Cost-effectiveness analysis in clinical trials with longitudinal data: filling in some gaps

Eleanor M Pullenayegum
Post-doc
Dept of Statistics and Actuarial Sciences
University of Waterloo

Why cost-effectiveness?

- Evidence-based medicine
  - Improved Treatment
  - Increased Cost!

- How to provide the best treatment within a finite budget
  - e.g. what should be funded by OHIP?

Background

- Clinical Trials
  - Estimate difference in mean effectiveness between trial treatment and control (ΔE)
  - Estimate difference in mean cost between treatment and control (ΔC)

- Balancing cost and effectiveness
  - How to deliver best effectiveness within a finite budget?
  - How to use the data

Cost-Effectiveness for Decision-Making

- Key concept is Willingness-to-Pay (λ)
  - How much is society willing to pay for a unit increase in effectiveness?
- Calculate Incremental Net Benefit

$$\mathrm{INB}(\lambda) = \lambda\Delta\mathrm{E} - \Delta\mathrm{C}$$

- Decision Rule
  - Adopt treatment if $\mathrm{INB}(\lambda)>0$
- In practice, INB is estimated so need to quantify uncertainty

Aside – other summary statistics

- Initially, decisions were based on the Incremental Cost-Effectiveness Ratio (ICER)

$$\mathrm{ICER} = \frac{\Delta \mathrm{C}}{\Delta \mathrm{E}}$$

- Adopt trial treatment if

$$\mathrm{ICER} \leq \lambda \text{ and } \Delta \mathrm{E} > 0 \text{ or}$$

$$\mathrm{ICER} > \lambda \text{ and } \Delta \mathrm{E} < 0$$

Censoring

- Calculating differences in means sounds straightforward
- In practice, patients will often withdraw from the trial early (**censoring**)
- Patients who survive longer are more likely to be censored
- Censoring is **informative**
  - i.e. if we use information only from patients who completed the trial, we will get biased results

Cost Data

- Heavily skewed, difficult to model parametrically
  - means are sensitive to choice of distribution
- Must estimate mean, not median
- Cost data are autocorrelated
  - costs at time of censoring are correlated with costs at end of study (**informative censoring)**

A non-parametric estimator

- If we had complete data

$$\widehat{\Delta C} = \frac{1}{n_1}\sum_{i:x_i=1} C_i - \frac{1}{n_0}\sum_{i:x_i=0} C_i$$

C_i = cost incurred by patient i

$x_i = 1$ if patient i receives treatment, 0 o/w

- With censored data, we use

$$\widehat{\Delta C} = \frac{1}{n_1}\sum_{i:x_i=1} \frac{\delta_i C_i}{\pi_i} - \frac{1}{n_0}\sum_{i:x_i=0} \frac{\delta_i C_i}{\pi_i}$$

δ_i = I[patient i's cost is not censored]

$\pi_i = P(\delta_i = 1 \mid T_i, x_i)$

T_i = survival time for patient i

Reduce loss of information

- Collect cost data at multiple time points
  - Time points a_0, a_1,…, a_K
  - C_{ik} is the cost incurred by patient i in interval $[a_{k-1}, a_k)$
- Estimated difference in mean costs for interval k

$$\widehat{\Delta C}_k = \frac{1}{n_1}\sum_{i:x_i=1}\frac{\delta_{ik}C_{ik}}{\pi_{ik}} - \frac{1}{n_0}\sum_{i:x_i=0}\frac{\delta_{ik}C_{ik}}{\pi_{ik}}$$

$$\delta_{ik} = \mathrm{I}[\text{patient i's cost in interval k not censored}]$$

$$\pi_{ik} = \mathrm{P}(\delta_{ik} = 1 \mid T_i, x_i)$$

- Total difference in mean costs: sum ΔC_k over k

A semi-parametric estimator

- Can also fit a regression model

$$E(C_{ik} \mid Z_i) = Z_i' \beta_k$$

- Usually $Z_{i1}=1$ $Z_{i2}=x_i$
- Could also include other covariates (e.g. age, gender)

$$\Delta C_k = \beta_{k,2};\ \Delta C = \sum_{k=1}^{K} \beta_{k,2}$$

- Estimate β by

$$\sum_{i=1}^{n} \frac{\delta_{ik}(C_{ik} - Z_i \beta_k) Z_i}{\pi_{ik}} = 0$$

Semi-parametric efficiency?

- Semi-parametric efficient estimator solves

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\delta_{ij}h_{eff}(Z_i)_{kj}(C_{ij}-Z_i{}'\beta_j)Z_i}{\hat{G}(T_i \wedge a_j; W_{ikj})}=0$$

$$\text{where } w_{ikl}(u)=\frac{h_{eff}(Z_i)_{kl}E(C_{il}\mid \mathcal{C}_{iu})}{G(u)}\otimes Z_i,\ h_{eff}(Z)=\left(\operatorname{var}(\mathbf{C}\mid Z)-\int_0^{\tau}\frac{E(\operatorname{var}(\mathbf{C}\mid \mathcal{C}_u)\mid Z)}{G(u)^2}dG(u)\right)^{-1},$$

$\mathcal{C}_{iu}=\sigma(Z_i, T_i I(T_i \le u), C_{ik}: a_k \le u), G(u)=P(\text{censored after time u})$

$\hat{G}(;w)$ is the Cox PH estimator for G using time-dependent covariates w

$\mathbf{C}=(C_{i1}, C_{i2},\ldots, C_{iK})'$

What were we thinking??!

- Why are we treating each interval separately?
- With 10 time intervals, we would be fitting 10 models (a total of 20 parameters)

$$E(C_{i1} \mid x_i) = \beta_{1,1} + \beta_{1,2} x_i$$

$$E(C_{i2} \mid x_i) = \beta_{2,1} + \beta_{2,2} x_i$$

$$E(C_{i3} \mid x_i) = \beta_{3,1} + \beta_{3,2} x_i$$

$$E(C_{i4} \mid x_i) = \beta_{4,1} + \beta_{4,2} x_i$$

$$E(C_{i5} \mid x_i) = \beta_{5,1} + \beta_{5,2} x_i$$

$$E(C_{i6} \mid x_i) = \beta_{6,1} + \beta_{6,2} x_i$$

$$E(C_{i7} \mid x_i) = \beta_{7,1} + \beta_{7,2} x_i$$

$$E(C_{i8} \mid x_i) = \beta_{8,1} + \beta_{8,2} x_i$$

$$E(C_{i9} \mid x_i) = \beta_{9,1} + \beta_{9,2} x_i$$

$$E(C_{i10} \mid x_i) = \beta_{10,1} + \beta_{10,2} x_i$$

“Usual” models

With longitudinal data, we would usually fit

$$E(C_{ik} \mid x_i) = \beta_1 + \beta_2 x_i$$

$$E(C_{ik} \mid x_i) = (\beta_1 + \beta_3 k) + \beta_2 x_i$$

$$E(C_{ik} \mid x_i) = (\beta_1 + \beta_3 k) + (\beta_2 + \beta_4 k) x_i$$

If we were feeling adventurous, we might use smoothing functions

$$E(C_{ik} \mid x_i) = \beta_1(k) + \beta_2 x_i$$

$$E(C_{ik} \mid x_i) = \beta_1(k) + \beta_2(k) x_i$$

CIDS Trial

- 659 patients at risk of cardiac arrest
- Randomised trial of amiodarone vs. implantable cardioverter defibrillator
- Cost data from first 430 patients
- Costs collected every 90 days for 6.4 years (so K=26)
- No censoring within first 10 intervals so can sum data within first intervals to give a total of K=17 intervals.

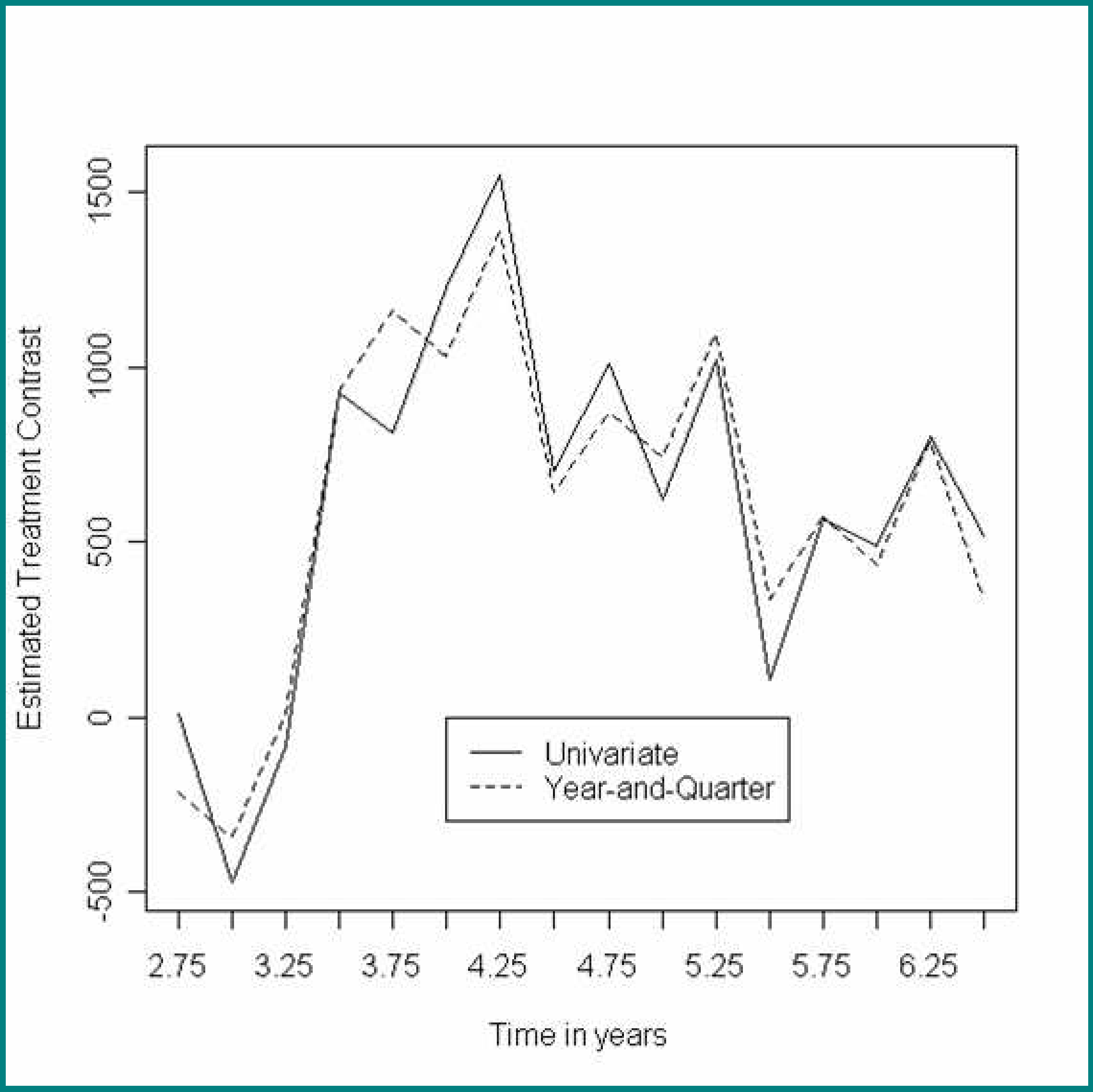
Models

- Initially use $E(C_{ik} \mid x_i) = \beta_{k,1} + \beta_{k,2} x_i$
  - The usual model for cost data (**stratified model**)
  - Estimate parameters using inverse-probability and the semi-parametric efficient estimator
- Then fit the model (for k>1) $E(C_{ik} \mid x_i) = \beta_1 + \beta_2 x_i$
  - The usual model for longitudinal data (**pooled model**)
  - Estimate parameters using inverse-probability and the semi-parametric efficient estimator

CIDS Results

Model & Estimator	Estimate	SE
Stratified IPW	48728	3970
Stratified Efficient	48300	3943
Pooled IPW	48517	3911
Pooled Efficient	40968	5869

D'OH!!



Estimated Treatment Contrast
-500
0
500
1000
1500
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
Time in years
Univariate
Year-and-Quarter

Revised Model

- Graph shows both overall trends and seasonal variation
  e.g. treatment effects larger in first quarter of the year
- Fit a revised model with treatment effect varying by year of study and quarter (for k>1).

$$
\begin{aligned}
E(C_{ik} \mid x_i) &= \alpha_{0k} \\
&+ x_i(I[k \le 4]\beta_1 + I[4 < k \le 8]\beta_2 + I[8 < k \le 12]\beta_3 + I[12 < k \le 17]\beta_4 \\
&\quad + I[k = 1,5,9,13,17]\gamma_1 + I[k = 2,6,10,14]\gamma_2 + I[k = 3,7,11,15]\gamma_3)
\end{aligned}
$$

Revised CIDS Results

Model & Estimator	Estimate	SE
Stratified IPW	48728	3970
Stratified Efficient	48300	3943
Pooled IPW	48517	3911
Pooled Efficient	40968	5869
Year-and-Quarter IPW	48726	3933
Year-and-Quarter Efficient	40514	2946

Key Points

- Where the problem originated affects how we think about it
- In this example we were thinking about estimating a difference in means rather than estimating the parameters from a model
- Developing a good model for the data is at least as important as using fancy mathematics
- Main contribution of my thesis was to apply very well known statistical ideas to a fairly well-studied estimation problem