**Project Description**

This repository contains statistical analysis scripts used in the study of associations between blood cell markers and depression status based on repeated physical examination records.

**Data:** This dataset includes **repeated measurements from participants** undergoing routine health check-ups. Each row represents one visit by a participant.

**Key variables include:**

| **Variable** | **Description** |
| --- | --- |
| dah | Unique identifier for each participant. All rows with the same dah refer to repeat visits by the same individual. |
| jcsj | Examination date (timestamp) of the physical check-up. |
| status | Binary indicator of depression diagnosis status at the next visit: typically coded as 0 = no depression, 1 = depression. |
| class | Ordinal categorical variable indicating depression severity at the next visit, with levels: no depression, mild, moderate, and severe. |
| Other columns | Include blood cell parameters and derived ratios (e.g., PLR, MLR, NLR), demographic data, and health behaviors (e.g., smoking, drinking). |

**indicator file (e.g., indicator\_name.xlsx)**

This file contains a list of biomarker variables analyzed in the study.

**Columns typically include:**

* Indicator name in Chinese
* English translation (optional)
* Abbreviations (e.g., PLR = Platelet-to-Lymphocyte Ratio)
* Units (if applicable)

**🧪 Analysis Overview**

The scripts in this repository perform the following types of analyses:

* **Poisson regression**: Used to model incidence rate and calculate dispersion statistics.
* **Mixed-effects models**: Account for within-person correlation in repeated measures.
* **Logistic regression**: Association of biomarkers with binary depression status.
* **Ordinal Logistic regression**: Associations with depression severity.
* **Trajectory modeling**: Latent class modeling to identify longitudinal biomarker patterns over time.