Pharma and the seven keys

Over the next decade, critical informatics technologies will revolutionize the drug industry.

BY RANDALL C. WILLIS

Skyrocketing development costs and changes in discovery paradigms—both technological and biological—are leading pharmaceutical companies to reexamine their business models. And nowhere is this better seen than in the application of information technology to various business functions.

"Information technology is at the heart of enabling the transformation of the pharmaceutical industry," says Steve Arlington, global pharmaceutical industry leader at IBM Business Consulting Services (BCS). "Now is the time for the industry to capitalize on the huge scientific achievements of the genomic era. To do that, companies need to invest in new technologies that will truly drive breakthrough growth and help them to differentiate themselves."

And Arlington warns that companies failing to respond to emerging market conditions will find their shareholder values will plummet. Thus, to assist companies in looking forward and understanding the new technologies that await them, IBM BCS analysts recently published a report, *Pharma 2010: Silicon Reality,* identifying seven key information technologies they believe will drive innovation in the pharmaceutical industry (*1*).

A new era

With the expected onslaught of generics in the near future and the steady shrinkage of operating and profit margins, the pharmaceutical industry is taking a much closer look at its information technology expenditures during the past two decades and is beginning to question the promised benefits. According to the IBM report, the poor return on investment may be the result of companies looking for technologies that perform more functions rather than those that help them make sense of the data they already possess. Other companies, however, might be facing problems resulting from poor strategic objectives, the inability to integrate applications seamlessly, or poor communication with technology end-users. Regardless of the source of the challenges, companies are becoming increasingly cost-conscious.

Unfortunately, this newfound frugality is running in the face of the slowly evolving biological paradigm of targeted treatment solutions, where health care focuses on the individual patient or specific subpopulations. This new therapeutic regimen has extensive (and potentially expensive) repercussions throughout the pharmaceutical value chain. Drug development will rely heavily on a wider spectrum of patient information derived using genomic, proteomic, metabolic, and epidemiologic technologies. Likewise, drug marketing and sales will significantly change as companies move from a "one-size-fits-all" model to "customerization." At each stage, the amount of information required and produced will be potentially overwhelming and, therefore, success in moving to the new model will rely heavily on the implementation of key technologies (Figure 1).

Selling the seven keys

In examining the pharmaceutical industry, the IBM analysts identified seven technologies they believe will dictate progress and success over the next decade. These are:

- petaflop and grid computing,
- predictive biosimulation,
- pervasive computing,
- smart tags or radiofrequency IDs (RFIDs),
- advanced storage solutions,
- process analytical technology (PAT), and
- Web-scale mining and advanced text analytics.

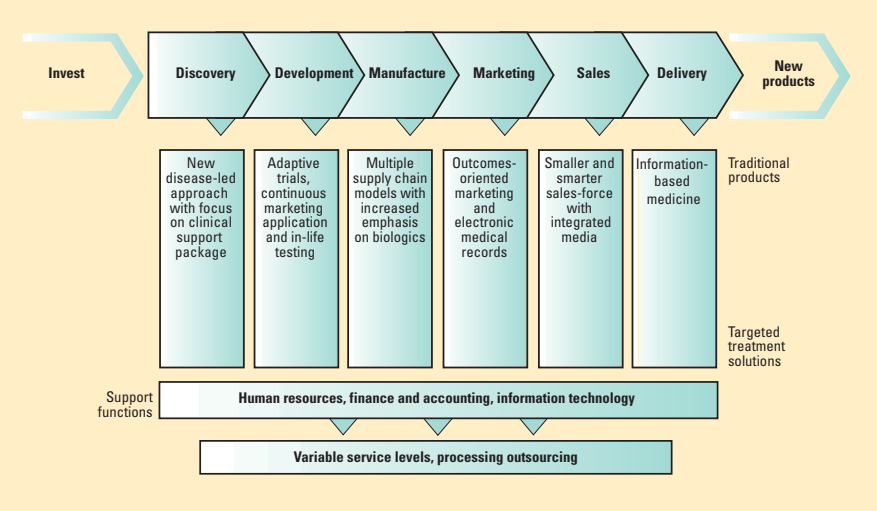

Figure 1. A new view. Financial and regulatory realities will require the pharmaceutical industry to develop a new business model, according to IBM analysts. (Adapted with permission from *Pharma 2010: Silicon Reality* by IBM Business Consulting Services.)

Computational power is often described in terms of the ability to perform rapid calculations or in floating operation points per second (flops). At present, the most powerful computers operate in the teraflop range, but IBM analysts believe this will not be enough to handle the complex calculations required by the pharmaceutical industry. Instead, they see the advent of petaflop computing as enabling drug developers to compute biomolecular simulations, such as enzymatic and protein-folding reactions. This need is behind IBM's impetus to produce Blue Gene, a supercomputer it expects

to complete in 2005 that will function at 6 times the speed in one-tenth the footprint of today's supercomputers.

Not all companies will be able to afford such powerful machines and are therefore more likely to invest in a system known as grid computing, where complicated calculations and functions are broken into discrete packages and distributed to a network of computers (*2*). Such systems will allow companies to harness the unused or underused computing power of idle computers both within a company and over the Internet. As an example of grid computing in action, the analysts offer the Smallpox Research Grid, which was established in February 2003 and relies on more than 2 million home computers to screen millions of potential drug molecules against eight smallpox protein models.

One of the challenges of modern computing in biology and pharmaceutical development is taking static flat data and transforming it into a dynamic model that shows how each component interacts. According to the IBM analysts, part of the problem stems from the fact that—aside from simply combining genomic, proteomic, and metabolic data into a comprehensive model of healthy and diseased tissues—it remains difficult for researchers to "distinguish changes that cause a disease from those that are caused by the disease."

This is where the predictive biosimulation efforts of academic institutions and companies such as Entelos and Physiome Sciences will become prominent (*3*). By developing in silico models, the analysts suggest, researchers can integrate relevant data, reproduce control systems, and simulate how the biological systems will respond to perturbation. Thus, they "can identify potential molecular targets and compounds as candidates for treating disease."

"Virtual patients will thus become the crash dummies of the life sciences industry," the analysts state. "They will improve target validation, reduce lead times and attrition rates, and help make testing in man very much safer—both in the laboratory and in life."

Pervasive arguments

Another critical component that will increase the likelihood of future pharmaceutical success, according to the analysts, will be pervasive computing, where everyone and everything involved in drug discovery, development, and marketing will be connected through an all-encompassing communications network. In part, such technologies will help companies ensure that they maintain accurate records of manufacturing and usage throughout development and testing, and thereby will allow them to better comply with regulatory requirements, such as the Sarbanes–Oxley Act of 2002. As proof, the IBM report describes the efforts of Aardex, a company that has developed a bottle monitor that records information about when a bottle is opened and thus indicates how well a patient is complying with his or her doctor's prescription.

"Virtual patients will thus become the crash dummies of the life sciences industry."

A rate-limiting step in the application of pervasive computing, however, will be bandwidth. The advent of networks incorporating wireless or WiFi technology will facilitate many computing systems, but transmission speeds and data encryption continue to be issues. The analysts, however, believe the development of ultrawide-broadband technology, which uses brief, low-power pulses of information across a wide spectrum, should alleviate some of these problems and allow the transmission of hundreds of megabytes per second over short distances.

Another critical component of pervasive computing is smart tags or RFIDs. These are microchips applied to products and carrying unique identification and other information as well as a small antenna to transmit this data to various handheld or built-in devices. More than glorified bar codes, RFIDs can typically hold hundreds of characters of information and will allow users to trace a product's life cycle. Aside from IBM, Sun Microsystems has invested much energy, time, and money in RFID implementation. And, as a sign of how rapidly RFIDs will become part of the industrial complex, Wal-Mart recently requested that by January 2005 its top suppliers apply smart tags to all pallets and cases shipped to the chain's distribution centers and stores.

Data storage is also a complicating issue for information technology, as standard magnetic media and optical storage systems (CDs and DVDs) are quickly becoming swamped by the sheer volume of data being produced in research centers throughout the world. But as the 1967 movie *The Graduate* foresaw, plastics is "the word." According to the IBM report, IBM Research has been working on a system called Millipede, which uses thousands of tiny needles to create nearly atom-sized marks on thin plastic films. Used as a binary medium, Millipede functions like a nanotech punch card but can be used for up to 10,000 read–write cycles and has a capacity of 1 terabyte per square inch.

Likewise, the IBM report indicates that storage area network (SAN) file servers, storage virtualization systems, and storage grids will become more critical as the need to access the burgeoning information increases. SAN file servers allow users to access information from any location by serving as a common language through which files are translated. Storage virtualization separates storage from data analysis, allowing users to establish virtual disk space in which to work. And, like grid computing, storage grids host data on both internal and external facilities while maintaining a virtual pool.

Proof in process

As mentioned earlier, a key driver of this move to seamless integration and large-scale data conservation is the various regulatory initiatives proposed or slated for the near future. For example, the FDA's PAT initiative recommends, and might shortly require, companies to monitor their manufacturing processes in real time, both in situ and in-line, rather than at specific checkpoints with aliquoted samples (*4*). The FDA's goal is to ensure that all manufacturing processes are stable throughout their runs and produce the desired products.

Although complying with PAT will be expensive in the short term, relying as it does on a complete overhaul of the QA/QC process, the new system is ultimately expected to save companies money by allow-

ing them to adjust or stop production faster when things go wrong. This would limit the amount of material destroyed to ensure that no faulty product makes it through the system. In addition to relatively straightforward replacement of analytical instruments, such as mass spectrographs and chromatographs, PAT compliance will require companies to implement real-time data analysis and decision-making software. Not surprisingly, the moment-by-moment data produced in each production run will push most current data storage systems to the limit.

Digging for data

Because a company can only produce so much data and relies on its core competencies, it is critical that it be able to access the plethora of information on the Internet and similar channels. Thus, companies are developing and using algorithms to mine data from disparate public sources, including e-mails, news groups, and bulletin boards, as well as more formal sources such as Medline. The IBM analysts compare Web-mining to drinking from a fire hydrant, because 80% of the data available on the Internet exists in unstructured formats, and the challenge of new data tools is sifting through this morass to find the salient nuggets.

"The IBM analysts compare Web-mining to drinking from a fire hydrant."

Scientists are already using data-mining tools such as these to identify distant linkages between new drug molecules or genes and various disease states or animal models, allowing them to map out complex relationships without necessarily reproducing experiments performed elsewhere. They hope to develop multiple markets for specific drugs or avoid unforeseen complications and adverse reactions that can complicate the clinical screening process and are responsible for many drugs being pulled from development.

Thus, say the analysts, while the pharmaceutical industry currently spends inordinate sums on information technology with little or no compensation, these new technologies should enable companies to differentiate their products and processes, and better allow them to function and succeed in the more complicated, heavily regulated environment in which they find themselves.

References

(1) *Pharma 2010: Silicon Reality;* IBM Business Consulting Services; May 5, 2004; available at www.ibm.com/bcs/pharma2010.
(2) Crafford, C. *Modern Drug Discov.* **2003,** *6* (11), 25–26.
(3) Willis, R. C. *Modern Drug Discov.* **2003,** *6* (2), 35–40.
(4) Schmidt, C. S. *Modern Drug Discov.* **2003,** *6* (6), 51–52. ■