HHS Public Access

Author manuscript

J Clin Transl Sci. Author manuscript; available in PMC 2019 June 05.

Published in final edited form as:

J Clin Transl Sci. 2018 October ; 2(5): 267–275. doi:10.1017/cts.2018.332.

Sustainability Considerations for Clinical and Translational Research Informatics Infrastructure

Jihad S. Obeid[1], **Peter Tarczy-Hornoch**[2], **Paul A. Harris**[3], **William K. Barnett**[4], **Nicholas R. Anderson**[5], **Peter J. Embi**[4], **William R. Hogan**[6], **Douglas S. Bell**[7], **Leslie D. McIntosh**[8], **Boyd Knosp**[9], **Umberto Tachinardi**[10], **James J. Cimino**[11], and **Firas H. Wehbe**[12]

[1]Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC.

[2]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA.

[3]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN.

[4]Regenstrief Institute, Inc. and Indiana University School of Medicine, Indianapolis, IN.

[5]Department of Public Health Sciences, University of California at Davis, Davis, CA.

[6]Department of Health Outcomes and Biomedical Informatics, University of Florida, Gainesville, FL.

[7]Department of Medicine, University of California Los Angeles, Los Angeles, CA.

[8]Research Data Alliance-US, Rensselaer Polytechnic Institute, Troy, NY.

[9]Institute for Clinical and Translational Science, University of Iowa, Iowa City, IA.

[10]Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, WI.

[11]Informatics Institute, University of Alabama at Birmingham, Birmingham, AL.

[12]Department of Preventive Medicine, Division of Health and Biomedical Informatics, Northwestern University Feinberg School of Medicine, Chicago, IL.

Abstract

A robust biomedical informatics infrastructure is essential for academic health centers engaged in translational research. There are no templates for what such an infrastructure encompasses or how it is funded. An informatics workgroup within the Clinical and Translational Science Awards network conducted an analysis to identify the scope, governance and funding of this infrastructure. After we identified the essential components of an informatics infrastructure, we surveyed Informatics leaders at network institutions about the governance and sustainability of the different components. Results from 42 survey respondents showed significant variations in governance and sustainability; however, some trends also emerged. Core informatics components such as

Corresponding author: Jihad S. Obeid, MD, 135 Cannon St., Suite 405, MSC200, Charleston, SC 29425., jobeid@musc.edu.

Disclosures

There are no conflicts of interests relevant to the issues examined and discussed in this publication.

The authors have no conflicts of interest relevant to the topic of this manuscript.

electronic data capture systems, electronic health records data repositories and related tools had mixed models of funding including, fee-for-service, extramural grants, and institutional support. Several key components such as regulatory systems (for example, electronic IRB systems, grants and contracts), security systems, data warehouses, and clinical trials management systems were overwhelmingly supported as institutional infrastructure. The findings highlighted in this report are worth noting for academic health centers and funding agencies involved in planning current and future informatics infrastructure, which provides the foundation for a robust, data-driven clinical and translational research program.

Keywords

translational research; infrastructure; biomedical informatics; CTSA; sustainability

INTRODUCTION

Academic Health Centers (AHCs) have invested significant resources during the past decade in building their research infrastructure in order to be competitive in the inevitably evolving landscape of translational research, genomics, and personalized medicine. Since its inception, the translational science roadmap proposed by the National Institutes of Health (NIH), highlighted the NIH's investment in an informatics infrastructure (1, 2). This roadmap culminated in the NIH's Clinical and Translational Science Award (CTSA) program that provided a substantial but partial funding for clinical and translational research infrastructure at institutions, the rest of which came from local institutional resources. Since the inception of the CTSAs, the requirements for the informatics cores—the organizational units responsible for support of informatics services within individual CTSAs—have evolved as defined within successive CTSA RFAs. Initially there was an emphasis on the creation of core informatics resources needed by clinical and translational researchers including, for example, research data warehouses and electronic data capture (e.g. "Biomedical Informatics resources, including critical information systems", "Biomedical informatics research activity should be innovative in the development of new tools, methods, and algorithms") (3). Over time, the emphasis shifted from the CTSA informatics cores building such resources (often with a combination of CTSA and institutional funding) to the assumption such core resources existed already and the goal was ensuring interoperability:

> "Biomedical informatics is the cornerstone of communication within C/D/Is [center, department, or institute] and with all collaborating organizations. Applicants should consider both internal, intra-institution and external interoperability to allow for communication among C/D/Is and the necessary research partners of clinical and translational investigators" (4).

Currently, the goal of the CTSA informatics cores is to leverage these resources for local investigators and national efforts to better support multi-institutional clinical trials (e.g. "Informatics is a high priority, overarching function that can transform translation at the CTSA hubs and in the CTSA network. Informatics resources, support, expertise, training, collaboration and innovation are critical to a successful translational research environment") (5). Additionally, the need for informatics and data science has become more pronounced in

recent years due to the emphasis on using electronic systems to improve efficiency for research studies (6), data-driven medicine (7), the learning health system (8), and precision medicine (9). Data-driven research strategies are critical in the discovery of new potential clinical interventions and advancing human health (10). In 2015, the National Library of Medicine (NLM) working group of the Advisory Committee to the NIH Director recommended that the NLM become the programmatic and administrative home for data science at NIH in order to address the research data science needs of the NIH (11). This resulted in several notices for request-for-information by the NLM aimed at soliciting input for its strategic planning (12, 13). Therefore, it is no surprise that AHCs recognize the need for investment in Biomedical Informatics (14). Virtually all AHCs that were funded by the CTSA program included a biomedical informatics component or core. However, there is no prescribed template for what a biomedical informatics program dedicated to research encompasses. There is no clear map of what are the building blocks, and what components of the infrastructure fall under the umbrella of Clinical and Translational Research Informatics (CTRI). There is a gap in our understanding of the boundaries of CTRI, as well as the different models of governance, maturity and support. As a result, this infrastructure varies significantly among different AHCs. Moreover, there has been no clear formula for how these components are funded initially and later sustained. As noted above there has been a shift in CTSA funding from establishing infrastructure to now focusing on leveraging infrastructure that is assumed to be in place, which makes it all the more timely to understand how AHCs are funding this vital informatics infrastructure.

This gap needs to be addressed so that AHCs can strategically plan for an informatics infrastructure that supports the broader local and nationwide translational research infrastructure as a whole. Therefore, a comprehensive characterization of CTRI structure and boundaries is critical. Toward this goal, we conducted a survey to assess the governance and sustainability models across informatics programs at CTSA-funded hubs. In this report, we outline the essential components of a CTRI infrastructure, and examine various avenues of support, governance and sustainability.

METHODS

The Informatics Domain Task Force (iDTF) of the CTSA network is comprised of leaders in research informatics at the respective institutions, who are engaged in day-to-day activities and governance issues related to research infrastructure. The iDTF membership includes at least one informatics representatives from all institutions in the CTSA network. A group of volunteers within the iDTF membership assembled to identify the scope of the infrastructure in question, and means of sustainability. The workgroup conducted monthly calls over a one-year period to identify the components of the CTRI infrastructure across CTSA funded institutions and the means by which these components are sustained. We used a simplified Delphi approach to identify a comprehensive list of the infrastructure components required for a robust translational research enterprise.

After the critical components were identified, we constructed a REDCap survey (15) to examine the local locus of control or ownership at each institution for each of the identified components.

The options for locus of control (LOC) or ownership included the following:

- Informatics: realized at an institution by an informatics core, service center, department or institute, including entities under the direction of a Chief Research Information (or Informatics) Officer (CRIO). Whatever the type of this entity, it usually gets some funding from the CTSA for some of the activities described below.
- Information Technology (IT) or Information Systems departments: typically directed by the Chief Information Officer (CIO) either at the health or hospital system, or at the university or college.
- Research office: for example, the Office of Research and Sponsored Programs, Clinical Research Office, usually under the governance of a vice president, provost or dean for research.
- Other: this was provided as an option for LOC that does not fit any of the above three.

Respondents were allowed to select one or more LOC options for each infrastructure category.

The options for avenues of funding or sustainability included the following:

- Institutional infrastructure
- Fee-for-service model
- Sustained grant support
- Other

Here again, respondents were allowed to select one or more funding mechanism for each infrastructure category. The survey link was sent to the rest of the iDTF members via email. The full survey is available in the supplementary material. To facilitate the reproducibility of this work, we made the REDCap data dictionary as well as the code used to analyze the survey results publicly available (16). The survey data was analyzed using R software for statistical computing v3.5.0 (17). The analysis primarily included descriptive statistics. In addition, we examined statistical correlations between responses. For the free text comments and questions, we manually categorized and reviewed all the text replies to look for themes across different questions. We provided some representative examples of the replies in the results.

RESULTS

Informatics Infrastructure Components

During the initial phase of the project, the workgroup focused on identifying the components of CTRI infrastructure. The team converged on a hierarchical list of 66 resource components that were grouped into 16 categories of components under 6 major headings (see table 1). A more detailed list of the 66 resources along with definitions and examples is available in the supplementary material (supplementary table 1).

Governance and Support of the Infrastructure

Representatives from 42 of 64 CTSA hubs (66%) responded to the survey. Tables 2 and 3 show the response totals for each category of components. The results are visualized in figures 1a and 1b as radar graphs. In an attempt to examine differences between responders and nonresponders, we examined each hub's funding level, which is likely a reflection of the magnitude of the overall NIH funding at a given hub. We extracted funding information from the NIH Research Portfolio Online Reporting Tools (RePORTER) (18). Nonresponders included both smaller and larger hubs; however, the mean annual funding level is $8,454,397 (Standard Error of Mean (SEM)=$785,893) for responders and $6,236,066 (SEM= $597,043) for nonresponders. A 2-tailed Student's t-test (assuming unequal variances) is 0.015. Therefore, a limitation of the study is that smaller programs may be under-represented.

Several research components were managed primarily by Informatics: electronic health records (EHR) data repositories/data warehousing (n=41, 98% of respondents); electronic data capture (EDC) (n=39, 93% of respondents); training, support and education (n=39, 93%); and extramural data collaborations (n=38, 90%). Virtually all component groups (including the ones mentioned above) had overlapping management by a combination of Informatics, Research Office, IT or other. Similarly, all had mixed funding through one or more of institutional funding, fee-for-service, grants or other; however, notably, over 90% of respondents indicated that regulatory compliance systems, IT infrastructure, security, EHR data warehousing, governance resources, grants and contracts systems, and clinical trials management systems (CTMS) were all funded primarily by institutional funds.

We computed a correlation matrix across all 128 variables using Pearson's r (16 component groups x 8 checkboxes, 4 LOC and 4 sustainability) (supplementary figure 1). We then filtered to those within components in order to detect associations and trends between LOC and sustainability. Several observations emerged from this analysis; for example, Informatics LOC was correlated with grant support across many components (e.g. regulatory, extramural collaborations, cyberinfrastructure, and innovation); LOC by IT, on the other hand, was correlated with institutional support. The data set and results are provided in supplementary file Supplementary_Tables_2_correlation_data.xlsx, including correlation analyses with p-values in sheet 3.

We conducted an analysis based on CTSA funding data from NIH RePORTER, as a proxy for overall NIH research funding at a given hub. We examined the association of between funding level and responses in the survey. We categorized the top 18 funded hubs (CTSA grant Total Cost > $7,000,000) in one group and the rest in another. There were no significant differences in sustainability models across those two groups; however, there were certain trends in higher-funded institutions; for example, more grant funding (67% vs. 42%, p=0.20) and institutional funding (94% vs. 79%, p=0.34) for education/training, more fee-for-service models for EHR data repositories (presumably from provision of data, 72% vs. 46%, p=0.16).

Contextual Feedback

There were over 250 comments provided by the 42 respondents in the free text fields under the headings of “general comments”, “locus of control notes”, and “sustainability options notes”. Several general comments from respondents relayed opinions that some categories (e.g. the broad categories of regulatory systems and data repositories) were not granular enough. Some respondents suggested other governance entities and loci of control for example Academic Informatics groups (which may or may not belong under the CIO) and Analytics teams under a Chief Analytics Officer (CAO) for data warehousing efforts, and administrative cores for research administrative systems. Several mentioned biobanking as largely sponsored by National Cancer Institute (NCI)funded entities (i.e. cancer centers, along with other smaller efforts across campuses). Three respondents mentioned a dedicated cancer center CTMS. In response to sustainability options, several respondents mentioned other sources of funding such as tuition fees for informatics education activities, and executive and industry sponsorship for CTMSs. Although extramural collaborations and research innovations were primarily grant funded, some proposed that those are good areas for strategic institutional investments.

DISCUSSION

A previous survey designed to examine adoption of clinical research information technology showed evidence of rapidly increasing adoption over the period of 2005 to 2011 in Information Technologies for Clinical Research, including research compliance systems, EDC, research data repositories, and other relevant infrastructure (19). However, the previous work did not address governance and sustainability of these technologies. Moreover, our current manuscript identifies several other components essential for a successful research informatics program (e.g. EHR research interface, education/training, methodological research/innovation) as deemed by our panel of experts. Of note, iDTF members and iDTF leadership (iDTF lead team) with support from CTSA Principal Investigators (PI) lead team identified the issues addressed by this survey as very important to get clarity on (through the chartering of the working group that led the study). An early version of this survey and the preliminary results were of such interest to iDTF members, that the survey was refined and invitation for participation was expanded to all hubs.

Structure and Governance

In order to fathom the breadth and boundaries of CTRI across diverse environments, the workgroup set out to identify various elements of the infrastructure, as summarized in table 1.

The major heading “Application and support of clinical and translational research” included by far the largest number of components, grouped into 10 subheadings, with a total of 44 components (supplementary table 1), which included several essential systems such as those used for regulatory compliance (e.g. Institutional Review Board (IRB) systems), data capture, data repositories and program evaluation. Not surprisingly, these systems varied significantly as to ownership and sources of funding; however, there were some expected trends. For example, regulatory systems (such as electronic IRB and grants and contracts

systems) tended to be under the control of the office of research and sustained as a core component of the institutional infrastructure. On the opposite side of the spectrum, components such as electronic data capture (EDC) systems and electronic health records (EHR) data repositories including implementation of self-service systems such as i2b2 (20), were primarily governed by Informatics and tended to have diverse sources of sustainability including institutional, fee-for-service and grant support.

The second major heading “Research collaboration” focused on extramural collaborations that involved systems that span data across multiple institutions, such as PCORnet (21) and i2b2/SHRINE (22, 23) projects. Such collaborations required the development and implementation of common data models and standardized terminologies that allow federated queries across multiple institutions. Aside from the underlying technology, these collaborations need to accommodate multiple levels of security and governance allowing delivery of de-identified information for cohort discovery to identified data after IRB approval for multi-site clinical trials. A few industrybased models such as TriNetX® and Flatiron Health, Inc. were given as examples of collaboration options using commercial systems. The majority of respondents (90%) stated that extramural collaborative projects were managed by informatics groups; however, about half of those stated co-management by other local groups such as the Research Office or IT department. Non-commercial collaborations were predominantly funded by grant support.

Cyberinfrastructure was identified as the third major heading and that includes the information technology backbone, which allows the necessary translational research tools to run. This included server and network hardware, software licenses for backend institution-wide systems and office-based products, along with an underlying robust security framework. These are controlled largely by an IT department, viewed as essential components of the institutional infrastructure, and funded as such. Almost half the respondents stated that some of the burden of research-specific systems are shared with the informatics teams.

The fourth major heading focused on oversight and governance resources specifically related to research IT and Informatics. This included two broad categories, one focused on serving on governance and steering committees and the other focused on regulatory operational support (i.e. resources needed to enforce the operational aspects of policies and regulations, for example, data release approval processes, support for audits and other digital workflow requirements). The majority (27 or 64%) of respondents stated that informatics personnel participated in these activities, but in most cases, they were shared by IT and research office personnel. In terms of sustainability this was viewed as an infrastructure requirement, however 10 respondents (24%) stated that in some cases this activity was supported by grants, presumably, grant-specific systems’ governance and oversight activities.

The fifth major heading addressed training and education. This included a variety of mechanisms: formal courses in informatics, seminars, workshops, training videos, and one-on-one training and support. Most of the respondents (93%) suggested that Informatics participated in or led training efforts. In several cases, this was viewed as part of the academic mission and supported by the institution or through academic programs and tuition

fees. Half the respondents indicated that training was grant supported, presumably either for training on grant-specific systems or partly via informatics training grants.

The sixth and last major heading covered “methodological research activities in informatics and innovation”. The rationale for including this category is that it is critical for driving grant funding. Moreover, informatics infrastructure often generates the data about clinical and translational research processes that it supports and lends itself to a continuous learning and improvement cycle based on examination of these data. Methodological research and innovation is essential in advancing the field of informatics, and data science including machine learning, natural language processing (NLP) and image processing. Novel methods in electronic data capture, electronic consents, and EHR-research pipelines, to name a few, are also necessary drivers for more efficient, less costly execution of clinical trials. Several components of an innovation engine were included as resources: dedicated informatics faculty, a student pipeline, software engineers, support staff, and support for rapid prototyping and implementation. These activities inherently belong in an academic informatics program. This rationale is based on the premise that a successful program requires a healthy publication output and significant investment in effort on grant proposals. Not surprisingly, these resources were primarily supported by grant activity as reported by the majority (81%) of respondents, several of whom mentioned pilot funding within their institutions as a potential source. However, more than half (69%) of the respondents indicated that this was also considered and supported as an institutional investment presumably for cultivating innovation and recruiting talent. Twenty-six percent reported innovation as part of fee-for-service activities. The example given was when the fee-for-service core is used for prototype development in support of specific projects or programs.

Sustainability Themes

Several components were overwhelmingly supported as institutional infrastructure, although under the control of different centers of authority, e.g. IRB support, grants and contracts, CTMS, EHR data repositories for research, cyberinfrastructure and communications systems. These are considered core components for any academic institution engaged in translational research. As such, they are sustained by institutional funds as part of operational costs. In some cases, where there is ongoing research or innovations in these areas, grant funding could be sought; for example, research into IRB reliance, or into novel granular patient-level security in data repositories. On the other hand, several components were in many cases deemed reliant on extramural funding, for example, extramural collaborations, and methodological research and innovation activities. Grant funding in these areas typically includes funding as a core or program on infrastructure grants such as the CTSA grant, as faculty and staff effort on other program grants, or research-specific efforts. The correlation analyses identified interesting and statistically significant (p-values < 0.05) associations between Informatics involvement and grant support across many components (e.g. regulatory, extramural collaborations, IT infrastructure, innovation, and security); whereas, IT control, was correlated with institutional support. Several sites are adding fee-for-service models to their portfolio of sustainability, especially in the areas of EHR data extraction from their EHR data repositories (n=24, 57% of respondents) and assistance in constructing electronic data capture tools (n=22, 52% of respondents). We project that such

fee-for-service will gain more momentum as extramural funding becomes tighter and the as pharma industry increases support for clinical research. A recent report identifies the different methods of using EHR systems for enhancing research recruitment and the levels of adoption at CTSA institutions (24). Most notable was the high adoption rates for brokered access to EHR data warehouses and the availability of self-service interfaces for direct access by researchers for deidentified exploratory analyses. Our data shows that such data repositories are governed and brokered primarily by Informatics and supported largely as an institutional investment reinforcing the importance of these technologies. Provision of an EHR data warehouse and related services for research purposes are no longer optional components for a robust translational research enterprise.

Contextual Comments

The comments expressed in the unstructured fields of the survey provided insight into areas not covered by the structured responses. As detailed in the results, that information reflects the variety of organizational structures at the representative CTSA institutions, as well as local views and alternative approaches to governance structures (e.g. CAO in data warehousing efforts, and cancer centers in biobanking efforts) and sustainability (e.g. tuition fees for educational activities, and industry sponsorship for CTMSs).

General Remarks

Given the variability in governance and funding models between sites, it is unlikely that one approach will be appropriate for all. This paper presents a set of alternatives that should be weighed by experienced informaticians based on local factors. We hypothesize that some of the variability is related to differences in organizational structures that we did not ask about in this survey, for example the presence of Informatics departments, centers or institutes and structure and scope entities led by Chief Information Officers and Chief Medical Information Offices, etc. Consensus on best practices was out of scope for this working group and white paper and could be explored in future work by other working groups in the iDTF. That said, common patterns in our results suggest there is some benefit to those patterns. Thus, strong trends are de facto evidence of best practices, e.g. central funding for data warehousing, regulatory systems, and CTMS, which may help those sites that do not have stable funding for these components make a stronger case for it. Nonetheless, the responses included several permutations of governance and funding, proposed new ones, such as cancer centers and industry sponsorship in the support of CTMS, and highlighted the theme of strategic institutional investments in various areas of the infrastructure.

A question emerged in response to the work in this paper as to where the CTSA community is at large—in terms of having a stable and sustainable enough environment to keep the existing resources going vs. focusing now mainly on innovation. The results of this paper suggest that although a rich set of resources have been put in place across CTSAs, there remain important gaps to address with regards to the ongoing maintenance and refinement of these resources in a sustainable fashion that is not dependent on ongoing infrastructure investments by NCATS. As discussed below, one of the key next steps is looking at best practices for charge-back approaches that can help with sustainability.

Measuring Maturity

Measuring maturity is a useful way to assess research IT services for institutional planning. The maturity model process provides a formal staged structure providing an assessment of institutional preparedness and capabilities (25, 26), which are critical factors in sustainability. Well developed in hospitals by HIMSS electronic medical record adoption model (27), and in university Information Technology by EDUCAUSE (28), these are now being applied to research IT through a process currently being developed through the Clinical Research IT Forum and by members of the Association of American Medical Centers (AAMC) Group on Information Resources (GIR) in which four of the authors here (Barnett, Anderson, Embi, Knosp) participate.

The maturity model process has two types of measures, a maturity model, or index, and a deployment model, or index. A maturity index measures organizational capacity to deliver a service, whereas a deployment index measures the degree to which an institution has formalized the implementation and sustenance of a particular technology.

These models or indices are typically measured at 5 levels of maturity:

- Level 1: Processes unpredictable and unmanaged
- Level 2: Managed and repeatable processes, but typically reactive
- Level 3: Defined, with standardized capabilities and processes
- Level 4: Quantitatively Managed, with measurement and formal control
- Level 5: Optimized, with regular assessments to improve processes

Although we were not able to undertake a formal maturity study based just on the two factors measured here (namely LOC and sustainability), the use of a formal maturity model process for CTRI services can play an important role in sustainability. Individual institutions can use the data presented here in the context of a maturity scale to better understand, and plan, their CTRI services.

Maturity models are useful in aligning institutional priorities and commitments, both operational and policy, with informatics services. This alignment can ensure appropriate executive commitment, as well as alignment of services with mission. Deployment indices are useful in understanding capabilities and gaps in particular technologies. These are critical in planning services and management commitments, to ensure more fully capable systems and capabilities that align with peers and partners for multi-site collaborations.

Limitations

Despite our efforts to balance between a long granular survey and a high-level short survey, respondents expressed their concerns in comments about a few broad categories that were difficult to address as a group. For example, the category of data repositories was deemed too broad; more granular subcategories included, an Enterprise Data Warehouse (as supported by IT), a Research Data Warehouse, i2b2, and data provision services (as supported by informatics). Data provision oversight is typically a collaboration between regulatory and informatics personnel, and is funded in part by the CTSA grant. For practical

reasons, a few important topics were left out of the survey, the most notable are: regulatory and compliance topics (e.g. Health Insurance Portability and Accountability Act (HIPAA)-related resources, de-identification of data) and services for unstructured data (i.e. NLP, image processing).

For brevity, we provided only binary options (checkboxes) in the survey; therefore, we could not quantitate percentages of support at each institution (for example fee-for-20 service vs. grant funding vs. institutional support), instead we only looked at combinations of answers. Finally, responders tended to have larger CTSA awards than nonresponders, suggesting that our data may be more representative of larger hubs than smaller hubs.

Future Directions

A more thorough exploration of capabilities to measure return on investment is needed, which was out of scope for the survey and this report. iDTF members, iDTF lead team and liaison to CTSA PI lead team have identified one of the most important follow-up questions to be to identify and share best practices in terms of charge-back models for provision of informatics services to clinical investigators, which is resulting in the chartering of a new working group.

Further work in this area could explore differences in structure at various institutions and potential best practices across the diverse environments, as well as more detailed models for charge-back on provided services. New activities not specifically addressed in this manuscript should be considered. For example, as data sharing gains momentum with the dissemination of the FAIR (Findability, Accessibility, Interoperability, and Reusability) guiding principles (29), new services could be provided, on a fee-for-service basis or as a collaborative effort with informaticians on grants, to help researchers comply with data sharing requirements. The next phase of this work should also consider where innovations in emerging technologies such as blockchain, data lakes, NLP and ontologies would fit in terms of funding and sustainability.

Regular assessments using institutional maturity indices or deployment indices can help determine institutional preparedness and application specific deployment capabilities.

The outcomes of these assessments can play an important role in supporting sustainability efforts by providing a competitive landscape review, which supports institutional competitive and collaborative investments. These assessments can also aid alignment between mission and technology effort, which will support appropriate institutional investment in technologies. By identifying the next steps in technology investment to support the mission, these models can also serve as a planning tool for institutions, particularly those with shared missions of their AHCs.

CONCLUSION

The components identified in this report can serve as a checklist for a comprehensive translational informatics infrastructure. However, the data clearly show that there is not a fixed blueprint for the governance and sustainability of the different components across

institutions. A better understanding of the factors that influence the different models, through further investigation, may mitigate potential risks to the long-term stability of the infrastructure, and elucidate which sustainability models serve best for return on investment.

The landscape and trends highlighted above are worth noting for AHCs and funding agencies planning for current and future informatics infrastructure, which constitutes the cornerstone for a robust clinical and translational research program and data-driven biomedical research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

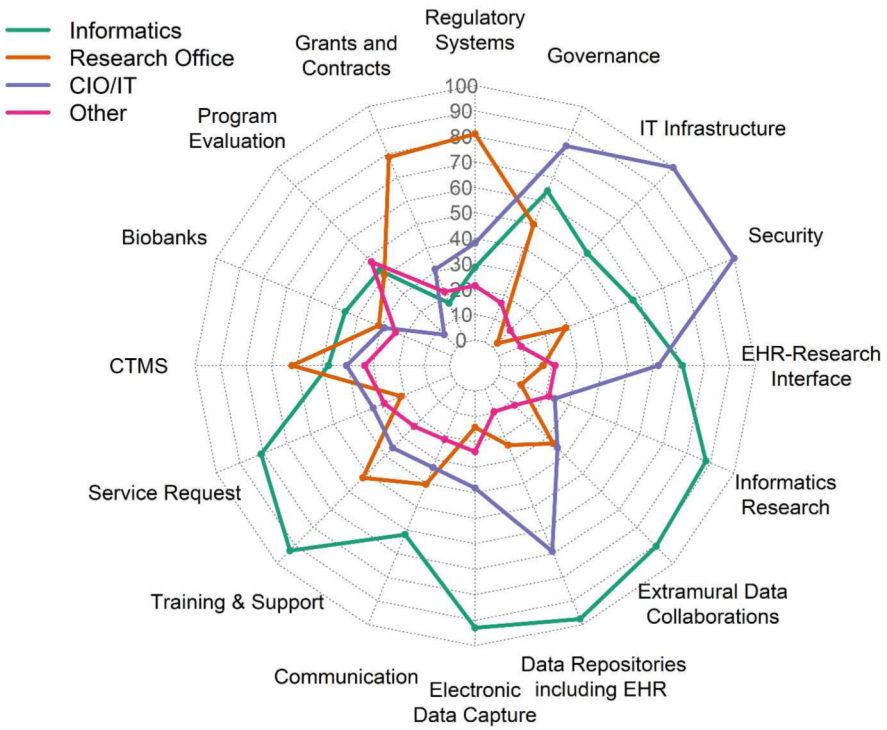
Acknowledgements

The authors wish to thank Robert R. Lavieri and Van K. Trieu of Vanderbilt University and Kate Fetherston of University of Rochester Medical Center for supporting the effort. This work would not have been possible without the support and responses of the individuals at each of the CTSA institutions who participated in the survey. This project was funded in whole or in part with Federal funds from the National Center for Research Resources and National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), through the Clinical and Translational Science Awards Program, grant numbers: UL1TR001450, UL1TR001422, UL1TR001860, UL1TR001108, UL1TR002001, UL1TR000445, UL1TR001427, UL1TR002537, UL1TR001102, UL1TR000448, UL1TR001876, UL1TR001111, UL1TR002373, UL1TR000423, UL1TR001436, UL1TR001425, UL1TR001417, UL1TR001881, UL1TR001105, UL1TR001857, UL1TR001863, UL1TR001445, UL1TR001079, UL1TR001412, UL1TR000433, UL1TR001070, UL1TR000457, UL1TR001082, U54TR001356, UL1TR001866, UL1TR000135, UL1TR001855, UL1TR001073, UL1TR001442, UL1TR000454, UL1TR001439, UL1TR000460, UL1TR001449, UL1TR000001, UL1TR001873, UL1TR001998, UL1TR001872, UL1TR000039, UL1TR002243, U54TR000123, and U24TR002260.

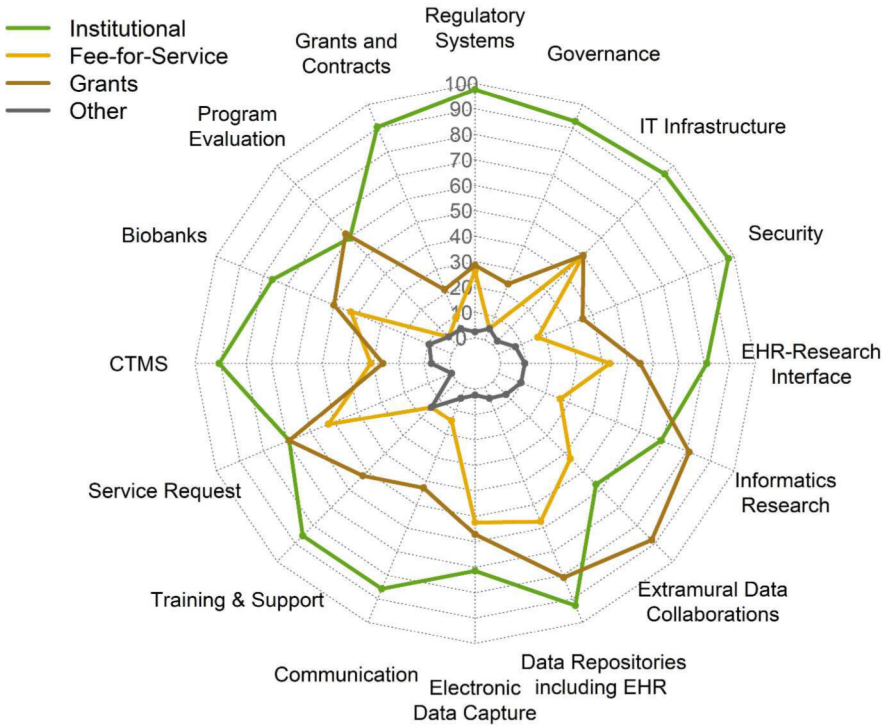
References

1. Zerhouni E, Medicine. The NIH Roadmap, Science 302, 63–72 (2003). [PubMed: 14526066]
2. Zerhouni EA, US biomedical research: basic, translational, and clinical sciences, JAMA : the journal of the American Medical Association 294, 1352–8 (2005). [PubMed: 16174693]
3. RFA-RM-06–002: Institutional Clinical and Translational Science Award (available at https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-06-002.html).
4. RFA-RM-09–004: Institutional Clinical and Translational Science Award (U54) (available at https://grants.nih.gov/grants/guide/rfa-files/rfa-rm-09-004.html).
5. PAR-15–304: Clinical and Translational Science Award (U54) (available at https://grants.nih.gov/grants/guide/pa-files/PAR-15-304.html).
6. Disis ML, Tarczy-Hornoch P, Ramsey BW, in Translational Research in Biomedicine, Alving B, Dai K, Chan SHH, Eds. (S. KARGER AG, Basel, 2012), vol. 3, pp. 89–97.
7. Staiger TO, Kritek PA, Luo G, Tarczy-Hornoch P, in Handbook of Anticipation, Poli R, Ed. (Springer International Publishing, Cham, 2017), pp. 1–21.
8. Friedman C, Rubin J, Brown J, Buntin M, Corn M, Etheredge L, Gunter C, Musen M, Platt R, Stead W, Sullivan K, Van Houweling D, Toward a science of learning systems: a research agenda for the high-functioning Learning Health System, J Am Med Inform Assoc 22, 43–50 (2015). [PubMed: 25342177]
9. Collins FS, Varmus H, A new initiative on precision medicine, N Engl J Med 372, 793–5 (2015). [PubMed: 25635347]
10. Collins FS, Reengineering translational science: the time is right, Sci Transl Med 3, 90cm17 (2011).
11. The National Library of Medicine Working Group of the Advisory Committee to the Director, Final Report (National Institutes of Health, 2015).

12. National Library of Medicine, Request for Information (RFI): Strategic Plan for the National Library of Medicine, National Institutes of Health (2016).
13. National Library of Medicine, Request for Information (RFI): Next-Generation Data Science Challenges in Health and Biomedicine (2017).
14. Califf RM, Berglund L, Principal C Investigators of National Institutes of Health, A. Translational Science, Linking scientific discovery and better health for the nation: the first three years of the NIH's Clinical and Translational Science Awards, Acad Med 85, 457–62 (2010). [PubMed: 20182118]
15. 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 42, 377–81 (2009). [PubMed: 18929686]
16. Wehbe F, Obeid J, Code for CTSA Sustainability Paper Figures (2017), doi:10.6084/m9.figshare.5340919.v2.
17. R Core Team, R: A Language and Environment for Statistical Computing. (2018) (available at https://www.r-project.org/).
18. NIH Research Portfolio Online Reporting Tools (RePORT) (available at https://report.nih.gov/index.aspx).
19. Murphy SN, Dubey A, Embi PJ, Harris PA, Richter BG, Turisco F, Weber GM, Tcheng JE, Keogh D, Current state of information technologies for the clinical research enterprise across academic medical centers, Clin Transl Sci 5, 281–284 (2012). [PubMed: 22686207]
20. 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), Journal of the American Medical Informatics Association : JAMIA 17, 124–30 (2010). [PubMed: 20190053]
21. The National Patient-Centered Clinical Research Network, PCORnet, the National Patient-Centered Clinical Research Network (2017).
22. Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, Kohane IS, The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories, J Am Med Inform Assoc 16, 624–30 (2009). [PubMed: 19567788]
23. McMurry AJ, Murphy SN, MacFadden D, Weber G, Simons WW, Orechia J, Bickel J, Wattanasin N, Gilbert C, Trevvett P, Churchill S, Kohane IS, SHRINE: enabling nationally scalable multi-site disease studies, PLoS One 8, e55811 (2013). [PubMed: 23533569]
24. Obeid JS, Beskow LM, Rape M, Gouripeddi R, Black RA, Cimino JJ, Embi PJ, Weng C, Marnocha R, Buse JB, for the Methods and Process and Informatics Domain Task Force Workgroup, A survey of practices for the use of electronic health records to support research recruitment, Journal of Clinical and Translational Science 1, 246–252 (2017). [PubMed: 29657859]
25. Crosby PB, Quality is free: the art of making quality certain (McGraw-Hill, 1979; https://books.google.com/books?id=0RxtAAAAMAAJ).
26. Radice RA, Harding JT, Munnis PE, Phillips RW, A programming process study, IBM Systems Journal 24, 91–101 (1985).
27. Healthcare Provider ModelsHIMSS Analytics - North America (available at https://www.himssanalytics.org/healthcare-provider-models/all).
28. Grajek S, The Digitization of Higher Education: Charting the Course http://er.educause.edu/articles/2016/12/the-digitization-of-higher-education-charting-the-course, EDUCAUSE Review (2016).
29. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, t Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B The FAIR Guiding Principles for scientific data management and stewardship, Sci Data 3, 160018 (2016). [PubMed: 26978244]



a. Locus of Control



b. Sustainability Model

Figure 1.

a) Distribution in percent of respondents for locus of control across component groups of infrastructure. b) Distribution in percent of respondents for sustainability models across different the component groups (n=42). Abbreviations: IT: Information Technology; CIO: Chief Information Officer; CTMS: Clinical Trials Management Systems; EHR: Electronic Health Records.

Table 1.

The groups of components and major headings that were identified by the workgroup.

Category Heading	Component Groups
Applications for Clinical and Translational Research	Research/Regulatory Compliance
	Service Request/Fulfillment
	Program Evaluation
	Grants and Contracts Systems
	Clinical Trials Management Systems
	Electronic Data Capture
	Biobanking systems
	Data Repositories and EHR data
	EHR systems research interface
	Communication
Research Collaboration	Extramural data collaborations
Cyberinfrastructure	Security
	IT infrastructure
Oversight and Governance	Governance resources
Training and Support	Education and training
Research and Innovation	Methodological Informatics research and innovation, faculty and other resources

Abbreviations: EHR: Electronic Health Records; IT: Information Technology.

Table 2.

Number of respondents (with %) for each category of components and their selections for locus of control: Informatics, Research Office, and/or IT Department. Total n=42.

Category of components	Informatics	Research office	IT Department / CIO	Other
Regulatory systems	12 (29%)	34 (81%)	16 (38%)	9 (21%)
Service Request	34 (81%)	9 (21%)	14 (33%)	12 (29%)
Program Evaluation	18 (43%)	17 (40%)	3 (7%)	20 (48%)
Grants and Contracts	7 (17%)	33 (79%)	13 (31%)	9 (21%)
CTMS	20 (48%)	26 (62%)	17 (40%)	14 (33%)
Electronic Data Capture	39 (93%)	6 (14%)	16 (38%)	10 (24%)
Biobanks	19 (45%)	13 (31%)	12 (29%)	10 (24%)
Data Repositories including EHR	41 (98%)	10 (24%)	29 (69%)	4 (10%)
EHR-Research Interface	30 (71%)	7 (17%)	26 (62%)	9 (21%)
Communication	26 (62%)	17 (40%)	14 (33%)	9 (21%)
Extramural Data Collab	38 (90%)	14 (33%)	15 (36%)	5 (12%)
Security	24 (57%)	12 (29%)	42 (100%)	4 (10%)
IT Infrastructure	22 (52%)	1 (2%)	42 (100%)	4 (10%)
Governance	27 (64%)	21 (50%)	35 (83%)	7 (17%)
Training & Support	39 (93%)	22 (52%)	15 (36%)	10 (24%)
Informatics Research	37 (88%)	4 (10%)	10 (24%)	9 (21%)

Abbreviations: IT: Information Technology; CIO: Chief Information Officer; CTMS: Clinical Trials Management Systems; EHR: Electronic Health Records.

Table 3.

Number of respondents (with %) for each category of components and their selections for sustainability options: institutional infrastructure, fee-for-service, and/or grant support. Total n=42.

Category of components	Institutional infrastructure	Fee-for-service	Grant support	Other
Regulatory systems	41 (98%)	11 (26%)	12 (29%)	1 (2%)
Service Request	29 (69%)	22 (52%)	29 (69%)	0 (0%)
Program Evaluation	25 (60%)	2 (5%)	26 (62%)	2 (5%)
Grants and Contracts	38 (90%)	4 (10%)	9 (21%)	2 (5%)
CTMS	38 (90%)	13 (31%)	11 (26%)	3 (7%)
Electronic Data Capture	30 (71%)	22 (52%)	24 (57%)	1 (2%)
Biobanks	32 (76%)	18 (43%)	21 (50%)	4 (10%)
Data Repositories including EHR	39 (93%)	24 (57%)	34 (81%)	2 (5%)
EHR-Research Interface	34 (81%)	18 (43%)	23 (55%)	4 (10%)
Communication	36 (86%)	6 (14%)	18 (43%)	2 (5%)
Extramural Data Collab	24 (57%)	18 (43%)	37 (88%)	3 (7%)
Security	41 (98%)	7 (17%)	15 (36%)	3 (7%)
IT Infrastructure	40 (95%)	21 (50%)	21 (50%)	1 (2%)
Governance	39 (93%)	2 (5%)	10 (24%)	2 (5%)
Training & Support	36 (86%)	6 (14%)	22 (52%)	6 (14%)
Informatics Research	29 (69%)	11 (26%)	34 (81%)	4 (10%)

Abbreviations: IT: Information Technology; CTMS: Clinical Trials Management Systems; EHR: Electronic Health Records.