Utah Cost of Crime

Sex Offender Treatment (Adults):
Technical Report

December 2012

THE UNIVERSITY OF UTAH

Utah Criminal Justice Center

COLLEGE OF SOCIAL WORK
COLLEGE OF SOCIAL & BEHAVIORAL SCIENCES
UTAH COMMISSION ON CRIMINAL AND JUVENILE JUSTICE
S.J. QUINNEY COLLEGE OF LAW
Utah Cost of Crime

Sex Offender Treatment (Adults):
Technical Report

Christian M. Sarver, M.S.W.  
Jennifer K. Molloy, M.S.W.  
Robert P. Butters, Ph.D.

December 2012

Utah Criminal Justice Center, University of Utah
In the United States (U.S.), more than 30,000 sex offenders participate in court mandated treatment programs annually (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2009). Offenders are required to participate in treatment for a range of offenses, including lewdness, exhibitionism, sexual assault, sexual abuse of a child, sodomy, and rape. Treatment is provided in both secure and community-based settings and falls, broadly, into the following categories: psychological interventions; drug therapies, either for the purposes of castration or psychological treatment; surgical castration; and educational programming. The majority of sex offender treatment programs rely on group-based, cognitive behavioral approaches grounded in social learning theories (Center for Sex Offender Management (CSOM), 2006). The most common treatment targets are victim empathy (87% secure, 93% community), denial (91% secure, 92% community), and intimacy/relationship skills (84% secure, 91% community) (McGrath et al., 2009). More recently, interventions are structured according to the principles of effective correctional services, which matches offenders to treatment based on their risk-level, criminogenic needs, and learning styles (Andrews & Bonta, 2006; Andrews et al., 2001; Bonta, 2001). In the U.S., interventions provided in secure settings range from five months to four years while community-based programs using treatment approaches last anywhere from eight months to the duration of the offender’s life (Daly, 2008). The majority of states that employ community-based approaches use strategies that build upon prison-based programming.

**Prior Research**

Research on the effectiveness of corrections-based sex offender treatment shows generally positive, though mixed, results. In a meta-analysis of 23 studies, Hanson et al. (2009) found that sex offenders who participated in treatment, either prison- or community-based, had a sexual recidivism rate that was 43% lower than untreated offenders (10.9% vs. 19.2%) and a general recidivism rate that was 34% lower (31.8% vs. 48.3%). Hanson’s analysis compared recidivism outcomes between programs that adhered to the principles of effective treatment (risk-needs-responsivity (RNR)) to those that did not and found that the former were associated with larger reductions in recidivism. This latter finding may shed light on the inconsistencies in sex offender treatment research. According to the RNR model, treatment targets should have an empirically-demonstrated relationship to recidivism; however, two of the three most common treatment targets in U.S. sex offender treatment programs (victim empathy and denial) are not associated with recidivism (CSOM, 2006).

Schmucker and Losel (2008) analyzed 69 studies and found that treated sex offenders had a sexual recidivism rate that was 37% lower (11.1% vs. 17.5%) than untreated offenders and a general recidivism rate that was 31% lower (22.4% vs. 32.5%). While this study compared heterogeneous interventions, moderator analyses showed that the only treatments associated with significant changes in recidivism rates were cognitive behavioral and hormonal therapies provided in community-based settings. While previous research confirms the efficacy of community-based programming (McGrath et al., 2009), this finding is difficult to interpret because prisoners in secure settings are generally higher risk than probationers (Schmucker & Losel, 2008). Aos, Miller, and Drake (2006) found that
cognitive-behavioral therapies, for both prison-based offenders (five studies) and probationers (six studies), were associated with lower recidivism rates when compared to untreated offenders. Aos and his colleagues found no effects on recidivism for behavioral (two studies) or insight-oriented (three studies) treatments. Earlier meta-analyses confirm the positive effect of treatment on recidivism (Alexander, 1999; Hall, 1995; Hanson et al., 2002; Polizzi et al., 1999) although several have been critiqued on the basis of methodological weakness within the primary studies (Losel, 2005). Hall’s (1995) analysis of 12 studies showed a 30% reduction in sexual offending for treated offenders, a finding that has been viewed with caution due to Hall’s inclusion of studies that used comparison groups comprised entirely of offenders who had refused, or dropped out of, treatment. The lack of methodological rigor in sex offender treatment research has contributed to controversies regarding whether or not such programs are effective. Two early meta-analyses found no positive effect on recidivism outcomes for treated offenders (Furby, Weinrott, & Black, 1989; U. S. General Accounting Office (USGAO), 1996). At the time of their study, Furby et al. (1989) concluded that meta-analysis was impossible because of the dearth of high quality research.

Problems in the research stem partially from public concern over sex offender recidivism, which results in most convicted offenders being sentenced to some form of treatment. Consequently, researchers struggle to find a sample of untreated offenders to use as a comparison group. Given low baseline rates of sex offending, treatment effects are difficult to discern if the study follow-up period is less than three years (Daly, 2008). Heterogeneity in study constructs—including for treatment targets, offender type, and outcome measures—further complicates interpretation of the effects of treatment.

Methods

Inclusion Criteria
A systematic review was conducted, in accordance with the protocol outlined by PRISMA (Moher, Liberati, Tetzlaff, & Altman, 2009), to identify studies for inclusion in this meta-analysis. The research team identified eligibility criteria for population, intervention, setting, outcome, and methodology (see Methods Report for a further explanation of the search strategy). Given Furby’s (1989) failure to identify studies of sufficient quality to conduct a meta-analysis, the researchers restricted the search to studies published between 1987 and 2011. Studies had to meet the following criteria to be eligible:

a) The study must evaluate a criminal justice intervention. Primary prevention programs and programs serving non-court involved populations were excluded. Interventions could be chemical, educational, therapeutic, or a combination of those strategies; surgical interventions were excluded. Studies that evaluated containment strategies for managing sex offenders were not eligible if they did not also include a treatment component.

b) Both experimental and quasi-experimental studies were eligible for inclusion. Quasi-experimental studies had to use matching or statistical methods to demonstrate equivalence between the intervention and comparison group. Treatment dropouts were not considered an appropriate comparison group;
comparison groups consisting of offenders who refused treatment were included only if the authors conducted analyses that demonstrated that the groups were similar.

c) Both the treatment group and the comparison/control group must consist of adult offenders (ages 18 years and older) convicted of a sex offense, which could include exhibitionism, sexual assault, sexual abuse, rape, and other offenses classified as sex crimes within local criminal statutes.

d) The study must include a post-treatment measure of recidivism—which could be arrest, conviction, or incarceration—as an outcome. Recidivism data from official sources was preferred, but studies using only self-report recidivism measures were also eligible. Offenses committed while the offender was in a secure facility were not included; however, recidivism during the time that a participant was on community supervision was included. Non-criminal outcome measures—such as measures of treatment targets—were excluded from this analysis.

e) The study must report quantitative results than can be used to calculate an effect size. Given the interest in recidivism, dichotomous data were preferred (e.g. odds ratios). If the study only included continuous measures, effect sizes were calculated and converted into odds ratios (Lipsey & Wilson, 2001) using log odds (see Methods Report).

Retrieving and Screening Studies

The initial literature search identified 3,750 citations, from which researchers pulled 159 studies for further evaluation. Full articles were screened by one researcher, which resulted in 29 studies that met inclusion criteria. Twenty-percent of the full articles (k=30) were double-screened for inclusion by a researcher; all disagreements were resolved through discussion between the two researchers. Six of the included studies were identified as follow-up reports on the same study population and excluded from the analysis. Outcomes were included for only one timeframe from each study. Because the cost model was based on three to five year recidivism rates, the outcome closest to this timeframe was used in the analysis. In total, there were 23 studies included in the final analysis (see Appendix A for PRISMA chart).

Extracting Data

The research team developed a detailed code sheet and manual, which included variables related to study quality, program characteristics, participant characteristics, and treatment variables (see Methods Report for a full description of coding variables). One author coded all of the included studies and entered the data into an Excel spreadsheet. Ten percent (10%) of included studies were double-coded (k=2), by a researcher; discrepancies were resolved through discussion. To assess study quality, the authors used a modified version of The Maryland Scale of Scientific Rigor (Aos, Phipps, Barnoski, & Lieb 2001; Gottfredson, MacKenzie, Reuter, & Bushway, 1997). Studies that received a rating lower than “3” (unmatched comparison group or no comparison group) were excluded. Where studies reported multiple measures of recidivism, researchers selected the broadest measure (e.g. arrest over conviction). Outcome data were collected on both sexual recidivism and general recidivism (including sexual). Too few studies included measures of violent recidivism and non-sexual recidivism to conduct a meta-analysis on those outcomes. Offenders who
dropped out of treatment prior to completion were analyzed as part of the treatment group (even if the authors reported results separately). Studies were classified as secure-based if the intervention took place in a prison, jail, or psychiatric hospital. Studies were classified as community-based if the offenders were on probation or parole at the time of treatment (most commonly this was in a community mental health clinic). Community-based programming that was provided as a planned aftercare component of secure programming was classified as secure-based treatment; community-based programming that was unrelated to prison treatment was classified as community-based, even if the offender had received treatment during incarceration. Studies were coded in terms of treatment type: cognitive-behavioral, behavioral, psychotherapeutic, chemical, combined, and other.

Analysis
Data were coded into an Excel spreadsheet, which allowed researchers to calculate descriptive statistics for the full sample. The authors then recoded variables, to condense data into comparable units wherein each study contributed only one effect size to each outcome measure, and entered those into Comprehensive Meta-Analysis (CMA, version 2). Using CMA, the authors assessed heterogeneity using the Q and I-squared statistics (see Results section). The Q statistic is a test of the null hypothesis: a significant value (p>.05) indicates that the variation between studies was greater than one would expect if the difference could be explained entirely by random error (Borenstein, Hedges, Higgins, & Rothstein, 2009). Because the Q statistic is not a precise measure of the magnitude of dispersion between studies, the authors conducted additional analyses to quantify the proportion of variance that could be attributed to differences in study characteristics (such as setting, population, and intervention). The I-squared statistic (values range from 0% to 100%) provides an estimate of how much of the variation between studies can be explained by random error: values near 0 indicate that all of the difference can be explained by random error. Values at 25%, 50% and 75% are, respectively, considered low, moderate, and large heterogeneity (Piquero & Weisburd, 2012; Sedgwick, 2012). Given the range of study characteristics present in this sample, a random effects model, which assumes variability between studies (Piquero & Weisburd, 2010), was used to generate a summary effect size for each outcome measure. All data was coded and transformed into odds ratios, with values above 1 indicating a negative treatment effect and values below 1 indicating a positive treatment effect (i.e. reduced recidivism rates for offenders who participated in treatment).

Results

Sample Characteristics
The majority of studies (k=12) were conducted in Canada. Four studies were unpublished technical reports, conducted by government or private entities, and the remaining studies were published in peer-reviewed journals. One study received a score of five out of five on study quality and the remaining studies (95%) received a score of three or four. The majority of studies (12) evaluated an intervention that was primarily cognitive-behavioral, four studies evaluated interventions that were primarily behavioral and seven evaluated multi-modal interventions that combined elements from different therapeutic orientations.
Both the comparison and intervention groups in all included studies consisted of male offenders. In the majority of studies, the mean age of participants was mid-thirties. Fourteen of the studies evaluated interventions conducted in a secure setting and nine studies evaluated interventions conducted in a community-based setting. The community-based treatment samples included three studies of probationers, three studies of parolees, and two studies of mixed groups. The follow-up period ranged from six months to twenty years. Twenty-one of the studies included a measure of sexual recidivism; twelve used a measure of general recidivism that included sex offenses; three used a measure of general recidivism that excluded sex offenses; six used a measure of violent recidivism; and two used a composite measure of sexual and violent recidivism. Total sample size ranged from 27 to 2,557 and the entire sample describes 4,778 offenders in treatment groups and 5,417 offenders in comparison groups (see Appendix D).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Frequency</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publication type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer-reviewed journal</td>
<td>19</td>
<td>(83)</td>
</tr>
<tr>
<td>Unpublished technical report</td>
<td>4</td>
<td>(17)</td>
</tr>
<tr>
<td>Book</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td><strong>Sample location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S.</td>
<td>7</td>
<td>(30)</td>
</tr>
<tr>
<td>Canada</td>
<td>12</td>
<td>(52)</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>(18)</td>
</tr>
<tr>
<td><strong>Methodological Quality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5: Random Control Trial (RCT)</td>
<td>1</td>
<td>(4)</td>
</tr>
<tr>
<td>4: High quality quasi-experimental(^1)</td>
<td>1</td>
<td>(4)</td>
</tr>
<tr>
<td>3: Quasi-experimental with testing or matching</td>
<td>21</td>
<td>(92)</td>
</tr>
<tr>
<td><strong>Dropouts enumerated</strong></td>
<td>3</td>
<td>(13)</td>
</tr>
</tbody>
</table>

\(^1\)Employed a quasi-experimental research design with a program and matched comparison group, controlling with instrumental variables or Heckman approach to modeling self-selection; May also include RCT with problems in implementation.

**Meta-analysis**

Sexual recidivism was examined in 21 studies (two of the studies only reported a composite measure of recidivism, that included but was not limited to sexual recidivism; those results are included in the discussion of general recidivism). In 14 of those, results favored treatment (seven were significant at p<0.05). The odds-ratios for sexual recidivism ranged from 0.06 to 2.58. The random effects mean odd-ratio was 0.67 (95% CI of 0.51 to 0.89, p<0.01), indicating that the treatment groups had significantly lower rates of sexual recidivism than the comparison groups. The Q test showed that the distribution of the effect sizes was significantly heterogeneous (Q=54.67, df=20, p<0.001), which was expected given the range of offenders, interventions, and treatment types included in the meta-analysis. The I-squared statistic (I²=63.42) indicated that a moderate amount of the variance can be attributed to actual differences between studies rather than random error.
Following the omnibus meta-analysis (see Appendix B), studies were grouped by setting and recidivism type for further moderator analysis.

**Sexual recidivism by setting.** Thirteen studies examined sexual recidivism following prison- or secure-based treatment, of which eight showed results that favored treatment (five significant at p< 0.05). The random effects mean odds-ratio was 0.73 (95% CI 0.53 to 0.99, p=0.05), indicating that the lower recidivism rate for the treatment group approaches significance. The Q test showed significant heterogeneity between studies (p<0.001). Eight studies examined sexual recidivism following community-based treatment and six showed results that favored the treatment group (three were significant at p< 0.05). The random effects mean odds-ratio was 0.47 (95% CI 0.23 to 0.97, p<0.05) indicating a statistically significant reduction in recidivism for the treatment group. The Q test revealed significant heterogeneity (p<0.05). The between groups Q test was not significant (Q=1.19, df=1, p=0.28), however, which indicates that there is no statistically significant difference between secure- and community-based treatment effects.

**General recidivism.** General recidivism (including sex offenses) was examined in 15 studies (see Appendix C). In 14 of those, results favored treatment (nine were significant at p< 0.05). The odds-ratios for general recidivism ranged from 0.20 to 1.23. The random effects mean odds-ratio was 0.59 (95% CI of 0.47 to 0.74, p<0.01), indicating that the treated offenders had significantly lower rates of sexual recidivism than the untreated offenders. The Q test showed that the distribution of the effect sizes was significantly heterogeneous (p<0.01), which was expected given the range of offenders, interventions, and treatment types included in the meta-analysis.

**General recidivism by setting.** Eight studies examined general recidivism following secure-based treatment and all showed results that favored treatment (five were significant at p<0.05). The random effects mean odds-ratio was 0.55 (95% CI 0.44 to 0.69, p<0.05). Sex offender treatment provided within a secure setting was associated with a significant reduction in general recidivism. The Q test was not significant (p=0.06), which means that the null hypothesis was rejected: on this measure, the studies appear similar. Seven studies examined general recidivism following community-based treatment and six showed results that favored the treatment group (two were significant at p<0.05). The random effects mean odds-ratio was 0.63 (95% CI 0.41 to 0.99, p=0.05) indicating that the reduction in recidivism for the treatment group approached significance. The Q test was significant (p<0.01). The Q test assessing between-group heterogeneity was not significant (Q=.318, df=1, p=0.57), which suggests that there is no significant difference between the effects of secure- and community-based treatment on general recidivism.

**Limitations**
The strength of a meta-analysis rests on the comprehensiveness of the search strategy. While the authors sought to identify all eligible studies, the possibility exists, and is in fact likely, that those efforts failed to identify all the extant research on sex offender treatment. In some cases, the researchers were unable to obtain studies that were identified as eligible evaluations. Further, the strength of a meta-analysis is dependent on the quantity and
quality of the available primary research. Overall, this sample contains few randomized studies and a high proportion of relatively weaker study designs, which may inflate the overall effects of treatment on recidivism. Finally, the studies included here reflect significant heterogeneity in terms of offenders, settings, dosage, study quality, and outcome measures. In many cases, the study authors did not provide sufficient information to allow for moderator analyses of relationship between those characteristics and treatment effect. Further, the relatively small sample of included studies made it difficult to conduct moderator analyses where sufficient information was available.

References


Included Studies
Note: The studies marked with an asterisk (*) were included in the analyses. Studies without an asterisk are eligible but statistically dependent.


APPENDIX A: Search Results

Search: Title and Abstract
Search Limiters: Date Range
(1987-2011), English

3,750 study abstracts reviewed

1. Exclude reviews, theoretical articles, and correlational studies
2. Exclude studies that do not have a comparison group
3. Exclude studies that evaluated containment strategies to manage offenders
4. Exclude studies conducted outside the U.S. or Canada that are not published in peer-reviewed journals.
5. Exclude dissertations

159 abstracts met inclusion criteria
Full text of all articles procured and printed for screening and review.

1. Criteria 1-5 above plus:
2. Must report on a quantitative outcome variable of recidivism
3. Must demonstrate equivalence between treatment and comparison groups

29 studies meet final inclusion criteria.
All studies double-coded for study quality

6 studies excluded for statistical dependence.

23 primary studies of Sex Offender Treatment included in meta-analysis
APPENDIX B: Table of Included Studies

<table>
<thead>
<tr>
<th>Author</th>
<th>Date</th>
<th>N in Each Group</th>
<th>Study Design</th>
<th>Sexual Recidivism</th>
<th>General Recidivism&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Treatment</td>
<td></td>
<td>Odds-Ratio 95% CI</td>
<td>Odds-Ratio 95% CI</td>
</tr>
<tr>
<td>Abracen et al.</td>
<td>2011</td>
<td>63</td>
<td>Matched</td>
<td>1.25, 0.37, 4.18</td>
<td></td>
</tr>
<tr>
<td>Bakker et al.</td>
<td>1999</td>
<td>238</td>
<td>Convenience</td>
<td>0.33, 0.19, 0.57</td>
<td></td>
</tr>
<tr>
<td>Barnoski</td>
<td>2006</td>
<td>432</td>
<td>Convenience</td>
<td>1.72, 0.81, 3.65</td>
<td>0.92, 0.54, 1.59</td>
</tr>
<tr>
<td>Duwe &amp; Goldman</td>
<td>2009</td>
<td>1020</td>
<td>Matched</td>
<td>0.67, 0.50, 0.90</td>
<td>0.69, 0.58, 0.83</td>
</tr>
<tr>
<td>Friendship et al.</td>
<td>2003</td>
<td>647</td>
<td>Convenience</td>
<td>0.93, 0.53, 1.62</td>
<td></td>
</tr>
<tr>
<td>Hanson et al.</td>
<td>1993</td>
<td>106</td>
<td>Convenience</td>
<td>1.60, 0.82, 3.09</td>
<td></td>
</tr>
<tr>
<td>Hanson &amp; Nicholaichuk</td>
<td>2000</td>
<td>245</td>
<td>Convenience</td>
<td>0.47, 0.29, 0.76</td>
<td>0.43, 0.30, 0.60</td>
</tr>
<tr>
<td>Hanson et al.</td>
<td>2004</td>
<td>403</td>
<td>Matched</td>
<td>1.07, 0.70, 1.61</td>
<td>0.98, 0.73, 1.32</td>
</tr>
<tr>
<td>Looman et al.</td>
<td>2000</td>
<td>89</td>
<td>Matched</td>
<td>0.29, 0.15, 0.55</td>
<td>0.56, 0.30, 1.07</td>
</tr>
<tr>
<td>Marques et al.</td>
<td>2005</td>
<td>172</td>
<td>Random</td>
<td>1.13, 0.61, 2.09</td>
<td></td>
</tr>
<tr>
<td>Marshall &amp; Barbaree</td>
<td>1998</td>
<td>68</td>
<td>Convenience</td>
<td>0.29, 0.11, 0.76</td>
<td></td>
</tr>
<tr>
<td>Marshall et al.</td>
<td>1991</td>
<td>23</td>
<td>Convenience</td>
<td>0.48, 0.15, 1.61</td>
<td></td>
</tr>
<tr>
<td>Marshall et al.</td>
<td>2008</td>
<td>94</td>
<td>Matched</td>
<td>0.21, 0.02, 1.97</td>
<td>0.31, 0.10, 0.98</td>
</tr>
<tr>
<td>McGrath et al.</td>
<td>1998</td>
<td>71</td>
<td>Matched</td>
<td>0.08, 0.01, 0.70</td>
<td>0.28, 0.10, 0.81</td>
</tr>
<tr>
<td>McGrath et al.</td>
<td>2003</td>
<td>105</td>
<td>Convenience</td>
<td>0.48, 0.25, 0.95</td>
<td>0.53, 0.30, 0.93</td>
</tr>
<tr>
<td>McGrath et al.</td>
<td>2007</td>
<td>104</td>
<td>Matched</td>
<td>0.86, 0.28, 2.64</td>
<td>1.23, 0.70, 2.16</td>
</tr>
<tr>
<td>Procter</td>
<td>1996</td>
<td>51</td>
<td>Matched</td>
<td>0.40, 0.07, 2.29</td>
<td></td>
</tr>
<tr>
<td>Rice et al.</td>
<td>1991</td>
<td>29</td>
<td>Matched</td>
<td>1.39, 0.49, 4.03</td>
<td></td>
</tr>
<tr>
<td>Rjuddis &amp; Timmerman</td>
<td>2000</td>
<td>56</td>
<td>Matched</td>
<td>2.58, 0.28, 23.83</td>
<td>0.80, 0.34, 1.92</td>
</tr>
<tr>
<td>Robinson</td>
<td>1995</td>
<td>189</td>
<td>Random</td>
<td>0.46, 0.19, 1.10</td>
<td></td>
</tr>
<tr>
<td>Wilson et al.</td>
<td>2005</td>
<td>60</td>
<td>Matched</td>
<td>0.26, 0.07, 1.01</td>
<td>0.52, 0.24, 1.10</td>
</tr>
<tr>
<td>Wilson et al.</td>
<td>2009</td>
<td>18</td>
<td>Matched</td>
<td>0.07, 0.00, 1.23</td>
<td>0.20, 0.07, 0.62</td>
</tr>
<tr>
<td>Zgoba</td>
<td>2005</td>
<td>495</td>
<td>Convenience</td>
<td>0.84, 0.49, 1.44</td>
<td>0.42, 0.27, 0.64</td>
</tr>
</tbody>
</table>

Total Sample = 10,195

<sup>1</sup>General recidivism includes sexual offenses