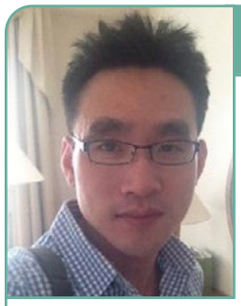
GENOMIC MARKERS IN CLINICAL DEVELOPMENT



HYUNJIN (GENE) SHIN, Principal Scientist, Takeda

Dr. Hyunjin (Gene) Shin is a principal scientist in the Translational Bioinformatics/Computational Biology Group at Takeda. Dr. Shin currently leads a number of preclinical and clinical projects related to biomarker identification and predictive modeling for patient stratification/precision medicine. He is also an expert in research of disease characterization and potential drug target identification using biomedical big data and machine learning. Gene received his B.S. in electrical engineering at Seoul National University, Korea in 1998. He then completed M.S. and Ph.D. in electrical and computer engineering at the University of Texas at Austin. In 2007, Gene joined Dr. X. Shirley Liu's lab at Dana-Farber Cancer Institute/Harvard School of Public Health as a postdoctoral research fellow. After his postdoctoral training, he moved to Takeda and since then, he has been working as a lead computational biologist and data scientist at the organization, supporting many translational research projects of developing therapeutics for cancer and other diseases.

How could genomic markers impact clinical development?

Developing a new drug is challenging and takes a lot of time, and we still have a very small chance of success. Biomarker development can be one solution for that. By enriching likely responders in the clinical trial, we could increase the chance of success through a biomarker ID. Genomics is important in oncology because cancer is a genetic disease. Therefore, I believe genomics data has the most useful information about a disease and drug response.

What are the challenges of integrating the genomic data into the workflow?

Different genomic data sets, for example, somatic mutation and gene expression, confer us very useful information about disease progression or drug response. However, I don't think we have a full understanding of the characteristics of each individual data type; somatic mutation and gene expression data provide different views on the molecular mechanisms underlying a disease and drug.. If we want to put them together in a single framework, we need to be careful about, for example, data normalisation, the pre-processing and proper computational framework - just trying to avoid any potential errors or bias in terms of the data integration.

How can machine learning be utilised for clinical development?

Machine learning has been widely used for many areas in biomedical and health sectors. Machine learning can help us identify hidden patterns. For example, if you want to do some patient stratification using a machine learning model using genomics data, the human perception and the abilities are limited, but machine learning helps us to find very subtle patterns even from very notoriously complex data sets. Therefore, machine learning is very important and useful for clinical development. However, we should also consider the regulatory paths of using machine learning. For example, I do not think the FDA currently have very clear guidelines for clinical development using machine learning.

What are currently the most important technologies impacting biomarker research?

I am mostly talking about oncology, as that is what my work mostly relates to. ctDNA is a revolutionary technology, allowing us to non-invasively measure mutation changes over the therapeutics. It can give us some ideas about longitudinal development of drug resistance; thus, it may lead us to a better decision on drug combination. Another, single Cell RNA Sequencing becomes increasingly important. Nowadays we have a growing number of clinical development programmes for immuno-oncology. In immuno-oncology, the ultimate goal of molecular profiling is to have gene expression patterns from multiple types of cells; and single cell RNA sequencing actually allows us to do that.

What are the three main takeaways of your presentation?

I think some people saw the possibility for rescuing unsuccessful clinical trials through patient stratification based on biomarker ID. There are a lot of clinical trials where the primary and secondary end point were not met. This time, I presented the biomarker identifications through machine learning; and this type of work can help us rescue some clinical trials by patient selection or a patient enrichment. Secondly, I could also show the pros and cons and the real face of clinical biomarker development, based on machine learning. I made sure to remind people of the hurdles and the challenges that we must address in order to use machine learning for clinical biomarker development.

Thirdly, data integration is very challenging. I presented some work related to data integration of multiple types of genomics data. While the work was successful, we still have a long way to go. We need to learn more.

For more Information on the Biomarkers Series US **CLICK HERE**