The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0305-5728.htm

Knowledge capture within the biopharmaceutical clinical trials environment

Martin Grossman
Bridgewater State College, Bridgewater, Massachusetts, USA, and
Stephen Bates
Perceptive Informatics, Inc., Waltham, Massachusetts, USA

Abstract

Purpose – The purpose of this paper is to provide an overview of knowledge capture in the biopharmaceutical industry, focusing primarily on the transition from paper-based to electronic data capture (EDC) systems.

Design/methodology/approach – The paper draws on biopharmaceutical industry literature and data from example clinical studies to describe the issues involved in transitioning to EDC in the clinical trials environment.

Findings – While electronic data capture systems provide greater efficiencies along the clinical trial supply chain, the industry is still far from achieving wide scale utilization of such technologies. The barriers to successful implementation are multifaceted, involving not only the information technology itself, but also user acceptance issues, lack of interoperability standards, and regulatory compliance. Major shifts in organizational culture and a unified effort within the industry will be necessary in order to derive full benefits from electronic capture systems in the future.

Research limitations/implications – This study was limited in that case data from only one company was used to supplement the literature review. Further research is warranted to better understand the factors that facilitate adoption of electronic knowledge capture systems in the biopharmaceutical industry.

Originality/value – While the need for knowledge management in the healthcare industry is indisputable, there has been remarkably slow progress in this area, and a dearth of research exploring implementation issues. The value of this type of inquiry is profound as it will help us better understand the issues in implementation and adoption, and ultimately to deliver more effective and safe drugs to the public in a more efficient manner.

Keywords Pharmaceuticals industry, Knowledge management, Trials, Drugs

Paper type Case study

Introduction

A number of critical factors, including an aging population, greater life expectancies and expanding markets into the developing world, have contributed to the phenomenal growth of the biopharmaceutical industry in recent years. Along with these external forces have come major advances in bioengineering, nanotechnology, and genomics, making the drug delivery process increasingly dependent on new and advanced information technologies. With a global drug market of $643 billion annually (IMS, 2007), the future would certainly seem bright for the biopharmaceutical industry. According to some industry reports, while the industry is still experiencing double digit profitability rates, the market is actually in decline. Several factors are at play, higher R&D costs, time delays and rising rates of downstream attrition, being among

VINE: The journal of information and knowledge management systems
Vol. 38 No. 1, 2008
pp. 118-132

0305-5728
DOI 10.1108/03055720810870932

the most critical (Bart, 2003; Carel and Pollard, 2003). Some estimates report that the costs of developing a new drug exceed $800 million and take an average of 11 years from creation to market, causing pharmaceutical companies to experience a pronounced "profitably gap" (Bart, 2003). As pressures on the industry continue to mount, quality and safety issues become a greater concern. The highly publicized recall of the COX-2 inhibitor Vioxx in 2002 is but one example of how the system can go awry.

In order to gain greater efficiencies along the drug discovery supply chain, biopharmaceutical companies will need to embrace more innovative methods. Perhaps the most important prerequisite to innovation in a knowledge-based economy is the ability to leverage the intellectual assets of the organization. Organizations around the world are turning to strategic knowledge management (KM) approaches to address this need. KM deals with the capture, creation, codification, sharing, access, application and re-use of knowledge. It is a paradox that while great technological strides have been realized in the biopharmaceutical sector over the last few decades, the utilization of KM has been slow to take hold in both the healthcare and clinical trials environments. Electronic data capture (EDC) is an emerging technology that can streamline the drug discovery process by allowing for the collection of clinical data in an electronic format, reducing much of the overhead and quality problems inherent in paper systems. While electronic data capture systems promise to modernize the clinical trial process, the industry is still far from achieving wide scale utilization of such technologies. The barriers to successful implementation are multifaceted, involving not only the information technology itself, but also user acceptance issues, lack of interoperability standards, and regulatory compliance. Major shifts in organizational culture and a unified effort within the industry will be necessary in order to derive full benefits from electronic capture systems in the future.

This paper provides an overview of data capture in the biopharmaceutical industry, focusing primarily on the transition from paper-based to electronic data capture (EDC) systems. The paper draws on biopharmaceutical industry literature review and data from example clinical studies to describe the issues involved in implementing EDC in the clinical trials environment.

Knowledge management in the drug discovery process

Well into its second decade as a unique discipline, KM continues to gain momentum within organizations across the world. KM has been adopted by such corporations as Accenture, Cable & Wireless, DaimlerChrysler, Ernst & Young, Ford, Hewlett Packard and Unilever (Rao, 2005) as well as by government and international development agencies (Malhotra, 2003). In the big pharma environment companies such as Bristol-Myers Squibb have embraced KM, realizing significant benefits as a result (Leavitt, 2002). In general, however, the biopharmaceutical industry has been slow to implement KM practices in the drug delivery context (Hodgson, 2001).

Effective KM requires several important processes involving the identification, acquisition and dissemination of knowledge to provide a strategic advantage. Dalkir (2005) distills these into the following three stages, which make up an "integrated" KM life cycle:

(1) Knowledge capture and/or creation.

(2) Knowledge sharing and dissemination.

(3) Knowledge acquisition and application (see Figure 1).

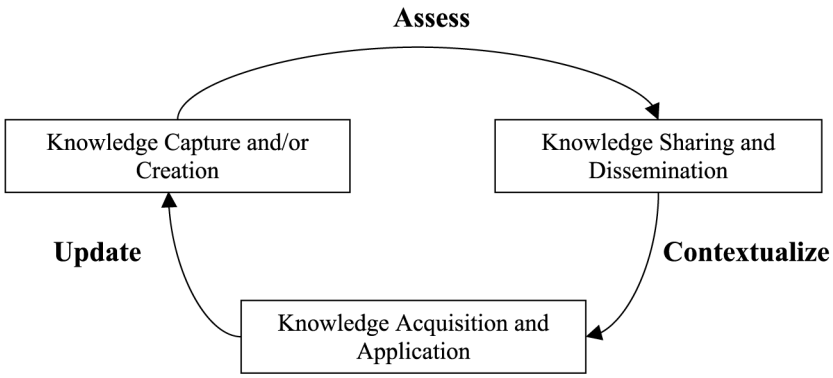

Figure 1.
Knowledge management cycle

Source: Dalkir (2005)

In this model, knowledge content is "assessed" to ensure that is of sufficient value before it transitions from the knowledge capture/creation stage to the knowledge sharing and dissemination stage. It is then "contextualized" as it flows to the next stage in order to be understood and used. Finally, it flows back to the first stage, as the knowledge content is "updated". Data from clinical trials are assembled and evaluated for safety and efficacy and used to help in the approval of new treatments, ultimately getting these medicines into the global marketplace to the patients who need them. It is heavily dependent on processes in each stage of the KM cycle. However, the early activities of the R&D pipeline, in which data is systematically captured, integrated and distributed, are particularly important as these determine quality of the knowledge as well as how efficiently it is disseminated and utilized by investigators. An important aspect of the clinical trials process is the recording of observations, patient behaviors, and environmental factors that might be perceived but not easily codified (i.e. tacit knowledge). Much information, for example, may be buried in laboratory notebooks, which contain not only specific details relating to experimental procedures, but also inferences and conclusions drawn from such activities (Carel and Pollard, 2003).

Effective KM becomes an even greater issue in clinical trials as newer biomedical techniques become more widespread, causing changes to the flow along the pipeline from the traditional linear to a more dynamic, iterative process where more is done in a parallel fashion (Carel and Pollard, 2003).

Data capture in clinical trials

There is much data that needs to be recorded in a clinical trial setting. The patient being observed possesses an untapped reservoir of usable information pertaining to the treatment under question. Quantitative data relating to bodily functions needs to be measured and recorded, as well as qualitative data based on observation. After the data is collected it needs to go through several waypoints as the information is validated, massaged and stored in a database, available for future retrieval and utilization. A unique aspect of this scenario is the requirement to validate all data as close to real-time as possible, as the source of the data (i.e. the patient), may not be able

to provide confirmation at a later date due to worsening health conditions or death. In this case, valuable data may be lost or rendered useless in a future analysis.

Traditional paper-based approach

In the traditional approach, a number of key actors are involved in capturing and recording data on paper-based artifacts. The typical workflow may involve a number of tasks performed in a linear manner (Sahoo and Bhatt, 2003):

- the Investigator records data on the case report form (CRF);
- the Monitor conducts the source data verification (SDV), collects and submits the CRF to the data management (DM) unit of the sponsor; and
- DM personnel complete the data entry, after seeking clarification from Investigators and staff and validating the accuracy of the entries. If problems occur, data clarification forms (DCF) are issued.

There are many inherent problems in such paper-based processes, including administrative costs, risks of damage and misplacement. As a result, response time and quality suffer in this scenario. Not counting the potential delay due to recruiting patients for the trial, a paper-based approach can add huge lag times to the clinical process, up to six to nine months (Sahoo and Bhatt, 2003).

Another drawback to the paper approach has to do with storage and retrieval. Important metadata, or data about the data, are not easily generated using paper systems. This limits the ability to search the data for future problem solving (Beyster *et al.*, 2005).

The promise of EDC

As biopharmaceutical companies become increasingly pressured to bring new products to market in a faster, more cost-efficient manner, the need for more innovative approaches becomes critical. Issues of quality and safety are also of great importance, increasing the frequency and length of clinical trials, and causing even more cost in terms of time and dollars. Over the last decade a growing number of companies have turned to EDC to alleviate some of these problems. Related to an older technology called Remote Data Entry (RDE), which were introduced in the 1980s and early 1990s, EDC started to gain recognition about a decade ago (Hyde, 1998, Pratt, 2007). The potential benefits of electronic approaches have been well documented. Kush (2006), for example, claims that EDC can reduce errors going from the investigator site to the biopharmaceutical company by between 70-80 percent and that it can cut database lock time to a matter of hours (as opposed to the four to eight weeks typical in paper method). Other advantages of EDC, extracted from a number of sources (i.e. Bart, 2003; Kush, 2006; Bushnell *et al.*, 2006; Sahoo and Bhatt, 2003), include the following:

- saving time and effort for the project team and monitors with regard to CRF retrieval, DCF generation, query resolutions, CRF transmittal, and courier tracking;
- elimination of initial process of data entry, data cleansing, and the manual query resolution process for the data management team;
- effective monitoring of the progress of the clinical trial from a central location;
- facilitation of quick reporting of events through web-enabled support systems;

- hastening of the time of drug development process and the reduction of the cost of drug development;
- increased business process throughput by automating data capture;
- secured indexed storage of original document images to legal admissibility standard;
- decentralized information capture; and
- reduction of paper chase by capturing paper forms and documents via a number of devices (e.g. scanners connected over the web, networked MFDs over web or WAN, remote fax machines, dedicated centralized production scanners).

While not abundant, there is empirical evidence to support the contention that electronic techniques provide time and cost advantages. Bart (2003) analyzes a number of early studies that compared EDC with paper-based approaches. Although varied in terms of context and experimental methodology, EDC revealed cost and times savings across all studies cited. In total, a cost savings factor of 5.8 was achieved.

Beyster *et al.* (2005) reports on a number of companies who have successfully implemented electronic capture systems and who have derived large benefits as a result. Wyeth eClinical, for example, saved millions of dollars and reduced trial duration by 18 months after implementing an integrated electronic Case Report Form (eCRF) system. Eli Lilly is another company who realized notable improvements (i.e. reducing the time between last patient visit and database lock to 24 hours) after implementing an internet-based solution during a clinical trial.

Bushnell *et al.* (2006) found that EDC versions of the instruments typically used in Irritable Bowel trials (e.g. IBS-QOL, EQ-5D, and WPAI:IBS) to be comparable to their paper-based equivalents in terms of internal consistency and test-retest reliability. More significant, the electronic versions had greater patient acceptability. Lauritsen *et al.* (2004) found that electronic and telephone diaries used for data capture in clinical trials improved quality over paper-based methods.

By some estimates, we have reached a "tipping point" with regard to EDC adoption. According to a recent report by Health Industry Insights (2007) companies have increased their EDC investments by about 6.5 percent annually. The report further predicts that investment in EDC solutions will increase at a 14.7 percent compound annual growth rate, to a total exceeding $3.1 billion, within the next four years.

EDC implementation issues

Optimistic projections notwithstanding, there are no guarantees that EDC will reach critical mass any time soon. There are many implementation challenges that have slowed adoption and which will need to be overcome. As of 2007, only one third of clinical trials are using EDC (Schaltenbrand, 2007). Of those who have implemented EDC initiatives, 70 percent have not achieved the expected benefits (Tyson and Dietlin, 2006). A wide array of technology, people and process-related factors have made EDC implementation a challenging proposition and one that needs to be carefully managed. The following section highlights the most important issues surrounding EDC implementation:

Information technology infrastructure – an important consideration is the technology itself. There are many EDC approaches, each involving different infrastructures (i.e. software, hardware and networking components). Depending on

the implementation, EDC may include fax, optical character recognition (OCR), interactive voice response systems (IVRS), speech recognition, remote data entry (RDE) and web-based technologies. Investigator sites, therefore, need adequate resources (e.g. PCs, fax machines, scanners, internet connectivity, broadband access, etc.) to be able to derive full benefit from the EDC implementation.

Issues such as software installation, server maintenance and customization need to be addressed, requiring the technical support and collaboration with IT departments, which are frequently understaffed and do not typically receive revenue in clinical trial contracts. Without monetary or other incentives, IT departments may tend not to become fully engaged in the EDC effort (Welker, 2007). External hosting has become an important trend in EDC implementation. An application service provider (ASP) model may help reduce dependencies on local organizations, leaving IT departments to concentrate on basic equipment, such as PCs, phone lines and scanners.

While IT certainly has an important role to play, a common mistake is to allow it to drive the implementation process. In many cases, if IT is perceived as the one driving the EDC project, end-user acceptance will be negatively impacted. It is critical to incorporate participation from all stakeholders, including study managers, data managers and monitors, involved in the implementation (Tyson and Dietlin, 2006).

Another challenge facing the industry is the increasing shortage of both qualified investigator sites and eligible patients for clinical trial research (Sung *et al.*, 2003). It is likely that the increasing demands of the clinical trial industry (i.e. larger patient populations, more stringent eligibility requirements, total number of competing clinical trials) have contributed to this shortfall. In response to this situation, pharmaceutical companies are increasingly conducting trials and recruiting patients in the developing world. Although there is a large and often eager population of both investigators and patients, the required digital infrastructure is often not up to par in such environments. Many investigator sites in the developing world lack even the most basic computer systems and Internet connections. Additionally, although many of the investigators are English speaking and technologically fluent, much of the ancillary personnel, such as study nurses and technicians, have limited English language capabilities and computer-based skill sets. Paper-based forms tend to prevail, as they are both easier to complete and to translate in this challenging environment.

Data standards – one of the biggest challenges in implementing new information technology (IT) in today's rapidly changing world is connecting all of the disparate system components together in a seamless way. For this, we need robust enterprise architecture. This is particularly true in the biopharmaceutical industry, where the lack of accepted data interchange standards has been an ongoing impediment to progress. The lack of interoperability creates serious disruptions in the clinical trials pipeline and is estimated to cause the industry in excess of $150 million a year (Vermeiren, 2007). The Clinical Data Interchange Standards Consortium (CDISC), a non-profit organization whose mission it is to develop interoperable data standards in the healthcare industry, has been working hard to ameliorate the problem via several initiatives.

The Submissions Data Standards (SDS), a set of standards from CDISC, were formulated to guide the organization, structure and format of standard clinical trial tabulation datasets which need to be submitted to regulatory agencies such as the FDA. A major problem in implementing EDC systems is the choice of the database

structure used with the specific EDC application and/or Clinical Data Management System. Implementing SDS at the start of the clinical trial is an important step in ensuring that the data are entered correctly initially and that future conversion is not required. Much time is gained by eliminating inconsistencies between the locked database and submitted data.

A number of other standards are prevalent in the healthcare industry, adding to the mix. Kim (2005) describes the maze of data standards prevalent in the healthcare and pharmaceutical research arenas. Common exchange and messaging standards in the industry include HL7, DICOM and NCPDP which deal with patient, radiological and prescription data respectively (see Table I). Other standards categories include:

- *Terminology* standards – provide specific codes relating to diseases, problem lists, allergies, medications and diagnoses included in paper charts and transcriptions (e.g. SNOMDED for clinical terms; ICD for medical diagnosis).
- *Document* standards – indicate what type of data is contained in a document and where it is located (e.g. CCR for inter-provider information communication such as patient identification information, medical history, current allergies and medications, care plan recommendations).
- *Conceptual* standards – allow data to be transported across disparate systems maintaining meaning and context (e.g. HL7 Reference Information Model).
- *Architectural* standards – define the processes involved in data storage and distribution (e.g. the Centers for Disease Control's NEDSS system to advance the development of efficient, integrated, and interoperable surveillance systems at the state and local levels facilitating sharing of appropriate data across jurisdictions).

Another trend, occurring simultaneously with EDC adoption in the clinical trials environment, is the increasing use of electronic medical records (EMR) systems to maintain patient data in physicians' offices and in hospitals. Unfortunately, the data in most EMR systems (which contains much of the same patient information captured in EDC systems), can not be used directly in the clinical trials environment due to incompatible data formats and also because the infrastructure is not governed by clinical research regulations (eClinical Forum and PhRMA EDC/eSource Taskforce, 2006). This results in much duplication of effort and is extremely inefficient. For example, it is not uncommon for information to be hand written in a patient chart, later entered into an EMR, and then printed on paper for hand transcribing from the EMR into an EDC. Or, in the other direction, data which has been collected in an EDC is backfilled into an EMR or patient chart to comply with regulations unique to the healthcare industry.

Acceptance issues – acceptance and adoption of new information technology is an ongoing challenge and a recurring theme in the information systems literature. Ericcson and Avdic (2003), offer a framework for KM system acceptance, based on the following three factors: perceived relevance; systems accessibility; and management support.

This framework is applicable to the clinical trials environment in which the capture of knowledge plays such a critical role. Users of a KM system need to understand how it adds value to their work and be able to integrate it into current work practices. Unless those entering data into an EDC system perceive some personal gain from

Data exchange/messaging standard	Description	Developer
Health Level Seven Messaging Standards Versions 2 and 3 (HL7 V2.x and V3)	Electronic message formats for clinical, financial and administrative data	Health Level Seven, available at: www.hl7.org
Digital Imaging and Communications in Medicine Committee (DICOM)	Format for communicating radiology images and data	National Electronics Manufacturers Association, available at: www.nema.org
Clinical Data Interchange Standards Consortium (CDISC)	Format for reporting data collected in clinical trials	Clinical Data Interchange Standards Consortium, available at: www.cdisc.org
National Council for Prescription Drug Programs (NCPDP)	Structure for transmitting prescription request and fulfillment	The National Council for Prescription Drug Programs, available at: www.ncpdp.org
Accredited Standards Committee X12 (ASC X12)	Electronic messages for claims, eligibility and payments	American National Standards Institute, Accrediting Standards Committee, available at: www.x12.org/x12org/index.cfm
Institute of Electrical and Electronics Engineers Standard 1073 (IEEE1073)	Messages for medical device communications	Institute of Electrical and Electronics Engineers Standards Association, available at: www.standards.ieee.org/sa/sa-view.html

Source: Adapted from Kim (2005)

Table I.
Key data exchange /messaging standards

participation, they are unlikely to be enthusiastic about the implementation. Welker (2007) considers the problem of user motivation to be a serious barrier to EDC implementation and stresses the importance of shifting the mindset, and culture, of the organization to facilitate acceptance. An important precursor to such cultural shifts is adequate training (Welker, 2007; Kush, 2006), not only in the specific technologies involved, but also in understanding why the system is necessary and how it impacts the big picture in terms of workflow. Tyson and Dietlin (2006) emphasize that fostering a "culture of ownership" is an important success factor, empowering users of the EDC system to implement necessary process changes when necessary.

Systems accessibility deals with such questions as when, where and how users gain access to the system. As mentioned previously, access to the necessary hardware, software and communications technologies is critical in order to be able to derive full benefit from the EDC implementation. The user of an EDC system is typically presented with an analogue of the paper-based CRF in a graphical user interface (GUI). It is important for the GUI to be well designed, enabling users to perform their data entry tasks without additional training or consultation of manuals. The chosen product needs to have an intuitive interface, providing adequate language and visual cues for the user

Finally, management support undergirds any successful KM implementation and is an integral success factor in EDC systems adoption. Welker (2007) cites the lack of adequate managerial involvement as a factor that impedes adoption of EDC in many clinical settings. Kush (2006) also stresses the importance of strong managerial capabilities such as planning, leadership, and use of metrics, as necessary to make EDC implementation successful.

Another important factor is ease of use. It is interesting to note that physicians will often prefer paper-based over electronic methods simply because they are more familiar and at times more convenient to use. Medical assessments are the cornerstone of the clinical trial industry. Physicians play key roles in determining how clinical trials are conducted. Faced with demanding time constraints, the study protocols that they author often focus exclusively on the immediate details pertaining to the trial (e.g. investigator conduct and results reporting) while ignoring the need for a later independent review of the study data (Sawhney *et al.*, 2006). Once the review process becomes imminent, physicians will often opt for paper-based systems because there is not enough time to work through the validation processes required in EDC systems. Additionally, the traditional use of subjective assessments (e.g. narrative descriptions of patient symptoms or describing oncology tumors simply as "worse" or "better," rather than providing linear or volumetric measurement) do not lend themselves as well to EDC systems. For these reasons, paper-based systems tend to be favored in many trials.

Regulatory compliance – the challenge of data capture has become more complex and demanding due to recent changes in the regulatory environment. The Health Information Portability and Accountability Act (HIPAA), for example, is legislation designed to protect the privacy, security and distribution of the health information of identifiable individuals. The areas of HIPAA that directly impact EDC implementation are privacy and security. Specifically, HIPAA mandates that Protected Health Information (PHI), regardless of the form of the information, must be fully "de-identified" before collection. This in itself is a potential problem since so little of the information collected in research clinics meets the standards of de-identification as set forth in the law (Outcome Sciences, 2004).

Validation is also an important issue that impacts EDC implementation, requiring strict attention to federally mandated regulations. Investigator sites need to be familiar with the FDA's Electronic Record, Electronic Signature Regulations (21 CFR Part 11) which requires that computer systems and electronic records provide for the following safeguards (Marks, 2004):

- validation of computerized systems to ensure the accuracy, reliability, integrity, availability and authenticity of required records and signatures;
- audit trail which is both computer generated and time-stamped;
- physical, logical and procedural security measures;
- access control – limited to authorized individuals with proper system training and for inspection;
- adherence to established policies and standard operating procedures; and
- retention of records to allow for retrieval.

EDC is a key element of the eClinical/knowledge management platform

EDC is a key element of the eClinical/knowledge management platformAn overall eClinical solution is critical knowledge management for biopharmaceutical drug development. It allows for the interchangeability and interconnectivity of data among standard components, such as Clinical Trial Management Systems (CTMS), Interactive Voice Response Systems and EDC, in the clinical environment. Perceptive Informatics, the technology business of PAREXEL, a global biopharmaceutical services provider, for instance, provides an eClinical platform, which brings data together to create a broader picture of the trial, resulting in improved quality, safety and decision-making.

As an integral part of this flexible eClinical infrastructure, data from EDC solutions are seamlessly married with data from other study-related sources such as e-patient diaries and hybrid, paper-based data capture systems. Study reports and portals provide real-time access to system metrics and clinical data.

EDC speeds data analysis timelines

Data capture methods can be evaluated using a number of different metrics, including accuracy and completion rate. By comparing two similar phase III oncology studies that had previously been completed by Perceptive Informatics, differences in the efficiencies of two capture methods are considered; paper-based and electronic-based (EDC).

For these studies, data were generated and gathered at investigator sites via physical assessments (e.g. radiography, blood work) to compose the entire collection of the patient's raw clinical data. This raw data was managed by the clinical sites, allowing clinical investigators to assess the patient's performance while on the drug under investigation. These assessments, in addition to the raw data supporting them, are collectively referred to as Clinical Dossiers. They were crucial to the independent review of study data (including the actual radiological exams) during these trials.

In both of the phase III studies, knowledge capture and transfer began at the investigator site with the submission of the collective raw data for each patient to the client's data management group. However, both studies differed in the methods used. Protocol A used the older, paper-based case report forms (CRFs) to compile the data. Protocol B used a web-based EDC system to compile the data. Once the raw data were submitted to the data management group and a patient was identified as complete and

slated for independent review, the data management group compiled all patient knowledge into the Clinical Dossier and transferred it for a facilitated independent review.

Once acquired and prior to the independent review, there was reconciliation between the data in the dossier and the data in the database. Any discrepancies noted through this reconciliation delayed the independent review, since the data management group now needed to develop a solution and revise the dossier.

Measuring the time from the first receipt of the dossier to the time in which the patient was ready for independent review is a simple and intuitive way to compare the efficiency of the two different methods used. It is not unreasonable to assume that the shorter the time between clinical dossier receipt and patient review readiness and the fewer the discrepancies, the more efficient a method is in capturing and transferring clinical trial knowledge.

For protocol A, which employed the paper delivery system, an average of 80 days passed between the delivery of the first dossier to the time that the patient was ready for independent review. For protocol B, which employed the EDC method, approximately 40 days passed between the delivery of the first dossier to the time that the patient was ready for independent review.

This data would suggest that the EDC method can cut down discrepancy resolution time by approximately 50 percent. While this data implies that EDC is superior in facilitating knowledge capture and transfer over the older paper-based method, we are still faced with many of the operational challenges discussed previously. The use of EDC in protocol A caused the sponsor to be extremely selective with the clinical sites and in many cases the sponsor had to subsidize the installation of a dedicated high speed internet line. The sponsor also had to place a significant emphasis on the training of staff in not only the functionality, but also in the regulatory considerations of EDC. Lastly, necessary fields collecting Serious Adverse Event and Adverse Event data were absent from the EDC, causing additional delays in both the data reconciliation and data cleaning. This suggests that while EDC was much more efficient than the older, paper-based method, a more proactive design of EDC to capture all aspects of the data would have yielded an even greater efficiency.

EDC is no substitute for training and monitoring

In one independent review of radiographic data (i.e. CT, MRI, and X-ray exams) to assess the efficacy of an experimental cancer treatment, the basic assessment centered upon testing whether this new compound would shrink existing cancerous lesions (i.e. tumors) and/or prevent the occurrence of new lesions. Both the investigator site and trained radiologists independently measured tumor diameters on medical imaging scans to determine change over time for existing lesions and simultaneously recorded the presence or absence of new lesions.

It was decided early on to use EDC to record their independent results. It was envisioned that EDC would help ensure consistency between the two reviews. The rationale for this assumption consisted of the following:

- *Both reviewers had the complete medical imaging dataset.* The investigator's EDC form prompted reviewers to enter exam data (i.e. modality and exam date) prior to review. A common occurrence with paper forms is that an assessment is submitted, yet the modality and date of the imaging exam in question cannot be

determined or is incorrect (e.g. exam date is after assessment date). When discrepancies arise, it is very difficult to ensure that the independent reviewer is looking at the same images as seen by the investigator. Using EDC, it is relatively easy to produce an ongoing reconciliation (via electronic reporting) of exams acquired at the investigator site and those that had had been received by the independent reviewer. Missing or erroneously data are more quickly corrected in-stream rather than nearer the final analysis (when all the paper forms eventually surface for review by the sponsor).

- *Review criteria edit checks.* The assessment criteria used for this study, Response Evaluation Criteria in Solid Tumors (RECIST), requires complicated algorithms of limited lesion selection areas, linear measurement, and percent change calculations over time that are very difficult to enforce with paper forms. Mathematical errors and lesion selection omissions/errors have typically been the norm with manual reporting. EDC systems were built to aid reviewers by providing the mathematical results and attempting to enforce proper lesion selection and follow-up reporting.

After a planned interim analysis was conducted consisting of approximately 120 patients, it was determined that the investigators and the independent reviewers disagreed on "new lesion" determination 37 times in this sample size alone. In all instances, the investigators had indicated the presence of a new lesion, while the independent reviewer did not confirm this conclusion. Although there are no published concordance expectations for this type of analysis, the amount of discrepancy seen here was completely unacceptable. Therefore, an additional independent team of physicians was established to produce a third party review of both datasets and determine the root causes of the discrepancies.

After analyzing all EDC data and imaging exams for the 36 patients the final results were as follows:

- A total of 23 cases consisted of miscellaneous errors by either the investigator or the independent reviewer;
- Five cases involved differences in the way the two reviewer/EDC systems interpreted disease assessments; and
- Nine were considered "borderline" new lesion and were very subjective.

A closer look at these results indicated that in 14 instances the requirements for the EDC forms were not consistent between the investigator and the independent reviewer systems or that the requirements were too vague (e.g. no minimum size was ever established for a new lesion). The remaining 23 instances of discordance were not as clear. In at least a few instances it was confirmed that investigators were utilizing information outside of the EDC reported results (i.e. patient symptoms, additional confirmatory medical tests, etc.) to determine new lesion assessments. Additionally, it is speculated that investigators were not completing the assessments personally but merely entering the data generated by others (probably subordinates or on-call radiologists). Under these circumstances existing lesions can often be erroneously reported as "new."

In summary, lack of coordination and explicit requirements of the two EDC systems, lack of training (or perhaps willful noncompliance), and lack of effective

monitoring contributed significantly to the overall poor concordance of these study results. Although in this instance EDC did play a significant role (it should be noted there were no mathematical errors observed among thousands of calculations), the technology alone cannot be counted on to supplant the standard clinical trial management system consisting of continuous training and monitoring.

Summary and conclusion

Through a review of the literature and application of specific case data, a number of important advantages of EDC surfaced, including improved data clarification/correction capabilities, improved quality/accuracy of data acquisition, automated data entry and submission, and greater time and cost efficiencies. EDC is growing at a rate of 6.5 percent annually and is beginning to receive widespread attention in the industry. There has been increased vendor activity in the EDC space and investments in associated technology and solutions are growing at double-digit rates, with projected sales in the billions within the next few years.

In spite of this dramatic growth, it is also evident from our investigation that while EDC has the potential to radically improve the clinical trial process, it still faces a number of significant hurdles, which may slow its adoption. As in any information technology, successful implementation involves more than simply mastering the technology itself. Barriers also include user resistance, lack of interoperability standards, and regulatory compliance.

Major shifts in organizational culture will be required to implement EDC successfully. Collaboration between the different stakeholders in the pipeline (e.g. sponsor, contract research organization, investigator site) is also crucial. The industry itself needs to come together to harmonize data standards.

There are a number of limitations in the approach used in this investigation and further research is warranted. Our analysis was limited to only one phase of the KM life cycle, the capture and/or creation stage. Also of vital importance is how knowledge is shared, disseminated, applied and retrieved in the clinical trials pipeline. Future research, focusing on these stages of the KM cycle, is necessary to better understand the limitations and problems of leveraging intellectual capital in the drug discovery process.

Case studies are inherently limited in scope. To get a broader sense of the barriers to EDC in this environment, future survey based and experimental studies are warranted. While EDC is no panacea, it represents an important advance and one that has the potential of improving the management of critical knowledge, translating ultimately to the delivery of safer, higher quality and more affordable health care.

References

Bart, T. (2003), "Comparison of electronic data capture with paper data collection: is there really an advantage?", *Business Briefing Pharmatech*, pp. 1-4.

Beyster, M.A., Hardison, D.H. and Lubin, G.M. (2005), "Improving clinical trials by implementing information technology: where will you be in three years?", Science Applications International Corporation, San Diego, CA, available at: www.saic.com/life-sciences/pdfs/mcv.pdf

Bushnell, D.M., Reilly, M.C., Galani, C., Martin, M.L., Ricci, J., Patrick, D.L. and McBurney, C.R. (2006), "Validation of electronic data capture of the irritable bowel syndrome: quality of life

measure, the work productivity and activity impairment questionnaire of irritable bowel syndrome and EuroQol", *Value in Health*, Vol. 9 No. 2, pp. 98-105.

Carel, R. and Pollard, J. (2003), "Knowledge management in drug discovery R&D", 3rd Millennium, Inc., Waltham, MA, available at: www.3rdmill.com/resources/3rd_Millennium_KM_in_Pharma_RD.pdf

Dalkir, K. (2005), *Knowledge Management in Theory and Practice*, Elsevier Butterworth-Heinemann, Burlington, MA.

Ericcson, F. and Avdic, A. (2003), "Knowledge management systems acceptance", in Coakes, E. (Ed.), *Knowledge Management: Current Issues and Challenges*, IRM Press, Hershey, PA.

Health Industry Insights (2007), "Electronic data capture (EDC) poised to disrupt life sciences industry", May, available at: www.healthindustry-insights.com/HII/getdoc.jsp?containerId = prUS20671907

Hodgson, J. (2001), "The headache of knowledge management", *Nature Biotechnology*, Vol. 19 No. 6, pp. 44-6, available at: www.pharmaknowledge.com/Nature%20Biotechnology%20article.htm

Hyde, A.W. (1998), "The changing face of electronic data capture: from remote data entry to direct data capture", *Drug Information Journal*, Vol. 32, pp. 1089-92.

IMS (2007), "IMS Health reports global pharmaceutical market grew 7.0 percent in 2006, to $643 billion", 20 March, available at: www.imshealth.com/ims/portal/front/articleC/0,2777,6599_3665_80560241,00.html

Kim, K. (2005), "Clinical data standards in healthcare", iHealth Reports, California HealthCare Foundation, Oakland, CA, July, available at: www.chcf.org/documents/ihealth/ClinicalDataStandardsIn HealthCare.pdf

Kush, R. (2006), "Electronic data capture: pros and cons", *BioExecutive International*, Vol. 2 No. 6, pp. S48-S52.

Lauritsen, K., Degl'Innocenti, A., Lene, H., Jørgen, P., Lytje, M.F., Clemmensen-Rotne, K. and Wiklund, I. (2004), "Symptom recording in a randomised clinical trial: paper diaries vs electronic or telephone data capture", *Controlled Clinical Trials*, Vol. 25 No. 6, pp. 585-97.

Leavitt, P.M. (2002), "The role of knowledge management in new drug development", American Productivity & Quality Center, available at: www.providersedge.com/docs/km_articles/Role_of_KM_in_New_Drug_Development.pdf

Malhotra, Y. (2003), *Measuring Knowledge Assets of a Nation: Knowledge Systems for Development*, Report of the *Ad Hoc* Group Meeting on Knowledge Systems for Development, , United Nations Headquarters, NYC, NY, 4-5 September, pp. 68-126.

Marks, R.G. (2004), "Validating electronic source data in clinical trials", *Controlled Clinical Trials*, Vol. 25 No. 5, pp. 437-46.

Outcome Sciences (2004), "Health information portability and accountability act (HIPAA): implications for data capture with special reference to web-based electronic data capture", available at: www.outcome.com/new_home/images/HIPAA_EDC_Outcome_Sciences.pdf

Pratt, T. (2007), "EDC: caught in a trap", *Applied Clinical Trials*, June, pp. 26-8.

Rao, M. (2005), "Overview: the social life of KM tools", in Rao, M. (Ed.), *Knowledge Management Tools and Techniques: Practitioners and Experts Evaluate KM Solutions*, Elsevier Butterworth-Heinemann, Oxford.

Sahoo, U. and Bhatt, A. (2003), "Electronic data capture (EDC): a new mantra for clinical trials", *Quality Assurance*, Vol. 10, pp. 117-21.

Sawhney, G., Chandler, B., Williams, K. and Chica, S. (2006), "The role of independent review in oncology trials", *Applied Clinical Trials*, May, available at: http://actmagazine.findpharma.

com/appliedclinicaltrials/Clinical + Trials/The-Role-of-Independent-Review-in-Oncology-Trials/ArticleStandard/Article/detail/324333

Schaltenbrand, N. (2007), "Large-scale electronic data capture: when size is not the only issue", Touch Briefings, London, available at: www.touchbriefings.com/pdf/2743/schaltenbrand.pdf

Sung, N., Crowley, W.F., Genel, M., Salber, P., Sandy, L., Sherwood, L.M., Johnson, S.B., Catnese, V., Tilson, H., Getz, K., Larson, E.L., Scheinberg, D., Reece, E.A., Slavking, H., Dobs, A., Grebb, J., Martinez, R.A., Korn, A. and Rimoin, D. (2003), "Central challenges facing the national clinical research enterprise", *The Journal of the American Medical Association*, Vol. 289 No. 10, pp. 1278-87.

Tyson, G. and Dietlin, T. (2006), "Realizing the promise of electronic data capture: a practical guide", available at: www.campbellalliance.com/articles/Realizing_the_Promise_of_Clinical_EDC_A_Practical_Guide.pdf

Vermeiren, K. (2007), "EDC and CDISC, the perfect combination for successful submission", available at: www.sgs.com/lss_edc_article_ict_sgs_120707_bis.pdf

Welker, J.A. (2007), "Implementation of electronic data capture systems: barriers and solutions", *Contemporary Clinical Trials*, Vol. 28, pp. 329-36.

eClinical Forum and PhRMA EDC/eSource Taskforce (2006), "The future vision of electronic health records as ESource for clinical research", available at: www.eclinicalforum.com/content/Knowledge/Articles/Future%20EHR-CR%20Environment%20Version%201.pdf

About the authors

Martin Grossman is an Assistant Professor in the Department of Management at Bridgewater State College. Dr Grossman received a DBA in Information Technology Management and a Masters in International Business from Nova Southeastern University and a BS in Computer Systems from Florida Atlantic University. He has over 20 years of IT experience in the communication and aviation industries. Dr Grossman's current research interests include knowledge management, e-learning, and technology innovation. Martin Grossman is the corresponding author and can be contacted at: mgrossman@bridgew.edu

Stephen Bates is Program Director of the Oncology/Cardiology Group at Perceptive Informatics, Medical Imaging. Stephen joined Perceptive in 2001. During the past six years, he has had extensive experience in a wide range of clinical trials and imaging techniques, including the performance of both oncology (RECIST/Cheson) and CNS volume measurement. Prior to joining Perceptive, Stephen was a US Navy and Nationally Certified Orthopaedic Technician. He managed the largest US Military orthopaedic surgical clinic in Asia, where his responsibilities included JCAHO compliance and mass casualty preparedness. Stephen graduated *summa cum laude* with a Bachelor of Science degree in Cell/Molecular Biology from Bridgewater State College and is currently pursuing a Masters Degree in Technology Management.