LEADER IN ADME/PK-PREDICTION SOFTWARE

Successfully validated software for prediction, simulation and optimization of human clinical ADME/PK with accuracy superior to laboratory methods

- ✓ Get accurate predictions directly from molecular structure
- ✓ Broad mechanistic coverage
- ✓ Reduce uncertainty and risks
- ✓ Increase data production and learning
- ✓ Add missing data and sanity-check available lab data
- ✓ Frontload and improve decision-making
- ✓ Save costs and time
- ✓ Reduce the use of animals and chemicals
- ✓ Get help from dedicated expertise

Contact:
Urban Fagerholm, CEO and Associate Professor at Prosilico AB
urban.fagerholm@prosilico.com; +46-70-1731302

Introduction

Replacement and reduction of animal studies (3R)

The goal is to reduce and replace animal studies/methods, optimally with computational (*in silico*) methods with better performance [EU directives; 1-4].

ADME/PK

ADME/PK is the Absorption, Distribution, Metabolism, Excretion/ Pharmacokinetics of compounds, such as drugs, drug candidates and other chemicals we are exposed to. ADME/PK determines their exposures, and thereby, influences their effect/side effect-profiles.

Laboratory methods and their limitations

Data from animal are being scaled to humans. Animal models have been, and still are, the golden standard for prediction of human clinical ADME/PK of drug candidates. It is estimated that such experiments take many months and cost >30,000-50,000 USD per substance in Drug Discovery & Development.

Cell models (*in vitro* models) are cost- and time-efficient alternatives. They are generally less accurate and have more narrow application domain (every other compound seem non-quantifiable/non-predictable) [5]. *In vitro* experiments take days to weeks and cost > 5,000-10,000 USD per substance (including compound synthesis and development of analytical methods).

Extreme maximum errors (from 1,000- to 1,000,000-fold) and systematic errors are major drawbacks with laboratory methods [6,7]. This implies potential safety risks in early clinical studies and failures. Laboratory variability (on average ca 3- to 10-fold, and maximum between ca 50- to 5000-fold) and retrospective cherry-picking of favorable results are other problems [7,8].

We have approximated that within the pharmaceutical industry at least 1,300 MUSD is spent each year on global ADME/PK screens and predictions.

***In silico* methods**

In silico methods, developed based on laboratory or clinical data, have also been developed by many, but results have overall not reached those of laboratory methods [9]. An important feature is valid compound-specific confidence intervals. This is generally lacking, which limits the interpretation of predictions for compounds of new and unknown chemical space. Other features that are normally missing are visual guidance how to optimize ADME/PK-characteristics, security solutions (to avoid disclosure of proprietary molecular structures) and blind external validation of model performance.

About PROSILICO

We are motivated by a strong and sincere willingness to drive continuous improvement and renewal within ADME/PK science. We believe that a technology shift and significant quality improvements are within reach. Together we form a group of highly merited scientists with a long experience from the pharmaceutical industry, university and IT-development.

PROSILICO was founded in 2014 by ex-AstraZeneca scientists Urban Fagerholm (CEO, model developer) and Sven Hellberg (vice CEO, model developer, head of data analysis). Later, Ola Spjuth (model developer) joined the company.

Urban Fagerholm, PhD and Associate Professor in Pharmacokinetics and Biopharmaceutics, and Clinical Psychologist, has worked more than 30 years in the pharmaceutical industry (preclinical and clinical pharmacokineticist at AstraZeneca in Södertälje, Sweden) and academia (Uppsala university), with focus on ADME/PK-prediction methodology and -mechanisms (particularly permeability). He has published ca 50 scientific papers and book chapters and 2 books. In 2009, he was (by International Biographical Centre, Cambridge, UK) nominated International Educator of the Year, elected to 2000 Outstanding Intellectuals of the 21th Century, and named member of Top 100 Professionals of the Year.

Sven Hellberg, PhD and Computational Chemist has worked more than 40 years in the pharmaceutical industry (incl. ex-manager in Pharmaceutical R&D at AstraZeneca in Södertälje, Sweden). Sven is highly experienced with 30+ years in QSAR, compound design and drug discovery, with skills and experiences including molecular modeling, pharmacophores, docking, structure-based design, fragment evolution, hit finding, lead identification, lead optimization, chemometrics, machine learning and pattern recognition. He has ca 50 scientific journal publications, 18 conference publications and 86 patent publications.

Ola Spjuth, PhD and Professor in Pharmaceutical Bioinformatics at Uppsala University, Sweden, has long experience from AI and machine learning and large-scale data analysis in computational pharmacology and toxicology. He has developed methods and tools to improve predictions of metabolism, ADME/PK, safety/toxicology, targets, MoA etc. Ola has a specific interest in confidence predictors and in particular Conformal Prediction for providing valid measures of a specific prediction's confidence. He has authored 100+ scientific publications.

PROSILICO has over 40 national and international customers, including major pharmaceutical companies and customers in the US, Japan, UK, Netherlands, France and Switzerland.

New Software - ANDROMEDA by Prosilico

This white paper describes PROSILICO's new, unique software for prediction, simulation and optimization of human clinical ADME/PK – ANDROMEDA by Prosilico. The software, available as SaaS (Software as a Service) and OnPrem (intranet installation), enables the following:

- Instant predictions of human clinical ADME/PK, exposures and doses of drugs, drug candidates, metabolites and other chemicals directly from molecular structure (one compound at a time or batches of compounds), with shown confidence intervals.
- Highlighting of major ADME/PK-obstacles and -mechanisms.
- Optimization of ADME/PK properties via visual guidance.
- Production of ADME/PK (mechanisms and estimates) for modeling of clinical data.
- Combination of *in silico* results and laboratory data for predictions and simulations.
- Replacement of comparative animal and *in vitro* studies and models, with improved predictive accuracy and breadth → reduced costs, time, uncertainty, environmental exposure to chemicals, and risks for development failures and risks for humans, increased data production and learning, and frontloaded and improved of decision-making).

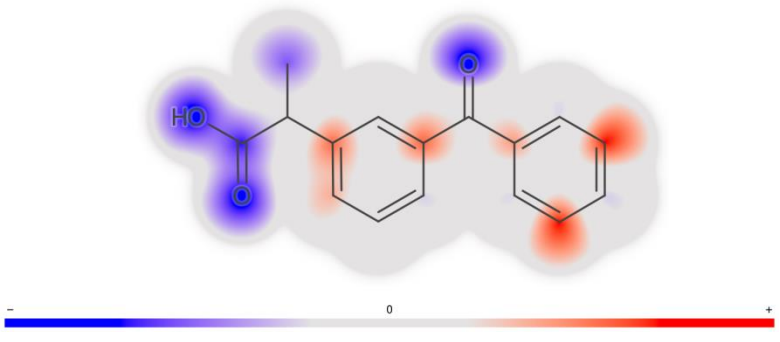
Features

Accurate predictions directly from molecular structure (instant results)	Optimization human clinical ADME/PK via visual guidance
Prediction of 30 human clinical ADME/PK-parameters (including e.g. *in vivo* dissolution, fraction absorbed, renal and biliary excretion, clearance, gut-wall metabolism, oral bioavailability, half-life, colonic uptake, extended-release potential, brain binding, BBB-permeability, absorption rate, CYPs, transporters, blood-to-plasma conc. ratio), exposure profiles and dose	Prediction of human clinical ADME/PK for batches of compounds
Prediction of human clinical ADME/PK for compounds problematic and out of reach for labs	Simulation of human clinical ADME/PK
Prediction of human clinical ADME/PK with valid compound-specific confidence intervals	Combination of *in silico* and laboratory data for human clinical ADME/PK-predictions
Available both as an online web-based product and local installation (e.g. for company intranet) with higher security	Main application domain – MW 100 to 700 g/mole; log D -8 to 8. Also useful for compounds with MW up to 1000-1500 Da, including macrocycles and PROTACs

The Technology

- **Unique databank** of curated, quality-checked human clinical ADME/PK-data, on which separate *in silico* models are developed.

- ***In silico* models developed using Conformal Prediction (CP)**, which is an innovative methodology for augmenting machine learning predictions with a valid confidence prediction [10].
  In contrast to standard machine learning methods that produces point estimates (such as a specific class for classification models or a real value for regression models), CP produces prediction intervals at a given confidence level provided by the user, and the size of the interval is dependent on how "strange" (nonconforming) the test compound is compared to the training compounds, but also the desired confidence of the prediction and the overall *efficiency* of the predictor owing to the training data. PROSILICO's CP-models are developed using **CPSign** by Aros Bio AB [11], which also produces colored signatures for molecular regions contributing to decreasing and increasing each ADME/PK-parameter (Figure 1).



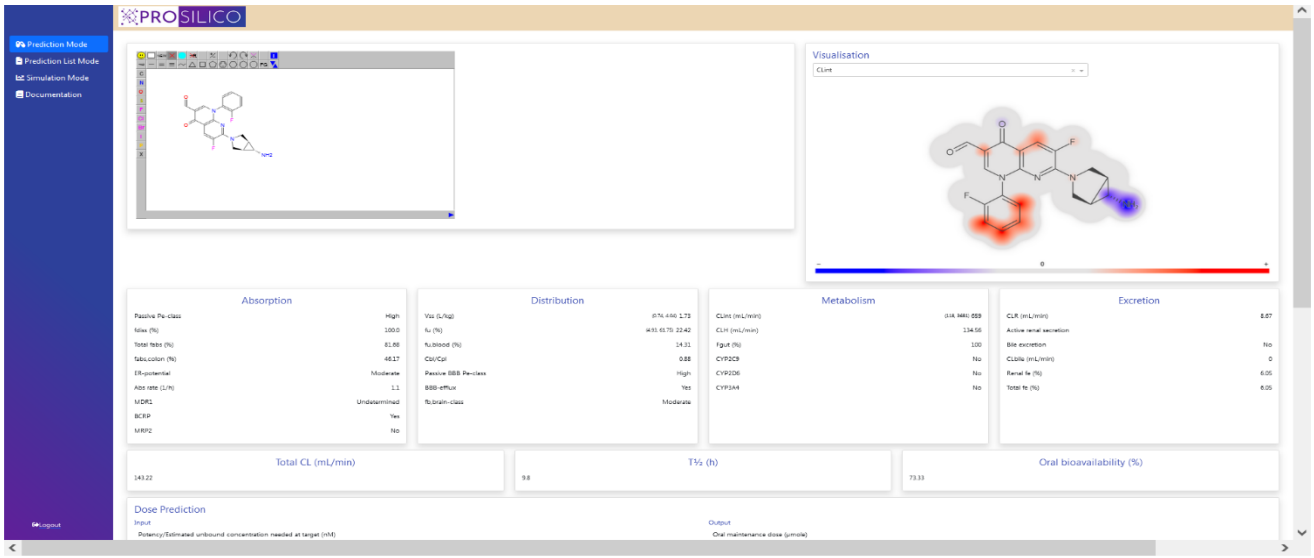
V_{ss} = 0.18 (0.07; 0.43) L/kg

Figure 1 Molecular structure and signatures for ketoprofen. Molecular regions contributing to decreasing (blue) and increasing (red) the ADME/PK-parameter volume of distribution (V_{ss}) are shown. The predicted V_{ss}-estimate and 70 % confidence limits (within parentheses) are shown to the right.

- **Integration of new, unique *in silico* models, algorithms and physiologically-based pharmacokinetic (PBPK) system** to generate prediction results for a wide set of primary and secondary ADME/PK-parameters in man for various types of small size chemicals (main molecular domain 100-700 g/mole), including drugs, drug candidates, metabolites.

- **The software ANDROMEDA by Prosilico** (available as SaaS (cloud/web) and OnPrem (intranet-installation) is based on the new, unique prediction models, algorithms and PBPK-model. Molecular structures are either imported as SMILES (one by one or as a list) or sketched in its molceular editor, and prediction results are generated instantly.

Presentation of Prediction Mode and Simulation Mode in ANDROMEDA by Prosilico



Figures 2 and 3 Prediction Mode (for *in silico* prediction and optimization of ADME/PK for single compounds) and Simulation Mode (for prediction and simulation of single compounds, including the possibility to incorporate lab data). The software also has a Prediction List Mode for prediction of ADME/PK for batches of compounds.

Performance & Validity

PROSILICO's prediction models have undergone full internal validations, head-to-head-comparisons *vs* laboratory methods and competitor *in silico* models, blind external validations and peer-reviews. Presented results for major ADME/PK parameters are from true prospective predictions (not from retrospective fits or from training sets).

Oral bioavailability

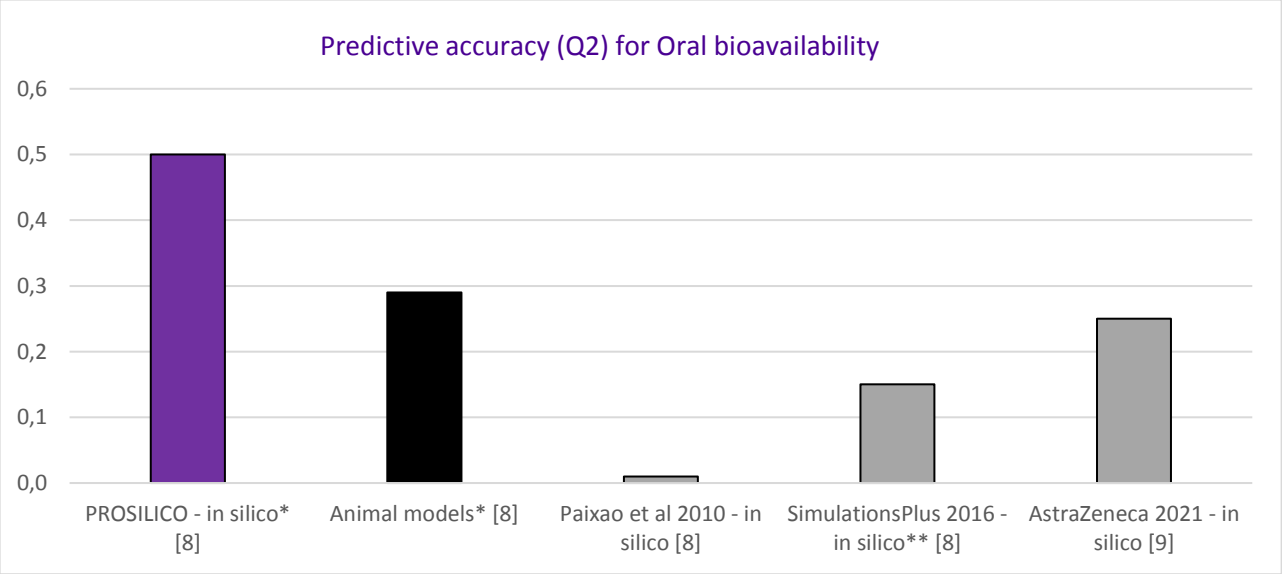


Figure 4 Predictive accuracy (Q^2) of different prediction models (including 3 competitors) for oral bioavailability in man. *Head-to-head comparison. **Compounds with limited permeability and solubility were excluded beforehand by the authors.

Volume of distribution

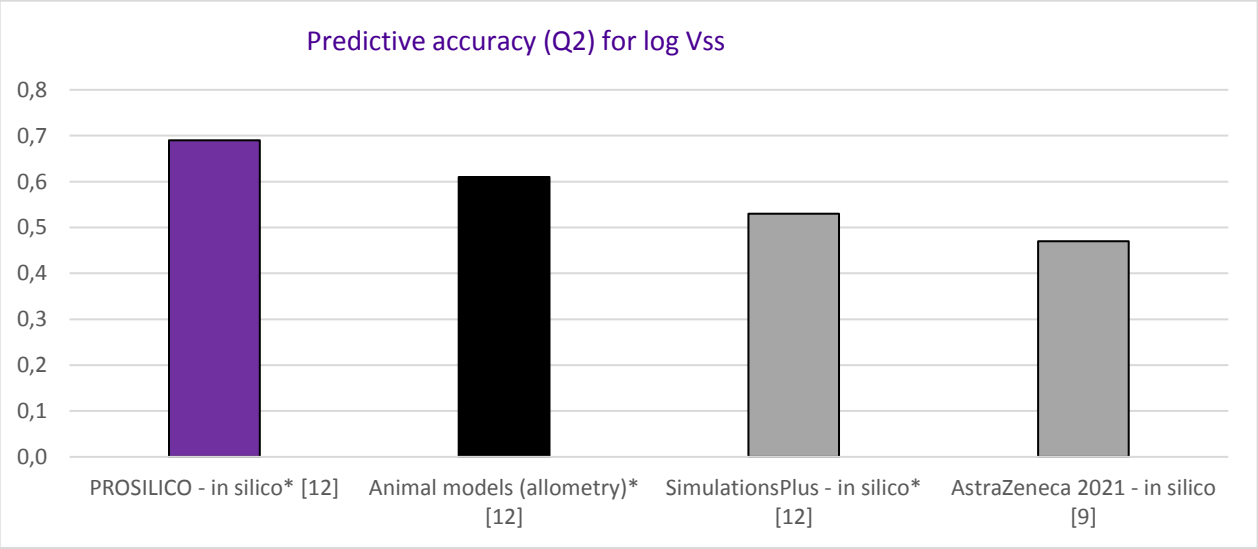


Figure 5 Predictive accuracy (Q^2) of different prediction models (including 2 competitors) for log steady-state volume of distribution (log V_{ss}) in man. *Head-to-head comparison.

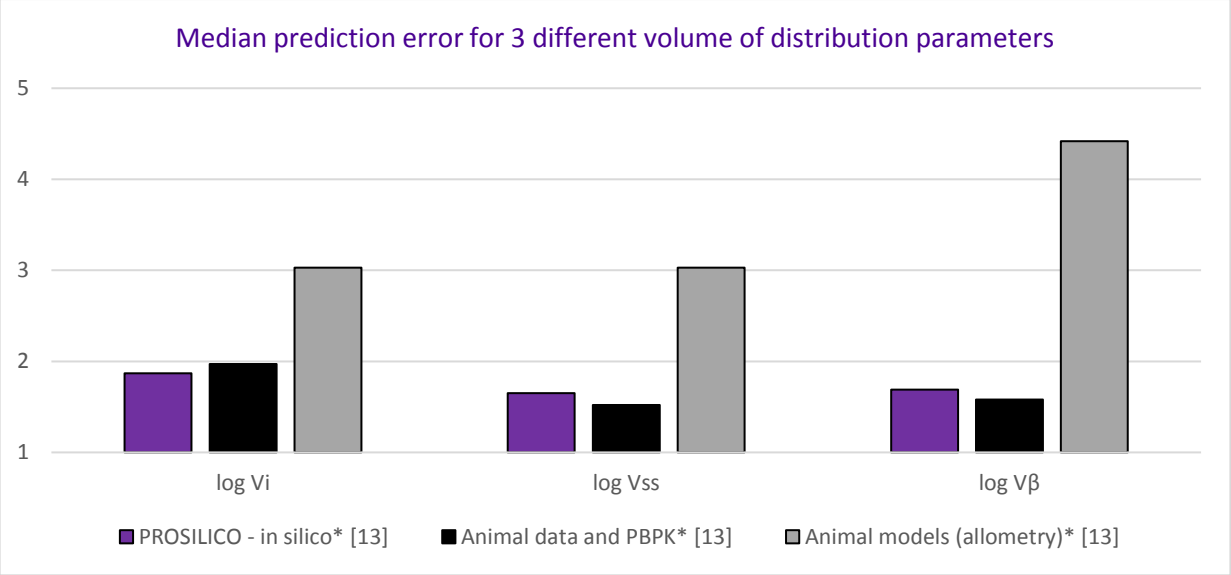


Figure 6 Predictive accuracy (Q^2) of different prediction models for different log volume of distribution parameters (initial, steady-state and terminal log V_{ss}) in man. *Head-to-head comparison.

Intrinsic metabolic clearance

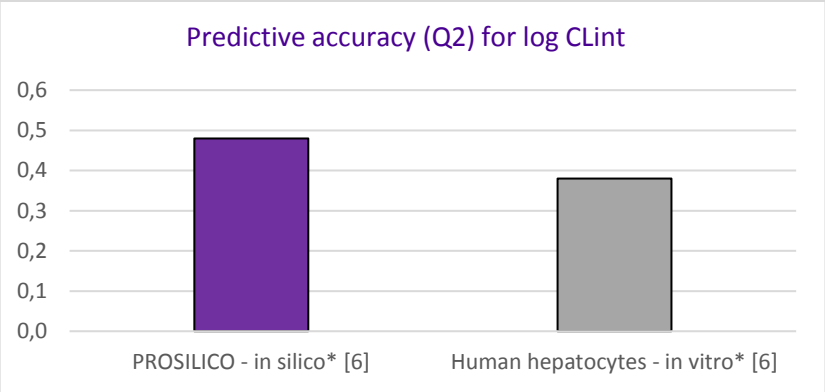


Figure 7 Predictive accuracy (Q^2) of two prediction models for log intrinsic clearance (log CL_{int}) in man. *Head-to-head comparison. The *in silico* method also predicted CL_{int} for compounds with non-quantifiable hepatocyte CL_{int} (about every other compound) well.

Unbound fraction in plasma *vs* laboratory variability

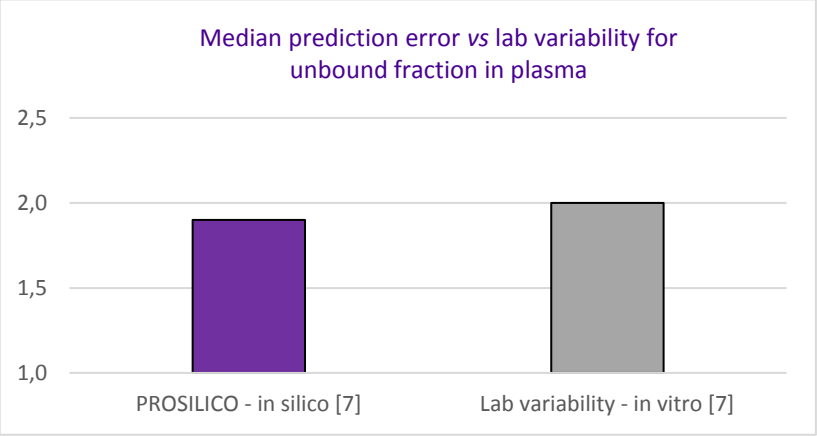


Figure 8 Median prediction error for *in silico vs* median variability at labs for unbound fraction in human plasma (f_u).

In vivo dissolution potential (maximum fraction of oral dose dissolved in GI fluids *in vivo*)

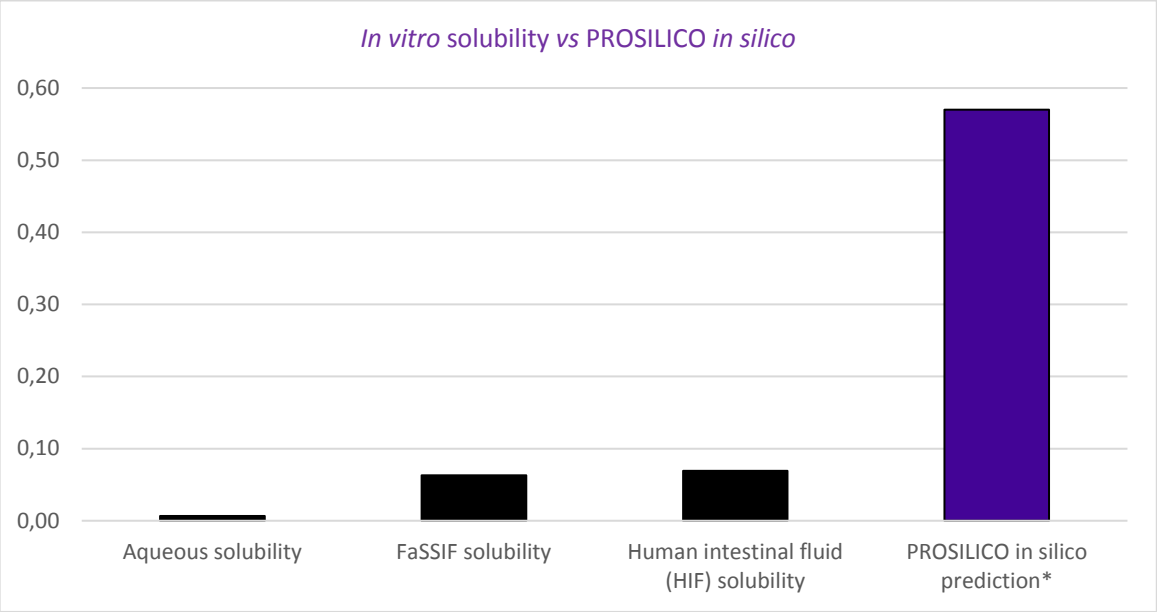


Figure 9 Correlation/fit (R^2) for *in vitro* methods *vs* predictive accuracy (Q^2) of *in silico* method.

Other parameters

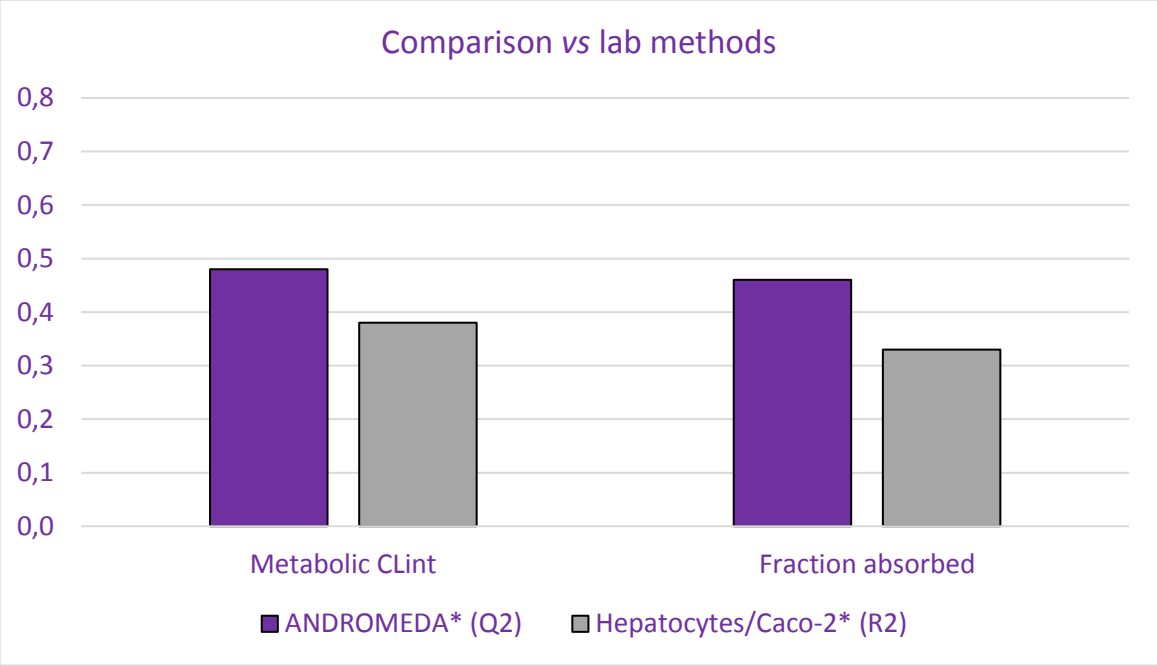


Figure 10 Correlation/fit (R^2) for *in vitro* methods (human hepatocytes [7] and Caco-2 [15]) *vs* predictive accuracy (Q^2) of *in silico* methods.

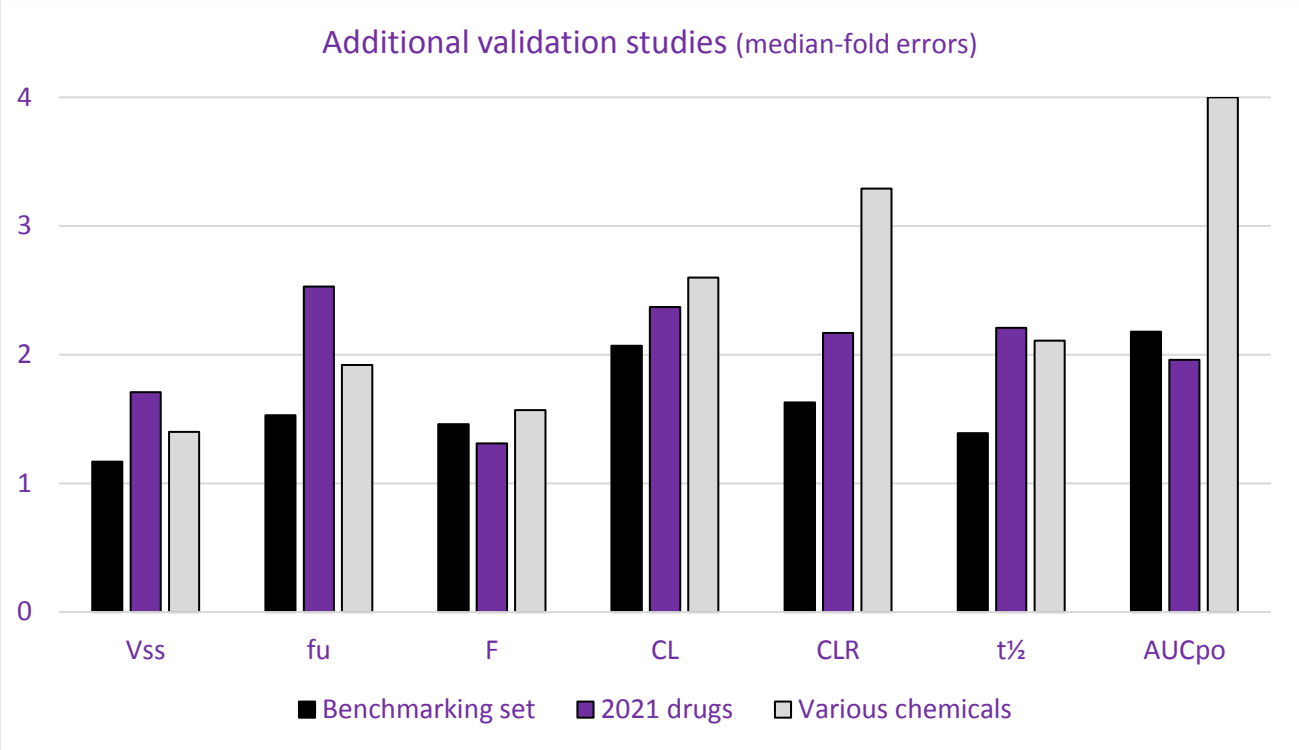


Figure 11 Median prediction error for *in silico* for a a) benchmarking set of 24 physico-chemically different compounds (log P -2.0 to 5.3, log D -5.0 to 4.8, 0 to 8 hydrogen bond donors, 1 to 14 hydrogen bond acceptors, polar surface area 8 to 246 Å^2; taken from [14]) [15], b) new small drugs marketed 2021 (n=28) [16 and various chemicals (n=65) [17].

The challenge with modern small drugs (multimechanistical PK) – ANDROMEDA covers and predicts them

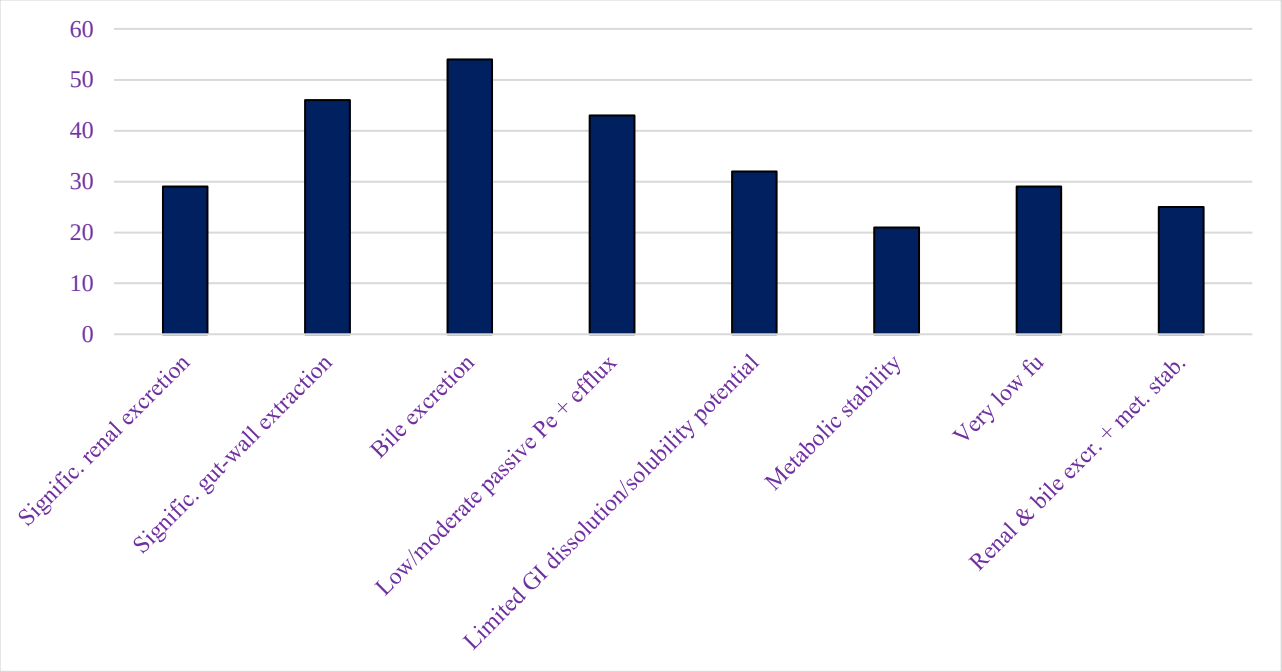
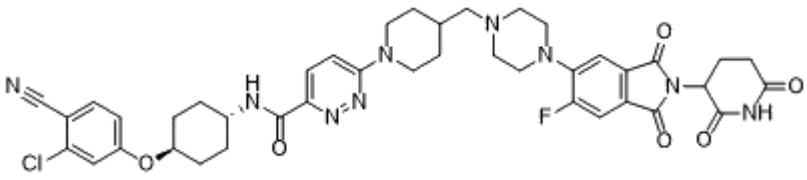


Figure 12 Percentages of new small drugs marketed in 2021 with certain predicted PK-characteristics [16].

***In silico* predictions of the gastrointestinal uptake of macrocycles in man using conformal prediction methodology** [18] Fagerholm, Hellberg, Alvarsson and Spjuth

The gastrointestinal uptake of macrocyclic compounds is not fully understood. Here we applied our previously validated integrated system based on machine learning and conformal prediction to predict the passive fraction absorbed (f_a), maximum fraction dissolved (f_{diss}), substrate specificities for major efflux transporters and total fraction absorbed ($f_{a,tot}$) for a selected set of designed macrocyclic compounds (n=37; MW 407-889 g/mol) and macrocyclic drugs (n=16; MW 734-1203 g/mole) *in vivo* in man. Major aims were to increase the understanding of oral absorption of macrocycles and further validate our methodology. We predicted designed macrocycles to have high f_a and low to high f_{diss} and $f_{a,tot}$, and average estimates were higher than for the larger macrocyclic drugs. With few exceptions, compounds were predicted to be effluxed and well absorbed. A 2-fold median prediction error for $f_{a,tot}$ was achieved for macrocycles (validation set). Advantages with our methodology include that it enables predictions for macrocycles with low permeability, Caco-2 recovery and solubility (BCS IV), and provides prediction intervals and guides optimization of absorption. The understanding of oral absorption of macrocycles was increased and the methodology was validated for prediction of the uptake of macrocycles in man.

Prediction of absorption of PROTAC ARV-110 in man



Passive permeability	Moderate	
Fraction dissolved	65	%
Fraction absorbed	27	%
Efflux	Yes	
CL/F	810	mL/min
Observed CL/F	343	mL/min
Oral bioavailability	20	%
Observed -"- in rodents	24-38	%

Additional points

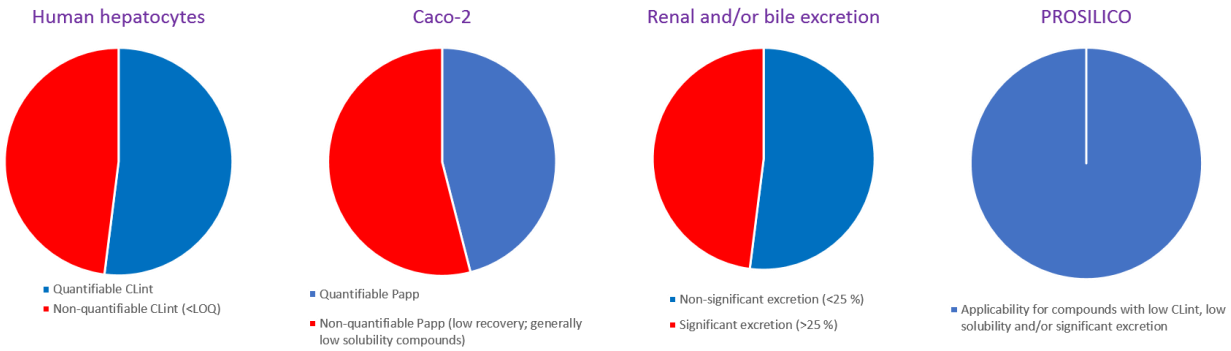


Figure 12 Application domains of *in vitro* methods [19,20] *vs* PROSILICO's *in silico* platform.

- Accuracy on par with or superior to comparable laboratory methods.
- Broader application domain than *in vitro* methods (e.g. ca 100- to 1000-fold lower limit of quantification compared to human hepatocytes and applicability for compounds with very low solubility and Caco-2 recovery and with significant excretion and efflux) (Figure 12).
- Overall, <2- to 3-fold median prediction errors in blind external validations by minor to major international pharmaceutical companies (>150 compounds), which is significantly below errors at labs [5].
- Markedly lower maximum errors compared to laboratory methods (implies improved safety in early clinical trials and reduced failure risk in drug discovery and development).
- Successful predictions of ADME/PK and exposure profiles in first-time-in-man studies.

Authority approval

ANDROMEDA by Prosilico was used as major source of preclinical ADME/PK in a CTA that was approved by German authority BfArM.

Customer quotations

"Very impressed!"
"Using an in silico system superior to IVIVE methods is highly interesting"
"Easy to understand and use"
"Recommended!"
"Convincing results!"
"The opportunity to bypass the limitations of in vitro methodologies is very attractive."
"Not only does it save money, it – even more importantly – helps us reduce animal testing when we can rely on it."
"Much faster and easier to use than competitor simulation software and no input data are needed"
"The data are impressive"
"Impressed that it predicts so many important PK-parameters"
"Good visualization"
"The prediction errors seem to be minimal"

References

1. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:SV:PDF
2. https://english.ncadierproevenbeleid.nl/latest/news/16/12/15/ncad-opinion-transition-to-non-animal-research
3. https://www.djurensratt.se/blogg/historiskt-beslut-i-eu-parlamentet-strategi-utfasning-av-djurforsoken-pa-vag
4. https://www.tidningencurie.se/debatt/brist-pa-riktlinjer-hindrar-alternativa-metoder-att-sla-igenom/
5. Poulin P et al. J Pharm Sci. 100:4127-4157, 2011.
6. Fagerholm et al. J Pharm Sci. July 2022.
7. Fagerholm et al. Xenobiotica 51:1095-1100, 2021.
8. Fagerholm et al. Molecules 26:2572, 2021.
9. Miljković F et al. Mol Pharmaceut. 10, 2021.
10. Alvarsson et al. J Pharm Sci. 110:42-49, 2021.
11. Aros Bio AB, 2021.
12. Fagerholm et al. Xenobiotica, Dec 2021.
13. Fagerholm et al. In manuscript
14. Sköld et al. J Med Chem. 49:6660-6671, 2006.
15. Fagerholm et al. Submitted, 2022.
16. Fagerholm et al. Submitted, 2022.
17. Fagerholm et al. Xenobiotica, 2022.
18. Fagerholm et al. J Pharm Sci 2022.
19. Stringer et al. Xenobiotica, 38:1313-1329, 2008.
20. Skolnik et al J Pharm Sci. 99:3246-65, 2010.