Journal of Criminal Justice 40 (2012) 60–71

Contents lists available at SciVerse ScienceDirect

Journal of Criminal Justice

Assessing the effectiveness of drug courts on recidivism: A meta-analytic review of traditional and non-traditional drug courts

Ojmarrh Mitchell [a,*], David B. Wilson [b], Amy Eggers [a], Doris L. MacKenzie [c]

[a] Department of Criminology, 4202 E. Fowler Ave. SOC107, University of South Florida, Tampa, FL 33620, USA
[b] Criminology, Law and Society, 4400 University Drive, MS 4F4, George Mason University, Fairfax, VA 22030, USA
[c] Law and Justice and Sociology, The Pennsylvania State University, 327 Pond Building, University Park, PA 16802, USA

ARTICLE INFO

Available online xxxx

ABSTRACT

Purpose: The objective of this research was to systematically review quasi-experimental and experimental evaluations of the effectiveness of drug courts in reducing offending.
Methods: Our search identified 154 independent evaluations: 92 evaluations of adult drug courts, 34 of juvenile drug courts, and 28 of DWI drug courts. The findings of these studies were synthesized using meta-analysis.
Results: The vast majority of adult drug court evaluations, even the most rigorous evaluations, find that participants have lower recidivism than non-participants. The average effect of participation is analogous to a drop in recidivism from 50% to 38%; and, these effects last up to three years. Evaluations of DWI drug courts find effects similar in magnitude to those of adult drug courts, but the most rigorous evaluations do not uniformly find reductions in recidivism. Juvenile drug courts have substantially smaller effects on recidivism. Larger reductions in recidivism were found in adult drug courts that had high graduation rates, and those that accepted only non-violent offenders.
Conclusions: These findings support the effectiveness of adult drug courts in reducing recidivism. The evidence assessing DWI courts' effectiveness is very promising but more experimental evaluations are needed. Juvenile drug courts typically produce small reductions in recidivism.

Introduction

In just two decades, drug courts grew from a solitary court in one jurisdiction to a movement with thousands of courts spread throughout the United States (Huddleston, Marlowe, & Casebolt, 2008; Huddleston & Marlowe, 2011). Drug courts not only have grown in number, but also in type. Originally, drug courts targeted adult, illicit substance users; yet, in recent years drug courts have been applied to non-traditional populations, namely, juvenile illicit substance users and repeat DWI offenders (Huddleston et al., 2008).

Despite the popularity of drug courts, their effectiveness in reducing recidivism remains ambiguous as several issues have not been sufficiently addressed. While many drug court evaluations find that participants have lower recidivism rates than non-participants, much of this research is methodologically weak. Most distressingly, prior syntheses have found that drug courts' effectiveness in reducing recidivism is weakest among the most methodologically rigorous evaluations (e.g., Belenko, 2001; U.S. General Accounting Office [GAO], 1997, 2005; Shaffer, 2011; Wilson, Mitchell, & MacKenzie, 2006). Another recurring, unresolved issue is the duration of drug courts' effects on recidivism. Prior reviews note that most evaluations track recidivism during the period of involvement in drug court or shortly thereafter; thus, drug courts' long-term effects on recidivism are unclear. A third important issue concerns the effectiveness of non-traditional drug courts (i.e., juvenile drug courts and DWI drug courts), as prior reviews have not distinguished these courts effects on recidivism from those of traditional, adult drug courts. Finally, it is unclear which drug court features are associated with greater effectiveness in reducing recidivism.

The purpose of this meta-analytic synthesis of drug court evaluations is to address these issues utilizing a much larger number of evaluations than used in previous reviews. Specifically, in this research, we synthesize drug court evaluations, while distinguishing evaluations of adult drug courts (i.e., drug courts targeting adult, illicit substance users) from evaluations of juvenile drug courts and DWI drug courts. We also thoroughly examine the long-term effects of drug court participation on recidivism. And perhaps most important we

* Corresponding author. Tel.: +1 813 974 8344.
E-mail addresses: omitchell@usf.edu (O. Mitchell), dwilsonb@gmu.edu (D.B. Wilson), aeggers@mail.usf.edu (A. Eggers), dlm69@psu.edu (D.L. MacKenzie).

0047-2352/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcrimjus.2011.11.009

examine variation in drug courts' effects by a number of moderator variables with a particular emphasis on moderators assessing methodological rigor and drug court characteristics.

Drug courts and their rise

The drug court model combines drug treatment with the legal and moral authority of the court in an effort to break the cycle of drug use and drug related crime (GAO, 1997; National Association of Drug Court Professionals, 1997). Briefly, a prototypical drug court operates as follows (National Association of Drug Court Professionals, 1997; Mitchell, 2011): Shortly after arrest, drug-involved offenders who appear to be eligible for drug court participation are identified and screened for program eligibility. Arrestees deemed eligible are offered entry into the drug court with an agreement that the charges against them will be reduced or dismissed upon successful program completion. Arrestees who agree to enter the drug court become drug court "clients." Once in the drug court, clients have their cases handled non-adversarially in one of two ways. In the "pre-plea" case processing method clients waive their right to a speedy trial and enter drug court; if they successfully complete court requirements, then their charges are dropped. In the "post-plea" case processing method, clients are admitted to drug court after conviction but before sentencing. Clients who successfully complete the program typically receive a sentence of time served or probation.[1] As a condition of program entry, drug court clients agree to abide by the court's demands, which typically include frequent urine testing, treatment attendance, and appearance before the court for status hearings. These status hearings are crucial as it is here that the drug court judge and clients converse directly, and it is in these hearings where judges in collaboration with other court actors most clearly use the authority of the court to motivate compliance. The court uses various rewards (e.g., praise, tokens of achievement, movement to the next phase of the program) and sanctions (e.g., increased treatment attendance or urine testing, short jail stays) to compel compliance to program requirements. Compliant clients advance through three or more, progressively less intense stages before completing the drug court, which typically takes at least one year. Ultimately, successful drug court clients are acknowledged at a formal graduation ceremony.

Drug court eligibility requirements vary across the thousands of jurisdictions operating such courts. In the majority of jurisdictions, however, eligibility is restricted to non-violent offenders with evidence of substance dependence (Belenko, 1998). Most commonly, non-violent offenders are defined as offenders neither charged with, nor previously convicted of, a serious violent offense. While not all jurisdictions restrict eligibility to non-violent offenders, the vast majority of drug courts do, in part because this criterion is necessary to be eligible for federal drug court funding.[2] Many courts also exclude arrestees charged with drug trafficking offenses, with three or more prior felony convictions, or with serious mental health issues (see e.g., Kalich & Evans, 2006). In the end, most eligible offenders are charged with drug or property offenses and have relatively few prior felony convictions.

It is important to emphasize that the program requirements for drug courts are often strict and clients are closely monitored for adherence to the demands of the program. Perhaps the best evidence of the rigor of the drug court model is the high percentage of drug court clients who fail to graduate from these programs. For example, a GAO survey of drug courts operating at the end of 1996 found that "about 48%" of drug court clients successfully completed the program (GAO, 1997, p. 56). Similarly, Belenko's (2001) review of drug court evaluations found an average graduation rate of 47% with a range of 36% to 60%. Thus, the best estimate of drug court graduation rates is just under 50%.

Taken together, the key components of drug courts are: (1) collaborative, non-adversarial, outcome driven court processing, (2) early identification of eligible offenders; (3) drug treatment integrated into criminal justice case processing; (4) urine testing; (5) judicial monitoring; and, (6) the use of graduated sanctions/rewards (National Association of Drug Court Professionals, 1997; Hora, 2002). These components combine to form individualized interventions that simultaneously provide drug treatment to drug abusing offenders and hold them accountable for their behavior.

In just two decades, drug courts have gone from a single court in one jurisdiction to an international movement with thousands of courts in operation. The most recent data indicate that there were over 2,400 drug courts in operation in the United States (Huddleston & Marlowe, 2011). Drug courts have also spread to other nations such as Canada, the United Kingdom, New Zealand, Australia, South Africa, Bermuda, and Jamaica (Berman & Feinblatt, 2005).

Not only have the number of drug courts increased, but also drug courts have increased in kind. Originally, drug courts were local courts that primarily served adult offenders with illicit substance abuse problems ("adult drug courts"). In recent years, however, drug courts for juvenile offenders and offenders charged with driving while under the influence (DWI) of alcohol have been opened and proliferated. Currently, there are 476 juvenile and 172 DWI drug courts in operation (Huddleston & Marlowe, 2011). Further, the drug court model also has begun to make inroads in federal and tribal jurisdictions (Huddleston et al., 2008).[3]

It is interesting, albeit common in the criminal justice system, that drug courts' initial expansion occurred without a solid body of empirical evidence establishing their effectiveness in reducing criminal behavior. In fact, an early review of the drug court literature conducted by the U. S. General Accounting Office (1997) concluded that the existing evidence was insufficient to draw any firm conclusion on the effectiveness of these programs with respect to recidivism. More specifically, the U. S. General Accounting Office (1997) identified several limitations of the 20 evaluations examined, including a failure to examine outcomes beyond program participation and a failure to use a comparison group design.

More recent reviews find stronger evidence of drug courts' effectiveness in reducing recidivism, but several issues prevent definitive conclusions. For example, Belenko (2001) reviewed 37 evaluations and drew a cautious but positive conclusion on the impact of drug courts on long-term drug use and criminal offending. Belenko, however, was critical of the dearth of evaluations that examined post-program drug use and other criminal behavior, noting that only six of the studies assessed the long-term effects of these programs. Further, Belenko noted that the number of evaluations of juvenile drug courts was too small to draw any conclusions about their effectiveness in reducing recidivism.

In 2005, the GAO updated its review of drug court evaluations (U. S. General Accountability Office, 2005). This review examined 27 evaluations of adult drug court programs that used comparison groups and this time concluded that the evidence indicates drug courts reduce recidivism in the period of time corresponding to the drug court treatment, but drug courts' effects on recidivism beyond this period and on drug use were questionable. While this systematic review is valuable, it didn't examine whether the effectiveness of drug courts varies systematically with methodological rigor and it only assessed the effectiveness of adult drug courts.

A larger systematic review conducted by Wilson et al. (2006) synthesized the findings of 55 evaluations of drug courts for adult and juvenile offenders. These authors tentatively concluded that drug court participants have lower rates of recidivism (drug and non-drug offending) than similar offenders who did not participate in drug courts. These findings held for evaluations that measured recidivism during and after program participation. Like the earlier reviews, these findings were tempered by the generally weak methodological rigor of the evaluations. These authors also noted that there were too few evaluations of juvenile drug courts to

draw conclusions about these courts effectiveness in reducing recidivism.

Most recently, Shaffer (2011) meta-analytically synthesized the results of 82 drug court evaluations. Shaffer's synthesis found that drug court participation was associated with a modest reduction in recidivism with a mean phi coefficient effect size of 0.09 (with a 95% C.I. of 0.08 to 0.10). This effect size translates into recidivism rates of 45.5% and 54.5% for drug court participants and non-participants, respectively. This review, however, did not examine whether these recidivism suppressing effects eroded over time, and it did not examine whether these effects varied systematically by type of drug court (i.e., adult, juvenile, or DWI drug court). Shaffer did examine the relationship between effect size and methodological rigor, but used a simple, dichotomous measure of methodological rigor that categorized evaluations as above or below the mean level of methodological rigor. Even with this measure of methodological rigor, Shaffer found that evaluations with high methodological rigor were associated with smaller reductions in recidivism.

Taken together, existing systematic reviews of drug court evaluations tentatively support the effectiveness of drug courts. Yet, many issues remain unresolved. First, it is still uncertain whether drug courts' recidivism suppressing effects are supported by the most methodologically rigorous evaluations. Second, the long-term effects of drug court participation are uncertain. And, third, it is unclear whether juvenile drug courts and DWI drug courts are effective in reducing recidivism. Fortunately, numerous new evaluations have been completed in recent years. The purpose of this research is to address these issues using this larger body of drug court evaluations.

Methodology

The population of evaluations eligible for this review was experimental and quasi-experimental evaluations of drug courts that utilized a comparison group. The criteria for inclusion were that: (a) the evaluation examined a drug court program[4]; (b) the evaluation included a comparison group that was treated in a traditional fashion by the court system (e.g., probation with or without referral to treatment)[5]; (c) the evaluations reported a measure of criminal behavior, such as arrest or conviction for some measurement period following the start of the program (the measure may have been based on official records or self-report); and (d) enough information was reported to compute an effect size.

The goal of the search strategy was to identify all evaluations, published or unpublished, meeting the above eligibility criteria. In order to achieve this objective, a multi-pronged search strategy was utilized. The search began by conducting a computerized keyword search of bibliographic databases. In particular, we conducted a search of the following databases: NCJRS, Criminal Justice Abstracts, Dissertation Abstracts, PsycINFO, Sociological Abstracts, Social Science Citation Index, Sciences Citation Index Expanded, Arts & Humanities Citation Index, Conference Papers Index, Ingentaconnect, C2 SPECTR, and CINAHL, as well as Google internet searches. The keywords used were: drug court (drug court*), DWI court (DWI court*), DUI court (DUI court*), evaluation, recidivism, re-arrest, and re-conviction. Each of the first three keywords was used in combination with each of the four latter terms.

We also searched for eligible evaluations by carefully reading retrieved studies and existing reviews of drug court research for unfamiliar evaluations. In particular, we reviewed the reference lists of existing reviews of drug court evaluations to identify eligible evaluations (Belenko, 2001; GAO, 2005; Latimer, Morton-Bourgon, & Chrtien, 2006; Marlowe et al., 2009; Shaffer, 2006, 2011). Likewise, many of the included evaluations reviewed prior drug court research. Unfamiliar studies referenced by eligible evaluations were also assessed for eligibility.[6]

Eligible studies were coded using coding forms that captured key features of the nature of the treatment, research participants, research methodology, outcome measures, as well as an effect size that coded the direction and magnitude of observed effects. Specifically, we utilized the odds-ratio effect size, as this type of effect size is the most appropriate effect size for dichotomous outcomes (Lipsey & Wilson, 2001). It is important to note that all effect sizes were coded in a manner such that *positive effect sizes indicate the treatment group had a more favorable outcome than the comparison group* (i.e., less recidivism). Last, two coders coded each study, and any discrepancies between coders were resolved by the lead author.

Our analyses of these effect sizes utilized the statistical approach outlined by Lipsey and Wilson (2001). In particular, we used the inverse variance method and assumed that the true treatment effects varied as a function of both measured (i.e., coded study features) and unmeasured differences between studies; that is, a random-effects model. Further, to determine which study features were associated with effect size we used meta-analytic analogs to analysis of variance and regression, assuming a mixed-effects model estimated via full-information maximum likelihood (Overton, 1998; Raudenbush, 1994).

This analytic strategy requires statistical independence (i.e., each effect size had to be coded from a unique research sample). We utilized several strategies to maintain the statistical independence. First, all evaluations (i.e., treatment/comparison contrasts) were cross-checked against one another to ensure that multiple studies reporting the results of the same evaluation do not contribute multiple estimates of program effects to any analysis. Second, in evaluations that report multiple measures of criminal behavior, we applied a set of selection criteria that created three data sets of effect sizes, with a particular evaluation contributing only one effect size to each of the data sets. In the first data set preference was given to effect sizes that: (1) were general (i.e., covered all offense types as opposed to being offense specific), (2) were based on arrest, (3) were dichotomous, and (4) followed sample members for 12 months.[7] We preferred effect sizes meeting these criteria, because such effect sizes were the most commonly reported outcome measure. The goal here is to calculate one effect size from each evaluation meeting these criteria to maximize the comparability of recidivism measures across evaluations.[8] Each independent evaluation contributed one, and only one, effect size to this "general recidivism" data set. This general recidivism data set served as the main data set in the analyses that follow.

We also created two other more specific data sets: one data set for measures of drug-related recidivism (e.g., arrest or conviction on drug charges), and another data set for measures of drug use (e.g., drug test results, self-reported drug use). In creating these data sets we had considerably fewer effect sizes to choose from. When multiple effect sizes were available for any of these data sets, preference was given to effect sizes that: (1) were more general (e.g., encompassed multiple types of non-drug offending, instead of one specific type of non-drug offending), (2) were dichotomous, and (3) followed sample members for 12 months. If an evaluation did not report one of these specific types of outcomes, then that evaluation did not contribute to the particular data set.

Results

Description of eligible studies

Our search found 370 potentially eligible studies. We retrieved 365 of the potentially eligible studies for further review. (The remaining five studies cannot be located.) Of the retrieved studies, 181 were eligible for this systematic review. Many of the eligible studies, however, utilized overlapping samples or the same sample in initial and follow-up evaluations of one drug court. These 181 studies yielded

154 independent evaluations of drug court programs. These 154 eligible and independent evaluations form the sample for this systematic review.

Table 1 provides descriptive information on the evaluations and drug courts examined in this review. Most of the drug court evaluations examined the effectiveness of adult drug courts. Ninety-two of the 154 evaluations (60%) assessed adult drug courts. Another 34 (22%) evaluations examined juvenile drug courts and the remaining 28 (18%) evaluations probed DWI (driving while intoxicated) courts. All but eight of the eligible evaluations examined U.S. drug courts. Four of the remaining evaluations examined adult drug courts in Australia. Two evaluations were of Canadian drug courts. One evaluation assessed a juvenile drug court in New Zealand, and another examined an adult drug court in Guam.

Based upon the Maryland Scientific Methods Scale (Farrington, Gottfredson, Sherman, & Welsh, 2002), evaluations were placed into four categories: (1) weak quasi-experiments, (2) standard quasi-experiments, (3) rigorous quasi-experiments, and (4) randomized experiments. Rigorous quasi-experiments typically used subject-level matching on key variables or propensity score matching. Standard quasi-experiments typically used either a historical comparison group that met drug court eligibility criteria constructed from archival data or a group of offenders who were eligible but not referred to the drug court program. The critical feature here is that the participants did not self-select into the drug court or comparison condition. Weak quasi-experimental designs typically involved comparing drug court clients to drug offenders who were eligible for participation in a drug court but declined participation ("refusers") or were referred to the drug court but were rejected by drug court administrators ("rejects"). Such designs have questionable internal validity because refusers and rejects are likely to differ on factors like pre-treatment motivation, perceived seriousness of drug problem, and self-efficacy, among many other potentially important factors. It's important to note that evaluations that used refusers or rejects as the comparison group but included efforts to minimize selection bias (e.g., controlled for many important predictors of recidivism) were given higher ratings depending on the nature of the efforts to minimize selection bias.

Overall, this body of literature is methodologically weak with few randomized evaluations of each type of drug court and only a modest number of rigorous quasi-experimental evaluations of adult drug courts and juvenile drug courts. Of the 92 evaluations of adult drug courts, only 3 (3%) were randomized experiments, another 20 (22%) were rigorous quasi-experiments; thus, approximately 25% of evaluations were categorized as relatively rigorous. Evaluations of juvenile and DWI drug courts were somewhat more rigorous with a greater proportion of evaluations categorized into the two most rigorous categories, 35% and 32% for juvenile and DWI drug courts, respectively. However, only one (3%) uncompromised randomized experimental evaluation of a juvenile drug court has been conducted and four (14%) such evaluations have been conducted on DWI drug courts Table 2.[9]

A common factor negatively affecting the methodological rigor of drug court evaluations was the comparability of the comparison group to the group receiving drug court treatment. The majority (56%) of adult drug court evaluations utilized a comparison group constructed from historical controls, clients who declined to participate in the drug court, or clients who were rejected by drug court administrators. These comparison groups allow for historical factors and selection bias to threaten the internal validity of the evaluations' results. Evaluations of juvenile and DWI drug courts were somewhat less likely to utilize these problematic comparison groups, 36% and 22%, respectively. Taken as a whole, the bulk of the existing evaluations are weak or standard quasi-experiments, which do not convincingly account for the possibility of selection bias and other threats to internal validity.

Another common concern among these evaluations is the short recidivism tracking periods utilized. Of the 92 evaluations of adult drug courts, 42 (46%) tracked recidivism of its sample for a maximum of 12 months or less. Likewise, 53% of juvenile drug court evaluations and 75% of DWI drug court evaluations tracked recidivism for a maximum of 12 months or less. Further, the recidivism tracking period most often overlapped completely or partially with the period of drug court treatment; this finding held for all three types of drug court evaluations.

In terms of programming, the three types of drug courts differed on several components. Adult drug courts were more likely to use the pre-plea method of disposition, dismiss charges upon graduation, and have three or fewer phases than juvenile and DWI drug courts. Typically, DWI drug courts took longer to complete than adult and juvenile drug courts, but the majority of DWI drug court participants graduated whereas less than the majority of participants in other kinds of drug courts graduated.

The evaluations exhibited relatively little variation in terms of sample characteristics. Overwhelmingly, the samples used in these evaluations were composed predominantly of males. The vast majority of courts restricted eligibility to non-violent offenders. Approximately 20% of samples had minor criminal history; this percentage was highest in evaluations of adult drug courts and lowest in evaluations of juvenile drug courts.

Mean effects by type of drug court

We calculated effect sizes that measure drug courts' effects on general recidivism, drug related recidivism, and drug use for each type of court. Table 3 displays the mean odds-ratio for each court

Table 1
Key features of evaluations

Variable	Adult Drug Court ($k^a = 92$) Frequency (%)	Juvenile Drug Court ($k = 34$) Frequency (%)	DWI Drug Court ($k = 28$) Frequency (%)
Methodological rigor rating			
Weak quasi-experiment	18 (20%)	4 (12%)	3 (11%)
Standard quasi-experiment	51 (55%)	17 (50%)	7 (25%)
Rigorous quasi-experiment	20 (22%)	11 (32%)	5 (18%)
Random experiment	3 (3%)	1 (3%)	4 (14%)
No information/unclear	0 (0%)	1 (3%)	9 (32%)
Type of comparison group used			
Declined/rejected	28 (30%)	5 (15%)	1 (4%)
Historical controls	24 (26%)	7 (21%)	5 (18%)
Eligible non-referred	9 (10%)	0 (0%)	1 (4%)
Comparable (randomization)	3 (3%)	3 (9%)	5 (18%)
Regular probation	15 (16%)	10 (29%)	2 (7%)
Other Non-eligible drug offenders	9 (10%)	1 (3%)	5 (18%)
Not reported/can't tell	4 (4%)	2 (6%)	9 (32%)
Overall attrition > 20%			
Yes	4 (4%)	4 (12%)	3 (11%)
No	87 (95%)	28 (82%)	25 (89%)
No information/unclear	1 (1%)	2 (6%)	0 (0%)
Differential attrition > 20%			
Yes	5 (4%)	3 (8%)	2 (7%)
No	87 (95%)	30 (88%)	26 (93%)
No information/unclear	1 (1%)	1 (3%)	0 (0%)
Max length of follow-up			
12 months or less	42 (46%)	18 (53%)	21 (75%)
12.01-24 months	23 (25%)	8 (24%)	2 (7%)
25.01-36 months	6 (7%)	2 (6%)	0 (0%)
36+ months	8 (9%)	4 (12%)	2 (7%)
No information/unclear	13 (14%)	2 (6%)	1 (4%)
Follow-up overlap with treatment period			
Complete overlap	26 (28%)	7 (21%)	5 (17%)
Partial overlap	46 (50%)	7 (21%)	14 (50%)
No overlap	17 (18%)	9 (26%)	8 (29%)
No information/Unclear	3 (3%)	11 (32%)	1 (4%)

[a] Number of evaluations.

Table 2
Key features drug courts and clients

Variable	Adult Drug Court ($k=92$) Frequency (%)	Juvenile Drug Court ($k=34$) Frequency (%)	DWI Drug Court ($k=28$) Frequency (%)
Method of disposition			
Pre-plea	21 (23%)	0 (0%)	0 (0%)
Post-plea	36 (39%)	18 (53%)	11 (39%)
Uses Both	13 (14%)	1 (3%)	1 (4%)
Not reported	22 (24%)	15 (44%)	16 (57%)
Charges Dismissed Upon Graduation			
Yes	34 (37%)	8 (24%)	1 (4%)
No	19 (21%)	3 (9%)	12 (43%)
Not reported	39 (42%)	23 (68%)	15 (57%)
Number of Phases			
Three or less	45 (49%)	7 (3%)	3 (11%)
More than three	21 (23%)	17 (50%)	10 (36%)
Doesn't use phases/Not reported	26 (28%)	11 (32%)	15 (54%)
Number of status hearings/month[a]			
Two or less	19 (21%)	7 (21%)	8 (29%)
More than two	16 (17%)	3 (9%)	0 (0%)
Not reported	57 (62%)	24 (71%)	20 (71%)
Minimum time to graduation			
Less than 12 months	20 (22%)	17 (50%)	4 (14%)
12 to 15 months	46 (50%)	13 (38%)	8 (29%)
More than 15 months	12 (13%)	0 (0%)	5 (18%)
Not reported	14 (15%)	4 (12%)	11 (39%)
Graduation rate[b]			
.00 to .25	20 (22%)	17 (50%)	4 (14%)
.26 to .50	46 (50%)	13 (38%)	8 (29%)
.51 to .75	12 (13%)	0 (0%)	5 (18%)
More than .75	14 (15%)	4 (12%)	11 (39%)
Not reported	20 (22%)	17 (50%)	4 (14%)
Single treatment provider			
Yes	15 (16%)	3 (9%)	0 (0%)
No	28 (30%)	8 (24%)	5 (18%)
Not reported	49 (53%)	23 (68%)	23 (82%)
Gender composition			
All male (90%+male)	1 (1%)	2 (5%)	1 (4%)
Mostly male (60-90% male)	77 (84%)	32 (94%)	19 (68%)
Approx. equal (59%-40% male)	7 (8%)	0 (0%)	0 (0%)
Not reported	7 (8%)	0 (0%)	8 (29%)
Offender type			
Only non-violent offenders	72 (79%)	18 (53%)	23 (82%)
Includes violent offenders	16 (17%)	4 (12%)	0 (0%)
Not reported	4 (4%)	12 (35%)	5 (18%)
Minor criminal history			
Yes	22 (24%)	2 (6%)	2 (7%)
No	48 (52%)	17(50%)	14 (43%)
Not reported	22 (24%)	15 (44%)	12 (50%)

[a] In first treatment phase of drug court.
[b] Graduation rate is calculated as proportion of *terminated* clients who completed program successfully (i.e., excludes those currently active in program).

Table 3
Mean random-effects odds-ratio by type of recidivism measure

Outcome	Mean ES	95% Confidence Interval Lower	Upper	Q	k	Tau^2
Adult drug courts						
General recidivism[a]	1.66*	1.50	1.84	442.19*	92	0.178
Drug recidivism[b]	1.70*	1.39	2.08	323.98*	42	0.368
Drug use	1.45	0.92	2.28	15.78*	4	0.165
Juvenile drug court						
General recidivism	1.37*	1.15	1.63	66.31*	34	0.105
Drug recidivism	1.06	0.69	1.63	29.65*	14	0.357
Drug use	1.50	0.67	3.34	2.05	3	0.359
DWI drug court						
General recidivism[c]	1.65*	1.35	2.02	78.40*	28	0.159
Drug recidivism[d]	1.59*	1.22	2.09	16.84	14	0.054
Drug use	1.87	0.34	10.23	5.02*	2	1.227

[a] The mean effect size is 1.57 (95% C.I. 1.43-1.72), when *three* large positive effect sizes were removed.
[b] The mean effect size is 1.46 (95% C.I. 1.28-1.67), when *two* large positive effect sizes were removed.
[c] The mean effect size is 1.63 (95% C.I. 1.33-1.99), when *one* large positive effect sizes was removed.
[d] The mean effect size is 1.57 (95% C.I. 1.20-2.04), when *one* large positive effect sizes was removed.
$^{+}$ $p<0.10$; * $p<0.05$.

type (adult, juvenile, and DWI) by outcome type. The forest plots for the general recidivism effects are shown in Figs. 1 through 3. The forest-plots show a clear pattern of evidence favoring drug courts with most evaluations observing effects favoring the drug court (88%, 70%, and 85%, for adult, juvenile, and DWI, respectively). The general recidivism measure is both statistically significant for all three court types and moderate to small in size (mean odds-ratio of 1.66, 1.37, and 1.65, for adult, juvenile, and DWI, respectively). Relative to a 50% recidivism rate in the control group (a typical value), these translate into recidivism rates for the drug court of 37.6%, 42.2%, and 37.7%.

The effects of these courts on drug related recidivism are very similar for adult and DWI drug courts with random effects odds-ratios of 1.70 and 1.65, respectively. However, for juvenile drug courts, the results on drug related recidivism outcomes were less encouraging. The mean odds-ratio was 1.06, which is essentially null. Thus, this evidence raises the possibility that juvenile drug courts are not effective in reducing drug related crime or that any effect that is produced is small.

Few evaluations assessed the effect of drug court participation on measures of actual drug use (i.e., urinalysis or self-reported drug use). We found only nine independent evaluations that reported useable post-program entry measures of drug use for both drug court participants and non-participants. Four of the nine effect sizes assessed the effect of participation in adult drug court programs, three were from evaluations of juvenile drug courts, and the two remaining effect sizes were from evaluations of DWI drug courts. For each type of court, the average effect size was positive, indicating reduced drug use for drug court participants in comparison to non-participants. However, because of the small number of drug use effect sizes, the average effect size for each type of court was not statistically significant.

Although these mean effect sizes indicate that drug court participants have, on average, lower rates of general recidivism than non-drug court participants, these findings need to be interpreted within the context of the methodological rigor of the evaluations. Table 4 displays the mean effect size by methodological features and type of drug court. For brevity's sake, this table reports only the mean random effects odds ratios for the general recidivism measure.

Of concern in these tables are the smaller effects for the more rigorous designs. Focusing on the adult courts, we find the smallest effects for the experimental designs. For adult drug courts, the mean odds-ratio for experimental designs is not statistically significant across the three experimental evaluations. Given the importance of these experimental evaluations, these findings deserve greater exploration. The three randomized experimental evaluations of adult drug courts included in this synthesis evaluated: (1) a drug court in Maricopa County (Deschenes, Turner, & Greenwood, 1995; Turner, Greenwood, Fain, & Deschenes, 1999); (2) Baltimore City's drug court (Gottfredson, Najaka, & Kearley, 2003; Gottfredson, Najaka, Kearley, & Rocha, 2006); and, (3) a drug court in New South Wales, Australia (Shanahan et al., 2004). All three evaluations reported findings on the effectiveness of drug court participation on measures of general and drug related offending in the *first 12 months after program entry*, which are the effects included in our primary analyses. The results for these experimental evaluations are inconsistent (heterogeneous), with two of the three evaluations finding modest positive results (odds-ratios of 1.65 and 1.82), and one evaluation (Deschenes et al., 1995) with a near null effect (1.06) on the general recidivism outcome measure. Thus, the finding that experimental evaluations of adult drug courts do not collectively find statistically

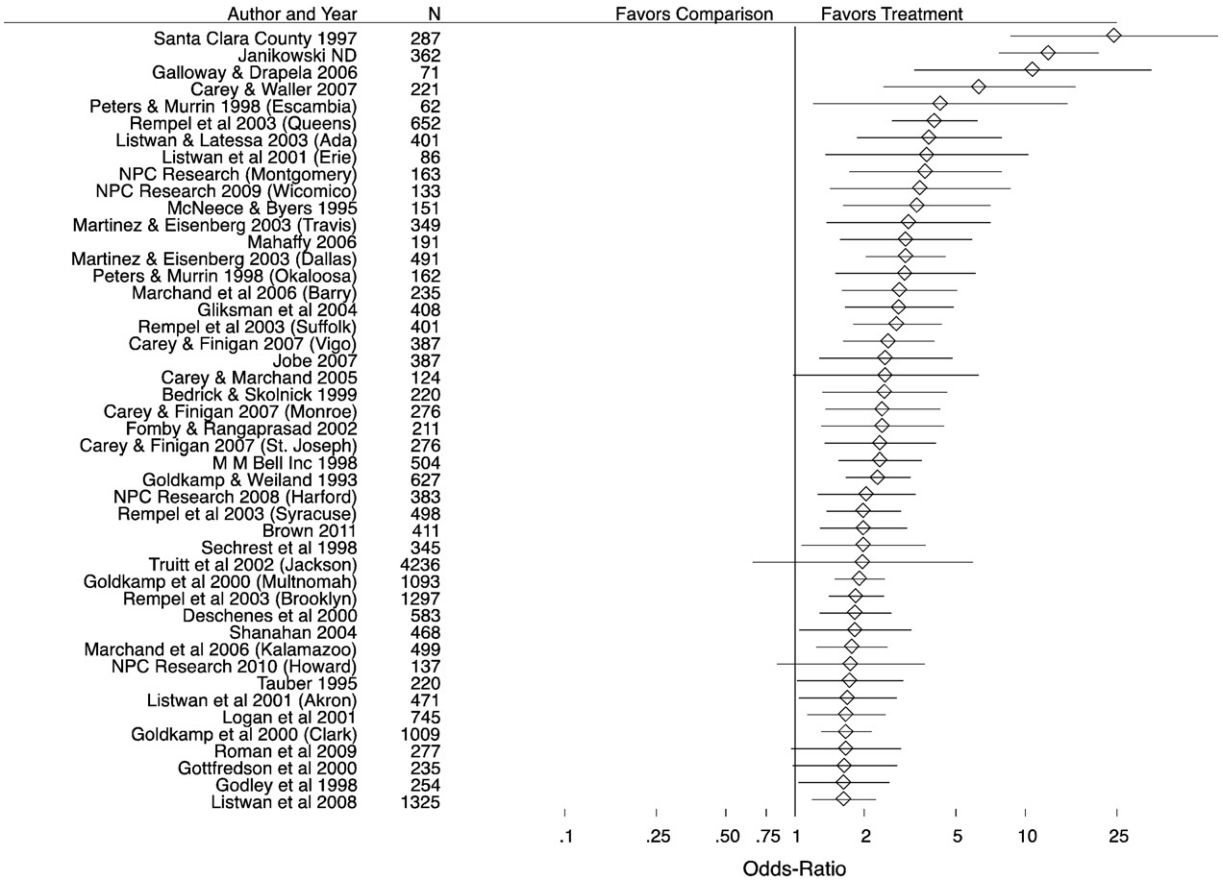

Fig. 1. Forest plot of general recidivism effect sizes for adult drug courts.

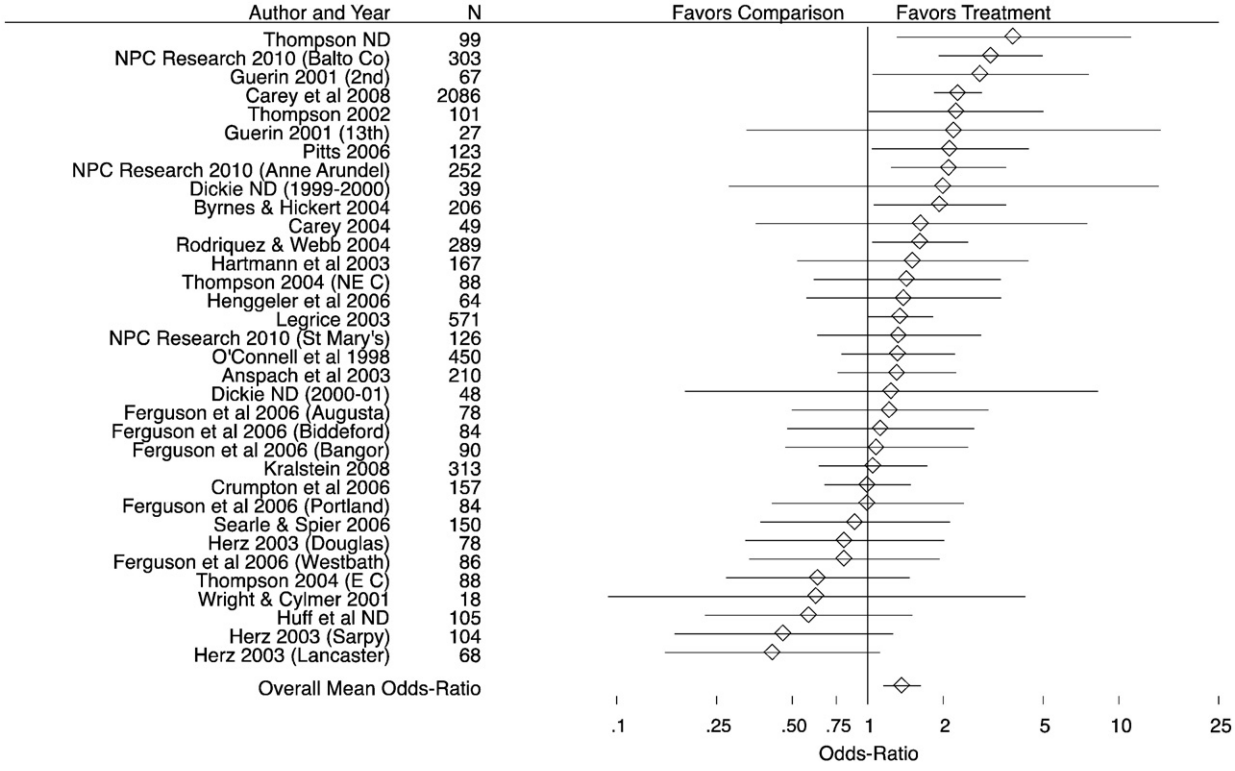

Fig. 2. Forest plot of general recidivism effect sizes for juvenile drug courts.

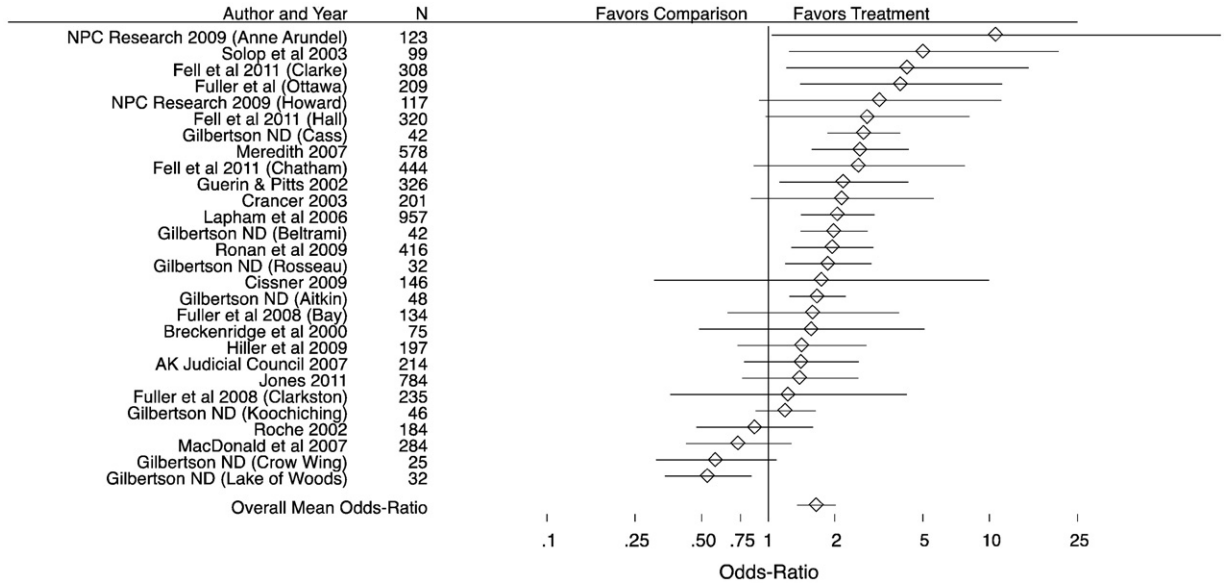

Fig. 3. Forest plot of general recidivism effect sizes for DWI drug courts.

significant reductions in recidivism is driven by inconsistent results across the three evaluations and the small number of experimental evaluations (i.e., low statistical power), which is evidenced by the rather large confidence intervals around the mean effect sizes.

All three experimental evaluations have unique characteristics; however, the Maricopa County evaluation was particularly unusual, in that the drug court participants were compared to a control group of offenders involved in a drug-testing program. In fact, the drug-testing control group was drug tested more often than the drug court group—a highly unusual finding in this body of research. This fact, among others, makes this evaluation problematic. We find that if this evaluation is removed for the analysis, the mean odds-ratio for experimental evaluations on the general recidivism outcome measure is 1.73 (95% C.I. of 1.18 to 2.53), which is statistically significant and larger than the effect for the rigorous quasi-experimental designs. Another complication with the evaluation of the Maricopa County program is that the results changed substantively in a follow-up evaluation that reported outcomes *three-years after program entry* (Turner et al., 1999). In this subsequent evaluation, participants in the Maricopa County drug court program had significantly less recidivism than the control group. If the effect size from this subsequent evaluation is used, then the mean odds-ratio for experimental evaluations is 1.65 (95% C.I. of 1.25 to 2.18), which is statistically significant and virtually identical to the mean effect size for all evaluations of adult drug courts. Thus, the results of the Maricopa County evaluation materially affect the size and statistical significance of the mean odds-ratio for experimental evaluations.

We draw three conclusions from the above analyses. First and foremost, all three experimental evaluations of adult drug courts find evidence of these courts' effectiveness in reducing recidivism. Two of the three evaluations find recidivism reductions in the first year after program entry, and the remaining evaluation finds recidivism reductions at three years post-program entry. Second, because the vast majority of adult drug court evaluations, even the most rigorous evaluations, find modest reductions in general recidivism, we believe that the evidence indicates that adult drug courts reduce recidivism. Third, the mean odds-ratio effect size measuring these courts effect on general recidivism appears to be approximately 1.65, which translates into an average recidivism rate of 38% for drug court participants, if we assume a 50% recidivism rate for non-participants.

For juvenile drug courts, we find that these courts have small effects on recidivism, especially in methodologically rigorous evaluations. The strongest evidence of the effectiveness of juvenile drug courts comes from weak quasi-experimental evaluations, as in these evaluations the general recidivism mean odds-ratio is relatively large (1.85) and statistically significant. However, the general recidivism mean odds-ratios are considerably smaller in evaluations with higher levels of methodological rigor. For standard quasi-experimental evaluations, the mean odds-ratio is 1.32 (95% C.I. of 1.07 to 1.62). The mean odds-ratios for strong (rigorous) quasi-experimental and experimental evaluations are similar in magnitude 1.22 and 1.39, respectively[10]; neither is statistically significant. These findings indicate that evaluations of juvenile drug courts reduce general recidivism, but the magnitude of these effects is smaller than that of adult drug courts. Thus, the most rigorous evaluations of juvenile drug courts indicate that these courts have small effects on recidivism and the magnitude of these effects is approximately an odds-ratio of

Table 4
General recidivism odds ratios by type of court and methodological Rigor rating

	Adult Drug Court		Juvenile Drug Court		DWI Drug Court	
	Mean (95% C.I.)	*K*	Mean (95% C.I.)	*k*	Mean (95% C.I.)	*k*
Methodological Rigor Rating	$Q_B^a = 2.35$		$Q_B = 3.14$		$Q_B = 21.76^*$	
Weak	1.97 (1.54-2.54)	18	1.85 (1.26-2.72)	4	1.99 (1.49-2.65)	3
Standard	1.63 (1.40-1.89)	51	1.32 (1.07-1.62)	17	2.56 (1.66-3.96)	7
Strong	1.37 (1.25-1.97)	20	1.22 (0.90-1.65)	11	1.99 (1.49-2.65)	5
Experimental	1.45 (0.80-2.62)	3	1.39 (0.50-3.85)	1	1.15 (0.79-1.67)	4

[a] $Q_B = Q$ between, which measures variation between category means in the meta-analytic analog to ANOVA.
$^+$ $p<0.10$; * $p<0.05$.

1.30, which translates into a 43.5% recidivism rate for drug court participants, if we assume a 50% recidivism rate for non-participants.

The substantive findings regarding the effectiveness of DWI drug courts are similar to those for adult drug courts. Just as with evaluations of adult drug courts: (1) the largest effects of DWI drug courts are found in methodologically weak evaluations, (2) collectively, experimental evaluations find small and non-statistically significant mean odds-ratios for general and drug related recidivism; and, (3) the mean odds-ratios for experimental evaluations are greatly influenced by one evaluation that found a negative effect of participation. In regards to the first point, the mean odds-ratio is moderate and statistically significant for the three quasi-experimental design categories, but relatively small and non-statistically significant for the four evaluations in the experimental category (1.27 with 95% C.I. of 0.87 to 1.85). Thus, quasi-experimental evaluations yield substantively different findings than experimental evaluations of DWI drug courts. In regards to the second point, experimental evaluations find a small and non-statistically significant difference in general recidivism between participants and non-participants. And finally, additional analysis of the four experimental DWI drug court evaluations reveals that one such experimental evaluation conducted by MacDonald, Morral, Raymond, and Eibner (2007) heavily influences the mean odds-ratio for the experimental evaluations. While the three other experimental evaluations of DWI courts found positive odds-ratios on the general recidivism measure ranging from 1.39 to 2.25, MacDonald et al.'s evaluation found a negative effect (odds-ratio of 0.73). If MacDonald et al.'s evaluation is omitted from the general recidivism analysis, then the mean odds-ratio becomes 1.58 (95% C.I. of 0.99 to 2.54) and has a *p*-value of 0.057. Thus, the MacDonald et al. evaluation is highly influential on the results of the experimental evaluations.

The evidence presented above finds considerable evidence of the effectiveness of drug courts, but the strength of this evidence varies by type of court. In regards to adult drug courts, over 90 independent evaluations have been conducted and the overwhelming majority of these evaluations find that drug court participants have less recidivism than non-participants. Further, experimental evaluations of these courts consistently find sizeable reductions in recidivism. Thus, the evidence indicates that adult drug courts are effective in reducing recidivism. Likewise, we characterize the evidence as cautiously supporting the effectiveness of DWI drug courts, because while quasi-experimental evaluations find strong and consistent indications that these programs reduce general and drug related recidivism, randomized experimental evaluations find a small, non-statistically significant reduction in recidivism. Yet, the findings from experimental evaluations of DWI drug courts are ambiguous in that the majority of these evaluations find positive effects but a single, influential evaluation with negative findings heavily influences the mean effect. Clearly, only additional evaluations using experimental methods can definitively resolve the remaining ambiguity surrounding the effectiveness of DWI drug courts. Evaluations of juvenile drug courts, especially more rigorous evaluations, consistently indicate that these courts have relatively small effects on recidivism.

Drug courts' long-term effects

An important issue in drug court research is whether the effects last long-term. Assessing drug courts' long-term effect, however, is challenging. There are two inter-related issues. The first is whether the observed pattern of positive results reflects a suppression effect. Many of the outcomes are examined during the course of drug court participation. It is possible that drug courts suppress offending behavior while someone is active in the program but that this effect disappears post-program once behavioral contingencies are removed. The second issue is simply whether observed effects continue long-term, such as three years after post-program.

We examined the first issue of suppression by coding whether the recidivism tracking period overlapped: (1) completely, (2) partially, or (3) not at all with the period of drug court participation. If drug courts' effects on recidivism are limited to the period in which participants are active in the court, then the mean effect size should be largest when the treatment and recidivism tracking periods overlap completely and the mean effect size should decrease as the amount of overlap between the treatment and recidivism tracking periods decreases. Table 5 displays the mean effect size for each of these categories and shows that the effects of drug courts remain post-program. That is, the positive results do not appear to be simply a temporary suppression effect.

We examined the second issue by computing mean effect sizes for different follow-up periods. Recall that evaluations most commonly reported recidivism rates 12 months after drug court entry (or termination). When effect sizes measuring recidivism at multiple time points were available, we preferred effect sizes based on this most common time period (i.e., 12 months) to facilitate between study comparisons. However, not all evaluations measured recidivism at 12-months and some evaluations reported recidivism at multiple time-points (e.g., 12, 24, and 36-months). Consequently, there is both within and between evaluation variation in the length of recidivism tracking period. We exploit both sources of variation. Between study variation was examined by calculating the mean effect size by

Table 5
Mean random effects odds ratios of traditional drug courts over time

	General Recidivism			Drug Recidivism		
Between evaluation comparisons	Mean	95% C.I.	*k*	Mean	95% C.I.	*k*
Recidivism measure overlaps with drug court	$Q_B = 1.66$			$Q_B = 1.77$		
Complete overlap	1.57	1.27-1.94	26	1.35	0.71-2.56	8
Partial overlap	1.77	1.51-2.06	47	2.04	1.42-2.94	26
No overlap	1.49	1.16-1.92	18	1.41	0.74-2.69	8
All effect sizes by follow-up length	$Q_B = 1.33$			$Q_B = 6.39^{+}$		
12 months or less	1.62	1.37-1.92	43	1.50	0.92-2.46	15
12.01-24 months	1.65	1.31-2.07	26	1.56	0.62-3.93	4
24.01-36 months	1.70	1.09-2.66	6	3.72	1.92-7.20	8
36+ months	2.22	1.53-3.22	8	1.67	0.84-3.31	7
Evaluations that measure recidivism at 12 and 24 months	Not tested			Not tested		
12 months	1.74	1.40-2.17	21	–––	–––	–
24 months	1.66	1.38-1.98	21	–––	–––	–
Evaluations that measure recidivism at 12, 24, and 36 months	Not tested			Not tested		
12 months	1.71	1.18-2.48	8	–––	–––	–
24 months	1.72	1.29-2.29	8	–––	–––	–
36 months	1.80	1.44-2.24	8	–––	–––	–

$^{+}$ $p<0.10$; * $p<0.05$.

length of recidivism tracking period. These findings (Table 5) show a roughly stable mean odds-ratio for the different follow-up periods. A complication with this analysis is the possible confounding of study features with follow-up length (e.g., evaluations reporting third-year recidivism measures may be more rigorous than those reporting first-year measures). To address that issue, we also examined within study variation by analyzing the subset of evaluations that reported results at both the 12 and 24 month follow-up period (21 studies) and the subset of evaluations that reported results at the 12, 24, and 36 month follow-up period (8 studies). As shown in Table 5, drug court effects remain remarkable stable over time with only a very slight decrease from 12-months through 36-months.

The analyses presented above support the conclusion that any effect adult drug courts have on recidivism is not limited to the short-term. Rather, the available research suggests that adult drug court participants have reduced recidivism during and after drug court treatment, and these effects appear to last at least three years post-drug court entry.

Features of the drug court

An important question concerns why some drug courts are more effective in reducing recidivism than others. Longshore et al. (2001) provide a useful conceptual framework for thinking about variation in drug courts. They hypothesize that the most effective drug courts: (1) use the courts' *leverage* (rewards and sanctions) to motivate offender change; (2) serve *populations* with less severe problems; (3) have high *program intensity*; (4) apply rewards and sanctions *predictably*; and, (5) emphasize offender *rehabilitation* as opposed to other court goals like speedy case processing. While a full test of this framework is beyond the scope of this research, we were guided by Longshore et al.'s predictions in coding relevant drug court features. Specifically, based on Longshore et al.'s work, we expected that drug courts with greater incentives for success are more effective in reducing recidivism. For example, courts that process cases using the post-plea method have greater leverage because offenders already have been convicted and face immediate sentencing upon failure in drug court. Likewise, drug courts that dismiss or expunge charges/convictions have greater leverage as successful offenders are able to minimize some of the collateral consequences of conviction. We also expected that courts with greater treatment intensity (e.g., more status hearings, longer minimum times to graduation) produce greater reductions in recidivism. Longshore et al. predict that courts serving less serious offenders produce larger reductions in recidivism than other courts; we tested this expectation by coding several features capturing population severity.

We used the coded moderator variables and the meta-analytic analog to ANOVA to test these predictions. The results of these analyses are presented in Table 6. Note that because of the relatively small number of evaluations of juvenile and DWI drug courts, these analyses are limited to effect sizes coded from adult drug courts[11]; and because several evaluations did not report the information of interest, missing data reduces the number of available effect sizes.

Few of these moderator variables revealed statistically significant relationships. In regards to the variables tapping court *leverage*, courts that dismissed/expunged charges upon graduation had larger effect sizes than other courts, but this was only statistically significant for the drug recidivism outcome. Likewise, for program *intensity*, only one measure was statistically associated with effect size; specifically, courts that had more than two status hearings per month in the earliest treatment phase (the typical number of hearings) had larger reductions in drug related recidivism than other courts, but this finding

Table 6
Mean random effects odds ratios of traditional drug courts by drug court features

	General Recidivism			Drug Recidivism		
Variable	Mean	95% C.I.	*k*	Mean	95% C.I.	*k*
Method of Disposition	$Q_B = 0.40$			$Q_B = 5.50^{+}$		
Pre-plea	1.74	1.37-2.19	22	1.52	0.89-2.60	13
Post-plea	1.60	1.35-1.89	38	1.52	0.91-2.54	13
Uses both	1.57	1.19-2.06	13	4.51	1.97-10.33	6
Charges dismissed upon graduation	$Q_B = 0.87$			$Q_B = 6.18^{*}$		
Yes	1.65	1.40-1.95	36	1.54	1.35-1.76	19
No	1.50	1.21-1.86	20	1.21	0.97-1.49	7
More than 3 phases	$Q_B = 0.29$			$Q_B = 3.18^{+}$		
Yes	1.60	1.29-2.00	21	1.07	0.51-2.21	8
No	1.69	1.45-1.96	50	2.06	1.38-3.06	25
More than 2 status hearings/month	$Q_B = 0.05$			$Q_B = 28.58^{*}$		
Yes	1.57	1.22-2.03	16	2.26	1.64-3.11	4
No	1.54	1.23-1.92	19	1.37	1.10-1.70	7
Minimum time to graduation	$Q_B = 1.07$			$Q_B = 3.03$		
Less than 12 months	1.58	1.29-1.94	22	1.54	0.82-2.86	9
12-15 months	1.77	1.54-2.03	47	2.03	1.41-2.92	25
More than 15 months	1.61	1.22-2.12	12	0.78	0.24-2.55	3
Offender type	$Q_B = 7.38^{*}$			$Q_B = 0.03$		
Non-violent only	1.68	1.51-1.86	70	1.63	1.19-2.24	34
Includes violent offenders	1.25	1.03-1.52	16	1.74	0.93-3.26	8
Minor criminal history	$Q_B = 2.91^{+}$			$Q_B = 0.15$		
Yes	1.80	1.48-2.20	23	1.52	0.75-3.07	8
No	1.51	1.33-1.72	50	1.75	1.18-2.60	27
Graduation rate	$Q_B = 12.79^{*}$			$Q_B = 10.18^{*}$		
.00-.25	1.91	1.38-2.64	11	1.40	0.98-2.01	5
.26-.50	1.41	1.20-1.64	39	1.29	1.06-1.57	17
.51-.75	2.17	1.65-2.87	14	2.48	1.48-4.17	2
Single treatment provider	$Q_B = 0.00$			$Q_B = 0.10$		
Yes	1.55	1.18-2.04	16	1.48	0.98-2.25	6
No	1.56	1.29-1.87	30	1.56	1.25-1.95	12
Publication Status	$Q_B = 0.19$			$Q_B = 0.51$		
Published	1.63	1.29-2.07	21	1.97	1.23-3.16	15
Unpublished	1.68	1.48-1.91	71	1.59	1.11-2.27	28

$^{+}$ $p<0.10$; * $p<0.05$.

was not replicated in the general recidivism analysis. Interestingly, for general recidivism, the analyses revealed that courts serving less *severe* clients (i.e., exclusively non-violent offenders and clients with minor criminal history) had statistically larger effect sizes than other courts. Overall, the fact that few of these moderator variables predicted effect size suggests that adult drug courts generally work and this effectiveness is largely robust to programmatic variations.

A common finding in drug court evaluations is that graduates are much less likely to recidivate than non-graduates. Based on this finding, we reasoned that courts with higher graduation rates have lower recidivism rates. Our analyses partially support this hypothesis; in that, the programs with the highest graduation rates had the largest effect sizes. However, contrary to our hypothesis, we find the smallest effect sizes among courts with graduation rates in the middle range—not the lowest graduation rates.

Additional sensitivity analyses

The majority of the evaluations included in this review were unpublished technical reports produced by government or private research entities. Such a large proportion of unpublished evaluations reduces the likelihood of publication bias affecting our estimates of drug courts' effectiveness. As shown in Table 6, the results for published and unpublished evaluations of adult drug courts were roughly similar. This finding is also true for juvenile and DWI courts (not shown in tables). As a test for other forms of publication selection bias, such as outcome selection bias, we performed the Duval and Tweedie's trim-and-fill analysis for each court type. The conclusions discussed above were robust even under the trim-and-fill model that tends to over-fill when there is substantial heterogeneity, as was the case here.

We also performed sensitivity analyses to determine the effect on our results of allowing a few authors to contribute many effect sizes to our analyses. Specifically, we flagged authors who contributed 10% or more of the effect sizes to any of the court specific analyses, and then re-ran the analysis without these effect sizes. For adult drug courts, only NPC research contributed 10% or more of the effect sizes ($k=15$). While the effect sizes from NPC had a larger mean odds-ratio effect size (1.94) than other evaluations (1.63), this difference was not statistically significant and the mean odds-ratio effect size for the other evaluations is virtually identical to the mean effect size with the NPC evaluations included. We performed similar analyses with the juvenile and DWI drug court effect sizes. These sensitivity analyses revealed that our results are substantively unchanged by multiple effect sizes from the same author(s).

Last, we assessed the applicability of our findings beyond the United States by examining the eight international evaluations. Seven of the eight evaluations examined adult drug courts (1 in Guam, 2 in Canada, and 4 in Australia). Six of these seven evaluations had positive general recidivism effect sizes, four of which were moderate in size (odds ratio greater than 1.60), and the mean general recidivism odds-ratio is 1.62 with a 95% confidence interval of 1.11 to 2.36, which is statistically significant. Notice that this mean general recidivism odds-ratio is virtually identical to the mean odds-ratio reported above for all adult drug court evaluations (1.66). Only one international evaluation of a juvenile drug court met our eligibility criteria; this evaluation assessed a juvenile drug court in New Zealand (Searle & Spier, 2006). This evaluation found a negative effect indicating that the juvenile drug court did not reduce recidivism. These international evaluations mirror our findings on adult and juvenile drug courts. Thus, our findings appear to accurately reflect existing evaluations of drug courts outside the United States.

Conclusions

Our results indicate that drug participants have lower recidivism than non-participants, but the size of this effect varies by type of drug courts. For adult drug courts, the average effect of participation is equivalent to a reduction in general recidivism from 50% to approximately 38% and a reduction in drug-related recidivism from 50% to approximately to 37%. And these reductions in recidivism persist for at least three years after program entry. Thus, the accumulated evidence suggests that adult drug courts are effective in reducing recidivism and the policy implication of this conclusion is that continued funding, development, and operation of adult drug courts is warranted.

For DWI drug courts, the magnitude of their effects on recidivism are comparable to those of the adult drug courts. Yet, because of the ambiguous findings of the most rigorous, randomized experimental evaluations, we believe that additional experimental evaluations of DWI courts are needed to more definitively demonstrate their effectiveness.

In contrast to the moderate effects of adult and DWI drug courts, juvenile drug courts have relatively small effects on recidivism. Our estimates indicate that the average effect of participation in a juvenile drug court is equivalent to a reduction in recidivism from 50% to approximately 43.5%. This average effect is more than 40% smaller than the average estimated effects of participation in an adult or DWI drug court. The question becomes: Why are juvenile drug courts less effective than other kinds of drug courts? Obviously, we cannot answer this question with certainty, yet two factors seem relevant. First, juvenile drug courts generally provide services to relatively high-risk offenders, whereas other kinds of drug courts typically exclude high-risk offenders. Second, juvenile drug courts appear to be less demanding interventions than adult drug courts, in that, drug testing and status hearings appear to be less frequent, and the period of program participation appears to be shorter in duration.

Beyond these general conclusions about the effectiveness of drug courts, it is important to emphasize that the estimated effects of drug court participation are highly variable. We attempted to explore the sources of this variability by applying Longshore et al.'s drug court framework. We found some evidence supporting the importance of *leverage* and *intensity* in that drug courts that dismissed charges and courts with more frequent status hearings had larger reductions in recidivism than other courts, but these findings were only statistically significant in the analysis of drug related recidivism. We believe that more primary research comparing the effectiveness of drug courts with varying features needs to be conducted to confirm these meta-analytic findings.

In support of Longshore et al.'s conception framework of drug courts, we find that programs with *less severe* populations are more effective in reducing general recidivism. Specifically, evaluations of programs that only allowed non-violent offenders to participate had larger reductions in general recidivism than other program evaluations. This finding holds in various sensitivity analyses. Because this finding conflicts with the risk principle from Andrews et al.'s well-known principles of effective intervention (see e.g., Andrews & Bonta, 1992), it is sure to be met with considerable controversy. Yet, given the strength and consistency of this finding, we believe that this finding deserves consideration and further empirical scrutiny.

Further, our finding that courts with violent offenders are less effective in reducing general recidivism seems to contradict findings from other drug court researchers who have found that violent offenders perform as well in drug courts as non-violent offenders (see e.g., Rossman et al., 2011; Saum, Scarpitti, & Robbins, 2001). Closer inspection, however, reveals that such findings examine a *different unit of analysis* than our meta-analytic research. In essence, these researchers' work concerns the recidivism of *individuals* with evidence of prior violence in comparison to non-violent drug court participants; whereas, our meta-analytic findings concern the relative reduction in recidivism between *courts* that allow violent drug court clients in comparison to other courts. These two questions are distinct and the answers to these questions need not match. As an example, consider the research of Saum et al. (2001), who as previously mentioned

found that drug court participants with evidence of prior violence exhibited comparable reductions in recidivism as non-violent drug court participants. This study examines the individual-level of analysis. At the court evaluation-level of analysis, we find that the court evaluated in Saum et al'.s research (coded here under Scarpitti, Butzin, Saum, Gray, & Leigey, 2005) had relatively small effects on recidivism in comparison to other drug court evaluations; in fact, this evaluation found that participants had more recidivism than non-participants. Rather than contradicting our finding, the results of Saum et al. buttress our conclusion that courts that accept violent offenders are less effective than other courts. In short, it is entirely possible that both sets of findings are correct; violent drug court participants do as well as non-violent participants in drug courts, and courts that accept violent offenders are less effective than other courts.

While our finding regarding the inclusion of violent offenders will be controversial, it is important to note that this finding supports not only Longshore and colleagues' theoretical perspective but also federal regulations that require drug courts to restrict eligibility to non-violent offenders in order to reach federal funds. Title V of the Violent Crime Control and Law Enforcement Act of 1994, for example, authorized federal funding of drug courts but prohibited the distribution of federal funds to drug courts that allow clients with current or prior serious violent convictions to participate (GAO, 1997, p. 41). Our results, while certainly not definitive, suggest that this restriction has merit.

One important issue, which goes beyond our current data, concerns whether drug courts restrict program eligibility in other ways, which may not be warranted. For example, courts exclude non-violent offenders currently charged or previously convicted of distribution/sales offenses from eligibility. Given that a large proportion of drug distribution/sales offenses are committed by drug users trying to support their expensive drug habits (see e.g., Johnson et al., 1985), such policies may exclude a substantial population of offenders who could benefit from drug court treatment. Likewise, many courts also exclude drug abusing offenders with extensive criminal histories and serious mental health issues. While such offenders are more of a risk to public safety, offering effective drug treatment to such offenders holds the promise of producing reductions in re-offending—which is precisely what the "risk principle" hypothesizes.

There is some evidence outside of the drug court context that suggests expanding the drug court model to broader populations of offenders can be effective. Perhaps the most prominent example of this is found in the research of Adele Harrell and colleagues who evaluated the National Demonstration of the Breaking the Cycle (BTC) project. This project applied an intervention based on the drug court model to *nearly all drug abusing offenders* arrested on felony charges in three sites (Tacoma, WA, Birmingham, AL, and Jacksonville, FL). In spite of partial program implementation, the evaluation found that participation in the BTC was associated with reductions in criminal behavior (Harrell et al., 2002; Mitchell & Harrell, 2006). BTC's findings are buttressed by a recent simulation analysis that indicates relaxing the eligibility criteria for criminal justice based drug treatment programs would substantially increase the number of offenders eligible for treatment, and this expansion would avert several million crimes that these offenders who otherwise commit (Bhati & Roman, 2010). Thus, one important issue for future drug court research concerns the effectiveness of these programs when applied to drug abusing offenders who traditionally have been excluded from participation.

Notes

1. Some courts use both methods of case disposition for different groups of offenders. For example, Delaware used the diversionary approach with youthful, less criminally involved offenders, and the post-plea approach with more serious offenders. Those on the diversionary track were required to participate in drug court for six to twelve months; whereas, the more serious offenders on the post-plea track were required to be involved for longer periods of time.

2. According to a 1997 GAO report (1997, p. 38), approximately 80% of drug courts operating at the end of 1996 received federal funds; and therefore, at least 80% of drug courts restrict eligibility to non-violent offenders.

3. The drug court model also has been applied outside of criminal courts; family drug courts are relatively new advents that handle family court issues (e.g., parental rights, allegations of neglect) in cases in which drug abuse is determined to be a factor.

4. The task of identifying drug courts was facilitated by the fact that nearly all of the interventions meeting this criterion self-identified as an evaluation of a "drug court." The only ambiguity regarding this criterion concerned speedy case processing drug courts (e.g., see Belenko, Fagan, Dumanovsky, & Davis, 1993) and evaluations of the Breaking the Cycle (BTC) demonstration project (e.g., see Harrell, Mitchell, Hirst, Marlowe, & Merrill, 2002). Speedy case processing drug courts were ruled ineligible because they focused on expedited case processing of drug cases, not substance abuse treatment with judicial monitoring of drug-involved cases. Evaluations of the BTC demonstration project were also ruled ineligible. While this intervention was based on the drug court model, a judge generally did not actively monitor clients and most clients, rarely, if ever had a status hearing. Perhaps the clearest indication that the BTC demonstration project was not a drug court program is the fact that a recent multi-site evaluation of drug courts used Pierce County's BTC program as a non-drug court comparison program (see Rossman et al., 2011).

5. Essentially, this criterion required all evaluations to have a comparison and the comparison group received no drug treatment or minimal drug treatment. Therefore, we excluded evaluations that compared drug court treatment to another drug treatment program of similar intensity. We also excluded evaluations in which the comparison group was comprised predominantly or solely of dropouts from the drug court as this design is particularly weak methodologically.

6. Further, we searched websites of several prominent research organizations. Specifically, we searched for relevant research reports on the following websites: NPC Research; RAND Drug Policy Research Center; The Urban Institute's website; and, the University of Cincinnati's School of Criminal Justice publications page.

7. When continuous outcome measures best fit the effect size selection criteria, the standardized mean difference effect size was used. And odds-ratio effect sizes and standardized mean difference effect sizes were combined using the methods developed by Hasselblad and Hedges (1995). Specifically, mean difference effect sizes were converted to the odds-ratio scale.

8. If no such effect size was available we selected the effect size that most closely matched these criteria. For example, property offenses were more general than violent offenses, effect sizes based on re-convictions were preferred over re-incarcerations, and effect sizes following sample members closest to 12 months were preferred over other effect sizes.

9. Three randomized experimental evaluations of juvenile drug courts exist; however, two of these had large attrition problems (Dickie, No Date-a, No Date-b); both evaluations had more than 50% total attrition and one had large differential attrition. We coded these compromised experiments as rigorous quasi-experiments.

10. It is important to note that if the two compromised experimental evaluations (Dickie, No Date-a, No Date-b) are rated as experiments, our findings are essentially unchanged: the mean odds-ratio for general recidivism is 1.44 and is not statistically significant.

11. Juvenile and DWI drug court moderator analyses are reported in Mitchell, Wilson, Eggers, & MacKenzie, (forthcoming).

References

Andrews, D. A., & Bonta, J. (1992). *The psychology of criminal conduct.* Cincinnati, OH: Anderson.

Belenko, S. (1998). Research on drug courts: A critical review. *National Drug Court Institute Review*, *1*, 1–26.

Belenko, S. (2001). *Research on drug courts: A critical review 2001 update.* New York: The National Center on Addiction and Substance Abuse.

Belenko, S., Fagan, J., Dumanovsky, T., & Davis, R. C. (1993). *New York City's special drug courts: Recidivism patterns and processing costs.* New York, NY: New York City Criminal Justice Agency, Inc.

Berman, G., & Feinblatt, J. (2005). *Good courts: The case for problem-solving courts.* New York: New Press.

Bhati, A., & Roman, J. (2010). Treating drug involved offenders: Simulated evidence on the prospects of going to scale. *Journal of Experimental Criminology*, *6*, 1–33.

Deschenes, E. P., Turner, S., & Greenwood, P. W. (1995). Drug court or probation: An experimental evaluation of Maricopa County drug court. *Justice System Journal*, *18*, 55–73.

Dickie, J. L. (No Date-a). *Summit County juvenile court drug court: Evaluation report July 1, 1999 - June 30, 2000* Unpublished manuscript, The Institute for Health and Social Policy, University of Akron, Akron, OH.

Dickie, J. L. (No Date-b). *Summit County juvenile court drug court: Evaluation report July 1, 2000 - June 30, 2001* Unpublished manuscript, The Institute for Health and Social Policy, University of Akron, Akron, OH.

Farrington, D. F., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In L. W. Sherman, B. C. Welsh, D. Farrington & D.L. MacKenze (Eds.), *Evidence-Based Crime Prevention* (pp. 13–21). London: Routledge.

Gottfredson, D. C., Najaka, S. S., & Kearley, B. (2003). Effectiveness of drug treatment courts: Evidence from a randomized trial. *Criminology and Public Policy*, *2*, 171–196.

Gottfredson, D. C., Najaka, S. S., Kearley, B. W., & Rocha, C. M. (2006). Long-term effects of participation in the Baltimore City drug treatment court: Results from an experimental study. *Journal of Experimental Criminology*, *2*, 67–98.

Harrell, A., Mitchell, O., Hirst, A., Marlowe, D., & Merrill, J. (2002). Breaking of cycle of drugs and crime: Findings from the Birmingham BTC demonstration. *Criminology and Public Policy*, *1*, 187–216.

Hasselblad, V., & Hedges, L. V. (1995). Meta-Analysis of Screening and Diagnostic Tests. *Psychological Bulletin*, *117*, 167–178.

Hora, P. F. (2002). A dozen years of drug treatment courts: Uncovering our theoretical foundation and the construction of a mainstream paradigm. *Substance Use & Misuse*, *37*, 1469–1488.

Huddleston, W., & Marlowe, D. B. (2011). *Painting the current Picture: A national report on drug courts and other problem-solving court programs in the United States.* Washington, D.C.: National Drug Court Institute.

Huddleston, C. W., III, Marlowe, D. B., & Casebolt, R. (2008). *Painting the current picture: A national report card on drug courts and other problem-solving court programs in the United States.* Washington, D.C.: National Drug Court Institute.

Johnson, B. D., Goldstein, P. J., Preble, E., Schmeidler, J., Lipton, D. S., Spunt, B., et al. (1985). *Taking care of business: The economics of crime by heroin abusers.* Lexington, MA: Lexington Books.

Kalich, D., & Evans, R. D. (2006). Drug court: An effective alternative to incarceration. *Deviant Behavior*, *27*, 569–590.

Latimer, J., Morton-Bourgon, K., & Chrtien, J. -A. (2006, August). *A meta-analytic examination of drug treatment courts: Do they reduce recidivism?* (Tech. Rep. No. rr06-7). Ottawa, Ontario: Department of Justice Canada.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Longshore, D., Turner, S. W., Morral, A., Harrell, A., McBride, D., Deschenes, E., et al. (2001). Drug courts: A conceptual framework. *Journal of Drug Issues*, *31*, 7–26.

MacDonald, J. M., Morral, A. R., Raymond, B., & Eibner, C. (2007). The efficacy of the Rio Hondo DUI court: A 2-year field experiment. *Evaluation Review*, *31*, 4–23.

Marlowe, D. B., Festinger, D. S., Arabia, P. L., Croft, R., Patapis, N. S., & Dugosh, K. L. (2009). A systematic review of DWI court program evaluations. *Drug Court Review*, *6*, 1–52.

Mitchell, O. (2011). Drug and other specialty courts. In M. Tonry (Ed.), *Oxford handbook of crime and criminal justice* (pp. 843–871). New York: Oxford University Press.

Mitchell, O., & Harrell, A. (2006). Evaluation of the breaking the cycle demonstration project: Jacksonville, FL and Tacoma, WA. *Journal of Drug Issues*, *36*, 93–114.

Mitchell, O., Wilson, D. B., Eggers, A., & MacKenzie, D. L. (forthcoming). Drug courts' effects on criminal offending for juveniles and adults: A Campbell Collaboration systematic review. Norway: The Campbell Collaboration.

National Association of Drug Court Professional (1997). *Defining drug courts: The key components.* Washington, D.C.: Bureau of Justice Assistance.

Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis of moderator variable effects. *Psychological Methods*, *3*, 354–379.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York: Russell Sage Foundation.

Rossman, S. B., Rempel, M., Roman, J. K., Zweig, J. M., Lindquist, C. H., Green, M., et al. (2011). *The multi-site adult drug court evaluation: The impact of drug courts.* Washington, D.C.: The Urban Institute.

Saum, C. A., Scarpitti, F. R., & Robbins, C. A. (2001). Violent offenders in drug court. *Journal of Drug Issues*, *31*, 107–128.

Scarpitti, F. R., Butzin, C. A., Saum, C. A., Gray, A. R., & Leigey, M. E. (2005). *Drug court offenders in outpatient treatment: Final report to National Institute of Drug Abuse.* Newark, DE: University of Delaware.

Searle, W., & Spier, P. (2006). *Christchurch youth drug court pilot: One year follow-up study.* Wellington, New Zealand: Ministry of Justice.

Shaffer, D. K. (2006). *Reconsidering drug court effectiveness: A meta-analytic review.* (Tech. Rep.). Las Vegas, NV: University of Nevada.

Shaffer, D. K. (2011). Looking inside the black box of drug courts: A meta-analytic review. *Justice Quarterly*, *28*, 493–521.

Shanahan, M., Lancsar, E., Haas, M., Lind, B., Weatherburn, D., & Chen, S. (2004). Cost-effectiveness analysis of the New South Wales adult drug court program. *Evaluation Review*, *28*, 3–27.

Turner, S., Greenwood, P., Fain, T., & Deschenes, E. (1999). Perceptions of drug court: How offenders view ease of program completion, strengths and weaknesses, and the impact on their lives. *National Drug Court Institute Review*, *II*, 58–81.

U.S. General Accounting Office (1997). *Drug courts: Overview of growth, characteristics, and results.* Washington, D.C.: United States General Accounting Office.

U.S. General Accounting Office (2005). *Adult drug courts: Evidence indicates recidivism reductions and mixed results for other outcomes.* Washington, D.C.: United States General Accounting Office.

Wilson, D. B., Mitchell, O., & MacKenzie, D. L. (2006). A systematic review of drug court effects on recidivism. *Journal of Experimental Criminology*, *2*, 459–487.