



# Evaluations of California Initiative to Advance Precision Medicine Projects, 2015-2018

TO THE CALIFORNIA LEGISLATURE  
JULY 2019

Evaluations of California Initiative to Advance Precision Medicine Projects, 2015-2018  
Report for the California Legislature  
July 2019

## Prepared by the Governor's Office of Planning and Research

Kate Gordon, JD, MCP  
Director

Elizabeth Baca, MD, MPA  
Deputy Director for Research and Education

Julianne McCall, PhD, MS  
Science Officer

Shannon Muir, PhD, MS  
Science Officer

Megan Varvais  
Science Communication Specialist and Administrator

Nichole Holm  
Policy Fellow

### Special acknowledgement for their assistance:

David R. Paquette, PhD  
Former Science Officer

Governor's Office of Planning and Research  
1400 Tenth Street  
Sacramento, California 95814  
Phone: (916) 322-2318

**Contents**

**EXECUTIVE SUMMARY .....1**

**BACKGROUND ..... 2**

**PROCESS OF EVALUATION ..... 4**

**EVALUATIONS OF DEMONSTRATION PROJECTS ..... 6**

1. California Kids Cancer Comparison ..... 6

    Scientific Background and Context..... 7

    Project Summary..... 8

    Project Achievements..... 9

    Expert Evaluation by Dr. Nikhil Wagle, MD .....12

    Assessment of How the Project Addressed Programmatic Goals .....13

    Conclusion ..... 14

2. Precision Diagnosis of Acute Infectious Diseases..... 16

    Scientific Background and Context..... 17

    Project Summary..... 19

    Project Achievements..... 19

    Expert Evaluation by Dr. Nancy Cox, PhD..... 22

    Assessment of How the Project Addressed Programmatic Goals ..... 24

    Conclusion ..... 27

3. Artificial Intelligence for Imaging of Brain Emergencies..... 28

    Scientific Background and Context..... 28

    Project Summary..... 29

    Project Achievements..... 29

    Expert Evaluation by Dr. Jack Tsao, MD, DPhil..... 31

    Assessment of How the Project Addressed Programmatic Goals ..... 31

    Conclusion ..... 33

4. Early Prediction of Major Adverse Cardiovascular Events Using Remote Monitoring..... 35

    Scientific Background and Context..... 35

    Project Summary..... 37

    Project Achievements..... 37

    Expert Evaluation by Dr. John Rumsfeld, MD, PhD ..... 40

    Assessment of How the Project Addressed Programmatic Goals ..... 41

    Conclusion ..... 44

5. Full Genome Analysis to Guide Precision Medicine ..... 45

Scientific Background and Context.....	45
Project Summary.....	46
Project Achievements.....	47
Expert Evaluation by Dr. Elaine Mardis, PhD .....	48
Assessment of How the Project Addressed Programmatic Goals .....	50
Conclusion .....	55
<b>6. Personal Mobile and Contextual Precision Health .....</b>	<b>56</b>
Scientific Background and Context.....	56
Project Summary.....	57
Project Achievements.....	57
Expert Evaluation by Dr. Bambang Parmanto, PhD.....	58
Assessment of How the Project Addressed Programmatic Goals .....	59
Conclusion .....	60
<b>7. Precision Medicine for Early Prostate Cancer: Integrating Biological and Patient Complexity Variables to Predict Treatment Response.....</b>	<b>61</b>
Scientific Background and Context.....	61
Project Summary.....	63
Project Achievements.....	63
Expert Evaluation by Dr. Mark Rubin, MD .....	67
Assessment of How the Project Addressed Programmatic Goals .....	68
Conclusion .....	71
<b>8. Precision Medicine for Multiple Sclerosis: Making It Work.....</b>	<b>72</b>
Scientific Background and Context.....	73
Project Summary.....	73
Project Achievements.....	74
Expert Evaluation by Dr. Harold Lehmann, MD, PhD.....	74
Assessment of How the Project Addressed Programmatic Goals .....	76
Conclusion .....	82
<b>CONCLUSION.....</b>	<b>84</b>
<b>APPENDIX A: EXPERT EVALUATORS.....</b>	<b>85</b>
<b>APPENDIX B: PUBLICATIONS, PRESS RELEASES, AND PRESENTATIONS GENERATED AS A RESULT OF CIAPM-FUNDED RESEARCH .....</b>	<b>93</b>

## Executive Summary

The field of precision medicine has two overarching goals: gaining a better understanding of disease mechanisms, and fostering a more modern and equitable approach to health care using technology and data.

Since then-Governor Brown and the California legislature launched the California Initiative to Advance Precision Medicine (CIAPM) in 2015, the program has funded, and seen through to completion, eight collaborative translational research projects across the state. The projects were competitively selected among public and private academic institutions with over 75 partners in industry, health care, academic, and nonprofit sectors. In total, the Governor's Office of Planning and Research awarded over \$11 million to support innovative demonstrations of how cutting-edge technologies and the use of data may improve disease prevention and medical interventions for Californians.

Many projects targeted specific disease areas, while others sought to improve tools that could be applied to a wide range of disorders and integrated into preventative, diagnostic, or therapeutic platforms. The diseases that were specifically addressed include pediatric cancer, infectious diseases, brain damage, heart disease, rare childhood genetic disorders, high blood pressure, depression, prostate cancer, and multiple sclerosis. The technologies that were created, developed, and refined include artificial intelligence, remote biosensors, automated genomic analysis, networked databases to foster data sharing, computation, and collaboration, a clinical portal to engage patients, and mobile health apps, among others.

In total, the research teams attracted over \$12 million in external matching funds and were sustained by in-kind institutional support valued over \$6.4 million. Twenty-five publications detailing study results have been released in scientific journals or are currently in preparation. Seventeen articles and press releases have highlighted the projects, including a *TIME Magazine* cover story, and over 70 presentations have been delivered to public and professional audiences. In ways specific and general, the work of the eight CIAPM projects has already helped shape the future of biomedical research and clinical practice.

To evaluate the quality and significance of the projects' scientific achievements and clinical contributions, the Office of Planning and Research facilitated assessments by expert scientists and clinicians recruited from outside the state. Evaluators were granted access to the research teams' materials and specific aims of the program to determine how well the projects fulfilled their proposed milestones and furthered the respective scientific fields as well as the state's ambitious goals. This report serves as a summary of their assessments, supported by contextual background on the scientific fields and project details.

# Background

## *The Emerging Field of Precision Medicine*

Concurrent advances in high-performance computing and biotechnology have enabled the collection of massive amounts of health-related data that is positioned to transform the prevention of disease, the delivery of health care, and the quality of targeted interventions beyond the one-size-fits-all approach of the 20<sup>th</sup> Century. Through the lens of modern research, the future of health care will see strategies shift from acute intervention and the management of illness to health assessments and proactive oversight and control of disease risks and prevention.

The U.S. President's Council of Advisors on Science and Technology in 2008 defined precision medicine as: "The tailoring of medical treatment to the individual characteristics of each patient... to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not". As the definition implies, the potential impact of precision medicine lies in its ability to navigate and direct health care decisions toward the most effective treatment for a given patient. Thus, the quality of care improves while diminishing unnecessary diagnostic testing and therapies.

While enhancing the efficiency and effectiveness of prevention and treatment strategies are overarching goals of this next step in health practice, the field also seeks to better understand the underlying mechanisms of disease onset and progression, treatment response, and health outcomes by incorporating and accounting for measurements of genetic, molecular, environmental, and lifestyle factors that contribute to a person's health and wellbeing. The rapid rise in DNA sequencing capacity, widespread adoption of electronic health records, and availability of mobile health devices have driven clinical implementation. In parallel to furthering the science, the field of precision medicine is also facilitating a culture change toward engaging individuals as active partners in health care and medical research.

## *Brief History of the California Initiative to Advance Precision Medicine*

California is home to many research institutions, health-oriented community organizations, and industries that are engaged at the cutting edge of precision medicine applications. The California Initiative to Advance Precision Medicine (CIAPM) was established by the State of California in 2015 and continues to be the only program of its kind nationwide. It was developed to help coordinate public, private, and non-profit partners to expand precision medicine approaches to underrepresented communities and foster the creation of new technologies and therapies to improve the health of diverse populations. The initiative brings together state precision medicine leaders and supports projects aimed at demonstrating the power and application of precision medicine to the people of California, with the ultimate goal of improving health outcomes, reducing health disparities, and elevating public health.

The California Legislature first allocated \$3 million in state funding to the Governor's Office of Planning and Research (OPR) to launch the program in 2015, which has since grown to

\$53 million in total from budget allocations through 2018. With numerous industry, academic, and nonprofit stakeholders already positioned and new ventures continually surfacing in the precision medicine arena, CIAPM actively fosters connections and unique partnerships between the state, universities, and nonprofit and private entities. The resulting knowledge network expands the state's capacity to address issues around health equity, inform investments in health care delivery, and support economic activity in the field of precision medicine throughout California.

The primary recipients of CIAPM grants are collaborative research teams selected through a peer-reviewed competitive process consistent with standards similar to the National Institutes of Health. For each Request for Proposals (RFP) cycle, out-of-state experts are recruited to serve on selection committees and assigned based their ability to assess scientific validity and potential contribution to the field of precision medicine. Selection committee members cover a range of perspectives, including community, patient engagement expertise, health disparity knowledge, health care, research, and industry.

Funding awards differ from those at the federal level and other sources in that OPR encourages and most recently requires research teams, who are based at nonprofit public and private scientific and clinical institutions, to partner with community organizations from the beginning of the project onward. In so doing, research results are oriented toward enhancing access to health care and developing effective treatments and preventative measures for a diverse population, reflective of California's diverse population.

The projects aimed to address one or more of the following focus areas, as set forth in Sections 65057 and 65058 of the California Government Code:

- The application of precision medicine to specific disease areas
- The challenges of system interoperability
- Economic analysis
- Standards for sharing data or protocols across institutions
- The federal and state regulatory environment
- The clinical environment
- Challenges relating to data, tools, and infrastructure
- The protection of privacy and personal health information
- The potential for reducing health disparities
- Methods and protocols for patient engagement

As of the writing of this report, eleven projects have been awarded across nine institutions, with academic, community, and industry partners numbering over eighty-five. Of those, eight projects concluded in December 2018 and subsequently have been evaluated by independent experts, as described in detail in this report.

## Process of Evaluation

As required by state law, expert evaluations followed the completion of the first rounds of funding. For this purpose, peers in the scientific community were recruited to assess each project's scientific, clinical, and technical accomplishments objectively within the context of the specialized line of research and the greater field of precision medicine. In addition, evaluations were meant to gauge how the work has addressed the goals of the program to advance precision medicine and benefit diverse communities across California.

The specific goals of the program, as it exists in the Governor's Office of Planning and Research (OPR), are listed in statute and actuated through competitive Requests for Proposals. Each awarded project was not expected, nor designed, to address each point, and only goals relevant to a project were considered by the evaluators.

1. Demonstration of the promise of precision medicine in a specific disease area, health issue, technology, or fundamental biological process
2. Use of existing patient data and other data sources
3. Efficient, effective data integration and analysis
4. Development of precision medicine capabilities
5. Development and implementation of participant engagement strategies
6. Impact for patients within the project timeframe and beyond
7. Reduction of health disparities within the project timeframe and beyond
8. Creation of an economic impact/value analysis
9. Development/realization of the project's clinical and commercial potential
10. Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications
11. Potential to scale and leverage multiple electronic health records systems and other platforms
12. Solutions to and insights about challenges, including interoperability, health disparities, privacy, participant engagement, consent, security, and ethical concerns
13. The development of partnerships/collaborations
14. Other contributions related to advancing precision medicine such as infrastructure development

### *Out-of-State Expert Evaluators*

Eight out-of-state expert evaluators were selected and recruited based on expertise. Each evaluator was subject to a conflict-of-interest screening, adapted from the standard process of the National Institutes of Health (NIH).<sup>1</sup> Following agreement to nondisclosure and confidentiality, each evaluator was assigned one primary project and up to two secondary projects. All evaluators are listed in Appendix A.

---

<sup>1</sup> NIH Conflict of Interest rules for initial peer review for grant applications, and technical evaluation of R&D contract proposals, based on federal regulations (42 CFR Part 52h) and presented in detail in NIH Guide Notices NOT-OD-13-010 and NOT-OD-14-069.



### *Evaluation Materials*

The following materials were provided to evaluators at the onset of the evaluation process:

- Request for Proposals
- Original proposals & milestones
- Final reports, prepared by project teams using a template
- Instructions for evaluators

### *Evaluation Process*

Primary evaluators drafted evaluative comments and reviewed the materials provided for secondary projects. OPR then convened evaluators by videoconference in designated groups of 1-2 experts for the purpose of gathering further details and addressing specific aspects of the evaluative comments. For those evaluators who indicated the need for additional information from research teams for a comprehensive evaluation, OPR facilitated the exchange of evaluators' follow-up questions and research teams' responses.

### *Evaluation Report*

This report is based on the research teams' final reports, original proposals, expert evaluations, evaluation videoconference discussions, and research teams' responses to evaluators' follow-up questions. Feedback on drafts of the report was solicited from the expert evaluators, research teams, and the Governor's Office.

Appendix A features the expert evaluators, and Appendix B lists all publications, press releases, and presentations generated as a result of CIAPM-funded research.

# Evaluations of Demonstration Projects

## 1. California Kids Cancer Comparison

Principal Investigator: Dr. David Haussler, UC Santa Cruz  
Project Period: September 1, 2015 – December 31, 2018

### *Research Team and Collaborators*

#### UC Santa Cruz

David Haussler, PhD  
Olena Morozova Vaske, PhD  
Isabel Bjork, JD, MSc, MA  
Rob Currie, MBA  
Holly Beale, PhD  
Ann Durbin  
Katrina Learned  
Ellen Kephart  
Jacob Pfiel  
Lauren Sanders  
Katrina Slater

#### Stanford University, Lucile Packard Children's Hospital, Stanford

Sheri L. Spunt, MD, MBA  
Norman J. Lacayo, MD  
Kara L. Davis, DO  
Alejandro Sweet-Cordero

#### University of British Columbia, BC Cancer Agency

Marco Marra

#### Children's Hospital Orange County

Leonard Senders  
Ashley Plant

#### Pacific Pediatric Neuro-Oncology Consortium (UC San Francisco)

Sabine Mueller

#### UC San Francisco

Alejandro Sweet-Cordero

#### Children's Mercy Hospital in Kansas City

#### Sanford University of South Dakota Medical Center

#### University of Michigan

#### University of Pittsburg

#### Alex's Lemonade Stand Foundation

#### Amazon Services

#### Azure

Jacob's Heart  
Key for a Cure  
Kids v Cancer  
Live For Others Foundation  
St. Baldrick's Foundation  
Team Finn  
Team G Foundation  
Unravel Pediatric Cancer  
DNAexus  
Microsoft  
NuMedii  
Seven Bridges  
George Kraw

## Scientific Background and Context

### *Pediatric cancer*

Compared to 50 years ago, standard treatment for childhood cancer has remained largely unchanged. In California, 500 of the 1,700 children annually diagnosed with cancer either do not respond to standard treatments or have no standard therapies available to treat their condition. Hard-to-treat cancers are categorized as “relapsed” or “refractory,” meaning they have resisted available treatments. For most of these patients, hospice is the only remaining option.

Whereas adult cancers are often linked to genetic mutations that are acquired during an individual's lifetime, such as through environmental exposures, pediatric cancers present with fewer mutations and are therefore more difficult to diagnose. Support for and advancement of adult cancer therapies continue to progress compared to pediatric cancers, due in large part to ample federal funding, the underlying biology of adult cancers, more established data sharing platforms, and the drug development pipeline. Childhood cancers are significantly less common, which ultimately limits clinical trials from recruiting sufficient numbers of patients and makes drug development a much slower process than for adult cancers. Pediatric cancers also tend to be more complex, which prevents a single therapy from benefitting large groups of patients. Of federal funding allocated to the National Cancer Institute, approximately 4% is annually designated for pediatric cancers.<sup>2</sup>

---

<sup>2</sup> <https://nationalpcf.org/facts-about-childhood-cancer/>

### *DNA-based therapies*

Cancer cells are characterized by uncontrolled cell growth that causes tumors to form and grow. This type of unchecked proliferation occurs because the cancer cells' genome (complete blueprint of genetic material, or DNA) has acquired specific mutations (changes in basic DNA units). Until 2017, cancer drugs were developed to target specific mutations based on where a tumor was located within the body (e.g., lung cancer). These targeted drugs work by interfering with the ability of cancer cells to grow or survive. They are effective as long as the patient's tumor has the most common mutation the drug is designed to target. Occasionally, a mutation known to occur in one tumor type is detected in a tumor in a different location of the body. For instance, a mutation typically found in lung cancer may be found in a brain tumor. Since the drug is targeted to prevent or reverse the negative effects of that mutation, it may help treat the brain cancer, even though it was originally developed for lung cancer. With access to advanced DNA sequencing technologies, the details of each person's cancer can help direct the treatment strategy.

### *RNA-based therapies*

Just as in healthy cells, each cancer cell contains the full genome but only uses a fraction of the total available genes to perform its specialized functions. Which genes are active or dormant distinguishes one cell type from another, like a muscle cell compared to a kidney cell. Scientists can uncover which genes are active by identifying the collection of RNA molecules present in a cell, which are single-helix copies of double-helix DNA. Consider the analogy of DNA as the full set of blueprints of a building site and RNA as the daily orders for a construction crew. The set of RNA molecules present in a cell is known as its 'transcriptome,' and the scientific field is called 'transcriptomics.' By comparing RNA molecules that are detected in a patient's tumor to those found in thousands of other tumors (comparative RNA analysis), scientists can ascertain how the molecular pattern of a patient's tumor resembles or differs from other tumors. This information can help guide the best treatment strategy based on the underlying mechanisms of tumor growth.

## **Project Summary**

For hard-to-treat cancers, recent clinical trials based on tumor genomics have had limited success, especially for children. On average, DNA analysis of pediatric cancers yields useful information for fewer than 10% of patients for whom standard treatment has been unsuccessful. The California Kids Cancer Comparison (CKCC) project sought to ameliorate cancer care for pediatric patients by leveraging two fundamental concepts: 1) instead of relying exclusively on genomic mutations in the tumor (DNA-based analysis), the research team employed an RNA-based approach; and 2) the team instituted large-scale computation to compare all RNA in a specific tumor with over 11,000 tumors from other patients ("Cancer Comparison"). Using these techniques, the team aimed to determine what is likely driving the uncontrolled growth of a patient's specific tumor and therefore identify new potential targets for therapy.

As a step toward incorporating RNA analysis in the clinic, the research team collaborated with ongoing clinical trials for children with cancer. While the trials looked for new treatment options based on tumor DNA, CKCC obtained each tumor's RNA data and analyzed it in several innovative ways. Using this data-driven comparative approach, the team identified new molecular information about the case, previously unavailable to the clinical team, in

100% of cases, exceeding their original goal of 20% for this initial study. Some of this information could be used for alternative treatment possibilities. The team then communicated its findings to the clinical trials and received feedback to develop effective communication strategies with clinicians.

To further test clinical efficacy of comparative cancer RNA analysis, the research team established a registry focused on clinical validation of the findings and optimization of patient/family engagement in medical decision-making. Supported by external funds, this work is ongoing and will be completed by the end of 2019.

## Project Achievements

Over the course of the funding support, the team 1) established a consortium of hospitals and research institutions, 2) collected de-identified patient RNA data from clinical trials, 3) suggested new treatment options based on data analyses, 4) reported findings to molecular tumor boards, 5) optimized computational approaches and created a large tumor data catalog, 6) engaged patients in the research process and clinical decision-making, 7) achieved data sharing goals, and 8) conducted an early-phase patient registry study to validate clinical utility. The team was also able to 9) attract substantial additional funding. More detail on each of these points is provided next.

### *Research consortium*

Building a collaborative network of researchers, clinicians, and hospital administrators was integral to the success of this project. By providing multiple channels of secure data sharing and consistent access to new and former patients, the research team was able to design their work to best support clinical practice. Core members of the consortium included UC Santa Cruz, UC San Francisco, Pacific Pediatric Neuro-Oncology Consortium, Children's Hospital of Orange County, British Columbia Cancer Agency, Stanford University School of Medicine, and Lucile Packard Children's Hospital. The research team also partnered with University of Michigan Comprehensive Cancer Center, Sanford University of South Dakota Medical Center, and Children's Mercy Hospital, Kansas City.

Advocacy groups and industry partners were also often included in meetings and research discussions. According to the researchers, "Partnerships with advocacy groups and industry have made our translational role possible."

### *Data management*

Genomic insights are dependent on large datasets of high quality data. However, data are rarely shared between clinical trials and among cohorts, hampering the ability to leverage the power of modern genomic technologies. Until now, efforts had exclusively focused on a narrow set of diseases, thereby limiting the scope of clinical relevance. Through its research consortium and additional partners, CKCC worked with multiple active clinical trials that were using a DNA-based approach to compile multiple sets of data to improve the ability to identify new treatment options for children with any type of cancer.

The team first developed data agreements with each of their clinical partners, ensuring data privacy and security and incorporating the preferences of each institution. Understanding that each institution has specific requirements for transferring and managing data was

essential to moving forward, since no single data transfer solution would have been accepted by all partners. In response, the CKCC team was open to facilitating a variety of data sharing options, including via physical hard disk drives and web-based interfaces.

In addition to DNA information, the partnering clinical trials also generated RNA data from each tumor. The RNA sequencing information from each tumor comprised a large data set that required transfer from the relevant clinical institution to the CKCC team for analysis. Technical challenges included the considerable size of the data and methods to ensure data security and patient privacy. To address these concerns, CKCC created the infrastructure needed through security-enhanced transfer mechanisms and a comprehensive directory. Patient privacy was safeguarded via de-identification of clinical data.

The CKCC team received both archived tumor data from former patients (legacy patients), which were used to validate, test, and optimize its analytical approaches, and from children actively participating in clinical trials (prospective patients). Data from prospective patients were analyzed to discover new treatment options.

#### *Analysis of data*

During the first phase of this project, the team received data from 168 total patients: 57 legacy cases and 111 prospective cases. Of those, 131 cases total passed quality control and were analyzed using the team's computational analysis tools. The RNA-based approach detected many molecular alterations associated with cancer and signified new treatment options in all cases. In comparison, the standard DNA-based approach employed in the clinical trials only yielded new treatment options for about one third of cases. To be a valid treatment option, the suggested target either had a matching therapy already approved by the U.S. Food and Drug Administration (FDA) for another condition (e.g., a different cancer), or a matching experimental therapy that was available as part of a clinical trial.

The CKCC project was designed as a research study, not a clinical study, which are held to other standards. All RNA data were provided by clinical partners, rather than collected directly by the research team. Some partners generated the RNA data via a Clinical Laboratory Improvement Amendments (CLIA)-certified manner. CLIA regulations mandate quality standards for laboratory testing of samples from humans for the purpose of diagnosis, prevention, or treatment of disease. The CKCC protocol therefore recommends that all findings are CLIA validated.

#### *Molecular tumor boards*

An important aspect of the CKCC project was the team's participation in molecular tumor board meetings, which are regular discussions between physicians, researchers, and other health care providers regarding individual patients' prognoses and treatment plans. The meetings allow clinical specialists and researchers to interpret and use results from genomic analyses to inform clinical decisions for better patient outcomes. Over the course of the project, CKCC presented to all four core partners and adapted to variable timeframes. When the request was urgent, the research team was able to conduct their analyses in as few as 24 hours and provide viable treatment options intended to help clinicians make informed treatment decisions. In total, findings from 28 legacy cases and 58 prospective cases were presented at nine molecular tumor board meetings in person. Additional remote meetings

were conducted via teleconferencing, and one-on-one follow-up discussions were arranged for specific cases, upon request.

Due to the novelty of the CKCC analyses for tumor board discussions, the research team solicited feedback from its partnering clinics and continually refined its strategy for effective communication. As a result, the research team expanded its reference compendium with additional pediatric clinical data (from study cases and public datasets), identified more clinical trials that were available to patients, clarified the type of findings (such as whether an analysis was conducted to confirm, enhance, or challenge findings from a research partner), and added quality control processes. Instead of the originally envisioned online tool that would allow clinicians to upload, analyze, and disseminate data and analyses, clinical partners preferred the delivery of findings in presentations and reports.

#### *Computational improvements and tumor data catalog*

Throughout the project, the research team refined their computational methods and developed tools that automate almost all steps of the process, facilitating faster and more reliable genome analyses. To support the comparative analyses, the team created a catalog of RNA data representing more than 11,000 tumors. Although challenging, the team acquired these data from public repositories and individual researchers by managing permissions, shared data agreements, and contracts that required substantial staff time and effort. Additionally, the preparation of data was difficult and time-consuming, as publicly available data files were very large, existed in various forms, and were occasionally wrongly labeled. According to the researchers, their “Reference Compendium work has met with remarkable success, resulting in the largest known compendium of RNA-Seq data for cancer, with over 11,000 tumors representing more than 70 types of cancer.” The number of pediatric tumors represented in the catalog increased by 44 percent by the end of the project.

#### *Patient engagement and data sharing*

The team prioritized engagement of patients and their families to more effectively meet their needs and satisfy their preferences. Throughout the project, the research team attended pediatric cancer advocacy events and provided educational presentations and materials to the patient community about genomic approaches to cancer therapy. According to the researchers, many families stated their desire to donate their children’s genomic data to research and found some solace knowing that their experience and data can contribute toward advancing cures to childhood cancers. Patient and family response assessment formed a central part of the second stage of this work, the clinical registry.

The tumor RNA catalog has been made publicly available, supplying high quality data from patients in this project and many others to other researchers and health care providers. This was a significant step in pediatric cancer research communities, considering how pediatric cancer data are tightly controlled. This resource is anticipated to be integral to future research toward childhood cancer cures and treatments.

#### *Clinical registry development*

To understand and optimize the clinical impact of RNA analysis and communication strategies to physicians and patients’ families, the CKCC team developed a clinical registry. The registry was meant to guide physicians’ use of RNA analysis in clinical practice and capture the relative significance of how the clinical environment, inclusive of molecular

tumor board and genomic reporting, can be modified to better engage patients and their families in decision-making. As of July 2019, twenty-four patients were enrolled, analyzed, and represented in presentations to clinical teams and researchers. By the end of the project, the team will add 6-12 more patients. The registry also incorporates evaluation of clinical decision-making through review of transcripts from molecular tumor board meetings, also called consensus meetings, and interviews with clinicians. These findings will shed light on how physicians view genomic profiling information in the context of available therapeutic options and what additional training is needed for physicians to feel comfortable handling this information. A final aspect of this work is a qualitative evaluation of interviews of patients and families to better understand decision-making and response to genomic analysis and how to make genomic-focused precision medicine more accessible to patients and their families.

All software is open-source, and the team plans to make all data and analyses available to the public to benefit future research. This aspect of the project will be completed at the end of 2019 with the support of additional, external funds that were raised to match state funds.

According to the researchers, “Some early feedback provides preliminary evidence of the clinical utility of our approach. For instance, we have had several examples where kids, whose data were analyzed as part of the CKCC study, went on to enroll in clinical trials suggested by CKCC analysis.”

#### *Leverage of CIAPM funds*

In addition to CIAPM funds, the CKCC project was made possible by various sources of support. In-kind resources in the form of supplies and time were supplied by the Computational Genomics Lab at UCSC’s Genomics Institute, the UCSC TumorMap team, and the UCSC Xena Browser. CKCC’s clinical research partners provided a large amount of RNA data at no cost from clinical studies they conducted that were funded by other grants. Supplemental financial support was provided by a grant of \$2.5 million from St. Baldrick’s Foundation to build on and expand the CKCC precision medicine effort through the UCSC Treehouse Childhood Cancer Initiative.

#### **Expert Evaluation by Dr. Nikhil Wagle, MD**

The CKCC team made significant progress towards establishing an approach for comparing tumor RNA from children with currently incurable cancers to RNA data from thousands of other tumors (adult and pediatric) in order to provide novel potential therapeutic options for pediatric cancer patients. According to the evaluator, “The team should be commended on significant accomplishments to advance precision medicine in pediatric cancers.” Major accomplishments over the reporting period include the following:

- Established innovative partnerships between a cutting-edge genomics research group at a research university (CKCC at UCSC) with clinical trials at multiple academic medical centers and hospitals
- Established data transfer infrastructure to conduct the study
- Established and refined the data processing and computational approach to conduct the study



- Developed a catalog of RNA data, consisting of data from more than 11,000 tumors, with 1,584 from pediatric cases, and making it available to other researchers
- Successfully conducted comparative RNA analysis from 98 prospective patients
- Provided potential new treatment options based on analyzed tumor samples
- Referred some patients to ongoing clinical trials based on RNA analysis
- Presented at molecular tumor boards and communicated results to clinical teams
- Established a clinical registry to optimize physicians' use of RNA analyses and better engage patients and their families in decision-making
- Developed new communication and educational tools for clinicians to use and communicate findings from RNA analyses

The main finding that the team's RNA-based approach to discover new molecular insights was successful in 100% of cases is a remarkable result. The team continues to confirm whether the CKCC recommendations were helpful, if they impacted decision-making, and, most importantly, if they improved outcomes for children with cancer. While others have tried RNA-based tests for cancer diagnostics without much success so far, the CKCC approach is novel in several ways that makes it very promising, including the creation and use of a large catalog of tumor RNA data and cutting-edge computational approaches. With additional future analyses and a continually changing landscape of data and diagnostic approaches, the thresholds for identifying a new treatment option may change and the overall success rate may decrease, but the 100% rate of new information about the case reported in this project is a promising sign that this RNA-based approach may provide significant benefit to patients.

### Assessment of How the Project Addressed Programmatic Goals

The project successfully addressed multiple goals of the CIAPM program, including 1) demonstrating the promise of precision medicine, 2) efficient and effective data management, 3) impact for patients, 4) engaging participants, and 5) developing quality partnerships. More detail on each of these points is provided next.

#### *Demonstrate the promise of precision medicine in a specific disease area*

According to the evaluator, "The project has established that data sharing and data interpretation from multiple centers can be performed in a centralized way, which is critical for the advancement of precision medicine. The infrastructure and tools established are an important first step in a pediatric oncology precision medicine program. It now remains to be seen if this will truly allow the implementation of effective precision medicine and improve treatment options and ultimately outcomes for children with refractory and resistant cancers. The improvement in diagnostic power, in particular, is an early sign that this approach might provide significant benefit to patients."

#### *Efficient and effective data management*

The project was highly successful in establishing a data sharing effort in which data from multiple clinical trials was aggregated, integrated, and analyzed in a centralized way. The infrastructure the team established and the computational and analysis tools they developed are an important step in a pediatric oncology precision medicine program. The project was also highly successful in leveraging existing data for the generation of the large RNA catalog, which was assembled from multiple public data sources and includes both

legacy and prospective cases obtained from partner institutions. Through the clinical registry, the team is also integrating genomic analysis with clinical interpretation and considerations for improved patient engagement in decision-making. Together, these resources and tools represent important precision medicine capabilities, and by making the RNA catalog publicly available, this work has the potential to accelerate advances in the field more broadly.

#### *Impact for patients*

The project had an impact on patients in that the CKCC team was able to provide clinical teams with new molecular tumor information to aid with patient care. As mentioned above, it remains to be seen if the CKCC recommendations impacted clinical decision-making and improved outcomes for children with cancer, but the high rate of finding new treatment options suggests that this RNA-based approach might provide significant benefit to patients.

#### *Participant engagement*

The development of new communication and educational tools for clinicians to better engage patients and their families in decision-making is a laudable achievement. Following the completion of the entire study in late 2019, the team will provide a quantitative description of treatment considerations, including factors raised by the clinicians with regards to family concerns and cultural barriers. The team did participate in several patient education forums, which is, at this time, an effective way to engage with patients.

#### *Development of partnerships*

The team was very successful in developing numerous meaningful partnerships for this project. The fact that they were able to identify multiple active clinical trials that were interested in collaborating with the CKCC team, sharing raw data, and receiving analytic results and treatment recommendations, signifies the strong demand in the field of molecular cancer diagnostics for novel approaches and the enthusiasm for the CKCC approach. The team has also worked effectively with several disease advocacy foundations to educate the public about their efforts.

#### *Potential for impact on precision medicine and on the State of California*

According to the evaluator, “With further development, the CKCC project has the potential to significantly impact the care of children with cancer in the State of California.” The team learned and internalized an important lesson regarding the readiness of patients and physicians to receive and understand the data generated by the RNA-based approach. All processes and best practices the team established—including the analysis approach, the automation of many of its steps, its accommodation to different regulatory environments, the expansion of the tumor data catalog, and the communications with patient care teams—will be important for advancing this data-driven precision medicine approach.

## **Conclusion**

CIAPM funding for this demonstration project enabled the CKCC genomics research team at UCSC to establish an innovative collaborative model with clinical trials at academic medical centers. By assembling a successful data transfer and analysis approach for large data sets, analyzing numerous patient samples, and establishing effective communication with clinicians, the team obtained evidence that its data-intensive RNA-based tumor analysis

approach is able to discover new treatment options when DNA-based analysis does not. Additionally, the team built the infrastructure to advance its innovative computational approach toward clinical testing. The CKCC approach has the potential to scale across pediatric and adult cancers to impact broad patient populations in California and beyond.

## 2. Precision Diagnosis of Acute Infectious Diseases

Principal Investigator: Dr. Charles Chiu, UC San Francisco  
Project Period: September 1, 2015 – December 31, 2018

### *Research Team and Collaborators*

University of California, San Francisco

Charles Chiu, MD, PhD  
Steve Miller, MD, PhD  
Joseph DeRisi, PhD  
Eric Chow, PhD  
Steven Hauser, MD  
Michael Wilson, MD  
Michael Geschwind, MD, PhD  
Jeffrey Gelfand, MD  
Felicia Chow, MD  
Jacob Appelbaum, MD, PhD  
Chaz Langelier, MD, PhD  
Kristen McCaleb, PhD

University of California, Los Angeles

Romney Humphries, PhD. D(ABMM)  
Jeffrey Klausner, MD, MPH  
Tara Vijayan, MD

University of California, Davis

Christopher Polage, MD  
Stuart Cohen, MD

University of California, Berkeley

Brent Fulton, PhD, MBA

California Department of Public Health

James Watt, MD  
Dongxiang Xia, MD, PhD  
Sharon Messenger, PhD  
Debra Wadford, PhD

Synapse, Inc.

Jonathan Hirsch, PhD  
Laurie Gomer, MBA

DNAexus, Inc.

Davis Shaywitz, MD, PhD  
Omar Serang, BS

Quest Diagnostics, Inc.

Rick Pesano, MD, PhD  
John Leake, MD, MPH

Illumina, Inc.

Mostafa Ronaghi, PhD  
Google, Inc.  
Sri Madabushi, PhD  
Abbott Laboratories, Inc.  
American Tissue Culture Collection  
Children's Hospital Colorado / University of Colorado in Denver, CO  
Children's Hospital Los Angeles  
Children's National Medical Center at Washington D.C.  
DNAnexus, Inc.  
Oxford Nanopore Technologies, Inc.  
Quest Diagnostics, Inc.  
St. Jude Children's Research Hospital in Vanderbilt, TN  
Syapse, Inc.  
University of California, San Diego  
University of Maryland  
U.S. Food and Drug Administration  
Zuckerberg San Francisco General Hospital and Trauma Center

## Scientific Background and Context

### *Infectious Diseases*

Diagnosing a disease quickly and accurately is essential for patients with acute infections. As many infectious diseases present similar signs and symptoms, laboratory tests have been developed to detect particular microbes that cause a patient's illness. Physicians commonly test the blood, urine, and other fluids and subject the patient's sample to microbiology laboratory tests for specific disease agents based on the patient's symptoms and history, such as recent travels. With tens of thousands of possible disease agents known, a physician must prioritize which disease agents are more likely to be responsible and select the appropriate diagnostic tests.

Infectious agents like bacteria, viruses, fungi, and parasites come in many forms and are therefore only detected by different diagnostic tests. Despite advances in diagnostics, many patients with suspected infections are prescribed antimicrobial therapies, such as antibiotics, rather than thoroughly tested, diagnosed, and treated for the underlying, specific disease agent. As a result, the current stock of available antimicrobial therapies are overused, increasing the risk of antimicrobial resistance.

For critically ill patients, when initial diagnostic tests do not reveal the cause of illness, additional testing is often costly and time-consuming, and may only provide vague or contradictory results. Health care decisions for hospitalized patients must therefore often

be made on the basis of limited and imperfect information, with an increased likelihood of death as many therapeutics may be tried without success.

In California, it has been estimated that 240,000 infections are annually contracted from hospital stays alone, costing the health care system approximately \$3.1 billion.<sup>3</sup> Many more infections are acquired in community settings.

### *Metagenomic next generation sequencing*

To help address this problem, the research team pioneered an approach that has the potential to identify thousands of disease agents in a single minimally invasive test. The state-of-the-art genome analysis approach is called metagenomic next generation sequencing (mNGS), where ‘meta’ represents all organisms, ‘genome’ encapsulates the entire genetic (DNA) material of an organism, and ‘sequencing’ refers to genome analysis. The technique can detect almost all known infectious agents in a single sample from a patient, quickly revealing the cause of bacterial, viral, fungal, and parasitic infections, many of which routinely elude physicians. Such information is also highly useful for ruling out an infection, as in the case of an autoimmune disease that may present similar symptoms.

This data-driven precision medicine test works by analyzing the genomic information present in a patient’s sample, such as cerebrospinal fluid (CSF) or blood. Because bacteria, viruses, fungi, and parasites also have DNA-based genomes and can be recognized by unique genetic signatures, infectious agents may be detected among total genomic information in a sample, which is mostly human.

DNA sequencing technologies have advanced at an exponential pace, dramatically lowering both speed and cost. While a genome analysis approach for diagnosing infections has long been theoretically possible, the process would have taken months and cost millions of dollars to diagnose a single patient one decade ago. The research team has so far optimized their system to offer the mNGS diagnostic test for \$2,200, providing results within a few days of sample collection.

Just prior to the awarding of funds, the research team and its collaborators had obtained evidence that supported the life-saving potential of the technique for critically ill patients. The mNGS test radically sped up the diagnostic process and led to a diagnosis when conventional tests had failed. For example, the team was presented with a case of a 14-year-old boy who was near death with brain inflammation that remained undiagnosed despite months of lab tests, expensive imaging technologies, and invasive procedures, such as brain biopsy. The mNGS test was able to identify the infectious agent, and the boy quickly recovered after targeted treatment.<sup>4</sup>

---

<sup>3</sup> <https://www.calhospital.org/infection-control>

<sup>4</sup> Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, and Chiu CY. (2014) Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *New England Journal of Medicine* 370:2408-2017.

## Project Summary

Advancements in clinical diagnostics have been made possible by faster and more cost-effective technologies, including highly innovative computational tools that facilitate rapid interpretation of large data sets. This development has the potential to fundamentally change the diagnostic approach for patients with infectious diseases by enabling a comprehensive analysis that can identify almost all infectious agents in a single test.

The research team previously developed a genome-based test for infectious disease diagnostics and obtained evidence that this comprehensive test can accelerate identification of infectious agents in previously difficult-to-diagnose cases. The test has the potential to save the lives of critically ill patients by enabling appropriate treatments in a timely fashion.

The goals of this project were to expand the use of metagenomic next generation sequencing mNGS for diagnostic tests of meningitis, encephalitis, sepsis, and pneumonia from a research lab to a routine clinical setting and to collect evidence for its clinical and economic utility as compared to conventional testing. Moving a diagnostic test into clinical practice is a complex process that includes 1) developing procedures that ensure accurate and reliable test results and confirming the performance of the test in a licensed diagnostic laboratory that is certified to federal standards (Clinical Laboratory Improvement Amendments, CLIA); 2) determining the test's clinical validity, that is, how well it identifies, measures, or predicts the presence or absence of a clinical condition; 3) creating processes for ordering, billing, and reimbursing the test; 4) educating the medical profession about the utility and use of the test; and 5) scaling the test to allow for large volume testing.

Toward their ultimate goal of facilitating accessibility to this precision medicine test to critically ill Californians and beyond, the team has also initiated a cost effectiveness analysis and made progress on multiple technical, logistical, and regulatory steps. With additional third-party funding in place, the team is well positioned to complete the project and, if successful, transform clinical diagnosis for infectious diseases.

## Project Achievements

Over the course of the project, the team 1) finalized its clinical-grade software and expanded it to a cloud computing platform, 2) confirmed the mNGS test in a clinical laboratory for diagnosing the causes of brain inflammation, 3) optimized the test for faster results, 4) confirmed the test in a clinical laboratory for diagnosing infections through blood plasma, 5) initiated clinical studies of critically ill hospitalized patients to compare conventional and mNGS-based approaches, 6) established and regularly convened a clinical microbial sequencing board (CMSB, a multidisciplinary group that meets to discuss complex patient cases), 7) started to evaluate the impact of the mNGS test on overall costs and clinical outcomes compared with conventional diagnostic approaches, and 8) pursued additional efforts to make the test widely available. More details of each achievement are provided in the remainder of this report.

### *Clinical-grade software*

This project leveraged prior computational work by the team to efficiently and rapidly identify infectious agents based on their genomic signature using existing public reference databases. The research-grade software is currently available as open source.

As part of the project, the team also enhanced and finalized the clinical-grade version of the software. In contrast to the research-grade software, analysis of complex genome data is automatic and results are provided in a format that can be interpreted by experts in conventional infectious disease diagnostics without prior computational expertise.

The clinical-grade software has been migrated to a secure internet-accessible (cloud) computing platform that maintains patient confidentiality, enabling deployment on laptops, computational servers, and in the cloud. In addition, a web portal has been developed that facilitates continuous 24/7 processing and visualization of mNGS assay data from any location. Taken together, the enhanced accessibility will support future use of the software by other laboratories, starting with the research team's partners at St. Jude's Children's Hospital (project collaborators), the U.S. Centers for Disease Control and Prevention (CDC), and Simecere Diagnostics, a clinical diagnostics company in China.

### *Diagnosing causes of brain inflammation*

In order to move beyond the research laboratory and offer the mNGS test as a clinical diagnostic test that can be ordered by physicians treating patients worldwide, the team confirmed the performance of the test in the UCSF Clinical Microbiology Laboratory, a licensed diagnostic laboratory that is certified to federal CLIA standards. While the test can in principle be used for different types of human tissue samples, clinical performance must be confirmed individually for each sample type. The team first pursued testing of CSF, which surrounds the brain, for diagnosis of brain inflammation. A summary of this CLIA validation study was published in 2017, and a more detailed article on the study has been published as a preprint, as listed in Appendix B.

To accomplish their goal, the team established the standard operating procedures for all steps of the mNGS test and then confirmed high sensitivity and accuracy of the test in their CLIA-certified laboratory. After establishing a test code and billing protocols, the team started offering this test as a clinically billable diagnostic reference test to physicians worldwide, a critical step toward making it broadly available.

### *Test optimization for faster results*

The research team conducted an optimization study to further reduce the amount of time necessary to obtain sequencing results. They compared two state-of-the-art assays from Illumina and Nanopore companies, finding that the Nanopore sequencing assay was able to provide results in shorter time: 8 hours versus 48-72 hours required for the other assay. A manuscript is currently in preparation to share the details of their study and results.

### *Diagnosing causes of blood infections*

Work toward demonstrating the sensitivity and accuracy of the mNGS test for analyzing blood plasma in the CLIA-certified laboratory progressed well, with many of the performance parameters completed by the end of the project. The work included validation of detection, host interferences, stability, accuracy, and reproducibility studies using approximately 200



plasma samples. In addition, the sequencing assay was automated for blood plasma, relieving staff time without sacrificing speed or quality. A manuscript is currently being prepared for publication.

### *Clinical studies*

In order to determine how well the mNGS test works to diagnose brain infections in routine clinical settings, the research team initiated a prospective study with clinical partners at four UC medical centers and five additional hospitals in California and nationwide. Within the first year of launching the clinical study, 190 patients were enrolled and their consent had been obtained using a standardized clinical protocol. While each center performed conventional infectious disease testing for each patient, at the time of enrollment, the causes of their brain inflammation remained unknown. Patient samples of CSF were sent to UCSF for analysis. The mNGS test was positive for an infectious agent in 37 cases. In 14 cases, the mNGS test yielded a diagnosis when extensive conventional testing failed, and in the other 23 cases, it confirmed the results that were eventually obtained using conventional testing. The team already published an example of a positive patient outcome using the mNGS test: the diagnosis of an unexpected hepatitis E infection in a patient with brain inflammation that led to early and effective treatment with the appropriate drug. The team has published additional examples demonstrating the utility of the clinical mNGS test for diagnosing various infectious causes of brain inflammation.

For the remaining 153 patients participating in the study, the mNGS test was negative. Based on clinical experience with brain inflammation, the comprehensive nature of the test, and its high sensitivity and specificity, the team argues that negative test results likely indicated non-infectious causes of illness. In response to surveys, participating clinicians commented that a negative mNGS test result gave them ‘confidence’ and ‘reassurance’ that in some cases prompted them to begin workup and/or treatment for non-infectious causes earlier than they would have without mNGS testing.

While the team continues to recruit additional patients for this study, the preliminary data suggest that mNGS testing can 1) accurately diagnose infections by confirming conventional test results, 2) diagnose infections when conventional testing fails, and 3) help rule out infection in cases where brain inflammation occurs due to other causes. Assessment of the overall clinical utility of mNGS testing—how well the test can provide information about diagnosis, treatment, or management of a disease—in comparison with conventional testing will be possible once all data from this study, including outcomes, have been analyzed.

The research team also initiated work toward a clinical trial for sepsis. Approximately 450 samples have been collected to-date, 100 of which are positive for a bacterial and/or viral infection. Once their validation of using blood plasma for sepsis is complete, the clinical trial will commence, likely with a pilot trial beginning with UCSF patients.

### *Clinical Microbial Sequencing Board*

Given the novel nature of the mNGS test for the diagnosis of infectious diseases, the team established a multidisciplinary CMSB, which is modeled after cancer tumor boards and consists of a group of researchers and clinicians who are experts in interpreting mNGS test results. This consult board held weekly teleconference meetings with the treating physicians at the participating medical centers to discuss individual patient cases, provide expert input

on the interpretation of mNGS test results, and aid in developing the best course of treatment for the patient. The participating clinicians were asked to provide feedback on their experience with the mNGS test during this study, and their comments have been very positive overall.

#### *Economic impact analysis*

The team is collaborating with a health economist at UC Berkeley, Dr. Brent Fulton, to conduct a cost/outcomes analysis to determine whether more rapid diagnoses based on mNGS testing will result in lower costs and improved clinical outcomes. One example is the above-mentioned case of a boy who underwent months of conventional lab tests, expensive imaging technologies, and an invasive brain biopsy without a definitive diagnosis. In his case, an earlier definitive diagnosis with the mNGS test would have significantly limited suffering and reduced costs by avoiding the need for multiple hospitalizations and unnecessary, expensive testing. Not all cases of suspected infections will so obviously benefit from mNGS testing, however, so the cost/outcomes study is designed to develop metrics to determine which patients with brain inflammation have the greatest potential to experience improved health outcomes and/or a reduction in health care utilization and expenditures as a result of using the test. This economic study is ongoing and depends on the completion of the clinical study, which continues to receive support from additional, external funding.

#### *Additional efforts to enable broad availability of the test*

To expand the use of this diagnostic test by licensing it to other clinical laboratories across the country, approval by the U.S. Food and Drug Administration (FDA) as a “breakthrough device” was granted in 2019. Since this pioneering diagnostic test relies on cutting edge approaches, that is, large-scale genome analysis and advanced computational tools that are not routinely regulated by the FDA, future discussions will shape the development of standards and best practices for the use of mNGS in the diagnosis of infectious diseases.

#### *Leverage of CIAPM funds*

The project benefited from additional sources of support. Co-investigator Steve Miller volunteered his time to the project, and UCSF laboratories contributed reagents and effort in-kind. Quest Diagnostics, Inc. and Abbott Diagnostics, Inc. made in-kind contributions of hundreds of human tissue samples for the clinical confirmation work, and the Chan-Zuckerberg BioHub contributed salary support for personnel. The project attracted substantial additional funding from various philanthropic sources, for a combined total of \$5.5 million, including from the Sandler Foundation, the William F. Bowes, Jr. Foundation, the Charles and Helen Schwab Foundation, the Steven and Alexandra Cohen Foundation, the George and Judy Marcus Innovation Fund, and the UCSF Medical Center Clinical Laboratories.

### **Expert Evaluation by Dr. Nancy Cox, PhD**

Transformative technological advances within the last decade have made large-scale genome analysis so much faster and cheaper that it is now in reach of clinical diagnostics and is starting to have an impact on diagnosing cancer and severe developmental defects. Bringing “next generation sequencing” (NGS) to infectious disease diagnostics is an exciting new application of this cutting-edge technology that has the potential to fundamentally

change the diagnostic approach for critically ill patients and accelerate diagnosis in previously difficult-to-diagnose cases.

The project is based on the team's previous work of developing an NGS-based diagnostic test that removes both the need for prior assumptions about the likely infectious agent and time-consuming lab cultivation methods to test for infectious agents. This was accomplished by developing innovative computational tools that can interpret large amounts of genome data across all organisms present in a sample (metagenomic) to identify the genomic signature of the infectious agent. The team has made impressive progress and has significantly advanced the mNGS test from the research lab toward clinical implementation. Specifically, the team's substantial achievements include the following:

- Confirming the performance of the mNGS test for diagnosing brain infections in a licensed diagnostic laboratory that is certified to federal standards (CLIA confirmation)
- Making considerable progress toward CLIA confirmation for mNGS-based diagnosis of blood infections
- Automating assay analyses for faster results
- Launching a patient study to assess the clinical and cost utility of the mNGS test in diagnosing brain inflammation
- Launching a clinically billable mNGS infectious disease diagnostic test available to hospitals worldwide
- Establishing and regularly convening a Clinical Microbial Sequencing Board to support and standardize interpretation of mNGS results in the clinical context of each patient

According to the evaluator, "They have made substantial progress in bringing next-gen technology into laboratory medicine for infectious disease. Broad application of this should both improve patient care, and in the long run, reduce medical costs. California may also benefit directly, if much of the early business comes to the Chiu lab and to the extent that they license the technology and software to others. Similarly, as federal funding for the development of new -omics biomarkers and diagnostics increases due to newly minted strategic plans, the state should benefit from the increased employment that comes with increased federal funding."

The evaluator added, "Achieving CLIA confirmation of an innovative diagnostic test that is fundamentally different from previous tests is impressive in this short period of time and represents a significant step toward moving it from a research setting toward clinical use. It is important to keep in mind, though, that the team has not yet determined the accuracy with which the test identifies the presence or absence of an infection—its clinical validity. One of the potential benefits of this comprehensive test is that it can be used as a "rule-out" test—i.e., a negative result (no infectious agent detected) can increase confidence in a non-infectious diagnosis, speed up the diagnostic process, and help ensure that patients receive the correct treatment as quickly as possible. However, it has not yet been determined how often the test returns a false negative result (no infectious agent detected when there was an infectious agent present). Preliminary data from their ongoing clinical study look promising, especially with regard to being able to identify infectious agents in cases where conventional testing has failed. Nonetheless, the team needs to be sure to clearly explain

these caveats when reporting results to treating clinicians, and the Clinical Microbial Sequencing Board they established is a great mechanism to do so.”

Regarding the research team’s acceleration of providing sequencing results and diagnostic recommendations, the evaluator made the following statement, “This is a quite significant achievement. Rapid diagnosis of infection is critical to outcomes in many conditions, and this project is bringing that into practice. A lot of the most difficult work has now been done, and it will be exciting to see the consequences of this research unfold. One of the intriguing outcomes may be a more accurate picture of the frequency of autoimmune encephalitis, which is poorly diagnosed and may be more common than has been appreciated. Failing to identify a pathogen with these more rapid and sensitive techniques may improve the ability to recognize this important disorder.”

Further, the evaluator commented, “This project’s primary advancements involve patient care and health care costs. There are challenges to cost/benefit analyses for new technologies because new technology costs are easily and immediately measurable (and usually higher than for existing technologies) while measurement of benefits takes some time to accrue. Nevertheless, the value of these particular advances are likely to be sufficiently high as to be broadly cost-effective within a relatively short period of time. The primary reason is that rapid diagnosis of pathogen causes of certain diseases is strongly associated with outcomes. Sepsis, one of their key target diseases, is among the 10 most expensive conditions at US medical centers and early understanding of the relevant pathogen and its profile of susceptibility and resistance to antibiotics is critical to favorable outcomes. With really expensive conditions, you need only modest clinical impact to have hugely important financial impact.”

Regarding the research team’s future plans, the evaluator stated, “Dr. Chiu is very likely to achieve stated future goals – this is a clear priority for the lab and the studies already completed provide the necessary background for achieving these goals. Over the longer term, this lab is likely to contribute to advances in pathogen laboratory testing more generally. I believe that the long-term goal of using RNA-Seq to understand host response in the context of the observed medical phenotype may indeed be particularly effective in making more immediate determinations of general class of disease process (viral, bacterial, autoimmune) with potential for further differentiation. The technologies they propose could substantially improve diagnosis and surveillance for diseases like autoimmune encephalitis that remain poorly characterized and almost certainly under-diagnosed. But it should be noted that the current strategic planning at NHGRI [the U.S. National Human Genome Research Institute] suggests that RNA-Seq and other next-gen –omics measurements that may be used in diagnostics and/or as biomarkers may enjoy considerable opportunity for federal funding in the near term, and Dr. Chiu and colleagues should be well positioned to champion their ideas for additional federal dollars.”

### **Assessment of How the Project Addressed Programmatic Goals**

The project addressed many of the goals of the CIAPM program, including 1) demonstration of the promise of precision medicine, 2) efficient, effective data integration and analysis, 3) development of precision medicine capabilities, 4) clinical and commercial potential, 5) potential downstream use of tools, measurements, and data, including open public

accessibility of generated data and publications, 6) potential to scale, 7) participant engagement, 8) impact for patients, and 9) development of partnerships. More detail on each of these points is provided next.

#### *Demonstration of the promise of precision medicine*

According to the researchers, “We completed validation of the first-ever clinical metagenomic sequencing test for diagnosis of infectious causes of meningitis and encephalitis, and demonstrated its utility in a prospective clinical trial that compares the mNGS test head-to-head with all conventional microbiological testing. This is a direct demonstration of the promise of precision medicine for precision diagnosis of infectious diseases.”

Regarding the second part of their project, they added, “the supplemental project resulted in expanded clinical validation of metagenomic sequencing to new sample types, including plasma and other body fluids, and to new, rapid sequencing technologies, such as nanopore sequencing.”

#### *Efficient, effective data integration and analysis*

Existing databases were critical to this study. The project utilized reference databases from the National Center for Biotechnology Information and the FDA, as well as existing and extensive metagenomic validation data previously developed. Data that were generated as a result of the first part of this project were used to train machine learning-based diagnostic and predictive models of infection and non-infectious illness. To better understand the economic impact, patient data and health care utilization databases were used, such as the Truven database.

According to the researchers, “We enabled complex metagenomic sequencing data to be rapidly and efficiently analyzed with a clinical-grade bioinformatics pipeline. Analysis is automatic and begins as soon as the sequencing files are generated, and results are provided in a format (including a graphical visualization suite) that enables data to be analyzed by laboratory directors without prior bioinformatics expertise. Clinical and laboratory patient data are integrated and discussed along with the mNGS test result during a weekly clinical microbial sequencing board.”

#### *Development of precision medicine capabilities*

According to the researchers, “we have generated an extensive database that includes mNGS data, both DNA and RNA library data, and patient clinical, laboratory, and radiographical data abstracted from the electronic medical record. Generation of this data will enhance our capability to mine this valuable dataset in the future. As an example, efforts are underway to perform machine learning analyses to generate predictive models for discriminating between bacterial, viral, and non-infectious causes of meningitis/encephalitis in patients. Similar sub-analyses of these datasets can be used for characterizing the microbiome, human host gene expression responses, and microbial antibiotic resistance that can be used to develop new diagnostic biomarkers for infectious diseases.”

#### *Clinical and commercial potential*

According to the researchers, “To our knowledge, this is the first-ever clinical metagenomic sequencing test for infectious disease diagnosis. Unlike all other microbiological tests, we

have shown that this test is capable of detecting the full spectrum of pathogens, including viruses, bacteria, fungi, and parasites, in a single assay. We have also shown in several instances that mNGS testing of patients in the study directly impacted clinical outcome. In addition, the sequence-based identification used by the mNGS assay applies to a variety of other clinical sample types and infectious syndromes.

There is significant commercial potential with the validated mNGS assay. In May of 2017, we launched the mNGS assay for pathogen identification in meningitis and encephalitis as a clinically billable diagnostic reference test available throughout the United States and worldwide. We have also identified codes for test reimbursement, [the first] for this kind of test. FDA approval [has] raise[d] the possibility of licensing the test to other laboratories and generating clinical revenue from testing. There is also significant commercial potential in the SURPI+ computational analysis pipeline, which can be licensed to other clinical and research groups for use in analyzing their data.

Our clinical assay is now being routinely offered to outside hospitals for reference lab testing (we have over 50+ clients to date, including the Mayo Clinic, ARUP, UCLA, Boston Children's Hospital, and Brigham and Young Hospital). We have received interest from venture capital firms and large biotechnology companies for licensing our assay and/or forming a startup. We also intend to make the SURPI+/SURPIrt software commercially available for use."

*Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications*

Several publications have been released or are in preparation, as listed in Appendix B. The pathogen detection software is currently trademarked and copyrighted and available for free to nonprofit organizations and academic/clinical laboratories, while also available to commercial entities via a license. In addition, the research team contributes all sequencing data to public reference or restricted access databases, depending on the level of sensitivity of the data.

*Potential to scale*

According to the researchers, "The project has vast potential to scale, given that we have only developed and implemented complete clinical infectious diseases assays for meningitis/encephalitis (from CSF) and sepsis (from plasma). We intend to scale clinical metagenomic sequencing to all body fluids, and also convey host response and antibiotic resistance information in the laboratory test results."

*Participant engagement*

According to the researchers, "The implementation of the study involved directly engaging both patients and their providers. Patients are consented and educated on the mNGS test. Providers are given immediate feedback on the results of the test via the real-time teleconferencing CMSB.

Infectious diseases are common and have impact across all segments of the population, disproportionately impacting the poor and underserved. The mNGS assay developed with [CIAPM] funding has the potential to improve workup and treatment of patients with infectious diseases, especially persons of low socioeconomic status and minorities who often have broader exposure to infectious agents. The patients enrolled in the PDAID

[Precision Diagnosis of Acute Infectious Diseases] study reflect a broad range of racial and ethnic diversity.”

#### *Impact for patients*

Of the initial cohort of patients enrolled in the study, 21% were found to be positive for an infectious agent. In 8% of cases, the mNGS test provided a diagnosis when conventional tests failed.

According to the researchers, “We anticipate that expansion of testing as enabled by clinical validation will produce more timely diagnoses in infected patients and improve clinical outcomes as well as reduce health care costs. Of note, we have found that a negative mNGS result may be particularly useful as a “rule-out” test for infections, leading to consideration of workup and treatment for non-infectious causes of a patient’s illness (for example, steroids and immunosuppressive agents for a patient with autoimmune encephalitis).”

#### *Development of partnerships*

The team was successful in assembling an impressive 9-hospital partnership for the clinical study, exceeding the originally proposed 3-hospital collaboration. The team also successfully leveraged industry contributions, including from Quest Diagnostics Inc., Abbott Diagnostics, Inc., and DNAnexus. The team also worked closely with the FDA to submit an application for “breakthrough device” designation and preliminary approval to license the test.

## **Conclusion**

The team achieved its goal of confirming the performance of the mNGS test and collected preliminary evidence that their approach can directly impact clinical care and improve patient outcomes. The recent FDA “breakthrough device” designation for diagnosis of infectious diseases will expand the scope and availability of the test. Toward their ultimate goal of making this mNGS test widely accessible to patients in California and beyond, the team is well on its way to completing clinical studies, and has initiated a cost/outcomes analysis and multiple technical, logistical, and regulatory steps to enable broad use of the test. This project represents a prime example of a data-intensive precision medicine approach with excellent potential to scale and has been effective in moving the approach from a research setting toward clinical use.

### 3. Artificial Intelligence for Imaging of Brain Emergencies

Principal Investigator: Pratik Mukherjee, UC San Francisco  
Project Period: February 1, 2017 – December 31, 2018

#### *Research Team and Collaborators*

University of California, San Francisco

Arash Afshinnik, MD  
Claude Hemphill, MD  
Nerissa Ko, MD  
Geoff Manley, MD, PhD  
Pratik Mukherjee, MD, PhD  
Esther Yuh, MD, PhD

University of California, Berkeley  
Jitendra Malik, PhD

Brain Trauma Foundation

Community Regional Medical Center in Fresno

Stanford University

TBI Endpoints Development Project

Transforming Research and Clinical Knowledge in Traumatic Brain Injury Consortium

UC Berkeley

Zuckerberg San Francisco General Hospital and Trauma Center

#### Scientific Background and Context

##### *Neurologic injuries*

Neurologic injuries such as stroke and Traumatic Brain Injury (TBI) affect 15 million Americans every year and account for 7% of annual health care costs. Moreover, irreversible brain damage can begin to occur within minutes of injury, and immediate diagnosis and evaluation options have not been available to aid this immediate critical window. To this point Computed Tomography (CT) scanning is used to diagnose brain injury and damage, but is limited to the ability and time of a radiologist to read and interpret these images. This subjectivity leaves not only a lack of rapid diagnosis methods, but also has yet to provide a reliable biomarker for studies of these brain injuries.

##### *Computed Tomography (CT) scanning*

CT scans take a series of X-ray images to recreate the image of bones, blood vessels, organs, and other tissues inside the body and is frequently used to identify internal injuries and indications of disease. Currently, the scans are subject to radiologist interpretation and “grading” which is subjective and time consuming.



### *Artificial Intelligence (AI)*

Artificial Intelligence is the employment of computer/machine-based technology simulations of intelligent human behavior to accomplish tasks. These often include tasks such as visual perception, decision-making, translation, and image recognition. AI technologies are created through many iterations of training a program to recognize input and perform tasks, making it more precise as it receives more data.

### **Project Summary**

Advances in image recognition, such as those employed by Google and Facebook to recognize faces in images has yet to be applied in the same manner to medical imaging. Automated recognition of injury in brain images could improve rapid detection of emergencies as well as reduce physicians' time constraints and subjectivity common in these diagnoses. The Artificial Intelligence for Imaging of Neurologic Emergencies team successfully applied state-of-the-art AI to automate CT scan evaluation. This technology demonstrated unprecedented improvement in time to diagnosis as well as cloud-based implementation, enabling medical professionals to use it anywhere in the world, and update a catalog of potential clinical biomarkers of neurologic injuries. This tool holds the potential to expedite treatment of irreversible damage and reduce subsequent long-term disability and death.

The goal of this project was to develop and train AI tools to recognize intracranial bleeding in a variety of acute brain injuries and hemorrhages, as well as implement this technology in a cloud-based manner to enable remote access from a variety of medical centers. After training the AI with more than 100,000 CT scans, they achieved detection of intracranial bleeding with greater than 99% accuracy, equivalent to the performance of board-certified radiologists. Additionally, they accomplished integration of this technology into a cloud-based platform, which can be used with several different CT scanning devices and protocols at many hospitals.

This project also recognizes the particular need for clinically effective, cloud-based, automated image analysis in underserved areas of the U.S. and developing world, as TBI and stroke disproportionately affect underprivileged populations. Provision of CT scanning equipment alone is not enough to meet the needs of these populations, as they are often still deficient in enough qualified radiologists to interpret the patient CT scans. Once approved for commercial use, the research team aims to make this technology available to meet the needs of Californians as well as patients around the world without rapid access to qualified diagnostic professionals.

To further test and apply their technology to integrative patient care, the team is working with a commercial partner to receive FDA approval for this device as a diagnostic test. Once approved, they will implement this technology with a focus on measurement of patient outcomes and application as the first quantitative biomarker of brain injury in clinical trials.

### **Project Achievements**

Over the course of the funding support, the team 1) demonstrated the promise of AI automated image analysis of neurologic injuries for application in routine clinical use,

research studies, and drug trials, 2) developed a computer vision AI tool to identify and quantify biomarkers of brain injury, 3) analyzed thousands of patient images with a greater than 99% accuracy of diagnosis, 4) optimized the AI tool for cloud-based, multi-scanner use, working towards FDA-approval for mass-system integration, and 5) aggregated data from multiple medical centers with various electronic health records. The team was also able to 6) attract additional funding. Further details on each of these points is provided below.

#### *AI in precision medicine image analysis*

Precision medicine requires obtaining as much data as possible to adequately make decisions about a patient and their potential diagnosis. While this is often limited by capacities of medical professionals to individually assess patient data, the team demonstrated the capacity of AI to overcome this human limitation and therefore analyze thousands of patient images to reach conclusive diagnoses.

#### *Development of an AI tool for neurologic injuries*

The team trained and tested a groundbreaking automated technology to analyze individual head CT scans for clinical purposes as well as detailed biomarkers of disease severity. In order to move beyond the standard simple and nonspecific classification of brain imaging scans as “normal” and “abnormal,” the team demonstrated their tool’s ability to reliably and objectively identify a wide spectrum of pathological lesions as well as their locations. This tool offers quantitative diagnostic information at an unprecedented scale.

#### *Optimization of a cloud-based platform*

Commercial implementation of this project is currently underway through work with an industry partner to utilize this technology in an FDA-approved manner in multiple clinical settings. Implementation of a cloud-based platform enables researchers and medical professionals to use this technology on multiple different CT scanning systems in both hospital and remote locations. This platform will enable rapid precise diagnosis for patients by eliminating barriers of access and delay to diagnosis during critical clinical periods and expediting time to treatment.

#### *Patient impact*

Application of this unique technology within the cloud extends the reach of this precision medicine tool to areas without access to enough trained radiologists. The team is aware that “In the U.S. and globally, TBI and stroke disproportionately affect the underprivileged.” The number of trained radiologists has not caught up with the proliferation of CT scanning instruments in developing areas. This leaves patients in these areas unable to have neurologic emergencies identified and treated. Creating even more need to implement this technology in all areas of need.

#### *Data management*

Data were obtained from over 100,000 patient CT scans and used to train the technology to recognize individual neurologic injuries. Once sufficiently trained, the technology was tested on thousands of existing image data from clinical repositories as well as multiple research centers.

The project also incorporated the first collaboration of neurological clinicians at UCSF and Community Regional Medical Center in Fresno, as well as UC Berkeley computer scientists.

The partnership expanded to include a leading industry partner and dozens more institutions from the Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) study.

#### *Leverage of CIAPM funds*

The team leveraged existing data and funds through collaboration with U.S. National Institutes of Health (NIH)-funded clinical research projects, computational infrastructure at multiple UCs, and multiple collaborating medical institutions. The team is pursuing additional long-term funding through NIH grant mechanisms.

#### **Expert Evaluation by Dr. Jack Tsao, MD, DPhil**

This project has made contributions in the domain of brain injury as well as the field of artificial intelligence and machine learning. While traditional methods use time consuming and costly trained, the expert evaluator commented, “these investigators were able to employ images reviewed throughout the course of care”, accounting for and correcting inaccuracies along the way. With the promise of equal or greater accuracy than that of a radiologist, the AI technology would directly impact the clinical treatment of patients with neurologic injury at the highest level of technical and scientific quality.

Specifically, the team’s substantial achievements include the following:

- Demonstrating the promise of precision medicine in technology through artificial intelligence neural networks
- Efficiently and effectively integrating data and analysis to design a precision medicine tool capable of interpreting CT scans with nearly the same specificity of a human radiologist
- Making progress toward FDA approval and commercialization for widespread deployment of technology in a variety of clinical settings
- Demonstrating potential to scale analysis compatibility with the two most widely used scanning systems, for mass integration in clinical care
- Collaborating with multiple clinical and computational institutions for joint design and implementation of the technology
- Demonstrating the potential to scale this precision medicine technology throughout California, nationally, and internationally

#### **Assessment of How the Project Addressed Programmatic Goals**

The project successfully addressed multiple goals of the CIAPM program including 1) demonstrating the promise of precision medicine, 2) efficient and effective data management, 3) development of precision medicine capabilities, 4) development/realization of the project’s clinical and commercial outcomes, and 5) potential to scale and leverage multiple electronic health records systems and other platforms.

#### *Demonstrate the promise of precision medicine in a specific area*

According to the researchers, “[The] project has demonstrated the promise of deep learning for automated image analysis of neurological emergencies in head CT scans and has

brought this technology to the threshold of clinical applicability for research studies, drug trials and routine clinical use.”

#### *Efficient and effective management of data*

According to the researchers, this project was successful in the integration, curation, and analysis of “detailed imaging data thousands of CT scans, which is a scale rarely achieved in the imaging literature, where most studies have a sample size of less than 100 scans.” They “leveraged existing data from thousands of patients available from clinical repositories as well as multicenter research studies.” Furthermore, the team demonstrated the promise of this technology to advance CT imaging by providing the first quantitative method of measuring neurologic injury, which up until this point was typically limited to “positive for abnormality” or “negative for abnormality.” Currently, the “FDA only recognizes these “normal” and “abnormal” head CT categories for patient stratification in TBI clinical trials, despite the fact that an ‘abnormal’ head CT spans a wide spectrum of pathological lesions, anatomic locations, and numbers and sizes of lesions.”

The data management from this team offers “more quantitative granular information than has been previously been possible for large sample sizes of head CT scans,” and may provide a step towards FDA recognition of these quantitative measurements beyond the existing “normal” and “abnormal” patient categories in TBI clinical trials.

#### *Development of precision medicine capabilities*

According to the researchers, “The deep learning technology we have developed offers automated individualized analysis of head CT scans for clinical purposes and also clinically relevant quantitative information that can serve as biomarkers for improved prognostic models of patient outcome in the most common neurologic emergencies, including TBI, ischemic and hemorrhagic stroke, and aneurysmal subarachnoid hemorrhage.”

#### *Development/realization of the project’s clinical and commercial potential*

The team described that they are “working with [their] industry partner for a cloud-based implementation of the deep learning technology that can be FDA cleared for routine clinical use beyond merely triaging suspicious scans for earlier radiologist interpretation, but also preventing misses of the emergency findings by the radiologist and treating physicians. [They] are also working toward improved prognostic models of TBI, hemorrhagic stroke and aneurysmal hemorrhage based on the additional quantitative information provided by the automated head CT analysis, which will hopefully better guide acute inpatient management of these prevalent disorders. In the U.S. and globally, TBI and stroke disproportionately affect the underprivileged.

There is a critical shortage of radiologists in underserved areas of the U.S. and developing nations, yet CT scanners are proliferating in the developing world. Cloud-based AI could rapidly identify emergencies and triage those scans for immediate attention by radiologists wherever available, and also be delivered at low cost due to vastly reduced local hardware needs and remote updating of algorithms. Therefore, continued work toward a clinically effective cloud-based implementation of our deep learning technology remains an important focus.”

### *Potential to scale and leverage electronic health records systems and platforms*

According to the researchers, “Imaging data utilized by this Demonstration Project came from many different medical centers utilizing different electronic health records systems. The ability to aggregate the scans across centers was aided by the common Digital Imaging and Communications in Medicine (DICOM) file format used universally for medical imaging. Continued scaling into the future will be facilitated by the ongoing development of more advanced imaging informatics standards such as the Brain Imaging Data Structure (BIDS).

The two initial publications from this project have already been made publicly available through the widely used arXiv server [listed in Appendix B] and future publications will also be made publicly available. The thousands of head CT scans and associated clinical and outcome data from the TRACK-TBI study will be made publicly available through the Federal Interagency Traumatic Brain Injury Research (FITBIR) repository, which is sponsored by the National Institutes of Health (NIH), the Department of Defense (DoD) and the Veterans Affairs (VA) administration.”

### *Potential for impact on precision medicine and on the State of California*

As stated by the researchers, “For the next 5 years, we hope to have achieved clinical implementation of the cloud-based platform at a global scale, including all residents of California. We also plan to generate automated quantitative biomarkers of neurological emergencies for large-scale precision medicine research studies and for clinical trials of experimental therapies. During this Demonstration Project, we have gained experience with aggregating, ingesting, curating, and analyzing large numbers of medical images from Northern California. We are currently developing a Standard Operating Procedure (SOP) for these steps that we are testing on the imaging dataset collected as part of the TRACK-TBI study presently being conducted in 18 Level I trauma centers from around the USA. Once perfected, this SOP can be disseminated via publication to the research community.”

The team overcame technological hurdles associated with accessing and interpreting a groundbreaking number of CT scans from multiple different electronic health record systems and CT scanners. As a result, they designed a technology capable of automating brain imaging analysis in a transferrable manner through many different health system platforms. They also maintained perspective regarding the importance of access to this technology through a cloud-based platform, as well as identification of clear biomarkers of brain injury in hospital and clinical study settings. Awareness of these perspectives throughout their study process enabled them to design a precision medicine tool with the capacity for implementation throughout California. Continued work towards proving equality and superiority of this technology to that of radiologists, as well as FDA device approval are important steps for the team to advance this automated precision medicine approach.

## **Conclusion**

CIAPM funding for this demonstration project enabled the Artificial Intelligence for Imaging Brain Injuries research team at UCSF to design an innovative tool and AI model for automating analysis of patient data towards a faster patient diagnosis and treatment plan. By working within multiple hospital systems and across multiple campuses to design this innovative precision medicine tool, the team created a transferrable model of automated image analysis capable of eliminating subjectivity in brain injury diagnosis. This AI image

analysis approach has the potential to scale automated image analysis to many different imaging modalities and for different parts of the body in a variety of different diseases to ultimately impact a wide range of patient populations in California and beyond.

## 4. Early Prediction of Major Adverse Cardiovascular Events Using Remote Monitoring

Principal Investigators: Dr. Brennan Spiegel, Dr. Noel Bairey-Merz, and Dr. Jennifer van Eyk; Cedars-Sinai Medical Center  
Project Period: January 1, 2017- December 31, 2018

### *Research Team and Collaborators*

Cedars-Sinai Medical Center  
C. Noel Bairey Merz, MD  
Jennifer Van Eyk, PhD  
Chrisandra Shufelt, MD  
Janet Wei, MD  
Margo Minissian, PhD, ACNP

University of California, Los Angeles  
Peipei Ping, PhD  
Corey Arnold, PhD

Agilent

AliveCor

Beckman Coulter

DocuSign

Fitabase

Fitbit

HealthLoop

Neoteryx

SCIEX

Tasso

Thermo Fisher Scientific

UC Los Angeles

### Scientific Background and Context

Cardiovascular disease is the leading cause of death for both men and women in California. Major Adverse Cardiac Events (MACEs), like myocardial infarction, stroke, and heart failure, result from complex biological and physiological factors as well as demographic and social determinants.

Many people experience a heart attack, stroke, or other complication of cardiovascular disease for reasons that are within the realm of being prevented: they were undertreated,

not taking their medicines, or not receiving the care they needed in the first place. Studies have shown that this occurs more often with younger women and racial/ethnic minorities. According to the researchers, “One reason for this is that early signs of disease can be easily missed, and also because people spend most of their life far away from a doctor or hospital where it is challenging to monitor disease progression.”

### *Ischemic heart disease*

While several lifestyle risk factors are solidly associated with ischemic heart disease (IHD), such as smoking, physical inactivity, poor nutrition, and obesity, interventions designed to induce positive behavior changes are often ineffective due to poor adherence and behavior substitutions. According to the researchers, “For individuals already diagnosed with IHD, secondary prevention focuses on controlling blood pressure, cholesterol, and often diabetes, which is managed within the confines of a health care provider’s clinical setting. Furthermore, algorithmic tools to predict cardiovascular events, such as the atherosclerosis cardiovascular disease risk calculator that considers age, sex, race, blood cholesterol, blood pressure, diabetes, and smoking status, are often only performed in hospital or clinic visits. Identifying early markers predictive for those individuals that are at an increased risk for developing cardiovascular events between visits to capture day-to-day or month-to-month changes in symptomology, biochemical biomarkers, or psychosocial behavior may potentially prevent catastrophic situations, improve access to care and reduce resource utilization.”

### *Remote monitoring*

Advances in digital medicine are facilitating new opportunities to perform continuous remote patient monitoring of patients’ health using experience sampling methodologies such as patient-reported outcomes (PROs) and direct measures of physical activity, such as heart rate and step counts. With continuous insight of participants in real-time and real-world situations, clinicians could potentially evaluate a patient and possibly intervene without needing to wait until the next in-office appointment.

Patient-reported informatics (PRI) is the umbrella term that includes step counts, active minutes, sleep parameters, heart rate, measures of physiologic stress, and other biometrics. While PRI data are derived from wearable devices, PRO data are generated from mobile health platforms and reflect self-reported levels of perceived stress, anxiety, depression, and health-related quality of life each week.

Biomarkers that circulate in the blood also serve an important role in predicting risk of cardiovascular events and should be considered alongside remote monitoring of PROs and PRIs. U.S. Food and Drug Administration (FDA)-approved micro-sampling devices are available and allow an individual to collect small volumes of blood remotely away from health care facilities. This sampling technique provides biochemical information on a semi-continuous manner. Once patients mail their samples to the hospital, researchers can use a laboratory technique called mass spectrometry to measure over 500 blood proteins representing a broad systemic response, such as inflammation, vascular reactivity, organ function, and fat content.



## Project Summary

To understand whether cardiovascular threats can be detected early enough for effective treatment or prevention, the research team pursued the question of whether physiological, biochemical, and psychosocial measurements could predict Major Adverse Cardiac Events (MACEs). They did this by recruiting 200 patients (aged 54-76 years) diagnosed with stable ischemic heart disease (IHD) and remotely monitoring them beyond the walls of the hospital using wearable biosensors for 12 months. Patients wore a specialized watch that measured activity, sleep, heart rate, and stress levels. Additionally, patients reported their levels of anxiety, depression, and quality of life using a smartphone or computer. To supplement the passive monitoring, patients periodically sent a small finger prick blood sample by mail, allowing doctors to measure over 500 different blood chemicals. By combining these different types of data, the researchers sought out a “signal in the noise” to better predict who may be about to have a heart attack or stroke. The team also measured how this approach could be covered by insurance companies and hospital payers.

## Project Achievements

As of the writing of this report, the research team continues to analyze the data collected during the 2-year project period. They provided preliminary results to validate the feasibility of their approach, optimize their coordination of partners and data collection, test their analytical methodology, and demonstrate early signs of how remote monitoring can represent complex health status.

### *Protocol adherence*

This study was dependent on patients’ regular cooperation in using the biosensors, supplying subjective PRO responses, and collecting and returning blood samples. Whereas other research programs have expressed difficulty, this study observed adherence rates up to 90% with the wearable sensor. The researchers credited their use of few devices for the higher than expected rate of adherence, which simplifies the task of remote monitoring for both researchers and patients. When a single device is capable of simultaneously recording multiple variables, researchers can be more confident in interpreting the patient’s state through the data, whether he/she is active, sedentary, asleep, or not currently using the device.

According to the researchers, “Over the course of the study, average adherence dropped from an initial 87.7% to 72% on the final day. The reduction in adherence is consistent with the so-called ‘Law of Attrition’ in similar eHealth studies, as patients will generally have high interest at the point of recruitment, but may gradually lose interest in, start to forget about, or become burdened by the study. Because the Fitbit provides data access in real time, gaps of adherence can be detected quickly, and reminders could be sent as a result, possibly improving adherence.”

Study participants demonstrated high adherence with home biomarker sampling using the blood collection device, which was an important finding to guide future remote precision monitoring studies using this or similar devices for chronic diseases. Greatest adherence was achieved when patients collected their samples from home and shipped the samples to the laboratory.

According to the researchers, “Despite some drop-off in usage with the sensor, Mitra© (home blood collection) device, and PRO app, the high overall adherence with the protocol was encouraging, and likely also related to close contact with the study coordinators who answered questions and provided coherent explanations about the role of using the wearable devices. This suggests that future remote monitoring programs will benefit from the ‘human touch,’ and that just offering technology without careful explanations and outside the context of a therapeutic relationship might not yield the same high levels of engagement observed in this study.” This component of the study was published in a peer-reviewed journal, listed in Appendix B.

### *Exploring the accuracy of remote sensors to represent clinical outcomes*

To better assess the validity of remote biosensors as surrogates for clinically important outcomes, the research team sought to evaluate correlations between the data from the wearable device (PRI) and how the patients were feeling about their health (PRO). The results were mixed and limited, in part, by the low number of participants, considering the complexity of the study design. Consistent with other studies, the preliminary findings supported links between physical activity (as measured by steps per day) and overall health, physical function, fatigue, and standard measures of cardiac health. Further analysis implied an association between physical activity and psychiatric outcomes, in that patients who sleep more and exercise less are at greater risk for developing symptoms of anxiety or depression.

During the study, one patient benefitted from targeted and fast treatment due to his access to monitoring devices at home. After noting chest pain, the 64-year-old man initiated remote monitoring that confirmed a cardiac abnormality and recommended visiting the Emergency Department. The device helped in detecting the problem and facilitating appropriate medical treatment.

### *Confirming feasibility of remote blood collection*

Compliance for home blood collection and submission for laboratory analysis was over 80%, similar to levels observed when conducted in hospital settings. According to the researchers, “This is among the first studies to demonstrate the feasibility and acceptability of remote home biomarker monitoring in the context of a precision medicine trial.” The research team drafted recommendations for using the Mitra® blood collection device and will soon make them available through a publication. Other researchers may apply this method to any chronic disease where biomarker measurements have clinical relevance and predictive value.

Patients were not only adhering to monthly blood collection and submission schedules, they were also collecting the blood using the right technique. Upon comparing the quality of blood samples that were collected remotely with those collected in-clinic, the researchers found that the quality of the samples were as good if not better when collected at home. On average, samples collected remotely were roughly 1.5 times more likely to be classified as good quality than samples collected by in-clinic staff. Other factors that predicted whether a blood sample was of good quality included lower patient-reported anxiety scores. According to the researchers, “the collection of blood samples at home by patients via microsampling device shows high potential for shifting blood sample collection opportunities to low-cost, patient-driven settings.”

### *Linking monitoring data with cardiac health*

Analyses to-date have addressed baseline links between biosensor data and risk of MACEs, while longitudinal time-series analyses are ongoing. Although most of the remote sensor measurements did not correlate with known cardiac health markers, some notable relationships did surface. Generic quality-of-life PROs and certain blood chemicals served as profiles that matched known biochemical markers of cardiac health, suggesting that future protocols could predict biochemical changes based on PROs or remote blood sampling.

Likewise, clear and strong associations were difficult to obtain between biosensor data and known biomarkers of cardiac health, but some did arise. While the researchers did not detect a relationship between sleep or heart rate with cardiac biomarkers, they did find value in step counts, a surrogate for physical activity. The more steps recorded per day, the lower the level of MACE biomarkers.

### *Cost-effectiveness model*

A cost-effectiveness model was developed to evaluate the potential health economic benefits or costs of remote patient monitoring in patients with stable IHD. The main outcome was cost per quality-adjusted life year, a unit often used in economic studies to assess value for money. In this case, the value gained through health monitoring and medical intervention is quality of life.

Using decision analysis software (TreeAge Pro 2017, TreeAge Software, Inc., Williamstown, MA), the research team employed a 10,000-patient simulation model to compare standard care versus three hypothetical intervention strategies applied selectively to “high-risk” patients identified via remote monitoring. The simulation followed the cohort of patients designed to be identical to the actual study over a 6-year period.

The personnel and materials costs were based on the current study. Hourly wages ranged from \$75 for nurses to \$10 for interns. Staff spent about 30 minutes with each participant to train them on the wearable biosensor devices and the home-based devices. The three devices utilized in the study amounted to \$256, with an average annual cost of services and replacement fees of approximately \$80. In total, the costs per patient for initiation and monitoring in the first year amounted to \$596. Afterward, the costs dropped to \$231 per patient for years 2-6, assuming no MACEs occurred.

According to both the researchers and the evaluator, the findings of the economic impact analysis component of the study were less important than the framework of the approach in determining whether remote monitoring could be cost effective. The models created by the research team will be included in a publication and may be used by others in future explorations of instituting remote monitoring programs for stable IHD and other chronic diseases.

### *Leverage of CIAPM funds*

The research team was successful in establishing partnerships with industry, which supplemented state funds with in materials, technical support, and other services. In addition, the host institution, Cedars-Sinai, contributed funds and waived indirect costs. In

total, the research team leveraged their award of \$1,423,261 to receive an additional \$1,062,351 in materials and services and \$446,000 of in-kind institutional support.

### Expert Evaluation by Dr. John Rumsfeld, MD, PhD

According to the evaluator, “This is an important project, in particular with regard to representing an *approach* to evaluation of digital health, biomarkers and patient-reported outcomes; and providing a *framework* for cost-effectiveness evaluations. The study is very preliminary in nature: largely a feasibility assessment and initial results that are ‘hypothesis-generating’ for future research. The investigators are to be congratulated on progress and preliminary findings thus far.” The evaluator also noted some limitations.

Specifically, the evaluator identified the following strengths:

- “This project represents high quality work; and the study team has achieved their stated goals to date.
- Very strong industry partnering; this type of collaboration is a model for research in the current era.
- Most digital health studies in the existing literature focus on single technologies; this study demonstrates feasibility for multiple-technology evaluations.
- Although a small cohort study, feasibility testing was accomplished related to a number of aspects including: digital health technologies/metrics, patient reported outcomes (PROs), and home testing (blood sampling).
- The pragmatic design/ approach, which is a framework for future digital health/biomarker research.
- The high adherence to the Mitra samples and PROs over the 12-week period, and overall ~90% adherence to the wearable sensors are a great reflection of patient engagement. This is important as digital health/precision medicine initiatives move forward, providing evidence that patient engagement with remote monitoring interventions can be high and sustained.
- This is a well written and useful framework for the approach to Cost Effectiveness analyses for these types of studies and technologies. With caveats noted below under limitations, agree this framework is a real contribution of this project to date, as it can be applied in future research.
- This study was set up well for hypothesis-generation to inform future research.
- Excellent patient engagement, which has been a challenge in previously published digital health studies.
- The high quality of the blood samples taken at home is a very important finding of this study, one worthy of highlighting.
- The PRO capture [rate was] very encouraging; especially with growing interest in capturing PROs as part of research outcomes as well as in clinical care under value-based payment models, etc.
- Excellent commitment to – and early success with – publications.”

The analysis of blood sample feasibility and adherence is a novel finding in the field. Future steps include assessing the value of such biomarker sampling, in the context of clinical management, clinical efficiency, and improved health outcomes. According to the evaluator,

“None of these are certain (much less cost effectiveness), and will require evidence in additional research.”

The project’s limitations include the small cohort of 200 patients and observational study design, as one may only draw preliminary relationships, rather than causal implications, between factors and conditions presented in this study. The evaluator also wanted to underscore that the cost-effectiveness exercise was hypothetical and not reflective of actual results as of yet. According to the evaluator, although the results of the economic analysis cannot be interpreted with confidence, “This does not detract from the framework/approach to the Cost Effectiveness analysis...”

To provide context for the whole study, the evaluator wanted to highlight the following statement by the researchers: “In many respects, we believe these pragmatic, procedural findings are as significant and generalizable, if not more so, than the individual quantitative relationships we found among datasets in this hypothesis-generating study.” In other words, the contribution of a robust approach to better understand remote monitoring and biosensor surrogates, supplemented with confirmations of real-world feasibility, is greater than the preliminary correlations between PRIs, PROs, and health outcomes.

### **Assessment of How the Project Addressed Programmatic Goals**

The project successfully addressed multiple goals of the CIAPM program, including 1) demonstrating the promise of precision medicine, 2) use of existing data, 3) efficient and effective data management, 4) developing capabilities for precision medicine, 5) engaging patients, 6) impact for patients, 7) progress toward reducing health disparities, 8) conducting an economic impact analysis, 9) developing clinical and commercial potential, 10) potential for downstream use, including public availability, 11) potential to scale and leverage multiple electronic health records systems and other platforms, and 12) establishing and utilizing partnerships.

#### ***Demonstration of the promise of precision medicine in a specific disease area, health issue, technology, or fundamental biological process***

The study evaluated several multidisciplinary models for conducting research into value-based clinical prediction medicine. While the focus of this research was ischemic heart disease, the researchers developed a framework for applying precision medicine techniques to any chronic, high prevalence, high resource-utilizing disease. A wide variety of technologies was tested, including wearable activity monitors, remote home-based monitors, smartphone applications for remote PRO assessment, and home-based biomarker analysis.

#### ***Use of existing patient data and other data sources***

Data from electronic health records (EHRs) were used to populate the analytic models. The digital infrastructure that supports EHRs, resulting from the U.S. HITECH Act, is serving as a backbone of modern American health care. The researchers were able to use the available data repository for basic lab values, lipid panels, and sociodemographic data to complement the new remote monitoring data obtained in this study.

### *Efficient, effective data integration and analysis*

Cedars-Sinai provided a biomedical infrastructure that facilitated efficient and effective data integration and analysis. The breadth and depth of analyses presented in this study are expansive for the award amount and duration and is a reflection of the research team's existing digital health monitoring, bioinformatics, and genomics platforms.

### *Development of precision medicine capabilities*

The research team achieved great strides in evaluating the feasibility and patient acceptability of previously developed precision medicine capabilities. They demonstrated that using a low-cost, consumer-facing activity monitor is capable of providing clinically relevant data. This is a novel contribution to the digital health literature, as it validates remote monitoring as a tool for clinical care of patients with chronic disease, such as patients with IHD.

In addition, the study demonstrated that compliance with remote biosensor monitoring can reach levels as high as 90% when patients have a clear understanding of what they are being asked to do, why they are asked to do it, and feel supported by the team overseeing the remote monitoring program.

Finally, the research team showed that home blood sampling is feasible and acceptable, and the blood samples collected at home are as good if not better in quality compared to samples collected in the clinic. This finding supports future studies aiming to expand research and clinical care beyond the walls of traditional care environments and into the community.

### *Development and implementation of patient engagement strategies*

Through this research, the team sought to better understand patients' perspectives to improve their quality of care. PROs were a vital data component of this study and reflective of the research team's orientation to engage patients in clinical precision medicine research. In addition, detailed protocols were developed for how to engage patients around a precise remote monitoring study (a publication on this topic is forthcoming).

### *Impact for patients within the project timeframe and beyond*

Two hundred adults (54-76 years old, 60% female, 29% non-white) diagnosed with stable IHD were engaged in this project. Each was supported by the research team with home monitoring devices, information, and clinical care. As referenced earlier in this report, remote monitoring with the Kardia home ECG device allowed one study participant to self-identify a cardiac abnormality. He subsequently reported to the Emergency Department, where the condition was confirmed and treated. This is an immediate positive result of this study. This event was drafted in a detailed case report and was accepted for publication in a peer-reviewed journal.<sup>5</sup> Notably, five other study participants used the AliveCor KardiaMobile measurements to direct their treatment plan. One subject scheduled a follow-up appointment with his Cardiologist as a result of the home measurement and was medically

---

<sup>5</sup> <https://www.journalmc.org/index.php/jmc/article/view/3151/2539>

treated as an outpatient. The four the other subjects had previously been diagnosed and used the device to monitor any changes.

*Reduction of health disparities within the project timeframe and beyond*

Although this study was not oriented toward this goal in particular, future studies may use this framework to enhance the care of patients with IHD, which is a condition with known racial, ethnic, and sex disparities.

*Creation of an economic impact / value analysis*

A primary aim of this study was to formally assess the cost-effectiveness, cost-utility, and projected budget impact of remote patient monitoring for IHD. The results of this study are preliminary and based on theoretical conditions that were informed by evidence from 26 studies and the current trial, such as the material costs of monitors, personnel wages, and data aggregation services. The research team intends to include their economic models in a future publication, which can potentially be applied to any chronic disease.

*Development/realization of the project's clinical and commercial potential*

The team partnered with seven California-based companies, as described previously. According to the evaluator, this was one of the project's greatest strengths as a model for other researchers in precision medicine.

*Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications*

The research team published two scientific articles and has two additional manuscripts in preparation (listed in Appendix B). Twenty-five presentations have been given about this project to public, scientific, and clinical audiences across the state, nation, and eight nations. The research team has also provided further validation of certain devices for precision medicine applications, including Fitbit Charge 2, AliveCor Kardia, and the Mitra® device.

*Potential to scale and leverage multiple electronic health records systems and other platforms*

This project tested several digital health platforms that align with the EHR. These include wearable biosensors and smartphone apps that have the potential to convey their data via HIPAA-compliant channels to clinical end-users through the EHR.

*Other contributions related to advancing precision medicine*

According to the researchers, "it is vital to expand the notion of precision medicine to include value-based clinical precision medicine. In many instances, precision medicine is considered to be largely a basic science, but this study demonstrates that it can be an applied clinical and health delivery science as well. In addition, the notion of precision resides not only on the 'front' side of care, where decisions are being made about how best to deliver care, but also on the 'back' side of care, where outcomes are measured. In other words, in order to know if a precise therapy is working, we also need precise outcome measures to evaluate the clinical benefit of the therapy. In this study, we have evaluated NIH PROMIS scores, for example, which are highly precise measures of patient perception of their care. We have also helped to validate the role of activity monitors like Fitbit as an important tool that captures patient outcomes in a clinically relevant manner. Finally, we have incorporated formal health economics models to help current and future policymakers

determine whether the results of this study might ever be applicable to clinical practice within the framework of budgetary and resource constraints.”

## Conclusion

By all accounts, this was an ambitious project from the beginning. It aimed to establish a sustainable remote biosensor monitoring program, validate a home-based blood sampling kit, understand the connections between biosensor data, blood-based biomarkers, and patient-reported health outcomes, develop an algorithm to predict which patients are at highest risk of adverse cardiac events, and conduct an economic impact analysis of remote monitoring of chronic heart disease.

The team made significant progress and shared its findings widely through multiple publications, press releases, and presentations to both public and professional audiences. The vast stock of data that were collected from continuous wearable biosensor devices, an array of over 500 biomarkers analyzed from blood samples, and smartphone app-based patient-reported health outcomes from 200 study participants will continue to be analyzed for several years.

This study has already contributed useful insights by validating a home blood collection technique in terms of the quality of the sample and the high compliance rate. Future remote monitoring programs would likely benefit from incorporating a similar “human touch” element with the technology, such as by providing careful explanations within a therapeutic context and access to support staff. The research team additionally demonstrated how a patient may use home-based devices to inform their decision to seek immediate medical care.

The research team is continuing its analyses and focusing in the near future on studying links between health outcomes and biosensor, biomarker, and patient-reported outcome data and their predictive value regarding cardiac events like stroke and heart failure. By providing a sustainable and high quality framework for remote patient monitoring, the research team has made great progress in moving the field of precision medicine closer to standard clinical practice.



## 5. Full Genome Analysis to Guide Precision Medicine

Principal Investigator: Dr. David Martin, Children’s Hospital Oakland Research Institute  
Project Period: March 1, 2017 – December 31, 2018

### *Research Team and Collaborators*

Children’s Hospital Oakland Research Institute (CHORI)

David Martin, MD  
Dario Boffelli, PhD

UC San Francisco (UCSF)

Puy-Yan Kwok, MD, PhD  
Ophir Klein, MD, PhD  
Anne Slavotinek, MD, PhD  
Joseph Shieh, MD, PhD  
Bryce Mendelsohn, MD, PhD  
Renata Gallagher, MD  
Jessica Tenney, MD  
Daniah Belefond, MD  
Hazel Perry, MD

UCSF Benioff Children’s Hospital Oakland

Art D’ Harlinque, MD

UC Berkeley

Steven Brenner, PhD  
Andrew Sharo  
Jingqi Chen, PhD

Human Longevity

Brandon Hunter, MS, MBA

GenomeOne

Illumina

### Scientific Background and Context

#### *Genetic disorders*

Genes are packages of DNA and compose the blueprint for the human body, from its one-cell stage on through the lifespan. A child inherits a set of genes from both parents, which undergo a process of combining at conception to form a whole genome. Occasionally, a gene will be defective, caused by a mutation in the DNA sequence. The mutated form of the gene may either be passed down from a parent, whose gene was also mutated, or take a novel form, such as from a mistake in the process of copying the DNA. Some mutations cause the gene to not work properly, which occasionally leads to a lifelong disability, such as in mobility or learning. Genetic disorders commonly present during infancy and

childhood, when the development of the human body is rapidly laying the foundation for the rest of a lifetime.

The extreme cases of pediatric disorders (referred to as “extreme phenotypes” in the scientific literature) are the subject of this study and are commonly seen by all major children’s hospitals. The medical resources required in diagnosing a child with a rare or extreme disorder are usually extensive and do not always lead to a resolution. Although some cases fall into recognizable syndromes with a known genetic basis, a great many do not.

This project aimed to provide a resolution for dozens of patients and their families while also developing a pipeline that will continue past the period of funding into what the research team considers, “a program of discovery, integrated with other centers, that will eventually resolve the genetic basis of most or all extreme [disorders].” Through these efforts, it will eventually be possible to elucidate how many disorders actually have a genetic basis.

### *Full Genome Analysis*

Methods of diagnosing genetic disorders have followed the trajectory of technological advancement. Most current clinical diagnostic tools that utilize genetic testing are only sensitive to mutations that are common and extensively studied. Access to an individual’s complete genome would enable clinicians to scan for mutations across all genes simultaneously for any disorder or condition that has a known genetic link. Sequencing the full genome of infants with extreme symptoms for whom other diagnostic methods have failed has shown that it can guide clinical decision-making in a high proportion of cases.

An example of this scenario featured a child born with small and immature muscles and other conditions that did not lead clinicians to any specific disorder. Over several years and many failed diagnostic tests, including focused gene sequencing, no diagnosis was achieved. The research team was involved in ultimately sequencing the full genome, which successfully identified a genetic basis for the child’s condition.

Full Genome Analysis (FGA) is a method of genome sequencing that was recently developed by a member of the research team to serve extreme cases like that described above. It differs from other sequencing methods in its comprehensiveness: whereas other methods (including whole exome sequencing and microarray techniques) only sequence minor sections of the genome that are already considered important based on current understanding, FGA takes stock of all genetic material without skipping any sections. The adoption of this method into routine clinical practice would require development of standards and a more complete catalog linking genetic variations with clinical outcomes.

## **Project Summary**

Reading the human genome to identify genetic causes of disorders is an extremely recent ability. Today, several methods are available to sequence and analyze a person’s genetic background, but they address the technological challenge differently. Most genetic tests available today narrowly focus on the most common causes of genetic disorders, like particular mutations. For disorders that are rare or have an unknown cause, a complete view of the genome is necessary.

In response to this need, the research team developed a method called Full Genome Analysis (FGA) that reads the genome in a more comprehensive manner than all other available techniques. The method serves two main purposes: 1) diagnosing clinical cases whose underlying mechanisms remain unidentified using other methods and 2) providing a tool to systematically discover genetic mutations previously unknown to cause disease.

This project brought together doctors and scientists to study whether FGA would be useful in the clinic. The multidisciplinary and multi-institute research team applied FGA to 45 pediatric cases of extreme undiagnosed disorders suspected to be genetic in nature. Most the children were from traditionally underserved backgrounds. The importance of involving communities of color cannot be understated, as current genetic references do not yet reflect California's demographic spread.

After sequencing the genomes of the affected children, their parents, and in some cases, their siblings, the new method was able to identify the likely cause of disorder for 40% of cases (18 total). In response to this success, the research team developed a standardized pipeline for the acquisition and delivery of FGA data for clinical decision-making.

The long-term goal of this work is to contribute genomic data to a catalog of all genetic mutations and variations that can cause human disease. By doing so, it would allow any clinician in the world the opportunity to diagnose even the rarest genetic disorders. Conducted with prevention in mind, FGA would allow clinicians to assess disease risk and potentially take therapeutic action before symptoms become extreme and biological damage has occurred.

## Project Achievements

As of the writing of this report, the research team continues to draft a manuscript that will be submitted for publication. Their results collected to date served 45 cases of extreme pediatric disorders, designed and built an analysis pipeline for clinical utility, and contributed to a growing catalog of genetic mutations associated with disease.

### *Patient recruitment*

The research team completed the planned recruitment of 50 cases for this study, based on considerations from an expert panel. The cases included 100 parents and 61 siblings. The medical records of all recruited cases were deidentified and recorded in a Phenotips database. This particular high-quality database was selected to be able to share and analyze aggregate genome data.

### *Full Genome Analysis*

FGA involves multiple computational steps, including the reading of DNA, comparison to reference DNA, and mapping of potential mutations. The research team completed sequencing of all 161 genomes from 50 cases, exceeding their goal of 150 genomes.

First-pass analysis examined 45 cases and produced a list of candidate mutations per case. The short list was examined by an expert panel of clinical geneticists to determine if it contained obvious candidate mutations that could explain the case's condition. The

remaining five cases were delayed due to the extended absence of a member of the research team. The five remaining cases will be analyzed using independent funds.

#### *In-depth computations and development of an analysis pipeline*

In-depth analysis focused on cases for which the first-pass analysis could not identify a candidate causal mutation. A customized pipeline was built to process samples and rank candidate mutations according to the patient's symptoms. In order to take full advantage of the increased information provided by FGA, the research team developed and largely finalized a pipeline, called "SVint," to focus on types of mutations that elude other genetic tests.

SVint includes several unique features specifically designed for this project, including an annotation tool. The pipeline prioritizes mutations into categories and also includes substantial capacity for data quality assessment.

#### *Clinical application of the pipeline*

The research team applied the custom pipeline to data from 35 families, exceeding their original milestone by 10 families. Preliminary reports have been drafted for all. After initial manual overview, they were able to identify candidate mutations for 12 families. The team continues to perform detailed reviews of existing biological evidence for those candidate mutations in order to provide full reports. They also plan to conduct in-depth analyses of the remaining cases.

#### *Diagnoses*

Combined results of first-pass analysis and in-depth analysis led to the diagnoses of 18 of 45 cases (rate of diagnosis: 40%). According to the researchers, "this result strongly supports the clinical utility of FGA." A complete description of the 18 causal mutations will be made public in the forthcoming publication. Of note, 3 of the 18 causal mutations are "structural variants," meaning, they would have likely been missed by the two alternative genetic sequencing methods (short-read genome resequencing and exome sequencing) that are commonly used in precision genomics.

Diagnoses were made following consultations with the referring clinician, who then used the information to design further options for treatment of each patient's disease. FGA results are not at present considered a clinical test and therefore are not included in the medical record. For this reason, several of the mutations identified by FGA were independently validated using a clinical test.

#### *Leverage of CIAPM funds*

In-kind resources, materials, and time were provided by the research team to supplement the awarded funds. Computational resources in excess of the supported funding level were also provided in kind, including sequencing of eleven additional genomes (provided in kind by Puy-Yan Kwok). Most personnel provided additional time in-kind.

### **Expert Evaluation by Dr. Elaine Mardis, PhD**

In evaluation of the scientific, clinical, and technical quality of the work, the evaluator made the following comments:

- “Overall, the scientific, clinical and technical quality of the work accomplished by this group is outstanding.
- Their full genome analysis (FGA) of nearly all the 50 families accrued is a technical tour de force due to the types of data generated and the integration challenges.
- Their clinical quality reveals around 40% of cases solved by identifying the causal [genetic mutation] for 18 cases, wherein some cases had fairly simple alterations and some that were significantly complex and would have been difficult to solve by sequencing and alignment approaches alone.
- The scientific quality of the work is high, and supports the hypothesis that more comprehensive analysis can give results in cases where straightforward analyses (sequence alignment, not assembly) and focused sequencing (targeted panels or exomes) is not capable of identifying the genetic underpinnings of disease causality.”

The evaluator found that the team accomplished its overall goals and most of its specific aims, providing the following comments:

- “The team accomplished or exceeded the stated accrual goals (50 families)
- The team did not complete the first-round analysis of all families, with five remaining to go through analysis (due to extended leave of one team member)
- The team analyzed 10 cases in excess of the number planned using in-depth analysis processes, demonstrating that their FGA approach was capable of solving complex structural variants that underlie disease causality according to the genes involved.” In assessing whether the overall outcome represents a significant achievement for the project, the evaluator submitted the following comments:
- “The overall outcome was a strong achievement for the project due to the scale and complexity of the proposed tasks.
- The overall outcome recapitulates the percentages reported by others in this type of study who are using whole genome sequencing methods, but not using FGA.
- The fact that the nature of some variant regions hasn’t been fully determined indicates there may be structural alterations to the genome that are still too complicated to be resolved (perhaps requiring long read sequencing, for example, which wasn’t used here).”

The evaluator noted additional achievements in addressing the goals of the program:

- ”This project has addressed the promise of technology-based approaches to solving difficult genetic diagnoses in children who otherwise would not have answers to the genetic basis of their disabilities. While this approach will rarely lead to treatment indications, it does provide answers to parents about why their child has difficulties, and what their chances are of having a similarly affected child.
- By using these advanced diagnostic approaches, time and money can be saved by insurance payers (including the state), in determining a diagnosis more quickly and with far fewer diagnostic assays required.
- Patient data is critical to the process they are following, including physician’s descriptions of the patient phenotype, any diagnostic tests that have already been run, family history information, and other information that should be in the EHR.

- Other data sources would be used as well, such as Online Mendelian Inheritance in Man (OMIM), Monarch Initiative, Matchmaker Exchange, ClinVar and so on.
- The integration of data, such as listed above, is critical to the process. [The team] integrated data from Bionano (map) and Illumina (sequence) to produce their genome assemblies.
- The project has been very successful in regard to developing infrastructure and expert teams to address the moving parts required for this important aspect of precision medicine.
- If expanded, the system that has been developed by the team would enable precision medicine-based approaches in other pediatric genetic institutions.
- The project supported a very balanced application to ethnic minority patients and their families.
- The project overall has potential impact for precision medicine, as it adds to the data supporting whole genome sequencing (and its variations) as impactful in a larger percentage of cases than is exome or targeted panel sequencing.
- The project overall could have an impact on precision medicine in the State of California by supporting reimbursement in cases that could benefit from this approach to diagnose the genetic disease causality. It could also position the State of California, and CIAPM in particular, as being thought-leaders in this area of pediatric precision medicine.
- It seems there are future plans to grow the partnership between UCSF and Benioff hospitals. Such partnerships across the state would be highly beneficial to pediatric genetics patients state-wide.”

Among the evaluator’s considerations of this research team’s future plans, she stated, “The future goals are a natural extension of the work already completed. It seems there is a needed piece missing, which is commitment from the hospital systems involved in the effort, so that the entirety of the service line in pediatric genetics for both hospitals is invested and active.”

### Assessment of How the Project Addressed Programmatic Goals

The project successfully addressed multiple goals of the CIAPM program, including 1) demonstrating the promise of precision medicine, 2) use of existing data, 3) efficient and effective data management, 4) developing capabilities for precision medicine, 5) engaging patients, 6) impact for patients, 7) progress toward reducing health disparities, 8) developing clinical and commercial potential, 9) realization of clinical potential, 10) potential for downstream use, including public availability, 11) potential to scale and leverage multiple electronic health records systems and other platforms, and 12) establishing and utilizing partnerships.

#### *Demonstration of the promise of precision medicine in a specific disease area, health issue, technology, or fundamental biological process*

According to the researchers, “The project was designed to address a common problem in pediatric medicine: a child with a severe and often complex set of symptoms that suggests a genetic disorder, but has no characteristics that clearly point to a known disorder. Such children often are the subjects of lengthy and expensive diagnostic journeys that make use

of any available clinical clues and technology; sometimes there is a positive outcome, but often the patients and their families are left with no clear understanding of the cause and prognosis of the disorder. There are many cases like these at our hospitals; the case selection committee that we established for this study, which included all of the clinical geneticists at UCSF Benioff Hospitals, considered many additional cases that could have been included in the study but had to be left out because more extreme cases were available. All of the cases selected for the study already had negative results for potential genes that might have explained their condition; many of the cases had also undergone [other genetic] analysis with negative results. Thus, all the recruited cases had exhausted available means of genetic diagnosis.”

The application of the team’s FGA method was highly successful: to date, they identified a likely causal mutation in 40% of the recruited cases. This result supports the utility of FGA as a test that can resolve difficult-to-diagnose cases while simplifying genetic testing with a single method. According to the researchers, “Most of the cases had been under the care of clinical genetics for several years – FGA returned a result in a few weeks.”

Many of the likely causal mutations are in genes that had not previously been directly implicated in a disease. This project, though limited to 50 cases, has expanded the number of conditions in which a diagnosis may be delivered.

#### *Use of existing patient data and other data sources*

Medical records of all recruited cases were collected and entered into a PhenoTips database, which facilitates secure storage, sharing, and data analysis. De-identified patient information included physicians’ descriptions of the patient’s symptoms, any diagnostic tests that had previously been conducted, family history, and other information that was included in the electronic health record (EHR).

Other data sources were also used in the analysis pipeline, including Online Mendelian Inheritance in Man and ClinVar.

#### *Efficient, effective data integration and analysis*

The FGA method relied on multiple datasets. Data were integrated from Bionano and Illumina to produce genome assemblies.

#### *Development of precision medicine capabilities*

This project has served as a platform to develop a comprehensive infrastructure capable of delivering diagnoses of genetic disorders on a routine basis. To carry out FGA, the research team assembled multiple teams with specific expertise:

- clinical geneticists evaluated cases for admission to the study and made the final call on the genetic diagnosis
- genetic counselors carried out case recruitment and sample collection
- sequencing personnel carried out the laboratory components of FGA
- bioinformatic personnel carried out the genome assembly, variant identification and variant prioritization

According to the researchers, “The project revealed the need for more effective bioinformatic tools for the analysis and reporting of FGA data; it also showed that close collaboration of computational biologists and clinical geneticists is required for the development of effective tools. For this project, we have developed a tool that integrates structural variants in the analysis of coding variants, and we have begun work to improve the prioritization of variants – the goal is to provide a manageable shortlist of potential causal variants to clinical geneticists for their evaluation. So far, we have focused on the simplest modes of inheritance (*de novo* dominant and homozygous recessive); work to expand the analysis to more complex modes of inheritance will be carried on outside this project.”

#### *Development and implementation of patient engagement strategies*

Patients were given written summaries of their sequencing results and copies of the preliminary reports to empower them to use this information elsewhere in their health care or to share with relatives, who may be similarly impacted by their inherited genomic backgrounds.

This project aimed to study a broad variety of cases, especially those that ordinarily would not have been referred to a genetics clinic. The two clinics involved in this study serve as the major pediatric referral centers for patients in Northern California and the Central Valley. Over 1,000 patients are annually assessed in the outpatient genetics clinics, and several hundred are admitted as inpatients, providing ample cases for recruitment.

Cases who have already been genetically tested have relationships with a clinical geneticist. According to the researchers, “Pediatric subspecialists (Genetics, Neurology, Neonatology, Pediatric Intensivists, Hematology, Cardiology, etc.) provide a referral network and support the needs of the community pediatricians and other subspecialists from these regions through satellite clinics, outreach efforts, education, telemedicine, and phone consultation.”

The research team met with key pediatric medical groups and clinics to discuss the study and solicit participants, targeting those serving minority populations with traditionally fewer opportunities to obtain genetic diagnosis services. The research team employed a robust pre- and post-test consent and put a counseling process in place for all genome-scale testing. According to the researchers, “Patients and families met one-on-one with a genetic counselor and/or a physician to discuss the goals of testing, possible outcomes of testing, use of data for research and related privacy matters, the types of unexpected or secondary findings that might transpire, how genetic test data can and cannot be used to determine insurance coverage, and how genomic data are stored. In particular, we discuss specific scenarios when secondary findings might be useful or troubling, and encourage patients to consider their personal thoughts about these possibilities.”

Agreement to testing was sought from patients who were able to provide it. When raw data were requested or generated on-site, additional information was provided, including about its storage and analysis. Patients and families were encouraged to ask questions, in-person interpreters were used when appropriate, and information brochures were available in Spanish and several other languages.



### *Impact for patients within the project timeframe and beyond*

At present, many patients suspected of having a genetic disorder undergo lengthy and inconclusive testing. This occurs because their set of symptoms do not point clearly to a known testable genetic cause, while some patients are not tested because a target is not apparent or because insurance providers deny coverage. FGA is capable of short-circuiting this process by providing a definite diagnosis for many patients after a single test.

For patients who do not receive a diagnosis from FGA, their data still contribute to longer-term efforts to identify the genetic basis of their disorder. Since the FGA test provides more information than all focused genetic tests that such a patient would normally undergo, the researchers expect that the cost of health care for these patients would lower, compared to the present process of an extensive workup and consultation without diagnostic resolution.

All of the cases recruited to this study had exhausted their options for genetic diagnosis using available tests. The cases for which FGA returned a causal mutation were given an explanation for their medical condition. In a few cases, this information also offered therapeutic options. As stated by the researchers, “this information confers no direct treatment benefit for the affected child, but it does provide an answer for the family, as well as the possibility of family screening and prenatal diagnosis. In some cases, the mutation is *de novo* (not present in either parent), which can provide reassurance that other children born to the parents are not at risk. Finally, information gained from this study will help to increase our understanding of the conditions, and may lead to the development of new therapeutic options.”

Since biomedical research will continue to link at-present-unknown genetic mutations with disorders, the genetic data collected via this project will be reanalyzed in the future against any new information that may emerge. For the 60% of study participants whose FGA test did not yield a diagnosis during the study period, future analyses could facilitate the identification of a causal mutation. According to the research team, “This could occur through the information shared in a database of other families affected by the same condition, or by advances in scientific understanding of the condition.”

### *Reduction of health disparities within the project timeframe and beyond*

According to the researchers, “Providing a state-of-the-art genetic test for children and families who might otherwise go without any genetic testing was a main goal of this project. The patient population at UCSF Benioff Children’s Hospital Oakland includes a large proportion of underserved East Bay groups; this is reflected by the ethnic composition of our recruited cases. Approximately one third of the recruited cases had been denied further testing by their insurance; enrollment in this program was their only opportunity for a genetic diagnosis.”

### *Realization of clinical and commercial potential*

Although the FGA test is currently a research procedure, some results have been validated in clinical assays.

*Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications*

The genetic and phenotypic data collected by this study will be deposited, in de-identified form, in scientific databases created to share information from multiple families affected by severe genetic disorders. The goal is to expedite construction of a comprehensive dataset that can be explored with machine learning tools and identify novel genetic variants that are causative of genetic disorders.

*Potential to scale and leverage multiple electronic health records systems and other platforms*

The research team and evaluator envision a future when the EHR may be engaged as a platform to evaluate cases that might benefit from re-analysis or promote cases for FGA analysis based on various data contained in the record.

*The development of partnerships / collaborations in your project*

At the local level, this project set out to form a team of investigators that covered core competencies in clinical genetics, FGA, and computational analysis of clinical genomes. This group was able to develop a comprehensive infrastructure that piloted the genomic testing of 50 children (161 genomes total, with parents and siblings included) and delivered genomic diagnoses to cases that had thus far exhausted available clinical testing options.

*Next steps*

In the future, the research team intends to connect with similar teams across the globe to compile genomic and symptom data to construct a comprehensive, shared database that can be explored with machine learning tools. The catalog of genetic and disease information could be made available to any clinician worldwide, thereby enhancing access to a cutting-edge diagnostic tool. As stated by the researchers, “We believe that this is the pathway to a complete set of causal genetic variants. We have already established collaborations with the Center for Mendelian Genomics at the University of Washington and the Decipher Developmental Disorders project. Our collaboration with Illumina will help us engage other projects with patient populations similar to ours.” Additional partners could include Genome One (Australia) and the University of Cambridge (UK).

Locally, the research team sees an opportunity to build a cohesive effort across UCSF and Benioff Children’s Hospital Oakland by integrating clinical care and basic and translational research. This could include a Pediatric Rare Disease Program, which would pioneer targeted genetic-based diagnostics and treatments for all children, focusing first on children with rare diseases and then expanding to pediatric patients with common illnesses, such as diabetes or asthma.

The research team also plans to collaborate with other UCSF groups who study common forms of diseases to develop new therapeutic approaches based on genetic mechanisms revealed by FGA and other tests. In the future, they could extend models elsewhere and scale partnerships with other institutions.

Additional topics that are relevant to the research team’s work include ethical concerns regarding genome sequencing of embryos (IVF and prenatal) and related issues. These topics could be explored by groups of researchers, members of patient communities and advocacy groups, clinicians, bioethicists, and others.

## Conclusion

Despite advances in genomic technologies, many children with rare or extreme symptoms still pass through the health care system undiagnosed. Until now, genetic tests were limited to particular sections of the human genome and only detected mutations that were already known to cause disease. The research team led by the Children's Hospital Oakland Research Institute has provided a new method that for the first time facilitates screening of the entire genome. By comparing the affected child's genomic sequence to the sequences of their parents, their siblings, and other database entries, researchers and clinicians are gaining a clearer view of which variations in the genetic code may be causing disease. In this way, the new test and computational pipeline contribute to diagnoses while also contributing new causal information for all persons presenting with a similar genetic mutation in the future. By expanding their work with international collaborations over time, the team is aiming to transform medical practice by fostering the development of a comprehensive catalog of genomic data that links every possible genetic variation with associated health impacts.

## 6. Personal Mobile and Contextual Precision Health

Principal Investigator: Nicholas Anderson, UC Davis  
Project Period: January 1, 2017 – December 31, 2018

### *Research Team and Collaborators*

#### UC Davis

Nick Anderson, PhD  
Madhan Dharmar, MD, PhD  
William Riedl, MS  
Michael Lehman  
Christopher Lambertus, MS  
Brad Pollock, MPH, PhD

#### UC San Francisco

Ida Sim, MD, PhD  
Meghana Gadgil, MD, MS  
Jason Satterfield, MD  
Mitch Feldman, MD  
Tim Satterwhite, MD  
Amy Lozano  
Sarah Chatfield  
Vijay Rayanker

#### UC Berkeley

David Lindeman, PhD  
Dan Gillette, MS

#### Overlap Health, San Francisco

David Haddad, MSc

### Scientific Background and Context

Mobile phones have the ability to transform health care because they can collect and deliver patient-generated data to health care providers without patients having to visit a clinic or office. Patients can also receive mobile alerts, such as a reminder to take medication, and record their medication adherence. Integrating this information with an electronic health record could, for example, link a medication change in a psychiatric patient to a change in mood in order to assess the effectiveness of a medication. Despite a promising future, several barriers need to be overcome before data from mobile health applications can be reliably linked with data from a patient's electronic health record where it can be used by care providers to inform clinical care. Mobile health applications also allow patients to track their own progress over time using simple visualizations within the applications, which increases patients' role in the monitoring and management of their own health. Nonetheless, it is not known if generation and possession of this knowledge leads to more favorable clinical outcomes.

Two conditions that lend themselves to mobile health applications are high blood pressure and depression. Patients who have or are at risk for high blood pressure can be given a blood pressure cuff that uses Bluetooth technology to wirelessly transmit blood pressure readings to an application on a mobile phone. For patients with or at risk for depression, questionnaires designed to measure a patient's current mood can be filled out using a mobile application. These mobile measurements can be taken much more frequently than a patient's weekly or monthly in-person appointment with a clinical care provider, which give the practitioner a much more granular view of the patient's day to day status.

Central to the discussion about mobile health applications is consideration for patient privacy and informed consent. According to the U.S. Department of Health and Human Services, the federal Health Insurance Portability and Accountability Act (HIPAA) establishes privacy standards for medical records and personal health information, and sets limits and conditions on the uses and disclosures of the information without patient authorization. Any mobile health application that collects, stores, or transmits specified health information is subject to HIPAA.

## Project Summary

This project aimed to develop and evaluate a HIPAA-compliant mobile health application designed for sending alerts and collecting, monitoring, and reporting patient-generated data about blood pressure, mood, movement (steps), and medication adherence. These data were to be integrated with data from a patient's electronic health record on a third-party platform, where they could be accessed by the patient's clinical care team. The researchers planned to include clinicians and volunteer participants in the initial design of the mobile applications, which were to be tested on 200 participants diagnosed with and medicated for high blood pressure or depression.

The project also planned to study how participants interacted with the health application by measuring the frequency of application use, average time spent interacting with the application and the visualization features.

## Project Achievements

Over the course of the project, the research team 1) established policy, privacy and legal frameworks for linking data providers, and clinical and personal data, 2) developed a digital platform to record and report patient blood pressure, mood, physical steps, and medication adherence, and 3) evaluated application engagement and care plan adherence, both quantitatively and qualitatively.

### *Linkage framework*

Because study participants were recruited from two separate health systems (UC Davis and UC San Francisco), a framework needed to be developed that could extract and aggregate data from the two different data environments. The research team contracted with Overlap Health, which provided a digital platform where information gleaned from either the UC Davis or UC San Francisco patient databases (using a Fast Healthcare Interoperability Resource, or FHIR) could be securely coupled with information that patients generate using their mobile devices, all in a cloud-hosted, HIPAA-compliant environment. Working with

Overlap Health, researchers were able to have their mobile health application run on both Android and Apple iOS systems.

#### *Application platform development*

The research team engaged a range of stakeholder groups throughout the design and implementation of the mobile health application. Researchers had to be sure that the application met privacy standards, accurately assessed clinical measurements (including the linkage with the wireless blood pressure cuff), and featured a user interface that was intuitive, engaging, and informative. The back end of the application was developed with the ability to monitor both individual- and cohort-level statistics, such as number of participants that engaged with the application on a given day, and the average session duration.

#### *Evaluating access and outcomes*

162 research participants that had been diagnosed with either high blood pressure (108) or depression (54), and had been prescribed a medication regimen, were recruited and screened for the study and underwent standard informed consent protocols. During the onboarding process, patients received assistance downloading the developed medical health application onto their personal Android or Apple iOS mobile phone. Over the six-month study period, patients were sent alerts via their phones, reminding them to measure their blood pressure, assess their current mood, and report whether or not they had taken their medication.

At the end of the six-month period, research participants went through a debriefing process, which included one-on-one interviews (27) and focus groups (5).

The clinical outcomes of the study are in the process of being analyzed. Additionally, in coordination with the UC Berkeley Data Science group, the research team is working to make this unique, full set of acquired data (including mobile, clinical, and usage data) available for training and validation purposes. These data will be used to develop analytic methods and design future interventional mHealth studies, through the use of three separate training modules on data quality, descriptive data structures, and statistical methods.

### **Expert Evaluation by Dr. Bambang Parmanto, PhD**

According to the expert evaluator, this project aimed to integrate and link data generated from patients in their daily lives with clinical data from health care providers to develop personalized models of high blood pressure and depression. “The Investigators correctly identified an area that has strong potential impacts and is highly significant for achieving precision health.” Previous research around the use of mobile devices in health care has mainly focused on data collection that is separate from a health system’s official electronic health record for a patient. Developing technologies that integrate mobile health data with clinical data is timely, and “it is commendable that CIAPM provided funding for this type of research.”

The evaluator identified some of the logistical challenges of the project, including that the project

- Takes place across multiple institutions.
- Involves existing electronic health record systems at UCSF and UC Davis, which requires institutional approval and coordination with the IT departments of the health systems.
- Involves a large number of participants, who also have to be using the existing electronic health record system.

## Assessment of How the Project Addressed Programmatic Goals

According to the evaluator, the project achieved two primary goals, as stated in the proposal: (1) development of a secure linkage between protected clinical data and patient-generated data from a mobile health application and (2) development of models for personalized digital monitoring and analysis for high blood pressure and depression.

The first goal was achieved by the development and implementation of secure linkage technology between the mobile patient data environment and the enterprise clinical data environment. The evaluator commented that, “The technology developed in this project can be categorized as ‘proof of product’, which is more advanced than ‘proof of concept’.” This contribution is significant given the limited time and the logistical challenges of the project. Moreover, the use of two enterprise clinical environments (UC Davis and UCSF) is significant as it shows that the system is scalable to other sites and other enterprise environments.

The second goal was achieved with the pilot study on the use of digital signatures for medication adherence in 162 participants. The evaluator commented that the second goal is also a significant achievement considering the eighteen-month study period.

### *Assessment of the impact on precision medicine*

The evaluator made the point that there is currently no business model for the integration of patient-generated data and clinical data, and referenced the researchers’ conclusion that institutional commitment from health systems would be necessary for such an integration to achieve its goals. The evaluator felt that the direct integration of all patient-generated data into the electronic health record might not be the solution because there is not a clear financial or clinical benefit for the health systems. The evaluator suggested that one potential business model would be to have a bi-directional flow of data between the electronic health record system and a disease/condition-specific database portal. In this scenario, the electronic health record could contain a summary of the patient-generated data, and the mobile health database would have the detailed information necessary to inform a clinical intervention. This bi-directional model may make the use of mobile health database more attractive to health systems.

### *Assessment of issues and obstacles*

The evaluator acknowledged that linking health system electronic records with patient-generated data is relatively novel for those who manage information technology for health systems. Further, the evaluator commented that it was not surprising that it took longer than expected (four months vs. one month) for the health systems to provide institutional approval.

### *Overall assessment and future goals*

According to the evaluator, the project achieved its main goal of “developing infrastructure for linking patient-generated data with enterprise electronic health record systems from two health system providers”. Additionally, the research team was able to achieve a “proof of product” stage in their technology development, allowing the researchers to successfully conduct a pilot study using the integrated system for tracking adherence to medications for newly diagnosed patients with hypertension and depression.

However, two milestones were not completed: convening a third advisory committee meeting, and engaging commercial partners and other stakeholders in further technology development. The former milestone was not achieved due to logistical issues, and is not hugely significant. That the latter milestone was not met was mostly due to difficulties involving the transfer of data between non-compatible systems, which required highly customized solutions. Digital systems that are more independent of institutional infrastructure will allow for adoption by a greater number of commercial or community partners.

### **Conclusion**

Over the course of this project, the research team designed, developed, implemented, and evaluated a web hosted, clinically linked, HIPAA-compliant system that links two major clinical systems with a private mobile health industry partner. The research team was able to produce and refine architecture, recruitment and engagement models, and new policy frameworks that may inform the expansion of capabilities to link clinical health information to patient’s personal health, even if the patient is geographically removed from the health care providers. To develop this patient-centric framework, the project used clinical electronic health records to simultaneously screen and recruit newly diagnosed and medicated high blood pressure and depression patients at two academic medical systems. All enrolled patients were provided with a mobile health application on their personally owned Android or iOS device, and clinically monitored for medication compliance, physical activity, and either blood pressure or depression symptoms through both unobtrusive as well as patient-provided data over a six-month period.

Now that major legal and system-compatibility issues have been resolved, linking electronic health record data with health applications that monitor other classes of patient-generated data should be relatively simple. Further work will need to be done to ascertain whether or not engagement with mobile health applications improved patient medication adherence or other clinical outcomes.



## 7. Precision Medicine for Early Prostate Cancer: Integrating Biological and Patient Complexity Variables to Predict Treatment Response

Principal Investigator: Sheldon Greenfield, MD, UC Irvine  
Project Period: January 1, 2017 – December 31, 2018

### *Research Team and Collaborators*

#### UC Irvine

Sheldon Greenfield, MD  
Sherrie Kaplan, PhD, MPH  
Edward Uchio, MD, FACS, CPI  
Hal Stern, PhD, MS  
Maozhu Dai, PhD Candidate  
Daniel Nguyen  
Sergio Gago, PhD

#### Cedars-Sinai Medical Center

Timothy Daskivich, MD, MSHPM

#### UC Los Angeles Medical Center

Mark Litwin, MD, MPH, FACS

#### Veterans Affairs Long Beach Healthcare

Greg Gin, MD, MS

#### Veterans Affairs Los Angeles

Isla Garraway, MD, PhD

#### Vanderbilt University

David Penson, MD, MPH, MMHC  
Daniel Barocas, MD, MPH, FACS

#### GenomeDx Biosciences

#### Ambry Genetics Corporation

## Scientific Background and Context

### *Prostate cancer*

Each year, 200,000 new cases of prostate cancer are diagnosed nationally, making it the most common cancer in men. The prostate is a small gland in the reproductive system that supports the health of sperm, and over a lifetime, 1 in 7 men will be affected by tumor growth. Like most cancers, its severity depends on the subtype and how soon a tumor is detected. While some types of prostate cancer grow slowly and do not pose a serious threat, other types can spread quickly beyond the prostate gland.

Following a biopsy of the cancerous tissue, clinicians use the Gleason score to evaluate the aggressiveness of prostate cancer cells, graded between 2 (nonaggressive) to 10 (very aggressive). Clinicians can also use a blood test to identify whether a prostate-linked

biomarker is present at higher concentrations than would be healthy. In addition, genomic testing and imaging exams are used to assess a patient's risk of a poor prognosis. Combined, this information contributes to determining the stage of the tumor, from I (when the cancer is still confined to the prostate) to IV (when the cancer has spread to other areas of the body). Patients are typically subjected to either radiation therapy, surgery, or if low risk, active surveillance. However, the currently available information is inadequate to predict the future for an individual.

### *Outcome prediction models*

Patients do not respond equally to treatments, even when cancerous tumors are similarly graded. An important goal toward improving outcomes is to be able to predict which treatment plan would be most effective for each individual. To many, this type of approach is one of the primary tenets of the field of precision medicine. Equipped with that knowledge, a patient may avoid investing time and resources toward ineffective treatments and recovery sooner.

To develop an outcome prediction model (also called "risk prediction model"), scientists and clinicians compile data that link patients' unique conditions with specific health outcomes. Differential access to care, chronic stress, greater physiologic burden from lower socioeconomic status, and environmental factors associated with a cancer diagnosis and disease progression have all been implicated as important contributors to disparities in prostate cancer care. Many researchers now suggest that a complex interplay between race/ethnicity, as well as other patient characteristics, including sociodemographics, health status, health habits, life context (including exposure to stressors and presence of social support), environmental characteristics, and certain biologic processes, such as inflammation, collectively place subgroups of the population at differential risk for suboptimal outcomes of prostate cancer. A complexity profile that integrates such diverse variables may considerably reduce disparities in prostate cancer care through personalized medicine.

Currently, physicians have no standardized method for determining potential benefit from treatment types, differentiating indolent from more aggressive tumor type, or predicting risk for recurrence/metastasis or early mortality. Treatment decisions must therefore be based on clinical experience, judgment, and patient preferences, and may lead to over treatment of lower risk patients, under treatment of higher risk patients, and subsequently to observed disparities in prostate cancer care. Even existing risk prediction models, such as the Cancer of the Prostate Risk Assessment (CAPRA), include only clinical severity markers, and do not incorporate genomic or other variables potentially contributing to risk for disease progression, as noted above. Individualizing risk profiles for targeted subgroups with distinct potential for benefit from specific treatment types is the first step in personalizing care and reducing disparities in prostate cancer.

Physicians can then use the model to guide future preventative and medical interventions for other patients that share similar characteristics. Traditional models have used data that include the age of the patient, blood-based biomarker level, Gleason score, and tumor stage. The research team recognized that the traditional set of data did not adequately identify subgroups of patients who are more or less at risk for poor outcomes, so they conducted a national study (the Comparative Effectiveness Analysis of Surgery and

Radiation, or CEASAR study) to develop a more granular model that accounts for patients' other health conditions, lifestyle habits, and social circumstances.<sup>6</sup>

The CEASAR Study sought to better understand whether prostate cancer severity and patients' personal conditions may be taken into account when predicting health outcomes after therapeutic radiation or surgery. The researchers concluded that men could be categorized for their relative risk and better navigated through therapy options based on their risk category.

## Project Summary

This project focused on improving clinicians' ability to predict whether certain treatments would be effective for prostate cancer patients from diverse backgrounds. The research team improved a prediction model they developed previously by adding more parameters, including data that further detailed patient characteristics and genetic indicators. To ensure that project outcomes matched the needs of diverse patient communities, the team followed a model of enhanced patient engagement. These activities and others promoted greater racial representation in the database for traditional underserved patient groups.

In addition to the improvements to the prediction model, funding was also applied to coordination efforts that engaged patients, scientists, and clinicians across five sites in Southern California, surveillance and treatment of patients, and the creation of a comprehensive, sustainable, and queryable database that applied the enhanced risk prediction variables.

As written by the research team, the primary aims of the study were to "1) assess the replicability of the risk prediction model developed for the CEASAR study in a more diverse patient sample; 2) assess the impact of adding established genomic classifiers to the risk prediction model for comparing treatment outcomes for patients within California; and 3) to generate [a prostate cancer patient] registry." The team designed their approach to be replicable to other disease areas.

## Project Achievements

The results of the project include the establishment of a collaborative network, creation of a leadership team and Citizen Scientist Committee, design and creation of a federated database, design of a patient survey and tools for data collection, abstraction of medical record information, patient recruitment, data sharing, collection, and quality control, and completion of numerous analyses.

### *Organization of a collaborative network of clinicians, scientists, and industry partners*

The research team was deliberate in how they formed and maintained a collaborative network composed of study clinicians, scientists, and industry partners. In addition to coordinating clinicians and scientists across all five study sites, two genomics companies,

---

<sup>6</sup> <https://www.sciencedirect.com/science/article/abs/pii/S0302283818301192>

GenomeDX and Ambrygen, joined the team and agreed to provide genetic tests free of charge for this study. GenomeDx also provided additional reports on the genomic profiles, called the Genome Resource Information Database reports.

#### *Creation of a leadership team and Citizen Scientist Committee*

The leadership team included researchers, stakeholders, patients, and advisors and served to direct the project. A Citizen Scientist Committee (formerly known as the Patient Stakeholder Advisory Committee) was also created to review the project as a whole and provide feedback and input as well as suggest best practices of how to disseminate information to patients. According to the researchers, “During these meetings, the research team discussed the current status of the project and received the committee’s insights on the measures, study outcomes, as well as their own personal experience within the clinical setting and treatment process.” Based on feedback, the researchers intend to maintain and again engage the Citizen Scientist Committee for future studies.

#### *Designing a federated registry*

The study leadership team finalized the study protocol that consisted of designing a “federated registry.” Such a framework would be a virtual consolidation of multiple repositories, allowing searches to be conducted across different collections. This master database was designed to allow the interoperability of data collected from all five study sites. This allowed the many partners to maintain institute-specific data policies without sacrificing participation in the study.

#### *Generation of key study variables and data collection instruments*

When creating the patient survey, the research team conducted focus groups to help finalize study measures, review the overall study design, and discuss future dissemination activities. The patient survey was modeled after a previous successful project, the CEASAR study.

The focus group consisted of physicians and researchers from the CEASAR study and the current research team. They reviewed the questionnaire and proposed modifications to key study measures. Based on input from the focus groups, the team added new measures (e.g., resilience) to improve their prediction models and to understand the spectra of patient outcomes within treatment groups representing a diverse population.

The final questionnaire included the following:

- Measures for quality of life for patients with prostate cancer
- Measures of health status, addressing physical function and mental health
- Overall health rating scale
- A measure of prostate cancer management confidence
- A measure of worry or concern related to prostate cancer
- A scale for measuring stress and resilience
- Patient demographics

The questionnaire was finalized and piloted to ensure its implementation was feasible. Following the baseline patient survey, the research team conducted 6-month and 12-month follow-up surveys. The 6-month questionnaire was particularly helpful for trajectory analyses, as it allowed the researchers finer data over time. Using these data, patients were

identified and categorized into risk groups. With the help of the UCI IT Department, the questionnaires were made available in an electronic version, in addition to four other modalities, including by phone call, in person, mail-in submissions, and electronic using the REDCap platform. All study sites had access to their own site's data.

#### *Medical record information*

Existing patient data were used from electronic medical records. The research team abstracted information including blood biomarker levels, Gleason scores, tumor stage, types of treatment, complications, genetic testing results, and other clinical variables.

To streamline this process, methods were adopted from the previous CEASAR study. Genomic data were collected using the Decipher test, which is clinically validated and CLIA-certified. The test analyzes 22 genes in biopsied prostate tissue to predict the probability of metastasis (when the tumor spreads to other parts of the body). In addition, the company GenomeDX provided additional reports, termed Genome Resource Information Database (GRID) reports. These analyses report on RNA values, covering approximately 46,000 genes and indicating activity relevant to prognostic and predictive capacity for prostate cancer. Another genetic test, ProstateNEXT (provided by Ambrygen), measures 14 genes that are known tumor markers. Of the panel of genes, BRCA2 has been most strongly linked to overall risk of prostate cancer pathology and recurrence rates. Adding these genetic tests strengthened the ability of risk prediction models to predict prostate cancer patients' outcomes.

#### *Patient recruitment*

The research team exceeded their original target of 600 and enrolled 651 patients across their five study sites. Of the 860 patients approached, only 209 patients declined, a 76% participation rate. Approximately 10 percent were Latino, 16 percent were African American, and 57 percent were White. The two VA patient samples represented traditionally underserved communities, while the majority of patients in the two UC sites and Cedars Sinai were Non-Latino whites.

Consistent with previous studies in the field, the research team found that African American patients were more likely to receive radiation therapy, compared to White and Latino patients.

#### *Data sharing, collection, and quality control*

The research team established a data sharing agreement with all five study sites to initiate the process of creating a master database/registry. At that point, data were consolidated. To ensure quality, the team conducted 20% reliability checks of the incoming data and followed-up with any missing or inaccurate information.

#### *Registry creation*

Data from the patient questionnaires (baseline, 6-month, and 12-month), genetic tests, and electronic medical records were finalized and consolidated into a master database/registry. Each study site maintained access to and control over their individual databases. The IT team worked to ensure that the databases were secure, protected, and easily sharable among the study sites as well as with their collaborating industry partners.

### *Replicability of the CEASAR study risk prediction model in a more diverse patient sample*

The researchers found that patient demographic and clinical characteristics differed by which type of treatment was prescribed. Patients receiving radiation therapy tended to be older, less educated, were less healthy, exhibited higher blood biomarker levels and Gleason scores, were more likely to be African American, presented with more depressive symptoms, and were more likely to report sexual dysfunction and urinary incontinence, compared with patients receiving surgery or being managed by active surveillance. These findings were similar to those previously observed in other studies. Comparatively, more African American men were represented in the current study and were being managed with active surveillance (43.2% versus 17% in the CEASAR study). According to the research team, this result reflects, “changing practice patterns.”

According to the researchers, “In general, variables used in the propensity score models showed comparable relationships to the prediction of surgery vs. radiation therapy... Only the effects of younger age, marital status and [overall health] on the probability of surgery vs. radiation differed between the two studies.”

Risk groups were formed based on the full set of variables included in this study and previous studies. Baseline scores were used to follow-up on sexual dysfunction and urinary incontinence for patients receiving surgery versus radiation.

Two risk groups (high and low) were formed and compared with groups from previous research. Using the two risk models, the researchers found that the treatment effects were comparable. The inclusion of genetic information seemed to contribute to the magnitude of treatment effectiveness for some comparisons, however, these results are preliminary and are the subject of ongoing analyses.

The above findings suggest that using propensity score methods to determine treatment effects offer comparable results across two studies that vary by patient characteristics, study sites, and changes over time in treatment strategies. According to the researchers, “Using the models, individual men can be counseled on their harms for either treatment, dependent on their baseline profile.”

### *Correlating genetic risk with patient characteristics and treatment*

Using the genetic test risk classifier, 50.2% of the sample was classified as low risk, 20.3% were intermediate risk, and 29.5% were high risk. Genetic test results varied by treatment type, race/ethnicity, and by risk group. Patients being managed by active surveillance were scored lower than either treatment group. Fourteen percent of patients on active surveillance were assigned high risk compared to 40 percent for either treatment group. African American and Latino men were generally at higher risk.

Although these analyses are preliminary, they warrant attention. For the overall sample, the researchers found that genetic scores were associated with Gleason score, adjusted for age, education and overall health. Genetic test results indicated less risk for those in the active surveillance group compared with those patients receiving surgery or radiation treatment.

The overall genetic test scores also correlated with patient-reported stress (higher stress levels were associated with greater genomic risk), measures of resilience (more resilience was seen in men scored at lower risk), and with risk scores that indicated tumor growth, 5-year metastasis, and 10-year mortality. African American men were found to have higher risk scores for tumor growth and mortality than Whites and Latinos. The researchers are currently evaluating the extent to which genetic tests contribute to the outcome prediction model.

#### *Additional preliminary findings*

The researchers found that African American men appear to present more genetic markers that promote tumor growth and risk for metastasis. These markers were also associated with the success of radiation therapy. Patient-reported stress was also associated with increased genetic risk for tumor metastasis, even when patient data were controlled for age, education, overall health, blood biomarkers, and Gleason scores. African Americans also demonstrated higher mean scores for stress compared with other demographic groups.

According to the researchers, “Consistent with recent studies, preliminary findings from our study suggest that more than one-fifth or as many as 40% of African American patients in our study may have been undertreated (i.e. have high Decipher risk score scores and are at moderate to high risk for metastasis).”

“Out of the individuals who have undergone germline genetic testing to date, we have identified prostate cancer specific pathogenic mutations in 5% of patients and variants of unknown significance in prostate cancer related genes (*PMS2*, *MSH2*, *MSH6*, *HOXB13* and *PALB2* genes) in another 15% of individuals. The pathogenic mutations that we detected involved genes like *ATM* and *BRCA1*. This corroborates with previous literature regarding the prevalence of these mutations in prostate cancer patients. We will utilize variants reported by the ProstateNEXT assay in our predictive model as risk factors of hereditary predisposition to prostate cancer. The integration of hereditary germline variants with tumor gene expression data for genes such as *ATM*, *NBN*, etc. provides novel opportunities to determine function and penetrance of germline variants, to improve molecular profiling and further improve risk prediction models. Additionally, these mutations can be therapeutically meaningful in prostate cancer given the potential of predicting durable response to certain drugs like PARP inhibitors in *BRCA* mutation positive patients.”

“These findings are preliminary and are only intended to suggest future analyses with additional data. However, they are suggestive of a possible biochemical / biobehavioral / genomic pathway for racial/ethnic differences in tumor progression and prostate cancer mortality that may have important implications for precision medicine and reduction of disparities in prostate cancer treatment and outcomes.”

#### **Expert Evaluation by Dr. Mark Rubin, MD**

According to the evaluator, “There are several significant accomplishments made by Dr. Greenfield’s team that are worth noting. First, although complex, the ability to use genomic data was clarified and the two partner companies have agreed to pay for all testing. Importantly, the investigators will have access to a more complete data set (called GRID) and not just the limited Decipher scores. Second, a leadership team with broad representation was formed. Third, a patient advocate committee was formed-this will be

important for broader acceptance and should help understand the communities' needs and concerns. Fourth, it appears that standard protocols were established. Fifth, the team dealt with the multifold manner of collecting survey data using different types of formats. The targeted enrollment of 600 was exceeded. It is reported that 76% of those approached agreed to participate. It is good to see in the tables provided that the team is starting to capture data related to the populations served...

It is good to see the generation of molecular data. In comparing disparities results, how will the team address the complexity of all contributing factors as they may play a role on the genetic-genomic results? For example, cultural bias toward certain types of therapy or attitudes toward health care.

In summary, this pilot project is making good headway. The importance of this project is that it takes on a specific disease type with experts in the field. It addresses an important issue surrounding health disparities and utilizes state of the art genomic (tumor) and germline genetic testing.”

### **Assessment of How the Project Addressed Programmatic Goals**

The project successfully addressed multiple goals of the CIAPM program, including 1) demonstrating the promise of precision medicine, 2) use of existing data, 3) efficient and effective data management, 4) developing capabilities for precision medicine, 5) engaging patients, 6) progress toward reducing health disparities, 7) developing clinical and commercial potential, 8) potential for downstream use, including public availability, 9) potential to scale and leverage multiple electronic health records systems and other platforms, 10) insights to overcome patient engagement challenges, and 11) establishing and utilizing partnerships. Each project was not expected, nor designed, to address each point. Only goals relevant to the project were considered.

#### *Demonstration of the promise of precision medicine in a specific disease area, health issue, technology, or fundamental biological process*

According to the researchers, “There still remains clinical uncertainty about the relative effectiveness of different treatment options for early stage prostate cancer since current risk factor models. The current methods that are used are patient age, Prostate Specific Antigen (PSA) level, Gleason score and tumor stage which do not adequately identify subgroups of patients at low or high risk for suboptimal outcomes. Much of this variability may be explained by variables not currently included in risk prediction models, such as aggressiveness of tumor type, as indicated by biomarkers and genomic classifiers, and patient characteristics, such as comorbidity, social circumstances, and overall health or life expectancy. This project was able to take these additional patient characteristics and create risk prediction models that can provide an in-depth look into an individual patient to help provide insight on the best treatment plan with optimal outcomes that is personalized for that individual.”

#### *Use of existing patient data and other data sources*

The researchers used existing patient data from electronic medical records regarding blood biomarkers, Gleason scores, tumor stage, types of treatment, complications, and other



clinical variables. This study added genomic data and information from patient questionnaires.

#### *Efficient, effective data integration and analysis*

According to the researchers, “Due to the multiple different types of data and information in general and from each site, we created an efficient and effective data integration system. We received questionnaire, EMR, and genetic data in different formats from each site due to the individual site’s IRB requirements and how the data are stored and distributed. Also, each data file (questionnaire, EMR, and genomic data) was coded somewhat differently by site requiring transformation to common metrics for the integrated file. Quality control was checked by the Project Manager and Dr. Sergio Gago. Dr. Gago and the IT team created the master queryable database, which has been used to create data for the final report and will continue to be used by the entire study team for further analysis.”

#### *Development of precision medicine capabilities*

According to the researchers, “The data from this study has generated a risk prediction model for patient harms that will be useful to aid physicians and patients in personalizing prostate cancer treatment decisions to maximize effectiveness, and choose the treatment optimal for individual patients. This will help advance the field of prostate cancer research as well as precision medicine by creating collaboration and shared decision making in treatment options between the physician and early prostate cancer patients.”

#### *Development and implementation of participant engagement strategies*

According to the researchers, “One of the goals for this study was to create a Patient Stakeholder Advisory Committee, or “Citizen Scientist Committee” to gain feedback regarding our current methods as well as gain insight in the best methods of dissemination of results. We were able to form this committee and have had multiple discussions with them, using their personal experiences and expertise in prostate cancer treatment to shape key project decisions including the current status of the project and its future for personalizing medical care, and insights on the research protocol, study measures, analyses and dissemination targets and methods. The committee was very eager to be a part of this project and played an invaluable role in its success. The members of the committee will continue to play an active role in dissemination of study results and further efforts to secure additional funding.”

#### *Reduction of health disparities within the project timeframe and beyond*

The results of this study illuminate “a path to reducing disparities among African American patients by identifying those at higher risk for aggressive tumors than would be predicted using clinical variables alone, and therefore in need of more aggressive treatment.”

#### *Development/realization of the project’s clinical and commercial potential*

According to the researchers, “The project, as noted, has considerable clinical implications and potential for optimizing the match between individual patients and the most effective treatment. The primary goal of the project was to provide doctors and patients, at the time of decision making, with the tools necessary to anticipate, on an individual level, both the benefits and the harms of the treatment that are major and disruptive of the quality of life. In this case, from a clinical and commercial standpoint, our potential includes estimating the

probability of cancer progression so that treatment and monitoring can be modified and improve treatment outcomes.”

*Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications*

According to the researchers, “If the data for the predictive models are feasible to collect, and if the models are automated and can be inserted into the electronic medical record, much like the cardiac calculator, then the goals of precision medicine for prostate cancer can be realized. The generic instruments can be used to generate risk profiles for other cancers as well and other diseases, as our prior research has demonstrated in diabetes. It is our notion that all of the products should be in the public domain and easily accessible. We are working with our IT department to develop a common-data sharing platform to potentially include our master database as well as the predictive models which can be distributed among commercial companies, universities, or even in clinics to help physicians and the patients make a personalized treatment plan.”

*Potential to scale and leverage multiple electronic health records systems and other platforms*

According to the researchers, “Our risk prediction models are based on the data collected in multiple diverse institutions at the present time. Both the epidemiology and the treatments for prostate cancer are evolving and thus the models will evolve over time. For example, during the time of this study, treatment has shifted in favor of active surveillance and the harms of intensive treatment are not relevant in these patients, but the probability of progression in the “untreated” patients is critical for monitoring, for addressing anxiety, and for treatment if it appears that the tumor is recurring or progressing. Therefore, the creation of the “federated registries” and leveraging multiple different electronic health record systems to sustain them over time is critical. Randomized trials will not be able to provide the necessary effectiveness information compared to dynamic registry data derived from high quality institutions. The data on disparity populations are particularly needed and only observations studies such as this one, can provide such data. Our collaboration crossing public/private institutions that uses multiple different electronic health record systems, produce the data necessary to individualize care and guide appropriate payments for the precise treatments that an individual should have.”

*Solutions to and insights about participant engagement challenges*

According to the researchers, they encouraged participant engagement by calling patients by phone and “sending consent forms through the mail to try to help streamline the process. Additionally, [they] identified a method that would help streamline the consenting/recruitment process, which [had] patients sign the consents electronically on REDCap. These methods helped increase enrollment speed and participant engagement, due to many patients living far away or who did not have scheduled appointments at their designated clinic scheduled. These were the only main challenges we faced during the length of this project.”

*The development of partnerships/collaborations*

According to the researchers, “Over the course of this project, we developed multiple partnerships that were key to the success and impact of the project as a whole. With the help of one of our site PIs, Dr. Ed Uchio we developed collaboration with GenomeDX who agreed to provide their genetic test (Decipher) to aid in the prediction of the probability of

metastasis in prostate cancer patients, and to cover the costs of their test for the patients enrolled in the study. The Decipher test costs between \$3000-\$5000 per test. This collaboration therefore resulted in-kind support of around 1.8 million dollars.

Additionally, we secured the collaboration of another genetic company, Ambry Genetics, a Konica Minolta Company, through a presentation of our project at a CIAPM conference. After some discussions, Ambry Genetics joined the research team and also agreed to cover the costs of their genomic test, ProstateNEXT, a 14-gene panel that tests the risk of developing prostate cancer. The collaboration of Ambry Genetics and their willingness to wave the cost of their test, ProstateNEXT, as well as the time and support from their team, provided an additional in-kind support of around \$1.8 million dollars for this project.”

## Conclusion

The goal of providing information about disease risk and predicted treatment efficacy to California men with early prostate cancer is within reach. The prediction models developed in this study are expected to continue to develop in capacity to serve individuals or small groups of patients with accurate information on the likelihood of intensive treatments working for them as well as the probability of cancer recurrence. Information derived from a combination of patient-reported, clinical, and genomics data would allow patients to select the optimal treatment based on their baseline scores at the time of diagnosis.

As the models are finalized, the research team plans to share their findings with physicians in southern California, beginning with the urologists and radiation oncologists at the five partner institutions. Due to the network developed over the course of this project, implementing the models into the clinic setting is likely to occur.

## 8. Precision Medicine for Multiple Sclerosis: Making It Work

Principal Investigators: *Project inception* Walter F. Stewart, PhD, MPH, Sutter Health;  
*Project completion* J.B. Jones, PhD, MBA, Sutter Health  
Project Period: March 1, 2017 – December 31, 2018

### *Research Team and Collaborators*

#### Sutter Health

J.B. Jones, PhD, MBA  
Sherry Yan, PhD  
Walter F. Stewart, PhD, MPH  
Christa A. Bruce, MPH  
Jacqueline Liu, BS  
Chelsea Lunders, MA  
Diamonne Mitchell, MPH  
Satish Mudiganti, MS  
Ivelina D. Popova, ITIL, ITSM  
Eilleen Sabino-Laughlin, MPH  
Zijun Shen, MS  
Talaya Sin, MA

#### UC San Francisco

Riley Bove, MD  
Stephen Hauser, MD  
Erica Schleimer  
Tanya Krishnakumar

#### Roche/Genentech

Laura Julian, PhD  
Jaclyn Wise

#### Sutter's Jordan Research and Education Institute

Joanna Cooper, MD  
Lynn Jehle, NP

#### Palo Alto Medical Foundation

Joe Lacy, MD  
Lee Greenwald, MD

#### Plain Language Health

Jen Pearce, MPA

#### National Multiple Sclerosis Society

Janelle Del Carlo

#### Sutter Philanthropy

#### Sutter Health & UCSF patient advisors

## Scientific Background and Context

### *Multiple sclerosis*

Multiple sclerosis (MS) is a chronic nervous system disease that affects the brain and spinal cord when the body's immune system mistakenly attacks healthy cells, causing neurodegeneration. Usually noticed first between the ages of 20 and 40, MS patients face decades of physical disability. Rather than a gradual worsening of symptoms, patients experience relapsing episodes of disability in between periods of relative health. As the disease progresses, windows of health diminish in frequency and duration, leaving the patient in a state of permanent disability.

The disease affects people very differently, presenting a challenge to health care providers in designing and adapting treatment strategies. It is at least three times more common in women than men and has historically been more often associated with Caucasians of northern European ancestry. Recent evidence from several studies, however, suggests that the incidence of MS may be higher in African Americans than Caucasians, possibly due to increasing awareness of MS and more accurate diagnoses. At Sutter Health in 2016, 7% of patients seeking MS care were African American and 7% were Latino.

## Project Summary

With great variability between patients with MS diagnoses, the research team set out to develop an interactive digital tool that would instantly combine the latest information relevant to a patient and display it during a medical appointment to facilitate clinical decision-making that aligns with the patient's experience, preferences, and priorities. This project builds on previous efforts that gathered data from large groups of MS patients to improve neurologists' ability to predict how MS may develop over time for a specific patient. This work adds the ability to incorporate real-time data from the patient's electronic health record (EHR) as well as patient-reported information about symptoms experienced between medical appointments. Monitoring symptoms in a consistent way and capturing that information to share with a patient's physician can lead to earlier and more effective treatment.

The platform developed through this project, neuroSHARE, is an interactive, digital health solution that supports access to relevant clinical and patient-reported data, predicts the future course of disease using data analytics, provides guidance for disease management aimed at slowing progression and addressing associated symptoms, and supports shared decision making between the patient and physician. The research team utilized well-established protocols to progressively refine neuroSHARE features to facilitate user-friendliness, usefulness, and user competence.

The web-based platform is launched from within the EHR and addresses the challenges of delivering precision care by 1) simplifying complex information by curating, organizing, visualizing, and emphasizing the most relevant details for easy use at the point-of-care and 2) supporting a patient's understanding, awareness, and involvement beyond what the current EHR is equipped to do. In addition to incorporating knowledge based on the latest research studies, neuroSHARE is designed to interpret data provided directly by patients

and ensure that the next medical appointment addresses any concerns and priorities directly.

Finally, the research team developed neuroSHARE in a way that makes it adaptable to a number of other neurological conditions that require ongoing care, including movement disorders, seizure disorders, chronic headache, stroke, and dementia.

## Project Achievements

The web platform neuroSHARE was successfully designed, developed, deployed, and evaluated in three real-world neurology practice settings. While analyses are ongoing, the following achievements reflect preliminary data and findings.

### Expert Evaluation by Dr. Harold Lehmann, MD, PhD

According to the evaluator, “Neuro-share is a project from Sutter Health designed initially around the management of Multiple Sclerosis. The overall goal was to make ‘precision medicine’ at the bedside, with patient data displayed so that guideline-informed shared decision making could be performed and so that patients understood better their condition and the implications of treatment options. The nature of ‘precision medicine’ refers to having as much relevant patient data available in the most interpretable way, along with guidelines or other relevant sources of knowledge available as well, to make treatment decisions as patient-specific as possible. The access to all this information is through a clinic Web-App called neuroSHARE, which provides a dashboard of the patient, whose central visualization shows the history of the patient over time, with treatments, outcomes, monitoring, and encounters all displayable (through a user’s choice), along with the patient’s concerns, guideline-based steps for disease-related care, comparison with a like cohort of patients, and computer-supported progress notes. The data in the app are informed by information supplied by the patient through a patient-facing electronic questionnaire.” The app was delivered to 4 neurology clinicians practicing in 3 different neurology clinics.

As the primary application is web-based, this implies interoperability between sites and different clinical settings, avoiding the need to re-engineer the interface at each site. At present, the researchers have provided qualitative results from the first 3 months of implementation.

According to the evaluator, “All technical milestones were met, except for the after-visit summary. Most implementation took longer than proposed. In particular, IRB (ethics board) and other contractual and regulatory approval took almost a year longer than anticipated.”

“Feasibility has been demonstrated, for both technical and workflow integration. That is, patients and providers could use the neuroSHARE environment in the course of routine clinical care.”

Regarding the research team’s use of existing patient data and effective data integration and analysis, they successfully delivered on their proposal to provide data and information in a visual platform that can be viewed by both patient and physician. According to the

evaluator, “The data display nicely integrates data about treatments (time span), outcomes (line graph), monitoring and visits (events) in a single graph.”

The evaluator commented that the data intended for decision support could have been more robust in its tailoring to the individual patient, but the results are recognized as preliminary and therefore may adapt toward greater orientation toward unique patient conditions.

A Stakeholder Advisory Group was established and composed of patients and health care providers at the start of the project. Following deployment of the digital platform, nearly 100 requests for modifications were provided to the design team, speaking to the level of user engagement.

Of the group of patients that were sent a clinical questionnaire, about one third of patients noticed; of those who did notice, 95% completed it. Over the course of the project, clinicians used neuro-SHARE about one third of the time. When interviewed, participating clinicians commented on their expectations of employing the software more frequently once the patients submitted pre-visit questionnaires and once the Virtual Cohort was implemented.

As stated by the evaluator, “Patients did report that shared use of the app affected the discussion between them: More information was discussed, more domains were covered, and there was a greater attention to the ‘big picture.’”

When asked whether this project provided insights to address challenges in precision medicine, the evaluator stated, “The investigators provide a solid list of barriers to the creation, evaluation, and adoption of technologies like these.”

Regarding the development of partnerships, the evaluator commented on the team’s ability to build a network “beyond the original partner of the National MS Society and UCSF, to include Palo Alto Medical Foundation, Plain Language Health, and Roche/Genentech (as well as some Sutter units: Research and Education Institute, East Bay Medical Foundation, Philanthropy).”

As per the adaptability of this digital platform beyond applications to patients with MS, the evaluator stated, “From what is displayed in the Final Report, the approach is conceptually readily generalizable. Some details would have to change (types of outcomes, treatments; types of guidance; definition of the Virtual Caseload), but the format appears generalizable. Whether patients with, and providers for, other disease would request conflicting features, remains to be seen, given the investigators’ commitment to user-based design. Whether the SHARE implementation has to be rigorously evaluated for each disease remains to be seen, as well.”

According to the evaluator, “The investigators bumped into institutional and regulatory barriers related to governance: was this a quality-improvement project? Was this IT development? Was this research? Different committees and agencies have at times conflicting guidance, which delays progress. This conflict will be met by every project aiming for precision medicine. While the investigators came up with a local solution, this set of barriers took up over half the original project time. Broader guidance and policy is needed.”

In terms of how the research team leveraged state funds, the evaluator noted their ability to obtain over \$330,000 from non-Sutter sources, \$250,000 from Sutter entities, and approximately \$100,000 in in-kind contributions from Sutter and UCSF in the form of investigator time and technology development.

## Assessment of How the Project Addressed Programmatic Goals

The project successfully addressed multiple goals of the CIAPM program, including 1) demonstrating the promise of precision medicine, 2) use of existing data, 3) efficient and effective data management, 4) developing capabilities for precision medicine, 5) development and implementation of patient engagement strategies, 6) impact for patients, 7) progress toward reducing health disparities, 8) developing clinical and commercial potential, 9) potential for downstream use, including public availability, 10) potential to scale and leverage multiple electronic health records systems and other platforms, 11) insights for institutional or regulatory policies and processes, and 12) establishing and utilizing partnerships. Each project was not expected, nor designed, to address each point. Only goals relevant to the project were considered.

### *Demonstration of the promise of precision medicine in a specific disease area, health issue, technology, or fundamental biological process*

The neuroSHARE platform demonstrates the integration of both technical and workflow perspectives to bring data to the point-of-care, when physicians and patients meet to discuss treatment plans and concerns. According to the research team, they have seen, “early indications from qualitative data collected during our run-in period that these data are being used in, and favorably influencing, the encounter itself. Whether these data have the expected impact on longer-term outcomes will be assessed in our six-month pilot evaluation.”

### *Use of existing patient data and other data sources*

The neuroSHARE platform used the following data and data sources:

- Clinical data: all data from Sutter Health’s EHR system were abstracted, including information regarding medications, laboratory tests, in-person appointments, referrals, progress notes, and others.
- Patient-reported data: Prior to an office visit with a clinician, patients completed a questionnaire about their MS diagnosis and experience. In addition to a sequence of metrics, the questionnaire asks patients to describe the concerns that they want to address during their next visit. These concerns are clearly emphasized within the neuroSHARE application to help guide the discussion during the office visit.
- Disease-specific data from an MS reference population: neuroSHARE allows a physician to display a “Virtual Caseload” within the application, based on patient data gathered in previous research. It allows the clinician and patient to consult the aggregated experience of like individuals to inform treatment choices.
- Imaging Data: A clinician has the option to display ordered and completed MRI scans within the platform. According to the researchers, “This feature of neuroSHARE is designed to significantly reduce the amount of time clinicians spend searching for, retrieving, and viewing a patient’s MRI images.”



- **Audit Data:** The audit file is a time-stamped record of each interaction with neuroSHARE or the EHR. The audit file records who accesses the solution, the specific features and functions accessed by a user, and the sequence and duration of the use of features. According to the researchers, “These audit data are essential to understanding how neuroSHARE is used (and therefore, how to evolve it), as well as to understand its impact (e.g., measuring whether physicians are spending less time searching for data in the EHR when they use neuroSHARE).”

#### *Efficient, effective data integration and analysis*

According to the researchers, “neuroSHARE accomplishes the integration of data at the point-of-care by using Application Programming Interfaces (APIs). APIs are a fast, efficient, secure way to exchange data between source systems (e.g., an EHR) in real-time and without the need to create duplicate stores of data. APIs allow applications like neuroSHARE to retrieve data from any API-enabled source system and integrate it into the visual display. We launched neuroSHARE with the ability to retrieve, via API, data associated with UCSF’s reference MS cohort, from Sutter’s EHR, and patient-reported questionnaire data that are retrieved via API from the questionnaire software platform. Longer-term, we plan to integrate data from additional sources, including remote sensors (e.g., movement data from an MS-focused mobile phone application via our collaboration with Roche/Genetech and their Floodlight Open mobile app study), social determinants of health, and information stored patient health records from non-Sutter sources. neuroSHARE is built on a platform designed to simplify the task of creating purpose-built applications that can take advantage of the wealth of data stored in source systems such as EHRs by orchestrating the requests for, and integration of, API-based data retrieval for display within the application.”

#### *Development of precision medicine capabilities*

This project demonstrated of a method for integrating MS precision medicine capabilities into everyday neurology care via a digital health application used by patients and clinicians. Through the retrieval, analysis, integration, and display of diverse data sources, neurologists are provided easy access to a wealth of data necessary to deliver precise care to MS patients. According to the researchers, “the pilot project has created a foundation for a diversity of additional precision medicine capabilities that we have started to (or soon will) pursue, including: 1) follow-up efforts to evaluate effectiveness of neuroSHARE in a large scale pragmatic trial following the completion of our pilot evaluation (supported by additional, external funding); 2) access to higher quality and more specific patient data, including data from a MS-focused mobile application, to be routinely integrated into the patient EHR and that can support ongoing collaborative research to enhance the UCSF cohort database; 3) infrastructure and tools that can be repurposed for advancing precision medicine in other areas of neuroscience (e.g., movement disorders); and 4) longer-term, the inclusion of precision medicine genomics research by leveraging Sutter’s biobanking effort to obtain serial blood samples on MS patients that can be combined with EHR data enriched by patient-generated data.”

#### *Development and implementation of participant engagement strategies*

As described by the research team, “We utilized participant engagement strategies in every phase of our project, including the design, development, and evaluation phases. neuroSHARE was co-designed with clinicians and patients. Both clinicians and patients understand how neuroSHARE should be designed to address the challenges inherent to

having or treating MS. During the development phase, we continued to include patients and clinicians in reviews of the design and the look-and-feel of neuroSHARE. In fact, we received voluminous feedback that led to a reorganization of the data and other features in the initial prototype. During the evaluation phase, we continue to involve patients and clinicians in semi-structured interviews and informal feedback sessions, all of which are critical to understanding the impact of neuroSHARE, as well as provide feedback on how to evolve the tool in subsequent iterations.

Additionally, at the inception of the project, we created a Stakeholder Advisory Group (SAG) consisting of patients with MS, leaders from the National MS Society, Sutter neurologists, researchers, and representatives from MS care centers. The SAG met on a quarterly basis throughout the project. The goal of the SAG was to ensure that neuroSHARE was responsive to the needs of both patients and clinicians. This advisory group has met five times, and has played a vital role in ensuring that we develop a solution that is health literate, useful, and meaningful. The SAG is also essential in helping us communicate in health-literate, patient-friendly ways; this was helpful during the design phase of the project, and will continue to help guide our work as we begin to share our results.”

#### *Impact for patients within the project timeframe and beyond*

According to the researchers, “Given the variability in disease progression, the episodic nature of symptoms, and the fact that patients, on average, seek MS care every six months, there will be limited direct impact to clinical outcomes in the six-month pilot period. As such, our focus is on the mediators of outcomes and patient experience. Results from early evaluation within the run-in period suggest that neuroSHARE is having a favorable impact on the patient experience. A high-level overview of key learnings about the early impact on the patient experience (including a patient’s understanding of the information conveyed, their awareness of disease trajectory, and their sense of involvement in care) during our run-in period are detailed below.

#### Key Learnings about Understanding of Information Conveyed During Encounter:

- The visual display makes the care transaction easier to understand and engage with. Patients reported that having visuals aided their recall, and that visuals, in addition to verbal inputs, make it easier to understand and engage in conversation.
- The visual display depicting the patient’s retrospective care and treatment experience is a powerful engagement tool. Patients described how seeing the full picture of their MS was a profoundly moving experience, and elicited empathy from their provider who took time to acknowledge the magnitude of the MS experience.
- From the patient perspective, the web-app appears to make the clinician more efficient; patients noticed that clinicians spent less time searching for information. Patients saw this as having peripheral benefits, including not being asked to recall details because the clinician can quickly find what they need, and having more time for discussion during the visit because less time is spent searching for information.

#### Key Learnings about Patient Awareness of Disease Trajectory:

- In order to understand where you are today, you have to understand where you were yesterday. Awareness of the longitudinal experience (the ‘big picture’) helps

answer the ‘why’ of where things are today, and having an easily digestible history of the condition can inform prospective thinking, including around medications.

- Having a visual representation of the comprehensive disease experience allows patients to be focused less on recall and more on helping to make sense of what’s ahead.
- Making information easily available increases likelihood of sharing - If you can make it easy for clinicians and patients to find and see data in the moment they want it, when it fits within the context they are discussing, it appears to lead to more sharing than may have otherwise happened.

#### Key Learnings about Patient Sense of Involvement in Care:

- Co-viewing information with the doctor puts doctors and patients on same page (‘looking at the same thing at the same time’) and promotes a conversation of equals (‘not a doctor telling me X, Y and Z’).
- We are seeing early indications that the joint viewing of information with their clinician may help some patients feel more involved in that moment, and for a subset, this may influence behavior after the clinical encounter ends. This will be explored in subsequent interviews.

Over the long term, successful adoption of neuroSHARE can lead to patients receiving evidence-based, state-of-the-art care through 1) earlier detection of disease progression and of poor treatment response; 2) switching to more appropriate disease modifying therapies; 3) improved symptom management including through referrals; and 4) better functional scores over time.”

#### *Reduction of health disparities within the project timeframe and beyond*

Analyses are underway to gauge the impact of neuroSHARE on demographic, gender and geographic disparities after six months of implementation. According to the research team, “through qualitative interview data collected from both clinicians and patients, we have very early indications that the combination of visual and auditory learning that neuroSHARE facilitates may lead to more information being retained by patients. If borne out by additional analyses, this has important implications for improving health literacy. We also have early evidence that MS disease type is being documented more often via neuroSHARE. This, too, if borne out during the full pilot evaluation, may improve our ability to understand how MS disease type influences disease progression, symptom experience, and outcomes.”

#### *Development/realization of the project’s clinical and commercial potential*

With its deployment across three clinical settings, neuroSHARE continues to be tested in real-world conditions. According to the researchers, “The underlying architecture is EHR-agnostic, instead relying on APIs to integrate data from diverse sources. We believe this approach is broadly applicable to all neurology practices, within or outside of Sutter. Our project was not designed to develop a solution that is ready for commercialization. However, we think that there is a market for solutions like neuroSHARE that address critical health care delivery challenges. Data from the pilot evaluation will help to determine if neuroSHARE can influence outcomes in a manner sufficient to justify a business case for commercialization of solutions like this.”

*Potential downstream use of tools, measurements, and data, including open public accessibility of generated data and publications*

Three study results have potential for downstream benefit: 1) the neuroSHARE platform was designed in an API-enabled ecosystem to be adapted for different purposes and therefore may be used by many others for other chronic diseases; 2) audit data and subsequent measures can be adopted more broadly by the digital health community to assess the impact of digital health solutions on clinician workflow; 3) the research team plans to broadly disseminate their findings related to the design, deployment, measurement, and staged evaluation of neuroSHARE for purposes of moving forward the science of translating precision care from the bench to the bedside.

*Potential to scale and leverage multiple electronic health records systems and other platforms*

In its current form, the neuroSHARE platform is optimized for API-based data transactions with Sutter's EHR system and additional data sources, such as the patient questionnaire software. According to the research team, "The underlying architecture was intentionally designed to support API-based transactions with alternative data sources that support this type of data exchange, all with the goal of scaling for organization-wide use. We have not yet tested this capability with any other EHR system. After we conclude our pilot evaluation, we will begin to explore what will be necessary to scale our current platform to accommodate other EHR systems and data sources; it is likely that any scaling efforts will require collaboration with an external organization that has the necessary capabilities for supporting enterprise-grade software applications."

*Insights for institutional, state, and federal regulatory processes and policies*

According to the researchers, "Deploying digital health solutions into real-world practice settings, where use is voluntary and not driven by a research protocol, requires a modified approach to evaluation that balances rigor and pragmatism. To that end, we have developed a model for a staged approach to evaluating digital health solutions in real-world settings. The initial stages focus on implementation, workflow integration and adoption, while later stages focus on changes to process and value outcomes once we achieve sufficient sample size and time in the field. This approach allows us to rapidly test and discard or evolve digital health solutions using concrete outcomes based on methods and designs suitable for each stage. The insights into what, when, and how to evaluate aspects of a digital health solution may be helpful to other projects focused on translating precision medicine to the point-of-care.

*Formalized processes, best practices, or partnerships*

According to the researchers, "Our partnership with the members of the Stakeholder Advisory Group (SAG) will continue as a formalized resource for our Neurology research program, which includes the continued evolution of neuroSHARE, as well as other neurology-focused projects. The SAG has proven to be an invaluable resource for including a diversity of viewpoints – clinician, patient, disease experts, etc. – in our work.

The primary reasons the SAG has been so effective are:

- The SAG allows each person to come together on equal footing, (i.e., the clinician's viewpoint was no more or less important than the patient's viewpoint).

- By bringing these stakeholder viewpoints together in a single forum, we were able to more quickly and effectively prioritize and make decisions.
- As part of instantiating the SAG, we developed a charter with a mission (“Create a forum where those who will ultimately use and benefit from MS-SHARE can shape its development”) and purpose (“Bring together a diverse group to provide input on content, design, evaluation and communication related to MS-SHARE”). We also defined each member’s responsibilities (e.g., show respect, direct and honest communication, listen to differing points of view, etc.).
- SAG meetings were/are facilitated by a neutral, non-Sutter facilitator with expertise in plain language communication (which helps bridge the gaps between the varied stakeholders).

Another important best practice that was used during this project is a user-centered design approach to developing and evolving neuroSHARE. We worked closely and frequently with patients and clinicians to understand how they currently go about doing the ‘jobs’ that are associated with having/treating MS and to engage them as co-designers in developing a solution (neuroSHARE) that we collectively believe will help them do these jobs more effectively. We believe that this co-design process is essential if we are to avoid the problem of throwing more technology at problems that are due, at least in part, to technology (e.g., the EHR and the problem of too much data and too little actionable information). The co-design process puts the emphasis on developing a solution to the jobs a person is trying to get done, rather than focusing on the technology and a set of features as the primary endpoint of the development effort. Because we have been focused on completing development of neuroSHARE and initiating the pilot evaluation, we have not yet disseminated this knowledge to other partners in California or beyond.

Health systems are like complex living organisms that have developed their own immune systems in order to survive in litigious, highly regulated environments. Initiatives at the legal, regulatory, or privacy frontier can look sufficiently threatening to provoke an immune response to immobilize the threat. The neuroSHARE project’s focus on using technology to take precision medicine from the ‘bench to bedside’ evoked this type of response. Technology and the use cases for which it can be employed to solve problems evolve much more rapidly than the systems put in place to mitigate risk. While an immune response safeguards the organization, it presents a challenge to project timelines as the issues raised by new use cases are identified and mitigation plans agreed upon.

Elements of our project (e.g., data sharing agreements, contractor agreements, privacy and security assessments, new technologies at the point of care, etc.) required review and approval by multiple compliance entities within our organization, including legal, privacy, information security, the institutional review board, and grants administration. For a complex digital health project such as ours, each entity had a legitimate purview, but it was not always clear how to move from identifying potential issues to approving an acceptable path forward.

The outputs from each entity are dependent on one another, even though each entity often works inside its own silo. New technologies and new use cases raise issues that rarely have a clear, black-or-white resolution. Rather, each entity has to wrestle with how to apply their existing compliance frameworks to new use cases. In our case, the decisions needed from

each entity were interdependent. The challenge was to develop a process for determining who makes the initial approval so the project can move forward, versus, for example, a scenario where privacy waits for a decision from legal while legal is waiting for privacy.

Research projects, with their own set of regulations and requirements, bring an even greater level of complexity. We found that the most effective way to solve these challenges was to bring all compliance entities together to identify potential issues and jointly wrestle with interdependent issues and identify a mutually-agreeable path forward. In response to neuroSHARE and similar digital health projects, Sutter has developed a team with this specific experience. Bringing this team together at the beginning of a complex, innovative technology project has had a significant impact on speeding up projects.

The other significant challenge we encountered is associated with the lack of existence of robust test data sets for use in development. Because external development partners cannot have access to real patient data, test data sets must be created. It is extremely difficult to create test data sets that reflect the volume and variability of real data. As a result, vendors work with limited data and test their code against limited scenarios. When code is transferred back to the system, we invariably encounter bugs associated with the difference between test data and real data. This slows down and increases the cost of testing; often, re-work is required by the vendor when bugs are identified, and some are not identified until code is in a production setting.

We minimized the impact of this challenge by allotting additional time for testing, planning our test cases carefully, and by having staff onsite during go-lives to rapidly address any edge case bugs that did not manifest until the application was used with real clinical practice data. Although we anticipate and plan for this challenge, test data are likely to remain a rate limiting step that slows the progress of translating precision medicine from bench to bedside.

Because we believe our experience is likely to occur in any health care system setting, we felt it especially important to capture our challenges, corrective actions, and lessons learned and disseminate them for the benefit of future bench-to-bedside efforts. We have detailed these challenges, the corrective actions we took, and the implications for future work.”

## Conclusion

The research team succeeded in developing a web-based digital platform to empower patients to communicate their concerns and priorities with clinicians, while also allowing clinicians access to information from different sources all in one space. The platform was rolled out across three clinical settings, and the research team is actively evaluating its effects on decision-making and health outcomes.

During this project, the research team staged an approach for evaluating digital health solutions in real-world settings that may be adapted to other chronic diseases. The initial stages target implementation, workflow integration, and adoption, while later stages address changes to process and value outcomes, allowing for the accumulation of sufficient sample size and time in the field. This approach allows researchers to rapidly test and either discard or evolve digital health solutions using concrete outcomes based on methods and

designs tailored to each stage. Preliminary findings suggest that neuroSHARE can be successfully integrated into real-world specialty care settings and favorably impact the patient's experience.

## Conclusion

Precision medicine is positioned to transform how Californians will receive tailored health care and maintain well-being. Just as important, this field is also rewriting the scientific approach to understanding health and disease. As technology usage and capacity continues to expand, the possibility to increasingly derive underlying mechanisms of disease prevention, management, and treatments seems boundless.

In its first three years, the California Initiative to Advance Precision Medicine has funded eleven projects across the state, of which eight have been completed that together demonstrate the potential of the multifaceted and quickly developing field of precision medicine. Indeed, they are projects like in this set that are helping to define the fields. Traditionally considered the application of “omics” data, such as genomics, metabolomics, and proteomics, the fields have since grown to encapsulate the application of nearly any type of data that helps to explain an individual’s uniqueness.

As data collection methods improve, researchers are turning their focus increasingly toward expanding the ability to glean useful information from data, and the critical issues of data privacy and security. Each project described in this report has leveraged data in a meaningful way to empower patients to engage in their own health care, pinpoint the cause of illness, automate imaging analysis and diagnostics, explore and recommend the most effective therapeutic options for an individual’s unique biological and lifestyle backgrounds, or provide medical support beyond the traditional walls of a clinic.

By the numbers, the research teams leveraged their state awards to attract approximately \$12 million in external matching funds and \$6.4 million in institutional in-kind support. Study results have been disseminated through 25 publications and draft manuscripts, 17 media articles and press releases, and over 70 presentations for public and professional audiences. By sharing innovative methods, frameworks, databases, and results, the projects have already begun to influence and change modern clinical approaches.

California’s precision medicine program’s emphasis on patient and community partnerships plays a critical role in fostering Governor Newsom’s and the California legislature’s mission to reduce health disparities across the state, by enhancing access to technology-enabled affordable health management systems and improving diverse representation in medical databases. Through genuine patient engagement strategies, research results are more likely to be aligned with equitable applications of precise prevention, diagnoses, and treatment.

The benefits of precision medicine are far from fully realized, but through CIAPM and other contributors, progress can be guided and shaped to ensure the outcomes support all Californians.



## Appendix A: Expert Evaluators



### Nancy J. Cox, PhD

Director, Vanderbilt Genetics Institute Director,  
Division of Genetic Medicine  
Mary Phillips Edmonds Gray Professor of Genetics,  
Vanderbilt University

Nancy J. Cox, PhD is a quantitative human geneticist leading a long-standing research program focused on identifying and characterizing the genetic component of common human diseases.

Dr. Cox earned a BS in Biology from the University of Notre Dame in 1978, a PhD in Human Genetics at Yale in 1982, and did post-doctoral research at Washington University and the University of Pennsylvania before joining the University of Chicago in 1987. She spent 28 years at the University of Chicago rising to Professor and Chief of the Division of Genetic Medicine before moving to Vanderbilt University in 2015 to become the Mary Phillips Edmonds Gray Professor of Genetics, inaugural Director of the Vanderbilt Genetics Institute, and Director of the Division of Genetic Medicine.

Dr. Cox is the past President of the American Society of Human Genetics (2016-2018), a Fellow of the American Association for the Advancement of Science, was part of the team that received the Landon Award in 2008 from the American Association for Cancer Research, and achieved the Leadership Award in 2010 from the International Genetic Epidemiology Society. Dr. Cox's current research is focused largely on integrating data on genome variation and genome function with electronic health records to push the next round of translation of genome discovery into health care. Currently funded research projects on which Dr. Cox is PI or co-PI include using these data integration approaches to analyze whole genome sequence data generated by the Centers for Common Disease Genomics, and developing the new Center of Excellence in Health Disparities for Personalized Medicine and Population Health at Vanderbilt.



**Harold P. Lehmann, MD, PhD**

Director of Research and Training,  
Division of Health Sciences Informatics  
Professor of Health Sciences Informatics,  
Johns Hopkins University

Dr. Harold Lehmann is a professor of Health Sciences Informatics and of Pediatrics at the Johns Hopkins University School of Medicine. He holds joint appointments in Health Policy and Management and in International Health at the Johns Hopkins Bloomberg School of Public Health.

Dr. Lehmann's areas of clinical expertise include adolescent medicine and general pediatrics, and he serves as the Director of Research and Training of the Division of Health Sciences Informatics. Dr. Lehmann received his BA from Columbia College, his MD from Columbia University College of Physicians and Surgeons, and his PhD from Stanford University.

Dr. Lehmann's research interests include medical informatics, evidence-based medicine, and decision making, biostatics, decision analysis, and Bayesian communications. Dr. Lehmann is a fellow of the American College of Medical Informatics and the American Academy of Pediatrics. He is a member of the Society for Medical Decision Making and the American Medical Informatics Association.



### Elaine Mardis, PhD

Nationwide Foundation Endowed Chair in Genomic Medicine  
Co-Executive Director,  
Institute for Genomic Medicine at Nationwide Children's Hospital  
Professor of Pediatrics,  
The Ohio State University College of Medicine

Elaine Mardis, PhD is Co-Executive Director of the Institute for Genomic Medicine at Nationwide Children's Hospital, the Nationwide Foundation Endowed Chair of Genomic Medicine, and a Professor of Pediatrics at The Ohio State University College of Medicine. Dr. Mardis joined Nationwide Children's Hospital in 2016.

Educated at the University of Oklahoma with a BS in Zoology and a PhD in Chemistry and Biochemistry, Dr. Mardis did postgraduate work at BioRad Laboratories. She was a member of the faculty of Washington University School of Medicine from 1993-2016. Dr. Mardis has authored over 350 articles in prestigious peer-reviewed journals and has written book chapters for several medical textbooks.

She serves as an associate editor of three peer-reviewed journals (*Disease Models and Mechanisms*, *Molecular Cancer Research*, and *Annals of Oncology*) and is Editor-in-Chief of *Molecular Case Studies*, published by Cold Spring Harbor Press. Dr. Mardis has given lectures at scientific meetings worldwide and was awarded the Morton K Schwartz award from the American Association for Clinical Chemistry in 2016. Dr. Mardis has been a member of the American Association for Cancer Research (AACR) since 2007, was the program committee chair for the 2018 AACR Annual Meeting, and is the AACR President-elect. She has been listed since 2013 as one of the most highly cited researchers in the world by Thompson Reuters.



### **Bambang Parmanto, PhD**

Director, Rehab Engineering Research Center on Information & Communication Technology Access  
Professor and Interim Chair,  
Department of Health Information Management, University of Pittsburgh, School of Health and Rehabilitation Sciences

Bambang Parmanto, PhD is a professor of Health Information Management. He received his BS in Computer Sciences from Bandung Institute of Technology in Indonesia and both a MS and PhD in Information Sciences from the School of Information Sciences at the University of Pittsburgh.

His research interests have been in developing technologies and in using advances in information technologies such as telehealth and mobile and wearable technologies to deliver adaptive and personalized interventions for individuals with chronic and complex conditions. He leads the Health and Rehabilitation Informatics research group at the University of Pittsburgh.

Dr. Parmanto is the director of Rehab Engineering Research Center on Information & Communication Technology Access and Principal Investigator of a large project on self-management for individuals with chronic and complex conditions using mobile health. He has led the development of software, applications, and multiple platforms for the implementation of telehealth for managing chronic health conditions, including: VISYTER (Versatile and Integrated System for Telerehabilitation), iMHere TeleWellness (interactive Mobile Health and Rehabilitation), SmartCAT (smartphone-based Child Anxiety Treatment), and Active2 (Active Two-Way Communication Platform for Supporting mHealth Applications). He is the recipient of the American Health Information Management Association's Triumph Research Award.



**Mark A. Rubin, MD**

Professor of Pathology of Laboratory Medicine,  
Weill Cornell  
Medicine Director, Department for BioMedical Research,  
University of Bern, Switzerland  
Project Leader, Precision Medicine Inselspital, Bern University  
Hospital, Switzerland

Mark Rubin, MD is a leader in the fields of prostate cancer biology and precision medicine as it applies to all cancers. Dr. Rubin's laboratory led a series of landmark studies defining distinct molecular features of prostate cancer, revealing pathways that are perturbed and drive different types of this cancer. Subsequently, Dr. Rubin's laboratory has been instrumental in establishing the mechanistic basis by which defined genomic alterations drive prostate cancers.

Dr. Rubin is involved in clinical trials aimed at evaluating the efficacy of molecular inhibitors in treating neuroendocrine prostate cancer. He is also developing novel drugs to target advanced prostate cancer. Dr. Rubin has translated many of his genomic discoveries into clinical tests that are currently patented and standardly used in the diagnosis and treatment of prostate cancer.

As the founding director of the Englander Precision Medicine Institute at Weill Cornell, he developed a cutting-edge genomics clinical lab and received the first New York State approval to use whole exome sequencing in the diagnosis and treatment of a broad variety of cancers. In May 2017, Dr. Rubin joined the University of Bern as Professor and Director of the Department for BioMedical Research and also as Project Leader for Precision Medicine at the University Hospital of Bern.



**John S. Rumsfeld, MD, PhD**

Chief Innovation Officer of the American College of Cardiology  
Professor of Medicine  
University of Colorado School of Medicine

Dr. Rumsfeld is an internationally recognized cardiovascular health services researcher, and a thought leader in health care innovation. His research has primarily centered on cardiovascular quality of care and outcomes, patient-reported outcome measures, and the association between depression/mental health and cardiovascular disease. He has more than 200 scientific publications and is a sought-after speaker nationally and internationally on the intersection of cardiovascular disease and depression. Dr. Rumsfeld is currently the Chief Innovation Officer for the American College of Cardiology, with a focus on digital health, advanced clinical analytics (big data / cognitive computing), and precision medicine.

He is currently a Professor of Medicine at the University of Colorado School of Medicine. Previously, he was National Director of Cardiology for the US Veterans Health Administration (VA) and Director of the VA's Clinical Assessment, Reporting, and Tracking Program, a national quality of care and patient safety program for the VA health system. In addition, he was previously the Chief Science Officer for the American College of Cardiology's National Cardiovascular Data Registries and is past-Chair of the American Heart Association's Quality of Care and Outcomes Research Scientific Council. Dr. Rumsfeld received his BS in Biology from UC Los Angeles, MD from the University of Chicago, and doctoral degree in epidemiology from the University of Colorado. He completed his internal medicine internship and residency at UC San Francisco and a cardiology fellowship at the University of Colorado.



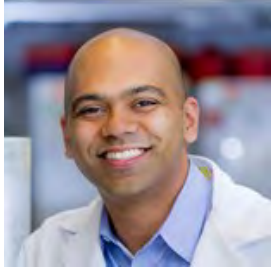
**Jack W. Tsao, MD, DPhil**

Professor of Neurology, Pediatrics, and Anatomy & Neurobiology  
Researcher at Children’s Foundation Research Institute at  
Le Bonheur Children’s Hospital, University of Tennessee Health  
Science Center

Chief of Neurology and Director of the Polytrauma/OIF/OEF Clinic  
Memphis Veterans Affairs Medical Center

Dr. Tsao received his BS in Biochemistry from Harvard College, a MS in Biochemistry from the University of Cambridge, England, a PhD in Physiology/Pharmacology from the University of Oxford, England, and MD from Harvard Medical School. Prior to finishing his MD, he was a postdoctoral fellow in at Johns Hopkins Hospital. He completed his Internal Medicine internship and Neurology residency at UC San Francisco and then began 14 years of active duty service in the U.S. Navy, where he was first stationed at Naval Hospital Jacksonville, Florida as the Neurology Department Head. While there, Dr. Tsao completed a Behavioral Neurology fellowship at the University of Florida. He was then assigned to the Uniformed Services University of the Health Sciences in Bethesda, MD for 4 years before being selected to become the inaugural Director of Traumatic Brain Injury (TBI) Programs for the U.S. Navy Bureau of Medicine and Surgery, where he managed Navy and Marine Corps TBI policy and programs for 6.5 years. While on active duty, Dr. Tsao was a member of the NATO Human Factors and Medicine - 193 Panel and served as Consultant Neurologist to the White House Medical Unit.

He is currently the Chief of Staff for the Chief of Naval Operations N81 Navy Reserve Unit, Navy Reserve Medicine Neurology Specialty Leader, and Senior Medical Officer for Operational Health Support Unit in Memphis. He has published 105 peer-reviewed articles and book chapters and edited books on TBI and teleneurology. His clinical research is focused on treatments for phantom limb pain in amputees, for which he was awarded the 2014 U.S. Navy Hero of Military Medicine, and the clinical effects of blast exposure and concussion. He is also past chairman of both the Government Services Section and the Practice Committee Telemedicine Work Group of the American Academy of Neurology. He is a fellow of both the American Academy of Neurology and the American Neurological Association.



### **Nikhil Wagle, MD**

Assistant Professor of Medicine, Harvard Medical School Associate Member, Broad Institute of MIT and Harvard Medical Oncologist, Dana-Farber Cancer Institute  
Deputy Director, Center for Cancer Precision Medicine, Dana-Farber Cancer Institute

Dr. Nikhil Wagle received his MD from Harvard Medical School and completed his residency training in Internal Medicine at Brigham and Women's Hospital, where he also served as chief medical resident, and completed his fellowship training in Hematology/Oncology in the Dana-Farber/Partners program. Dr. Wagle leads a translational research program in the field of breast cancer genomics and precision cancer medicine. The major goals of his work are to better understand the biology of metastatic breast cancer and to develop new ways to overcome or prevent drug resistance in patients with advanced breast cancer. Ultimately, his research aims to identify characteristics of tumors that might improve clinical decision-making for patients with advanced cancer. He also leads The Metastatic Breast Cancer Project ([mbcproject.org](http://mbcproject.org)), a nationwide direct-to-patient research initiative that engages patients with advanced breast cancer through social media and seeks to empower them to accelerate cancer research through sharing their samples and clinical information. The project's outreach program, developed in collaboration with advocacy organizations and patients, serves to connect thousands of patients around the U.S. with metastatic breast cancer, allowing them to participate regardless of where they live.

Dr. Wagle has received grant funding from the Susan G. Komen Foundation, the Breast Cancer Research Foundation, the V Foundation, the Breast Cancer Research Alliance, and the Department of Defense. He is a recipient of several awards and honors, including a Young Investigator Award from the Conquer Cancer Foundation of the American Society of Clinical Oncology, the 2013 Landon Foundation-American Association for Cancer Research (AACR) INNOVATOR Award for Research in Personalized Cancer Medicine, and the inaugural 2016 AACR NextGen Award for Transformative Cancer Research. His work has been published in the *New England Journal of Medicine*, the *Journal of Clinical Oncology*, *Cancer Discovery*, and *Nature Medicine*.



# Appendix B: Publications, Press Releases, and Presentations Generated as a Result of CIAPM-Funded Research

## 1. California Kids Cancer Comparison (CKCC)

### *Publications*

- Learned K, Durbin A, Currie R, et al. Barriers to accessing public cancer genomic data. *Scientific Data*, 6, 98 (2019), <https://www.nature.com/articles/s41597-019-0096-4>.
- Vaske O, Bjork I, Salama S, et al. RNA-Seq comparisons for difficult-to-treat pediatric cancer patients: results of a California Initiative to Advance Precision Medicine Demonstration Project. [In preparation]
- Vaske O and Haussler D. Data sharing for pediatric cancers. *Science* 2019; 363(6432):1125
- Pfeil J, et al. Hydra: a Bayesian non-parametric pipeline for discovering cancer gene expression subtypes. In review at *Bioinformatics* (July, 2019)
- Beale H, et al. Framework for determining accuracy of RNA sequencing data for gene expression profiling of single samples. In review at *Clinical Chemistry* (July, 2019)
- Sanders L, et al. Comparative RNA-Seq analysis aids in diagnosis of a rare pediatric tumor. 2019. *Molecular Case Studies*. Accepted.
- Newton Y, Rassekh SR, Deyell RJ, et al. Comparative RNA-Sequencing Analysis Benefits a Pediatric Patient With Relapsed Cancer. DOI: 10.1200/PO.17.00198 *JCO Precision Oncology* - published online April 19, 2018. <http://ascopubs.org/doi/pdf/10.1200/PO.17.00198>
- Newton Y, Novak AM, Swatloski T, et al. TumorMap: Exploring the Molecular Similarities of Cancer Samples in an Interactive Portal. *Cancer Res.* 2017;77(21):e111-e114.

### *Press Releases*

- Share The Data: David Haussler Envisions A New Future For Pediatric Cancer: <https://treehousegenomics.soe.ucsc.edu/2018/02/22/share-the-data-david-haussler-envisions-a-new-future-for-pediatric-cancer/>
- California Awards UCSC Precision Medicine Funds for Childhood Cancer Research: <https://treehousegenomics.soe.ucsc.edu/2018/01/08/california-awards-ucsc-precision-medicines-funds-for-childhood-cancer-research/>
- Learning about Sarcomas: <https://treehousegenomics.soe.ucsc.edu/2018/07/31/learning-about-sarcomas/>

### *Presentations*

- Olena M. Vaske. “Comparative RNA sequencing benefits pediatric cancer patients” invited speaker, Greehey Children’s Cancer Research Institute Seminar Series, San Antonio, TX, December 8, 2017.
- David Haussler. “Odyssey into the human genome,” invited speaker, special seminar, Connecticut College, New London, CT. September 2018.
- Jacob Pfeil. Poster from ASCO, June 2, 2018 Chicago, IL “Gene Expression Analysis for Improved Subtyping of High-Risk Neuroblastoma”.

- Lauren Sanders. Poster Presentation, International Symposium on Pediatric Neuro-Oncology, Denver, CO, July 2, 2018. “Characterizing pediatric high-grade glioma with comparative gene expression analysis and preclinical organoid models”
- Lauren Sanders. Abstract, Association of Clinical Oncology, Chicago, IL, June 1, 2018 “Comparative gene expression analysis for identifying clinically relevant overexpressed genes in childhood brain tumors.”
- Sofie Salama presented at Children’s Oncology Group (COG) Data Sharing Standards meeting on Treehouse methods and available data for sharing, April 10, 2018.
- Olena M. Vaske took part in the COG meeting in October, 2018, where she participated in the discussions of the Leukemia and Lymphoma Society’s data sharing priorities.
- Olena M. Vaske. Treehouse Childhood Cancer Initiative. NCI Kids First Childhood Cancer Workshop, National Institutes of Health, Bethesda, MD, March 12, 2018.
- Jacob Pfeil. Poster Presentation and Travel Award, TGen Pediatric Precision Oncology Conference, Scottsdale, AZ, March 5, 2018 “Gene Expression Analysis for Improved Subtyping of High-Risk Neuroblastoma.”
- Lauren Sanders. Poster Presentation, TGen Pediatric Precision Oncology Conference, Scottsdale, AZ, March 5, 2018 “Comparative gene expression analysis methods for identifying therapeutically relevant genes in childhood brain tumors.”
- David Haussler. “Genomics at UCSC,” invited speaker, UC's Council for Vice Chancellors for Research (COVCR), University of California, Santa Cruz, CA. March 2018.
- Childhood Cancer Awareness month: participated in local awareness events, including Kidrageous in Watsonville on September 23rd and Monterey on October 8th, and provided social media content about pediatric cancer efforts each day to increase social media presence throughout the month - September 2018.
- Sofie Salama and Isabel Bjork. California Initiative to Advance Precision Medicine, Advisory Committee and Public Comment Session, September 2018.
- Olena M. Vaske. Harnessing big data to help children with cancer. Santa Cruz Biden Cancer Community Summit, Santa Cruz, CA, September 21, 2018.
- David Haussler. “Odyssey into the human genome,” invited speaker, Santa Cruz Rotary meeting, Santa Cruz Rotary, Santa Cruz, CA. August 2018.

## 2. Precision Diagnosis of Acute Infectious Diseases

### *Publications*

- Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G. Validation of Metagenomic Next-Generation Sequencing Tests for Universal Pathogen Detection. Archives of pathology & laboratory medicine. 2017. Epub 2017/02/09. doi: 10.5858/arpa.2016-0539-RA. PubMed PMID: 28169558.
- Mongkolrattanothai K, Naccache SN, Bender JM, Samayoa E, Pham E, Yu G, et al. Neurobrucellosis: Unexpected Answer From Metagenomic Next-Generation Sequencing. Journal of the Pediatric Infectious Diseases Society. 2017. Epub 2017/01/08. doi: 10.1093/jpids/piw066. PubMed PMID: 28062553.
- Chiu CY et al., “Fatal human case of mosquito-borne St. Louis encephalitis infection diagnosed by clinical metagenomic sequencing, California, 2016”, 2017, (under review at Emerging Infectious Diseases).

- Murkey J, et al., “Hepatitis E virus-associated meningoencephalitis in a lung transplant patient diagnosed by clinical metagenomic sequencing”, 2017, (under review at Open Forum Infectious Diseases).
- Naccache, et al., “Clinical validation of a metagenomic next-generation sequencing assay for diagnosis of meningitis and encephalitis from cerebrospinal fluid”, 2017, (in preparation).
- Wilson, et al., “Clinical utility and comparative effectiveness of a metagenomic next-generation sequencing assay for diagnosis of neurological infections”, 2017, (in preparation).

#### *Press Releases*

- *Time Magazine* cover story, “The World is Not Ready for the Next Pandemic”, May 4, 2017.
- “At TEDMED, Human Drama Comes First. And That’s a Good Thing”, <https://www.kqed.org/futureofyou/300317/at-tedmed-human-drama-comes-first-and-thats-a-good-thing>
- “Innovating to Fight Epidemics: Latest Ideas From TEDMED”, <https://blog.nationalgeographic.org/2017/02/07/innovating-to-fight-epidemics-latest-ideas-from-tedmed/>
- “The Hopeful Future of Precision Medicine”, <https://blog.tedmed.com/hopeful-future-precision-medicine/>
- “The elitist sheen at TEDMED makes some scientists uneasy. Still, they flock to it”, <https://www.statnews.com/2016/12/08/tedmed-scientists-uneasy/>

#### *Presentations*

- Fortune Brainstorm Health 2017 Conference, “How to Save the Planet in a Post-Antibiotic World” (the PDAID study was discussed), [https://www.youtube.com/watch?v=bmuWJIPFAPA&list=PLS8YLn\\_6PU1IQD6xqwcVF\\_PvXX2IjAdWc&index=44](https://www.youtube.com/watch?v=bmuWJIPFAPA&list=PLS8YLn_6PU1IQD6xqwcVF_PvXX2IjAdWc&index=44)
- TEDMED 2016 Meeting, “What if Next-Generation Sequencing Could Help Us Diagnose Mysterious Infectious Illnesses?”, Palm Springs, CA, December 1st, 2016 <http://another-option.com/tedmed-2016-speakers-look-toward-the-futurein-fighting-infectious-diseases-and-impeding-the-aging-process/>
- 2015 ASM Conference on Rapid NGS Bioinformatic Pipelines for Enhanced Molecular Epidemiologic, Investigation of Pathogens, “SURPI: Deep Sequencing of Infectious Disease”, Washington DC, September 26th, 2015 speaker.
- 2015 NIST FDA Workshop on Standards for Pathogen Detection via Next-Generation Sequencing, “Establishing Standards for Microbial Detection via Next-Generation Sequencing”, October 27th, 2015.
- 2015 Illumina Bay Area Users Group 2015 Meeting, “Metagenomic Next-Generation Sequencing for Infectious Disease Diagnosis”, Burlingame, CA, December 10th, 2015 speaker.
- 2016 University of Washington Virology Meeting, “Next-Generation Sequencing Approaches for Viral Diagnosis and Discovery in Human Infections”, Seattle, CA, January 8th, 2016 speaker.
- 2016 Precision Medicine World Conference (PWMC) 2016 Meeting, “Point-of-Care and Genomic Approaches for Precision Medicine in Infectious Diseases”, Mountain View, CA, January 25th, 2016 speaker.

- 2016 AGBT (Advances in Genome Biology and Technology) 2016 Meeting, “Metagenomic Next-Generation Sequencing for Diagnosis of Infectious Diseases”, Orlando, CA, February 11th, 2016 speaker.
- 2016 NGS Dissemination Meeting at the Washington DC Press Office, “Next-Generation Sequencing for Patient Care in Clinical Microbiology”, Washington DC, February 19th, 2016 speaker.
- 2015 FOGM (Future of Genomic Medicine) Meeting, “Genomics for Infectious Diseases: Where do we go from here?”, San Diego, California, March 3rd, 2016 speaker and panel member.
- 2016 Association of Medical Microbiology and Infectious Diseases (AMMI) Annual Meeting, Special Lecture, “Next-Generation Sequencing for Diagnosis of Infections”, Vancouver, BC, Canada, March 31st, 2016 plenary speaker.
- 2016 Future of Genomic Medicine IX, “Metagenomics Technologies for Infectious Disease Diagnosis”, San Diego, CA March 3rd, 2016 speaker.
- 2016 23rd International Molecular Medicine Tri-Conference, “NGS Assays for Diagnosis of Infectious Diseases”, San Francisco, CA, March 7th, 2016 speaker.
- 2016 Medimmune Conference, “Genomics-Based Diagnostic Testing for Infectious Diseases”, Gaithersburg, MD, May 4th, 2016 speaker.
- 2016 Mt. Sinai Hospital Grand Rounds, “Metagenomic Sequencing: a New Paradigm for Infectious Disease Diagnostics”, June 1st, 2016, speaker.
- 2016 American Society for Microbiology, Special Presentation for Bill Gates & Advisors, “Genomic-Based Testing in Point-of-Care Settings”, June 16th, 2016, speaker.
- 2016 Molecular Diagnostics for Infectious Disease Conference, “Metagenomic Next-Generation Sequencing for Diagnosing Infectious Diseases”, August 17th, 2016, speaker.
- 2016 Fondation Mérieux Conference, “Metagenomic Sequencing for Antibiotic Resistance Characterization Directly from Clinical Samples”, Lyon, France, October 26th, 2016, speaker.
- 2017 Precision Medicine World Conference (PWMC).
- 2016 Meeting, “Clinical Next-Generation Sequencing for Infectious Diseases Diagnosis”, Mountain View, CA, January 23rd, 2017, speaker.
- 2017 IMPAACT Laboratory Science Meeting, “Precision Diagnosis of Acute Infectious Diseases”, Washington, D.C., January 25th, 2017, speaker.
- 2017 24th International Molecular Medicine Tri-Conference, “Emerging Assays for Infectious Diseases in Diagnosis and Outbreak Surveillance”, San Francisco, CA, February 5th, 2017, speaker.
- 2017 Presentation at Medimmune, “Metagenomic Sequencing for Diagnosis of Acute Infectious Diseases”, Mountain View, CA, April 20th, 2017, speaker.
- 2017 Illumina Spring Microbiology Symposium, “Precision Diagnosis of Acute Infectious Diseases by Metagenomic Sequencing”, Emeryville, CA, April 26th, 2017.

### 3. Artificial Intelligence for Imaging of Brain Emergencies

#### *Publications*

- Kuo WC, Haene C, Yuh EL, Mukherjee P, Malik J (2018). PatchFCN for intracranial hemorrhage detection. arXiv: 1806.03265

- Kuo WC, Haene C, Yuh EL, Mukherjee P, Malik J (2018). Cost-sensitive active learning for intracranial hemorrhage detection. arXiv: 1809.02882

#### *Presentations*

- Kuo WC, Haene C, Yuh EL, Mukherjee P, Malik J. “Cost-sensitive active learning for intracranial hemorrhage detection” Medical Image Computing and Computer Assisted Intervention (MICCAI) Granada, Spain (Sept. 2018).

## 4. Early Prediction of Major Adverse Cardiovascular Events Using Remote Monitoring

#### *Publications*

- Early Prediction of Atrial Fibrillation-Atrial Flutter Using Remote Patient Monitoring Journal of Medical Cases Evaluating Utility and Compliance in Patient-Based Ehealth Study Using Continuous-Time Heart Rate and Activity Trackers. *Journal of the American Medical Informatics Association*.
- A Machine Learning Approach to Classifying Self-Reported Health Status in a cohort of Patients with Heart Disease using Activity Tracker Data. *Journal of Biomedical and Health Informatics*.
- Protocol, Design, and Methodology of a Remote Monitoring System in Ischemic Heart Disease: The Prediction, Risk, and Evaluation of Major Adverse Cardiac Events (PRE-MACE) Study. *Digital Medicine*.

#### *Press Releases*

- Cedars-Sinai Release: How Medicine and Technology May Protect Your Health Before Disease Strikes <https://www.cedars-sinai.edu/Precision-Health/How-Medicine-and-Technology-May-Protect-Your-Health-Before-Disease-Strikes.aspx>
- Cedars-Sinai Release: Fighting Heart Problems Before They Happen <https://blog.cedars-sinai.edu/fighting-heartproblems-before-they-happen/>
- Fitabase Release: Precision Medicine Research & Fitbit <https://www.fitabase.com/how-itworks/customers/cedars-sinai-brennan-spiegel/>
- Next Generation Patient Experience Release: Cedars-Sinai Medical Center and UCLA Researchers Use Fitbits to Aid Remote Patient Monitoring Programs <https://patientexperience.wbresearch.com/cedarsinai-ucla-researchers-fitbit-strategy-remotepatient-monitoring-ty-u>

#### *Presentations*

- Western Medical Research Conference, January 24-26, 2019, Evaluating Mitra® Microsampling Devices for Remote Monitoring of Apolipoproteins in Patients at Risk for Cardiac Events, Oral presentation, International Human Proteome Organization (Targeted Proteomics in Women’s Health), Orlando Florida, September 2018. Mitra® Microsampling Devices in Remote, Longitudinal Monitoring of Apolipoprotein B/Apolipoprotein A-I in Patients at Risk for Cardiac Events.
- 2018 American Society for Mass Spectrometry (ASMS) conference, San Diego, CA. Feasibility of Mitra® Microsampling Devices in Remote, Longitudinal Monitoring of Apolipoprotein B/Apolipoprotein A-I In Patients at Risk for Cardiac Events.
- 2018 MSACL (Mass Spectrometry Application to the Clinical Laboratory), Palm Springs, CA. Early Prediction of Cardiac Events via Remote Monitoring Coupled with an Apolipoprotein Targeted MRM Panels.

- 2019 MSACL (Mass Spectrometry Application to the Clinical Laboratory), Palm Springs, CA, March 2019. Relationship Between Patient-Reported Outcomes and Cardiac Biomarkers: The Prediction, Risk, and Evaluation of Major Adverse Cardiac Events (PRE-MACE)
- American College of Cardiology's 68th Annual Scientific Session, March 16-18, 2019. Quality Control Recommendations for MRM Proteomics on Dried Blood Spots Using Mitra® Microsampling Device
- 2019 American Society for Mass Spectrometry (ASMS) conference, Atlanta, GA Case-Control Study: Expanded Proteomics and Lipidomic Profiling for Early Prediction of Adverse Cardiac
- 2019 American Society for Mass Spectrometry (ASMS) conference, Atlanta, GA, Poster: Evaluation of Design: Early Prediction of Major Adverse Cardiovascular Event Surrogates Using Remote Monitoring with Biosensors, Biomarkers, and Patient-Reported Outcomes
- 2019 ASMS Conference, Atlanta, GA, Poster: Early Prediction of Major Adverse Cardiovascular Event Surrogates Using Remote Monitoring with Biosensors, Biomarkers, and Patient-Reported Outcomes: Mitra® Microsampling Device Compliance
- Mobilizing Medicine: Medicine and Me (and everyone else) Fusion Personalized Medicine Cancun, Mexico March 2017. Jennifer Van Eyk.
- Proteome centric precision health- changing medicine? Proteome Forum, Potsdam, Germany April 2017 Jennifer Van Eyk
- Post-translational modifications and cardiac disease. North American International Society of Heart Research (ISHR), New Orleans, June 2017 Jennifer Van Eyk.
- Changing the Course and Impact of Chronic Disease: Personalizing Medicine. Plenary Talk, International HUPO, Dublin, September 2017 Molecular and Cellular Proteomics award lecture. Jennifer Van Eyk.
- Personalized Medicine: medicine and me. American Association of Clinical Chemistry Meeting, San Diego, Oct 2017 Jennifer Van Eyk
- Lessons learnt in precision medicine. American Heart Association, Nov 2017 Jennifer Van Eyk.
- Mobilizing Medicine: Individualized surveillance and intervention of major adverse cardiac events (MACE), Journal of Proteomics Research Symposium on Innovation Pittcon 2018, Orlando, Feb 2018 Jennifer Van Eyk.
- Medicine and Me: Individualization of remote monitoring. Switzerland Proteomics Society, Keynote Talk, April 2018 Jennifer Van Eyk.
- Medicine and Me: A route to individualized monitoring and intervention. Keynote, Canadian Proteomics Society, April 2018 Jennifer Van Eyk.
- Changing the course and impact of chronic disease: personalizing medicine, Keynote Talk Jennifer Van Eyk.
- Medicine and Me: impacting chronic disease, Queenstown Molecular Biology Meeting, Keynote Talk, Queenstown, New Zealand, Sept. 2018 Jennifer Van Eyk.
- Personalized medicine? Atherosclerosis assessment via continuous patient-centric monitoring, Society for Heart and Vascular Metabolism Meeting, Keynote Talk, South Carolina, Sept 2018 Jennifer Van Eyk.
- Semi-automation of CESI-MS workflows: probing dynamic proteome. International CESI meeting. Plenary Talk. Amsterdam, Netherlands, Oct 2018.

- Proteomic Dissection of CVD: Mechanisms and Biomarkers, Session on Big Data. Keynote Talk Annual American Heart Association meeting, Chicago, USA. Nov 2018 Jennifer Van Eyk.
- Route to cardiology clinic: implication to personalization medicine. Proteome Society of India (PSI) meeting, Pune, India, Keynote Talk, Dec 2018. Jennifer Van Eyk.

## 5. Full Genome Analysis to Guide Precision Medicine

### *Publications*

- Manuscript in preparation to publish the methods and results of the study, including a description of the analysis pipeline.

## 6. Personal Mobile and Contextual Precision Health

### *Publications*

- The research team has documented the project at the architectural, protocol, and analytics levels and openly provides this work to any interest parties within academic medical centers.
- Publications of study results are currently in preparation.

### *Press Releases*

- UC Davis – Mobile health data may offer new pathways to better patient care, Dec 2016
- Release of full deidentified data sets and application code expected in 2019

### *Presentations*

- Anderson, N, Riedl, A, Dharmar, M, Gadgil, M, Lehman, M, Lindeman, D, Gillette, D, Pollock, B, Haddad, D, Sim, I, Linking Clinical Events to Patient Provided mHealth Data AMIA Summit, San Francisco, CA, March 2018
- Gadgil M, Sim, I, Anderson, N, Using Patient-Reported and Mobile Health Data in Practice: Focus on Hypertension and Depression, Society of General Internal Medicine, 2018, Denver, Colorado
- PERCEPT -Patient engaged digital Health. Clinical and Translational Science Digital Health Symposium, April 2017, Sacramento, CA

## 7. Precision Medicine for Early Prostate Cancer: Integrating Biological and Patient Complexity Variables to Predict Treatment Response

### *Publications*

- Abstract: “Association of race and socioeconomic status with prostate cancer genomic risk classifier: Implications for precision medicine in prostate cancer” at the American Society of Human Genetics conference in San Diego, California.
- Abstract: “Predicting 3-year Individual Patient Reported Outcomes in a Prostate Cancer Observation study: Fulfilling the aims of the IOM Comparative Effectiveness Research Report” for the 2019 Academy Health meeting in Washington, D.C.

#### *Presentations*

- Precision Medicine World Conference in Ann Arbor, Michigan

### **8. Precision Medicine for Multiple Sclerosis: Making It Work**

#### *Publications*

- Publication in preparation

#### *Press Releases*

- EHR Intelligence: Sutter Health awarded \$1.2 million by California Initiative to advance precision of care for multiple sclerosis (11/23/16)
- News Medical: Life Sciences - Sutter Health awarded \$1.2 million by California Initiative to advance precision of care for multiple sclerosis (11/23/16)
- Multiple Sclerosis News Today: Partnership Receives \$1.2M from California Initiative to Advance Precision Medicine for MS (11/28/16)

#### *Presentations*

- AMIA 2018 Informatics Summit: *Can Informatics Return the Joy of Medicine? A Real World Pilot in Neurology*, Joshua N. Liberman, PhD, James B. Jones, PhD, MBA, Joanna A. Cooper, MD, Walter F. Stewart, PhD, MPH (panel presentation & discussion)
- ECTRIMS 2017: R. Bove, E. Schleimer, P.-A. Gourraud, A. Lizee, J. VanLeeuwen, C. Bruce, J.B. Jones, K. Rankin, S. Sanders, W.F. Stewart, S. Hauser. Making precision medicine work at the bedside without a universal health record: the MS BioScreen experience. ECTRIMS Online Library. Bove R. Oct 10, 2018; 229248 (Abstract/Poster presented at ECTRIMS 2017)