

Supplementary Material –R Code for Reproducible Age–Period–Cohort, Incidence Rate Ratio, and Trend Analyses.

The following commands illustrate the general approach used. All inputs correspond to study-specific data frames, and variable names are as described in the Supplementary Methods.

```

# Load required packages
library(Epi)           # For APC modeling
library(MASS)          # For negative binomial regression
library(AER)           # For overdispersion testing
library(ggplot2)       # For plots
library(dplyr)         # For data manipulation
# -----
# Age-Period-Cohort (APC) Analysis
# -----
# Prepare data: age, period, cohort, cases, population
apc_data <- data.frame(
  age = ...,
  period = ...,
  cohort = ...,
  cases = ...,
  population = ...
)
# Fit Poisson APC model with offset
apc_model <- glm(cases ~ factor(age) + factor(period) + factor(cohort),
  family = poisson(link = "log"),
  offset = log(population),
  data = apc_data)
# Test for overdispersion
dispersiontest(apc_model) # If overdispersion -> refit with quasi-Poisson
apc_model_qp <- glm(cases ~ factor(age) + factor(period) + factor(cohort),
  family = quasipoisson(link = "log"),
  offset = log(population),
  data = apc_data)
# Alternatively, Negative Binomial
apc_model_nb <- glm.nb(cases ~ factor(age) + factor(period) + factor(cohort) +
  offset(log(population)),
  data = apc_data)
# Extract estimated RRs and 95% CIs
exp(cbind(Estimate = coef(apc_model_nb), confint(apc_model_nb)))
# -----
# Incidence Rate Ratios (IRRs) for Supplementary Table 7
# -----
library(emmeans) # Required for calculating marginal means and stratum-specific IRRs

# NOTE: The data must be aggregated into 9-year periods (1996-2004, 2005-2013, 2014-2022)
# The data frame subsets (data_lymphoma_9y, data_cns_9y, data_sns_9y) must contain:
# 'Cases' (raw integer counts), 'Age_Group', 'Period', and 'log_PYAR' (offset).

# --- 1. Lymphoma Model (Negative Binomial) ---
# Used for overdispersed data where convergence was achieved.
model_lymphoma_9y <- glm.nb(
  Cases ~ Age_Group * Period + offset(log_PYAR),
  data = data_lymphoma_9y
)

# --- 2. CNS Model (Standard Poisson) ---
# Used due to severe data sparsity preventing Negative Binomial convergence.
model_cns_9y <- glm(
  Cases ~ Age_Group * Period + offset(log_PYAR),

```

```

    data = data_cns_9y,
    family = poisson(link = "log")
)

# --- 3. SNS Model (Standard Poisson) ---
# Used due to severe data sparsity preventing Negative Binomial convergence.
model_sns_9y <- glm(
  Cases ~ Age_Group * Period + offset(log_PYAR),
  data = data_sns_9y,
  family = poisson(link = "log")
)

# --- 4. Extraction of Stratum-Specific IRRs and 95% CIs ---

# Define the reference rate (Rate for Age Group 0-4 y, Period 1996-2004)
# Note: The 'emmeans' function automatically calculates model-based rates (type="response").

# Example extraction for Lymphoma:
emm_lymphoma <- emmeans(model_lymphoma_9y, ~ Age_Group | Period, type = "response")
ref_rate_lymphoma <- as.data.frame(emm_lymphoma)[1, 3] # Assumes first row is reference
irr_lymphoma_table <- as.data.frame(emm_lymphoma) %>%
  mutate(
    IRR = rate / ref_rate_lymphoma,
    Lower_95CI = asymp.LCL / ref_rate_lymphoma,
    Upper_95CI = asymp.UCL / ref_rate_lymphoma
  ) %>%
  select(-rate, -SE, -df, -asymp.LCL, -asymp.UCL)

print(irr_lymphoma_table) # -----
# Mann-Kendall Trend Test
# -----
library(Kendall)
mk_test <- MannKendall(apc_data$cases)
mk_test
# -----
# Visualization
# -----
# Example: IRR with 95% CI bands
ggplot(apc_data, aes(x = period, y = cases/population)) +
  geom_line() +
  geom_ribbon(aes(ymin = lower_CI, ymax = upper_CI), alpha = 0.2) +
  theme_minimal()

```