

PEER REVIEWED

Jules T. Mitchel, MBA, PhD | Yong Joong Kim, MS | Joonhyuk Choi, MS
Vadim Tantsyura, MS, MA | Douglas Nadler, MS | Imogene Grimes, PhD
Silvana Cappi, MSc, MBA | Philip T. Lavin, PhD | Kirk Mousley, MSEE, PhD

The Impact of Electronic Data Capture on Clinical Data Management

Perspectives from the Present into the Future

With the proper EDC toolbox, . . . the clinical group may now be able to . . . deploy a full EDC study in days rather than weeks or months.

Electronic data capture (EDC)-based clinical trials offer operational and cost-effective approaches for ongoing data entry via the Internet for clinical sites; medical monitoring; monitoring by clinical research associates (CRAs), including initial review of data in the home office and then performing source document verification at the study site; identification of potential errors by data management; and determination of the status of the clinical trial by project management.¹⁻⁴ Ten years ago, Kelly and Oldham⁵ discussed the challenges of implementing EDC and the potential advantages in clinical development. Kuchenbecker and colleagues⁶ foresaw the emerging role of Internet technologies for data acquisition and predicted the eventual common use of EDC in the pharmaceutical industry. Banick⁷ predicted that with EDC, time to database lock could be reduced by 43% and queries by 86%. The pros and cons of EDC have recently been presented,⁸ as well as the challenges for implementing EDC clinical trials.⁹

Starting a Study

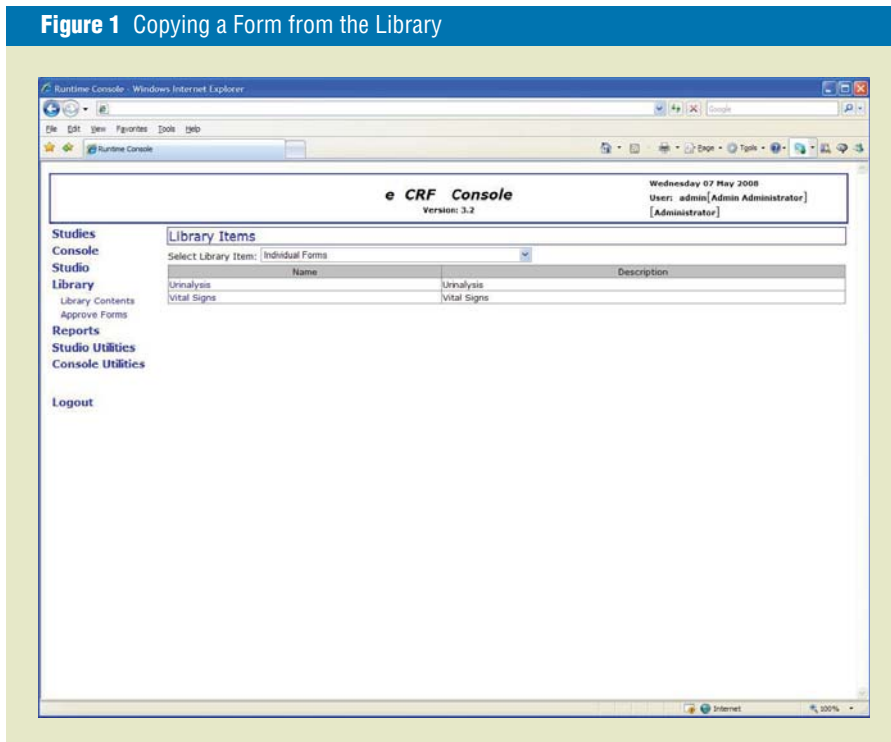
All planning and implementation of EDC must be done prior to enrollment of the first patient. Data entry screens, online edit-check specifications, and the annotated case report form (CRF) can, and must be, completed prior to patient enrollment. Basically, there is no luxury in an EDC trial to put off those tasks, as can be done with a paper-based CRF trial (often with unintended negative consequences). Also, there must be upfront and full integration in the design of the trial with clinical research, data management, and biostatistics to assure that the data entry process is user-friendly for the clinical sites and that the exported database structure is compatible with the planned statistical analysis.

Once forms and their associated edit (validation) checks are created, and assuming that a company adopts and enforces standards, forms and form elements can easily be reused for other studies through a library system. With some EDC systems, a new study can be created by merely invoking the “copy” function, which effectively clones the established system to create forms for the next trial (see Figure 1).

In order to initiate a study once the copy function is invoked, one needs only to add roles, users, and sites, which can also come from the library. Basically, do this once with agreed-upon standards for forms, variables, and edit checks, and the EDC study may already be 80% complete once a study is copied from the library.

With the proper EDC toolbox, which should include a form generator and programmerless edit checks, the clinical group may now be able to cre-

Figure 1 Copying a Form from the Library



Without having to learn a new programming language, traditional SAS programmers can write the batch edit checks compatible with an EDC system that uses SAS directly. The program code can identify which forms to display within the query system for each batch edit check, and supply the appropriate error message. With online batch edit checks, the field monitor is able to trigger the batch edit checks and assess them, similar to the management of a normal range or edit check generated at the time of data entry (see Figure 2).

Clinical Trial Oversight

The data manager acts as a bridge between field operations and biostatistics. Although the processes may differ between companies, the concept is somewhat universal. Statisticians want “clean” data, which data management must deliver. In EDC, there is a dramatic drop in the types of data errors found in paper-based CRF studies, such as out-of-range values or missing data.

In EDC, there is a dramatic drop in the types of data errors found in paper-based CRF studies, such as out-of-range values or missing data.

Whether the explanation provided by the site for the condition of the data is acceptable is another matter, and that is where the data manager jumps in. The data manager can view all edit-check activities and can generate queries directly to the clinical sites. This happens in real time; so when it is time to lock the database, there can be a very high expectation that the data are clean. Some data management tasks, such as providing reports or notifying the medical monitor about serious adverse events, may still require a phone call; these tasks may

ate the CRF forms and, together with data management, deploy a full EDC study in days rather than weeks or months. The learning curve is not steep, allowing for paper-competitive implementation on the first EDC study and accelerated implementation on subsequent studies. The experience level of the staff required to perform the implementation no longer rises to the level of a software developer or an EDC expert, although experience is a plus.

Query Management

Query management has also changed dramatically, with all outstanding queries and edit-check resolutions in an EDC trial being only a click away. With EDC, the entire query management process can be handled within the website, with no paper queries. Although queries must still be generated, they can now be managed via a web interface, rather than paper forms in ringbinders. Queries can be resolved in minutes rather than weeks, assuming that the site is responsive to the query.

Queries can be resolved in minutes rather than weeks, assuming that the site is responsive to the query.

Traditionally, batch edit checks, or “potential queries,” are generated by data management using statistical analysis software (SAS) programs or equivalent software. Then, on a regular schedule, these edit checks are run and distributed to either the CRAs or the data managers for resolution. Once resolved through the query process or given an “OK as is” designation, the edit checks are manually deactivated. Ongoing through a clinical trial, this process can be very labor intensive.

With EDC, batch edit checks—written in SAS, procedural language/structured query language, or other software—can be integrated with the electronic query system of the study. The EDC system can run the edits and display the results of those edits through a discrepancy review screen.

Figure 2 Running SAS Batch Edit Checks within EDC

Start Date/Time (dd/mon/yyyy hh:mm)	Stop Date/Time (dd/mon/yyyy hh:mm)	Run By
05 Feb 2008 14:49	05 Feb 2008 14:49	IH-T.IHMONITOR
05 Feb 2008 14:46	05 Feb 2008 14:46	IH-T.IHMONITOR
05 Feb 2008 10:34	05 Feb 2008 10:34	IH-T.INHOUSEMONITOR
05 Feb 2008 10:30	05 Feb 2008 10:30	IH-T.INHOUSEMONITOR
05 Feb 2008 10:22	05 Feb 2008 10:22	IH-T.INHOUSEMONITOR
05 Feb 2008 10:14	05 Feb 2008 10:15	IH-T.INHOUSEMONITOR
05 Feb 2008 10:09	05 Feb 2008 10:10	IH-T.INHOUSEMONITOR
05 Feb 2008 09:50	05 Feb 2008 09:50	IH-T.INHOUSEMONITOR

inherently be automated through management reporting or real-time e-mail notification capabilities.

With the advent of EDC, there is now an overlap in the ability of the CRAs, data managers, statisticians, and project managers to see each other's processes and workflow. For example, in a simple management report, the status of monitoring of the clinical site's data entry is available to all: Data managers, project managers, and CRAs can all see how many patients have their data locked and signed electronically (see Figure 3).

In EDC trials, as CRAs take the time to review the data prior to the monitoring visit, they can be much more knowledgeable about the status of the trial at the time of the monitoring visit. Issues such as missing data, illogical data, misspellings, and incorrect terminologies/acronyms can be reviewed offsite and then confirmed at the time of source document review. The monitor can also have the authorization to prevent the site from changing data after the monitoring visit. Of course, all of these tasks can be reversed at any time prior to database lock. Data management can also help the monitor by providing alerts to data issues as they are entered.

Management reports (or built-in workflow) can be used to close out a study, by confirming when all forms are monitored and locked, and when it is time for the investigators to sign the CRF electronically. Once all of these tasks are accomplished, the study can be locked. Prior to locking the data for a patient, a final check ensures that there are no unresolved edit checks and queries. When the patient record is locked, eSignatures can be invoked.

Discussion

With EDC, data entry and data modification responsibilities have shifted from data management to the CRA and site personnel. This allows data management personnel to focus on other "value added" activities. With EDC, data are entered only once by those who should know the data the best (i.e., the clinical study site). The site coordinator who enters the data needs to have access to the patient source records and must be permitted to make updates to the data per the site standard operating procedures.

The issue of "who does the entry" at the site has a large bearing on the success of an EDC study. As long as this task is not "farmed out" to a data entry

clerk who is not familiar with the patient, there should be only occasional human errors, such as typographical or transcription errors. The CRA now assumes some audit functions, and even some of the roles of data management. In fact, the CRAs are probably in the best position to make their job more effective by using the EDC system for review before making site visits. The statisticians are able to get cleaner data earlier in the process, and senior management appreciates the lack of significant delay between the point of last patient/last visit and database lock.

However, since EDC represents a paradigm shift, sponsors must be aware that training and job descriptions must be adjusted as workloads are redefined; in fact, the entire process of data management must be reconsidered. The paper-based data management system does not translate task-by-task into the EDC system. It is essential to reidentify the hand-offs and to introduce quality gates appropriate for the EDC system where site personnel, monitors, and data management share responsibilities for the quality of the data. To the extent that the EDC application can enforce workflow, this "new" teamwork becomes easier to implement. EDC represents an opportunity for career development and for improved job satisfaction through increased team interactions and more control over meeting timelines.

Since EDC represents a paradigm shift, sponsors must be aware that training and job descriptions must be adjusted as workloads are redefined.

When designed properly, EDC can facilitate data management processes from CRF generation to monitoring of the clinical data and integration of edit checks. The ability to reduce the time to database lock removes a timeline

Figure 3 Project Management Report—Data Entry Status

ABC		PROTOCOL DEMOECRF - CLIENT JANE MONITOR			5 Feb 2008 Ver 1.00			
PHARMACEUTICALS		MAIN	REGISTRATION	QUERY/EDIT CHECK	MANAGEMENT REPORTS	STATUS BY SUBJECT		
SUBJECT LIST								
(As of Tuesday, February 05, 2008, 1:11:30 PM)								
No	Site No.	Screening No.	Subject No.	# of Forms Entered	# of Pages Entered	# of Pages Reviewed In-Field	# of Pages Reviewed In-House	Subject Locked
1	99	99S019 (JJ)	9917	6	6	2	2	NO
2	99	99S029 (AA)	9901	13	14	1	0	NO
3	99	99S039 (LS)	9906	6	7	3	1	YES
4	99	99S049 (AA)	9910	5	5	3	0	NO
5	99	99S059 (WW)	9905	7	7	1	0	NO
6	99	99S069 (JY)		4	4	1	0	NO
7	99	99S079 (AB)	9907	11	12	3	0	NO
8	99	99S089 (AA)	9915	4	4	1	0	NO
9	99	99S179 (WW)	9926	6	6	0	0	NO
10	99	99S189 (LS)	9988	5	5	2	0	NO
11	99	99S209 (TT)	9949	4	4	1	0	NO
12	99	99S219 (JT)	9982	4	4	2	1	YES
13	99	99S339 (ZZ)		2	2	0	0	NO
14	99	99S359 (LS)	9987	4	4	1	0	NO
15	99	99S459 (LS)	9945	6	6	1	0	YES

database cleanup, and database lock, leaving more time for statistical analyses, final study reports, regulatory submissions, and, ultimately, reduced time to market launch. However, companies must sort through the multitude of EDC vendors to identify the software that is most compatible with their internal processes and be willing to restructure and take the appropriate steps to redo their workflow and assure the appropriate resource allocations.

EDC-enabled data management process standardization will become a primary focus for data managers in the next few years. EDC allows the globalization and standardization of data management operations, and remote EDC training will become the norm. The relative importance of database security will also increase with the emergence of EDC.

EDC can help clean and lock data faster than traditional paper CRF systems. Clinical trial professionals must adopt new processes, embrace standardizations, and learn to respond more quickly to management reports in addressing issues as they arise. This will help the clinical data management department become more effective in doing its job.

Acknowledgement

The authors would like to thank Joyce Hays, MS, chief executive officer of Target Health Inc., for reviewing the manuscript.

References

1. Mitchel J, You J, Lau A, et al. 2000. Paper versus web: a tale of three trials. *Applied Clinical Trials* August: 34-5.
2. Mitchel J, You J, Kim YJ, Lau A, et al. 2003. Internet-based clinical trials - practical considerations. *Pharmaceutical Development and Regulations* 1: 29-39.
3. Mitchel J, You J, Lau A, et al. 2003. Clinical trial data integrity: using Internet-based remote data entry to collect reliable data. *Applied Clinical Trials* March (Supplement): 6-8.
4. Mitchel J, Jurewicz E, Flynn-Fuchs K, et al. 2005. The role of CRAs in the development and implementation of Internet-based clinical trial applications: new career opportunities. *The Monitor* 19(4): 17-21.

stress, as statisticians and medical writers do not need to make up the delays to database lock.

As EDC prices drop (hopefully) and scalability improves, the size of the study should not be a reason why EDC is used or not used. EDC systems have now undergone many Food and Drug Administration audits with no adverse outcomes delaying or invalidating approval. Moreover, significant time and cost savings have evolved from study start to database lock and final report by eliminating double-key data entry; having an integrated query and online/offline edit-check system; doing electronic monitoring; and providing for eSignatures.

Contract research organizations and sponsors have developed EDC-specific processes to implement EDC for every stage of clinical development in every therapeutic area. EDC supports standardization, which can help set up studies faster. When a single EDC system is selected for a program, the cloning of one study to facilitate quick startup of a similarly structured companion study will improve efficiency of the startup of later studies. More importantly, it will facilitate a move toward common standards for a single program.

Conclusion

The pharmaceutical, biotechnology, and medical device industry, as well as academia and the government, have all begun to adopt EDC as a new data management tool. EDC acceptance is strong; there are very few instances where users have gone back to paper-based data collection. Though the goal of data management will not change—i.e., to assure “clean” data at the end of the study—there is no doubt that data management processes will evolve with the use of EDC systems.

EDC acceptance is strong; there are very few instances where users have gone back to paper-based data collection.

When EDC is managed properly, there is reduced time for study startup,

5. Kelly MA, Oldham J. 1997. The Internet and randomised controlled trials. *International Journal of Medical Informatics* 47: 91-9.
6. Kuchenbecker J, Dick HB, Schmitz K, et al. 2001. Use of Internet technologies for data acquisition in large clinical trials. *Telemedicine Journal and e-Health* 2001: 73-6.
7. Banik N. 1998. Evaluation of EDC versus paper in a multinational asthma trial. Presented at the DIA European Data Management Meeting, Berlin, October 1998.
8. Kush R. 2006. Electronic data capture—pros and cons. *BioExecutive International Supplement Series*, June 2006.
9. Kush R, Bleicher P, Kubick W, et al. 2003. eClinical trials: planning & implementation. *Thomson CenterWatch*, 2003. **ACRP**

Jules T. Mitchel, MBA, PhD, is president and cofounder of Target Health Inc. His more than 25 years of experience in the pharmaceutical industry includes development of drugs, biologics, devices, and diagnostics involving participation in numerous FDA meetings, preparation of regulatory submissions, study reports, and product development plans. He has held positions at Ayerst Laboratories (now Wyeth), Pfizer Laboratories, and Pfizer Consumer Health Care, and he can be reached at jules.mitchel@targethealth.com.

Yong Joong Kim, MS, is the senior director of application development and data management at Target Health Inc. He has the overall responsibility of development of the Target eCRF® and other software products at

Target Health. Previously he worked at the Rockefeller University as an SAS programmer/system analyst for 10 years. He can be reached at YKim@TargetHealth.com.

Joonhyuk Choi, MS, has served as director of software development for Target Health Inc. for the past seven years. He is one of the lead architects of the Target eCRF® system and is responsible for developing the company's product strategy and architectural direction. He can be reached at jhchoi@targethealth.com.

Vadim Tantsyura, MS, MA, has 15 years of extensive engineering, information technology, data management and project management experience, including eight years of pharmaceutical experience at Pfizer, Omnicare, Clinimetrics, Bristol-Myers Squibb, and Regeneron. He is currently the director of data management at Regeneron Pharmaceuticals, where he has built the clinical data management team and led the efforts that culminated in the first Regeneron BLA approval in February of 2008. He can be reached at vadim.tantsyura@regeneron.com.

Douglas Nadler, MS, is associate director of statistical services at Regeneron Pharmaceuticals, Inc. He has been in the pharmaceutical/biotech industry for nine years and can be reached at douglas.nadler@regeneron.com.

Imogene Grimes, PhD, is vice president of data sciences strategic services at PAREXEL International Corporation. Previously she was vice president, statistics, data management, and informatics at Regeneron, where she championed the approval of Regeneron's first BLA. She has 25 years of experience in the pharmaceutical industry, including a decade in major pharmaceutical companies

(Glaxo and Pfizer). She can be reached at imogene.grimes@parexel.com.

Silvana Cappi, MSc, MBA, is executive director global biometrics within clinical and nonclinical research and development at Ferring Pharmaceuticals International PharmaScience Centre in Copenhagen, Denmark. During her time as head of the Global Biometrics Department, it successfully implemented electronic data capture, CDISC SDTM and ADaM standards, and an outsourcing policy (with the selection of preferred providers), and Ferring's global clinical database, and contributed to the successful delivery of the company's first electronic submission (eCTD). She can be reached at silvana.cappi@ferring.com.

Philip T. Lavin, PhD, is executive chairman, Averion International Corp. He has served as a faculty member at the Harvard School of Public Health and Harvard Medical School and has successfully supported many PMAs, BLAs, and NDAs with more than 40 direct FDA product approvals. Over the past 20 years, he has served on multiple FDA advisory panels. He can be reached at Philip.Lavin@averionintl.com.

Kirk Mousley, MSEE, PhD, president of Mousley Consulting, Inc., has directed efforts in computer application design and development, clinical database design, data editing/cleaning, and submissions. His work has involved numerous database applications, clinical data management systems, and electronic data capture applications. He has 20 years of computer systems experience in the consulting, education, telecommunications, and aerospace fields, and he can be reached at kirk@mousley-consulting.com.