

# THE INFORMATION SYMPHONY

data orchestration for a changing industry

*managing data on a platform for innovation*

A leading ballet conductor<sup>1</sup> recently told me of courses which he runs for senior business leaders. He allows them to conduct an orchestra: first, very strictly, with close direction applied to each instrument; he then shows them how to conduct the same score but harnessing the artistic talent of each group of players, encouraging active listening to each other – and then adding a conductor's overall interpretation. This model of active collaboration between multi-skilled players on a managed platform provides a fascinating and revealing insight that has great relevance in the life sciences, particularly in today's world of change.

Enough has been written about the demise of the blockbuster model. What is genuinely exciting about today's pharmaceutical world is the response to today's environment; the changing faces of pharmaceutical discovery, trials, regulation and clinical practice; the importance of clinical outcome over regulatory approval and the ability for stratified clinical trials which put NDAs into the reach of privately funded biopharma.

Most roads to the future of drug R&D involve – or rely upon – collaboration, a term that is easy to say and much, much harder to achieve. Active collaboration is required between internal groups, between internal and increasingly externalised, globalised partners in risk-share or discreet relationships. It is breaking out between biopharma and the newly empowered clinical stakeholders in real-world translational medicine; it is being developed in the new worlds of diagnostics and sequencing and also between regulators and payers.

In many cases these are communities newly introduced to one another and used to working in relative isolation. Bringing them together to revolutionise an industry in less than five to 10 years is radical but it is happening right now. In an environment that is information rich, context-heavy and IP dependent, effective collaboration is only possible if the information they share is accurate, rapidly available and consumable. Modern, managed data ecosystems are intrinsic to the success of this revolution. Fortunately there is light at the end of the tunnel combining scalable domain-aware data systems with the extensibility of the cloud.

The challenges faced by the global life sciences and pharmaceutical industries are well known: spiralling R&D costs, regulatory bottlenecks and patent expiries. These challenges are well reviewed in the excellent 'The Creative Destruction of Medicine' by Eric Topol, Director of the Scripps Translational Science Institute. However, today's

**By Chris Molloy**

challenges reflect an industry more in flux than meltdown. While the blockbuster model may be at the end of its useful life and a number of significant patent cliffs have been leapt in recent months, the rise of new business models and more outcome-driven, patient-centred approaches to R&D and clinical practice are providing the industry with new routes to revival.

### **New business models in R&D**

Pharma, contract research and biotech have, until relatively recently, had well understood, traditional roles to play. Mega-pharma would provide the high capacity internal research capability and selectively harvest innovation from biotechnology; they would also outsource what activities could not be provisioned internally with the aim of providing a solid pipeline of clinical candidates to its well established development and commercialisation groups. All was well.

Although R&D groups delivered process improvement and optimisation, individual areas were able to remain relatively siloed. In addition, process optimisation and innovation often pull in different directions.

Recently, however, there has been significant pressure for both greater efficiency and innovation. Chris Viehbacher, Sanofi's CEO, recently said the firm has realised that, "major groups are not great sources of innovation<sup>2</sup>." In a presentation at CED Life Science Conference in Raleigh, he continued: "On average, studies have shown that if you spend a dollar on research and development it will return 70 cents<sup>3</sup>."

An additional challenge comes from the drive towards precision therapies directed toward smaller focused clinical cohorts. Using adaptive trialling and defined, biomarker-driven selection, late-stage clinical trials can be performed on an unprecedentedly small number of patients. This brings the possibility of biotech companies developing assets through to NDA without the need for licensing at clinical proof-of-concept (or earlier) stages.

These pressures have led to the development of more flexible, externalised and agile interpretations of the previous pharma model.

### **Internal collaboration: improve efficiency**

The requirement to make internal groups work more effectively across interfaces is a shared goal across R&D. The traditional 'stage-gated' approaches to interdisciplinary working are being rapidly dismantled in favour of a more integrated, multidisciplinary approach. This recognises that

in order to progress projects quickly there needs to be constant and effective collaboration between multiple groups all the time and in real-time, not in a traditional serial manner. This is a very similar concept to the cellular manufacturing theories that have been used in other high-tech industries for many years. However it is a critical change because it brings with it the understanding that each group working on a particular project must each have real-time access to the same high-quality information, enabling everyone to actively access the 'single point of truth' that makes collaboration really work.

What is changing is the document-driven approach to working in favour of a data-centric one. The production and consumption of reports – which are inevitably a distilled version of high-context, data-rich experiments – do not provide the effective communication necessary to make the process of science work effectively. They reflect human effort in compressing information, transmitting that information and then – in many cases – having to supply further context on demand. Document stores, like paper laboratory notebooks, are often archives of information: poorly – if at all – searchable and not effective as genuine tools of scientific communication.

The requirement to communicate and store documents will not go away, however the consumption of data is increasingly about getting real-time access to a personalised view of the data each researcher needs to do their job, or contribute to the discussion.

### **External collaboration: harnessing innovation**

The issues of internal collaboration are highly magnified in the ever-expanding environment of externalised R&D. As pharma embraces externalisation and makes itself increasingly 'virtualised' the ability to communicate effectively with a network of suppliers, partners and IP producers becomes even more critical. In essence this is a mode of global data manufacture and needs to be treated as such.

The approach to externalisation has been broad. The move 10 years ago to allowing corporate screening collections to be sent to third parties signalled a disruptive change in the way R&D was to be executed. It was followed by the use of contract research companies (CROs) on a fee-for-service basis and limited early-stage collaborations with biotech and academic parties. This in turn was superseded by the so-called 'big-brother' deals where pharma semi-funded third-party

organisations while retaining options to R&D pipelines. As available clinically-developable assets begin to reduce, the balance of power can shift away from the pharma companies to the suppliers who are now able to demand significant risk-share arrangements allowing them to benefit in any potential upside.

That said, the fee-for-service market is global and growing<sup>4</sup>. This article will not deal with the economic arguments for or against, but it is notable that high quality CROs (such as AIT) are now being established in the USA. Extending from this expanding market are some exciting new information-based business models such as Assay Depot ([www.assaydepot.com](http://www.assaydepot.com)) which is pioneering brokerage platforms for the world's available assays and techniques. This model also reflects the power of aggregation of useful data, of which much more later on.

Externalisation has now also extended, through initiatives such as Innovative Medicine Initiative (IMI<sup>5</sup>, Pfizer's CTI<sup>6</sup> and Lilly's PD<sup>2 7</sup>) and other approaches to models that would have been heretical in the past but have been pioneered with other industries, namely pre-competitive and open innovation where ideas and even compound/biological assets can be shared with third parties<sup>8</sup>.

These changes represent a number of steps forward for innovation, and can be seen as being a managed crowdsourcing of invention. However, the success of the complex relationships formed, broken or extended relies very heavily on the ability to harness the flow of information across the ecosystem and harvest the IP that will be its product. Each relationship should be able to provide complete capture of the data, information and ultimately knowledge that drives the productivity of R&D. In many cases this is still being handled in an ineffective manner that can lead to loss of IP and an ineffective use of externalised resources.

### **New technologies**

The 12 years since the sequencing of the human genome have not delivered on the promise of precision medicine. There has been no shortage of great science and insight into the role and regulation of the genome and environment. The understanding of epigenetics, investigation in the 'dark matter' of DNA and even the realisation that the genome itself can modify through life is a treasure-trove for science. However the combination of genomics, epigenetics, transcriptomics, metabolomics and so on represents an extremely complex, multivariate confounding and highly variable model for drug discovery, as Eric Lander (Director

of the Broad Institute) when asked in 2011 reflected: "How simple did you think it would be<sup>9</sup>?" The rapid development of science in this area has developed an understanding in researchers, regulators and life sciences professionals that the human is an extremely complex dataset and the information generated on it should be treated as such.

The avalanche of data across all R&D areas generated by the newest technologies – whether it be 3-D screening, higher definition nuclear magnetic resonance (NMR) or next generation sequencing (NGS) provide specific challenges. Not only is there a great deal more than ever before but the nature of the data, as with the disciplines that produces it, is diverse.

### **Outcomes**

Beyond R&D the world of clinical medicine is undergoing a volte-face from being about the drug, or the clinical intervention to being about the patient and the outcome. There is increasing clinical evidence than many therapies approved after extensive highly recruited clinical trials have limited efficacy real world of the clinical patient population. This is affecting the ideas and behaviours of regulators and pharmaceutical CEOs. In 2009 Severin Schwan (CEO, Roche) remarked on the uncertain success of many marketed pharmaceuticals: "Imagine a car that starts only half the time and whose brakes often don't work<sup>10</sup>." This thinking can also be seen in the approach of governments. In 2008, Professor Sir Michael Rawlins, Chairman of the UK National Institute for Health and Clinical Excellence (NICE), said that large-scale randomised clinical trials "long regarded as the 'gold standard' of evidence, have been put on an undeserved pedestal. Their appearance at the top of 'hierarchies' of evidence is inappropriate; and hierarchies, themselves, are illusory tools for assessing evidence. They should be replaced by a diversity of approaches that involve analysing the totality of the evidence-base<sup>11</sup>."

### **Translational medicine**

The ability to understand disease and the variability of response to treatment in the real world is a vital part of the renaissance of the life sciences model. For many years there has been talk of the concept of bench-to bedside approaches to R&D: a linear concept that starts with basic research and ends with approval. What real world represents is the logical extension to this model, delivering the more holistic, iterative, bench-to bedside-to-bench that enables the identification of

true unmet clinical need, more direct measures of treatment outcome through biomarkers and better patient stratification.

Selection of, and access to, real-world patient cohorts will be able to deliver more targeted, precision therapies but also presents real challenges.

Firstly there are many new stakeholders in this new ecosystem including academic medical centres, primary care physicians and – most importantly of all – the patients themselves.

Academic medical centres such as King's Health Partners (UK) and Dana Farber (USA) are a vital part of this new ecosystem, providing high levels of academic insight and access to real world patients. At King's Health Partners, for example, researchers directed by Professor Peter Parker, PhD, FRS<sup>12</sup> are using combinations of unique data integrations, improved process, patient consenting and biobanking to drive highly-collaborative patient-centric research projects. Secure use of consented clinical data, tissue and sample data across thousands of patients as well as research information is critical. He says: "We are delivering the best treatment choices today and the best opportunities for tomorrow by making better use of patient data across our partnership. [It] will actively and positively impact patient care in the development of innovative cancer treatments and prognostics." Using this approach researchers can gain insights into outcomes of past and existing patient cohorts as well as identify new patients for longitudinal studies.

### Digital human

The convergence of technology, data capture and medical practice is highly disruptive. As Eric Topol says: "Medicine is about to go through its biggest shakeup in history<sup>13</sup>." The increased availability of image, genomic and proteomic information is an important part of this 'real-world' puzzle. Whether these are used to track or provide prognoses of disease or to direct therapy, the so-called 'companion diagnostics', they represent elements of the patient pathway from presentation to outcome. Organisations such as Quest and Genomic Health are making extensive use of genomic and real-world patient technologies to identify and validate effective therapies.

### 'Big data'?

How does this seemingly diverse set of approaches all come together? Like musical notes on a page the answer in many cases is through data. In a recent report on the future of the pharmaceutical

industry, Ernst & Young stated that "information is the currency of Pharma 3.0<sup>14</sup>" and that companies must improve their ability to "extract value out of large volumes of data from diverse, unfamiliar sources".

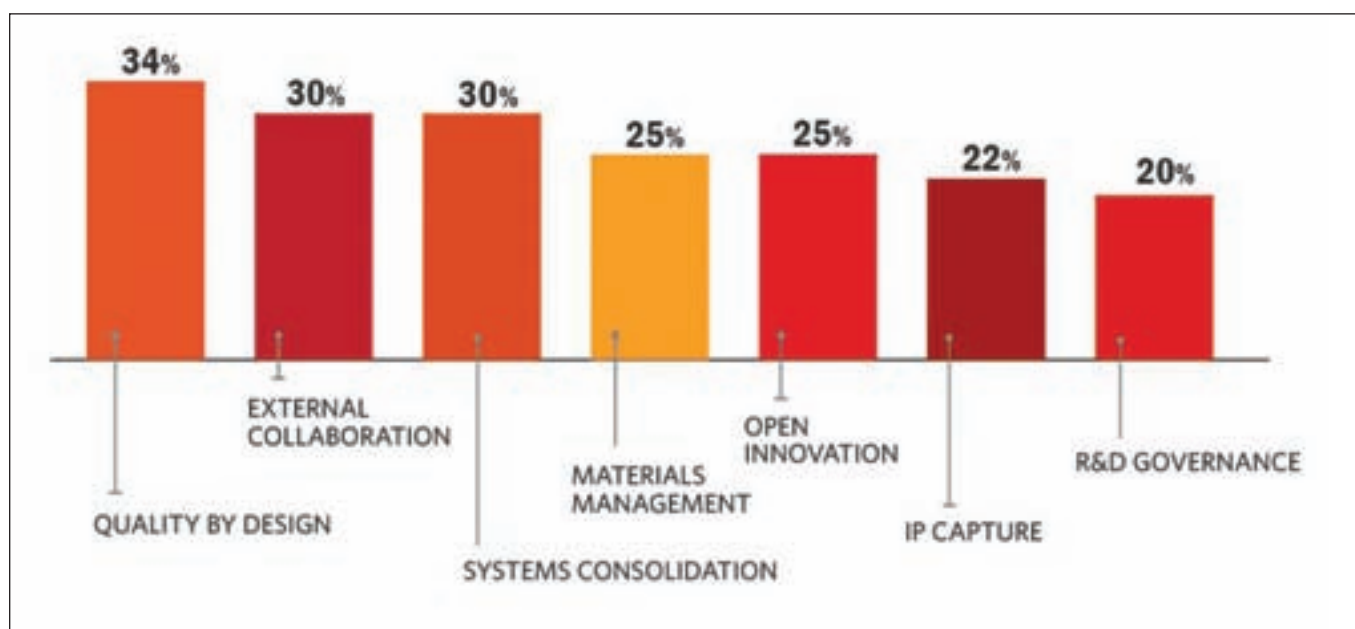
Data and information management is not something new to the pharmaceutical industry: it has pioneered the adoption of informatics technologies such as laboratory information management systems (LIMS), Laboratory Execution Systems (LES), medicinal chemistry electronic laboratory notebooks (ELN)s and more recently multidisciplinary ELNs, an excellent review recently provided by Peter Boogard and Pat Pijanowski<sup>15</sup>. However it also labours under the weight of a multitude of in-house 'point solutions' to accommodate for past gaps in vendor offerings or (perceived) company-specific processes. The positively disrupted pharmaceutical environment needs more joined up informatics.

There is plenty of energy around 'big data' but the term remains nebulous and, like the cloud can be a catch-all which either appears to solve everything, or be so ill defined as to be meaningless. The idea of 'Big Data' does not deliver directly to the business. One has to look at how data will influence business improvement.

### Data enables collaboration

R&D exists in a data ecosystem, with each researcher and group reliant upon many others to generate the data (and thereby information) needed to progress projects or drive innovative thinking. This data cannot be easily passed via static or even electronic documents. For effective use and scientific challenge information consumers must be able, should they have need, to drill down to past interpretations into the originating data and be able to access the context behind the data. They should be able to bring together – by simple queries – the information from multiple sources to make decisions and should be able to see the same 'single point of truth' as every other researcher who has access to that information. They should also be able to do this in real-time. This is a concept shared with most enterprise resource planning (ERP) systems and is a basic principal.

In an externalised network the same concepts of timeliness and ease of integration of data remain true but with an even greater emphasis upon security. It is vital that each collaborating party only sees the information to which they have access: it that the security privileges should match the collaboration agreement.



### Data enables IP

Data, and the information it generates, are the bedrock of the IP of any life sciences company, but the ground is shifting here too. Firstly, IP emphasis is turning away from pure ownership of the chemical (or even biological) substance towards its therapeutic use. Secondly, what represents valuable IP is also changing, to the point where methods of analysis, manufacture, markers and measurement are now equally as important as API in a holistic treatment regime. Finally, as the America Invents Act becomes law, the concept of 'First to Invent' is being superseded by the 'First to File' principle.

What does all this mean? It means that rapid access to enterprise data and the context in which it was generated is now more important than ever. It requires R&D organisations to capture their raw data, means of calculation and context, ideas and innovations across multiple disciplines and across collaborative networks. It needs to reside in places where it can be accessed and integrated quickly and filed before anyone else irrespective of the date the work was done.

### Data enables quality

The impact of quality-by-design, intrinsic in other manufacturing industries is now a feature of intense interest across pharma and made more so by recent increases in FDA warnings. This area is expertly covered by Industrial Lab Automation<sup>16</sup>. Yet again the importance of the structured access to current and historical data, and the use of data

generated in one area to enhance the quality of output in another, is explicit.

Another area where quality is able to be improved is within the collaborative process where data quality issues become important. Just as with social media where a 'like' or 'dislike' can influence the interpretation of a message, it is important to be able to actively curate data for quality so that – as part of the food chain of data – it can be quality assured and weighted appropriately. This is an important part of the work in using the multivariate data within real-world clinical study.

### Data enables real world clinical study

The volume and complexity of data is the defining feature of real world clinical study and the future of healthcare. It is axiomatic in all areas that data must be usable in order to be useful yet this is the critical feature of the success of this endeavour. In recent strategic proposals from both the NIH and the Human Genetics Strategy Group, the importance of managing genomic data to enable collaboration and improved insights was key.

Genomic Health uses a combination of gene expression technologies to build sensitive and quantitative gene expression profiles as the foundation for new oncology diagnostics. With increasing volumes of high value data being generated, the company has been developing a complete data management layer to underpin its R&D process, enabling researchers to find real-time information from across the organisation and drive wiser decision-making.

**Figure 1:** Survey responses 'R&D informatics focus areas for 2012'

Bringing together secured clinical health record data with tissue and biomarker research information in a way that is consumable has been the enabling feature of initiatives such as that at Kings Health Partners, Dana Farber and others. The secure infrastructures and understanding of ethical and patient concerns is also critical. Until recently no production-ready systems were available to support this research. The availability of the Enterprise Translational Medicine System (ETMS) in 2011 represents a breakthrough in this important area.

### The response

The case for more enterprise use of data is one that is well recognised. Werner Boeing (CIO, Roche Diagnostics) believes that: “[By 2016] the CEO of a pharma company should be uncomfortable if the CIO is not present at a key strategy meeting. That’s how important IT needs to be for pharma strategy<sup>17</sup>.”

R&D organisations are looking to improve their information flow. A survey of 682 researchers by IDBS and Scientific Computing in 2011 asked what issues were on the minds of R&D organisations. The response was telling (see **Figure 1**). Collaboration models and quality feature highly but the drive to reduce the number and complexity of siloed systems remains in the minds of executives, research leaders and IT staff.

Many inside and outside pharma look to the cloud as a solution. It is certainly part of it. The ability to provide extensible computing power across a distributed collaborative network is well-established. However, ‘big’ is only beautiful if it is well managed. The solution of ‘put it all in the cloud’ is far more easily said than achieved and critical to making the cloud work is to have scalable platform applications capable of structuring and curating the data within the environment and upon which applications, collaborators and other data sources can feed. This combination of cloud-scale and system-control was one taken by Shire plc<sup>18</sup>.

R&D organisations are looking to make data and processes interoperable across their organisations – bench-bedside-bench. This requires a thorough review of how information flows across the collaborative space, both internally and externally, and how the organisation is going to consume real-world clinical and translational data. In most cases they are looking to reduce niche solutions in favour of more enterprise platforms capable of multidisciplinary working. The multiple systems that data often flow through to be consumed define a highly

fragmented, and often legacy, landscape of niche chemistry notebooks, LIMS, LES and manufacturing quality assurance systems.

Ron Shoup, Executive Director at AIT Bioscience, explains: “For several years we had successfully used a LIMS to build study designs, document method parameters and consolidate instrument data into final reports. While this gave us a direct view of a study as would be placed in a final report, some critical information remained in paper format. For example, important data relating to the validation of instruments and software, staff training records, QA audits, metrology data, and information surrounding reagents remained outside our digital environment. We therefore looked for a way to augment our existing informatics footprint and move towards a fully digital environment.”

What can now be delivered for the business user is single-source, secure access to their entire data landscape just in time and at the point of use; the ability to integrate third party systems – including external content and newsfeeds – and to be able to develop new ways of interacting safely with the data to drive even greater levels of innovation.

In a similar way as pharmaceutical companies are now focusing on the treatment not just being about the compound, the data solution is no longer just about the lab (the bench), it is about the whole enterprise and being joined up across the organisation, harnessing everyone’s talent.

A new generation of R&D informatics system is required: one that extends across disciplines and across collaborative boundaries, one that orchestrates data and can apply the business rules that define how it gets used. More importantly the platform should enable the applications that use the data to be interoperable, across desktop, browser and mobile environments. In short, a data platform for the new pharmaceutical industry. This encompasses and integrates authentic cross-functional enterprise ELNs but also a much more holistic use of data and information as an asset. The important news is that this now exists.

As Jay Galeota (SVP Strategy and BD, Global Human Health, Merck & Co) said: “Pharma is now in the information business and its part of everyone’s job function<sup>19</sup>.”

### Summary

Bain and Co stated that: “The blockbuster business model that underpinned ‘Big Pharma’s’ success is now irreparably broken. The industry needs a new approach<sup>20</sup>.” The world of finance may have lost some of its gleam over recent years, and

## Informatics

### References

- 1 <http://www.benpope.com/Site/Welcome.html>.
- 2 PharmaTimes Feb 21, 2012.
- 3 Fierce Biotech, Feb 16, 2012.
- 4 Nature Reviews Drug Discovery 10, 561-562 (August 2011).
- 5 <http://www.imi.europa.eu/>.
- 6 [http://www.pfizer.com/files/research/partnering/cti\\_brochure\\_9x12\\_v12single.pdf](http://www.pfizer.com/files/research/partnering/cti_brochure_9x12_v12single.pdf).
- 7 <https://openinnovation.lilly.com/dd/>.
- 8 Nature Biotechnology 29,1063–1065(2011).
- 9 Technology Review January/February 2011.
- 10 The Economist, December 12, 2009.
- 11 Rawlins M. DE TESTIMONIO, On the evidence for decisions about the use of therapeutic interventions, THE HARVEIAN ORATION 2008.
- 12 Head of the Division of Cancer Studies and R&D Lead for the Integrated Cancer Centre.
- 13 The Creative Destruction of Medicine, Eric Topol 2012 (<http://createdestructionofmedicine.com/>).
- 14 Progressions: building Pharma 3.0, Ernst & Young, 2011.
- 15 GIT Laboratory Journal 11-12, 2011, p14-16.
- 16 Scientific Computing, December 2011, p9-11.
- 17 Progressions: building Pharma 3.0, Ernst & Young, 2011.
- 18 <http://www.idbs.com/Data-Management-News/press-release/10APR14.asp>.
- 19 Progressions: building Pharma 3.0, Ernst & Young, 2011.
- 20 WINDHOVER INFORMATION INC. Vol. 21, No. 10.
- 21 Nature Reviews Drug Discovery 11, 91-94, February 2012)

not have all the answers, but it, like music, still has much to teach the life sciences sector about the use of data as a capital asset. Domain-aware information platforms sitting on the cloud enable them to orchestrate their institutional knowledge and put it at the fingertips of decision-makers and analysts in real-time.

It is true that life sciences R&D must become more productive and more innovative – in fact more innovative than it has ever historically been – to reverse the current trend of cost-per-marketed pharmaceutical, now estimated at almost \$4 billion. The good news is that the changes being made in targeting disease may be starting to work. Last year the US Food and Drug Administration's (FDA) Center for Drug Evaluation and Research (CDER) approved 24 new molecular entities and six new biologics. The approval of 30 new therapeutics is the most since 2004<sup>21</sup>. The extension of pharma across small molecules, large molecules, diagnostics, biosensors and into real-world clinical study and 'community' patient engagement recognises the commitment to integrated treatment portfolios and the journey towards the digital healthcare revolution.

In life sciences, where our goals are for better outcomes for more patients we need access to high context, high quality, integrated data from basic research through to clinical practice. This platform thinking, enterprise-wide and collaborative, will play a major part in developing a symphony from the complex dataset that is human biology. It can also do much to orchestrate a new movement in pharma R&D.

DDW

---

*Chris Molloy is VP, Corporate Development, for IDBS, a leading global provider of enterprise R&D and Healthcare informatics. Chris was formerly COO of MerLion Pharmaceuticals and has more than 20 years' experience in pre-clinical research and clinical development across large pharma and biotech as well non-executive and advisory positions in life science recruitment and specialist IT companies.*