

Information Sharing and Privacy with Personal Medical Records

A Research Report

Allan Friedman
David Lazer
Kennedy School of Government
Harvard University

Executive Summary

To better understand the role of electronic health records in healthcare delivery, we must understand how medical information moves through the healthcare system. This report summarizes the theoretical underpinnings of a model of patient data flow, and presents the results of a small field study designed to understand the mechanics of patient data flow in practice. Since privacy is critical to the acceptance of electronic records, special attention is paid to limiting the scope of data access. As a patient moves through the healthcare system, each medical or administrative step might require some amount of patient data to be accessed by the appropriate stakeholder. Limiting access to data may prevent care providers from properly aiding patients; failing to limit access to data can compromise patient privacy. Despite wide variation in practice, enough common structure runs through access patterns inside the private healthcare system to make generalizable modeling possible. Two perspectives of modeling of patient data flow are presented, one focusing on a wider scope of privacy issues, the other aimed at providing empirically grounded quality of care metrics. Further steps are identified to build a more complete, empirically grounded model.

December 1, 2006

Motivation

The healthcare sector has undertaken a massive, long-term effort to fully integrate information technology into service provision through personal medical records (PMR). Putting all personal health information in digital form on a network can allow for increased efficiency in healthcare delivery, but also raises serious privacy concerns if access to data is not controlled. This project examines the tradeoff between privacy and quality of care.

To better understand how to protect patient data, we must understand how the information to be protected moves through the healthcare system. This paper explores how to represent the dynamics of patient information access demand. Healthcare provision relies on the interdependence between large numbers of fragmented providers, while organizations themselves are balkanized by differences in functionality and process. As a patient moves through the healthcare system, each medical or administrative step might require some amount of patient data. With a comprehensive model of information access, appropriate models and policies can be developed to minimize data exposure while ensuring quality care.

Building a model of knowledge flows in patient care requires unpacking the medical care provision systems. Information flows, or access patterns, are the result of practitioners seeking out data to make decisions. Who needs what data? Practitioners must be examined in their roles (physician, administrator, technician), as well as their position in a organizational network of providers. A model must include organizational boundaries, as well as professional behavior and relationship with the patient. The type of information is relevant, as is the reason for accessing it. These factors together can be used to understand patterns of legitimate data use that any privacy regime must incorporate. Full policy analysis must address the consequences of overly broad data restrictions, preventing medical care and critical decisions. Incomplete information could lead to damaging medical errors, such as adverse drug interactions.

This approach will not only inform questions of medical privacy, but also be broad enough to offer insight into personal medical records in general. The

information architecture (in Figure 1) will depend on efficient interactions across organizational and technical boundaries. Understanding patterns of data access will help target potential bottlenecks and obstacles, and possibly identify new information system needs. Enabling physicians to get the data they need will improve efficiency; enabling the right data to flow across the medical system will reduce errors stemming from ignorance or information overload.

Understanding patient data

The adoption of digital data management in the healthcare sector has been slow, fragmented and largely lacking in coordination. Recent efforts have shifted from a standardized, centralized electronic health record model to what Halamka et al (2005) call “coordinated decentralization.” Traditional electronic health records are hospital-based systems with unique legacy features and limitations. The vision of a universal electronic healthcare record with full, interlinked data for everyone has proven unrealistic. Policymakers now speak of a regime of Personal Medical Records (PMR). These record systems may be thought of as virtual records, linking disparate local databases through a common interface. One such proposal can be seen in Figure 1. The desired structure should be “patient-centric” or organized around the individual, rather than the institution. Absent national standards or coordination, healthcare data today can be found in modern interactive systems, legacy hospital databases, or isolated paper records (Grossman 2006).

Despite the attention being paid to the need for proper patient information management, relatively little empirical work documents how patient information moves through the healthcare system. Young (2005) comments “we know too little about patient pathways to [model] them as freestanding entities.” One study found that 90% of information exchanges between practitioners in an Emergency Department were interpersonal, rather than through a formal information source (Coiera et al 2002). Inside the clinical environment, information transactions are embedded in ephemeral artifacts (charts, whiteboards) and organizational processes such as rounds and end-of-shift handover procedures (Van Eaton et al 2004, Bossen 2002, Gurses & Xiao 2006). Lorence and Churchill argue that the oral and paper-based system

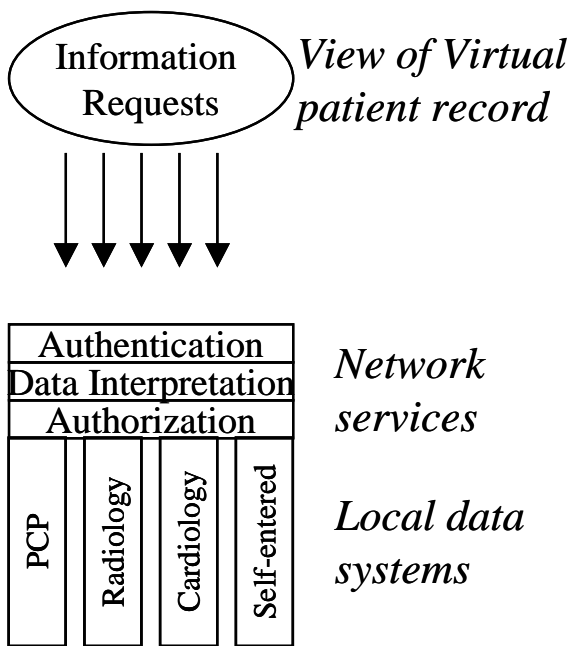


Figure 1 – A simplified example of a Personal Medical Record architecture. Data is maintained in disparate databases, including data the patient has personally chosen to include. The records are combined by a common interface to present a coherent record while still maintaining privacy and security.

has already shaped many aspects of healthcare delivery, and will ultimately dictate much of the architecture of digital infrastructures. Communication across organizational boundaries will help reduce the problem of shadow files in multiple locations, which may contain unique information that must be directly and specifically requested if it is not to be ignored (Giuse 2003).

Understanding how patient data flows through the medical system requires the integration of several different dimensions of data. The scope of the data use has traditionally been focused at one of several levels. Inside the ED or hospital ward, clinical data is particularly time sensitive, and integration into workflow is the primary priority (e.g. Eaton et al 2004, Gurses & Xiao 2006). A more macro view might span several different provider units such as the physician, laboratory, home care, etc. (e.g. Wang et al 2004). This scope can be narrowed to look at individuals inside a given physicians office, or expanded to encompass institutional management and payment. Non-medical personnel may still conduct transactions with medical data by actually executing physician instructions, in addition to carrying out

basic administrative management. It is important to note that some administrative data can be sensitive in and of itself. For example, one's mere presence in an HIV clinic, Ob/Gyn offices or genetic counselors sends a fairly strong signal about health status.

Data usage patterns also have time attributes that must be considered. Clearly, some medical data, such as blood type, are enduring features of a data record, but what about hourly reports from clinical monitoring equipment? Time sensitive data may indicate a more critical medical situation, but could pose no less of a privacy threat. Krol and Reich (1999) also stress the difference between acute and chronic conditions, which would have different data demands. The final question of scope is the actual data consumer: should data be conceived as flowing through roles (e.g. an anesthesia resident) or individuals (e.g. Alice). Treating data handlers as individuals increases the complexity of the model. The benefit of such a step is that it allows us to examine the interleaved networks of multiple roles in the healthcare system, including hierarchical relationships, potential for privilege abuse, etc. Data generation and spread through time and topological space must be understood medically, organizationally and technically.

The type of data is also relevant to the function of any records system. A particularly important distinction is that between machine-readable data and free form text. The former allows an intelligent system to process data, while the latter allows the entering party (physician, administrator, or even patient) greater flexibility. The tension between these two demands has been cited as a major sticking point to standard development (Ferranti et al 2006). Progress has been made at building a common data standard (UMDS 2003); encouraging physicians to use standardized forms or data entry tools rather than the traditional annotated charts. Data that can be moved easily across networks and interpreted with software agents are more likely to be used.

Finally, data flow is further complicated by emergent system behavior. Altering the flow of data through the medical system will likely have unforeseen consequences. In one example cited in medical literature, the initial lag in data entry from a written admission form to networked information system delayed rapid treatment in a group practice

(Templeton et al 1983). The doctors and staff had data in a more flexible, accessible fashion, but did not have the immediate access they needed to the basic data. In another case, the centralized record system of the Department of Defense was initially single-threaded to eliminate synchronicity faults, resulting in a system so slow that primary care physicians on the system were forced to eliminate routine checkups (Philpott 2006). That is, a simple attempt to safeguard the quality of information across access points dramatically reduced the access capacity. Similarly, the introduction of wireless alert pagers in a large hospital system disrupted the established system of crisis escalation, overloading attending physicians and creating information overload (Reddy et al 2003). While residents used to handle the basic crises, the pagers systems did not have any capacity to target alerts, so more senior physicians were involved where they previously had not been needed. In general, the integration of data infrastructure can disturb existing process in addition to providing benefits of increased connectivity. Each of these examples documents a single occurrence, so one must be wary of extrapolating, but they make compelling evidence that relatively small changes in the access patterns of a patient's data could have unanticipated consequences on the larger system dynamics.

Patient Data in an Organizational Context

Patient information is accessed over the course of medical care provision, which takes the form of a set of organizational processes. Organizational processes and their relationship with information technology has been studied more generally under the framework of workflow analysis (Georgakopoulos et al 1995). This level of analysis focuses on functional units, and how they fit together to accomplish a specific set of tasks. We can treat the patient as a task model to be followed to completion through steps in the medical system. This approach offers a framework to begin to understand patient data flow by focusing on the flow of the patient through the medical system.

The trail of a patient through care provision has been referred to as "patient path" (1980). Specifically, it is often used to denote a regime of established steps to treat a patient (Every et al. 2000). Analyses of these patient paths have been used in evidence-based medicine to quantitatively measure the best course of care, and the morbidity and mortality effects of

deviation from that path (see e.g. Rosenberg 1995, Sacket 2000).

The workflow approach incorporates the interdependence of the technical and the organizational aspects of healthcare. Individuals, processes and technology combine to establish how the system reacts to a given task. Workflow focuses on the task, rather than the patient, allowing us to model instances when the patient need not be present. Documents such as patient progress reports flow across actors, proceeding from thought to finalized, official record as individuals fulfill their duties (Payne and Graham 2006). For example, a report might be drafted by an intern, and must then be supplemented and approved by a resident or an attending. The workflow model determines who can do what to a given object (Malameteniou et al. 1998, Poulymenopoulou et al. 2003). This is not a study of care provision, but a process management approach of the tasks and data involved.

A data access-centered conceptualization of a workflow, then, is a set of rules about access to interact with a patient's data. This is conceptually distinct from data about a patient's status (e.g. Berler et al 2005), since it shifts attention from the immediate care provision. Instead, it rests on the assumption that care provision requires patient data, and produces more data, each of which involves interacting with the patient data record. Malameteniou and Vassilacopoulos (1998) argue that virtual patient record management can be understood through a workflow approach. This paper argues that a successful model is a balance of workflow and data management rules. An initial understanding for workflow in the clinical environment is developed from discussion with a range of practitioners described below

Preliminary Field Work

Goals and methods

The initial fieldwork was designed to obtain a general picture of information use by the various stakeholders involved in healthcare delivery. In order to build a quantitative model, the basic parameters must be understood, and the suitable set of questions must be designed to insure appropriate measurements and consistent responses. To that end, multiple

Table 1 - Participants in the preliminary field research (* from the Partners Healthcare System)

| Field | Job | Org |
|--------------------|-----------------------|-----------------------------------|
| PCP * | Physician | Hospital-based practice |
| PCP | Physician | Small, independent practice |
| Emergency Dept. | Physician | ED in medium-sized hospital |
| Geneticist * | Physician | Clinic in large research hospital |
| Surgeon | Physician | Hospital-based practice |
| Dermatology * | Resident | Large Research Hospital |
| Endocrinology | Administrator | Small, independent practice |
| Pathology Lab * | Administrator | Large research hospital |
| Administration * | VP, CIO | Large health system |
| Administration * | Chief Privacy Officer | Large health system |
| Healthcare Finance | VP, Development | Health Business Solutions |
| Insurance | Senior Administrator | Large Pharmaceutical Insurance |

practitioners were interviewed in an open-ended fashion, with discussion guided to cover how the role interacted with the patient: how and why the patient would interact with the practitioner, the source and scope of information involved in the event, and what other parties were incorporated into the patient process over the course of the transaction. Respondents were probed to discuss the range of circumstances under which they might interact with the patient and patient data. Where applicable, respondents were encouraged to talk about a “typical patient” or propose a small set of archetypal patients representing their regular interaction. Each interview offered two contributions to the project:

- A patient-centric ego-network centered on the respondents position in the healthcare system, textured with details on data type, quantity, and time requirements.
- Insight into the local process of patient data acquisition and use, as well as deviations from an ideal model of the patient path through the system, highlighting organizational and technical complexities.

Interview subjects represented a diverse sample of the healthcare provision sector, detailed in Table 1. The range of input encompasses variation in role, as well institutional factors including size of organization and sophistication of information systems. Most interviews lasted between 40 and 70 minutes, divided between telephone and physical discussions.

Summarized Findings

The basic pattern for medical provision follows a simplified path of entry (either patient-motivated or through referral), followed by a consultation, based on which diagnostic tests are ordered (See Figure 2). The care provider could offer treatment, refer the patient to another physician, or end the case. Of course, there are many complexities to be introduced, but this is the baseline model for treatment. A member of a surgery team would often, but always, interact with the patient beforehand: they would have patient information nonetheless. A patient might be receiving care for two distinct ailments from two different physicians, who may or may not communicate with each other, but each would still follow this basic pattern.

Another common feature across medical situations was the perception that there was usually a single locus of control, or a “hub” that coordinates patient care, like the set of transactions in Figure 3. For example, when the PCP refers a patient to a specialist, the specialist then collects the relevant information, interacts with the necessary parties, takes the required action and then (ideally) reports back to the PCP upon completion. Again, there were multiple exceptions raised, but the general case seemed to follow this approach. When assigned a patient through the triage process in the ED, a physician acts as the hub to coordinate care, but has no more responsibilities after the admit/discharge decision. Generally, the providers all saw either themselves or a single specific alter acting as the hub during the care process. This understanding of patient care would support an approach of collecting data from individual practitioners, since it lends itself to a practitioner-centric network. Of course, since this conclusion was reached by interviewing practitioners, further validation is necessary.

A third important insight drawn from the interviews is the distinction between inpatient and outpatient processes. The looping decision processes made by the hub actor described above is particularly relevant for the decentralized treatment model of outpatient care. Processes inside the hospital have several key

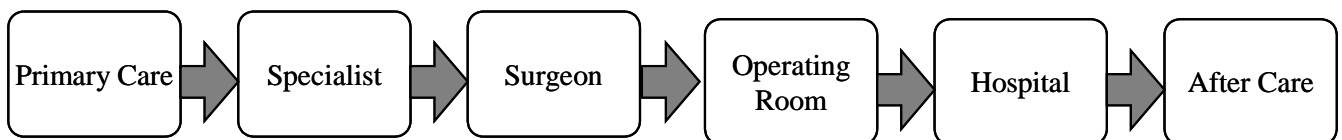


Figure 2 - Representing generic medical care needs as a simple path.

features. First, all the hospitals discussed by practitioners in this survey had some form of electronic medical record, even if it is purely an in-house system. The strongest data protection mentioned was a restriction on access across hospitals inside a larger system, with open access to every record inside. At the same time, patient data is instantiated as the medical chart, which follows the patient through the institution and can serve as a communication medium between care providers about the current status and treatment details. Lastly, the context of inpatient treatment was described as qualitatively different. Multiple care providers deal with the patient, more data is generated through more extensive monitoring processes, and the cost of consuming hospital resources puts time at a premium. Stays longer than 24 hours will necessarily involve interacting with different individuals fulfilling the same role of nurse, resident, etc. Patient data is referenced more frequently: during daily rounds by one or more physicians, when nurses and residents monitor status, and when care is transferred across shifts. Moreover, much of this data is ephemeral, and only relevant to measure period-to-period progress, needing no long term storage. At the end of hospitalization, relevant information should be culled from the accumulated records and given to the new “hub” for follow-up care, as well as being placed into some permanent patient-specific medical record.

For clinic-based ambulatory care, patient history was one of the most commonly referenced needs for data acquisition when beginning a relationship. All providers noted that they usually needed to supplement any history presented. For providers in the PCP role, getting an initial patient history for the first visit requires extensive research, often including mail-in or online questionnaires. Specialists with specific data needs reported an even harder time getting the patient information they need.

An endocrinologist might need details from a previous hospitalization, which would require two steps: relying on the patient to fill out the forms necessary to request a record release, and hoping the hospital records system is efficient and reliable. Both of these steps “are harder than one would think.” A geneticist could face an even steeper challenge, since they may need information from people other than the patient. The challenges of location and persuasion, particularly with the social stigmas on genetic defects, can present “an insurmountable dilemma.” Given the prominence of patient history in PCP and specialized diagnosis processes, it is interesting to note that the complete patient history was not always offered to the patients’ other care providers.

Along with patient history and physical examination, tests from pathology, radiology, and other labs are a major source of information in patient care. This information, of course, is generated by individuals who study samples and generate images of the patient, and pass the results back to the requesting practitioner. The nature of these lab tests varies widely, depending on the organizational context as

well as the technical sophistication. From an information modeling perspective, the level of interaction and information sharing between sponsoring physician and lab affects information flow. Neither primary-care physician regularly sent patient medical data to the lab with the sample, although the outside lab used by the small practice required extensive personal information:

Observations from Practitioners
The medical process can be seen as a path through care organizations.
Actors serve as hubs, forming form zones of data control through organizational control.
Different trends and processes are used in- and outpatient care.
The data requirements of labs and imaging vary with organizational practice and the clinical difficulty of interpreting test results.
There is regular but straightforward interaction with the billing component of the medical industry; physicians little direct involvement.

name, birth date, social security number, patient phone number and insurance information. The hospital-affiliated practice used only name and internal identification number. Both specialists had some in-house capacity to run tests, and both reported having a more involved relationship with the testing laboratories. Examples were offered of conditions where the laboratory analysis requires an

understanding of the patients' condition. At the cytogenetics lab examined, on the other hand, the problem was insufficient data: lab administrators found that a very large minority of requisition forms failed to include relevant patient information, required administrative details, or even the nature of the biological sample or the desired test. Lab personnel have to track down the requisitioning physician. If that is infeasible, and the patient is inside the hospital, they may access the patient's electronic hospital record, but even this is not always successful. The administrative side of the lab must have access to patient medical data to verify adequate patient data to perform the test. The technicians performing the tests do not themselves use any administrative data containing patient identifying data, although the data are present on all paper forms in the lab. Lab results were reported by fax, but only after verifying that the receiving fax machine met confidentiality standards of access control and security. In cytogenetics, "a good number" of cases require follow-up discussion between the requisitioning physician and the physician in charge of the lab who verifies technicians' analyses.

The interviewers seldom discussed their interactions with billing and financial administration. The literature confirms a diminished role of payers in the provision in medical care, as the "gatekeeper" function grows less popular. Instead, insurers will target care providers with specific incentives to minimize cost, and may only require approval for a select group of very expensive procedures with close substitutes. Medical information is still sent to the payer, of course, but that transfer is largely a simple one-way reporting transaction with very little potential for feedback. In mapping the flow of patient data, the financial component exists as a largely independent entity. Additionally, none of the practitioners discussed data collection or use for long term process studies or outcome-based medicine. The administrators in charge of data management highlighted the difficulties in obtaining such data, even in modern settings, and noted the relative lack of incentives and political difficulties in collecting the relevant information.

Finally, there were a number of examples where it became clear that a simple model of information protection would not be complete. A drug interaction warning system, for example, must notify providers

that combining a new prescription with an existing could pose a health risk, even if the practitioner does not have permission to know about the first prescription. One practitioner commented that inpatient chart was used to leave notes for other physicians, as an effective high-latency communication channel. Unfortunately, occasional poor penmanship and illegible signatures confused caregivers and required further investigation to figure out who meant what. Another physician narrowly caught a patient who was not suitably prepared for surgery because the anesthesiologist did not know about a certain treatment that would require specific precautionary steps. The information had been transmitted, but in the process of transferring the information into the pre-op files, the treatment was not included because it was not on the hospital formulary. With one exception, the medical providers volunteered strong feelings on the importance of direct, doctor-to-doctor communication, usually by phone. These immediate communications allowed for richer communication and discussions or questions. They are often very quick, seeking to clarify a diagnosis or a test result, or to assess the need of further action. The primary care physician often had to be contacted to get patient information, and a paging mechanism was usually employed.

Components of an effective model

The underlying goal of a model of patient information flow is to examine the tradeoff between privacy and patient care, but a model that is designed as broadly as possible might inform other aspects of patient data use in personal medical record systems. The interviews above help build an understanding of how data are used, and what must be measured to accurately represent this data usage. It is important to note that the preliminary fieldwork reports that many of these components are only evident locally: many transactions do not leave an electronic trail in extant data systems which might be mined for modeling.

The preliminary analysis suggests that the practitioner would be a good unit of analysis. In addition to being the initiator of patient data interactions, it allows for greater flexibility in empirical measurement techniques. Each actor that dealt with a patient's data should be part of the patient data flow model, with a description of that interaction mapped into as formal a representation as possible. While an exact taxonomy

of patient data is not immediately clear from this preliminary work, the data used by each actor can be broken down into administrative data which might tie the medical record to a specific person, and medical data which refers exclusively to some medical condition. This latter category could be broken down into further components, although different hierarchical topologies might be developed based on purposes of clinical care (who would use what) or data sensitivity (the potential harm of an information leak). An effective model would combine both.

Sources of variation

Even in a highly stylized model of medical care, there are two distinct sources of variability in any observed behavior. First, patients with similar conditions might actually require different care based on personal reactions, complications and further diagnostic efforts. Thus, a physician might treat two patients with the same complaint differently out of the natural biological variation. Perhaps more important to this study, however, is the variation in care deriving from the technical and organizational capacities of the care environment. The patient path and data flow might be very different across practitioners in distinct practices than it would be in a major research hospital. Any modeling strategy must be careful not to conflate these two sources of variation.

Outcome metrics

The model should allow the user to gain insight about quantities of interest. We focus on patient data because of privacy concerns. Privacy has proven notoriously difficult to quantify (Wathieu & Friedman 2005, Solove 2005). Since the privacy threats discussed above deal with unwarranted access to patient information by an “inside user” then one approach is simply to measure the number of users who have access to the data. In a perfect world, that number should be exactly equivalent to the number of people who need access to the data. In an environment without any privacy protection whatsoever, that number is equal to the number of authenticated users of the information system in which the data is stored.

The chief priority of any medical system, however, should be quality of medical care. Patient privacy systems may interfere with medical care if the appropriate person cannot receive the necessary data. A very basic measure of quality of care, then, could be the inverse of privacy: the number of providers who need a piece of patient data but cannot access it. As the interviews confirmed, however, lack of needed information can have a very small impact on care (the physician’s assistant must

call another office to verify a test result) or a very large impact on care (ignorance of parallel treatment leads to a fatal drug interaction). Applying empirical studies of patient care decisions from evidence-based medicine might be able to offer some insight, since many studies measure the morbidity and mortality rates of deviating from the ideal path. Even absent precise clinical error data, failure to have access to patient data measurably increases adverse drug effects.

The need for this type of information requires an in-depth method of data compilation. This type of data cannot be gleaned from a simple examination of extant digital medical records. Beyond the privacy concerns of examining patient records, relatively few deployed systems have detailed information about the physician accessing the data, requiring synchronization with the hospital personnel files. More importantly, the records are not modular enough to detail who needed what data, nor do they encompass the more subtle data collection and sharing mechanisms described above. To develop a rich enough data set, qualitative data should be used to build a flexible simulation model.

Simulating patient data flow and privacy

Good simulations abstract a very complex world into a simpler model and enable the user to “experiment” on this world by varying the parameters and rules. The goal of this project is to model patient data flow and understand how changing the workflow rules would affect privacy and quality of care. This model must be as empirically grounded as possible to have external validity. While medical care has been studied at the level of macroeconomic simulation, process

Critical concepts in modeling

Appropriate measures of privacy and quality are needed.

Organizational practice, technical capacity and medical practice are all related, yet each should have a distinct effect if properly understood.

models of the full delivery system have proven more difficult. With the unit of analysis at the individual level, a complete model of an entire healthcare system would be very complex, and could span thousands of individuals. Moreover, we describe above how most patient interactions with the system are part of a set “trajectory” across different providers, be they administrators, physicians or support actors.

Based on the preliminary fieldwork, two general perspectives might inform a set of questions about data flows and privacy. They both build on the paradigm of following the patient path through the care delivery system. The first focuses on the organizational components, while the latter provides a mechanism for measuring desirable outcomes more readily. Both approaches acknowledge that modeling a complete healthcare system is too broad a scope to maintain a useful level of detail. Note that they are not mutually exclusive, but rather emphasize different modeling goals.

The first step of the underlying modeling framework is to focus on a particular subset of care. Each approach involves building a set of representations and patient cases and understanding how a given organizational process would treat these cases. This would involve further targeted empirical research to infuse the models with quantitative data. On top of these organizational models of patient care, we place a set of workflow constraints for the patient

information, and then measure the resulting privacy and quality of care outcomes.

Probabilistic Network Flow Model

The first approach is to use practitioner expertise and experience to derive a range of possible options for a given condition. Consulting with physicians, a specific condition or set of conditions is selected. A description of a hypothetical patient is developed to that extent that a physician would have a specific diagnosis at each level of interaction. The case would then be presented to an entry-point actor, either as a primary care role or an emergency role. This actor would be asked to reflect on the steps they would take to treat such a case, and the likelihood of each course of action. Any action that involves another actor would expand the scope of the study to such an actor. Essentially, absent a specific patient to follow, we make one up and follow their hypothetical path. Actors can also be asked to estimate how often such a case would come before them, and whether they might have a set of options to choose from, to create a range of potential patient pathways. Of course, the patient information required of each action would also be noted. Together, this information would create a model not unlike Figure 4. If multiple entry points are selected, they can be integrated, with probabilistic weights being assigned to each branch.

Since every node has a set of data needed, information access rules can be imposed on this model.

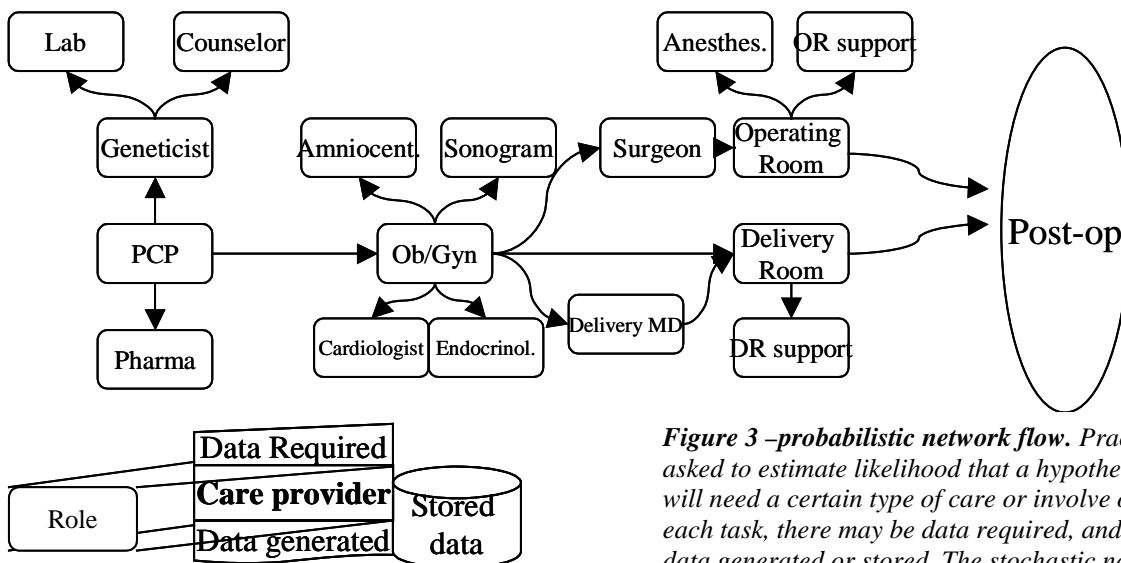


Figure 3 –probabilistic network flow. Practitioners are asked to estimate likelihood that a hypothetical patient will need a certain type of care or involve others. For each task, there may be data required, and possibly data generated or stored. The stochastic nature of the data builds a weighted network, which can be used to assess the seriousness of incomplete data or the harms of lax access control.

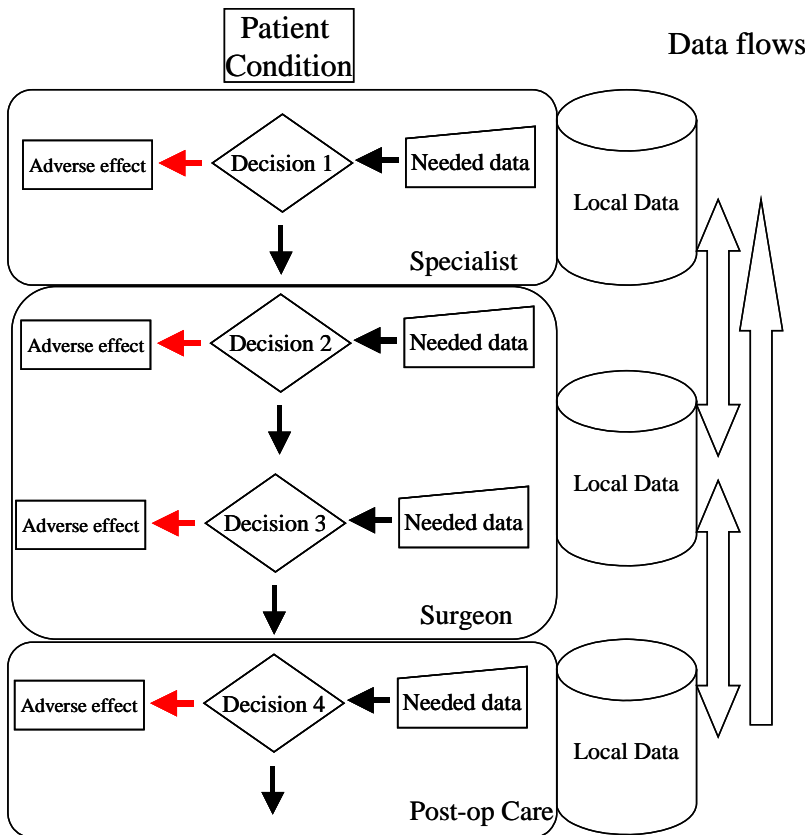


Figure 4: **Focusing on patient outcomes.** Actors make sets of decisions in an empirically studied case. Those without the necessary information cannot make the right choice, and the patient will suffer the adverse effects. Note in this case, information does not flow from the specialist to post-op care, creating the potential for an ill-informed decision.

Organizational conditions would yield the number of others who might have access to the patient information, whether in a large hospital, a small practice, or a hypothetical personal medical record system. These assumptions of data availability could be further subject to technical or legal constraints about who could access what. By varying these assumptions, some practitioners in the network flow model may lose access to patient information, negatively impacting quality of care. This model is constrained by a lack of patient outcome information, so lack of necessary data is the only proxy for diminished quality of care.

Incorporating patient outcomes

The major flaw with the first model is that it does not offer insight into actual patient care metrics. Since recent evidence-based medicine research has attempted to provide accurate and universal measurements for quality of care in a given situation, the second model uses this data. Specifically, many

quantitative studies of patient care identify an optimal clinical pathway, and then measure the consequences of deviating from that pathway, most often in likelihood of death. Mortality rates are computed for every step of the pathway. Tying this into patient data, failure to have the requisite information would make it impossible to make the correct decision to stay on the pathway. Thus, any attempt to constrain information flow might have an empirically grounded consequence in patient health. To build an informed choice model, the clinical pathway defined in the literature must be mapped into the health organization to be studied, as in Figure 4. This would consultation with physicians, as well as senior administrators with a holistic view of the system.

Once an organization-specific representation has been developed, a set of data access assumptions can be imposed on the clinical path, similar to the study of the probabilistic network flow described above. While the set of access patterns is unlikely to be as

detailed as the first model, if a certain actor cannot access needed information, it is possible that they may not be able to make the correct decision. Mortality rates from the empirical study can then be applied to derive more complete metrics of patient care.

A spectrum of modeling options

This second approach stresses the value of wellness metrics that are, after all, the ultimate goal of any care provision system. On the other hand, this information is taken from empirical medical literature, which can constrain the scope dramatically. If detailed research only exists on a certain patient path, then a faithful mapping into simulation could not incorporate detailed empirical observations in a new study without losing some measure of validity. Every layer of organizational complexity further hinders a faithful representation of the evidence-based findings into the computational model. Yet ignoring patient outcomes completely would not fully inform the research questions about the relationship between care and

data access. There exists a spectrum between flexibility and fidelity to established empirical work. The patient outcome model can be relaxed to accommodate observed details about medical organization. Alternatively, the literature can be used to make grounded assumptions about the human costs of poor data sharing in the data flow model. Any productive modeling strategy must find a balance between accurate and useful models of organizational data sharing, and appropriate use of scientific finding about proper patient care.

Of course, every research method has limitations. Even if the model is carefully implemented from valid qualitative data, there are several shortcomings to this approach. The model can test counterfactuals across organizational and information policy options, but is limited in the medical complexities it can incorporate. The complexities of information flow are limited to those anticipated by research subjects, since the model does not have a built-in capacity to reflect emergent behavior or feedback from how each practitioner would react to changes in data flow rules. These are assumed to be constant across the different conditions. Moreover, the long term effects cannot be measured from such short-range diagnostic questions. The benefits of electronic medical information may be manifest as patients repeatedly interact with the medical system under a variety of different conditions, yet the information-sharing models presented here focus on short-term acute care.

Next Steps

For both strategies discussed above, the first need is to identify the clinical context to be modeled. This critical decision must be made with extensive consultation with experienced physicians, as well as public health experts with experience in measuring treatment outcomes. Several of the contacts made in the preliminary research phase have expressed interest in participating in this decision. The difficulties of the selecting the appropriate treatment vector cannot be underestimated. Several experts strongly recommended a clearly constrained diagnostic option set. In addition, incorporating patient outcomes will draw extensively from the evidence-based medicine literature as well. If this decision can be made while considering other factors or policy interests, this would increase the overall relevance of the project.

With a clear idea of clinical demands, the modeling approach will be selected to best capture the unique features of the clinical context while still maintaining generalizability. The scope of the model does not have to be fully fixed at this stage, but a fairly complete set of model components will be detailed, including data and practitioner typologies. At this stage, for example, the level of administrative support to be included in the model will be selected. During this process, commercial and academic process-modeling software will be examined to see if they will meet the needs of the project. These will be compared to the costs (time and money) of building a simulation model to exact specifications from scratch.

The third major task before data collection will be the specification of environments to be studied. If possible, a comparison between a large healthcare provider such as Partners and a smaller, less centralized system could offer insights about the role of the institution in affecting outcomes. Existing relationships should facilitate this selection process. Following that, a set of questionnaires will be developed to secure quantitative responses from practitioners about their expectations of patient care delivery. These questionnaires should be field tested, then distributed based on a pattern appropriate to the model. Experience in the preliminary fieldwork has informed this researcher that a face-to-face interaction with the practitioner will have a higher success rate and offer more thought out answers than written correspondence.

Venue selection is particularly important for policy analysis. Most major research-driven hospitals already have some level of electronic health record (Jha et al 2006). These systems are evolving and improving, but several hospital administrators agreed that the basic data policies for in-house data use have already been established. The more diffused medical care delivery systems, outside the major hospital systems, have been much slower to adopt portable information systems that enable data sharing. Recent policy efforts have pushed the concept of Regional Health Information Organizations (RHIOs) to link diverse practices, clinics and primary and secondary regional hospitals. Given the decentralized organizational context of RHIOs, developing an understanding of good data handling practices should be a research priority.

More broadly, this project can form the foundations of a study in the efficacy of benefits of electronic health records. Beyond immediate questions of privacy, a data access model can inform the design and implementation of personal medical record systems. Such a model can also contribute to the policy questions of how to collect information for input into evidence-based medicine studies, as well as how to best leverage the findings of such studies to improve patient care.

References

- Berler A, Pavlopoulos S, Koutsouris D. Using key performance indicators as knowledge-management tools at a regional health-care authority level. *IEEE Transactions on Information Technology in Biomedicine*. 9:2, 2005.
- Bossen, C. The Parameters of Common Information Spaces: The Heterogeneity of Cooperative. Work at a Hospital Ward. *Proceedings of the 2002 ACM conference on Computer supported cooperative work*.2002.
- Brailsford SC, VA Lattimer P Tarnaras and JC Turnbull. Emergency and on-demand health care: modelling a large complex system. *The Journal of the Operational Research Society*.55. 2004
- Cameron, D and IG Jones. John Snow, the broad street pump and modern epidemiology *International Journal of Epidemiology*, 12:4, 1983
- Canfield, K. Clinical resource auditing and decision support for computerized patient record systems: A mediated architecture approach. *Journal of Medical Systems*. 18:3 1994.
- Cheng CHF and Levitt RE (2001) Contextually changing behavior in medical organizations. Proceedings of the 2001 Annual Symposium of the American Medical Informatics Association, 2001.
- Cohen, M. A., Hershey, J. C., and Weiss, E. N., Analysis of capacity decisions for progressive patient care hospital facilities. *Health Services Research*. 15(2) 1980.
- Coiera, EW, RA Jayasuriya, J Hardy, A Bannan and M Thorpe. [Communication loads on clinical staff in the emergency department](#). *Medical Journal of Australia* 176 (9): 415-418. 2002.
- Every, NR, Hochman, J, Becker, R, Kopecky, S and Cannon, C. Critical Pathways: A Review. *Circulation* 101 (461-465) 2000.
- Ferranti,J.M. R. C. Musser, K. Kawamoto, and W. E. Hammond. The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis. *Journal of the American Medical Informatics Association* 13(3) 2006.
- Giuse, DA. Supporting communication in an integrated patient record system. Proceedings of AMIA Symposium, 1065, 2003.
- Grossman, J. Personal Interview, May 8, 2006.
- Gurses AP, Xiao Y. A systematic review of the literature on multidisciplinary rounds to design information technology. *J Am Med Inform Assoc*. 13:3 2006.
- Jun JB, Jacobson SH, Swisher JR. Applications of discrete-event simulation in health care clinics: a survey. *Journal of Operational Research*; 50:2, 1999.
- Krol M, Reich DL.Object-oriented analysis and design of a health care management information system. *Journal of Medical Systems* 1999 Apr;23(2):145-58.
- Lorence, Daniel P. and Richard Churchill. "Incremental Adoption of Information Security in Health-care Organizations: Implications for Document Management" *IEEE Transactions on Information Technology in Biomedicine* 9:2, 2005.
- Malamateniou, F, G. Vassilacopoulos and P. Tsanakas. "A Workflow-based approach to virtual patient security." *IEEE Transactions on Information Technology in Biomedicine* 2:3. 1998.
- Payne, TH and G Graham. Managing the Life Cycle of Electronic Clinical Documents. *Journal of the American Medical Informatics Association* 13(4) 2006

Philpott, Tom. New medical records system a headache for doctors. *Stars and Stripes*. April 13, 2006. (online at <http://www.estripes.com/article.asp?section=104&article=34709>)

Poulymenopoulou, M; Malamateniou, F; Vassilacopoulos, G. Specifying **workflow** process requirements for an emergency medical service. *Journal of Medical Systems* 27 (4) 2003.

Reddy, MC. W Pratt, DW McDonald, MM Shabot. Challenges to Physicians' Use of A Wireless Alert Pager, Proceedings of the American Medical Informatics Assoc Symposium (AMIA2003), 2003

Rosenfeld, L and Moreville, P. *Information Architecture for the World Wide Web, Second Edition*. Cambridge, MA: O'Reilly Media, 2002.

Shortliffe, EH "The evolution of electronic medical records," *Academic Medicine*, vol. 74, pp. 414-419, 1999.

Sibbel, R., Urban, Ch.: Agent-Based Modeling and Simulation for Hospital Management, in: Saam, N. J., Schmidt, B. (eds.), *Cooperative Agents, Applications in the Social Sciences*, Dordrecht, Boston, London 2001, S. 183-202

J Templeton, M Bernes, M Ostrowski. The impact on a private group practice of converting to an automated medical information system. *Journal of Medical Systems* 7:4, 1983

Van Eaton, Erik G; Horvath, Karen D; Lober, William B; Pellegrini, Carlos A. Organizing the transfer of patient care information: the development of a computerized resident sign-out system. *Surgery*, 136 (1) 2004

Wang, Maisie; Lau, Christopher; Matsen, Frederick A; Kim, Yongmin. Personal health information management system and its application in referral management. *IEEE Transactions on Information Technology in Biomedicine* 8 (3) : 287-97 2004

Young, Terry. "An Agenda for Healthcare and Information Simulation" *Healthcare Management Science* 8, 2005.