

# 疾病在种族群体中分布的综合分析报告

## Comprehensive Analysis Report: Disease Distribution Across Ethnic Groups

基于 24,873 名个体队列的统计分析  
Statistical Analysis Based on a Cohort of 24,873 Individuals

分析日期 | **Analysis Date:** July 24, 2025  
样本量 | **Sample Size:** 15,233 (有效样本 | Valid samples)  
疾病类型 | **Disease Types:** 8 种主要疾病 | 8 Major diseases  
种族群体 | **Ethnic Groups:** 15 个群体 | 15 Groups

### 统计方法 | **Statistical Methods:**

卡方检验/Fisher 精确检验、Benjamini-Hochberg 校正、  
标准化患病率比、主成分分析

Chi-square/Fisher's exact tests, Benjamini-Hochberg correction,  
Standardized Prevalence Ratios, Principal Component Analysis

Contents

1	执行摘要   <b>Executive Summary</b>	2
1.1	主要发现   Key Findings . . . . .	2
2	研究人群特征   <b>Study Population Characteristics</b>	3
2.1	样本量分布   Sample Size Distribution . . . . .	3
2.2	年龄分布   Age Distribution . . . . .	3
3	总体疾病患病率   <b>Overall Disease Prevalence</b>	3
3.1	人群总体患病率（含 95% 置信区间）  Population-Wide Prevalence (with 95% Confidence Intervals) . . . . .	4
4	按种族群体分类的疾病患病率   <b>Disease Prevalence by Ethnic Group</b>	4
4.1	青光眼   Glaucoma . . . . .	4
4.2	年龄相关性黄斑变性   Age-related Macular Degeneration . . . . .	4
4.3	外周动脉疾病症状   Peripheral Artery Disease Symptoms . . . . .	5
5	统计分析   <b>Statistical Analysis</b>	5
5.1	假设检验   Hypothesis Testing . . . . .	5
5.2	多重检验校正   Multiple Testing Correction . . . . .	6
6	年龄调整分析   <b>Age-Adjusted Analysis</b>	6
6.1	标准化患病率比（SPR）  Standardized Prevalence Ratios (SPR) . . . . .	6
7	疾病相关性和共病模式   <b>Disease Correlation and Comorbidity Patterns</b>	6
7.1	疾病相关性   Disease Correlations . . . . .	6
7.2	共病负担   Comorbidity Burden . . . . .	7
8	主成分分析   <b>Principal Component Analysis</b>	7
9	临床意义和建议   <b>Clinical Implications and Recommendations</b>	7
9.1	针对性筛查的高优先级群体   High-Priority Groups for Targeted Screening . .	7
9.2	研究局限性   Study Limitations . . . . .	8
9.3	未来研究建议   Recommendations for Future Research . . . . .	8
10	结论   <b>Conclusions</b>	9

## 1 执行摘要 | Executive Summary

本报告对 24,873 名个体队列中疾病在不同种族群体间的患病率进行了严格的统计分析。分析涵盖八种主要健康状况：青光眼、年龄相关性黄斑变性、白内障、糖尿病（报告或血液检测以及临床标准两种定义）、缺血性心脏病、外周动脉疾病症状和早期呼吸道疾病。

This report presents a rigorous statistical analysis of disease prevalence across ethnic groups in a cohort of 24,873 individuals. The analysis examines eight major health conditions: Glaucoma, Age-related Macular Degeneration, Cataract, Diabetes (both reported/blood test and clinical criteria), Ischemic Heart Disease, Peripheral Artery Disease Symptoms, and Early Life Respiratory Disease.

### 1.1 主要发现 | Key Findings

1. 样本分布 | **Sample Distribution:** 研究人群高度不平衡，种族群体 1 占样本的 98.42% (n=14,992)，而其他 14 个种族群体合计仅占 1.58%。

The study population is highly imbalanced, with Ethnic Group 1 comprising 98.42% of the sample (n=14,992), while 14 other ethnic groups represent only 1.58% combined.

2. 统计学显著性 | **Statistical Significance:** 应用 Benjamini-Hochberg 多重检验校正后，只有外周动脉疾病症状在不同种族群体间显示出统计学显著差异（校正后  $p = 0.0037$ ）。

After applying Benjamini-Hochberg correction for multiple testing, only Peripheral Artery Disease Symptoms showed statistically significant differences across ethnic groups (adjusted  $p = 0.0037$ ).

3. 高风险群体 | **High-Risk Groups:** 尽管小样本量限制了统计效力，但几个种族群体显示出较高的疾病患病率：

- 种族群体 8：青光眼、年龄相关性黄斑变性和早期呼吸道疾病风险较高
- 种族群体 3：年龄相关性黄斑变性和缺血性心脏病风险较高
- 种族群体 11：尽管样本量小，但多种疾病患病率明显较高

Despite small sample sizes limiting statistical power, several ethnic groups showed elevated disease prevalence:

- Ethnic Group 8: Higher risk for Glaucoma, Age-related Macular Degeneration, and Early Life Respiratory Disease
- Ethnic Group 3: Higher risk for Age-related Macular Degeneration and Ischemic Heart Disease
- Ethnic Group 11: Notably high prevalence of multiple conditions despite small sample size

4. 年龄调整分析 | **Age-Adjusted Analysis:** 标准化患病率比显示，年龄分布差异部分解释了某些种族差异，特别是种族群体 8 的外周动脉疾病症状。

Standardized Prevalence Ratios revealed that age distribution differences partially explain some ethnic disparities, particularly for Peripheral Artery Disease Symptoms in Ethnic Group 8.

2 研究人群特征 | Study Population Characteristics

2.1 样本量分布 | Sample Size Distribution

分析样本总数：15,233 名个体（排除种族群体数据缺失后）  
Total analyzed sample: 15,233 individuals (after excluding missing ethnic group data)  
按样本量分类的种族群体 | Ethnic Group Classification by Sample Size:

- 大型群体 (n ≥ 100)：2 个群体 | Large groups (n ≥ 100): 2 groups
- 中型群体 (30 ≤ n < 100)：1 个群体 | Medium groups (30 ≤ n < 100): 1 group
- 小型群体 (n < 30)：12 个群体 | Small groups (n < 30): 12 groups

Table 1: 详细种族群体分布 | Detailed Ethnic Group Distribution

种族群体	样本量	百分比	Ethnic Group	Sample Size	Percentage
1	14,992	98.42%	1	14,992	98.42%
8	141	0.93%	8	141	0.93%
3	39	0.26%	3	39	0.26%
11	24	0.16%	11	24	0.16%
12	8	0.05%	12	8	0.05%
13	8	0.05%	13	8	0.05%
其他 (9 个群体)	21	0.14%	Others (9 groups)	21	0.14%

2.2 年龄分布 | Age Distribution

队列平均年龄：51.3 ± 9.2 岁  
Mean age of cohort: 51.3 ± 9.2 years  
年龄分层显示不同种族群体间存在显著差异，因此需要进行年龄调整分析以进行准确比较。  
Age stratification revealed significant variations across ethnic groups, necessitating age-adjusted analyses for accurate comparisons.

3 总体疾病患病率 | Overall Disease Prevalence

### 3.1 人群总体患病率（含 95% 置信区间） | Population-Wide Prevalence (with 95% Confidence Intervals)

Table 2: 总体疾病患病率 | Overall Disease Prevalence

疾病	患病率	95% CI	病例数	Disease	Prevalence
外周动脉疾病症状	25.32%	24.64%-26.02%	3,857	Peripheral Artery Disease	25.32%
白内障	12.82%	12.30%-13.36%	1,953	Cataract	12.82%
缺血性心脏病	9.30%	8.84%-9.77%	1,416	Ischemic Heart Disease	9.30%
糖尿病（报告或血检）	6.39%	6.02%-6.79%	974	Diabetes (Reported/Test)	6.39%
糖尿病（临床标准）	3.56%	3.28%-3.87%	543	Diabetes (Clinical)	3.56%
早期呼吸道疾病	1.89%	1.69%-2.12%	288	Early Respiratory Disease	1.89%
青光眼	1.21%	1.05%-1.39%	184	Glaucoma	1.21%
年龄相关性黄斑变性	0.91%	0.77%-1.08%	139	Age-related Macular Degeneration	0.91%

## 4 按种族群体分类的疾病患病率 | Disease Prevalence by Ethnic Group

### 4.1 青光眼 | Glaucoma

总体患病率：1.21% | Overall prevalence: 1.21%

显著发现 | Notable Findings:

- 种族群体 8：2.84%（95% CI: 1.11%-7.07%） - 比总体人群高 2.3 倍  
Ethnic Group 8: 2.84% (95% CI: 1.11%-7.07%) - 2.3x higher than overall population
- 种族群体 1：1.20%（95% CI: 1.04%-1.39%） - 与总体人群相似  
Ethnic Group 1: 1.20% (95% CI: 1.04%-1.39%) - Similar to overall population
- 多个种族群体显示 0% 患病率，但小样本量限制了解释  
Multiple ethnic groups showed 0% prevalence, though small sample sizes limit interpretation

### 4.2 年龄相关性黄斑变性 | Age-related Macular Degeneration

总体患病率：0.91% | Overall prevalence: 0.91%

显著发现 | Notable Findings:

- 种族群体 11：8.33%（2/24 例） - 比总体人群高 9.2 倍 \*  
Ethnic Group 11: 8.33% (2/24 cases) - 9.2x higher than overall population\*
- 种族群体 3：2.56%（95% CI: 0.45%-13.18%） - 比总体人群高 2.8 倍  
Ethnic Group 3: 2.56% (95% CI: 0.45%-13.18%) - 2.8x higher than overall population

- 种族群体 8: 2.13% (95% CI: 0.73%-6.07%) - 比总体人群高 2.3 倍  
Ethnic Group 8: 2.13% (95% CI: 0.73%-6.07%) - 2.3x higher than overall population
- \* 由于样本量小需谨慎解释 | *\*Interpret with caution due to small sample size*

4.3 外周动脉疾病症状 | Peripheral Artery Disease Symptoms

总体患病率: 25.32% | Overall prevalence: 25.32%

显著发现 | Notable Findings:

- 种族群体 19: 75.00% (3/4 例) - 比总体人群高 3.0 倍 \*  
Ethnic Group 19: 75.00% (3/4 cases) - 3.0x higher than overall population\*
- 种族群体 11: 33.33% (8/24 例) - 比总体人群高 1.3 倍 \*  
Ethnic Group 11: 33.33% (8/24 cases) - 1.3x higher than overall population\*
- 种族群体 1: 25.51% (95% CI: 24.81%-26.22%) - 与总体人群相似  
Ethnic Group 1: 25.51% (95% CI: 24.81%-26.22%) - Similar to overall population
- 种族群体 8: 7.09% (95% CI: 3.84%-12.73%) - 比总体人群低 0.3 倍  
Ethnic Group 8: 7.09% (95% CI: 3.84%-12.73%) - 0.3x lower than overall population

重要提示 | **Important Note:** 这是唯一在多重检验校正后显示统计学显著差异的疾病。  
This was the only disease showing statistically significant differences after multiple testing correction.

5 统计分析 | Statistical Analysis

5.1 假设检验 | Hypothesis Testing

对每种疾病在不同种族群体间进行卡方检验（或小期望计数时使用 Fisher 精确检验）。  
Chi-square tests (or Fisher’s exact tests for small expected counts) were performed for each disease across ethnic groups.

Table 3: 原始 p 值 | Original p-values

疾病   Disease	p 值   p-value
外周动脉疾病症状   Peripheral Artery Disease Symptoms	0.00046
缺血性心脏病   Ischemic Heart Disease	0.032
糖尿病（临床标准）   Diabetes (Clinical Criteria)	0.058
早期呼吸道疾病   Early Life Respiratory Disease	0.105
年龄相关性黄斑变性   Age-related Macular Degeneration	0.187
糖尿病（报告或血液检测）   Diabetes (Reported or Blood Test)	0.382
白内障   Cataract	0.451
青光眼   Glaucoma	0.993

5.2 多重检验校正 | Multiple Testing Correction

应用 Benjamini-Hochberg 错误发现率校正：  
Benjamini-Hochberg False Discovery Rate correction was applied:

校正后 p 值 | **Adjusted p-values:** 只有外周动脉疾病症状保持显著性（校正后 p = 0.0037）  
Only Peripheral Artery Disease Symptoms remained significant (adjusted p = 0.0037)

6 年龄调整分析 | Age-Adjusted Analysis

6.1 标准化患病率比（SPR） | Standardized Prevalence Ratios (SPR)

年龄调整揭示了重要见解：  
Age adjustment revealed important insights:

Table 4: 主要种族群体的标准化患病率比 | Standardized Prevalence Ratios for Major Ethnic Groups

种族群体   Ethnic Group	疾病   Disease	SPR	95% CI
群体 8   Group 8 (n=141)	外周动脉疾病   Peripheral Artery Disease	0.28	0.13-0.51
	青光眼   Glaucoma	2.33	0.63-5.95
	年龄相关性黄斑变性   AMD	2.39	0.49-6.98
群体 3   Group 3 (n=39)	缺血性心脏病   Ischemic Heart Disease	2.38	0.96-4.90
	年龄相关性黄斑变性   AMD	3.48	0.09-19.36

7 疾病相关性和共病模式 | Disease Correlation and Comorbidity Patterns

7.1 疾病相关性 | Disease Correlations

Pearson 相关分析显示疾病间存在适度关联：  
Pearson correlation analysis revealed modest associations between diseases:  
最强相关性 | **Strongest Correlations:**

- 糖尿病（报告） ↔ 糖尿病（临床）:  $r = 0.294$   
Diabetes (Reported) ↔ Diabetes (Clinical):  $r = 0.294$
- 白内障 ↔ 缺血性心脏病:  $r = 0.094$   
Cataract ↔ Ischemic Heart Disease:  $r = 0.094$
- 白内障 ↔ 青光眼:  $r = 0.087$   
Cataract ↔ Glaucoma:  $r = 0.087$



- 白内障 ↔ 糖尿病（临床）:  $r = 0.079$   
Cataract ↔ Diabetes (Clinical):  $r = 0.079$

7.2 共病负担 | Comorbidity Burden

多种疾病共存分析:  
Analysis of multiple disease occurrence:

Table 5: 主要种族群体的共病分布 | Comorbidity Distribution in Major Ethnic Groups

疾病数量   Number of Diseases	种族群体 1   Ethnic Group 1 (n=14,992)		种族群体 8   Ethnic Group 8	
	人数   Count	百分比   %	人数   Count	百分比   %
无疾病   No diseases	8,255	55.06	89	63.1
1 种疾病   1 disease	4,836	32.26	38	26.9
2 种疾病   2 diseases	1,441	9.61	12	8.5
3 种以上疾病   3+ diseases	460	3.06	2	1.4

8 主成分分析 | Principal Component Analysis

对具有足够样本量的种族群体疾病模式进行 PCA 分析显示:  
PCA of disease patterns across ethnic groups with adequate sample sizes revealed:

- **PC1 (72.01% 方差):** 主要由糖尿病患病率驱动，与呼吸道疾病呈负相关  
**PC1 (72.01% variance):** Primarily driven by diabetes prevalence and inversely by respiratory disease
- **PC2 (27.99% 方差):** 主要由外周动脉疾病症状驱动  
**PC2 (27.99% variance):** Primarily driven by peripheral artery disease symptoms

这表明不同种族群体间存在两种主要的疾病聚集模式。  
This suggests two major patterns of disease clustering across ethnic groups.

9 临床意义和建议 | Clinical Implications and Recommendations

9.1 针对性筛查的高优先级群体 | High-Priority Groups for Targeted Screening

基于升高的标准化患病率比:  
Based on elevated standardized prevalence ratios:



Table 6: 高风险群体筛查建议 | Screening Recommendations for High-Risk Groups

种族群体   Ethnic Group	建议筛查疾病   Recommended Screening	备注   Notes
种族群体 8   Ethnic Group 8	青光眼、年龄相关性黄斑变性   Glaucoma, Age-related Macular Degeneration	考虑遗传或环境因素   Consider genetic/environmental factors
种族群体 3   Ethnic Group 3	缺血性心脏病、年龄相关性黄斑变性   Ischemic Heart Disease, AMD	需要进一步调查   Warrants further investigation

9.2 研究局限性 | Study Limitations

1. 样本量不平衡 | **Sample Size Imbalance:** 98.42% 的参与者属于种族群体 1，严重限制了少数群体比较的统计效力
- 98.42% of participants belong to Ethnic Group 1, severely limiting statistical power for minority group comparisons
2. 小样本量 | **Small Sample Sizes:** 15 个种族群体中有 12 个  $n < 30$ ，使可靠的统计推断具有挑战性
- 12 of 15 ethnic groups have  $n < 30$ , making reliable statistical inference challenging
3. 多重比较 | **Multiple Comparisons:** 尽管进行了校正，大量检验增加了虚假发现的风险
- Despite correction, the large number of tests increases the risk of spurious findings
4. 缺失数据 | **Missing Data:** 原始样本中约 38.8% 缺少种族群体数据
- Approximately 38.8% of original sample had missing ethnic group data

9.3 未来研究建议 | Recommendations for Future Research

1. 过度抽样 | **Oversampling:** 未来研究应对少数种族群体进行过度抽样以获得足够的统计效力
- Future studies should oversample minority ethnic groups to achieve adequate statistical power
2. 遗传学研究 | **Genetic Studies:** 调查可能解释特定种族群体疾病风险升高的遗传标记
- Investigate genetic markers that may explain elevated disease risk in specific ethnic groups
3. 环境因素 | **Environmental Factors:** 检查可能导致种族差异的社会经济、饮食和生活方式因素

Examine socioeconomic, dietary, and lifestyle factors that may contribute to ethnic disparities

4. 纵向设计 | **Longitudinal Design:** 实施前瞻性队列研究以建立时间关系

Implement prospective cohort studies to establish temporal relationships

5. 验证研究 | **Validation Studies:** 在种族群体代表性更好的独立队列中重复发现

Replicate findings in independent cohorts with better ethnic group representation

## 10 结论 | Conclusions

这项综合分析揭示了不同种族群体间疾病患病率的显著差异，尽管解释受到严重样本量不平衡的限制。外周动脉疾病症状显示了种族差异的最有力证据，种族群体 1 显示较高患病率，而种族群体 8 在年龄调整后显示保护效应。

This comprehensive analysis reveals significant variations in disease prevalence across ethnic groups, though interpretation is limited by severe sample size imbalances. Peripheral Artery Disease Symptoms showed the most robust evidence of ethnic variation, with Ethnic Group 1 showing higher prevalence and Ethnic Group 8 showing protective effects after age adjustment.

虽然几个少数种族群体在各种疾病中显示出显著升高的患病率，但这些发现需要在更大、更平衡的队列中进行验证。识别出的高风险群体，特别是种族群体 8 和 3，需要有针对性的公共卫生干预和对潜在机制的进一步调查。

While several minority ethnic groups showed dramatically elevated prevalence for various conditions, these findings require validation in larger, more balanced cohorts. The identification of high-risk groups, particularly Ethnic Groups 8 and 3, warrants targeted public health interventions and further investigation into underlying mechanisms.

研究人群中种族群体 1 的强势主导地位突出了更具包容性的研究设计的关键需求，以充分代表种族多样性，从而有效解决健康差异问题。

The strong predominance of Ethnic Group 1 in the study population highlights the critical need for more inclusive research designs that adequately represent ethnic diversity to address health disparities effectively.

**报告编制信息 | Report Information**

报告编制日期 | **Report Date:** July 24, 2025  
分析参与者总数 | **Total Participants:** 15,233  
有效种族群体 | **Valid Ethnic Groups:** 15  
分析疾病数量 | **Diseases Analyzed:** 8

**统计方法 | Statistical Methods:**

卡方检验/Fisher 精确检验配合 Benjamini-Hochberg 校正、  
标准化患病率比、主成分分析  
Chi-square/Fisher's exact tests with Benjamini-Hochberg correction,  
Standardized Prevalence Ratios, Principal Component Analysis