INTERVIEW

Dr. Scott Lusher

Director Strategy & Communication, Netherlands eScience Center

FACILITATING DATA-DRIVEN DECISIONS IN DRUG DEVELOPMENT

Technology can support the use of "big data" for drug design and optimization, but not without a shift in mentality.

ELSEVIER

The question driving the drug design cycle will need to change from "which is the next compound in the series to synthesize?" to "how do we evaluate the more complex and conflicting data resulting from each iteration of the cycle?

"It is becoming increasingly obvious that the rapidly growing amounts of data that should inform drug discovery do not fit current design workflows. This situation will need to change."

Dr. Scott Lusher's experience lies in knowledge discovery, chemoinformatics and computer-assisted drug design. He began his career at Unilever Research but, interested in the drug discovery process, he moved later to the pharmaceutical industry (Organon, Schering-Plough and Merck). Dr. Lusher worked for 10 years on molecular rationale for new chemical entity selection and design and the implementation of molecular informatics in early research. Dr. Lusher is now a director at the Netherlands eScience Center, an organization developing new IT solutions in support of methods that better reflect and leverage the new, data-laden nature of scientific research. He is especially interested in exploring data challenges of different disciplines to apply lessons learned and solutions in other fields. In this interview, Dr. Lusher shares his perspective on the challenges facing precision medicine to assimilate and interpret the growing amount of data that should inform design decisions.

PRECISION MEDICINE AIMS TO INTEGRATE DATA ON PATIENT CHARACTERISTICS AND DATA ON TARGET-SPECIFIC PARAMETERS TO DELIVER MAXIMAL THERAPEUTIC BENEFIT TO A NARROW SUBGROUP OF PATIENTS. HOW DOES THE INCREASED AMOUNT OF DATA NEEDED TO SUPPORT THIS GOAL IMPACT THE WORK OF MEDICINAL CHEMISTS?

Initially, the added information probably does little to change the day-to-day work of a medicinal chemist when designing a drug. The complexity of outcomes that a medicinal chemist has to bear in mind when designing the compound is already quite significant. The only way to practically account for all desired pharmacokinetic and pharmacodynamic properties is to reduce them to a set of design criteria that can be measured by available assays. Complex drug-target and off-target interactions must be whittled down to standardized, measurable parameters that fit into equally standardized workflows that we hope reflect the potential clinical activity of the compound. Where the added information is likely to make a difference in early discovery is in the validation of targets, assay design and selection of design criteria. We all have great hopes that precision medicine approaches will ensure that we spend more time chasing criteria that better reflect clinical outcomes and safety in patient populations or sub-populations. However, this may well result in more complex criteria, and that is where the challenge will arise. Translating pharmacotherapy outcomes — efficacy and safety of the drug — into quantifiable parameters often results in conflicting criteria and drug design becomes an exercise in tradeoffs and compromises to find the right balance. This situation is exacerbated as additional input expands the list of design criteria. The challenge of precision medicine to the medicinal chemist is that the industrialization of drug discovery and development has augmented our capacity to rapidly measure multiple chemical and biological properties of new chemical entities. It is now astoundingly easy to collect thousands to millions of data points. What has not been improved sufficiently is the process to interpret, evaluate and make sense of those data. Our capacity to evaluate this growing volume, complexity and diversity of data and make decisions based on it, has not kept pace. It is becoming increasingly obvious that the rapidly expanding amounts of data that should inform drug discovery do not fit current design workflows. This situation will need to change.

WHAT WOULD BE A BETTER APPROACH TO FACILITATE THE DEVELOPMENT OF PRECISION DRUGS?

Discussion is already underway about shifting to a research and development process based on network biology, where the focus is not the impact a drug has on a single target but rather on the overall biological network where the target is rooted. Coupled to this approach will likely be designed polypharmacology, either as a single drug that impacts the function of multiple targets in a network, or a formulated mixture of two or more compounds, each targeting one of a handful of relevant network components. Network biology and polypharmacology will not, however, make drug design easier. On the contrary, tailoring the therapeutic window of precision drugs under this approach will require answering more complex questions, will require profiling more compounds, will require developing newer and more diverse chemical and biological assays. A team working in a therapeutic area will need to prioritize a manageable set of design criteria from a growing list of desired properties, and the trick will be anticipating the right criteria to prioritize. This means that the question driving the drug design cycle will need to change from "how do we design the next compound?" to "how do we evaluate the more complex and conflicting data resulting from each cycle iteration?" The development process will not be easier, but it will produce better drugs.

WHAT CHANGES DO YOU ANTICIPATE ARE NECESSARY TO BEST "EVALUATE MORE COMPLEX AND CONFLICTING DATA"?

The upshot of this new focus for the design cycle will be that data evaluation will take longer, bringing decision-making and compound synthesis out of phase. This is already a complaint about computational chemistry: it often can't keep up with synthesis and testing. So, in addition to finding ways of managing, sharing, and visualizing more complex data, changes must focus on enabling decisions to keep pace with compound synthesis.

To this end, all data must be openly available at all times to all team members and traceable to their source. This requires improved lab practices, systematic documentation of not only results but also protocols and any deviations from standard protocols, and a data management system that is updated regularly and easy to use so that generated data end up on a shared platform, rather than on USB sticks or in binders stashed in a drawer. This may sound trivial, but in reality is often done poorly. To achieve this, there must be incentive to enter data into the system as soon as it is generated and to control data accuracy. Increased transparency about when and how data are used is a very successful incentive. This is because the vast majority of data are entered into databases by scientists who rarely have cause to view those data again. They are not involved directly in the design of new compounds and using the data they generate. Increase their awareness of how others, like medicinal chemists, use their data to make decisions about design, and the quality and completeness of their data will improve dramatically. Knowing that their data matter is incentive to improve their sharing practices.

The nature of project meetings will also need to change. Solid evaluation and decision-making will require the input of everyone in the team, including those generating the data. Interpretation of

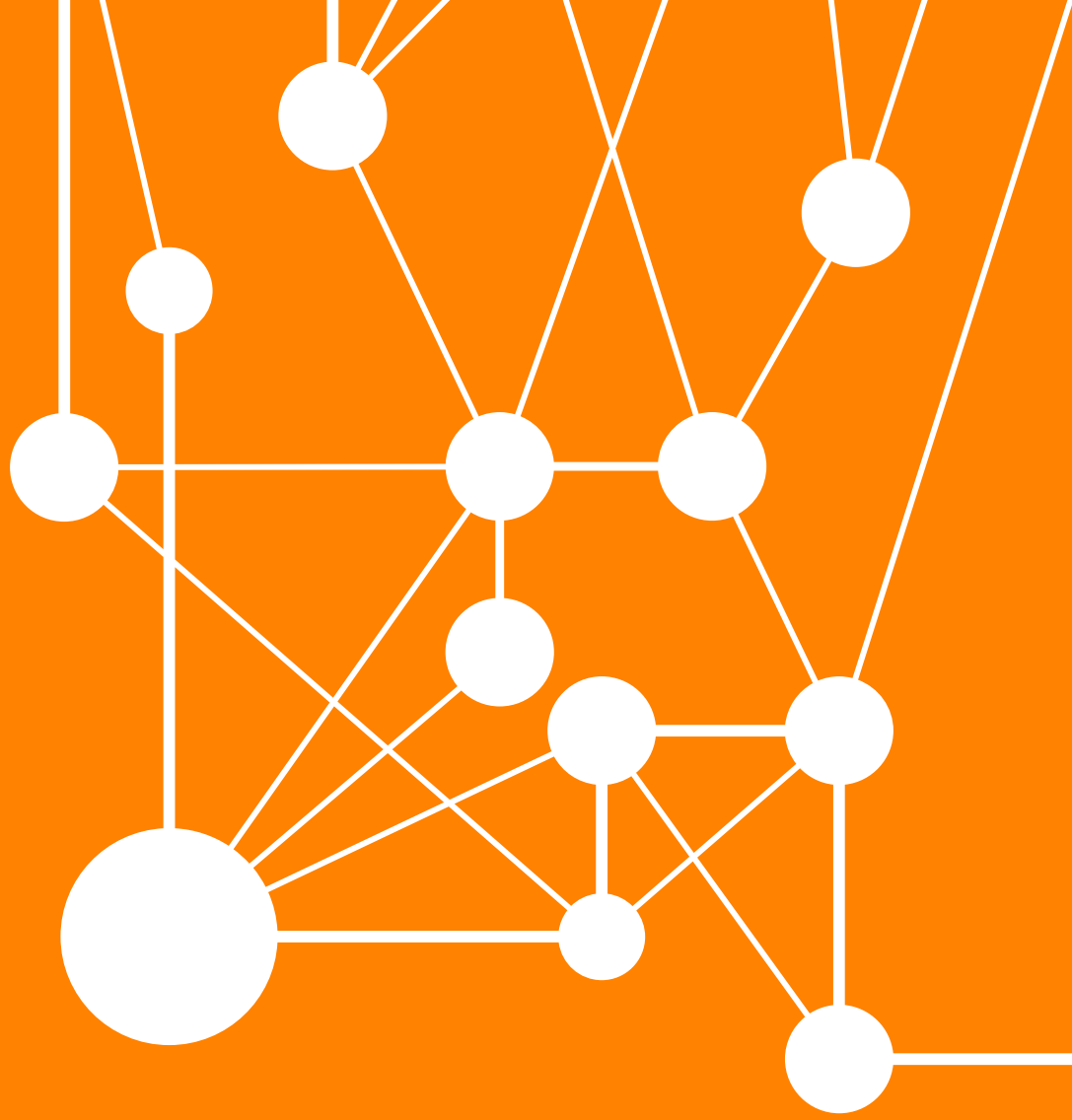
the data generated by team members should be done as a group, with everyone looking at the same, most up-to-date data. This has two effects. First, it creates an environment where all data are accessible and cannot be easily ignored — that is, it promotes data-based decisions. Second, giving data center stage at meetings highlights the importance of the work done, creates room to discuss the "story" behind the data, and incentivizes people to know the data before they come to a meeting.

WILL NEW TECHNOLOGICAL DEVELOPMENTS PLAY A BIG ROLE IN IMPLEMENTING THESE CHANGES?

Technology will naturally play a role as a tool for data recording, analysis and sharing. With the right technology, data can be recorded directly from instrument readouts, automatically analyzed and continually updated so results of standard analyses are always available. Technology will also help with new and creative ways of visualizing data. Presentation needs to become more data-centric — along the lines of spreadsheets that enable filtering, browsing, sorting and comparison — and current forms of visualization will not suffice as the number of assays increases and as our understanding of relevant biological pathways deepens.
However, the features and analysis power of any technology is secondary to the interpretation of data, and this is still done by people. A change in mentality is also required. Medicinal chemists, biologists, pharmacologists, and all other team members will need to work with, share, trust, and interpret data differently. With increased amounts and complexity of data comes a latent danger that team members become overwhelmed and, instead of exploring and interpreting data, fall back on old dogma or intuition to make decisions. There is also a propensity to look for data that support an idea, rather than allowing the objective interpretation of data to guide work. A team in that situation can have all the resources in the world to produce more data, but they will ignore it. What needs to happen is that generated data are used to answer the questions "what have we learned?" and "why do we want to make the next compound?" To arrive at answers, each team member must become a knowledge worker. Currently, each team member has a specific task: the medicinal chemists do the design and synthesis, the biologists perform bioassays, the toxicologists search for potential adverse effects, and so on. Those boundaries will need to vanish. While each individual brings his or her expertise to the group, all must be versed in every discipline encompassed by the team and all must work at the interface of these fields. Finally, more data specialists will be necessary, both to manage data flow (storage, updates, visualization, etc.) and to support increasingly complex analyses, statistics and modeling.

We need better ways of analyzing data. Speed can be afforded by automating standard analyses and continually updating visuals and models in support of decisions. However, deeper analyses need to be done and they simply require time. We require new mechanisms to truly understand the generated data, to make connections between data points, to elucidate and capture meaningful patterns. It is not uncommon that small biotech companies working on drug discovery, design and optimization move with more agility through the design cycle. This is partly because smaller companies do not operate within a restrictive industrialized framework and thus, have the time and the autonomy as a team to explore data, find connections, make mistakes and act upon their findings.

I also think that industry would do well in reasserting the importance of the team. Oversight should naturally remain in hands of management, but concrete decisions about the direction of a compound series are best made by all hands — the team — that generated and know their data.



Visit elsevier.com/rd-solutions
or contact your nearest Elsevier office.

ASIA AND AUSTRALIA
Tel: +65 6349 0222
Email: sginfo@elsevier.com

JAPAN
Tel: +81 3 5561 5034
Email: jpinfo@elsevier.com

KOREA AND TAIWAN
Tel: +82 2 6714 3000
Email: krinfo.corp@elsevier.com

EUROPE, MIDDLE EAST AND AFRICA
Tel: +31 20 485 3767
Email: nlinfo@elsevier.com

NORTH AMERICA, CENTRAL AMERICA AND CANADA
Tel: +1 888 615 4500
Email: usinfo@elsevier.com

SOUTH AMERICA
Tel: +55 21 3970 9300
Email: brinfo@elsevier.com

May 2015