Estimating the causal effect of HAART in the Swiss HIV cohort study

Jonathan Sterne
Department of Social Medicine,
University of Bristol UK
and
CEBU, MCRI,
Royal Children’s Hospital Melbourne

Collaborators: Matthias Egger, Miguel Hernán, James Robins, Bruno Ledergerber, Kate Tilling and the Swiss HIV Cohort Study
Outline

• Treatment of HIV infection
  – The Swiss HIV Cohort Study
• Time-dependent confounding
• Marginal structural models for causal inference in cohort studies
• Estimation of the causal effect of HAART in the Swiss HIV Cohort Study
  – Comparison with results from standard methods
Results
The % of patients who progressed to AIDS or death was lower with combination therapy (6%) than with dual therapy (11%; hazard ratio, 0.50; 95 % CI, 0.33 to 0.76; P=0.001). The hazard ratio for mortality was 0.43; 95 % CI, 0.19 to 0.99; P=0.04.

Conclusions
Treatment with indinavir, zidovudine, and lamivudine as compared with zidovudine and lamivudine alone slows the progression of HIV-1 disease in patients with 200 CD4 cells or fewer per cubic ml.
Figure 1. Kaplan-Meier Estimates of the Proportion of Patients Who Did Not Reach the Primary Study End Point of AIDS or Death.
Why do we need causal inference from HIV cohort studies?

• Limited information on clinical endpoints is available from RCTs
• Limited information on long term outcomes from RCTs
• Patients assessed in RCTs were sick, compared to the patients being treated now
• Large number of different drug combinations – few compared directly in RCTs
SHCS: some key numbers (Jan 2003)

- 12,340 patients
- 93,132 follow up visits
- 36,486 diagnoses
- 11,257 AIDS events
- 4,199 deaths
Standard approach to confounding

Control for confounding via stratification (Mantel-Haenszel methods) or using regression models
If we control for C, we will estimate only the indirect effect of T on D.
A covariate is a *time-dependent confounder* for the effect of treatment on outcome if:

1. past covariate values predict current treatment
2. current covariate value predicts outcome

Example:

1. people with low CD4 are more likely to get HAART
2. Low CD4 is a risk factor for AIDS and death

If, in addition, past treatment predicts current covariate value then standard survival analyses with time-updated treatment effects will give biased treatment effect estimates

For example, CD4 count predicts HAART and HAART raises CD4 counts
Inverse-probability weights to control confounding

• Usual approach to control of confounding in epidemiology:
  – Stratify, then combine results across strata
  – Use regression models

• Alternative:
  1. Fit a logistic regression model to estimate the probability of treatment (propensity scores), given an individual’s confounders. Then, either
  2. Stratify into quintiles based on the propensity scores, or
  3. Estimate the treatment-disease association using inverse-probability of treatment weights and use robust standard errors to
Deriving a model for HAART in the Swiss HIV Cohort Study

• Follow ups should be every 3 months, but there may be up to 5 measurements in a month, or there may be long gaps:

• Final dataset has one record (line) for every month of observation, containing:
  – first CD4/RNA/haemoglobin measurement
  – Indicators for whether the patient was treated with monotherapy / dual therapy / HAART during the month
  – Indicator for whether the patient progressed to AIDS/death

• 35 patients for whom the order of HAART date and AIDS date was unclear were excluded
Final dataset for analysis

• Follow up from on 1 January 1996 to 1 November 2002
• Patients on HAART at baseline or with AIDS before 1996 were excluded
• One record per month of observation
  – 189,200 person-months of observation
  – 500 patients progressed to AIDS or death
    • 264 events (0.31%) in 89,900 person-months not on HAART
    • 236 events (0.23%) in 104,300 person-months on HAART
Marginal structural models for causal inference

- Introduced by Robins et al. (1999)
- Stage 1: estimate each subject’s probability being treated at each time, using logistic regression
- Stage 2: use these to derive \textit{inverse probability of treatment weights} – defined as the inverse of each subject’s probability of his or her treatment history at each time
IPT weights

Notation:

\[ A(k) = \text{indicator for treatment at time } k \]

\[ L(k) = \text{value of the vector of risk factors at time } k \]

\[ \bar{L}(k - 1), \bar{A}(k - 1) = \text{treatment and covariate histories up to time } (k-1) \]

\[ iptw_i(t) = \prod_{k=0}^{t} \frac{1}{pr(A(k) = a_i(k) \mid \bar{A}(k - 1) = \bar{a}_i(k - 1), \bar{L}(k) = \bar{l}_i(k))} \]

Derived by estimating \( \Pr(A(k)=1) \) using a pooled logistic regression model (equivalent to a Cox model).
# IPT weights

Problem: large variation in the IPT weights leads to wide confidence intervals

```
. summarize weightaart, detail

weighthaart

Percentiles Smallest
1%  1.010219  1.000961
5%  1.040537  1.000985
10% 1.092107  1.001009       Obs  189215
25% 1.379526  1.001078       Sum of Wgt.  189215
50% 14.46719                      Mean           32.85658
75% 35.70359        2003.06     Largest      Std. Dev.   76.3925
90% 77.69395        2003.06     Variance  5835.813
95% 124.4434        2003.06     Skewness  10.41052
99% 307.8376        2003.06     Kurtosis  183.4149
```
Stabilised weights

Problem: large variation in the iptw weights lead to wide confidence intervals

Solution: stabilised weights

\[ sw_i(t) = \prod_{k=0}^{t} \frac{pr(A(k) = a_i(k) \mid \bar{A}(k-1) = \bar{a}_i(k-1), V = v_i)}{pr(A(k) = a_i(k) \mid \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) = \bar{l}(k))} \]

\( V = \) vector of time-independent covariates (included in \( L(0) \))
Censoring

Censoring is dealt with in an analogous way:

$$sw_i^*(t) = \prod_{k=0}^{t} \frac{pr(C(k) = 0 \mid \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1), V = v_i, T > k)}{pr(C(k) = 0 \mid \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) = \bar{l}(k), T > k)}$$

Final weight for subject i at time t is:

$$sw_i(t) \times sw_i^*(t)$$
Predictors of HAART (1)

<table>
<thead>
<tr>
<th>Current RNA group</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;400</td>
<td>1.40 (0.84,2.36)</td>
</tr>
<tr>
<td>400-1,000</td>
<td>0.75 (0.42,1.35)</td>
</tr>
<tr>
<td>&lt;10,000</td>
<td>0.86 (0.54,1.36)</td>
</tr>
<tr>
<td>&lt;100,000</td>
<td>0.90 (0.62,1.31)</td>
</tr>
<tr>
<td>≥100,000</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lagged RNA group</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;400</td>
<td>0.18 (0.11,0.31)</td>
</tr>
<tr>
<td>400-1,000</td>
<td>0.44 (0.25,0.78)</td>
</tr>
<tr>
<td>&lt;10,000</td>
<td>0.64 (0.41,1.01)</td>
</tr>
<tr>
<td>&lt;100,000</td>
<td>0.90 (0.62,1.31)</td>
</tr>
<tr>
<td>≥100,000</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CD4 group</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-49</td>
<td>1</td>
</tr>
<tr>
<td>50-99</td>
<td>0.84 (0.40,1.76)</td>
</tr>
<tr>
<td>100-199</td>
<td>0.51 (0.24,1.06)</td>
</tr>
<tr>
<td>200-349</td>
<td>0.24 (0.11,0.53)</td>
</tr>
<tr>
<td>350-499</td>
<td>0.17 (0.08,0.39)</td>
</tr>
<tr>
<td>500-749</td>
<td>0.12 (0.05,0.29)</td>
</tr>
<tr>
<td>≥750</td>
<td>0.11 (0.04,0.32)</td>
</tr>
</tbody>
</table>
### Predictors of HAART (2)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC stage B event ever</td>
<td>2.57 (1.65,3.99)</td>
</tr>
<tr>
<td>CDC stage B event in last month</td>
<td>1.27 (1.02,1.56)</td>
</tr>
<tr>
<td>Lagged monotherapy</td>
<td>2.34 (1.78,3.07)</td>
</tr>
<tr>
<td>Baseline monotherapy</td>
<td>0.76 (0.60,0.95)</td>
</tr>
<tr>
<td>Lagged dual therapy</td>
<td>1.17 (0.92,1.50)</td>
</tr>
<tr>
<td>Baseline dual therapy</td>
<td>1.17 (0.95,1.45)</td>
</tr>
<tr>
<td>Transmission group:</td>
<td></td>
</tr>
<tr>
<td>MSM</td>
<td>1</td>
</tr>
<tr>
<td>Heterosexual</td>
<td>0.87 (0.76,0.99)</td>
</tr>
<tr>
<td>IDU</td>
<td>0.60 (0.53,0.68)</td>
</tr>
<tr>
<td>Other</td>
<td>1.11 (0.87,1.43)</td>
</tr>
</tbody>
</table>

Other factors, including year after 1996, were also predictive of starting HAART.
Counterfactuals and causal inference

• Causal inference: compare what actually happened to a patient with what would have happened if – contrary to fact – they had received the alternative treatment
  – i.e. compare actual outcome with counterfactual outcome

• In RCTs, randomisation ensures that patients in the two groups are comparable, so we can use outcomes in the placebo group as a surrogate for the counterfactual outcomes in the treatment group
Marginal structural model

- Can be considered to be a causal model, in the sense that it compares what happens given your treatment history, to what would have happened in other situations
  - analogous to conducting an RCT each month, among patients still not on HAART

- Assumption: *no unmeasured confounders*

- Pooled logistic regression (equivalent to a Cox model), controlling for baseline covariates and baseline hazard, weighted by stabilised weights

\[
\text{logit Pr}[D(t) = 1 \mid D(t-1) = 0, \bar{A}(t-1), V] = \gamma_0(t) + \gamma_1 A(t-1) + \gamma_2 V
\]
## Effect of HAART, using different methods

<table>
<thead>
<tr>
<th>Model</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marginal structural model</td>
<td>0.22 (0.15,0.34)</td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
</tr>
<tr>
<td>HAART</td>
<td>0.83 (0.65,1.06)</td>
</tr>
<tr>
<td>Contr. baseline covariates</td>
<td>0.73 (0.55,0.97)</td>
</tr>
<tr>
<td>Time-updated</td>
<td></td>
</tr>
<tr>
<td>HAART</td>
<td>0.75 (0.61,0.92)</td>
</tr>
<tr>
<td>Contr. baseline covariates</td>
<td>0.42 (0.33,0.54)</td>
</tr>
<tr>
<td>Contr. time updated covariates</td>
<td>0.82 (0.61,1.10)</td>
</tr>
</tbody>
</table>
Other literature

• Cole et al. (AJE 2003;158:687-694) used MSMs to examine the effect of HAART on time to AIDS or death
  – 1,500 men and women from two American cohort studies, of whom 382 developed AIDS or died
  – Hazard ratio 0.54 (robust 95% CI 0.38 to 0.78)
  – But measures recorded in the cohort study were not the same as those used by clinicians

• Other published applications include:
  – Effect of zidovudine on mortality in HIV-infected men (Epidemiology 2000;11: 561-570)
Conclusions

• MSMs provided plausible estimates of the effect of HAART that could not be derived using standard methods
  – Context for concerns about adverse effects of HAART

• MSMs are a useful way to estimate treatment or exposure effects in the presence of time-dependent confounding:
  – Can be fitted using standard methods, and are intuitive because of analogy with RCTs
  – Focus on need to record reasons for treatment/exposure as completely as possible – strength of the Swiss HIV cohort study
  – Not a replacement for RCTs, but may supplement information from RCTs when these are difficult/impossible to do

• Need for more applications, particularly risk factors rather than treatments
References
(see James Robins’ site at www.biostat.harvard.edu/~robins/research.html)


