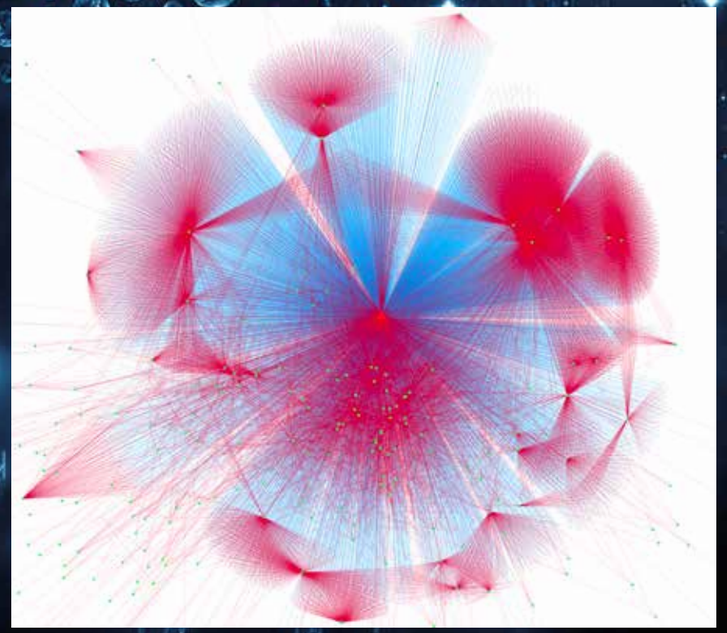
Case Study

Artificial intelligence and machine learning: revolutionising drug development

Artificial intelligence (AI) is revolutionising all stages of drug discovery and development from drug design to pivotal clinical trials – and beyond. With deep learning, machines are starting to mimic the activity of human neurons in the brain – to create, in effect, an artificial 'neural network.' We are taking an AI approach to a growing number of assays deployed during the drug discovery phases, reducing costs and speeding up the supply of data to our scientists. We are also using AI to help us analyse and interpret huge amounts of data from imaging studies of pathological samples and from biomarker research, to match the right drug to the right patient. In our clinical trials, AI is enabling us to continuously monitor incoming safety data and alert our scientists to safety signals that need attention.

By integrating the latest AI advances into all our drug development programmes, we are helping our scientists to harness the power of 'big data' to deliver potentially life-changing medicines to patients most likely to benefit.



Predicted compound bioactivity space and inferred target interactions

> "The energy, motivation and enthusiasm at DMTA Hackweek was phenomenal – we worked relentlessly every day and evening for the entire week. It looked almost impossible to build an automated DMTA platform from scratch within one week but we did it. There is a lot of work ahead of us but this is the first step on the path to fully leveraging the capabilities of emerging laboratory automation technology and machine learning."
>
> **Michael Kossenjans, Head of iLAB, Gothenburg**

Drug discovery: where human and machine readable data coexist

Through machine learning, our state-of-the-art drug discovery robots and other automated devices can adjust their activities in response to the data they receive, enabling us to work more quickly and efficiently.

AI-driven automation is helping us address some of our biggest challenges in chemistry with the aim of speeding up the entire Design-Make-Test-Analyse (DMTA) cycle of compound synthesis, and facilitating rapid, unbiased decision making.

Our novel, DMTA platform is the first step on the path to fully leveraging the capabilities of emerging laboratory automation technology and machine learning to speed up constructing and testing hypotheses in drug discovery. It is used to iteratively improve the overall profile of therapeutically relevant compounds. For an average project, hundreds of DMTA cycles are required to find the compound that meets the criteria for a candidate drug. When the steps in these cycles are done manually, they can take several weeks, but our ambition is to reduce the time between compound design and receipt of test data from four to six weeks, to less than five days.

In 2017, we held a DMTA "Hackweek" at which a small highly-focussed group of scientists from all our research sites came together with external experts to combine their scientific and technical expertise to build a first prototype 'DMTA machine'. In a designated innovation laboratory in Gothenburg, they worked solidly for five days to revolutionise the way we do drug discovery. With the spirit of 'hackers' the team overcame numerous challenges, connecting a mix of hardware and software accumulated during more than two decades. The outcome was a prototype machine that could run the entire DMTA cycle for an ongoing research project in less than two hours.

They say a journey of a thousand miles starts with a single step! With this simple prototype system, we are now developing the machine learning needed to optimise the potency of a new compound, predict different routes for synthesis and make the automation more sophisticated so that we can make more complex molecules and generate more screening data.

Quantum computing: Speeding up structural chemistry to find the molecule that matters

Establishing the chemical 3D-structure of potential new medicines is a key element of drug discovery and development because size and shape matter. They affect many different characteristics, including interactions with biological systems and the way molecules pack together to form materials needed for successful formulation.

Yet, rather like a parachutist who needs to make multiple jumps into the Swiss mountains to find the deepest valley, we have to look again and again to evaluate all possible shapes of molecules to find the low energy conformations we need to optimise our medicines.

The emerging field of quantum computing has the potential to make a dramatic impact on how we approach this problem. A quantum computer can simultaneously explore all the possible conformations of our compounds and, given suitable criteria, converge on the best possible option in a single operation.

By tuning the criteria, it is possible to locate a range of high-quality solutions. In the near-term, quantum computing is limited in the amount of information it can process, and we still have to do some post-evaluation analysis on the solutions it offers, using existing accurate methods on standard computers. The future application of machine learning is expected to take quantum computing to the next level. Our current hybrid approach looks set to bring the most relevant solutions within our grasp and help us make the best choices of chemical structure.

Using AI to liberate our IMED scientists

In discovery, we are also applying AI to make existing processes more efficient, and to turn data into knowledge. We are using AI to reliably predict the results of routine assays, such as human plasma-protein binding (hPPB) tests, to liberate our scientists and give them more time to focus their passion for science on problems that will give AstraZeneca an even greater competitive edge.

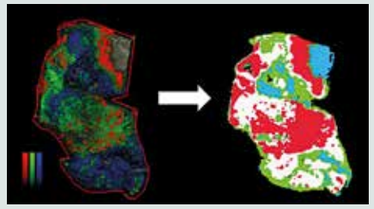
The hPPB assay, developed in Drug Safety and Metabolism, is used to help us understand how a potential drug molecule is distributed within a patient. We are working with world-leading partners to use the latest advances in AI to predict the results with a high degree of confidence.

We are currently assessing the utility of AI for safety screens, protein production, image analysis and the design of clustered regularly interspaced short palindromic repeats (CRISPR) gene editing. In the months and years ahead, we expect to use AI to transform any area in the discovery process where we collect data and turn that data into knowledge.

In 2017, a testament to the success of machine derived efficiency was the development of a virtual screening tool FastVS. The new “Google-like” web-based tool, developed in collaboration with OpenEye Scientific Software, reduces the time to search and score entries in numerous large molecular databases, from hours to seconds, optimising the process of drug discovery.

Big data analysis helps move traditional pathology into the 21st century

In a science-driven environment, the ability to rapidly identify and learn from signals and patterns in the data we generate is key to building knowledge and influencing future scientific direction and discovery. To achieve this, we need to access and integrate our large and diverse data sets in a usable format.



Graph theory-based clustering algorithm of MSI data to map drug distribution and metabolism in a tumour mouse model

In the past, evaluation of comprehensive gene expression, protein and metabolic data from an individual organ was hampered by our limited computational capacity for data analysis. For the first time, AI is making it possible to tackle ‘big data’ and analyse all these endpoints and their spatial relationship to each other.

We are using mass spectrometry imaging (MSI) to spatially map molecules to their cellular localisation in biological samples and sections of tissue, such as those taken for pathology assessment. This comprehensive data-rich spatial information is a fantastic opportunity to link the tissue microenvironment with drug localisation, efficacy and safety. However, existing data mining approaches place great demands on computer systems and we are limited to analysing small, single data sets. To address this we have developed new computational algorithms that enable accurate and efficient segmentation of large amounts of MSI data to improve our capacity to learn across multiple endpoints, as described in our recent publication in *Analytical Chemistry.*

This has enhanced our capacity to accurately quantify molecular changes in tissue and organ specific regions and mine data for increasingly complex spatial relationships. Critical to our success was the close collaboration between researchers in Drug Safety and Metabolism and external experts in computer and pathological sciences.

Looking ahead, we plan to combine deep learning algorithms with image analysis to accelerate evaluation of animal models of chronic kidney disease, generating more robust data for downstream multispectral image analysis. This will increase quantitative analysis speed, confidence and reproducibility of data and allow integrated multimodal image mining for detecting biological relationships and consequences. In 2017, our scientists were welcomed into Cancer Research UK’s (CRUK) Grand Challenge Team for their proposal to map the tomography of tumours – a Google map approach to mapping cancer. This has the potential to take pathology – one of the most traditional safety disciplines – firmly into the 21st century and beyond.

Harnessing AI to connect the right drug to the right patient

Tissue biomarkers are increasingly used to match the right drug to the right patient. However, current technology involves manual scoring of images by pathologists – a process that is subjective, time consuming and complex. To overcome this challenge we turned to AI, and developed a novel deep learning algorithm to automatically score tissue biomarkers, using a process called digital pathology.

In a proof-of-concept study of 71 patient tumour samples, we showed that AI can automatically score human epidermal growth factor receptor 2 (HER2), a well-established biomarker in breast cancer. The algorithm also identified samples at risk of misdiagnosis, demonstrating its potential to make tissue biomarker scoring faster, simpler and more precise.

In 2017, we presented this exciting work at two world leading scientific congresses and published it in *Scientific Reports.*

We are continuing to use the latest science to lead the way in similar advances, through collaborations with the best academic partners. Using a combination of cloud computing and the latest graphics processing unit (GPU) hardware, we intend to make automated analysis of digital pathology images a high throughput process, and to incorporate AI algorithms into the development of diagnostics tests. We aim to use the power of AI to affect patients’ lives, matching targeted medicines even more precisely to those most likely to benefit.

Watcher: monitoring drug safety, 24/7

In our early phase clinical trials, Watcher – a novel AI-based decision support system – is continuously monitoring incoming safety data and alerting scientists to safety signals that need consideration and action.

Watcher is an innovative AI driven alert system for clinicians and scientists involved in clinical trials, which embeds clinical decision rulesets in logic. It is a key component of our iDecide research programme – a five year collaboration between AstraZeneca, the University of Manchester Institute of Cancer Sciences, the Centre for Cancer Biomarker Sciences and the Christie National Health Service Foundation Trust. The digital Experimental Cancer Medicine Team (digitalECMT), based within CRUK Manchester Institute, delivers iDecide, working directly with patients to develop new ways to enable better and earlier clinical trial decisions that directly benefit the patient.

Watcher complements REACT 4, another iDecide tool, which collates and visualises safety, efficacy and biomarker data from Phase I and Phase IIa trials and is currently deployed in over 140 studies at AstraZeneca. REACT 4 enables clinical interpretation on-demand and depends on the system being actively used. However, Watcher’s continuous monitoring allows insights and notifications to be generated when a signal emanates, without manual intervention.

In the future, we plan to enhance Watcher with clinical rulesets and machine learning to augment and extend its current functionality. These developments will enable it to be brought into the patient’s home in conjunction with point-of-care devices, further empowering our patients to self-monitor while participating in our clinical trials.