Behavioral Crisis Response and the Role of Data

Tracking outcomes for co-responder and diversion programs

06.22.2026

Executive summary

Agencies are being asked to change how they respond to people in crisis, and to prove the new approach works. Both halves of that demand are hard, but the second one is where most programs fall short.

Co responder teams, mobile crisis units, and diversion programs have spread quickly since the 988 Crisis Line launched. The idea behind them is sound: pair officers with behavioral health professionals, or route the call to clinicians entirely, so that a person in crisis gets care instead of a cell. Early results from some programs are genuinely impressive. Results from others are flat or unclear.

That mixed picture is the problem this paper addresses. It is tempting to read the inconsistent evidence as a verdict on the programs themselves. The better reading is that the field has a measurement problem. Programs define success differently, capture data inconsistently, and often cannot follow a person past the initial call. Without solid, standard outcome data, an agency cannot prove its program works, cannot improve it, and cannot defend its funding. This paper lays out what to measure, why the current data falls short, and how consistent tracking turns a promising program into a case an agency can stand behind.

A new model under a bright spotlight

The way America responds to mental health crises is being rebuilt in real time.

In 2020, federal law created the 988 Suicide and Crisis Lifeline, which launched in July 2022. Its purpose was to route mental health crisis calls away from the default of police and jail and toward behavioral health services. That shift drove a surge of interest, new legislation, and significant funding for alternatives to a standard police response. Mobile crisis teams became a core piece of the new system, alongside co responder models that pair officers with clinicians. This matters because the old default served no one well. Most police encounters with people in crisis are not about danger. They are welfare checks and domestic disputes. In one study of these calls, only about one in ten ended in arrest, while most ended in a transport to a hospital or mental health unit. The people involved needed care, not custody. The new models exist to deliver that.

But the spotlight cuts both ways. Because these programs are new, publicly funded, and politically visible, they are under pressure to prove their worth from day one. Councils want to know the money is working. The public wants to know the approach is safe and fair. And agencies want to know they are actually helping. Answering any of those questions requires data most programs are not capturing well.

The promise: fewer arrests, more care

Start with what these programs are built to achieve, because the goals define what you have to measure.

A co responder or diversion program aims to change the outcome of a crisis call in a few specific ways. It aims to divert the person away from jail and away from an unnecessary emergency room visit. It aims to link them to community based mental health services and, crucially, to get them to actually engage with that care. It aims to reduce the use of force, because a clinician on scene changes how the encounter unfolds. And it aims to reduce repeat crisis calls over time, by addressing the underlying need instead of just clearing the call.

When these programs work, the results can be striking. A statewide co responder evaluation covering more than 25,000 calls found that almost all of them, around 98 percent, ended without an arrest, and involuntary holds dropped substantially. Across multiple studies, use of force fell by somewhere between roughly 20 and 50 percent where co responder teams were involved. Other research found fewer repeat police contacts and more referrals into community mental health services. Federal analysis has long held that mobile crisis teams are effective at diverting people from psychiatric hospitalization and are better than hospitalization at linking people to ongoing outpatient care.

Those are real, meaningful outcomes. They are also not the whole story.

The evidence is real, and it is mixed

An honest paper has to say this plainly: the research on these programs is not uniformly positive. The results vary, sometimes sharply.

Some studies found minimal or no statistically significant change in the outcomes they measured. Others found improved links to services but no matching drop in hospitalization or

repeat crises. One rigorous economic study, built on a randomized trial, found no cost savings from a co response program at all. In that study, the people who received a co response actually had higher behavioral health and emergency costs over the following year, likely because they were now connected to care they had not been getting before.

That last finding is worth sitting with, because it shows how slippery "success" can be. Higher outpatient costs could mean the program failed to save money, or it could mean the program did its job by connecting people to treatment they needed. The same number reads as failure or success depending on what you decided to measure and why.

This is the pattern across the literature. Outcomes vary by program design, by jurisdiction, by which measures were used, and by how long researchers followed up. A program that looks like a triumph on one metric looks like a wash on another. And that inconsistency is not mainly a sign that the programs do not work. It is a sign that the field does not yet measure them in a consistent, disciplined way.

Why the mixed picture is really a measurement problem

Look closer at why the evidence is so scattered, and a data problem comes into focus.
There is no shared definition of success. "Co responder program" is not one thing. Some pair an officer with a clinician. Some send clinicians alone. Some dispatch from 911, others from 988, others self dispatch. With so many models flying under the same banner, and each measuring different outcomes, comparing them is nearly impossible. National efforts are only now trying to standardize the basic vocabulary so programs can be compared at all.

The data lives in too many places. A single crisis call touches the police records system, dispatch, emergency medical services, the hospital, the jail, and one or more community providers. Each holds a piece of the outcome. Almost no agency can see all those pieces together, so the full arc of what happened to the person is invisible.

Follow up rarely happens. The outcomes that matter most, like whether the person engaged with care and whether they avoided another crisis, unfold in the weeks and months after the call. But most programs capture only the disposition at the scene and then lose track. Without follow up data, you can measure what you did but not whether it worked.

Capture is an afterthought. When documenting an outcome is slow or optional, it does not happen consistently. The record ends up partial, which makes any analysis built on it shaky. Put these together and the mixed evidence makes sense. The programs may well work. We often just cannot tell, because we are not measuring them well enough to know.

What a crisis program should actually measure

If the problem is measurement, the fix starts with deciding what to measure and capturing it every time. A serious crisis response program tracks outcomes at three stages.

At the scene. What was the disposition? Was an arrest avoided? Was an involuntary hold avoided? Was force used, and if so, how much? How long did the call take? These are the immediate outcomes, and they are the easiest to capture because the responder is right there.

The handoff. Was the person referred to a specific service? Was a warm handoff made, or just a phone number handed over? Diversion only counts if it connects to something real, so the quality of the linkage matters, not just the fact of a referral.

The follow up. Did the person actually engage with the service they were referred to? Did they have another crisis call in the following weeks or months? This is the hardest data to get and the most important, because it is the only thing that shows whether the intervention changed the trajectory or just changed the paperwork.

Alongside these, a program should track the population it serves and watch for disparities, so leaders can see whether the program reaches everyone fairly. The point is not to drown responders in fields to fill out. It is to define a tight, meaningful set of outcomes and capture them consistently on every contact, so the data adds up to something real.

From scattered notes to a defensible case

Consistent outcome data does three things a pile of anecdotes cannot.

It proves the program works, or shows honestly where it does not. When a council asks whether the money is justified, "we handled 1,200 crisis calls and diverted 94 percent from arrest, with force used in under 3 percent and two thirds engaging with follow up care" is an answer. "We think it is going well" is not. In an era of tight budgets and visible scrutiny, the program that can show its outcomes is the program that keeps its funding.

It drives improvement. When you can see that a certain type of call keeps ending in a repeat contact, or that one referral partner never results in engagement, you know where to fix the program. Without the data, you are guessing.

It builds trust. Transparent, consistent outcome reporting, including on fairness across the community, is how an agency earns public confidence in a new and closely watched approach. None of this is possible when the data is trapped in separate systems and half captured. It requires a single place to record every crisis contact and its outcome in a consistent form, connected to the rest of the agency's operations. This is the role a dedicated behavioral crisis response tool plays. Command HQ's application for this work, REACH, is designed to capture crisis contacts and their outcomes in a structured, consistent way and turn the accumulated record into a clear picture of what the program is actually achieving.

A practical place to begin

You do not need a research grant to measure your program well. Start here.

- **Define your outcomes first.** Pick a tight set of measures across the scene, the handoff, and the follow up. Write down exactly what each one means so everyone records it the same way.
- **Make capture fast and mandatory.** Build the outcome fields into the responder's normal documentation so recording them takes seconds and never gets skipped.
- **Close the follow up loop.** Set up a way to check whether referred individuals engaged with services and whether they had a repeat contact. This is the data that proves impact.
- **Review the numbers on a schedule.** Look at your outcomes monthly, watch for disparities, and use what you find to adjust the program.
- **Report out.** Turn the data into a clear, regular report for leadership, funders, and the public.

What gets measured gets funded

Behavioral crisis response is one of the most important shifts in modern public safety. Done well, it gets people care instead of custody, reduces force, and eases the load on jails and emergency rooms. The early evidence for the best programs is real.

But the field's evidence is mixed, and the mixed picture is mostly a measurement failure, not a program failure. Programs that cannot define success, capture it consistently, and follow it past

the initial call will never be able to prove their worth, improve their results, or defend their budgets. The programs that can will do all three.

The work of crisis response happens on the street. The case for it is made in the data. An agency that takes measurement as seriously as it takes the response will be the one still running its program, and improving it, years from now.