

# CLARA: an integrated clinical research administration system

Jiang Bian,<sup>1,2</sup> Mengjun Xie,<sup>3</sup> William Hogan,<sup>1</sup> Laura Hutchins,<sup>4</sup> Umit Topaloglu,<sup>1,2</sup> Cheryl Lane,<sup>2</sup> Jennifer Holland,<sup>5</sup> Thomas Wells<sup>6</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/amiajnl-2013-002616>).

<sup>1</sup>Division of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

<sup>2</sup>Research Systems, Information Technology, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

<sup>3</sup>Department of Computer Science, University of Arkansas at Little Rock, Little Rock, Arkansas, USA

<sup>4</sup>Department of Hematology and Oncology, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

<sup>5</sup>Office of Institutional Review Board, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

<sup>6</sup>Research Support Center, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

## Correspondence to

Dr Jiang Bian, Division of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, 4301 W Markham St #633-1, Little Rock, AR 72205I, USA; [jbian@uams.edu](mailto:jbian@uams.edu)

Received 27 December 2013

Revised 5 April 2014

Accepted 8 April 2014

Published Online First

28 April 2014



CrossMark

**To cite:** Bian J, Xie M, Hogan W, et al. *J Am Med Inform Assoc* 2014;**21**:e369–e373.

## ABSTRACT

Administration of human subject research is complex, involving not only the institutional review board but also many other regulatory and compliance entities within a research enterprise. Its efficiency has a direct and substantial impact on the conduct and management of clinical research. In this paper, we report on the Clinical Research Administration (CLARA) platform developed at the University of Arkansas for Medical Sciences. CLARA is a comprehensive web-based system that can streamline research administrative tasks such as submissions, reviews, and approval processes for both investigators and different review committees on a single integrated platform. CLARA not only helps investigators to meet regulatory requirements but also provides tools for managing other clinical research activities including budgeting, contracting, and participant schedule planning.

## INTRODUCTION

Research administration is crucial to an organization's research infrastructure. Its efficiency problem has been ever growing in biomedical research institutions, where management of research activities (especially those involving human subjects) is surprisingly complex. One of the first-order jobs of biomedical research administration is to ensure research regulatory compliances, which is primarily undertaken by the institutional review board (IRB).<sup>1–2</sup> Traditional paper-based IRB review process is cumbersome and error-prone, which can significantly delay research progress. An electronic IRB administration system can relieve administrative burdens and improve the administration process by providing a paperless method for submission, tracking, and review of information. However, over its entire lifecycle, a clinical research project often involves a variety of other regulatory and administrative review entities besides the IRB, which are often not covered by existing off-the-shelf or in-house-developed electronic IRB systems.<sup>3–7</sup> More importantly, research administration in reality handles both assurance of regulatory compliances and other aspects of clinical research such as budgeting, contracting, and post-approval regulatory monitoring, which are often missing in existing research administration systems or scattered around many different information and workflow systems.

We developed the Clinical Research Administration (CLARA) platform, an open-source, web-based, scalable, and standards-compliant research administration system that covers the full spectrum of managing clinical research activities at

the University of Arkansas for Medical Sciences (UAMS). All resources, including the source code and documentation, can be found at <https://github.com/bianjiang/clara>.

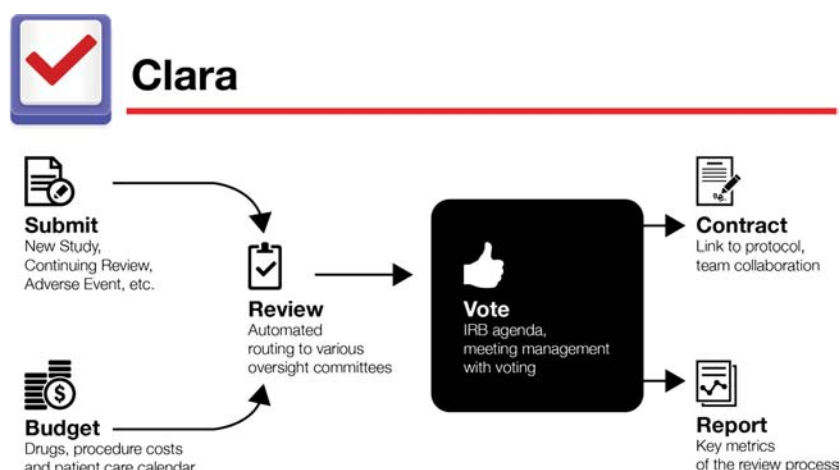
## BACKGROUND

The development of CLARA is supported by the Translational Research Institute (TRI, UAMS) funded by UAMS's Clinical and Translational Science Award (CTSA).<sup>8</sup> One of TRI's goals is to improve research infrastructure at UAMS so that investigators can conduct human-based research more smoothly and efficiently. Before CLARA, UAMS developed the Automated Research Information Administrator (ARIA) system to support IRB administration in 2000, and the Clinical Research Information Management System (CRIMSON) in 2007 to extend research administration beyond the IRB. Although ARIA and CRIMSON systems improved the efficiency of IRB and other review bodies, they suffer from a number of technological and design deficiencies, such as (1) scalability issues of back-end databases, (2) data inconsistency and quality issues, (3) slow system performance and bad user experience, and (4) lack of support for data extraction and reporting. These deficiencies in ARIA and CRIMSON, which—along with similar problems—also exist in other commercial and open-source software,<sup>3–7</sup> drove the critical design decisions for CLARA.

## METHODS AND TECHNOLOGIES

On the one hand, CLARA is expected to assist research administration at every step of the regulatory compliance process, such as review, approval, and tracking of human subject research protocols. On the other hand, CLARA is also designed to help investigators manage their research logistics, such as planning and building a visit-by-visit participant clinical schedule and generating a detailed clinical trial budget. **Figure 1** shows a high-level overview of the CLARA system. The objectives for CLARA are achieved through the following key features and components: (1) a standards-compliant user authentication and role-based access control model; (2) an integrated platform that supports collaboration and communications across regulatory and administrative bodies; (3) a flexible reporting module to support a wide variety of data extraction requirements; (4) auditing capabilities; (5) an extensible interface engine for connecting to other clinical and research systems; (6) an extensible version and change control system; (7) a study calendar-like budgeting tool; (8) IRB agenda management and live on-line IRB meeting support; and

**Figure 1** A system overview of the Clinical Research Administration (CLARA) platform.



(9) a set of tools and metrics for benchmarking clinical research administration workflows. CLARA, built with open software architecture and modular designs, has used a number of the latest software development technologies to realize the above objectives.

### Flexible data model and system architecture for extensibility and scalability

To achieve flexibility, scalability, and platform independence, CLARA adopts a document-oriented data model similar to previous models<sup>9–11</sup> in which each research protocol is a collection of forms in XML format. The semistructured XML data model is platform-independent and has great flexibility, which can accommodate changes to the data-modeling needs (eg, adding or removing a question from a submission form template) without disrupting existing applications. On the other hand, data elements stored in XMLs are still structured and therefore can be easily queried and analyzed. However, CLARA uses a production-level relational database (ie, Microsoft SQL Server<sup>12</sup>) for actual data storage and management to leverage the maturity of, and past investments on, relational database technologies. By storing and managing XML data in SQL Server, CLARA can enjoy many benefits offered by a mature relational database management system, such as abundant knowledge and expertise in data management, readiness and reliability of data backup, recovery, and replication infrastructure. In addition, Microsoft SQL Server, since its 2005 version, provides extensive support<sup>13, 14</sup> for XML data types such as built-in methods for querying and modifying XML instances as

well as a B-tree-based XML indexing framework, which helps CLARA to deliver high system performance.

CLARA is built as a web-based enterprise application using a number of modern and mature web development technologies including the Java-based Spring Framework,<sup>15</sup> Java Persistent API (JPA), with Hibernate database abstraction<sup>16</sup> and Apache Tomcat<sup>17</sup> web container. To ensure performance and scalability, load balancing was given special attention in the system design. In addition, to deploy proper database indexes, we use Nginx<sup>18</sup>—a high-performance reverse proxy implementation—to distribute workloads across different servers.

### Responsive and intuitive web-based user interface (UI)

CLARA adopted responsive web design approaches to provide users with a smooth and consistent browsing experience—easy reading and navigation throughout the system—across different end devices from PCs to mobile devices. CLARA uses JavaScript and cascading style sheets extensively, with UI design frameworks such as Sencha Ext JS,<sup>19</sup> jQuery,<sup>20</sup> and Twitter's Bootstrap<sup>21</sup> to ensure UI consistency across devices (eg, web forms in CLARA can be readily printed without any disruption in readability).

To make the data entry process easier and more user-friendly, CLARA follows best practices in UI design including: (1) using wizards to guide investigators through a submission process step-by-step; (2) offering timely in-place help information; (3) embedding branching logics among questions and sections so that only necessary questions pertaining to the context are presented; (4) performing real-time client-side data validation and integrity checks; and (5) applying clear categorization and easy

**Figure 2** A wizard-based user interface design for a typical data entry form in Clinical Research Administration (CLARA).

**Table 1** Summary of research studies processed in CLARA from April 22, 2013 to April 1, 2014

Form type	Submitted forms (n)	Approved/completed forms (n)	Days from submission to completion (average)
Initial modification (migration)	1008	987	2.0
New submission	373	268	33.0
Modification/amendment	1030	904	12.0
Continuing review	848	660	11.0
Human subject research determination	264	262	2.0
Reportable new information	48	41	8.0
Study closure	157	151	2.0
Contract	279	210	41.0

CLARA, Clinical Research Administration.

navigation between questions and sections (figure 2 shows a typical data entry form in CLARA).

### Flexible and expressive workflow engine

Clinical research administration workflows are inherently complex and highly dynamic with many different scenarios. In CLARA, we have developed a flexible rule-based workflow engine. Workflow rules can be described easily and expressively in a semistructured XML format (online supplementary figure S1 illustrates a sample review workflow and online supplementary figure S2 shows a partial XML snippet defining the workflow rules). Moreover, we adopted an event-driven service-oriented architecture in the workflow engine where event messages (eg, an approval action) are broadcast and event brokers can react accordingly (eg, an email service handler can send notifications to related bodies and an audit handler can create an audit trail). Such an event-driven system design minimizes the effort for accommodating new requirements and eases the maintenance of existing code structure when adding new functionalities.

**Figure 3** The clinical study calendar planning and budgeting tool in Clinical Research Administration (CLARA). In CLARA, a study budget is structured so that it can be turned into a detailed participant schedule and used in a clinical trial management system. The schedule of a clinical study in CLARA can normally be divided into multiple phases (eg, screening, treatment, and follow-ups), each phase can be split into multiple arms (eg, arm A as the drug group, and arm B for the placebo group), and each arm can be further divided into cycles and visits. Study activities such as office visits and laboratory procedures are attached to each study visit, coded with the current procedural terminology (CPT) codes, and priced accurately according to University of Arkansas for Medical Sciences (UAMS's) research hospital and professional fee schedules.

## RESULTS

We piloted CLARA with the Cancer Institute at UAMS in February 2013, and CLARA replaced ARIA and CRIMSON on April 22, 2013. Table 1 summarizes the research studies that have been processed through CLARA since then. An automated data-migration process has been established to transfer ARIA and CRIMSON study records into CLARA. However, the data elements in CLARA are vastly different from those in ARIA and CRIMSON. Therefore, an initial modification form is required for all open studies migrated from the old systems, which allows the investigators to fill in missing information and fix errors in previous submissions. A total of 1083 open studies have been migrated into CLARA; as of April 1, 2014, 987 (91.14%) of them have had the data re-entry process completed. Although we have not conducted a formal survey, general feedback from investigators and administrative bodies is very positive—for example, ‘the new system provides an integrated platform to communicate with investigators and other review entities’.

One unique and highly regarded feature of CLARA is its study planning and budgeting tool. It not only enables investigators to plan and estimate various costs of a clinical research study in a detailed and patient-oriented manner, it also empowers the hospital administration to conduct in-depth analyses for billing reconciliation prior to committing valuable resources to a research study. Figure 3 demonstrates a sample budget of a fairly complex randomized clinical trial with multiple phases and arms.

Another key feature of CLARA is that it provides a flexible and comprehensive reporting system to evaluate various aspects of the clinical research workflow at UAMS (online supplementary figure S3 presents a typical workflow benchmark report). These reports not only provide insights into how well a clinical research workflow is performing overall, but also help the administration to identify specific workflow deficiencies with quantitative evidence.

## DISCUSSION AND CONCLUSIONS

A number of factors contribute to the success of CLARA in UAMS's clinical research administration enterprise. CLARA is an integrated platform that covers the full lifecycle of research

Procedure	Cost	Amount	Price	Residual	With in 28 Days: 1 Days	With in 14 Days: 1 Days	With in 7 Days: 1 Days	Consent/Pool: 1 Days	Consent: 1 Days	Total (by charge type)	RHS	CHMS	CL
03010 ELECTROCARDIOGRAM REPORT	\$30.00	\$30.00	\$30.00	\$0.00	C	-	-	-	-	\$30.00			
07120 CT Scan-Body Scan	\$500.01	\$500.01	\$500.01	\$0.00	C	-	-	-	-	\$500.01			
07177 CT Scan-General	\$670.71	\$670.71	\$670.71	\$0.00	C	-	-	-	-	\$670.71			
03307 TTE W/DOPPLER COMPLETE (H: \$415.01)	\$533.01	\$533.01	\$533.01	\$0.00	C	-	-	-	-	\$533.01			
09054 OFFICE/OUTPATIENT VISIT NEW (H: \$134.90)	\$396.90	\$396.90	\$396.90	\$0.00	C	-	-	-	-	\$396.90			
03009 ELECTROCARDIOGRAM TRACING (H: \$33.00)	\$63.00	\$63.00	\$63.00	\$0.00	C	-	-	-	-	\$63.00			
06415 ROUTINE VESICUPUNCTURE (H: \$3.15)	\$3.15	\$3.15	\$3.15	\$0.00	-	C	-	-	-	\$3.15	\$3.15		
04100 ASSAY OF PHOSPHORUS (H: \$6.85)	\$6.85	\$6.85	\$6.85	\$0.00	-	C	-	-	-	\$6.85			
03730 ASSAY OF MAGNESIUM (H: \$9.67)	\$9.67	\$9.67	\$9.67	\$0.00	-	C	-	-	-	\$9.67			
08034 BRUNOASSAY TUNOR CA 125 (H: \$30.04)	\$30.04	\$30.04	\$30.04	\$0.00	-	C	-	-	-	\$30.04			
05020 COMPLETE CBC W/AUTO DIFF WBC (H: \$11.22)	\$11.22	\$11.22	\$11.22	\$0.00	-	C	-	-	-	\$11.22			
04753 CHRONIC GONADOTROPIN ASSAY (H: \$10.85)	\$10.85	\$10.85	\$10.85	\$0.00	-	-	C	-	-	\$10.85	\$10.85		
02980 ASSAY OF CREATININE (H: \$7.39)	\$7.39	\$7.39	\$7.39	\$0.00	-	C	-	-	-	\$7.39			
04520 ASSAY OF UREA NITROGEN (H: \$5.70)	\$5.70	\$5.70	\$5.70	\$0.00	-	C	-	-	-	\$5.70			
02247 BILIRUBIN TOTAL (H: \$7.24)	\$7.24	\$7.24	\$7.24	\$0.00	-	C	-	-	-	\$7.24			
04450 TRANSFERASE (AST) (SODT) (H: \$7.47)	\$7.47	\$7.47	\$7.47	\$0.00	-	C	-	-	-	\$7.47			
04480 ALANINE AMINO (ALT) (SODT) (H: \$7.63)	\$7.63	\$7.63	\$7.63	\$0.00	-	C	-	-	-	\$7.63			
04070 ASSAY ALKALINE PHOSPHATASE (H: \$7.47)	\$7.47	\$7.47	\$7.47	\$0.00	-	C	-	-	-	\$7.47			
02310 ASSAY OF CALCIUM (H: \$7.44)	\$7.44	\$7.44	\$7.44	\$0.00	-	C	-	-	-	\$7.44			



study administration, from design and planning of a research study to reporting and post-approval monitoring. Moreover, CLARA promotes and facilitates collaboration and communication among different review bodies as well as interaction between investigators and research administration entities. Another key element to CLARA's success is its ability to quickly adapt to changes in policy, process, and requirement without disrupting existing system services. We achieved such flexibility by using a non-conventional document-based data model (ie, semistructured XML). The reason for using XML instead of a relational database is threefold. (1) Each submission form can be naturally modeled as a document, and storing all information about a submission form in one place makes retrieving easy. In contrast, retrieving all relevant information about a form from a relational database often requires writing complicated SQL queries and an application layer code. (2) A version controlling a submission record, especially when data elements are spread across many tables, is difficult to implement in a relational database. (3) From our experience in developing ARIA and CRIMSON, system requirements change frequently. For example, since CLARA went live, we have had 311 change requests, of which 35 require modifications to the data structure and 108 are related to workflow changes. These changes have been rapidly implemented in CLARA because of the use of an XML data model. In contrast, structural changes are difficult to execute in a relational database without disrupting existing data and services.

The development of CLARA also provides an opportunity for informatics consultation concerning system security and compliance issues, data collection and key reporting requirements, and data validation methods and procedures. Moreover, the process of developing integrated research administration tools prompted a number of institutional regulatory and administrative entities to reconcile their current processes and practices for efficiency improvement. Through the development of CLARA, the IRB administration office at UAMS has eliminated a number of unnecessary workflows and reduced revision cycles between investigators and the administration. Moreover, the UAMS administration (especially the TRI) has formed a work group to leverage the detailed and comprehensive process metrics generated by CLARA and use Lean Six Sigma concepts<sup>22</sup> to find clinical research workflow deficiencies and identify potential improvement opportunities.

Last but not least, its interoperability makes CLARA powerful for data integration and attractive to its users. For example, CLARA interfaces with UAMS's hospital and professional billing systems to provide investigators with an accurate billing plan. The billing schedules created in CLARA can be readily transformed and pushed into the open-source Comprehensive Research Informatics Suite (CRIS, UAMS)<sup>23</sup> as study calendar templates for patient tracking and study activity management. Furthermore, CLARA also interfaces with electronic medical record (EMR) and hospital information systems such as Epic<sup>24</sup> to create research study definitions for tracking and billing research subjects.

Compared with other clinical research administration tools, CLARA is unique in three main aspects. First, unlike existing IRB systems such as IRBNet<sup>25</sup> and Click IRB<sup>26</sup> whose sole purpose is IRB administration, CLARA enables regulatory and administrative entities to communicate and collaborate on the same platform, which increases review efficiency. Second, unlike SPARC<sup>27</sup> and WebCAMP<sup>28</sup> which are mainly used to request, review, and track research core services, CLARA's budgeting tool allows investigators to create a template of a detailed participant

study schedule that can later be integrated into downstream clinical trial management and EMR systems. CLARA is close to Quali Coeus in terms of its goals and functions.<sup>29</sup> However, Quali Coeus covers more aspects of clinical research, such as conflict of interest, while CLARA is specialized in clinical trials with a patient-orientated budgeting and study-planning tool, which is missing in Quali Coeus. Third, the open-source nature and its modular design make CLARA highly customizable and easy to deploy.

**Acknowledgements** The work describe in this manuscript is supported by award UL1TR000039 through National Center for Advancing Translational Sciences (ie, formerly UL1RR029884 through the NIH National Center for Research Resources). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

**Collaborators** Yu Fan, Matt Baker, Yuan Jiawei.

**Contributors** JB designed the overall system architecture and led the development effort. JB and MX wrote the manuscript. UT and WH contributed to the design and implementation of the system, and helped with system integration. JH, LH, CL and TW contributed to implementation of critical system components and guided the developers through the clinical research workflow. All authors reviewed, edited, and approved the submitted version of this article.

**Funding** The work was supported by award UL1TR000039 through National Center for Advancing Translational Sciences (ie, formerly UL1RR029884 through the NIH National Center for Research Resources).

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- Green LA, Lowery JC, Kowalski CP, *et al.* Impact of institutional review board practice variation on observational health services research. *Health Serv Res* 2006;41:214–30.
- Vick CC, Finan KR, Kiefe C, *et al.* Variation in Institutional Review processes for a multisite observational study. *Am J Surg* 2005;190:805–9.
- Huron Education. Click IRB 7.0: A Complete IRB Solution Incorporating Huron's Industry-Leading HRPP Toolkit. [http://www.clickcommerce.com/cc/resource/doc/DOC00004984/ClickIRB\\_2012\\_FINAL.pdf](http://www.clickcommerce.com/cc/resource/doc/DOC00004984/ClickIRB_2012_FINAL.pdf) (accessed 30 Sep 2013).
- Kuali Foundation. Kuali Coeus. <http://www.kuali.org/kc> (accessed 30 Sep 2013).
- ResearchDataware, LLC. IRBNet-InnovativeSolutions. [Online]. <https://www.irbnet.org/release/index.html> (accessed 30 Sep 2013).
- Duke University Health System. Duke Medicine eIRB. [Online]. <https://eirb.mc.duke.edu/eirb/> (accessed 30 Sep 2013).
- Northwestern University. eIRB. [Online]. <https://eirb.mc.duke.edu/eirb/> (accessed 30 Sep 2013).
- University of Arkansas for Medical Sciences. Translational research institute. [Online]. <http://tri.uams.edu/> (accessed 30 Sep 2013).
- MongoDB, Inc. Document Databases. [Online]. <http://www.mongodb.com/document-databases> (accessed 30 Sep 2013).
- MongoDB, Inc. MongoDB. [Online]. <http://www.mongodb.org/> (accessed 30 Sep 2013).
- The Apache Software Foundation. Apache CouchDB. [Online]. <http://couchdb.apache.org/> (accessed 30 Sep 2013).
- Microsoft. SQL Server. [Online]. <https://www.microsoft.com/en-us/sqlserver/product-info.aspx> (accessed 27 Jan 2013).
- Microsoft. XML Support in Microsoft SQL Server 2005. [Online]. [http://technet.microsoft.com/en-us/library/ms345117\(v=sql.90\).aspx](http://technet.microsoft.com/en-us/library/ms345117(v=sql.90).aspx) (accessed 1 Apr 2014).
- Microsoft. Implementing XML in SQL Server. [Online]. [http://technet.microsoft.com/en-us/library/ms189887\(v=sql.105\).aspx](http://technet.microsoft.com/en-us/library/ms189887(v=sql.105).aspx) (accessed 1 Apr 2014).
- GoPivotal, Inc. Spring Framework. [Online]. <http://projects.spring.io/spring-framework/> (accessed 30 Sep 2013).
- The JBoss Community Team. Relational Persistence for Java and .NET. [Online]. <http://www.hibernate.org/> (accessed 30 Sep 2013).
- The Apache Software Foundation. Apache Tomcat. [Online]. <http://tomcat.apache.org> (accessed 30 Sep 2013).
- The Nginx Community. Nginx. [Online]. <http://wiki.nginx.org/Main> (accessed 30 Sep 2013).
- Perry Jason. O'Reilly Media, S.I. 2013. ISBN 1449361978.
- McFarland David. JavaScript&jQuery. O'Reilly, Sebastopol, Calif. 2011. ISBN 1449399029.
- Twitter Inc. Bootstrap. [Online]. <http://getbootstrap.com/> (accessed 30 Sep 2013).
- George M. *Lean six sigma: combining six sigma quality with Lean production speed, combining six sigma quality with Lean production speed.* McGraw Hill Professional, 2002.

- 23 University of Arkansas for Medical Sciences. Comprehensive Research Informatics Suite (CRIS). [Online]. <http://tri.uams.edu/research-resources-services-directory/biomedical-informatics-data-services/clinical-trials-management/> (accessed 30 Sep 2013).
- 24 Epic Systems Corporation. Epic. [Online]. <http://www.epic.com/> (accessed 30 Sep 2013).
- 25 Research Dataware, LLC. IRBNet. [Online]. <https://www.irbnet.org/release/index.html> (accessed 27 Jan 2013).
- 26 Huron Education. Click IRB. [Online]. <http://www.clickcommerce.com/> (accessed 27 Jan 2013).
- 27 Medical University of South Carolina. SPARC Request—Services, Pricing, & Application for Research Centers. [Online]. [http://academicdepartments.musc.edu/sctr/sparc\\_request/](http://academicdepartments.musc.edu/sctr/sparc_request/) (accessed 27 Jan 2013).
- 28 Weill Cornell Medical College. WebCAMP. [Online]. <ftp://ctsc.med.cornell.edu> (accessed 27 Jan 2013).
- 29 Kuali Foundation, INC. Kuali Coeus. [Online]. <http://www.kuali.org/kc> (accessed 27 Jan 2013).