



Evaluation 101 for Researchers

Kate Kissick and Charlene Zil

NADCP Conference

June 2019



EVALUATION 101 FOR RESEARCHERS

What to Collect and How to Measure it



OVERVIEW

What type of evaluation am I doing?

What do I need?

Who's included?

- Program Samples
- Comparison Groups

Questions



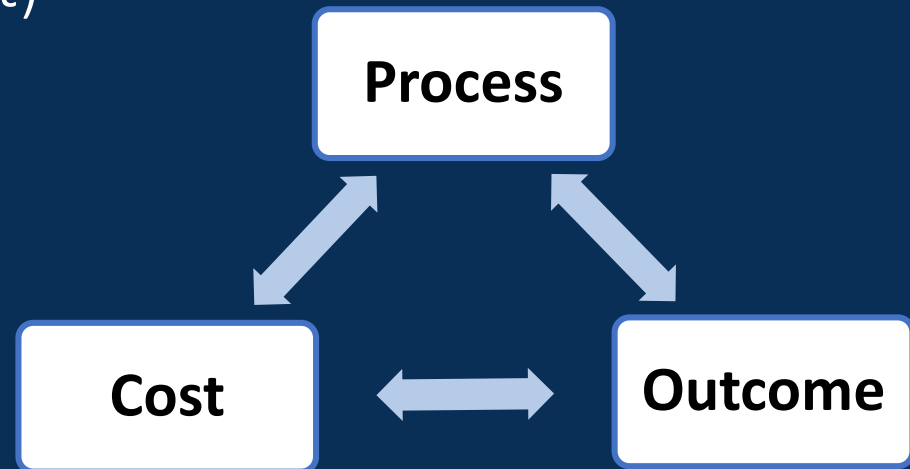
WHAT TYPE OF EVALUATION AM I DOING?



WHAT'S EVALUATION?

Three main areas of evaluation:

- Process (program improvement)
- Outcome (within program and outside program)
- Cost (cost-benefit)



COMMON RESEARCH QUESTIONS

Process

- Is the program maintaining model fidelity?

Outcomes

- Are participants successfully completing the program?
- Do programs reduce recidivism?

Cost

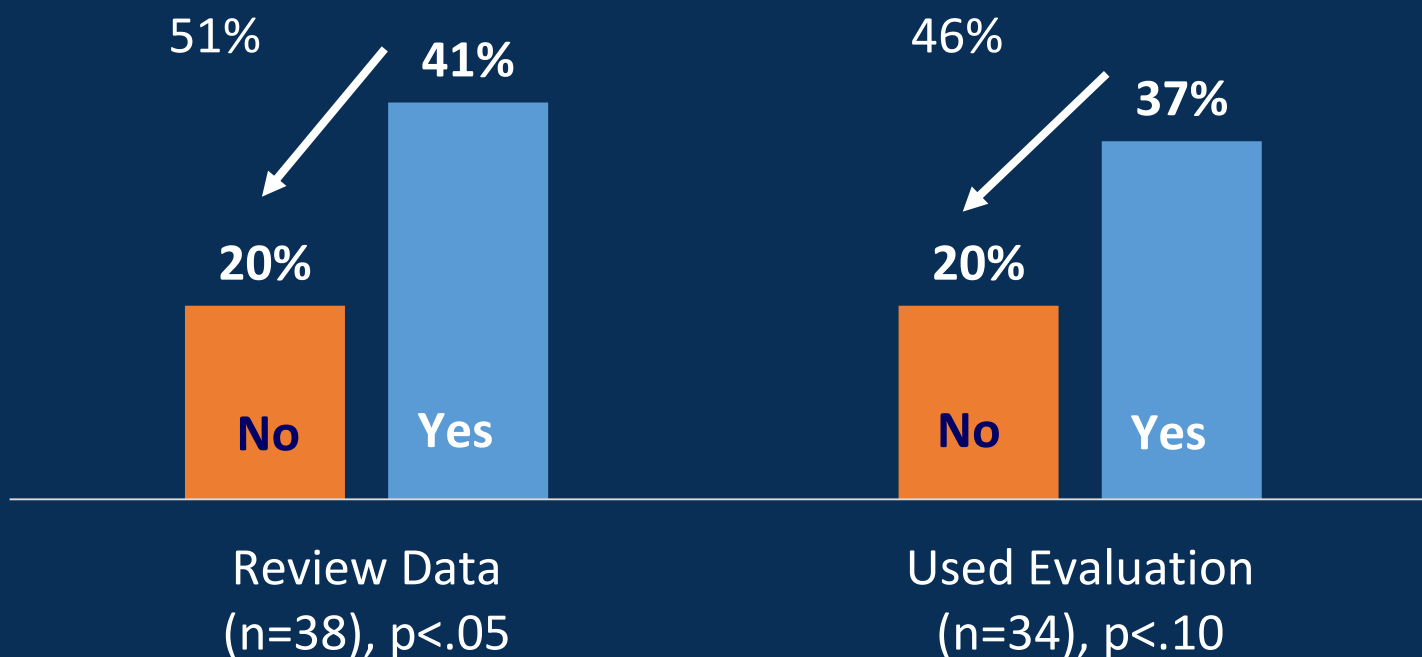
- What are the costs and savings of the program?

TYPES OF OUTCOME EVALUATIONS

Characteristics	Within Program	Outside Program
Performance Indicators	Services Received Grad Rate Time to Completion Factors Leading to Grad	Rearrests Reincarceration Subsequent Tx Social Services & Health Care Utilization
Comparison Group?	Not Needed	Needed!
Methods	Internal Monitoring or External Evaluation	External Evaluation
Purpose	Improve Program	Demonstrate Program Effectiveness

WHY EVALUATION MATTERS

Possible Reductions in Recidivism



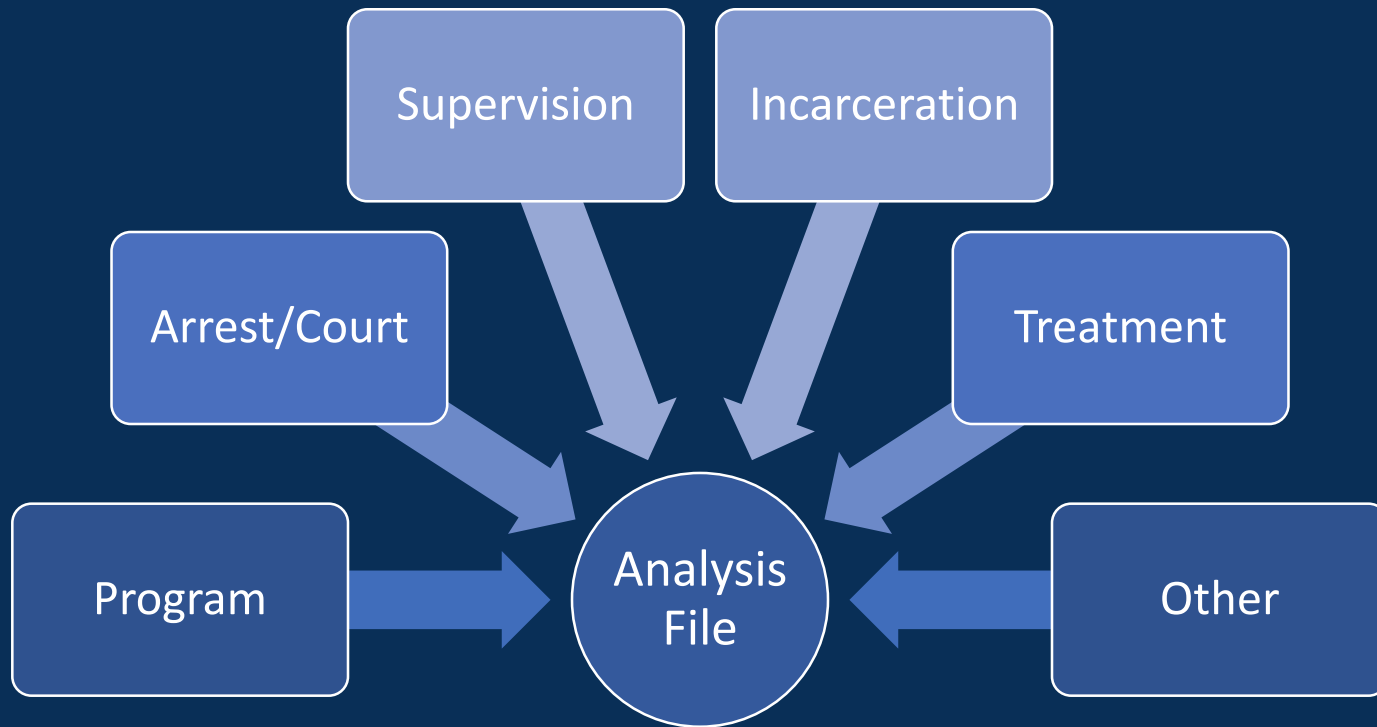
Carey, Finigan, & Pukstas (2008);
Carey, Mackin, & Finigan (2012)

WHAT DO I NEED?

All the data!!



ADMINISTRATIVE DATA SOURCES



EVALUATION MAP

Data Type	Within Program	Outside Program
Program Service Data	✓	✗
Arrests and/or Court Records	?	✓
Incarceration: prison & jail	✗	✓
Supervision: parole & probation	✗	✓
Treatment	✓	✓
Other	✓	?

PROGRAM DATA

File Structure

- Multiple, linked files, with flat or stacked data, depending on content (one or multiple rows per person)

Necessary Info

- Demographics
- Important dates (referral, entry, exit)
- Program completion/discharge status
- Dates of services and activities

IMPORTANT INFO IN PROGRAM DATA

- ✓ Name
- ✓ DOB
- ✓ Gender
- ✓ Race/ethnicity
- ✓ Program entry/
exit dates
- ✓ Program status
- ♥ Other unique IDs:
SSN, SID, etc.
- ♥ Referral information
(arrest date, court
case number)
- ♥ Risk/needs scores
- ♥ Substances used

PROGRAM SERVICE DATA

Incentives, sanctions, treatment responses

- Dates
- Type/Description
- Duration, if applicable

Status review hearings

- Dates
- Attendance

Drug tests

- Dates
- Type: urine analysis, EtG, PBT, oral swab, etc.
- Results: negative, positive, no-show, dilute, etc.

ARREST/COURT RECORDS

File Structure

- Stacked on both dates of arrest and charge (multiple rows per arrest, per person)

Necessary Info

- Dates of arrests (or case filings)
- Charge descriptions

Additional Available Court Info

- Court dates
- Case dispositions
- Sentences



SUPERVISION

File Structure

- Stacked on probation or parole episodes (multiple rows per person)

Necessary Info

- Start dates
- End dates (actual or projected)
- Supervision type
- Agency or location

INCARCERATION

File Structure

- Stacked on jail or prison episodes (multiple rows per person)

Necessary Info

- Start dates
- End dates (actual or projected)
- Facility or location

TREATMENT

File Structure

- Stacked on treatment events/episodes (multiple rows per person)

Necessary Info

- Dates of treatment (or start & end dates)
- Duration of treatment
 - Could be days for inpatient/residential
 - Could be hours for outpatient appointments
- Type or modality
- Completion/discharge status

TREATMENT

Treatment Modalities

- Inpatient
- Detox
- Residential/transitional
- Group outpatient
- Individual outpatient
- Medication Assisted Therapy (MAT)
- Assessment, case management activities

OTHER DATA

Child welfare

- Out of home placements, foster care
- Abuse/neglect maltreatment

License records

Wrap around services

WHO'S INCLUDED?

Selecting your program and comparison group samples



PROGRAM SAMPLE: WHO'S INCLUDED

Program Sample Types

- Census (for program impact)
- Graduates (descriptive only)
- Graduates vs. non-graduates
- Cohorts
- Other sub-populations of interest

PROGRAM SAMPLES

Sample Type	What's the Focus?	Considerations
Census	Program Impact	Longevity of program, Significant process modifications
Graduates	Descriptive Only	Graduates don't exist in comparison group!
Graduates and Non-Graduates	Service Improvement	Program status definitions
Cohorts	Policy Changes	Implementation windows, judge or team turnover, capacity
Other sub- populations	Group Disparities	Group sizes

PROGRAM SAMPLES: CENSUS

Who's Included?

- Everyone*

*Who's Not?

- False starters
- Multiple entries

Considerations

- Implementation window
- Outcome window



PROGRAM SAMPLES: GRADS AND NON-GRADS

Who's Included?

- Non-actives (e.g., graduates, non-graduates)

Who's Not?

- Actives
- Transfers
- Deaths

Considerations

- Opportunity window (minimum time to complete)



COMPARISON GROUPS

Purpose

- To answer the question, “Is the program effective?”
- To learn whether the program is effective, there has to be something to compare
- Effective, compared to what?
- A comparison group is the baseline

COMPARISON GROUPS

Without a comparison group, you have to make *assumptions* about what would have happened:

Examples:

Assumption #1: Participant would have had prison sentence, so we save the time/cost of that sentence

Assumption #2: Participant would have had children in foster care or in a shelter, so we saved those foster care resources

COMPARISON GROUPS

Problems with these assumptions:

- Offender does not always get full sentence or get sentence at all
- Offender almost never actually serves the full time of sentence
- Child does not always end up in foster care
- If child is in foster care or shelter, it is very difficult to predict how long they will stay there

COMPARISON GROUPS

Another common mistake: comparing program graduates to program failures

- Both groups received the program. A comparison group needs to be those who did NOT receive the program so we can determine if the program makes a difference
- By virtue of successfully completing the program (or not), the two groups are inherently different



COMPARISON GROUPS

Exception: you can compare graduates to non-graduates when you want to know what is different between those who graduate and don't

- Example: if more men than women graduate, it may be an indication that the program needs gender-specific services
- Example: if those who had shorter time between arrest and program entry were more likely to graduate, the program may want to look at ways to decrease this time

GOOD COMPARISON GROUPS

- Should represent what would have happened if there was no program
- Ideally has individuals who are exactly like the individuals who participated in the program but were not referred and did not participate

SELECTING A COMPARISON GROUP

Goal:

Find a similar group of individuals who would have been eligible, but did not participate in treatment court, and have similar demographics and criminal histories as the program group

COMPARISON GROUPS: STUDY DESIGN

Observational

- Researcher does not assign individuals to groups
- Observes/studies what naturally occurs

Experimental

- Researcher intervenes and assigns individuals
- Random assignment

Quasi-Experimental

- Non-randomized design (e.g., matching)

Kim, K. (July 2016). Overview of Study Designs. UC Davis. Accessed from:
https://www.ucdmc.ucdavis.edu/ctsc/area/biostatistics/Documents/StudyDesigns_July2016_Kim_Final.pdf



COMPARISON GROUPS: TIME FRAME

Contemporary

- Similar political and community context
- May not have enough individuals not referred to program to comprise a comparison group

Historical

- Larger pool of individuals to draw from
- Changes in court over time
- May have different laws, judges, political climate etc. impacting outcomes

COMPARISON SAMPLES

Who's Included?

- People with similar charges and demographics

Who's Not?

- Ineligible individuals
- Voluntarily withdrawals (opt-outs)
- Non-graduates (terminated participants)

Considerations

- Outcome window
- Location (jurisdiction)



SELECTING A COMPARISON GROUP

“Comparison Pool” vs. “Matched Comparison Group”

Potential Pool:

offenders who did not participate in treatment court, and went through the business-as-usual court process; treatment court eligibility is unknown

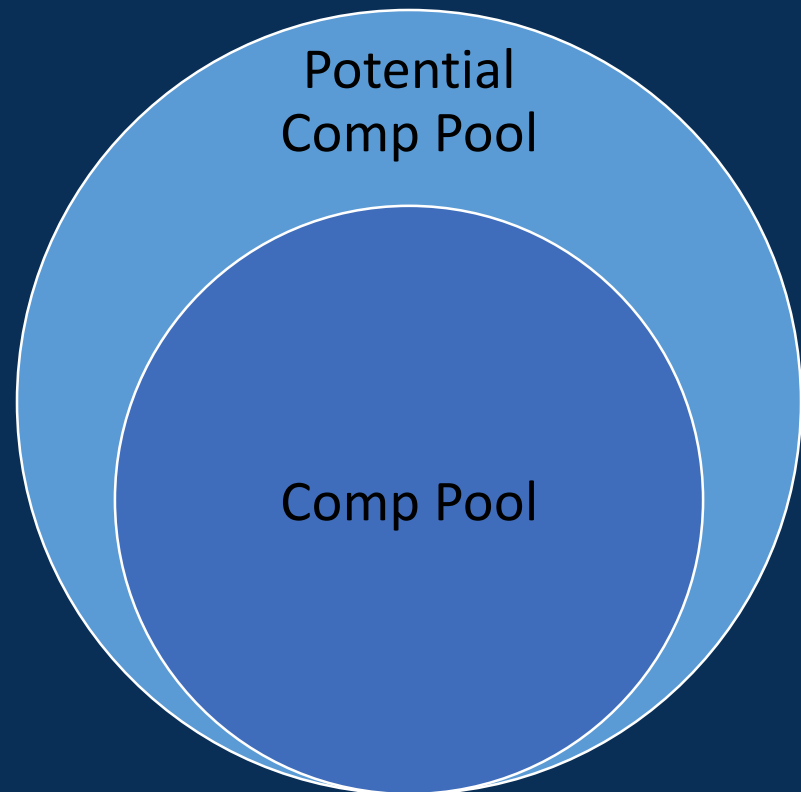


Potential
Comp Pool

SELECTING A COMPARISON GROUP

“Comparison Pool” vs. “Matched Comparison Group”

Comparison Pool:
business-as-usual
offenders who have index
events that would have
qualified them for
treatment court
(i.e., eligible!)

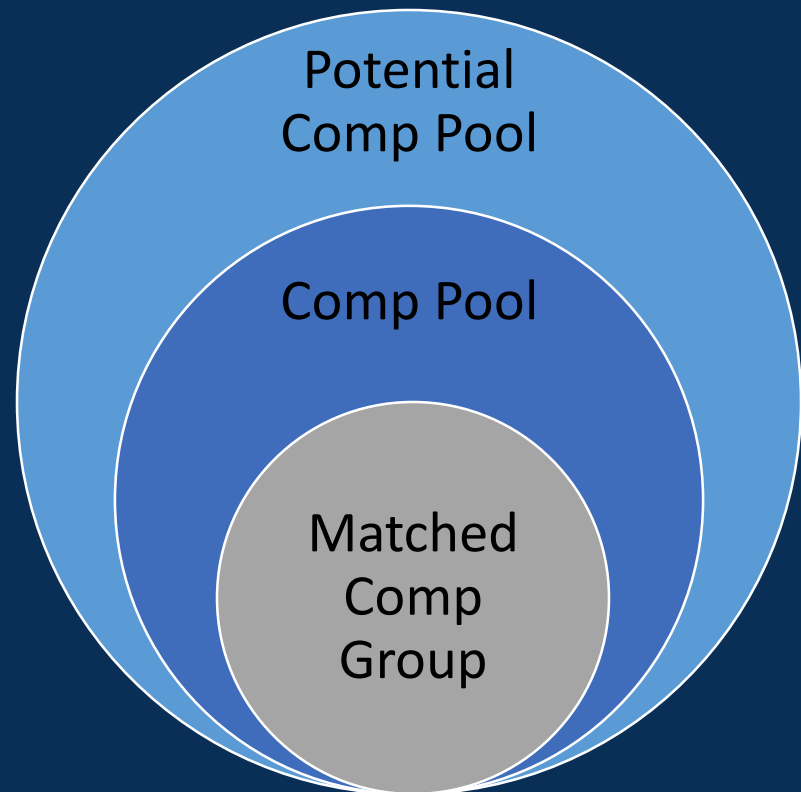


SELECTING A COMPARISON GROUP

“Comparison Pool” vs. “Matched Comparison Group”

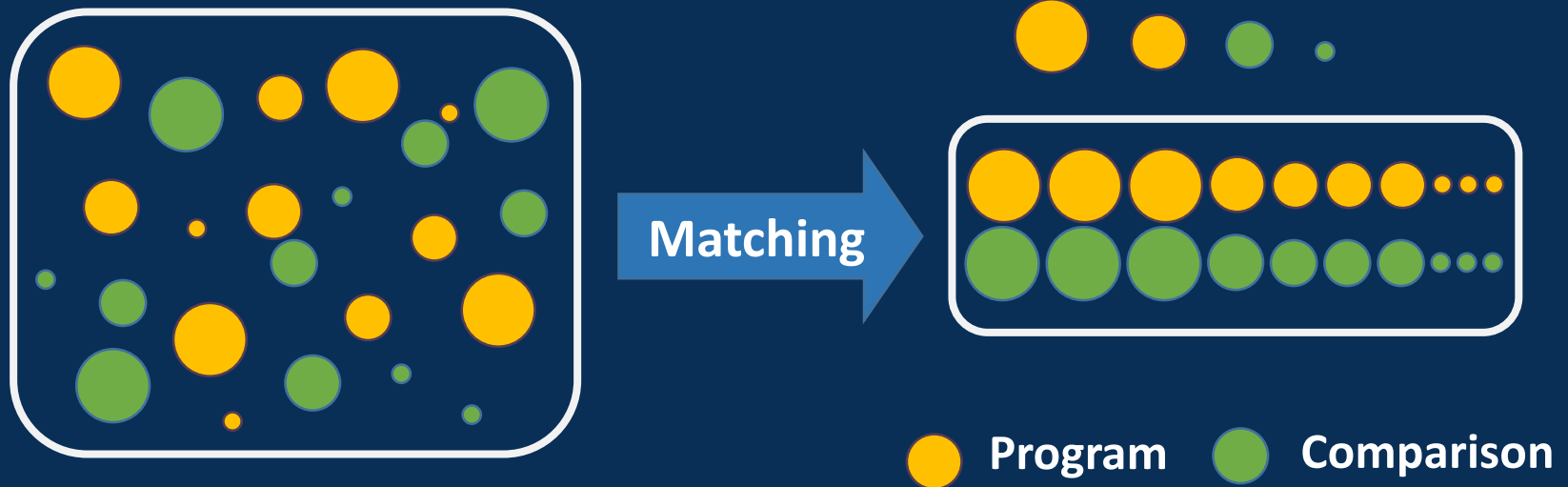
Matched Comparison Group:

eligible offenders who have been matched to the program group based on criminal history, demos, and other observable background characteristics



SELECTING A COMPARISON GROUP

“Comparison Pool” vs. “Matched Comparison Group”



SELECTING A COMPARISON GROUP

1. Start by exploring program group

- Date range of program entries and exits
- Date range of eligible index arrests
- Index event types: probation violations, new arrests, release from incarceration
- Index event charges: e.g., felony drug charge, felony property charge, etc.
- Other stated program criteria

SELECTING A COMPARISON GROUP

2. Based on the program exploration, define inclusion and exclusion criteria

Example:

- Arrest occurring between 2006 and 2015
- Had a drug or property charge
- Had a felony charge
- Located in Multnomah county

SELECTING A COMPARISON GROUP

3. Identify other criteria used in matching or group assignment, for example:

- Number of prior arrests
- Risk/need assessments
- Demographics (e.g., age, race, sex)

4. Select a matching method to identify comparison group members with similar backgrounds

QUESTIONS?



CONTACT INFORMATION

Kate Kissick

Kissick@NPCresearch.com

Charlene Zil

Zil@NPCresearch.com

