

Deliver on the ‘fail early’ pharmaceutical mantra

Collaborative, interdisciplinary data sharing holds the key to reducing late-stage toxicity failures

By Shikha Varma-O'Brien, Ph.D.,
Accelrys Inc.

Every pharmaceutical researcher knows that when it comes to drug discovery, “fail early, fail cheap” is the goal. No company wants to discover a major problem after millions of dollars and years of research have been invested, once clinical trials are well under way, or, as in the unfortunate cases of **Vioxx** and **Rezulin**, after a drug has been put on the market.

Despite this understanding, late-stage toxicity failures continue to vex the pharmaceutical industry. Researchers need more effective ways to pinpoint absorption, distribution, metabolism, and excretion (ADME)-related issues earlier in the discovery process, and more collaborative, interdisciplinary data sharing techniques, combined with advanced approaches to predictive analysis, offer great potential in this area.

Enable parallel optimization

During a typical drug-discovery project, the chemistry department creates potential compounds, biologists run assays to measure activity levels, and only after low potency molecules have been screened out are the top candidates tested for toxicity. The highly sequential and isolated nature of this research method poses several key problems. First, toxicity failures are discovered far too late in the process – and

a positive result on the eve of clinical trials can easily take an organization back to square one with millions of dollars in research down the drain. Second, toxicity data is disconnected from the work of chemists and biologists, which means that organizations often miss out on key insights that could be gained if research was done in parallel, across scientific disciplines. For example, a low-potency molecule that also has a very low toxicity risk might be worth investigating further. But in a typical discovery scenario, such a candidate would have already been discarded by the biology department before toxicologists were even brought into the process.

Parallel optimization, i.e., research that enables the investigation of compound structure, activity, and toxicity simultaneously, offers great promise when it comes to understanding toxicology implications earlier in the discovery process, but this can only be made possible through cross-disciplinary collaboration. All too often, domain experts like chemists, biologists, informatics experts, and toxicologists work in isolation from one another – they are evaluated and compensated differently, and use their own processes, systems, and semantics.

Successful parallel optimization requires that this disjointed way of doing things changes. More streamlined information sharing and better data integration are needed to make this happen. This is where new advances in scientific information management have an important role to play.

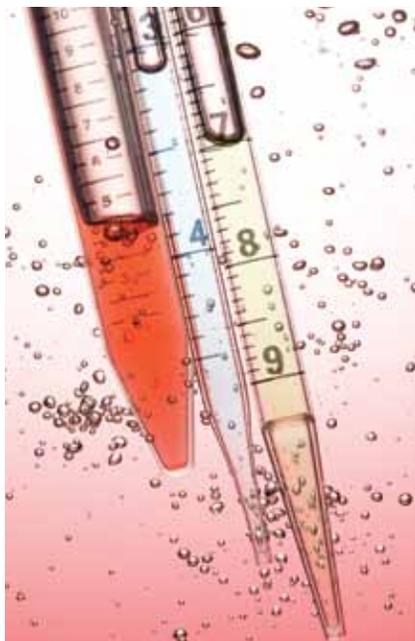
Unlock, integrate critical research sources

As any pharmaceutical researcher knows, the issue of data integration is a thorny one. Data generated by a single chemist or biologist, much less an interdisciplin-

ary group, is often spread across a diverse array of formats, applications, and proprietary systems, such as unstructured text documents saved in an electronic lab notebook or images generated by a microscope. The resulting volume is enormous – data may span thousands or even millions of possible compounds, assay results, and more. Pinpointing the information most relevant to ADME is akin to finding a needle in a haystack. This challenge is especially acute in today's global pharmaceutical environment, where data silos also stretch across geographic boundaries and time zones.

Fortunately, technologies like service-oriented architecture are changing this by enabling a more unified approach to managing complex scientific information related to drug candidates. A Web services-based IT foundation can support the integration of multiple sources of data in a “plug and play” environment, unlocking experimental results previously marooned within disciplinary, system, application, and departmental silos, and making the entire knowledge base of both successes and failures more accessible. As a result, researchers can quickly zero in on important toxicity information, use it to predict potential undesirable outcomes, and avoid wasting time by repeating experiments that have already been tried by another department or project team.

Beyond internal data sources, being able to access publicly available information – such as relevant literature or toxicology databases – is equally important. This is exemplified by author and researcher Christopher Lipinski's “Rule-of-Five,” which has helped scientists around the world make better predictions about potential failures in orally administered compounds. Imagine where the industry would be today if he had only shared this breakthrough analysis technique with his immediate team.



Researchers would do well to take a lesson from Dr. Lipinski and make it a point to share key insights – about what succeeded, what failed, and why – not just within their own organizations, but also with the larger pharmaceutical community. The work Dr. Ann Richards has done with the Distributed Structure-Searchable Toxicity Database Network is a strong example of this. When organizations can leverage all the knowledge available to them – both structured and unstructured, proprietary and public – in a smarter, more efficient way, everyone benefits.

Less one-off experiments, more predictive analysis

Countless variables can affect a drug candidate's toxicity and trigger undesirable side effects. But running one-off, trial-and-error type experiments on every possible compound can quickly become extremely costly and time consuming. During the last 50 years, the pharmaceutical industry has accumulated a wealth of information that can help organizations figure out what is likely to fail. The trick is being able to use this data to speed the discovery process and uncover potential ADME issues before late-stage preclinical or clinical trials. This requires moving beyond data access; organizations also

Toxicity data is disconnected from the work of chemists and biologists, which means that organizations often miss out on key insights that could be gained if research was done in parallel, across scientific disciplines.

need to explore analysis techniques, such as computational toxicology modeling, that can manipulate complex scientific information and enable researchers to replace expensive one-off experimentation with a more predictive and far less costly approach.

For example, a report published recently by **Bayer Schering Pharma AG** details the use of computational toxicology modeling for *in silico* prediction of Ames mutagenicity. By using a solution that employed a highly sophisticated Bayesian categorization model in conjunction with a foundation for scientific information management, Bayer Schering was able to combine large data sets from both public and internal sources in order to predict mutagenicity with greater ac-

curacy and sensitivity, and thus reduce the need for multiple experimental assays.

When using advanced analysis tools like the one described in the Bayer example, the key thing to remember is that a model is only as good as the data that goes into it. Thus, finding solutions capable of unifying advanced statistical methods with a broad spectrum of rich data sources is critical. By making a tighter structure-activity-toxicity connection, researchers can better understand toxicity at the structural level and make faster progress.

Editor's note: This is one of a series of occasional guest columns. Shikha Varma-O'Brien joined Accelrys Inc. (accelrys.com) in 2000 as an applications scientist specializing in pharmacophore modeling, QSAR, and structure-based design methods. As an applications scientist, Dr. Varma-O'Brien has worked on several computational drug-discovery projects with many pharmaceutical and biotechnology companies. After managing a team of applications scientists, she transitioned into product management. Dr. Varma-O'Brien presently holds the position of director of product management and marketing for modeling and simulations tools and has particular interest in data mining and structure/fragment-based design methods. Dr. Varma-O'Brien received her Ph.D. from Wayne State University, where she studied structural biology and biophysics.