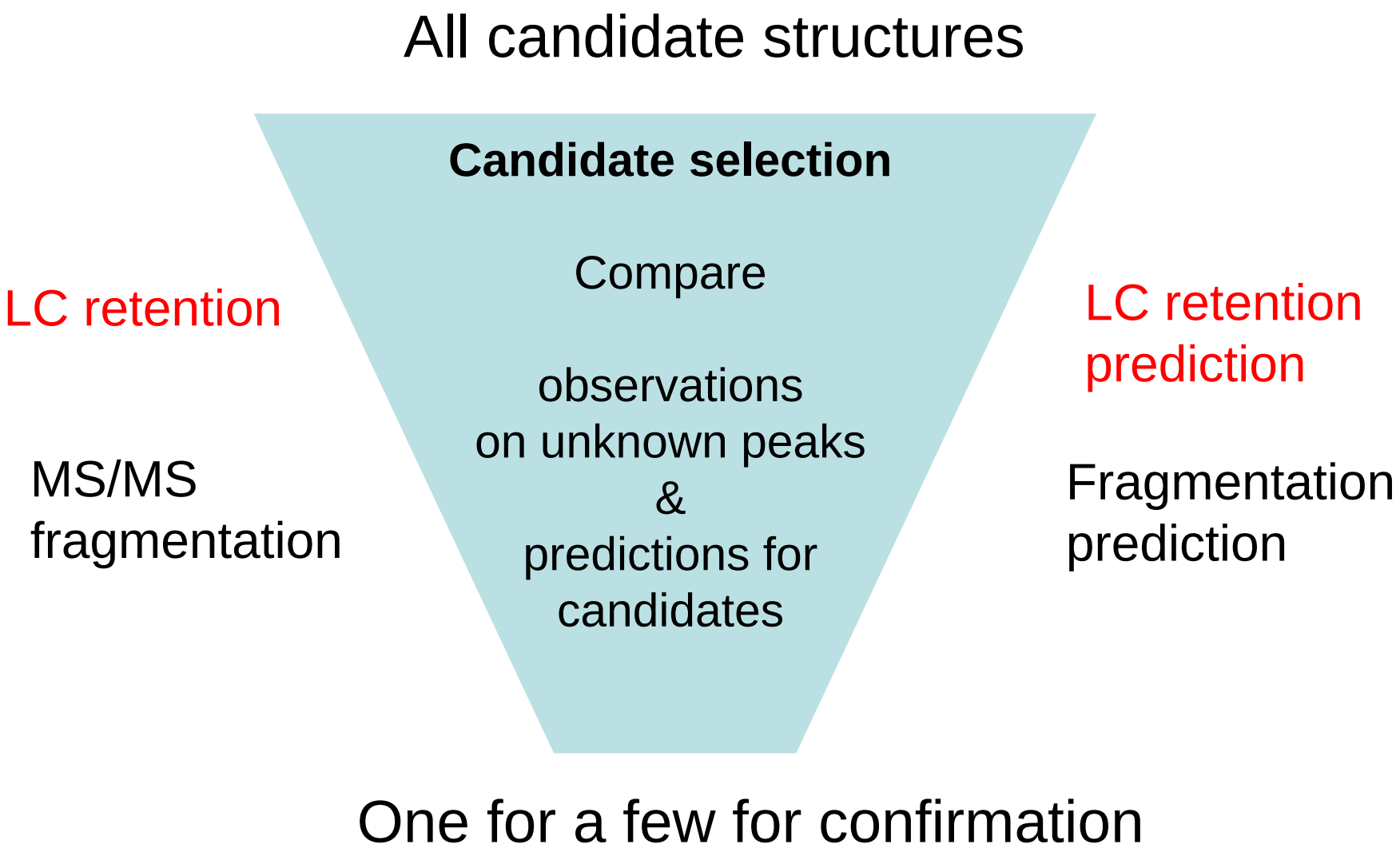
Comparison of approaches for retention prediction Useful in LC-HRMS Nontarget-Screening

Martin Krauss, Janek Dann, Meng Hu, Werner Brack

Department Effect-Directed Analysis,
Helmholtz-Centre for Environmental Research – UFZ, Leipzig

Retention prediction in structure elucidation



Retention prediction

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

Retention time = f (eluent properties (composition, gradient, pH, I),
molecular structure, stationary phase chemistry,
temperature)

If we use always the same system, eluent, column:

Retention time = f (molecular structure)

Build a QSRR model!

Many different approaches in the literature

Héberger, 2007, J. Chromatogr. A 1158, 273. (Review)

Retention prediction models: conceptual view

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

High accuracy

"Small domain" models

Alkylic acids
RT = f (chain length)

Alkylic acids
RT = f (chain length, branching, topology…)

Single parameter

Poly-parameter

100s of pharmaceuticals
RT = f (log K_{OW})

100s of pharmaceuticals
RT = f (phys.-chem., topological descriptors, …)

Lower accuracy

"Large domain" models

Better accuracy?

Retention prediction and retention indices

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

“Is my unknown at 12.3 min your unknown at 18.6 min?”

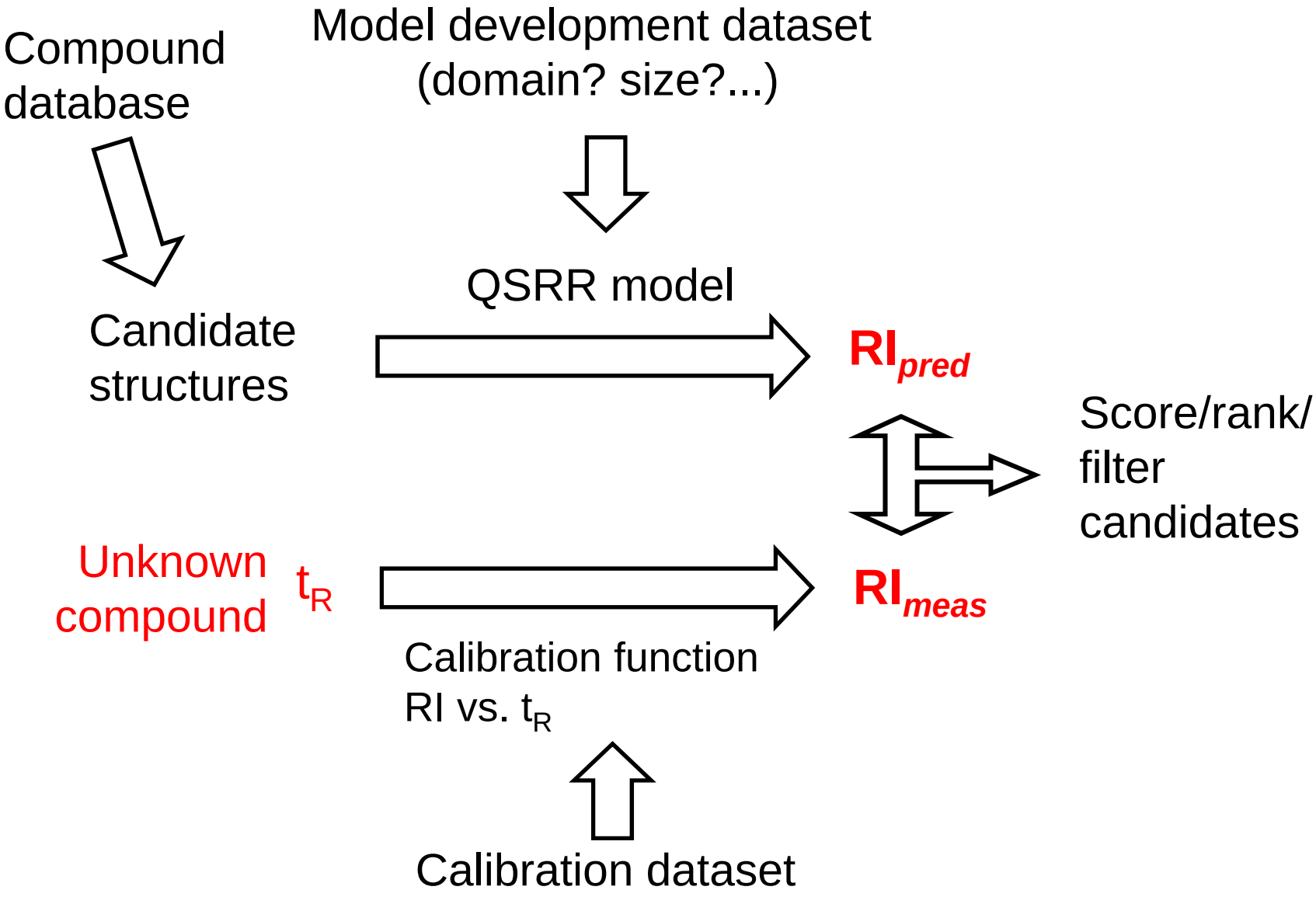
- Comparison among different LC methods, working groups
- Storage in shared databases
- Consistency/validation of LC methods over years

$\Rightarrow$ Normalizing or indexing retention times RI

GC(-MS) (Kovats‘) Retention Index: simple & quite accurate!

LC(-MS): ???

Using retention prediction in practice



Comparing different approaches (342 compounds)

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

Chromatographic Hydrophobicity Index (CHI)

Linear Solvation Energy Relationship ("Abraham equation")

Dept. Effect-Directed Analysis, UFZ

Ulrich et al. (2011) J. Chromatogr. A 1218, 8192.
based on Valko et al. (1997) Anal. Chem. 69, 2022

Retention time index (RTI)

Based on similarity of C18 retention and octanol-water partitioning

Thomas Letzel & coworkers, TU München

Nitroalkane retention index (NARI)

RI on homologue series of n-nitroalkanes

Grant Lab, Univ. of Connecticut

Hall et al., (2013) J. Chem. Inf. Model 52, 1222
Implemented in MolFind software,
Menikarachchi et al., (2013) Anal. Chem. 84, 9388.

CHI: Retention prediction by LSER

$$CHI = aA + bB + sS + eE + vV + c$$

Substance descriptors

Regression coefficients (system descriptors)

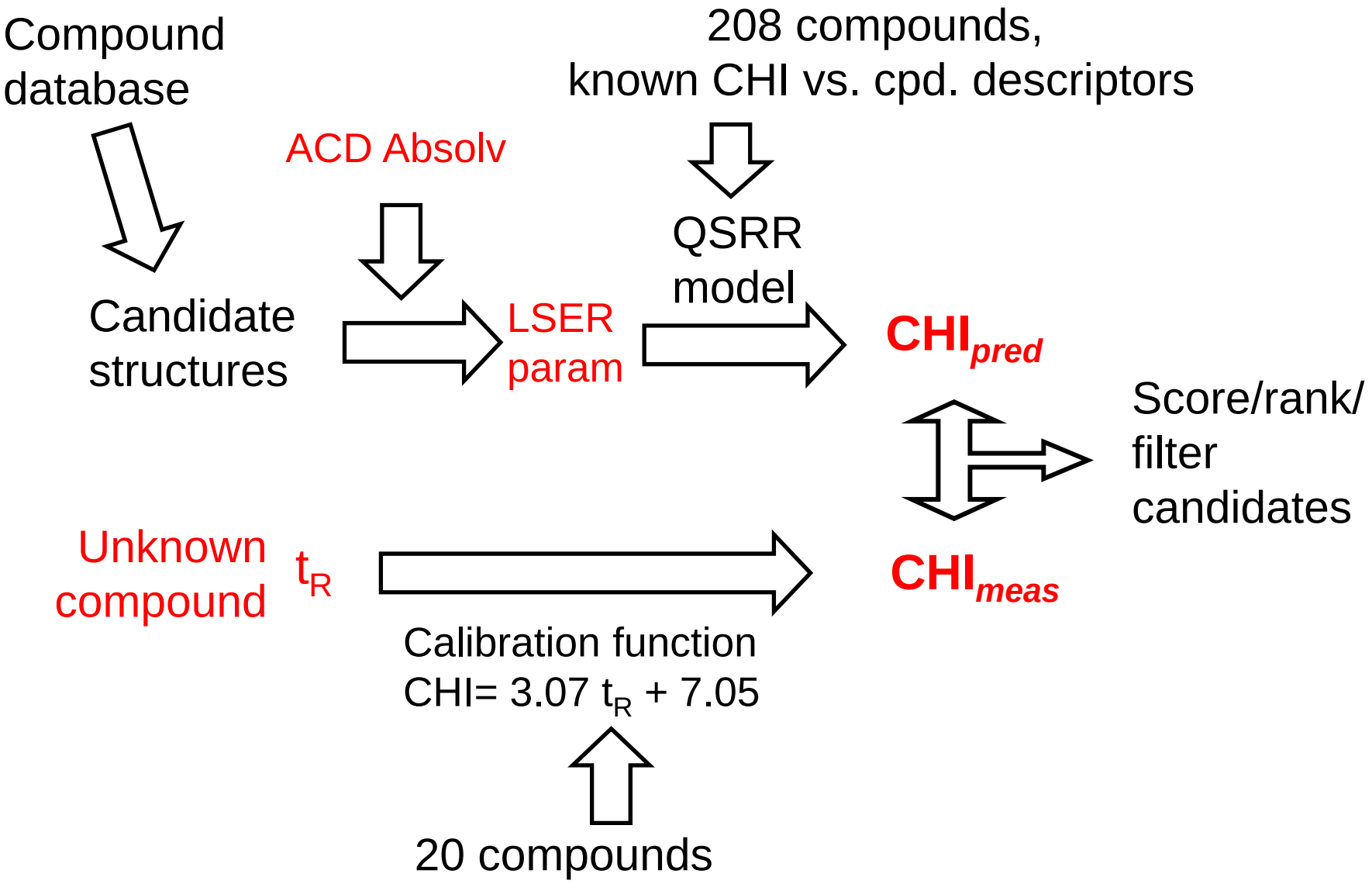
CHI ≈ Percentage of organic modifier required to elute the analyte in a linear gradient (Valko et al., 1997, Anal. Chem. 69, 2022.

CHI is (nearly) independent from gradient setup and column dimensions, but different for ACN and MeOH

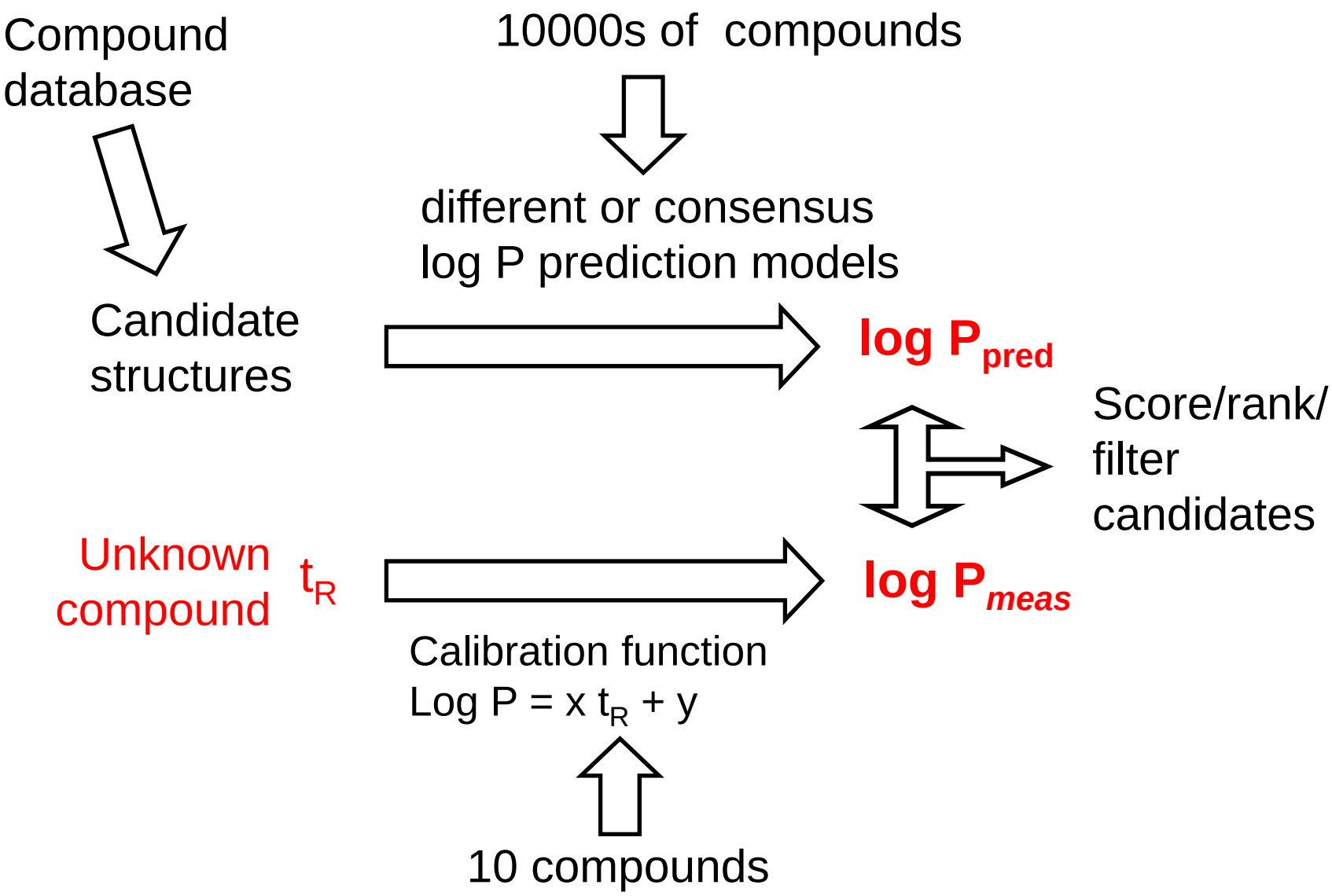
CHI is very similar for different (endcapped) C18 columns

Model does not account for ionic interactions and intramolecular H bonding

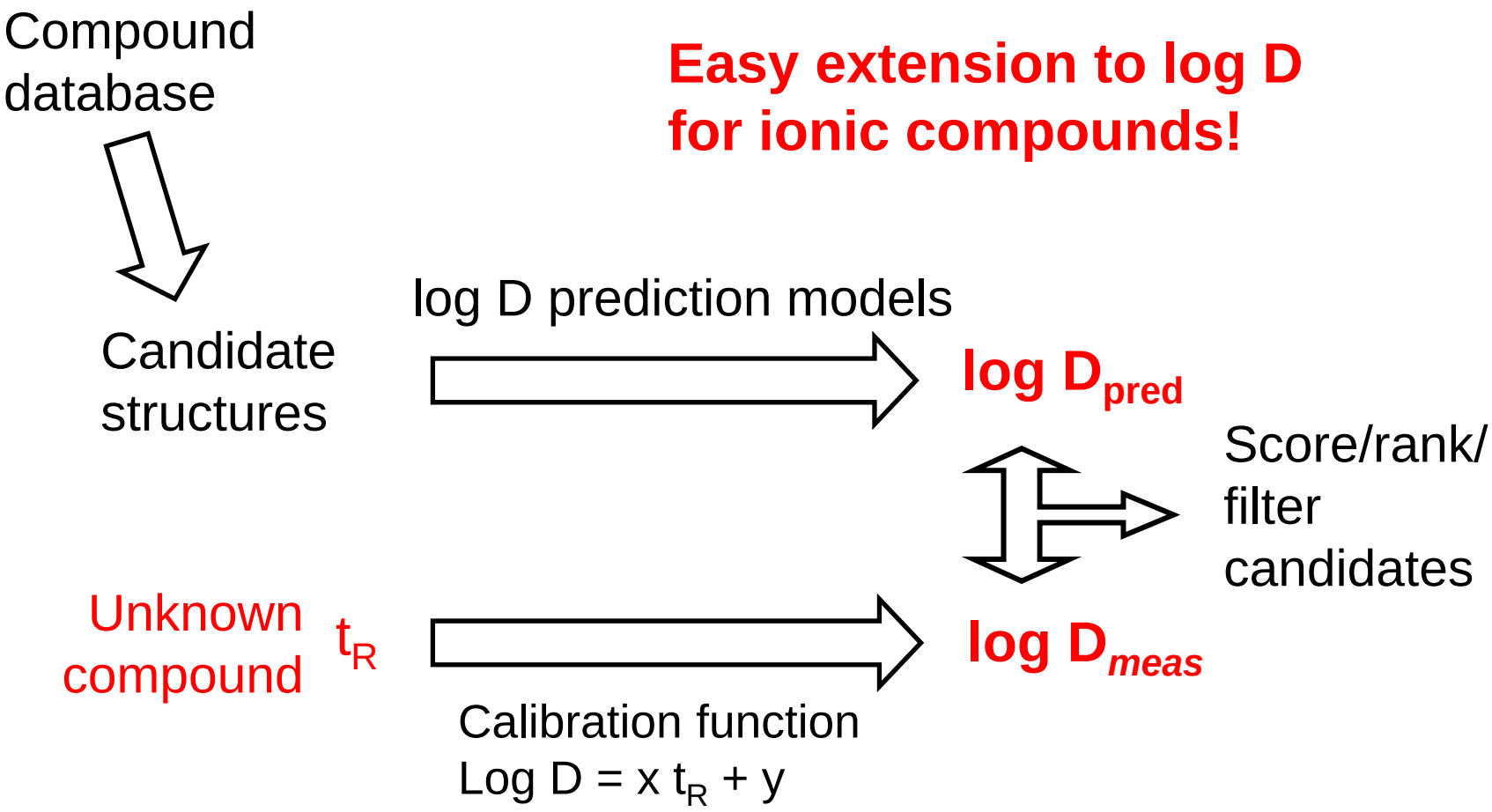
CHI: Retention prediction by LSER



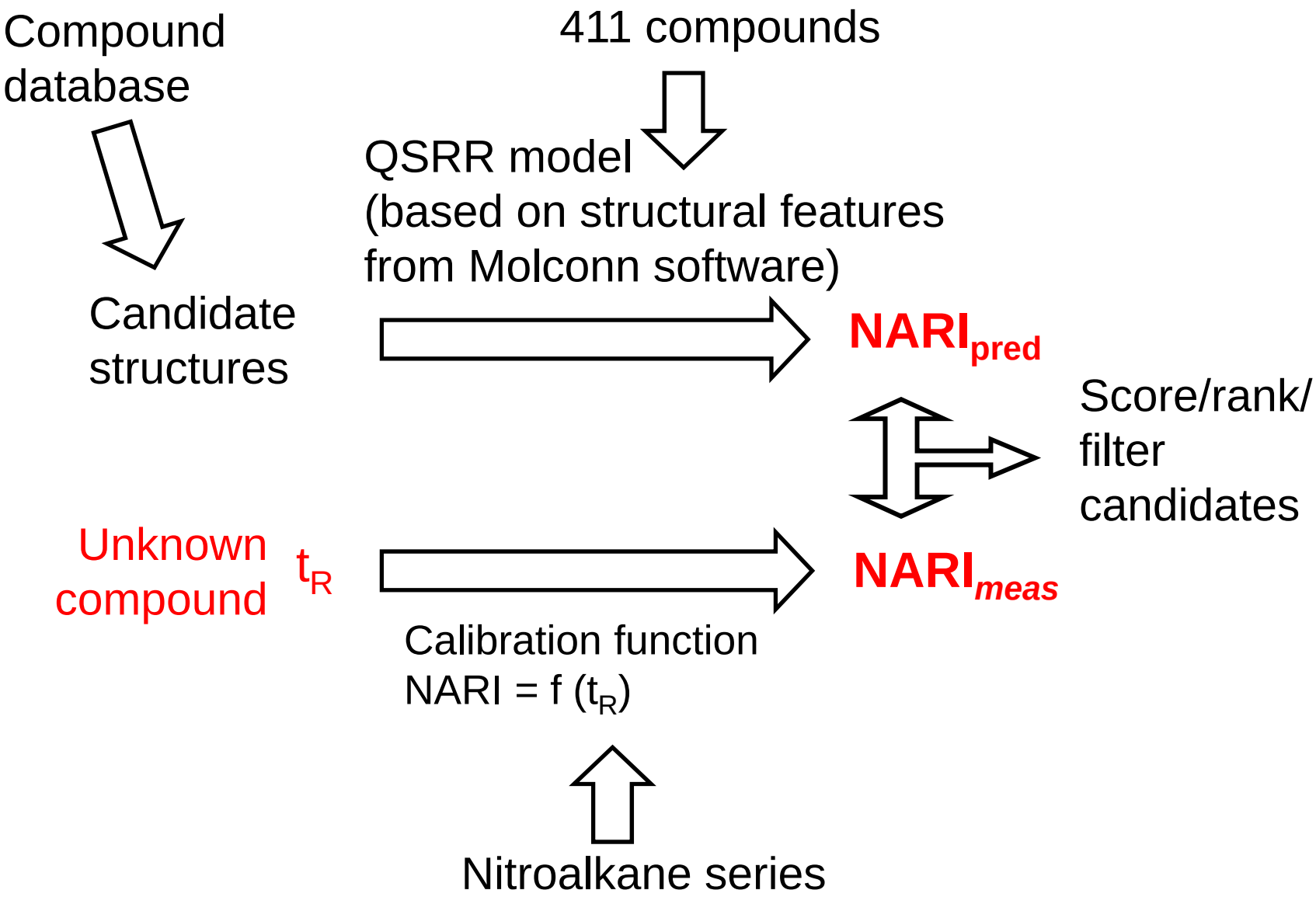
RTI: The beauty of simplicity



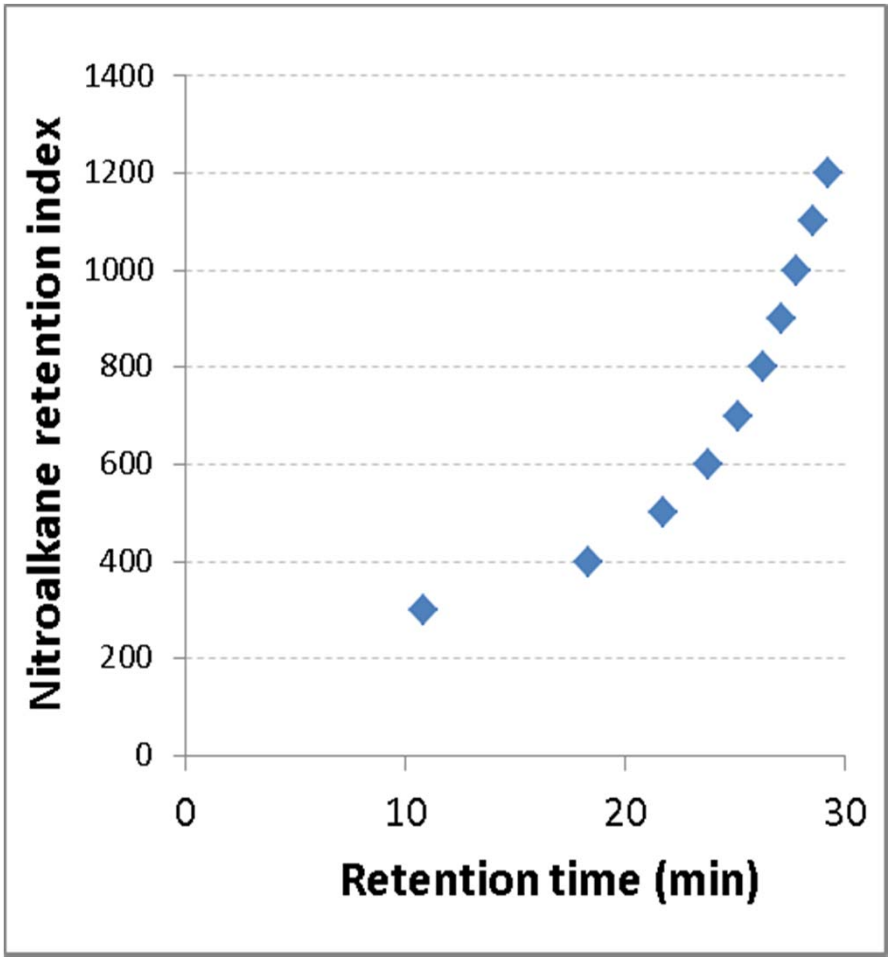
RTI: The beauty of simplicity



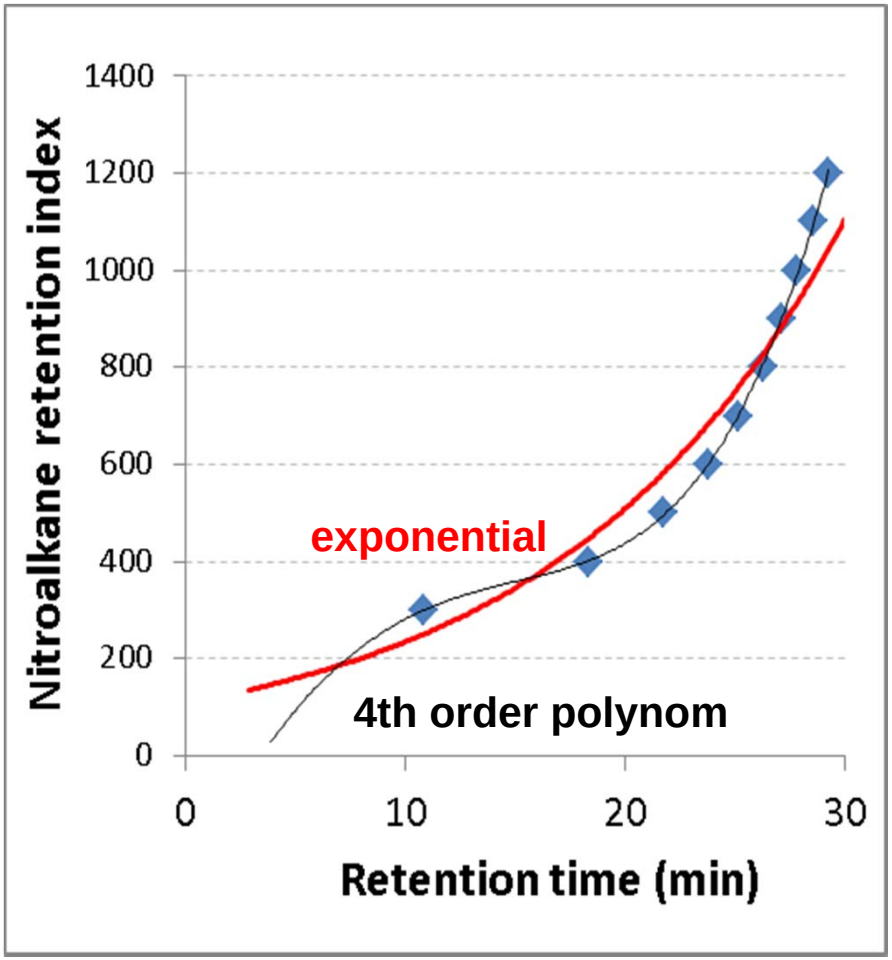
NARI: Kovats' for LC?



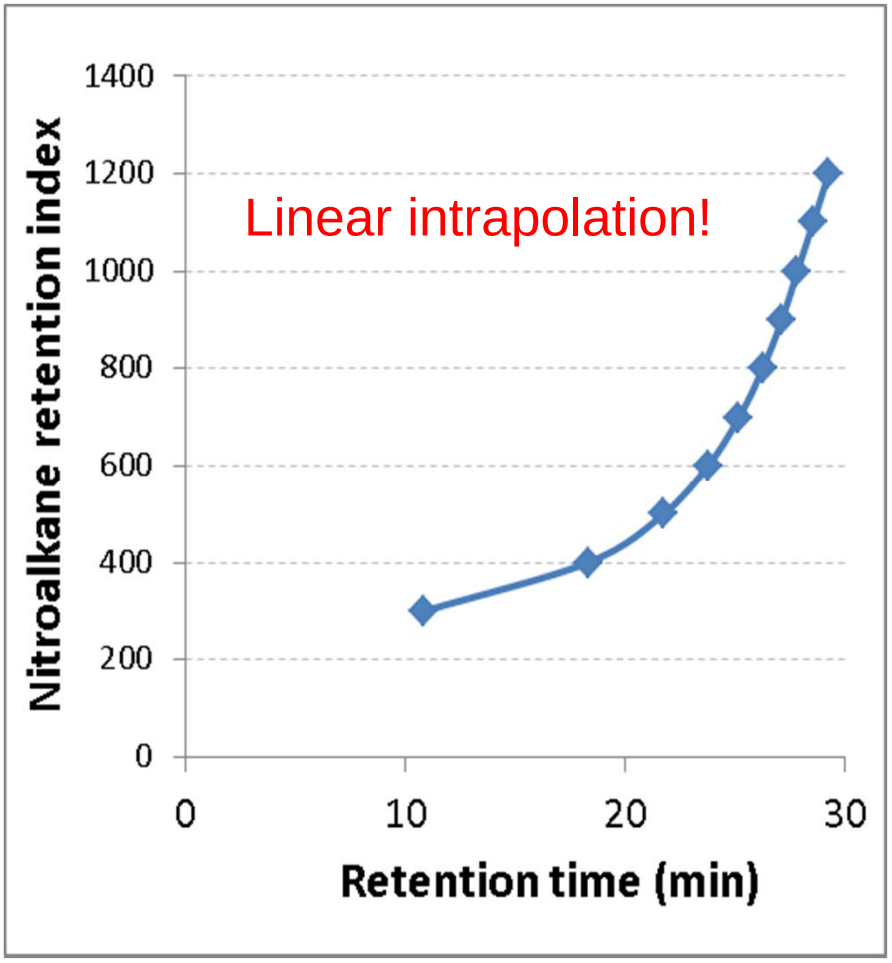
LC method calibration issues

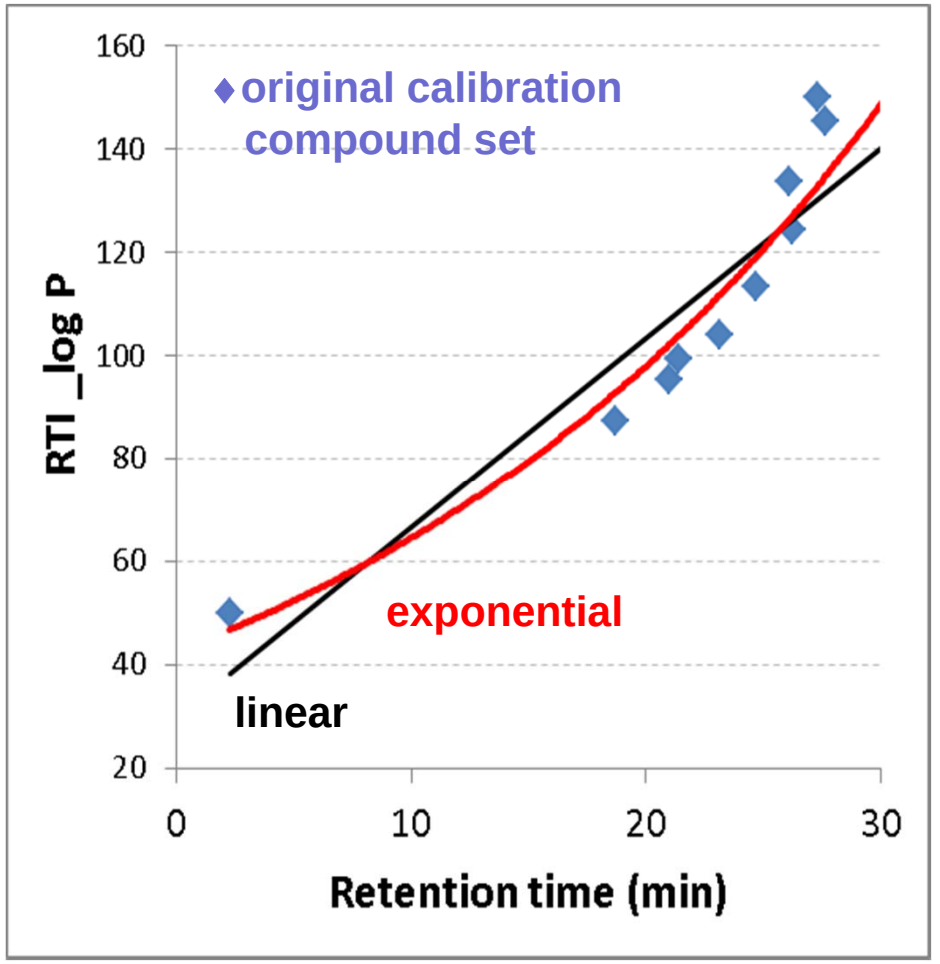


LC method calibration issues



LC method calibration issues

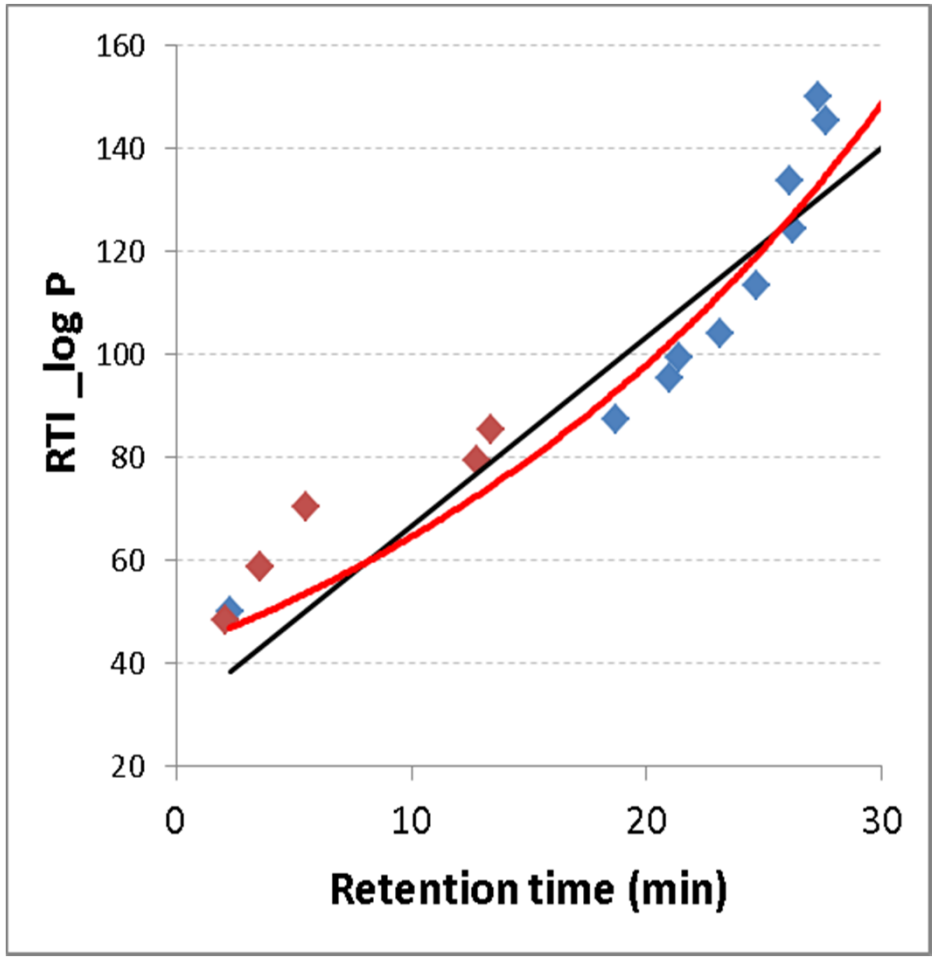




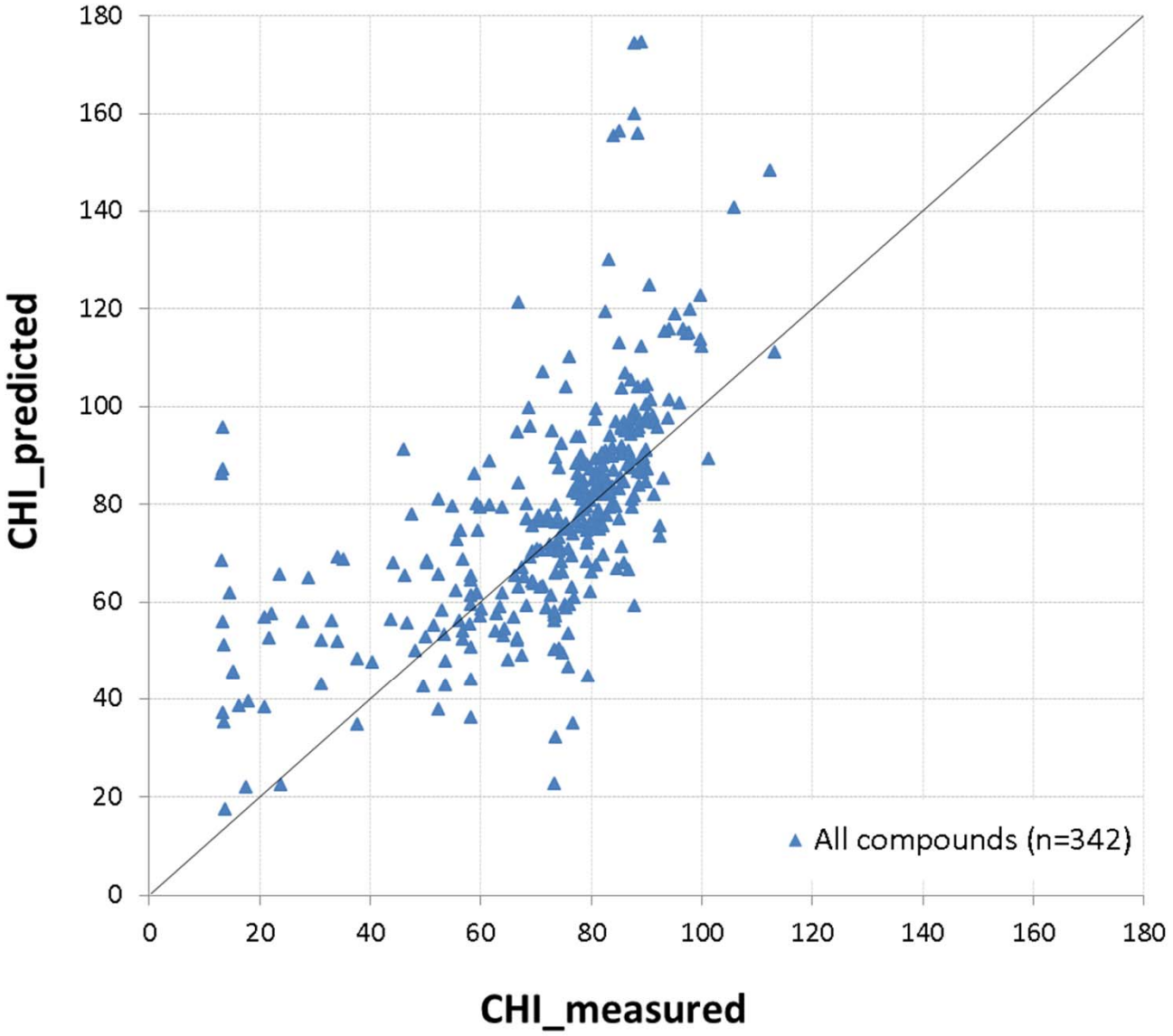
LC method calibration issues

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH – UFZ

Finding additional calibration compounds might be difficult…



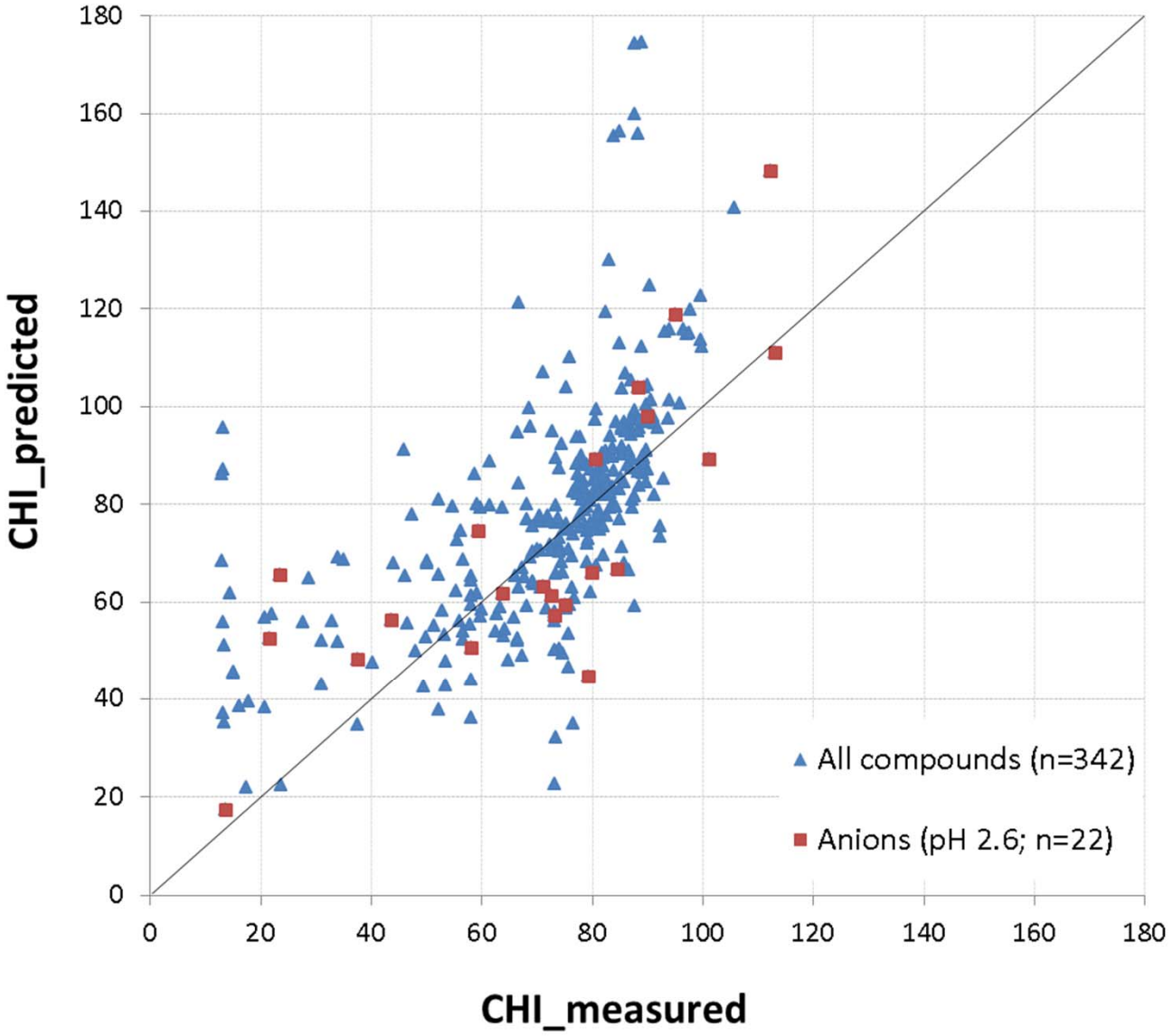
CHI model



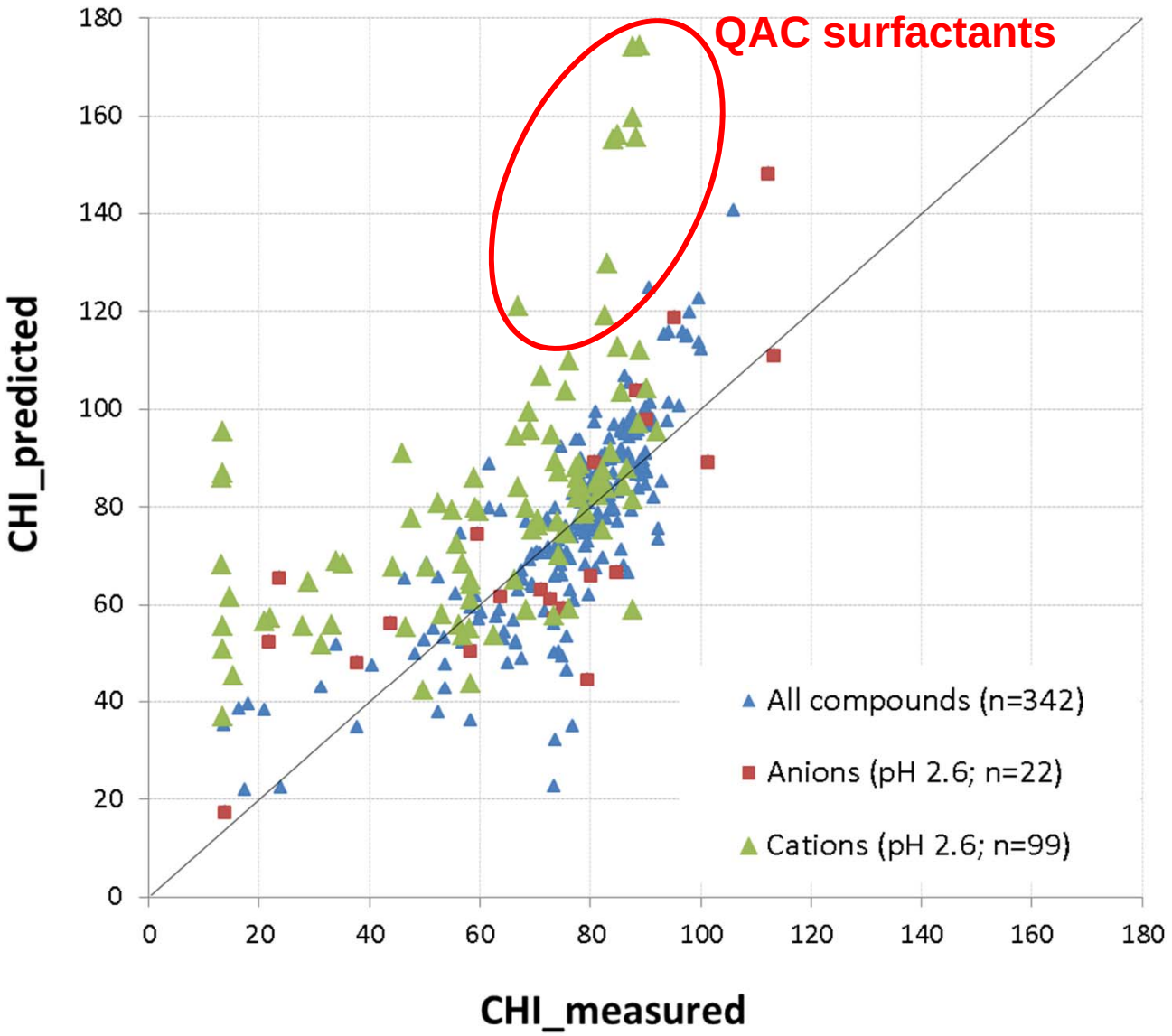
MAE 15.9

R^2 0.43

CHI model


CHI_predicted
CHI_measured
0
20
40
60
80
100
120
140
160
180
All compounds (n=342)
Anions (pH 2.6; n=22)

CHI model



Neutrals

MAE 11.9
(≈ 3.9 min)
R^2 0.63

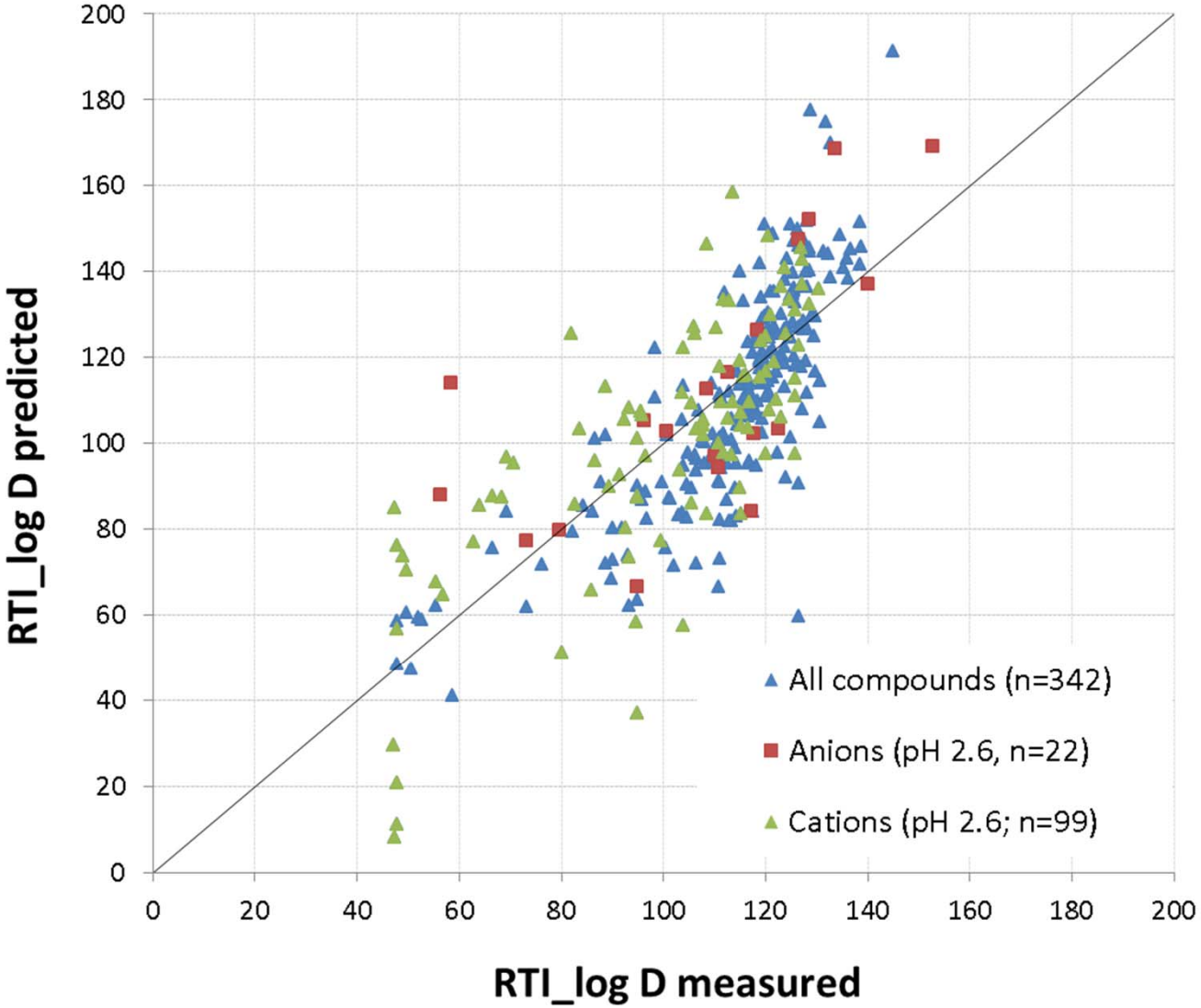
Anions

MAE 15.8
(≈ 5.1 min)

Cations

MAE 24.9
(≈ 8.1 min)

RTI model



Neutrals

MAE 12.3
(≈ 3.9 min)
R^2 0.63

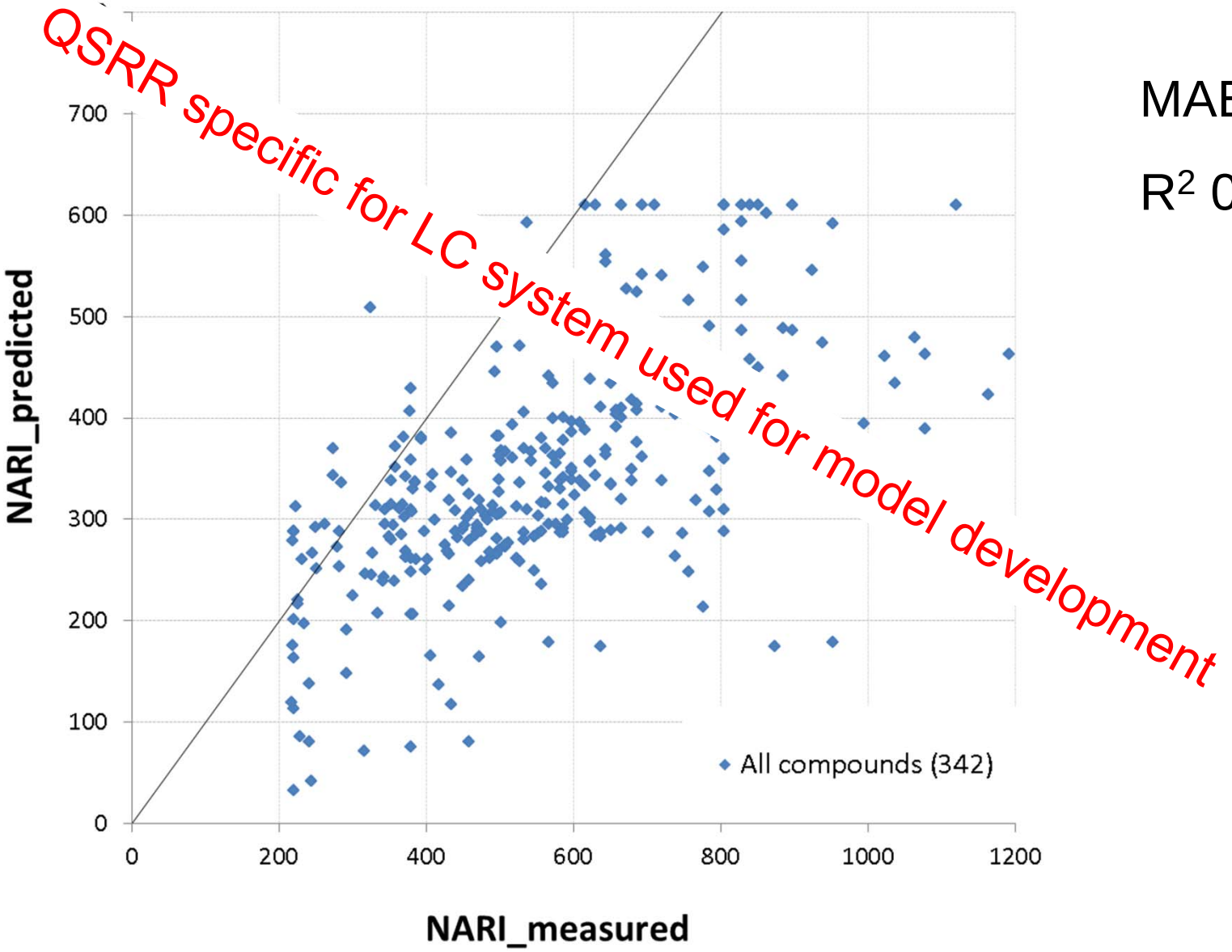
Anions

MAE 20.8
(≈ 6.6 min)

Cations

MAE 16.4
(≈ 5.2 min)

NARI model



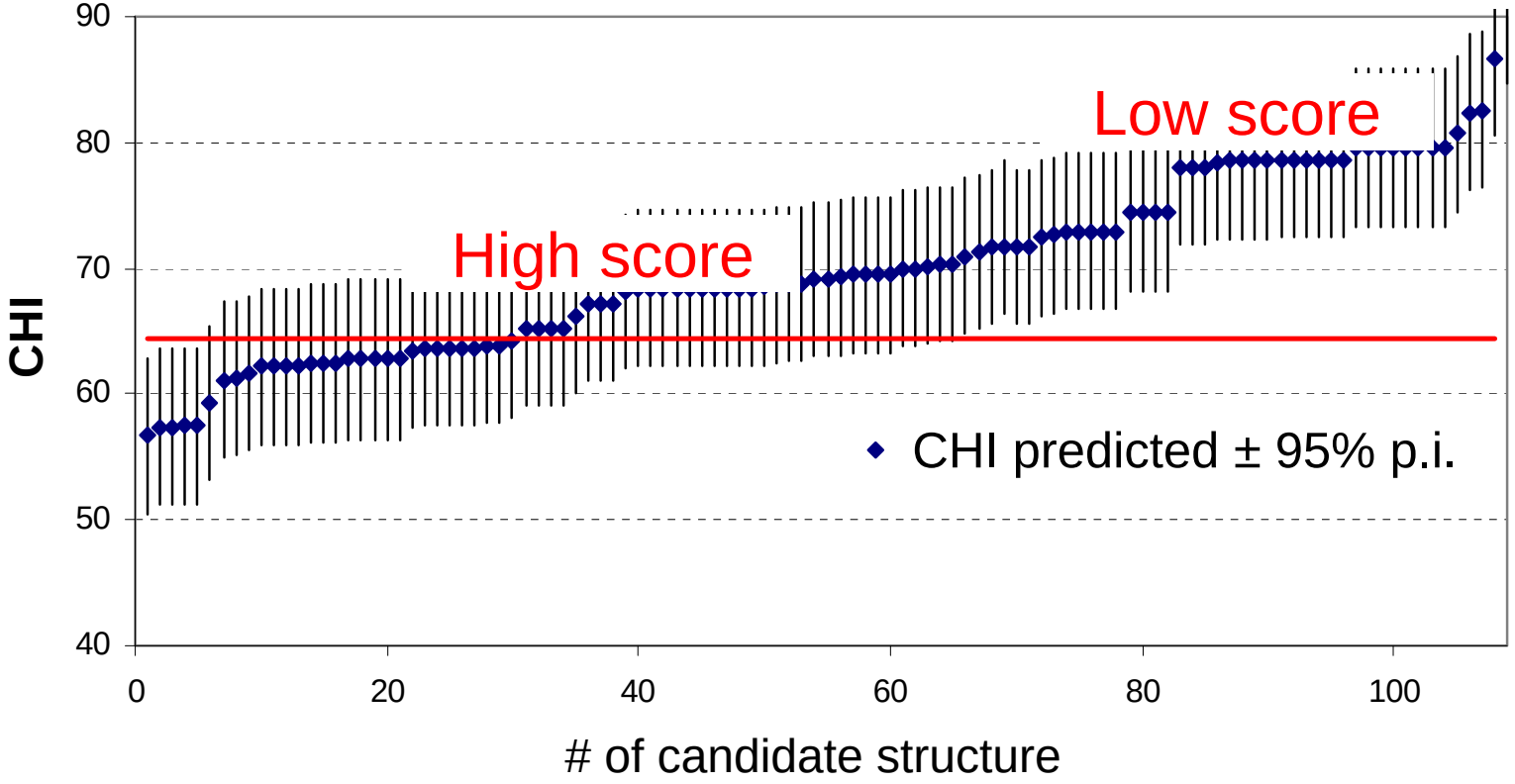
MAE 251

R^2 0.36

Candidate selection: not all about accuracy

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

Unknown t_R 19.5 min ⇒ CHI 64.4



Conclusions

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH - UFZ

Retention prediction models

- Are often far from accurate (hydrophilic, ionic& tricky compounds)
- Nevertheless provide useful information for confirmation of suspects or candidate exclusion
- A conceptually correct model might not be better than an analogy model

⇒ Be aware of compound domains the model is applicable

⇒ Be aware of prediction uncertainty

We have to develop some guidance on that!