BIOLOGICAL ACTIVITY PROFILING AS A TOOL FOR VIRTUAL SCREENING

Vladimir Poroikov, Dmitry Filimonov, Tatyana Gloriozova, Alexey Lagunin, Olga Tarasova

Orekhovich Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences, Moscow, Russia

Introduction

Since the majority of pharmaceuticals have a pleiotropic action exhibiting both pharmacotherapeutic and adverse/toxic effects, profiling of their biological activity *in silico* allows selecting the most relevant screens for particular compounds and filtering out substances that might exhibit hazard effects.

Based on the freely available information about biologically active compounds (PubChem, ChEMBL, DrugBank, etc.) new computational tools for estimation of biological activity have been developed (see below). The applied methods vary widely from the relatively simple pairwise chemical similarity assessment to more sophisticated ligand-based or target-based approaches.

Our group published the first study describing an approach to provide chemists with the information about the most relevant targets/assays for their compounds [1-3], and additional computational tools with similar functionality have been developed in other labs more recently as well.

No systematic comparison of the accuracy and predictivity of these web-services has been performed yet. Therefore, the aim of our study was to analyze the relative predictive power of the available services for predicting the biological activity profiles based on information about new pharmaceuticals approved by US FDA in 2011 [4].

Materials and methods

Criteria for selection of web-services, to be evaluated:

- Multiple biological activities (biological activity spectra) are predicted in one run.
- "Ready for prediction" (completely pre-trained computational systems).
- Freely accessible (not necessary to purchase a license).
- The approaches are described in papers published in the peer-reviewed journals.

The selected web-services:

ChemSpider http://www.chemspider.com/ (Reid D. et al. *JCAMD*, **2008**, 22: 479).
CPI-DRAR http://cpi.bio-x.cn/drar/ (Yang L. *PLoS One*, **2010**, 5: e9568).
PASS Online http://pharmaexpert.ru/passonline (Lagunin A. et al. *Bioinformatics*, **2000**, 16: 747).
SEA http://sea.bkslab.org/ (Keiser M.J. et al. *Nat. Biotechnol.*, **2007**, 25: 197.
SuperPred http://bioinformatics.charite.de/superpred/ (Dunkel M. et al. *NAR*, **2008**, 36: W55).

ChemSpider (LASSO), PASS, SuperPred (CDK similarity) - ligand-based approaches.

CPI-DRAR, SEA (five different knowledge bases) - target-based approaches.

Criteria for selection of validation set:

- Newly published data.
- Compounds from the validation set reveal various biological activities.
- Compounds from the validation set belong to the diverse chemical classes.
- Biological activity of the compounds has been investigated in detail.

2011 FDA drug approvals

The US FDA approved 30 new therapeutics last year, including 11 first-in-class agents.

NATURE REVIEWS | DRUG DISCOVERY VOLUME 11 | FEBRUARY 2012 | 91

From 30 drug substances approved by FDA in 2012, 6 are biopreparations and 24 are synthetic drug-like substances (New Chemical Entities).

Two compounds have been excluded from further consideration due to the peculiarities of chemical structure and biological action: Gadobutrol - BBB imaging gadolinium-based contrast agent; Ioflupane is iodine isotope I^{123}.

The list of drug substances with their biological activities, which was used as a validation set, is given below.

Generic Name (trade name)	Indication	Properties	Generic Name (trade name)	Indication	Properties
Spinosad (Natroba)	Head lice	Causes neuronal excitation in insects	Fidaxomicin (Dificid)	Clostridium difficile-associated diarrhoea	RNA polymerase inhibitor
Vilazodone (Viibryd)	Major depressive disorder	Selective serotonin reuptake inhibitor and 5-HT1A partial agonist	Ezogabine (Potiga)	Partial-onset seizures	Potassium channel opener
Azilsartan (Edarbi)	Hypertension	Angiotensin II type 1 receptor antagonist	Indacaterol (Arcapta neohaler)	COPD	Long-acting β2-adrenergic receptor agonist
Roflumilast (Daliresp)	COPD exacerbations	Phosphodiesterase 4 inhibitor	Rivaroxaban (Xarelto)	Prophylaxis of deep vein thrombosis in hip and knee replacement surgery	Factor Xa inhibitor
Gadobutrol (Gadavist)	Blood–brain barrier imaging agent	Gadolinium-based contrast agent	Ticagrelor (Brilinta)	Thrombotic cardiovascular events in patients with acute coronary syndrome	P2Y12 platelet inhibitor
Gabapentin enacarbil (Horizant)	Moderate-to-severe restless legs syndrome	Voltage-activated calcium channel inhibitor	Vemurafenib (Zelboraf)	BRAF-positive unresectable or metastatic melanoma	BRAF inhibitor
Vandetanib (Caprelsa)	Unresectable or metastatic medullary thyroid cancer	VEGF, EGFR and RET inhibitor	Icatibant (Firazyr)**	Hereditary angioedema	Bradykinin B2 receptor antagonist
Abiraterone (Zytiga)	Metastatic castration-resistant prostate cancer	CYP17 inhibitor	Crizotinib (Xalkori)	ALK-positive advanced or metastatic NSCLC	ALK inhibitor
Linagliptin (Tradjenta)	Type 2 diabetes	Dipeptidyl peptidase 4 inhibitor	Deferiprone (Ferriprox)	Transfusional iron overload due to thalassaemia syndromes	Iron chelator
Boceprevir (Victrelis)	HCV genotype 1	NS3/4A protease inhibitor	Clobazam (Onfi)	Seizures associated with Lennox–Gastaut syndrome	Benzodiazepine
Rilpivirine (Edurant)	HIV-1 infection	Non-nucleoside reverse transcriptase inhibitor	Ruxolitinib (Jakafi)	Intermediate or high-risk myelofibrosis	JAK1/JAK2 inhibitor
Telaprevir (Incivek)	HCV genotype 1	NS3/4A protease inhibitor			

Diverse chemical classes of 22 drug substances from the validation set:

Antibiotic macrolide	Purines	Triazolopyrimidines
Benzofurans	Azabicyclohexanes	Triazolopyrimidines
Benzimidazoles	Substituted benzonitriles	Peptides
Pyridines	Cyclopentapyrroles	Pyridines
Tetraazacyclododecanes	Macrolide antibiotics	Benzodiazepines
Substituted cyclohexaneacetic acid	Fluorobenzylamines	Pyrrolopyrimidines
Quinazoline	Quinolines	
Androstanes	Oxazolidinones	

Validation Criteria.

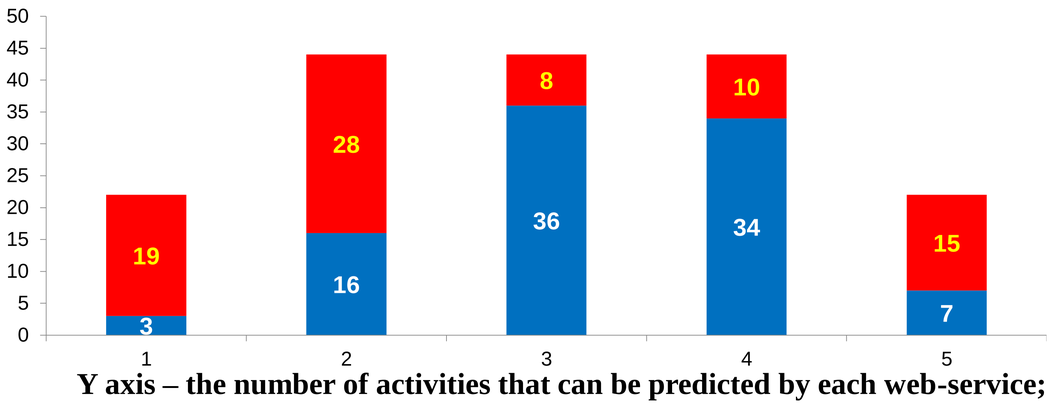
Since no one compound has been tested against all possible kinds of biological activity, the only parameter that can be used to characterize the relative predictivity of web-services is:

Sensitivity = TP/NA,

where TP – true positives; NA – number of "actives".

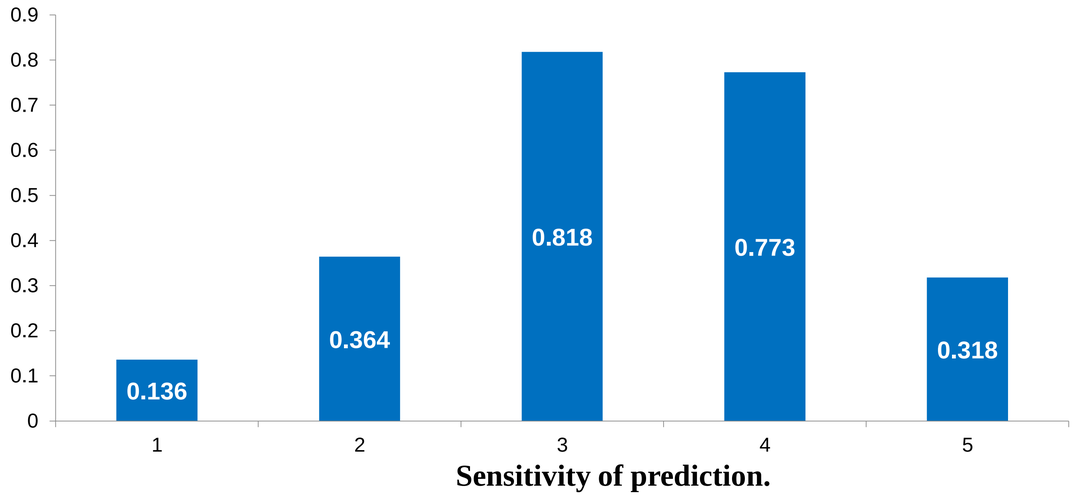
Results

Validation results are presented in diagrams below (1 – ChemSpider; 2 – CPI-DRAR; 3 – PASS Online; 4 – SEA; 5 – SuperPred).



Y axis – the number of activities that can be predicted by each web-service;

blue - the number of predicted activities; red - the number of unpredicted activities.



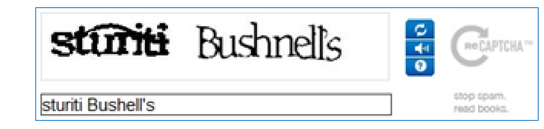
Sensitivity of prediction.

Computation time:

ChemSpider ~10 min; CPI-DRAR ~500 h; PASS Online ~10 min; SEA* ~120 h; SuperPred ~10 min.

---

*Effective time, including all false attempts due to the "Captcha" protection:



Conclusions.

1. Predictivity of five web-services for estimating biological activity profiles of drug-like compounds given in descending order: PASS Online > SEA > CPI-DRAR > SuperPred >> ChemSpider.
2. Computational time given in ascending order: PASS Online ≈ SuperPred ≈ ChemSpider << SEA << CPI-DRAR .
3. Aggregation of results provided by different web-services increases the sensitivity to 0.955 (42/44).

Acknowledgement. The work was partially supported by RFBR grant No. 12-07-00597.

References:

1. Lagunin A. et al. *Bioinformatics*, 2000, 16: 747-748.
2. Filimonov D.A. and Poroikov V.V. In: *Chemoinformatics Approaches to Virtual Screening*. Cambridge (UK): RSC Publishing, p.182-216.
3. http://www.pharmaexpert.ru/PASSOnline
4. Mullard A. *Nature Reviews Drug Discovery*, 2012, 11: 91-94.