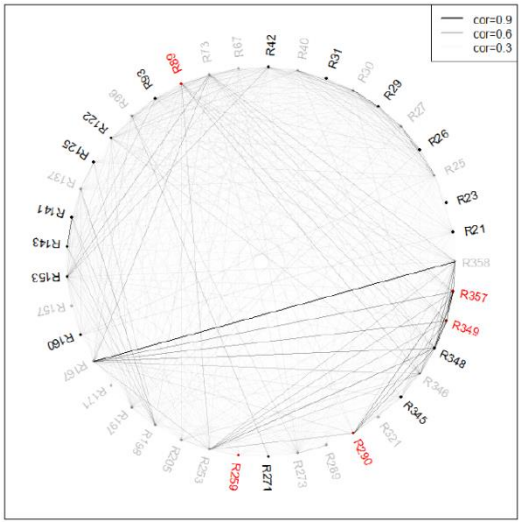
- Dirty (background etc.)
- Noisy
- Redundant (multi-collinearity)
- Variations in chemical shift



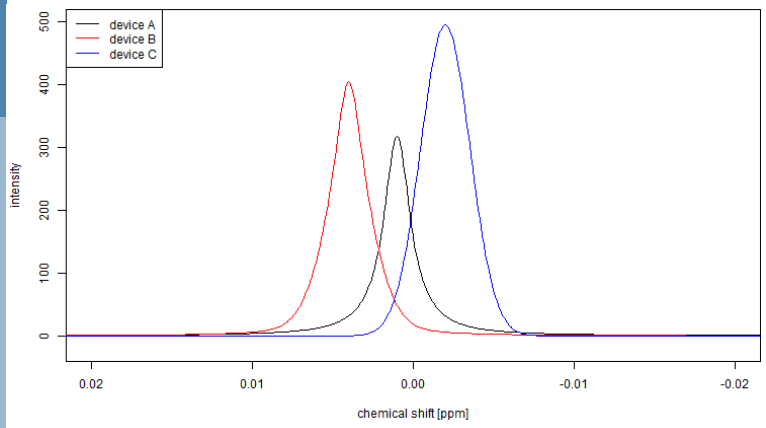
Noise, Shifts, Overlaps

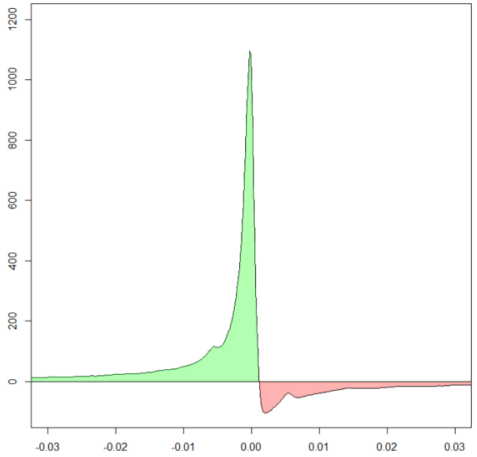


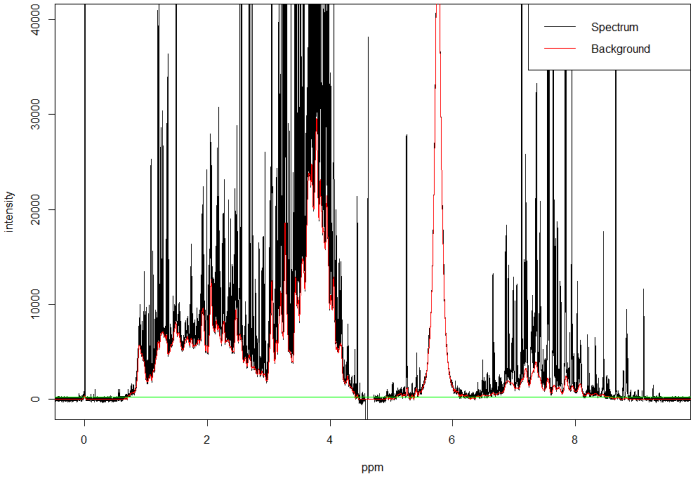
Complex Correlation Structure

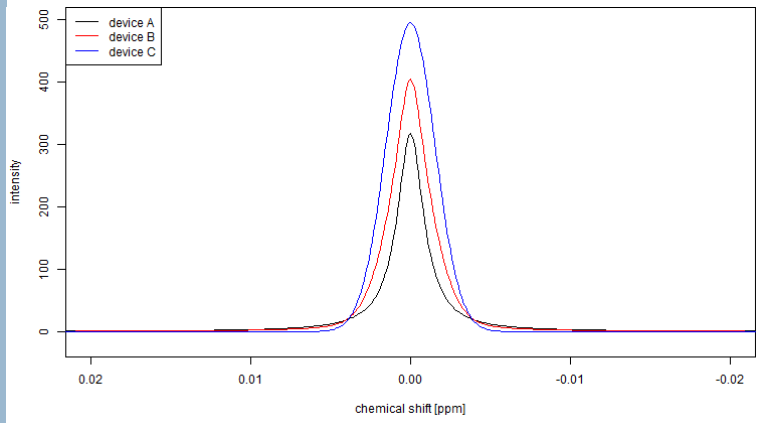


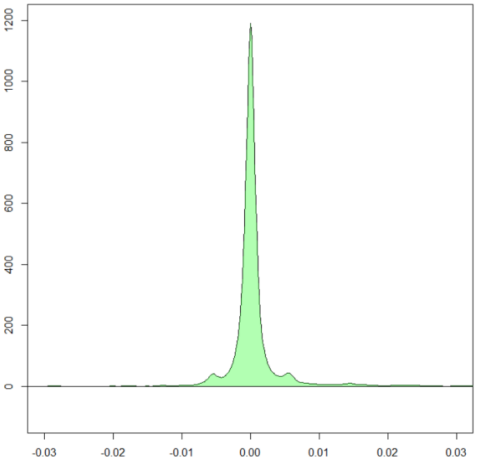
NMR data needs to be processed

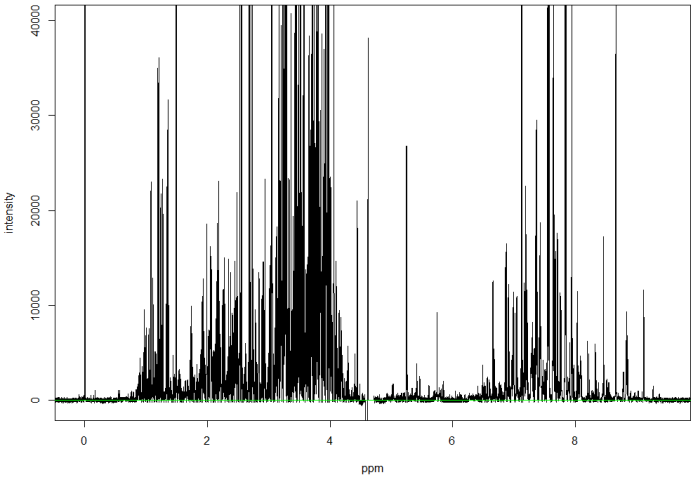










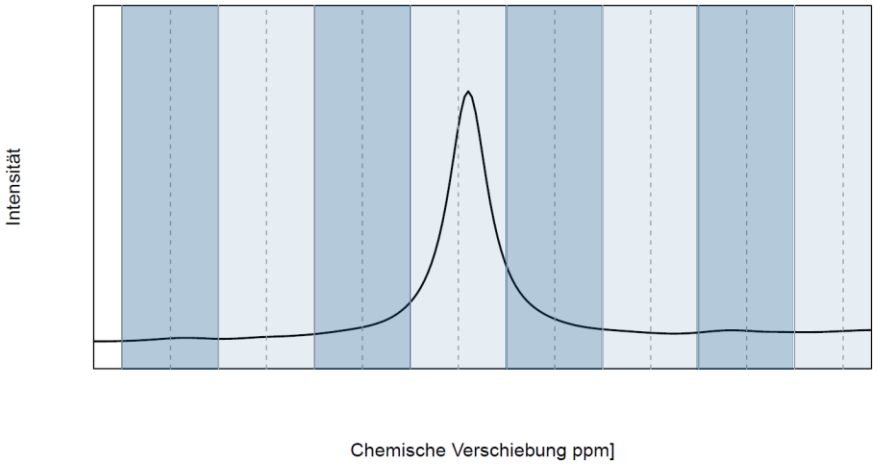


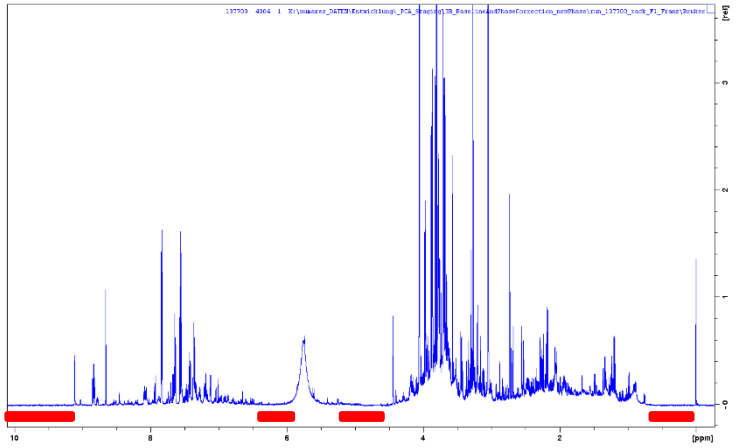
REFERENCING

PHASE CORRECTION

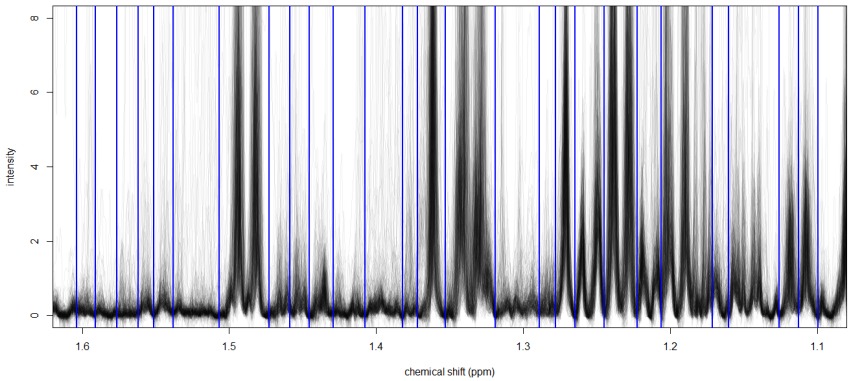
BACKGROUND TREATMENT

NMR data needs to be processed (cont.)

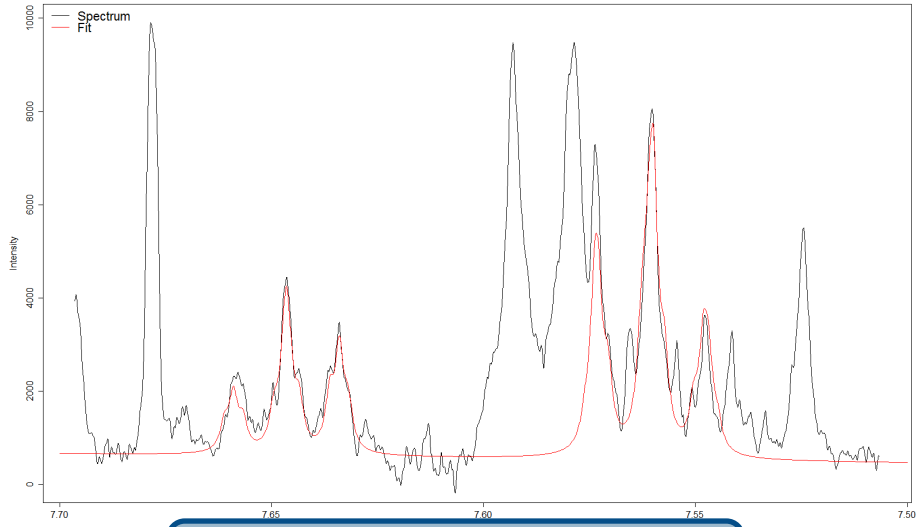




NOISE REGIONS

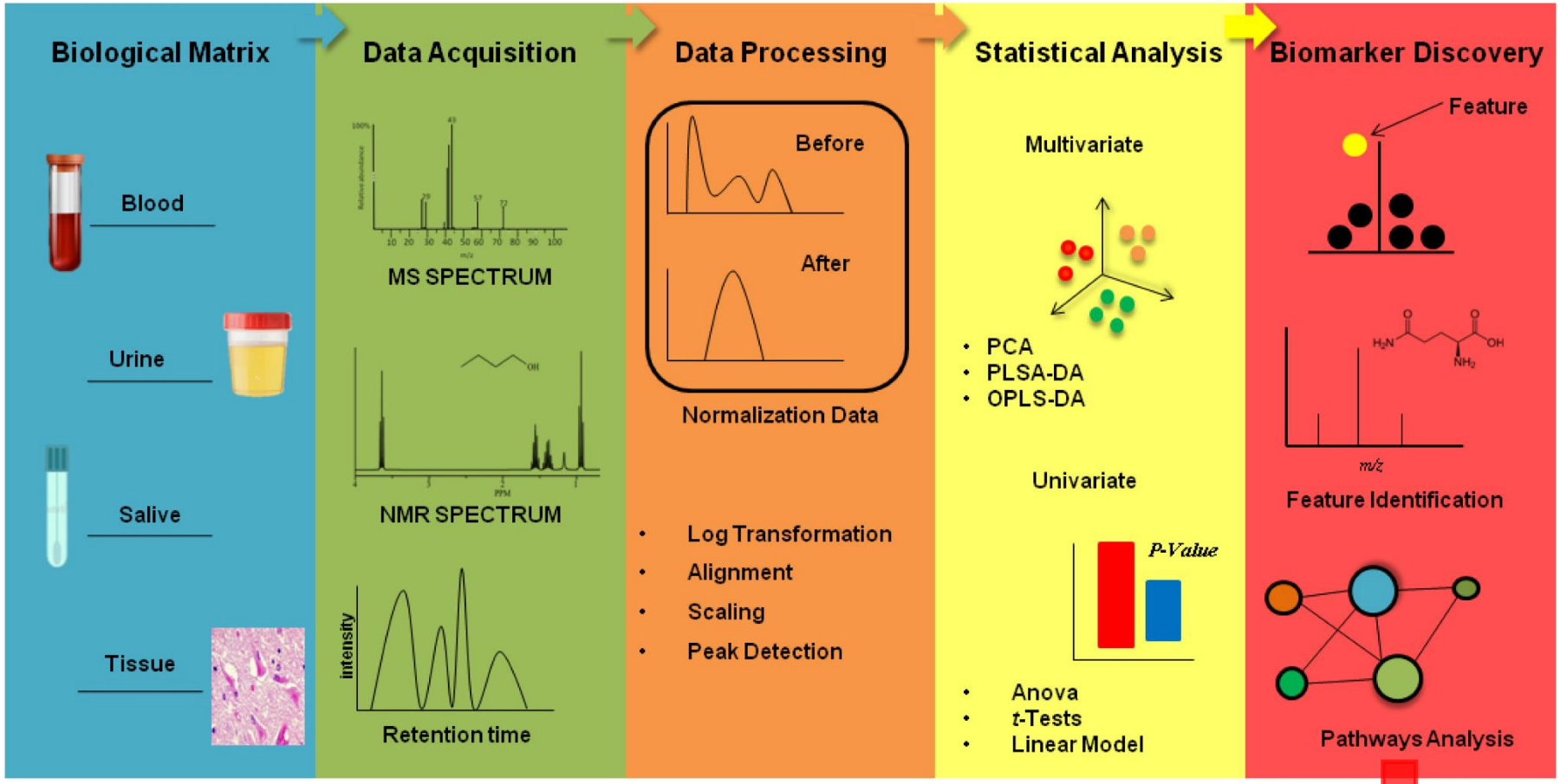


BINNING



SPECTRAL FITTING

Metabolomics workflow



Source: https://www.scielo.br/j/aem/a/yqQ9KFBbVwJPsjdWDwMkSwB/?lang=en#

Translation into clinical practice

Analysis of metabolomics data

[Corsaro et al. 2022; Debik et al. 2021; Pomyen et al. 2020]

numares HEALTH

Data (Pre-) Processing

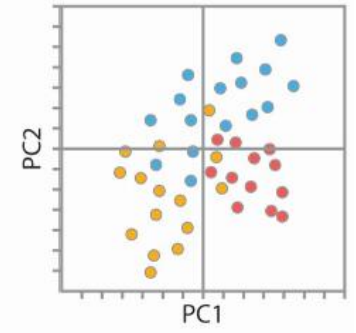
- Phasing
- Baseline correction
- Binning
- Normalization
- Scaling
- Transformation
- Outlier detection
- Missing value imputation
- Etc.

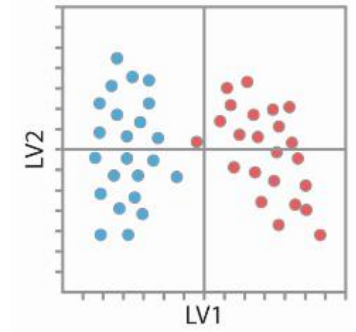
Data Analysis

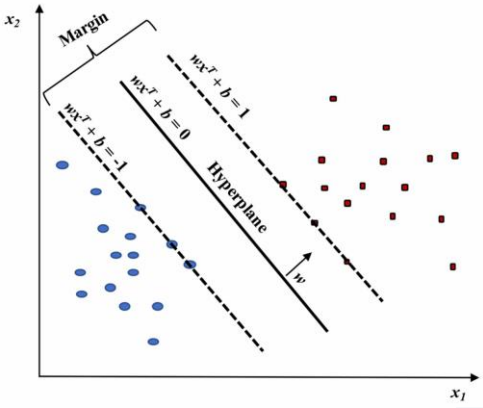
- **Univariate testing:** t-Test, ANOVA, chi-square …
- **Dimensionality reduction**: PCA, PLS
- **Clustering**: k-means, SOMs
- **Regression**: PCR, PLS
- **Classification**: (O)PLS-DA, SVM, RF, logistic regression
- Etc.

Biological interpretation

- Pathway analysis: ORA, FCS



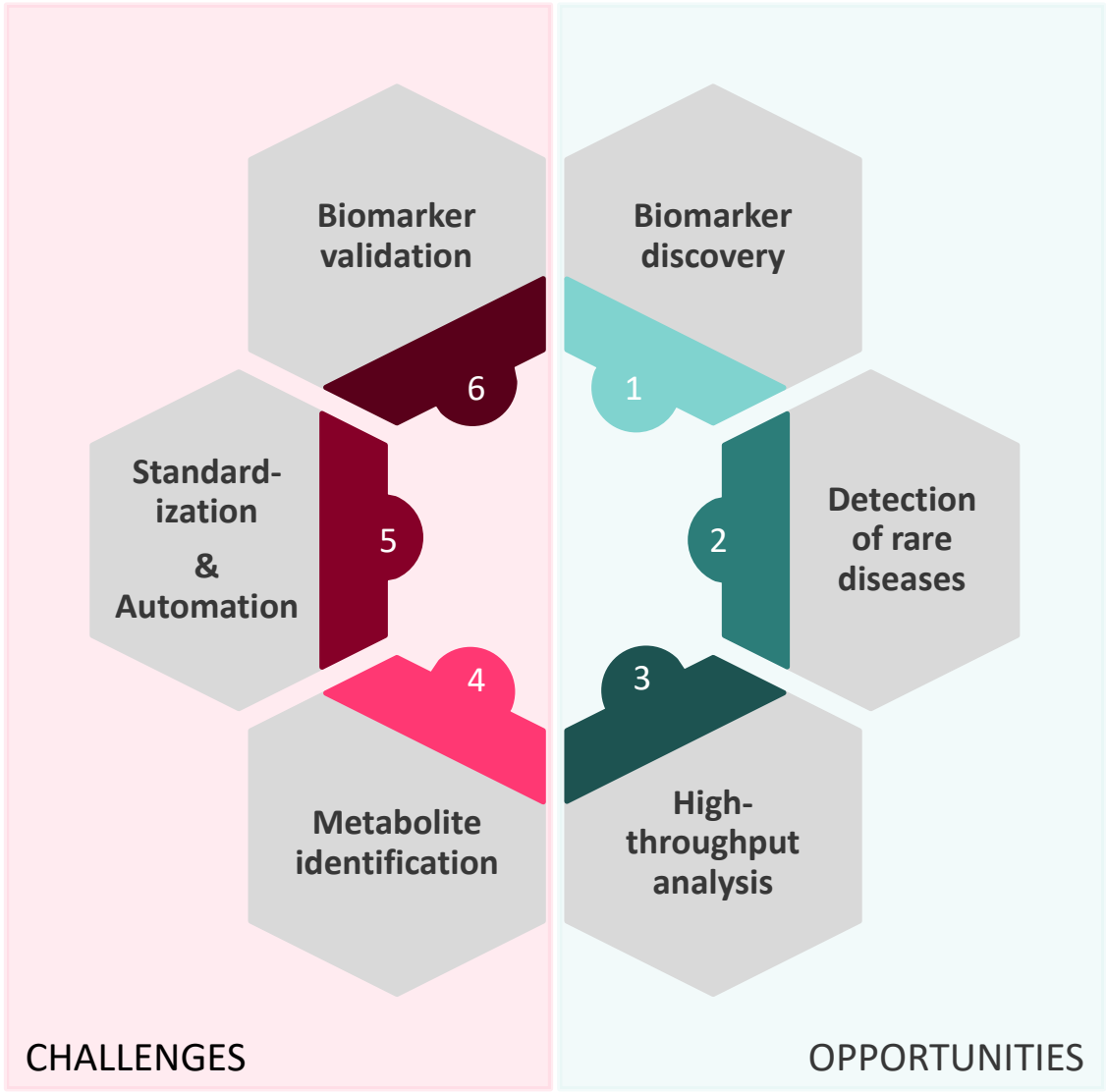




Deep Learning: application still limited, in particular data pre-processing (CNNs!), more comprehensive applications in other omic domains

Opportunities and Challenges of Clinical Metabolomics

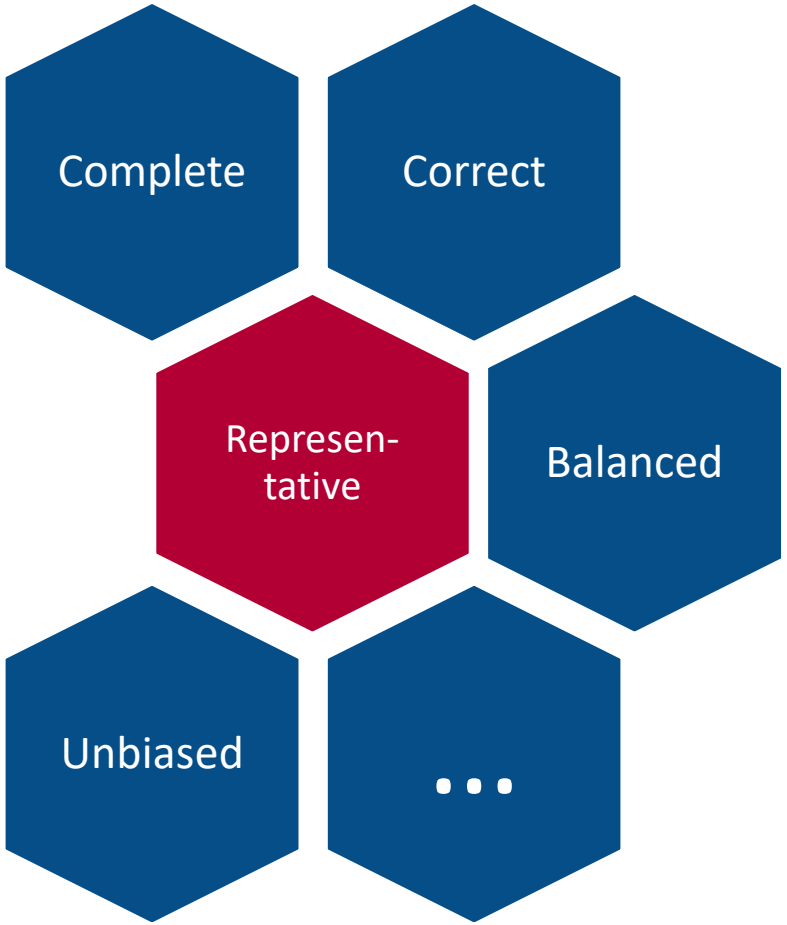
[Gonzalez-Covarrubias et al. 2022; Tebani et al. 2016; Zhang et al. 2015]



Overall challenge: necessity for good data

Garbage in, garbage out …

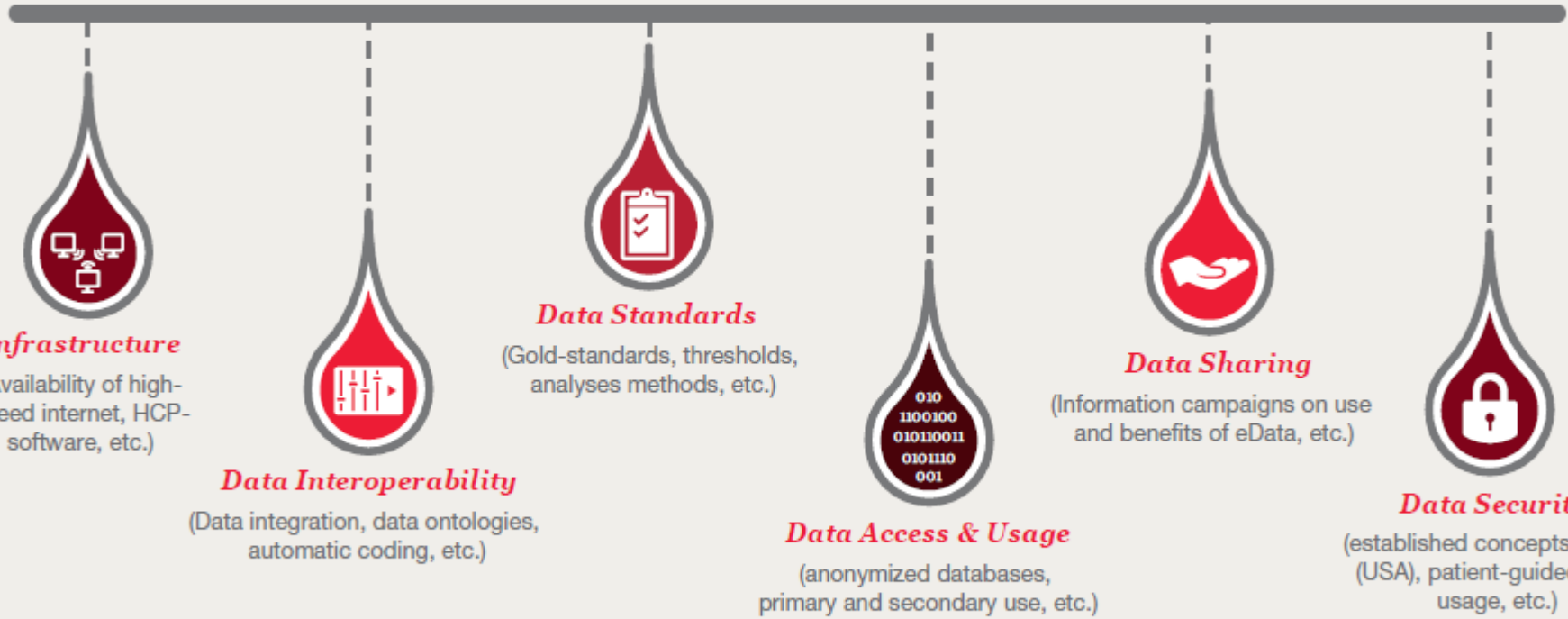
[Source: xkcd]



Big Data versus Smart Data

Benefits and Challenges from AI in healthcare

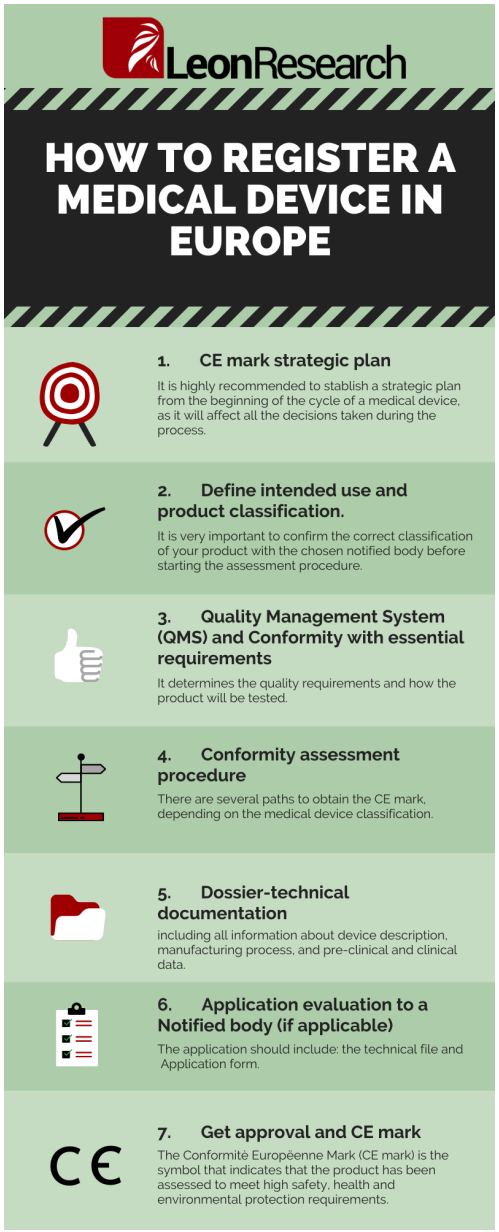
Data are the enabler but also the issue



HIPAA refers to the Health Insurance Portability and Accountability Act (US), which has created national standards to protect individuals' medical records and other personal health information

[Source: Sherlock in Health, *How Artificial Intelligence may improve quality and efficiency whilst reducing healthcare costs in Europe*; pwc, June 2017

Translation into clinical practice (Registration and Regulations)



CE FDA

Europe	USA

• requires CE mark!
• QMS necessary (ISO 13485)
• Conformity assessments (CE mark certifies that medical device complies with MDR/IVDR)
• Regulations: MDR ((EU) 2017/745), IVDR ((EU) 2017/746)
• 4 risk classes according to MDR (risk I, IIa, IIb, III)
• 4 risk classes according to IVDR (A, B, C, D)
• Notified body (e.g. TÜV)
	• Responsible: FDA
	• Types of submission:
	  • Premarket Notification (510k)
	  • Premarket Approval (PMA)
	  • De Novo

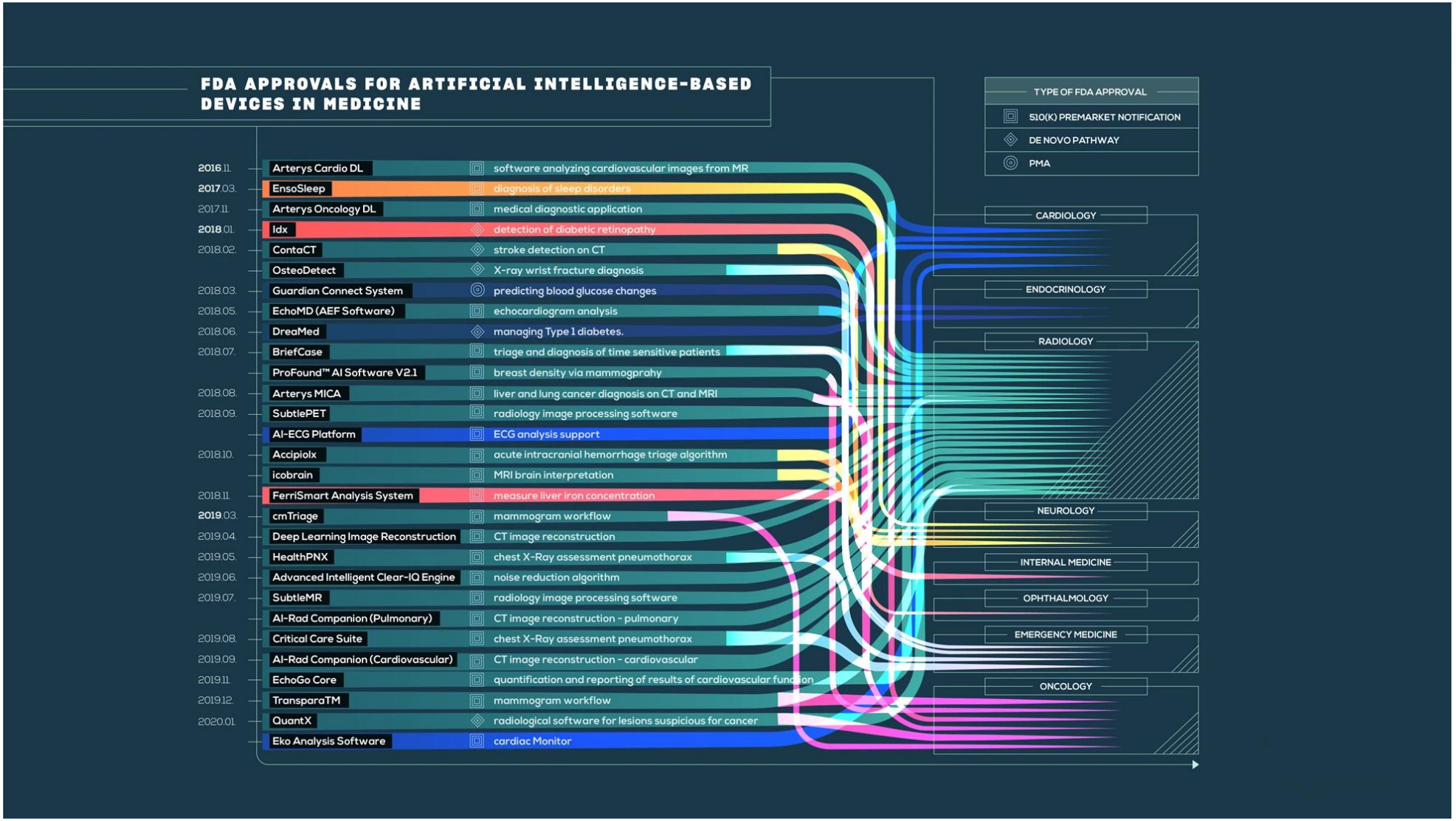
[Source: https://leonresearch.com/medical-devices-europe-mdr-2017-745/]

FDA-Guideline: Good Machine Learning Practice for Medical Device Development (GMLP, October 2021)

Good Machine Learning Practice for Medical Device Development: Guiding Principles	
1 Multi-Disciplinary Expertise Is Leveraged Throughout the Total Product Life Cycle	2 Good Software Engineering and Security Practices Are Implemented
3 Clinical Study Participants and Data Sets Are Representative of the Intended Patient Population	4 Training Data Sets Are Independent of Test Sets
5 Selected Reference Datasets Are Based Upon Best Available Methods	6 Model Design Is Tailored to the Available Data and Reflects the Intended Use of the Device
7 Focus Is Placed on the Performance of the Human-AI Team	8 Testing Demonstrates Device Performance During Clinically Relevant Conditions
9 Users Are Provided Clear, Essential Information	10 Deployed Models Are Monitored for Performance and Re-training Risks are Managed

FDA-approved AI-based medical technologies

https://medicalfuturist.com/fda-approved-ai-based-algorithms/



[Source: Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med*. 2020;3:118. Published 2020 Sep 11. doi:10.1038/s41746-020-00324-0]

References

- Corsaro et al. *NMR in metabolomics: from conventional statistics to machine learning and neural network approaches*. Appl. Sci. **2022**, 12(6)
- Debik et al. *Multivariate analysis of NMR-based metabolomic data*. NMR in Biomedicine **2021**, 35(10)
- Gonzalez-Covarrubias V, Martinez-Martinez E, Del Bosque-Plata L. *The potential of metabolomics in biomedical application*s. Metabolites **2022**, 12(2)
- Pomyen et al. *Deep metabolome: applications of deep learning in metabolomics*. Comput Struct Biotechnol J. **2020**
- Tebani et al. *Clinical Metabolomics: the new metabolic window for inborn errors of metabolism investigations in the post-genomic era*. Int. J. Mol. Sci. **2016**, 17(7)
- Zhang et al. *Metabolomics for biomarker discovery: moving to the clinic*. Biomed Res Int. **2015**
- Zhang et al. *Mass spectrometry-based metabolomics in health and medical science: a systematic review.* RSC Adv. **2020**

numares HEALTH

Thank you for your attention

Dr. Sindy Neumann
Head of Biostatistics, numares AG

sindy.neumann@numares.com