

慢性乙型肝炎治疗效果分析 - 完整方法学报告

目录

- 研究设计概述
- 数据预处理
- 倾向性匹配方法
- 统计分析方法
- 分析结果解读

1. 研究设计概述

1.1 研究目的

评估中西医结合疗法（治疗组）与单纯抗病毒治疗（对照组）在慢性乙型肝炎患者中的疗效差异。

1.2 研究类型

回顾性队列研究（Retrospective Cohort Study）

由于非随机分组，存在选择偏倚风险，需采用倾向性匹配方法控制混杂因素。

1.3 变量定义

变量类型	变量名称	说明
分组变量	treatment	1=治疗组（中西医结合），0=对照组（单纯抗病毒）
协变量	性别、年龄、饮酒史、吸烟史、是否合并脂肪肝、基线肝功能指标	用于匹配
主要结局	ALT 变化、AST 变化	肝细胞损伤标志物
次要结局	GGT、ALP、白蛋白、胆红素、肝硬度值、血小板	其他肝功能指标

1.4 临床正常范围

指标	正常范围	单位	临床意义
ALT	7-40	U/L	越低越好
AST	13-35	U/L	越低越好
GGT	7-45	U/L	越低越好
ALP	35-100	U/L	越低越好
白蛋白	40-55	g/L	越高越好
胆红素	3.4-17.1	μmol/L	越低越好
肝硬度值	2.5-7.0	kPa	越低越好

2. 数据预处理

2.1 数据读取与清洗

2.1.1 数据格式要求

原始数据：data/data.csv
格式：CSV 文件，第 1 行为说明行（跳过），第 2 行为列名

2.1.2 数值清洗函数

对于包含特殊字符的数值（如"小于 0.05"），提取数值部分：

$$\text{clean}(x) = \begin{cases} x & \text{if } x \in \mathbb{R} \\ \text{extract_number}(x) & \text{if } x \in \text{String} \\ \text{NaN} & \text{otherwise} \end{cases}$$

2.1.3 治疗分组编码

$$\text{treatment} = \begin{cases} 1 & \text{if 治疗方案} = 1 \text{ (中西医结合)} \\ 0 & \text{if 治疗方案} = 2 \text{ (单纯抗病毒)} \end{cases}$$

2.2 缺失值处理

策略：完全病例分析（Complete Case Analysis）

删除协变量存在缺失值的样本：

$$\mathcal{D}_{clean} = \{i: X_{ij} \neq \text{NaN}, \forall j \in \text{Covariates}\}$$

2.3 变量定义

2.3.1 变化值计算

对于每个结局指标，计算治疗前后变化值：

$$\Delta Y_i = Y_{i,12\text{月}} - Y_{i,\text{基线}}$$

其中： $-\Delta Y_i > 0$ ：指标升高（对于 ALT/AST 等，表示恶化） $-\Delta Y_i < 0$ ：指标下降（对于 ALT/AST 等，表示改善）

2.3.2 基线状态分类

根据正常范围上限 U 和容忍度 $\tau = 2$ ：

$$\text{Status}_{i,\text{baseline}} = \begin{cases} \text{异常偏高} & \text{if } X_{i,\text{基线}} > U + \tau \\ \text{正常} & \text{if } L - \tau \leq X_{i,\text{基线}} \leq U + \tau \\ \text{异常偏低} & \text{if } X_{i,\text{基线}} < L - \tau \end{cases}$$

3. 倾向性匹配方法

3.1 为什么需要匹配？

在非随机化研究中，治疗组和对照组的基线特征可能存在系统性差异（**选择偏倚**），导致：

$$E[Y(1)|T = 1] - E[Y(0)|T = 0] \neq E[Y(1) - Y(0)]$$

其中 $Y(1)$ 和 $Y(0)$ 分别为潜在结果， T 为治疗分配。

匹配的目的是使两组在协变量分布上相似，从而：

$$E[Y(1)|T = 1, X] - E[Y(0)|T = 0, X] \approx E[Y(1) - Y(0)|X]$$

3.2 粗化精确匹配（CEM）

3.2.1 方法原理

CEM（Coarsened Exact Matching）是一种非参数匹配方法，核心步骤：

Step 1: 协变量粗化

对连续协变量 X_k ，按分位数划分为 J 个区间：

$$C(X_{ik}) = j \quad \text{if} \quad Q_{j-1} < X_{ik} \leq Q_j$$

其中 Q_j 为第 j 个四分位数（当 $J = 4$ 时）。

Step 2: 层定义

每个样本的层标识为所有粗化协变量的组合：

$$S(i) = (C(X_{i1}), C(X_{i2}), \dots, C(X_{ip}))$$

Step 3: 精确匹配

只保留同时包含治疗组和对照组的层：

$$\text{Valid Stratum } s \Leftrightarrow n_{1s} > 0 \text{ and } n_{0s} > 0$$

其中 n_{ts} 为层 s 中治疗组别 t 的样本数。

3.2.2 本研究设置

参数	设置值
匹配协变量	性别、年龄、是否合并脂肪肝、基线 ALT、基线 AST
粗化区间数 J	4
匹配方式	层内精确匹配

3.2.3 CEM 的优势

- 1. **模型无关性**：不依赖倾向性评分模型的正确设定
- 2. **完全平衡**：匹配后协变量在组间完全平衡
- 3. **透明性**：匹配过程直观可解释

3.3 倾向性评分匹配（PSM）

3.3.1 倾向性评分定义

倾向性评分是给定协变量 X 下接受治疗的概率：

$$e(X_i) = P(T_i = 1|X_i) = \frac{\exp(\beta^T X_i)}{1 + \exp(\beta^T X_i)}$$

使用逻辑回归模型估计：

$$\log\left(\frac{e(X)}{1 - e(X)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

3.3.2 最近邻匹配

对每个治疗组样本 i ，在对照组中找到倾向性评分最接近的样本 j ：

$$j^*(i) = \arg \min_{j \in \text{Control}} |e(X_i) - e(X_j)|$$

3.3.3 卡尺 (Caliper) 约束

只有当距离小于卡尺阈值时才匹配成功：

$$\text{Match}(i, j) \Leftrightarrow |e(X_i) - e(X_j)| \leq c \cdot \sigma_e$$

其中 $c = 0.2$, σ_e 为倾向性评分的标准差。

3.4 匹配效果评估

3.4.1 标准化均值差异 (SMD)

评估匹配前后协变量平衡性：

$$\text{SMD}_k = \frac{\bar{X}_{k,1} - \bar{X}_{k,0}}{\sqrt{\frac{S_{k,1}^2 + S_{k,0}^2}{2}}}$$

其中：- $\bar{X}_{k,t}$ ：组别 t 中协变量 k 的均值 - $S_{k,t}^2$ ：组别 t 中协变量 k 的方差

判断标准：|SMD| < 0.1 表示良好平衡

4. 统计分析方法

4.1 基线可比性检验

4.1.1 正态性检验

使用 Shapiro-Wilk 检验判断数据分布：

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

- H_0 ：数据服从正态分布
- 若 $p > 0.05$ ，不拒绝正态性假设

4.1.2 连续变量比较

正态分布 → 独立样本 t 检验：

$$t = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_0^2}{n_0}}}$$

自由度 (Welch 校正)：

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_0^2}{n_0}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_0^2/n_0)^2}{n_0 - 1}}$$

非正态分布 → Mann-Whitney U 检验：

$$U = n_1 n_0 + \frac{n_1(n_1 + 1)}{2} - R_1$$

其中 R_1 为治疗组的秩和。

4.1.3 分类变量比较

期望频数 ≥ 5 → 卡方检验：

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

其中 $E_{ij} = \frac{R_i \times C_j}{N}$

期望频数 < 5 → Fisher 精确检验：

$$p = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}}$$

4.2 肝功能差异分析

4.2.1 全人群分析

比较两组变化值的差异：

$$H_0: \mu_{\Delta,1} = \mu_{\Delta,0} \quad \text{vs} \quad H_1: \mu_{\Delta,1} \neq \mu_{\Delta,0}$$

组间差异：

$$D = \Delta \bar{Y}_1 - \Delta \bar{Y}_0$$

95% 置信区间：

$$CI_{95\%} = D \pm t_{0.975, df} \times SE_D$$

其中：

$$SE_D = \sqrt{\frac{S_{\Delta,1}^2}{n_1} + \frac{S_{\Delta,0}^2}{n_0}}$$

4.2.2 分层分析 (消除天花板效应)

按基线状态分层后独立分析：

$$D_{\text{stratum}} = \Delta\bar{Y}_{1,\text{stratum}} - \Delta\bar{Y}_{0,\text{stratum}}$$

分层策略： - 全人群 - 基线异常患者 - 基线正常患者 - 排除双正常患者（基线和终点均正常）

4.3 非劣效性检验

4.3.1 检验原理

验证治疗组效果不劣于对照组超过预设界值 δ ：

$$H_0: \mu_1 - \mu_0 \geq \delta \quad \text{vs} \quad H_1: \mu_1 - \mu_0 < \delta$$

4.3.2 非劣效界值设定

基于临床意义，设定为正常范围上限的 25%：

指标	正常上限	非劣效界值 δ
ALT	40 U/L	10 U/L
AST	35 U/L	8.75 U/L

4.3.3 判断标准

计算差异的 95%置信区间：

$$CI_{95\%} = [D - t_{0.975} \times SE_D, D + t_{0.975} \times SE_D]$$

非劣效判定：

$$\text{Non-inferior} \Leftrightarrow CI_{\text{upper}} < \delta$$

即：95%CI 上限 < 非劣效界值

4.3.4 结果解读

情况	结论
CI 上限 < 0	优效（治疗组更好）
$0 \leq$ CI 上限 < δ	非劣效
CI 上限 $\geq \delta$	结果不确定

4.4 组内前后比较

4.4.1 配对t 检验

比较同一组内治疗前后的变化：

$$t = \frac{\bar{d}}{S_d/\sqrt{n}}$$

其中 $d_i = Y_{i,后} - Y_{i,前}$, \bar{d} 为差值均值。

4.4.2 Wilcoxon 符号秩检验

非参数替代方法：

$$W = \sum_{i:d_i>0} R_i^+$$

其中 R_i^+ 为正差值的秩。

4.5 肝硬度值分析

4.5.1 组内比较

$$H_0: \mu_{\text{基线}} = \mu_{12\text{月}} \quad (\text{配对 t 检验})$$

4.5.2 组间比较

$$H_0: \Delta_{\text{治疗组}} = \Delta_{\text{对照组}}$$

使用独立样本 t 检验和 Mann-Whitney U 检验双重验证。

4.5.3 分层分析

按肝硬度基线状态分层： - 基线异常 (>7.0 kPa) - 基线正常 (≤7.0 kPa)

4.6 治愈速度分析

4.6.1 复常率计算

基线异常患者中，终点恢复正常的比例：

$$\text{复常率} = \frac{n_{\text{基线异常} \rightarrow \text{终点正常}}}{n_{\text{基线异常}}} \times 100\%$$

4.6.2 复常率组间比较

使用 Fisher 精确检验比较两组复常率：

构建 2×2 列联表：

	复常	未复常
治疗组	a	b
对照组	c	d

$$p = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

4.6.3 维持率计算

基线正常患者中，终点仍保持正常的比例：

$$\text{维持率} = \frac{n_{\text{基线正常} \rightarrow \text{终点正常}}}{n_{\text{基线正常}}} \times 100\%$$

4.7 相关性分析

4.7.1 Pearson 相关系数

适用于正态分布的连续变量：

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

4.7.2 Spearman 秩相关系数

适用于非正态分布或有序变量：

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

其中 d_i 为两变量秩次之差。

4.7.3 相关强度判断

r	
0.00 - 0.19	极弱/无
0.20 - 0.39	弱
0.40 - 0.59	中等
0.60 - 0.79	强
0.80 - 1.00	极强

4.8 多因素回归分析

4.8.1 模型设定

普通最小二乘（OLS）回归：

$$\Delta Y_i = \beta_0 + \beta_1 \cdot \text{Group}_i + \beta_2 \cdot X_{i,\text{基线}} + \beta_3 \cdot \text{Age}_i + \beta_4 \cdot \text{Sex}_i + \beta_5 \cdot \text{脂肪肝}_i + \varepsilon_i$$

其中： - ΔY_i ： 结局变量的变化值 - Group_i ： 治疗组别（1=治疗组， 0=对照组） - β_1 ： 治疗效应（核心参数）

4.8.2 参数估计

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

4.8.3 置信区间

$$CI_{95\%}(\beta_1) = \hat{\beta}_1 \pm t_{0.975,n-p} \times SE(\hat{\beta}_1)$$

其中：

$$SE(\hat{\beta}_1) = \sqrt{\hat{\sigma}^2 (X^T X)^{-1}_{11}}$$

4.8.4 模型解释力

决定系数 R^2 ：

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2}{\sum_i (Y_i - \bar{Y})^2}$$

5. 分析结果解读

5.1 结果汇总框架

分析结果层次
第 1 层：基线可比性 <ul style="list-style-type: none">└ 验证两组基线是否平衡（$P \geq 0.05$ 表示可比）
第 2 层：全人群分析 <ul style="list-style-type: none">└ 变化值组间比较（是否有差异）└ 非劣效性检验（是否不劣于对照组）
第 3 层：分层分析 <ul style="list-style-type: none">└ 基线异常亚组（消除天花板效应）└ 基线正常亚组（验证稳健性）
第 4 层：多因素分析 <ul style="list-style-type: none">└ 控制混杂因素后的独立效应

5.2 统计显著性判断

P 值范围	显著性标记	解读
P < 0.001	***	极显著
P < 0.01	**	高度显著
P < 0.05	*	显著
P ≥ 0.05	ns	不显著

5.3 临床意义评估

5.3.1 效应量 (Effect Size)

Cohen’s d:

$$d = \frac{\bar{X}_1 - \bar{X}_0}{S_{pooled}}$$

Cohen’s d	效应大小
< 0.2	微小
0.2 - 0.5	小
0.5 - 0.8	中等
> 0.8	大

5.3.2 临床相关性

即使统计学显著，也需评估： 1. 变化幅度是否达到临床意义阈值 2. 是否影响患者预后或生活质量 3. 成本效益比是否合理

5.4 证据强度评估

证据级别	标准
★★★★★ 强	全人群和分层分析均显著，多因素分析支持
★★★★☆ 较强	全人群显著，分层分析支持
★★★☆☆ 中等	仅部分亚组显著
★★☆☆☆ 弱	趋势性差异 (0.05 < P < 0.10)
★☆☆☆☆ 极弱	无统计学差异

附录 A：代码实现对照

分析步骤	对应函数	所在文件
------	------	------

分析步骤	对应函数	所在文件
数据预处理	load_and_preprocess_data()	data_preprocessing.py
CEM 匹配	perform_cem_matching()	comprehensive_analysis.py
PSM 匹配	perform_psm_matching()	comprehensive_analysis.py
基线可比性	baseline_comparability_test()	comprehensive_analysis.py
肝功能差异	liver_function_difference_analysis()	comprehensive_analysis.py
非劣效性检验	non_inferiority_test()	comprehensive_analysis.py
肝硬度分析	liver_stiffness_analysis()	comprehensive_analysis.py
治愈速度	cure_speed_analysis()	comprehensive_analysis.py
分层分析	stratified_analysis()	comprehensive_analysis.py
相关性分析	correlation_analysis()	comprehensive_analysis.py
多因素回归	multivariate_regression_analysis()	comprehensive_analysis.py

附录 B：关键公式速查

B.1 描述统计

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

B.2 假设检验

$$t = \frac{\bar{X}_1 - \bar{X}_0}{SE}, \quad \chi^2 = \sum \frac{(O - E)^2}{E}$$

B.3 置信区间

$$CI_{95\%} = \hat{\theta} \pm z_{0.975} \times SE(\hat{\theta})$$

B.4 相关系数

$$r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

B.5 回归系数

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

附录 C：软件与版本

软件/库	版本	用途
Python	3.13	编程环境
pandas	-	数据处理
numpy	-	数值计算
scipy	-	统计检验
scikit-learn	-	机器学习/匹配
statsmodels	-	回归分析
matplotlib	-	图表绑制

附录 D：参考文献

1. Iacus SM, King G, Porro G. Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*. 2012;20(1):1-24.
2. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
3. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*. 2010;25(1):1-21.
4. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*. 2011;46(3):399-424.

文档版本： 1.0

创建日期： 2025 年 12 月 22 日

分析工具： Python (pandas, numpy, scipy, sklearn, statsmodels, matplotlib)