Medicaid DM Programs:

How to measure and improve success.

Disease Management Colloquium

June 28, 2004

Thomas W. Wilson, PhD, MPH, DrPH
Trajectory Healthcare, LLC
10663 Loveland-Madeira Rd, #210
Loveland, Ohio 45140
twilson@trajectory-inc.com
513.289.3743

TRAJECTORY™
It's About Time℠

Extension of ROI article

Wilson T. Evaluating ROI in State DM Programs. Robert Woods Johnson's State Coverage Initiative Program, 2003. www.statecoverage.net.

New Problems Faced by DM Medicaid.

- The magnitude of ROI will likely drop with the emergence of credible and valid methodologies.
  - *Source: Al Lewis, Disease Management Purchasing Consortium International, 6/27/04*
- The Florida DM Medicaid Experience seems to bear this out.
  - *Source:* http://www.oppaga.state.fl.us/reports/health/r04-34s.html
- As the DM movement evolves, it must be prepared to improve its' value.
  - *How?*

Solution:
Improvement Strategies

- **Choices: Reduce Cost or Improve Impact.**
- **What is can we correct?**
  - Better choice of the optimal population.
  - Reduce what we are doing that does not work.
  - Increase what we are not doing well that does work
- **How? #1: Ability to Distinguish what works from what does not work.**
  - *Rely on Evidence-based practices*
  - *Add to Evidence-based practices.*
- **How? #2: Hypothesis-driven Action**
  - *Intelligent Action*

ORGANIZATION

- I) DEFINITION OF PRAGMATIC EPIDEMIOLOGY
- II) IMPACT | CAUSALITY ASSESSMENT
  - METRICS
  - EQUIVALENCE
  - COMPARABILITY
- III) HYPOTHESIS-DRIVEN ACTION
  - With and without strong EBM.

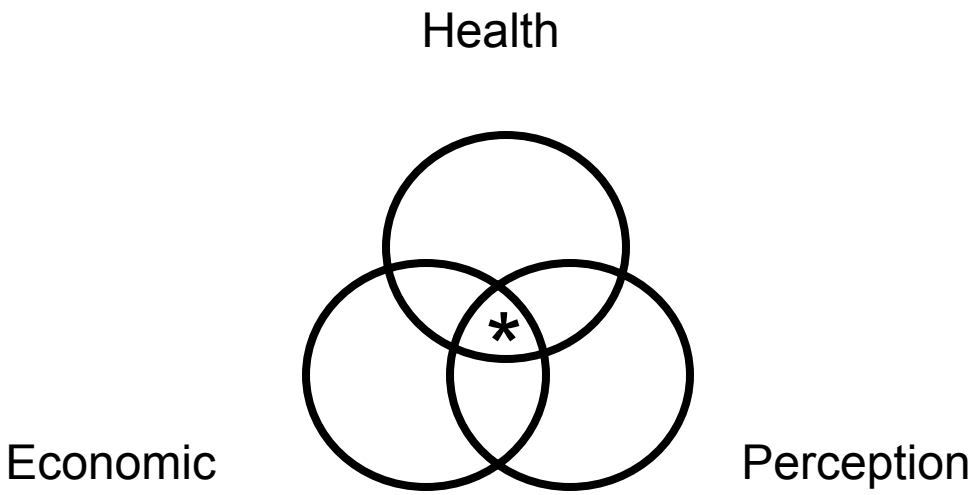
I) Definition

Pragmatic Epidemiology:
Epidemiology of Value

The *scientific* study of the distribution and determinants of health-related *value* in defined populations, and the application of this study to the control of health-*related value* problems.

“Value”: Operational Definition

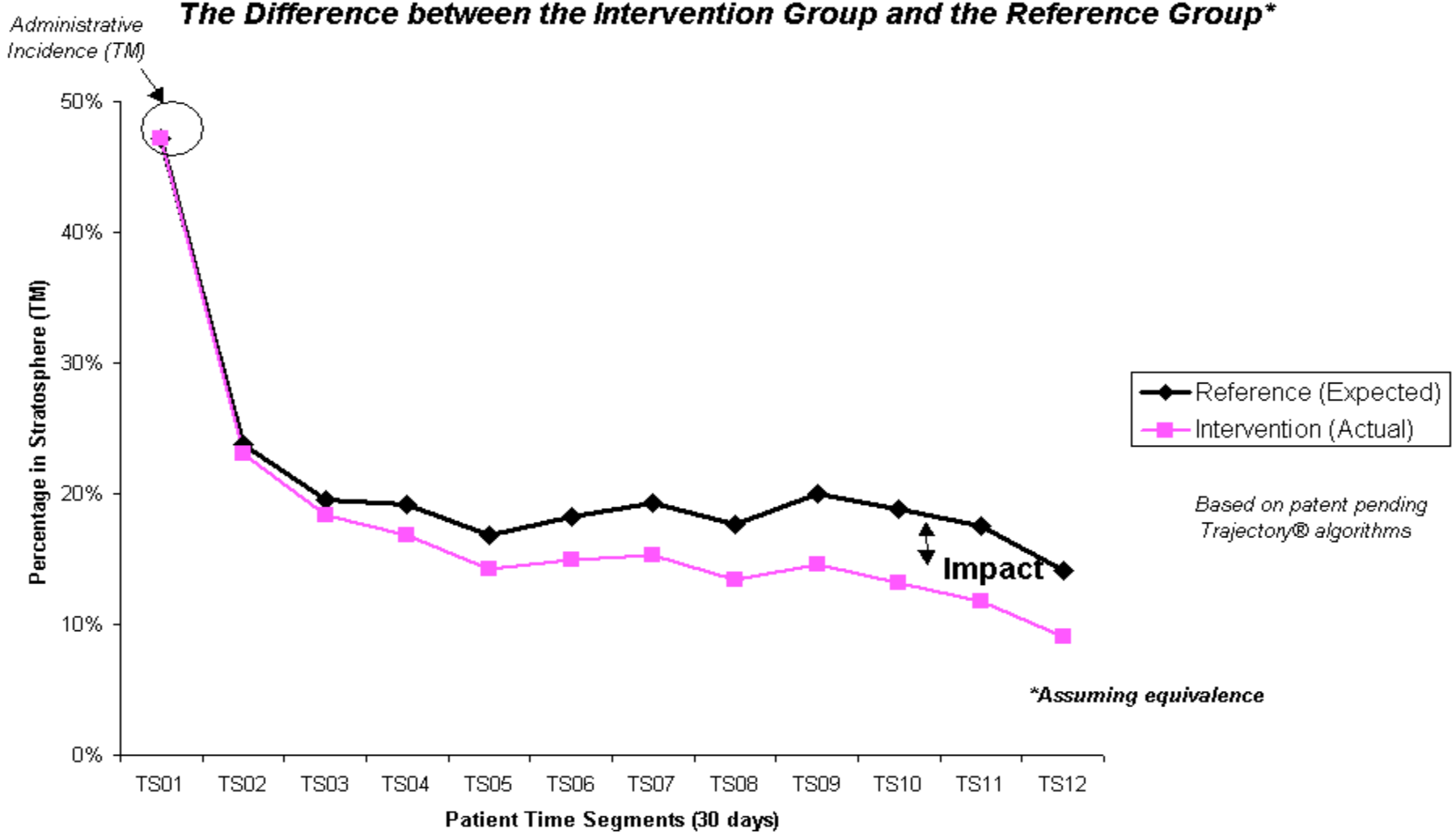
Person/Population



II) IMPACT | CAUSALITY ASSESSMENT

IMPACT

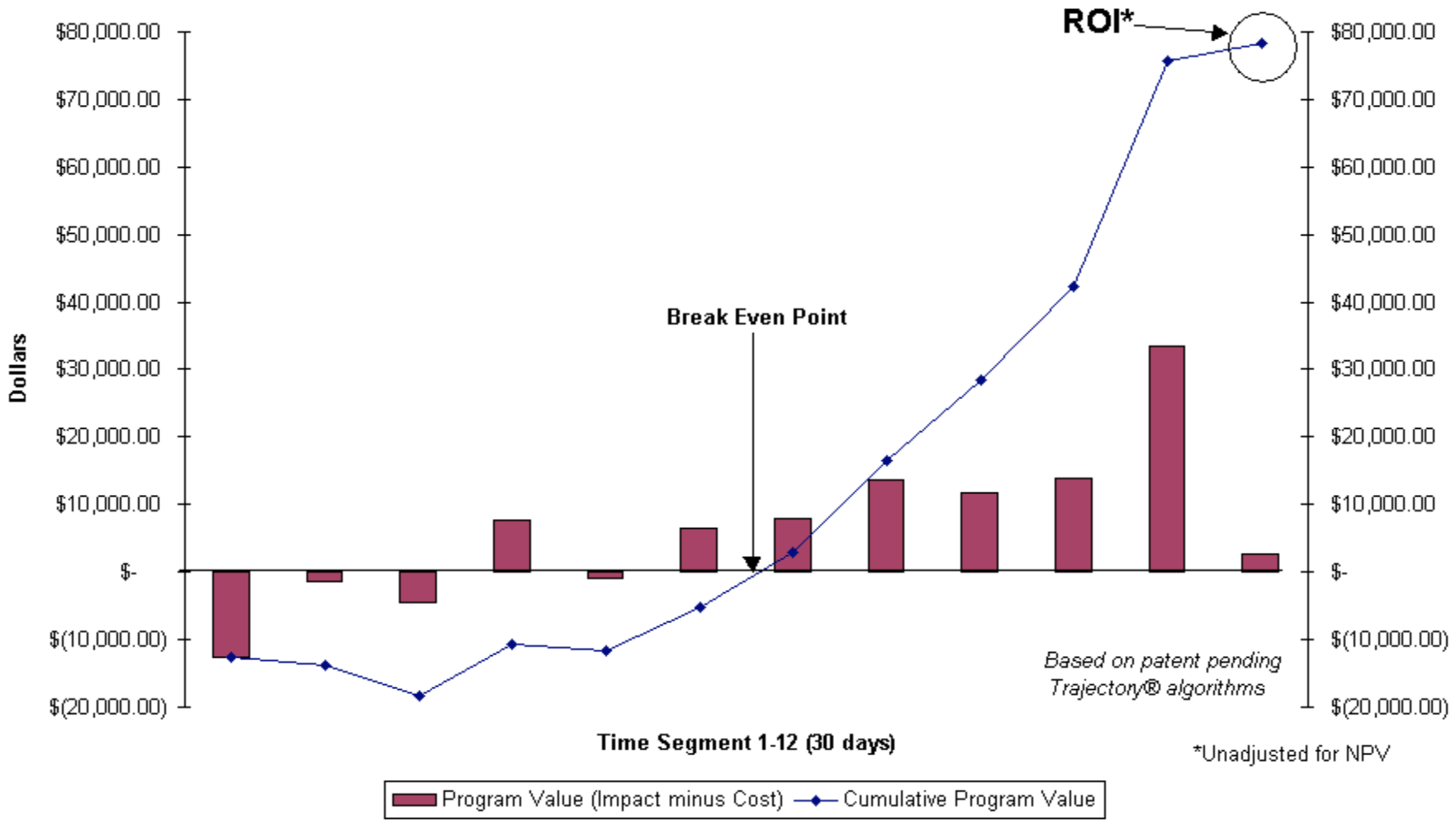
The Difference between the Intervention Group and the Reference Group*



R.O.I.

Impact minus cost of solutions

Value (Impact - Cost) Per Time Segment and Cumulative Program Value (ROI) over All Time Segments



B) THE EXPECTED

Q: How do we credibly determine the "expected"?

A: We need an "equivalent" reference group

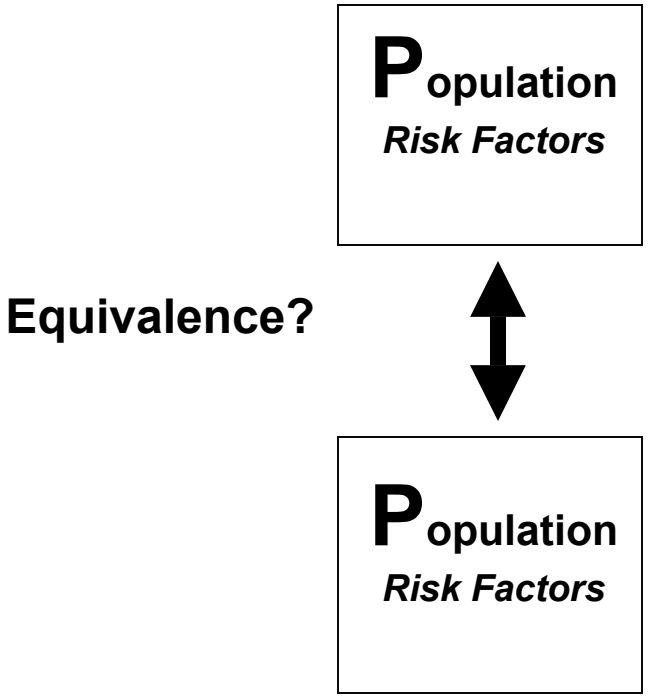
Many ways to answer this question

… some are very good, some are very bad.

- Patient's experience last year?
- Participants compared to non-participants?
- Published literature (inferred reference)?
- External "matched" control group?
- Field-based randomized trial??
- Double-blind, randomized control trial

EQUIVALENCE OF POPULATIONS

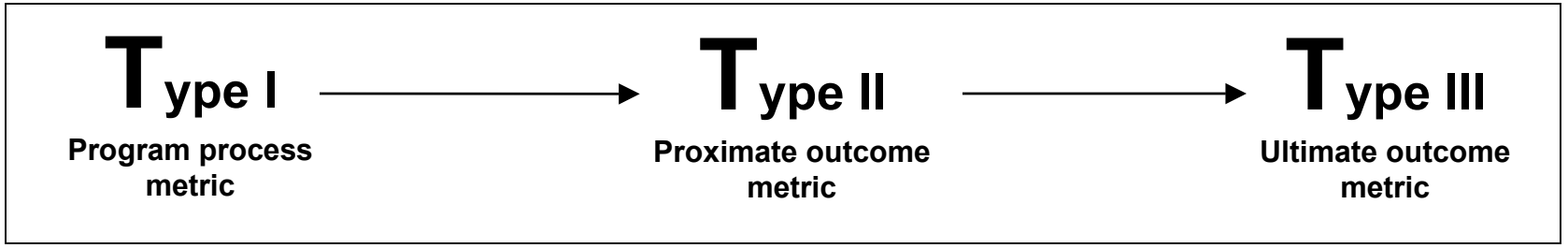
Disease Management Population

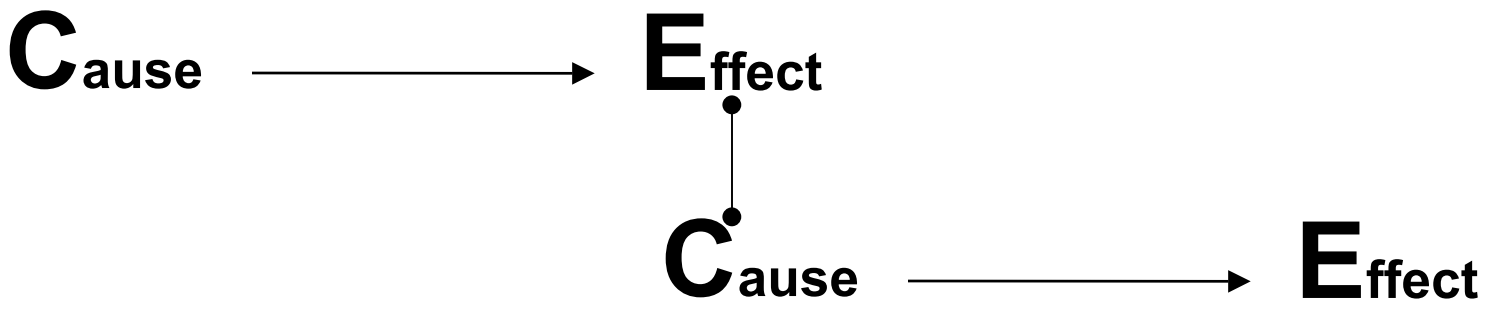


Reference Population

Intervention Pathway:

Metrics & The Scientific Equation: Cause & Effect (ad infinitum)





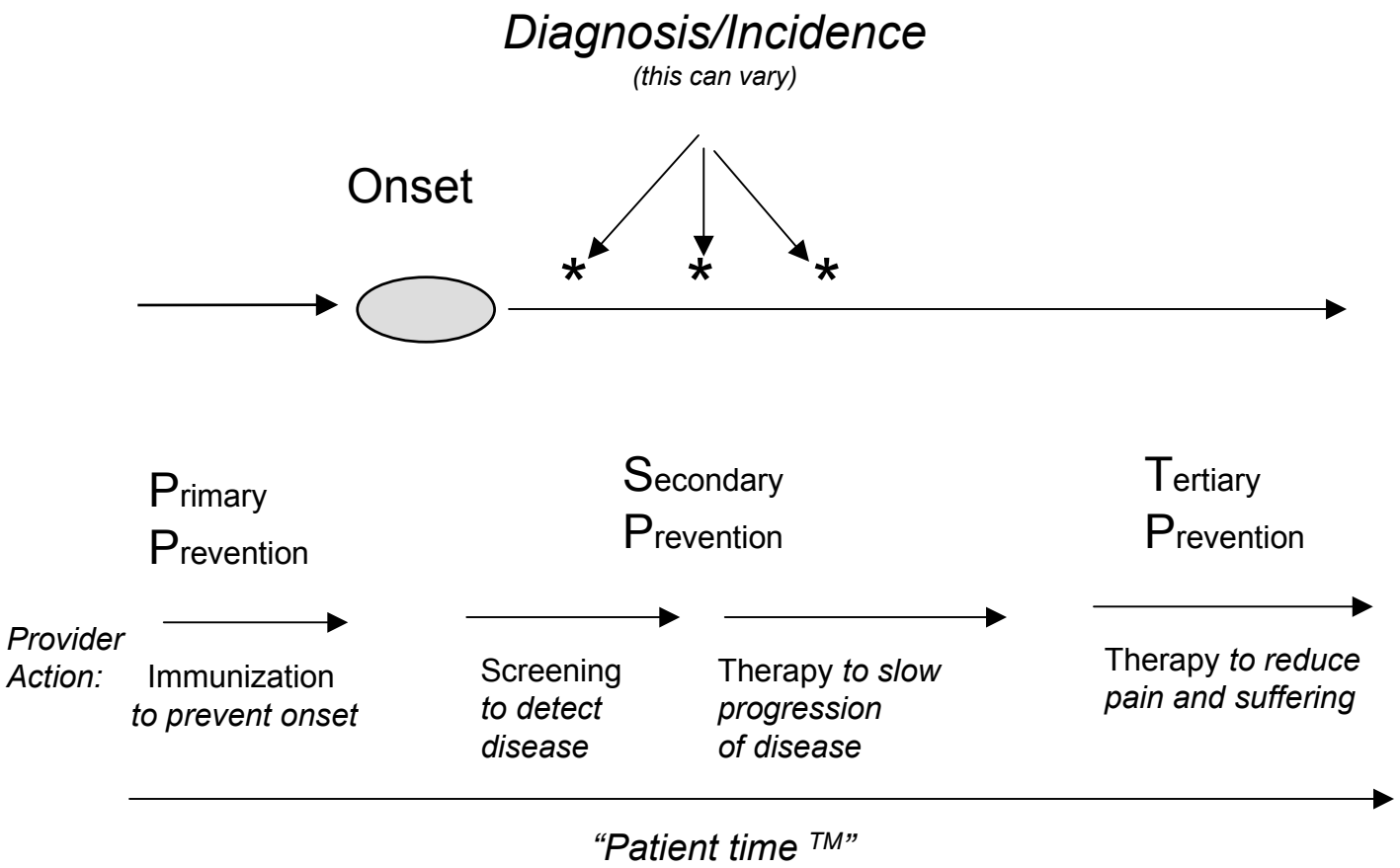
METRICS: What is a Type I, Type II, and Type III Metric?

Type I: CAUSES
- DM Program

Type II: PROXIMATE IMPACT
- Screening, Compliance, etc.

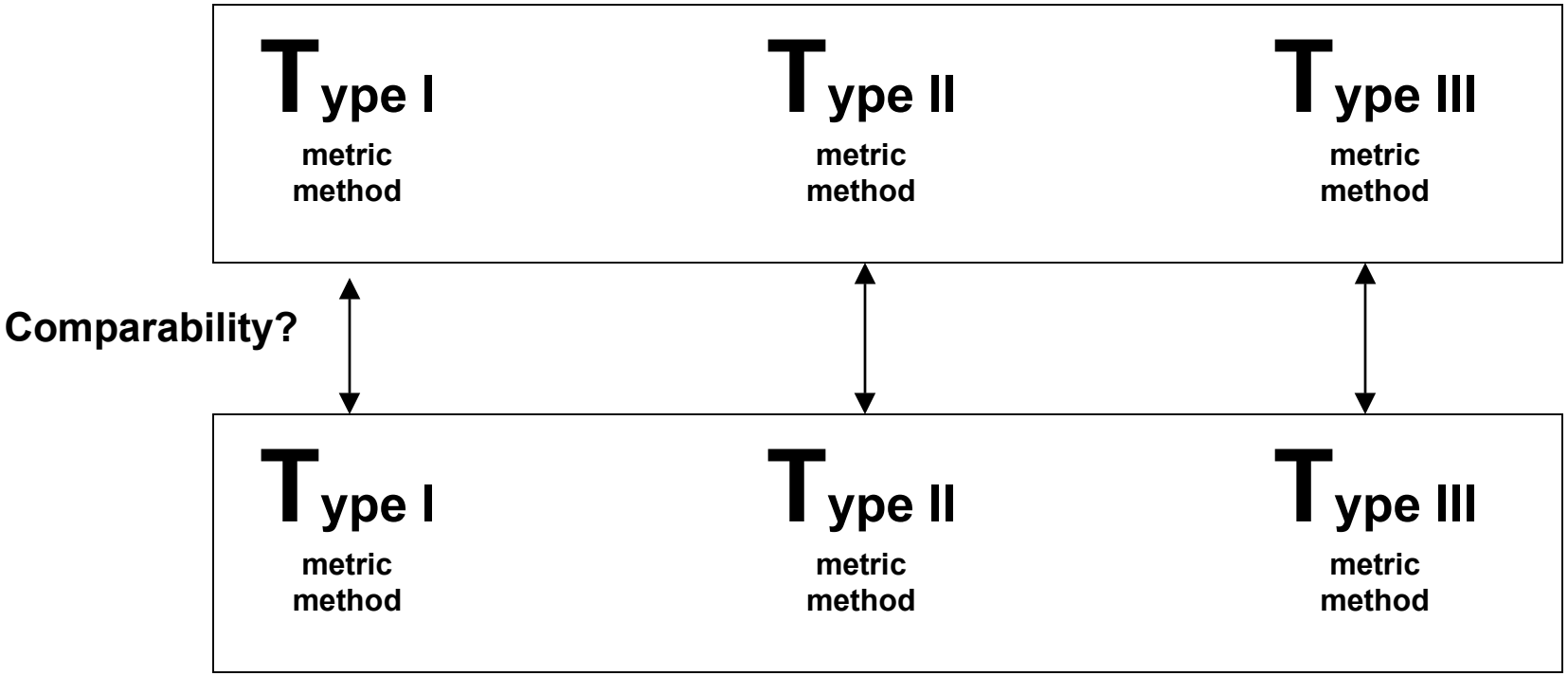
Type III: ULTIMATE IMPACT
- Health: Incidence, Prevalence, QALY, Biol.
- Economic: Claims Payments, Admissions
- Perception: Patient & Provider Satisfaction

Natural History of Health/Disease & Health Care Provider Response



COMPARABILITY OF METRICS

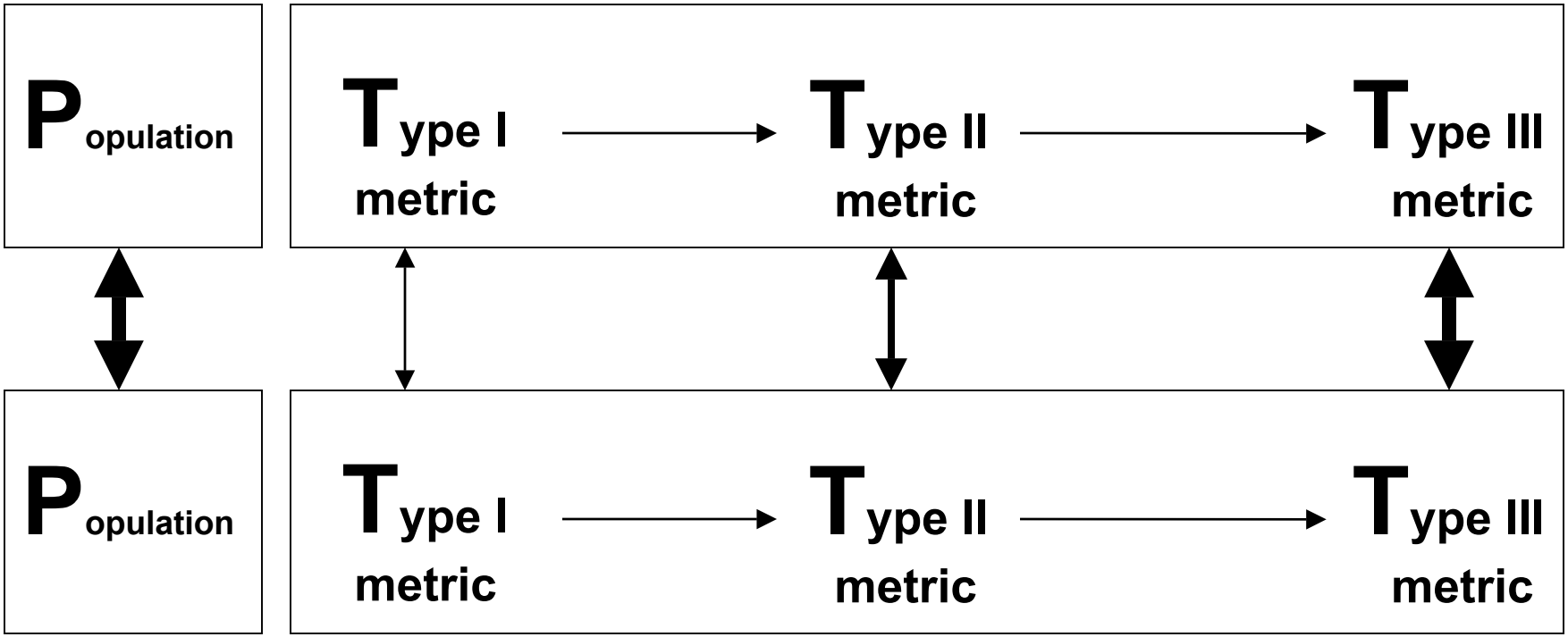
Disease Management Population



Reference Population

EQUIVALENCE | IMPACT

Disease Management Population



Reference Group

B) Study Designs (selected)

I. Post-Only

II. Benchmark

III. Pre-Post Type Designs (Quasi-Experimental)

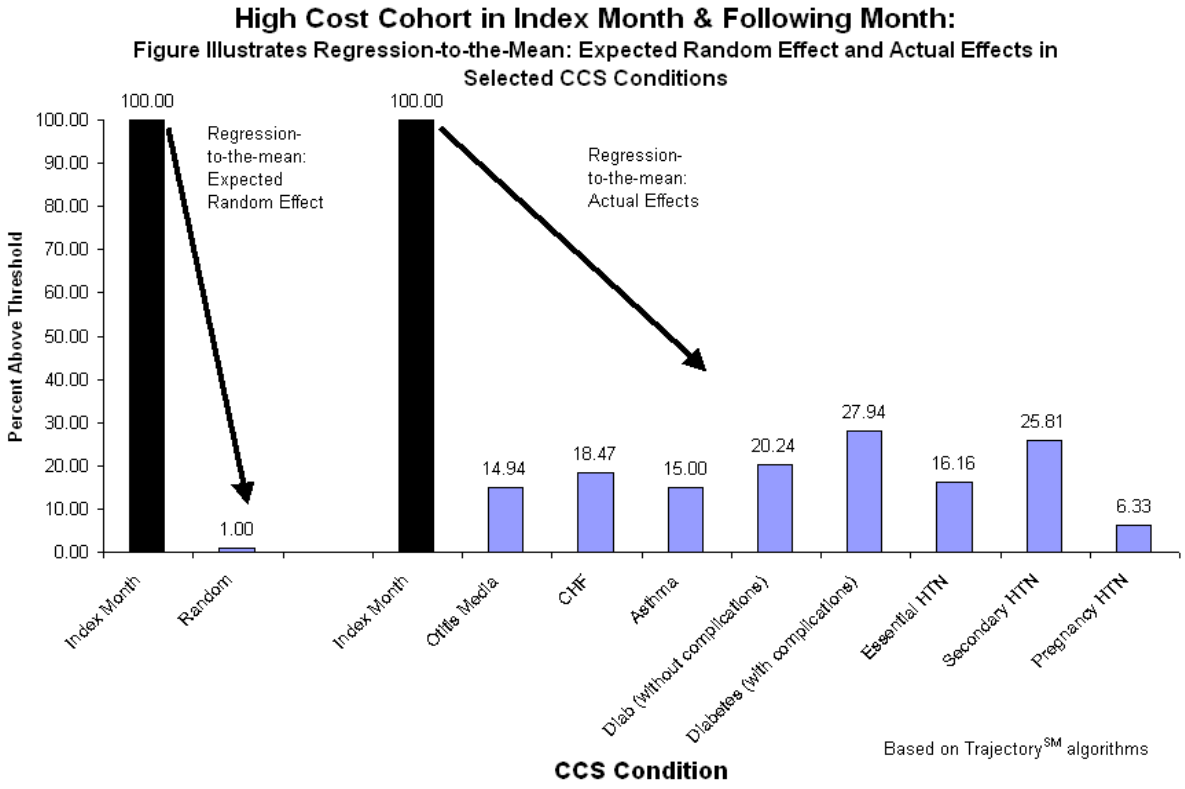
IV. Follow-up / Cohort

I. Post-Only

- Where results of a panel are not compared to any reference group.
  - i. Patient Selection:
    - Population-based
    - Referral, Outlier, etc. based

  - *What to watch out for: Design most likely to be misinterpreted if regression-to-the-mean and the "natural history of disease" are not taken into account. This is especially true on patients selected because they are outliers.*

"How I Learned to Stop Worrying and Love Regression to the Mean"



Adapted from Wilson TW. "How I Learned to Stop Worrying and Love Regression-to-the-mean." Presented at American Association of Health Plans--Building Bridges Conference, April 2002.

II. Benchmark

- A) Where results of a case series are compared to a national benchmark (e.g. HEDIS)

- B) Where case series (population and results) are compared with results, based upon an equivalent population, from a study from well-designed peer-reviewed journal.
  - Sources:
    - *Rich MW, Beckham V, Wittenberg C, Leven CL, Freedland KE, Carney RM. A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure. N Engl J Med 1995; 333:1190-5.*
    - *Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 329:977-986, 1993*

Benchmark (con't)

- C) Predictive Modeling: Where case series (population and results) are compared with results that are "predicted" to occur from a predictive modeling algorithm based upon a "beta weights" from another population.
  - *Sources:*
    - *Brindle P, Emberson J, Lampe F, Walker M, Wincup P, Fahey T, Ebraham S. Predictive Accuracy of Framingham coronary risk score in British men: prospective cohort study. British Medical Journal. 2003; 327: 1267-1271.*
    - *Fairman KA, Motherall BR. Do Decision Analytic Models Identify Cost-Effective Treatments? A Retrospective Look at Helicobacter Pylori Eradication. Journal of Managed Care Pharmacy. 2003 9(5): 323-333.*

III. Pre-Post (Quasi-Experimental)

- **Types:**

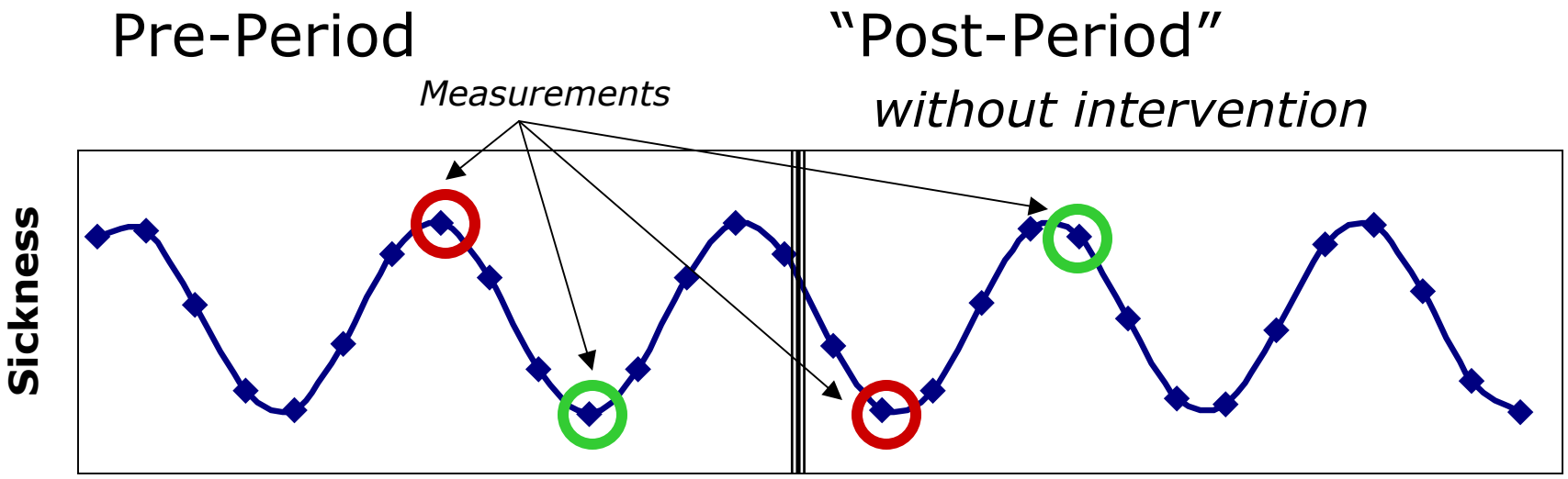
  a) **Classic Pre-Post:** results on same patients from a prior time period ("patient as their own control").

  b) **Time-series**: results on patients in multiple time periods prior to intervention compared to same patients in multiple time periods AFTER the intervention.

  What to watch out for: "Lost to baseline." Metric comparability. Natural history of disease could render the pre period a poor predictor of the post period in the patients studied.

Pre-Post Design:

Past is NOT Prologue: *Another Situation where equivalence is not achieved (if you're Red or Green)*

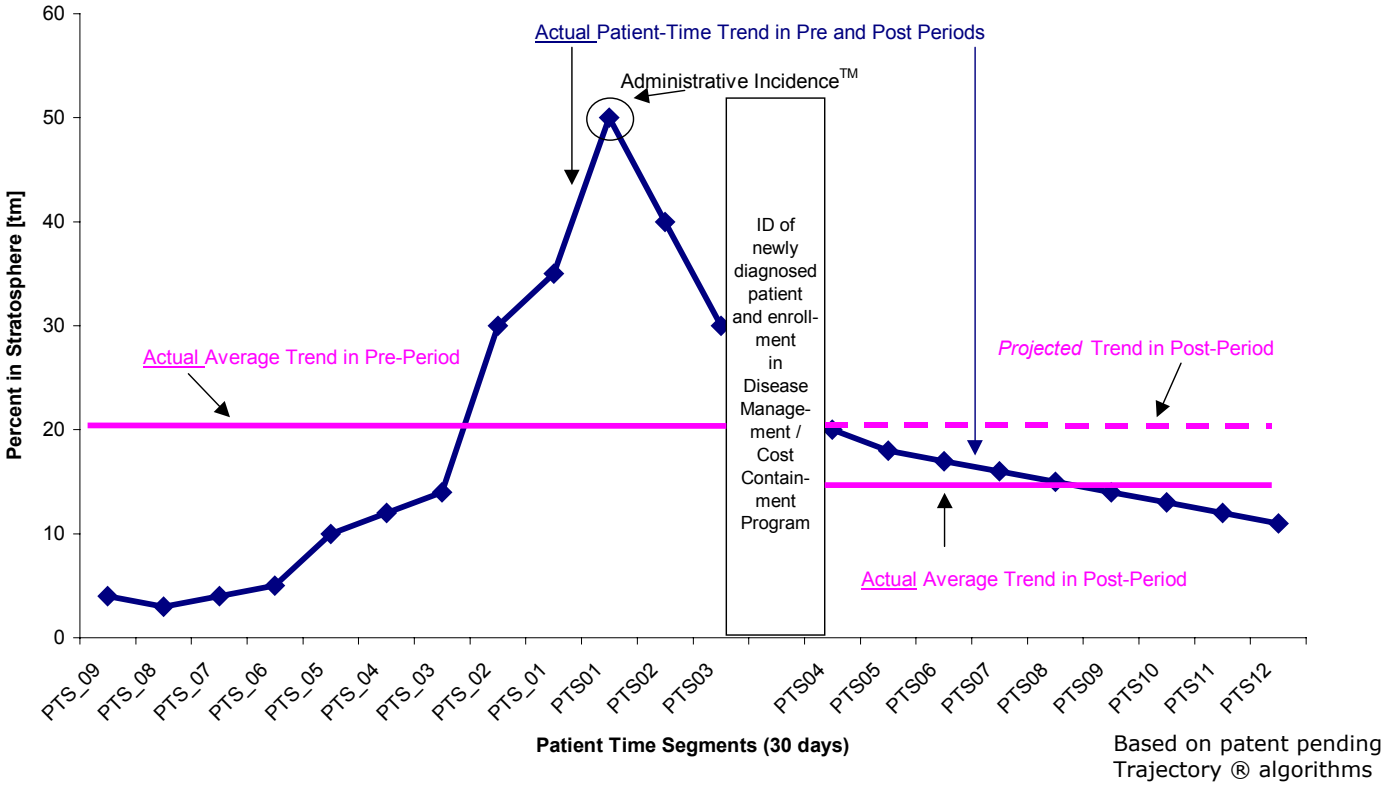


Spurious Progression: ***Measured at low end of cycle in pre period and high end in post period***

Spurious Regression: ***Measured at high end of cycle in pre period and low end in post period.***

Time Series (multiple pre-post time segments)

Equivalence Assumption is problematic



Based on patent pending Trajectory ® algorithms

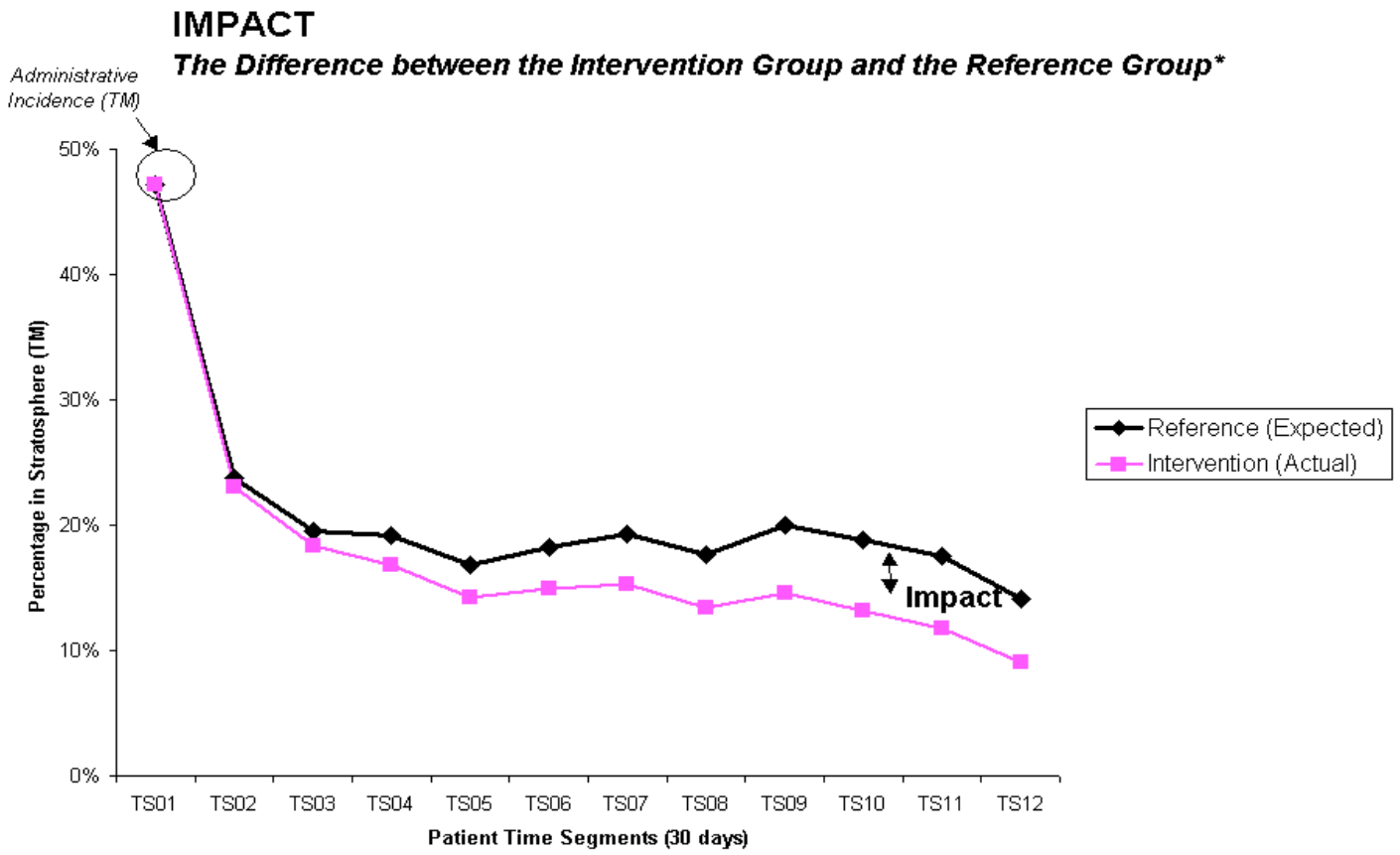
Prior Slide Legend

- The prior slide shows the percent of the defined population that are high cost ("The Tipping Point" or the "Stratospheres" ™) in 30 day patient time segments.
- The highest point is the "administrative incidence"™ -- this is the point at which each individual is initially identified with the condition (in this case, it is the first time the diagnosis for CHF appears in a calendar year).
  - Administrative incidence refers to two kinds of people:
    - 1) True incident cases (from the perspective of health)
    - 2) Unknown if case is incident or prevalent
      - This can be parsed out if we allow for a clinically relevant duration of time when the patient is "disease free" (ie. enrolled, but no claims-based evidence of disease).
- Before that identification point is the *retrospective* "Patient Time ™" trend, after that point is the *prospective* "Patient Time" trend.
  - The prospective patient time trend represents a true prevalent cases.

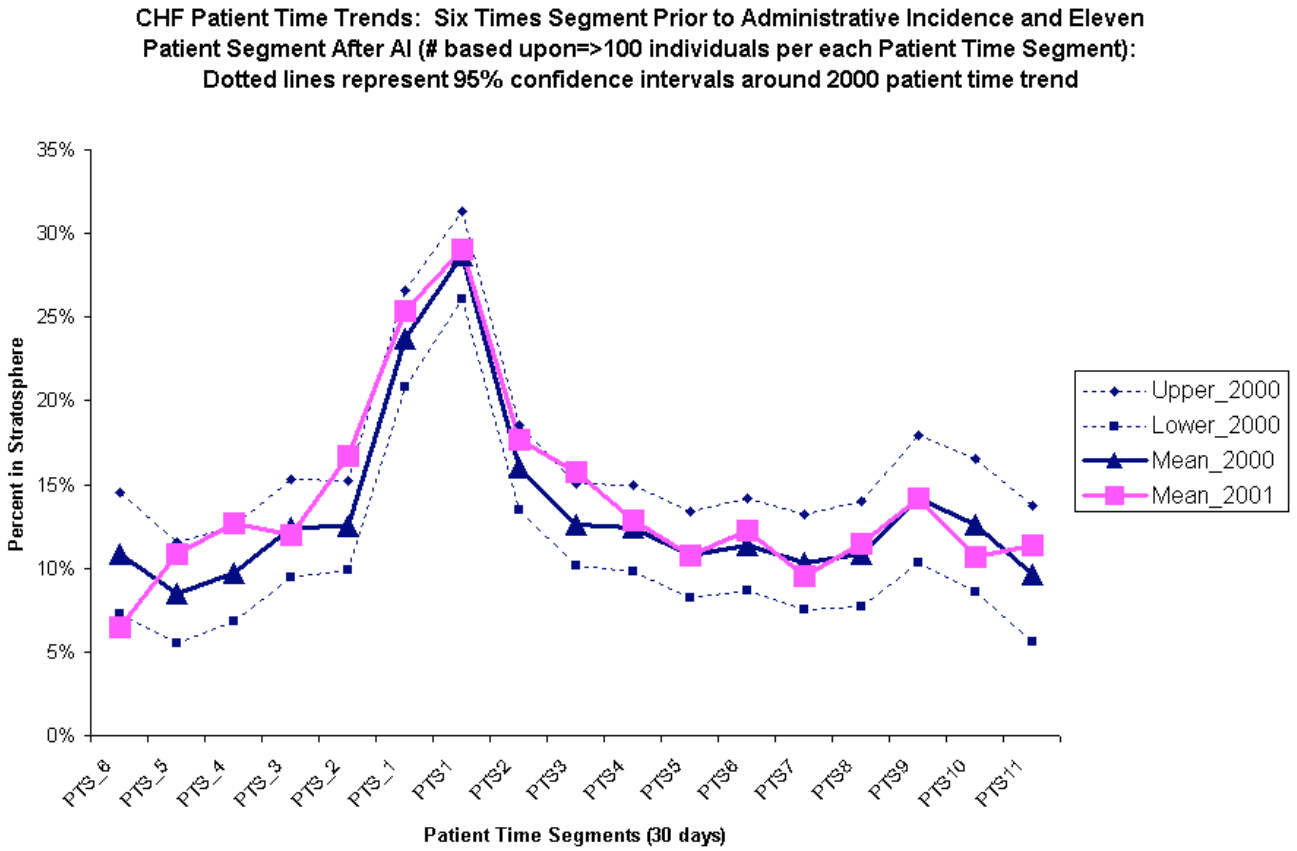
IV) Follow-up: Observational & Experimental

- Observational: Where DM program uses "naturally" occurring variation in "exposure" and observes of "outcomes" prospectively.
- Experimental: Randomization by Group (Place or Time) or Individual
- *What to watch out for: Selection bias. Was the reference group equivalent to the intervention (exposed group) at the beginning and throughout the study? No issues with temporal ambiguity or ecological fallacy.*

Follow-Up Design:
But are the two groups equivalent?



Predictability



Prior Slide Legend

- The prior slide shows the percent of the defined population that are high cost ("The Tipping Point" or the "Stratospheres" ™) in 30 day patient time segments in 2000 and 2001.
- This was done in a managed care population with no DM in either year.
- The 2000 Patient Time™ pattern is used as "hypothetical" predictor of the 2001 Patient Time trend.
- The results show that the prior trend was a good predictor of the the post trend

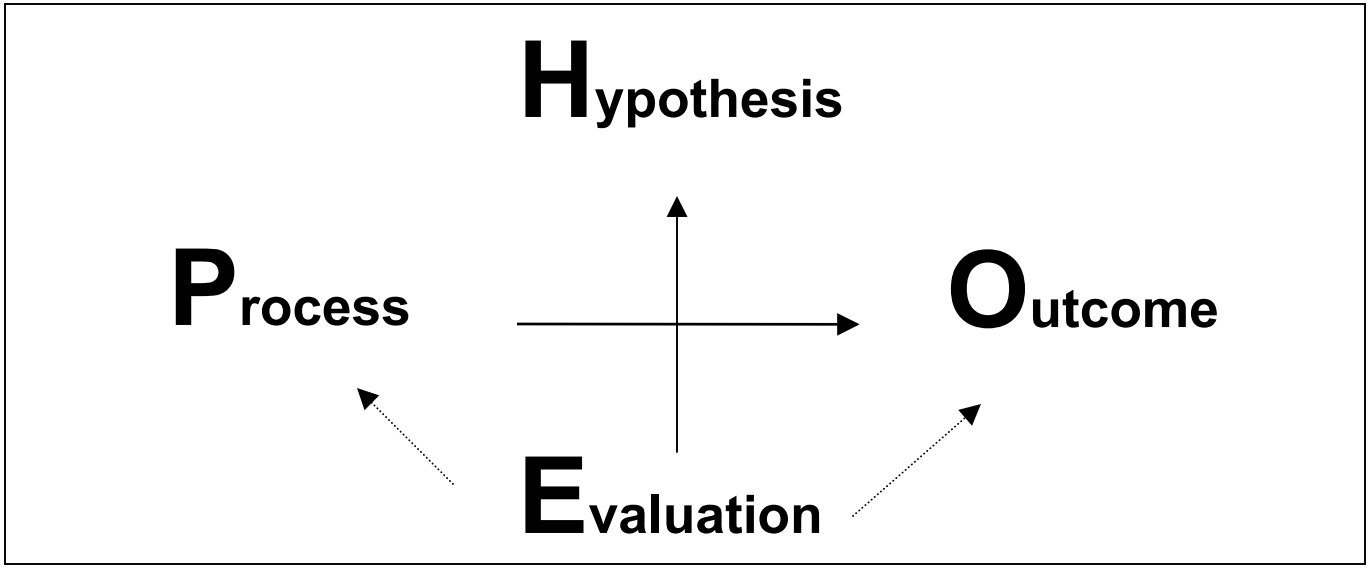
III) HYPOTHESIS-DRIVEN ACTION

What to Do Where Evidence-Based Medicine Doesn't Lead?

- "Estimates of the fraction of physician's care decisions that are supported by unambiguous clinical trial evidence ranges from **11 percent to 65 percent** depending on specialty and care setting.
  - A strong case can be made that these estimates are upper bound, since the studies focus on major decisions only and not the full range of care decisions—such as whether to hospitalize a patient or consult with another specialist – that are made in any complex treatment regimen."
  - *Source: Ginsburg PB, Nichols LM. The Health Care Cost-Coverage Conundrum: The Care We Want vs. The Care We Can Afford. Annual Essay 2002-3. Center for Studying Health System Change. (www.hschange.org CONTENT/616/?words=cost-coverage)*

There's H.O.P.E.™

™

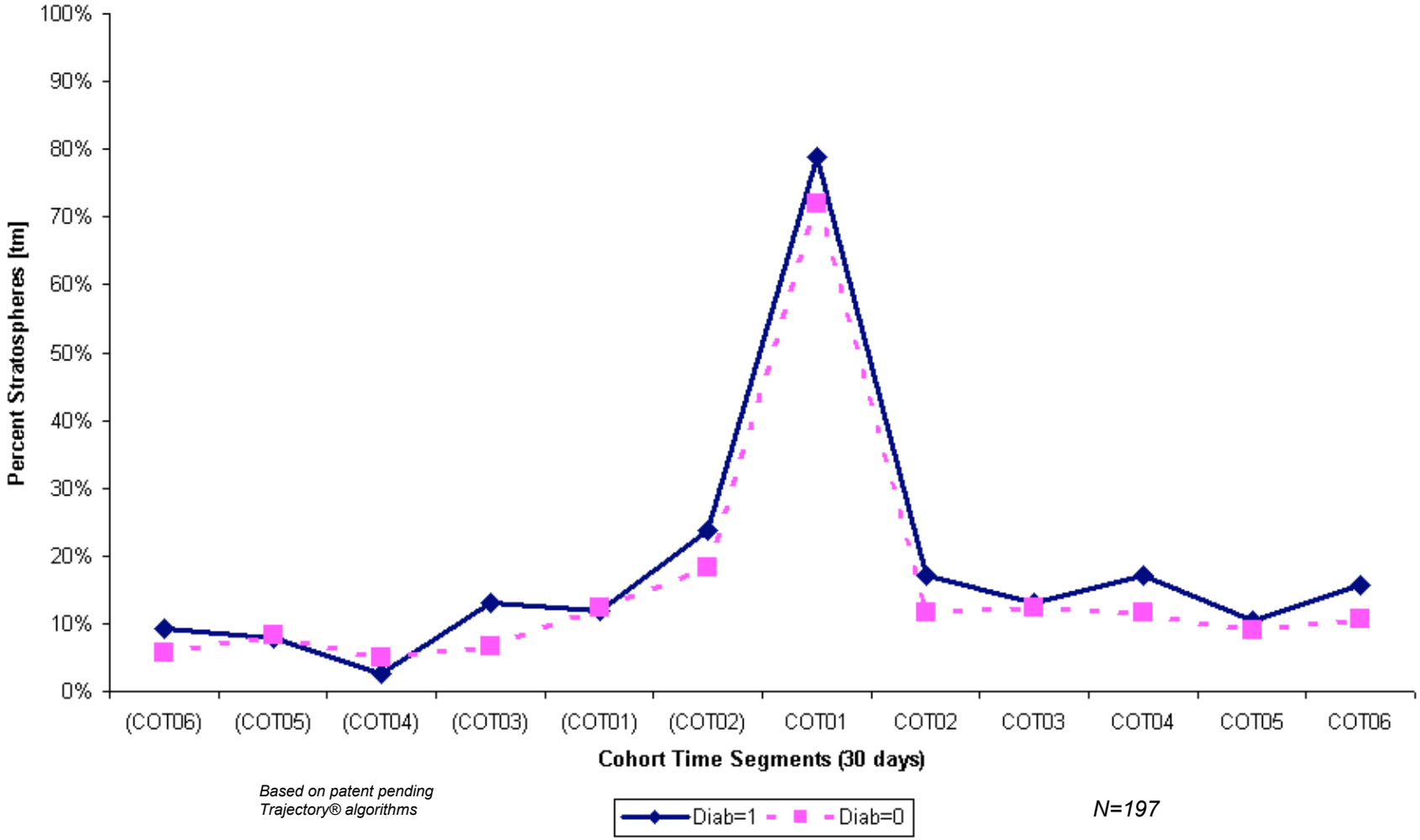


Cholera & Dr. John Snow

- **Setting: Cholera Epidemic - 19th Century London**
  - *"Evidence-based medicine" was not a guide*
  - Agent (did not know what caused cholera)
  - Host (no anti-cholera drug)
  - Environment (maybe …)
- **Study**
  - 1853, Dr. John Snow
  - Empirical-based medicine: Observational
- **Findings**
  - Numerous cases associated with the "Broad Street Pump"
- **Action**
  - Locking the pump
- **Implications**
  - Incomplete information. Yet, Intelligent & Effective Action

Current Example (real data)

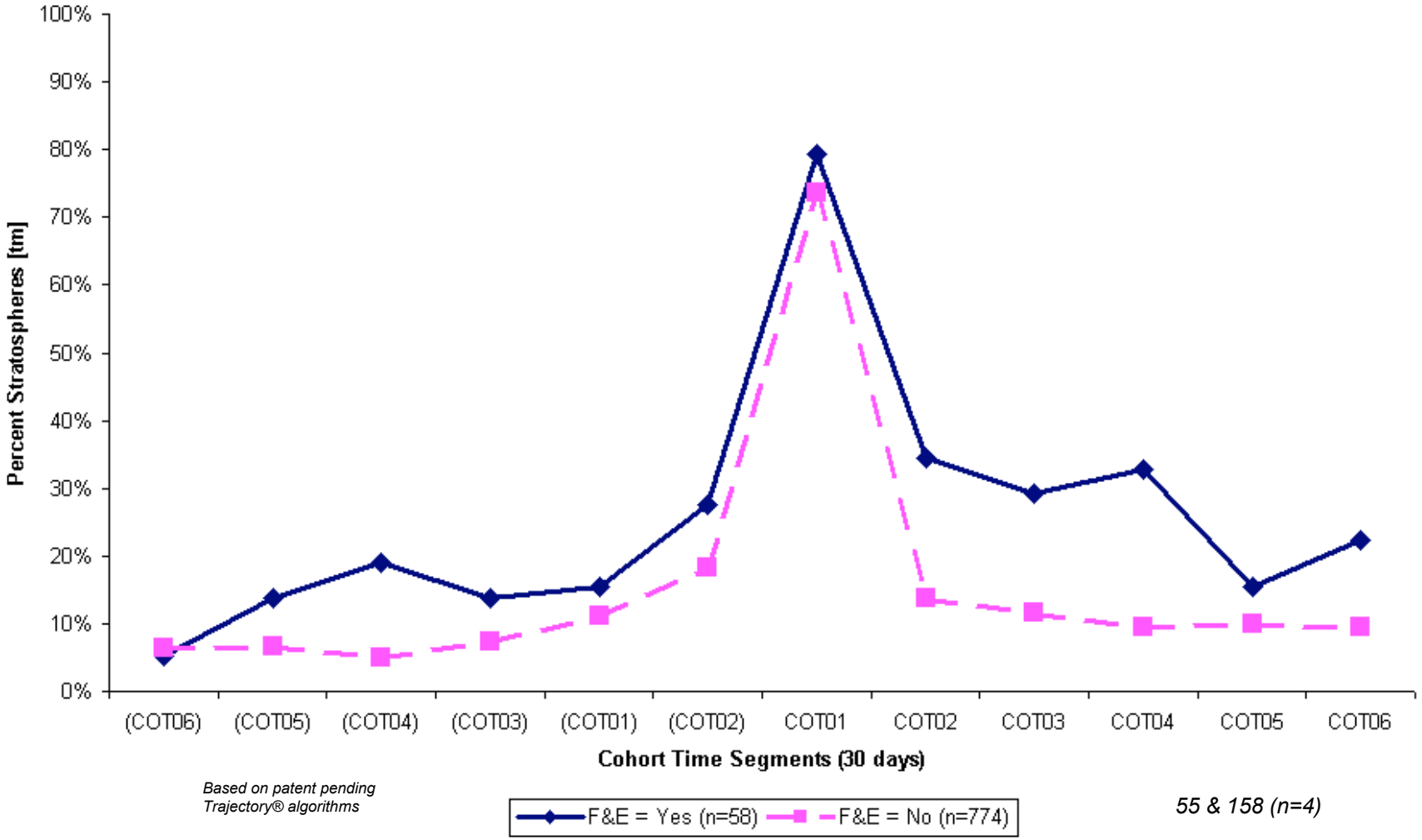
Diabetes Ascertained by Either Primary Dx on Claim (n=20) or Self-Reported (n=70) [both were n=14]



Based on patent pending Trajectory® algorithms

N=197

Cost Trends over Cohort Time: Stratification based on Fluid and Electrolyte Disorders Category (CCS #55)



Based on patent pending Trajectory® algorithms

55 & 158 (n=4)

Wrap-up

- Pragmatic Epidemiology and Principles of Causality
- Approaches for dealing with non-equivalence
- Approaches for efficient actions to reduce costs:
  - How do we "lock the pump" today?
  - How do we assess the effectiveness of our action? Can we really expect "certainty" in observational studies?
  - How do we improve to meet the new challenges facing DM?

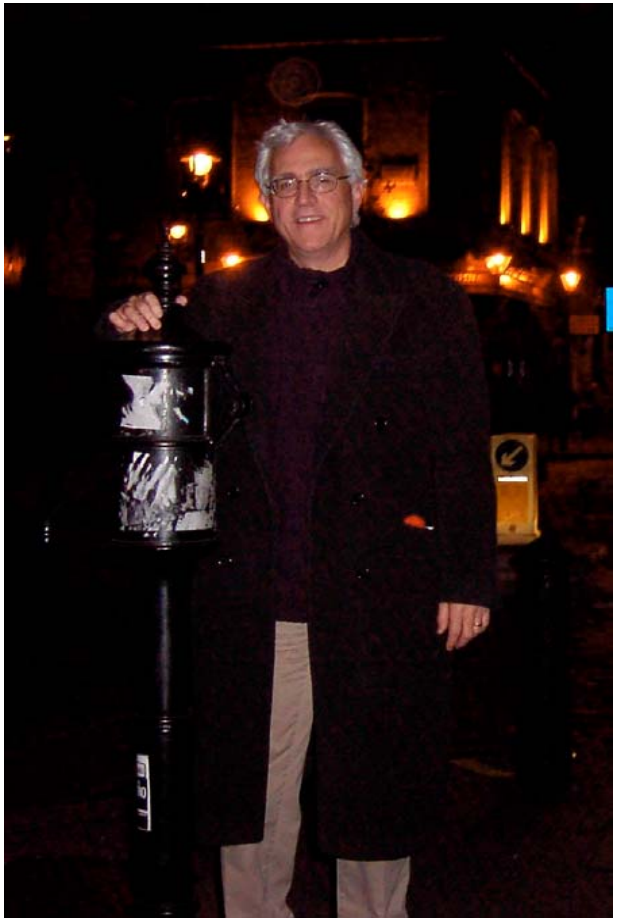
References

Wilson T, MacDowell M. Framework for Assessing Causality in Disease Management Programs: Principles. Disease Management, Fall 2003.

Wilson T. Evaluating ROI in State DM Programs. Robert Woods Johnson's State Coverage Initiative Program, 2003. *www.statecoverage.net.*

NMHCC Workgroup: Gruen J, Nash D, Wilson T, Lewis A, Fetterolf D, Thar W, Popiel R, Patel M. "Crossing the Measurement Chasm: Evaluating Disease Management Measurement Methodologies." Presented at the 8th annual Disease Management Congress, San Diego, California, September 30, 2003 (www.jacksonmediagroup.com/DMC) [in press: Joint Commission Journal of Quality and Safety)

The Pump



The John Snow Pub

Photo Credit: David Allison, Falling Leaves Press, 2002