



Published in final edited form as:

Clin Trials. 2009 February ; 6(1): 67–75. doi:10.1177/1740774508100983.

A centralized informatics infrastructure for the National Institute on Drug Abuse Clinical Trials Network

Jeng-Jong Pan, PhD^a, Meredith Nahm, MS^b, Paul Wakim, PhD^a, Carol Cushing, RN^a, Lori Poole, BS^c, Betty Tai, PhD^a, and Carl F. Pieper, DrPH^d

^a National Institute on Drug Abuse, Center for the Clinical Trials Network, Bethesda, MD

^b Duke Translational Medicine Institute, Durham, NC

^c Duke Clinical Research Institute, Durham, NC

^d Biostatistics and Bioinformatics Department, Duke University Medical Center, Durham, NC

Abstract

Background—Clinical trial networks were created to provide a sustaining infrastructure for the conduct of multisite clinical trials. As such, they must withstand changes in membership. Centralization of infrastructure including knowledge management, portfolio management, information management, process automation, work policies, and procedures in clinical research networks facilitates consistency and ultimately research.

Purpose—In 2005, the National Institute on Drug Abuse (NIDA) Clinical Trials Network (CTN) transitioned from a distributed data management model to a centralized informatics infrastructure to support the network's trial activities and administration. We describe the centralized informatics infrastructure and discuss our challenges to inform others considering such an endeavor.

Methods—During the migration of a clinical trial network from a decentralized to a centralized data center model, descriptive data were captured and are presented here to assess the impact of centralization.

Results—We present the framework for the informatics infrastructure and evaluative metrics. The network has decreased the time from last patient-last visit to database lock from an average of 7.6 months to 2.8 months. The average database error rate decreased from 0.8% to 0.2%, with a corresponding decrease in the interquartile range from 0.04%–1.0% before centralization to 0.01%–0.27% after centralization. Centralization has provided the CTN with integrated trial status reporting and the first standards-based public data share. A preliminary cost-benefit analysis showed a 50% reduction in data management cost per study participant over the life of a trial.

Limitations—A single clinical trial network comprising addiction researchers and community treatment programs was assessed. The findings may not be applicable to other research settings.

Conclusions—The identified informatics components provide the information and infrastructure needed for our clinical trial network. Post centralization data management operations are more efficient and less costly, with higher data quality.

Keywords

clinical trial networks; data management; informatics infrastructure; clinical research informatics; electronic data capture

Introduction

In order to effectively and efficiently translate drug abuse research into improved drug addiction treatment, the National Institute on Drug Abuse (NIDA) established the National Drug Abuse Treatment Clinical Trials Network (CTN). The CTN is a national network of addiction researchers and community-based treatment programs whose mission is to conduct multisite clinical trials of behavioral, pharmacological, and integrated behavioral/ pharmacological treatment interventions of therapeutic effect in community-based treatment settings. The infrastructure of the CTN is also available for non-network researchers to use as a platform for conducting research.

Following the recommendation of the 1998 Institute of Medicine study report and modeled after the National Cancer Institute's Community Clinical Oncology Program [1], the CTN awarded its first six Nodes in the fall of 1999. In a hub-and-spoke model, the Node comprises a regional research training center located in a major academic center and associated community treatment programs. Each center is affiliated with five to ten primary community treatment programs. As of August 2008, 16 Nodes and 240 affiliated community treatment sites across 35 states and Puerto Rico participate in the CTN. The principal investigators from these Nodes have supported more than 30 protocols and have enrolled and randomized more than 9000 study participants. More information on the NIDA CTN can be found at <http://www.nida.nih.gov/ctn/>.

From 1999 to 2005, the CTN used a distributed model for the data management of its trials, shown in Figure 1. The distributed model was originally chosen to increase local engagement and data quality through regional ownership and accountability. In this model, (Figure 1) each Node maintained a local data management center, which built and maintained a data acquisition system, developed case report forms (CRFs), and managed the data collected for each protocol where a community treatment program from the Node was participating as a site. The number of centers participating in a single trial ranged from one to four. As the CTN grew—conducting several trials simultaneously—so did the number of community treatment programs from different Nodes participating in the trials. A single data management center was likely involved in several trials at any one time. Every month, each Node data management center transferred the cumulative clinical data for each trial to NIDA's Central Data Repository, where the data for each study were checked for validity, transformed if needed, compiled, stored, and provided to the lead investigator. Additionally, a separate contractor received data from the Central Data Repository and performed coding of adverse events. The process routinely took two months from the time that data were collected until data were cleaned and present in the compiled data.

In June 2003, the CTN Steering Committee acknowledged that the data management activities were complicated, expensive, and inefficient, that efforts were being duplicated at multiple sites, and that a significant amount of effort was spent coordinating the routine activities of the network between the multiple centers. Contemporaneously, the newly appointed NIDA Director formed a CTN Program Review Committee to evaluate the mission, scope and vision, achievement of goals, and operational efficiency of the CTN. In its final report, submitted in early 2004, the committee recommended that a single or limited number of coordinating centers be established to manage all stages of protocol training, implementation, and data monitoring. Similarly, the published literature on clinical research networks contains many examples and descriptions of central data coordinating centers [2–19]. The transition described in this report mirrored others [20–23] where an initially decentralized model was replaced by a single center. We found no recent papers advocating a decentralized approach.

Methods

The CTN created a central Data and Statistics Center, contractually responsible for data management and statistics for CTN trials. Specific responsibilities are described in detail in RFP N01DA-5-2207, Data and Statistics Center for the Clinical Trials Network. Creation of a Data and Statistics Center provided the opportunity for centralization and for the CTN to construct an informatics infrastructure across all trials supporting the clinical trial life cycle.

A virtual organization, the CTN is composed of numerous organizations with specific roles. Individual system users are members from Nodes, community treatment programs, committees, Data and Safety Monitoring Boards, the Clinical Coordinating Center, the Data and Statistics Center, and the NIDA Center for the Clinical Trials Network (CCTN) office. Based on the functional needs of the network, the CTN informatics infrastructure shown in Figure 2 was created. Components were designed to support the virtual CTN organization in conducting a portfolio of clinical trials. Data collection and management systems provide a consistent way to collect, store, and manage clinical trial data. A trial management and reporting system maintains project plans and tracks progress. Trial progress reports communicate project status and assist in identifying and managing operational problems in trial conduct. A document management system in place prior to the Data and Statistics Center and not further described here provides collaborative workspaces and document-sharing functionality accessible via the Internet. A core group of statisticians and statistical programmers at the Data and Statistics Center provide statistical design and analysis support for trial design, data and safety monitoring, primary analyses, and secondary and pooled analyses for the CTN membership. Data are archived in a central repository, and sharing is accomplished through a public data sharing web site. We focus here on the systems used for data collection and management. Systems supporting data collection, management, and analysis are hosted and maintained by the Data and Statistics Center. Each of the systems used in the CTN informatics infrastructure is described below.

Data collection and management

A web-based electronic data capture (EDC) system is used to collect and maintain CRF data from all CTN trials (Figure 3). The EDC system currently in use is the InForm™ system (Phase Forward, Inc., Waltham, MA). Important features of the system include: support for the configuration of multiple trial databases without changing the underlying application, role-based security, enforced work-flow, on-screen consistency checks during entry, and Title 21 CFR Part 11 compliant technical controls. The data quality achieved with our implementation of the InForm EDC system has been reported by Nahm et al. [24].

The CTN also collects data directly from participants via questionnaires. The Clinical Research Information System (MindLinc, Durham, NC) is used to capture patient-reported outcome data and clinician-administered interviews. Patient-reported outcome data are captured via web-based forms completed by the participants. The system maintains subject-level security through a challenge question and answer. Referential integrity with the CRF data is maintained through the patient and visit identifiers, and enforced by both the Clinical Research Information System and InForm™ system.

The CTN also uses an interactive voice response system for participant randomization. The interactive technology is necessary in the clinic setting where randomization numbers are needed quickly during a visit. The CTN has used a vended system hosted by ALMAC (Almac Clinical Technologies, Yardley, PA). Interfaces to the electronic patient-reported outcome and EDC systems were developed to support real-time randomization and auto population of the EDC system based on responses to questionnaires. Interfacing the three systems also prevents

errors in a key field—the subject identifier—by obviating manual data entry of the subject identifier in three different systems.

Training and the Learning Management System

The Data and Statistics Center implemented the IBM Lotus Learning Management System to provide on-demand training for clinical sites and trial personnel and to decrease training-related costs. However, use of the system was voluntary. As of August 2007, there were fewer than 20 users of the system, and use was discontinued. We attribute this low number to training and associated documentation within the CTN being the responsibility of the Node and site investigators, most of whom use internal systems to document training.

Help desk

Use of web-based clinical trial systems requires a help desk to support users in resolving technical issues. The help desk is staffed from 8:00 a.m. to 8:00 p.m. Eastern Standard Time and employs an automatic call distribution system and after-hours paging system. Call logs are used to determine training needs, monitor system effectiveness, and to plan for future system revisions and updates. The help desk supports the use of EDC systems in four ongoing trials. The steady-state call volume averages 100 contacts a month, including both e-mail and phone calls. The top reasons that sites contact the help desk are to request account creation and to have a password reset.

Clinical trial and portfolio management

In order to understand and monitor the progress of each protocol, and with significant input from the CTN administrative office, the Data and Statistics Center developed monthly trial progress reports. Initially, the reports were static and available monthly. Currently, the enrollment and data status information is available real-time through the EDC system, with custom weekly data status reports to provide work lists to manage outstanding data and queries.

The monthly static reports are comprehensive and include three major parts: 1) summary of active trials; 2) a report on each active trial; and 3) a summary report on both closed and active trials. The key items of site and trial performance monitored in this report include: basic protocol information and timeline, recruitment (actual and expected by site), demographics (sex, race, ethnicity, age), data quality based on queries (by site), data quality based on data audits (by site), quality assurance monitoring regulatory (by site), availability of primary outcome measure, treatment exposure (attendance at active treatment sessions), attendance at follow-up visits, and other protocol implementation issues. The Trial Progress Report uses color-coded flags to reflect the performance of overall and protocol-specific items by site. Table 1 provides the criteria for site performance measurement used in the reports.

Data repository and data sharing

Dissemination and data sharing are important to the CTN, and a matter of NIH policy [25]. A public data share for CTN data was created and can be accessed at www.ctndatashare.org. The data share is built using open source software (www.plone.org). The data from all CTN trials conducted to date have been formatted in the submission data tabulation model standard (www.cdsc.org) and are available via the data share site. Use of standards of this sort assures that like data are stored in the same variable and that data from different trials can be easily combined for pooled analysis. In addition, data prior to anonymization and de-identification are maintained in a secure, central NIDA data repository.

Results and discussion

In 2006, the NIDA CTN was selected as one of 29 best practice clinical research networks by an independent research group leading the Inventory and Evaluation of Clinical Research Networks. The CTN was specifically recognized in the following outcome areas: 1) effectiveness in changing clinical practice; 2) internal interactivity within networks; and 3) use of informatics to advance network goals.

The CTN informatics infrastructure has increased information exchange, knowledge dissemination, and the ability to manage trials consistently. A key metric of translation is the time from last patient-last visit to database lock. Through centralization, the CTN has decreased the time for data lock from an average of 7.6 months to 2.8 months. The average database error rate decreased from 0.8% to 0.2%. Database quality also became more consistent, with a corresponding decrease in the interquartile range from 0.04%–1.0% before centralization to 0.01%–0.27% after centralization, as shown in Table 2. Centralization has provided the CTN with integrated and automated clinical trial status reporting, better management of the trials, and the first standards-based clinical trials public data share—a key strategy to increase the scientific return on investment. In addition, according to a preliminary cost-benefit analysis of one Node, the data management cost for a participant over the life of a trial has been reduced by 50%.

Transition from a decentralized to a centralized model

The decentralized model required each Node to maintain separate but equal infrastructure. Although the promise of a centralized informatics platform was impressive, the transition was a challenging phase for the network. There were many difficulties. With centralization, the network investigators were asked to relinquish local control of data management activities as trials migrated from existing data centers to the Data and Statistics Center. The migrated trials were ongoing, and migration activities added to the complexity of managing the trials. Users from community treatment programs needed to learn new systems and processes, further straining the CTN during the migration period. The Data and Statistics Center struggled to manage the migrations, establish infrastructure, manage ongoing data collection, and start up three new trials simultaneously. Migrating data resulted in additional workload for each trial in terms of data transformation, database construction, data entry, and quality control. The characteristics of the migrating trials are shown in Table 3, and the migration metrics are provided in Table 4.

At the onset, four ongoing clinical trials were selected to migrate to the Data and Statistics Center. Two of the trials used paper forms, and two used web-based EDC systems. To reduce the duration of funding both the existing data centers and the new Data and Statistics Center, the timeline for migration from the Nodes was set at three months. The scope of the migration included transferring data processing responsibilities, migrating existing data, building and validating data entry systems and associated edit checks, and reproducing existing trial management reports. Significant difficulties were presented by one of the two EDC trials to migrate. A custom-built system was in use for the trial, and the Data and Statistics Center commercial system was not able to replicate some of the functionality in the custom-built system. In addition, instead of direct entry from clinic charts, CTN sites completed paper CRFs and entered the data into the existing EDC system from the paper CRFs. The commercial EDC system screen structure displayed one question/answer set per row, and could not visually match the existing paper CRFs with multiple items per row and nested numbered lists. The visual differences increased the cognitive load on data entry staff, thus increasing the potential for error and causing great concern. The EDC system differences were considerable and

significant for trial sites to manage during an ongoing trial. For these reasons, and trial-related timing, the CTN decided to abort the migration of the second EDC trial.

Supporting the continuity of trial status reports also presented challenges. Where a decentralized model exists, it is possible that people are reporting information from varied sources, that there are multiple systems of record for some reporting items, that reports created by different individuals or organizations unknowingly use different denominators, and that there is no database from which to pull some of the information. The CTN was using Excel templates to generate detailed status reports of site start-up, enrollment, data collection, and data quality. Each lead investigator for a trial was responsible for maintaining a report according to a template and providing the report on monthly calls. Much of the data reported were not present in the Data and Statistics Center's standard data status reports used for other trials. It took two months for the data center and trial teams to understand and agree on the reporting needs, including making infrastructure decisions about data sources and what information would be reported via each mechanism.

As a contract, systems developed and used by the data center must be 21 CFR Part 11 compliant, as well as meet Federal Automated Information System Security Program requirements as documented in the Department of Health and Human Services Handbook, and Office of Management and Budget Circular A-130 Appendix III "Security of Federal Automated Information Systems." For example, user-level log-in and role-based security were necessary on all systems. Roles and associated work-flow and system privileges had to be agreed upon prior to configuring trial-specific systems. Because the Nodes are funded through Cooperative Grant Agreements, systems developed and used by the local data management centers did not have to meet these requirements. These additional requirements slowed system start-up in comparison to previous practice, and the systems meeting the requirements were perceived by the network as cumbersome to use.

When our migrations began, information on data structures and formats, number of data sources, and reporting specifications was not obtainable. This resulted in operational surprises and delays. A migration of trial operations is a project in and of itself and should be managed as such. We recommend that a migration plan be developed with all stakeholders and include:

- Identification of all data sources and field-level specifications from each
- A quantitative assessment of current data entry, cleaning status, and volume of incoming data per month
- A plan for handling queries outstanding at the time of migration
- A timeline for building, transferring, and implementing the destination data system
- Examples of all reports in use and field-level specifications for the reports
- Needs or functional analysis for systems to be replaced
- A timeline for data migration (time until data entry and cleaning are current)

Such a plan should be managed by a dedicated project manager.

Data standards were not in place in the CTN; thus, the migration meant incurring a second round of database set-up costs. Additionally, while data re-entry was less expensive than extensive transformation and validation, the impact on trial timelines and operations was significant, as seen in Table 4. This was further affected by lack of resources to support migration and start-up of new trials. Migrations require equal and more immediate resources than the start-up of a new trial. As was the case here, migration cost is increased where there are multiple local data centers, multiple external data sources, and disparate data formats.

Additionally, based on our experience, when trials have less than a year or two left to completion, the advantages from centralization are not likely to be realized within the life of the trial.

A sustained, healthy clinical trial infrastructure is a critical but often overlooked factor in documenting successful trial outcomes. Because funding for government-supported clinical trial networks is allocated for three to five years and is usually renewed through a competitive process, it is conceivable that a different contractor or grantee will be in place after a renewal funding cycle. Thus, networks must have informatics support that can sustain the continuity of operations through changes in network membership and contractors. To this end, it is essential that stewardship be expected of contractors, grantees, and members. For example, commercial off-the-shelf or open-source software and standards should be used to increase the likelihood that the systems can be obtained and implemented by the next contractor. Contractors and project officers should ensure that software licenses are transferable to the next contractor. In the limited instances where software must be developed de novo, the software should run on commonly supported Relational Database Management Systems and operating systems, and be unconditionally available to the next contractor.

Limitations

The model presented here has limitations in its applicability to other research settings. This evaluation included one network working in the area of substance abuse treatment and may not be applicable to other therapeutic areas. In addition, this work describes creation of a central data and statistics center under contract to the NIH. Circumstances and requirements are likely to be different for networks funded by grant mechanisms, private funding, and foundations. The migration metrics presented here were for the migration of several trials simultaneously and are not applicable to metrics that may be expected for managing the migration of a single trial.

Conclusion

The NIDA CTN is a virtual organization, persisting over time and conducting multiple ongoing trials. As such, the network requires infrastructure to ensure consistent performance, to generate scientific data of the highest integrity, and ultimately to meet the goal of improving the treatment of addiction. Centralization and the resulting infrastructure have increased the timeliness of data as well as data quality while decreasing the data management costs. Establishing an informatics infrastructure with the core components presented here took two years, a team that at times numbered over 50 people, and many steps.

Acknowledgments

This work was partially supported by NIDA Contract No. HHSN271200522071C.

The authors wish to acknowledge those who have contributed to the NIDA CTN, especially the principal investigators and staff at the regional research training centers and community treatment centers. Their dedication to addiction treatment, participative leadership, and open engagement has created a productive and improvement-oriented research endeavor. The authors also appreciate the comments from reviewers and the editorial assistance of Amanda McMillan, both of which contributed to the improvement of this paper.

Abbreviations

CCTN	Center for the Clinical Trials Network
CRF	case report form
CTN	clinical trials network

EDC	electronic data capture
NIDA	National Institute on Drug Abuse

References

1. Lamb, S.; Greenlick, MR.; McCarty, D. Bridging the Gap Between Practice and Research. Washington, D.C: Institute of Medicine, National Academy Press; 1998. p. 113
2. Dzik W. The NHLBI Clinical Trials Network in transfusion medicine and hemostasis: an overview. *J Clin Apher* 2006;21:57–59. [PubMed: 16619228]
3. Albert R. COPD Clinical Research Network. The NHLBI COPD clinical research network. *Pulm Pharmacol Ther* 2004;17:111–112. [PubMed: 15123218]
4. Palevsky P, O'Connor T, Zhang J, Star R, Smith M. Design of the VA/NIH Acute Renal Failure Trial Network (ATN) study: intensive versus conventional renal support in acute renal failure. *Clin Trials* 2005;2:423–435. [PubMed: 16317811]
5. Goss C, Mayer-Hamblett N, Kronmal R, Ramsey B. The Cystic Fibrosis Therapeutics Development Network (CF TDN): a paradigm of a clinical trials network for genetic and orphan diseases. *Adv Drug Deliv Rev* 2002;54:1505–1528. [PubMed: 12458158]
6. Kephart D, Chinchilli V, Hurd S, Cherniack R. Asthma Clinical Trials Network. The organization of the Asthma Clinical Trials Network: a multicenter, multiprotocol clinical trials team. *Control Clin Trials* 2001;22 (6 Suppl):119S–125S. [PubMed: 11728618]
7. Pogask R, Boehmer S, Forand P, Dyer A, Kunselman S. Asthma Clinical Trials Network. Data management procedures in the Asthma Clinical Research Network. *Control Clin Trials* 2001;22 (6 Suppl):168S–180S. [PubMed: 11728622]
8. Heart Special Project Committee. Organization, review and administration of cooperative studies (Greenberg Report): a report from the Heart Special Project Committee to the National Advisory Heart Council, May 1967. *Control Clin Trials* 1988;9:137–148. [PubMed: 3396364]
9. Rouff J, Child C. Application of quality improvement theory and process in a national multicenter HIV/AIDS clinical trials network. *Qual Manag Health Care* 2003;12:89–96. [PubMed: 12747132]
10. Weisdorf D, Carter S, Confer D, Ferrara J, Horowitz M. Blood and Marrow Transplant Clinical Trials Network (BMT CTN): addressing unanswered questions. *Biol Blood Marrow Transplant* 2007;13:257–262. [PubMed: 17317578]
11. March J, Silva S, Compton S, et al. The Child and Adolescent Psychiatry Trials Network (CAPTN). *J Am Acad Child Adolesc Psychiatry* 2004;43:515–518. [PubMed: 15100557]
12. Olola C, Missinou M, Agane-Sarpong E, et al. Medical informatics in medical research—the Severe Malaria in African Children (SMAC) Network's experience. *Methods Inf Med* 2006;45:483–491. [PubMed: 17019501]
13. Love M, Pearce K, Williamson A, Barron M, Shelton J. Patients, practices, and relationships: challenges and lessons learned from the Kentucky Ambulatory (KAN) CaRESS clinical trial. *J Am Board Fam Med* 2006;19:75–84. [PubMed: 16492009]
14. Doepel L. NIAID funds pediatric AIDS clinical trials group. *National Institute of Allergy and Infectious Diseases. NIAID AIDS Agenda*. 1997 Mar 12;
15. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). NIDDK RFA DK 02-033: Hepatotoxicity Clinical Research Network. [Accessed August 6, 2007]. Available at <http://grants.nih.gov/grants/guide/rfa-files/RFA-DK-02-033.html>
16. Wilson D, Dean M, Newth C, et al. Collaborative Pediatric Critical Care Research Network (CPCCRN). *Pediatr Crit Care Med* 2006;7:301–307. [PubMed: 16738501]
17. Escolar D, Henricson E, Pasquali L, Gorni K, Hoffman E. Collaborative Translational research leading to multicenter clinical trials in Duchenne muscular dystrophy: the Cooperative International Neuromuscular Research Group (CINRG). *Neuromuscul Disord* 2002;12:S147–S154. [PubMed: 12206809]
18. Zuspan S. The Pediatric Emergency Care Applied Research Network. *J Emerg Nursing* 2006;32:299–303.

19. Canner P, Gatewood L, White C, Lachin J, Schoenfeld L. External monitoring of a data coordinating center: experience of the National Cooperative Gallstone Study. *Control Clin Trials* 1987;8:1–11. [PubMed: 3568691]
20. Pinol A, Bergel E, Chaisiri K, Diaz E, Gandeh M. Managing data for a randomized controlled clinical trial: experience from the WHO Antenatal Care Trial. *Paediatr Perinat Epidemiol* 1998;12 (2 Suppl): 142–155. [PubMed: 9805727]
21. Almercery Y, Wilkins P, Karrison T. Functional equality of coordinating centers in a multicenter clinical trial: experience of the International Mexiletine and Placebo Antiarrhythmic Coronary Trial (IMPACT). *Control Clin Trials* 1986;7:38–52. [PubMed: 3956211]
22. Schmidt JR, Vignati A, Pogash R, Simmons V, Evans R. Web-based distributed data management in the Childhood Asthma Research and Education (CARE) Network. *Clin Trials* 2005;2:50–60. [PubMed: 16279579]
23. Cronin-Stubbs D, DeKosky S, Morris J, Evans D. Promoting interactions with basic scientists and clinicians: the NIA Alzheimer’s Disease Data Coordinating Center. *Stat Med* 2000;19:1453–1461. [PubMed: 10844710]
24. Nahm ML, Pieper CF, Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. *PLoS ONE* 2008;3:e3049. [PubMed: 18725958]
25. U.S. Department of Health and Human Services Web site. Final NIH Statement on Sharing Research Data. [Accessed on August 6, 2007]. Available at http://grants.nih.gov/grants/policy/data_sharing/

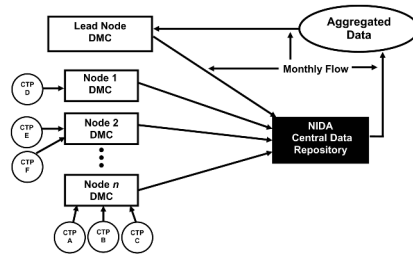


Figure 1. CTN distributed model for data acquisition and management between 1999 and 2005. DMC = Data Management Center; CTN = clinical trials network; CTP = community treatment programs; NIDA = National Institute on Drug Abuse.

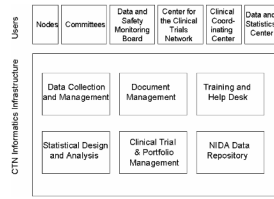


Figure 2. CTN informatics infrastructure components and user community. CTN = clinical trials network; NIDA = National Institute on Drug Abuse.

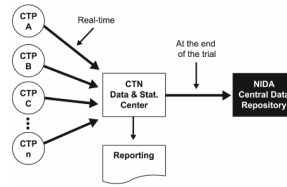


Figure 3. CTN centralized model for data acquisition and management from 2005 to present. CTN = clinical trials network; CTP = community treatment programs; NIDA = National Institute on Drug Abuse.

Table 1

Criteria for key trial performance measures

	Definition of measurement	Green	Yellow	Red^d
Recruitment	Percent = number of actual randomizations divided by number of randomizations expected at this point in time	Actual number of participants randomized is more than 80% of expected	Actual number of participants randomized is between 60% and 80% of expected	Actual number of participants randomized is below 60% of expected
Data quality from audits	Percent = number of data discrepancies divided by number of data fields audited	Error rate is less than 0.5%	Error rate is between 0.5% and 1%	Error rate is greater than 1%
Protocol monitoring^b	N/A	There are no critical issues	Critical issues not related to safety or research integrity exist, and their resolution is initiated or in process	Critical issue resolution has not been initiated, or there are critical issues (whether initiated or not) related to safety or research integrity
Regulatory	N/A	All active CTPs have up-to-date IRB and FWA approval	One or more active CTPs have IRB or FWA approval expiring within the next 30 days	One or more active CTPs do not have up-to-date IRB or FWA approval
Availability of primary outcome data	Percent = number of primary outcome measures with values divided by total number of primary outcome measures expected with values	Percent of primary outcome measures with values is more than 80%	Percent of primary outcome measures with values is between 60% and 80%	Percent of primary outcome measures with values is below 60%
Treatment exposure	Average percent of treatment sessions attended	Average proportion of treatment sessions attended is more than 80%	Average proportion of treatment sessions attended is between 60% and 80%	Average proportion of treatment sessions attended is below 60%.
Follow-up visits	Average percent of follow-up visits attended	Average proportion of follow-up visits attended is more than 70%	Average proportion of follow-up visits attended is between 50% and 70%	Average proportion of follow-up visits attended is below 50%
Other protocol implementation issues^c	N/A	The default is green unless there is one or more significant implementation issues	N/A	One or more significant implementation issues

^aPink color is used for trials that had completed the last follow-up at the last site.^bThis criterion is not cumulative; it reflects protocol monitoring reports received during the prior month only.^cIncludes protocol violations, training, and other issues.

CTP = community-based treatment program; FWA = Federal-wide Assurance; IRB = Institutional Review Board.

Table 2

Database quality before and after centralization

Before centralization	N=60 audits			
<i>Average error rate</i>	<i>Min</i>	<i>Max</i>	<i>IQR</i>	<i>Median</i>
0.81%	0.00%	10.00%	0.04–1.0%	0.22%
After centralization	N=44 audits			
<i>Average error rate</i>	<i>Min</i>	<i>Max</i>	<i>IQR</i>	<i>Median</i>
0.19%	0.00%	0.96%	0.01–0.27%	0.09%

Table 3

Characteristics of migrated trials

Characteristic	Average	Range
Number of unique pages	117	58–176
Total number of pages	370	211–605
Number of query rules (edit checks)	1352	516–2376
Number of data integrations per trial	1	0–2
Number of data centers per trial	2	1–4
Number of total expected pages entered at time of migration	31,298	21,079–56,536 (30–80%)
Total number of pages	96,090	33,335–184,265
Total queries generated & resolved	13,637	2960–23,147
Number of months to next planned analysis at time of migration	4	1–8

Table 4

Metrics of migration

Metric	Average (months)	Range (months)
Time from migration decision to database specifications received	0.6	0.3–1.0
Time from database specifications received to data entry system ready	3.5	1.6–8.5
Time from data entry system ready to data validation checks ready	1.9	0–3.5
Time from data validation checks ready to reports ready	4.2	2.5–7.0
Total duration of system development	8.5	3.0–15.5
Months without reports	8.1	5.6–11.4
Time from database specifications received to all data received	3.6	2.3–4.6
Time from migrated data received to all migrated data entered	3.8	3.0–5.1
Time from last patient last visit to database lock	2.8	2.3–3.3