Data-Driven Drug Discovery and Molecular Informatics

Fact Sheet (current as of 2019) – Group of Dr Andreas Bender, Centre for Molecular Informatics, Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, United Kingdom
Group: www.ch.cam.ac.uk/group/bender/, Personal: www.andreasbender.de, Email: ab454@cam.ac.uk

- Applying and developing data analysis/machine learning/artificial intelligence methods in the chemical biology and drug discovery areas, across chemical (structure) and biological domains (transcriptomics, imaging data, animal toxicity readouts, human adverse events, …), for understanding compound activity and disease subtyping, repurposing, improving toxicity/safety & supporting compound selection/design
- Currently about 15 group members (ca. 9 PhD students, 3 postdocs, academic visitors etc.)
- PI experienced in both pharmaceutical industry (Lead Discovery Informatics Group at Novartis, Cambridge/MA) and academia (Leiden/Cambridge; ~200 publications, ~£5m grant income in last 10 yrs)
- Funding from public sources (ERC, EU/Horizon2020, BBSRC, EPSRC, CEFIC, …) as well as pharmaceutical, chemical, and consumer goods industries (BASF, Eli Lilly, Johnson&Johnson, AstraZeneca, Unilever, Aboca, Evotec, Roche, GlaxoSmithKline, ...)

Main Research Areas (list not exhaustive)

- Mode-of-action analysis and target prediction
  - Mode-of-action analysis ('target prediction') based on chemical structure and –omics readouts; methods development and applications to various phenotypic readouts (related both to efficacy and toxicity)
  - Modelling bioactivities on target families (e.g. kinases, proteases, others)
- Modelling mixtures of bioactive compounds and traditional medicines
  - Compound combination/mixture modelling (oncology, antibacterials, …), based on both chemical structure and biological readouts (eg transcriptomics data)
  - Traditional medicines/natural products (Traditional Chinese Medicine, Ayurveda, …)
- Integrating chemical/biological data for characterizing compound action and disease, repurposing, …
  - Using gene expression/RNA-Seq data for compound selection and repurposing (methods development; prospective validation in oncology (pancreatic cancer, blastomas, etc.), malaria, stem cell differentiation, and others)
  - Integrative mode of action analysis using joint chemical/biological data
  - Patient stratification (using eg SNP/sequencing data in Inflammatory Bowel Disease, more generally)
- Toxicity prediction
  - Linking early-stage readouts (from safety profiling, ToxCast data etc.) to organism-level toxicity (histopathology readouts, adverse events, etc.)
  - Understanding toxicity readouts and toxic mechanisms based on 'omics (transcriptomics) and imaging data (for risk assessment, both in pharmaceutical and consumer goods context)

Future Aims and Collaborations Envisaged

- Our strengths are in *understanding and analysing chemical and biological data*, and in the *generation of computational models for compound selection, optimization, and improved understanding of both efficacy- and toxicity-related effects, both from the compound and the patient side*
- We are experienced in handling a wide range of data from the chemical and biological/medical domains, worked extensively with industry, and aim for practical utility and experimental validation of our models
- Please don't hesitate to get in touch if you are interested in future collaborations:
  Andreas Bender (ab454@cam.ac.uk, +44 1223 762 983)