



**Date:** 5 October 2012

## **Headline:** Philip Morris International, IBM Announce IMPROVER Systems Biology Verification Challenge Winners

**Byline:** Uduak Grace Thomas

This week, Philip Morris International and IBM, organizers of the Industrial Methodology for Process Verification in Research, or IMPROVER, initiative, announced the winners of the first of a series of planned challenges aimed at developing a transparent and robust process for verifying systems biology methods used in commercial settings.

This first challenge, called the Diagnostic Signature Challenge, was to assess and verify computational methods used to classify clinical samples based on gene-expression data in psoriasis, multiple sclerosis, lung cancer, and chronic obstructive pulmonary disease.

The overall best performer was a team from Wayne State University; the Center for Computational and Systems Biology, or CoSbi, a joint venture between the University of Trento and Microsoft Research, placed second overall; and ETH Zurich placed third.

The CoSbi team was also the best performer in the multiple sclerosis sub-challenge; researchers from Israel's Bar-Ilan University scored top marks in the lung cancer challenge; a team from the Institute of Systems Biology performed best in the psoriasis challenge; and scientists from University of California, Los Angeles, took home the award for the COPD challenge.

Each of the teams presented their work during a two-day symposium held in Boston earlier this week.

According to the organizers, about 54 teams participated in the Diagnostic Signature Challenge.

"It has been wonderful to see how IMPROVER has captured the attention of the scientific community," Manuel Peitsch, PMI's vice president of biological systems research, said in a statement. "I believe we are off to a great start for the IMPROVER project and we hope to see many more exciting developments in the months and years to come."

Ajay Royyuru, director of the computational biology center at IBM's Watson Research Center, pointed out that "there is an inherent difficulty in achieving verification in the life sciences as the data is always complex and methods available for reviewing, reproducing, and verifying have limitations."

However, IMPROVER “marks an important step towards developing an accepted process by which scientific inferences can be confidently verified,” he said.

John Quackenbush, a professor of computational biology and bioinformatics at the Dana-Farber Cancer Institute and one of the meeting’s keynote speakers, noted in a statement that “projects like IMPROVER are important because they challenge scientists to identify systems biology methods that are best able to make meaningful predictions and, hopefully, improve outcomes for patients.”

IMPROVER was launched in March, building on similar challenge initiatives in the academic world — such as the IBM-led DREAM (Dialogue on Reverse Engineering Assessment and Methods) project and the CASP (Critical Assessment of Protein Structure Prediction) initiative — but intended to focus on the verification of processes that would be of use in the industrial setting (BI 3/9/2012).

DREAM, for example, tries to “ask questions that are ...important for the specific data that are being generated in academia [and] that are of intellectual value across the spectrum in academia in systems biology,” IBM’s Stolovitzky explained to BioInform.

DREAM lacks “an organic view of research pipelines as, for example, there are in industry,” where companies map out pipelines that are unique to their products and usually include several “assumptions,” he said.

IMPROVER’s aim is to “de-risk?” each of these assumptions “by verifying [its] validity using challenges,” Stolovitzky said.

IMPROVER’s challenges “are designed in a way that they become increasingly complex, they build on one another,” Peitsch told BioInform. “In other words, to do the second challenge, you probably need to be pretty good at the first challenge.”

The overall aim of the Diagnostic Signatures Challenge was to verify whether it is possible to extract useful diagnostic signatures for each of the four diseases, the organizers explained.

Participants trained their methods on existing public data from resources such as the Gene Expression Omnibus prior to applying them to the test dataset, which was developed independently, Stolovitzky explained.

Generally speaking, “we had a spectrum of performances,” he said. “In one of the challenges, we observed that we couldn’t verify our basic hypothesis that transcriptomics was enough for phenotype. In the other ones, basically, we did verify that it is possible.”

For example, in the case of psoriasis, “there was a team that had a perfect prediction,” he said. On the other hand, it was more difficult to predict whether a patient with multiple sclerosis was relapsing or in remission, he said.

It’s also worth noting that teams “took very different approaches” to answering the questions posed by the challenge, according to Hugh Browne, R&D scientific spokesperson for PMI.

Some adopted “a purely data-driven approach,” while others had “a more mechanistic view,” and a third category combined both methods, he said. This suggests that “a combination of different methods that takes on board the biology but also the computational aspects” is probably the best way to address these challenges.

IMPROVER's organizers already have plans for future challenges in the works.

Next up will be the Species Translation Challenge — expected to launch in the second quarter of 2013 — which is intended to explore the predictive value of data from different species in order to improve current understanding of biological processes.

Specifically, participants will have to identify a function that maps measurements derived from systematic perturbations in one species to another, as well as evaluate translatability between species — human and rat in this case.

Peitsch explained that contestants will be provided with intracellular response data collected at different time points from human and rat bronchial epithelial primary cells that will be treated with roughly 60 different drug molecules.

“The challenge will be a combination of questions,” he said. “We will give [participants] part of the rat data [and] part of the human data and ask them to basically predict the other species' response.” More information about the challenge will be provided at a later date as the exact details are still being worked out, PMI's Browne said.

Looking even further ahead, “the last challenge will be a multifaceted challenge around COPD and the intention there is to share with the community [the] animal data [and] clinical data as well as a ... verified biological network that describes the mechanisms of COPD in both the human and its cognate emphysema in mice,” Peitsch said.

Meanwhile, the best performing methods in each category, as well as lessons learned from IMPROVER's first outing will be published in a special edition of the Systems Biomedicine journal at a later date as well as through other yet-to-be determined outlets, Stolovitzky said.

-- ENDS --