A BOILED-Egg to Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules

Cheminformatics Workshop – EPFL
September 9, 2016

Antoine Daina & Vincent Zoete
Molecular Modeling Group
SIB Swiss Institute of Bioinformatics, Lausanne – Switzerland
The fate of an oral drug in the body

Oral dosage

Liberation

Active ingredient in solution

Absorption

Distribution

Metabolism

Excretion

Pharmacokinetics (ADME)

Pharmacodynamics

Interaction with the therapeutic target (protein)

Therapeutic effect

Active ingredient in solution
Clinical impact of ADME

**In vitro:**
- dose $\rightarrow$ concentration $\rightarrow$ effect

**In vivo:**
- dose $\rightarrow$ ADME $\rightarrow$ concentration $\rightarrow$ effect
Failure for oral drugs in clinical development

- Estimation of ADME at early steps of drug discovery for PK profiles
- Development of computer models to predict ADME from chemical structures

1964 - 1985 (phases 2/3)
R.A. Prentis et al. Br. J. Clin. Pharm. 1988
T. Kennedy Drug Discov. Today 1997

2000 - 2010 (phases 2/3)
M.J. Waring et al. Nat. Rev. Drug Discov. 2015
J. Arrowsmith & P. Miller Nat. Rev. Drug Discov. 2013
Physicochemical parameters and passive absorption

- **The Egan Egg**
- **Great practical utility**
  - 500+ citations\(^1,2\)
- **Hard to reproduce**
  - closed-source descriptors
  - partly undisclosed dataset
- **Purely descriptive**
  - Delineation of the physicochemical space of well-absorbed drugs.

1. W.J. Egan et al. J. Med. Chem. **2000**
2. W.J. Egan et al. Adv. Drug Deliv. Rev. **2002**
On the way of laying our own egg

• Dataset: from Shen et al.\textsuperscript{1}
• Structures: neutralized and translated into SMILES
• Descriptors: WLOGP, in-house implementation Wildman and Crippen\textsuperscript{2} Log $P$
  TPSA, topological polar surface area from Ertl et al.\textsuperscript{3}

• Clear signal ✔

• Build the predictive model i.e. finding the classifying ellipse.

\begin{itemize}
  \item well absorbed
  \item poorly absorbed
\end{itemize}

\begin{itemize}
\item 1. J. Shen, J. et al. J. Chem. Inf. Model. \textbf{2010}
\item 2. S.A. Wildman & G.M. Crippen J. Chem. Inf. Comput. Sci. \textbf{1999}
\item 3. P. Ertl et al. J. Med. Chem. \textbf{2000}
\end{itemize}
Methodology for the optimization of the ellipse

- Curation of **training set**\(^{1,2}\) for **passive** absorption
  - 567 well-absorbed molecules
  - 93 poorly absorbed molecules
- Translated in SMILES, neutralized, aromatized

**Parameters** geometric definition of the ellipse

**Monte Carlo (MC) optimization**:
1. generate initial ellipse and **evaluate** for classification
2. random change \(\rightarrow\) new ellipse to evaluate
3. use Metropolis criterion to keep or reject the new ellipse
4. steps 2. and 3. repeated 100'000 times

100'000 parallel runs \(\rightarrow\) \(10^{10}\) ellipses to evaluate for classification

**Evaluation**: \(\text{Score} = MCC = 0.05 \times \text{surface}\)

Matthew’s Correlation Coefficient

\[
MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}
\]

from -1 to +1 (perfect prediction), near 0 for random classification.

---

1. J. Shen, *et al.* *J. Chem. Inf. Model.* 2010
2. D. Newby *et al.* *Eur. J. Med. Chem.* 2013
Best classification ellipse for passive absorption

- MCC = 0.70
- Accuracy = 93%
- In accordance with usual guidelines for good absorption
- Consistent with Egan Egg

Egan Egg

$0.0 \text{ Å}^2 < \text{TPSA} < 142.1 \text{ Å}^2$

$-2.3 < \text{WLOGP} < 6.8$
10-fold crossvalidation

- $\text{MCC}_{\text{CV}} = 0.65$ / $\text{Accuracy}_{\text{CV}} = 92\%$
- Internal robustness ✔
Another important physiological barrier: BBB

- **Blood-Brain Barrier (BBB)** is an effective **shield** protecting the brain:
  - **physical** barrier
e.g. tight junctions preventing paracellular penetration
  - **biochemical** barrier
enzymatic activities and active efflux
e.g. P-glycoprotein 1 (P-gp) pumping out

- **Passive diffusion** through BBB is the **major route** for drugs to access the brain from the bloodstream.¹

- Fundamental for the **distribution** of central-acting drugs or reversely for limited **unwanted effects** of peripheral drugs (or any molecule)

---

¹. L. Di et al. Drug Discov. Today 2012
Heatmap for passive diffusion through BBB

- 156 permeant molecules (log $BB > 0$, ●) and
- 104 non-permeant molecules (log $BB < 0$, °) taken from Brito-Sanchez et al.¹

- Clear signal ✔
- Narrower space ✔
- Optimization of a different ellipse
  - same method
  - same descriptors

¹ Y. Brito-Sánchez et al. Mol. Inf. 2015
Best classification ellipse for passive BBB permeation

- MCC = 0.79
- Accuracy = 90%
- Consistent with common guidelines for brain penetration

Final model

- MCC = 0.79
- Accuracy = 90%
- Internal robustness ✔

10-fold crossvalidation

- MCC\text{CV} = 0.75 / ACC\text{CV} = 88%
- Internal robustness ✔
BOILED-Egg: merging both models

1. A. Daina & V. Zoete, ChemMedChem 2016

- Same referential
- Brain Or Intestinal Estimate permeation
- Yolk and white not mutually exclusive
BOILED-Egg to track drug optimization path

- Exercise: Lead optimization of BCR-ABL tyrosine kinase inhibitors to anti-cancer Ponatinib\(^1\).
- Successive PD and PK optimization steps.

1. W.S. Huang *et al.* *J. Med. Chem.* 2010
**BOILED-Egg to map new FDA drugs**

- Exercise: map 46 NCEs accepted by the FDA in 2014 and 2015
- Colored according to usage label
- 83% consistency for route of administration and in/out « white »

- 6 false negatives (FN)
  - Daclatasvir
    - Good absorption
    - Active transport $^{1}$ OCT1

- 1 false positives (FP)
  - Olodaterol
    - FA 10-20%
    - P-gp efflux $^{2}$

---

1. FDA, Clinical pharmacology reviews 2014
2. Health Safety Regulation (Aus), 2014

---

- Oral administration
- Other routes only (i.v., inhalation, transdermal, ...)
- CNS indication
The BOILED-Egg within the SwissADME.ch web tool

- **SwissADME.ch** a free web tool to calculate pharmacokinetic, druglikeness and related parameters
  - Standardized input and enhanced analysis
  - Possibility to embed prediction of active transport
SwissADME one-panel-per-molecule output

• SVM model
  RBF kernel
  15 SwissADME descriptors
  10-fold CV

• Training: 564 substrates and 469 non-substrates\(^1\)
• Test: 215 substrates and 200 non-substrates\(^2\)
• Performance
  \(\text{Accuracy}_{\text{CV}}\) : 72%
  \(\text{Accuracy}_{\text{EXT}}\) : 89%

---

1. J. Mak et al. J. Cheminfor. 2015
2. V. Poongavanam, Bioorg. Med. Chem. 2012
SwissADME graphical output

- SwissADME include an enhanced version of the BOILED-Egg, including
  - estimation of P-gp substrate (active efflux)
  - Interactive analysis capabilities
SwissADME: a module of SwissDrugDesign

- **SwissADME:**
  - Not published or communicated yet.
  - Promising start of 40+ jobs per day, on average.
  - Referenced in VLS3D.com or click2drug.org

- **Interoperability** with other tools of SwissDrugDesign

---

V. Zoete et al. J. Chem. Inf. Model. 2016

D. Gfeller et al. Nucl. Acids Res. 2014
SwissDrugDesign: SIB initiative in CADD

Structure-based
- SwissDock.ch
- SwissSidechain.ch
- SwissParam.ch

Disease-related genomics
- SwissTargetPrediction.ch

Target identification and validation
- SwissSimilarities.ch

Hit finding
- SwissBioisostere.ch

Hit to lead
- SwissADME.ch

Lead optimization
- Clinical development

Preclinical development
Thank you