Clinical Genotyping: Near-Term Considerations for System-Wide Implementation, Operations, Adoption, and Data Management for Improving Drug Safety

Michael D. Kane, Ph.D.*

Assistant Professor of Computer and Information Technology. Lead Genomic Scientist, Bindley Bioscience Center, Discovery Park. West Lafayette, Indiana. Purdue University (765) 494-2564 mdkane@purdue.edu

expertise: genomic technology, bioinformatics, pharmacology

John A. Springer, M.S.

Assistant Professor of Computer and Information Technology. West Lafayette, Indiana. Purdue University (765) 496-7582 jaspring@purdue.edu

expertise: databases, data management, data warehousing

Jon E. Sprague, R.Ph.,Ph.D.

Dean, College of Pharmacy, Professor of Pharmacology. Ohio Northern University. Ada, Ohio (419) 772-2277 j-sprague@onu.edu

expertise: drug safety, pharmacology, pharmacy education

*corresponding author; bio-sketches of authors are included.

Preface

This manuscript describes the rationale and overall system-side behavior of a near-term clinical genotyping information system (both DNA analysis and data management) involving an early implementation in pharmacogenomics and drug safety assurance. The challenges to implementing a successful clinical genotyping system are described, and how the benefits of a focused, near-term system for drug safety assessment overcome the logistical and operational challenges that perpetually hinder the development of a societal-scale clinical genotyping system. *The long-term goal of implementing a drug safety program that utilizes genomic data is to improve patient care and safety while at the same time facilitating the movement of clinical genotyping from bench to bedside.* The authors are established collaborators in genomic and pharmacological research, and are eager to contribute to the development of a successful clinical genotyping system.

Introduction

The utilization of a patient's genetic data to aid diagnostic and prognostic healthcare represents the ultimate achievement of 50 years of genomic research. The technology to recognize this vision has emerged, and continues to evolve. In the future, patient-specific genomic data will be derived before birth and include an exhaustive sampling of genomic information. This genetic data will be periodically updated throughout a patient's lifetime on a tissue-specific basis in order to screen for genetic changes conferring age-related diseases. The patient's genotypic data will further be integrated with dedicated databases/warehouses harboring genetically-linked health and adverse drug response risk that will be utilized at the point-of-care for patient-specific therapeutic interventions. Yet, the path to this future in genomic-based healthcare is obscured by several independent factors that must be recognized and overcome to fully exploit genomic content in human healthcare. The authors have identified the following categorical hindrances to a societal-scale implementation of clinical genomics:

1) **High-Throughput DNA Analysis Technology: Costs, Data Standards and Future Technologies.**
2) **Information Management: Access, Security and System Structures.**
3) **Genomics & Genetics Education: Physicians, Pharmacists, Nurses and Consumers.**
4) **Point-of-Care Utilization of Genomics: Physician's Office, Hospital, Pharmacy and Consumer.**
5) **Capitalism & Pharmaceuticals: Risks and Returns on Investment in Genomic-based Laboratories and Information Systems.**
6) **Electronic Health Record Management and Utilization.**
7) **Translational Research: Establishing Linkages Between Allelic Information and Healthcare Outcomes.**

This RFI response discusses operational solutions to some of the issues listed above, and a summary of each issue is provided in the introduction. Many of these issues are addressed (directly or indirectly) later in the manuscript as near-term operational systems are described and rationalized. Ultimately the authors are interested in the many aspects of a implementation of a near-term system for clinical genotyping including (1) the information management system and data standards required for system implementation, (2) the interface between DNA analysis biotechnology and clinical genotyping information system, (3) the management of costs and opportunities to insure the success of a clinical genotyping system, (4) and the educational impact of the post-genome era on pharmacy students and other professions within healthcare.

1) High-Throughput DNA Analysis Technology: Costs, Data Standards and Future Technologies. Certainly there are numerous analytical methods for DNA analysis that support both SNP (single nucleotide polymorphisms) discovery and SNP detection (SNP discovery and detection represent distinct analytical challenges that are not described in this manuscript), and competition within the biotechnology industry continues to advance these capabilities from all relevant perspectives (cost, throughput, data quality, ease-of-use, etc.). Yet at the heart of a large-scale clinical genomics implementation is an information management system that can accommodate many different analysis methods (including new biotechnologies that will emerge in the future) through the development of a group of scalable data standards for genomic information. Although significant advances in biotechnology are occurring, the data standards for sharing genomic data will precede genotyping in the clinic.

2) Information Management: Access, Security and System Structures. Included in this manuscript is the rationale for an information system that categorically separates SNP data relevant to drug safety from SNP data relevant to general health outcomes. The authors feel strongly that this distinction is key to implementation of a successful clinical genotyping system. By categorically separating SNPs relevant to drug safety from SNPs linked to other health outcomes and SNPs with no known linkages (it is recognized that there is some small overlap in this distinction), consumers can: (1) ***understand*** how their own genomic data is being utilized and gain trust in these systems, (2) ***indicate*** how their own genomic data is managed and who can gain access to these categorical data sets, and (3) ***provide*** a rationale for security that is dependent upon the category of the data. For example, drug safety data may be more easily accessed by worldwide healthcare institutions and pharmacies since these data may be needed in an emergency for an injured traveler. In contrast, other SNP categories are stored much more securely and are NOT shared across institutions. This concept assumes that consumers will be (1) able to control access to their genotypic information and (2) SNPs inherent to drug safety are far less likely to serve (or be abused) as indicators of general health for an individual.

3) Genomics & Genetics Education: Physicians, Pharmacists, Nurses and Consumers. Given the very recent advances in human genomic knowledge and biotechnology methods, it is not feasible to assume that physicians, pharmacists, nurses, and other professionals within the healthcare industry harbor sufficient knowledge to translate raw genomic data to information relevant to health outcomes. This being said, the authors have made two conclusions regarding the near-term future of clinical genomics. First, essentially all genomic data will be filtered into categorical definitions and the known (or

potential) impact of a given SNP will be presented to the healthcare professional (described below in Table 1). For example, if a patient is prescribed a drug where an adverse response has been associated with one or more specific genotypes, then the patient's electronic health record (EHR) will simply indicate that the patient is "at risk for an adverse response due to genomic information" and make a recommendation to choose an alternate drug (and provide an alternate drug if one is available) and/or reduce the dose of the drug. This specific example, as well as many others, has recently been demonstrated in the clinical literature. Drugs are metabolized endogenously by a series of enzymes collectively referred to as the cytochrome P-450 system. These enzymes are further characterized into sub-groups named CYP1A1, CYP2D6, etc. Meur et al., (2006) demonstrated that the metabolic activity and oral clearance of the immunosuppressant, sirolimus, is significantly decreased in patients with CYP3A5*3 single-nucleotide polymorphism and further suggested that prior dose adjustments should be made in patients with this SNP. However, the technology for routinely implementing such a dose adjustment does not currently exist. The goal of this RFI is to rectify this current short-coming. Secondly, the authors have concluded that the initial commercially-viable implementation of a clinical genomics system will involve drug safety issues and be administered through pharmacy prescription systems. This initial implementation of a pharmacogenomic system will utilize SNPs that have an established link to drug safety outcomes and therefore can include information-based guidance to patients harboring SNPs relevant to drug safety (i.e. decision support for both the physician and pharmacist), exploit a prescription/dispensing system that is already guided by an information system, and inherently does NOT involve SNPs poorly linked to disease risk and/or does not provide insight on how a physician or pharmacist should alter treatment. Furthermore,

this near-term implementation will provide a cultural shift in pharmaceutical drug development whereby new drug indications can require genomic screening to ensure safety and efficacy, ultimately involving clinical drug development (phase I-IV) to be limited to patients with specific SNP genotypes to increase the overall safety and efficacy of new drug entities.

4) Point-of-Care Utilization of Genomics: Physician's Office, Hospital, Pharmacy and Consumer. This issue continues the rationale for a near-term implementation of clinical genomics in drug safety by allowing pharmacists to be the proprietors of genomic information and exploiting the interconnectivity of pharmacy information systems to allow access to patient genomic information across the country. As many more SNPs are ultimately derived for each patient, a more secure healthcare information system will include SNPs relevant to disease predisposition as they are established through translational research. As discussed later, the translational research that involves linking known SNPs to healthcare outcomes will be facilitated through the use of the near-term implementation genotyping system.

5) Capitalism & Pharmaceuticals: Risks and Returns on Investment in Genomic-based Laboratories and Information Systems. The implementation of a drug safety clinical genomic system provides the best overall return on investment for the healthcare community in the near-term. This is because the system utilizes SNPs that have an established link to drug safety outcomes and therefore can include information-based guidance to patients that possess SNPs relevant to drug safety (i.e. decision support for both the physician and pharmacist), exploit a prescription/dispensing system that is

already guided by an information system, and provide a cultural shift in pharmaceutical drug development whereby new drug indications can require genomic screening to increase the overall safety and efficacy of new drug entities.

6) Electronic Health Record Management and Utilization. The utilization of an EHR is a very new concept in healthcare, and although the benefits of an information management system that utilizes EHRs is essentially well recognized, the implementation of EHR systems has not been widely successful to date. Therefore it may be too ambitious to assume that the inclusion of clinical genomic data in current EHR systems will be inherently successful. The authors feel that the overall usefulness and impact of genotypic information in the clinic (from both the consumer and healthcare provider perspective) must precede a wide-spread system implementation. This further rationalizes the system described below, where SNPs relevant to drug safety as utilized in the pharmacy (and pharmaceutical industry) represents the ideal introduction of genotypic information in our healthcare system.

7) Translational Research: Establishing Linkages Between Allelic Information and Healthcare Outcomes. This logistical barrier to the overall impact of genotypic information in the clinic involves a disparity between discovering (or uncovering) linkages between known SNPs and human health, which requires a large collection of known SNPs from a wide variety of patients (including their health records within one or more data standards), and a method upon HOW to rationalize the collection of known SNPs from a wide variety of patients. In other words, statistically significant linkages between known SNPs and health outcomes can only be achieved if a large collection of

SNPs from normal and 'diseased' patients is available for data mining. Furthermore, this requires that the disease-relevant information and other meta data types be available within data standard formats to allow for data mining, which is the fundamental structure of an EHR. The authors feel strongly that the near-term drug safety system that integrates known SNPs with prescription drug indications will facilitate the acquisition of many other known SNPs that are NOT relevant to drug safety for the purposes of epidemiological research. In other words, patients undergoing genotyping for drug safety will have the option (ideally with incentives) to be genotyped for thousands of other known SNPs within their own genome to facilitate health outcomes research, ultimately to benefit themselves and society. As discussed below, this will involve an anonymous contribution of SNP and EHR data to a specialized data management system dedicated to identifying SNP-based risk assessment through the discovery of statistically significant linkages to other health outcomes such as diabetes, cancer, mental disorders, age-related disorders, etc. This concept gives rise to an oversight committee that governs data mining and statistical methods to establish "accepted" links between SNPs and health outcomes, and "approves" new linkages as they are discovered, proven and published.

Data management can be viewed along two perspectives, where the overall concept of "informational hierarchy" is used to describe both data concepts and data schemas (moving left to right in Table 1), which then define levels of information access (privacy & security) and levels of bioinformatics knowledge (raw biotechnology data to DNA sequence to protein sequence to physiological effect). This informational hierarchy (Table 1) is also organized vertically (top to bottom) to depict data transformations from raw data (biotechnology and DNA analysis data) into usable information (bioinformatics) and comprehensible knowledge (impact on human health).

Table 1.

Information Hierarchy within a Clinical Genotyping Information System

Conceptual Perspective	Data Schema Perspective	Access & Privacy Perspective	Bioinformatics Perspective
Data	Physical Data	Laboratory	Raw Data (DNA)
Information	Conceptual Data	Bioinformatics Statistics Epidemiology	DNA Sequence Bioinformatics Genomics & Ontology
Knowledge	User View	Healthcare Data Management Applications	Protein Sequence Enzyme Biochemistry Receptor Biochemistry
Comprehension	Consumer View	Patient	Physiological Effect Impact on Healthcare

CLINICAL GENOTYPING for DRUG SAFETY

System-Wide Operations

Patient-Controlled Access

The ethical concerns to genotyping in the clinic, which are also applicable to electronic health records in general, are essentially privacy and security. The benefits of incorporating genotyping (genetic information) in therapeutics and medicine are questioned when the risk of 'information abuse' is considered. For example, a patient may be unwilling to utilize the benefits of genotyping if they fear that their employer and/or insurance provider can utilize the same information to (accurately or inaccurately) predict the patient's future health status. This dilemma involves both societal and genetic components. At the genetic level, the validity of extrapolative health assessment based solely on genotypic data has not been broadly established, and is limited to a few known genetic diseases. Therefore any long-term claims to health status for the majority of the

population would be invalid at this point in time. Yet, it should be noted that *the risk of adverse drug response based on known SNPs in drug metabolism enzymes has been established (see table 2), and represents the near-term benefit to clinical genotyping.*

Table 2*.
Representative Drug Metabolizing Enzymes Associated with Adverse Drug Response

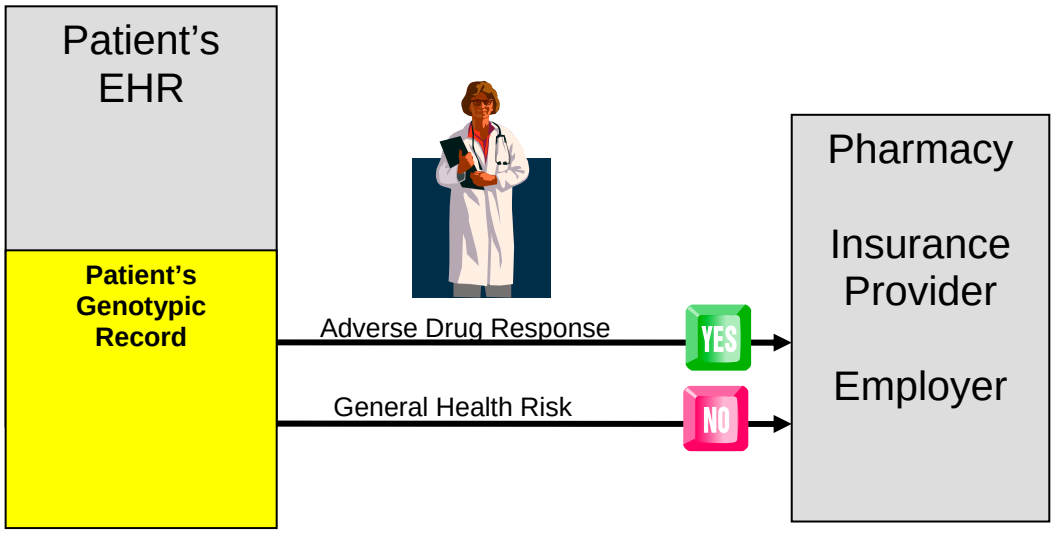
CYP Family	Allele	Nucleotide Change	Enzyme Activity Change	Impact on AUC**	Common Drugs Affected
1A2	CYP1A2*1C	-3860 G>C	Decreases (Nakajima, 1999)	Increases	Aminophylline Betaxolol Caffeine Flutamide Propranolol
2C9	CYP2C9*3A	1075 A>C	Decreases (King, 2004)	Increases	Amiodarone Carvedilol Fluoxetine Glimepiride Warfarin
3A4	CYP3A4*18A	878 T>C	Increases (Dai, 2001)	Decreases	Atorvastatin Carbamazepine Clarithromycin Diltiazem Losartan

*This is not an exhaustive representation of known SNPs relevant to drug metabolism and safety, and only three representative CYP families, a single SNP example, and 5 potential drugs affected are presented in this table for conservation of space. **AUC is the area under the curve of a drug's concentration in human plasma over time.

Furthermore, it is important to note that the use of the term SNP (single nucleotide polymorphism) herein includes nucleotide base substitutions and single base deletions/substitutions within the human genome. In addition, knowledge of this predisposition does not represent association with other health risks. Thus knowledge of the risk of adverse drug response is a benefit to the patient, employer and insurance provider since overall healthcare costs would be minimized by avoiding adverse drug reactions. Allowing the patient to control external access to their genotypic data within this categorical distinction (e.g. "adverse drug response risk" data access = yes; "general

health risk" data access = no) will positively contribute to the adoption and success of genotyping in the clinic (See figure 1), which is a natural artifact of utilizing the hierarchy described in Table 1 (above).

Figure 1. The patient controls outside access to genotypic data based on how the data is used (categorically). In this example, the patient has allowed access to SNP data corresponding to adverse drug response risk, yet prohibited access to SNP data known (or unknown) to be relevant to overall disease and general health risk. Note that the patient's genotypic record is regularly updated with new genetic data from the RISK database, and the categorical limited access to the health record will allow access to new adverse drug response RISK data.



Human Genotypic Database

The genotypic data derived from large numbers of people and patients will be crucial to establishing new genetic links to health risk. In this paradigm, it is assumed that a large cohort of patients will be genotyped across thousands of known single nucleotide polymorphisms that include SNPs that have established links to the risk of adverse drug responses and disease, as well as SNPs that currently have no known association with human health outcomes. The discovery of one or more SNPs that is associated with a specific phenotype or disease risk requires that a large Human Genotypic Database (HGD) is derived from individual genotypic records, which includes all other aspects of their health records. For example, the discovery of SNPs that are linked with

cardiovascular disease involves a statistical comparison of SNPs between a large group of patients experiencing cardiovascular disease and a large control (disease free) group. In practice, this involves the derivation of a HGD where the patient identifiers have been removed (achieving privacy through anonymity) that include both genotypic and overall health information for each person, which is a natural artifact of utilizing the hierarchy described in Table 1 (above).

RISK Database

The incorporation of health "risk" data, which is the known risk associated with each SNP position, into a patient's genotypic record must be temporary and periodically updated to reflect new discoveries and linkages. This dynamic component to the electronic health record reflects the fact that future discoveries may link known SNPs to one (or more) health outcomes, and in the absence of an updatable risk component a patient's genotypic record will become outdated and thus underutilized. For example, a patient may have data on a specific genotype (SNP or set of SNPs, in a specific genomic location) that, to date has been considered benign and represents no known risk, yet new research findings have determined that the SNP constitutes some level of health risk. Therefore, the most recent date and method by which an individual patient genotypic record has been updated to insure (1) that the most timely genotypic risk and population frequency data has been incorporated into the record and (2) insure that outdated genotypic records are updated (this assumes an application automatically updates the record, and utilizes a time/date stamp to manage updates) (Figure 2). This notion of state is easily handled by the database in the information system.

Figure 2. Example of the SNP-specific risk components of a patient's genotypic data, and how it may change using updates that reflect new discoveries from linkage studies. Note that this figure is an amalgam of different data types to demonstrate the behavior of the system, and is NOT meant to represent how information is presented to the healthcare professional and/or patient.

A. John Doe's genotypic component of the EHR.

```
[Patient_ID: John Doe]
[Most Recent Update: 11/6/2008]
[Most Recent Update Method: Genosoft 5.6]
[Most Recent Update Source: NIH Human SnipRisk Database, release 263.2]
{Genotype Data}
[SNP Position: ID_1] [Zyg: A, A] [Popul_Freq: 22%] [Risk: Unknown]
[SNP Position: ID_2] [Zyg: T, G] [Popul_Freq: 09%] [Risk: Unknown]
[SNP Position: ID_3] [Zyg: C, C] [Popul_Freq: 78%] [Risk: Unknown]
[SNP Position: ID_4] [Zyg: T, T] [Popul_Freq: 94%] [Risk: Unknown]
[SNP Position: ID_5] [Zyg: G, G] [Popul_Freq: 48%] [Risk: Unknown]
[SNP Position: ID_6] [Zyg: A, C] [Popul_Freq: 32%] [Risk: Cardio, LOW]
[SNP Position: ID_7] [Zyg: C, T] [Popul_Freq: 70%] [Risk: Unknown]
[SNP Position: ID_8] [Zyg: T, T] [Popul_Freq: 29%] [Risk: Unknown]
.
.
.
```

B. The EHR Update using the RISK database incorporating a new discovery at SNP Position "2". In this example, SNP position "2" has been linked to an adverse drug response risk.

```
[Patient_ID: John Doe]
[Most Recent Update: 11/6/2008]
[Most Recent Update Method: Genosoft 5.6]
[Most Recent Update Source: NIH Human SnipRisk Database, release 263.2]
{Genotype Data}
[SNP Position: ID_1] [Zyg: A, A] [Popul_Freq: 22%] [Risk: Unknown]
[SNP Position: ID_2] [Zyg: T, G] [Popul_Freq: 09%] [Risk: Drug, HIGH]
[SNP Position: ID_3] [Zyg: C, C] [Popul_Freq: 78%] [Risk: Unknown]
[SNP Position: ID_4] [Zyg: T, T] [Popul_Freq: 94%] [Risk: Unknown]
[SNP Position: ID_5] [Zyg: G, G] [Popul_Freq: 48%] [Risk: Unknown]
[SNP Position: ID_6] [Zyg: A, C] [Popul_Freq: 32%] [Risk: Cardio, LOW]
[SNP Position: ID_7] [Zyg: C, T] [Popul_Freq: 70%] [Risk: Unknown]
[SNP Position: ID_8] [Zyg: T, T] [Popul_Freq: 29%] [Risk: Unknown]
.
.
.
```

Clearly the management of a genotypic RISK database becomes useful as the central source for determining SNP-specific risk will be managed separately and must be subject to scientific and regulatory oversight. This genotypic risk database will include all known

SNPs, and their known frequency within the population in the human genome along with all known health risk information associated with each SNP.

Genotypic Testing Considerations

The deployment of genotyping technology in the clinic requires that results from laboratory tests (regardless of the genetic assay platform) be effectively managed for the benefit of patients and the general population. Unlike most laboratory tests used in the clinic, the results of genotyping tests must be stored in a patient-specific database (utilize patient identifier) due to the large number of potential data points (SNPs) from a single test, as well as contribute to population-scale database (anonymous identifier). Clearly, the first application of genotyping technology is aimed at surveying drug metabolism enzymes to identify patients that are deficient in drug metabolism activity, which leverages knowledge that specific SNPs are known to confer this phenotype and testing will be limited to these SNPs. The overarching logic to this approach is that a specific SNP is first associated with a clinically-relevant phenotype, and then deployed as a clinical test. Yet the association of known SNPs with clinically-relevant phenotypes can (and must) also be determined retrospectively. The population-scale database will reflect the growth of both the number of patients (people) contributing genotype database, and the number of SNPs assayed from each person's genome, and ultimately represent a resource linking genetics with public health informatics. In this approach a collection of known SNPs is assayed and stored in a population-scale database, which also includes (anonymous) data from the patient's healthcare record. This provides a resource (database) to discover linkage between specific SNP(s) and clinically-relevant phenotypes, ultimately linking genotypic data to specific phenotypes.

The data captured from clinical genotyping must include patient identification, genotypic data, and other aspects associated with patient-specific sampling, but also accommodate the integration genotypic data not collected in earlier genotyping tests, information about the testing method, quality control data, as well as the emergence of new technologies involved in testing and data management. Finally, the data must be integrated with a supporting (dynamic) database system that communicates health risks associated with each genotype. Given that the emergence of disease and drug adversity risk with each genotype may be dependent on other genotypic/phenotypic factors, or may simply not yet be known or fully understood, the conversion of genotypic data to health risk must be separate from the patient genotypic data record. The following is a list of data requirements for the genotypic data record;

1) Patient Identifier
2) Sample Source/Tissue
3) Age of Patient at Sampling
4) Genotypic Data
5) Genotyping/Laboratory Method
6) Quality Control Method
7) Ethnicity, Gender and Existing Genetic Considerations
8) Most Recent Date (and Method) of RISK Data Integration

In addition to the patient's identifier, data must include the source of the genetic material being tested (#2). Potential genetic factors may be tissue specific, such as genetic variability associated with oncogenesis (e.g. normal tissue vs. cancerous tissue), which are certainly crucial, if not the motive, for genotyping. In addition, contaminating genetic material (e.g. bacterial, contaminating human genetic material) may be present in skin samples or mucosal secretions may be considered as a component of the quality control methods (#6), and can be captured in the sample source data. Additionally, the age of the patient is needed for genotypic comparisons made for the patient later in life (#3). As

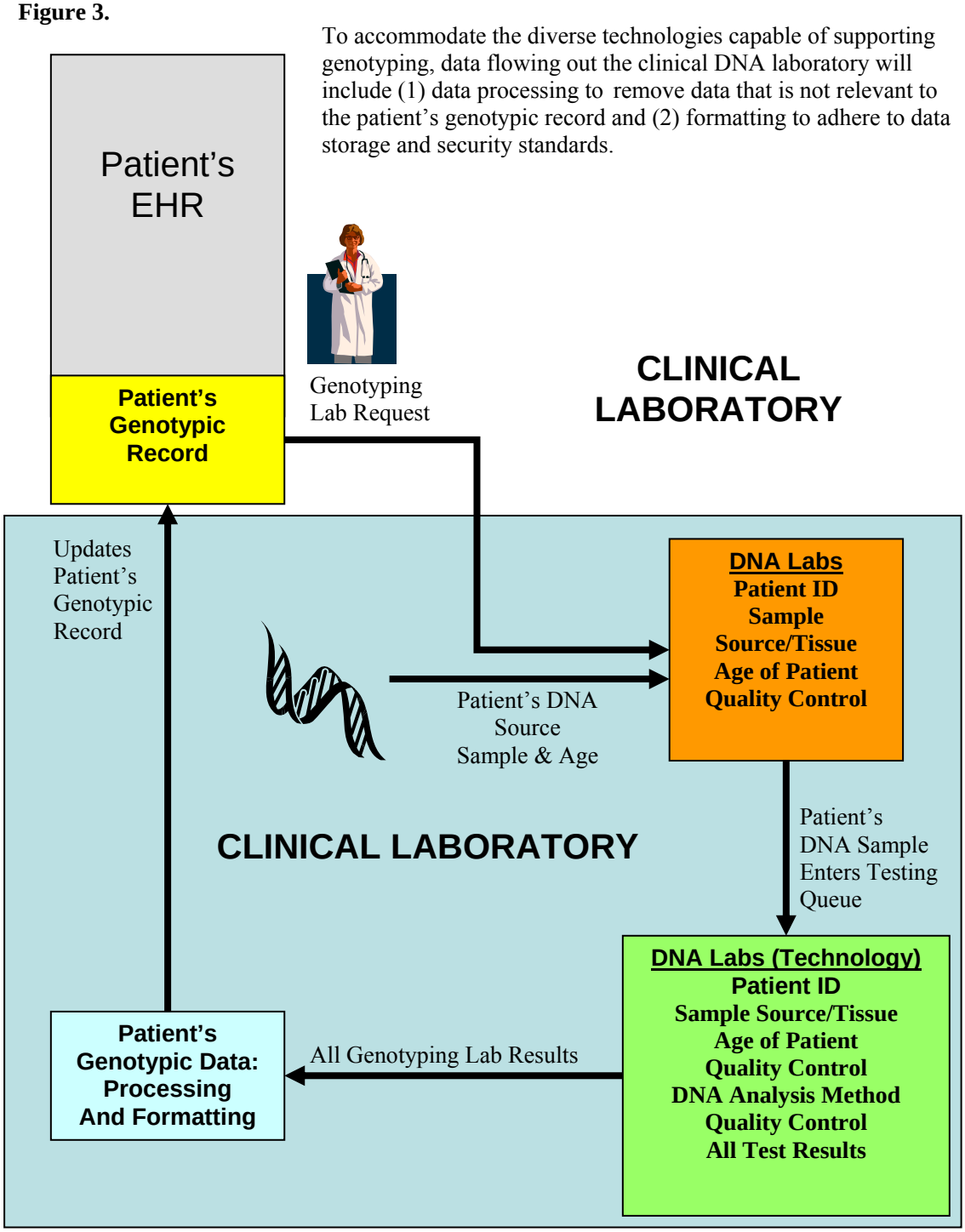
mentioned earlier, many methods for genotyping already exist and the emergence of new technologies in this arena is certain. Therefore the method used for a specific data collection/test must be captured, as well as the testing laboratory, personnel involved, and any other relevant information about the location and technology employed. The methods employed to insure the sample and the laboratory test was performed correctly will contribute to a quality control determination, and will utilize both genomic sequence and assay standards added to the sample under investigation. Knowledge of an existing genetic condition, such as trisomy 21, will result in triploid data (rather than the expected diploid data) for all genotypic data derived from genetic material on chromosome 21. Finally, given the proposed paradigm that allows the genotypic record to be updated with new risk information, the date of the most recent comparison between the patient's genotypic record and the risk database must be stored (in the patient's record) to insure risk assessment is based on all data available (#8).

Genotypic Data Standards and Data Sources

The data relevant to a patient's genotype will include nucleotide base identification and zygosity at each SNP position, and could include flanking genomic sequence information (depending upon the technology employed). For example, using DNA microarray technology for genotypic screening will be essentially limited to homozygous or heterozygous data for a given SNP position, while genotypic data derived from direct DNA sequencing will provide potentially hundreds of bases of DNA flanking one or more SNPs, which represents a large string of DNA sequence that can be captured. The genotypic data capture must be recognized within the context of the technology or

method utilized, and the method or technology utilized must be identified within the genotypic data record (see Figure 3).

Figure 3.

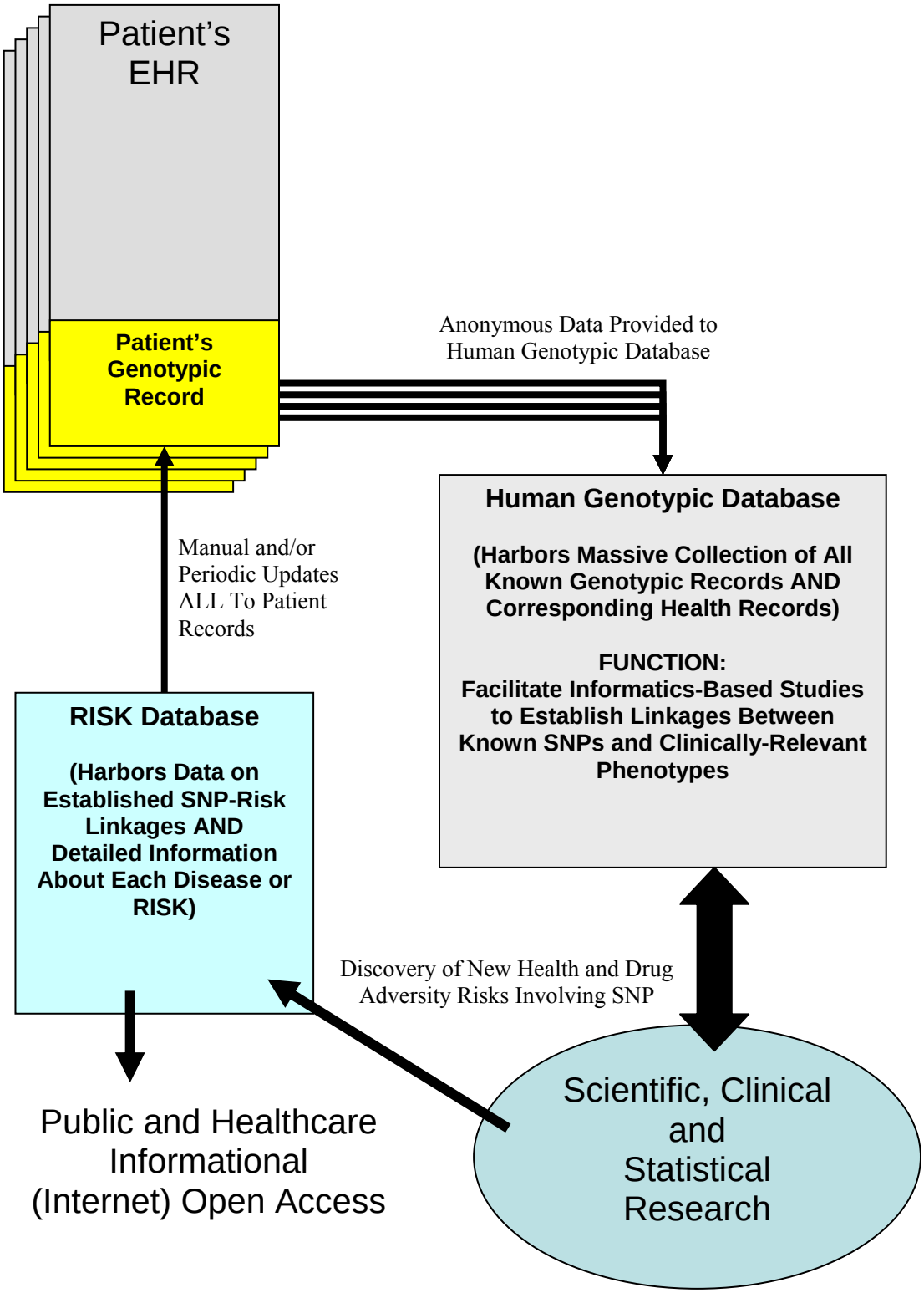


This is not meant to infer that any given method is more sensitive or specific, but rather that results are sometimes technology or method dependent. This is somewhat analogous to the utilization of positron emission tomography (PET) and magnetic resonance imaging (MRI), where results from both tests provide similar insight into the phenotype (phenomena), yet the actual laboratory results are derived from distinct methods. In the case of DNA sequencing, or genotypic data derived from more data rich sources, the DNA sequence data must be pared down to the SNP(s) that are present (maintained) in the database of risk linkages. Thus the method of genotyping includes both a categorical description of the biotechnology component (in this case, capillary electrophoresis) and a raw data analysis component (conversion of fluorescent-specific peaks to DNA sequence, and elimination of DNA sequence that does not constitute SNP data). Instances where a given patient harbors a rare genetic condition that is not amenable to SNP-level data must be considered as additional information of the patient, and not a component of a system-wide genotypic data record format.

Genotypic Information System

The general architecture of the clinical genotyping information system is represented in figure 4. The process of DNA testing is described in Figure 3, ultimately deriving or updating patient-specific genotypic data. Once the patient's record has been updated, the data is available for contribution to the Human Genotypic Database (HGD). As mentioned earlier, the HGD represents a key source for human genetic research capable of establishing new levels of risk to all known SNPs.

Figure 4. System-wide integration of the patient's EHR and its genotypic components, the Human Genotypic Database, research activities, the RISK Database, and the Internet.



In addition, once the patient's record has been updated the system accesses the RISK database to determine if the patient's updated SNP profile includes specific genotypes associated with a known health risk. Some level of overall health risk will be established, which will likely include categorical classifiers such as either "common" (benign or unknown risk), "drug" (adverse drug risk) or "health concern" (some level of overall health risk). These categorical definitions of risk will likely have a simple quantitative component (e.g. low, moderate or high risk) that will be used by the clinical system to flag the attention of healthcare workers and other system components.

Conclusions

Many factors will influence if and how people will derive their genotypic information including: genotyping test costs, privacy and ethics, as well as the overall cost-benefit of genotyping information. The cost-benefit of genotypic information is dependent upon the rigor of predicting clinically-relevant phenotypic traits based on SNP data. Definitive genetic testing may be tenuous given that every nucleotide in the genome is (theoretically) subject to variance, yet the current strategies for genetic testing are limited to testing for the most common mutations that are known to confer a health risk. For example, there are over 900 mutations in the human genome shown to cause cystic fibrosis (CF), yet most genetic testing laboratories limit their testing to the 6 most common mutations, and have predictive success rate of 90% in Caucasians (Gregg, 2002). Using current genetic testing systems, it is not feasible to test for all known mutations that cause CF given (1) the benefit of predicting or diagnosing CF from a genetic test does not justify the costs associated with testing hundreds of known mutations from a patient's sample, and (2) that there is a chance that a (rare) specific

polymorphism, which has not yet been characterized, can cause CF and would not be detected in a large-scale genetic testing screen. It can be expected that any genotyping strategy will be sensitive to false-negative results given that rare SNPs that are not tested under a given genotyping screen may confer a health risk phenotype.

Deriving sufficient patient information for a large-scale clinical genotyping system will initially involve a large population of patients with mature healthcare records that contain information regarding age-related conditions and diseases, where patient-specific genomic information can be added upon sampling/testing. Ideally, a near-term implementation of clinical genotyping will involve the *addition* of patient-specific genomic data to an *existing* healthcare information management system. Certainly there are many established healthcare groups and systems that are well positioned to benefit from the proposed near-term clinical genomics systems, and partnering with one or more of these groups will both (1) leverage the data and resources inherent to that system and (2) reduce implementation costs by reducing system redundancies. For example, the Veterans Administration (VA) hospital's healthcare information management system allows for patients to be screened for drug-drug interactions, patient allergies, past medical history etc. Incorporation of the genomic data base into this type of healthcare information management system would allow pharmacists point of care access to genetic information that will be beneficial in making therapeutic decisions. The VA system further has a limited drug formulary and a captive patient population that lends itself well to beta-testing the clinical genomic system. By starting with a small population, we can then move to the large-scale clinical genomic system to be implemented not only in hospitals but other pharmacy practice settings. *In conclusion, the implementation of a drug safety program that utilizes genomic data to improve patient care and safety while*

at the same time facilitating the movement of clinical genotyping from bench to bedside will improve general healthcare outcomes.

Literature Cited:

Meur Y, Djebli N, Szelag J, Hoizey G, Toupance O, Rerolle J, Marquet P (2006). CYP3A5*3 influences sirolius oral clearance in de novo and stable renal transplant recipients. Clin Pharmacol Ther 80:51-60.

Nakajima M, Yokoi T, Mizutani M, Kinoshita M, Funayama M, Kamataki T. (1999) Genetic polymorphism in the 5'-flanking region of human CYP1A2 gene: effect on the CYP1A2 inducibility in humans**.** J Biochem (Tokyo). 125(4):803-8.

King BP, Khan TI, Aithal GP, Kamali F, Daly AK. (2004) Upstream and coding region CYP2C9 polymorphisms: correlation with warfarin dose and metabolism. Pharmacogenetics. 14(12):813-22.

Dai D, Tang J, Rose R, Hodgson E, Bienstock RJ, Mohrenweiser HW, Goldstein JA. (2001) Identification of variants of CYP3A4 and characterization of their abilities to metabolize testosterone and chlorpyrifos. J Pharmacol Exp Ther. 299(3):825-31.

Gregg AR, Simpson JL. Genetic screening for cystic fibrosis. Obstet Gynecol Clin North Am. 29(2): 329-40 (2002).

NAME	POSITION TITLE
Kane, Michael David	Assistant Professor, Department of Computer and Information Technology. Lead Genomic Scientist, Bindley Bioscience Center. PURDUE UNIVERSITY

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Illinois State University, Normal, IL	BS	1988	Biological Sciences
Purdue University, West Lafayette, IN	PhD	1996	Molecular Pharmacology
Parke-Davis Pharmaceuticals, Ann Arbor, MI	Post-Doc	1998	Genomics & Neuroscience

A. Positions and Honors.

Positions and Employment

05/88-08/91 **Research Associate**, Pharmacokinetics and Drug Metabolism, Abbott Laboratories, Abbott Park, IL.
08/91-08/96 **Doctoral Student & Research Fellow**
08/96-08/98 **Post-Doctoral Fellowship,** Parke-Davis Pharmaceuticals.
08/98-07/00 **Senior Scientist**, Microarray Technologies Facility, Parke-Davis Pharmaceuticals.
07/00-07/02 **Vice President, Research and Development**, Genomic Solutions, Inc.
07/02-08/04 **President and Co-Founder**, Nucleico, LLC.
08/04-present **Assistant Professor**, Department of Computer and Information Technology.
08/04-present **Lead Genomic Scientist**, Bindley Bioscience Center, Purdue University.

Awards and Other Professional Activities:

1988 Abbott Laboratories Presidential Award Recipient.
1995 Purdue Research Foundation Fellowship Recipient.
1995 American Association of Pharmaceutical Scientists, Graduate Student Research Competition Award Recipient, Indianapolis Chapter.
1995 Sigma Xi Scientific Research Society, Purdue University Chapter, Graduate Student Research Award Recipient.
1996 Jenkins/Knevel Excellence in Research Award Recipient, School of Pharmacy and Pharmacal Sciences, Purdue University.
2006 Outstanding Faculty in Discovery Award, College of Technology, Purdue University
2006 Seed for Success Award, Purdue University

B. Selected peer-reviewed publications:

1. Kane, M.D. and Brewer, J.L. An Information Technology Emphasis in Biomedical Informatics Education. *Journal of Biomedical Informatics* 40: 67-72, 2007.

2. Kim, H., Kane, M.D., Kim, S., Dominguez, W., Applegate, B.M, and Savikhin, S. A Molecular Beacon DNA Microarray System for Rapid Detection of E. coli O157:H7 that Eliminates False Signal Risk. *Biosensors and Bioelectronics* 15:1041-1047, 2007.

3. Kemp, M.Q., Liu, W., Thorne, P.A., Kane, M.D., Selmin, O., Romagnolo, D.F. Induction of the transferrin receptor gene by benzo[a]pyrene in breast cancer MCF-7 cells: Potential as a biomarker of PAH exposure. *Environmental and Molecular Mutagenesis* 47: 518-526, 2006.

4. Gough, E.S. and Kane, M.D. Evaluating Parallel Computing Systems in Bioinformatics. Third International Conference on Information Technology: New Generations, IEEE Computer Society, 2006.

5. Hansen, R. and Kane, M.D. Critical Performance Variables in Genomic Sequence Analysis in Cluster Computing Architectures. TeraGrid '06: Advancing Scientific Discovery, 2006.

6. Kane, M.D. Technical Issues in DNA Microarray Production and Utilization: Impact on Clinical Research. *Expert Review of Molecular Diagnostics* 5(5): 649-654, 2005.

7. Kane, M.D. Aligning experimental design with bioinformatics analysis to meet research objectives. *Cytometry* 47: 50-51, 2002.

8. Kane, M.D. Screening for xenobiotic-mediated P450 induction using a dedicated DNA microarray protocol. *Journal of Clinical Ligand Assay* 24: 149-151, 2001.

9. Sprague, J.E., Worst, T.J., Haynes, K., Mosler, C.R., Nichols, D.E. and Kane, M.D. The pharmacodynamic characterization of an antisense oligonucleotide against monoamine oxidase-B (MAO-B) in rat brain striatal tissue. *Cellular and Molecular Neurobiology* 21: 53-64, 2001.

10. Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D, and Madore, S.J. Assessment of the sensitivity and specificity of oligonucleotide microarrays. *Nucleic Acids Research* 28: 4552-4557, 2000.

C. Genomics/Bioinformatics Patents & Book Chapters

Patents

1. Kurnit, D.M. and Kane, M.D. Systems, Methods and Compositions for Detection of Human Papilloma Virus in Biological Samples. 2006.
2. Kane, M.D. and Applegate, B.M. A DNA Microarray System for Rapid Detection of Foodborne Pathogens. 2006.
3. Kane, M.D., Nagel, A.C., and Dombkowski, A.A. Compositions and Systems for Identifying and Comparing Expressed Genes (mRNAs) in Eukaryotic Organisms. (US patent #6,955,876), 2000.

Book Chapters (invited)

1. Kane, M.D., Farquhar, I., Summers, K.H. Clinical Information Systems Development: Integrating Genomics and Bioinformatics within Healthcare Information Systems. In: Farquhar, I. and Summers, K. (eds) *The Value of Innovation: Impacts on Health, Life Quality, and Regulatory Research* (in press).

2. Kane, M.D. Introduction to Gene Expression Profiling with DNA Microarray Technology. In: Hardiman, G. (ed) *Microarray Methods and Applications: Nuts & Bolts, Volume 2*. DNA Press (in press).

3. Kane, M.D., Lin, H-S, Dombkowski, A.A., Diakiw, A., Nagel, A.C., Flexon, A., Leifheit, M., Sprague, J.E. and Hollenberg, P.F. Accelerating Preclinical Drug Development Using a Microarray-Based Method for Detecting Oxidative Enzyme (P450) mRNA Induction and Other Metabolic Genes in the Rat. In: Hardiman, G., (ed). *Microarray Methods and Applications: Nuts & Bolts*. DNA Press p. 231-248, 2003.

4. Nagel, A.C., Thompson, J.D., Marino, M.A., and Kane, M.D. Spanscript: A Novel Method for Deriving Non-Redundant 3' cDNA Libraries for DNA Microarray Production. In Hardiman (Ed.), *Microarray Methods and Applications: Nuts and Bolts*, DNA Press p. 155-170, 2003.

5. Kane, M.D., Dombkowski, A.A., and Madore, S.J. The Emerging Utility of Oligonucleotide Microarrays. In: Lu, Q. and Weiner, M., editors. *Gene Cloning and Expression Technologies*. Eaton Publishing. p. 537-547, 2002.

NAME	POSITION TITLE
Springer, John Andrew	Assistant Professor Department of Computer and Information Technology Purdue University

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Taylor University, Upland, IN	BS	1995	Mathematics
Indiana University, Bloomington, IN	MS	2002	Computer Science
Indiana University, Bloomington, IN	PhD	2007 (Exp.)	Computer Science

A. Positions and Activities

08/95-07/98	**Consultant**, Andersen Consulting, Indianapolis, IN
08/98-01/99	**Systems Analyst**, Eli Lilly and Company, Indianapolis, IN
01/99-08/99	**President**, Infinite Loop Consulting, Indianapolis, IN
08/99-05/04	**Associate Instructor**, Indiana University, Bloomington, IN
05/04-07/06	**Implementation Engineer**, Bostech Corporation, Indianapolis, IN
07/06-present	**Assistant Professor**, Department of Computer and Information Technology, Purdue University, West Lafayette, IN

B. Selected peer-reviewed publications

Martin, R., Robertson, E., & Springer, J. (2004). Architectural Principles for Enterprise Frameworks. In *CAiSE 2004 Workshops in connection with The 16th Conference on Advanced Information Systems Engineering* (pp. 151-162). Riga, Latvia: Faculty of Computer Science and Information Technology, Riga Technical University.

Springer, J. (2003). An Exploration of Enterprise Architecture Framework Views. In *Proceedings of the 1st International Conference on Enterprise Information Systems (ICEIS) Doctoral Consortium (DCEIS-2003)* (pp. 46-48). Angers, France: ICEIS Press.

C. Book Chapters

Martin, R., Robertson, E., & Springer, J. (2005). Architectural Principles for Enterprise Frameworks: Guidance for Interoperability. In P. Bernus & M. Fox (Eds.), *Knowledge Sharing in the Integrated Enterprise: Interoperability Strategies for the Enterprise Architect* (pp. 79-91). Boston, MA: Springer Boston.

NAME	POSITION TITLE
Jon Eric Sprague	Dean & Professor of Pharmacology

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Ferris State University, Big Rapids, MI	B.S.	1989	Pharmacy
Purdue University, West Lafayette, IN	Ph.D.	1994	Pharmacology /Toxicology

A. Positions and Honors.

2006-present Dean & Professor of Pharmacology, School of Pharmacy, Ohio Northern University, Ada Ohio

2006 Chair & Professor of Pharmacology, Virginia College of Osteopathic Medicine, Virginia Tech University, Blacksburg, Virginia

2004-2006 Chair & Associate Professor of Pharmacology, Virginia College of Osteopathic Medicine, Virginia Tech University, Blacksburg, Virginia

1998-2004 Associate Professor of Pharmacology, Department of Pharmaceutical & Biomedical Sciences, School of Pharmacy, Ohio Northern University, Ada, Ohio

1995-1998 Assistant Professor of Pharmacology, Department of Pharmaceutical & Biomedical Sciences, School of Pharmacy, Ohio Northern University, Ada, Ohio

1997-2000 Visiting Scientist, Parke Davis, Neurological and Neurodegenerative Diseases, Ann Arbor, MI

1994-1995 Visiting Assistant Professor of Pharmacology, Department of Pharmacology and Toxicology, School of Pharmacy and Pharmacal Sciences, Purdue University, West Lafayette, Indiana

1991-1995 Staff Pharmacist, Wells Yeager Best Pharmacy, Lafayette, Indiana

1989-1991 Staff Pharmacist/Assistant Manager, Hook Drugs, Columbia City, Indiana

B. Selected peer-reviewed publications (mentored students names are underlined).

1. Peterson AL, Gilman TL, Banks ML, Sprague JE. Hypothyroidism alters striatal dopamine release mediated by 3,4-methylenedioxymethamphetamine (MDMA, Ecstasty). Synapse. 59(2):317-319; 2006.
2. Rusyniak DE, and **Sprague JE**. Toxin-induced hyperthermic syndromes. Med. Clin. N. Am. 89(6):1277-1296, 2005. (invited review).
3. **Sprague JE**, Moze P, Caden D, Rusyniak DE, Holmes C, Goldstein DS, Mills EM. Carvedilol reverses hyperthermia and attenuates rhabdomyolysis induced by 3,4-methylenedioxymeth-amphetamine (MDMA, Ecstasy) in an animal model. Crit Care Med. 2005 Jun;33(6):1311-6.
4. **Sprague JE**, Nichols DE. Neurotoxicity of MDMA (ecstasy): beyond metabolism. Trends Pharmacol Sci 26(2):59-60, 2005.
5. Rusyniak DE, Tandy SL, Hekmatyar SN, Mills E, Smith DJ, Bansal N, Maclellan D, Harper ME, **Sprague JE**. The role of mitochondrial uncoupling in 3,4-methylenedioxymethamphetamine mediated skeletal muscle hyperthermia and rhabdomyolysis. J Pharmacol Exp Ther. 313(2):629-39; 2005.
6. **Sprague, J.E.**, Mallet, N.M., Rusyniak, D.E., and Mills, E. UCP3 and thyroid hormone involvement in methamphetamine-induced hyperthermia. Biochem. Pharmacol. 68(7):1339-1343, 2004.
7. Hrometz, S.L., Brown, A.W., Nichols, D.E., and **Sprague, J.E**. 3,4-methylenedioxymethamphetamine (MDMA,ecstasy)-mediated production of hydrogen peroxide in an in vitro model: the role of dopamine, the serotonin-reuptake transporter, and monoamine oxidase-B. Neurosci Lett. 367(1):56-59, 2004

8. **Sprague, J.E**., Brutcher, R.E., Mills, E. M., Caden, D., and Rusyniak, D.E. Attenuation of 3,4-methylenedioxymethamphetamine (MDMA, Ecstasy) induced rhabdomyolysis with α1- plus β3-adrenoreceptor antagonists. Br. J. Pharmacol. 142(4):667-670, 2004.
9. Rusyniak, D.E., Banks, M.L., Mills, E. M., and **Sprague, J.E**. Dantrolene use in MDMA (Ecstasy) mediated hyperthermia. Anesthesiology. 101(1):263-264, 2004.
10. Rusyniak, D.E., Scruggs, S.L., Kamendulis, L.M., Sprague, J.E., and Klaunig, J.E. The effects of ecstasy (MDMA) on rat liver bioenergetics. Acad. Emer. Med. 11(7):723-729, 2004.
11. Mills, E.M., Banks, M.L., **Sprague, J.E**., and Finkel, T. Uncoupling the agony from ecstasy. Nature. 264:403-404, 2003.
12. **Sprague, J.E.**, Banks, M.L., Mills, E., Rusyniak, D.E. 3,4-methylenedioxymethamphetamine (MDMA, ecstasy) induced changes in striatal temperature and uncoupling protein (UCP) expression. Soc. Neurosci. Abst. 2003.
13. **Sprague, J.E.,** Banks, M.L., Cook, V.J., and Mills, E. Hypothalamic-Pituitary-Thyroid axis and sympathetic nervous system involvement in the hyperthermia induced by 3,4-methylenedioxymeth-amphetamine (MDMA, Ecstasy). J. Pharmacol. Exp. Ther. 305(1):159-166, 2003.
14. **Sprague, J.E**., Preston, A.S., Leifheit, M., and Woodside, B. Hippocampal Serotonergic Damage induced by MDMA (Ecstasy): Effects on Spatial Learning. Physiol. Behav. In Press.
15. Kane, M.D., Hsia-lien, L., Dombkowski, A.A., Diakiw, A., Nagel, A.C., Flexon, A., Leifheit, M., **Sprague, J.E**., and Hollenberg, P.E. A Genomic-Based Method for Detecting P450 mRNA Induction: Implications for Xenobiotic Screening and Drug Development. Drug Development Research. In Press.
16. Kanthasamy, A., **Sprague, J.E.,** Shotwell, J.R., and Nichols, D.E. Unilateral infusion of a dopamine transporter antisense into the substantia nigra protects against MDMA-induced serotonergic deficits in the ipsilateral striatum. Neuroscience. 114(4):917-924; 2002.
17. Waites, N.S., and **Sprague, J.E.** Pharmacology and Toxicology of Ecstasy. U.S. Pharmacist. June 2002.
18. Falk, E.M., Cook, V., Nichols, D.E and **Sprague, J.E.** Antisense targeted at MAO-B attenuates rat striatal serotonergic neurotoxicity induced by MDMA. Pharmacol. Biochem. Behav. 72:617-622; 2002.
19. **Sprague, J.E.**, Leifheit, M., Selken, J., Milks, M.M., Kinder, D.H., and Nichols, D.E. In vivo Microdialysis and Conditioned Place Preference Studies in Rats are Consistent with Abuse Potential of Tramadol. Synapse. 43:118-121, 2002.
20. **Sprague, J.E**. Teaching Cardiac Arrhythmias: A focus on pathophysiology and pharmacology. Amer. J. Pharm. Educ. 65:169-177, 2001.
21. **Sprague, J.E**.; Worst, T.J.; Haynes, K.; Mosler, C.R.; Nichols, D.E.; Kane, M.D. The pharmacodynamic characterization of an antisense oligonucleotide against monoamine oxidase-B (MAO-B) in rat brain striatal tissue. Cellular Molecular Neuro Bio. 21(1):53-64, 2001.
22. **Sprague, J.E.;** Harrod, A.D.; Teconchuk, A.L. The pharmacology and abuse potential of ephedrine. Pharmacy Times. 64(5):72-80, 1998.
23. **Sprague, J.E**.; Everman, S.L.; Nichols, D.E. An Intregrated hypothesis for the serotonergic axonal loss induced by 3,4-methylenedioxymethamphetamine. Neurotoxicol. 19(3):427-442, 1998.
24. **Sprague, J.E**.; Johnson, M.P.; Schmidt, C.J.; Nichols, D.E. Studies on the mechanism of p-chloroamphetamine neurotoxicity. Biochem. Pharmacol. 52(8):1271-1277; 1996.
25. Maickel, R.P.; **Sprague, J.E**. Role of ACTH fragments in alcohol consumption. In: Stress, Gender, and alcohol-seeking Behavior. Editors: W.A. Hunt & S. Zakhari. NIH Publications, 1995.
26. Marona-Lewicka, D.; Rhee, G.S.; **Sprague, J.E**.; Nichols, D.E. Psychostimulant-like effects of para-fluoroamphetamine. Eur. J. Pharmacol. 287:105-113; 1995.
27. Marona-Lewicka, D.; Rhee, G.S.; **Sprague, J.E**.; Nichols, D.E. Reinforcing effects of certain serotonin-releasing amphetamine derivatives. Pharmacol. Biochem. Behav. 53(1):99-105; 1996.
28. **Sprague, J.E**.; Nichols, D.E. The monoamine oxidase-B inhibitor L-deprenyl protects against 3,4-methylenedioxymethamphetamine-induced lipid peroxidation and long term serotonergic deficits. J. Pharmacol. Exp. Ther. 273(3):667-673;1995.
29. **Sprague, J.E**.; Nichols, D.E. Inhibition of MAO-B protects against MDMA-induced neurotoxicity in the striatum. Psychopharmacol. 118:357-359;1995.
30. **Sprague, J.E**.; Huang, X.; Kanthasamy, A.; Nichols, D.E. Attenuation of 3,4-methylenedioxy-

methamphetamine (MDMA)-induced neurotoxicity with the serotonin precursors tryptophan and 5-hydroxytryptophan. Life Sci. 55(15):1193-1198;1994.

31. **Sprague, J.E**.; Sharp, D.; Maickel, R.P. Cholinergic regulation of ethanol consumption in rats. Res. Commun. Chem. Pathol. Pharmacol. 15(1&2):1-8;1994.
32. **Sprague, J.E**.; Maickel, R.P. Effects of stress and Ebratide (Hoe-427) on free-choice ethanol consumption: comparison of Lewis and Sprague-Dawley rats. Life Sci. 55(11):873-878;1994.
33. Zabik, J.E.; **Sprague, J.E.**; Binkerd, K. Central and peripheral components of the inhibitory actions of 5HTP on ethanol consumption in the rat. Pharmacol. Biochem. Behav. 47(3):547-551;1994.
34. **Sprague, J.E**.; Fitzgerald, A.L.; Suckow, M.A.; Grigdesby, C.F.; Uetrecht, R.L.; Maickel, R.P. Pharmacological studies of centrifugation-induced analgesia. Pharmacol. Biochem. Behav. 46:911-915;1993.
35. Zabik, J. E.; **Sprague, J.E**.; Odio, M. Interactive dopaminergic and noradrenergic systems in the regulation of thirst in the rat. Physiol. Behav. 54(1):29-33;1993.