Pitfalls to Avoid in Observational Studies in the ICU

Rob Fowler, MDCM, MS(Epi), FRCP
Department of Medicine & Critical Care Medicine
Sunnybrook Hospital
Institute of Health Policy, Management and Evaluation
University of Toronto
rob.fowler@sunnybrook.ca
How we conceptualize the research marathon

Anecdote → Case Series → Observational Studies → Pilot Trials → Large RCTs
Anecdote  Case Series  Observational Studies  Pilot Trials  Large RCTs

However, some questions won’t (can’t) be (best) answered with RCTs and we must rely on other methods...
...beware of the Sudenkuoppa
Common Wolf Pits in Observational Studies?

1. Confounding, confounding, confounding!
2. Bias by Indication (very common for ICU patients)
3. Selection bias
4. Effect modification
5. Immortal time bias
6. Time-dependent Factors
7. ... and so many more Wolf Pits
Common Wolf Pits in Observational Studies?

1. Confounding, confounding, confounding!
2. Bias by Indication (very common for ICU patients)
3. Selection bias
4. Effect modification
5. Immortal time bias
6. Time-dependent Factors
7. ... and so many more Wolf Pits
Confounding

Caused by an extraneous variable in a statistical model that correlates (directly or inversely) with both the dependent variable (the outcome) and the independent variable (the predictor).

- receipt of treatment
- outcome
- AKI age comorbidity
Potential Mechanisms to **mitigate** Confounding?

- Create a *path diagram*
- **Stratification** of patients into many subgroups*
- **Matching*** of patients on *key* variables
- **Multivariate*** adjustment
- **Propensity*** scores
- **Time-dependent*** models
- Consider **Hierarchical models***

... **NONE will be sufficient**

... Do an **RCT**!
The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients

Objective.—To examine the association between the use of right heart catheterization (RHC) during the first 24 hours of care in the intensive care unit (ICU) and subsequent survival, length of stay, intensity of care, and cost of care.

Design.—Prospective cohort study.


Subjects.—A total of 5735 critically ill adult patients receiving care in an ICU for 1 of 9 prespecified disease categories.

Main Outcome Measures.—Survival time, cost of care, intensity of care, and length of stay in the ICU and hospital, determined from the clinical record and from the National Death Index. A propensity score for RHC was constructed using multivariable logistic regression. Case-matching and multivariable regression modeling techniques were used to estimate the association of RHC with specific outcomes after adjusting for treatment selection using the propensity score. Sensitivity analysis was used to estimate the potential effect of an unidentified or missing covariate on the results.

Results.—By case-matching analysis, patients with RHC had an increased 30-day mortality (odds ratio, 1.24; 95% confidence interval, 1.03-1.49). The mean cost (25th, 50th, 75th percentiles) per hospital stay was $49,300 ($17,000, $30,500, $56,600) with RHC and $35,700 ($11,300, $20,600, $39,200) without RHC. Mean length of stay in the ICU was 14.8 (5, 9, 17) days with RHC and 13.0 (4, 7, 14) days without RHC. These findings were all confirmed by multivariable modeling techniques. Subgroup analysis did not reveal any patient group or site for which RHC was associated with improved outcomes. Patients with higher baseline probability of surviving 2 months had the highest relative risk of death following RHC. Sensitivity analysis suggested that a missing covariate would have to increase the risk of death 6-fold and the risk of RHC 6-fold for a true beneficial effect of RHC to be misrepresented as harmful.

Conclusion.—In this observational study of critically ill patients, after adjustment for treatment selection bias, RHC was associated with increased mortality and increased utilization of resources. The cause of this apparent lack of benefit is unclear. The results of this analysis should be confirmed in other observational studies. These findings justify reconsideration of a randomized controlled trial of RHC and may guide patient selection for such a study.

MANY CARDIOLOGISTS and critical care physicians believe that the direct measurement of cardiac function provided by right heart catheterization (RHC) (also known as pulmonary artery catheterization) is necessary to guide therapy for certain critically ill patients1-9 and that such management leads to better patient outcomes.1 While the benefit of RHC has not been demonstrated in a randomized controlled trial (RCT), the popularity of this procedure and the widespread belief that it is beneficial make the performance of an RCT difficult. Physicians cannot ethically participate in such a trial or encourage a patient to participate if convinced the procedure is truly beneficial. The most recent attempt at an RCT was stopped because most physicians refused to allow their patients to be randomized.5

In the absence of RCTs of RHC, observational studies have been used to evaluate its effectiveness. The relative risk of death has been found to be higher in the elderly6 and in patients with acute myocardial infarction19 who were managed with RHC. In a study of acute myocardial infarction in Medicare patients, hospitals with higher than expected use of RHC had higher than expected mortality.12 In observational studies, results have varied, and the data are inconclusive. This lack of consistency and the variability of results at individual institutions make the interpretation of observational studies difficult.
The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients

METHODS: A propensity score for RHC was constructed using multivariable logistic regression. Case-matching and multivariable regression modeling techniques were used to estimate the association of RHC with specific outcomes after adjusting for treatment selection using the propensity score.

RESULTS: By case-matching analysis, patients with RHC had an increased 30-day mortality (odds ratio, 1.24; 95% confidence interval, 1.03-1.49).

These findings were all confirmed by multivariable modeling techniques. Sensitivity analysis suggested that a missing covariate would have to increase the risk of death 6-fold and the risk of RHC 6-fold for a true beneficial effect of RHC to be misrepresented as harmful.
Is It Time to Pull the Pulmonary Artery Catheter?

Pulmonary artery (PA) catheterization is performed daily in hospitals around the world. The hemodynamic data derived from PA catheterization are used to determine, monitor, and modify therapy in critically ill and/or hemodynamically unstable patients, including those with acute myocardial infarction (MI), congestive heart failure (CHF), and multiorgan failure, and they are frequently used to monitor patients with heart disease who undergo cardiac or noncardiac surgery.

Pulmonary artery catheterization (PAC) is commonly performed to determine the clinical probability of pulmonary embolism (PE) using the Wells criteria and therapeutic management using the following risk-based algorithm:

- **Class I** indications for PA catheterization include:
  - Patients with CHF who have a high risk of adverse outcomes with treatment
  - Patients with septic shock
  - Patients with acute liver failure
  - Patients with acute respiratory failure

- **Class IIa** indications include:
  - Patients with severe sepsis
  - Patients with multiorgan failure

- **Class IIb** indications include:
  - Patients with severe septic shock
  - Patients with severe acute respiratory failure

- **Class III** indications include:
  - Patients with mild sepsis
  - Patients with nonseptic shock

The use of PAC in critically ill patients has been associated with increased mortality and increased utilization of resources. The apparent lack of benefit is unclear. The results of this analysis should be confirmed in other observational studies. These findings justify reconsideration of a randomized controlled trial of RHC and may guide patient selection for such a study.
Where are the Wolf Pits?

What are the inevitable challenges with this observational study ..... what are the biases and forms of confounding?
Where are the Wolf Pits?

What are the inevitable challenges with this observational study - what are the biases and forms of confounding?

....confounding by indication

But, in 1996, “Propensity Scores” were the new statistical magic for the field, and could fix all challenges of confounding....
Pulmonary-Artery versus Central Venous Catheter to Guide Treatment of Acute Lung Injury

The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network*

RESULTS

The groups had similar baseline characteristics. The rates of death during the first 60 days before discharge home were similar in the PAC and CVC groups (27.4 percent and 26.3 percent, respectively; \( P = 0.69 \); absolute difference, 1.1 percent; 95 percent confidence interval, −4.4 to 6.6 percent), as were the mean (±SE) numbers of both ventilator-free days (13.2±0.5 and 13.5±0.5; \( P = 0.58 \)) and days not spent in the intensive care unit (12.0±0.4 and 12.5±0.5; \( P = 0.40 \)) to day 28. PAC-guided therapy did not improve these measures for patients in shock at the time of enrollment. There
In critical care, it is very difficult (*read, nearly impossible*) to validly estimate the influence of a ‘treatment’-related variable on outcome with, usual effect sizes, irrespective of the statistical technique (so far). There is usually too much bias by indication (among others...).
It’s July 2009. An H1N1 pandemic has been declared. In two
dataset of many hundred patients, in Canada, 91% of the
critically ill received neuraminidase inhibitors; in Mexico 78%.

Reviewers ask you to to estimate the influence of
neuraminidase inhibitors on survival. Can you, should you, do
this? Is there any way to try to approach it?
Immortal Time Bias

**Immortal time bias:** a span of time in the observation or follow-up period of a cohort during which the outcome under study could not have occurred

**Here for H1N1,** patients must have survived long enough to have a diagnosis confirmed (2-3 days) **AND** to have been given oseltamivir ... those that died early didn’t have the opportunity to get the treatment, therefore the drug looks more effective than it is ... there is an *immortal time bias*
How about Corticosteroids?
Corticosteroid Treatment in Critically Ill Patients with Pandemic Influenza A/H1N1 2009 Infection
Analytic Strategy Using Propensity Scores

Sung-Han Kim1, Sang-Bum Hong2, Sung-Choeil Yun3, Won-II Choi4, Jong-Joon Ahn5, Young Joo Lee6, Heung-Bum Lee7, Chae-Man Lim2, and Younsuck Koh2; for the Korean Society of Critical Care Medicine H1N1 Collaborative*

1Department of Infectious Diseases, 2Department of Pulmonary and Critical Care Medicine, and 3Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, University of Ulsan College of Medicine, Seoul; 4Keimyung University School of Medicine, Daegu; 5University of Ulsan Hospital, Ulsan; 6Ajou University School of Medicine, Suwon; and 7Chonbuk National University Medical School, Jeonju, Republic of Korea

Rationale: Administration of adjuvant corticosteroids to patients with pandemic influenza A/H1N1 2009 (pH1N1) may reduce inflammation and improve outcomes.

Objectives: To assess the effect of adjuvant corticosteroid treatment on the outcomes of critically ill patients with pH1N1 infection treated in an intensive care unit (ICU) in 2009 to determine whether adjuvant corticosteroids used post-admission are beneficial and measure the impact on short- and long-term outcomes.

Early Corticosteroids in Severe Influenza A/H1N1 Pneumonia and Acute Respiratory Distress Syndrome

Christian Brun-Buisson1,2,3, Jean-Christophe M. Richard4, Alain Mercat5, Anne C. M. Thiebaut3,6, and Laurent Brochard1,2,7, for the REVA-SRLF A/H1N1v 2009 Registry Group*

1Université Paris Est-Crétteil and INSERM U955, Créteil, France; 2Assistance Publique-Hôpitaux de Paris, GH Henri Mondor, Service de réanimation médicale, Créteil, France; 3Inserm U657; Pharmacoepidemiology and Infectious Diseases, Institut Pasteur, Paris, France; 4Service de réanimation médicale, CHU de Rouen, UPRES EA 38, Rouen, France; 5Département de Réanimation Médicale et Médecine Hyperbare, CHU d’Angers, Angers, France; 6EA4499, Université de Versailles Saint-Quentin, France; and 7Department of Intensive Care Medicine University Hospital of Geneva and University of Geneva, Switzerland

Rationale: Despite their controversial role, corticosteroids are often administered to patients with adult respiratory distress syndrome (ARDS) secondary to viral pneumonia.

Objectives: To analyze the impact of corticosteroid therapy on outcomes of patients having ARDS associated with influenza A/H1N1 pneumonia.
Log-Rank test $P = 0.01$

**Figure 1.** Survival curves stratified by adjuvant steroid treatment. Sixty-five patients with pandemic influenza A/H1N1 2009 who received steroid treatment (dotted line) are compared with 65 matched control subjects who did not receive steroid treatment (solid line).

---

**TABLE 2. ANALYSIS OF ADJUVANT CORTICOSTEROID TREATMENT ASSOCIATED WITH 90-DAY MORTALITY IN 245 PATIENTS ADMITTED TO INTENSIVE CARE UNIT WITH PANDEMIC INFLUENZA A/H1N1 INFECTION**

<table>
<thead>
<tr>
<th></th>
<th>Crude</th>
<th>Adjusted Mortality</th>
<th>Propensity-matched</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients (n = 245)</td>
<td>3.76 (2.19–6.44) $&lt;0.001$</td>
<td>2.20* (1.03–4.71) 0.04</td>
<td>2.63† (1.43–4.82) 0.002</td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARDS (n = 136)</td>
<td>2.44 (1.22–4.87) 0.01</td>
<td>1.80† (0.69–4.69) 0.23</td>
<td>2.28‡ (0.84–5.05) 0.10</td>
</tr>
<tr>
<td>Non-ARDS (n = 109)</td>
<td>6.59 (2.73–15.89) $&lt;0.001$</td>
<td>3.49‡ (0.89–13.81) 0.07</td>
<td>3.33** (0.75–14.89) 0.11</td>
</tr>
<tr>
<td>Nonsteroid users (n = 194)</td>
<td>2.73 (1.43–5.20) 0.002</td>
<td>1.68†† (0.66–4.24) 0.27</td>
<td>2.29‡‡ (1.01–5.20) 0.047</td>
</tr>
</tbody>
</table>
Figure 1. Survival curves stratified by adjuvant steroid treatment. Sixty-five patients with pandemic influenza A/H1N1 2009 who received steroid treatment (dotted line) are compared with 65 matched control subjects who did not receive steroid treatment (solid line).

<table>
<thead>
<tr>
<th>Crude</th>
<th>OR (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.76 (2.19–6.44)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>2.44 (1.22–4.87)</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>6.59 (2.73–15.89)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>2.73 (1.43–5.20)</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Figure 1. Survival curves stratified by adjuvant steroid treatment. Sixty-five patients with pandemic influenza A/H1N1 2009 who received steroid treatment (dotted line) are compared with 65 matched control subjects who did not receive steroid treatment (solid line).
Figure 1. Survival curves stratified by adjuvant steroid treatment. Sixty-five patients with pandemic influenza A/H1N1 2009 who received steroid treatment (dotted line) are compared with 65 matched control subjects who did not receive steroid treatment (solid line).

<table>
<thead>
<tr>
<th>Crude</th>
<th>Adjusted Mortal</th>
<th>Propensity-matched</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adjusted OR (95% CI)</td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td></td>
<td>Adjusted OR (95% CI)</td>
</tr>
<tr>
<td>3.76 (2.19–6.44)</td>
<td>2.20* (1.03–4.71)</td>
<td>2.63† (1.43–4.82)</td>
</tr>
<tr>
<td>2.44 (1.22–4.87)</td>
<td>1.80‡ (0.69–4.69)</td>
<td>2.28§ (0.84–5.05)</td>
</tr>
<tr>
<td>6.59 (2.73–15.89)</td>
<td>3.49‖ (0.89–13.81)</td>
<td>3.33** (0.75–14.89)</td>
</tr>
<tr>
<td>2.73 (1.43–5.20)</td>
<td>1.68‖ (0.66–4.24)</td>
<td>2.29‡‡ (1.01–5.20)</td>
</tr>
</tbody>
</table>
What does trial data tell us?
Efficacy and Safety of Corticosteroids for Persistent Acute Respiratory Distress Syndrome

The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network*

METHODS

We randomly assigned 180 patients with ARDS of at least seven days’ duration to receive either methylprednisolone or placebo in a double-blind fashion. The primary end point was mortality at 60 days. Secondary end points included the number of ventilator-free days and organ-failure–free days, biochemical markers of inflammation and fibroproliferation, and infectious complications.

RESULTS

At 60 days, the hospital mortality rate was 28.6 percent in the placebo group (95 percent confidence interval, 20.3 to 38.6 percent) and 29.2 percent in the methylprednisolone group (95 percent confidence interval, 20.8 to 39.4 percent; P = 1.0); at 180 days, the rates were 31.9 percent (95 percent confidence interval, 23.2 to 42.0 percent) and 31.5 percent (95 percent confidence interval, 22.8 to 41.7 percent; P = 1.0), respectively.
Patients who get sicker tend to get ‘kitchen sink’ therapies (we want to ‘do something’)

However, we only usually ADJUST* for differences between patients when they arrive in ICU, not when they get sicker & get kitchen sink^
METHODS: In an observational cohort study of adults with H1N1pdm09-related critical illness from 51 Canadian ICUs, we investigated predictors of steroid administration and outcomes of patients who received and those who did not receive corticosteroids. We adjusted for potential baseline confounding using multivariate logistic regression and propensity score analysis and adjusted for potential time-dependent confounding using marginal structural models. (read, new statistical magic, 2016)
After adjustment for **time-dependent** variables (worsening in patients who then are prescribed steroids) we see **NO** influence of steroids on survival/mortality.