**SIGNIFICANCE.**

Population heterogeneity is a significant obstacle to cancer treatment. Patients with the same cancer-of-origin can vary greatly in the molecular mechanism of carcinogenesis and prognosis. Precision medicine holds promise to improve treatment decisions, and ultimately prognosis for each patient. Classifying cancer patients based on their molecular profile at different omics levels reveals comprehensive insights into precision medicine. There is a body of literature on classifying cancer patients by high throughput omics data, but methodology is still lacking for clustering cancer patients using multi-omics data while taking into account the potentially critical interactions among omic measures. *Our research will address this gap by developing the first methodology to identify cancer subtypes in the light of cancer molecular profile including mutation status, mRNA expression, methylation level, microRNA expression level, protein expression level, and their interactions.*

*This contribution will be significant because:*

* Increasingly available omics data provide unprecedented opportunity to identify the cancer subtypes. The proposed method, MGMCluster, can integrate any omics data to comprehensively classify cancer subtype.
* Successful classification of cancer patients is the corner stone of precision medicine. Comparing to other methods that classify cancer subtypes by integrating multi omics data, MGMCluster identifies cancer subtypes in a more accurate manner since it examines not only the status of the molecular entities of all omic measures but also the interactions among the entities of the same and different omic measures.
* Discovering similarities between subtypes of different cancers has previously enabled the repositioning of therapies. Our method will potentially open up a large number of new candidates for this approach.
* MGMCluster will provide knowledge of the heterogeneity and the different molecular causes of cancers.

At the completion of this study, we will provide open access software to the research community. This software will allow researchers to identify the subtypes of their patients by integrating their own omics data. Correct identification of cancer patient subtype will enable precision medicine. Moreover, MGMCluster will shed light on the common molecular causes of the same cancer subtypes.

**INNOV ATION.**

Most current cancer subtype clustering approaches only utilize a single type of omics data. Patients with similar profile at some omics levels but distinct profiles at other omics level may be mistakenly clustered together. Only a few methods are available for integrating multi omics data to cluster cancer subtypes. However, none of these integrative methods takes the interaction between omic measures into consideration.

*The proposed research is innovative because:*

* This is the first research that studies the cancer subtypes based on both interactions among the molecular entities of the same and different omic measures and the status of all entities.
* Our method incorporates the first graphical model that captures the joint distribution of multi omics data. This state of the art mixed graphical model allows us to model the status of each molecular entity as well as the interactions among them.
* Repurposing existing cancer treatment to treat some subtypes of other cancers that share similar molecular profile provides a new perspective of drug repositioning.

The novel application of Mixed Graphical Models in clustering empowers integrative identification of cancer subtypes. Our novel method, MGMCluster, extends the definition of cancer patients’ similarity from the conventional static molecular profile similarity to similar static molecular profile and dynamic interactions. This in turn facilitates cancer subtype clustering by deeper understanding of carcinogenesis and pathology(?).