Outcomes Measurement & Evaluation in Population Health using the ACG System

Dr. Stephen Sutch
Associate, Johns Hopkins University
ssutch1@jhu.edu
Dr. James Barrett, Paul Molyneux
Johns Hopkins Healthcare Solutions

JOHNS HOPKINS M E D I C I N E

POPULATION HEALTH ANALYTICS

Wednesday, 19 June 2024

RISK ADJUSTMENT

- Controlling for factors that may affect outcomes
- Case-mix adjustment
  - e.g. ACG actuarial cells, Age, Sex
- Segmentation and stratification: compare outcomes within defined strata
  - Stratified Sampling
  - Patient Need Groups (11 PNGs), Modifiers, e.g. age ≥ 65
- Propensity score matching (PSM) - quasi-experimental method
  - Observational study, Mimics randomization, Creates matched-pair controls

OUTCOMES IN QUALITY ASSESSMENT

- **Outcomes do not directly assess quality of performance.** They only permit an inference about the quality of the process
- The degree of confidence in that inference depends on the strength of the predetermined causal relationship between process and outcome.

Data Needs

- Because the relationship between process and outcomes is a probability, it is necessary to collect an appropriately large number of cases before one can infer if care is better or worse or meets specified standards.

Time Window

- Outcome measurement requires specification of the appropriate time window which is the time when outcome differences caused by degrees of quality in health care are most manifest.

## OUTCOMES FROM DIFFERENT PERSPECTIVES

- Clinical Perspective
- Patient Perspective
  - Subjective health status
  - Quality of life
  - Satisfaction
- Societal Perspective
  - Utilization
  - Cost

- Measures: Structure – Process – Outcome
(Donabedian, A, 1988)

- Outcome indicators of quality are more comprehensible to patients and the public than indicators of the process of technical care.
- However, they can cause misunderstanding by the public if the problem of multiple causation is not understood.

Other Considerations
- Availability
- Completeness
- Accuracy
- Susceptibility to manipulation
- Information about delayed outcomes

Data collection timeline
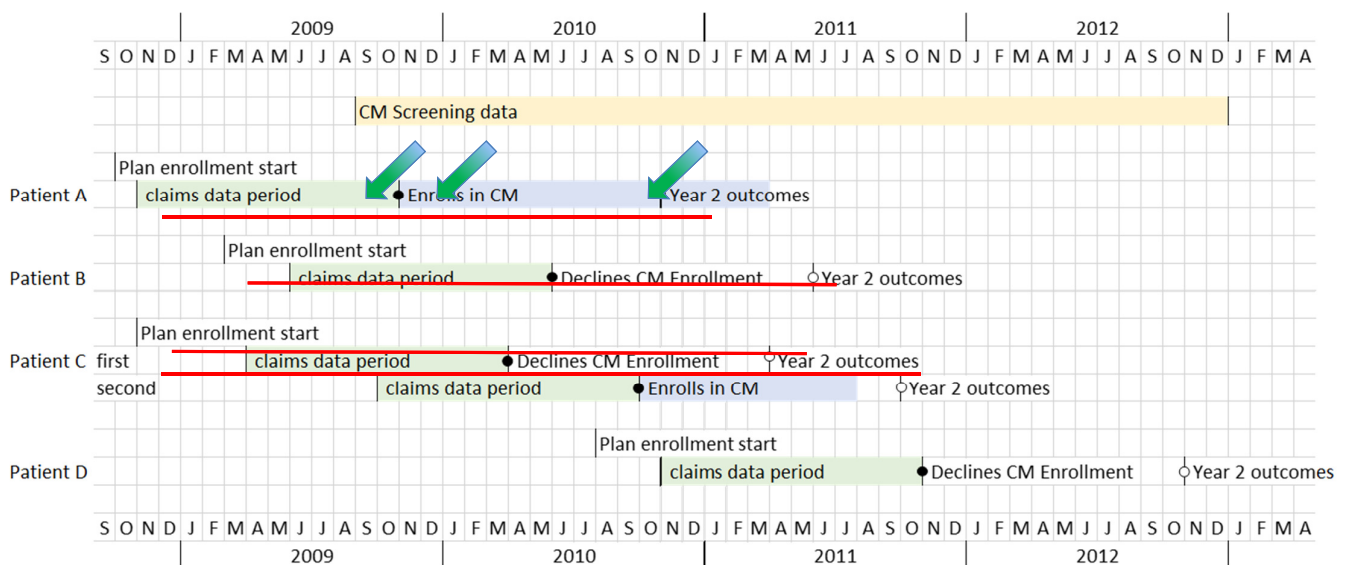
- Want to know the participants' outcome with and without treatment
- Participants differ from non-participants
- Objective: find a large group of individuals who match the participants in all relevant pre-treatment characteristics
- Therefore difference (if well selected) can be attributed to the program
- With multiple characteristics to control for, suggested use of propensity score – e.g. Probability of participation in the program given the pretreatment characteristics

PROPENSITY SCORE MATCHING (PSM)
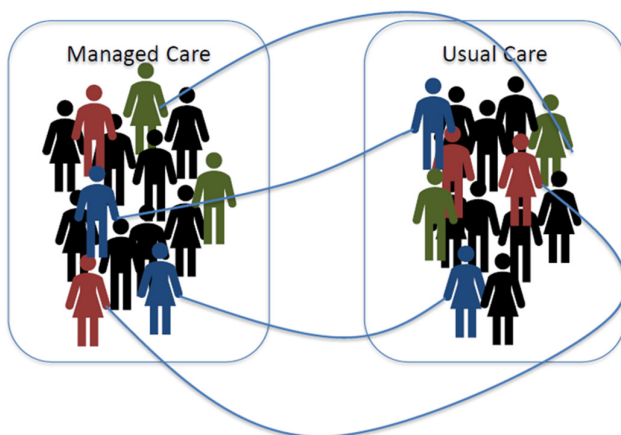
- Score each patient, data prior to enrolment
- Managed Care to Usual Care matching ("counterfactual")
- Nearest Score
- Can be paired or multiple e.g. 1-3, 1-4
- Follow-up and measure outcomes e.g. 6mth, 12mth, 24mth
- Compare results

JOHNS HOPKINS ACG MODELS
(WHAT CAN WE USE AS A PSM?)

- **Concurrent risk**
  - Age-gender, Local/reference ACG concurrent
  - Concurrent risk (regression-based)
- **Predictive cost risk**
  - Predicted cost
  - Rank/Reference probability
  - Persistent high user
  - High risk unexpected pharmacy cost
- **Hospitalization risk**
  - Inpatient admission
  - Readmission, ICU, Extended stay
- **Others**
  - Mortality Risk, Emergency Admission (UK)
  - Hotspotters (NL)

Cost and Utilization Outcome, Year 2

- Propensity Score matching (PSM) applied to reduce the risk of selection bias
  - Estimation of the Propensity Score
  - Matching algorithm, treatment / non-treatment pairs
  - Check matching quality and treatment effects
- Two matching methods used
  - **Nearest Neighbor**, each treatment case the control is chosen that had the closest propensity score (probability of enrollment in case management). Ensures a control match for each treatment case, but does not ensure the absolute difference in scores are close
  - **Caliper** method allows for a minimum absolute difference in prevalence to be specified, but does result in data loss where a match cannot be found within the specified absolute difference
- (Coca-Perraillon, 2006)

Propensity Score matching - Year 2 Outcome Measures

Medicaid Health Plan	PSM - Near Neighbor (n=4662)				P value*
	Not-Enrolled (n=2331)		Enrolled (n=2331)		
	mean	CI	mean	CI	
Total Cost $	30,718	(28906-32531)	26,644	(24809-28479)	0.002
Inpatient hospitalization	1.0854	(1.003-1.167)	0.828	(0.751-0.905)	0.000
Emergency Visits	3.2986	(3.071-3.526)	2.6319	(2.461-2.803)	0.000

- Employee plan mean cost reductions (Near Neighbor, Caliper)
  - -$4486.86 (18.3%), -$4186.91 (17.1%)
- Medicaid plan
  - -$4074.07 (13.3%), -$1372.66 (4.6%)
- Family Health plan
  - -$2458.51 (7.3%), -$2604.29 (7.7%)

Propensity Score matching
Year 2 Total Cost by sub-group

Plan/Sub-Group	N	Total Cost (CI)	N	Total Cost (CI)	P value*
		Not-Enrolled		Enrolled	
		PSM - Near Neighbor (n=4662)			
Medicaid Health Plan					
<2 Chronic Conditions	342	14,939 (11901-17978)	559	13,305 (11353-15257)	0.374
0-2 Major ADGs	751	24,708 (21778-27638)	828	20,688 (18286-23091)	0.037
3-4 Major ADGs	821	34,632 (31606-37658)	674	33,614 (29927-37302)	0.675
5+ Major ADGs	417	46,779 (41448-52110)	270	55,127 (46356-63899)	0.110

- Medicaid plan mean cost reductions (Near Neighbor, Caliper)
  - <2 Chronic cond. -$1634.77 (10.9%), -$1480.85 (9.7%)
  - 2+ Chronic cond. -$4019.44 (16.3%), -$3468.01 (14.1%)
  - 3-4 Major ADGs   -$1017.33 (2.9%), -$1274.02 (13.7%)
  - 5+ Major ADGs   +$8348.21 (17.8%), +$11,288.15 (25.7%)

- Establish measures and data collection from the outset, not retrospectively
- Decide on randomised study, or casemix adjust population cohorts
- Is there an obvious comparison population (Intervention v Control)
- Matched pairs create a population similar to those in managed care ("Intervention group")
- Creation of a risk score or probability, assigned pre-enrolment

- Consider the time frame (time window), is it absolute (same months), or did individuals/groups join at different times
- Follow up measurement at specific time periods
- Lost to study – how do you measure/adjust when individuals leave the study or intervention
- Compare outcome measures of different groups
- Create strata of sub-groups to better understand impact e.g. PNGs

- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.

- Coca-Perraillon, M. (2007). Local and global optimal propensity score matching. *SAS Global Forum 2007: Statistics and Data Analysis,* Paper 185-2007.

- Kleinman, K. (2010). Examples of tasks replicated in SAS and R: Example 7.35: Propensity score matching.

- …

- Wed 14:00-5:30 Room 3 - Outcomes Measurement & Evaluation in Population Health using the ACG System

- Thu 15:30–16:30 Room1 - The Added Value of Using Primary Care Data in Population Health Management

- Fri 09:30–11:00 Room 1 - Examining person level social determinants of health recorded in routinely collected healthcare data: insights into effects on healthcare utilisation

PCSI 2024
Bled, Slovenia
28–31 May, 2024

JOHNS HOPKINS MEDICINE

Thank You

Questions?

ssutch1@jhu.edu

Additional Analyses

- Distribution of utilization in 5ADG groups, are outliers leading to higher costs?
    - +18,463.00 / +11,288.15 / +1,451.33
    - +4,645.50 / - 913.17 / +7,720.49 Simple IQR trim
- Chronic condition count = 0, who are these patients? Referral only?,EDCs, ADGs
    - Year 2 CC+Count 1.45 / 1.85 / 3.15
- Check on underlying assumptions on PSM, re unmeasured effects
- TRIPOD: Set of recommendations for reporting the results of multivariate predictive model development and/or validation.
- CART analysis