

# Methods to Predict Patient Survival Based on Gene Expression Data

Eric Bair, Ph.D.  
University of California

## **Abstract:**

An important goal of DNA microarray research is to develop tools to diagnose cancer more accurately based on the genetic profile of a tumor. There are several existing techniques in the literature for performing this type of diagnosis. Unfortunately, most of these techniques assume that different subtypes of cancer are already known to exist. Their utility is limited when such subtypes have not been previously identified. Although methods for identifying such subtypes exist, these methods do not work well for all data sets. It would be desirable to develop a procedure to find such subtypes that is applicable in a wide variety of circumstances.

Even if no information is known about possible subtypes of a certain form of cancer, clinical information about the patients, such as their survival time, is often available. In this study, we develop some procedures that can utilize both the gene expression data and the clinical data to identify subtypes of cancer and use this knowledge to diagnose future patients. These procedures were successfully applied to several publicly available data sets.

We will present diagnostic procedures that can accurately predict the survival of future patients based on the gene expression profile and survival times of previous patients. This has the potential to be a powerful tool for diagnosing and treating cancer.