HHS Public Access

Author manuscript

Amyotroph Lateral Scler. Author manuscript; available in PMC 2018 February 13.

Published in final edited form as:

Amyotroph Lateral Scler. 2009 ; 10(5-6): 374–377. doi:10.3109/17482960802378998.

Web-based data management for a phase II clinical trial in ALS

Richard Buchsbaum[1], **Petra Kaufmann**[2], **Alexandra I. Barsdorf**[2], **Rachel Arbing**[1], **Jacqueline Montes**[2], and **John L.P. Thompson**[1,2] **For The QALS Study Group**

[1]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, USA

[2]Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Abstract

The objective was to report on the creation, features and performance of a web-based data management system for a two-stage phase II randomized clinical trial of Co-Enzyme Q10 in ALS. We created a relatively comprehensive web-based data system that provided electronic data entry; patient management utilities; adverse event reporting, safety monitoring, and invoice generation; and standardized coding for medications and adverse events. In stage 1, clinical sites submitted 7207 forms reporting on 105 patients followed for 10 months. Less than 0.7% of submitted forms contained errors. At the time of the delivery of the analysis data set, only four errors remained unresolved. Data were available quickly, with a median time from event to data posting of two days. The data set was locked and the analysis data set produced nine days after the final patient visit. A survey of trial personnel yielded generally positive feedback, with 75% of respondents wishing to use a similar system in the future. Given sufficient resources, a comprehensive web-based data management system can meet the need for clean, available data in clinical trials in ALS and similar diseases, and can contribute significantly to their efficient execution.

Introduction

Clinical trials for ALS (as well as other neuromuscular disorders) pose several challenges. First, ALS is considered an orphan disease: less than 200,000 patients are affected in the United States at any given time [1]. Secondly, increasing disability often limits patients' ability to travel to a clinical trial site [2]. Finally, the typically fatal outcome of the disease can lead to additional compliance and follow-up difficulties [3]. As a result, clinical trials consortia typically have many sites, each enrolling only a few patients. In order to address these challenges, researchers are creating new and innovative trial designs. Rapid access to clean and reliable data, always crucial, is particularly important in these designs.

Web-based data management systems have been widely applied in clinical research in recent years [4–6]. Most of these systems have been largely devoted to electronic data capture (EDC). However, a more comprehensive approach to data management that combines EDC with intensive data quality monitoring, protocol support and a variety of other trial-enhancing utilities [7–9] may be of benefit in ALS trials, given the demands noted above.

The clinical trial of high-dose Coenzyme Q10 in ALS (QALS) is a close collaboration between two academic principal investigators, one clinical and one statistical, each with their own grant, staff and infrastructure. Together the two teams designed a two-stage, phase II trial [10]. The first stage is a dose selection stage, requiring 105 patients randomized to placebo or to one of two doses of CoQ10. Stage 2 of the trial recruits an additional 80 patients randomized to either placebo or the selected dose of CoQ10.

In response to requests from the clinical investigators, the data management team designed a web-based data management system to meet the following specific goals:

- support and enforce adherence to the trial protocol;
- encourage data completeness by identifying and limiting missing data;
- maximize data quality from the outset, to minimize subsequent queries and delay;
- require that data be submitted quickly, so that safety monitoring and monitoring of loss to follow-up can take place in something approaching real time;
- make analysis data sets available as quickly as possible, so that the second stage of the trial could begin with minimum interruption; and
- provide an intuitive user interface for investigators and coordinators.

Methods

The QALS web site is the hub through which most trial information flows. From the site trial personnel can download trial documents such as the latest protocol and procedure manual; print out study newsletters, tracking logs and regulatory forms; generate automatic invoices for site payment, based on submitted data; and manage most aspects of patient recruitment and follow-up, from visit scheduling to data collection and entry. A number of the site's pages can be viewed online as Supplementary material to this paper, at http://www.informaworld.com/.

Access to the site is by password-protected logon, which determines the level of access provided to the site, and thereby to QALS data. All transmissions are encrypted using a 128-bit security certificate. Personal medical information (PMI) sent from clinical sites to the SAC is minimal, and strictly limited to data required for actual analysis.

Patients progress through the QALS protocol in a series of visits, each of which generates one or more data items or procedures. Data for each item are collected on a paper form called a worksheet, then entered by the clinical site into a web page called an electronic CRF (eCRF), which mirrors as closely as possible the layout of the corresponding paper worksheet. Worksheets are dynamically generated for each item. For example, any medication whose consumption is reported as ongoing on a Concomitant Medications form is carried over to the worksheet for the next visit. This both prompts the clinical site to enquire as to the status of the medication, and eliminates the need for re-entry. Data entered into an eCRF may be saved for future editing by clicking a 'Save' link, or posted by clicking

a ‘Save and Post’ link. Posted data are considered stable and can no longer be edited directly by a site.

The layout of the eCRFs prevents many data entry errors through the use of lists, drop-down lists and single-entry option groups. Skip patterns have been largely eliminated by design, and those that remain are enforced by browser-based JavaScript automation. Invalid or out-of-range values are flagged upon entry, and cannot be saved. All fields are required; ‘missing value’ codes are entered when data are unavailable. Such requirements are enforced only at posting, however, so that data entry personnel may save partially completed eCRFs and return to them later.

Since free-text data can be difficult to monitor and analyze, they have been reduced to a minimum. Medications and adverse event diagnoses are captured as text on the worksheets, for ease of use and speed of collection, but then entered into their respective eCRFs with the help of standardized electronic dictionaries. Medications are recorded using the National Drug Code (NDC) directory, obtained from the Food and Drug Administration and augmented with a few nutritional supplements and other treatments relevant to ALS. Adverse events – including serious events (SAEs), non-serious events (AEs) and events making up patients’ medical histories – are recorded using the National Cancer Institute’s Common Terminology Criteria for Adverse Events (CTCAE) version 3.0, augmented with a very few specific events especially relevant to ALS.

The trial protocol is enforced by the web interface, which prevents sites from downloading the worksheets for an upcoming visit until all the data for the previous visit have been posted. Adverse event data entry is divided into two stages, initial report and resolution, which are entered and posted separately. This encourages prompt recording of adverse events while providing a mechanism for recording the resolution of long-lasting events.

In addition to EDC, the QALS data management system supports many other aspects of the trial. For site investigators and coordinators, the QALS web site provides a table of all screened patients, indicating their status and their progress through the trial protocol. A visit schedule can be entered and printed for each patient, and a calendar indicates upcoming patient visits with links to the patient’s data. Shipments of biospecimens are recorded via the web site; when shipments occur, the central laboratory is automatically notified. Once specimens are processed, results are entered though a secure page accessible only by laboratory personnel. In the event of medical emergencies, patient treatment assignment can be obtained via the web site.

Several portions of the web site are unavailable to clinical sites and can be accessed only by authorized users. The safety monitor has secure access to a safety data page, with dynamically generated tables of rates and frequencies of adverse events by body system, and links to data on individual serious events. A diagnosis confirmation page allows the clinical principal investigator to confirm or overturn patient diagnoses for screened patients. A site invoice page generates payment documents for each site, based on posted data.

Data quality web pages provide dynamically generated tables for real-time monitoring of trial data. Trial staff can monitor upcoming visits to ensure collection of the primary

endpoint; visits for which the date window has closed but for which no data has been entered, to ensure timely data entry; patients with discrepant pill counts, to reduce medication non-compliance; patients skipping two or more visits, to reduce loss to follow-up; unexpected SAEs possibly related to the study drug, which require reporting to the FDA; and several others.

The web site generates automatic email notifications to trial personnel when an unexpected SAE possibly related to study drug occurs, when recruitment nears completion, and when recruitment for each stage is complete. Some of these materials can be viewed online as Supplementary material to this paper, at http://www.informaworld.com/.

Results

The results described are as of the completion of stage 1 of QALS.

Recruitment for stage 1 began on 19 April 2005, and the first patient was randomized on 29 April 2005. The final patient for stage 1 was randomized on 9 February 2006. The final study visit for stage 1 was conducted on 29 November 2006. While great effort was made to prevent it, two patients (1.9%) were lost to follow-up.

Full follow-up for a patient consisted of 75 items performed over 12 visits, for a potential total of 7875 eCRFs. As of the close of recruitment for stage 1, 6632 visit-based items were completed. The remaining items were not performed due to skipped visits, patient death, or loss to follow-up. In addition, 533 AEs and 42 SAEs were reported, bringing the total number of eCRFs posted to 7207.

Most errors are corrected as part of data entry or prevented outright by the design of the forms. Throughout stage 1, 49 eCRFs (0.7%) were detected to contain an error at posting. Table I lists the frequency of each error. No item had more than one error. At the time of the delivery of the analysis data sets, all but four errors (0.06% of submitted forms) were resolved.

The last study visit contributing final efficacy outcome occurred on 7 November 2006, and the data were posted the same day. The efficacy data set was delivered to the study statistician on 16 November, nine days later. The last visit collecting final safety data occurred on 29 November 2006, and the data were posted on 4 December 2006. The safety data set was delivered to the study statistician on 6 December, two days later. The median interval between data collection and data posting for visit procedures was 2 days (mean 8.9, SD 104.6). The median interval between data collection and data posting for adverse events was 53 days (mean 99.4, SD 110.2).

A voluntary, anonymous on-line survey was administered to gather user opinion of the system. Although the response rate was relatively low (12 of 19 sites, or 63%), feedback was generally positive. Nine of 12 respondents (75%) agreed or agreed strongly with the statement “I would like to use this or a similar system in future clinical trials with which I am involved”.

Discussion

The web-based data management system created for QALS succeeded in supporting the implementation of the innovative, adaptive design of the trial. It allowed for data analysis to begin within days of the last subject completing the trial's first stage.

Although loss to follow-up occurred, the rate was lower than other recent ALS trials [3]. We submit that the QALS data management system contributed to the low rate of loss to follow-up, because it allowed the investigators to monitor skipped visits, late visits, telephone visits and other indicators of possible follow-up problems in real time. On the basis of the experience in stage 1, the system has been modified to alert investigators when data entry is delinquent. This allows investigators to more quickly distinguish between skipped visits and delayed data entry, enhancing the ability to detect patients truly at risk for loss to follow-up.

An issue arose in the handling of data error queries. In some cases, protocol violations were listed as data errors, when in fact the data were correct and protocol violations had occurred. Future work should be carried out to distinguish between values that are impossible (and therefore necessarily erroneous) and those that are possible but in violation of the protocol.

Web-based systems can increase confidence in clinical trials. Data encryption (similar to standard e-commerce practice) as well as the minimization of PMI transmission, can allay patient fears regarding confidentiality. Prompt availability of data, timely data entry, and a low upfront error rate also improve the ability of investigators to monitor the trial, allowing for rapid detection and correction of problems such as patient drop-out, loss to follow-up, and missing data. All these measures can increase recruitment, protocol adherence, and patient retention in a field where these are critical.

However, the system had a number of conditions for success. It was the result of a close collaboration over several months between investigators, statisticians, coordinators, programmers and evaluators. We anticipate that a significant up-front investment in time and effort will be required in similar trials. Also, the system requires a robust infrastructure, including server computers, ongoing maintenance, and technical support for clinical sites. Finally, appropriate funding was provided by the sponsor (NIH-NINDS). It is clear in retrospect that this was cost-effective, but not all sponsors or agencies are as aggressive in making this investment.

Many clinical trials merely use electronic data collection (EDC), restricting computerized data management to the collection of data for analysis [11]. The QALS system was more comprehensive. Other web-based systems have been still more elaborate. Some have enhanced security through the use of real-time code generation [4]; others have more elaborate management infrastructure [12] and still others have implemented completely paper-free data collection [13]. We did not consider these more elaborate features appropriate or cost-effective for this relatively small phase II trial, but they may be highly appropriate for larger, phase III trials.

Given the multiplicity of options, it is clear that one of the major tasks in creating such a system is to create the right mix of features for each project. When combined with careful

and rigorous design of trial procedures, data sets, and data collection instruments, web-based data management systems can contribute dramatically to the speed and reliability of both conduct of and analysis for clinical trials in ALS and related disorders.

Acknowledgments

This work was supported by Federal grants NINDS ROI NS 48555 (RB, RA, JLPT), NINDS R01 NS 48125 (PK, AB, JM) and K12RR017648 (PK), and an Irving Research Scholar Award (PK). ***Declaration of interest:*** The authors report no conflict of interest.

References

1. Sorenson EJ, Stalker AP, Kurland LT, Windebank AJ. Amyotrophic lateral sclerosis in Olmsted County, Minnesota, 1925–1998. Neurology. 2002; 59:280–2. [PubMed: 12136072]
2. Norris F, Shepherd R, Denys EUK, Mukai E, Elias L, Holden D, Norris H. Onset, natural history and outcome in idiopathic adult motor neuron disease. J Neurol Sci. 1993; 118:48–55. [PubMed: 8229050]
3. Thompson JL, Levy G. ALS issues in clinical trials. Missing data. Amyotroph Lateral Scler Other Motor Neuron Disord. 2004; (Suppl 5):48–51. [PubMed: 15512872]
4. Mitchell R, Shah M, Ahmad S, Rogers AS, Ellenberg JH. Adolescent Medicine Trials Network for HIV/AIDS Interventions. A unified web-based query and notification system (QNS) for subject management, adverse events, regulatory, and IRB components of clinical trials. Clin Trials. 2005; 2:61–71. [PubMed: 16279580]
5. Lallas CD, Preminger GM, Pearle MS, Leveillee RJ, Lingeman JE, Schwope JP, et al. Internet based multi-institutional clinical research: a convenient and secure option. J Urol. 2004; 171:1880–5. [PubMed: 15076298]
6. Wisniewski SR, Eng H, Meloro L, Gatt R, Ritz L, Stegman D, et al. Web-based communications and management of a multicenter clinical trial: the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) project. Clin Trials. 2004; 1:387–98. [PubMed: 16279277]
7. Reboussin D, Espeland MA. The science of web-based clinical trial management. Clin Trials. 2005; 2:1–2. [PubMed: 16279572]
8. Litchfield J, Freeman J, Schou H, Elsley M, Fuller R, Chubb B. Is the future for clinical trials internet-based? A cluster randomized clinical trial. Clin Trials. 2005; 2:72–9. [PubMed: 16279581]
9. Smith KS, Eubanks D, Petrik A, Stevens VJ. Using web-based screening to enhance efficiency of HMO clinical trial recruitment in women aged 40 years and older. Clin Trials. 2007; 4:102–5. [PubMed: 17327250]
10. Levy G, Kaufmann P, Buchsbaum R, Montes J, Barsdorf A, Arbing R, et al. A two-stage design for a phase II clinical trial of coenzyme Q10 in ALS. Neurology. 2006; 66:660–3. [PubMed: 16534103]
11. Pavlovic I, Miklavcic D. Web-based electronic data collection system to support electrochemotherapy clinical trial. IEEE Trans Inf Technol Biomed. 2007; 11:222–30. [PubMed: 17390992]
12. Jacobs BP, Bent S, Tice JA, Blackwell T, Cummings SR. An internet-based randomized, placebo-controlled trial of kava and valerian for anxiety and insomnia. Medicine (Baltimore). 2005; 84:197–207. [PubMed: 16010204]
13. Lopez-Carrero C, Arriaza E, Bolanos E, Ciudad A, Municio M, Ramos J, Hesen W. Internet in clinical research based on a pilot experience. Contemp Clin Trials. 2005; 26:234–43. [PubMed: 15837443]

Appendix: QALS Study Group

Columbia University Clinical Coordinating Center (CCC), New York, NY: P. Kaufmann, MD, MSc; A.I. Barsdorf, MA; J. Andrews, MD; D. Vecchio, MS; K. Bednarz, RD; G.

Harrington, RN, BSN; J. Montes MA, PT; V. Battista RN, BS; P.H. Gordon, MD; H. Mitsumoto, MD. *Columbia University Statistical Analysis Center (SAC), New York, NY:* J.L.P. Thompson, PhD; A. R. Tierney, MPH; G. Levy, MD, MS; B. Levin, PhD; R. Arbing, MSc, MPH; R. Buchsbaum. *Baystate Medical Center, Springfield, MA:* C. Armon, MD; G. Baquis, MD; E. Reich, RN. *HFA ALSA-Certified ALS Center/Berman Center for Outcomes and Clinical Research, Minneapolis, MN*: E. Tiryaki, MD, W. S. David, MD, PhD; Cathy Heuer, RN, BSN; S. Conn, RN; S. Swanson, PT. *California Pacific Medical Center, San Francisco, CA:* R. G. Miller, MD; J. Katz, MD; D. Forshew, RN, BSN; D. Scholtz, MD; S. Champion; S. McDade, PT. *Cleveland Clinic, Cleveland, OH:* E. P. Pioro, MD, PhD; K. Kelly, MSN, CNS; R. Kolb. *Drexel University-Hahnemann Campus, Philadelphia, PA:* A. Deboo, MD; S. Feldman, PT; T. Heiman-Patterson, MD. C. Barr, RN. *Massachusetts General Hospital, Boston, MA:* M. Bellanich; P. Butsch; J. Eckenrode; L. S. Krivickas, MD; D Pulley; B. Traynor, MD; M. E. Cudkowicz, MD, MSc; R. J. Walsh, MD; A. Caraganis; K. Kruczek; A.B. DiBernado, MD. *HFA ALSA-Certified ALS Center/Berman Center for Outcomes and Clinical Research, Minneapolis, MN:*, MD; W. David, MD, PhD; C. Heuer, RN, BSN; S. Conn, RN; S. Swanson, PT. *Northwestern University, Chicago, IL:* R. Sufit, MD; P. Casey, MS; S. Manes, DPT. *State University of New York Upstate Medical, Syracuse, NY:* J. Shefner, MD, PhD; J. Taft, MPAS, PA-C; M. L. Watson, BS, RRT; C. Lu; S. Strodel; M. Grosso, PT, DPT, RPA-C. *University of Arkansas, Little Rock, AR:* S. A. Rudnicki, MD; D. Fewell, LPN. *University of California at San Francisco, San Francisco, CA:* C. Lomen-Hoerth, MD, PhD; D. Forshew, RN, BSN; V. Vanderpool; C. Villierme; E. Garwood; L. Squire. *University of Chicago, Chicago, IL:* K. Rezania, MD; R. P. Roos, MD; B. Soliven, MD; E. Shaviers. *University of Colorado, Denver, CO:* J. Cumming, H. Neville, MD; Y. Rollins, MD, PhD; S. P. Ringel, MD; B. Oskarsson, MD; T. Gamage, PT; B. Hammack, PhD; B. Hewitt. *University of Kansas, Kansas City, KS:* R. J. Barohn, MD; A. R. Dick, MD; A. McVey, MD; A. Stuckey, CRT; L. Herbelin, BS; D. Gibson; M. Walsh. *University of Kentucky, Lexington, KY:* E. J. Kasarskis, MD, PhD; C. Shrestha, MPH; M. Thompson; V. Bellamy; R. Gleason, PT, MS; R.A. English, PT. K. Vanderpool, RN, BS; D.G. Taylor; S. Wells. *University of Texas Health Science Center:* C. E. Jackson, MD; E. Ensrud, MD; P. P. Kittrell, MSN, RN; D. A. Myers, PT. *University of Vermont, Burlington, VT: C. Potter* R. Tandan, MD; J Gardner; K. Moxley, BSN, RN; D. Cutura; M.C. Jones, MAMC. *Washington University, St. Louis, MO*: B. Abrams, RN: A. Pestronk, MD; J. M. Florence, DPT; J. R. Schierbecker, MHS, PT; C. Wulf, BA; B.C. Malkus, PT, MHS. *Yale University School of Medicine, New Haven, CT:* S. P. Novella, MD; J. M. Goldstein, MD; L. Marshall; A. Toenjes, RN, MSN, APRN.

Table I

Frequency of error types in submitted data.

Error	n
Date of last dose prior to SAE is after date of onset.	16 (32.7%)
Date of onset for adverse event is before patient was randomized.	12 (24.5%)
End date for concomitant medication is before medication start date.	4 (8.2%)
Date of onset for adverse event is after date of resolution.	4 (8.2%)
Assessment of Eligibility indicates that FVC is > = 60% of predicted (Q3), yet the Forced Vital Capacity form (Q1–4) contradicts this.	4 (8.2%)
Date of data collection is after the date that the eCRF data was posted.	3 (6.1%)
Assessment of Eligibility indicates that disease duration is < 5 years (Q5), yet the neuromuscular exam form (date of onset) contradicts this.	3 (6.1%)
Start date for condition in medical history is after date of end date.	2 (4.1%)
Start date for condition in medical history is after patient was randomized.	1 (2.0%)