Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Cynthia Davison August 1, 2018

Genetic Data Integration:
A Model for Clinical Implementation
And Intervention Research

By

Cynthia Davison
Master of Public Health

Applied Public Health Informatics

_____
Laura M. Gaydos, Committee Chair

_____
J Mark Conde, Committee Member

Genetic Data Integration:
A Model for Clinical Implementation
And Intervention Research

By

Cynthia Davison

Master of Business Administration
Georgetown University
1985

Thesis Committee Chair: Laura M. Gaydos, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Applied Public Health Informatics
2018

Abstract

Genetic Data Integration:
A Model for Clinical Implementation
And Intervention Research

By Cynthia Davison

Opioid addiction is a multifactorial condition for which there is growing evidence of a genetic contribution and thus vulnerability to abuse. Concurrent with the growth of this epidemic, is significant advancement in genomic sequencing, cloud-based services, predictive analytics, big data storage and retrieval, and emerging computing technologies - a concomitant growth that is enabling an ever-widening scope of genetic inquiry and thus application. Given this technological landscape, we can construct a model that integrates genetic data into a cloud-based platform, facilitates use of next generation and emerging applications, and contributes to the growth of evidence-based treatment and genomic knowledge. As a foundation for development, the MeTree study platform, developed by Duke University and sponsored by the National Institutes of Health, serves as a starting point.

As currently constructed, the IGNITE (Implementing GeNomics In pracTicE) MeTree project platform offers elements from which to architect a clinico-genomic decision support platform that incorporates modern technologies. The MeTree platform collects family health history (FHH) data, links to a patient electronic health record (EHR) database, and provides clinical decision support (CDS) to providers and patients using guidelines-based recommendations for individuals at risk of developing common chronic diseases. It supports the SMART-on-FHIR standard, an HL7 data access and management platform, and, thus, it can employ the SMART-on-FHIR genomic profile for incorporating genetic data into an EHR system. Adoption of that standard creates the opportunity for an application-driven, microservices architected model for disease study and clinical use, the original concept behind SMART-on-FHIR development. Since prior studies suggest that dopamine receptor genes are prime candidates for the study of genetic variants and their effects on opioid dependence vulnerability, the investigation and inclusion of specific abuse-related variant information along with EHR, and FHH data has the potential to expand the current understanding of genetic addiction vulnerabilities, and genotype-phenotype associations. And, from a public health perspective, the inclusion of genetic data and application of predictive analytic techniques as presented in this model, can point the way for intervention strategy development for many multivariate disease types, and for the future implementation and practice of precision medicine.

Genetic Data Integration:
A Model for Clinical Implementation
And Intervention Research

By

Cynthia Davison
Master of Business Administration
Georgetown University
1985

Thesis Committee Chair: Laura M. Gaydos, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Applied Public Health Informatics
2018

Introduction

Opioid addiction is a multifactorial condition for which there is growing evidence of a genetic contribution and thus vulnerability to abuse.[1-5] Concurrent with the growth of this epidemic, is significant advancement in genomic sequencing, cloud-based services, predictive analytics, big data storage and retrieval, and emerging computing technologies - a concomitant growth that is enabling an ever widening scope of genetic inquiry and thus application.[6] These advances have implications for greater understanding and identification of our genetic identities and stand poised to shape individual health care delivery in a new and revolutionary way. Given this technological landscape, we can construct a model that integrates genetic data into a cloud-based platform, facilitates use of next generation and emerging applications, and contributes to the growth of evidence-based treatment and genomic knowledge. As a foundation for development, the MeTree study platform, developed by Duke University and sponsored by the National Institutes of Health, serves as a starting point.

As currently constructed, the IGNITE (Implementing GeNomics In pracTicE) MeTree project platform offers application elements from which to build a clinico-genomic decision support platform that incorporates modern technologies. The MeTree platform collects family health history (FHH) data, links to a patient electronic health record (EHR) database, and provides clinical decision support (CDS) to providers and patients using guidelines-based recommendations for individuals at risk of developing common chronic diseases. It supports the SMART-on-FHIR standard,[7] an HL7 data access and management platform, and, thus, it can employ the SMART-on-FHIR genomic profile for incorporating genetic data into an EHR system. Adoption of that standard creates opportunity for an application-driven, microservices architected model for disease study and clinical use, the original concept behind SMART-on-FHIR development.[8] Since prior studies suggest that dopamine receptor genes are "prime candidates for the study of genetic polymorphisms and their effects on opioid dependence vulnerability," [5] the investigation and inclusion of specific abuse-related variant information along with electronic health record, EHR, and FHH data has the potential to expand the current understanding of genetic addiction vulnerabilities, and genotype-

phenotype associations. And, from a public health perspective, the inclusion of genetic data and application of predictive analytic techniques, can point the way for intervention strategy development for many multivariate disease types, and for the future implementation and practice of precision medicine.

Background

Opioid substance abuse disorders are complex, multifactorial conditions. Beginning in the 1990's, with the debut of OxyContin and the proliferation of medical marketing campaigns, health officials began recording a staggering increase in morbidity and mortality associated with prescription opioid use, and morphine and heroin. Today, more than 115 Americans die every day after overdosing on opioids, according to the latest figures from the Centers for Disease Control and Prevention (CDC). And, more than 630,000 people have died from a drug overdose between 2000 and 2016, a majority due to opioid abuse. [9] In 2015, it is estimated that 2 million people in the US suffered from substance abuse disorders related to prescription opioid pain relievers. [10] And, in that year alone, costs due to the epidemic were pegged at $504 billion, or 2.8 percent of GDP,[11] a figure that accounts for healthcare spending, and loss of productivity. Of particular significance to clinical practice, studies show that prescription opioids are a pathway to illegal drug use. It is estimated that about 20 to 29 percent of patients who are prescribed opioids abuse them, and, that about 80 percent of heroin users misused prescription opioids first. [10]

Compounding this epidemic is a limited, and arguably inadequate, understanding of effective treatment and prevention strategies. Given that opioid medications offer health benefits as well as present risks, the prevention and reduction of prescription drug misuse presents a major challenge. [12] Prevention strategies are more "restrained," as a result, and less subject to evaluation. [12] The CDC cautions that opioids should be used only when the benefits for pain relief and function are expected to outweigh risks.[13] Yet, the organization's 2016 Guideline for Prescribing Opioids for Chronic Pain, gives no clear guidance on how to measure that tradeoff. The twelve recommendations in the report are voluntary and, sadly, as noted, the clinical scientific evidence informing the recommendations is low in qualityS.[13] In sum, the report suggests that non-opioid

therapy is the preferred strategy, and more research is necessary to fill in critical evidence gaps.

It is worth noting that there are specific risk assessment tools currently in clinical use for measuring individual abuse vulnerability, though reception appears mixed. In large part, these are survey tools, designed to assess behavioral patterns based on patient responses in order to determine risk. The most frequently recommended instruments for risk assessment screening include: Opioid Risk Tool (ORT), the Structured Clinical Interview for DSM-IV (SCID), and the Current Opioid Misuse Measure (COMM), and the Screener and Opioid Assessment for Patients with Pain-Revised (SOAPP-R). Together, these form a disparate group of questionnaires and patient administered surveys. (The ORT survey, for example, is a six-item questionnaire designed to predict the risk of problematic drug-related behaviors. A score of 8 or higher, given that one question can have multiple responses, is considered a high risk for opioid misuse).[14] The CDC finds fault overall with these tools due to a lack of available research to support clinical usage, and, specifically with the ORT and SOAPP-R tools, for inconsistency in predicting opioid abuse or misuse. In one study cited; for ORT, sensitivity was 0.58 and specificity 0.54 ; for SOAPP-R, sensitivity was 0.53 and specificity 0.62.[13] Not surprisingly then, their usefulness in determining patient abuse vulnerability, and guiding medical prevention and treatment, is characterized at one end of the spectrum as no better than 50%, [15] or, the flip of a coin.

Opioid Abuse Research

To date, research on opioid abuse underscores its multifactorial nature and indicates associations of dependence behavior with a number of factors including; family history of substance dependence, socioeconomic status, physical health, and specific genetic influences related to pain perception and dopamine receptor interactions. [5,16] The relationship is complex and can involve comorbidity with mental health and environmental factors,[17] the extent of which varies and depends on the individual. Overall, these factors in combination invite further scrutiny.

In studies, opioid abuse research has shown that:

- In terms of demographics, individuals who are younger, white, male, Alaskan/Am. Indian, and live in rural vs. urban settings, are more at risk.[18] People aged 18 to 25 years have the highest rates of chronic nonmedical use of analgesic opioids and heroin use. [18] And, opioid analgesic death rates are highest among people aged 45-54 years. [18] Interestingly, the rate of heroin abuse is highest among 25-34 year-olds. [18]
- Risk factors for opioid abuse include: mental health issues, poor physical health, history of substance abuse, family instability, and socio-economic level.[19] For young adults in particular, studies point to external factors as strong influencers in the development of substance abuse. The presence of a family member with an abuse disorder can significantly influence the childhood development and likelihood that a young person will struggle into adulthood, with emotional, behavioral, or substance use problems. [20] And, the disorder can manifest in young adults as early antisocial behavior, depression and anxiety. [21]
- Heritability also appears to be a factor in studies that do not include genomic information. Several studies of twins have indicated that a large fraction, about half, of the risk for opioid addiction is genetic in origin. [22,23] In one study, relatives of probands, or starting points in a family tree, that have opioid disorders were ten times more likely to have opioid abuse disorders themselves. [5,20]
- In genetic studies, dopamine and opioid receptor genes have emerged as "prime candidates in the study of genetic polymorphisms" and their effects related to opioid abuse and dependency. [5,24] These genes are associated with pain perception and reward behavior. Neurotropic related genes have also been implicated. And, while genetic influences and effects are complicated, variations in four genes; DRD2, OPRM1, OPRD1, and BDNF, consistently, have shown associations with substance abuse. [5] These genes encode receptors and the signaling molecules that play important roles in opioid abuse disorders. [5]

In consideration of the multivariate nature of opioid addiction, its human toll, and the need for intervention and prevention strategies, a technologically advanced data system should enable providers, researchers, and policy makers, to draw on multiple data

sources, in conjunction with electronic health records, as inputs for sophisticated analysis and development of clinical guidelines in which to assess and treat patients. Technological advances, currently in various stages of convergence, adoption and use, can make this happen, as envisioned, in a cloud-based EHR database system.

Technological Advance and Convergence

Today, the convergence of technological advances in cloud services and emerging computing technologies, standardized data messaging, and advanced analytic and data storage capabilities provide the opportunity for genetic data integration into the clinical care setting, and a foundational reset in the delivery of clinical medicine and the development of evidence-based guidelines for the prevention and treatment of disease. Technological advancements for the proposed model touch on; 1) the utilization of next-generation sequencing (NGS) for identifying variants and increasing DNA knowledge, 2) the inclusion of genomic data into an EHR system through a standardized messaging application programming interface (API), in particular, SMART-on-FHIR, 3) advanced, scalable data storage methodologies, adaptable to increasing payloads and relational structures, 4) the statistical capabilities of advanced software for the investigation and evaluation of multiple data sources; genetic, FHH, and EHR data as a basis for a clinical support application, 5) cloud computing and storage systems to house, organize and operate the above, and finally, 6) the emergence of cloud microservices systems which facilitate application-based access to an EHR database system. In combination, these technologies will shape the proposed solution as follows:

1) Genetic data extraction and processing using next-generation sequencing techniques has significantly decreased the cost, time and effort of earlier methods, particularly, the Sanger method, considered the progenitor of genetic sequencing and its benchmark methodology.[6] These technologies, characterized by high throughput and massive parallel sequencing capabilities, have led to new gene discoveries and greater availability of variant information on a mass scale. Leveraging this molecular data through linkages with EHR and environmental data is building a bridge between genomic bioinformatics and clinical informatics. [25]

2) Advances in health care applications have demonstrated the potential to link and facilitate data flow from disparate systems and data sources. Specifically, SMART-on-FHIR, an HL7 open-source, standards-based application, presents a cutting-edge, interoperable platform designed to standardize data flow and to allow third-party applications to operate across different healthcare EHR systems. Launched in 2010, the SMART (Substitutable Medical Applications, Reusable Technologies) application offers the functionality through which interchangeable applications can access authorized pieces of an EHR. [26] Underneath it, FHIR (Fast Healthcare Interoperability Resources) provides the data format and standard, using widely adopted medical terminologies for coding data; LOINC, SNOMED, ICD-10, for organizing and storing the data by bits of information, or resources. The application is extensible. It operates through discrete data elements for research and retrieval and is able to accommodate new discoveries. SMART-on-FHIR's use of common internet protocols HTTP RESTful, and JSON or XML for data representation, also empowers the platform's interoperability. Of note, the FHIR genomic resources profile is currently in draft form. [27]

3) Today, advanced data platforms and storage technologies can enable secure access and storage for genetic data in massive quantities. These database tools remove many of the previous barriers for Big Data, or genetic information, through provisioning clusters of database servers with distribution and replication capabilities. Characteristic of this technology are horizontally scaled servers designed to manage information storage, updating, retrieval, and scaling as data needs change. (Of note, a current trend in data storage embraces the notion of "polyglot persistence," or the use of different data storage technologies, SQL or NoSQL, to manage varying data storage needs. This methodology releases APIs calls from having to conform to relational data models and enables use of schema-less data format such as Cassandra, or Mongo.) [28]

4) Today, cutting-edge statistical computing software such as R enables researchers to apply statistical tools and learning equations to complex and multi-variable datasets and generate statistical associations and predictive learnings. Clustering methods, or techniques for datasets without associated response variables, can reveal unknown subgroups and hierarchical relationships among large groups of genetic data.

And, through stepped up linear regression approaches, modeling validation methods and algorithm selection processes, [29] researchers can reveal data associations and potential predictive relationships that can point to new treatment insights.

5) The emergence of cloud-based platforms and services facilitates greater data linkage and analytic opportunities for using Big Data; genomic and EHR. Through advanced infrastructure and application software technologies, these off-site server and network platforms offer modern computing capabilities and data system scalability for large and increasing data inputs. Plus, they offer backup data services, and data and system security. As they evolve, cloud platforms are being significantly shaped by emerging technologies. [30]

6) The development of microservices technology has the potential to upend current cloud computing services architecture. In contrast to monolithic architectural structures where software is tied to an EHR database, a microservices environment facilitates an applications-based approach to EHR access and simplifies cloud architecture through a system of interoperable and interchangeable parts. Characteristic of this technology is deployment of individual, and isolated, containerized application units,[31] that share a common operating system, and are linked in complex patterns relating to business needs. Docker, for example, is an open-source microservices application that creates, deploys, and operates self-contained execution environments, or software containers, similar in manner to a virtual machine (VM). Essentially, a microservices architectural system moves software manufacture and installation approach from a product model to a business services model. [32]

Proposed Solution

The proposed solution design for this project integrates genetic data information relating to opioid addiction into a cloud-based application platform using as a foundation the IGNITE study MeTree application platform. The goal of the project is to further the understanding and development of opioid prevention and treatment strategies for clinicians and public health policy-makers, and to serve as a platform for broader use in investigating diseases from a more modern and multifactorial inclusive perspective. The

development and evolution of this system will serve as a model for the integration of genomic information into an EHR system, and for the data's use in the research and development of clinical guidelines.

As conceived for this project, the genetically integrated IGNITE MeTree platform model will address the challenges of implementing genomics into clinical practice and developing evidence-based intervention treatment strategies. These issues include:

- Point-of-care integration of genetic test results with EHRs and clinical decision support tools
- Standardized genotyping platforms for use in clinical practice
- Modernized storage of genetic data in analyzable formats
- Development of CDS capabilities
- Adaptation of clinical strategies as technology develops
- Ongoing assessment of outcomes data in relation to scientific disease understanding and technological change

The MeTree Platform

The foundation for this proposal is the IGNITE MeTree platform. Created through the Geomedical Connection, a collaboration between Duke University, University of North Carolina at Greensboro and the Cone Health System, this platform integrates family health history (FHH) information, EHR data, and a clinical decision support application, MeTree, into a primary care practice tool. [33] The MeTree risk assessment application collects FHH data directly through a patient portal and provides guidelines-based recommendations for individuals at risk of developing common chronic diseases. Algorithms are used to determine individual risk levels. At present, the application calculates risk scores and provides clinical decision support for breast cancer, ovarian cancer, colon cancer, thrombosis, coronary artery disease, aortic aneurysm, ischemic cerebrovascular disease, type diabetes, and hereditary liver diseases.[7] It is based on the HL7 Virtual Medical Record standard and can be integrated with medical records that support the SMART-on-FHIR method. It is also Epic-based. [7]

For purposes of this project, the MeTree platform offers these foundational elements;

- Alignment IGNITE Metree study platform is a National Human Genome Research Institute (NHGRI) formerly approved project. Given the mission of NHGRI, to improve health of all humans through advances in genomics research, the addition of genomic data, it appears, would fit comfortably into its framework. In this sense, the proposed model is aligned with the purpose of the NHGRI, and presumably its design would meet approval for genetic based study.

- Data Integration The Metree Network links EHR data, family health history and MeTree, a risk assessment application. Patient identifier linkages are established between data types and sources, and the storage framework enables data retrieval utilizing SMART-on-FHIR API tools. (Specifically, the MeTree FHH application stores family health history data separately from EHR data stores and technically can retrieve data on-demand from the EHR database). In this respect, genetic integration using the FHIR genomic profile could be achieved.

- Clinical Integration MeTree runs algorithms based on current clinical guidelines and provides risk-stratification prevention strategies integrated and delivered into the clinical workflow. The application calculates risk scores (Gail, Framingham, etc.) and "indicates when new technologies have clinical utility." [7] From a technical point of view, the MeTree application is integrated with the EHR database, via patient identifier, rather than acting as a standalone tool. [7] Guidelines, or metadata, based on opioid abuse related variants in relation to EHR and environmental factors could be integrated into the Metree CDS, as data analysis evolves.

- Interoperability MeTree supports SMART-on-FHIR, an HL7 standard that can provide integration across data channels and application access. Prototypes have been developed for a clinico-genomic resource application, an API that will be required for linking genomic variant sequencing information from a laboratory database to an EHR system. Feasibility of the SMART-on-FHIR Genomics API

has been demonstrated and adopted by the HL7 Clinical Genomics Workgroup [34] and by Duke University, sponsor of the MeTree network. Duke University has successfully developed a SMART-on-FHIR compatible server infrastructure at the Duke Health Network that integrates SMART-on-FHIR technologies and enables "plug-and-play" API linkages. [35]

- Research Utility: The IGNITE MeTree network contains embedded open-source applications such as R, SQL server, an Apache HTTP Server, and a Linux operating system. The value in using open-source is that it invites innovation versus being trapped in a legacy system, and any upgrades and changes extend from open contributions that are evaluated and input on a community level. As a result, the project is untethered from the restrictions and control of a single vendor. As genetic discoveries evolve, software flexibility and adaptability most likely will grow in importance. Security issues arising from open-source software, however, will need to be managed.

- Security Given previous project approval, it can be assumed that patient privacy and security provisions are in place and in line with HIPAA and other research study regulations. And, presumably, IRB approval was obtained for the initial platform use. As envisioned, this project would be an extension of its predecessor.

- Ease of Use The MeTree survey instrument is a patient-facing, web-based, interface for standardized collection of FHH data. Patients submit data on their own time thereby streamlining the process of gathering FHH data. Specifically, the MeTree application collects data on: diet, exercise, smoking and other clinical data in addition to family health history. It also supplies clinical decision support and patient recommendations on "20 cancers, 14 hereditary cancer and cardiovascular syndromes, and 21 other conditions." [7] The application, as conceived for this study, would also deliver risk stratification findings, in conjunction with FHH data, on opioid abuse vulnerability.

- Quality Clinical accuracy and validation in programming, output, coding, and algorithmic use has been performed by genetic counselor review. [36]

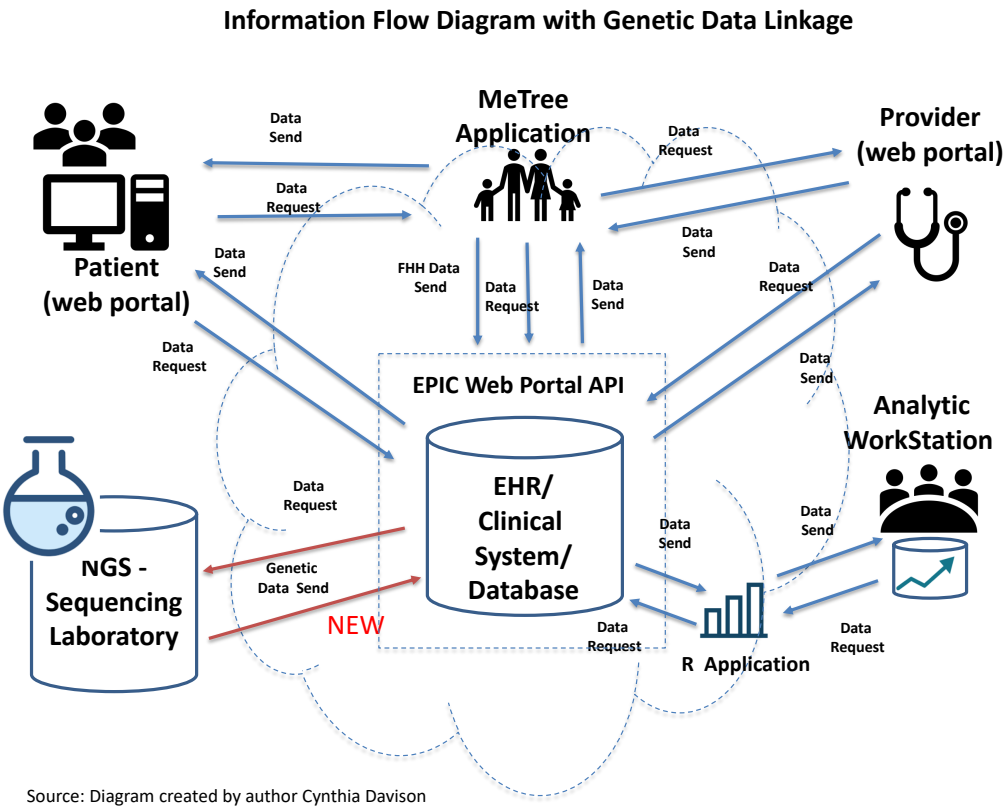
Technical Proposal

As envisioned, the proposed design will provide a cloud platform for the investigation of associations between behavioral, familial, genetic, and clinical factors in advancing disease understanding and intervention strategies. Specifically, the proposed platform will incorporate; API connectivity utilizing SMART-on-FHIR technology, advanced statistical analysis tools such as R, and expanded data and storage capabilities in a cloud-based platform that will provide system security, scalability, availability and reliability. Additionally, use of SMART-on-FHIR technology will enable construction of a microservices and systems enterprise architecture conducive to the use of emerging "plug-and-play" applications. The model will consist of these elements and the ability to:

- Integrate genetic data into an EHR system using utilizing a standardized ontology and messaging platform
- Link disparate data elements, including family health history, to expand the current understanding of prevention and treatment strategies.
- Provide the analytical tools i.e.; R software, to build decision support output for clinical use that reflects advanced technology and analytic techniques.
- Provide a platform for ongoing study of preventative and treatment strategies in clinical delivery for opioid addiction, and other multifactorial diseases and conditions.
- Provide a data platform which is scalable.
- Provide a platform which is in compliance with security requirements of HIPAA, and that adheres to IRB stipulations.
- Deliver a platform which aligns with the mission of the NHGRI.

Information Flow Diagram (IFD)

Figure 1. Information Flow Diagram with Genetic Data Linkage

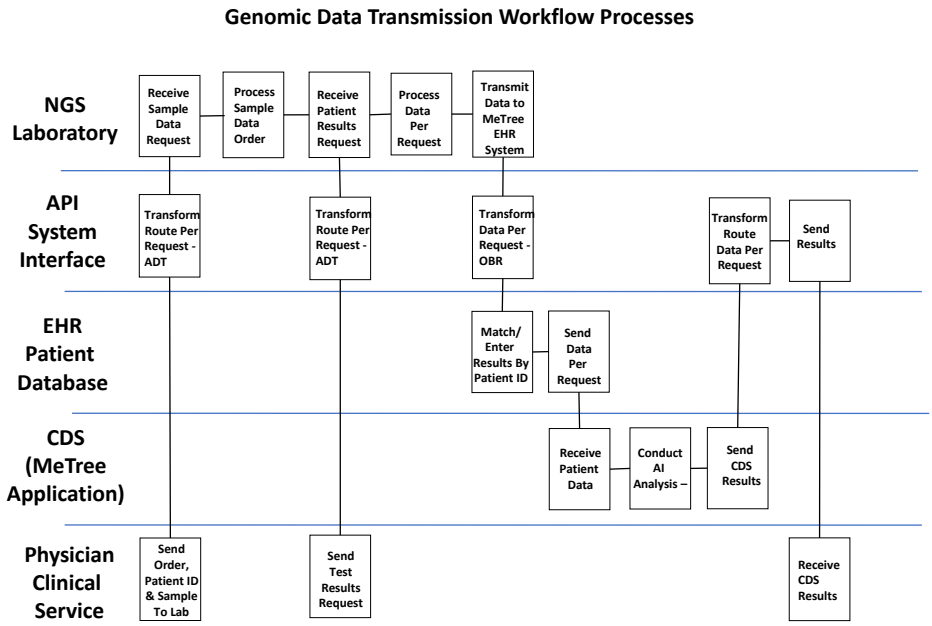


Source: Diagram created by author Cynthia Davison

An Information Flow Diagram presents a visual illustration of how data and information moves, or flows, through a system. As envisioned, the IGNITE MeTree platform model will incorporate a new data input for patient genetic findings, and thus require APIs for linking the sequenced laboratory data to the platform's electronic health record (EHR) database. The platform will also require additional and scalable data storage capacity, and capabilities for retrieval, statistical analysis, and reporting into the MeTree clinical decision support (CDS) application. Business processes that utilize genetic data will require on-demand data retrieval, and receipt of the data in a standardized format. The diagram above illustrates data flow between the EHR database, the genetic sequencing laboratory, application systems, and direct user stakeholders.

Task Flow Diagram

Figure 2. Task Flow Diagram for Genetic Data Transmission Workflow

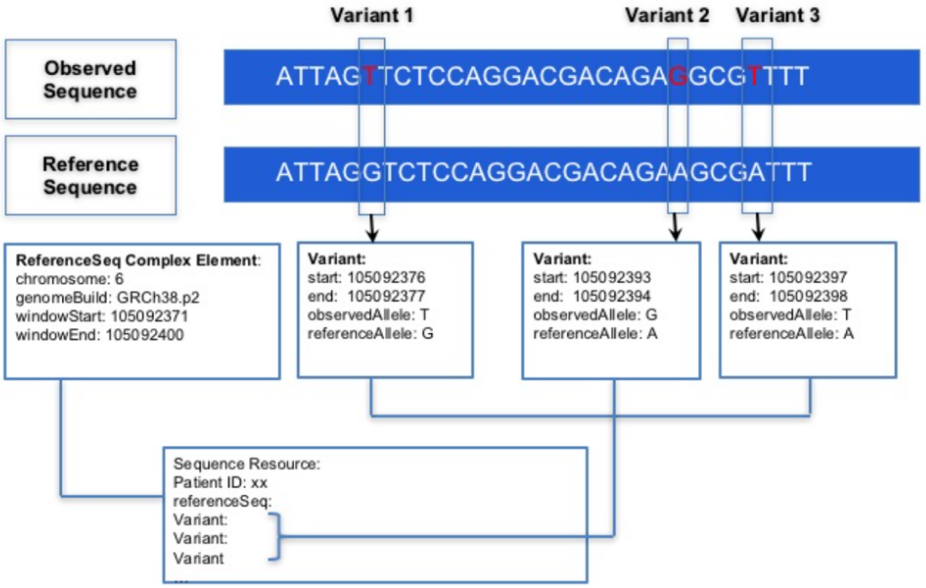


Source: Diagram created by author Cynthia Davison

The Task Flow Diagram illustrates specific steps in a process, and, as such, it is a representation of the various process tasks and their interrelationships. The linking and inclusion of genetic data to the envisioned MeTree model project system will require new processes for data communication and flow. Above is a swim lane diagram that parses the tasks for managing the transference of genetic data from the sequencing laboratory to the envisioned MeTree and EHR database network upon a physician request. Prior to task development, processes for user authorization and authentication will need to be established as well as those for data validation involving data transmission, and data storage. Checksum is one example used for this. The data also will require HL7 encryption standards in motion and at rest. And, use of the data system will also require privacy and security provisions in accordance with HIPAA and other regulations.

SMART-on-FHIR Genetic Data Transmission

Figure 3. HL7 - SMART on FHIR - Genetics Data Resource



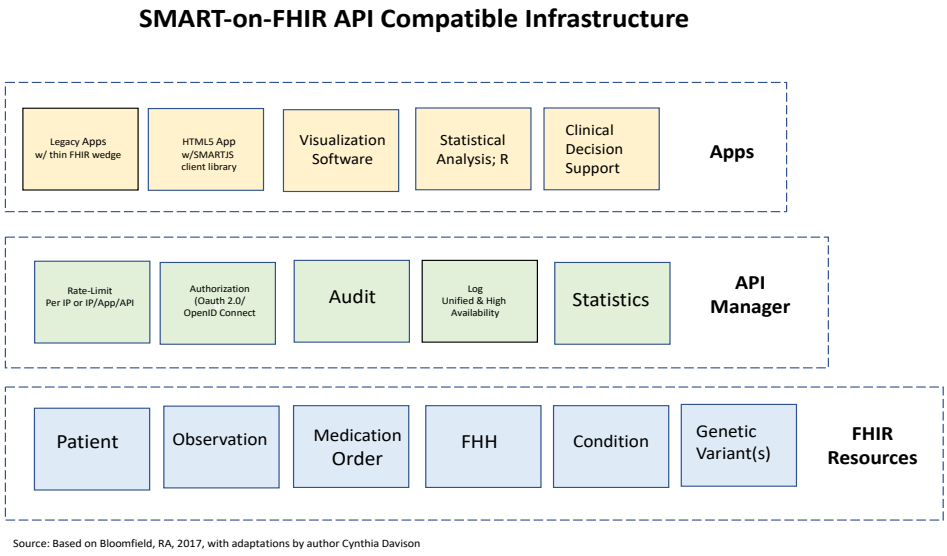
Source: Hl7.org, 10.8 Genomic Implementation Guidance, 2017
https://www.hl7.org/fhir/genomics.html

A major issue in integrating genomic findings into an EHR and clinical decision-making involves the lack of standard nomenclature for genetic variants and by extension, standardized messaging, and thus interoperability from the laboratory to the EHR database. Adoption of a standards-based ontology through which to integrate genetic data and clinical systems is a first step in data and clinic-based integration. Given that, the SMART-on-FHIR genomics profile, an open-source, observation-based and extensible application designed for next-generation sequencing, conceptually, can be employed. The depiction above illustrates the genetic resource for next-generation sequencing and recording of sequence id/string and detected variants, as presented by hl7.org. [27] Notably, the modular construction inherent in FHIR (Fast Healthcare Interoperability Resources), as seen in the depiction, enables rich payload able to handle expressive queries and development of substitute applications. Utilizing this transmission system, the proposed solution is capable of linking with genomic tests and any future

pharmacogenic testing. Of note, FHIR currently faces some adoption challenges stemming from content disagreement, and issues with multiple versions. [27]

SMART-on-FHIR API Infrastructure

Figure 4 SMART-on-FHIR API System Infrastructure

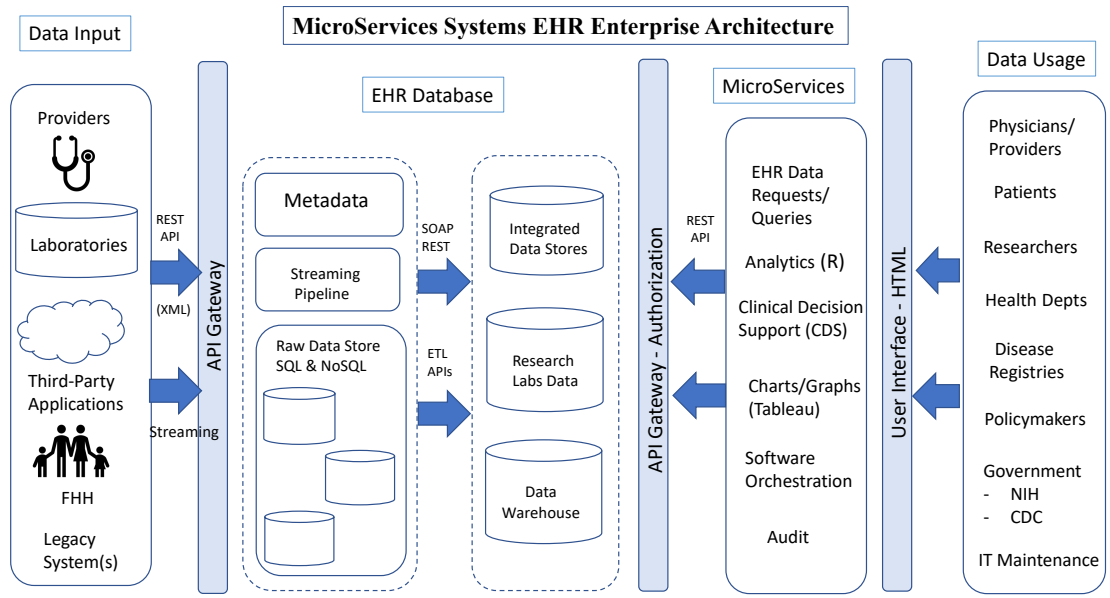


SMART-on-FHIR technology is a standardized framework for the development of interchangeable healthcare applications in the healthcare arena. [8] In 2014, Duke University, sponsor of the IGNITE Metree project, implemented SMART-on-FHIR into the University's EHR Epic-based system. [35] To date, implementation, designed to enable seamless data access using applications, is considered a success. The diagram above illustrates how the SMART-On-FHIR resources infrastructure, configured for the Duke project, could be utilized for the proposed model. The infrastructure enables; 1) use of interchangeable APIs, such as the envisioned Metree application, (top row), 2) an API layer for authorization, auditing, and analytics (middle), and, 3) data coded in FHIR-compatible format, retrievable by resource (bottom). [35] This configuration serves as an illustration of the workability of a standards-based API system, and, in effect, a "plug-and-play" solution, in which SMART-on-FHIR compatible applications are key enablers for database access. Implementation of the SMART-on-FHIR API would involve

mapping laboratory and patient data to FHIR resources, and implementing the API authorization scheme supported by SMART - OAuth 2.0 and OpenID Connect. Currently, medical services providers, Cerner, Epic, and Allscripts, all provide sandboxes for SMART-on-FHIR development. [37]

Enterprise Architecture

Figure 5. Enterprise Architecture – MicroServices Environment



Source: Diagram created by Cynthia Davison based on model in; ABN AMRO's Data Integration Architecture, ABN AMRO Developer Blog, Mar 13
https://medium.com/abn-amro-developer/abn-amros-data-integration-architecture-f33506a211c3

Emerging microservices have the potential to radically alter the focus and build of cloud-based systems. In contrast to monolithic architectures, whereby applications are specifically tied and coded to a database system, a microservices architecture is an evolutionary model that organizes system software around business capabilities, [38] and structures applications as units of "loosely coupled" and collaborating services. [38] Characteristic of this architecture is deployment of containerized, interchangeable application units which share a common OS, and specific APIs that define and modulate business processes according to set patterns. An API gateway serves as an authorization entry point. Database protocols and ontologies are standardized for API connectivity. Notably, Docker, is a premier microservices application for container deployment in this arena. The diagram above illustrates how SMART-on-FHIR technology can enable a

"plug-and-play" microservices environment through standardizing database protocols, and providing a platform technology for application processing apart from the EHR database. In this respect, SMART-on-FHIR becomes a critical enabler toward adoption of a microservices architecture strategy.

Conclusion

This proposal focuses on the development of a clinical research model to fully integrate genetic, familial health, behavioral and medical data, and provide a platform for the development of evidence-based interventions for clinical use. The proposed model utilizes SMART-on-FHIR technology in conjunction with a microservices architecture to create a platform that enables integration of genetic data into an EHR system, and facilitates data accessibility for FHIR compatible applications. As such, it provisions the clinical decision support application, MeTree, when deployed as a containerized application, to pull needed data through an authorized and SMART-on-FHIR access capability. The hope is that as it evolves, the ability of the proposed model to connect disparate data sources and empower ongoing predictive analysis can lead to better intervention models and guidelines for opioid abuse disorder, and for many other diseases.

Looking ahead, however, implementation of the model will raise multiple technical issues. As conceived, the model addresses barriers to genomic data integration into an EHR database, yet it does not identify possible implementation complications, or issues surrounding the deployment of microservices and business process patterns. Additionally, the SMART-on-FHIR genetics resource, as discussed, is still in draft stage though structural integration has been successfully achieved as demonstrated by the Duke University API infrastructure project.

Further, the harvest and use of genetic information will present important ethical and moral considerations - and dilemmas. These will revolve around privacy issues and fears of retaliation from employers, insurance companies, or others, due to exposure of genomic information. Issues will also touch on the privacy boundaries of physician-patient relationships and the information genetic data can yield. For instance, should

providers be obligated to inform patient relatives of a positive predictive test? [39] And, by extension, what is the responsibility of the patient to inform a relative of an inheritable condition? Currently, genetic data and patient privacy is protected within the confines of the medical profession and other HIPAA specified entities. In the commercial sphere, comparatively, it is relatively unregulated – as of yet.

In conclusion, it is clear that advances in genetic sequencing and computing services have the potential to revolutionize medical care and understanding. For public health, leveraging these technologies will lead to greater knowledge of disease and intervention strategies. The solution proposed here offers a cloud-based model that, as envisioned, can apply to research and development of evidence-based treatment for many different types of illnesses stemming from genetic, environmental and behavioral factors, all, or in part. In this respect, the future is boundless.

Sources

1. Randesi M, van den Brink W, Levran O, et al. Dopamine gene variants in opioid addiction: comparison of dependent patients, nondependent users and healthy controls. *Pharmacogenomics.* 2018;19(2):95-104.
2. Cheng Z, Zhou H, Sherva R, Farrer LA, Kranzler HR, Gelernter J. Genome-wide Association Study Identifies a Regulatory Variant of RGMA Associated With Opioid Dependence in European Americans. *Biological psychiatry.* 2018.
3. Donaldson K, Demers L, Taylor K, Lopez J, Chang S. Multi-variant Genetic Panel for Genetic Risk of Opioid Addiction. *Annals of clinical and laboratory science.* 2017;47(4):452-456.
4. Berrettini W. A brief review of the genetics and pharmacogenetics of opioid use disorders. *Dialogues in clinical neuroscience.* 2017;19(3):229-236.
5. Mistry CJ, Bawor M, Desai D, Marsh DC, Samaan Z. Genetics of Opioid Dependence: A Review of the Genetic Contribution to Opioid Dependence. *Current psychiatry reviews.* 2014;10(2):156-167.
6. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell.* 2013;155(1):27-38.
7. Duke Center for Applied Genomics & Precision Medicine. MeTree Software. https://precisionmedicine.duke.edu/researchers/precision-medicine-programs/risk-assessment/family-history/metree-software. Accessed May 26, 2018.
8. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc.* 2016;23(5):899-908.

9. Centers for Disease Control and Prevention. Understanding the Epidemic. https://www.cdc.gov/drugoverdose/epidemic/index.html. Accessed May 26, 2018.
10. NIDA. Opioid Overdose Crisis. 2018; https://www.drugabuse.gov/drugs-abuse/opioids/opioid-overdose-crisis. Accessed July 7, 2018.
11. Council of Economic Advisers. Council of Economic Advisers Report: the Underestimated Cost of the Opioid Crisis,. https://www.whitehouse.gov/briefings-statements/cea-report-underestimated-cost-opioid-crisis/. Accessed May 26, 2018.
12. SAMSHA. Preventing Prescription Drug Misuse: Overview of Factors and Strategies. 2016; https://www.samhsa.gov/capt/sites/default/files/resources/preventing-prescription-drug-misuse-overview.pdf. Accessed July 7, 2018.
13. Dowell D. HTM, Chou R.,. Guideline for Prescribing Opioids for Chronic Pain — United States, 2016. *MMWR* Vol 65: Centers for Disease Control and Prevention:1–49.
14. Cheattle M. Opioid Prescribing and Monitoring; How to Combat Opioid Abuse and Misuse Responsibly. *Practical Pain Management* 2017; https://www.practicalpainmanagement.com/resource-centers/opioid-prescribing-monitoring/risk-assessment-safe-opioid-prescribing-tools. Accessed July 4, 2018.
15. Meshkin B, Lewis K, Kantorovich S, Anand N, Davila L. Adding Genetic Testing to Evidence-Based Guidelines to Determine the Safest and Most Effective Chronic Pain Treatment for Injured Workers. *International Journal of Biomedical Science : IJBS.* 2015;11(4):157-165.
16. Diatchenko L, Fillingim RB, Smith SB, Maixner W. The phenotypic and genetic signatures of common musculoskeletal pain conditions. *Nature reviews Rheumatology.* 2013;9(6):340-350.
17. National Institute on Drug Abuse. Opioid Overdose Crisis. https://www.drugabuse.gov/drugs-abuse/opioids/opioid-overdose-crisis. Accessed May 26, 2018.
18. Jones C, Center for Disease Control and Prevention,. The Opioid Overdose Epidemic; Epidemiology and Trends. 2013.
19. Substance Abuse and Mental Health Services Administration (SAMHSA). Preventing Prescription Drug Misuse: Understanding Who Is at Risk, May 2016, CAPT Decision-Support Tools.
20. Lander L, Howsare J, Byrne M. The Impact of Substance Use Disorders on Families and Children: From Theory to Practice. *Social work in public health.* 2013;28(0):194-205.
21. Chambers RA, Taylor JR, Potenza MN. Developmental neurocircuitry of motivation in adolescence: a critical period of addiction vulnerability. *Am J Psychiatry.* 2003;160(6):1041-1052.
22. Tsuang MT, Bar JL, Harley RM, Lyons MJ. The Harvard Twin Study of Substance Abuse: what we have learned. *Harvard review of psychiatry.* 2001;9(6):267-279.
23. Kendler KS, Karkowski LM, Neale MC, Prescott CA. Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Arch Gen Psychiatry.* 2000;57(3):261-269.

24. Volkow ND, Wang GJ, Fowler JS, et al. Brain DA D2 receptors predict reinforcing effects of stimulants in humans: replication study. *Synapse.* 2002;46(2):79-82.
25. Roy S, LaFramboise WA, Nikiforov YE, et al. Next-Generation Sequencing Informatics: Challenges and Strategies for Implementation in a Clinical Environment. *Arch Pathol Lab Med.* 2016;140(9):958-975.
26. G. M. What is SMART on FHIR? 2017; https://www.redoxengine.com/blog/what-is-smart-on-fhir/. Accessed July 7, 2018.
27. hl7.org. 10.8 Genomic Implementation Guidance. 2017.
28. Sadalage PJ, Fowler M. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence.* Upper Saddle River, NJ: Addison-Wesley; 2013.
29. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning with Applications in R.* New York: Springer; 2013.
30. Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Gener Comp Sy.* 2009;25(6):599-616.
31. Fernandez E, Yoshioka N, Washizaki H, Syed M. Modeling and Security in Cloud Ecosystems. *Future Internet.* 2016;8(2):13.
32. Erich F, Amrit C, Daneva M. *Report: DevOps Literature Review.* 2014.
33. Orlando LA, Wu RR, Beadles C, et al. Implementing family health history risk stratification in primary care: impact of guideline criteria on populations and resource demand. *American journal of medical genetics Part C, Seminars in medical genetics.* 2014;166c(1):24-33.
34. Alterovitz G, Warner J, Zhang P, et al. SMART on FHIR Genomics: facilitating standardized clinico-genomic apps. *J Am Med Inform Assoc.* 2015;22(6):1173-1178.
35. Bloomfield RA, Jr., Polo-Wood F, Mandel JC, Mandl KD. Opening the Duke electronic health record to apps: Implementing SMART on FHIR. *Int J Med Inform.* 2017;99:1-10.
36. Orlando LA, Buchanan AH, Hahn SE, et al. Development and validation of a primary care-based family health history and decision support program (MeTree). *North Carolina medical journal.* 2013;74(4):287-296.
37. smarthealthit.org. ITdotHealth; SMART Decisions. 2017; https://smarthealthit.org/wp-content/uploads/SMART-ITdotHealth2017-Executive-Summaries.pdf. Accessed July 4, 2018.
38. microservices.io. What are microservices? http://microservices.io. Accessed July 7, 2017.
39. Fulda KG, Lykens K. Ethical issues in predictive genetic testing: a public health perspective. *Journal of Medical Ethics.* 2006;32(3):143-147.