Application of Technology ■

Electronic Clinical Trial Protocol Distribution via the World-Wide Web: A Prototype for Reducing Costs and Errors, Improving Accrual, and Saving Trees

Lawrence B. Afrin, MD, Valarmathi Kuppuswamy, BE, Barbara Slater, BS, Robert K. Stuart, MD

Abstract Clinical trials today typically are inefficient, paper-based operations. Poor community physician awareness of available trials and difficult referral mechanisms also contribute to poor accrual. The Physicians Research Network (PRN) web was developed for more efficient trial protocol distribution and eligibility inquiries. The Medical University of South Carolina's Hollings Cancer Center trials program and two community oncology practices served as a testbed. In 581 man-hours over 18 months, 147 protocols were loaded into PRN. The trials program eliminated all protocol hardcopies except the masters, reduced photocopier use 59%, and saved 1.0 full-time equivalents (FTE), but 1.0 FTE was needed to manage PRN. There were no known security breaches, downtime, or content-related problems. Therefore, PRN is a paperless, user-preferred, reliable, secure method for distributing protocols and reducing distribution errors and delays because only a single copy of each protocol is maintained. Furthermore, PRN is being extended to serve other aspects of trial operations.

■ **J Am Med Inform Assoc.** 1997;4:25–35.

Clinical trials are essential for progress in medicine, but current operational methods typically are inefficiently mired in the movement of great volumes of paper, a costly approach from many perspectives. Poor accrual rates also contribute to high costs by prolonging trials; poor mechanisms for making community physicians aware of trials at nearby centers contribute to the accrual problem. We developed a Web-based system, the Physicians Research Network (PRN), to facilitate intramural paperless protocol distribution. Also, PRN aims to boost accrual by facilitating community physicians' inquiries about available trials. This paper describes the process of developing PRN and the summary usage data from the first four months of operation. Finally, PRN is compared to other Web-based protocol distribution systems, and plans for the expansion of PRN to other aspects of clinical trials are described.

Affiliation of the authors: Division of Hematology/Oncology, Department of Internal Medicine, Medical University of South Carolina, Charleston, SC.

This research was partially done under a cooperative agreement between the National Institute of Standards and Technology's Advanced Technology Program (under the Healthcare Information Infrastructure Technology contract, number 70NANB5H1011) and the HealthCare Open Systems and Trials, Inc., consortium.

This work was previously published in abstract form and presented at the Spring 1996 Congress of the American Medical Informatics Association, Kansas City, Missouri, June 7, 1996.

Correspondence and reprints: Lawrence B. Afrin, MD, Division of Hematology/Oncology, Room CSB903, Medical University of South Carolina, 171 Ashley Avenue, Charleston, SC 29425.

Received for publication: 7/3/96; accepted for publication: 8/26/96.

Background

Operational methods for clinical trials are well standardized and have changed little over the last 25 years.[1,2] Within each trial center, significant operational resources are expended on protocol distribution

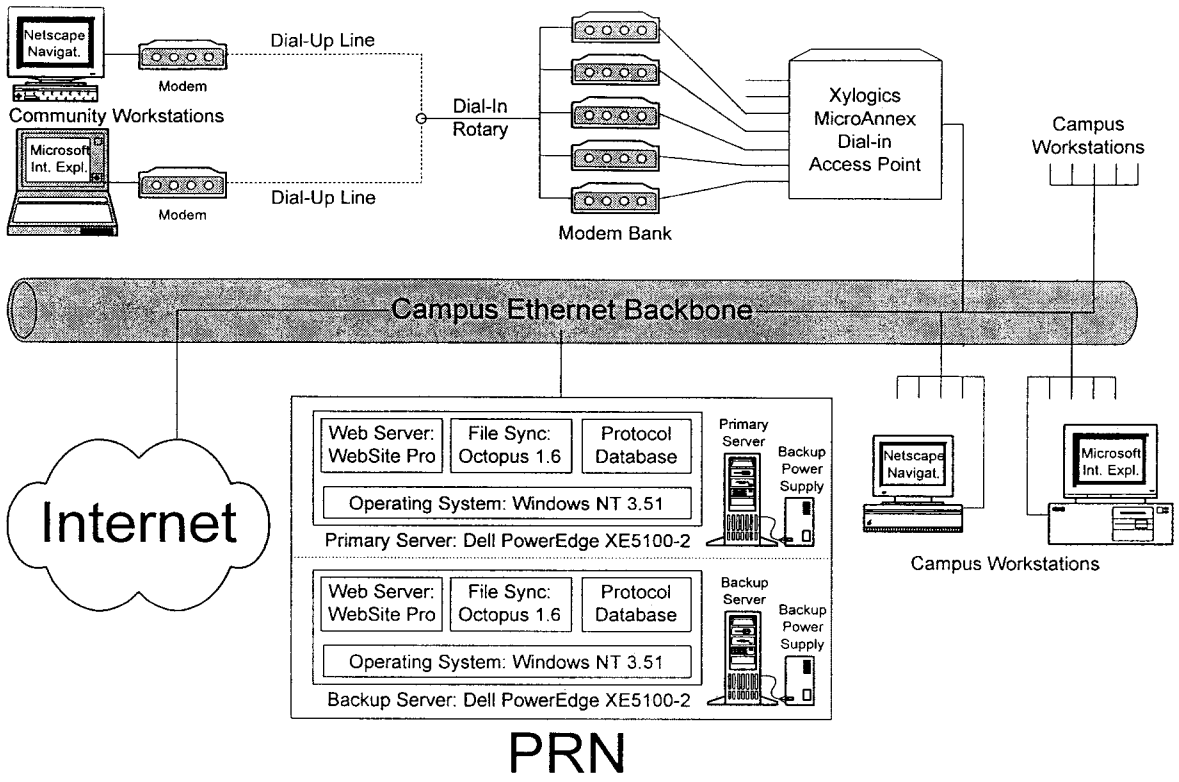

Figure 1 PRN Production System Components. Workstations running Web browsers (e.g., Netscape Navigator, Microsoft Internet Explorer) have high-speed campus Ethernet access or slower dial-up access (via the Xylogics MicroAnnex remote communications server) to the PRN Web, comprised of identically configured, Intel Pentium-based primary and ''hot spare'' backup servers from Dell (on different subnets) running O'Reilly & Associates' WebSite as a Web server package on top of Microsoft Windows NT 3.51. Octopus Technologies' Octopus package keeps the backup protocol database in sync with the primary one. If the primary server fails, Octopus immediately switches the backup into service; users' access to PRN is affected only by workstation, workstation subnet, or backbone failures, not by server or server subnet failures.

to all points where this information is needed. For multicenter trials, protocols and their updates are first distributed to main members, and redistribution to a subnetwork of affiliate centers may also ensue. Intergroup trials add another level of scaling.

Protocol distribution involves a veritable blizzard of paper. Information in protocol version #N can be made obsolete by version #N+1 soon after the last copy of version #N is distributed. Current paper-based protocol distribution methods cause many problems: *delays* in protocol information reaching regulatory bodies and front-line personnel; *errors* in document copying and distribution; and significant labor and materials *costs*.

Furthermore, despite substantial resources invested in trials, trial progress usually is slow due to poor accrual rates, typically less than 3% of all potentially eligible patients at time of diagnosis.[3–6] Physicians are generally aware that outcomes are usually improved by participation in trials[7–12] but frequently do not know of trials available at nearby centers.[3,6,13–16] Improved protocol distribution methods and heightened awareness of trials should speed trials and decrease errors and costs. Electronic systems likely will be the foundation of these improvements because of their speed and other factors. The National Cancer Institute's PDQ system was an early effort in this area, providing summary information about cancer trials.[17]

Previous electronic systems for distributing trial information have suffered from limitations in information quantity and quality, poor accessibility, and difficult interfaces for end users.[18–20] Many newer systems are based on the World-Wide Web (e.g., institutional systems such as at Thomas Jefferson University[21] and the National Institutes of Health,[22,23] and systems of broader scope such as CenterWatch[24]). While easier to use, some new systems still suffer from many of the same problems, such as infrequent updates and availability of only summary information about available trials. A few oncology cooperative

groups, such as the SouthWestern Oncology Group (SWOG), now have Web-based methods for distributing protocols in their entirety,[25] although the methods used have some limitations (see Comparisons, below).

We developed a new Web-based protocol distribution system (PRN) that addresses these issues of information quantity, quality, and accessibility. We used the Medical University of South Carolina's (MUSC) Hollings Cancer Center's Clinical Trials Program as a testbed, although the system's design is not specific to cancer trials. Also involved in the testbed were physicians and other staff from selected South Carolina practices. These practices had a history of referring patients for trials at our cancer center, and they were interested in easier methods of identifying available trials and referring patients.

When the project was begun in September 1994, no academic or commercial systems meeting our design objectives existed (a situation we feel persists today). Also, no advanced Web development tools existed (e.g., HTML editors or middleware for interfacing Web servers and databases), limiting the sophistication of the system's initial design (an issue currently being addressed with the addition of such features as version control). On the positive side, however, our institution's extensive information technology (IT) infrastructure already was widely Web-enabled or Web-capable, eliminating the potential problem of inadequate intramural access to the protocol library.

Design Objectives

Ultimately, PRN intends to be a comprehensive set of automation tools for clinical trial operations. For the initial phase of development, we focused on protocol distribution. The IT field is littered with examples of proprietary document distribution systems that have fared poorly over time for many reasons (e.g., FrameMaker, InterLeaf, BookReader, Envoy, Common Ground, and Replica). We wanted to avoid their deficiencies and develop an open-standards–based, platform-independent system that would be applicable to all types of trials on all scales (e.g., single-center trials, multi-center trials, intergroup trials). Ease of maintainability and use, reliability, security, speed, and cost-effectiveness also were considered important. Because methods of version control meeting our design objectives were not available early in the system's development, we were able to provide users with a guaranteed current version of each protocol but not any older versions. With new tools recently created for Web development purposes, this deficiency is now being addressed.

Based on trends in the development of the World Wide Web over the year prior to project inception, we felt a Web-based system would be the best distribution platform, allowing a trial center to eliminate all hardcopy distribution and maintain only a single, central, fully hyperlinked copy of each protocol, thus giving all staff involved in trials rapid access to any part of any protocol. Furthermore, community physicians' workstations could be configured for ''one-click'' access (via modem) into the protocol library's home page, providing an easy means of discovering available trials and facilitating rapid eligibility inquiries through semi-automatic means (described below).

System Description

Web Server and Protocol Library Construction

Initially, a small prototype system was developed as a feasibility study of whether Web technology could be used in the fashion proposed. As attention shifted to production system design issues such as reliability and security, an optimal set of production system components was identified (Fig. 1). All hardware and software components were selected for their reliability and are protected by a variety of backup, monitoring, and alarm systems. The servers, which are on different subnets, are configured so that, if the primary server ever fails, the backup server automatically and immediately begins functioning as the primary server. Thus, users are not affected by server failures or failures of the primary server's subnet. (Global campus network failures still render PRN inaccessible, though additional failsafes are being developed, as mentioned below.) Security has been implemented via O'Reilly & Associates' (Sebastopol, CA) WebSite Web server facilities for restricting access by user identity and location; further security is provided by a separate user identity check at the Xylogics (Burlington, MA) MicroAnnex dial-up entry point as well as by other mechanisms.

Due in part to the novelty of the system, we were unable to obtain electronic sources for most of the protocols. Therefore, via the process shown in Figure 2, we loaded all of our active oncology protocols (for local, industry, and group trials) from their paper sources into HyperText Markup Language (HTML) format in the server, with each protocol section represented as a single HTML page. A procedure for handling protocol updates has been in place in the trials program from the beginning of the PRN project to ensure that protocols already loaded are maintained to be current. Protocols and updates are processed within 24 hours of receipt (or regulatory approval, if

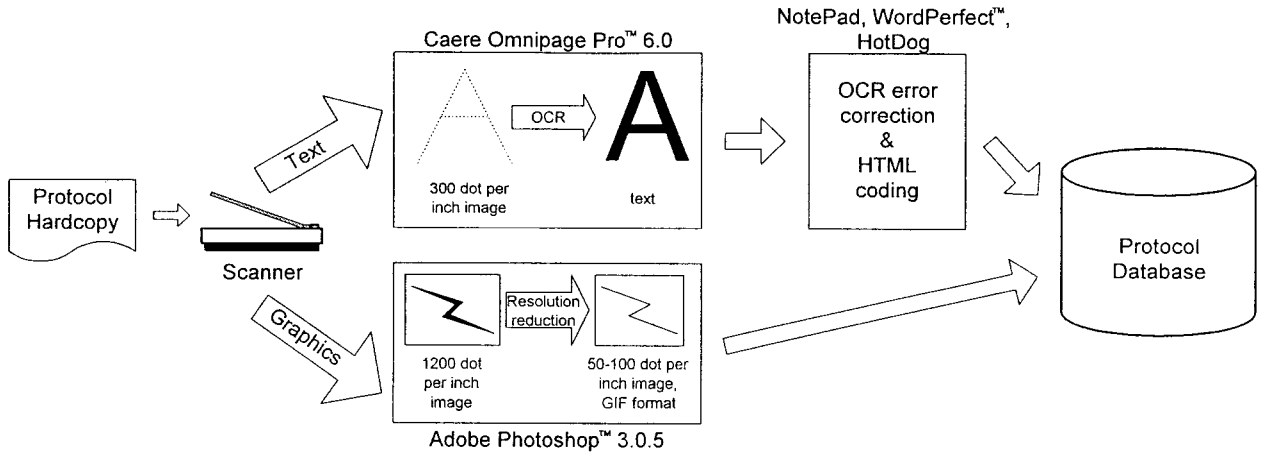

Figure 2 Protocol Loading Process. Protocol hardcopy is fed through an Envisions Env24Pro scanner. Scanned images of text are converted by the Optical Character Recognition process into text, after which editing tools are used to correct OCR errors and insert HyperText Markup Language (HTML) coding to instruct the browser how to display the text. HTML coding includes hypertext linking of all of the protocol's internal references (e.g., literature references, "See section . . ." references, etc.). Graphics are handled separately from text. Each graphic image's resolution is reduced to a reasonable lower limit of readability to decrease downloading time when the image is being retrieved through a (relatively low-speed) dial-up connection to the Internet. Graphics are stored (separately from their associated HTML pages) in Graphics Interchange Format files, displayable by all graphics-capable Web browsers. The HTML pages reference the GIFs as in-line graphics so that, when the page is displayed, associated graphics are automatically shown, too. In the PRN web, each protocol section comprises a single HTML page regardless of the section's hardcopy length. A protocol's HTML pages and GIF files are stored in a unique directory allocated for the protocol in the Windows NT file system. The protocol "database" is this set of protocol directories. Protocol indices, including the full-text search index, are updated after a protocol has passed review.

required). As a safety net, monthly cross-checking between IRB approval records and the protocol library is also performed to ensure the clinical trials program has not neglected to inform the PRN library manager of a protocol or an update. Because of the limitations in current OCR technology, introduction of errors into the electronic protocols was a significant concern, which we addressed by mandating a careful primary manual review of OCR output, particularly numerical data such as all dosages. After a primary review of each protocol, a random sample of ten paragraphs and ten hyperlinks is again reviewed by a second reviewer, who for training reinforcement purposes notifies the primary reviewer of each error caught. Errors of more significance than obvious typographical errors trigger a more thorough secondary review. Each Web page produced includes instructions in how to notify the protocol library manager if an error is found. Of note, reviewers found frequent typographical errors in the source hardcopy protocols. In the future, access to electronic protocol source documents will significantly decrease the 4 hours of labor required on average to completely load a protocol when starting from the hardcopy. We already are seeing increased cooperation on this issue from authors of local protocols. Both our IRB and our cancer center trials program, which has its own pre-IRB protocol review and approval process, are now considering establishing electronic protocol submission as a prerequisite for the regulatory review and approval process.

Several indices into the protocol database were created, including a disease-oriented index, an index by protocol source, an index by protocol identification number, and a searchable index of the full text of all protocols. While standardized vocabularies were not used in the disease-oriented index because of the relatively few diseases studied by the protocols in this pilot, such an approach would be valuable as PRN is expanded to include more protocols and to specialties other than cancer. Due to the aforementioned absence of advanced Web-development tools during the early development of PRN, the protocol indices currently are hard-coded in HTML. New tools facilitating platform-independent integration of Web servers and database management systems are now allowing us to develop a database-based approach to protocol library management. This approach provides much greater flexibility and easier library maintenance through such features as dynamic index creation, version control, and a Web interface to library management functions.

All personnel needing access to protocol documents were identified, assigned access codes, and trained in use of the Netscape Web-browsing software.

Provision of PRN Access to Selected Community Practitioners

For the pilot, three group practices (consisting of nine oncologists, three general practitioners, and six ancillary staff) in our institution's referral cachement area agreed to serve as test sites for community-based access to the server. For the different computing environments (Windows 3.1, Windows 95, Macintosh, OS/2) used by these physicians, we designed a set of packages consisting of inexpensive or free software that provides a modem-equipped computer access to the Internet; a mouse click on a displayed icon establishes the connection, launches Netscape, and displays the "home page" (principal entry point) for PRN. All security checkpoints but one are negotiated automatically; the user is not aware of such processing as long as he or she is operating in an authorized environment. The only checkpoint the user sees is a request for the PRN access code when proceeding beyond the home page. The local authentication procedures currently used may be replaced or supplemented by more secure verification/authentication servers in the future.

Each community-based user was required to sign an agreement prohibiting release of protocol information and ad hoc clinical use of protocol treatments. Then, PRN access was installed on desktop, laptop, office, and home computers as desired by the practitioners. We also developed a method (Fig. 3) for the community user to inquire via PRN whether a patient is eligible for a trial or whether a trial is available for a patient. Traditionally, the community user would be deterred from such inquiries by the typical need for several phone calls to identify the appropriate data manager or investigator. With PRN, the user makes the inquiry and lets the system identify and notify the appropriate staff member, who, desirous of improving accrual, quickly contacts the user to further discuss the inquiry. This approach clearly is not as sophisticated as current state-of-the-art eligibility determination systems. For two principal reasons, the PRN eligibility inquiry mechanism was purposely not designed as a fully automated approach (i.e., matching submitted patient characteristics against a database of eligibility criteria and immediately informing the inquirer of an "eligible" or "ineligible" result). First, we believed eligibility determination occasionally requires human discretion, and we did not want an automated process that might reject a potential trial candidate through overly strict interpretation of eligibility criteria. Second, we believed there was valuable goodwill to be gained through the system's promising a referring provider a rapid contact in follow-up to an inquiry and then keeping that promise.

Access logfiles maintained by WebSite and the MicroAnnex were processed and analyzed using WordPerfect, Microsoft Excel, and Visual Data.

Status Report

Production hardware and software cost $48,443, but this included approximately $10,000 in material obtained specifically for development, not production, purposes. (Also of note is that the processor, memory, and storage capacities purchased were much higher than needed just for the pilot described in this study. The specifications were derived from necessarily very rough estimates of system demand if all of the Medical University of South Carolina's clinical trials programs were to make use of PRN.) From September 1994 to January 1996, three workers (VK, LA, and BS)

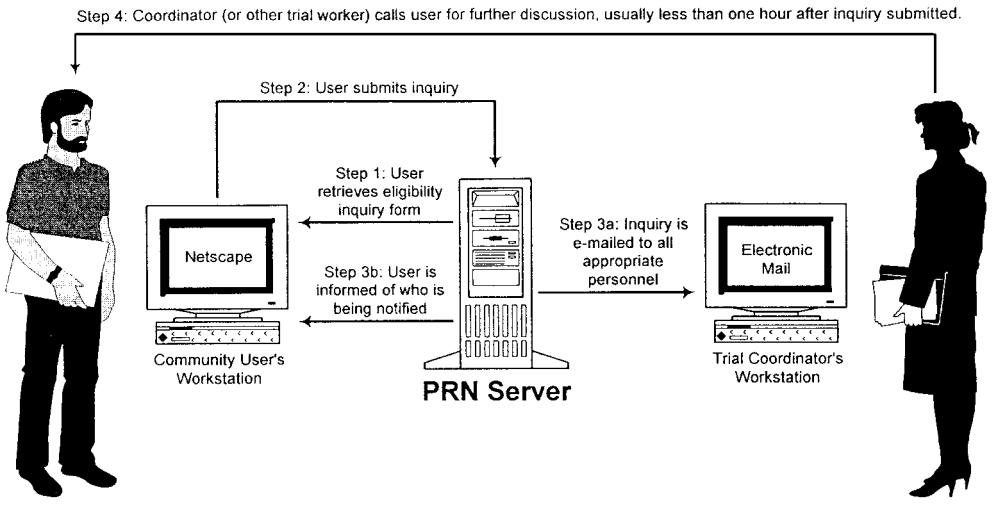

Figure 3 PRN trial eligibility inquiry process.

loaded 147 protocols in 14 disease categories (e.g., 20 protocols for breast cancer, 49 for lymphohematopoietic malignancies, 17 for gastrointestinal malignancies, 13 for lung cancer) into the protocol database; most were input by a trained, part-time student worker (VK) who spent 581 hours on the task from January 1995 to January 1996. A full-time PRN library/system manager (BS) was hired in December 1995 and also assisted in the effort. Labor costs for the complete protocol loading effect totalled $24,492 (including fringe benefits; disproportionately weighted by physician LA from September 1994 through January 1995). Additional labor costs for system development (all LA, 50% effort) through January 1996 totalled $42,500 (including fringe benefits; again disproportionately higher than normal Web development costs because of the typical salary range of the developer's primary occupation).

In comparing costs of PRN versus the older, paper-based system, we note that total labor previously spent within the adult oncology clinical trials program on photocopying and distributing protocols amounted to 1.0 full-time equivalent at a cost of approximately $25,000 including fringe benefits. After the roll-out of PRN, that FTE in its entirety was redirected to other tasks. Mean photocopier usage decreased 59%, from 14,957 pages per month in a 3-month period prior to roll-out to 6,188 pages per month in a similar period following roll-out, for a projected annual savings of over $2,200. Precise determinations of cost-effectiveness, which would include such factors as savings from decreased photocopier wear and tear and benefits from increased availability of protocol documents, were not performed in this study.

The production system was installed and activated in early June 1995 as we were nearing completion of loading all the adult non-gynecologic oncology protocols into the database. After testing, we began to roll out the system to all oncology trials personnel, officially "going live" on October 1, 1995. From August 1995 to February 1996, we also set up a total of 9 computers and 18 users in three community practices for access to the system (most of these users at one large practice in February). Modem speeds ranged from 2,400 bits per second (bps) to 115,200 bps; even at 2,400 bps, PRN response time was acceptable due to the near complete absence of graphical content. There has been no known failure of server hardware or software to date.

From September 1994 through January 1996, the PRN Web server handled a total of 46,788 "hits," or requests for information. The majority of the hits were from project staff for testing purposes. In our analysis of actual production usage, we excluded hits on graphic images (served separately from the pages containing them) and hits from development and testing staff, and we also excluded activity prior to roll out of the production system to regular users. After these exclusions, logfile analysis from June 1995 through January 1996 showed 2,719 hits on protocol pages. Hits initially averaged only 19 per week but acutely increased to a mean of 144 per week (Fig. 4) when the trials program, on November 1, 1995, destroyed the ten paper sets of these protocols previously maintained at various sites around the institution.

A total of 213 users had been trained and given access codes as of January 1996 (Table 1, Table 2). We believe the categories of personnel represent what would be expected to be a typical spectrum of production system users. In particular, the community practices were selected to include a large, sophisticated, trials-oriented urban oncology group practice, a more typical small, rural oncology practice, and a small rural hospital-based general group practice; all practices had preexisting referral relationships with our cancer center and expressed interest in having access to the PRN protocol library. Community users had made no eligibility inquiries via PRN as of January 1996. Access logs show 84 of the 213 authorized users (39%) used the system at least once. Twelve users (5%) performed a remote, dial-up access at least once. The Micro-

Table 1 ■

Breakdown of PRN Users and Usage; Types

Type of User	Number of Registered Users	Number of Actual Users	Percentage of Total System Use by Actual Users (%)
Physician	41	20	36
Nurse	60	31	15
Clinical trials office personnel	21	9	15
Pharmacist	66	18	17
Miscellaneous	25	6	17
Totals	213	84	100

Terms: Physician: any user with an MD, PA, or NP degree involved in the care of our cancer trials patients. Nurse: any similarly involved office- , clinic- , or hospital-based user with an RN or LPN degree. CTO personnel: any user employed in the Clinical Trials Office of our cancer center or the CTOs of the participating community practices. Pharmacist: any pharmacist or pharmacy resident involved in the care of our cancer trials patients. Miscellaneous: any user not in any other user category (e.g., administrators).

Annex dial-up server recorded 951 accesses for a total access time of 427 hours since it was first activated at the end of July 1995. As the MicroAnnex provides access not to PRN specifically but to the Internet at large (via the MUSC network) and does not log a user's specific activity on the Internet, it is not possible to say what portion of the 427 hours was spent examining protocol material from the PRN web. Three MUSC physicians accounted for 30, 35, and 85 hours of dial-up use (over 49, 89, and 108 accesses), and 3 community-based physicians accounted for 22, 23, and 55 hours of use (over 59, 66, and 128 accesses). Physician-users usually spend 20–25 minutes online in each dial-up session.

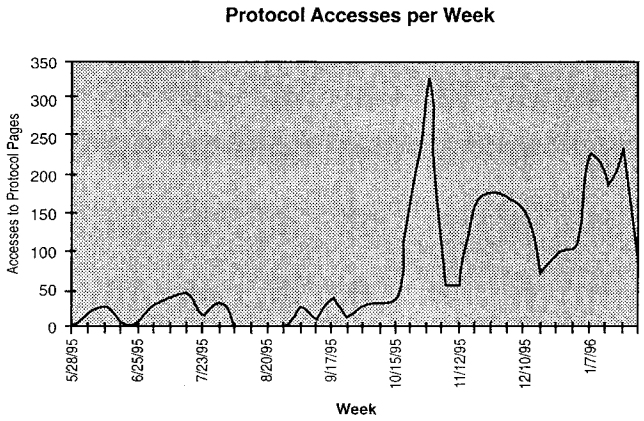

Figure 4 Protocol accesses per week.

Discussion

Current, accurate protocol information must be conveniently and quickly available to all personnel involved in clinical trials. Standard clinical trial methods provide for paper-based protocol distribution, causing delays, errors, and significant expenditures. We felt replacement of the paper-based system with an electronic one based on the Internet would be helpful.

The Web today is the principal method for publishing and retrieving information across the Internet[27] and is an advantageous environment for protocol distribution because only a single master copy of a protocol need be maintained. Furthermore, security can be applied so that only authorized personnel can retrieve information on protocols whose authors request restricted distribution.

Round-the-clock accessibility to protocol information is critical. Protection of protocol information against unauthorized access is equally important. Thus, reliability and security were paramount considerations in the process of designing our system. Operational simplicity, flexibility, high-speed performance, and cost were secondary but still significant design issues.

Table 2 ■

Breakdown of PRN Users and Usage: Locale

User Locale	Percentage of Registered Users (%)	Percentage of Actual Users (%)	Percentage of Total System Use by Actual Users (%)
MUSC-based user	85	95	97
Community-based user	15	5	3
Totals	100	100	100

Security

Numerous security features provide more protection against unauthorized access to protocol information than current paper-based systems. No other Web sites have links to PRN, and Web indexing robots are prohibited access; authorized users' browsers are set up with bookmarks to reach PRN. Community users must sign the agreement described above (see Methods), and they are funneled through a secure access point to reach the PRN home page. All users are scrutinized by the server for location; access attempts from unauthorized locations are rejected. Current measures provide more security than in the paper environment; we have not yet determined a need to encrypt PRN transmissions, though encryption will be desirable when PRN is enhanced to manage patient-specific data (such as for on-line trial data collection).

Simplicity

All aspects of the system's daily operation were designed to be as simple and automated as possible to avoid needing a trained computer technician to manage the system and to facilitate our ability to package the system and distribute it to other centers interested in pursuing the same path. To this end, Windows NT Server provides a familiar filesystem and graphical user interface. WebSite provides convenient Web server management tools. All PRN documents are stored as simple files within the native Windows NT filesystem, rather than involving more complex database packages. Images are stored in formats that can be displayed by all Web browsers. We use only very basic HTML coding, which adheres closely to the defined HTML standard.[28] All routine system maintenance tasks are highly automated, such as generating new user access codes, backing up disk storage devices onto tape, generating system usage reports, and archiving access log files.

Room for improvement exists in the process for loading protocols into PRN: loading from paper form by document scanning and OCR requires at least twice as much time as loading from electronic form. The labor required to build the protocol database is directly related to the number of protocols to be loaded and their form (paper versus electronic). As previously mentioned, improvement is also possible in the procedures for managing the protocol library; newly available Web development tools are assisting our efforts in this area. An increasingly sophisticated protocol library management system will facilitate future enhancements such as automated tracking of patients through their study calendars (including automated order entry), Web-based data reporting and collection systems, and development of a layman-oriented ''parallel'' library of information on available protocols.

Performance

Server performance has been excellent, with periodic monitoring showing reserve capacity sufficient to accommodate 2,000 times the current activity. Furthermore, should server performance fall unacceptably as workload increases, the system can accommodate additional and/or faster processors, memory, storage, and network hardware.

Performance is more of an issue at the user level, where older workstation hardware that is fundamentally inadequate to support fast Web browsing can be frustrating. For the user who solely accesses PRN, the only significant delay on such systems is the initial 15–30 seconds for launching Netscape. From then on, performance is always adequate because PRN documents are short and almost always free of graphic images. Performance in low-end systems usually benefits most by increasing workstation memory, followed by upgrading to a faster processor if necessary. Faster network connections are not needed for PRN but may improve access to other, graphics-intensive Web sites.

Costs and Savings

Materials costs for implementing PRN today for a single trials program would be less than we encountered due to steadily falling hardware and software prices and lower processor, memory, and storage capacities than we required to handle the estimated needs of the many trials programs we projected will eventually use our pilot system. Labor costs stem almost exclusively from loading and maintenance of the protocol database, tasks which can be accomplished by secretarial-level personnel. In our experience, the production operation has rarely required attention from technical personnel. Still, a PRN implementation for a single trials program might not be cost-justified for a program smaller than MUSC's cancer clinical trials program.

Although a formal cost–benefit analysis was not a part of this study, the production operation of the PRN pilot does not appear to cost significantly more than the previous paper-based protocol distribution operation. Hardware difficulties requiring expensive repairs could have altered this situation if we had encountered such problems. If PRN is managed as a shared resource (hardware, software, and personnel) among multiple trials programs (e.g., perhaps operated by the institution's IRB), then the cost per program could be reduced. After the initial protocol library is built for any one trials program, only a portion of the library manager's time is needed to maintain that library, allowing a single manager to maintain the libraries of multiple programs. Furthermore, the resulting per-program cost is potentially recoverable through industry trial case management fees or other sources of indirect funding used to support clinical trial operations. (PRN might even help generate higher per-case fees by being represented as an effort to market industry trials to community physicians—and thus attempting to improve industry trial accrual.) We believe much of an institution's protocol maintenance cost will be eliminated if cooperative groups and industry trial organizations embrace the PRN concept (see Expansion and the Future, below).

Outcomes

Results from early use of PRN are encouraging but must be considered preliminary in view of the short period of data collection. In the first 4 months of production use, we received no negative feedback from internal or external users despite frequent, repeated polling. There were no complaints of large sections missing from, or outdated in, the protocols, a significant change from the paper-based system. The 1.0 FTE spent previously in protocol copying and distribution was redirected to more productive tasks, but system management did require an additional 1.0 FTE. Photocopier usage dropped 59%, although this was offset by hardware and software costs. Users of protocol documents felt more secure that the version of the protocol they were retrieving from the server was completely current, reaping slight additional labor savings by not having to confirm the accuracy of the material being reviewed before acting on it.

Although we demonstrated the technical feasibility of providing non-computer–sophisticated users ''one-click'' access to the Web from a spectrum of personal computing equipment, we believe the chief outcome of the extramural deployment was that these physi-

cians and their staff actually used the access regularly. The lack of eligibility inquiries and accrual impact from the extramural testbed was not surprising, given the limited scope of the deployment, the short evaluation period, and the novelty of on-line inquiries to users long used to making inquiries via a series of telephone calls.

The disparity between MicroAnnex usage and PRN Web usage by the community practitioners suggests they were using their dial-up access method more as a means of general access to the Internet than to PRN specifically. We have encouraged this behavior, as we believe that practitioners familiar with information resources on the Internet will be able to use this knowledge to deliver improved patient care.

Other than some logistical delays in scheduling inservice training sessions and community installations, no significant problems were encountered in the pilot. The technology has functioned with no known failures to date.

Although any document, paper or electronic, that was originally composed by human hands is subject to the introduction of human error, we were very concerned about the possibility of introducing additional error during the paper-to-electronic-document conversion process. We implemented procedures we felt would ensure accuracy to an acceptable degree. To date, there have been no reports from users of errors in our electronic protocol library. Even if errors (original or new) are present in the electronic protocol, the numerous existing checks and balances in our system (the attending physicians, fellows, trial coordinators, pharmacists, and nurses and the institutional procedures they are all supposed to follow) would likely cause gross errors to be caught before a patient could come to harm or would allow possibly harmed patients to be identified to facilitate follow-up evaluation. However, it is unlikely that the error rate (original or new) will ever be zero or that the error detection rate (original or new) will ever be 100%. Further discussion of these matters is beyond the scope of this paper. However, it is clear that assimilation of a protocol into the PRN library from the protocol's original electronic source would be far preferable to our current method, not only for improving accuracy but also for reducing labor.

Data regarding PRN continue to be collected; a more rigorous assessment will be performed after the system has been in production use for a year.

Comparisons

As noted previously, some centers are using the Web to distribute summary information about their available clinical trials, and there is at least one organization attempting to develop a Web-based clearinghouse of summary information on trials available at a number of centers.[24] However, when we recently reviewed this database, it was far from complete and was already having problems with information currency.

The SWOG protocol library became available on the Web within a few months of ours.[25] More recently, other cooperative groups also have begun using the Web to distribute protocols to member sites.[29–31] However, these groups bundle each protocol into a single file typically stored in plain-text (ASCII) format (typically 50–150 kilobytes in size) or Portable Document Format (PDF) (typically 1–3 megabytes in size). The former requires the user to download the entire file to see the desired section or page. The latter previously required the same approach, but in PDF's latest incarnation using byte-ranging technology, the user now can download just the page(s) of interest. However, because it represents each hardcopy page as a distinct electronic page completely identical in appearance to the hardcopy original, the PDF-coded protocol can require more time to browse than an HTML-coded protocol in which each section of the protocol exists as a single HTML page (typically 1–10 kilobytes in size). Also, most cooperative groups' protocol libraries currently do not feature internal reference hyperlinking, a feature we believe facilitates users' efforts to find specific items of information. Our library also supports full-text searching, a feature that might be difficult to implement if poorly legible hardcopy originals are stored in PDF files. (Text that cannot be recognized by PDF technology as such is instead stored in the PDF file as a graphic bitmap and thus cannot be included in a full-text search index; however, this limitation is irrelevant if all protocols are available from electronic masters and legibility is not an issue.)

We felt that minimizing the time to download and peruse a protocol section would facilitate a model in which, every time the user needs protocol information, he or she can reference very promptly the lone master version of the protocol (guaranteed accurate and current) instead of a previously downloaded and questionably trustworthy copy. Although in its current pilot implementation PRN's group and industry protocols are not truly ''master versions,'' in the group implementation the user of a group protocol at the local center would retrieve the true master protocol from the *group* server, not the *local* server (see below). Because of the speed advantages of HTML over PDF, together with the immaturity and uncertain future of PDF technology when we started this project in 1994, we chose HTML coding at the time and still feel

HTML is the preferable choice for this application. However, the Web's fundamental reliance on Multipurpose Internet Mail Extensions (MIME) typing for notifying a browser of a downloaded document's format allows PRN to simultaneously accommodate different electronic protocol coding formats for different protocols.

Expansion and the Future

There are many opportunities for expanding PRN to clinical trial operations in areas other than oncology and in other institutions. The system can serve all community practitioners interested in contributing patients to studies. Also under development are version control mechanisms, allowing access to older versions of the protocol when necessary, and additional protocol distribution failsafe mechanisms, such as automated fax-on-demand and text-to-voice systems.

Although PRN eliminates an individual trial center's need for multiple protocol copies, it is also very applicable to the multicenter trial environment (e.g., cooperative groups such as the Cancer and Leukemia Group B and SWOG). To make our pilot a practical system, we had to load into it all of our oncology protocols: local protocols, industry protocols, and cooperative group protocols. Clearly, if this same system is implemented at another institution, then that institution, too, would have to load all of its local, industry, and group protocols. This duplication of industry and group protocols is the antithesis of the Web. If each industry sponsor and group were to implement PRN, a group's server would contain all of the group's protocols, an industry sponsor's server would contain all of the sponsor's protocols, and the trial center's server would only have to contain the center's locally originated protocols. Then, each center's protocol index would link listings for local protocols to the local database, listings for group protocols to the appropriate group databases, and similarly for industry protocols, all but eliminating protocol copying and distribution. Each time a protocol update is necessary, only the single copy in the protocol sponsor's server need be updated. Security arrangements are more involved, since it would be desirable for a user's access code for the local server to also work for group and industry servers. This arrangement requires a system of trusted authentication servers. While such a system is already available in the Windows NT environment through its domain trust model, a platform-independent solution would be preferable.

In summary, we have described a new, relatively inexpensive electronic clinical trial protocol distribution system based on global technology standards, with several advantages over the old, paper-based system, including increased trials program efficiencies. Although it is not documented in this study, we believe electronic protocol distribution via the Web has the potential for decreasing error rates in disseminated information and increasing accrual. We have had a single oncology trials center pilot that has been well accepted by users, and we are pursuing several avenues of enhancement and expansion. In these times of increasing importance of clinical trials but decreasing funding for them, the PRN system is an effective method for helping to address this conflict.

We acknowledge the support of the Medical University of South Carolina's (MUSC) Environmental Hazards Assessment Program (funded under the U.S. Department of Energy's Assistance Instrument #DE-FG01-92EW50625), MUSC's Center for Computing and Information Technology, and Microsoft Corporation.

References ■

1. Jenkins J, Hubbard S. History of clinical trials. Semin Oncol Nurs. 1991;7:228–34.
2. Friedman L. The NHLBI model: a 25 year history. Stat Med. 1993;12:425–31.
3. Mansour EG. Barriers to clinical trials. Part III: Knowledge and attitudes of health care providers. Cancer. 1994;74:2672–5.
4. Avis FP, Ellenberg S, Friedman MS. Surgical oncology research: a disappointing status report. Ann Surg. 1988;207:262–6.
5. Fisher B. On clinical trial participation. J Clin Oncol. 1991;9:1927–30.
6. Farrar WB. Clinical trials: access and reimbursement. Cancer. 1991;67:1779–82.
7. Richardson JL, Myrtle R, Solis JM, et al. Participation of community medical oncologists in clinical research trials. Prog Clin Biol Res. 1986;216:269–80.
8. Winn RJ, Miransky J, Kerner JF, et al. An evaluation of physician determinants in the referral of patients to cancer clinical trials in the community setting. Prog Clin Biol Res. 1984;156:63–73.
9. Wenger NK. Why community physicians should encourage their patients to participate in randomized clinical trials. Circulation. 1978;58:963–4.
10. Stiller CA. Centralised treatment, entry to trials and survival. Br J Cancer. 1994;70:352–62.
11. Stiller CA, Draper GJ. Treatment centre size, entry to trials, and survival in acute lymphoblastic leukemia. Arch Dis Child. 1989;64:657–61.
12. Hussain M, Wozniak A, Valdevieso M, et al. Cancer clinical trials: reasons for poor accrual. Proceedings of the Annual Meeting of the American Society of Clinical Oncologists. 1992;11:A1428 (abstr).
13. Winn RJ. Obstacles to the Accrual of Patients to Clinical Trials in the Community Setting. Semin Oncol. 1994;21:112–7.

14. Fleming ID. Clinical trials for cancer patients: The community practicing physician's perspective. Cancer. 1990;65: 2388–90.
15. Long DG. Clinical trials: A family physician's perspective. Cancer. 1991;67:1798–9.
16. Bateman M, Kardinal CG, Lifsey D. Barriers to minority recruitment: implications for chemoprevention trials. Proceedings of the Annual Meeting of the American Society of Clinical Oncologists. 1993;12:A472 (abstr).
17. Hubbard SM, Henney JE, DeVita VT, Jr. A computer data base for information on cancer treatments. New Engl J Med. 1987;316:315–8.
18. Hubbard SM. Information systems in oncology. In: DeVita VT, Jr., et al. (eds). Cancer: Principles and Practice of Oncology, 4th Edition. Philadelphia: JB Lippincott Co., 1993; 2582.
19. Shaw DJ, Czaja RF. User interactions with the PDQ cancer information system. Bull Med Libr Assoc. 1992;80:29–35.
20. Fare C, Ugolini D. The PDQ (Physician Data Query), the cancer database, in oncological clinical practice. Cancer Treat Rev. 1991;18:137–43.
21. http://www.jci.tju.edu/jcc/JCCProt.html
22. http://text.nlm.nih.gov
23. http://wwwici.nci.nih.gov/prot/dismenu.html
24. http://www.CenterWatch.com/
25. http://www.oo.saci.org
26. http://www.yahoo.com/Computers_and_Internet/Software/Data_Formats/HTML/
27. http://www.cc.gatech.edu/gvu/stats/NSF/merit.html
28. htp://www.yahoo.com/Computers_and_Internet/Software/Data_Formats/HTML/HTML_3_0/1
29. http://ecog.dfci.harvard.edu/
30. http://www.acr.org/rtog/rtog_intro.html
31. http://pog.ufl.edu/