Welcome to Connection

This is the inaugural issue of a newsletter for the Stanford Center for Biomedical Informatics Research, or BMIR. At BMIR we connect data to health. That’s why we’ve named this publication Connection. Our goal is to keep all our friends, colleagues, and supporters informed about BMIR’s activities and identify opportunities for possible interaction and collaboration.

This newsletter is arriving at a difficult time for everyone. In this unprecedented situation, BMIR faculty members are working effectively and applying their skills to the COVID-19 pandemic, as you will read below.

A lot has been happening in BMIR outside of our response to COVID-19 as well. This issue reports on some of the highlights of the past few months.

We hope you find Connection informative and helpful. I encourage you to connect with us by visiting our website or contacting me directly at musen@stanford.edu. I also invite you to join us at an upcoming BMIR Research Colloquium.

Mark Musen. MD, PhD

Director, Stanford Center for Biomedical Informatics Research

Grant from Chan Zuckerberg Initiative Will Enhance BMIR’s Protégé Software

A team of biomedical informatics experts at BMIR is constructing the next generation of Protégé, a free, open-source ontology editor that provides a framework for building intelligent systems and controlled terminologies.

A grant from the Essential Open Source Software for Science program of the Chan Zuckerberg Science Initiative is supporting the Protégé work that is being led by Mark Musen, MD, PhD. The enhancements will make Protégé easier to maintain and extend as well as making it easier for third parties to contribute to the software's code base.

The Protégé platform, which BMIR launched as an open-source product in 1999, gives a growing user community a suite of tools for knowledge-based applications.

Today, more than 350,000 scientists worldwide use Protégé, and many users create plug-ins that add new functionality to the software.

COVID-19 Research

There is a lot of excitement about using data science and modeling to forecast how COVID-19 will spread, and BMIR is at the forefront of that work.

Jonathan Chen, MD, PhD, for example, is collaborating with the CDC to synthesize clinical trends in COVID-19 and related respiratory illness to inform national monitoring and policy goals by assimilating Stanford electronic health record data.

Nigam Shah, MBBS, PhD, is leading Stanford’s effort to continuously profile anonymized data from patients screened for SARS-CoV-2 to watch for shifting trends in presenting symptoms of patients, test positivity rates, the age distribution of positive cases and hospitalization rates as well as length of stay.

Manisha Desai, PhD, has pivoted the Quantitative Sciences Unit to launch new clinical trials as well as observational studies and quality improvement work. She and Tina Hernandez-Boussard, PhD, are collaborating on two other projects: predicting best strategies for scarce resource allocation during the COVID-19 pandemic and modeling outcomes of COVID-19 patients.
BMIR Team Uses Gene Signature to Aid in Earlier TB Diagnosis

Purvesh Khatri, PhD, and his team at BMIR have identified a three-gene signature that accurately diagnoses tuberculosis, predicts progression from a latent infection to active TB six months before clinical diagnosis, and reflects the response to treatment. The team of researchers directly compared published gene signatures for patients with active tuberculosis and evaluated their performance against the target product profiles from the World Health Organization and the Foundation for Innovative New Diagnostics.

The signature has been licensed for commercialization, and efforts are under way to develop a test to be administered at the point of care. To date, the genetic signature is the only method that can diagnose TB at such an early stage using a commercially available platform. Work is under way to move this research forward into clinical translation.

BMIR Team Develops Tool to Help with Lung Cancer Diagnosis and Prognosis

Andrew Gentles, PhD, and collaborators from BMIR and elsewhere at Stanford have developed a resource for determining the diagnosis and prognosis of patients with non-small-cell lung cancer (NSCLC).

Many existing therapies target lung cancer cells, but there are many downsides, which is what prompted this work.

While therapies can be effective, they can be imprecise, and healthy cells can get caught in the crossfire. Also, cells can become resistant to treatment if the cancer mutates over time. In addition, immunotherapies can have a dramatic effect for some patients, but these newer therapies still target a limited set of cells. Therefore, disrupting the “cross-talk” between different cell types could provide alternative avenues to attack cancer.

Dr. Gentles and his BMIR team worked with Stanford researchers Sylvia Plevritis, PhD, in Biomedical Data Science and Max Diehn, MD, PhD, in Radiation Oncology to identify and validate prognostically unfavorable associations between certain types of cells and prognostically favorable associations between other types of cells. That led to development of a resource for identifying many of these types of interactions, with the potential that some could be good candidates for therapeutic targets in lung cancer.

Study Uses Machine Learning to Help Reduce Unnecessary Tests

It’s estimated that up to half of all hospital lab tests are medically unnecessary, which is why BMIR’s Jonathan Chen, MD, PhD, led a study to systematically identify low information inpatient laboratory tests on an individual basis.

Dr. Chen and his multi-disciplinary team of researchers from Stanford, UCSF, and the University of Michigan used machine learning algorithms on electronic medical records to predict the results of hospital diagnostic tests. The goal was to develop a systematic way to identify tests that are unlikely to give new information before the tests are ordered.

As an example, Dr. Chen cites the hemoglobin A1c test, which provides a marker of elevated blood sugar. He says those tests are sometimes conducted so closely together that it’s physiologically impossible for the value to change.

“We used the best-performing machine learning models to analyze tens of thousands of patient records. That led us to predict normal results for dozens of the most common laboratory tests,” Dr. Chen says.

The findings suggest that low-yield diagnostic testing is common. The researchers generalized that machine learning algorithms could identify hundreds of thousands of potentially wasteful lab tests, which incur direct costs and may cause indirect harm to patients.
BMIR researchers have made significant contributions to efforts to curb two aggressive forms of cancer: glioblastoma and hepatocellular carcinoma.

Olivier Gevaert, PhD, and his BMIR lab partners are part of a multi-institutional team that developed AMARETTO, a data-driven platform for diagnostic, prognostic, and therapeutic decision-making in cancer.

The Gevaert Lab developed the initial algorithm for AMARETTO, which serves as a framework for software tools for network biology and medicine. In addition to using AMARETTO to find treatments for glioblastoma and hepatocellular carcinoma, scientists are using the platform in efforts to discover and validate novel drug targets in a wide range of other human cancers.

The BMIR group is collaborating with Nathalie Pochet, PhD, and her lab at Harvard and the GenePattern team led by Jill Mesirov, PhD, at UC-San Diego. AMARETTO is reaching a wider audience and is now available both as an R package, a GenePattern module and a Jupyter notebook, an open source web application.

Extensive functionality has been added to interpret the output of AMARETTO by providing links to molecular pathways, clinical and imaging phenotypes, and validation of the predictions using perturbation experiments from the Library of Integrated Network-Based Cellular Signatures (LINCS) Program of the NIH. Taken together, AMARETTO has now been enabled for use by a wide range of researchers—including basic and clinical scientists—at Stanford, Harvard, UCSD, and UC-San Francisco.

The research was the subject of “Informatics Technology for Cancer Research Program Drives and Fosters Community of Cancer Informatics Researchers: An *AMARETTO Tool Success Story,” which Dr. Pochet posted on the National Cancer Institute’s Cancer Data Science Pulse blog.

More information on AMARETTO, including a copy of Dr. Pochet’s blog post, is available by contacting BMIR at contact-BMIR@stanford.edu.

Using artificial intelligence, Tina Hernandez Boussard, PhD, and her team at BMIR have found that patients taking the most common form of antidepressants experience less pain relief from the most widely prescribed opioids.

The researchers developed a machine-learning approach that predicts patients’ post-operative pain at both hospital discharge and eight weeks later. They found that selective serotonin reuptake inhibitors (SSRIs), a common form of antidepressant, reduced the effectiveness of the most frequently prescribed opioids such as hydrocodone and codeine.

“This work is part of our search for ways to combat the opioid epidemic by managing pain while reducing opioid exposure,” Dr. Boussard said.

The research received support from two prestigious sources. The National Institutes of Health National Library of Medicine awarded an R01 grant to Dr. Boussard, and the Digital Health Cooperative Research Centre in Australia gave additional funding.

The activities of the Boussard Lab are part of BMIR’s overall efforts to unlock the value of the unstructured data in electronic health records so the data can be used for predictive analytics and potential translation for clinical care.

An article on the research, “Predicting Inadequate Postoperative Pain Management in Depressed Patients: A Machine Learning Approach” was published in *PLOS ONE* and is available by contacting BMIR at contact-BMIR@stanford.edu.
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.