Understanding Retrospective vs. Prospective Study designs

Andreas Kalogeropoulos, MD MPH PhD
Assistant Professor of Medicine (Cardiology)
Emory University School of Medicine

Emory University – Center for Faculty Development and Excellence
School of Medicine – Office of Faculty Development
Clinical Research Boot Camp
February 28, 2014
Disclosures

- Support: NHLBI U10 (HF Network)
- Support: American Heart Association (SDG)
- American Society of Echocardiography: Research Committee
- American College of Cardiology: NCDR Research & Publications Subcommittee
Objectives

1. Outline basic study designs in clinical research
2. Put clinical study designs into practical perspective
3. Discuss how to leverage existing data at Emory to support your clinical projects and grants
Clinical Study Design Jargon

❖ “Exposure”??
   o Risk Factor (observational) – may be “protective factor” also
     • Discrete
     • Continuous
   o Intervention (experimental or observational)
     o Always discrete

❖ “Outcome”??
   o Outcome as in outcome
   o Any measurable parameter of clinical interest
     • “Surrogate”

❖ “Unit”??
   • In clinical research, human subjects
   • In healthcare/policy research, can be facilities/institutions too
Basic Concept #1: Control of Exposure

- **Interventional ("Experimental")** = you control the factor of interest (the "exposure")
  - Assign an *intervention* (drug, device, diet, policy etc.) to one or more groups to see what happens

- **Observational** = you just observe what happens
  - … but remember, doesn’t mean you do not perform research procedures (e.g. contact, blood draws, tests, questionnaires etc.)!!
Basic Concept #2: Timing

- Retrospective vs. Prospective

- Retrospective:
  - The outcome has already happened (by the time of study design)!
  - Practically, you just dig into data (~EHR)
  - Can only be observational

- Prospective:
  - Interventional (has to be prospective)
  - Observational
Interventional

- Parallel
- Crossover
Interventional Designs – Parallel

- What’s special about “random”?
- Why do we need controls?
- Parallel Group Randomized Controlled Trials (RCTs)
  - Phase III – outcomes (FDA approval)
  - Phase II – “mechanistic” (surrogates)
  - Phase I – mostly safety
**Original Investigation**

**Effect of Progenitor Cell Mobilization With Granulocyte-Macrophage Colony-Stimulating Factor in Patients With Peripheral Artery Disease**

A Randomized Clinical Trial

Joseph Poole, MD, PhD; Kreton Mavromatis, MD; José N. Binongo, PhD; Ali Khan, MD; Qunna Li, MSc; Mohamed Khayata, MD; Elizabeth Rocco, BS; Matthew Topel, MD; Xin Zhang, MS; Charlene Brown, RN; Matthew A. Corriere, MD; Jonathan Murrow, MD; Salman Sher, MD; Stephanie Clement, MD; Khuram Ashraf, MD; Amr Rashed, MD; Tarek Kabbany, MD; Robert Neuman, MD; Alanna Morris, MD; Arshad Ali, MD; Salim Hayek, MD; John Oshinski, PhD; Young-sup Yoon, MD; Edmund K. Waller, MD, PhD; Arshed A. Quyyumi, MD

**Poole et al, JAMA 2013**

**DESIGN, SETTING, AND PARTICIPANTS** In a phase 2 double-blind, placebo-controlled study, 159 patients (median [SD] age, 64 [8] years; 87% male, 37% with diabetes) with intermittent claudication were enrolled at medical centers affiliated with Emory University in Atlanta, Georgia, between January 2010 and July 2012.

**INTERVENTIONS** Participants were randomized (1:1) to received 4 weeks of subcutaneous injections of GM-CSF (leukine), 500 μg/day 3 times a week, or placebo. Both groups were encouraged to walk to claudication daily.

**MAIN OUTCOMES AND MEASURES** The primary outcome was peak treadmill walking time (PWT) at 3 months. Secondary outcomes were PWT at 6 months and changes in circulating PC levels, ankle brachial index (ABI), and walking impairment questionnaire (WIQ) and 36-item Short-Form Health Survey (SF-36) scores.
Interventional Designs – Crossover

- Randomized – usually in two groups
- Sequential – trajectory
- Before-After = small sample size (efficient)
- By definition “mechanistic”
  - Assess treatment effects on surrogate endpoints
- Crossover Randomized Before-After Studies
  - Why controls?
  - Why crossover?
  - More solid design for short-term treatment effects
Stimulus Intensity in Left Ventricular Leads and Response to Cardiac Resynchronization Therapy

Venkata V. Bavikati, MD; Jonathan J. Langberg, MD; B. Robinson Williams, III, MD; Danesh Kella, MD; Michael S. Lloyd, MD

Background—Increased left ventricular (LV) stimulus intensity has been shown to improve conduction velocity and cardiac output. However, high-output pacing would shorten device battery life. Our prospective trial analyzed the clinical effects of high- versus low-output LV pacing.

Methods and Results—Thirty-nine patients undergoing initial cardiac resynchronization therapy device implantation with bipolar LV leads were assigned to 3 months of either high-output LV pacing (Hi) or low-output LV pacing (Lo) in a randomized, blinded crossover fashion. Hi and Lo settings were determined with a rigorous intraoperative protocol specific to each patient. Clinical and echocardiographic data were obtained at randomization, at 3 months, and a subsequent 3 months after crossover. Mean age was 66.4 ± 9.8 years, and mean QRS duration was 159.3 ± 23.1 ms. Compared to baseline, both arms had significant improvements in Minnesota Living With Heart Failure score (given as mean [95% confidence interval]) (baseline versus Lo: 43.3 [35.5 to 51.1] versus 21.3 [14.6 to 28.0], P < 0.01; baseline versus Hi: 43.3 [35.5 to 51.1] versus 23.6 [16.1 to 31.1], P < 0.01) and 6-minute walk distance (baseline versus Lo: 692 ft [581 to 804] versus 995 ft [876 to 1114], P < 0.01; baseline versus Hi: 699 ft [585 to 813] versus 982 ft [857 to 1106], P < 0.01). Although both Hi and Lo arms had some echocardiographic parameters that significantly improved compared to baseline (baseline end-diastolic diameter 5.7 cm [5.5 to 6.0] versus Lo 5.5 cm [5.1 to 5.8], P < 0.01; baseline end-systolic diameter 4.9 cm [4.6 to 5.3] versus Hi 4.7 cm [4.3 to 5.0], P < 0.05), there were no significant differences observed when comparing the Hi- versus Lo-output arms.

Conclusions—Low-output LV pacing with a relatively narrow safety margin above capture threshold affords significant improvement from baseline and is clinically equivalent to high-output LV pacing. These data support a strategy of minimizing the programmed LV safety margin to increase battery life in cardiac resynchronization therapy devices.


Key Words: cardiac resynchronization • congestive heart failure • pacing
Basic Concept #3: Directionality

- Observational
  - Cross-sectional
  - Case-control
  - Cohort
Cross-Sectional = No Direction

different groups

- group 1
- group 2
- group 3

compared at one time

Image: http://education-portal.com
Cross-Sectional – Uses and Abuses

- Excellent for descriptive purposes (think surveys)
- Horrible for inference (temporality principle)
- Uses: “Statement of the problem”
Case-Control = Backward

The study begins by selecting subjects based on disease (cases).

Exposed

Unexposed

Review records

Exposed

Unexposed

Review records

No disease (controls)
Case-Control – Uses and Abuses

- Excellent for rare outcomes
- Can only examine one outcome
- Can examine multiple exposures
- Takes sophistication to sell 😊
- “Hypothesis-generating”
- Modest strength for inference (prone to bias)
Cohort = Forward

Study begins → Outcomes

Exposed group

Disease

No disease

Unexposed group

Disease

No disease

Image: U of Ottawa
Cohort – Uses and Abuses

- Excellent for common outcomes
- Can examine multiple exposures
- Can examine multiple outcomes
- “Hypothesis-generating”
- Strongest observational design
- … but still affected by various biases (“unobserved confounding”)
Hierarchy of Inference

Opinion

Observational
  Retrospective

Association

Observational
  Prospective

Causation

Interventional
Can do retro—or pro-spectively!!

Observational

- Cross-sectional
  - Case-control
  - Cohort
Retrospective: Pros and Cons

- **Pros:**
  - [Relatively] inexpensive process (database, data collection, analysis)
  - Quick results

- **Cons:**
  - Missing data on both sides (potential bias)
  - Definitions adapted to circumstances
  - Unmeasured confounders (afterthoughts)
  - Harder to sell 😞

- **Uses?**
Retrospective: Example

Kalogeropoulos et al, JACC 2009
Retrospective: Uses

<table>
<thead>
<tr>
<th>Variable</th>
<th>All (n = 860)</th>
<th>Event (n = 109)</th>
<th>Non-Event (n = 751)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yrs</td>
<td>59.2 ± 13.4</td>
<td>51.3 ± 15.1</td>
<td>59.4 ± 15.4</td>
<td>0.041</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>68.6</td>
<td>74.3</td>
<td>66.7</td>
<td>0.105</td>
</tr>
<tr>
<td>Race, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>52.4</td>
<td>54.1</td>
<td>51.6</td>
<td>0.966</td>
</tr>
<tr>
<td>Black</td>
<td>44.6</td>
<td>44.1</td>
<td>44.9</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3.1</td>
<td>2.8</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>29.6 ± 7.9</td>
<td>29.7 ± 7.7</td>
<td>29.3 ± 8.0</td>
<td>0.119</td>
</tr>
<tr>
<td>Ischemic etiology, %</td>
<td>38.2</td>
<td>42.2</td>
<td>36.9</td>
<td>0.264</td>
</tr>
<tr>
<td>NYHA functional class</td>
<td>2.5 ± 0.7</td>
<td>2.9 ± 0.7</td>
<td>2.4 ± 0.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Left ventricular ejection fraction, %</td>
<td>58.2 ± 7.9</td>
<td>55.9 ± 6.1</td>
<td>60.0 ± 6.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Devices, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defibrillator</td>
<td>30.3</td>
<td>18.3</td>
<td>34.2</td>
<td>0.002</td>
</tr>
<tr>
<td>Biventricular pacemaker</td>
<td>3.4</td>
<td>7.3</td>
<td>2.1</td>
<td>0.014</td>
</tr>
<tr>
<td>Combined</td>
<td>37.3</td>
<td>41.5</td>
<td>36.0</td>
<td>0.360</td>
</tr>
<tr>
<td>Systolic blood pressure, mmHg</td>
<td>154.2 ± 18.7</td>
<td>150.2 ± 18.4</td>
<td>153.5 ± 18.7</td>
<td>0.002</td>
</tr>
<tr>
<td>Heart rate, beats/min</td>
<td>78 ± 14</td>
<td>82 ± 16</td>
<td>76 ± 13</td>
<td>0.008</td>
</tr>
<tr>
<td>Sodium, mEq/l</td>
<td>137.4 ± 3.4</td>
<td>138.9 ± 3.3</td>
<td>137.9 ± 3.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Potassium, mEq/l</td>
<td>4.1 ± 0.6</td>
<td>4.5 ± 0.6</td>
<td>4.0 ± 0.6</td>
<td>0.033</td>
</tr>
<tr>
<td>Creatinin, mg/dl</td>
<td>1.4 ± 1</td>
<td>1.5 ± 1.4</td>
<td>1.4 ± 0.6</td>
<td>0.003</td>
</tr>
<tr>
<td>Blood urea nitrogen, mg/dl</td>
<td>24.7 ± 18.4</td>
<td>28.5 ± 15.4</td>
<td>21.5 ± 16.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Glucose, mg/dl</td>
<td>110.1 ± 66.9</td>
<td>118.7 ± 46.6</td>
<td>110.5 ± 72.3</td>
<td>0.033</td>
</tr>
<tr>
<td>Cholesterol, mg/dl</td>
<td>129.3 ± 38.6</td>
<td>158.4 ± 42.7</td>
<td>129.6 ± 37.3</td>
<td>0.011</td>
</tr>
<tr>
<td>Triglycerides, mg/dl</td>
<td>94 ± 94</td>
<td>95 ± 12</td>
<td>94 ± 94</td>
<td>0.582</td>
</tr>
<tr>
<td>Hemoglobin, g/dl</td>
<td>13.3 ± 1.8</td>
<td>13.3 ± 1.8</td>
<td>13.3 ± 1.8</td>
<td>0.003</td>
</tr>
<tr>
<td>White blood cells, 10^3/mm^3</td>
<td>8.3 ± 2.2</td>
<td>7.7 ± 2.0</td>
<td>7.4 ± 2.3</td>
<td>0.045</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>24.5 ± 7.5</td>
<td>26.4 ± 8.0</td>
<td>23.6 ± 7.3</td>
<td>0.001</td>
</tr>
<tr>
<td>Albumin, g/dl</td>
<td>3.5 ± 0.8</td>
<td>3.4 ± 0.7</td>
<td>3.5 ± 0.7</td>
<td>0.043</td>
</tr>
<tr>
<td>Comorbidities, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>63.0</td>
<td>56.6</td>
<td>65.4</td>
<td>0.124</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>39.3</td>
<td>39.3</td>
<td>39.3</td>
<td>1.000</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>44.7</td>
<td>48.6</td>
<td>44.0</td>
<td>1.000</td>
</tr>
<tr>
<td>Depression</td>
<td>20.8</td>
<td>37.6</td>
<td>20.9</td>
<td>0.003</td>
</tr>
<tr>
<td>Chronic lung disease</td>
<td>20.3</td>
<td>35.6</td>
<td>20.6</td>
<td>0.003</td>
</tr>
<tr>
<td>Medications, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE inhibitor or ARB</td>
<td>90.8</td>
<td>90.3</td>
<td>93.4</td>
<td>0.111</td>
</tr>
<tr>
<td>Betablockers</td>
<td>36.9</td>
<td>74.3</td>
<td>97.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aldosterone antagonists</td>
<td>46.3</td>
<td>52.3</td>
<td>44.3</td>
<td>0.193</td>
</tr>
<tr>
<td>Diuretics</td>
<td>87.8</td>
<td>95.4</td>
<td>85.3</td>
<td>0.004</td>
</tr>
<tr>
<td>Angiotensin</td>
<td>4.8</td>
<td>4.8</td>
<td>0.3</td>
<td>0.314</td>
</tr>
<tr>
<td>Diuretics</td>
<td>52.3</td>
<td>75.1</td>
<td>46.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Statins</td>
<td>43.9</td>
<td>38.5</td>
<td>45.5</td>
<td>0.222</td>
</tr>
<tr>
<td>Antiarhythmics</td>
<td>27.2</td>
<td>33.9</td>
<td>25.1</td>
<td>0.083</td>
</tr>
</tbody>
</table>

*Available in 406 of 445 (91.2%) patients; † available in 314 of 448 (70.0%) patients; ‡ available in 406 of 445 (91.2%) patients; § available in 429 of 445 (95.9%) patients.

ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; NYHA = New York Heart Association.
Retrospective: How to

- System, system, system:
  - Clear definitions of populations and timeframe
  - Clear definitions of outcomes
  - Collect as much data and definition elements as possible (hard to go back!) – and be quantitative!
  - Think publication: how am I going to defend this?
  - Design database carefully
  - Find people to help (it’s a lot of work!)
  - Now you can start
Prospective: Pros and Cons

- **Pros:**
  - High quality data
  - Future proof (for blood biomarkers)
  - Strong validity
  - Can accommodate novelty

- **Cons:**
  - Expensive process (=you need funding!)
  - Takes time to design (!!) & conduct

- Uses?
Prospective – Large Cohorts

Inflammatory Markers and Incident Heart Failure Risk in Older Adults

The Health ABC (Health, Aging, and Body Composition) Study

Andreas Kalogeropoulos, MD,* Vasiliki Georgiopoulou, MD,* Bruce M. Psaty, MD, PhD,† Nicolas Rodondi, MD, MAS,‡ Andrew L. Smith, MD,* David G. Harrison, MD,* Yongmei Liu, MD,§ Udo Hoffmann, MD, MPH,∥ Douglas C. Bauer, MD,¶ Anne B. Newman, MD, MPH,№ Stephen B. Kritchevsky, PhD,§ Tamara B. Harris, MD, MS,** Javed Butler, MD, MPH,* for the Health ABC Study Investigators

Atlanta, Georgia; Seattle, Washington; Lausanne, Switzerland; Winston-Salem, North Carolina; Boston, Massachusetts; San Francisco, California; Pittsburgh, Pennsylvania; and Bethesda, Maryland

Objectives
The purpose of this study was to evaluate the association between inflammation and heart failure (HF) risk in older adults.

Background
Inflammation is associated with HF risk factors and also directly affects myocardial function.

Methods
The association of baseline serum concentrations of interleukin (IL)-6, tumor necrosis factor-α, and C-reactive protein (CRP) with incident HF was assessed with Cox models among 2,610 older persons without prevalent HF enrolled in the Health ABC (Health, Aging, and Body Composition) study (age 73.6 ± 2.9 years; 48.3% men; 59.6% white).

Results
During follow-up (median 9.4 years), HF developed in 311 (11.9%) participants. In models controlling for clinical characteristics, ankle-arm index, and incident coronary heart disease, doubling of IL-6, tumor necrosis factor-α, and CRP concentrations was associated with 29% (95% confidence interval: 13% to 47%; p < 0.001), 46% (95% confidence interval: 17% to 84%; p = 0.001), and 9% (95% confidence interval: −1% to 24%; p = 0.087) increase in HF risk, respectively. In models including all 3 markers, IL-6, and tumor necrosis factor-α, but not CRP, remained significant. These associations were similar across sex and race and persisted in models accounting for death as a competing event. Post-HF ejection fraction was available in 239 (78.6%) cases; inflammatory markers had stronger association with HF with preserved ejection fraction. Repeat IL-6 and CRP determinations at 1-year follow-up did not provide incremental information. Addition of IL-6 to the clinical Health ABC HF model improved model discrimination (C Index from 0.717 to 0.734; p = 0.001) and fit (decreased Bayes information criterion by 17.8; p < 0.001).

Conclusions
Inflammatory markers are associated with HF risk among older adults and may improve HF risk stratification.

(J Am Coll Cardiol 2010;55:2129–37) © 2010 by the American College of Cardiology Foundation

Kalogeropoulos et al, JACC 2010
Prospective – Homegrown Cohorts

Right Ventricular Function with Standard and Speckle-Tracking Echocardiography and Clinical Events in Adults with D-Transposition of the Great Arteries Post Atrial Switch

Andreas P. Kalogeropoulos, MD, Anjan Deka, MD, William Border, MBChB, MPH, Maria A. Pernetz, RDCS, Vasiliki V. Georgiopoulou, MD, Jawad Kiani, MD, Michael McConnell, MD, Stamatios Lerakis, MD, Javed Butler, MD, MPH, Randolph P. Martin, MD, and Wendy M. Book, MD, Atlanta, Georgia

This work was partially supported by an American Society of Echocardiography Echo Investigator Award (2008) and an Emory Heart & Vascular Board grant titled “Novel Markers and Outcomes in Heart Failure.”

Methods: Sixty-four adults with D-transposition of the great arteries and prior atrial switch (mean age, 29 ± 6 years; 22 women; mean right ventricular [RV] fractional area change, 22.9 ± 7.5%; 31 with pacemakers at baseline) and no histories of heart failure or ventricular tachycardia were prospectively evaluated. Global longitudinal strain (GS), global systolic strain rate (GSRs), and global early diastolic strain rate (GSRe) of the right ventricle were measured using speckle tracking from apical views and compared with standard parameters of RV function (fractional area change, tricuspid annular plane systolic excursion, tissue Doppler velocities, and isovolumic acceleration) for association with and potential prediction of clinical events, defined as incident stage C heart failure or ventricular tachycardia.
The Cohort Is Your Friend

- Second only to RCTs in quality of evidence
- Obviously pro- is better than retro-spective
- Large retro-cohorts are feasible in Emory (CDW)
- Small to mid-size pro-cohorts are feasible in Emory
- Multipurpose databases
- Start a study or get the data from somewhere??
  - Depends on the purpose
Prospective Cohort: Existing DB

• NIH cohort or RCTs databases
• **Pro:** the data are there!
• **Con:** …but maybe not exactly the data you want
• **Pro:** faster publication cycle
• **Con:** …but politics not always straightforward

• **Purpose:** prove concept with “highly visible” data
• **Drawback:** you not gonna get any money in 2014
Prospective Cohort: Create the Data

- The real thing!
- Fundable (by definition)
- Doable in Emoryland – for clinical researchers:
  - Lots of prevalent disease
  - “Outcomes” (services) research

- **Purpose:** real portfolio
- **Drawback:** it’s a lot of work
Basic Clinical Study Design - Summary

- Interventional studies: you test something (=prospective)
- Observational: risk factors, biomarkers, models etc.
  - Retrospective vs. Prospective
    - Interventional (e.g. RCT)
    - Prospective observational (e.g. cohort)
    - Retrospective observational (e.g. CDW)
When do you need IRB approval?

- Human Subjects Research: Always!
- Any prospective study: participant needs to sign up!
- ICF waived only if ALL data exist by the time of request
- External databases: you may still need IRB approval
- In any case, you can always ask 😊
Questions?