Data, Evaluation and Outcomes: *What You Should Collect and Why*

Shannon Carey, Ph.D., NPC Research
Charlene Zil, MPA, NPC Research

Overview

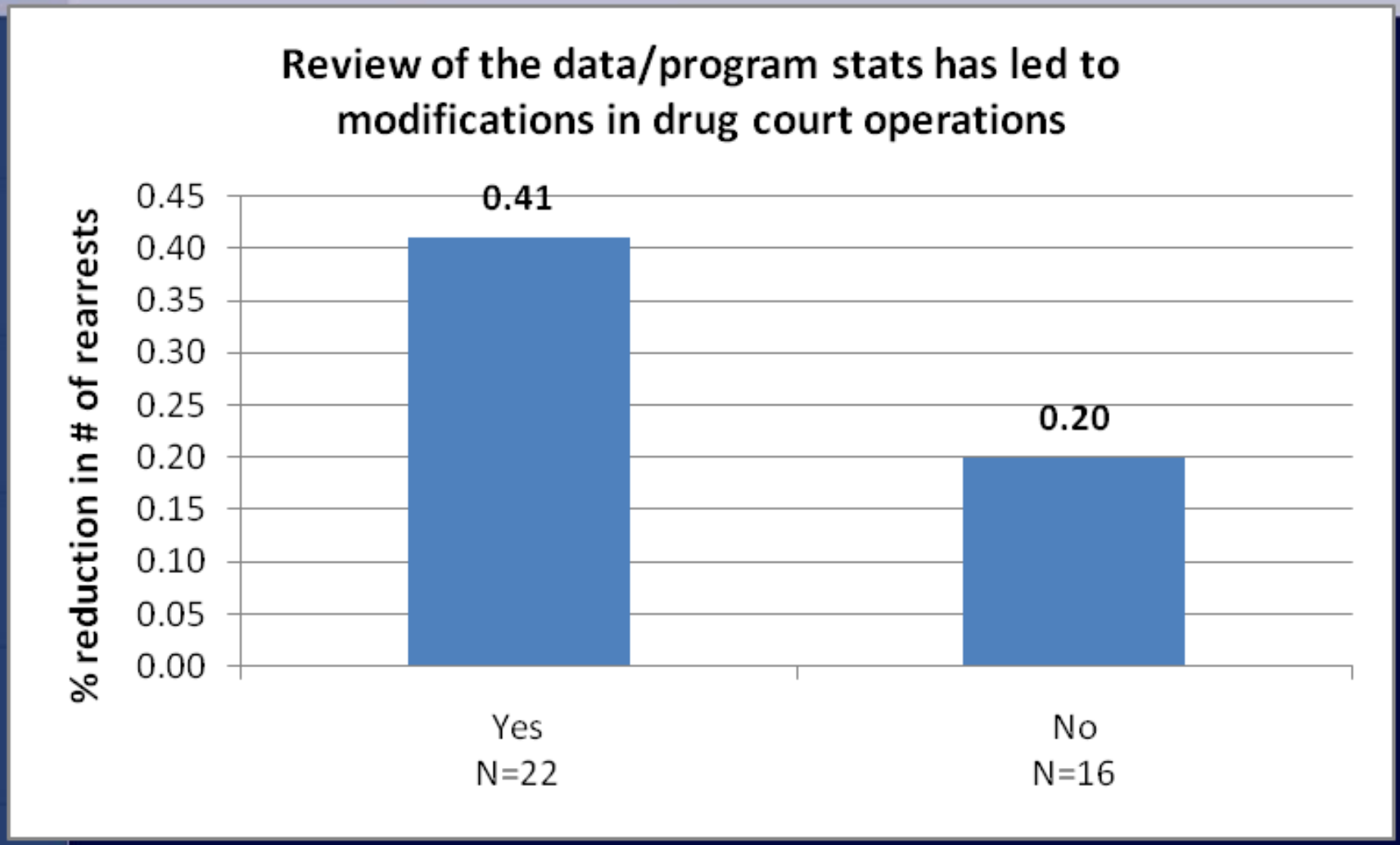
- What are the new standards in this area?

  Data and Evaluation - Thursday morning 8am

- Why do evaluation?
- What is evaluation? (Process, Outcome, Cost)
- What data should you collect?
- Why should you collect data?
- (If time) How can you use evaluation ?

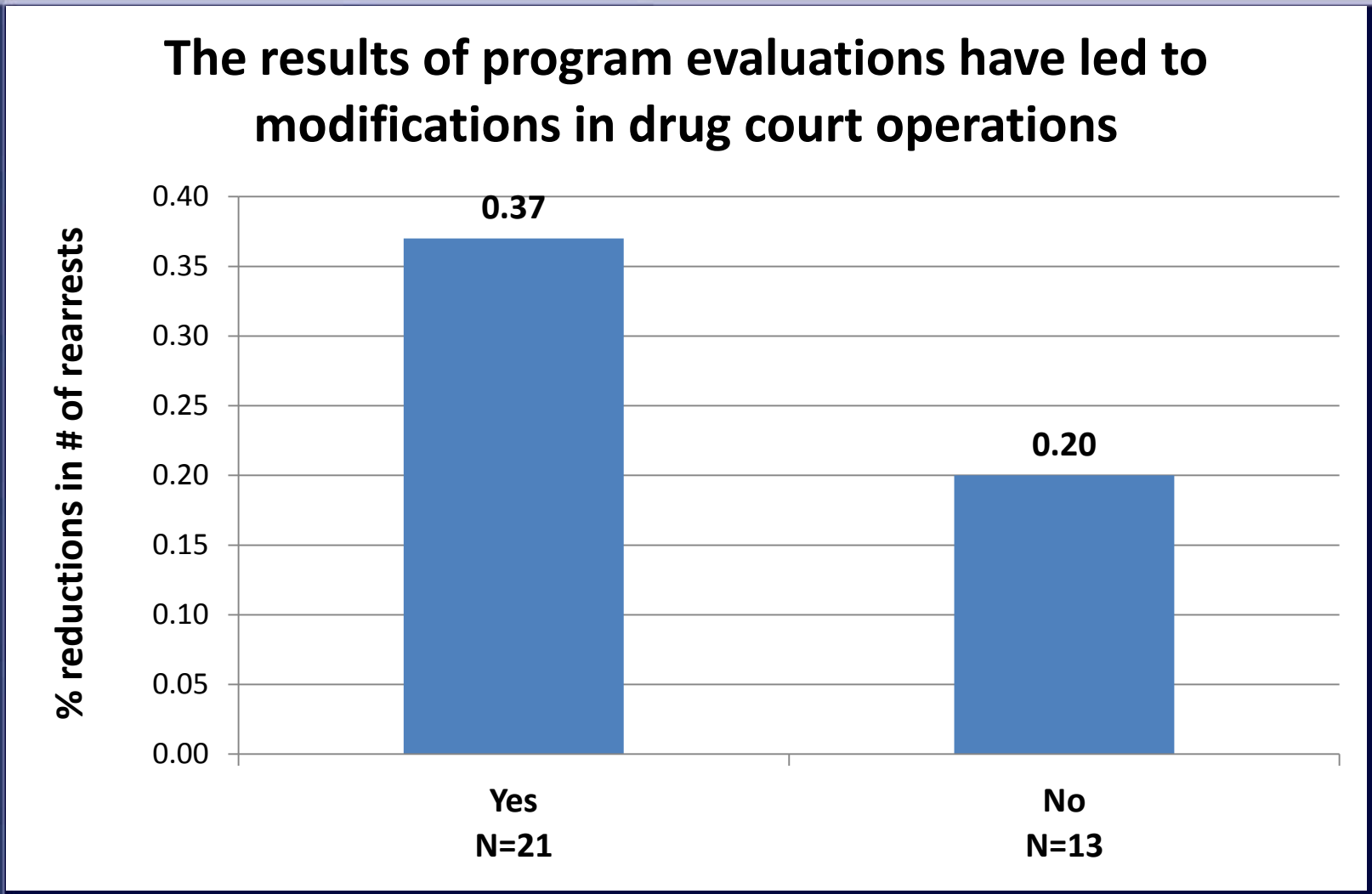
Why do Evaluation?

Drug Courts Where Review of the Data and/or Program Statistics Led to Modifications in Program Operations had 105% greater reductions in recidivism



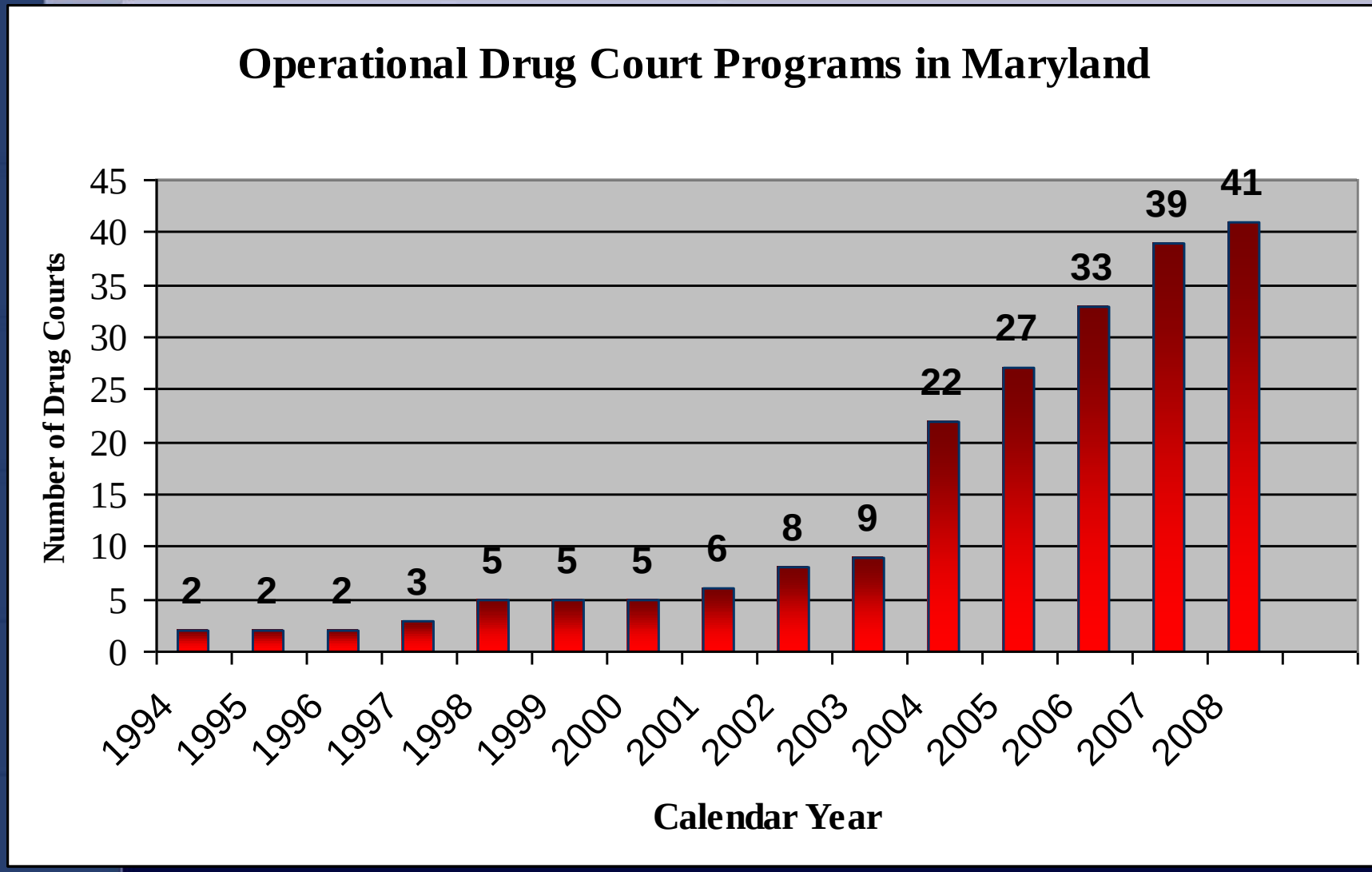
Note: Difference is significant at $p<.05$

Drug Courts that used program evaluations to make modifications in drug court operations had 85% greater reductions in recidivism

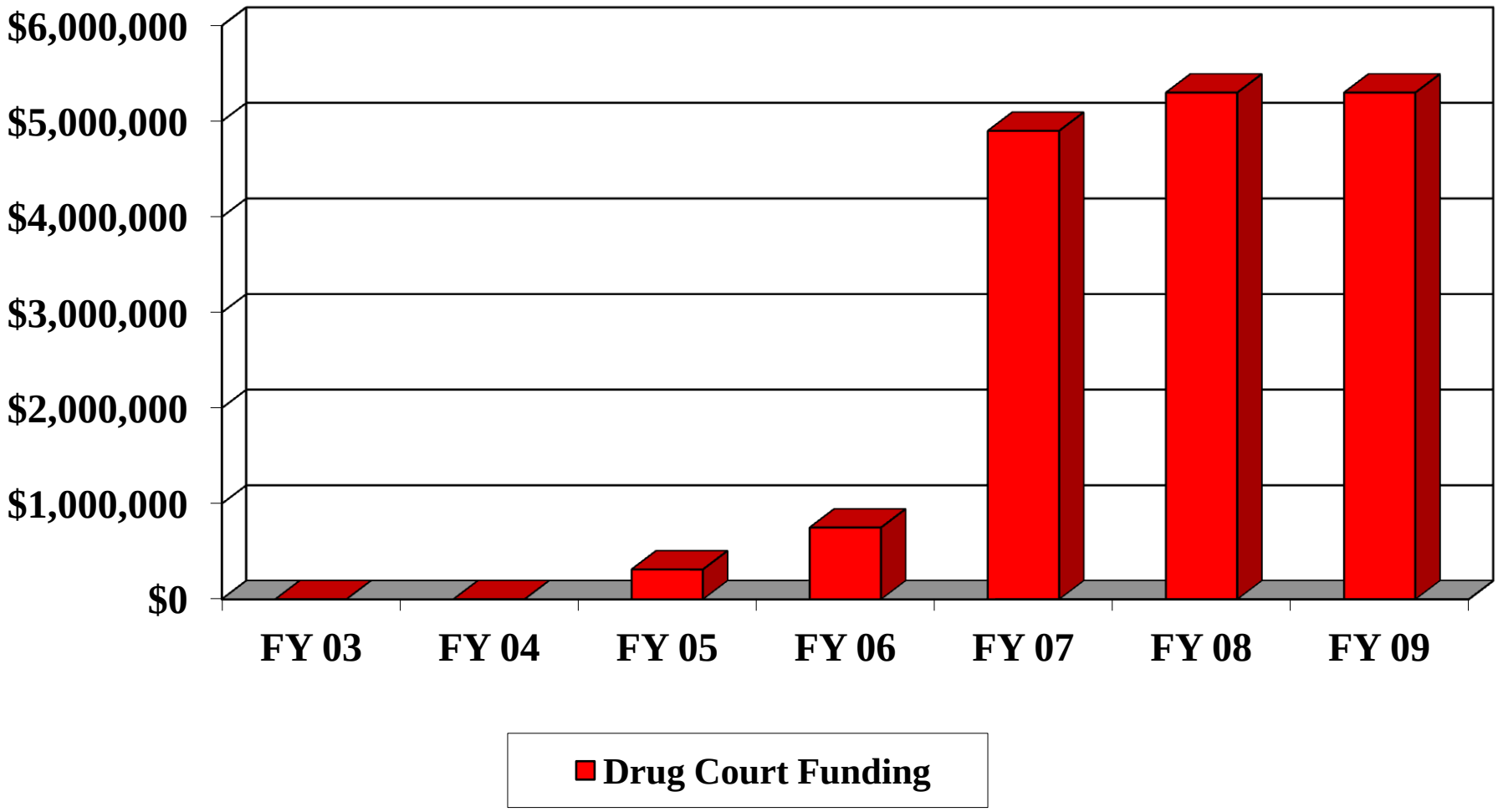


Note: Difference is significant at $p<.10$

History of Drug Courts in Maryland



History of Drug Court Funding in Maryland



What is Evaluation?

- General definition: systematic efforts to collect and use program information for multiple purposes, including program improvement, program accountability, program management, and program development.

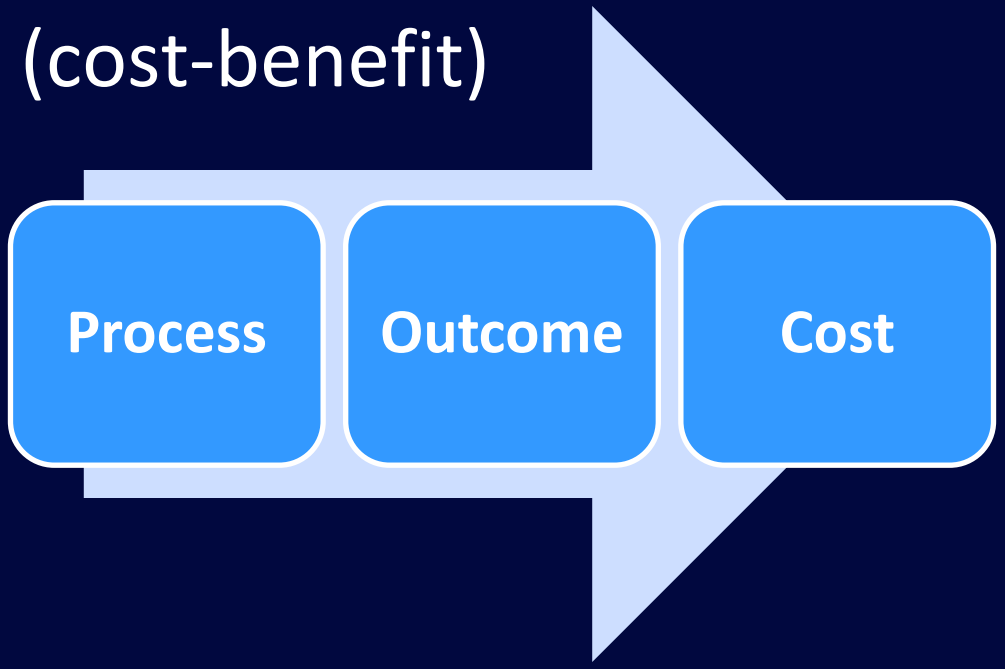
What is Useful Evaluation?

- Evaluation should help to inform and improve programs as they develop, and not focus only on whether the programs “worked” or “didn’t work.”
- Evaluation should involve a collaborative approach with the program sharing in the process and in understanding the results

Evaluation

Three main areas of evaluation:

- Process (program improvement)
- Outcome (impact)
- Cost (cost-benefit)



Process Evaluation

Purpose: Examine program policies & procedures to:

- Determine how the program was implemented and if it was implemented as intended
- Learn whether and how well the program is following the intended model (best practices).

Why do process evaluation?

Benefits:

- Useful Information about program functioning
- Allows an assessment of the reasons for successful or unsuccessful performance
- Provides information for replicating the program in another site
- Contribute to program improvement
- Increase effectiveness for participants
- Better Outcomes, Better Cost-Benefits

Process Methods

Program self review

- Study research on best practices

  Team members can take turns bringing articles and other resources

  Watch short videos or webinars at policy meetings (see NDCRC.org and drugcourtonline.org)

- Self-review program practices

  Ask for feedback from team members and participants (what's working for them, what's challenging, suggestions for improvement) – interviews, surveys, focus groups

  Review policy manual – does manual describe best practices, are you following manual?

Process Methods

Peer Review and outside evaluation

- Program survey
- Interviews with program staff
- Document review
- Site visit/observations
- Focus groups with participants

- Review of program participant information from databases & paper files
- Analyze results (summarize data)
- Interpret results – best practices

Outcome and Impact Evaluation

Purpose: Determine whether the program has improved participant outcomes during and after participation

Outcome and Impact Evaluation

- Are services delivered as planned? (If not, why not?)
- Graduating participants as intended?

  (What is program graduation rate? What is different about those who graduate and those who do not?)

  [This can be done internally]
- Is the program having the intended impact?
  - Are participants being re-arrested less often?
  - Are participants spending fewer days in jail?

  [Difficult to do internally]

Outcome and Impact Evaluation

Outcome Evaluation: Within Program
(services received, grad rate, completion in intended time-frame, factors that lead to graduation)

(No Comparison Group Needed)
Internal/self-monitoring AND outside evaluation

Impact Evaluation: Outside/After Program
(recidivism, subsequent treatment, social services, health care)

(Comparison Group Needed)
Difficult to do internally– research/evaluation expertise needed

Internal Outcome Evaluation Methods

- Enter data consistently into a database (paper files are useless for this)
- Best if you have system that can run reports for you
  (See if an IT person can create reports)
- Examine participant information in a variety of ways with reports
  (e.g., graduation rate; compare grads to term on gender, ethnicity; what phase do drop outs tend to drop out)

External Evaluation Methods

- Obtain access to data (IRB, MOU)
- Collect data
- Prepare datasets (examine/understand, put in format for analysis, match with other data)
- Select and match comparison group
- Analyze data
- Interpret results

Example·of·why·it·takes·so·long·to·clean·and·prepare·data·for·analysis:¶

¶

Below·is·a·frequency·table·for·all·offenses·coded·as·"violations/misc"·in·the·report.·These·19,631·charges·are·from·the·original·raw·dataset·and·represent·5%·of·all·charges·we·received·from·the·DOJ.¶

¶

OffenseDescription¤	Frequency¤	Percent¤	Valid·Percent¤	Cumulative·Percent¤
SEE·COMMENT·FOR·CHARGE¤	5816¤	29.6¤	29.6¤	29.6¤
1203.2(A)·PC-PROBATION·VIOL:REARREST/REVOKE¤	3869¤	19.7¤	19.7¤	49.3¤
1203.2·PC-PROBATION·VIOL:REARREST/REVOKE¤	3718¤	18.9¤	18.9¤	68.3¤
3455·PC-POST·RELEASE·COMMUNITY·SUPV·VIOLATION¤	1914¤	9.7¤	9.7¤	78.0¤
3454·PC-PRCS:FLASH·INCARCERATION¤	1218¤	6.2¤	6.2¤	84.2¤
2620·PC-PROCEEDINGS:PRISONER·PRODUCED¤	896¤	4.6¤	4.6¤	88.8¤
777·WI-MODIFY·COURT·ORDER:COMMITMENT/PROB/ETC¤	430¤	2.2¤	2.2¤	91.0¤
DNA·PALM·COLLECT:NO·ARREST¤	317¤	1.6¤	1.6¤	92.6¤
1767.3·WI-REVOKE/ETC·PAROLE:CYA¤	255¤	1.3¤	1.3¤	93.9¤
602·WI-FAIL·TO·OBEY·ORDER·OF·JUVENILE·COURT¤	193¤	1.0¤	1.0¤	94.9¤
3000.08·PC-VIOLATION·OF·PAROLE¤	169¤	.9¤	.9¤	95.7¤
6253·PC-WORK·FURLOUGH·HOLD/TRANSFER:CDC¤	152¤	.8¤	.8¤	96.5¤
3151·WI-VIOLATION·OUTPATIENT·STATUS¤	55¤	.3¤	.3¤	96.8¤
2621·PC-PRISONER:MATERIAL·WITNESS¤	51¤	.3¤	.3¤	97.1¤
1203.3·PC-PROBATION·REVOKED/ETC¤	42¤	.2¤	.2¤	97.3¤
40302(A)·VC-FAIL·PROVIDE·CDL/ID:WHEN·ARRESTED¤	42¤	.2¤	.2¤	97.5¤
853.8·PC-FAILURE·TO·APPEAR·WARRANT¤	33¤	.2¤	.2¤	97.7¤
23103.5·VC-PLEA·TO·23103·IN·LIEU·OF¤	32¤	.2¤	.2¤	97.8¤
ILLEGAL·ALIEN/IHP·DEPORTATION¤	25¤	.1¤	.1¤	97.9¤
3060·PC·SUSPEND/REVOKE·PAROLE¤	22¤	.1¤	.1¤	98.1¤

Outcome/Impact Methods

- Obtain access to data (IRB, MOU)
- Collect data
- Prepare datasets (examine/understand, put in format for analysis, match with other data)
- Select and match comparison group
- Analyze data
- Interpret results

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 it to. Effective compared to what?

A comparison group is the baseline. It tells us what would have happened if there had been no program.

Comparison groups

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

Common Examples:

- Assumption 1: Participant would have had prison sentence so saved the time/cost of that sentence.
- Assumption 2: Participant would have had children in foster care or in shelter care 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 the sentence.
- Child does not always end up in foster care.
- If child is in foster care or shelter care, it is very difficult to predict how long s/he will stay there.

Comparison groups

Common Mistake: Comparing program graduates to program failures

"It is not surprising that succeeders succeed and failures fail"

---John Goldkamp

Comparison groups

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 or not completing the program, the two groups are inherently different.

Comparison groups

You can compare graduates to terminated *when you want to know what is different between those who graduate and those who do not* :

- Example: If more men graduate than women, it may be an indication that the program needs gender specific services
- Example: If those who had a 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 who did not participate.

Audience Poll

1. Comparison group options

2. Go on to the data that needs to be collected?

Good comparison groups

<u>Random Assignment</u>

"*Gold Standard*" in research and evaluation

- Individuals who have been determined to be eligible for drug court are randomly assigned (flip a coin) to participate in drug court or "business-as-usual"
- Called a *Control Group*

Good comparison groups

Benefits of random assignment:

- Should eliminate any differences between the two groups that could affect outcomes (e.g., criminal history, ethnicity, motivation, age)
- Can be reasonably certain that any difference in outcome is due to the program

Good comparison groups

Drawbacks of random assignment

- Difficulty getting program, judge or team to agree
- Must begin project in real time and wait for outcomes to occur (2 to 3 years)
- This condition does not exist in reality (it is an artificial condition and may not represent true circumstances)
- Groups may not end up comparable unless sample is very large
- Very hard to ensure fidelity to the randomization process

Good comparison groups

Quasi-experimental design:

Historical Comparison Group

- Individuals who were eligible for the drug court before the drug court was implemented.
  - ✓ Benefit: Motivation is less of an issue
  - ✓ Benefit: Larger sample size
  - ✓ Benefit: Reasonably easy to find appropriate individuals

Good comparison groups

Quasi-experimental design:

Historical Comparison Group

- Drawback: Changes in court system over time
- Drawback: Takes time to select (paper files)
- Drawback: May not have all the information necessary to determine exact eligibility

Good comparison groups

Contemporary Comparison Groups
(Same time as program participant sample)

<u>Waiting lists</u>

Eligible for the program but no room

- ✓ Benefit: Very similar to participants
- ✓ Drawback: Generally very small sample (or doesn't exist).

Good comparison groups

Contemporary Comparison Groups
(Same time as program participant sample)

<u>"Slipped through the cracks"</u>

Eligible for the program but were not referred

- ✓ Benefit: Less issue about motivation
- ✓ Benefit: Current "business-as-usual" contemporary
- ✓ Drawback: Can be difficult to find

How do you decide which comparison to use?

- ✓ What resources do you have available for evaluation?
- ✓ What data is available?
- ✓ Is the program too new for an outcome or impact evaluation?
- ✓ How much time do you have?
- ✓ What are you willing to do?

It is not always necessary or feasible to evaluate all outcomes for every program.

Final note on comparison groups

Do not compare graduates to the comparison group

Common for programs to want to look only at their graduates (Graduates are often considered the true product of the program)

- Terminated participants receive program services and use program resources too.
- Terminated participants are also a product of the program.

Final note on comparison groups

Do not compare graduates to the comparison group

- ✓ Graduates = Apples; Terminated = oranges
- ✓ Comparison group has people who would have graduated (apples) had they participated in the program as well as people who would terminate (oranges). (But we don't know which ones.)
- ✓ Comparing graduates to the whole comparison group is comparing "apples" to "apples and oranges."

What Data Should We Collect?

Depends on the purpose: (there is a lot of overlap)

1. **Case Management**
2. **Evaluation**
3. **Self-monitoring**

Case Management Questions

Who drug tested positive yesterday?

(drug test dates and results)

Whose color came up for a drug test today?

(drug test dates and colors in advance)

Who missed a treatment session?

(dates of treatment sessions for each participant)

Who complied with all program requirements for the past two weeks?

Case Management Questions

Program staff are more likely to collect data if the data is useful to them in their every day work in the program

Evaluation Questions

What are your program goals?

- **Reduce recidivism?**
- **Reduce drug use?**
- **Employment?**
- **Education?**
- **Family Reunification?**

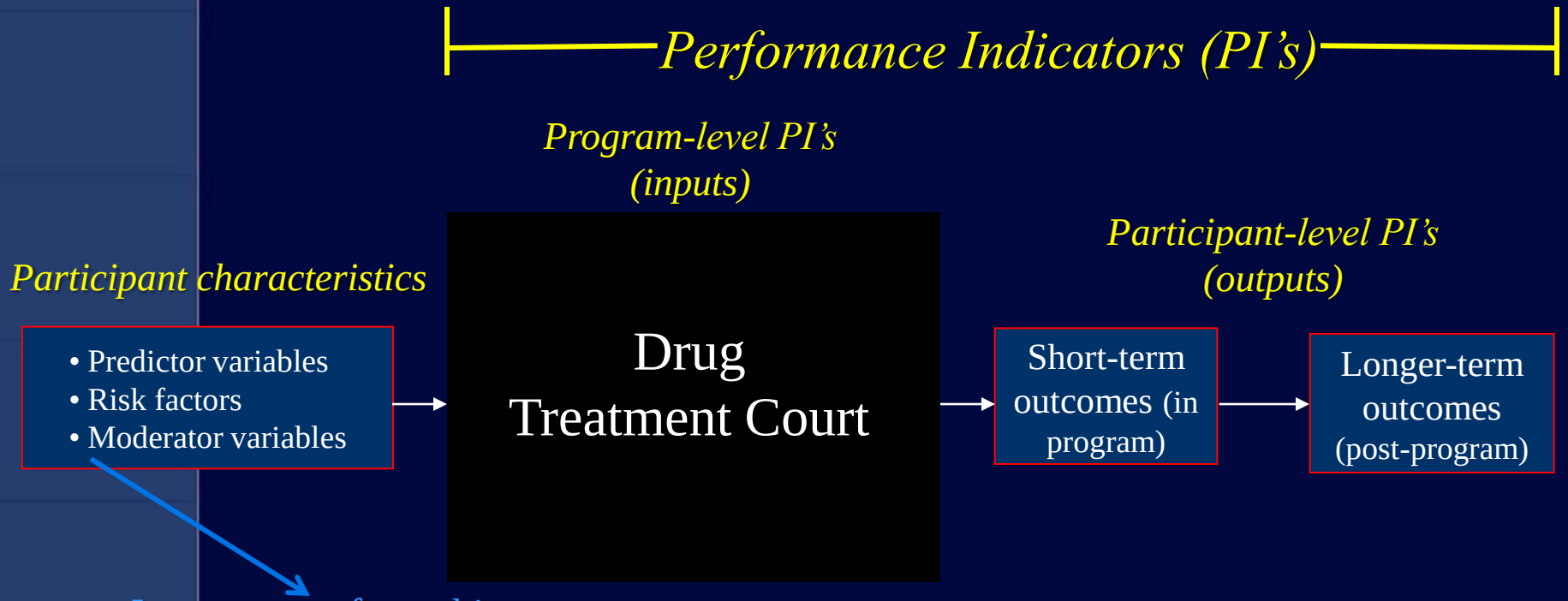
Think about the questions you might have about your participants and your program

Self-Monitoring Questions

- What is different about those who graduate and those who don't?
- When in our process do participants struggle (and drop out)?
- Are participants receiving the services indicated by their assessment?

Think about the questions you might have about your participants and your program

What Data Should We Collect?



Importance of matching:

High risk → intensive supervision

High need → intensive treatment

. . . and vice versa!

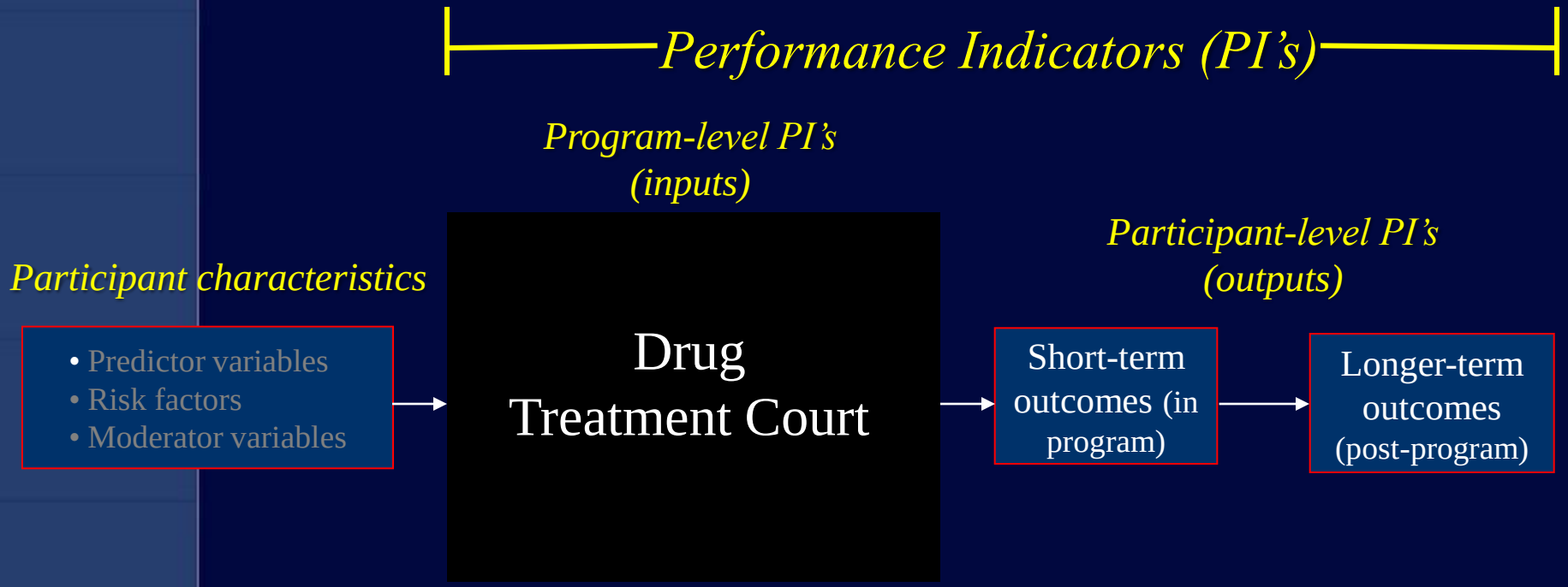
Risk Factors

- Current age (< 25 years)
- Gender
- Delinquent onset (< 16 years)
- Substance abuse onset (< 14 years)
- Prior convictions
- Prior rehabilitation failures
- Prior violence
- Family history of crime or addiction
- Criminal or substance abusing associates
- Antisocial Personality Disorder
- Psychopathy

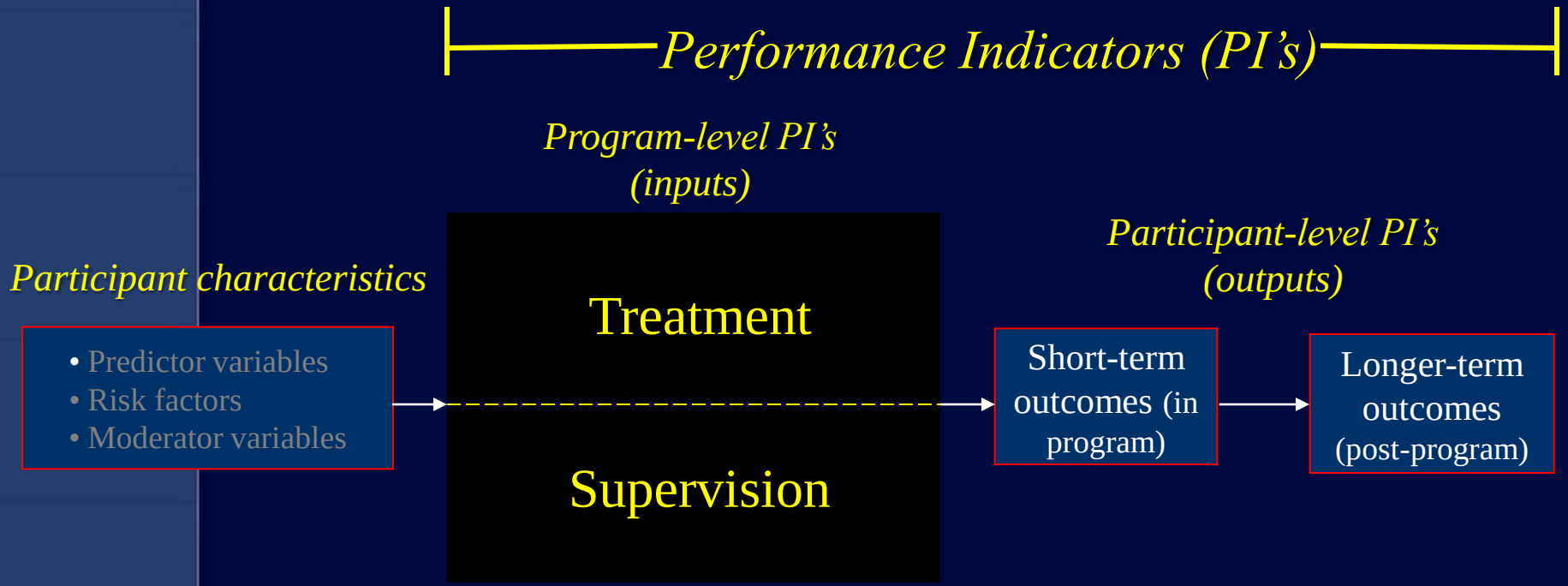
Need Factors

- Substance dependence or addiction
- Severe mental illness + substance abuse
- Lack of employment or educational skills
- Executive / frontal cognitive dysfunction
- Traumatic brain injury (TBI) + substance abuse

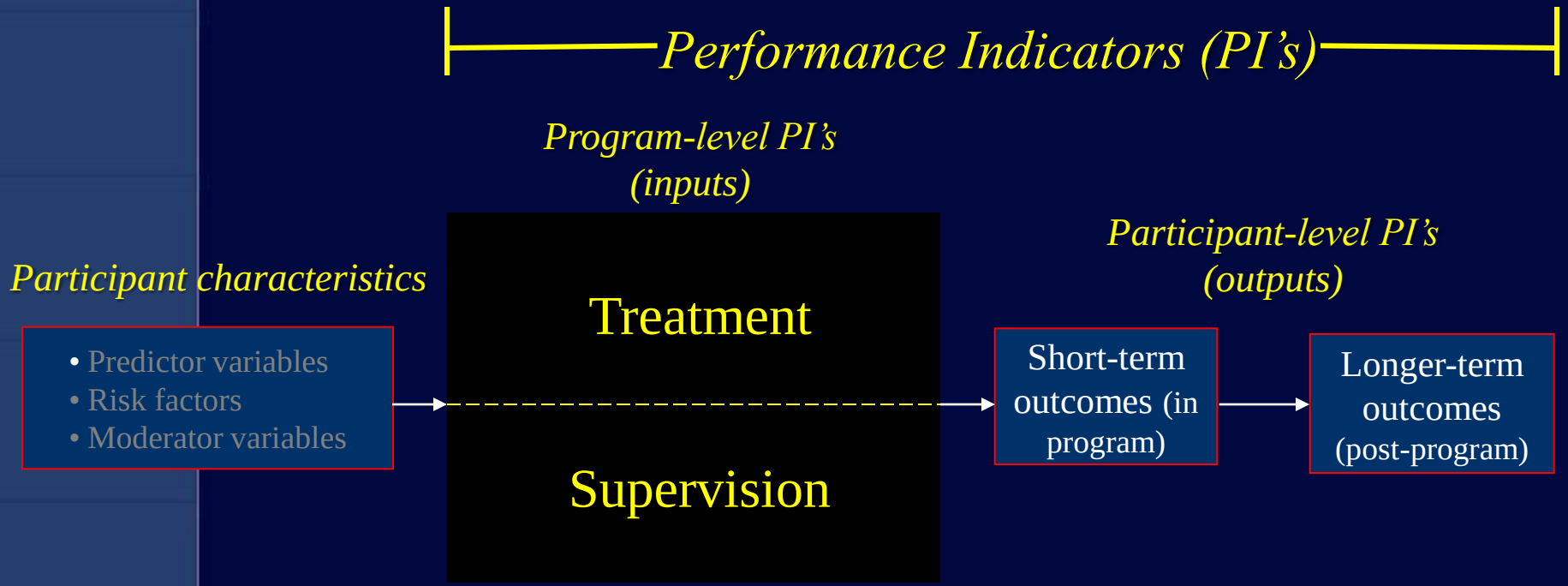
What Data Should We Collect?



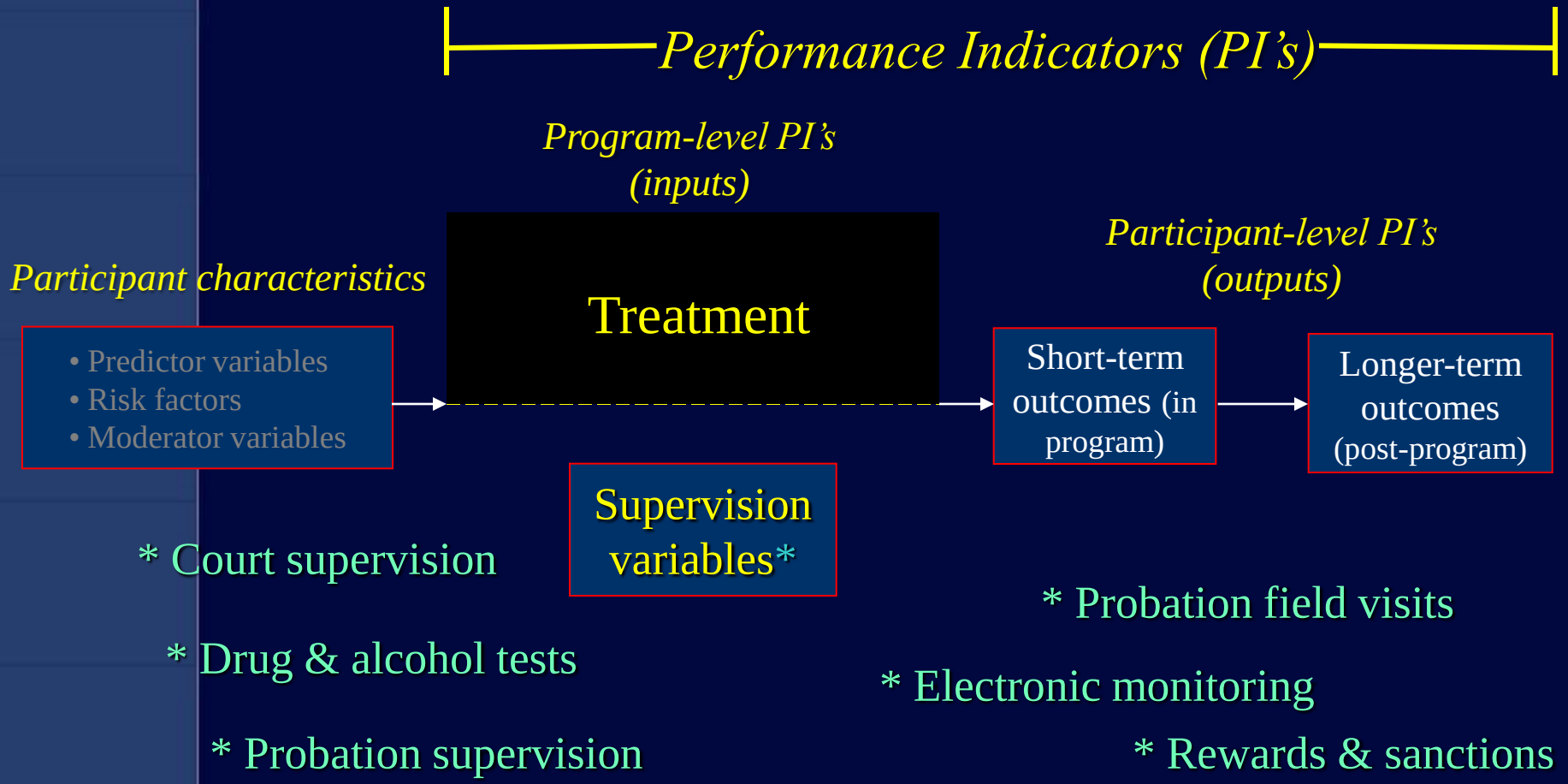
What Data Should We Collect?



What Data Should We Collect?



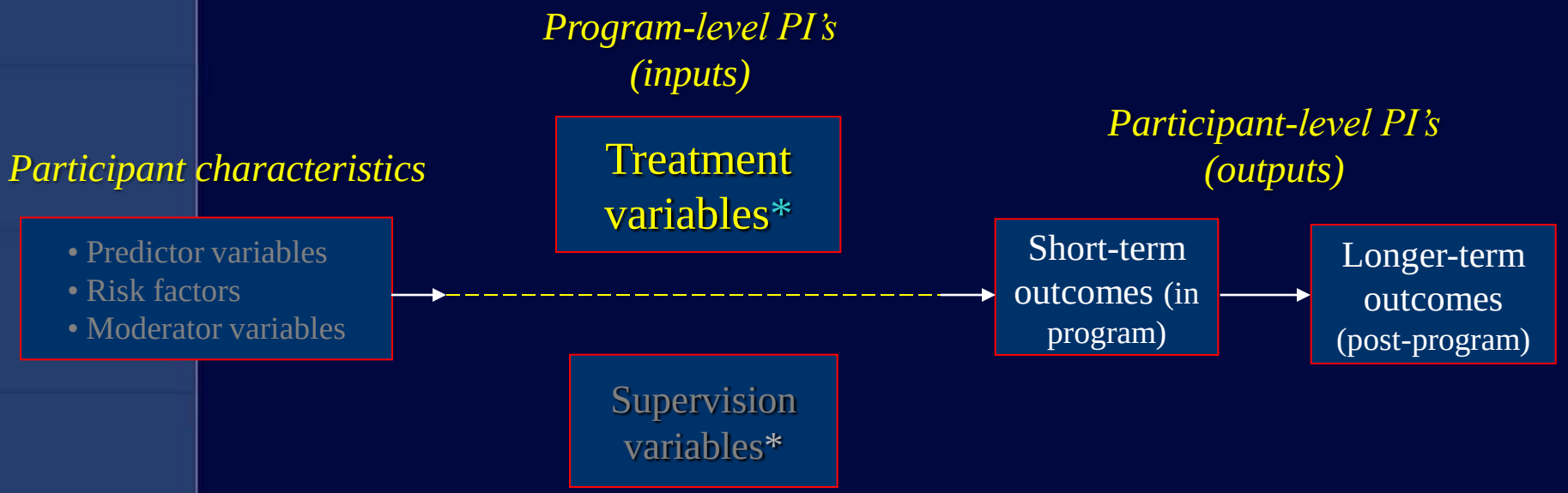
What Data Should We Collect?



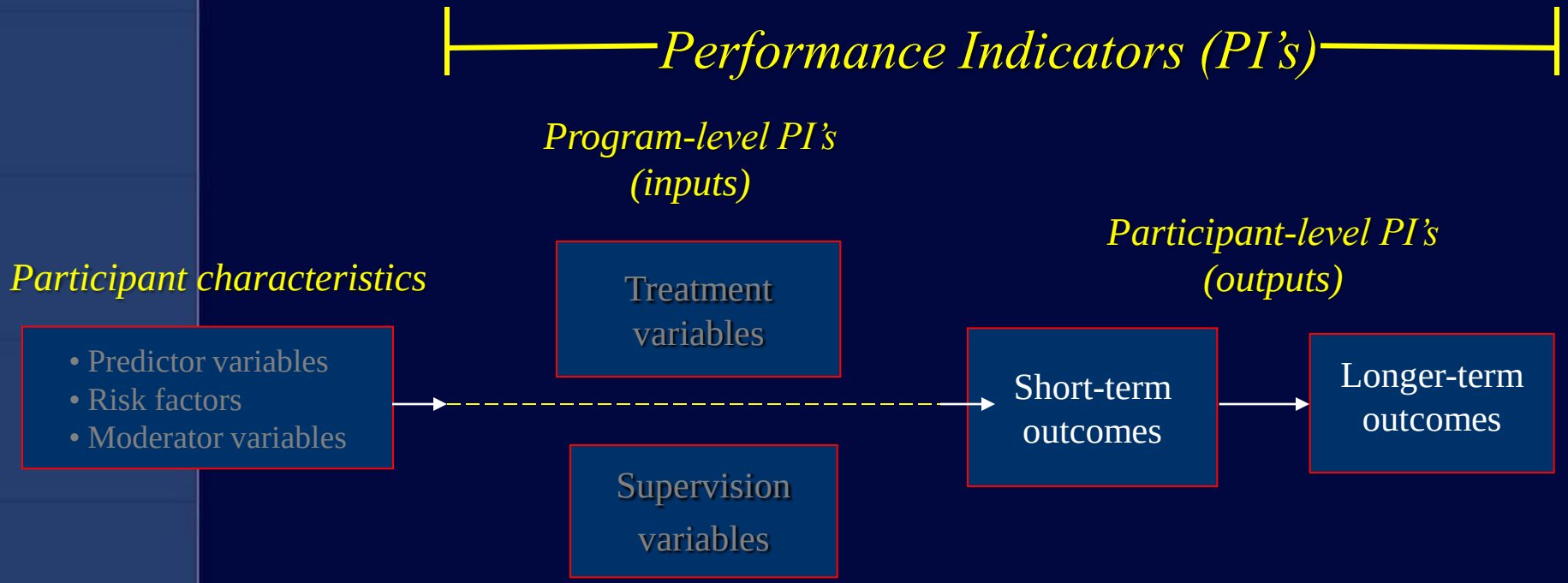
What Data Should We Collect?

* Substance abuse treatment
* Restorative justice services
* Complementary treatment and social services

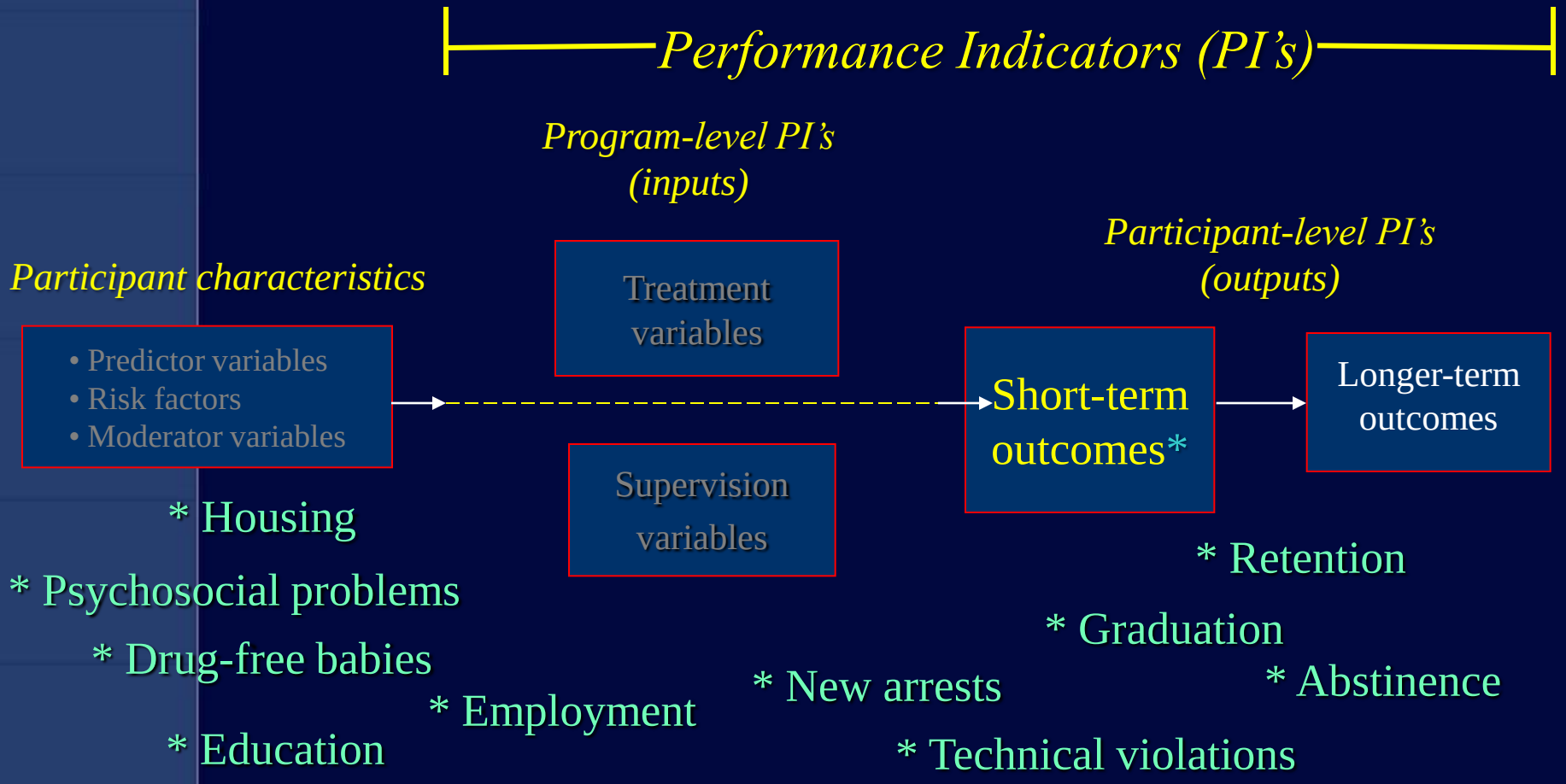
Performance Indicators (PI's)



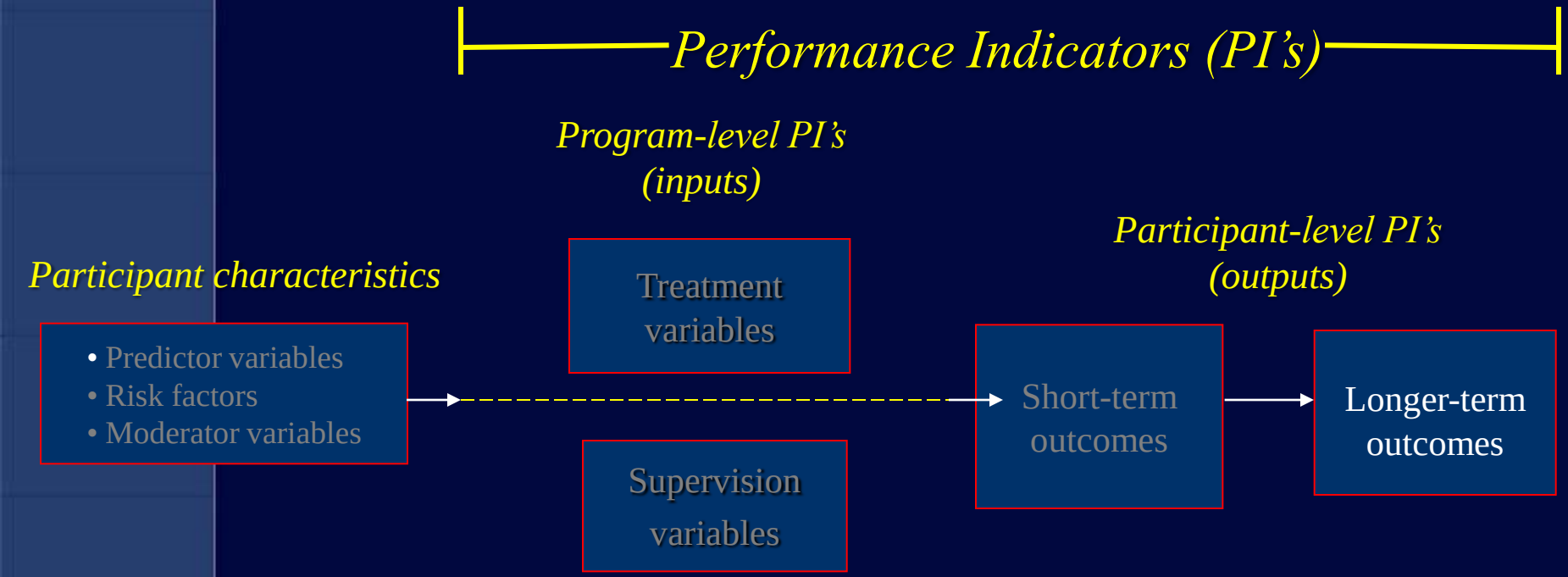
What Data Should We Collect?



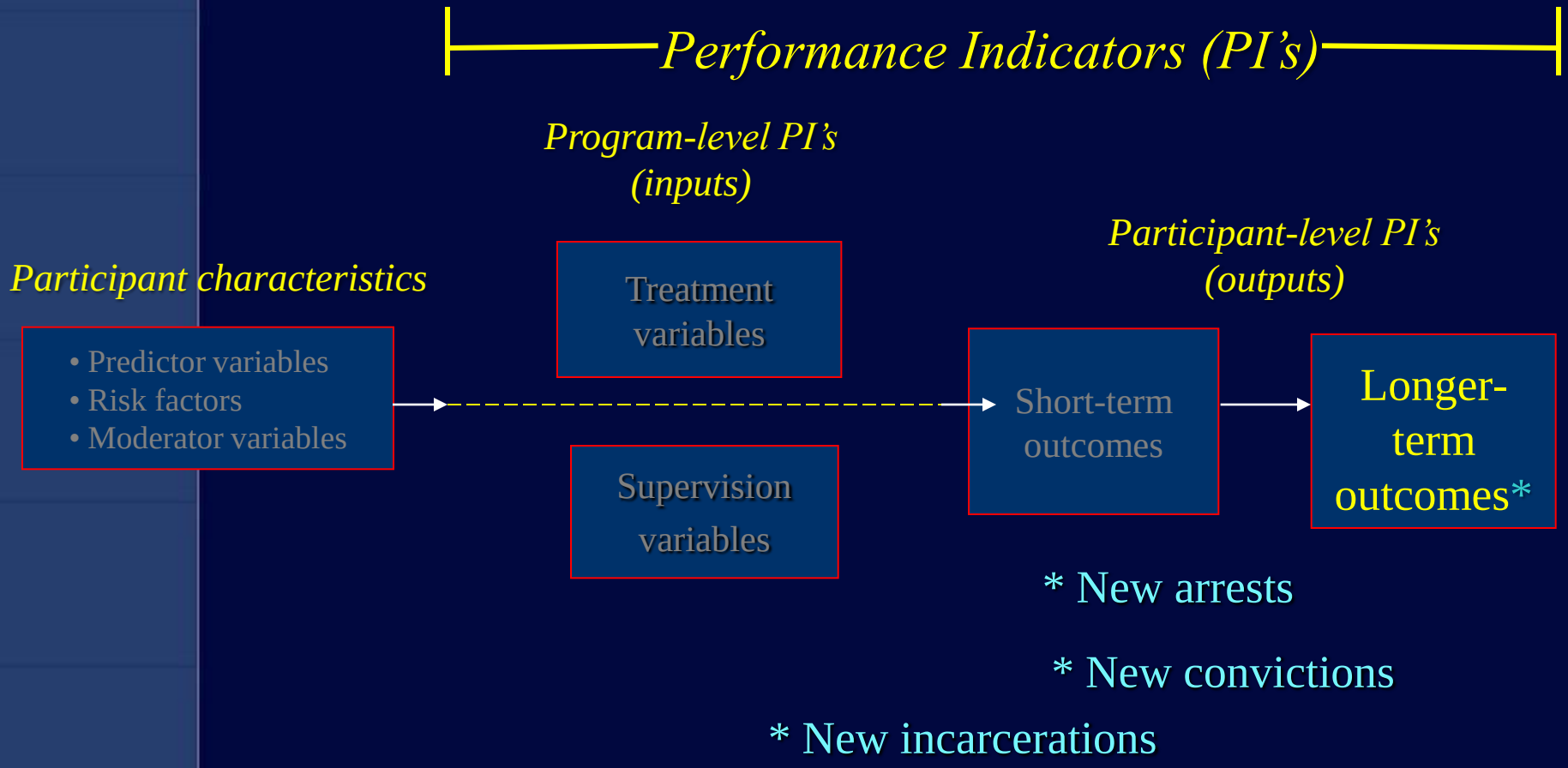
What Data Should We Collect?



What Data Should We Collect?



What Data Should We Collect?



Data Needed
(See Handout)

- Identifiers (so we can find people in other databases)*
- Demographics*
- Drugs of Choice (Primary and secondary)*
- Risk-Need Score
- Program entry and exit dates*
- Date of eligible arrest and court case number
- Date of referral to AODTC program*
- Program status on exit (Graduated/Terminated/etc.)*
- If terminated, reason for termination
- Dates of entry into each phase*
- Dates of UAs (and other drug tests)*
- Dates of positive UAs (and other drug tests)*

*crucial data

Data Needed (Continued)

- Dates of drug court appearances*
- Dates of services received*
- General treatment issues
- Rewards and Sanctions (Dates, types and duration)*
- Non-compliant behavior (Dates, types)
- Aftercare services (Dates and types)
- Employment status at entry and exit
- Education status at entry and exit
- (For Juvenile) School attendance status at entry and exit
- (For Dependency and juvenile) Out-of-home placement and re-unification during program

*crucial data

Impact (and Cost Related) Data

(generally not collected by the program but IMPORTANT for evaluation – so find out source)

- Subsequent treatment episodes (after program)
- Dates of re-arrest after entering the drug court program*
- Dates of DUI re-arrests after entering the program*
- Probation start and end dates
- Jail/Detention entry and exit dates
- Prison start and end dates
- Social and health services information
- (For juveniles) School related data such as completion status
- (For Dependency and Juvenile) Out-of-home placement and re-unification

Cost Evaluation

Cost-effectiveness **analysis calculates the cost of a program and then examines whether the program led to its intended positive outcomes. (Outcomes are not "costed." (e.g., for every $1 spent there is a 10% reduction in recidivism)**

Cost-benefit **evaluation calculates the cost of the program and also the cost of the outcomes, resulting in a cost-benefit ratio. (e.g., for every $1 spent on the program, $7.50 is saved in outcomes.)**

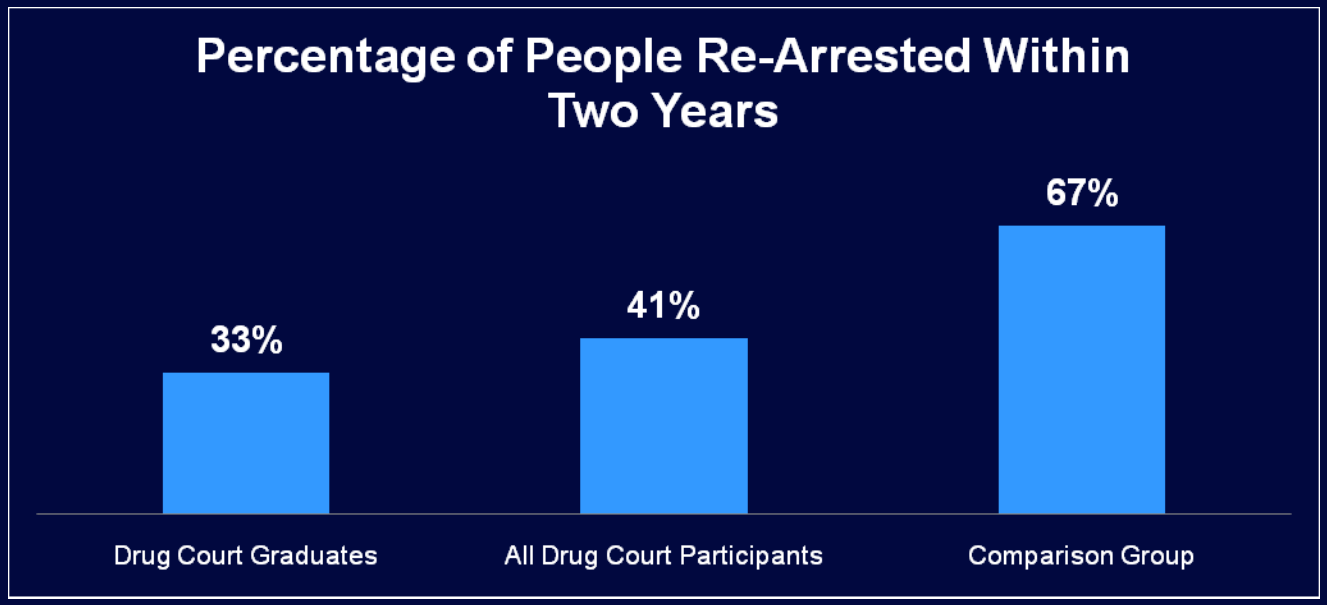
Cost Evaluation

Cost Evaluation questions:

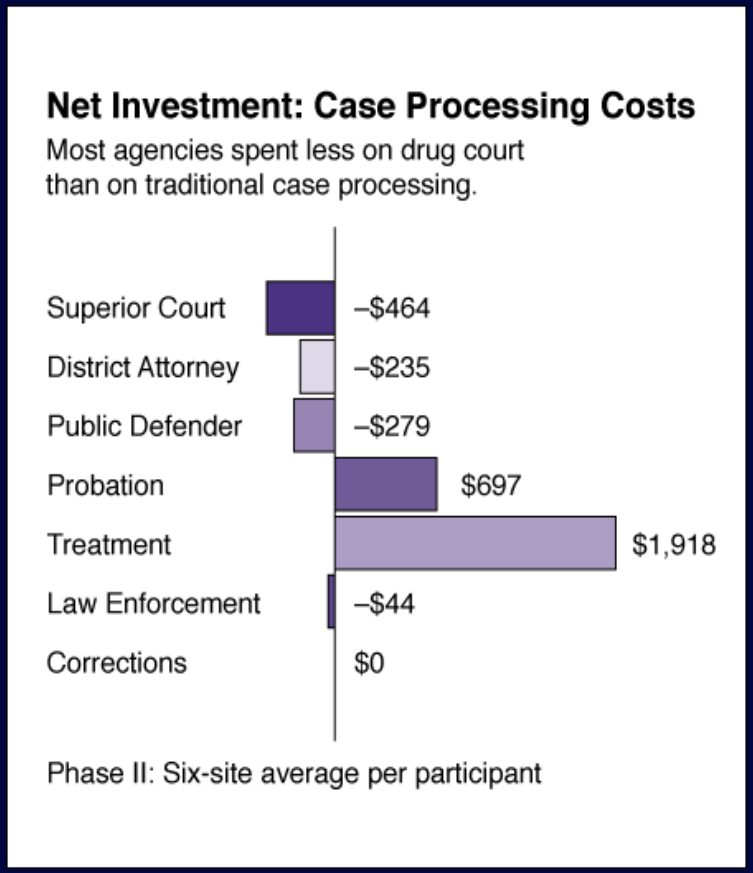
1. What does it cost to run the program?
2. What are the benefits?
3. How will taxpayers be impacted by decisions to add, expand, eliminate or shrink these programs?
4. What are the financial implications of problem solving court approaches to meeting public service needs?
5. What does it cost the different agencies that contribute resources to the program? What are the benefits to these agencies?

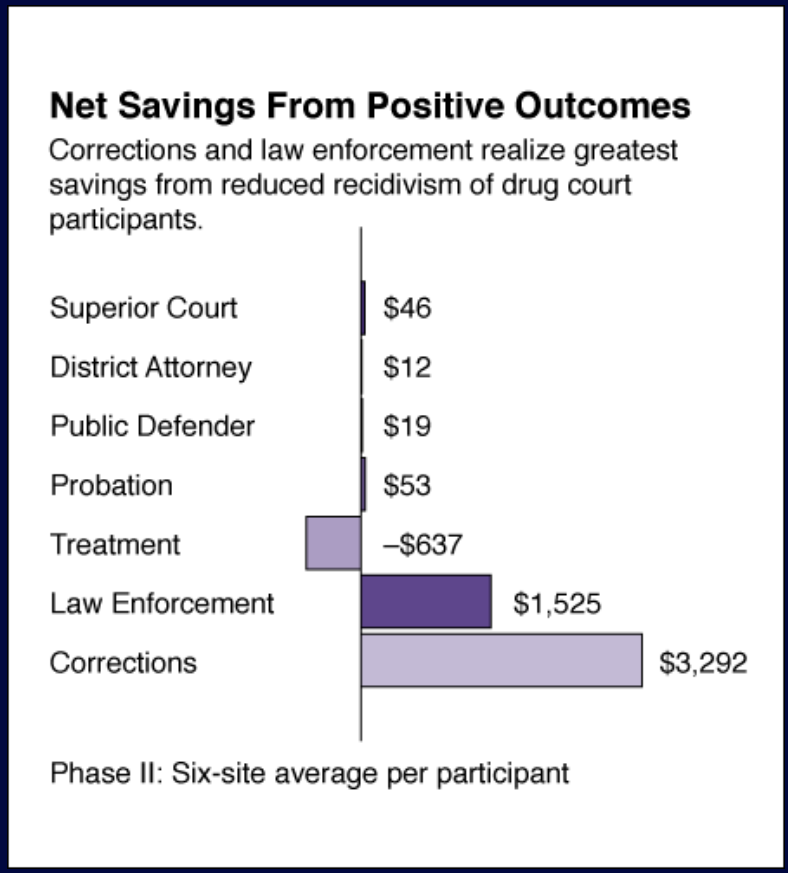
How do you use evaluation results to get funding?

- Fact sheets (1 or 2 page quick summary of positive outcomes – e.g., cost savings)
- Executive summaries
- Anecdotes from graduates
- Graduates in person



Results
Costs by Agency





NPC RESEARCH
www.npcresearch.com

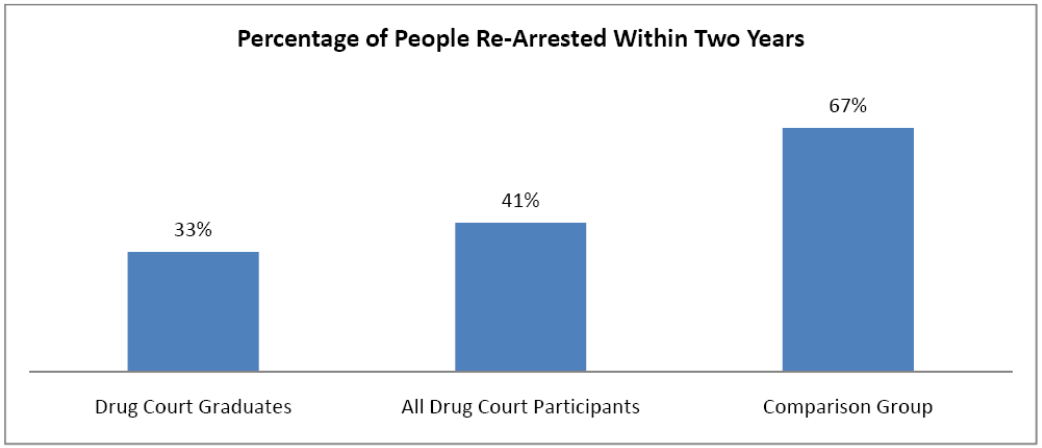
San Mateo County (California) Drug Court – South Cost Study Fact Sheet

The following statistics are part of a multi-site evaluation of the costs and benefits of California's drug courts. This fact sheet is a component of Phase III, statewide launch phase, of a research effort to develop a statewide methodology for assessing the benefits and costs of drug courts in the State of California. The aim of this effort is to produce a validated methodology to conduct inexpensive cost-benefit studies on an ongoing basis of drug courts throughout the state. As a part of this effort, a web-based tool was created – the Drug Court Cost Self-Evaluation Tool (DC-CSET) – which drug courts statewide can use to help determine their own costs and benefits. For more information on this study and other drug court studies go to www.npcresearch.com and www.courtinfo.ca.gov/courtadmin/aoc.

Recidivism Rate

San Mateo County Drug Court - South participants had the following recidivism (re-arrest) rates over a two-year period.[1] Re-arrests include any type of arrest (but not including traffic citations).

Figure A: Two-Year Re-Arrest Rate for San Mateo County Drug Court – South: Graduates, All Participants and Comparison Group

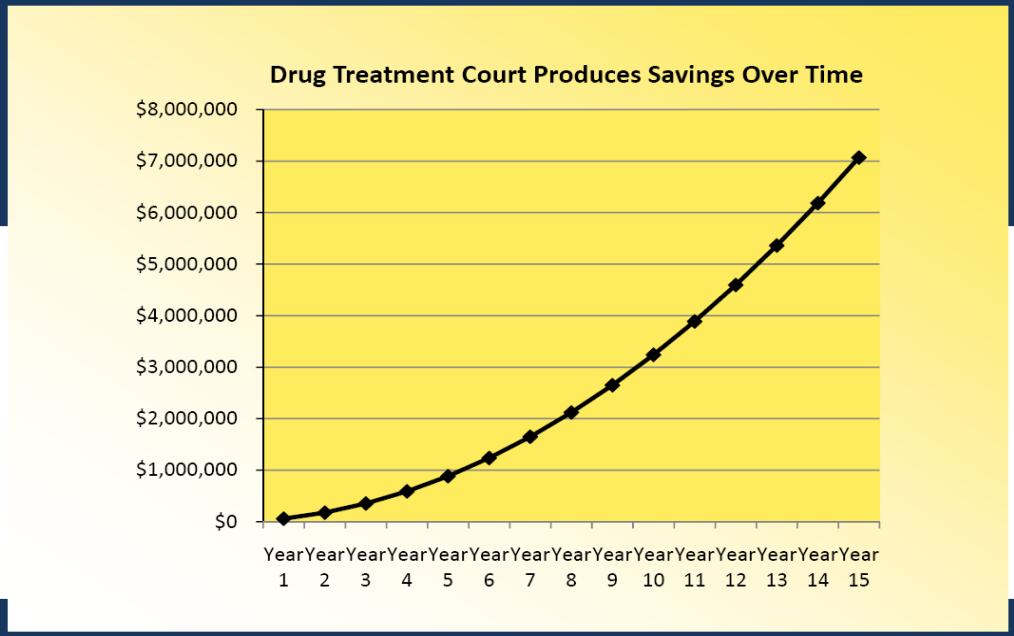


Costs and Benefits Due to Recidivism

Table A shows the average number of recidivism events (e.g., re-arrests, jail time, prison time) per offender for drug court graduates, all drug court participants (regardless of graduation status) and the comparison group over 2 years after drug court entry.[2]

❖ What individual and programmatic factors influence recidivism?

- **Younger participants and individuals with more arrests prior to DTC participation were more likely to re-offend.**
- When controlling for the other variables, the odds of re-offending decreased 7% with each additional year older the participant was at DTC start and increased 48% with each additional prior arrest.
- An optimal length of stay in the program—approximately 13 to 14 months—was associated with no cumulative subsequent arrests.



❖ Do drug treatment courts save money in the long run?

YES. Overall, the DTC results in cost savings, especially for program graduates. The DTC program had criminal justice system outcome cost savings of **$2,945 per participant** after 10 years. Outcome cost savings were **$46,207 per graduate** after 10 years, so there is a clear benefit to the taxpayer in working to engage offenders and helping them successfully complete the DTC program. Overall, these results demonstrate that the DTC program uses fewer criminal justice system resources than traditional court processing.

The chart above illustrates actual and projected cost savings over time.

NPC RESEARCH
4380 SW Macadam Ave., Ste. 530 | Portland, OR 97239
503.243.2436 | www.npcresearch.com

Show the Human Side

Before DC

After DC

Who do you tell?

- **The Media:**

  To get the word out to the general public (e.g., community activities)

- **The Legislature**:

  Senate Health & Welfare, House & Senate Judiciary Committees, Senate & House Appropriations, House Institutions & Correction

- **State Policy Executives**:

  Secretary of the Agency of Human Services, Deputy Commissioner of Health, Commissioner of Mental Health, Commissioner of Corrections

Contact Information

Charlene Zil
zil@npcresearch.com

Shannon Carey, Ph.D.
carey@npcresearch.com

www.npcresearch.com