MEDINFO 2001
V. Patel et al. (Eds)
Amsterdam: IOS Press

The Cancer Informatics Infrastructure (CII): An Architecture for Translating Clinical Research into Patient Care

John S. Silva[a], Marion J. Ball[b], Judith V. Douglas[c]

[a] *Director, Office of Informatics, National Cancer Institute, Bethesda, MD, USA*
[b] *Adjunct Professor, Johns Hopkins University School of Nursing, Baltimore, MD, USA*
[c] *Adjunct Faculty, Johns Hopkins University School of Nursing, Baltimore, MD, USA*

Abstract

Today, the clinical trial process remains slow and paper-based. The creation of a Cancer Informatics Infrastructure (CII) can provide the architectural base across the continuum of cancer research and cancer care. Recommendations of a Long Range Planning Committee identified near-term activities for the Office of Informatics at the National Cancer Institute (NCI). These include participating in national standards development; fostering oncology-related terminology and standards, e.g., Common Data Elements (CDEs); and leveraging mainstream informatics and Internet technologies, using the successful Internet model that focuses on facilitating stakeholder participation, sponsoring the CII rather than subsidizing it, and providing a test bed as well as an infrastructure. Diffusion tactics include extending the CII concept beyond its "early adopters" to the wider community through recommendations for the near-term and development of a major document defining next-phase activities.

Keywords:

Cancer Care; Cancer Research; Knowledge Environment; Clinical Trials; Cancer Informatics Infrastructure (CII)

Introduction

Today, despite the accelerating rate of drug development, clinical trials remain slow. Clogged by excessive paperwork and slowed by redundant processes, clinical trials are burdensome for physicians, patients, researchers, and other stakeholders. Information exists in multiple forms in multiple locations, making it difficult to conduct and assess clinical trials activity. Far too many anecdotes document the "hassle factors" posed by clinical trials, and are validated by staggering statistics on the paper generated over the clinical trial lifecycle. Far more serious are the impact these inefficiencies have upon the translation of clinical research into patient care. A cancer informatics infrastructure (CII) is needed to support the practice of evidence-based medicine, the evolution of the science that underlies cancer treatment, and the new consumerism in healthcare.

Methods and Materials

In July 1998, national experts convened in the first of four plenary sessions to advise the Director of the Office of Informatics on how to take advantage of existing and emerging technologies to support the work of the National Cancer Institute. Supported by the work of a smaller "tiger team," the Long Range Planning Committee (LRPC) completed its report in March 2000. The report, *Translating Cancer Research into Cancer Care* [1], urged the creation of a Cancer Informatics Infrastructure (CII) to expedite clinical trials and ultimately the full range of cancer research both inside and outside NCI. Since that date, work has continued to move the report's recommendations forward and thereby facilitate the translation of cancer research into cancer care.

Results

Conceptual Model

The CII concept provides an architectural base for moving information across the continuum of cancer research: basic, clinical, translational, and population-based research. It is a model that brings together

- the mission, processes, and work culture of cancer research, care, and policy
- the stakeholders and "consumers" of research and care
- the tools and technologies (including hardware and software) that people use to do the work.

It is based on the premise that the NCI should build only those capabilities that are specific to its needs: common data elements, research building blocks, and tools to support the conduct of cancer related research. Although the CII will first be implemented within the context of NCI clinical treatment and diagnostic trials, over time it will have broad application to NCI's extensive research mission from basic to applied observational, population-based research.

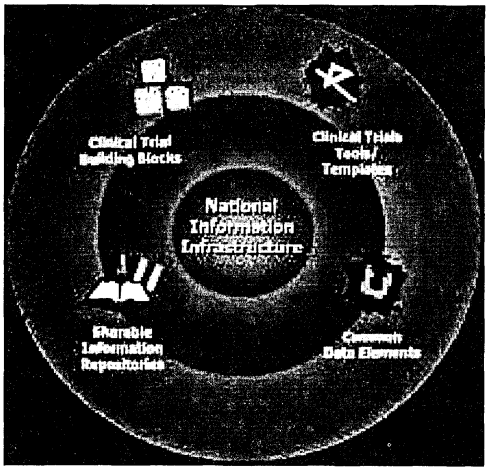

Figure 1 – Cancer Informatics Infrastructure Model

Although the specifics of how it might be applied to those diverse settings will need to be addressed with specific efforts, the model stresses interoperability among technologies and collaboration among communities to develop and share relevant knowledge about cancer, as shown in Figure 1. For NCI, this means that the CII can provide totally new ways to collaborate, such as the linkages among basic biologists, mouse researchers, genomic researchers and clinicians studying human cancers in the Mouse Models of Human Cancer Consortium. For organizations outside NCI, this means that the CII can provide links among local systems that were heretofore incompatible.

The model maps to a future-states vision for 2004, wherein the CII translates clinical trials results into clinical care, and care results drive future research. Common processes and tools expedite information exchange, and access to information supports all stakeholders – patients and physicians, investigators, trial managers, and payers – as they make vital decisions affecting the course of cancer treatment and research [2]. In this future state, the benefits accrue to the individual patient while simultaneously improving the standard of cancer care for all [3].

More specifically, the information and knowledge provided by the CII will function throughout the clinical trial lifecycle shown in Figure 2. Simply put, the CII will make it possible for *patients and their physicians* to access up-to-date medical information, maintain patient-centric records, and be partners in shared decision making. At the same time, *investigators* will be newly able to design and obtain approval for a trial in 60 days, rapidly accrue patients into trials, and populate research databases using clinical data. *Trial managers* will have the capability to minimize time from scientific concept to first patient accrual, maximize patient participation in cancer clinical trials, exchange information to optimize effective studies, and facilitate translation of clinical trial results into cancer care. One last key stakeholder group, *payers*, will be able to provide high quality cancer treatment for their members due to their newfound ability to make a solid business case for participating in clinical trials.

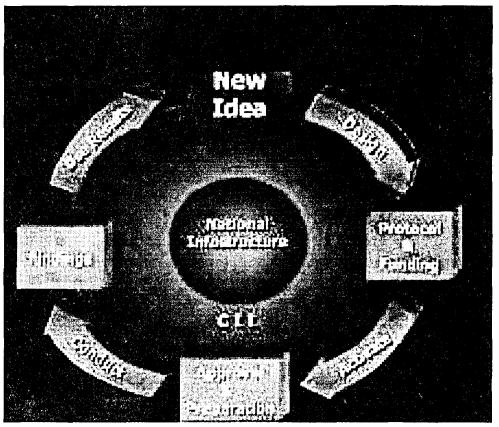

Figure 2 – Clinical Trial Lifecycle

Discussion

The realization of these capabilities depends upon the CII. In turn, creation of the CII depends upon progress in four key areas, each of which was targeted by the Long Range Planning Committee for specific activities in the near term:

- Formulate the role of the National Cancer Institute in the national standards development process
- Convene a national advisory meeting on oncology-related terminology and standards
- Focus informatics efforts on demonstration and evaluation projects that enhance NCI's ability to carry out its mission, by building on ongoing mainstream informatics initiatives and Internet technologies
- Develop a process to strategically and tactically diffuse the product and concepts from the above three areas and activities throughout the cancer community.

National Standards Development

Initiatives are underway at NCI seeking consensus from relevant stakeholders on proposed information standards unique to the CII enterprise. These acknowledge the need to coordinate oncology-specific standards with broader-based efforts, including those conducted by standards development organizations (SDOs) such as HL7 and SNOMED, umbrella organizations such as ANSI HISB and ISO TC 215, and larger communities of interest such as the Food and Drug Administration (FDA), the pharmaceutical industry, the National Library of Medicine's Unified Medical Language System (UMLS), and most importantly practicing scientists and clinicians.

Such efforts require a dynamic approach. Content management is critical given the need to continuously revise common data elements to reflect changes in the science. Configuration management – for example, appropriate content maintenance, version control, and seamless integration of updates – is also essential to make the implementation of standards as "transparent" and effortless as possible for the entire cancer community. Appropriate linkages and participation can assist in supporting oncology-relevant standards, and NCI has an important role to play here.

Oncology-Related Terminology and Standards

Implementation of the CII requires commonality to be developed and disseminated across the cancer community. Here the NCI has launched two key initiatives: Common Toxicity Criteria (CTCs), now web-based, and Common Data Elements (CDEs), now being developed. Both the CTCs and CDEs are cancer-specific standards, unique to NCI. As such, they must be distinguished from broader-based, oncology-relevant standards set by SNOMED, HL7, XML and the like. This distinction is critical. NCI's role vis-à-vis broad-based standards requires linkages with and participation in the national standards development process, whereas NCI's responsibility for cancer-specific standards will no doubt be much more intensive and as such must be defined across the entire lifecycle.

Formal change management processes are essential to CDEs and other oncology-related terminology. The adoption and implementation of newly developed CDEs will result in increased needs for monitoring adherence and providing technical assistance to support all stakeholders. Processes must ensure that terminological standards evolve in parallel with and support of clinical and research needs, including those of the pharmaceutical industry and the FDA. This requires developing guidelines and resource materials that formalize best practices for CDE development and eliciting clinical input from designated "champions" early in each round of terminological development. Work done on CDEs in spiral CT for lung cancer provides a model for such efforts [4].

To succeed, terminologies must develop associated meta-knowledge and meta-data about each term to provide linkages between terms, logical contexts for terms, and specifications on instantiating and using terms to clarify interrelationships. The CDE database schema and information model must be described in detail to facilitate comparison to other models and terminologies based on meta-data standards. (Such public review is essential to efforts to harmonize meta-data repositories.) Building on existing CDE "categories" and employing knowledge representation techniques like semantic networks and description logic will ultimately result in a rich, consistent information model for CDEs. In like manner, creating detailed data dictionary entries for new data elements, including non-textual data (e.g., imaging), will minimize variations.

In the effort to establish CDEs as a *de facto* standard for oncology data collection, NCI now makes them available for free on its website. To further encourage the dissemination and use of CDEs (and other CII technology), new functionalities can be provided that

- Enhance the web-enabled interface to CDEs, making it easier for new users and users outside the cooperative groups
- Automate the connection between the CDE web site and research systems using CDEs, by encouraging projects at different phases of the cancer lifecycle to download and incorporate CDEs into data collection systems for clinical trials, possibly through funding supplements to existing trials
- Enhance download formats, including case report form (CRF) templates, draft database designs, and XML, enabling users to search for and download relevant CDEs and evolving the CDE resource into a meta-data repository.

Mainstream Informatics and Internet Technologies

Implementing the CII is a complex and long-term task, but most of the technologies and applications required to support it are available now. Creation of the CII requires a set of common infrastructure services, such as medical informatics standards and tools, digital libraries, collaboration tools, security services, and electronic transaction support. For maximum impact, the CII must exploit existing and emerging technologies and capitalize on initiatives now underway both inside and outside NCI.

As the principal stakeholder for the long-term interests of the cancer-trials community, NCI is positioned to emphasize investments that help evolve standards or scale-up deployment of the CII. When new standards are proposed, such as common data elements or CDEs, the NCI should, from the outset, assure that the designs allow smooth transitions to exploit anticipated developments in health care and in information technology. In special instances, investments can target generic information technologies that play a critical role in NCI infrastructure development. (It may be appropriate to co-manage these investments with other agencies that serve in a more primary role in technology-development.)

A second key role for NCI is in buying down the risks of creating and adopting new technologies. For example, NCI should make targeted investments to assist early adopters in evaluating CII technologies or to assess how new standards and processes introduce or eliminate barriers to efficient trials management and broad participation.

For the CII to succeed, specific investments need to address compelling near-term needs, such as "bootstrapping" new efforts in the development of standards and technology. These investments must be made in a manner that is consistent with the long-term CII vision. The focus on near-term needs grounds the CII in the baseline of present practice and potentially entails adaptations to the CII vision. By getting involved at very early stages, NCI can exert greater leverage with its investment and ensure that their

informatics investments are consistent with overall strategy, although the need to explore diverse approaches to particular problems may remain.

Leveraging ongoing informatics and Internet technology efforts will maximize NCI's return on investment while making it possible to address NCI-specific needs. For example, CDE standards developed under NCI sponsorship should be integrated into mainstream framework efforts such as HL7.

Mediation of community standards-development efforts is no easy task, but the Internet Engineering Task Force (IETF) offers a proven model for national initiatives like as the CII. The governing body for Internet standards since the early 1970s, IETF has succeeded in building a national-scale community process to support an evolving collection of standards and capabilities.

In the rapidly evolving environment of cancer research and treatment, as in the world of the Internet, the only constant can be a set of principles that make up the process model. Within the IETF model, these include

- Providing mechanisms to facilitate stakeholder participation
- Leveraging sponsorship rather than subsidizing the entire CII
- Providing both a test bed and an infrastructure.

This latter item is especially important, as it provides for the development of technology prototypes that can be directly evaluated as candidate approaches.

Standards Development

The CII can benefit from ongoing efforts in other sectors. Collaboration with Radiology and Pathology can ensure that their ongoing efforts to create digital libraries for large-scale multimedia records are compatible with the CII. In like manner, the CII can leverage work done on the Guidelines Interchange Format (GLIF) to involve multiple stakeholders in developing suites of building blocks. In addition, the CII can exploit e-commerce, where emerging business models support electronic transactions between parties. Standard practices in e-commerce, notably business-to-business applications, have brought multiple legacy systems together. Similar mechanisms will enable basic researchers to collaborate with clinical researchers and result in the more effective use and re-use of knowledge in their own legacy systems. They will also allow for linkages with individuals and entities outside NCI, from patients to ancillary care providers.

Technology Development

Like standards development, technology development needs to capitalize on work done by other federal agencies in the areas of scientific collaboration and research, including the Department of Energy, National Science Foundation, and Defense Applied Research Projects Agency. National initiatives are underway to provide security services needed to protect patient privacy and confidentiality. Work by the Computer-Based Patient Record Institute (CPRI) and in conjunction with the Health Insurance Portability and Accountability Act (HIPAA) addresses the policy and technology issues critical to the CII and the patient-centric data it will include. The CII should leverage this work rather than develop services independently.

Diffusion throughout the Cancer Community

Although the concept of CII has its supporters, these visionaries and early adopters do not constitute a majority in the cancer community. Thus, a change-oriented diffusion process is needed to illustrate the impact of the CII to "early majority pragmatists."

The three target areas—national standards development, cancer-related standards, and mainstream informatics and Internet technology—contain a number of specific projects and actions that are highly relevant to the current user community, e.g., Centers, groups, NCI, etc. The products developed here by "early adopters" can be disseminated outward to their more conservative colleagues. Inclusion of the CII in NCI announcements and project work will assist in diffusion efforts.

Conclusion

A major document advancing the CII concept will define next-phase activities and critical issues acknowledged by the Long Range Planning Committee in its report. This document will widen the diffusion process, by virtue of its wide-ranging authorship and its publication within an internationally recognized informatics series [5].

These efforts on behalf of the CII will strengthen the role of the National Cancer Institute as a leader and "heat seeker" in the area of clinical cancer trials.

References

[1] Long Range Planning Committee (LRPC). Final Report: *Translating Cancer Research into Cancer Care.* Washington DC: National Cancer Institute, March 2000.

[2] LRPC. Appendix 2: summaries and scenarios. In: *Translating Cancer Research into Cancer Care*; pp. 21-36.

[3] Silva JS. Fighting cancer in the information age: an architecture for national scale clinical trials. *MD Computing* 1999 (May/June): 16(3), 43-4.

[4] Langlotz CP. Appendix 3: CDE development model-spiral CT in lung cancer. In: *Translating Cancer Research into Cancer Care*; pp. 37-45.

[5] Silva JS, Ball MJ, Chute CG, Langlotz CP, Niland J, Scherlis WL, Douglas JV, eds. *Cancer Informatics.* New York: Springer-Verlag, in press.

Address for correspondence

John S.. Silva, M.D..
Jc-silva-md@worldnet.att.net