

中国药科大学本科生毕业论文（设计）开题报告

学生姓名	卜洋洋	学 号	2020210606	专 业	应用统计学（生物统计方向）
指导教师姓名	房婧雅	指导教师职称	讲师	实习单位名称	中国药科大学理学院生物统计系
课题名称	复杂疾病个体特异性因果网络构建及临界点预测				
课题性质	<input checked="" type="checkbox"/> 基础研究 <input type="checkbox"/> 应用课题 <input type="checkbox"/> 设计型 <input type="checkbox"/> 调研综述 <input type="checkbox"/> 理论研究				
开题报告内容：（包括拟研究或解决的问题、采用的研究手段及文献综述，不少于 2000 字）					
<p>一 研究背景与文献综述</p> <p>复杂疾病（如流感、癌症和神经系统疾病等）的发生和发展往往伴随着多尺度的生物学动态变化，呈现出高度的异质性和非线性演进特征，即存在从健康状态或亚健康状态向病理状态的急剧突变点。识别这种关键点对于疾病的早期干预和精准医疗至关重要。</p> <p>近年来，随着组学技术的发展，大量计算方法被用于监测复杂疾病的动态临界点。Chen 等^[1]提出了动态网络生物标志物（Dynamic Network Biomarker, DNB）方法，该方法基于拓扑网络变化，能够在疾病突变点到来之前提供早期预警信号。研究表明，即便在样本量较少的情况下，DNB 方法仍然能够通过高通量数据检测关键状态的转变。该方法已在多个疾病模型中得到验证，并成为关键状态检测领域的重要工具。Liu 等^[2]进一步扩展了 DNB 理论，提出了一种基于单一样本的关键状态检测方法。这一方法能够从个体水平识别疾病前状态，即在症状尚未显现时预测疾病进展，为精准医疗提供了新的视角。除了 DNB 方法之外，Chen 等^[3]采用隐马尔可夫模型（Hidden Markov Model, HMM），通过状态转移概率识别复杂生物系统的相变临界点。Chen 等^[4]利用时间差分网络（Temporal Differentiation Networks）分析生物系统在突变点前后的网络变化结构。他们发现，随着疾病进展，分子网络的拓扑结构会发生显著的动态重构，而这一特性可以作为预测疾病恶化的重要信号。Liu 等^[5]提出了一种景观动态网络生物标志物（landscape Dynamic Network Biomarker, <i>l</i>-DNB）方法，该方法结合分子网络的稳态景观分析，能够在单样本水平上检测复杂疾病的关键状态。他们的研究结果显示，<i>l</i>-DNB 方法不仅能够预测疾病突变点，还能够识别影响疾病进展的核心基因模块。Liu 等^[6]进一步改进隐马尔可夫模型，实现了基于单样本的隐马尔可夫模型，这一方法克服了单样本无法精准识别疾病前状态的问题。Han 等^[7]提出了一种基于单样本节点熵（Single-Sample Node Entropy）的动</p>					

态网络模型，用于检测癌症在恶化前的关键状态。该方法利用蛋白相互作用网络，结合节点的熵值计算，识别癌症进展中的关键基因，并预测癌症恶化的临界点。Zhong 等^[8]提出了一种有向网络排序评分（Directed-network Rank Score），这一方法结合网络动力学与统计排序，能够更精准地捕捉生物系统的突变点，在细胞命运分析中展现了较高的准确性。与此同时，Zhong 等^[9]还引入了样本特异性网络熵（Sample-Specific Causality Network Entropy, SCNE）方法，通过推断个体水平的因果关系网络，计算关键状态的变化情况。SCNE 方法在单细胞数据分析中表现优异，尤其适用于高维、小样本数据集。

上述这些计算方法在特定疾病当中具有应用价值。Chen 等^[10]将 DNB 方法应用于乳腺癌 MCF-7 细胞分化的研究中，成功识别了乳腺癌细胞的早期分化状态。Chen 等^[11]采用 DNB 方法预测流感爆发，并基于 2009-2016 年日本病例数据验证了其有效性。研究结果表明，DNB 方法可以在流感大规模传播前提前 4 周发出预警信号，为公共卫生部门制定应对策略提供了宝贵的信息。Liu 等^[12]研究了乳腺癌的内分泌耐药性，并利用 DNB 方法检测耐药发生的关键状态。他们发现，某些基因突变在乳腺癌细胞获得耐药性之前就已发生，而 DNB 方法能够准确预测这一转变过程，为个性化药物调整提供了数据支持。Sun 等^[13]将 *l*-DNB 方法应用于肝癌进展的临界状态检测。研究发现，低级别和高级别肝硬化是肝癌发展的关键节点。这一发现有助于优化肝癌的早期干预策略，提高患者的生存率。

二 研究目的与意义

研究目的：本课题旨在构建复杂疾病的个体特异性因果网络，并预测疾病恶化的临界点，以揭示复杂疾病发生与发展的关键动态机制。通过单细胞转录组学（scRNA-seq）数据结合蛋白质-蛋白质相互作用（PPI）网络，利用统计因果推断技术推导个体水平的动态分子网络，解析基因表达的波动性及其因果关系变化。基于此，识别疾病从稳定状态向病理状态演进的关键节点，筛选动态网络生物标志物（DNB），为疾病的早期诊断、精准干预和个体化治疗提供新的理论支撑和技术工具。

研究意义：

1. 促进复杂疾病精准医学的发展

目前，复杂疾病（如流感、癌症和神经系统疾病等）研究大多停留在群体水平，忽视了个体间的生物学差异。本研究通过构建个体特异性因果网络，从个体层面解析疾病动态变化规律，

有助于精准识别个体化的疾病进展模式，为精准医疗提供数据支撑。

2. 拓展疾病早期预测与干预的新策略

传统的疾病研究主要依赖静态基因表达分析，难以捕捉疾病发展的关键转折点。本研究基于单细胞组学数据，结合因果推断方法，能够动态监测疾病状态的变化，并识别疾病恶化的临界点，为早期预测和个体化干预提供新思路。

3. 推动生物网络建模和因果推断技术的发展

本研究结合因果推断、动态网络分析及多组学数据融合，构建个体特异性分子因果网络，不仅拓展了网络生物学在复杂疾病研究中的应用，也为系统生物学、网络医学等领域提供新的方法论支持。

4. 发现关键动态生物标志物，优化个体化治疗方案

通过挖掘疾病恶化过程中的关键基因和分子网络，本研究能够筛选出具有潜在临床价值的动态生物标志物，为靶向治疗、药物开发及个性化诊疗方案设计提供科学依据，推动疾病精准防治的发展。

三 研究内容与方法

本课题围绕“复杂疾病个体特异性因果网络构建及临界点预测”展开，结合单细胞转录组学（scRNA-seq）数据，蛋白质-蛋白质相互作用（PPI）网络、统计因果推断方法和动态网络分析，探索个体化疾病进展的分子机制，并识别疾病恶化的临界点。本研究预采用一个模拟数据集与五个真实数据集：利用微分方程生成基因调控网络，主要评估统计方法的有效性与稳定性；以流感感染单细胞数据（GEO: GSE30550），结直肠癌单细胞数据（GEO: GSE161277），CPTAC 数据库中胶质母细胞瘤单细胞数据作为癌症与神经系统疾病的建模数据集；以 TCGA 数据库中结直肠癌（COAD/READ）和胶质母细胞瘤（GBM）转录组数据作为模型验证数据集。具体研究方法与技术路线如下：

step 1. 获取上述单细胞数据集，数据来源于 GEO 数据库（<https://www.ncbi.nlm.nih.gov/geo/>）与 TCGA 数据库（<https://portal.gdc.cancer.gov/>）；

step 2. 数据预处理

- 质量控制：去除低质量细胞（如基因检测数过低或者线粒体基因表达比例过高的细胞）；
- 归一化与标准化：采用 log-normalization 或 SCTransform 进行数据标准化处理；
- 降维分析：采用 PCA（主成分分析）或者 UMAP/t-SNE（非线性降维）进行数据

降维，以便可视化和后续分析；

- 细胞类型注释：利用已知的细胞标志物（如 SingleR、CellMarker）进行自动化或人工注释，以明确不同细胞群体的功能特性；

step 3. 个体特异性网络构建

- PPI 网络构建：通过 SRING、BioGRID、HPRD 等数据库获取 PPI 信息，以构建并增强基因调控网络；
- 因果推断建模：采用贝叶斯网络、格兰杰因果模型、因果结构学习等方法推导基因之间的因果关系，计算个体特异性因果网络，评估因果关系可信度；

step 4. 关键动态网络生物标志物筛选

- 基因表达变异性分析：计算单细胞水平的基因表达噪声，识别关键波动基因；
- 网络中心性分析：计算网络中心性指标，识别在因果网络中占主导地位的基因；
- 动态网络生物标志物检测：采用 DNB 方法，识别疾病进展过程中的突变点；

step 5. 疾病临界点预测

- 隐马尔可夫建模：采用 HMM 预测疾病从稳定状态到病理状态的转变点；
- 时间序列分析：观察关键基因表达在疾病进展过程中的变化趋势；

step 6. 生物学验证

- 生存分析：在验证数据集上分析关键生物标志物在患者生存预测中的价值；
- GO/KEGG 功能分析：分析关键生物标志物的生物学功能

四 预期结果

本研究将突破传统基于群体水平的复杂疾病研究模式，构建精准、高分辨率的个体特异性因果网络，实现对疾病动态演变的深度解析。通过单细胞组学数据整合、因果推断建模与网络动态分析，本研究不仅有望揭示复杂疾病的核心分子机制，还将为个体化诊断、疾病早期预警及精准治疗提供科学依据。

参考文献

- [1] CHEN L, LIU R, LIU Z P, 等. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers[J/OL]. Scientific Reports, 2012, 2(1): 342. DOI:10.1038/srep00342.
- [2] LIU R, YU X, LIU X, 等. Identifying critical transitions of complex diseases based on a single sample[J/OL]. Bioinformatics, 2014, 30(11): 1579-1586. DOI:10.1093/bioinformatics/btu084.
- [3] CHEN P, LIU R, LI Y, 等. Detecting critical state before phase transition of complex biological systems by hidden Markov model[J/OL]. Bioinformatics, 2016, 32(14): 2143-2150. DOI:10.1093/bioinformatics/btw154.
- [4] CHEN P, LI Y, LIU X, 等. Detecting the tipping points in a three-state model of complex diseases by temporal differential networks[J/OL]. Journal of Translational Medicine, 2017, 15(1):

217. DOI:10.1186/s12967-017-1320-7.
- [5] LIU X, CHANG X, LENG S, 等. Detection for disease tipping points by landscape dynamic network biomarkers[J/OL]. National Science Review, 2019, 6(4): 775-785. DOI:10.1093/nsr/nwy162.
- [6] LIU R, ZHONG J, YU X, 等. Identifying Critical State of Complex Diseases by Single-Sample-Based Hidden Markov Model[J/OL]. Frontiers in Genetics, 2019, 10: 285. DOI:10.3389/fgene.2019.00285.
- [7] HAN C, ZHONG J, HU J, 等. Single-Sample Node Entropy for Molecular Transition in Pre-deterioration Stage of Cancer[J/OL]. Frontiers in Bioengineering and Biotechnology, 2020, 8: 809. DOI:10.3389/fbioe.2020.00809.
- [8] ZHONG J, HAN C, WANG Y, 等. Identifying the critical state of complex biological systems by the directed-network rank score method[J/OL]. Bioinformatics, 2022, 38(24): 5398-5405. DOI:10.1093/bioinformatics/btac707.
- [9] ZHONG J, TANG H, HUANG Z, 等. Uncovering the Pre-Deterioration State during Disease Progression Based on Sample-Specific Causality Network Entropy (SCNE)[J/OL]. Research, 2024, 7: 0368. DOI:10.34133/research.0368.
- [10] CHEN P, LIU R, CHEN L, 等. Identifying critical differentiation state of MCF-7 cells for breast cancer by dynamical network biomarkers[J/OL]. Frontiers in Genetics, 2015, 6[2025-02-26]. <http://journal.frontiersin.org/Article/10.3389/fgene.2015.00252/abstract>. DOI:10.3389/fgene.2015.00252.
- [11] CHEN P, CHEN E, CHEN L, 等. Detecting early-warning signals of influenza outbreak based on dynamic network marker[J/OL]. Journal of Cellular and Molecular Medicine, 2019, 23(1): 395-404. DOI:10.1111/jcmm.13943.
- [12] LIU R, WANG J, UKAI M, 等. Hunt for the tipping point during endocrine resistance process in breast cancer by dynamic network biomarkers[J/OL]. Journal of Molecular Cell Biology, 2019, 11(8): 649-664. DOI:10.1093/jmcb/mjy059.
- [13] SUN Y, ZHAO H, WU M, 等. Identifying critical states of hepatocellular carcinoma based on landscape dynamic network biomarkers[J/OL]. Computational Biology and Chemistry, 2020, 85: 107202. DOI:10.1016/j.compbiolchem.2020.107202.

学生签名： 年 月 日

指导教师意见：

指导教师签名： 年 月 日

实习单位意见：	
负责人签名：	年 月 日

年 月 日

填写说明

1.指导教师意见：由毕业论文（设计）指导教师填写，包括对文献综述的评语，对本课题的深度、广度及工作量的意见及对论文（设计）结果的预测；

2.实习单位意见：由校外实习单位或校内论文指导教师所属学院（处、部）签署意见。