Real-World Designs and Implications for Causal Inference
Real World Designs and Implications for Causal Inference

David Price

Professor of Primary Care Respiratory Medicine, University of Aberdeen
Director, Observational & Pragmatic Research Institute, Singapore
Director, Optimum Patient Care, Australia and UK
Pragmatic Trial Issues

- TWICS Trial
- Lessons From Other Trials:
  - ELEVATE Trial
  - MAGNIFY Trial
Effect of Theophylline as Adjunct to Inhaled Corticosteroids on Exacerbations in Patients With COPD
A Randomized Clinical Trial

Graham Devereux, PhD; Seonaidh Cotton, PhD; Shona Fielding, PhD; Nicola McMeekin, MSc; Peter J. Barnes, DSc; Andrew Briggs, PhD; Graham Burns, PhD; Rekha Chaudhuri, MD; Henry Chrystyn, PhD; Lisa Davies, FRCP; Anthony De Soyza, PhD; Simon Gompertz, MD; John Haughney, FRCP; Karen Innes, MSc; Joanna Kaniewska, PhD; Amanda Lee, PhD; Alyn Morice, FRCP; John Norrie, MSc; Anita Sullivan, PhD; Andrew Wilson, PhD; David Price, FRCPG

National Institute for Health Research

University of Aberdeen
**Hypothesis:** Addition of ‘low dose’ theophylline to inhaled steroid therapy in COPD reduces the risk of moderate-severe COPD exacerbation a year of treatment.

**Design:** Pragmatic multi-centre double-blind randomised placebo-controlled trial.

**Primary Outcome:** Total number of moderate-severe exacerbations during the year of treatment

**Inclusion criteria**
- ≥ 40 years old
- Smoking history of at least 10 pack years
- Predominant COPD diagnosis (FEV1/FVC<0.7)
- Current ICS use (irrespective of LABA and/or LAMA use)
- ≥2 exacerbations
- Clinically stable with no COPD exacerbation for at least 4 weeks

**Exclusion criteria**
- A predominant respiratory disease other than COPD
- Current theophylline use
- Drugs interacting with theophylline and/or ↑plasma theophylline

---

**Recruitment & consent** (clinical inc spirometry, QoL)

**Theophylline, 200mg od/bd**

**Placebo, od/bd**

**6 month follow-up visit**
- clinical (inc spirometry), QoL, HE data

**12 month follow-up visit**
- clinical (inc spirometry), QoL, HE data

Dosing based on IBW and smoking status

Primary outcome data supplemented with data from GP/hospital records


Presented at ATS International Conference 2019
RCTs vs RWE Studies: Population vs Controlled design

Population

- Broad
  - Managed as...
    - Clinical diagnosis
      - Confirmed diagnosis
        - Registration RCTs
          - Long term phase III
          - Pragmatic RCTs

- Narrow
  - Highly controlled
    - Pragmatically controlled
      - Observational

Ecology of care

- Constrained
- Free

Observational studies

TWICS

## TWICS Trial: Intention to Treat: Primary Outcome, COPD Exacerbations

<table>
<thead>
<tr>
<th></th>
<th>Theophylline</th>
<th>Placebo</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N with outcome data</td>
<td>772</td>
<td>764</td>
<td>1536 (98%)</td>
</tr>
<tr>
<td>Person years of FU</td>
<td>748</td>
<td>742</td>
<td>1490</td>
</tr>
<tr>
<td>Total number of exacerbations</td>
<td>1727</td>
<td>1703</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Theophylline</th>
<th>Placebo</th>
<th>Unadjusted Rate ratio (95% CI)</th>
<th>Adjusted Rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exacerbations: mean (SD)</td>
<td>2.24 (1.99)</td>
<td>2.23 (1.97)</td>
<td>1.00 (0.91 - 1.09)</td>
<td>0.99 (0.91 - 1.07)</td>
</tr>
</tbody>
</table>

Adjusted for: age, gender, pack years, number of exacerbations in previous 12 months, COPD treatment, long term antibiotics, recruiting setting, centre as a random effect.
# Strengths, Limitations, and Issues

## Strengths

- Pragmatic: replicated use of low-dose theophylline in clinical practice
- Ability to use placebo
- 121 geographically dispersed recruitment sites
- 60% of patients identified in primary care
- Minimal inclusion criteria
- Infrequent study assessments, no monitoring of blood theophylline levels
- No change in routine care or routine care setting

## Limitations

<table>
<thead>
<tr>
<th>Limitations</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>26% of participants ceased taking study drug</td>
<td>No measure of adherence</td>
</tr>
<tr>
<td>✓ Similar % in a small trial of low-dose theophylline</td>
<td></td>
</tr>
<tr>
<td>✓ % similar between groups</td>
<td></td>
</tr>
<tr>
<td>✓ Offset by 10% over-recruitment, 98% follow-up rate</td>
<td></td>
</tr>
<tr>
<td>Patient reported exacerbations</td>
<td></td>
</tr>
<tr>
<td>✓ Patient recall has been shown to be good over 12 months</td>
<td></td>
</tr>
<tr>
<td>✓ Validation exercise showed good concordance</td>
<td></td>
</tr>
<tr>
<td>Definition of exacerbation as treatment with steroids ± antibiotics</td>
<td></td>
</tr>
<tr>
<td>✓ May underestimate short lived mild exacerbations</td>
<td></td>
</tr>
<tr>
<td>✓ No effect seen on QoL or health status</td>
<td></td>
</tr>
</tbody>
</table>


Presented at ATS International Conference 2019
Pragmatic Trial Issues

• TWICS Trial
• Lessons From Other Trials:
  ○ ELEVATE Trial
  ○ MAGNIFY Trial
Comparison of Anti-leukotriene and Inhaled Corticosteroids

ICS are the most effective preventer drug in short-term use for those with perfect inhaler technique, 15% reversibility, substantial lung function impairment, not smoke-exposed and good adherence.


LTRA as an Asthma Therapy: Re-assessing the Guideline Statement in Light of Real-life Research

<table>
<thead>
<tr>
<th>LTRA</th>
<th>ICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taken as tablet</td>
<td>Taken via inhaler</td>
</tr>
<tr>
<td>No training required</td>
<td>Patient must learn how to use inhaler</td>
</tr>
<tr>
<td>Higher adherence in studies(^1)</td>
<td>Lower adherence found</td>
</tr>
<tr>
<td>Peak action of onset within 24hrs</td>
<td>More gradual effects over several weeks</td>
</tr>
<tr>
<td>Effective in non-allergic disease</td>
<td>Selective response</td>
</tr>
<tr>
<td>Treats nose and lung in one go</td>
<td>No efficacy in nose</td>
</tr>
<tr>
<td>More effective in smokers</td>
<td>Less effective in smokers</td>
</tr>
</tbody>
</table>

ELEVATE Trial: Design

ICS: inhaled corticosteroids; LTRA: leukotriene receptor antagonists; SABA: short acting β₂-agonists

RCTs vs RWE Studies: Population vs Controlled design

- Broad Population
  - Managed as...
    - Clinical diagnosis
    - Confirmed diagnosis

- Narrow Population
  - Registration RCTs
    - Long term phase III

- Ecology of care
  - Highly controlled
  - Pragmatically controlled
  - Observational

- Constrained
- Free

ELEVATE

# ELEVATE Trial: Demographics and Drop Out Rates

Comparison to Other Studies (ELEVATE Step 2, GOAL\(^1\))

<table>
<thead>
<tr>
<th></th>
<th>ELEVATE(^1)</th>
<th>GOAL(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Step 2; N=306</td>
<td>Strata 1; N=1098</td>
</tr>
<tr>
<td><strong>Sex (% Female)</strong></td>
<td>51%</td>
<td>57%</td>
</tr>
<tr>
<td>**Age ***</td>
<td>45.8 (16.4)</td>
<td>36.3 (15.6)</td>
</tr>
<tr>
<td><strong>Quality of Life (Juniper AQLQ 1, worst, to 7)</strong></td>
<td>4.74 (1.04)</td>
<td>4.4 (1.00)</td>
</tr>
<tr>
<td>**Lung Function ***</td>
<td>86 %PPEF</td>
<td>77 %PFEV(_1)</td>
</tr>
<tr>
<td>**Percent reversibility ***</td>
<td>8.9% (9.86)</td>
<td>22% (12.2)</td>
</tr>
<tr>
<td><strong>Smokers – current</strong></td>
<td>21.9%</td>
<td>9.5%</td>
</tr>
<tr>
<td><strong>Drop out rate</strong></td>
<td>4.0%</td>
<td>15.4%</td>
</tr>
</tbody>
</table>

*Mean (SD); GOAL: The Gaining Optimal Asthma control Study; AQLQ: asthma quality of life questionnaire; NR: not reported; NA: not applicable, %PPEF: percent predicted peak expiratory flow; %PFEV\(_1\): percent predicted forced expiratory volume in one second

ELEVATE Trial: Main Results

- Inhaled corticosteroid
- LTRA

Lessons from ELEVATE

• The danger with this type of study is crossover. This was handled by performing 2 types of analyses:
  o **Intention-to-treat (ITT) analysis** – patients analysed as per their initial assigned group.
  o **Per protocol analysis** - excludes patients who deviated from the protocol.

• Poor and differential adherence, coupled with treatment crossover, potentially biases results toward equivalence.

• However, both poor and differential adherence rates are realities of real-world prescribing and thus part of the treatment effect.

• The 2-year time point better reflects the real-world effectiveness of the therapy chosen initially.
Pragmatic Trial Issues

• TWICS Trial
• Lessons From Other Trials:
  ○ ELEVATE Trial
  ○ MAGNIFY Trial
No Association Between Adherence and Clinical Outcomes

Objective: To determine whether an individualised problem-solving intervention improves ICS adherence and asthma outcomes.

Design: Participants were randomised to either problem-solving (PS, n=165) or asthma education (AE, n=168) over 3 months and followed up for another 3 months. Adherence was measured by an electronic inhaler monitor.

Inclusion criteria:
- Moderate-to-severe asthma
- ≥ 18 years old
- On ICS treatment for asthma
- With reversible airflow obstruction
MAGNIFY Technology: Propeller Health Add-on and Mobile App

Take medications on time, build the habit.
**Aim:** To evaluate the impact of an enhanced adherence package (dual bronchodilator + add-on + app) on time to treatment failure and other clinical outcomes in exacerbating COPD patients with poor adherence to mono or dual therapy over one year.

**Study Design:** A pragmatic, cluster randomised, multicentre trial over one year. Practices will be randomised to either:
- Adherence Support Arm: provision of adherence technology devices, or
- Standard of Care: continue usual routine clinical care

**Primary outcome:** Time to treatment failure

**Secondary outcomes:** Adherence based on Rx refill records over 12 months, moderate/severe exacerbations at 12 months and total exacerbations at 12 months
Utilising EHR resulting from a Quality Improvement Program in UK primary care

OPC Quality Improvement Program in 800+ GP practices

Optimum Patient Care Research Database (OPCRD)
- 8 million patients
- 800+ general practices in the UK
- Median duration of follow up: 15 years
- All primary care healthcare contacts
- Most secondary care data
- All prescribing
- Coded

Identification of Patients

EHR baseline + outcome data

Optimum Patient Care Research Database (OPCRD). Available from: https://opcrd.co.uk/
Real-World Designs and Implications for Causal Inference
Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes

Real-World Designs and Implications for Causal Inference

Vincent Willey, PharmD, BCACP
Principal Scientist
HealthCore
July 11, 2019
AIRWISE Trial Methodology

Study design

• Planned enrollment of 3200 COPD patients
  – Anthem and non-Anthem health plan members

• Randomized (1:1)
  – Dual bronchodilator therapy
  – Triple therapy

• Open label

• 12 month follow-up

• Community-based physician sites
  – Use of the administrative claims data to identify practices with the highest density of potential study patients

• Analysis
  – Non-inferiority
  – Primary endpoint: Time to first moderate or severe COPD exacerbation
Work With Existing Data To...

**Develop the protocol**
- Gaps in care
- Current treatment patterns
- Inform inclusion/exclusion criteria, sample size calculation and endpoints

**Perform protocol feasibility**
- Sites
- Patients

**Recruit sites and patients**
### “Preparatory” Retrospective Cohort Study

<table>
<thead>
<tr>
<th>Count</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>260,750</td>
<td>patients with ≥1 claim for COPD (index date), aged ≥40 years, and ≥12-month pre-index enrollment</td>
</tr>
<tr>
<td>53,484</td>
<td>newly diagnosed patients with COPD with ≥24-month post-index enrollment with ≥1 more COPD claim</td>
</tr>
<tr>
<td>14,293</td>
<td>patients (27%) had ≥1 spirometry claim in 12 months before or 12 months after index date</td>
</tr>
<tr>
<td>4130</td>
<td>screened medical record population with confirmed spirometry results</td>
</tr>
<tr>
<td>1505</td>
<td>patients with spirometry-confirmed COPD FEV₁ / FVC &lt;0.7</td>
</tr>
</tbody>
</table>


COPD = Chronic Obstructive Pulmonary Disease; FEV₁ = Forced Expiratory Volume in 1 second; FVC = Forced Vital Capacity
“Preparatory” Retrospective Cohort Study

GOLD = Global Initiative for Chronic Obstructive Lung Disease; ICS = Inhaled Corticosteroids
LAMA = Long-Acting Muscarinic Antagonists; LABA = Long-Acing Beta Agonists
AIRWISE Inclusion/Exclusion Criteria

Inclusion Criteria

• COPD diagnosis as defined by the study physician
  — no spirometry required
• Currently on one of the following COPD medication therapies
  — LAMA monotherapy
  — LABA monotherapy
  — ICS/LABA
• Physician determination that patient is not controlled on current medication therapy
• Age ≥ 40 years

Exclusion Criteria

• Currently on LAMA/LABA or triple therapy (ICS/LAMA/LABA)
• Contraindications to any of the study drugs
• Documented diagnosis of current asthma
• Pregnancy or nursing
Real-World Designs and Implications for Causal Inference

Join the conversation with #RWEregulatory
Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes: Real World Designs and Implications for Causal Inference

11 July 2019

Mark Levenson, Ph.D.
Center for Drug Evaluation and Research
Food and Drug Administration
By now we heard:

A randomized trial becomes an observational study on Day 2.
Figure 1. Cumulative Hazard Rates for the Primary Outcome of Stroke or Systemic Embolism, According to Treatment Group.

Connolly, et al. 2009, Dabigatran versus Warfarin in Patients with Atrial Fibrillation
### Table 4. Discontinuation of the Study Drug, Adverse Events, and Liver Function According to Treatment Group.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dabigatran, 110 mg (N=6015)</th>
<th>Dabigatran, 150 mg (N=6076)</th>
<th>Warfarin (N=6022)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number of patients (percent)</td>
<td>number of patients (percent)</td>
<td>number of patients (percent)</td>
</tr>
<tr>
<td><strong>Study-drug discontinuation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued at 1 yr†</td>
<td>862 (15)</td>
<td>935 (16)</td>
<td>608 (10)</td>
</tr>
<tr>
<td>Discontinued at 2 yr†</td>
<td>1161 (21)</td>
<td>1211 (21)</td>
<td>902 (17)</td>
</tr>
<tr>
<td><strong>Reason for discontinuation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient’s decision</td>
<td>440 (7.3)</td>
<td>474 (7.8)</td>
<td>375 (6.2)</td>
</tr>
<tr>
<td>Outcome event</td>
<td>192 (3.2)</td>
<td>164 (2.7)</td>
<td>130 (2.2)</td>
</tr>
<tr>
<td>Serious adverse event‡</td>
<td>163 (2.7)</td>
<td>166 (2.7)</td>
<td>105 (1.7)</td>
</tr>
<tr>
<td>Gastrointestinal symptoms¶</td>
<td>134 (2.2)</td>
<td>130 (2.1)</td>
<td>38 (0.6)</td>
</tr>
<tr>
<td>Gastrointestinal bleeding</td>
<td>58 (1.0)</td>
<td>80 (1.3)</td>
<td>54 (0.9)</td>
</tr>
</tbody>
</table>
Graham et al. 2019, Comparative Stroke, Bleeding, and Mortality Risks in Older Medicare Patients Treated with Oral Anticoagulants for Nonvalvular Atrial Fibrillation.
Estimands (ICH E9R1 Draft)

- **Population**: patients targeted by the scientific question
- **Variable (endpoint)**: obtained for each patient, that is required to address the scientific question
- **Intercurrent events**: Events that occur after randomization
- **Summary**: population-level summary for a comparison between treatment conditions
Addressing Intercurrent Events

• Treatment policy strategy (think ITT)
• While on treatment strategy (think per protocol)
• Composite strategy
  
• Hypothetical strategy
• Principal stratum strategy
  
• Others approaches
  – Fixed follow-up time.
Real-World Designs and Implications for Causal Inference
Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes:

Real World Designs and Implications for Causal Inference

Jesse A. Berlin, ScD
VP and Global Head of Epidemiology
Johson & Johnson
11 July 2019
General comments

• How do you know whether treatment effects vary across subgroups unless you include the subgroups
  • Those using the medication in practice (beyond the limited populations typically included in registration studies)
  • FDA draft guidance on increasing diversity in clinical trials

• What about a kind of compromise?
  • Do the larger study, with broad entry criteria, e.g., include elderly, those with comorbidities, etc.
  • Plan the PRIMARY analysis to focus on the narrower cohort, but now you have randomized evidence on the other populations

• Consider sub-studies
  • Validation (and applying “correction factors”)
Causal inference methods to assess safety upper bounds in randomized trials with noncompliance

Yiting Wang¹, Jesse A Berlin², Jose Pinheiro³ and Marsha A Wilcox¹

Causal inference

Abstract
Background: Premature discontinuation and other forms of noncompliance with treatment assignment can complicate causal inference of treatment effects in randomized trials. The intent-to-treat analysis gives unbiased estimates for causal effects of treatment assignment on outcome, but may understate potential benefit or harm of actual treatment. The corresponding upper confidence limit can also be underestimated.

Purpose: To compare estimates of the hazard ratio and upper bound of the two-sided 95% confidence interval from causal inference methods that account for noncompliance with those from the intent-to-treat analysis.

Methods: We used simulations with parameters chosen to reflect cardiovascular safety trials of diabetes drugs, with a focus on upper bound estimates relative to 1.3, based on regulatory guidelines. A total of 1000 simulations were run under each parameter combination for a hypothetical trial of 10,000 total subjects randomly assigned to active treatment or control at 1:1 ratio. Noncompliance was considered in the form of treatment discontinuation and cross-over at specified proportions, with an assumed true hazard ratio of 0.9, 1, and 1.3, respectively. Various levels of risk associated with being a non-complier (independent of treatment status) were evaluated. Hazard ratio and upper bound estimates from causal survival analysis and intent-to-treat were obtained from each simulation and summarized under each parameter setting.

Results: Causal analysis estimated the true hazard ratio with little bias in almost all settings examined. Intent-to-treat was unbiased only when the true hazard ratio = 1; otherwise it underestimated both benefit and harm. When upper bound estimates from intent-to-treat were <1.3, corresponding estimates from causal analysis were also <1.3 in almost 100% of the simulations, regardless of the true hazard ratio. When upper bound estimates from intent-to-treat were >1.3 and the true hazard ratio = 1, corresponding upper bound estimates from causal analysis were >1.3 in up to 66% of the simulations under some settings.

Limitations: Simulations cannot cover all scenarios for noncompliance in real randomized trials.

Conclusion: Causal survival analysis was superior to intent-to-treat in estimating the true hazard ratio with respect to bias in the presence of noncompliance. However, its large variance should be considered for safety upper bound estimates especially when the true hazard ratio = 1. Our simulations provided a broad reference for practical considerations of bias–variance trade-off in dealing with noncompliance in cardiovascular safety trials of diabetes drugs. Further research is warranted for the development and application of causal inference methods in the evaluation of safety upper bounds.
Causal inference

• Various approaches to controlling confounding and/or addressing non-compliance
• Instrumental variables (from economics)
  • Related to exposure but NOT to outcome, except through exposure
• **KEY POINT**: Randomization is the perfect instrumental variable
Wang and colleagues methods

• Simulations to reflect *randomized* cardiovascular safety trials of diabetes drugs (could be done in EHR system!)

• Focus on upper bound estimates relative to 1.3, based on regulatory guidelines (*non-inferiority*)

• Hypothetical trial of 10,000 total subjects *randomly* assigned to active treatment or control

• Treatment discontinuation and cross-over

• Considered compliers, always takers, never takers, assuming *increased risk of CV outcome with non-compliance*