

# **Center for Causal Modeling and Discovery of Biomedical Knowledge from Big Data**

**The University of Pittsburgh at Pittsburgh**

PIs: Gregory F. Cooper, Ivet Bahar, and Jeremy Berg

Grant Number: 1-U54HG008540-01

Much of science consists of discovering and modeling causal relationships that occur in nature. Increasingly big data are being used to drive such discoveries. There is a pressing need for methods that can efficiently infer causal networks from large and diverse types of biomedical data and background knowledge. This center of excellence will develop, implement, and evaluate an integrated set of tools that support causal modeling and discovery (CMD) of biomedical knowledge from very large and complex biomedical data. We also plan to actively share our knowledge, methods, and tools with others, through an innovative set of training and consortium activities. In the past 25 years, there has been tremendous progress in developing general computational methods for representing and discovering causal knowledge from data, based on a representation called causal Bayesian networks (CBNs). These methods have been applied successfully in a wide range of fields, including medicine and biology. While much progress has been made in the development of these computational methods, they are not readily available, sufficiently efficient, nor easy to use by biomedical scientists, and they have not been reconfigured to exploit the increasingly Big Data available for analysis. This Center will make these methods widely available, highly efficient when applied to big datasets, and easy to use. The proposed Center will provide a powerful set of concepts and tools that accelerate the discovery and sharing of causal knowledge derived from very large and complex biomedical datasets. The approaches and products emanating from this center of excellence are likely to have a significant positive impact on our understanding of health and disease, and thereby on the improvement of human health. PUBLIC HEALTH RELEVANCE: This center of excellence will develop, implement, and evaluate an integrated set of tools that support causal modeling and discovery (CMD) of biomedical knowledge from very large and complex biomedical data. The approaches and products emanating from this center of excellence are likely to have a significant positive impact on our understanding of health and disease, and thereby on the improvement of human health.