THE ADVANCEMENT AND DEMOCRATIZATION OF MEDICAL RESEARCH IN CANADA THROUGH THE DEVELOPMENT AND VALIDATION OF RANDOMIZED-REGISTRY TRIALS

E. Marc Jolicoeur MD MSC MHS FRCP-C FACC FS桦AI
Associate Professor of Medicine
Université de Montréal, Canada

Montréal Heart Institute
marc.jolicoeur@icm-mhi.org
Financial disclosures

Consultant – Neovasc Inc
Research Grants (significant): AstraZeneca, Boston Scientific, St-Jude Medical
Dr. Jolicoeur is supported by research grants from the FRSQ and the CIHR
“Born On Pessimism, Grow On Skepticism, Mature On Optimism & Die On Euphoria”

John Maynard Keynes, economists
The 6th principle

Canada Health Act

Universality
Accessibility
Portability
Comprehensiveness
Public administration

DATA USAGE
Clinical Registry Trial

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research?

Michael S. Lauer, M.D., and Ralph B. D’Agostino, Sr., Ph.D.
Clinical Registry Trial

+ $50 per participant over costs for already existing and paid-for registries

Excellent enrollment!

Rapid Randomization in the TASTE Trial, with Enrollment of Most Patients Receiving Primary Percutaneous Coronary Intervention (PCI). Adapted from the Institute of Medicine (www.iom.edu/-/media/Files/Activity%20Files/Quality/VSRT/LST%20Workshop/Presentations/Granger.pdf). The incremental cost of the Thrombus Aspiration in ST-Elevation Myocardial Infarction in Scandinavia (TASTE) trial was $300,000, or $50 for each participant who underwent randomization.
SWEDHEART registry

- 75 variables on 2nd prevention after 12M
- 106 baseline variables for patients with ACS
- 150 variables for PCI patients
- 100 variables for patients undergoing heart surgery

Data capture Web-based registry with all data registered online by caregivers

Baseline data

- National Cause of Death Register
- National patient registry
- Diagnoses at discharge
- Other data
- National Registry of Drug prescriptions
Health administrative databases in Canada

- National Cause of Death Register
- Discharge Abstract Database (CIHI) / Med-ECHO (RAMQ)
- Outpatient and emergency visits
- Fee-for-service billing
- Provincial Registry of Drug prescriptions
- International Classification of Diseases ICD CODES
- Electronic medical records
What is a RCT really?

1. Table 1 that summarizes baseline characteristics
2. Serious adverse events
3. Individual endpoints / outcomes
4. Drugs received at baseline and changed post-randomization
ARE RRTS AS RELIABLE AS CONVENTIONAL RCTS TO CAPTURE EVENTS?
Medical records → Health administrative databases → Provincial Registry of Drug prescriptions

Discharge Abstract Database (CIHI) / Med-ECHO (RAMQ)
Past experience with discharge abstract database (DAD)

<table>
<thead>
<tr>
<th>Medical condition</th>
<th>Agreement</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myocardial infarction</td>
<td>95%</td>
<td>(0.91, 0.97)</td>
</tr>
<tr>
<td>Angina</td>
<td>64%</td>
<td>(0.59, 0.72)</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>84%</td>
<td>(0.61, 0.64)</td>
</tr>
</tbody>
</table>

Past experience with validation of health administrative databases using medical records has shown us that health administrative databases are reliable to identify certain acute diseases but likely underreport chronic comorbid conditions.

Past experience with validation of prescription claim databases.

- A public prescription claim database is one of the most accurate means of determining drugs dispensed to patients:
- The percent of missing or out of range data in the prescription claims database of the Regie de l’assurance maladie du Quebec (RAMQ) to be small (0.4%)*
- Similar conclusions reached with the Ontario Drug Benefit (ODB) program **

Past experience with validation of prescription claim databases.

• However, coverage is a limitation, as < 50% of patients are covered by public drug insurance in most provinces: Elderly; The welfare recipients; Those not insured through their employer
  – (these patients may not be completely representative of the overall population)

• Information is only available for drugs dispensed from community pharmacies: drugs dispensed during hospital / nursing home stays are not typically captured. Compliance is not directly assessed.
The WOSCOPS investigators may have been the first trial to match participant’s data with public records to obtain an extended follow-up.

Other examples recently followed:
1- WHI trial
2- DCOR trial
3- CLARICOR trial
The Women’s Health Initiative validation studies

- The WHI reported a series of validations studies on the reliability of Medicare claims to detect cardiovascular outcomes using trial’s data as a reference.
- WHI compared hormone therapy vs. placebo in post-menopausal women from 1993 to 1998.
The CMS database identified some myocardial infarctions “missed” by the Women’s Health Initiative adjudication process.

Inversely, the adjudication process identified some myocardial infarctions “missed” by the CMS database.

Good agreement between WHI-adjudicated outcomes and Medicare claims for the diagnosis of myocardial infarction (κ, 0.71–0.74) and excellent for coronary revascularization (κ, 0.88 to 0.91).

<table>
<thead>
<tr>
<th>Event-based analysis</th>
<th>SN</th>
<th>SP</th>
<th>PPV</th>
<th>NPV</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke (any diagnosis code; definition 1)</td>
<td>82.0</td>
<td>99.7</td>
<td>84.6</td>
<td></td>
<td>0.83</td>
</tr>
<tr>
<td>Stroke (principal diagnosis; definition 2)</td>
<td>75.3</td>
<td>99.8</td>
<td>87.5</td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>Hemorrhagic stroke (definition 3)</td>
<td>75.3</td>
<td>&gt;99.9</td>
<td>84.3</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Ischemic stroke (definition 4)</td>
<td>82.2</td>
<td>99.7</td>
<td>79.5</td>
<td>99.7</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person-level analysis</th>
<th>SN</th>
<th>SP</th>
<th>PPV</th>
<th>NPV</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke (any diagnosis code; definition 1)</td>
<td>88.0</td>
<td>99.7</td>
<td>87.5</td>
<td>99.7</td>
<td>0.87</td>
</tr>
<tr>
<td>Stroke (principal diagnosis; definition 2)</td>
<td>81.0</td>
<td>99.8</td>
<td>89.6</td>
<td>99.5</td>
<td>0.85</td>
</tr>
<tr>
<td>Hemorrhagic stroke (definition 3)</td>
<td>80.6</td>
<td>&gt;99.9</td>
<td>87.2</td>
<td>99.9</td>
<td>0.84</td>
</tr>
<tr>
<td>Ischemic stroke (definition 4)</td>
<td>88.4</td>
<td>99.6</td>
<td>82.7</td>
<td>99.8</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Results are for events when WHI medical records were available for adjudication. Event-based analyses use a 7-day match window. $\kappa$ indicates kappa statistic; NPV, negative predictive value; PPV, positive predictive value; SN, sensitivity; SP, specificity; and WHI, Women’s Health Initiative.
Clearly, a validation and a calibration of the definitions for outcomes must done before health administrative databases can be used for study monitoring in Canada.

In the WHI study, AMI, revascularization and stroke were good, but PVD was not well correlated.

The Good Clinical Practice (GCP) guidelines mandate trial monitoring and typically rely on on-site monitoring. If our validation effort is satisfactory, the use of health administrative databases may become a (much cheaper) accepted standard for trial monitoring.
The Canadian Randomized Registry Trial Initiative

National Coordination Team
Marc Jolicoeur (MHI)
Jack Tu (ICES)
PJ Devereaux (PHRI)
Jean-Lucien Rouleau (UdeM)
Jean-Claude Tardif (MHI)
Lindsay Jacobi Cadete (assistant)

Team ethics
Don Willison (UofT)
Ray Saginur (UofO)
Trudo Lemmens (UofT)

Knowledge users
Laurie Lambert (INESSS)
Claudia Blais (INSPQ)

Industry Collaborators
Shurjeel Choudri
(Pharmaceutical)
Bayer Canada

Charlie Sciortino
(medical devices)
St-Jude Canada

BC
Elaine Chong
Bev Holmes
Graham Wong

ALBERTA
Michael Hill
Larry Svenson
Finlay McAlister

MANITOBA
Salah Mahmud
Terry Klassen
P. Nickerson

ONTARIO
PJ Devereaux
Salim Yusuf
Jackie Bosch
Janice Pogue
Reza Rezai
Michael Schull

SASKATCHEWAN
To be determined

QUEBEC
Marc Jolicoeur
Robyn Tamblyn
James Brophy
Jean-Claude Tardif

ATLANTIC
Adrian Levy
Jafna Cox
The Canadian Randomized Registry Trial Initiative

- **AIM I** - To compare the rates of clinical outcomes (adjudicated outcomes and SAEs) recorded from Canadians who participated in contemporary large conventional RCTs to those obtained during the same time period from the same participants using administrative health databases.

- **AIM II** - To assess the quality of reporting of other important information obtained administrative health databases, such as demographic characteristics, past medical history, pharmacological profiles and other aspects of patient follow-up.
<table>
<thead>
<tr>
<th>Study</th>
<th>Outcomes</th>
<th># Pts with HIN numbers available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ON</td>
</tr>
<tr>
<td>PURE (~6Y)</td>
<td>Death (CV), MI, Stroke, HF, Cancer</td>
<td>3445</td>
</tr>
<tr>
<td>POISE (30D)</td>
<td>Death (CV), MI, cardiac arrest</td>
<td>1743</td>
</tr>
<tr>
<td>POISE-2 (6M)</td>
<td>Death, MI</td>
<td>1968</td>
</tr>
<tr>
<td>HOPE + HOPE-2 (&gt;5Y)</td>
<td>Death (CV), MI, Stroke, HF, Cancer</td>
<td>2411</td>
</tr>
<tr>
<td>DREAM + EPI-DREAM (&gt;6Y)</td>
<td>Death (CV), MI, Stroke, HF, Cancer</td>
<td>2640</td>
</tr>
<tr>
<td>Total=19,246 Pts for all provinces combined</td>
<td></td>
<td>12207</td>
</tr>
</tbody>
</table>
How to do it?

• **Step 1** – Match the trial participant’s file and the patient’s medical record

• **Step 2** – Retrieve all medical services claims and their respective ICD-9/10 codes from health administrative databases

• **Step 3** – Display ICD codes into standardized health problems diagnoses

• **Step 4** – Link the events (outcomes/SAE) to the ICD codes

• **Step 5** – Compare the estimated intervention effects and attempt bootstrap replication
Matching for endpoints and SAEs

Trial
- Date of Randomization
- Date of event as reported in the trial ± 7d
  - Hospital admission date
  - Clinical diagnosis DAD: ICD 9 code = 410

Health databases
- Serious adverse Event that is not an endpoint: major bleeding ± 7d
  - Hospital admission date
  - Clinical diagnosis DAD Code in clusters: ICD code Y and T + DAD type 9 = T81.0

Matches
- Matched endpoint and diagnoses
- Matched SAEs and Diagnoses
- Matched Endpoint, + cause of death

Adjudicated Endpoint: myocardial infarction

Adjudicated Endpoint + death: congestive heart failure leading to cardiac death

Date of Last follow-up

Time 0

Time final
Serious adverse events

As per ICH guidelines, a serious adverse event (SAE) is defined as any untoward medical occurrence that at any dose:

- Results in death; or
- Is life-threatening; or
- Requires inpatient hospitalization or prolongation of existing hospitalization; or
- Results in persistent or significant disability/incapacity; or
- Is a congenital anomaly/birth defect; or
- Is a medically important event
In the case of an adverse effect in therapeutic use (SAE related to a drug),

The purpose of the diagnosis cluster for adverse effect in therapeutic use is to link the drug(s), medication(s) or biological substance(s) causing the adverse effect to the specific adverse effect(s) with which it is associated.

In the current codification, codes Y40-Y59 relate for drug(s), medicament(s) or biological substance(s) causing the adverse effect, and the adverse effect in question.
AIM 2

To assess the quality of reporting of other important information obtained from provincial and national administrative health databases, such as demographic characteristics, past medical history, pharmacological profiles and other aspects of patient follow-up.
Patient-specific baseline characteristics and past medical history
[ How to generate a table 1 ]
The list of co-morbid conditions will be obtained after integrating many sources of information, including diagnostic codes from medical services claims and single-indication dispensed drugs. ICD-9/10 codes will be mapped and displayed into standardized health problems diagnosis and use to generate table I.
## Patient-specific baseline characteristics and past medical history

[How to generate table 1]


<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>Medical record review</th>
<th>Index admission + 5 preceding years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1099</td>
<td>55.5</td>
</tr>
<tr>
<td>Diabetes without complications</td>
<td>425</td>
<td>21.4</td>
</tr>
<tr>
<td>Previous AMI</td>
<td>412</td>
<td>20.8</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>404</td>
<td>20.4</td>
</tr>
<tr>
<td>COPD</td>
<td>340</td>
<td>17.2</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>308</td>
<td>15.5</td>
</tr>
<tr>
<td>Chronic renal disease</td>
<td>304</td>
<td>15.5</td>
</tr>
<tr>
<td>Acute renal disease</td>
<td>139</td>
<td>7.1</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>106</td>
<td>5.4</td>
</tr>
<tr>
<td>Diabetes with complications</td>
<td>102</td>
<td>5.2</td>
</tr>
<tr>
<td>Cardiogenic shock</td>
<td>87</td>
<td>4.4</td>
</tr>
<tr>
<td>Neoplasm</td>
<td>60</td>
<td>3.0</td>
</tr>
<tr>
<td>Dementia</td>
<td>54</td>
<td>2.7</td>
</tr>
<tr>
<td>Rheumatic disease</td>
<td>44</td>
<td>2.2</td>
</tr>
</tbody>
</table>

AMI, acute myocardial infarction; COPD, chronic obstructive pulmonary disease.

*Validated comorbidities not shown in Table 2 were assessed according to medical record review (ulcer disease, mild liver disease, diabetes with complications). †Historical information was not collected because these conditions were assumed to have been present for several years.
Foreseeable limitations

• Ethical considerations if not specifically consented beforehand
• Patients excluded or ineligible to public health care system (1-2%): migrants, young, members of the Canadian Forces and RCMP, federal correctional facilities
• Interprovincial transfers
• Direct vs. probabilistic medical record linkage
Clinical Registry Trial

**Pros**
- Democratization of research process
- Ideal for trials testing interventional strategies or niche indications (no industry support...)
- Results are readily generalizable
- Immediate cost-effectiveness analysis possible

**Cons**
- Quality of data can be variable
- Meta-registry required (Orwellian state)
- Legal and ethical issues to be worked out
- In total contradiction of GCP...