Achievable Steps Toward Building a National Health Information Infrastructure in the United States

William W. Stead, MD, Brian J. Kelly, MD, Robert M. Kolodner, MD

Abstract Consensus is growing that a health care information and communication infrastructure is one key to fixing the crisis in the United States in health care quality, cost, and access. The National Health Information Infrastructure (NHII) is an initiative of the Department of Health and Human Services receiving bipartisan support. There are many possible courses toward its objective. Decision makers need to reflect carefully on which approaches are likely to work on a large enough scale to have the intended beneficial national impacts and which are better left to smaller projects within the boundaries of health care organizations. This report provides a primer for use by informatics professionals as they explain aspects of that dividing line to policy makers and to health care leaders and front-line providers. It then identifies short-term, intermediate, and long-term steps that might be taken by the NHII initiative.

Received for publication: 09/03/04; accepted for publication: 11/05/04.

The Institute of Medicine (IOM) has raised the bar on assessing and improving the quality of health care and health in the United States. The landmark 2001 IOM report described “the chasm between what we know is good quality care and what the norm is in practice” and laid out a compelling vision that would “increase the likelihood of desired health options consistent with current professional knowledge.” This emphasis on quality extended the focus beyond that of the previous IOM report, which concentrated on the unacceptable rate of medical error and its significant impact on patient safety, to highlight improvements that could have even broader effects on improving health. At the same time, escalating costs of health care, along with the increasing number of people without access to it, are significant challenges facing the policy, medical, and business communities.

How does the United States resolve these health care issues? There is a growing advocacy for a systematic approach to better managing clinical data as a critical foundation leading toward more cost-effective health care. The National Health Information Infrastructure (NHII) is an initiative in progress with the goal of linking information into workflow to help individuals and their health teams make better health and clinical decisions so that they do the right thing and only the right thing. Properly executed, such a comprehensive, knowledge-based network could be the key to engaging patients in their care, providing care when and where needed, supporting processes that avoid error, and reducing administrative cost.

Clearly, health care is an information business, and it needs an appropriate information and communication infrastructure. A pure technology solution, however, will not be enough, and it is even possible for today’s health care information systems to cause as many problems as they cure. Rather, success will require a careful selection of elements that will actually help the user in each practice situation. Some parts are best left undone until the technology gets better. Equally important, standards are needed to foster a common language so that systems can effectively exchange information. At the same time, standards are challenging because practice models are evolving and biologic diversity exceeds today’s information models.

Our goal, therefore, is to foster a better understanding of which approaches are likely to work on a large enough scale to have the intended beneficial national impacts and which approaches are better left to smaller projects within the boundaries of health care organizations. This report is intended as a primer for use by informatics professionals as they explain aspects of that dividing line to policy makers and to health care leaders and front-line providers. It then goes on to suggest how the NHII initiative might put the building blocks in place to support development of a health care information and communication infrastructure over time. This report was presented in part as the keynote for the architecture track of a 2003 consensus conference on NHII sponsored by the Department of Health and Human Services. It was expanded through discussion at a 2004 workshop for U.S. Senate Staff at the Vanderbilt Center for Better Health.

Our focus is on the pragmatic steps that might be taken at a national level to move forward with the development and implementation of a health care information and communication infrastructure.

- We begin by defining five interrelated components of the health care information and communication infrastructure.

People often use these terms interchangeably, and clarity about their differences and overlap is important when discussing these topics.

- Next, we attempt to explain some of the challenges the United States faces in the NHII initiative and discuss approaches that might work on a national basis. We suggest, for example, that the NHII will sit beside today’s electronic medical record systems, leverage their content, and make it easier to share information among these systems and stakeholders.

- We conclude by outlining the potential building blocks that the NHII initiative can put in place in the short term, together with those that can be accomplished over a longer period. Where possible, we suggest assignment of responsibility.

Components of the Health Care Information and Communication Infrastructure

Healthcare data are complex and used by various stakeholders for different purposes. The National Committee for Vital and Health Statistics defines three primary dimensions through which healthcare information can be viewed:

- The patient view or Personal Health Dimension,
- The Health Care Provider Dimension,
- The community or Population Health Dimension.10

As illustrated in Figure 1, often the same information is used for multiple purposes and by different stakeholders. While we recognize that considerable healthcare data are currently stored on paper or other media, we focus our discussion on electronic data only. Five key components will serve as building blocks for a framework for healthcare information and communication infrastructure:

- Electronic Medical Record Systems (EMRSs)
- Electronic Health Records (EHRs)
- Personal Health Records (PHRs)
- Standards
- Data Interchange Capabilities

Electronic Medical Record System

As a result of the volume and diversity of healthcare data, information systems were developed and used to better manage healthcare information and improve healthcare quality and safety since the early 1970s. These pioneering efforts led to many of today’s EMRs.

An EMR automates aspects of clinical practice, such as placing a care provider order, recording a clinical note, or capturing administrative functions such as scheduling and billing. A patient’s electronic medical record (EMR) is generated as a by-product of these clinical and administrative functions. It often lives within the specific EMR that created it and is unique to that system. In that case, the EMR’s meaning is clear only to that specific EMR, since the record is constructed with terminology and data structures particular to that system.

Electronic Health Record

In contrast, we use the term Electronic Health Record (EHR) to refer to any information in electronic form about a person that is needed to manage and improve their health or the health of the population of which they are a part. An EHR is a superset of an EMR and totally includes it. To meet this vision, an EHR might collect information as appropriate from across the healthcare system (i.e., from several EMRs) and a variety of personal information sources. The EHR exists outside of any particular EMR and, therefore, has a consistent meaning.

Personal Health Record

The term Personal Health Record (PHR) refers to a personal electronic collection of information. The PHR might include the patient’s own record of their progress and changes they made in therapy plus their electronic copies of information from their providers. They might also decide what parts to share with their providers or health care plan.

As illustrated in Figure 2, there often is significant overlap in the content and functionalities of these records and record systems. There are equally significant differences in purpose and ownership to consider as decisions are made about the NHII strategy.

Standards

We use the term standards very broadly to refer to agreements about how to do something where coordinated action is needed. Such agreements allow for choice where variety or flexibility is important. For example, information systems with very different user interfaces may exchange information with each other if they capture data with the same terminology. In other words, standards are the middle ground. They provide more freedom than requiring everyone to use the same system or process. They are more restrictive than those that allow everyone to do things their own way. To date, standards development efforts have focused on messaging standards and terminologies:

- Messaging standards provide a standard format for exchange of a transaction, such as an order, among systems.
- Terminology standards provide a standard vocabulary for representing a concept, such as a specific orderable item, for use across systems.

Figure 1. The three dimensions of health care data.10
To meet the objectives of the NHII initiative, the United States will need additional standards, including ones dealing with authentication and security, health care delivery processes, representations of medical knowledge that are interpretable by a computer, and interfaces to software components.

**Data Interchange Capabilities**

Finally, data interchange refers to the service of exchanging information among sites and systems. Since the United States does not have a national patient identifier, a data interchange begins with a mechanism—such as a master patient index or its equivalent function—to determine which records relate to a single person. Such a mechanism allows for the display of information at one site obtained from another site or the aggregation of certain data such as laboratory results or medications in a central database for display at any site. The former is fairly straightforward, but the latter requires implementation of messaging and terminology standards and agreement on additional technical standards for exchanging information.

In either case, a mechanism is needed to authorize access to data by a person at a site other than the place it originated. Such a service, therefore, is as much about governance, trust, and common terminology as it is about technology.

**Challenges to be Addressed through the NHII Strategy**

The NHII initiative faces challenges greater than those faced by any prior infrastructure initiative because of the complexity of health care processes, patient data, and biomedical knowledge. The strategies for the NHII need to be crafted to overcome these challenges.

**The Challenge of Scope and Timing**

If the NHII is to achieve its goals, it must support multiple dimensions, namely, personal health, health care provider, and population health. While it would be national in scale, its implementation must support a health care culture that values local control and individual choice. These two assumptions suggest that a single, massive, national information system will not accomplish everything the United States expects from the NHII. Even if one system could provide the needed functionality and adaptability, it is not practical to construct and maintain such a system given the state of the art of the technology as described below in the section on hardware-software development mismatch.

Accordingly, much of the implementation of the NHII will be local. But some form of a national level framework is essential to coordinate local work, so that the pieces work together to achieve the larger goal. Only when a national framework of governance, policy, health care processes, and architecture are in place, can the United States deploy and implement the technologies and systems to support the high-quality, cost-effective health care system envisioned in the IOM reports. It is beyond the scope of this report to explore each of the legs of such a framework. However, every decision will be a balance between control and adaptability, as shown below in the section on appropriate use of standards.

The NHII strategy must reflect both short- and long-term time horizons. The combination of quality, access, and cost problems requires progress in three to five years. While this time frame is too short to achieve the vision of a fully functional health care information and communication infrastructure, many gains are possible. The United States needs an approach in which early steps provide benefit while establishing a foundation for subsequent ones. Smaller steps also provide for trial and error with a minimum of extra costs or delays. If each step provides benefit and learning, the NHII will ultimately deliver on the vision.

Finally, the NHII will be built in an environment of constant change: The explosion in biomedical knowledge is accelerating. At the current rate of publication, a person who finishes school knowing everything, and reads two articles every night, becomes 1,225 years behind at the end of the first year. At the same time, the nature of health care policy and practice is expected to change as the United States figures out how to cross the quality chasm. These sea changes will be dwarfed by the transformation in practice that will occur when it becomes commonplace to individualize risk assessment and treatment based on genomics. The strategy for the NHII needs to accommodate this rate of change. Large fixed systems will not be able to adapt and support this dynamic environment.

**The Challenge of the Hardware–Software Development Mismatch**

Many people think that it is easy to solve difficult problems with computer solutions. This perception stems from the rapid improvement we continue to see in the performance of computer hardware. Processing power doubles approximately every 18 months. We see similar or greater increases in storage capacity and in network bandwidth. It is now

*The journals indexed by the National Library of Medicine publish over 550,000 articles/year. If we assume this accounts for 60% of what is published, the biomedical literature grows by over 900,000 articles per year.*
possible to assemble a high performance super-computer for less than $100,000, clustering inexpensive personal computers with commodity network components.

In contrast, software development productivity has been increasing at only 5% per year. The traditional approach to development of a health information system involves extensive application programming. An analyst documents, for example, each step in a practice process, such as making an appointment or recording the elements of a diagnosis. A programmer then represents the practice process in thousands of lines of computer code. Poorly designed software lacks the flexibility to be configured to fit different, better work processes. It can even force people to adapt their processes to the application or prevent them from evolving to a more effective practice without having to rewrite the entire application.

The issue for a health care information and communication infrastructure is that despite advances in designing and writing software, large health care application programming projects remain problematic. The Chaos studies performed by the Standish group have shown that the majority of software projects (all industries) are over budget and delivered late.

Specifically, software applications often become a bottleneck in a health care organization trying to rapidly improve its processes, large ongoing investment notwithstanding.

This mismatch between the performance of hardware and software developers (Figure 3) suggests that a different model is needed for creating an information and communication infrastructure that will ultimately improve health care. The question becomes how to leverage the raw power of hardware and general purpose software to exploit informatics techniques for:

- communication—accessing information wherever it is
- data processing—finding information related to a question regardless of source
- visualization—presenting information in ways that ease human assimilation

Assuming there are proper protections on privacy in place, the Internet might provide a better model for the health information and communication infrastructure than commercial enterprise resource planning systems.

The Challenge of Appropriate Use of Standards

As for standards, each may take one of two forms: prescriptive and reference. Prescriptive standards grew out of the physical world where something fits or does not. The gauge of a train track or an electrical plug, for example, specifies exactly what must be done in a designated situation to the level needed to achieve a fit. They are appropriate where the need for compatibility is greater than the need for flexibility, and they can lay the foundation for explosive growth in the productive use of a technology.

One clear example of the success of prescriptive standards is the World Wide Web. The Internet was used primarily by government and academia from the early 1970s to 1993 when Tim Berners-Lee invented three simple standards—URL, HTTP, and HTML. These standards are prescriptive. The URL defines the address on the network. HTTP defines how to transport the information on the network. HTML specifies how to display material. Although these standards are prescriptive, they specify as little as possible. They say absolutely nothing about the material to be displayed, only how it is displayed. From this combination of limited prescriptive standards with freedom of content arose the World Wide Web, browsers, and the explosive growth and use of the Internet for personal and business use throughout the world, transforming global communications and commerce in a few short years.

Contrast that success with two examples of the inhibiting and potentially disastrous effects that a lack of standards can have. First, the lack of a common standard for fire hydrants prevented firefighters in nearby cities from using their fire hoses to help Baltimore during the devastating fire in 1904 that destroyed almost 2,500 buildings. Second, the absence of a standard for high-definition television signals delayed the release and adoption of such a technology in the United States for 10 years until a standard was adopted in 1996.

Prescriptive standards, however, are good only if they can, in fact, be implemented. Prescriptive standards can slow down adoption of a technology if they cover too much detail. Standards that are developed to support one type of business process may not optimally support others. A clinical note system that yields universally standardized medical data, for example, would be wonderful from a data retrieval perspective, but may be so slow to use that health care workers cannot accomplish their daily clinical tasks.

As for reference standards, this is a new form of standard made possible by computers, which allows them to speak a common language, thus facilitating information exchange. A single framework of information is created allowing multiple levels of specificity. "Tagging" data to that framework provides a common language and also specifies the format in which the data will be expressed: a text document, an Excel spreadsheet, an image format such as TIF. Each tag then describes a "category of interest," which in the example below, is "body part," interspersed with the actual name of the data element, which in the example is the word "limb."
A properly constructed framework of information can evolve over time as knowledge advances. Continuing our example, the framework could expand to include categories of interest for limb type and body side. (see below).

One can begin to use a reference standard as soon as the major categories at the top of the framework are understood. No meaning is lost because the raw data are retained in addition to the tag. Precision increases over time as the framework is fleshed out.

The Challenges for Medical Informatics

In addition to prescriptive and reference standards, there are informatics tools to help facilitate data exchange by bridging among various vocabularies and terminologies. The National Library of Medicine’s Unified Medical Language System (UMLS) is the best example of such a tool. Instead of trying to build yet another standard vocabulary, the UMLS Metathesaurus brings together more than 3 million terms representing 1 million distinct concepts from more than 100 source vocabularies. The relationship among terms and concepts is explicitly mapped, creating relationships among the terms that do not exist in any of the source vocabularies.

The challenge for medical informatics, therefore, is to provide information structures and tools to enable adoption of “good enough” standards soon enough so that the NHII can rapidly grow and expand while accommodating change and variability. Spreading a technology before the standards are established could lead to advances that do not achieve the full benefits for society that might otherwise be attained. Waiting is equally unacceptable.

On the other hand, like the pervasive changes wrought by the Internet after the establishment of just three standards by Tim Berners-Lee, a transformation of health care can occur if the country establishes minimum standards for the generation, capture, manipulation, and exchange of health-related information and the governance and informatics tools to support their evolution.

A Careful Course to a National Health Information Infrastructure

A national health information infrastructure, in practice, will serve more as an enabler that facilitates the convergence of information with security and privacy, rather than an information system in itself, or even a federation of information systems.

The development of the Interstate Highway System may provide the best analogy. The master plan defined the ultimate reach of the system, features such as limited on and off ramps, adequate capacity, and funding. It did not specify the type of vehicle, its fuel, its purpose, or the content it might carry.

Similarly, to move forward rapidly with a national health information infrastructure, its scope should be limited to things that must be done collectively to permit information sharing. Initially, then, it might have four key functions:

1. First, providing standards to make the meaning of content in local systems increasingly explicit over time, reducing the work and cost of implementing local systems as a by-product;
2. Second, facilitating direct access to information external to local systems if permitted by patient consent or mandated by public health law;
3. Third, providing practice guidelines and other decision support information in a form that is easy to incorporate into local systems;
4. Finally, steering governance of the NHII and the sharing of information among systems through development of guidelines and policies.

The Interstate Highway System is, once again, a good analogy. It reused toll roads that had already been built close to the new standard where practical. It largely left local roads intact. This latter strategy provided the “last mile,” allowing early sections to be used as soon as they opened, while also providing the fallback for emergency bypass during accidents. As widely acceptable rules for how to drive on the Interstate system and methods of punishing noncompliance were developed and enforced, local road design and laws were influenced over time.

As the NHII evolves, it will lead to a decoupling of the management of the information we want to share about patients from the local systems that automate practice in hospitals and clinics. StarChart, Vanderbilt University Medical Center’s EMR, is a working local example of how this might be accomplished. At Vanderbilt, all sources of patient data, systems such as Admitting or Laboratory, note capture tools or transcription, and image capture such as Picture Archiving and Communication, externalize the information either as a text report with a standard header or as that same header with a pointer to an image. This information is assembled by StarChart into an EMR that is directly accessible by browser or application programming interfaces as appropriate.

Time Line for Evolution

It is easier to grasp the scope of the NHII if it is seen in terms of building blocks layered over time. The early building blocks can leverage current technology and know-how; later ones will require advances in basic computer science and informatics research.

Near Term

The pragmatic steps, which can be taken immediately to accelerate the building of the NHII, deal with policy and governance issues as well as standardization and data interchange approaches.

Governance and Policy

Successful health care EMRS implementations have safeguards in place to effectively and efficiently allow appropriate health care providers access to patient health care information while ensuring that a patient’s health care data will only be used in accordance with pre-established policies. These policies are currently set by the organization implementing the system and business associates bound to them by contract. As we move toward sharing information on a
regional or national basis, new governance mechanisms and policies are needed.

The general view is that this is a solvable problem, as demonstrated by the successful sharing of information among multiple hospitals and providers across regions. Lessons learned from these early regional efforts such as those in Santa Barbara, Indianapolis, and the Department of Veterans Affairs and the Department of Defense could serve as a roadmap that can be exploited to form the basis for national and private sector governance policies.

For now, EMRS implementation to private physician practices has been slowed by the cost of the technology, because it often is the payers who receive greater benefit from such a system than the providers. This barrier is most pronounced for small practices. Increased incentives for providers, however, combined with the creation of safe harbor legislation to allow health care institutions to offer information technology solutions to providers, should help ameliorate this current situation.

Standardization and Data Interchange

To facilitate the efficient interchange of data in the future, there are five areas of standardization that should be finished quickly. They will provide immediate benefit.

First, rather than seek unanimous agreement on what to call various drugs and medical products—or map their local names to a single standard—they could be identified with a computer readable tag, according to a reference standard at the point of manufacture. If tagged with their “information content,” each new local system could simply use that tag instead of creating its own local list of identifiers. With minimal modification, existing systems could pass such information along to the next system, even when unable to use it internally. In time, meaning would be consistent across systems. This approach is analogous to the one used by the grocery industry to handle inventory and check out. Applied to medical products:

- Drugs. Packaged drug products are currently identified by a National Drug Code (NDC), which identifies the manufacturer, the product, and the package size. The first is assigned by the FDA; the other two by the manufacturer. Currently, each combination of manufacturer, drug substance, strength, dosage form, formulation, and package size has a unique NDC, making it impossible to determine the drug, dosage form, and strength for clinical decision support. Having that information explicit in the identifier, or providing a database linking that information explicitly to the NDC, could solve the problem. RxNorm and the Veterans Administration’s NDF-RT are steps in this direction.
- Laboratory results: The Logical Observation Identifiers Names and Codes (LOINC) standard is used increasingly to identify names of laboratory results. Each laboratory, however, now decides what to put in each of the elements of a LOINC tag. For example, the standard says which element contains the analytic method; the laboratory in the mapping process selects the term to describe their method. If the auto-analyzers prepopulate such elements with a standard term at the time of manufacture, the problem might be eliminated.

Second, payers should provide explicit computer interpretable reimbursement logic. Currently, contract terms are manual and subject to interpretation. Both providers and payers have teams who decide how to translate them into computer logic and other teams that work the problems when the resulting systems disagree. As local systems begin to use payer rules to edit transactions on the front end, it might become practical to pay on billing, freeing up both provider and payer staff from transaction processing for quality improvement work.

Third, the National Library of Medicine (NLM) should be asked to serve as the entity for coordinating and distributing computer readable terminologies and information standards. The success with SNOMED CT paves the way for additional national licenses, and the UMLS shows how the various sources may be coordinated. These efforts should be expanded. In addition, funds provided to the NLM could also be used to develop tools to ease incorporation of these standards into provider information systems. For example, the tools that have been developed to integrate vocabularies into the UMLS might be used by providers to map their local vocabularies to the UMLS as a step toward phasing out their local vocabularies.

Fourth, providers should be asked to handle eligibility inquiries via standard message-based mechanisms, so that any provider system can interface inexpensively and quickly with payer systems. The same approach, used for years by automatic teller machines to allow our credit cards to be recognized around the world, could be used to check patient eligibility. The reduction in lookup and keying of this information will provide the financial incentive for providers to upgrade their systems to use the interfaces.

Fifth, all systems that process patient data should make that information available for export as text reports, with tags according to a document architecture, such as the HL7 Clinical Document Architecture. If an unstructured text patient note in an EMRS was appropriately tagged with the universal header for patient notes, it would be easy to query this (or any other) system and display the notes. This minor modification—essentially another report—is all that is needed for another system to display it.

Intermediate Term

Additional steps could provide benefit in the intermediate term if begun today. These steps might be taken in the context of regional demonstration projects, such as those funded by the Agency for HealthCare Research and Quality (AHRQ). They involve new practice patterns and functionality not yet in general use. These functions can be easily handled by today’s technology:

- Smart card technology for positive patient identification to avoid the overhead and privacy concerns of a uniform identifier while providing a key for a patient to use to grant access to their records and to log encounters with providers.
- Personal health records, including Web-based, self-charting tools, to allow a patient to track their own progress and provide direct input into their medical record.
- De-identified data for surveillance and quality monitoring exported from various existing systems. Application Programming Interfaces might be inserted to allow them to directly accommodate queries from external systems for clinical documentation.
- Practice guidelines published with tags according to a documented architecture to permit direct incorporation into
various systems, so when the data listed within the header for a diagnosis shows a specific condition or disease, such as Diabetes Mellitus, practice guidelines for that condition could be automatically displayed.

Long-Term
Over the longer term, the NHII might include certain central services for use across the health care system. These services would be fundamentally different from today’s systems, separating the management and access of key information from the systems that automate the various facilities that make up the health care system.

The starting point might be services to support authentication that a patient or provider is who they say they are, together with services to authorize or consent to information access. To deliver on the long-term vision, however, the United States needs basic computer science research focused in three areas:

- Technologies for extracting, managing, and protecting health data. For example, how do programs extract structured data from text, image, and voice? How do they manage data captured across a lifetime and guarantee that it will remain accessible as technology evolves? How do they relate genomic and clinical data? How do they scrub data from the systems that have touched it, if the patient decides to revoke access?
- Technologies for increasing human capacity by reducing the cognitive and work effort required to perform a task.
- Practical approaches to increasing the scale of current health information and communication tools. For example, how do systems adapt to varying workflow requirements? How do they handle the explosion in biomedical knowledge? How do they accommodate the additional complexity of data interchange?

On a parallel track with these technology and computer-science-centric efforts, it will be important to focus biomedical informatics research on new publication types and knowledge representation models that make fresh knowledge directly actionable by both humans and computers. For example, if a new drug–drug interaction is suspected, could it be recorded in a national databank in addition to publication as a case report?

Conclusion
The goal—an improved health care system resting on an information and communication infrastructure—is achievable. The United States will not reach this goal by maintaining the status quo. Action is necessary at all levels, by federal and state government, regions and communities, and health care organizations. The complexity of the task of implementing infrastructure increases as the scope changes from organization, to region, to country. Accordingly, approaches that work in one organization are unlikely to work for the country as a whole. The NHII initiative should focus on a few critical items, such as establishing key standards and health care policies, which must be done on a national scale to make it easier to exchange information among regions and organizations. In parallel, regional- and organization-based efforts to share information across the care continuum need to begin in earnest. Best practices and successes need to be reproduced. If successful organizational and regional undertakings are aligned with national strategy, the collective efforts will incrementally build toward an effective health care information and communication infrastructure.

References


