ELSEVIER

editorial

Steve Carney

Target validation

I suppose it could be my age, but I was thinking just the other day about the way people went on about hi-fi in the late 70s and early 80s. It was not uncommon to find a huge slab of concrete sitting on 4 squash balls in people's living rooms, upon which sat their turntable (generally a Linn Sondek if someone had gone to all that trouble); in the 80s, I remember a fad in which people would use green felt tip markers to mark the edges of their CDs as they claimed this improved sound quality and ambience. Thankfully, these topics seem to have been largely abandoned as dinner party conversation topics. You may ask, what have these ramblings to do with the science of drug discovery? Well, I recall the mantra in hi-fi from that age was 'Rubbish in, rubbish out', which even today in the field of drug discovery is particularly apposite.

Please allow me to elaborate; in drug discovery, one of the earliest stages of the process is target identification and validation. If the selection of the target to investigate is in some way flawed, then irrespective of the quality of the process used to produce hits, leads and so on, the final compound will not be suitable as a quality drug candidate, hence, rubbish in, rubbish out.

The process of target identification

I suppose the fundamental question is: how are targets identified? In general, they would come from ideas from research scientists, either within academia or industry. They will arise as a result of something that a scientist has done in the lab, or through data mining or something that they have read. The generation of an hypothesis at this stage will attempt to link a biological molecule (or molecules) to a pathologic process, i.e. does modulation of that particular molecule result in a beneficial result in disease pathology. If the target is the one that is already available in the laboratory, this initial part of the process may be achieved through in-house experimentation. Alternatively, the hypothesis may be developed through a thorough review of the scientific literature. I do not have figures on this, but my guess would be that the majority of hypotheses underpinning new projects have their birth in scientists' minds and the pages of journals. A case must then be made to allow the appropriate allocation of resource to test such an hypothesis. Thus begins the process of target validation.

The process of target validation

Possibly the two most precious commodities in the pharmaceutical industry are scientists and time. Allocating both to a speculative project represents a significant investment in that idea and, therefore, it is essential that the process is completed as thoroughly as possible and in as short a time as possible. To achieve this, it is important to establish a plan in which go and no-go points are established. This is important in that scientists are experts in designing experiments, but perhaps often are not so good at letting go. Selling your idea to your peers can be a daunting experience, but the better supported your idea, the easier (in theory) this should be.

In these early stages, it is unlikely that you will have massive amounts of experimental data, unless it has arisen out of a programme that is already running, or there are scientists who perhaps have too much blue sky on their hands. I would propose that the majority of early stage ideas in the Pharmaceutical industry come from speculative ideas suggested, or inspired, by the literature read by the scientist. It takes the trained eye (or brain) to pull together associations from disparate findings in the literature to synthesize a working hypothesis. This hypothesis can then be expanded from this kernel by further examination of the scientific literature and can form the basis of a project intended to establish the validity of your idea. This is, in essence, the cornerstone of all scientific method.

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2011.11.010

Tools for preparing a literature precedent for your project

Initially, therefore, one of the most important parts of the process is establishing a literature precedent for your idea; one of the first steps in target validation. Even with all of the tools available to the scientist, trawling through the literature is a time-consuming business and it is no surprise that companies have developed tools to help the researcher in this search. These tools really fall into a relatively small number of categories: manual review; automatic indexing; knowledge databases derived from manual curation; internally developed systems and the new kid on the block, Target Insights, which examines co-occurrence of key words and terms within articles when data mining. There are a number of criteria that are of importance to all of these approaches with respect to value and utility and these include: just how comprehensive is the search? How current is the search and/or database? How reliable are the searches with respect to concept indexing and taxonomy? Does the programme search full text, or just abstracts? Do the respective systems lend themselves to user-friendly interfaces?

Manual approaches

I guess every scientist has gone through the process of manually searching the literature, by whatever means they are most familiar with. This has the advantage of being cheap on first examination, however, when you consider that this takes away one of the most valuable resources of the Pharmaceutical industry from their key role – making drugs, perhaps it is a very false economy. Also, the scientist will not have time to go through the full text of the articles that he identifies, so generally the information derived will be only from the abstracts of these articles.

Automatic indexing

Moving on from this approach are the systems that rely on automatic indexing, such as those produced by Temis and Luxid. These have the advantage over manual searches in that they (by definition) index automatically and hence are more current. The systems are more reliable than manual searches but the quality can vary. Obviously, since the system is automated, the content is current. Such systems also search in abstracts but also some full text, hence an improvement upon the manual approach.

Curated approaches

Knowledge databases based upon a curated approach such as Current Data or Biosys in many respects have much in common with the automated indexing products, but obviously they rely upon a curated approach. This introduces an element of expert opinion and experience into the search process, but it suffers in that this process takes time and hence the database suffers from being less current than the automated approaches.

In-house development

Self developed systems generally will be developed according to the users' particular needs, as a result it would not be surprising to find that they are generally as good as, or better than the commercially available automated and curated systems. Of course, developing an in-house system solution takes time, resource and money and therefore may be an approach that is not adopted by all.

Targeting Insights

Finally, a new product, targetinsights, has recently come on the market. It has several advantages over all of the other previously mentioned systems. It draws upon a comprehensive database of Medline, Embase and conferences and, being automated, is highly current. The system is very reliable, and allows the use of off-the-shelf and bespoke taxonomies. The content indexing is quality assured and all of the recall is from full-text articles; the relation indexing is one of this product's strengths through its use of co-occurrence.

Let us face it, drug discovery is a risky business and everyone involved in it should be attempting to reduce that risk wherever possible. At the target identification and validation stage, the risk is in committing resource to a programme, whose chances of success are unknown. The better and more comprehensive the information derived from literature searches, the lower the risk. The ability to connect disparate information relating gene:protein:pathway:-pathology, increases confidence and lowers risk. Products capable of performing both tasks well decrease risk synergistically.

Effective literature searching can help in hypothesis generation and can strengthen a proposal significantly. Obviously, it can never, on its own, deliver a project, but it can allow ranking of hypotheses and, hence, permits advancing only the most credible projects. By developing the most credible cases in the shortest time, both risk and cost are reduced and scientists are freed up to do the thing scientists do best – science. Taking these factors into account, the question you should ask is not: can we afford to have these systems, but can we afford not to have them.

Steve Carney
Editor, Drug Discovery Today