

在另一篇文章，Spencer等人利用基于机器学习数据挖掘的方法，通过追溯性分析治疗方案的效果，建立了一个肺损伤的预测模型。此外，这种模式中的生物标记提取自预期的临床试验，并且在多个时间收集血清样本。他们利用一种3-way蛋白质组学方法来筛选差异表达以及与放射性肺炎（radiation pneumonitis, RP）相关的蛋白质。他们提出用蛋白质组学方法来鉴定RP预测的生物标记，可以消除追溯物理模型中的信息缺口。Wang等人的文章基于线性规划和利用Kappa统计及模糊聚类整合来分析大的功能簇，从而构建了一个单一的基因网络。最后，在Loganantharaj和Chung的文章中，他们介绍了一种综合蛋白质-蛋白质间的信息、途径信息和阵列分析数据集之间的联系，以鉴定重要基因和潜在信号转导网络的方法^[3]。

第三类由两篇涉及医疗信息和时空数据分析进展的文章组成。在他们的文章中，Chen等人开发了癌症病人的病情预测系统。该系统可更准确地用于预测结果或癌症病人的存活率。Song等人的文章提出了另一个方法，可以分类基因短暂表达的曲线。这个曲线中，个体表达轨迹是纵向数据，同时有一个多变的方差和协方差。这个方法主要基于广义的混合模型，并被细菌的基因表达所例证^[3]。



八、展望

癌症作为一种表观遗传失调，可以由不同的遗传和蛋白质谱来描述，这与癌症的发生和进展相联系。一些修饰，如基因组改变的表达模式、基因组失稳和体细胞突变，可以被用来区分正常细胞和癌细胞，并且这些知识可以专门应用于诊断和治疗癌细胞。在这种情况下，生物信息学方法不仅是一个生物标记识别的重要工具，而且也促进了对整个肿瘤发展进程的理解。

“组学”科学的发展涉及大量不同来源的数据整合。现在，从高通量的基因、基因变异、基因表达和蛋白质组学得来的数据，不断积累在公共数据库中，而且作为必然结果，生物信息学技术也将不断发展，且能够帮助检查这些复杂的过程和网络数据集。因此，生物信息学分析可用于处理数据整合，并且可提供很多关于癌症的信息。目前，它已经用于癌症的临床实践，如癌症的诊断和治疗。生物信息学的建立就是该方面定量研究的自然结果。

根据Pusztai与其合作者的说法，上述方法在临床应用之前，科学家和临床医生将需要克服一些挑战。虽然癌症的遗传表达谱能够促进癌症的个性化治疗，但初步的观察资料将需要得到验证，并且癌症潜在靶标的真实性也将需要验证。正如作者所说：“比较药物的临床发展和推测性的诊断标记与现在发表的实验结果都是一些临床标记发现的实验资料，它们正处于令人鼓舞的I/II阶段，这说明它们的临床应用和评估这些标记的准确度还需要更复杂的、更随机的阶段III标记验证。”

很明显，生物信息学研究在癌症的应用方面仍处于早期阶段，尽管有很多困难，在后基因组时代，它已经成为一个必需的技术，可以协助和改善癌症治疗。我们预计，新的技术进步将对我们治疗癌症做出巨大的贡献。