Chapter 9
Clinical and Translational Research Informatics Education and Training

Peter J. Embi and Philip R. O. Payne

Role of Informatics in Clinical and Translational Science

The modern biomedical research domain has experienced a fundamental shift towards integrative clinical and translational research. This shift has been manifested in a number of ways, including the launch of the NIH Roadmap initiative [1–3] that has resulted in the creation of the Clinical and Translational Science Award (CTSA) program [3], as well as the rapid growth of high-throughput biomolecular technologies and corresponding bio-marker-to-phenotype mapping efforts [4]. A commonly reported thread in a broad variety of reports and commentaries concerned with this evolution focuses on the challenges and requirements related to the collection, management, integration, analysis, and dissemination of large-scale, heterogeneous biomedical data sets [5–8]. However, well-established and broadly adopted theoretical and practical frameworks intended to address these needs are still lacking in the biomedical informatics knowledge base [7, 9–11]. Instead, the development and execution of integrative clinical or translational research is significantly limited by the propagation of "silos" of both data and expertise.

A critical need in overcoming such barriers to the efficient, timely, and impactful conduct of clinical and translational research is the development of a biomedical and informatics workforce educated and trained to make contributions both by leveraging informatics capabilities to accelerate biomedical research and to advance basic and applied science in the field of biomedical informatics itself.

P. J. Embi (✉)
Regenstrief Institute and Indiana University, Indianapolis, IN, USA
e-mail: pembi@regenstrief.org

P. R. O. Payne
Washington University in St. Louis, St. Louis, MO, USA
e-mail: prpayne@wustl.edu

E. S. Berner (ed.), *Informatics Education in Healthcare*, Health Informatics,
https://doi.org/10.1007/978-3-030-53813-2_9

As the conduct of clinical and translational research is an information-intensive task, much work at the intersection of biomedical informatics and biomedical research is needed and has, in fact, been ongoing. Indeed, in recent years, the application of biomedical informatics principles, approaches and tools to the conduct and support of clinical and translational research has evolved. The result is the emergence of two complementary biomedical informatics sub-disciplines that have arisen in response to the unique challenges and opportunities facing research, namely Translational Bioinformatics (TBI) and Clinical Research Informatics (CRI). While definitions vary, we will define these two sub-disciplines as follows:

- **Translational Bioinformatics** (**TBI**) is the sub-discipline of biomedical informatics concerned with the development of storage, analytic, and interpretive methods to optimize the transformation of increasingly voluminous biomedical data into what has been called P4 medicine (predictive, preventive, personalized and participatory) [4, 12, 13]
- **Clinical Research Informatics** (**CRI**) is the sub-discipline of biomedical informatics concerned with the development, application, and evaluation of theories, methods and systems to optimize the design and conduct of clinical research and the analysis, interpretation and dissemination of the information generated [5].

Given that these domains of TBI and CRI are both complementary and critical to the conduct of clinical and translational research these two sub-disciplines can collectively be referred to as **Clinical and Translational Research Informatics** (**CTRI**), and this overarching sub-domain of biomedical informatics is what we will focus on in this chapter.

As depicted in Fig. 9.1, the combined sub-domain of CTRI overlaps with, and complements, the related, but distinct, informatics sub-domains concerned with aspects of basic and early translational science (e.g. bioinformatics), clinical practice (e.g. clinical informatics), and public and population health (e.g. public health informatics). This range of domains has been referred to as the translational research spectrum with multiple points of translation as shown in Fig. 9.1. As such, it is evident that CTRI spans the T1 and T2 ends of the translational research spectrum.

Challenges and Opportunities of CTRI

Management of Heterogeneous Data Sets

The ability to collect and manage heterogeneous data sets with increasing levels of dimensionality is a significant challenge. The dissemination and adoption of advanced information management platforms that will allow researchers and their staff to focus on fundamental scientific problems rather than practical informatics needs are critical to reducing the burden of managing large multi-dimensional data sets [7, 10, 14, 15]. Central to the ability to realize this opportunity is the imperative

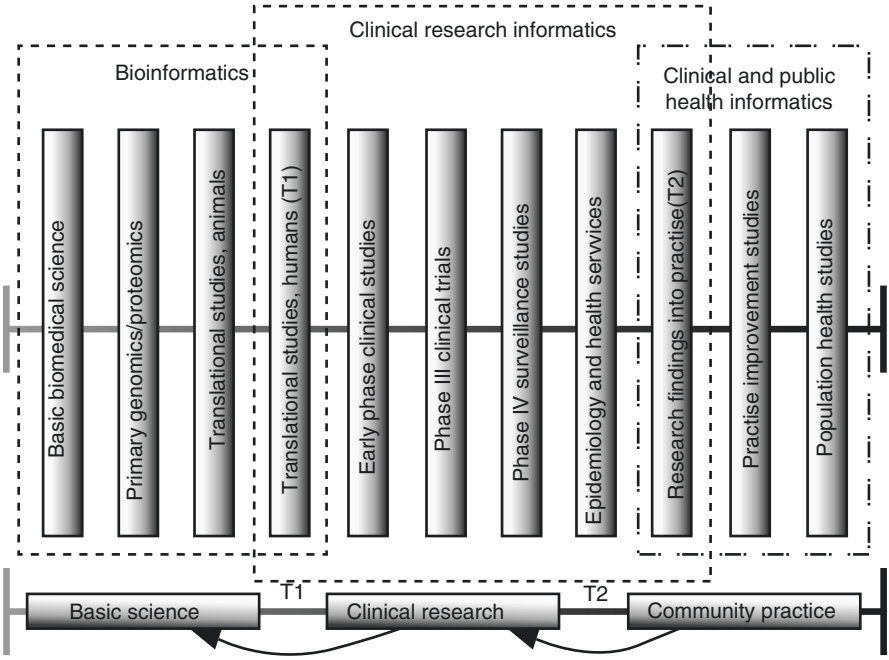

Fig. 9.1 Illustration of types of research across which CTRI is focused, and the relationships between CTRI and the other sub-domains of translational bioinformatics, clinical informatics, and public health informatics. These relationships also parallel the focus areas and methodologies associated with the clinical and translational science paradigm, including the commonly referred to T1 and T2 blocks in translational capacity (where the T1 block is concerned with impediments to the translation of basic science discoveries into clinical studies, and the T2 block with the translation of clinical research findings into community practice) (Reprinted with permission from BMJ, which holds the copyright, from Embi and Payne [5])

that the semantics of such data be well understood and made actionable relative to such operations [16–18].

Appropriate Methods and Tools

The need for knowledge-anchored methods and tools intended to enable the discovery, query, and integration of local distributed data, information, and knowledge resources is critical. This challenge is particularly pressing in multi-disciplinary team-science programs. The challenge is compounded by the fact that knowledge needed to discover, query, and integrate heterogeneous data, and information is often spread over a variety of sources [16]. The utilization of knowledge sources by scientific end-users is significantly hampered due to a lack of easy-to-use tools for knowledge resource discovery and information retrieval. Development of such tools is an opportunity for informatics.

Workflow Facilitation

The provision of systematic and extensible platforms capable of expediting workflows for knowledge integration and analysis is critical to discovery science paradigms. The challenge in facilitating workflow is exacerbated by the lack of availability of systematic data and knowledge "pipelining" tools that are capable of supporting the definition and reuse of computational workflows incorporating multiple source data sets, contextual knowledge sources, intermediate data analysis steps and products, and output types [19, 20].

Workforce Needs

As illustrated by the challenges and opportunities facing the CTRI sub-domain, there exists "…a major need to educate informaticians, clinical research investigators/staff, and senior leadership concerning the theory and practice of CTRI. Such education was thought to be necessary to ensure appropriate expectation management; adoption/utilization of CTRI related methods or tools; and the allocation of appropriate resources to accomplish organizational aims" [5].

Such programs enable the creation of a critical pipeline of experts and thought leaders needed to drive CTRI as a discipline, expanding the current state of clinical and translational research informatics education in general.

Indeed, for the reasons stated above and due to significant progress in recent years, CTRI has emerged as a distinct discipline in its own right. Initiatives such as the CTSA program noted above have helped to galvanize the CTRI community and drive important work in CTRI with the goal of advancing clinical and translational science.

Amidst these ongoing efforts and the progress that has recently been made in CTRI both nationally and internationally, it is recognized that the numbers of IT, informatics, and research professionals trained in CTRI is quite small and inadequate to support the advancements needed if we are to reap the benefits promised by this field.

In order to develop an adequately trained workforce with expertise in the critically important and emerging domain of CTRI, a range of new programs have been under development in recent years. Such programs enable the creation of a critical pipeline of experts and thought leaders needed to drive CTRI as a discipline, expanding the current state of clinical and translational research informatics education in general. Those involved recognize that there are multiple levels of education and training needed to expand the research informatics workforce, including a variety of related but distinct programs that will serve audiences needing different levels of training/educational intensity based upon their career goals and job-requirements. These range from short tutorials, to intensive courses, to certificate programs, to formal training culminating in Masters or PhD level education in CTRI.

Table 9.1 Educational program applicability by learner stage/role

	Tutorial	Multi-week course	Certificate program	Master's degree (or PhD)
Student/resident, clinicians, faculty, leadership	**X**	x		
Investigators, research staff, or informatician liaisons	x	**X**	x	
Informatician, investigator, or research staff who will use or support research informatics		x	**X**	x
Informatician with research informatics career focus			x	**X**

X most applicable, *x* possibly applicable

The different levels of education for learners at varying levels of intensity based upon their stage of training, their role in the research and informatics/IT enterprise, and their career goals guide such program development. A description of the varying types of learners and the related types of training that would likely be relevant/of interest to such groups of learners is depicted in Table 9.1. As the chart depicts using different size marks, learners in each category on the left may opt for more or less intensive training, but we have indicated with the large "**X**" those offerings we think most appropriate to each type.

To date, such programs are few and far between. However, there are some being delivered at the time of this writing, such as: (1) in-person and online "short courses" in CRI; (2) CRI online training programs; (3) Certificate programs in "Clinical and Translational Research Informatics" via online, distance-learning. A discussion of the curricular content areas will follow, but first we will lay out the different types of education and training opportunities that tend to dominate the current CTRI landscape.

Tutorials and Short Courses

In order to provide a basic understanding of clinical and translational research informatics to a wide audience including students, clinicians, research personnel and even institutional leaders who may not require or be interested in more intensive and lengthy programs of study, some research informatics "short-courses" or tutorials have been developed. Such courses typically consist of a truncated subset of information from a more intensive weeks-long research informatics course, such as the one described below and are delivered both online and in-person.

One such example is a 3-hour tutorial offered at national informatics professional meetings. The goal of such a program is to familiarize the groups listed with the basic concepts, goals, and utility of biomedical informatics approaches as they

relate to advancing both the generation of evidence (i.e. through research as well as through common data collection, subject recruitment, and other activities) and the translation of research knowledge into practice.

Driven by the recognition of the importance of education and training focused on research informatics to ensure optimal use of information resources and capabilities across the research enterprise, some have also developed and deployed formal educational programs specifically focused in the CTRI space. One such example was a clinical and translational research informatics online training program developed by Embi and colleagues in collaboration with the American Medical Informatics Association's (AMIA) 10 × 10 initiative [11]. This 10 × 10 program, which was conducted from 2011–2016, provided students with an intensive survey of the field of CTRI delivered mostly via distance-learning, with a concluding face-to-face session that took place at an AMIA national meeting. Using state-of-the-art asynchronous distance education resources and techniques, the program incorporated multiple modes of learning and participant interaction including weekly voice-over-PowerPoint lectures, threaded discussion forums, online knowledge assessments, and a class project that is presented during a concluding face-to-face session. The audience included: (a) investigators interested in learning more about CTRI's relevance to clinical/translational research, (b) informaticians who are interested in strengthening their knowledge of CTRI as a subdomain of biomedical informatics, and (c) other students interested in the domain, such as those from the biotechnology or pharmaceutical industry, government, etc. A typical course schedule/curriculum for the 10 × 10 program is depicted in Table 9.2.

Certificate Programs

Beyond tutorials and short courses, there exists another level of training programs in CTRI that lead to granting of formal university-based certificates and that are often delivered via online/distance-learning offerings. The programs typically draw upon and leverage courses from CTRI tracks of Masters or PhD level curricula, though some are stand-alone. Typical certificate programs include a five-course series consisting of core courses and tracks with the ability for partial customization to suit learners (Fig. 9.2). Courses include such titles as: Introduction to Biomedical Informatics; Clinical and Translational Research Informatics; Decision Analysis and Cost Effectiveness Analysis; and Quality Improvement and Patient Safety; Introduction to Bioinformatics; Computational Genomics; Data Modeling and Database Design; IS/IT Architecture; JAVA Programming for the Enterprise; and Introduction to Research Methods and Biostatistics.

Sometimes, those who start off with the short-course option, will transition to the certificate to gain further knowledge. Typically, enrollees work with their advisor to determine whether they should pursue a T1 or T2 focused program of coursework, or "Track," based upon their interests, background, and career goals. That is, those who are interested in either a T1 (research informatics as applied toward the T1 end

Table 9.2 Curriculum for clinical research informatics 10 × 10 course

Week	Competencies (at the conclusion of this session students will be able to:)
1. Course overview and general biomedical informatics principles	Discuss the goals of the course
	Discuss basic principles of biomedical and health informatics including health system architectures, evaluation, etc.
	Discuss definitions of biomedical informatics and of the clinical research informatics subdomain of biomedical informatics
	Discuss the major challenges and opportunities facing the CRI domain.
2. Overview of clinical research	Discuss the definitions and types of clinical research and the related areas of translational research
	Discuss basic principles of clinical research including the research process, aspects of study design, data collection and analysis, etc.
3. Informatics applications in clinical research, part 1	Discuss the application of research-specific informatics approaches and tools in clinical research
	Discuss the uses of general informatics systems as applied to clinical research
	Discuss informatics methods and tools applied to research hypothesis development
4. Informatics applications to clinical and translational research, part 2	Discuss informatics methods and tools applied to protocol development
	Discuss informatics methods and tools applied to patient recruitment
	Discuss informatics methods and tools applied to adverse event surveillance and pharamcovigiliance
	Discuss informatics methods and tools applied to dissemination and utilization of research findings
5. Research data collection, management and analysis	Discuss current best practices and principles for data collection, management and reporting
	Discuss methods and tools applied to research data collection
	Discuss methods and tools applied to data analysis and reporting
6. Enterprise systems in CRI	Discuss principles and practice of research database and data warehouse development
	Discuss the key elements and features of clinical trial management and electronic data capture systems
7. Data and knowledge standards in CRI	Discuss the importance of standards, terminologies and models in biomedical informatics
	Discuss ontology and model initiatives in CRI
8. Regulatory and ethical issues in CRI	Discuss key issues in privacy, confidentiality and research oversight relevant to CRI practice
	Discuss key ethical considerations in research informatics
	Discuss key principles and tools for trial registration and results dissemination

(continued)

Table 9.2 (continued)

Week	Competencies (at the conclusion of this session students will be able to:)
9. Translational research informatics, and CRI-BMI overlaps	Discuss the applications of informatics principles of translational science (both T1 and T2)
	Discuss the overlap of clinical research informatics and related domains of clinical informatics, translational bioinformatics, and public health informatics
10. Review major CRI initiatives and future directions	Discuss major national and international initiatives driving the CRI Agenda
	Discuss key CRI directions for the future

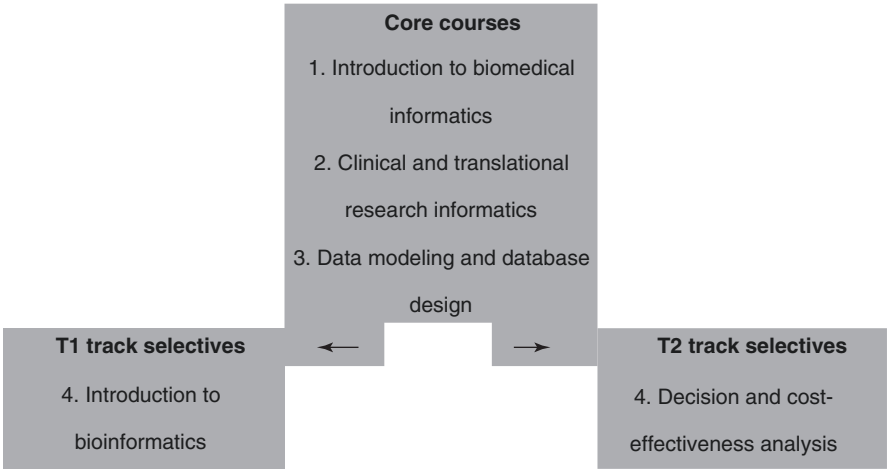

Fig. 9.2 Example curricula for certificate program enrollees, branched into T1 and T2 example tracks

of the translational spectrum) or T2 (research informatics as applied toward the clinical/population health end of the translational spectrum) emphasis, follow a customized "track" focused more so on bioinformatics or clinical informatics respectively, as appropriate. Figure 9.2 demonstrates example curricula for each track students might pursue.

Masters, PhD and Fellowship Programs in CTRI

For those who will focus on CTRI as their primary area of emphasis in a biomedical informatics career, formal training at the masters, doctoral, or fellowship level is appropriate. Training programs have been developed with just such a focus, and provide exploration of exemplary data, information, and knowledge management challenges and opportunities that exist as the intersection of biomedical informatics

and both clinical and translational science. Such programs tend to offer a foundation in biomedical informatics, with an emphasis on issues unique to the CTRI subspecialty.

Lessons Learned

When viewed in a holistic manner, the preceding CTRI-focused training landscape and its historical evolution serve to elucidate three important lessons learned, as enumerated below.

Tailoring the Focus of the Curriculum for Different Learner Roles

One key lesson learned by the CTRI community has been that there are a variety of types of individuals who require training and expertise with regards to the domain. For example, some individuals seek training in order to support or enable their ability to serve as CTRI practitioners, wherein they might be responsible for the development, management, and support of various technology platforms and interventions targeting the clinical and translational science domain. Other individuals may seek training in order to inform their pursuit of innovative and novel scientific studies concerned with biomedical informatics theories and methods that may serve to address the clinical and translational research information needs. Finally, individuals in leadership or decision making roles (e.g., policy makers, etc.) may seek training in CTRI in order to inform their analysis and understanding of critical policy, financial, and socio-technical issues with relevance to clinical and translational research that they may need to address.

Each of these types of individuals requires a different type of training, which can be generally differentiated based upon: (1) its breadth (coverage of domain) vs. depth (level of detail); (2) its degree of theoretical vs. application-level orientation; and (3) its focus on different aspects of the research cycle and translational spectrum. For example, clinical and translational researchers who are not primarily CTRI practitioners may need training that has significant depth and application-level orientation with a moderate level of breadth in CTRI, focusing on the particular research areas they will be responsible for in their professional research setting. On the other hand, informaticians who want to specialize as CTRI investigators or researchers may need both a broad and theoretical grounding in the field with a high degree of depth into CTRI areas so as to ensure that they possesses a rigorous, strategic, and methodological understanding of the domain. Finally, policy or decision makers may need a great deal of breadth of understanding of the field, with an equal treatment of theory and applications-level foci, and a low level of depth. All of the

aforementioned scenarios illustrate and continue to argue for highly tailored approaches to the design and delivery of CTRI training based on audience type and composition.

Differentiation of Acculturation vs. Training to Determine Type of Instruction

A related lesson that emerged from the ongoing development of CTRI training programs is that it is important to differentiate among the various CTRI roles and how these roles influence the needs of such individuals for either acculturation or training in the field. This differentiation will influence the type of course that is offered. In such a context, acculturation can be thought of as the process of gaining a "survey" level of understanding of the salient issues surrounding a domain, without gaining the theoretical and/or applied skills necessary to pursue practice or research in that area.

In contrast, training is more concerned with the preparation of individuals to actually pursue practice or research in an area. In the CTRI domain, given the diversity of potential stakeholders, there is a corresponding need for both types of education. For example, principal investigators of clinical or translational research programs may need to be acculturated to understand basic concepts and trends in CTRI so that they can efficiently interact with CTRI professionals, but do not necessarily need to gain a deeper level of understanding of underlying theories and methods. In contrast, individuals in the practitioner or investigator roles, as described in the preceding lesson learned, will need a far greater level of understanding regarding the field, necessitating in-depth training. To generalize, acculturation is a type of training need that can likely be achieved via seminars, workshops, and brief tutorials, while training likely requires formal degrees, coursework, or certificate programs, to name a few of many options.

Need for Alignment with Cross-Cutting and/or Foundational Biomedical Informatics Theories and Methods

Finally, as the maturation of CTRI training and education has progressed, it has become increasingly apparent that such efforts need to more carefully and systematically align competencies and curricula with cross-cutting and/or foundational biomedical informatics theories and methods. It is only though such alignment and harmonization that the emergent CTRI community and its members can benefit from historical and empirically evidenced trends in the broader biomedical informatics community (thus realizing the primary advantage of history, namely the ability to learn from it). For example, CTRI investigators and practitioners who seek to explore how EHR platforms can be leveraged to support/enable clinical trial

recruitment can and should learn from, and apply, the lessons learned as well as basic theories and methods associated with the clinical informatics community's pursuit of advanced clinical decision support and guideline delivery systems. As such, curricula and education/training programs targeting such CTRI focused individuals need to "interweave" such cross-cutting or foundational knowledge into evolving and CTRI-specific competencies and coursework.

Conclusion

The field of CTRI is advancing rapidly, and there is a great and growing need to educate and train a range of personnel in the theories, methods, resources and regulatory and ethical issues unique to the CTRI domain. As an emergent and rapidly evolving sub-discipline of biomedical informatics, CTRI can extend core theories, methods, and historical lessons from the parent field. Because the CTRI workforce is growing at an accelerated rate, both education and training programs need to continue to develop and be evaluated in a similarly rapid manner. The ongoing efforts such as those illustrated above are beginning to address these educational and training needs to address this important area.

Key Take-Away Points

- CTRI researchers and educators should capitalize on the theories, methods and activities in the broader biomedical informatics domain.
- Multiple training and education scenarios exist to satisfy the need of the CTRI workforce, including formal coursework and degrees, as well as more topical workshops, distance education, and certificate programs.
- The delivery of CTRI education and training can and should be tailored to meet the variable needs, stakeholders, and roles incumbent to the clinical and translational science community, which by necessity also requires the differentiation of training versus acculturation to the field.

References

1. Zerhouni E. Medicine. NIH Roadmap Sci. 2003;302(5642):63–72. https://doi.org/10.1126/science.1091867.
2. Zerhouni EA. Clinical research at a crossroads: the NIH roadmap. J Investig Med. 2006;54(4):171–3.
3. Zerhouni EA. Translational and clinical science–time for a new vision. N Engl J Med. 2005;353(15):1621–3. https://doi.org/10.1056/NEJMsb053723.
4. Butte AJ. Medicine. The ultimate model organism. Science. 2008;320(5874):325–7. https://doi.org/10.1126/science.1158343.
5. Embi PJ, Payne PR. Clinical research informatics: challenges, opportunities and definition for an emerging domain. J Am Med Inform Assoc. 2009;16(3):316–27. https://doi.org/10.1197/jamia.M3005.

6. Chung TK, Kukafka R, Johnson SB. Reengineering clinical research with informatics. J Investig Med. 2006;54(6):327–33.
7. Payne PR, Johnson SB, Starren JB, Tilson HH, Dowdy D. Breaking the translational barriers: the value of integrating biomedical informatics and translational research. J Investig Med. 2005;53(4):192–200.
8. Sung NS, Crowley WF Jr, Genel M, Salber P, Sandy L, Sherwood LM, Johnson SB, Catanese V, Tilson H, Getz K, Larson EL, Scheinberg D, Reece EA, Slavkin H, Dobs A, Grebb J, Martinez RA, Korn A, Rimoin D. Central challenges facing the national clinical research enterprise. JAMA. 2003;289(10):1278–87.
9. Pate C, Chin D, Lutz S, Beard N, Isgur B, Loring J, Carlos RR. Research rewired. Merging care and research information to improve knowledge discovery. PricewaterhouseCoopers Health Research Institute; 2008.
10. Casey K, Elwell K, Friedman J, Gibbons D, Goggin M, Leshan T, Stoddard R, Tate D, Vick P, Vincent J. A broken pipeline: flat funding of the NIH puts a generation of science at risk. National Institutes of Health: Bethesda; 2008.
11. Hersh W, Williamson J. Educating 10,000 informaticians by 2010: the AMIA 10×10 program. Int J Med Inform. 2007;76:377–82.
12. Butte AJ. Translational bioinformatics: coming of age. J Am Med Inform Assoc. 2008;15(6):709–14. https://doi.org/10.1197/jamia.M2824.
13. Butte AJ. Translational bioinformatics applications in genome medicine. Genome Med. 2009;1(6):64. https://doi.org/10.1186/gm64.
14. Maojo V, Garcia-Remesal M, Billhardt H, Alonso-Calvo R, Perez-Rey D, Martin-Sanchez F. Designing new methodologies for integrating biomedical information in clinical trials. Methods Inf Med. 2006;45(2):180–5. https://doi.org/10.1267/meth06020180.
15. Ash JS, Anderson NR, Tarczy-Hornoch P. People and organizational issues in research systems implementation. J Am Med Inform Assoc. 2008;15(3):283–9. https://doi.org/10.1197/jamia.M2582.
16. Payne PR, Mendonca EA, Johnson SB, Starren JB. Conceptual knowledge acquisition in biomedicine: a methodological review. J Biomed Inform. 2007;40(5):582–602. https://doi.org/10.1016/j.jbi.2007.03.005.
17. Richesson RL, Krischer J. Data standards in clinical research: gaps, overlaps, challenges and future directions. J Am Med Inform Assoc. 2007;14(6):687–96. https://doi.org/10.1197/jamia.M2470.
18. Ruttenberg A, Clark T, Bug W, Samwald M, Bodenreider O, Chen H, Doherty D, Forsberg K, Gao Y, Kashyap V, Kinoshita J, Luciano J, Marshall MS, Ogbuji C, Rees J, Stephens S, Wong GT, Wu E, Zaccagnini D, Hongsermeier T, Neumann E, Herman I, Cheung KH. Advancing translational research with the semantic web. BMC Bioinformatics. 2007;(Suppl 3):S2. https://doi.org/10.1186/1471-2105-8-s3-s2.
19. van Bemmel JH, van Mulligen EM, Mons B, van Wijk M, Kors JA, van der Lei J. Databases for knowledge discovery. Examples from biomedicine and health care. Int J Med Inform. 2006;75(3–4):257–67. https://doi.org/10.1016/j.ijmedinf.2005.08.012.
20. Oster S, Langella S, Hastings S, Ervin D, Madduri R, Phillips J, Kurc T, Siebenlist F, Covitz P, Shanbhag K, Foster I, Saltz J. caGrid 1.0: an enterprise grid infrastructure for biomedical research. J Am Med Inform Assoc. 2008;15(2):138–49. https://doi.org/10.1197/jamia.M2522.