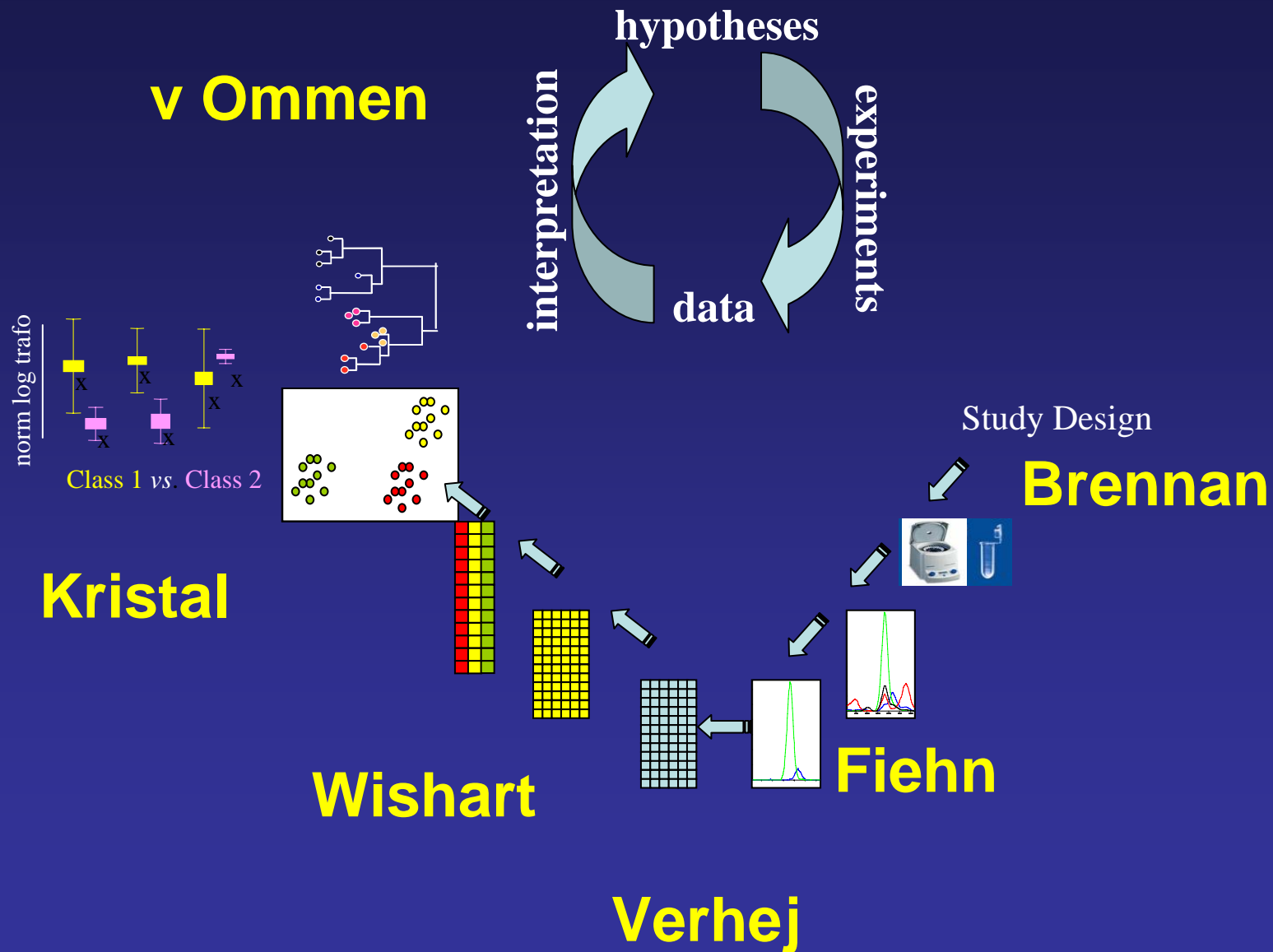
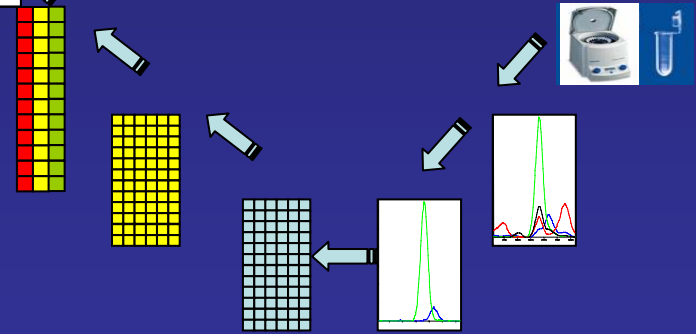
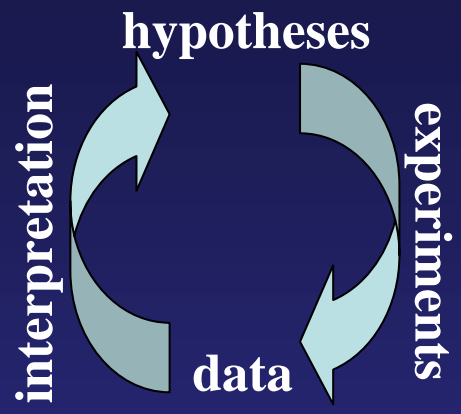
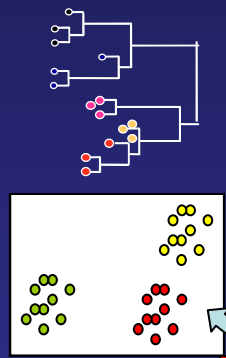
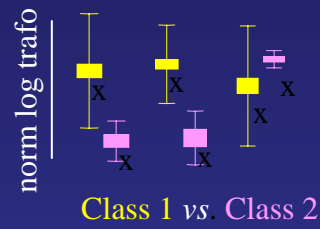


Summary Discussion

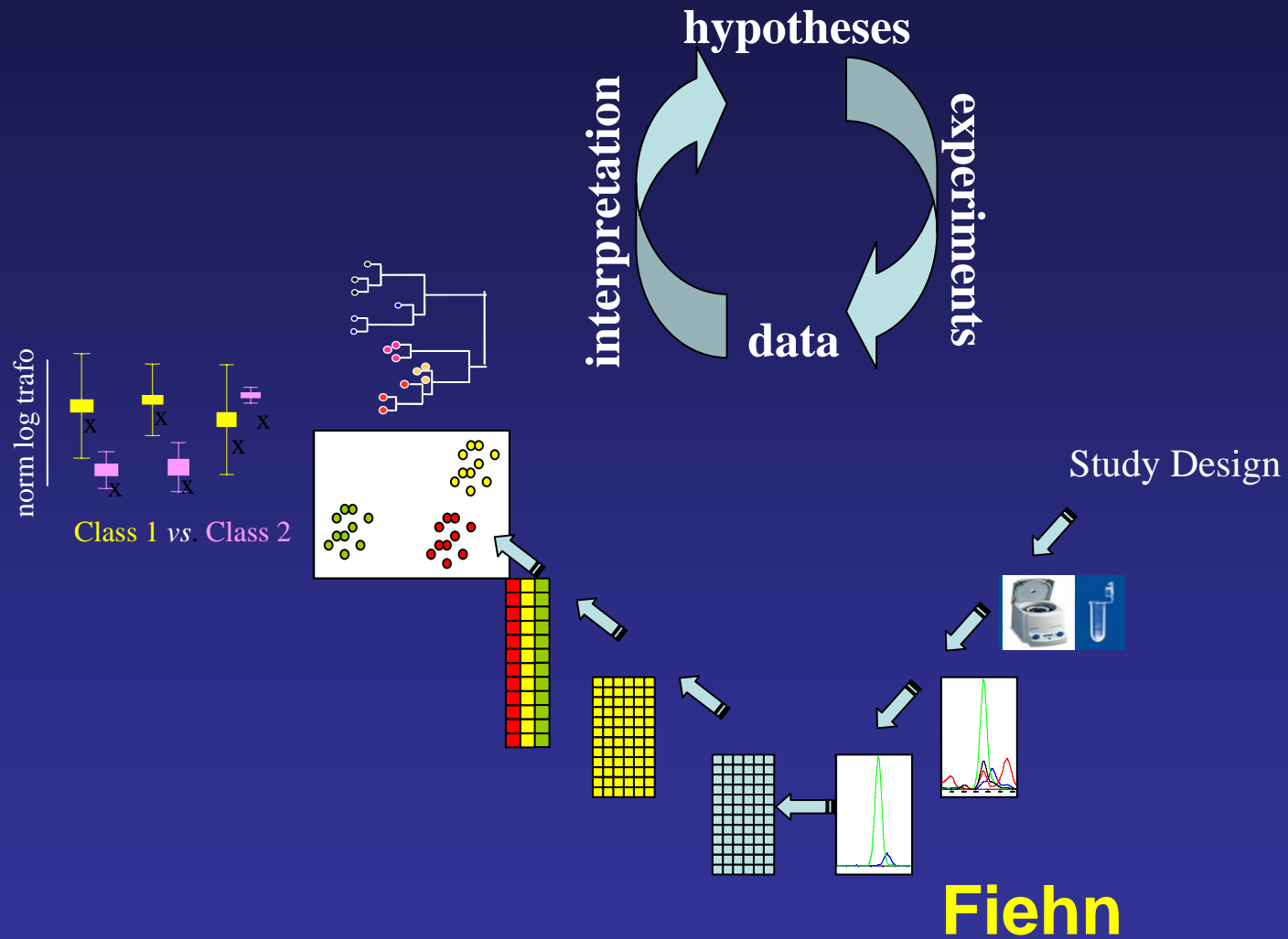




Brennan: Sampling strategy

- **Sampling Time** **Urine 24h**
- **Ben and Ivana make standard matrices**
No optimal method
- **Blood samples:** serum, plasma
pre-processing
Exchange of SOPs
Small round robin...
- **Storage:** short term storage
long term storage
freeze thawing
No large effect !
- **Quality control and validation** **Critical (LC-MS) !**
- **Faecal samples:** preparation and potential uses in nutrition research

Summary Discussion



Fiehn: MS

- **GC, LC most useful**
- **type of MS less critical than separation or interface**
- **DB: open source**
- **Spectral Librarys : far too few, not public**
- **Multi-target vs unbiased**

GC with RI

GC injector

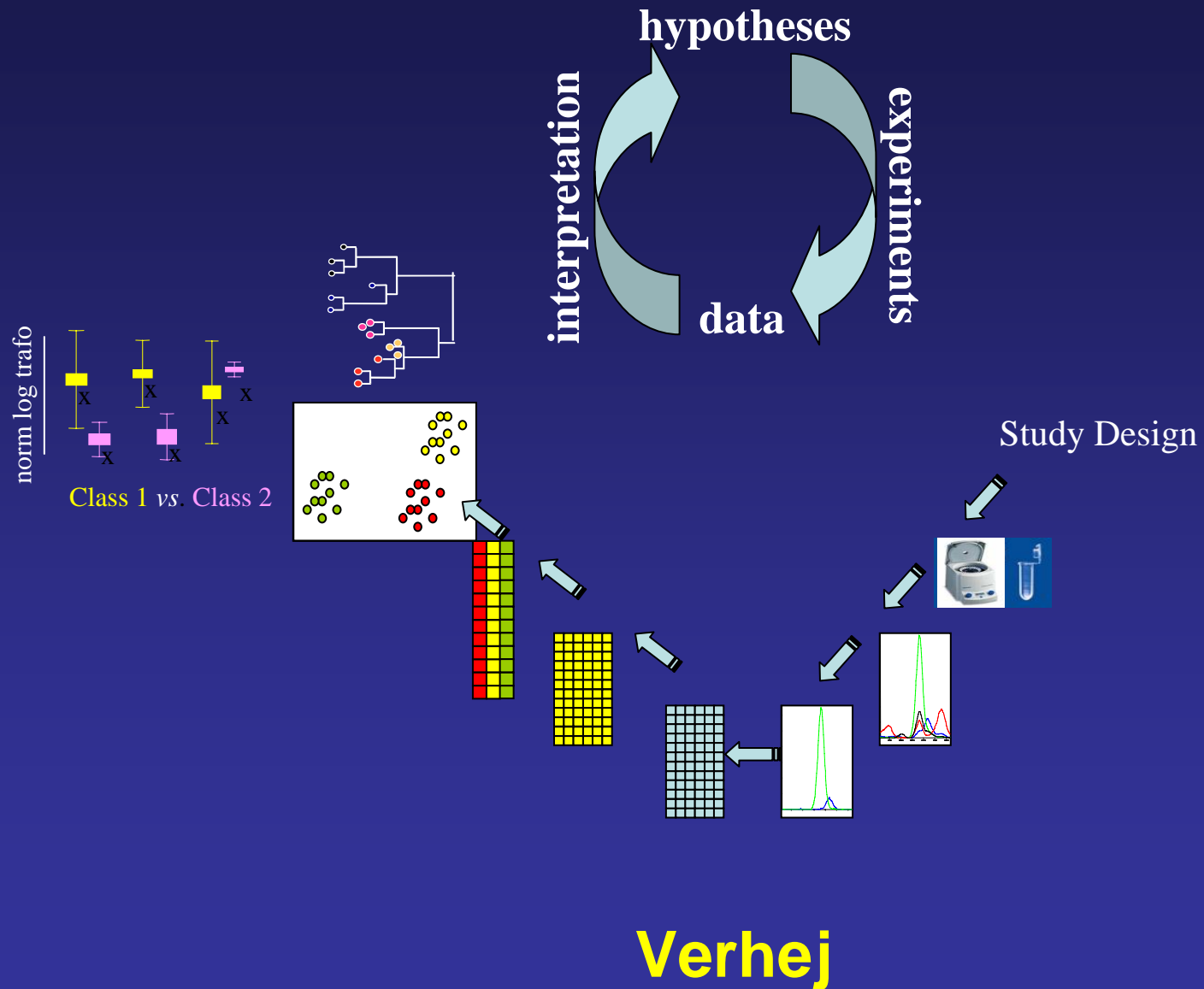
ESI suppression

Comparisons across studies

Society Letter to NIH, support by NuGO

Validation greatly benefits from known (expected) compounds

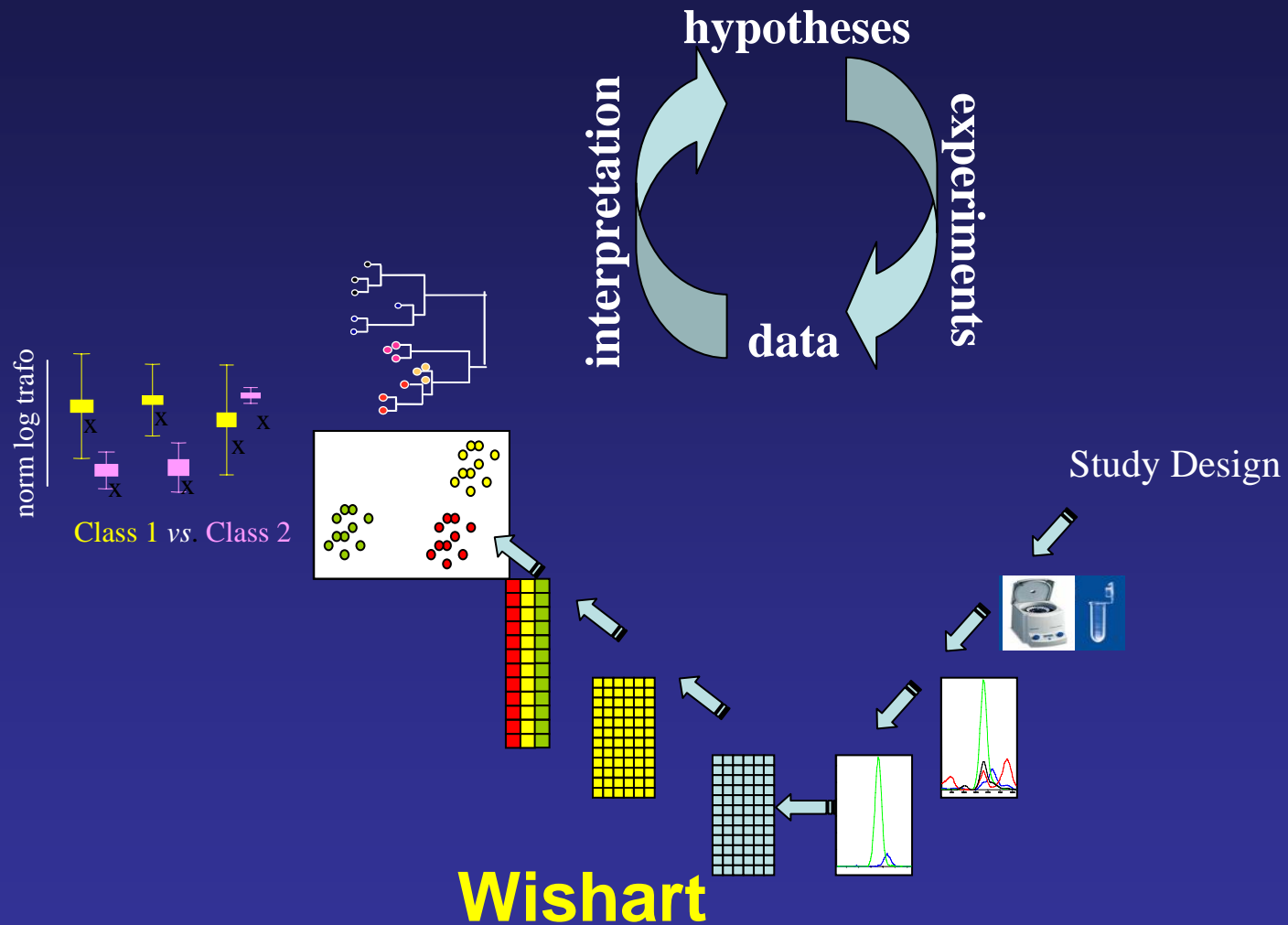
Summary Discussion



Verheij: Data extraction

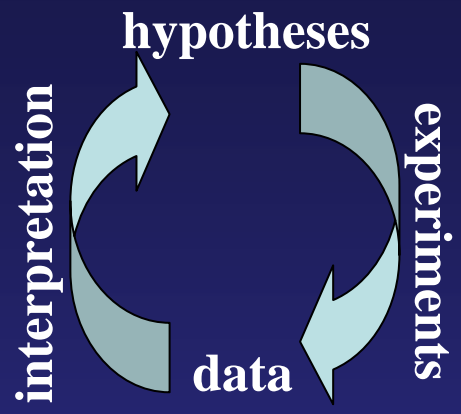
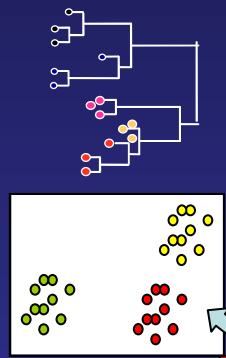
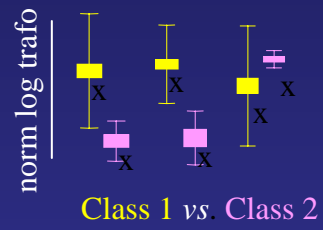
- **Many tools**
Impress (prop.), vendor, XCMS and mzMine
- **tools suffer similar problems**
Missing, split peaks
Baseline artifacts
- **deconvolution**
- **Benchmarking needed**
Public and downloadable data ?
- **email will be sent with questions and input requests and a synopsis**

Summary Discussion



Wishart: compound ID

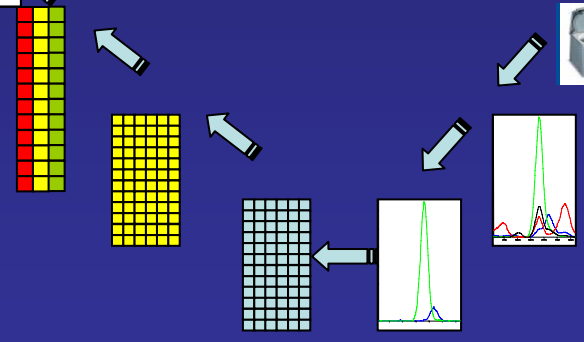
- **ChemSpider and HMDB** **Powerful query tools**
- **elemental composition** **7 golden rules, accurate MS**
- **authentic standards**
- **virtual library of enzyme transformations of exogenous cmpds**
- **chemical derivatives for separation, quant and ID**
- **Benchmarking needed** **50 standard mix as grid and for QC**



Study Design



Kristal



Kristal: statistics and informatics

➤ **“The Problem”**

**Overfitting, eg. PLS-DA
education necessary
perhaps in Boston**

➤ **Gold standard**

**Independent biological
replicate study**

➤ **Secondary route**

Permutation tests

➤ **Tell what you do.**

➤ **Be suspicious (‘mental overfitting’).**

➤ **Understand what you do (curse of
dimensionality, software programs are too easy)**

Informatics: An example classification workflow

Data Validation, Data Normalization,
Missing Data Decisions, Inclusion/Exclusion Criteria



Subgroups, Class-specific models



Outlier removal \leftrightarrow scaling \leftrightarrow transformations

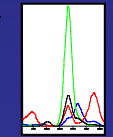
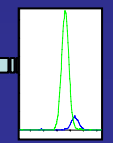
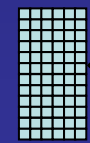
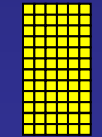
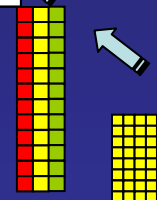
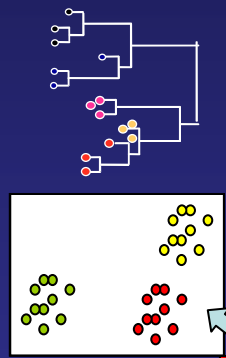
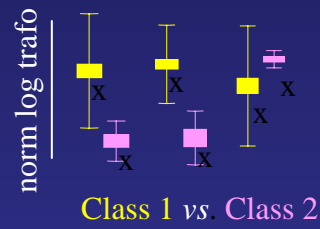


Unsupervised: Clustering SOMs PCA
Supervised: kNN SIMCA PLS PLS-DA Random Forest
Machine learning: Neural Nets GAs GPs

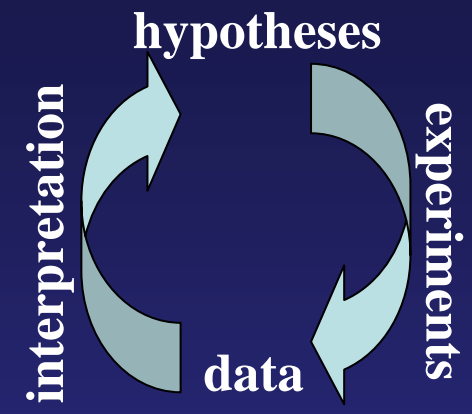


Overfit tests, Internal validation, optimization,
External validation, optimization, 2° validation

v Ommen



Study Design



v Ommen: connect results to knowledge

➤ **support NuGO wiki**

**Easy editable
basic information from HMDB**

➤ **pathway tools**

Pathvisio.org

➤ **tool kit depository**

NuGO net pages

➤ **Nutritional Phenotype DB**

→ **core group**

→ **architecture**

→ **embedding**

→ **funding**

➤ **wish list and list of currently used tools is long !**