

# Current State of Information Technologies for the Clinical Research Enterprise across Academic Medical Centers

Shawn N. Murphy, M.D., Ph.D.<sup>1,2,3</sup>, Anil Dubey, M.D., M.S.<sup>1,3,4</sup>, Peter J. Embi, M.D., M.S.<sup>5</sup>, Paul A. Harris, Ph.D.<sup>6</sup>, Brent G. Richter<sup>7</sup>, Fran Turisco, M.B.A.<sup>8</sup>, Griffin M. Weber, M.D., Ph.D.<sup>9,10</sup>, James E. Tchong, M.D.<sup>11</sup>, and Diane Keogh<sup>1</sup>

## Abstract

Information technology (IT) to support clinical research has steadily grown over the past 10 years. Many new applications at the enterprise level are available to assist with the numerous tasks necessary in performing clinical research. However, it is not clear how rapidly this technology is being adopted or whether it is making an impact upon how clinical research is being performed. The Clinical Research Forum's IT Roundtable performed a survey of 17 representative academic medical centers (AMCs) to understand the adoption rate and implementation strategies within this field. The results were compared with similar surveys from 4 and 6 years ago. We found the adoption rate for four prominent areas of IT-supported clinical research had increased remarkably, specifically regulatory compliance, electronic data capture for clinical trials, data repositories for secondary use of clinical data, and infrastructure for supporting collaboration. Adoption of other areas of clinical research IT was more irregular with wider differences between AMCs. These differences appeared to be partially due to a set of openly available applications that have emerged to occupy an important place in the landscape of clinical research enterprise-level support at AMC's. *Clin Trans Sci* 2012; Volume 5: 281–284

**Keywords:** clinical research informatics, clinical Research Forum, health informatics

## Introduction

As the conduct of clinical research becomes increasingly complex and information-intensive, academic medical centers (AMCs) are investing and growing their information technology (IT) infrastructure to be specifically dedicated to clinical research. Initiatives like the Clinical and Translational Science Award (CTSA) Program have, in recent years, focused attention on and fostered innovation and development in research IT and informatics capabilities at local and national levels.<sup>1</sup>

Even as larger and increasingly complex research efforts are needed to advance science, and the availability of growing repositories of clinical data become available through the implementation of electronic health records, compliance with state and federal security and privacy regulations and other factors increase the complexities inherent in conducting research.<sup>2,3</sup>

Therefore, although significant progress has been made to advance research and the related research IT infrastructure, myriad challenges and unrealized opportunities face the research IT enterprise, particularly with regard to support for clinical and translational research. In 2011, we undertook a survey of representative AMCs on behalf of the Clinical Research Forum (CRF) IT Roundtable to explore the current state of readiness with regard to research IT and to compare current findings to those of the prior surveys conducted on this topic. Four general areas were investigated:

1. The use of IT in research compliance, such as conflicts of interest, research budgeting, and reporting to the Institutional Review Board (IRB);
2. The use of IT for electronic data capture (EDC) requirements related to clinical studies and trials of different size;
3. The use of data repositories for the repurposing of clinical care data for research; and,

4. The IT infrastructure needs and support for research collaboration and communication.

Two such studies were conducted by The CRF, an organization comprised of academic health centers, professional organizations, and industry partners, whose goal is to sustain and expand the research capabilities within academic institutions across the United States.<sup>4,5</sup> As part of the CRF, the CRF IT Roundtable focuses on assessing, promoting, and evaluating the research-related IT capabilities of AMCs across the United States. In recent years, the CRF IT Roundtable has conducted two surveys similar to this one to describe the state of the USA AMC IT infrastructure and capabilities to which we compared the present survey.

## Methods

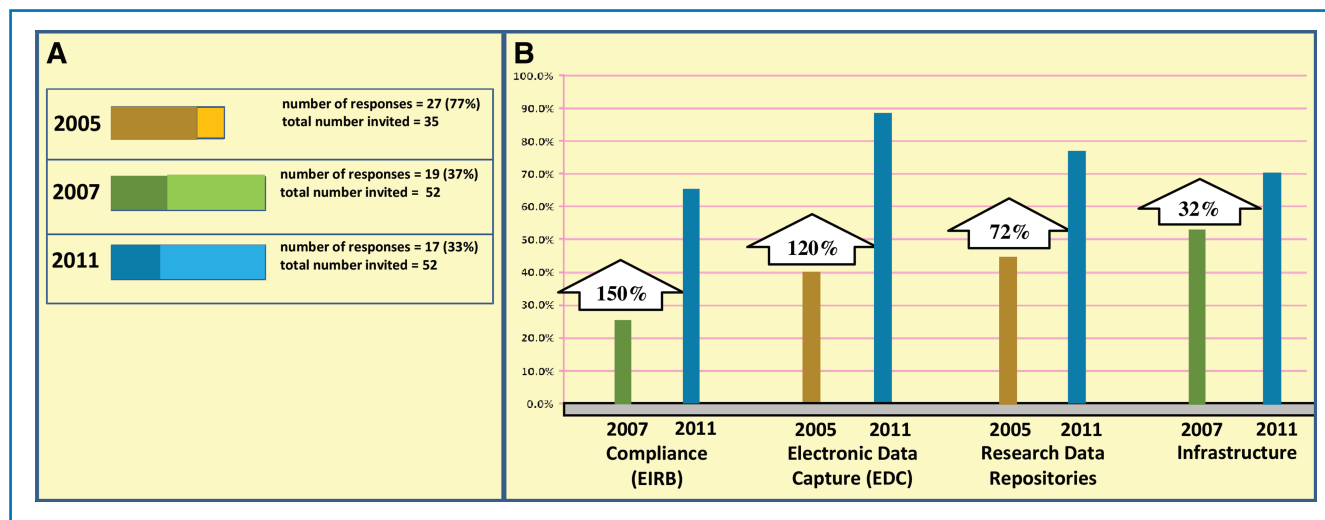
A survey committee comprised of representatives from selected member organizations, CRF administration and sponsors was formed to develop the 2011 survey. The group began by reviewing the CRF 2005 and 2007 survey instruments. The committee selected questions that were deemed by consensus to be still relevant in 2011 and would allow the Forum to measure systems implementation progress, investigate how organizations were addressing common challenges, and understand enterprise governance and support approaches for clinical research IT.

Two committee members who had worked on the prior studies prepared the initial survey. The survey was revised by the committee before pilot testing. Based on testing, minor modifications were made including giving subjects the opportunity to complete the survey in more than one sitting (due to the possibility that one person could not complete all questions at each site without consulting with others at the institution), and the ability to print the questions so the printout could be used as a data collection

<sup>1</sup>Partners HealthCare System, Inc., Research Computing, Charlestown, Massachusetts, USA; <sup>2</sup>Department of Neurology, Harvard Medical School, Boston, Massachusetts, USA; <sup>3</sup>Laboratory of Computer Science, Massachusetts General Hospital, Boston, Massachusetts, USA; <sup>4</sup>Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA; <sup>5</sup>Departments of Biomedical Informatics and Internal Medicine, and the Center for Clinical and Translational Science, The Ohio State University, Columbus, Ohio, USA; <sup>6</sup>Department of Biomedical Informatics, Vanderbilt University, Nashville, Tennessee, USA; <sup>7</sup>Enterprise Research IS and Informatics, Brigham and Women's Hospital, Massachusetts General Hospital, and Partners HealthCare System, Inc., Boston, Massachusetts, USA; <sup>8</sup>Aspen Advisors, LLC, Morrison, Colorado, USA; <sup>9</sup>Information Technology, Harvard Medical School, Boston, Massachusetts, USA; <sup>10</sup>IMBIO, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA; <sup>11</sup>Duke Translational Medicine Institute, and the Department of Medicine, Duke University, Durham, North Carolina, USA.

Correspondence: Shawn N. Murphy (snmurphy@partners.org)

DOI: 10.1111/j.1752-8062.2011.00387.x



**Figure 1.** Comparison of response rates and responses regarding adoption of major categories of research IT infrastructure between the current (2011) and previous (2005 and 2007) surveys. (A) It demonstrates the response rate difference. (B) It depicts percentage increases for each category.

worksheet. The full survey, consisting of 38 questions in six sections, can be accessed at [www.clinicalresearchforum.org](http://www.clinicalresearchforum.org)

The survey was compiled and delivered via Web using the online service, SurveyMonkey. Invitations were delivered via email to the IT Roundtable representative from each of the 51-member organizations. It was available online for 2 months and up to three reminder e-invitations were sent to all prospective subjects. Participation was voluntary.

Once the survey was closed, all response data were exported into MS Excel for descriptive analyses by selected committee members. In addition to analyzing for key findings and outcomes, the data were also descriptively compared with data from the previous 2005 and 2007 surveys conducted by the group. Finally, several organizations that were unable to complete the survey online were contacted by telephone.

## Results

Seventeen of 51 member organizations submitted complete surveys during the study period, for a 33% response rate. Respondent organizations were generally reflective of the overall CRF membership in terms of size and National Institutes of Health (NIH)-funding level; geographically, respondents were more likely to be from the East Coast and Midwest than from other USA regions.

Figure 1 compares the current results with those from the two surveys in 2005 and 2007. Figure 1(A) shows the number of responses relative to the total number of invited responses for each of the three surveys (2005, 2007, and 2011). It is not surprising that the response rate was similar for the online surveys done in 2007 and 2011 and much higher for the 2005 survey, which was conducted via one-on-one conference calls. The data needed to populate this graph came from the current survey and published results from the prior surveys.<sup>4,5</sup>

Figure 1B depicts changes over time in the percentage of respondents who have implemented elements of functionality pertaining to the general areas of research compliance (compliance), EDC, clinical data repositories (research repositories), and general clinical research computing infrastructure (infrastructure). To make such comparisons across time (and across surveys), an

attempt was made to match a measurement from the current survey corresponding to an element of functionality from each of the areas above to the most similar respective measurement in the prior studies of 2005 and 2007.<sup>4,5</sup> When a measurement in the current survey matched a corresponding measurement in both the 2005 and 2007 surveys, only those from the 2007 study were used.

For research compliance, electronic IRB submission and processing was a common element of functionality that was measured among all the studies (2005, 2007, and 2011).<sup>4,5</sup> For the purpose of comparison, Figure 1B compares the current (2011) results with the corresponding results of the 2007 study only.

For EDC, the subcategory of “EDC for investigator initiated studies” measured in the current survey was the broadest and most inclusive definition, and thus was the best comparator for the 2005 measure “EDC applications for clinical trials.”<sup>4</sup>

For clinical data repositories (research repositories), the subcategories of “receiving clinical care data” and “store and archive data,” both of which had the same results with regards to fraction of respondents with completed installations, were matched with the measurement of the fraction of respondents with completed installations of a “patient data warehouse” application from the 2005 study.

Figure 2 presents the percentage of respondents in the current survey who have completed the implementation of various elements of functionality contained within each of the general areas mentioned earlier, along with the names of the open-source solutions and the most commonly used solutions (commercial or open source) mentioned by respondents who had completed such implementations. Note that not all respondents completed all the questions in each of the sections of the survey, but for any question the number of responses was never less than 16.

## Discussion

Substantial progress has been made by AMCs to accelerate the adoption and usage of IT infrastructure in support of their research enterprises over the past several years, with most respondents now reporting implementation of systems that at least have the potential to support clinical research activities. Among those, systems that relate to administrative, financial,

	% Respondents with completed installations	Open source solution(s)	Most frequent solution used	
<b>Research Compliance</b>	Conflict of Interest	76.5%	CaBIG	Click Commerce
	Clinical Research Budgeting	35.3%	None	Click Commerce
	Electronic IRB	64.7%	Atlassian JIRA, caBIG	Click Commerce
<b>Electronic Data Capture (EDC)</b>	Investigator-initiated studies	87.5%	RED Cap	RED Cap
	Multi-center studies EDC - less than 1000 participants	81.3%	RED Cap	RED Cap
	Large Studies EDC - greater than 1000 participants	68.8%	RED Cap	Inform/Phase Forward
<b>Research Repositories</b>	Identify patients for clinical trials	70.6%	i2B2	i2B2
	Aggregate clinical research data	62.5%	i2B2	i2B2, Microsoft Amalga
	Advanced searching and analytics	50.0%	i2B2	i2B2, SAS, SPSS, R
<b>Infrastructure</b>	Content collaboration (e.g. Wikis, Content Management solutions)	76.5%	Confluence	Sharepoint
	Communication/social networking (e.g. Listserv, discussion boards)	88.2%	Listserve	Listserve

**Figure 2.** Percentage of respondents with completed installations, open-source solutions cited by respondents, and most commonly cited solution cited by respondents (commercial or open source) for key elements of functionality in the categories of research compliance, electronic data capture, research repositories, and infrastructure.

and regulatory compliance appear to be the main areas of focus among the organizations surveyed, with 88% having such systems implemented or in progress.

It is not uncommon for the implementation of such systems to take years to complete and sometimes longer for institutional units/departments to fully retire so-called “shadow systems” that are often paper-based systems. Among the reasons for the often slow adoption of such systems is the complexity inherent in completely changing the workflow of a busy central IRB office, while also getting buy-in from leadership to enforce the use of an online application for all human subject protocols by all investigators and their administrative staff. Helping institutions in making progress is the fact that the vendor market has matured over the past few years, as evidenced by 69% of the installed base using a commercial product.

New research IT systems have also appeared in the landscape over the past few years, apparently in response to regulatory and other environmental requirements and needs. For instance, the recent focus on academic Conflict of Interest management by AMCs might be one motivator behind a new category of systems, those regarding regulatory portfolio management. Enterprise adoption of systems tracking Conflict of Interest in the 2011 survey shows 77% of responders have fully implemented a solution on an enterprise basis. This is one category where the vendor market has not made significant inroads, with 65% of 2011 survey respondents indicating use of homegrown solutions. Additional areas that continue to move forward are tracking required training for investigators and systems to support the clinical trials billing process, such as online clinical research budgeting and sponsors’ billing. Overall, it appears that issues with a regulatory focus or

that could pose potential compliance issues are getting substantial attention by many organizations. The discussions at the Forum elaborated that much of what is driving this appears to be fear of legal penalties for noncompliance.

The use of clinical data repositories to support clinical research also appears to have arisen as a critical area of emphasis in recent years, with substantial growth in such repositories noted since 2007. This is not surprising given that such resources can be important at many stages of the research lifecycle. Such activities begin with tasks preparatory to research such as feasibility assessment and hypothesis generation, including cohort identification for participant recruitment, and often conclude with the delivery of the phenotypic and/or genotypic information required for a wide range of clinical research studies. Although the value proposition for such systems is high,<sup>6</sup> it is well recognized that the need to customize the solution for particular Electronic Medical Record (EMR) implementations make implementation a highly customized and resource-intensive process. Indeed, homegrown solutions were noted in about 50% of the solutions. The use of biorepositories for research<sup>7</sup> was also explored in the survey. Seven of the 17 sites had installed a biorepository, but no consensus existed on software for managing the biorepository.

One area with the largest growth between 2005 and 2011 is seen in the area of enterprise-wide EDC solutions, which grew from 40% to 82%. Many institutions reported a tiered approach to supporting EDC at the enterprise level, employing a suite of products to cover investigator-initiated single-site studies, and small (<1,000 patients) and large (>1,000 patients) multisite studies. From other studies, we know that factors influencing the choice of EDC products for individual studies include overall

study budget for EDC, time for study setup, end-user training materials and ease of use, data export capacity, regulatory compliance requirements (FISMA, 21 CFR 11), and site/user management capacity for multicenter study support.<sup>8</sup> It was clear in the Forum discussions that the low maintenance cost of the REDCap EDC solution made it particularly appealing.

One of the most significant findings of the current survey is the adoption of open-source or license-free applications, with the greatest amount of adoption seen in the infrastructure and collaboration areas. Although most organizations in 2011 continue prior trends toward adopting “best-in-class” applications from different sources, the current survey reveals a change from the past in that “best-in-class” applications increasingly were from the “open” class of software. Prevailing “open” systems tend to be those led by large, often publicly funded, consortiums and they carry free licensing. Likely contributing to the success of such efforts is the fact that they are often community-led by recognized domain experts. Examples of systems that led among respondents in the various domains of research activity include: in EDC/Surveys, REDCap<sup>9</sup>; in clinical data repositories, i2b2<sup>10</sup>; and in collaboration tools, Confluence (Atlassian, Sydney, Australia).<sup>11</sup> Research networking websites, which create investigator profiles using both public data, such as PubMed papers, and internal administrative systems, such as faculty activity reports, are gaining popularity among research institutions as mechanisms to facilitate the formation of new collaborations.<sup>12</sup> Discussions at the Forum suggested it is likely that the increased speed of adoption is due to several factors, including: (1) the low cost of initial entry; (2) new institutional funding of centralized services, often through the NIHCTSA; and (3) software created by domain experts fluent in the security and compliance requirements.

Another finding of interest relates to trends in non-research-specific IT services such as email, storage, technical computing, networking, and data sharing. Such services continue to be of great importance to researchers, and our survey indicates that they continue a trend toward increased centralization among many but not all organizations. Drivers for such centralization likely include not just a greater demand for reliable services, but also escalating regulatory requirements that often serve to encourage organizations to address security and compliance due to the wider enforcement of government security standards across state and federal agencies. One such example regulation is The Health Insurance Portability and Accountability Act (HIPAA) Security Rule of 2004<sup>13</sup> and its recent modifications<sup>14</sup> that have heightened its privacy components. Another example is the Federal Information Security Management Act (FISMA)<sup>15</sup> which has led AMCs to have to put in place policies and oversight to ensure that the research systems and processes of federal agencies and/or sponsors are in compliance with federal regulations. Indeed, the burden of liability is leading many AMCs to develop centralized and more cohesive support models to enable clinical research in a compliant manner.

The limitations of the survey include the overall response rate with only 17 respondents. The complexity of the survey almost certainly contributed to the poor response rate, although we did succeed in getting a representative sample. Surveys in themselves have the inherent limitation of not actually directly studying the items in question, but for this spatially diverse investigation a direct study was impractical. Finally, we did not measure the direct benefit of these implementations, although prior publications have addressed this issue.<sup>6</sup> Forum discussions did, however, offer

opinions on the highest areas of return on investment. These included secondary use of healthcare data to qualify patients for clinical trials, use of EDC systems to manage clinical trials, and management of biosamples for genomic research.<sup>7</sup>

## Conclusion

The past 7 years have seen a substantial increase in the amount and type of research information systems' adoption by AMCs. The availability of more robust and available vendor-based and “open-source” solutions, coupled with new research initiatives (e.g., CTSA) and regulatory requirements, appear to be contributing to these advances. Chief information officers and clinical research leaders at academic institutions should find these results of interest in terms of informing their understanding of the current state-of-the-art in clinical research IT infrastructure. Given the pace over the past several years, progress in this domain can be expected to continue to increase. This progress can ultimately be expected to improve the pace and quality of clinical and translational science, although additional efforts to integrate resources across the clinical and research enterprises will ultimately be needed to further facilitate research. Future studies will confirm whether or not these infrastructure developments have the intended effect.

## Acknowledgments

Preliminary data from this survey were presented in oral form at the annual meeting of the CRF IT Roundtable in April 2011 in Washington DC. We thank E. Sasha Paegle, Microsoft for supporting this research.

## References

- Zerhouni EA, Alving B. Clinical and translational science awards: a framework for a national research agenda. *Transl Res*. 2006; 148(1): 4–5.
- Ness RB. Influence of the HIPAA Privacy Rule on health research. *JAMA*. 2007; 298(18): 2164–2170.
- Embi PJ, Payne PR. Clinical research informatics: challenges, opportunities and definition for an emerging domain. *J Am Med Inform Assoc*. 2009; 16(3): 316–327.
- Baird J, Turisco F. Service-oriented architecture: a critical technology component to support consumer-driven health care. *AHIP Cover*. 2005; 46(3): 74, 76, 78 passim.
- DiLaura R, Turisco F, McGrew C, Reel S, Glaser J, Crowley WF, Jr. Use of informatics and information technologies in the clinical research enterprise within US academic medical centers: progress and challenges from 2005 to 2007. *J Investig Med*. 2008; 56(5): 770–779.
- Nalichowski R, Keogh D, Chueh HC, Murphy SN. Calculating the benefits of a Research Patient Data Repository. *AMIA Annu Symp Proc*. 2006; 1044.
- Murphy S, Churchill S, Bry L, Chueh H, Weiss S, Lazarus R, Zeng, Q, Dubey A, Gainer V, Mendis M, et al. Instrumenting the health care enterprise for discovery research in the genomic era. *Genome Res*. 2009; 19(9): 1675–1681.
- Franklin JD, Guidry A, Brinkley JF. A partnership approach for electronic data capture in small-scale clinical trials. *J Biomed Inform*. 2011; 44(Suppl 1): S103–S108.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009; 42(2): 377–381.
- Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, Kohane I. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc*. 2010; 17(2): 124–130.
- Harris ST, Zeng X. Using wiki in an online record documentation systems course. *Perspect Health Inf Manag*. 2008; 5: 1.
- Gewin V. Collaboration: Social networking seeks critical mass. *Nature*. 2010; 468: 993–994.
- HHS. HHS Standards for Privacy of Individually Identifiable Health Information; Final Rule: 45 CFR Parts 160 and 164 2002.
- HHS. Health Information Technology for Economic and Clinical Health (HITECH) Act, Title XIII of Division A and Title IV of Division B of the American Recovery and Reinvestment Act of 2009 (ARRA) (Pub. L. 111–5). 2009.
- United States Office of Management and Budget (OMB), Federal Information Security Management Act of 2002 (“FISMA”, 44 U.S.C. § 3541, et seq.) is a United States federal law enacted in 2002 as Title III of the E-Government Act of 2002 (Pub.L. 107–347, 116 Stat. 2899). 2002.