Meet the Panel

Moderator:
- Murray Krahn (University of Toronto)

Speakers:
- Jaime Caro (McGill University)
- Claire deOliveira (CAMH)
- Kristian Fillion (McGill University)
Session Overview

Agencies such as CADTH and INESSS have the mandate to assess health technologies in order to make them available for Canadians. Economic evaluations are a valuable tool that is used, yet many of them are carried out in situations of competing needs, uncertainty and limited knowledge. New approaches are intended to decrease the uncertainty of these evaluations and to optimize the access of new technologies, but their utilization by decision-makers might be suboptimal.
Opportunities to improve the access to new technologies, Development of new methods in health economics and their adoption in the decision-making process: DICE Simulation

J. Jaime Caro  MDCM FRCPC FACP
Adj. Prof Medicine & Epidemiology and Biostatistics, McGill University, Montreal
Chief Scientist, Evidera

(Discretely-Integrated Condition Event)
Cancer example

Happen at a point in time

Events

Cancer → Progression-free → Progression

Condition

Happen over time
What is DICE?

A modeling technique that conceptualizes the decision-analytic problem in terms of two fundamental aspects:

**Conditions**
- Aspects that persist over time
- Have levels, which can change & affect events
- Many conditions can be present at once
- Interested in time spent at a given level (value)

**Events**
- Aspects that happen at a point in time
- Can affect the level of a condition or other events
- Many can happen, at any time
- Interested in number that happen (and when)
The essentials of DICE

- **List of conditions**
  - Name (unique)
  - Level (at a given point in time)

- **List of events**
  - Name (unique)
  - Time of occurrence

- **Consequences of each event**
  - For itself (recurrence?)
  - For other events
  - For conditions

- **Discrete-integrator**
  - Read conditions list
  - Read event list
  - Process event consequences in sequence
  - End simulation & report results.
DICE Structure

1. **List of conditions**
   - **Vital status**
     - **Alive**
     - **Death**
   - **Treatment**
   - **Level**

2. **List of events**
   - **Death**
     - **Events**
       - **Name**
         - **Start**
           - Never
         - **Death**
           - Formula_{SoC}

3. **Consequences of each event**
   - **Start**
     - don’t happen again
     - select treatment = SoC
     - estimate time of death
     - set vital status = alive
   - **Death**
     - Accrue QALYs
     - Add up costs

- **Discrete-integrator**
  - Read conditions list
  - Read event list
  - Process event consequences in sequence
  - End simulation & report results.
DICE Implementation - software

- **List of conditions**
  - Name (unique)
  - Level (at a given point in time)

- **List of events**
  - Name (unique)
  - Time of occurrence

- **Consequences of each event**
  - For itself (recurrence?)
  - For other events
  - For conditions

- **Discrete-integrator**
  - Read conditions list
  - Read event list
  - Process event consequences in sequence
  - End simulation & report results.

---

**Software**

- FORTRAN
- C++
- Excel
- Python
- R
- VB.net
**DICE Excel® Implementation**

1. **List of conditions**
   - Vital status
   - Treatment

2. **List of events**
   - Death
   - Start

---

**Conditions**

<table>
<thead>
<tr>
<th>Name</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vital Status</td>
<td>Alive</td>
</tr>
<tr>
<td>Treatment</td>
<td>SoC</td>
</tr>
<tr>
<td>QALYs</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>

**Events**

<table>
<thead>
<tr>
<th>Name</th>
<th>Time To Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>Never</td>
</tr>
<tr>
<td>Death</td>
<td>Formula_{SoC}</td>
</tr>
</tbody>
</table>

*Excel "tables"*
# DICE Excel® Implementation

## 1. List of conditions
- Vital status
- Treatment

## 2. List of events
- Death
- Start

## 3. Consequences of each event
- Start
  - don’t happen again
  - select treatment = SoC
  - estimate time of death
  - set vital status = alive
- Death
  - Accrue QALYs
  - Add up costs

### Conditions

<table>
<thead>
<tr>
<th>Name</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vital Status</td>
<td>Alive</td>
</tr>
<tr>
<td>Treatment</td>
<td>SoC</td>
</tr>
<tr>
<td>QALYs</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>

### Events

<table>
<thead>
<tr>
<th>Name</th>
<th>Time To Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>Now</td>
</tr>
<tr>
<td>Death</td>
<td>Never</td>
</tr>
</tbody>
</table>

### Event

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>Start</td>
<td>Never</td>
</tr>
<tr>
<td>Condition</td>
<td>Treatment</td>
<td>SoC</td>
</tr>
<tr>
<td>Event</td>
<td>Death</td>
<td>Ln(1-rand())/-hazard</td>
</tr>
<tr>
<td>Condition</td>
<td>Vital status</td>
<td>Alive</td>
</tr>
</tbody>
</table>
Getting the expressions to do something

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>Start</td>
<td>Never</td>
</tr>
<tr>
<td>Condition</td>
<td>Treatment</td>
<td>SoC</td>
</tr>
<tr>
<td>Event</td>
<td>Death</td>
<td>(\ln(1 \text{rand}())/-\text{hazard})</td>
</tr>
<tr>
<td>Condition</td>
<td>Vital status</td>
<td>Alive</td>
</tr>
<tr>
<td>Output</td>
<td>QALYs</td>
<td>0</td>
</tr>
<tr>
<td>Output</td>
<td>Costs</td>
<td>0</td>
</tr>
</tbody>
</table>

Do

```vbnet
ThisEventName = UCase(EventTypes(Range("NextEvent"), 3))
ThisEvent = Range(EventTypes(Range("NextEvent"), 3) & "Rge")
For i = 1 To UBound(ThisEvent, 1)
    If ThisEvent(i, 3) <> "" Then
        Range("CalcCell") = "=" & ThisEvent(i, 3)
        Range(ThisEvent(i, 2)) = Range("CalcCell").Value
    End If
Next i
Range("CalcCell") = "" 'Housekeeping
Loop Until ThisEventName = "END"
```
Fatal cancer if untreated
- Both risk of progression and of death depend on age, sex and the level of a biomarker
- SoC achieves remission but eventually progresses
  - Main side-effect: neutropenia
- New treatment, *MetaMin*, increases PFS and OS
  - Main side-effect: anaphylaxis
- At progression, SoC switches to second line, *MetaMin* continues

Interest is in LYG, QALYs, Costs, ICER.
Cohort Markov approach

<table>
<thead>
<tr>
<th>Time (m)</th>
<th>Disease-free</th>
<th>Relapsed</th>
<th>Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1,000</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1000-P</td>
<td>1000 x $P_{rel}$</td>
<td>0 x $P_{death/Rel}$</td>
</tr>
<tr>
<td>2</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>300</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
## DICE Cohort Markov version

**Start**
- **Disease-free: 100%**
- **Transition: cycle**

**Transition**
- **Update Disease-free, Relapsed, Death**
- **Select next event**
- **Transition, End**

**End**
- **Report all results**

### Conditions Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease-free</td>
<td>100</td>
</tr>
<tr>
<td>Relapsed</td>
<td>0</td>
</tr>
<tr>
<td>Death</td>
<td>0</td>
</tr>
</tbody>
</table>

### Transition event Table

<table>
<thead>
<tr>
<th>Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Death</td>
<td>Death+ $P_{death} \times \text{Relapsed}$</td>
</tr>
<tr>
<td>Condition</td>
<td>Relapsed</td>
<td>$\text{Relapsed} \times (1-P_{death}) + \text{Disease-free} \times P_{rel}$</td>
</tr>
<tr>
<td>Condition</td>
<td>Disease-free</td>
<td>Disease-free $\times (1-P_{rel})$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event</th>
<th>Transition</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>Transition</td>
<td>Now + cycle</td>
</tr>
<tr>
<td>Event</td>
<td>End</td>
<td>If ((Time $\geq$ TimeHorizon), Now, Never)</td>
</tr>
</tbody>
</table>

### Constants Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TimeHorizon</td>
<td>10</td>
</tr>
<tr>
<td>Never</td>
<td>99,999</td>
</tr>
<tr>
<td>Now</td>
<td>0</td>
</tr>
<tr>
<td>$P_{rel}$</td>
<td>.015</td>
</tr>
<tr>
<td>$P_{death}$</td>
<td>.001</td>
</tr>
<tr>
<td>Cycle</td>
<td>1</td>
</tr>
</tbody>
</table>

- **Compact specification**
  - Doesn’t change if longer time horizon
  - Or shorter cycles
    - Can make these variable!

- **Expressions written only once reducing error**

- **No programming!**

- **Transparent, easy to grasp**
### DICE Microsimulation version

#### Conditions Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Initial Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease</td>
<td>Disease-free</td>
</tr>
</tbody>
</table>

#### Profiles Table

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>BioM</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>Male</td>
<td>124</td>
</tr>
<tr>
<td>45</td>
<td>Female</td>
<td>216</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

#### Transition event Table

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Disease</td>
<td>If (Disease=&quot;Disease-free&quot;,if(rand()&lt;P_{rel},“Relapsed”, “Disease-free”),If(rand()&lt;P_{death},“Dead”, “Relapsed”))</td>
</tr>
<tr>
<td>Event</td>
<td>Transition</td>
<td>Now + cycle</td>
</tr>
<tr>
<td>Event</td>
<td>End</td>
<td>If ((Time&gt;=TimeHorizon), Now, Never)</td>
</tr>
</tbody>
</table>
DICE Time-to-Event Model

**Transition event Table**

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start event Table</td>
<td></td>
<td>If (Disease ≠ “Disease-free” if (rand()) &lt; P, “Relapsed” otherwise)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>Relapse</td>
<td>-ln(1 - rand()) / hazard_rel</td>
</tr>
<tr>
<td>Event</td>
<td>Death</td>
<td>-ln(1 - rand()) / hazard_otherDeath</td>
</tr>
</tbody>
</table>

**Relapse event Table**

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Disease</td>
<td>Relapsed</td>
</tr>
<tr>
<td>Event</td>
<td>Death</td>
<td>Min(-ln(1 - rand()) / hazard_death/rel, Death)</td>
</tr>
<tr>
<td>Event</td>
<td>Treatment switch</td>
<td>Now</td>
</tr>
</tbody>
</table>

**Treatment switch event Table**
DICE outputs

Conditions Table

<table>
<thead>
<tr>
<th>Name</th>
<th>Initial Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease</td>
<td>Disease-free</td>
</tr>
</tbody>
</table>

Relapse event Table

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assigned Item</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Disease</td>
<td>Relapse</td>
</tr>
<tr>
<td>Event</td>
<td>Death</td>
<td>-ln(1-rand())/hazard_{death/relapse}</td>
</tr>
<tr>
<td>Event</td>
<td>Treatment switch</td>
<td>If (Treatment=&quot;SoC&quot;, Now, Never)</td>
</tr>
</tbody>
</table>
Advantages & limitations

- Very flexible & natural
- Can combine cohort, individual & time-to-event approaches
- Transparent, simple to communicate
- Standard framework (easy to learn)
- Less error-prone
- Enables structural sensitivity analysis
- Straightforward to review
- Fast to create, easy to modify

- Excel is slow
- No individuals, interactions
- No resources, queues
  - (lacking experience, validation, publications)
What can DICE be used for?

- Economic analyses
- Epidemiologic modeling
- Patient Flow
- RCT Simulation
- Portfolio optimization
- MCDA
DICE transforms the way we develop models

**Old way**
- Pick a technique (e.g., Markov)
- Adapt decision problem To selected technique
- Spend wks/m programming
- **Final Model**
  - Complex
  - Huge
  - Tricky to verify
  - Hard to explain
  - Forget changing structure!

**New way**
- **Focus on decision problem**
- Design model to suit decision problem
- Implement DICE in days
- **Final Model**
  - Straightforward
  - Compact
  - Easy to verify
  - Simple to explain
  - Change structure any time
Models and Methods in Health Economics and the Decision-Making Process

Claire de Oliveira, M.A. PhD
Introduction

- there are scarce health care dollars
- policy makers who fund and organize health care struggle to provide patients with latest therapies, given limited financial resources
- accurate cost estimates are required to estimate burden of care
- also, important input for health technology assessment (e.g. economic evaluations)
- thus, need to measure costs appropriately

objective of talk

- discuss new costing methods using administrative health care databases and provide examples
Road map of talk

- cost estimation using administrative data
  - required data elements
- cumulative cost functions
- methods to estimate cumulative costs
  - inverse probability weighting (IPW)
  - phases of care approach
    - net cost method
- modelling costs through regression analysis
Cost estimation using administrative data

• provinces can link population-based registries, clinical, and administrative data to measure costs
• existing costing methods follow CADTH guidelines

• measuring costs generally requires 2 types of data elements:
  • utilization (how many resources are used)
  • unit cost (how much each resource costs)
• data elements not always available in databases

• once we have cost data, we can estimate cumulative cost functions
Cost estimation using administrative data

- **both** utilization and unit costs in data:
  - physician services (billings data)
  - outpatient prescription drugs (claims data)

- utilization + **estimates** of unit costs:
  - inpatient hospitalizations → CIHI DAD
  - ED visits → CIHI NACRS

- **only** utilization (+ unit costs from other sources):
  - home care (in many provinces)
  - other care
Cumulative cost functions

- Variable of interest in health care costing studies is typically mean health cost (i.e. incidence-based costs)
  - Cumulative cost from index event over some interval
- Rate of cost accumulation tends to increase around index events, such as hospitalizations and death
- Theoretically could follow all participants until death, but death rarely observed for every patient due to short study horizons
- Portion of health care cost that is unobserved may be especially important → costs tend to rise dramatically in pre-death period
Cumulative cost functions

- typically focus on mean total costs for restricted time period → 2 major issues

- 1) among patients who die, death drives up costs in period before death, BUT costs may also be driven down because no costs are accrued after death
- to deal with this, consider death as a terminal event
- subjects will accrue costs until they die, or until they reach time horizon of analysis → complete case
- in each case, patients no longer accumulate relevant costs
Cumulative cost functions

2) however, some patients are not complete cases

a) portion of relevant health costs for these participants will be unobserved → data are right censored before outcome of interest has occurred (i.e. observation that ends prematurely)

b) study may enroll patients over a period of time but discontinue follow-up on a fixed calendar date

in both cases, censoring is random, and observed health care costs represent the lower limit of relevant costs

to adjust cumulative cost estimates for censoring → develop function that describes way in which data are censored, use reweighting
**Inverse Probability Weighting (IPW)**

- IPW reweights each complete case so that it represents not only itself but also some number of incomplete or censored cases.
- Cumulative cost of each patient who died or reached full period of observation must represent cost of that patient and of censored cases that would have been observed had there been no censoring.
- Number of censored cases represented by complete case at observation time $t$ is proportional to probability of that case being censored.
- Costs for complete cases with short follow-up should be weighted less than cases with longer observation period, accounting for higher probability of censoring with longer observation periods.
Inverse Probability Weighting (IPW)

advantages
• deals with censoring
• can be expanded within regression framework to control for covariates

disadvantages
• when evaluating covariate effects, cannot distinguish effects on cost accumulation from effects on survival
• does not account for differential rates of healthcare cost accumulation near death
Phases of care approach to estimate costs

- alternative method to estimate cumulative costs (variant of incidence cost approach)
- useful to estimate lifetime costs or costs in presence of heavy censoring
- defines time periods of interest following diagnosis where costs may differ across periods; typically 3 phases:
  - initial care phase;
  - continuing care phase;
  - terminal care phase
- used many times with net cost method, where net costs = cost difference between patients and controls
  - typically done with matching (hard and/or propensity score matching)
Phases of care - examples

FIGURE 1. Monthly costs of care for colorectal cancer patients by length of survival.

## Phases of care - examples

<table>
<thead>
<tr>
<th>Tumour Site</th>
<th>Phase, estimated cost (95% CI)</th>
<th>Pre-diagnosis (3 months)</th>
<th>Initial (6 months)</th>
<th>Continuing (annual)</th>
<th>Terminal (12 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head and neck</td>
<td>$595 ($326–$865)</td>
<td>$19,702 ($19,691–$19,714)</td>
<td>$5151 ($5143–$5159)</td>
<td>$37,346 ($37,332–$37,360)</td>
<td></td>
</tr>
<tr>
<td>Esophagus</td>
<td>$818 ($455–$1180)</td>
<td>$41,567 ($41,539–$41,596)</td>
<td>$5491 ($5474–$5509)</td>
<td>$54,354 ($54,336–$54,371)</td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>$3381 ($2906–$3855)</td>
<td>$21,355 ($21,325–$21,384)</td>
<td>$11,954 ($11,937–$11,971)</td>
<td>$30,265 ($30,242–$30,289)</td>
<td></td>
</tr>
<tr>
<td>Pancreas</td>
<td>$1892 ($1468–$2315)</td>
<td>$29,979 ($29,950–$30,008)</td>
<td>$6296 ($6272–$6319)</td>
<td>$54,152 ($54,138–$54,167)</td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td>$553 ($331–$774)</td>
<td>$4649 ($4635–$4664)</td>
<td>$4005 ($3998–$4012)</td>
<td>$18,494 ($18,479–$18,509)</td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>$637 ($375–$899)</td>
<td>$8394 ($8391–$8397)</td>
<td>$5017 ($5015–$5020)</td>
<td>$17,391 ($17,385–$17,397)</td>
<td></td>
</tr>
<tr>
<td>Renal</td>
<td>$1503 ($1111–$1895)</td>
<td>$14,950 ($14,936–$14,964)</td>
<td>$3991 ($3981–$4002)</td>
<td>$38,292 ($38,274–$38,309)</td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>$1548 ($1192–$1904)</td>
<td>$33,241 ($33,227–$33,225)</td>
<td>$6563 ($6546–$6581)</td>
<td>$72,463 ($72,444–$72,483)</td>
<td></td>
</tr>
<tr>
<td>Lymphoma</td>
<td>$1484 ($1125–$1843)</td>
<td>$17,831 ($17,820–$17,842)</td>
<td>$6276 ($6268–$6285)</td>
<td>$59,202 ($59,182–$59,222)</td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td>$1325 ($1006–$1645)</td>
<td>$18,214 ($18,194–$18,233)</td>
<td>$8035 ($8024–$8045)</td>
<td>$74,857 ($74,837–$74,877)</td>
<td></td>
</tr>
</tbody>
</table>
Phases of care approach to estimate costs

advantages

- phase-specific costs can be combined with survival data to estimate average lifetime costs from diagnosis to death
- actual costs do not have to be observed over entire period between diagnosis and death (no need to deal with censoring)
- average phase-specific costs can be estimated from only a few years of cost data \(\rightarrow\) can use only recent years of cost data, which reflect current patterns of care
- makes use of data more efficiently than a cohort approach
Phases of care approach to estimate costs

disadvantages

- makes assumptions about cost patterns, particularly during continuing care phase
- lifetime cost estimates difficult to compare among different groups due to difficulties in adjusting for differences in covariates
- need to be careful when discounting costs
- patients can contribute data to multiple phases → non-ignorable correlations between phase-specific cost estimates
- nonetheless, easily understood and straight forward estimation
Modelling costs through regression analysis

• once complete cost data available → can model costs; this can be challenging
• typically, individual-level cost data:
  • large proportion of zeroes (non-users)
  • strongly skewed distribution
  • long right-hand tail
• non-normality of data typically due to small proportion of individuals who account for large share of costs
• error term typically exhibits high degree of heteroskedasticity
• relationship between costs and regressors and/or the correct model specification generally not linear

Camh
Modelling costs through regression analysis

- generalised linear model (GLM) has become preferred strategy for modelling health care costs when there are unknown forms of heteroskedasticity

- advantages
  - predictions made on raw cost scale, thus no retransformation is required
  - allow for heteroskedasticity through choice of distributional family

- disadvantage
  - can suffer substantial loss in precision in face of heavy-tailed, log scale residuals or when variance function is misspecified

- nonetheless, GLM is generally recommended
Modelling costs through regression analysis

- which distribution to use?
- use **modified Park test** (Manning and Mullahy, 2001)
  - chooses **family distribution** based on most likely appropriate variance function
  - idea underlying: GLM distribution should reflect relationship between variance and mean
- researcher chooses link function
  - **link function** specifies shape of conditional mean function and characterises how mean is related to set of covariates
- usually family distribution is either Gamma or Poisson
- given skewness of cost data, typically chose log link
Modelling costs through regression analysis

• how to handle zeroes?
• no zero mass problem → GLM
• zero mass problem → two-part model
  • 1. pr(any expenditures) using full sample (estimate with logit or probit model)
  • 2. level of expenditures (conditional on y>0) using subsample with y>0 (estimate with appropriate continuous model → usually GLM)
Thank you.

**Contact information:**
Claire de Oliveira, M.A., PhD
claire.deoliveira@camh.ca
Confounding by Drug Formulary Restriction in Pharmacoepidemiologic Research

Kristian B. Filion PhD FAHA

Departments of Medicine and of Epidemiology, Biostatistics, and Occupational Health
McGill University
Lady Davis Institute, Jewish General Hospital
Montreal, Quebec, Canada
Disclosures

• This study was funded by an operating grant from the Canadian Institutes of Health Research (CIHR).

• No relationships to disclose.
Introduction

• Pharmacoeconomic studies
  • Often rely on model inputs obtained from pharmacoepidemiologic research, especially when considering drug effectiveness and safety in a real-world setting.
  • Are affected by the limitations of the studies used to estimate model inputs.

• Although much attention has focused on traditional epidemiologic biases (i.e., selection bias, information bias, confounding), the impact of biases particular to pharmacoepidemiology should also be considered.
Formulary Restrictions

• The prescription of newly-marketed medications is often limited by formulary restrictions.

• Recent analyses have revealed that formulary restrictions can introduce important exposure misclassification in administrative databases due to incomplete data capture.

• However, the potential consequences of confounding due to drug formulary restrictions in pharmacoepidemiologic research remain incompletely understood.
Objective

To illustrate this potential bias using the example of fluticasone/salmeterol combination therapy (Advair©), an oral inhaler used for the treatment of asthma and chronic obstructive pulmonary disease (COPD), whose use is restricted in the province of Quebec, Canada.
Study Population

- Quebec’s administrative databases
  - Hospitalizations, physician visits, and vital statistics for Quebec residents
  - Drugs dispensed for those who participate in the provincial drug plan (restricted to those with coverage)

- Retrospective cohort study of new users of respiratory medications
  - Any bronchodilator, inhaled corticosteroid, cromone, or anti-leukotriene (January 1, 1990 - December 31, 2005)

- Identified new users of fluticasone/salmeterol combination and excluded:
  - Aged < 18 years
  - < 1 year of continuous drug coverage
  - Dispensed fluticasone and salmeterol as 2 separate drugs on the same day in the year before cohort entry
Formulary Restrictions

• Fluticasone/salmeterol (inhaled corticosteroid/long-acting β2 adrenergic receptor agonist) combination therapy became restricted in October 2003
  - Over-use as 1st line therapy by general practitioners
  - Restricted to patients with 1) asthma or other reversible obstructive diseases who remained poorly controlled despite their use of an inhaled corticosteroid; or 2) patients with moderate or severe COPD with an exacerbation in the last year despite regular use of a long-acting bronchodilator

• Classified according to initial dispensation:
  - Liberal period: September 1999 to September 2003
  - Restricted period: January 2004 to October 2006
  - Excluded: October 2003 to December 2003
Endpoints and Follow-up

• Primary outcome:
  - First hospitalization for respiratory causes

• Secondary outcomes:
  - Hospitalization for any cause
  - All-cause mortality

• Censoring:
  - Death (hospitalization endpoints)
  - End of continuous drug coverage (7-day grace period)
  - Departure from the database
  - End of the 12-month follow-up
  - End of the study period (March 31st, 2007)
Statistical Analyses

- Prescription rate and rate of new use:
  - Interrupted time-series analyses with autocorrelation parameters included in segmented regression models
  - Predicted rate in the absence of formulary change

- 12-month event rates:
  - Estimated using Poisson regression
  - Periods compared using six Cox proportional hazards models with increasing level of statistical adjustment
## Statistical Analyses (Con’t)

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude</td>
<td>None</td>
</tr>
<tr>
<td>Age- and sex adjusted</td>
<td>Age, sex</td>
</tr>
<tr>
<td>Partially-adjusted model 1</td>
<td>+ hospitalization for asthma, COPD, or pneumonia in the year before cohort entry</td>
</tr>
<tr>
<td>Partially-adjusted model 2</td>
<td>+ <strong>measures of overall</strong> (number of hospitalizations for any cause in the year before cohort entry, number of physician visits in the year before cohort entry, number of prescription drugs dispensed in the prior year, Charlson comorbidity index) <strong>and respiratory health</strong> (a history of asthma, COPD, or pneumonia) in the prior year, as well as <strong>non-respiratory drugs</strong> (narcotics, NSAIDs, non-topical antibiotics, oral corticosteroids, and statins)</td>
</tr>
<tr>
<td>Partially-adjusted model 3</td>
<td>+ <strong>respiratory drugs</strong> (bronchodilators, short-acting beta-agonists, long-acting beta-agonists, long-acting beta-agonists with inhaled corticosteroid, short-acting beta-agonists with anticholinergic combination, cromoglycates, inhaled corticosteroids, inhaled corticosteroid/bronchodilator, leukotrienes, and xanthines)</td>
</tr>
<tr>
<td>Fully-adjusted</td>
<td>+ <strong>prescribing physician specialty</strong> (respirologist, general practitioner, and other), dispensing during the <strong>summer</strong> months (April – September)</td>
</tr>
</tbody>
</table>
Individuals available in RAMQ database
(n=1,410,211)

Age-eligible individuals
(n=1,029,393)

Fluticasone/salmeterol dispensing among eligible individuals
(n=1,618,818)

- Excluded: Individuals aged < 18 years during study period (n=380,818)

Individuals receiving dispensing of fluticasone/salmeterol during the study period with no previous dispensing of fluticasone/salmeterol in the prior year
(n=108,879)

- Excluded: Dispensing not in the study period (n=323,434)
- Dispensing of fluticasone/salmeterol within the prior year (n=1,186,505)

New users of fluticasone/salmeterol in the study period
(n=90,103)

- Excluded: Less than one year of continuous drug coverage before dispensing (n=9,990)

New users of fluticasone/salmeterol included in study population
(n=77,212)

- Excluded: Dispensing of both fluticasone and salmeterol on the same day in the year prior (n=12,768)
- First dispensing of fluticasone/salmeterol between September 30th, 2003 and January 1st, 2004 (n=54)
- No follow-up (n=69)
Prescription Rate

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-47.1 (-333.61, 239.4)</td>
</tr>
<tr>
<td>Time in Months</td>
<td>82.5 (76.66, 88.3)</td>
</tr>
<tr>
<td>Indicator for Restriction</td>
<td>-1068 (-1213.29, -923.1)</td>
</tr>
<tr>
<td>Months Since Restriction</td>
<td>-70.7 (-80.7, -60.7)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Note: Denominator = all patients prescribed a respiratory drug.
Rate of New Use

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>133.7 (76.1, 191.3)</td>
</tr>
<tr>
<td>Time in Months</td>
<td>7.1 (5.1, 9.2)</td>
</tr>
<tr>
<td>Indicator for Restriction</td>
<td>-444.3 (-496.7, -391.8)</td>
</tr>
<tr>
<td>Months Since Restriction</td>
<td>-4.2 (-7.4, -1.0)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.95</td>
</tr>
</tbody>
</table>
## Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Liberal (n = 72,154)</th>
<th>Restricted (n = 5,058)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>62.1 (17)</td>
<td>64.0 (17)</td>
<td>1.9 (1.4, 2.3)</td>
</tr>
<tr>
<td>Male</td>
<td>29,647 (41)</td>
<td>2,139 (42)</td>
<td>1 (0, 3)</td>
</tr>
<tr>
<td><strong>Respiratory History</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical History in the Prior Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>23,666 (33)</td>
<td>2,475 (49)</td>
<td>16 (15, 18)</td>
</tr>
<tr>
<td>COPD</td>
<td>31,747 (44)</td>
<td>2,476 (49)</td>
<td>5 (4, 6)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>9,863 (14)</td>
<td>717 (14)</td>
<td>1 (-1, 2)</td>
</tr>
<tr>
<td>Hospitalization in the Prior Year for Respiratory Cause</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>2,961 (4)</td>
<td>397 (8)</td>
<td>4 (3, 5)</td>
</tr>
<tr>
<td>COPD</td>
<td>8,078 (11)</td>
<td>1,103 (22)</td>
<td>11 (9, 12)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>2,273 (3)</td>
<td>316 (6)</td>
<td>3 (2, 4)</td>
</tr>
</tbody>
</table>

Abbreviations: COPD = chronic obstructive pulmonary disease. Data are presented as mean (standard deviation) or count (%).
## Patient Characteristics (Con’t)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Liberal (n = 72,154)</th>
<th>Restricted (n = 5,058)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measures of Disease Burden</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization for Any Cause in the Prior Year</td>
<td>14,937 (21)</td>
<td>1,523 (30)</td>
<td>9 (8, 11)</td>
</tr>
<tr>
<td>Physician Visits in the Prior Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-5</td>
<td>16,125 (22)</td>
<td>877 (17)</td>
<td>-5 (-6, -4)</td>
</tr>
<tr>
<td>6-10</td>
<td>18,617 (26)</td>
<td>1,241 (25)</td>
<td>-1 (-3, 0)</td>
</tr>
<tr>
<td>11-15</td>
<td>13,520 (19)</td>
<td>935 (18)</td>
<td>0 (-1, 1)</td>
</tr>
<tr>
<td>16+</td>
<td>23,892 (33)</td>
<td>2,005 (40)</td>
<td>7 (5, 8)</td>
</tr>
<tr>
<td>Charlson Comorbidity Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>17,985 (25)</td>
<td>739 (15)</td>
<td>-10 (-11, -9)</td>
</tr>
<tr>
<td>1</td>
<td>34,662 (48)</td>
<td>2,575 (51)</td>
<td>3 (1, 4)</td>
</tr>
<tr>
<td>2</td>
<td>8,616 (12)</td>
<td>650 (13)</td>
<td>1 (0, 2)</td>
</tr>
<tr>
<td>≥ 3</td>
<td>10,891 (15)</td>
<td>1,094 (22)</td>
<td>7 (5, 8)</td>
</tr>
<tr>
<td><strong>Prescribing Physician Specialty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Practitioner</td>
<td>59,101 (82)</td>
<td>3,488 (69)</td>
<td>-13 (-14, -12)</td>
</tr>
<tr>
<td>Respirologist</td>
<td>10,173 (14)</td>
<td>1,361 (27)</td>
<td>13 (12, 14)</td>
</tr>
<tr>
<td>Other</td>
<td>2,880 (4)</td>
<td>209 (4)</td>
<td>0 (0, 1)</td>
</tr>
</tbody>
</table>
Patient Characteristics (Con’t)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Liberal (n = 72,154)</th>
<th>Restricted (n = 5,058)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior Medication Use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Unique Prescription Drugs</td>
<td>9.3 (5)</td>
<td>11.6 (5)</td>
<td>2.3 (2.1, 2.4)</td>
</tr>
<tr>
<td>Respiratory Medications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronchodilators</td>
<td>8,256 (11)</td>
<td>1,513 (30)</td>
<td>19 (17, 20)</td>
</tr>
<tr>
<td>Cromoglycates</td>
<td>117 (0)</td>
<td>5 (0)</td>
<td>0 (0, 0)</td>
</tr>
<tr>
<td>Inhaled corticosteroids</td>
<td>31,339 (43)</td>
<td>3,581 (71)</td>
<td>27 (26, 29)</td>
</tr>
<tr>
<td>Leukotrienes</td>
<td>3,522 (5)</td>
<td>432 (9)</td>
<td>4 (3, 4)</td>
</tr>
<tr>
<td>Beta-agonists</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LABA</td>
<td>5,827 (8)</td>
<td>1,198 (24)</td>
<td>16 (14, 17)</td>
</tr>
<tr>
<td>LABA with Inhaled Corticosteroid</td>
<td>168 (0)</td>
<td>237 (5)</td>
<td>5 (4, 5)</td>
</tr>
<tr>
<td>SABA</td>
<td>35,198 (49)</td>
<td>3,716 (73)</td>
<td>25 (23, 26)</td>
</tr>
<tr>
<td>SABA/Anticholinergic Combined</td>
<td>13,451 (19)</td>
<td>1,828 (36)</td>
<td>18 (16, 19)</td>
</tr>
<tr>
<td>Xanthines</td>
<td>4,915 (7)</td>
<td>325 (6)</td>
<td>0 (-1, 0)</td>
</tr>
</tbody>
</table>

Abbreviations: LABA = long-acting beta-agonist; NSAIDs = non-steroidal anti-inflammatory drugs; SABA = short-acting beta-agonist. Data are presented as mean (standard deviation) or count (%).
# Hospitalization Rates

<table>
<thead>
<tr>
<th>Period</th>
<th>Number of Events</th>
<th>Number of Person-Years</th>
<th>Crude</th>
<th>Age- and Sex Adjusted‡</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hospitalizations for Respiratory Causes:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted</td>
<td>1,020</td>
<td>3,889</td>
<td>26.2 (24.7, 27.9)</td>
<td>25.1 (23.5, 26.8)</td>
</tr>
<tr>
<td>Liberal</td>
<td>10,001</td>
<td>53,537</td>
<td>18.7 (18.3, 19.1)</td>
<td>18.9 (18.4, 19.5)</td>
</tr>
<tr>
<td><strong>Hospitalizations for Any Cause:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted</td>
<td>1,248</td>
<td>3,783</td>
<td>33.0 (31.2, 34.9)</td>
<td>32.1 (30.3, 34.1)</td>
</tr>
<tr>
<td>Liberal</td>
<td>14,378</td>
<td>51,490</td>
<td>27.9 (27.5, 28.4)</td>
<td>28.5 (27.8, 29.2)</td>
</tr>
<tr>
<td><strong>All-Cause Mortality:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted</td>
<td>274</td>
<td>4,359</td>
<td>6.3 (5.6, 7.1)</td>
<td>5.1 (4.5, 5.8)</td>
</tr>
<tr>
<td>Liberal</td>
<td>2,610</td>
<td>58,126</td>
<td>4.5 (4.3, 4.7)</td>
<td>4.0 (3.7, 4.2)</td>
</tr>
</tbody>
</table>

†Rates are expressed as events per 100 person-years.  
‡Estimated for a male at the mean age of 62 years.
## Restricted vs Liberal

<table>
<thead>
<tr>
<th></th>
<th>Crude</th>
<th>Age- and Sex-Adjusted</th>
<th>Partially-Adjusted Model 1</th>
<th>Partially-Adjusted Model 2</th>
<th>Partially-Adjusted Model 3</th>
<th>Fully-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitalizations for Respiratory Causes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.41 (1.32, 1.51)</td>
<td>1.33 (1.25, 1.42)</td>
<td>1.05 (0.98, 1.12)</td>
<td>0.87 (0.81, 0.93)</td>
<td>0.79 (0.74, 0.84)</td>
<td>0.78 (0.73, 0.83)</td>
</tr>
<tr>
<td>Hospitalizations for Any Cause:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.19 (1.12, 1.26)</td>
<td>1.13 (1.07, 1.20)</td>
<td>0.96 (0.90, 1.02)</td>
<td>0.85 (0.80, 0.90)</td>
<td>0.83 (0.78, 0.88)</td>
<td>0.82 (0.77, 0.87)</td>
</tr>
<tr>
<td>All-Cause Mortality:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.40 (1.24, 1.59)</td>
<td>1.28 (1.13, 1.45)</td>
<td>1.10 (0.97, 1.25)</td>
<td>1.03 (0.91, 1.17)</td>
<td>0.98 (0.86, 1.13)</td>
<td>0.97 (0.84, 1.11)</td>
</tr>
</tbody>
</table>

Data are presented as HR (95% CI). The liberal period serves as the reference period for all analyses.
Strengths and Limitations

• Strengths:
  - True measure of association is known
  - Impact on prescribing practices assessed
  - Restricted to new users

• Limitations:
  - COPD added as approved indications during study period
    - Sensitivity analyses excluded patients hospitalized for COPD
  - Residual confounding due to missing smoking status
  - Restricted period: a modest number of new users
  - Generalizability to other formulary restrictions or other jurisdictions?
  - Did not assess the impact of the policy on clinical outcomes
Conclusions

• Drug formulary restrictions can result in substantial and unexpected confounding.

• In the case of fluticasone/salmeterol, crude estimates indicate that there was an increased risk of hospitalizations for respiratory causes during the liberal period relative to the restricted period but a statistically significant decreased risk during the restricted period following statistical adjustment.

• To ensure that study results are valid, formulary restrictions should be considered during the design and analysis of pharmacoepidemiologic studies.
Confounding by drug formulary restriction in pharmacoepidemiologic research†

Kristian B. Filion¹,²,³*,‡, Maria Eberg¹ and Pierre Ernst¹,²

¹ Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, Quebec, Canada
² Department of Medicine, McGill University, Montreal, Quebec, Canada
³ Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada

ABSTRACT

Purpose The potential consequences of confounding due to drug formulary restrictions in pharmacoepidemiologic research remain incompletely understood. Our objective was to illustrate this potential bias using the example of fluticasone/salmeterol combination therapy, an oral inhaler used to treat asthma and chronic obstructive pulmonary disease, whose use is restricted in the province of Quebec, Canada.

Methods We identified all new users of fluticasone/salmeterol in Quebec’s administrative databases and classified those who received their initial dispensing of fluticasone/salmeterol between 1 September 1999 and 30 September 2003 as users from the liberal period and those who received it between 1 January 2004 and 31 October 2006 as users from the restricted period. The primary outcome was time to first hospitalization for respiratory causes within 12 months of cohort entry.

Results Our cohort included 72,154 new users from the liberal period and 5,058 from the restricted period. Compared with use during the liberal period, use during the restricted period was associated with an increased rate of hospitalization for respiratory causes (crude hazard ratio [HR] = 1.41, 95% confidence interval [CI] = 1.32, 1.51). Subsequent adjustment for age, sex, and hospitalization for respiratory causes in the previous year attenuated the association (HR = 1.05, 95%CI = 0.98, 1.12). Further adjustment for other potential confounders resulted in a lower rate during the restricted period (HR = 0.78, 95%CI = 0.73, 0.83).

Conclusions Formulary restrictions can result in substantial and unexpected confounding and should be considered during the design and analysis of pharmacoepidemiologic studies. Copyright © 2015 John Wiley & Sons, Ltd.
Acknowledgements

• Collaborators
  - Pierre Ernst
  - Maria Eberg

• Personal Support
  - CIHR New Investigator Award

• CIHR Operating Grant
Thank You!

Kristian.filion@mcgill.ca