



The Role of Clinical Information Systems in Facilitating the Personalization of Medicine

In response to:

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Request for Information (RFI): Improving Health and Accelerating Personalized Health Care Through Health Information Technology and Genomic Information in Population- and Community-based Health Care Delivery Systems

Mark Hoffman, Ph.D.
Director of Genomics Strategy
Cerner Corporation

mhoffman1@cerner.com

(816) 201-3291

2800 Rockcreek Parkway
Kansas City, MO. 64117

Contents

PART 1 - PERSONALIZED MEDICINE – PROVIDER PERSPECTIVE	5
The Technological Environment for Personalized Medicine	5
Trends in diagnostic technology	5
Clinical Information Systems - Survey	7
Clinical Information Systems:	10
Technical considerations relevant to personalized medicine	10
Software architecture	10
Security	13
Interface messaging	14
Codification of diagnostic results	15
Software Application Capabilities	16
PART 2 – PERSONALIZED MEDICINE – DATA ANALYSIS	19
General Considerations for Clinical Data Warehouses	19
Clinical data sets to support personalized medicine	21
Non-clinical data management for Personalized Medicine	22
PART 3 - DELIVERING PERSONALIZED MEDICINE	24
Executive Summary	25
Reference List	26
APPENDIX 1 – CERNER PROFILE	27

Introduction

The personalization of medicine is the process through which clinical decisions are based on complex information known about the person whose health is being managed. While clinicians have always taken personal information about the patient into consideration when making decisions, recent advances in biotechnology, imaging capabilities, information technology and basic medical practice have promoted the widespread belief that the practice of medicine in the near future will be much more "personalized" than it has been in the past. This is primarily due to the rapid increase in the amount of information that can be known about an individual.

Personalized medicine is often used synonymously with pharmacogenetics or pharmacogenomics. It has long been recognized that human genetics contributes to the variability between persons in how well they respond to a medication. Pharmacogenetics is the analysis of the contributions of single genes to drug response. For example, variations in the NAT2 gene influence patient response to the antimalarial drug isoniazid. Likewise, approximately 7% of the population has mutations in the CYP2D6 gene which render them non-responsive to codeine. Technological advances which have permitted the evaluation of hundreds, even thousands of genes, in a single test run have opened the door for systemic analysis of gene expression patterns in response to external stimuli, such as medication. This system-based approach is called pharmacogenomics.

While both pharmacogenomics and pharmacogenetics are among the most significant forces driving personalized medicine, many other trends are also making significant, if less publicized, contributions to the personalization of medicine. High-precision imaging modalities, especially Magnetic Resonance Imaging (MRI), are increasingly available and provide much richer information to the clinician. Outcomes analysis using large data sets has provided insights into the most appropriate decisions for various populations.

In addition to new diagnostic advances, the holistic consideration of a "person" from the clinical systems perspective is an important, but often overlooked aspect of delivering "personalized medicine". For example, medical information is notoriously fixed in place – a single patient can have three or more entirely separate electronic medical records in the city in which they reside. The current state offers few means, other than the patient's memory and tenacity, to ensure that these records are synchronized and consistent. This can create either an inconvenience (repeated entry of the same information) or a hazard (lack of awareness of a new medication allergy across care-providers). Other non-diagnostic dimensions to personalized medicine include sociological factors that influence how, when or where a patient may prefer to receive care and considerations that improve the ease and comfort through which medical care is delivered (improving the chances of a person seeking appropriate care).

In addition to the provider view of personalized medicine, data captured from clinical information systems can be very useful in assessing the severity and frequency of genetic-based adverse reactions to medications. Likewise, large clinical data sets can help identify adverse clinical outcomes that warrant further genetic investigation.

Clinical information systems can contribute to both enabling the care provider to take advantage of these advances and to the generation of new discoveries that further enhance the personalization of medicine. This report summarizes some of the key drivers behind the increased personalization of medicine, the current state of the art in clinical information systems and emerging trends that will determine the future utilization of clinical information systems in the delivery of personalized medicine.

Part 1 - Personalized Medicine – Provider Perspective

A fundamental goal of personalized medicine is transforming the interaction between the healthcare provider and the patient. Personalized medicine has inaccurately been contrasted with "population-based medicine". In reality, the technological forces driving the personalization of medicine allow the stratification of a patient into many more populations than was previously possible. These technological forces primarily are based on the development and availability of new diagnostic tools. For the information generated by these tools to be useful and clinically relevant, insights into the effects of membership in a particular population on clinical outcome are still required. The provider who can master the complexities of these changes will be able to offer a personalized clinical experience.

The Technological Environment for Personalized Medicine

The traditional diagnostic tools available to a physician consisted primarily of vital signs, a limited catalog of laboratory tests and relatively low resolution imaging (X-rays). Most physicians believed that the information generated by these tests could adequately be managed with paper based charts, reports and films (a belief that is still fairly common).

The continued reliance on out-dated information management technology is in many ways a rate-limiting barrier to the adoption of personalized medicine. Physicians who rely on pen and paper know that their information management system would be overwhelmed by the volume and complexity of information needed to manage the diagnostic results that will drive personalized medicine.

Trends in diagnostic technology

The greatest driving force behind the belief that medicine is becoming "personalized" has been the rapid advances in diagnostic technology, especially in biotechnology. There are more than 1000 clinically accepted genetic tests (www.genetests.org). The variety of results generated by these tests ranges from a single "normal / abnormal" finding to potentially many dozens or even hundreds of findings in a single assay.

The contributions that clinical information systems can make to personalized medicine are best understood in the context of the multiple technologies that are contributing to the rapid increase in the variety of diagnostic technologies available in clinical practice.

Genomics

- Pharmacogenetic testing – Increasingly, diagnostic laboratories are offering genetic tests that provide information about how well a person may well respond to a particular medication or which contribute to the dosing algorithm applied to a patient. Many laboratories have begun to offer genetic tests for the cytochrome p450 genes. These genes influence the response to more than 50 common medication and are among the most likely to be clinically significant (6).
- Increased catalog of genetic tests – while the growth in the number of diagnostic genetic tests resulting directly from the Human Genome Project (HGP) has not yet met expectations, there is continued progress in defining the process for introducing new genetic tests.
- Gene expression (microarray analysis) – high density gene expression analysis is primarily used in discovery, though there are a limited number of diagnostic labs actively working to incorporate gene expression analysis into their clinical offerings.
- Increased precision of molecular diagnostic tests – the increasing utilization of DNA sequencing to replace mutation panels has led to a need for data management and interpretation capabilities to ensure that these results provide clinical value.

Other biotechnology advances

- Proteomics – while analysis of single proteins of diagnostic value has been a common practice for decades, the systemic analysis of entire networks of proteins has not yet reached clinical practice. Like genomics, this emerging field has very significant potential.
- Metabolomics – the application of systems analysis to metabolites is another emerging field but has not yet been widely translated into active clinical practice.

Imaging

- MRI – Magnetic Resonance Imaging is rapidly becoming a mainstream and widely available technology. The high resolution diagnostic images allow physicians to explore anatomic changes in much greater detail.

- Precision imaging probes – A growing library of radiologically labeled diagnostic probes are utilized to gain new insights into dynamic health conditions (abnormalities that are not detectable using static visual analysis).

Basic Medical practice

- "Traditional tests" – the application of evidence based medicine to the utilization and interpretation of traditional tests has also provided the means to improve the accuracy and delivery of patient care.
- Infectious disease – many of the methods described above, especially nucleic acid-based technologies, have played a role in a revolution in how infectious diseases are detected, classified and managed. The management of HIV, for example, has provided many examples of the integration of genomic information into the appropriate selection of medication (1).

Each of these advances in diagnostic technology generates results that must be stored and presented in the Electronic Medical Record.

Clinical Information Systems - Survey

A survey of the variety of models available for developing and delivering clinical information systems will help set the scene for a discussion of both the benefits and anticipated challenges related to the use of CIS systems in support of personalized medicine.

A partial list of the components that comprise a comprehensive clinical information system is included below:

Registration – the means to enter a person into the system, capture basic demographics about the person, the nature of their specific visit and billing information

Electronic Medical Record (EMR) – the data repository storing results, orders, admissions, discharges, visits and all of the other detailed information about the interactions between a person and the system.

Electronic Chart – the means by which a provider (physician, nurse, technologist and other roles directly involved in patient care) interacts with the system, especially the EMR. For example, within the Cerner system this application is branded as "PowerChart™".

Ordering system – the technology through which orders for medications, procedures, tests and other activities are captured. Often a decision support engine is integrated with the ordering system, providing opportunities to intercept orders that are contraindicated by other information in the system.

Departmental systems – these provide functionality tailored to unique needs of the various departments in the clinical enterprise – the laboratory, pharmacy, Emergency Department and radiology department, for example, each have unique workflows and data requirements.

Billing system – manages the generation of financial statements, invoices, interactions with the diverse group of payors and other transactions.

A variety of models exist through which the components listed above can be associated. The models described below are representative examples, many hybrid models exist. In the interest of accuracy, only Cerner will be specifically named.

1. **Unified Architecture.** The strategy adopted by Cerner has been to design the components of a Clinical Information System to work from a common architecture and technology platform. Some key attributes of this approach are:
 - a. **Visibility** – information entered at one point in the clinical workflow, registration for example, is visible to all in the system with appropriate security access. This promotes personalized medicine by promoting a more positive patient experience and by reducing the risks associated with gaps in information.
 - b. **Transparency** – interactions with the system can be monitored, process inefficiencies or policy violations can more easily be identified and addressed.
 - c. **Reduce duplicate effort** – as a patient is moved through various areas, the need to recapture fundamental information, such as medication allergies. In addition to promoting a more efficient process, the reduction in effort also reduces patient frustration with repeated requests for the same information.
 - d. **Ease of support** – through the use of common core technologies, the training required to support the system is transferable across multiple disciplines.

- e. **Consensus** – the decision to adopt this approach requires that an organization have either initial consensus among the many stakeholders or evolving consensus (in which a core is installed and additional components are layered in).
2. **Interfaced.** This strategy involves the installation of components that operate in "stand-alone" mode and are connected via software interfaces to share information. Attributes of this approach include:
- a. **Limited visibility** – only information that can be successfully transferred via an interface is visible across modules. Through the "siloeing" effect created by this approach, personalized medicine is impaired as the care providers do not have a holistic perspective of the patient.
 - b. **Limited transparency** – performing analysis across disciplines and across systems becomes very cumbersome.
 - c. **Duplication of effort** – Any information that cannot be transferred across an interface must be re-entered multiple times, introducing opportunities for error.
 - d. **Difficulty of support** – A single change requires updating multiple systems. This approach can make it more difficult to cross-train personnel.
 - e. **Local decisions** – each stakeholder can make decisions based on their particular preferences. The institution must either provide support for integration or manage independent islands of narrow systems.

For each of these models, there are two primary means of implementing a Clinical Information System:

Develop internally – many organizations have developed systems that are 100% customized to their requirements. Often these systems perform very well until major technology advances require significant updates. Internal turnover among the staff, especially those who developed the systems, creates a high risk for the continuity of support.

Purchase commercially – Commercial systems are available that offer either the unified architecture (Cerner) or a set of interfaced solutions.

In reality, most healthcare organizations adopt a hybrid approach – they may have a core, unified system for their EMR and registration activities, for example, but interface to systems for departmental capabilities or billing.

Clinical Information Systems:

Technical considerations relevant to personalized medicine

Clinical information systems incorporate many sophisticated technical capabilities. Only those of direct relevance to personalized medicine are discussed below.

Software architecture

Clinical information systems have evolved in disparate manners. Some designs are more oriented to support personalized medicine than others. For example, systems began with a focus on billing and then layered in other functionality. These systems were designed with a transactional orientation and may lack the clinical emphasis needed to support personalized medicine. Other systems were designed to support the continuity of information related to the person about (and for) whom transactions are undertaken. While there are many architectural considerations that are involved in clinical information systems, some that are particularly significant for personalized medicine are:

Electronic Master Person Index (EMPI). EMPI is the ability to assign a unique identifier to a person and then use the identifier in multiple clinical delivery settings to promote the continuity of care across a wide geography. Seemingly obvious identifiers, such as Social Security Number, are inadequate as they are not reliably unique. The use of an EMPI is the basis from which many strategies related to the portability of healthcare information are derived.

In order to support the portability of medical information across providers and thus "personalize" the patient experience, an EMPI is a basic requirement. Cerner has implemented EMPI with national healthcare organizations and with organizations such as Detroit Medical Center and Aurora Healthcare in Wisconsin.

Decision support engine. A contemporary clinical information system should include an embedded decision support engine that allows users to build and maintain rules. The Arden syntax (7) provides the basis from which many decision support systems evolved. In general, clinical rules require triggers, evaluation criteria and responsive actions.

- **Triggers:** Events such as ordering a medication, opening a chart or other actions.

The system evaluates this order against patient results using a decision support rules engine.

Logic Section

If Using Templates Direct Scripting

the triggering request contains an order whose primary mnemonic is abacavir... whose ordering physician OPT QUALIFIER in OPT ORDDOC and FUTURE USE FUTURE USE OPT LIST

Action Section

Action : Using Templates Direct Scripting

A1 Send cancel/order alert **Discern Alert** stating HIV genotype analysis has indicated that the viral strain found in this patient may be resistant to Abacavir. Please review the results and reconsider the order. Alternative orders are listed below. with secondary button named OPT BUTTON NAME that links to OPT ITRI. ADDRESS disabled the override

The patient has a genotype result indicating resistance to Abacavir, so the system generates an on-screen alert:

Discern

Discern Alert

Patient : Hogan, Lucy

Order : abacavir

HIV genotype analysis has indicated that the viral strain found in this patient may be resistant to Abacavir. Please review the results and reconsider the order. Alternative orders are listed below.

Cancel previous order for abacavir

Add new order(s) for:

didanosine

stavudine

lamivudine

zalcitabine

OK

One of the key findings from this project was that discrete, codified molecular results would have greatly simplified the design and maintenance of these decision support rules (see below, coding systems). This project also provided insights into the level of granularity that is practical for use in decision support, for example building rules that evaluate allele or interpreted results is more practical than rules based on SNP findings.

Security

Many approaches to managing the security of genetic test results (and others relevant for personalized medicine) are currently available within commercial systems. The technical complexity of these capabilities makes sophisticated support for security one of the key advantages of a professionally developed and supported system. Examples of the security methods that are relevant include:

User authentication – in order to perform activities, users should be authenticated with a signon and password. Certain activities, such as the issuing of an electronic signature, should require re-authentication.

Organization security – only personnel appropriately affiliated with an organization have the ability to access results from patients affiliated with the same organization

Role-based security – only personnel in a role for which it is appropriate for them to view sensitive results have access to those results. For example, registration clerks and physicians may both work within the same system, but only physicians have access to patient results. Within a role, the classification system can be as simple or complex as deemed appropriate by an institution.

Task security – in addition to controlling which results a person has access to, the system can limit what actions may be performed by given personnel. For example, it may be appropriate for a laboratory technologist to review a result but they may not be able to correct the result.

Audit trail – every transaction should generate an auditable event (ie "Dr. Smith opened John Doe's electronic chart at 9:30am"). This is often a shortcoming of systems implemented through off-the-shelf software.

An active subject of discussion is whether or not genetic test results require security over-and-above those applied to other sensitive test results. For example, HIV test results are highly sensitive, inappropriate access to these results could result in harm to the patient.

All security implementation issues require considerable flexibility from clinical information systems suppliers. The approach taken by Cerner, for example, is to provide the capability to support the methods described above but allow each client site to determine whether and how to implement them. This allows for variance in interpretation of regulations and guidelines between user communities. It also provides flexibility for international users who may not be subject to the same regulatory or security guidelines.

Perhaps the most unique security concern related to personalized medicine is that not only do genetic test results have implications for the patient, but also for their immediate family members. Controlled exposure of these results to physicians caring for another family member could be very helpful in providing personalized care, but the risks are considerable. For example, many genetic tests results identify discrepancies between actual and believed paternity. A clinician could inadvertently disclose this discrepancy and trigger a cascade of unintended issues. Family members may not wish for their parents or children to become aware of a sensitive genetic test result (with Huntington's disease being the most appreciated example). Until guidance related to the controlled exposure of results for family members is issued, most clinical information systems suppliers are likely to be very cautious about including this capability in their systems, despite the significant benefits that could be provided.

Interface messaging

The messaging mechanism through which clinical systems communicate is Health Level 7[®] (HL7[®]), an international standard. Because most clinical information system implementations use a hybrid approach, the ability to accurately and completely transfer complex diagnostic findings between systems is very important to the continued development of personalized medicine. The format of the messages sent between systems is defined by HL7, which holds regular meetings of stakeholders. From the perspective of a commercial Clinical Information System supplier, the following points should be considered:

Currently, many commercial systems rely on HL7 version 2.7. Many initiatives, including the National Cancer Institute's Cancer Bioinformatics Grid (caBIG) and the HL7 genomics special interest group, have placed heavy emphasis on utilizing or enhancing the emerging HL7 standard – 3.0 (or RIM). The model for version 3.0 has only recently reached the stability required for commercial suppliers to make technology changes to support the new version(8;10). While version 3.0 has many appealing attributes and is likely to eventually become core to commercial systems, the practical adoption of this version could take a considerable time and should not be an assumption.

Limitations in interface messaging standards could become a rate-limiting deficit that impairs the ability of disparate clinical information systems to transfer information that is important to the delivery of personalized medicine.

Codification of diagnostic results

Codification of information promotes standardization, allows collaborators to exchange data with limited need for data mapping, allows distribution of standardized decision support rules and allows deeper semantic meaning to be captured without further specification. Medical vocabularies can be used to codify information about clinical orders and the results generated by diagnostic procedures or clinical observations. The Logical Observation and Identification Codes (LOINC) system, for example, is useful for standardizing information about many test orders (4). To the extent that personalized medicine will rely heavily on the use of pharmacogenetic testing, LOINC has incorporated some codes related to pharmacogenetic testing. For example:

```
"21661-4" "CYP2D6 gene mutation analysis del" "Prid" "Pt" "Bld/Tiss" "Nom" "Molgen" "CYTOCHROME
P450;CEBRISOQUINE HYDROXYLATION" "MOLPATH.MUT"
"21662-2" "CYP2D6 gene.1-BP del" "Arb" "Pt" "Bld/Tiss" "Ord" "Molgen" "CYTOCHROME
P450;CEBRISOQUINE HYDROXYLATION" "MOLPATH.MUT" "
```

The content of LOINC related to pharmacogenetic testing is relatively sparse and lacks the semantic relationships that would help support complex queries.

Another medical vocabulary that is widely used is the Systematic Nomenclature of Medicine (SNOMED®). The current version of SNOMED (CT) is structured as an ontology, with hierarchical relationships between terms. SNOMED is oriented toward the codification of results. However, when reviewed for concepts relevant for pharmacogenetics or genetics in general, SNOMED has relatively few concepts.

In the bioinformatics community, there are a variety of resources that are relevant to human genetics – the Gene Ontology (GO)(2), dbSNP (9) and others. These systems, however, were primarily designed to address research questions rather than to assist with the codification of clinically significant diagnostic results.

In order to address the gap in vocabularies, Cerner initiated the Clinical Bioinformatics Ontology (CBO)(3). The CBO provides coded values that are associated through a semantic network of relationships. The granularity of the CBO is intended to fall between dbSNP, which provides a code for every known SNP in the human genome, and the medical vocabularies cited above, which have a representative sampling. Clinical significance is a key determinant for whether or not a concept is included in the CBO. The CBO is maintained through a well defined curation process that applies consistent standards and naming conventions.

The CBO is relevant to advancing personalized medicine for the following reasons:

The CBO promotes the exchange of data through the use of common coded values.

It promotes the development of standardized decision support rules that can be packaged and distributed.

The CBO includes concepts for a wide variety of genes involved in drug metabolism, including: CYP2D6, CYP2C9, CYP2C19, VKORC1, TPMT and others. These concepts can be used to codify results that are captured in laboratory information systems or through other applications.

The CBO is distributed through an open license from www.clinbioinformatics.org in textual and RDF formats.

Software Application Capabilities

The software architecture topics described above are the platform upon which the front end applications used by clinicians are based. A variety of considerations are required to update these applications to better meet the needs of personalized medicine.

1. Order entry applications

- a. **Recognize "genetic duplicates"**. Unlike many laboratory tests, clinical genetic tests need only to be performed once, unless new variants are identified. Order entry applications should incorporate logic capable of intercepting test orders that are likely to yield duplicate information.
- b. **Medication orders in the pharmacogenetic context**. As demonstrated through the HIV pharmacogenetics project, medication orders can be intercepted based on a variety of clinical parameters, including pharmacogenetic test results. Consideration of user concerns, such as over-use of alerts, must be carefully weighed during the implementation of such alerts. As with security policies, the adoption of these rules will vary greatly among institutions and must be provided in a manner that is flexible. The use of pharmacogenomic information to evaluate medication orders can (and should) be combined with other information known about the patient, for example clinical information systems can also capture information about medication allergies and apply that information to the evaluation of medication orders.

2. Laboratory Systems

Laboratory Information Systems (LIS) were designed to support the data resulting from clinical pathology, surgical pathology, cytology, microbiology and blood bank work. The data generated in support of personalized medicine is quite different from the data in the traditional laboratory disciplines, as is the workflow used to generate diagnostic results.

- a. **Move from text to discrete results.** Most molecular diagnostic laboratories utilize text oriented pathology systems to capture and report their findings. Shifting to systems that are capable of storing molecular findings in a discrete format will offer these laboratories the ability to expose results to decision support logic, to perform data mining and analysis and to exercise greater control over the presentation and communication of the results.
- b. **Optimize workflow.** Molecular diagnostic laboratories have a unique workflow. Extending laboratory information systems to support this workflow will offer these laboratories the means to operate more efficiently.
- c. **Device communication.** A key challenge in the molecular diagnostic laboratory is transmitting orders from a clinical system to a molecular diagnostic device and/or retrieving results from that device. This is primarily due to the origin of these devices in the research community, for which formal medical device interface (MDI) specifications were not required.

The Cerner Millennium Helix™ solution was developed to address these concerns. By integrating with other laboratory modules, such as clinical pathology, it enables laboratories to better manage cross-discipline testing and promote a holistic perspective to support personalized medicine.

3. Presentation of results to clinician

- a. **Electronic chart.** Separation of the granular findings from the interpretative report is a key design requirement for presenting sophisticated diagnostic results to the clinician. Currently few clinicians are trained in the interpretation of SNP results, gene expression assays or other tests and rely heavily on the interpretative report. The nuances of precision and resolution between various testing methodologies are not always clear to the recipient of a molecular diagnostic report. One advantage of electronic review of diagnostic reports is that the results can be

associated with hot-links to external reference sites from which the recipient can learn more about the gene, disease or test.

- b. **Printed report.** Similar issues apply to the receipt and understanding of printed reports, but without the ability to automatically link out to external resources.
- c. **Decision support.** A key means to convey results in a timely manner is through the interception of an event which may require careful review of results. For example, a physician ordering mercaptopurine may be unaware that the patient had been genotyped to detect mutations in the TPMT gene. The system could invoke an alert with a link to a results viewer, forcing the ordering physician to examine the results and their interpretation before placing an order.

4. Image Management

- a. **Storage and retrieval.** High resolution images, whether showing MRI results or microarray findings, have significant storage requirements. Ensuring that these images are archived appropriately is an important consideration for personalized medicine.
- b. **Annotation.** Annotating a clinical image with appropriate metadata is an important consideration in making complex diagnostic information useful to the clinician.

Part 2 – Personalized Medicine – Data Analysis

Analyzing clinical data to support personalized medicine can be accomplished with two general goals.

1. To gain insights related to the operational delivery of care
2. To perform research

The goal of the data analysis will determine the security and privacy measures that must be taken. For example, operational analysis can be useful for strategic planning to determine how many test kits need to be ordered for the molecular diagnostics laboratory or to predict when traffic through the genetic clinic is likely to be heaviest. Within an organization, this type of analysis can routinely be performed within the constraints of the general security policies of the organization.

Research involving clinical data generally requires approval from an Institutional Review Board (IRB). If the data used has been completely scrubbed of patient identifiable information conforming with HIPAA guidelines, consent may not be required, otherwise patient consent is required to use individual information. In order to support consented research, a system must have the ability to control access to data for persons who have revoked their consent.

General Considerations for Clinical Data Warehouses

Both the operational and research data warehouses involve common processes and technologies.

1. **Extraction of data.** This is necessary because running research queries against a patient care system can impede performance and thus potentially create risks.
 - a. Data extraction can occur in "pulses" of HL7 messages in near-real-time or through periodic batch extractions of data. A common practice is to perform extractions during low activity times (midnight for example).
 - b. Depending on the intended use of the data, identifiers can be scrubbed (or excluded from the extraction).

Ideally, data captured in a clinical information system will have the following attributes:

Discrete – data is capture in numeric format or from a defined list of codified values.

Units of measure are consistently applied

Codified using consistent terminology – Discrete results are associated with a coding system that is used consistently between data contributors.

Corrected results clearly delineated from original results.

Few unambiguous results – the system clearly captures whether or not a question was asked versus defaulting to "no".

Unfortunately, clinical data warehouses do not often receive data that possesses these attributes. Significant effort must be invested in data mapping, reconciliation, removal of duplicate fields and other pre-analytical work before the researcher can reach meaningful conclusions using the data.

2. **Import data into data warehouse.** Often a data warehouse will have a schema that is optimized to support queries (rather than transactions). This import process reconciles the differences between data models, removes duplicate entries and performs other technical transformations.
3. **Query data.** The researcher can query the data. Access to data sets can be controlled through security policies.

A data warehouse designed to support studies related to personalized medicine will face many of the challenges inherent in any clinical data repository:

Data mapping. If multiple contributors submit data, disparate naming conventions must be reconciled to a common core. Some challenging issues that are inherent in genetic analysis:

- Reconciling variations in genetic panel composition. If one contributor includes 4 polymorphisms in her panel and another includes 5, how can the findings appropriately be compared?
- Do the contributing systems handle "not tested" in the same manner, for example does a blank field indicate a "normal" result or that the assay wasn't performed?
- The genetic tests performed to support patient care are significantly different from those that are useful in evaluating drug candidates.

Synchronization between parent system and data warehouse.

- If a result is corrected in the contributor system, how much time passes before that correction is conveyed to the data warehouse?

Analysis.

- The analysis of genetic data requires tools with specific capabilities designed to support genetic data sets. Correlating genetic data with other clinical information requires a considerable degree of flexibility.
- The analytical platform should support both exploratory (visual) and statistical analysis of data.

Clinical data sets to support personalized medicine

Clinical data sets have an important role to play in supporting the growth and adoption of personalized medicine.

Assess severity of genetic risk. Many of the widely anticipated benefits of prospective pharmacogenetic testing are based on anecdotal examples. Proxy data in a clinical warehouse can assist in determining the frequency of adverse events. This information, combined with the known population frequency of pharmacogenetic polymorphisms, can be useful in assessing the severity and frequency of genetically based adverse events.

Identify areas for genetic analysis. Clinical data sets can be mined for proxy data indicating adverse events. Preliminary stratification of patients in these data sets may be useful in determining whether or not there is a genetic contribution to the adverse event.

Assess clinician behavior. In a system in which orders are completely tracked, clinician behavior can be tracked. For example, how often do physicians order a particular medication even if that order is contraindicated by lab results? And in a multi-contributor environment comparative analysis can be performed to determine relative compliance. For the few institutions offering pharmacogenetic testing as part of routine patient care, this data can be valuable in determining how clearly results are communicated.

Example: The HIV Insights data warehouse captures anonymized information about HIV patients, including their viral load levels, HIV genotypes and medication orders. Retrospective analysis of this data revealed that a statistically significant group of physicians ordered

antiretroviral medications that were contraindicated by the HIV genotype (Uy, Hoffman, Baker, et. al. - manuscript submitted).

These examples should show how, with proper expectations, a clinical data warehouse can indeed support the further development of personalized medicine.

A representative clinical data warehouse

A representative multi-contributor system is the Cerner HealthFacts™ clinical data warehouse. This system receives anonymized data feeds from more than 50 clinical sites throughout the United States. The data can include diagnostic laboratory results, clinical orders, billing information and other discrete data captured by clinical information systems. HealthFacts data has been used to support a variety of analyses and is the basis for patient safety profile reports that are delivered to contributing organizations. These reports provide retrospective analyses indicating how often medication orders were placed despite contraindicating lab results or other information (age, gender etc.).

Non-clinical data management for Personalized Medicine

The development of new medications and the process of recognizing genetic contributions to medication response is generally managed through the clinical trials process. In a clinical trial, the trial sponsor can include tests and analysis that are not routinely performed in the delivery of patient care, such as comprehensive cytochrome p450 analysis, in order to gain a deeper insight into the response to the medication or therapy for which the trial is performed. Because the trial sponsor can stipulate data capture requirements as one of the objectives that must be met by participating investigators, they are much more likely to receive consistent and comprehensive data. Often clinical trials are performed in the same setting in which "routine" patient care is provided, creating the need to utilize common technology as often as possible while also creating the necessity to clearly differentiate trial participants from other patients.

The integration of clinical trials capabilities into the information systems utilized in the delivery of patient care is a key advance that will promote the development of personalized medicine. For example, the integration of patient recruitment capabilities based on a defined set of eligibility requirements into the overall clinical information will improve patient recruitment. Likewise, the clear indication within an electronic charting system that a patient is enrolled in a clinical trial will help promote the safe management of that patient as they undergo treatment. Cerner has developed and deployed many of these capabilities within the PowerTrials™ solution.

Knowledge bases, such as PharmGKB, (5), can provide reliable and detailed reference information about drug-gene associations. PharmGKB is primarily oriented toward the requirements of the research community.

Biospecimen registries/repositories

Increasingly biomedical researchers are building large collections of stored biological materials for population studies. These repositories have a set of common requirements:

Enrollment – In general, a patient must consent to contribute a sample to the repository. At the time this consent is provided, the patient is enrolled in the system.

Consent management – the system must manage the revocation of consent. It must also adequately manage variability in the scope of use of the sample.

Sample management – typically such systems provide inventory management capabilities as well as the ability to easily designate a sample for aliquoting or transfer to another investigator.

Sample annotation – annotation of the sample with demographic and clinical information enhances the value of the sample. For example, if users can search for samples that qualify for a complex set of criteria, that provides maximum value to their research efforts.

Biospecimen repositories should be integrated with clinical systems in order to maximally enrich the annotation of the specimens.

Part 3 - Delivering Personalized Medicine

Many of the challenges related to the ultimate realization of personalized medicine are rooted in the clinical culture and in policy.

How can the Standard of Care for a particular condition be updated in a timely manner to reflect emerging biomedical knowledge?

How can clinicians be trained to incorporate both the technology and philosophy of personalized medicine into their practice?

To what extent should prospective tests be reimbursed?

Can a consensus be reached on the clinical significance of pharmacogenetic findings?

To what extent should personal information that could impact that healthcare of another be shared?

What are the risks of applying emerging knowledge to the delivery of patient care?

What are the risks of not applying emerging knowledge to the delivery of patient care?

What are the implications of personalized medicine for genetic counseling?

While the answers to many of these questions will be determined by policy makers, educators and researchers; clinical information systems can help provide supporting data and can provide the means to intervene in the clinical process. Discrete data capture, leading to more comprehensive and accurate data analysis, can help gain insight into the population issues involved in personalized medicine. Analysis of the cost of prospective testing versus the cost of adverse outcomes will help payors determine the value of prospective testing. Automated decision support systems, deliberately maintained and updated, can serve as the vehicle for ensuring consistent use of complex guidelines. Easy access to on-line reference information can enable clinicians to stay current or to quickly investigate relevant supplemental information. Together, these capabilities will make significant contributions toward the realization of medicine that is more personal.

Executive Summary

We recommend the following topics and priorities for further consideration by AHIC.

A clearinghouse for the distribution of guidelines related to the clinically acceptable use of pharmacogenetic information in patient care should be developed. These guidelines should be distributed in machine-readable format to promote adoption and inclusion in clinical information systems.

The privacy and security considerations related to exposing healthcare information collected about one person to decision support logic utilized to benefit another should be formally evaluated.

EMPI methods should be used to promote the portability of personal medical information. National standards which contribute to the adoption of EMPI should be a priority.

Recommendations related to HL7 and support for personalized medicine should consider the practical implications and either begin with 2.7 or conduct the work in parallel (2.7 and 3.0) if they are to be feasibly implemented in the majority of currently installed clinical information systems.

Priority should be given to communicating results generated by current diagnostic practices (allele specific PCR and diagnostic sequencing) should be a priority before addressing technologies (gene expression microarrays, proteomics) that are not yet widely utilized in diagnostic practice.

Pilot projects designed to gather information comparing clinical outcomes of patients for which prospectively collected pharmacogenetic information was used in designing their care to a control group should be funded. For example, the length of stay for surgical patients whose pain management regimen is guided by CYP2D6 genotype versus a control group could be examined using a clinical data warehouse.

Pilot projects comparing the various methods of communicating complex information to non-specialist clinicians should be funded.

Reference List

1. Durant J, Clevenbergh P, Halfon P, et al. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet*. 1999;353:2195-9.
2. Harris MA, Clark J, Ireland A, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004;32 Database issue:D258-61.
3. Hoffman M, Arnoldi C, Chuang I. The clinical bioinformatics ontology: a curated semantic network utilizing RefSeq information. *Pac Symp Biocomput*. 2005;139-50.
4. Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. *J Am Med Inform Assoc*. 1998;5:276-92.
5. Klein TE, Altman RB. PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base. *Pharmacogenomics J*. 2004;4:1
6. Nebert DW. Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet*. 1999;56:247-58.
7. Pryor TA, Hripcsak G. The Arden syntax for medical logic modules. *Int J Clin Monit Comput*. 1993;10:215-24.
8. Schadow G, Mead CN, Walker DM. The HL7 reference information model under scrutiny. *Stud Health Technol Inform*. 2006;124:151-6.
9. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29:308-11.
10. Smith B, Ceusters W. HL7 RIM: an incoherent standard. *Stud Health Technol Inform*. 2006;124:133-8.

The contents of this response are intended for informational purposes only. The descriptions of general capabilities are not intended to be considered as formal descriptions of Cerner technologies.

Appendix 1 – Cerner Profile

Improving Healthcare Delivery...All Together

Cerner is the leading U.S. supplier of healthcare information technology solutions that optimize clinical and financial outcomes. Around the world, health organizations ranging from single-doctor practices to entire countries turn to Cerner for our powerful yet intuitive solutions.



Cerner offers clients a dedicated focus on healthcare, an end-to-end solution and service portfolio, and proven market leadership.

Healthcare Focus

Cerner's undivided focus on healthcare spans more than a quarter-century. The consistent leadership of our three founders cultivates enduring success in executing against a single mission: to connect the appropriate persons, knowledge and resources at the appropriate time and location to achieve the optimal health outcome.

Our 7,300 associates, including nearly 900 clinicians, develop clinical, financial and knowledge solutions in 36 facilities in 12 countries. With our cumulative investment in research and development expected to exceed \$2 billion by 2010, Cerner is the industry's most aggressive R&D investor. Our focus on research promotes year-over-year innovation that steadily extends our clients' technology investments.

End-to-End Solution and Service Portfolio

Cerner offers more than 60 solutions that span the healthcare continuum. Our broad and deep functionality satisfies individual, departmental and organizational imperatives. Following a structured design and implementation process that delivers predictable outcomes, Cerner's managed services help clients realize value at the lowest possible total cost of ownership. State-of-the-art intellectual and technology resources maximize performance, security and availability, guaranteeing uptime of at least 99.9 percent.

Proven Leader

Cerner is committed to forming aligned, collaborative relationships that foster our clients' long-term success. We focus on enhancing existing solutions and services while creating new ones that help health organizations overcome the industry's most pressing challenges. As a result, our clients have:

Decreased denials by 36 percent

Somerset Medical Center; Somerville, N.J.

Shortened code-to-bill time by 30 percent

Penn State Milton S. Hershey Medical Center; Hershey, Pa.

Accelerated medication delivery time by 70 percent

Children's Hospital and Regional Medical Centers of Seattle; Seattle, Wash.

Reduced pharmacy dispensing errors by 30 percent

Baystate Health System; Springfield, Mass.

Decreased claims turnaround from 1-2 months to 10-14 days

Creekwood Women's Care; North Kansas City, Mo.

All together, Cerner clients and associates are improving healthcare delivery around the world.