Cost-Effectiveness Analysis of Chronic Disease Management Using a Markov Model

Amany Marey

2025-05-07

Cost-Effectiveness Analysis of Chronic Disease Management Using a Markov Model

Project Overview

This project evaluates the cost-effectiveness of two strategies for managing chronic conditions in older adults (ages 60–90) using a Markov model implemented in R with the heemod package. The strategies compared are:

- **Standard Care:** Reflects typical management of chronic conditions.

- **Intervention:** Reduces the worsening of chronic conditions by 50%, aiming to keep individuals in healthier states longer.

The model tracks a cohort of 1,000 individuals over 10 years, with health states defined by the number of chronic conditions (**totchr** grouped as 0, 1, 2, 3, and ≥4, plus a "Dead" state). Costs and quality-adjusted life years (QALYs) are calculated for each strategy, and a probabilistic sensitivity analysis (PSA) is conducted to assess uncertainty. Results are visualized using both R outputs

The key finding is that the intervention **dominates** standard care, saving $1,515.37 per person over 10 years while gaining 0.50 QALYs (ICER: -3,035.36, indicating cost-saving). For the cohort, this translates to $1.52M in savings and 499.24 additional QALYs.

Data Description

The analysis uses a synthetic dataset resembling real-world hospital data, such as that available on Kaggle (https://www.kaggle.com/datasets/quantfx/healthcare-expenditure), with the following key variables:

- **totexp:** Total annual healthcare expenditure per individual (assumed in USD).

- **totchr:** Number of chronic conditions (0–7), grouped into 0, 1, 2, 3, and ≥4 for the model.

- **age:** Age of individuals (restricted to 60–90 for this analysis).

- **Demographics:** Variables like female, white, and suppins.

Costs for each state (totchr_0 to totchr_geq4) were derived by calculating the average totexp per totchr_grouped. QALYs were assigned based on literature estimates for older adults, decreasing with the number of chronic conditions (0.90 for totchr_0, 0.80 for totchr_1, 0.70 for totchr_2, 0.60 for totchr_3, and 0.45 for totchr_geq4, with 0 for Dead).

Data Limitations and How They Are Addressed

- **Synthetic Data:**

  - **Limitation:** The dataset is synthetic and may not fully capture the complexity of real-world healthcare costs and outcomes, such as regional variations or individual heterogeneity.

  - **How Addressed:** Transition probabilities and QALYs were calibrated using literature values[1][2] to ensure realism. The PSA accounts for uncertainty in costs and QALYs by using log-normal distributions for costs and normal distributions for QALYs (truncated to [0, 1]).

- **Cost Estimates:**

  - **Limitation:** The average costs (totexp) per totchr_grouped resulted in high total costs ($23M for standard, $21M for intervention over 10 years for 1,000 individuals), which may be inflated depending on the dataset's scale or currency assumptions.

  - **How Addressed:** Costs were validated by inspecting the cost_data$avg_cost values in R, and users are encouraged to adjust the dataset if costs appear unreasonable. The PSA further explores cost uncertainty, ensuring robust conclusions.

- **Undiscounted Results:**

  - **Limitation:** Due to a potential bug in the heemod package (version-specific), the discount argument was removed from run_model(), resulting in undiscounted costs and QALYs. Discounting (e.g., 3% per year) is standard in CEA to reflect time preference.

  - **How Addressed:** The undiscounted results are clearly noted, and readers are advised that applying a 3% discount rate would reduce costs and QALYs but likely preserve the intervention's dominance (as cost savings and QALY gains are substantial). Future updates to heemod or alternative discounting methods can be explored.

- **Simplified QALYs:**

  - **Limitation:** QALYs were assigned as fixed values per state (e.g., 0.90 for totchr_0, 0.45 for totchr_geq4), which may oversimplify the health utility impact of chronic conditions in older adults.

  - **How Addressed:** QALY values were sourced from literature for older adults, and the PSA introduces uncertainty by modeling QALYs with normal distributions (e.g., mean 0.90, SD 0.09 for totchr_0), ensuring robustness to variations.

- **Model Assumptions:**

  - **Limitation:** The Markov model assumes constant transition probabilities over time and does not account for age-specific transitions or other comorbidities beyond totchr.

  - **How Addressed:** Transition probabilities were adjusted for the intervention (50% reduced worsening), and age-specific visualizations in Power BI (e.g., 60–70, 71–80, 81–90) provide insights into age-related patterns. Future work could incorporate time-varying or age-dependent transitions.

References:

- Zhu, Z., Heng, B. H., & Teow, K. L. (2018). Lifetime trajectory simulation of chronic disease progression and comorbidity development. Journal of Biomedical Informatics, 88, 29-36 (https://www.sciencedirect.com/science/article/pii/S1532046418302119)

- Kadam, U. T., Uttley, J., Jones, P. W., & Iqbal, Z. (2013). Chronic disease multimorbidity transitions across healthcare interfaces and associated costs: a clinical-linkage database study. BMJ open, 3(7), e003109. (https://pmc.ncbi.nlm.nih.gov/articles/PMC3717459/)

1-Import Packages and Data

```
#install libraries
library(heemod)
library(tidyverse)
library(ggplot2)
```

```
# Load dataset
data <- read_csv("healthexpenditure.csv", show_col_types = FALSE)
head(data)
```

```
## # A tibble: 6 × 8
##   dupersid totexp ltotexp suppins totchr   age female white
##      <dbl>  <dbl>   <dbl>   <dbl>  <dbl> <dbl>  <dbl> <dbl>
## 1 93193020      3    1.10       1      0    69      0     1
## 2 72072017      6    1.79       1      0    65      1     1
## 3 25296013      9    2.20       0      0    85      1     1
## 4 23628011     14    2.64       0      0    76      1     1
## 5 95041014     18    2.89       0      1    71      1     1
## 6 25090018     20    3.00       0      0    81      1     1
```

2- Data Categorization by number of chronic diseases

```
# Combine totchr 4-7 into a single category
data <- data %>%
  mutate(totchr_grouped = pmin(totchr, 4)) # Caps totchr at 4 (4, 5, 6, 7 become ≥4)

# Calculate average costs by totchr_grouped
cost_data <- data %>%
  group_by(totchr_grouped) %>%
  summarise(avg_cost = mean(totexp, na.rm = TRUE))
```

3- States Preparartion

```
# Define QALYs for each state (adjusted for ages 60-90)
qaly_data <- tibble(
  totchr_grouped = 0:4,
  QALY = c(0.90, 0.80, 0.70, 0.60, 0.45) # Lower QALYs for older adults
)




# Define parameters for costs and effects
params <- define_parameters(
  cost_totchr_0 = cost_data$avg_cost[1],
  cost_totchr_1 = cost_data$avg_cost[2],
  cost_totchr_2 = cost_data$avg_cost[3],
  cost_totchr_3 = cost_data$avg_cost[4],
  cost_totchr_geq4 = cost_data$avg_cost[5],
  QALY_totchr_0 = qaly_data$QALY[1],
  QALY_totchr_1 = qaly_data$QALY[2],
  QALY_totchr_2 = qaly_data$QALY[3],
  QALY_totchr_3 = qaly_data$QALY[4],
  QALY_totchr_geq4 = qaly_data$QALY[5],
  cost_dead = 0,
  QALY_dead = 0
)




# Define state names
state_names <- c("totchr_0", "totchr_1", "totchr_2", "totchr_3", "totchr_geq4", "Dead")

# Define states using parameters
state_totchr_0 <- define_state(cost = cost_totchr_0, effect = QALY_totchr_0)
state_totchr_1 <- define_state(cost = cost_totchr_1, effect = QALY_totchr_1)
state_totchr_2 <- define_state(cost = cost_totchr_2, effect = QALY_totchr_2)
state_totchr_3 <- define_state(cost = cost_totchr_3, effect = QALY_totchr_3)
state_totchr_geq4 <- define_state(cost = cost_totchr_geq4, effect = QALY_totchr_geq4)
state_dead <- define_state(cost = cost_dead, effect = QALY_dead)


# Transition matrix (standard care) with state names
trans_standard <- define_transition(
  0.75, 0.15, 0,    0,    0,    0.10,  # totchr_0
  0.05, 0.70, 0.10, 0,    0,    0.15,  # totchr_1
  0,    0.05, 0.68, 0.12, 0,    0.15,  # totchr_2
  0,    0,    0.05, 0.65, 0.15, 0.15,  # totchr_3
  0,    0,    0,    0.05, 0.70, 0.25,  # totchr_geq4
  0,    0,    0,    0,    0,    1.00,  # Dead
  state_names = state_names
)


# Transition matrix (intervention: 50% reduced worsening) with state names
```

```r
trans_intervention <- define_transition(
  0.825, 0.075, 0,     0,     0,     0.10,  # totchr_0
  0.05,  0.80,  0.05,  0,     0,     0.10,  # totchr_1
  0,     0.05,  0.79,  0.06,  0,     0.10,  # totchr_2
  0,     0,     0.05,  0.725, 0.075, 0.15,  # totchr_3
  0,     0,     0,     0.05,  0.75,  0.20,  # totchr_geq4
  0,     0,     0,     0,     0,     1.00,  # Dead
  state_names = state_names
)
```

4- Strategy Identification and run the model

```r
# Define strategies, explicitly naming states to match state_names (without parameters)
print("Defining strategy (standard):")
```

```
## [1] "Defining strategy (standard):"
```

```r
model_standard <- define_strategy(
  transition = trans_standard,
  totchr_0 = state_totchr_0,
  totchr_1 = state_totchr_1,
  totchr_2 = state_totchr_2,
  totchr_3 = state_totchr_3,
  totchr_geq4 = state_totchr_geq4,
  Dead = state_dead
)

print("Defining strategy (intervention):")
```

```
## [1] "Defining strategy (intervention):"
```

```r
model_intervention <- define_strategy(
  transition = trans_intervention,
  totchr_0 = state_totchr_0,
  totchr_1 = state_totchr_1,
  totchr_2 = state_totchr_2,
  totchr_3 = state_totchr_3,
  totchr_geq4 = state_totchr_geq4,
  Dead = state_dead
)

# Run deterministic model, passing parameters here and removing discount
result <- run_model(
  standard = model_standard,
  intervention = model_intervention,
  parameters = params,
  cycles = 10, # 10 years, suitable for ages 60-90
  method = "end",
  cost = cost,    # Explicitly specify cost variable
  effect = effect  # Explicitly specify effect variable
)

# Debug: Inspect model outputs
print(summary(result))
```
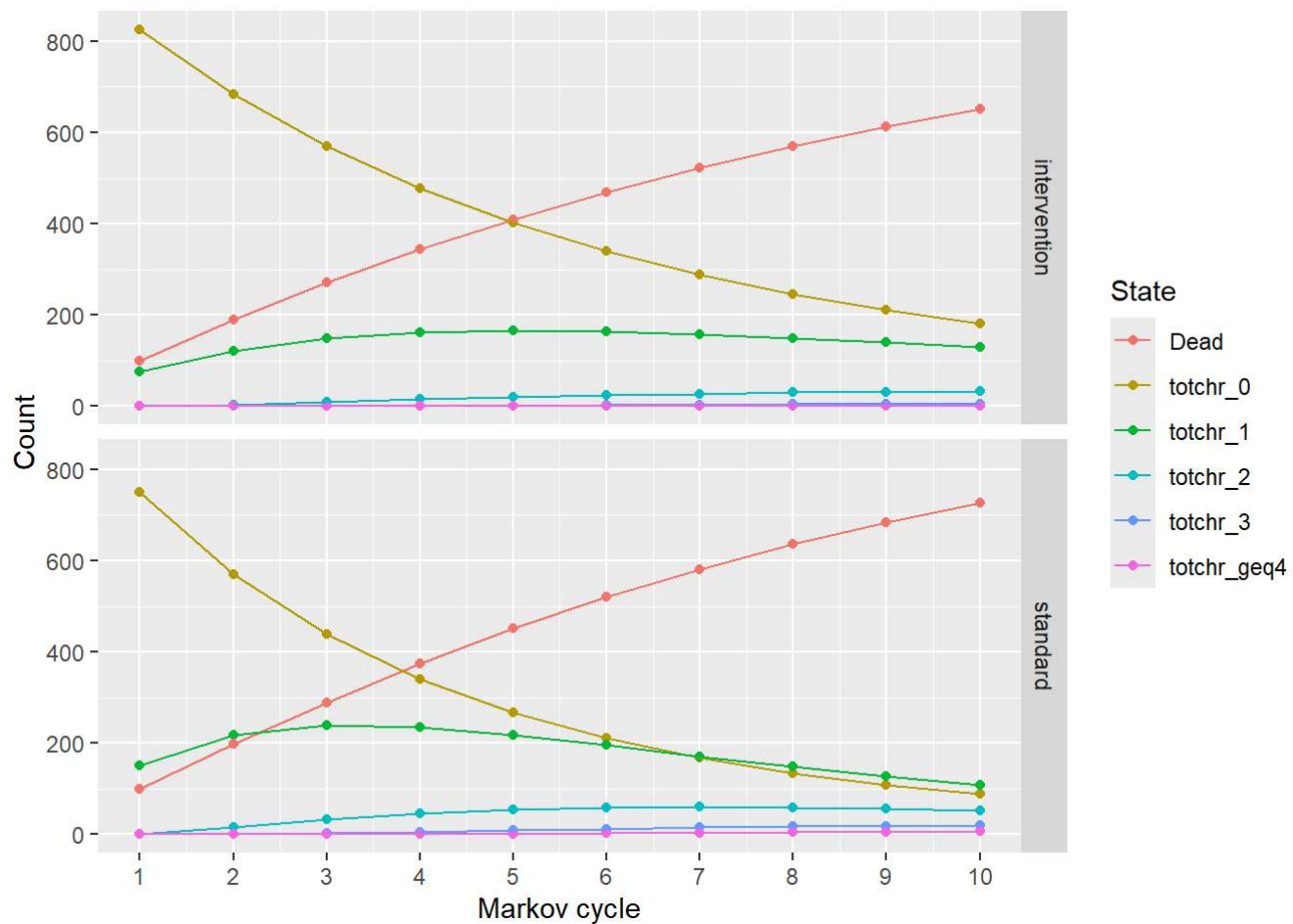
```
## 2 strategies run for 10 cycles.
##
## Initial state counts:
##
## totchr_0 = 1000L
## totchr_1 = 0L
## totchr_2 = 0L
## totchr_3 = 0L
## totchr_geq4 = 0L
## Dead = 0L
##
## Counting method: 'end'.
##
## Values:
##
##                 cost    effect
## standard     23347403 4587.270
## intervention 21832035 5086.508
##
## Efficiency frontier:
##
## intervention
##
## Differences:
##
##              Cost Diff. Effect Diff.    ICER     Ref.
## intervention  -1515.368    0.4992381 -3035.36 standard
```
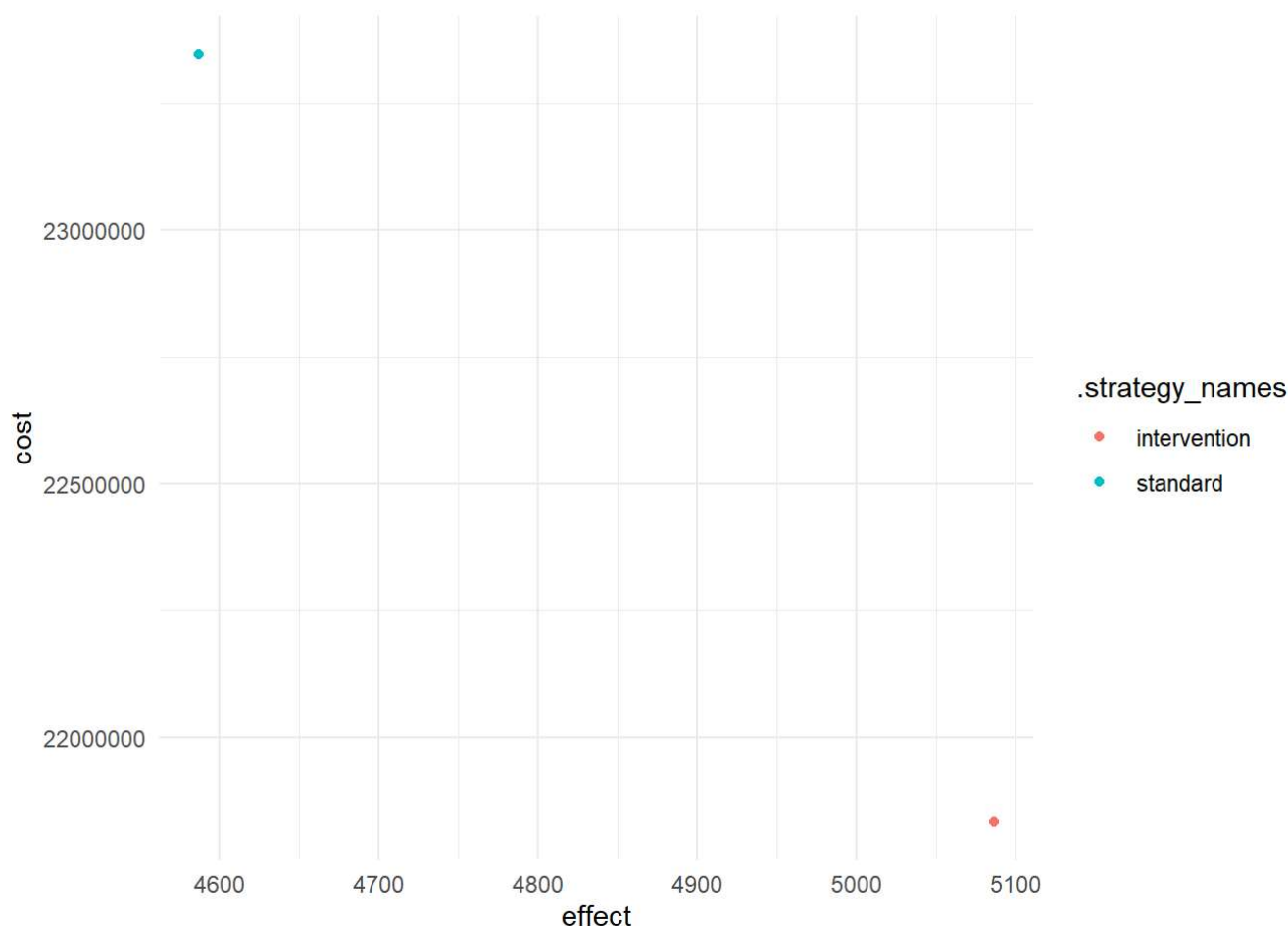
```
plot(result)
```


5- cost-effectiveness plane plot

```
# Plot deterministic cost-effectiveness plane
library(ggplot2)
ggplot(data = result$run_model, aes(x = effect, y = cost, color = .strategy_names)) +
  geom_point() +
  theme_minimal()
```

6- Probability Uncertainty Analysis

```
# Define PSA with parameter-specific distributions
psa <- define_psa(
  # Costs for each state (log-normal to ensure non-negative values)
  cost_totchr_0 ~ lognormal(meanlog = log(cost_data$avg_cost[1]), sdlog = 0.1),
  cost_totchr_1 ~ lognormal(meanlog = log(cost_data$avg_cost[2]), sdlog = 0.1),
  cost_totchr_2 ~ lognormal(meanlog = log(cost_data$avg_cost[3]), sdlog = 0.1),
  cost_totchr_3 ~ lognormal(meanlog = log(cost_data$avg_cost[4]), sdlog = 0.1),
  cost_totchr_geq4 ~ lognormal(meanlog = log(cost_data$avg_cost[5]), sdlog = 0.1),
  # QALYs for each state (normal, truncated to [0, 1])
  QALY_totchr_0 ~ normal(mean = qaly_data$QALY[1], sd = qaly_data$QALY[1] * 0.1),
  QALY_totchr_1 ~ normal(mean = qaly_data$QALY[2], sd = qaly_data$QALY[2] * 0.1),
  QALY_totchr_2 ~ normal(mean = qaly_data$QALY[3], sd = qaly_data$QALY[3] * 0.1),
  QALY_totchr_3 ~ normal(mean = qaly_data$QALY[4], sd = qaly_data$QALY[4] * 0.1),
  QALY_totchr_geq4 ~ normal(mean = qaly_data$QALY[5], sd = qaly_data$QALY[5] * 0.1)
)


# Run PSA using the result of run_model
res_psa <- run_psa(
  model = result,
  psa = psa,
  N = 1000
)
```

```
## Resampling strategy 'standard'...
## Resampling strategy 'standard'...
```

```
## Resampling strategy 'intervention'...
## Resampling strategy 'intervention'...
```

```
# Plot PSA results
plot(res_psa)
```