

## 中医病案可视分析

胡晓璇<sup>1,2)</sup>, 彭苏元<sup>3,4)</sup>, 侯海晶<sup>1,2)</sup>, 杨霓芝<sup>1,2)</sup>, 吕永杰<sup>3)</sup>, 周亮<sup>3)\*</sup>

<sup>1)</sup> (广州中医药大学第二临床医学院 广州 510120)

<sup>2)</sup> (广东省中医院肾内科 广州 510120)

<sup>3)</sup> (北京大学健康医疗大数据国家研究院 北京 100191)

<sup>4)</sup> (北京大学公共卫生学院 北京 100191)

(zhoul@bjmu.edu.cn)

**摘要:** 中华中医药是传统文化宝库, 研究及继承中医医案是更好地对其内容进行传承与创新的方式之一. 提出中医病案可视分析方法, 对多变量、多模态时变病案数据和药物对应的现代医学症状高维数据进行分析. 提出流图、降维图、实验室检查指标图等多关联视图和交互关联更新构成的可视分析方法, 支持医学专家运用中医整体观和辨证论治理论分析病案; 根据视觉感知原理设计方便对比的可视映射和交互操作, 探索不同中药对应现代医学症状的相似程度, 展示中西医结合辨病特点. 医学专家对名医的3个不同类型肾脏病例的病案进行分析, 证明提出的可视分析方法在中医传承教学、挖掘核心处方进行新药方设计等方面具有潜力.

**关键词:** 可视分析; 中医医案分析; 多变量多模态时变数据; 医学症状数据  
**中图分类号:** TP391.41 **DOI:** 10.3724/SP.J.1089.2021.19259

## Visual Analysis of Traditional Chinese Medicine Health Records

Hu Xiaoxuan<sup>1,2)</sup>, Peng Suyuan<sup>3,4)</sup>, Hou Haijing<sup>1,2)</sup>, Yang Nizhi<sup>1,2)</sup>, Lyu Yongjie<sup>3)</sup>, and Zhou Liang<sup>3)\*</sup>

<sup>1)</sup> (*The Second Medical College, Guangzhou University of Chinese Medicine, Guangzhou 510120*)

<sup>2)</sup> (*Department of Nephrology, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou 510120*)

<sup>3)</sup> (*National Institute of Health Data Science, Peking University, Beijing 100191*)

<sup>4)</sup> (*School of Public Health, Peking University, Beijing 100191*)

**Abstract:** Traditional Chinese medicine is a profound source of Chinese culture, and studying health records is an effective means for traditional Chinese medicine inheritance and advancement. A visual analysis method is proposed for traditional Chinese medicine health records to analyze multivariate, multimodal, time-varying health record data, and studying medicines in high-dimensional symptom spaces. With multiple linked views composed by a flow chart, dimensionality reduction plots, and lab test plots, aided by brushing-and-linking interactions, the visual analysis method supports medical experts to practice the holistic view and the theory of syndrome differentiation in the analysis. A perception-inspired comparative visual mapping and interaction is designed to investigate the integration of traditional Chinese medicine and modern medicine. The analysis of three cases of various kidney diseases treated by a famous doctor demonstrates that proposed method is promising in traditional Chinese medicine inheritance, and mining core prescriptions to design new ones.

收稿日期: 2021-07-05; 修回日期: 2021-08-27. 基金项目: 国家中医药管理局杨霓芝全国名老中医传承工作室(国中医药人教教育便函[2016]167号); 国家重点研发计划(2018YFC1704100); 北大百度基金(2019BD017, 2020BD032); 北京市高精尖学科建设项目(BMU2019GJJK001). 胡晓璇(1986—), 女, 硕士, 主治医师, 主要研究方向为中医药防治慢性肾脏病; 彭苏元(1990—), 女, 博士, 助理研究员, 主要研究方向为慢性肾脏病的慢病管理、大数据、人工智能等前沿技术在医疗保健领域的应用; 侯海晶(1978—), 女, 博士, 副主任医师, 硕士生导师, 主要研究方向为中医药防治慢性肾脏病; 杨霓芝(1948—), 女, 学士, 教授, 主任医师, 博士生导师, 主要研究方向为中医药防治慢性肾脏病; 吕永杰(1997—), 男, 学士, 主要研究方向为数据分析; 周亮(1986—), 男, 博士, 副研究员, 博士生导师, 论文通讯作者, 主要研究方向为可视化、视觉感知.

**Key words:** visual analytics; traditional Chinese medicine health record analysis; multivariate multimodal temporal data; medical symptom data

运用现代科学解读中医药原理加深对中医药作用的理解,具有重要的科学研究和医疗实践意义. 中医的辨证论治方案符合精准医疗、个体化诊疗的特点,而因此产生的病案是多变量、多模态时变数据,具有可视化和可视分析研究价值. 本文运用中医病案可视分析,对中医病案理-法-方-药进行交互式探索、分析,揭示名医的诊治思路.

中医病案包括疾病的发生、发展、转归,以及进行检查、诊断和治疗等医疗活动过程的记录,即复杂的多变量、多模态时变数据. 传统方法通过阅读病案和结合教材教学等方式对病案进行分析,总结名医经验. 现有的数据驱动的计算机分析方法普遍采用数据挖掘方法分析中医病案<sup>[1-4]</sup>. 然而,这些方法具有以下局限性:挖掘过程没有人的参与,难以形成思维映射,内化成知识;难以进行多模态数据联合分析;无法对多次门诊病案进行分析来研究时变信息.

本文的可视分析方法避免了这些局限. 可视化专家与医学专家紧密合作,经过多次迭代,设计了针对中医病案中的多模态时变信息和中药的多维度数据信息2个方面进行分析的方法. 可视分析方法由多关联视图,即病案时变流图和实验室检查指标图组成的时变视图,支持多模态时变可视分析;多维药物数据降维对比可视化视图允许分析者了解药物基于现代医学症状的相似程度,探索不同患者的诊疗方案的异同. 同时,本文方法采用关联更新技术,将分析人员的专门技术引入分析过程,医学专家能够充分利用自己的知识提出假设,通过交互方式验证假设.

本文创新点如下.

(1) 提出中医多次门诊病案可视分析方法.

(2) 设计了使用前注意视觉元素的图标和交互对比透镜,在降维可视化中进行多属性多病案比较.

(3) 展现高维数据在个体化诊疗的应用场景.

利用基于交互式网页的可视分析工具,医学专家对名医治疗3个不同类型肾脏病的患者的多次门诊病案进行了分析,进一步加深了对中医整体观理论和辨证论治的认识,以及对名医诊治思想的理解. 这些用例证实了多模态数据联合可视化方法的有效性,显示了其在中医传承研究中的潜力.

## 1 相关工作

### 1.1 数据驱动的中医科研

目前,中医科研研究多采用中药、症状等高频分析、系统聚类和关联规则等,对古籍、疾病和名医等经验进行数据挖掘<sup>[1-2]</sup>,建立中药-症状的对应关系网络数据库<sup>[3]</sup>;或采用网络药理数据库进行名医经验等挖掘工作<sup>[4]</sup>,其挖掘结果多以柱状图、信息表格和关联规则网络展示图展现. 既往的数据挖掘结果以平面展示,且需医学专家后续对其解读分析,或以既往专家的文献研究、经验观点进行印证,明确挖掘结果的可靠性.

### 1.2 医疗可视分析

医学是可视化和可视分析的重要应用领域. 相关研究有医疗中应用的交互式信息可视化系统的全面综述<sup>[5]</sup>、医疗数据的分析和规划方法的综述<sup>[6]</sup>、当前医疗数据的可视化和可视分析进展介绍<sup>[7]</sup>,以及对医疗可视分析的展望<sup>[8-9]</sup>.

现代医学的电子病案可视分析方法已有较长历史,早期有方法专注于展示单个患者的电子病例,以展示其医疗历史<sup>[10-11]</sup>. 从事件序列角度对电子病历进行可视分析是当前的重要方法<sup>[12-15]</sup>. 近年来,已有方法对电子病案的事件序列进行可视分析,并对未来事件进行预测,以辅助医疗决策<sup>[16]</sup>. 然而,这些方法都不能对中医病案进行有效的分析,它们只能可视化零散的药物使用情况,不能分析中医病案中众多药物的功能、组合原理,展示处方随时间的变化过程. 而中医的重要特点是贯穿诊疗始终的“整体观”和“辨证论治”思想,需要还原治疗疾病的全过程. 当前针对中医诊疗的可视化主要是针对药物共现和关联分析的呈现,现有工作中尚未发现针对中医病案的可视分析方法.

专门的药物处方可视化和可视分析与本文工作相关. 通过将处方中的3个参与者(医生、药物、患者)进行拆分,再使用交互可视分析进行关联,可以进行大规模处方数据的分析<sup>[17]</sup>. 药物数据中往往涉及具有大量属性的表格数据,研究者提出在医生熟悉的表格环境下将数据以项目表格和特征表格的方式进行可视化<sup>[18]</sup>.

### 1.3 多维时变数据可视化

本文的可视分析方法与多维度可视化相关.

降维方法是多维度可视化的常用方法, 通过将高维数据映射到低维(往往是二维), 在低维空间中进行可视化. 降维方法有线性<sup>[19]</sup>和非线性<sup>[20]</sup>2 大类, 全面的降维方法对比可参见相关文献[21]. 其中, 非线性方法能更好地在降维结果中保留多维空间中的特征. 本文选择了当前流行的、性能优秀的、一致的流形近似和投影降维方法(uniform manifold approximation and projection, UMAP)<sup>[22]</sup>, 这种方法类似于  $t$ -分布和随机近邻嵌入法( $t$ -distributed stochastic neighbor embedding,  $t$ -SNE)<sup>[23]</sup>, 尽量保留数据在原始多维空间中的近邻关系, 克服了  $t$ -SNE 的一些缺陷, 并提高了计算效率.

多变量时变数据可视化在病案可视化中起到重要作用, 其形式主要有折线图、堆叠柱状图、河流图<sup>[24]</sup>或流图等. 经过可视化专家与医学专家的反复讨论, 本文选择了平滑的流图进行可视化, 兼顾可视化传递信息的准确性和美观性<sup>[25]</sup>.

专家要求在降维视图中同时可视化 2 个类型的变量, 因此, 本文设计了利用前注意视觉元素的新视觉映射. 前注意视觉元素指在很短时间内不需大脑搜索就能被观察到的视觉元素<sup>[26]</sup>. 心理学和可视化领域<sup>[26]</sup>对前注意视觉元素做出了鉴别, 研究了前注意视觉元素组合的规则. 例如, 使用颜色和纹理对多维数据进行可视化<sup>[27]</sup>. 近来, 通过感知实验验证的特别设计的前注意视觉元素已被运用在高动态范围标量数据可视化<sup>[28]</sup>和虚拟现实环境中的可视化领域<sup>[29-30]</sup>.

## 2 背景和设计

目前中医药病案数据研究主要针对名医经验传承模式, 也有各类数据挖掘方法, 其中大部分基于挖掘疾病-症状-药物等, 可体现中医治病 2 大特点: 整体观和辨证论治. 整体观即人体本身的统一性、完整性和人体与自然界的相互关系; 辨证论治即通过望闻问切, 综合认识疾病和治疗疾病的基本原则<sup>[31]</sup>.

本文将利用多模态数据联合可视化的方法, 进一步加深对中医整体观理论和辨证论治的理解和认识, 在探索名医经验的同时, 为中医个性化诊疗提供应用场景. 经过可视化专家与医学专家紧密合作, 反复深入讨论, 以及需求分析, 明确可视分析方法的设计目标和用户需求. 即需要达到的目标如下.

**R1.** 展现名医诊疗中的中医整体观特点.

**R2.** 研究名医诊疗中的辨证论治特点研究名医诊疗中的辨证论治的特点.

**R3.** 快速识别名医对药物的认识.

**R4.** 展示名医对疾病的认识, 体现诊治特点.

**R5.** 视觉元素设计应符合中医理论的意象.

使用迭代式原型设计流程, 全过程中可视化专家与医学专家紧密配合, 最终得出可视分析方法. 期间, 研究者花费大量时间沟通、理解对方学科的概念和方法, 进行过多次头脑风暴讨论, 研究如何实现设计目标 and 需求. 时变病案数据、中药多维数据的最终可视化均是从多个替代设计方案中经医学专家试用选出的.

本文花费大量时间研究如何将药物进行分类和进行颜色映射. 最初, 试图通过用户进行交互式分类, 并为此设计并实现了多关联套索工具. 然而, 医学专家试用后认为此方法过于费时费力, 并且不能展现病案中的名医对药物的认识. 故而将药物分类改为按照《中药学》教材分类<sup>[32-33]</sup>和名医对药物的认识分类作为默认类别(R3, R4). 由于中医传承模式的不同, 《中药学》教科书理论框架与医生(特别是专病诊疗经验丰富的名医)在诊疗活动中的实际运用和阐释发挥存在差异, 因此设计了降维视图中多变量对比图标. 其中, 将药物认识按照对应症状、四气五味(中药药性理论的基本内容之一, 按照中药的性质和滋味进行分类)分别进行投影降维, 并进行可视化设计. 名医在治疗过程通常基于对疾病病机及药物特性的认识开具处方, 以更多元地展示教材药物分类与名医对药物认识的区别. 降维图中空间显示为中药对应的现代医学症状, 其数据来源于 SymMap(Symptom mapping)中医药证候关联数据库<sup>[3]</sup>. 本文将 3 则病案所有出现的中药在 SymMap 系统中 47 味中药及对应的 387 个西医症状建立降维图的空间数据(R4). 另外, 本文还将药物分类对应颜色, 设计了符合中医理论的色表(R5).

在整个设计过程中, 医学专家提出需要解决的中医问题, 尽可能通俗地解释其对应的病案数据; 可视化专家提出可能相关的可视化和交互方法, 反向促进、激发医学专家提出新的中医概念和理论并进行分析.

## 3 本文方法

根据需求设计的可视分析方法由多关联视图组成, 通过关联更新进行交互分析. 多关联视图如

图 1 所示, 由病案时变流图、实验室检查指标图、带有属性对比可视化且能通过交互透镜进行邻域对比的中药多维数据降维图以及药物属性表格组成. 病案 1 以尿检异常(血尿)为主诉, 以尿液检查中的“尿潜血”作为观察指标; 病案 2 以尿检异常(蛋白尿)为主诉, 以尿液检查中的“尿蛋白”作为观察指标; 病案 3 以肾功能异常为主诉, 以血液肾功能检查的“血肌酐”作为观察指标, 该指标是反映肾脏功能的指标, 目前临床使用相对广泛.

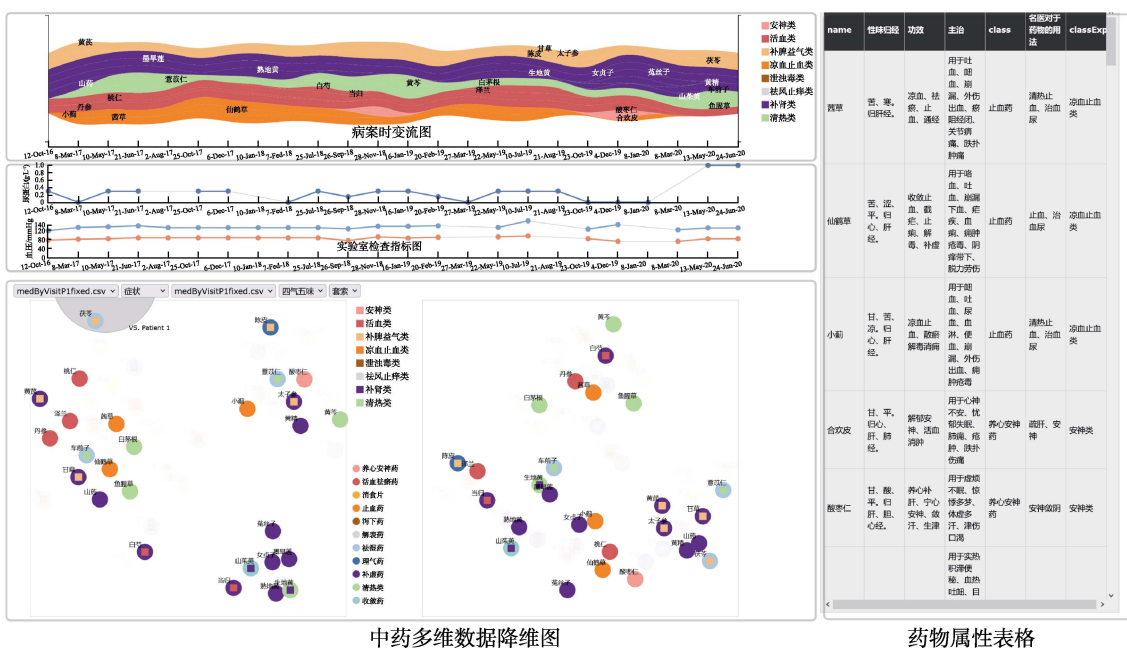
### 3.1 病案时变视图

病案时变视图由病案时变流图和实验室检查指标折线图组成, 数据来源于多次随诊门诊病案资料, 其横坐标为时间(诊次的时间排序), 流图纵

坐标为每诊次使用的药物组合(即药方), 折线图纵坐标为实验室检查数据值. 通过多次门诊病案在时间域的变化情况, 展示患者整体患病情况(R1), 并通过时变数据流图体现了患者不同诊次随症状而改变的诊疗情况(R2).

#### (1) 病案时变流图

图 2 所示病案时变流图是病案时变视图的核心. 最初, 可视化专家提出使用离散的堆叠柱状图可视化时变的药物处方数据, 以保证准确性; 但医学专家认为这样破坏了药物随时间的变化趋势, 不利于展示整体观(R1)和辨证论治(R2)特点, 并且其美观性较差. 因此, 最终方案使用平滑变化的流图, 辅以交互式显示距离当前时间最近门诊的准



中药多维数据降维图  
图 1 中医医案可视分析系统总览

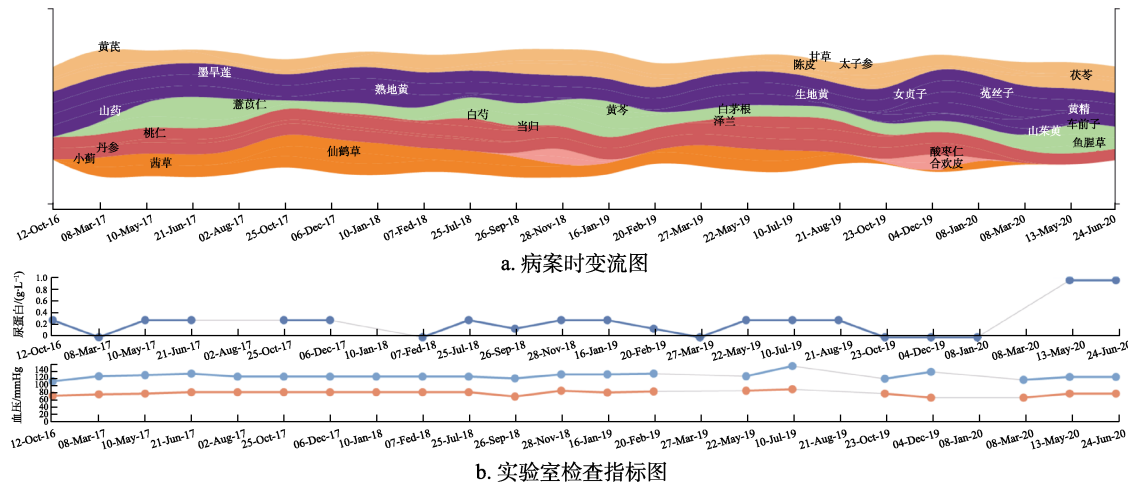


图 2 病案 1(血尿)

确药量的方式, 兼顾 R1 和 R2 的设计要求。

案时变流图直观地展现了同一患者各次门诊的处方所用中药随时间的变化。在流图中, 横轴为诊次时间排序, 由于门诊访问时间间隔呈非线性, 使用线性时间进行可视化效果不佳, 无法体现中医根据病情的个性化诊疗特点。例如, 患者病情稳定时处方可能长期无变化或变化较小; 而对患者病情短期进展较快的情形, 由于可视化绘制空间不成比例, 可能无法被分析者正确感知并进行分析。纵轴为处方中各中药的药量, 其使用了平滑的流图设计。为了保证数据呈现的准确性, 分析者选中任意河流位置, 距离该位置最近的门诊处方中的药量会显示在流图旁边, 以兼顾可视化的准确性和美观性。医学专家认为, 这种设计配合合理的药物分类着色, 可以直观地体现患者整体患病情况(R1)。

根据设计需求, 流图需要能展示整类药物的时变信息, 药物名称应直接显示在流图上, 方便查找。因此, 将流图中的药品按照名医分类进行排序, 同一类药物处在流图的相邻河流中。采用一种简单的自适应算法放置药品名称。首先, 记录按门诊次数的每条河流最宽处(可能有多个位置)中心横纵坐标, 并计算该条河流宽度的中位数。若该河流宽度中位数不接近河流最宽处, 则从记录的河流最宽处位置中随机选择, 定为中药药名标签位置; 若宽度中位数与河流最宽处接近, 说明河流宽度变化不大, 在整个河流长度上随机生成药品名称标签的横坐标位置, 并计算相应河流中部纵坐标位置。

对于中药药名标签颜色, 计算标签所在的河流对应颜色的 CIELab 亮度值  $L$ 。若  $L \leq 50$ , 将标签设为白色; 否则设为黑色, 以获取较好的对比度。得到的病案时变流图如图 2a 所示。该方法可以通过时变数据流图体现患者不同诊次随症状而改变的诊疗情况(R2)。

### (2) 实验室检查指标图

患者的疾病相关实验室检查指标被绘制成折线图。经医学专家确认, 需要的实验室检查指标有尿蛋白(尿液检查)、血肌酐(肾功能检查)、收缩压和舒张压(血压)4 种数据。尿蛋白、血压控制情况为慢性肾脏病进展的独立危险因素; 而血肌酐是反映肾功能情况的直接指标, 目前临床使用相对广泛。由于病案中存在缺失值——患者并非每次门诊都检测所有实验室检查指标, 在绘制时使用圆点代表检测的实验室检查指标, 2 个相邻门诊

测用粗实线连接圆点, 用细虚线连接 2 次检测间的缺失值, 代表疾病变化趋势, 如图 2b 所示。在呈现整体趋势的同时, 清晰、无歧义地可视化检测值和缺失值, 向分析者准确地提供中医医案的完整信息, 更加直观地体现患者症状、实验室检查指标的改变与诊疗药物组方之间的关系。

### (3) 药物分类色表设计

本文设计了符合中医理论的色表(R5)。中医理论将活血类映射成红色, 清热类映射为绿色, 祛风止痒映射为白色(浅灰表示), 泻浊毒类映射为褐色, 补肾类映射为紫色。五色(青、赤、黄、白、黑)对应中医五行理论(五脏, 即肝、心、脾、肺、肾), 将这些概念与符合视觉感知的色表设计结合并经过反复实验, 得到了如图 3 所示的色表。



图 3 符合中医概念和视觉感知的色表

## 3.2 药物信息视图

为体现名医对药物和疾病的独特认识, 并加深对中医辨证论治个体化诊疗的理解(R2, R3, R4), 本文设计了图 1 所示药物信息视图。药物信息视图以降维可视化为基础, 通过专门设计的图标和交互透镜达到对比不同病案、对比教材与名医对药物的分类、药-证结合的分析目的。

### (1) 降维可视化

将所有药物对应的全部非重复症状和四气五味属性分别组成高维空间(387 维和 23 维), 这样, 每种药物可看做由 387 维( $K=387$ )布尔向量构成的症状空间点  $P_s$ , 以及由 23 维( $M=23$ )布尔向量构成的中医属性空间点  $P_a$ , 即

$$\begin{cases} P_s = (s_1, s_2, \dots, s_K), & s_i \in \{0, 1\} \\ P_a = (a_1, a_2, \dots, a_M), & a_i \in \{0, 1\} \end{cases}$$

本文设想过使用基于矩阵的热力图可视化多维中药数据。然而, 经讨论后认为, 此种方法无法呈现中药的聚类效果, 限制了交互方式, 并且也不容易添加额外信息进行多病例对比。最终, 本文选用了降维方法, 利用图标(glyph)绘制可视化中药信息, 并在实验对比了  $t$ -SNE 和 UMAP 这 2 种降

维方法后, 选择了 UMAP 进行降维<sup>[22]</sup>. 将药物对应的症状空间位置  $P_s$  和中医属性空间位置  $P_a$  分别降维至二维空间, 得到相应低维空间位置  $P'_s$  和  $P'_a$ . 2 个降维映射分别为

$$\begin{cases} \pi_s: \mathbb{R}^K \rightarrow \mathbb{R}^2, P_s \rightarrow p_s \\ \pi_a: \mathbb{R}^M \rightarrow \mathbb{R}^2, P_a \rightarrow p_a \end{cases}$$

降维后的数据将以多属性对比图标散点图方式呈现.

## (2) 多属性对比图标

由于需要快速对比名医和教材对药物的分类 (R3), 因此需要在降维视图的数据点上同时呈现相应的 2 种属性——专家分类和教材分类, 并使用恰当的视觉映射快速区分 2 个属性. 独特的视觉属性会激发前注意处理机制, 使人对相应图形进行快速区分<sup>[26]</sup>. 为此, 本文设计了如图 4 所示前注意视觉元素构成的多属性对比图标, 设计图标为中心对称, 以充分利用颜色, 减少不对称性对视觉搜索的影响.

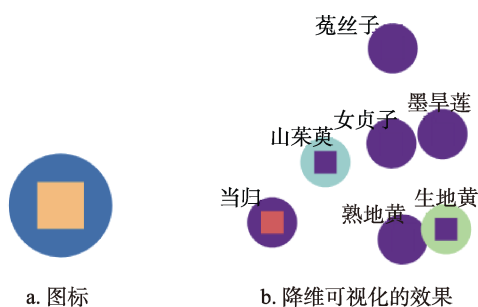


图 4 利用前注意视觉元素设计的可以同时显示 2 种分类的图标

将名医分类和教材分类映射成不同尺寸、大小的视觉元素如图 4a 所示, 方形作为内侧代表名医分类, 同心圆形作为外侧代表教材分类. 由于颜色是前注意视觉元素, 因此在其他条件不变的情况下, 颜色差异能被快速识别——当内外侧颜色不一致时, 人能迅速地将其与其他图标区分; 同时, 人能专注于某种视觉属性进行搜索<sup>[26]</sup>, 图标内外侧颜色可分别作为名医和教材分类中不同类别快速区分的视觉属性. 使用不同形状代表不同分类, 分析者不需额外记忆尺寸信息, 减少了脑力负担; 同时, 由于分析者理解可视化的是药物属性的 2 种分类, 不会将 2 种形状误认为药物的 2 种属性. 图 4b 显示了降维可视化中相邻药物的对比图标效果. 例如, 当归(内红外紫)和一系列紫色图标具有不同的名医分类、相同的教材分类; 山茱萸(内紫外青)、生地黄(内紫外绿)和紫色图标有相同的名医分类、

不同的教材分类. 此外, 专注于名医分类(内侧方形), 当归因颜色不同容易被区分; 专注于教材分类(外侧圆形), 山茱萸和生地黄容易和其他药物区分. 此外, 图标设计也符合中国哲学“天圆地方”的理念, 颜色选取与中医五脏对应、与五色理论相吻合 (R5).

## (3) 邻域对比透镜

为分析名医对不同疾病 (R4)、不同症状 (R2) 的认识, 需要比较不同患者的用药特点. 降维视图提供通过 2 个降维图灵活地进行全局对比, 可选择不同患者、不同属性(药物的四气五味属性或现代医学症状属性). 然而, 全局对比在观察降维图局部变化时注视区域需要进行跨空间反复跳动 (2 个降维图相同位置药品对比需横跨一个降维图的宽度), 分析效率低, 易造成分析者疲劳.

为此, 本文设计了如图 5 所示邻域对比透镜, 分析者只需将透镜放置在感兴趣的位置, 即可对比当前病案与目标病案(需要对比的病案). 其中, 高不透明度代表目标病案中出现的药品, 低不透明度代表当前病案中出现的药品, 在图标中心用 X 表示在目标病案中出现的当前病案中未出现的药品.

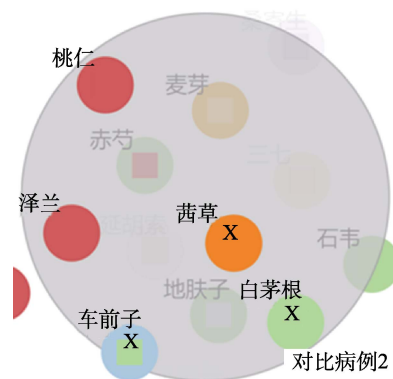


图 5 邻域对比透镜对比当前病案与目标病案

## 3.3 药物信息表格和关联更新

根据设计需求, 本文方法为分析者提供作为参考的药物信息表格, 显示药物的名称、性味归经、功效、主治、教材分类、知名中医分类和知名中医对药物的用法的文字信息. 分析者在可视化视图中进行探索, 得出假设后可在信息表格中进行查阅验证; 反之, 阅读表格信息, 分析者也能建立新的假设, 并在可视化视图中进行验证. 本文可视分析方法中多个视图通过关联更新进行联系, 如在病案流图中, 选中河流对应的药物会在降维视图中高亮显示, 在套索模式下会更新并按流图

中所示进行分类。

### 4 用例分析和用户体验

本文选取了名医诊治 3 名慢性肾脏病患者的病案作为用例，其中病案 1 症状以尿检异常(血尿)为主，病案 2 症状以尿检异常(蛋白尿)为主，病案 3 以肾功能异常(肾衰竭)为主，通过多属性、多病案比较和可视化分析方案，更直观地体现高维数据在个体化诊疗的应用场景。

#### 4.1 时变分析

从流图展示的处方分析，可以发现病案 1(图 2)、病案 2(图 6a)、病案 3(图 6b)均存在 3 种颜色(黄色、紫色、红色)贯穿始终，这 3 种颜色对应的药物类别为补脾、补肾和活血类，揭示名医治疗肾病以补脾补肾活血为治法为主，且 3 种颜色贯穿治疗全过程，提示名医对肾病认识的整体观，以补脾补肾的特性指导肾脏病治疗。

从图 2 发现，清热类药物、凉血止血类药在病案 1 中始终被使用，且对比其他 2 个病案药物也具有独特性。

在图 6a 针对蛋白尿尿检异常的诊治中，可以发现使用降蛋白尿类中药明显增多，与折线图所

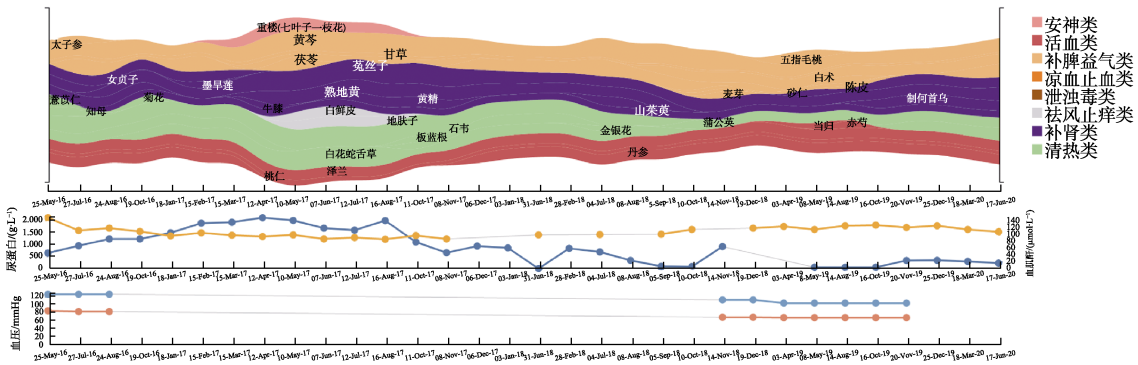
表示的尿检异常的实验室检查指标相吻合。同时，在该病案的早期清热类药物比重大，后期清热类药物使用比重减少(流图中绿色部分)。另外，该患者在治疗过程中使用祛风止痒类中药(流图中白色部分)，可以推测患者出现皮疹或是风热类症状，这个推测可以在图 7b 相应的药物空间域中找到相应症状范畴的药物对比，有助于探索名医对该症状药物使用的特性，更好地传承名医经验。

图 6b 所示为病案 3 肾功能异常(肾衰竭)病案流图，对比病案 1 和病案 2，可以发现清热类药物(绿色部分)在就诊早期曾被使用，但后期停用，而出现泄浊毒类药物(褐色部分)，可明显发现名医对于肾衰竭的认识，随着疾病的发展，在扶正基础上早期以利水祛湿清热为治法，后期浊毒明显，更多改用泄浊毒类中药。

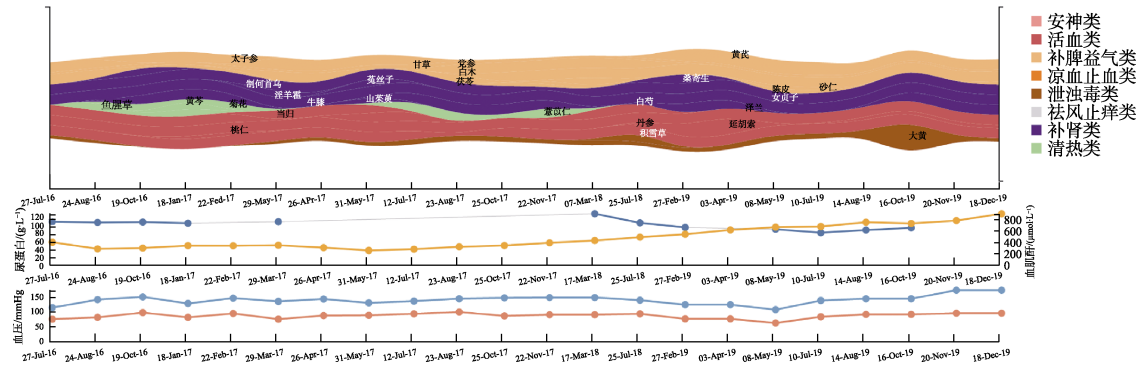
从以上 3 个流图可以推测，名医治疗慢性肾脏病以补脾补肾活血为基本法，根据病案症状的区别辨证论治，并基于病机认识，随着病情的进展调整处方的变化。可视分析可以很好地展示中医整体观和辨证论治的理论。

#### 4.2 降维图对比分析

如图 7 所示，在针对病案 1 的降维图中，可快速通过图标的颜色差异来识别名医和教材对中药



a. 病案 2 (蛋白尿)



b. 病案 3 (肾功能异常)

图 6 时变分析

理解的差异: 茯苓、陈皮、黄芪、太子参、甘草、白芍、山茱萸、当归和生地黄, 这几味中药均为名医对于药物的认识有别于药典药物属性的独特认识, 在病案 2 和病案 3 中均可发现同样的结果. 基于降维图快速识别名医对药物的独特认识, 有助于医学专家快速、直观地发现名医和教材对于药物

和症状的不同认识, 更深刻地领会中医个体化诊疗的研究内涵. 名医认识药物对于疾病的作用基于最早的药典和教材, 随着其临床经验的累积, 可发现药物在治疗肾病中的特性, 并基于现代药理研究结果, 在选用药物的分类时具有个人经验特色, 这也是传承名医经验的内容之一.

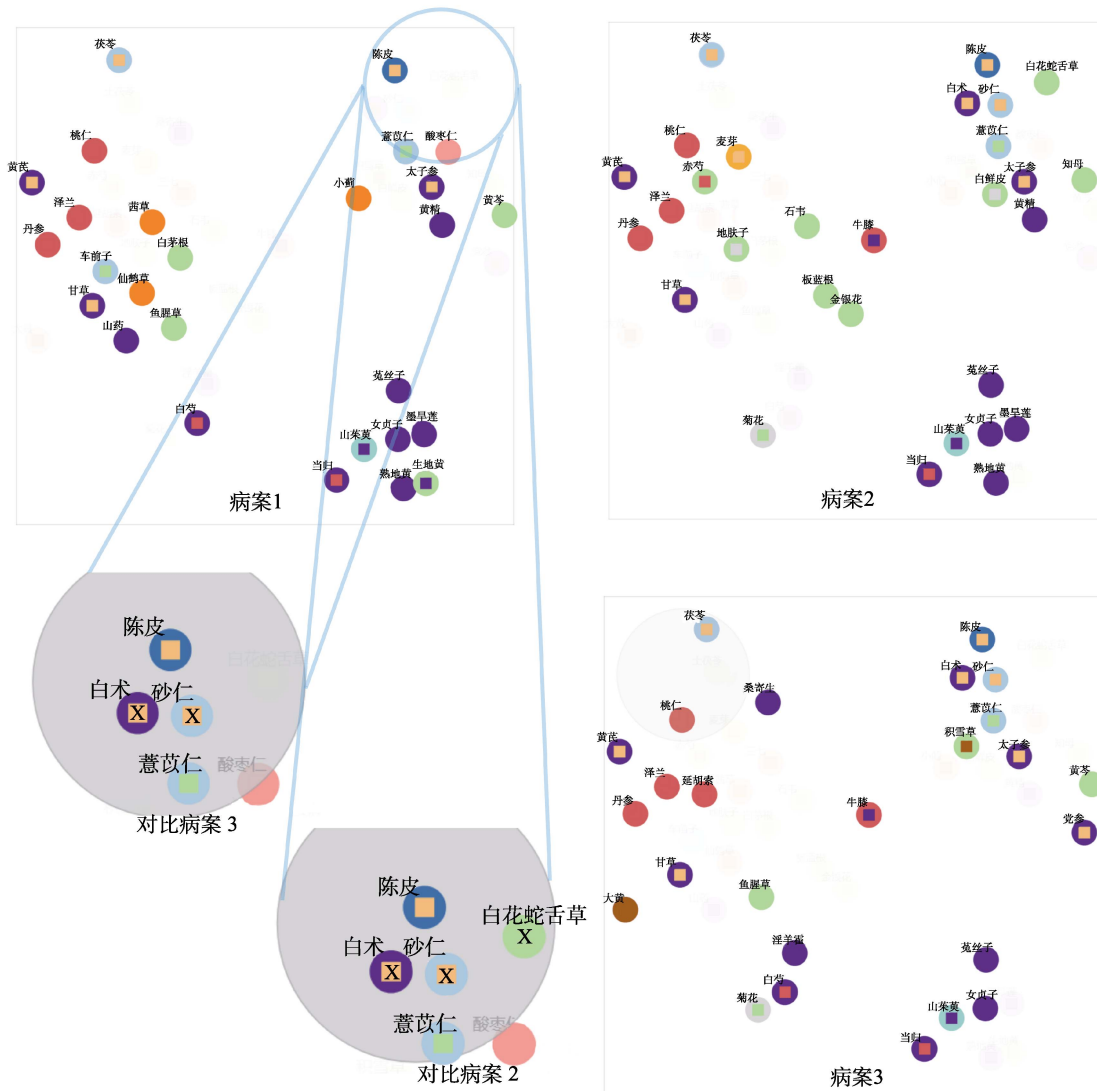


图 7 降维图分析

### 4.3 邻域对比透镜交互分析

使用邻域对比透镜可以发现在降维图的同空间范畴的现代医学症状. 病案 1 和病案 3 使用的药物类别呈现不一致, 对病案 1 中陈皮对应区域使用邻域对比透镜, 病案 1 仅有一味药物(陈皮), 病案 3 该区域对应 3 味药物(陈皮、白术、砂仁), 病案 2 该区域对应 4 味药物(陈皮、白术、砂仁、白花蛇舌草). 这一结果揭示, 名医治疗同一类现代西医症状, 不同病案患者使用不同药物, 甚至不同类别的药物(药物的颜色区别). 这体现了中医病证结合个

体化诊疗的内涵, 为后续研究名医经验提供了方向.

### 4.4 用户体验

可视化专家对使用了本文可视分析方法的医学专家进行了访谈, 访谈要点总结如下.

(1) “现有的对于中医医案的研究缺乏对病案的完整还原, 无法体现中医独特的疾病整体观与辨证论治过程, 可视分析以时态作为变量, 可充分展示中医医案的全貌, 从中可以窥视医者在治疗疾病特点、如何把握疾病的动态演变. 例如, 对于病案 3 应用可视分析工具的流图, 发现名医治疗肾

衰患者中,观察患者毒素逐渐累积的趋势,治疗对策从清热泻浊到泄浊毒的比例变化,把握致病的外在因素调整处方变化”。

(2)“借助可视分析发现名医治疗肾病存在异病同治、同病异治的特性,使后辈更快地理解与掌握医案理法方药,更好地传承其经验”。

(3)“另外,运用可视分析工具分析平日的医案可帮助我们快速总结自身经验,建议开发者再进一步进行系列病案多名患者的可视分析,有助于对比有效病案与治疗欠佳的病案区别点,帮助我们更快成长”。

(4)“可视分析以时态为变量的多维数据分析,可以还原医案进行主客观相结合的研究,更好地体现中医医学的独特性”。

## 5 结 语

本文提出了中医病案可视分析方法,通过病案流图和实验室检查指标图,首次引入时变作为研究变量,还原病案的整体性,可增强对病机的认识.通过将中药和现代医学症状关联的高维空间进行可视化,使用专门设计的多属性对比图标、邻域对比透镜,分析者可以快速、交互式地识别名医对中药的独特认识.

本文提出的可视分析方法为继承知名专家的学术思想提供便捷方式.后期可根据可视分析结果进行药物药理研究,挖掘基于专家认识的药对、新药研发等提供一定信息基础研究,发现潜在经验,提示进一步的传承研究.

未来期望能对本文方法进行改进和更深入的研究.当前的方法只对少量病案进行分析,具有一定的局限性.本文期待更多的病案、病种研究,或广泛的知名中医专家经验的挖掘研究.在可视分析方法上,考虑设计病案全局差异对比、多个病案同时对比的工具;使用自动标签排布算法避免标签(如中药名)和图形重叠.本文期望让更多医学专家使用该可视分析方法,获取专家反馈,进行用户研究.

## 参考文献(References):

- [1] Hu Xiaoxuan, Zhang La, Li Zhibin, *et al.* Analysis on medical records of Huang Chunlin for characteristics of chronic kidney disease treatment by data mining method[J]. World Science and Technology-Modernization of Traditional Chinese Medicine, 2013, 15(5): 969-974(in Chinese)
- [2] Peng S Y, Liu X S, Li Y, *et al.* TCMISS-based analysis on the regularity of ancient TCM prescriptions for chronic renal failure[C] //Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. Los Alamitos: IEEE Computer Society Press, 2014: 144-149
- [3] Wu Y, Zhang F L, Yang K, *et al.* SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping[J]. Nucleic Acids Research, 2019, 47(D1): D1110-D1117
- [4] Duan Xiaojun, Tan Ping, Hou Haijing, *et al.* Clinical experience of treating chronic renal failure with Chinese medicine by national renowned doctor Yang Nizhi[J]. Lishizhen Medicine and Materia Medica Research, 2019, 30(10): 2503-2505(in Chinese)  
(段小军, 谈平, 侯海晶, 等. 全国名老中医药专家杨霓芝教授治疗慢性肾衰竭经验[J]. 时珍国医国药, 2019, 30(10): 2503-2505)
- [5] Rind A, Wang T D, Aigner W, *et al.* Interactive information visualization to explore and query electronic health records[J]. Foundations and Trends in Human-Computer Interaction, 2013, 5(3): 207-298
- [6] Kosara R, Miksch S. Visualization methods for data analysis and planning in medical applications[J]. International Journal of Medical Informatics, 2002, 68(1-3): 141-153
- [7] Cao Nan. Information visualization and visual analysis applications in intelligent medicine[J]. Communications of the China Computer Federation, 2019, 15(3): 10-16(in Chinese)  
(曹楠. 信息可视化及可视分析在智慧医疗领域的应用[J]. 中国计算机学会通讯, 2019, 15(3): 10-16)
- [8] Caban J J, Gotz D. Visual analytics in healthcare: opportunities and research challenges[J]. Journal of the American Medical Informatics Association, 2015, 22(2): 260-262
- [9] Gotz D, Borland D. Data-driven healthcare: challenges and opportunities for interactive visualization[J]. IEEE Computer Graphics and Applications, 2016, 36(3): 90-96
- [10] Cousins S B, Kahn M G. The visual display of temporal information[J]. Artificial Intelligence in Medicine, 1991, 3(6): 341-357
- [11] Plaisant C, Mushlin R, Snyder A, *et al.* LifeLines: using visualization to enhance navigation and analysis of patient records[M] //The Craft of Information Visualization, 2003: 308-312
- [12] Wang T D, Plaisant C, Quinn A J, *et al.* Aligning temporal data by sentinel events: discovering patterns in electronic health records[C] //Proceedings of the 26th Annual CHI Conference on Human Factors in Computing Systems. New York: ACM Press, 2008: 457-466
- [13] Wang T D, Plaisant C, Shneiderman B, *et al.* Temporal summaries: supporting temporal categorical searching, aggregation and comparison[J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 1049-1056
- [14] NiklasNorén G, Hopstadius J, Bate A, *et al.* Temporal pattern discovery in longitudinal electronic patient records[J]. Data Mining and Knowledge Discovery, 2010, 20(3): 361-387

- [15] Monroe M, Lan R J, Lee H, *et al.* Temporal event sequence simplification[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2227-2236
- [16] Jin Z C, Cui S Y, Guo S A, *et al.* CarePre: an intelligent clinical decision assistance system[J]. *ACM Transactions on Computing for Healthcare*, 2020, 1(1): Article No.6
- [17] van der Corput P, Arends J, van Wijk J J. Visualization of medicine prescription behavior[J]. *Computer Graphics Forum*, 2014, 33(3): 161-170
- [18] van der Corput P, van Wijk J J. Exploring items and features with  $F^F$ ,  $F^I$  - tables[J]. *Computer Graphics Forum*, 2016, 35(3): 31-40
- [19] Cunningham J P, Ghahramani Z. Linear dimensionality reduction: Survey, insights, and generalizations[J]. *The Journal of Machine Learning Research*, 2015, 16(1): 2859-2900
- [20] Lee J A, Verleysen M. *Nonlinear dimensionality reduction*[M]. Heidelberg: Springer, 2007
- [21] van der Maaten L, Postma E, van den Herik J. *Dimensionality reduction: a comparative review*[R]. Tilburg: Tilburg University, 2009
- [22] McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction[J]. *The Journal of Open Source Software*, 2018, 3(29): 861
- [23] van der Maaten L, Hinton G. Visualizing data using  $t$ -SNE[J]. *Journal of Machine Learning Research*, 2008, 9(11): 2579-2605
- [24] Havre S, Hetzler B, Nowell L. ThemeRiver: visualizing theme changes over time[C] // *Proceedings of the IEEE Symposium on Information Visualization*. Los Alamitos: IEEE Computer Society Press, 2000: 115-123
- [25] Byron L, Wattenberg M. Stacked graphs-geometry & aesthetics[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2008, 14(6): 1245-1252
- [26] Healey C, Enns J. Attention and visual memory in visualization and computer graphics[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(7): 1170-1188
- [27] Healey C G, Enns J T. Large datasets at a glance: combining textures and colors in scientific visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 1999, 5(2): 145-167
- [28] Zhou L, Rivinius M, Johnson C R, *et al.* Photographic high-dynamic-range scalar visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(6): 2156-2167
- [29] Krekhov A, Krüger J. Deadeye: a novel preattentive visualization technique based on dichoptic presentation[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(1): 936-945
- [30] Krekhov A, Cmentowski S, Waschk A, *et al.* Deadeye visualization revisited: investigation of preattentiveness and applicability in virtual environments[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(1): 547-557
- [31] Sun Guangren. *Traditional Chinese medicine basic theory*[M]. Beijing: China Press of Traditional Chinese Medicine, 2017(in Chinese)  
(孙广仁. 中医基础理论[M]. 北京: 中国中医药出版社, 2017)
- [32] Gao Xuemin. *Chinese medicine*[M]. Beijing: China Press of Traditional Chinese Medicine, 2004(in Chinese)  
(高学敏. 中药学[M]. 北京: 中国中医药出版社, 2004)
- [33] National Pharmacopoeia Committee. 2020 (No.1) People's Republic of China pharmacopoeia[M]. Beijing: China Medical Science Press, 2020(in Chinese)  
(国家药典委员会. 中华人民共和国药典(一部, 2020 年版)[M]. 北京: 中国医药科技出版社, 2020)