Visualizing Healthcare System Dynamics in Biomedical Big Data

Griffin Weber, Harvard Medical School; *Nicholas Benik, Harvard Medical School*; *Katy Borner, Indiana University*; *Nicholas Brown, Harvard Medical School*; *Daniel Halsey, Indiana University*; *Isaac Kohane, Harvard Medical School*; *Daniel O'Donnell, Indiana University*

Data in observational databases, such as electronic health records and administrative claims records, reflect both the pathophysiology (PP) of patients as well as their interactions with the healthcare system, which we call Healthcare System Dynamics (HSD). For example, the result of a laboratory test, such as a white blood cell (WBC) count, is a measure of the patient; but, the test was performed at that particular date and time because the patient visited her clinician, and the clinician determined that it was necessary to order the test. Investigators typically focus only on the test result, but the HSD component of the data often has greater predictive power. For example, counter intuitively, patients with an abnormal WBC test result at 3pm have higher three-year survival rates than patients with a normal WBC test result at 3am. This is because clinicians usually do not order laboratory tests at 3am unless they think the patient is very sick. The time of day of a clinical encounter, the day of the week, the amount of time since the last encounter, and total amount of data that has been recorded about a patient (the patient's "fact count") are just a few of the HSD concepts that are important in predicting different types of patient outcomes. In this BD2K Targeted Software Development project, we created an ontology and data visualization plugins for a widely used clinical research platform, i2b2, to help investigators learn about HSD and incorporate it into their studies. More broadly, this project illustrates, though data visualizations, how understanding the processes that lead to the generation of biomedical big data can be just as important as the data themselves.