Keeping Our Focus on COVID-19

With emergency use authorization of COVID-19 vaccines last December, it was expected that by now the pandemic would no longer be front page news. Instead, the world remains focused on research related to messenger RNA, and BMIR is part of that. On page 3 of this issue, you can read about some exciting research being performed by the Khatri lab. Their work contributed greatly to a recent study, published in Nature, that gave powerful support to the need for the current two-dose sequence of the Pfizer-BioNTech COVID-19 vaccine. The study was based on an interdisciplinary systems immunology approach to identify a global map of complex interactions between cell–cell, cell–environment, protein–protein, and protein–DNA interactions. The work, which required the Khatri lab to cross analyze vast amounts of collected data, is the epitome of what we do in BMIR.

Mark Musen, MD, PhD
Director, Stanford Center for Biomedical Informatics Research

BMIR Physician Data Scientist Is Developing Recommender Engines to Facilitate Medical Decision Making

After Jonathan Chen, MD, PhD, completed his medical training he realized the great variation in how patients are diagnosed and treated.

“A century ago, a single doctor could know it all and do it all because we didn’t really know how to treat that much. Since then, exponential advances in medical science have enabled us to take better care of people, but it has also led to tens of thousands of diagnoses, tests, and drugs. We are well past the point where the complexity of modern medicine exceeds the capacity of the unaided expert mind,” said the BMIR physician-scientist.

Because of that, Chen’s current research direction is focused on trying to discover and distribute the latent knowledge embedded in clinical data through informatics and computational tools.

“Medicine turns out to be more complex than you might expect. It has to do with the scarcest resource, access to professional expertise through a clinician’s time. I think that’s where informatics has a huge role to play. It’s really the only practical solution to the escalating overall complexity of medicine, and the only way to effectively deal with the iron triangle of health care cost, quality, and access,” he added.

Billions of people worldwide need access to medical expertise, not just in underdeveloped areas. Patients can have their care compromised while waiting months for an expert medical consultation, even in well-populated settings in industrialized countries.

“Not only that, but health care expenditures in the U.S. are more than double most other developed countries, and it’s only growing as a share of our economy,” Chen said.

His mission is to change that by developing fully digital consultation systems. One technique he is working on is recommender algorithms for medical decision making, which is based on technology that Amazon, Netflix, and other e-commerce and streaming services use to recommend purchases, books, or movies based on what “others like you” have enjoyed.

In Chen’s vision, physicians will use recommender engines that comb real-world clinical data to help them make key decisions based on steps other clinicians have taken with similar patients, empowering individuals with the collective experience of the many.

Chen is currently using his artificial intelligence algorithms in collaboration with Stanford’s Departments of Medicine and Pathology, and he expects to expand that work into value-based care initiatives, cost-saving reinvestment programs, and other areas.

In the near future he expects decisions based on recommender engines to be commonplace in exam rooms, helping physicians become better at what they already do and making medical practice a more consistent, universal experience for everyone.
Although her particular expertise is in data science, Summer Han, PhD, is motivated to work with investigators who are making an impact in clinical practice.

That became evident in 2015, when Han began a key collaboration with Oncology Division Chief Heather Wakelee, MD. In Wakelee’s lung cancer clinic were survivors who were developing a second primary lung cancer (SPLC). Because it was not clear who was at high risk for developing SPLC, Wakelee wanted help in developing strategies for advising long-term lung cancer survivors.

Han, a principal investigator in the Quantitative Sciences Unit (QSU), worked with Wakelee to develop a pioneering SPLC prediction model based on factors such as age, sex, race, treatment, histology, stage, and extent of disease. They evaluated the clinical utility of the model by calculating its net benefit in varied risk thresholds for screening and in 2017 published the results of a study in the *Journal of Clinical Oncology*.

The authors noted that the model did not include other key information on genetics, smoking, and other factors that could contribute to the prediction model for SPLC. That led Han to seek, and receive, a National Cancer Institute R37 MERIT Award for an Early Stage Investigator. The five-year grant is supporting research to identify the genetic, clinical, and environmental determinants for SPLC, to assess an individual's risk of developing SPLC, and to evaluate efficient lung screening strategies for SPLC to help inform the development of consensus screening guidelines for lung cancer survivors.

As a primary investigator running her own lab, she relies on the multidisciplinary strengths of biostatisticians, epidemiologists, and medical doctors on her team as she pursues studies with practical ramifications. As an example, Han had developed a mathematical model that she wanted to convert into a web-based risk assessment tool to aid clinical decision making for lung cancer patients and survivors. Results to date led to the publication of *Development and Validation of a Risk Prediction Model for Second Primary Lung Cancer* in the July 13, 2021 issue of the *Journal of the National Cancer Institute*.

In practice, Wakelee and other oncologists at Stanford Health Care plan to use this app in counseling prospective patients.

“I don’t want the results of my work to just sit in the literature. I want to be involved in activities that have an impact on clinical practice at Stanford and beyond,” she said.

Machine Learning Technique Leverages Unrelated Data to Study Rare Cancers

For scientists studying rare diseases, it is usually difficult to build good predictive models because limited sample sizes and highly selective features result in a shortage of data.

Olivier Gevaert, PhD, has shown that a technique used in machine learning can be used to overcome a shortage of data in rare cancers by relying on data from more common cancers.

Gevaert was among five authors of a December 2020 article in *Nature Communications* that demonstrated how a meta-learn-
The lab of Purvesh Khatri, PhD, collaborat-
ed on the first systems analysis of an mRNA
vaccine and shed light on the powerful
boost gained from the second dose of the
Pfizer-BioNTech COVID-19 vaccine.

Khatri and colleagues described a study that
was published in the July 12 issue of Nature.
The Pfizer mRNA vaccine belongs to a new
class of vaccines, and the research team
wanted to learn what molecular changes
occurred in the body with the vaccine. The
study was an outgrowth of their previous
work, published in Science, that the re-
searchers performed to model infection with
COVID-19, using the same technology and the same analysis methods, known as a
systems immunology approach.

Systems immunology is an interdiscipli-
ary approach that uses high-throughput
technologies and computational methods
that can be applied to identify a global map
of complex interactions between cell–cell,
cell–environment, protein–protein, and
protein–DNA interactions.

Following analysis of a large amount of data
generated, the researchers observed few in-
flammatory changes after a first dose of the
vaccine. But after a second dose, there were
10 to 100-fold increases in inflammatory
responses in specific immune cell types.

“That underscores the importance of not
skipping the second dose, because all the
changes that we observed were coming after
just 24 hours of receiving the second dose,”
Khatri said.

“We also observed a very profound biologi-
cal insight—that a specific subset of innate
immune cells, called monocytes, have mem-
ory too,” he said.

A subset of monocytes constituted only
0.01% of all circulating blood cells after
the first dose of the Pfizer vaccine. But after
the second dose of the Pfizer vaccine, their
numbers expanded 100-fold to account
for a full 1% of all blood cells. In addition,
their disposition became less inflammatory
and more intensely antiviral, which indi-
cates their ability to protect broadly against
diverse viral infections.

The research also showed no evidence of
auto antibody generation in any of the sub-
jects, which suggests that the vaccine does
not lead to the immune response attacking
the body itself.
OntoPortal Alliance Enables Use of BioPortal Technology in Many Application Areas

The Stanford Center for Biomedical Informatics Research (BMIR) uses advanced research techniques to discover, apply, translate, and organize data that make a difference for health and health care. With its expertise in clinical and translational informatics research and biostatistics, the division works to uncover new ways to advance personalized medicine and to enhance human health and wellness.

Collaboration is in our DNA. We are excited about the prospect of working with other experts who share our goal to connect data to health and medicine. We encourage you to contact Mark Musen, Director of BMIR (musen@stanford.edu), to learn more.