Journal of Aggression, Conflict and Peace Research

An examination of potential biases in research designs used to assess the efficacy of sex offender treatment
Anthony Beech Nick Freemantle Caroline Power Dawn Fisher

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

An examination of potential biases in research designs used to assess the efficacy of sex offender treatment

Anthony Beech, Nick Freemantle, Caroline Power and Dawn Fisher

Professor Anthony Beech is a Professor in Criminological Psychology at School of Psychology, University of Birmingham, Birmingham, UK. Professor Nick Freemantle is based at Primary Care and Population Health, University College London, London, UK. Dr Caroline Power is a Forensic Psychologist at Psychology Department, Partnerships in Care, Essex, UK. Dr Dawn Fisher is the Head of Psychology at St. Andrews Healthcare, Birmingham, UK and The University of Birmingham, Birmingham, UK.

Abstract

Purpose – The purpose of this paper is to examine the potential biases in research designs used to assess the efficacy of sex offender treatment.

Design/methodology/approach – In all, 50 treatment studies ($n = 13,886$) were examined using a random effects meta-analysis model.

Findings – Results indicated a positive effect of treatment for both sexual (OR $= 0.58$, 95%, CI 0.45-0.74, $p < 0.0001$), and general recidivism (OR $= 0.54$, 95%, CI 0.42-0.69, $p < 0.0001$), indicating that the likelihood of being reconvicted after treatment was around half compared to no treatment. RCTs showed no significant effect for sexual or general, recidivism. Significant effects were found for non-RCT designs (i.e. incidental cohort, completers vs non-completers designs). Assignment based on need (i.e. giving treatment to those who were high-risk) indicated a negative effect of treatment.

Practical implications – The results highlight the importance of considering study design when considering treatment efficacy.

Originality/value – The current research reports studies identified up until 2009, and examined both published, and unpublished, research originating from a variety of samples employing a random effects model. Consequently, it can be argued that the results are both original and are reflective not only of identified studies, but are also representative of a random set of observations drawn from the common population distribution (Fleiss, 1993). The results of the study suggest that what is required in future research is methodological rigour, and consistency, in the way in which researchers measure the effectiveness of sexual offender treatment.

Keywords Meta-analysis, Treatment, Sex offenders, Biases, Random effects models, Research designs

Paper type Research paper

Introduction

Demonstrating treatment effectiveness is probably the most contentious issue in the field at the present time. Research studies can be interpreted very differently and experts regularly disagree upon the value of studies given the methodological flaws often present within this type of research (Hanson, 1997). Therefore, there is little consensus as to whether treatment works (see Brooks-Gordon *et al.*, 2006; Harkins and Beech, 2007; Marshall and Marshall, 2007, for discussions of this issue). There are a number of factors that have contributed to the lack of accord. The main issue is centred on the confidence in which we can say that the wide mixture of study, and quality, design have the potential to demonstrate that therapy has an impact. Some have argued (e.g. Harris *et al.*, 1998; Quinsey *et al.*, 1993; Rice and Harris, 1997, 2003), that unless a randomized control trial (RCT) design is employed (i.e. where

participants are assigned by chance to treatment/control, and where pre-existing differences between the two groups will be randomly distributed), the question of whether treatment works, in terms of reducing recidivism rates, can never be properly answered.

The perceived scientific rigour of the RCT approach has led to the assertion that this is the gold standard design for evaluating any type of treatment (see e.g. Egger *et al.*, 2005). This is not a new assertion, as nearly 80 years ago, Fisher (1935) noted that, "It may be said that the simple precaution of randomization will suffice to guarantee the validity of the test of significance, by which the result of the experiment is to be judged." (p. 21). Despite the perceived scientific rigour provided by RCTs they are rarely applied within criminal justice settings, or even for that matter in sex offender treatment studies, even though this is the strongly preferred design by the Association for the Treatment of Sexual Abusers (ATSA), see www.atsa.com/ppRCTuse.html and www.atsa.com/ppRCTimplement.html for a thorough exposition of the implementation of such a design.

RCT designs, however, are not without their own associated problems. For example, for this design to be effectively implemented, treatment for some individuals will have to be deliberately withheld to create a "non-intervention" group. This raises ethical and legal issues for major institutional systems, given the possible consequences of denying a high-risk sexual offender treatment (Harkins and Beech, 2007; Marshall and Marshall, 2007). In addition, unless a sample is sufficiently large, random allocation cannot adequately ensure that treatment/control groups are reasonably balanced (Marques and Murphy, 2004). As regards the latter problem, techniques such as "minimization" (which aims to reduce the imbalance between the number of individuals in each treatment group over a number of factors) can be used to help improve the efficacy of the approach (Bracken, 2001). Hence, more typically, when evaluating sex offender treatment, efficacy designs have included: incidental cohort designs where control groups may have been taken from a different time period, or from a sample that could not be offered treatment, or possibly because they were released prior to the implementation of the treatment programme under study (e.g. Hanson *et al.*, 1993; Proctor, 1996); comparing volunteers undertaking treatment with treatment refusers (Barnes and Peterson, 1997); comparing treatment completers vs treatment drop-outs (Seager *et al.*, 2004); and assignment based on "need". In these designs it is clear that there are potential confounds that will influence the outcomes of such studies in terms of inflating differences between those undertaking treatment compared to controls. Some argue that non-RCT designs are so inherently biased that they should not be used at all.

As regards to the potential biases in these designs, changes in rates of sexual offending over time can have effects upon incidental cohort designs, where the recidivism rates of treated offenders are compared to an untreated cohort of offenders released in previous years (e.g. Hanson *et al.*, 2004). For example, Mishra and Lalumière (2009), and the Minnesota Department of Corrections, have noted that recidivism rates have declined since the 1990s. In the latter case this has been attributed to the State's increase in the length and intensity of post-release supervision (Minnesota Department of Corrections, 2007). These findings have clear implications in this design in that if cohorts are used that are not contemporaneous with the treated individuals, an artificial difference may be found irrespective of whether treatment has been undertaken.

For volunteer/treatment refusal designs, motivation and subsequent selection bias (whether candidates really participated voluntarily, or were referred by the associated administration dealing with them) have been shown to confound treatment outcome (Jones *et al.*, 2006). In completers/drop-out designs, Pelissier (2007) found that higher levels of motivation and education, and beginning treatment within three months of initial imprisonment, were associated with treatment completion. Therefore, the control groups are much more likely to be less well educated, unmotivated, and hence possibly more risky. Beyko and Wong (2005) found that non-sexual criminogenic needs (i.e. aggression, rule-violating behaviours), and responsivity factors (i.e. denial, lack of motivation) significantly differentiated programme non-completers from completers, particularly with rapists.

Assignment on the basis of need designs, by definition, means that higher risk individuals are offered treatment in preference to lower risk controls. Here the best that may be expected in treatment is a reduction in recidivism rates of those high-risk individuals to those of low

risk controls. Hence, it is imperative that risk levels are also taken into account, which is not often the case (see Friendship *et al.*, 2003, where risk level (high, medium-high, low-medium, and low), based on Static-99 scores were taken into account). Therefore, such factors must be considered when evaluating whether an intervention was found to be successful or unsuccessful.

But how is treatment typically evaluated in the field? Meta-analysis is recognized as a useful tool to do this, as it is the process by which a number of studies' results are combined in order to yield an overall weighted average statistic (Egger *et al.*, 2005), hence, adding stronger statistical weight than single studies. Kenworthy *et al*. (2004), conducted a meta-analysis of nine identified RCTs ($n = 567$). Results ranged from one study demonstrating no benefit of psychodynamic treatment, and the potential for harm at 10 years, to another indicating that a cognitive-behavioural treatment (CBT) approach resulted in reduced re-offending. A systematic review of the same nine studies, by Brooks-Gordon *et al.* (2006), concluded that CBT reduced re-offence at one year, but increased re-arrest at ten years. Hence, merely relying on RCT outcomes suggests somewhat inconclusive evidence for treatment. Therefore, it would seem necessary to look for the effectiveness of treatment using these, as well as other treatment designs.

Three meta-analytic studies in the 1990s (Hall, 1995a; Alexander, 1999; Gallagher *et al.*, 1999) have assessed the effectiveness of treatment across a wide range of experimental designs. All have been criticized by Hanson *et al.* (2002) for methodological reasons, including the fact that there is too much method variance across studies, and the inclusion of studies where bias could be expected (such as comparing treatment completers vs drop-outs). Hanson *et al.* (2002) conducted a meta-analysis with the aim of addressing these methodological flaws, and included all credible studies of psychological treatment, irrespective of treatment design. To be included in the Hanson *et al.* analysis, studies had to have a control group, incorporating either those who had received no treatment, or alternatively those who had received treatment that was judged to be inadequate or inappropriate. In all, 43 studies were included in this analysis ($n = 9,534$), with an average follow-up of 46 months. A significant effect of treatment was found, 12.3 percent for treated compared to 16.8 percent for untreated samples. A similar effect of treatment was demonstrated for general recidivism, 27.9 percent for treated samples compared to 39.2 percent for untreated samples.

Lösel and Schmucker (2005), in a meta-analysis of a wider range of studies ($k = 69$; $n = 22,181$), including surgical castration and hormonal medication, identified a positive effect of treatment with treated offenders (11.1 percent) found to recidivate at a significantly lower level than control groups (17.5 percent). However, there was no measurable outcome difference between RCT and other designs. Variables that were found to be associated with larger treatment effects included: small sample size of studies; age homogeneity of participants; outpatient vs institutional treatment; and the researchers' affiliation with the programme. It was hypothesised by Lösel and Schmucker that the latter finding is due to there being a more thorough implementation and monitoring, and where researchers were involved in delivering the programme as well.

Better quality studies should of course be more valid than results from poorer quality designs; however, quality scores have generally not been a good predictor of study results in the general epidemiological field (Greenland and O'Rourke, 2001). Hence, consideration of this in the corrections field does seem to be useful. Hanson *et al.* (2009) have examined the effects of treatment in 23 studies ($n = 6,746$) that met a basic set of criteria for quality of design. Here, all studies were rated on the extent to which they adhered to the risk, need, and responsivity (RNR) principles of the "What Works" literature (Andrews and Bonta, 2010). Hanson *et al.* found that the sexual recidivism rate in untreated samples was 19 percent, compared to 11 percent in treated samples. Studies that adhered to all three RNR principles were found to produce recidivism rates that were less than half of the recidivism rates of comparison groups. Studies that followed none of the RNR principles had little effect in reducing recidivism levels.

In 2007 the Collaborative Outcome Data Committee (CODC), issued a document on the quality of research designs (CODC Guidelines). Although these guidelines state a clear preference for RCTs, as do current ATSA guidelines (see www.atsa.com/ppRCTuse.html), the authors point out that it is highly unlikely a "definitive" study will ever provide a conclusion to the on-going debate within the field. Therefore, the CODC guidelines suggest that it is only through the accumulation of

results from diverse research methodologies that a more definitive conclusion can be drawn regarding the effectiveness of treatment. Given the biases noted above, the aim of the current research is to provide a meta-analysis of the outcome of different treatment designs in order to assess their relative outcomes (giving an indication of their relative levels of bias), which have not been generally reported in previous outcome studies. We would note that we have deliberately included studies where such biases are likely to have occurred in order to begin to quantify such effects. We have also gone back to all original studies correcting for some errors that have appeared in other published research. Hence, we hope to provide here the current best set of data on treatment studies that have been conducted.

Method

Study selection

The initial selection of studies used was obtained from those reported by Hanson *et al.* (2002) ($k = 43$) given the comprehensive nature of this study. This involved data extraction from the Hanson *et al.* paper, the reference list, and also contacting authors and researchers in order to obtain original papers detailed within this research. This data set was extended by conducting a new literature search to identify suitable studies published after May 2000, or studies published before May 2000, which were not included in the Hanson *et al.* (2002) study. Since the publication of the Hanson *et al.* (2002) study a number of previously included studies had been updated. It was therefore necessary to include a number of these studies for the purpose of the current research. Some of the original studies from the Hanson *et al.* meta-analysis were therefore excluded and replaced with updated versions. For example, Marques *et al.* (1994) was replaced with the more recent publication of the same study by Marques *et al.* (2005). Personal communication with authors was also utilized for identifying studies with different references/authors, where these were duplicates of the same sample.

Search terms applied within the Hanson *et al.* (2002) meta-analysis were utilized and additional search terms were generated for the purposes of this study (see the Appendix for a list of search terms used within the research). Using the specified list of search terms, internet computer searches of a range of relevant databases were conducted (also detailed in the Appendix). Manual and Internet computer searches were conducted of relevant journals, which are also listed in the Appendix. A search of other additional sources of information was also carried-out. For example, manual searches of relevant reference lists were conducted and relevant articles were followed up. Searches were also conducted of the internet sites of relevant institutions, organizations and departments of corrections (each of which are again listed in the Appendix). Additionally, experts within the field of sexual offender treatment were contacted in order to account for any unpublished data or studies, and to identify any data that may have been missed through other search methods. Titles were scanned for initial relevance after studies were identified. If a study appeared to be related to the research question, the abstract was read in detail, and if identified as relevant, the full report was accessed and analysed for suitability against the chosen inclusion/exclusion criteria detailed below.

Inclusion criteria

Studies had to include a sample of sexual offenders (i.e. those convicted of a sexual offence) and include recidivism as an outcome measure in primary studies. The dependent variable of recidivism ranged from lapse behaviour (i.e. a single occurrence of an undesirable activity) to incarceration. Studies focusing exclusively on other measures (e.g. personality, hormone levels) were excluded. The following criteria also had to be met to be included in the study.

Comparison procedure. Studies had to incorporate a control group design, that is, the study contained a comparison of the recidivism rates (including sexual or general) of a sample of treated sexual offenders with a comparison group of untreated sex offenders. This comparison sample could either be an untreated control group or a group of offenders receiving treatment deemed to be inappropriate, inadequate or that differed from the evaluated programme in content, intensity and specificity.

Recidivism criteria. The same recidivism criteria must have been used for both treatment and control groups in order for equal comparisons to be made. Additionally, recidivism rates must have been reported for approximately the same follow-up period. In terms of follow-up periods, most research studies provided an average length of follow-up. Where this was not provided (e.g. a minimum and maximum length of follow-up was provided) a median length of follow-up was calculated. Additionally, if different follow-up periods were reported for both treatment, and comparison groups, the average of these two figures was calculated. A decision was also made to include analyses of both sexual and general recidivism within the meta-analysis to account for the evidence suggesting that sexual offenders will often re-offend with a non-sexual offence as opposed to a sexual offence (Hanson and Bussière, 1998). It was hoped that this decision would also partially help to address the issue of plea-bargaining and the masking of sexual reconvictions through official recording of data (Quinsey *et al.*, 1993).

Sample size. The combined sample size used by each primary study had to consist of at least five participants (i.e. five individuals in each group).

Country of origin. No restrictions were made as to where the studies were conducted. However, only studies reported in English were included, with the exception of one study reported in French by Martin (1998), as per the Hanson *et al.* (2002) study.

Time of publication. Those studies available before May 2000 were obtained via the Hanson *et al.* (2002) meta-analysis. The search therefore aimed to obtain studies conducted as of May 2000. However, any studies conducted before this arising within the search were also reviewed for eligibility.

Type of treatment. Studies were categorized according to the type of treatment administered: cognitive-behavioural (CBT); systemic therapy which is targeted at identified youth and family problems within and between the multiple systems in which family members are embedded; psychodynamic; behavioural treatment; unknown; and mixed. Studies using medical interventions alone were excluded, unless the medical intervention was administered alongside a psychological intervention.

Coding procedure for treatment designs

Research designs were coded using the same categories as Hanson *et al.* (2002). To this end the original authors of the Hanson *et al.* (2002) meta-analysis were contacted in order to access detailed study coding. All studies were coded by two of the authors (A.B. and C.P.). Studies were then assigned to one of six research design categories, as follows.

Randomised controlled trial (RCT). This category is where a randomly assigned treated group of offenders was compared to a randomly assigned group receiving no treatment, or alternative treatment deemed to be inappropriate or inadequate.

Incidental cohort. This category incorporated a group of treated offenders being compared against a group of offenders receiving no treatment or alternate treatment who had been incidentally assigned to these conditions. For example, studies were included within which comparison groups of sexual offenders were drawn from the following: offenders released before the implementation of treatment programmes; offenders matched using official information sources; offenders who received an earlier version of the treatment programme; offenders who received no treatment; offenders receiving treatment judged to be lower in quality (e.g. programmes were unavailable, or offenders had insufficient time left on their sentence to complete a treatment programme).

Assignment based on need. Studies in which those assigned to treatment based on need were compared to those deemed not to need treatment (i.e. higher risk individuals, or those with a high level of problems indicative of higher risk were offered treatment, lower risk individuals were not. Here, comparisons were made between higher risk "treated" compared to lower risk "untreated" individuals.

Completers vs drop-outs. Here studies were included which compared treatment completers to those who had dropped out of treatment.

Attendance vs refusers. Here, studies were included where any treatment attenders (including drop-outs) were compared to those who refused to partake in treatment.

Completers vs non-completers. Only one study was included here as it used a combination of dropouts, refusers and treatment termination as the control group, and compared these with those who had completed treatment.

The categories outlined above, therefore, followed a general structure in order to separate those studies from which pre-existing group differences would not be expected (e.g. RCTs); those in which group equivalence was not assured, but where reasonable steps had been taken to ensure some level of equivalence (e.g. incidental assignment/matched controls); and studies in which differences may reasonably be expected (e.g. through the inclusion of treatment drop-outs).

Index of treatment effectiveness: odds ratio (OR)

The starting point of any meta-analysis involves the selection of a summary statistic or effect measure (Egger *et al.*, 2005). Egger *et al.* (2005) recommend that all measures of effect should be accompanied by confidence intervals. For the current meta-analysis, the OR statistic was employed, which is defined as the ratio of the odds of an event occurring in two groups. ORs are recommended with the use of dichotomous data (Fleiss, 1994; Lipsey and Wilson, 2001). ORs are also noted to be relatively unaffected by arbitrary design features such as the proportion of offenders in the treatment and comparison groups, or the overall recidivism rate (Fleiss, 1994). ORs were derived with the use of 2×2 tables detailing the recidivism data outcomes of both the treatment and comparison groups. ORs were then calculated as follows (from Fleiss, 1994):

$$OR = \frac{\text{recid}t/\text{nonrecid}t}{\text{recid}c/\text{nonrecid}c}$$

Recid$t$ refers to the number of recidivists in the treatment group, nonrecid$t$ is the number of non-recidivists in the treatment group, recid$c$ is the number of recidivists in the comparison group and nonrecid$c$ is the number of non-recidivists in the comparison group. ORs cannot be calculated should there be no events in either of the groups, therefore, following the recommendations of Fleiss (1994), in such events a value of 0.5 was added to each cell of the 2×2 contingency table when a cell was empty. This method therefore enabled the analysis of empty cells. If a study reported a different set of results for different offender types or risk groups, effect sizes were calculated separately and then averaged to a single effect size. Interpretation of ORs were as follows: a value of 1.0 would indicate no difference between the groups being compared; values below 1.0 are indicative of treatment having a positive effect; values above 1 indicate treatment as having a negative effect.

Data synthesis

StatsDirect (www.camcode.com) was used to perform the data synthesis. There is no fixed set of rules to be employed when deciding which model of meta-analysis to apply (e.g. Fleiss and Gross, 1991). Due to the nature of the current research and the likely chance of significance between study variability in treatment effects, the decision was made to employ a random effects model, as suggested by DerSimonian and Laird (1986). A random effects model makes the assumption that for each study, the true effects are random observations drawn from a common population distribution, and there is no single "fixed" treatment effect (Egger *et al.*, 2005). The particular version of the random effects model from StatsDirect was a non-iterative analysis estimated with weights to reflect unequal variances. A random effects model leads to larger confidence intervals and relatively more weight being given to smaller studies, in the presence of heterogeneity (systematic differences in treatment effects between studies). In order to test for the likely potential of a high level of heterogeneity between individual studies, a *Q* test of homogeneity was applied (Hedges and Olkin, 1985), whereby a statistically significant result is indicative that between trial variability is more than would be expected by chance alone. We would also note that with the employment of a random effects model there is less likelihood of overstating the precision of the results, which it is suggested can be the case with fixed effects models (Schmidt *et al.*, 2009).

In addition we undertook a meta-regression analysis using full random effects numerical simulation techniques. The advantages of the full effects approach is that it considers between study variability in treatment effects as a parameter to be estimated with uncertainty from the data, unlike standard methods which consider the observed heterogeneity to be the true heterogeneity (Smith *et al.*, 1995). Although empirically Bayesian in approach, the analysis was conducted using non-informative priors (the principle of indifference) to replicate the assumptions in standard meta-analysis. In this hierarchical meta-regression analysis we: replicated the conventional analysis of the effects of treatment on sexual recidivism across all the studies; undertook a meta regression analysis examining the effect of randomization against other designs examining treatment effects.

Data description

A total of 50 studies were included within the meta-analysis, of these: four were RCTs; 25 were incidental assignment studies; five examined assignment based on need; seven compared treatment completers vs treatment dropouts; eight compared treatment completers vs treatment refusers; one study compared treatment completers vs a combination of dropouts, refusers and treatment termination. From the studies, 13,886 offenders were examined, of which 7,037 had received treatment, and 6,849 belonged to a comparison group (i.e. did not receive any treatment, or had not completed treatment). The dates of studies ranged from 1976 through to 2009. The total sample of each individual study collected ranged from between 14 and 2,557. The majority of studies were based on American ($k = 24$) or Canadian ($k = 15$) samples, with samples also derived from the UK ($k = 8$), New Zealand ($k = 1$), Holland ($k = 1$) and Australia ($k = 1$). The majority of studies focused on adult male sexual offenders, however, five studies focused on adolescent treatment. The type of treatment offered within each study selected was categorized as CBT ($k = 39$), psychotherapy ($k = 5$), systemic ($k = 2$), behavioural ($k = 2$), unknown ($k = 1$), mixed ($k = 1$). The "mixed" study (Washington Institute for Public Policy, 1998) was reported by authors to consist of a combination of treatment techniques including group therapy, psycho-educational classes, behavioural treatment, drama therapy and family involvement. As well as differences in study design, individual studies also varied in terms of the way in which recidivism was defined. Definitions included those studies that looked specifically at reconviction, those using re-arrest, those looking at broader definitions of recidivism, such as readmission to institutions, and one study which examined recidivism using the presence and commission of relapse behaviours. Follow-up periods averaged at 56.11 (SD = 34.14) months, or 4.68 (SD 2.85) years for the studies in which these were reported ($k = 44$). Data for all studies are shown in Table I.

Results

Overall comparisons were made across the studies using ORs for the type of design, for both sexual and general recidivism.

Analysis of sexual recidivism by study design

Overall, a significant advantage for treated vs control groups was demonstrated for sexual recidivism (OR = 0.58, 95% CI 0.45-0.74, $p < 0.0001$). Figure 1 shows sexual recidivism rates, ORs, confidence intervals and the heterogeneity statistic, for each specific study design.

It can be seen from Figure 1 that, when studies were compared by treatment design, the following pattern emerged. No significant advantage was identified for the four studies employing an RCT design (OR = 0.52, 95%, CI 0.15-1.84, $p = 0.31$), with the confidence intervals suggesting that this design has the potential to work well or not at all. Between-study variability in treatment effect was substantial ($p = 0.0005$). For the 25 studies employing an incidental cohort design, the results indicate a highly significant reduction in sexual recidivism (OR = 0.61, 95%, CI 0.450.84, $p = 0.002$) between treated and untreated participants. Although there was significant variability between the studies ($p < 0.0001$), the confidence intervals indicate that for most studies employing this design, a treatment effect is likely to be found in over 95 percent of any

Table I Details of all studies examined in the analyses

Study	Year	Treatment Type	Study design	Sample size		Recidivism rates (sexual)		Recidivism rates (any)	
				Treatment	Comparison	Treatment	Comparison	Treatment	Comparison
Alberta Hospital (Studer and Reddon; Clelland et al.)	1998	CBT	Completers vs dropouts	194	157	8	13	31	53
Allam	1998, 1999	CBT	Incidental (matched on risk)	153	74	5	6	20	31
Bakker et al.	1999	CBT	Incidental (retrospective cohort)	238	283	21	54		
Barnes and Peterson	1997	CBT	Attendance vs refusal	147	138	4	12	18	36
Berlin et al.	1991	Psychotherapy	Completers vs dropouts	257	206	13	31		
Borduin et al. (a)	1990	Systemic	RCT	8	8	1	6	2	7
Borduin et al. (a)	2009	Systemic	RCT	24	24	2	11	7	6
Bremer (a)	1992	Psychotherapy	Completers vs dropouts	66	27	12	3		
(Clearwater) Nicholaichuk et al.	2000	CBT	Incidental (matched controls)	296	283	43	94	95	99
Cooper (a)	2000	Unspecified	Incidental (current)	64	25	5	1	1	2
Craissati and McClurg	1997	CBT	Incidental (group vs individual therapy)	43	14	2	0		
CS/RESORS	1991	CBT	Incidental	335	181	17	7		
Dwyer	1997	CBT	Completers vs dropouts	125	55	8	9		
England-Aytes et al.	2001	CBT	Incidental	170	149	1	7	11	11
Florida	1989, 1977	Psychotherapy	Need	39	199	2	10	5	36
Friendship et al.	2003	CBT	Incidental (matched on risk)	647	1910	17	54	84	315
Guarino-Ghezzi and Kimball (a)	1998	CBT	Incidental	33	25	0	1	10	12
Hall	1995b	CBT	Attendance vs refusal	24	6	0	1	5	3
Hanson et al.	2004	CBT	Incidental	403	321	56	45	189	161
Hanson et al.	1992, 1993	Psychotherapy	Incidental	106	89	37	30	64	54
Hedderman and Sugg	1996	CBT	Incidental	133	191	6	17	5	38
Huot	1999	CBT	Attendance vs refusal	92	159	15	30		
JJ Peters	unpublished	Psychotherapy	RCT	148	83	21	6	81	50
Kramer	1985	Behavioral	Incidental	37	19	2	6	5	9
Lab et al. (a)	1993	CBT	Need	46	109	1	4	12	22
Lindsay et al.	1998	CBT	Incidental	7	7	0	4		
Lindsay and Smith		Psychotherapy							
(La Macaza) Martin	1998	CBT	Incidental	65	56	4	12	9	26
Marques et al.	2005	CBT	RCT	259	225	57	45		
Marshall and Barbaree	1988	CBT	Completers vs dropouts	68	58	9	29		
Marshall et al.	1991	CBT	Incidental	17	23	4	8		
McGrath et al.	2003	CBT	Completers vs refusers	56	90	3	27	17	32
McGuire	2000	CBT	Attendance vs refusers	54	14	1	1	2	1
Missouri	1989	CBT	Attendance vs dropouts	105	156	5	20	24	64
Perkins	1987	CBT	Incidental	62	12	20	2	30	7

(continued)

Table I

Study	Year	Treatment Type	Study design	Sample size		Recidivism rates (sexual)		Recidivism rates (any)	
				Treatment	Comparison	Treatment	Comparison	Treatment	Comparison
(Pinel) Proulx et al.	1998	CBT	Completers vs dropouts	117	55	29	13	43	31
Proctor	1996	CBT	Incidental (retrospective)	54	54	2	8		
Rattenbury	1986	Unspecified	Need	69	69	15	10	27	25
RHC Pacific	1995	CBT	Need	38	29	3	0	13	6
Smiley and Mulloy									
Rice et al.	1991	Behavioural	Need	51	85	26	24	33	46
(RTC Ontario) Looman et al.	2000	CBT	Incidental	95	95	25	30		
Ruddijs and Timmerman	2000	CBT	Incidental (concurrent)	56	56	3	1	17	46
(Saskatchewan) Nicholaichuk	1996	CBT	Incidental	26	35	3	1		
Scalora and Garbin	2003	CBT	Incidental	76	118	2	35		
Schweitzer and Dwyer	2003	CBT	Incidental	196	164	6	8	20	15
Seager et al.	2004	CBT	Completers vs non-completers	109	37	5	12		
Taylor	2000	CBT	Incidental	702	142	59	11		
(Twin Rivers) Gordon and Packard	1998	CBT	Attendance vs refusal	209	97	4	8	31	24
(Warkworth) Seto and Barbaree	1999	CBT	Attendance vs refusal	312	85	19	5	41	13
Washington	1998	Mixed	Attendance vs refusal	321	306	35	43	64	159
Worling and Curwen (a)	2000	CBT	Incidental	85	46	10	6		

Note: (a) denotes study of adolescent sexual offenders

**Figure 1** Sexual offence recidivism effectiveness by treatment design
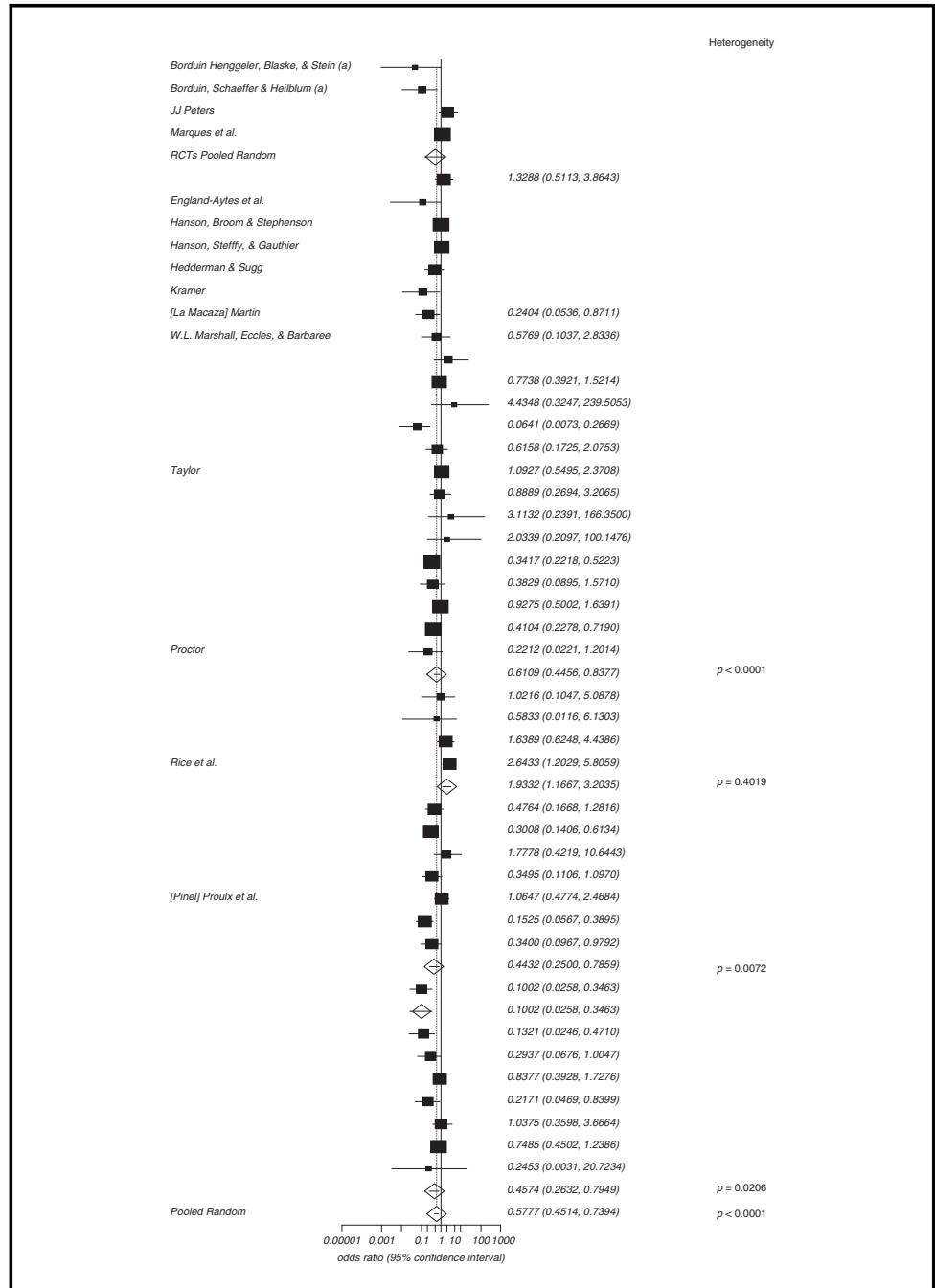
studies of this type. For the five studies comparing assignment based on need, a significantly higher rate of sexual recidivism was found in the treated group compared to offenders considered not to need treatment (OR = 1.93, 95%, CI 1.17-3.20, $p = 0.01$). Variability between studies was not significant. For the seven studies examining treatment completers vs drop-outs, sexual recidivism was shown to be significantly lower in those completing treatment compared to those who had not (OR = 0.44, 95%, CI 0.25-0.79, $p = 0.005$). Again, variability between studies was more than would be expected by chance alone ($p = 0.007$). For the seven studies examining treatment completers with treatment refusers, sexual recidivism was shown to be significantly lower in those completing treatment (OR = 0.46, 95%, CI 0.26-0.79, $p = 0.006$). Between-study variability in treatment effect was substantial ($p < 0.05$). Only one study (Seager *et al.*, 2004) was

identified which examined comparisons of treatment compared with a mix of dropouts ($n = 17$), refusals ($n = 19$), and treatment termination ($n = 1$), in terms of sexual recidivism, again recidivism was shown to be lower in those completing treatment (OR = 0.10, 95%, CI 0.03-0.34).

Analysis of general recidivism by study design

A total of 32 studies reported general recidivism data, where a significant advantage for treated vs control groups was demonstrated for general recidivism (OR = 0.54, 95%, CI 0.420.69, $p < 0.0001$). Figure 2 shows general recidivism rates, as well as ORs, CIs, and the heterogeneity statistic by each study design.

Figure 2 General offence recidivism effectiveness by treatment design
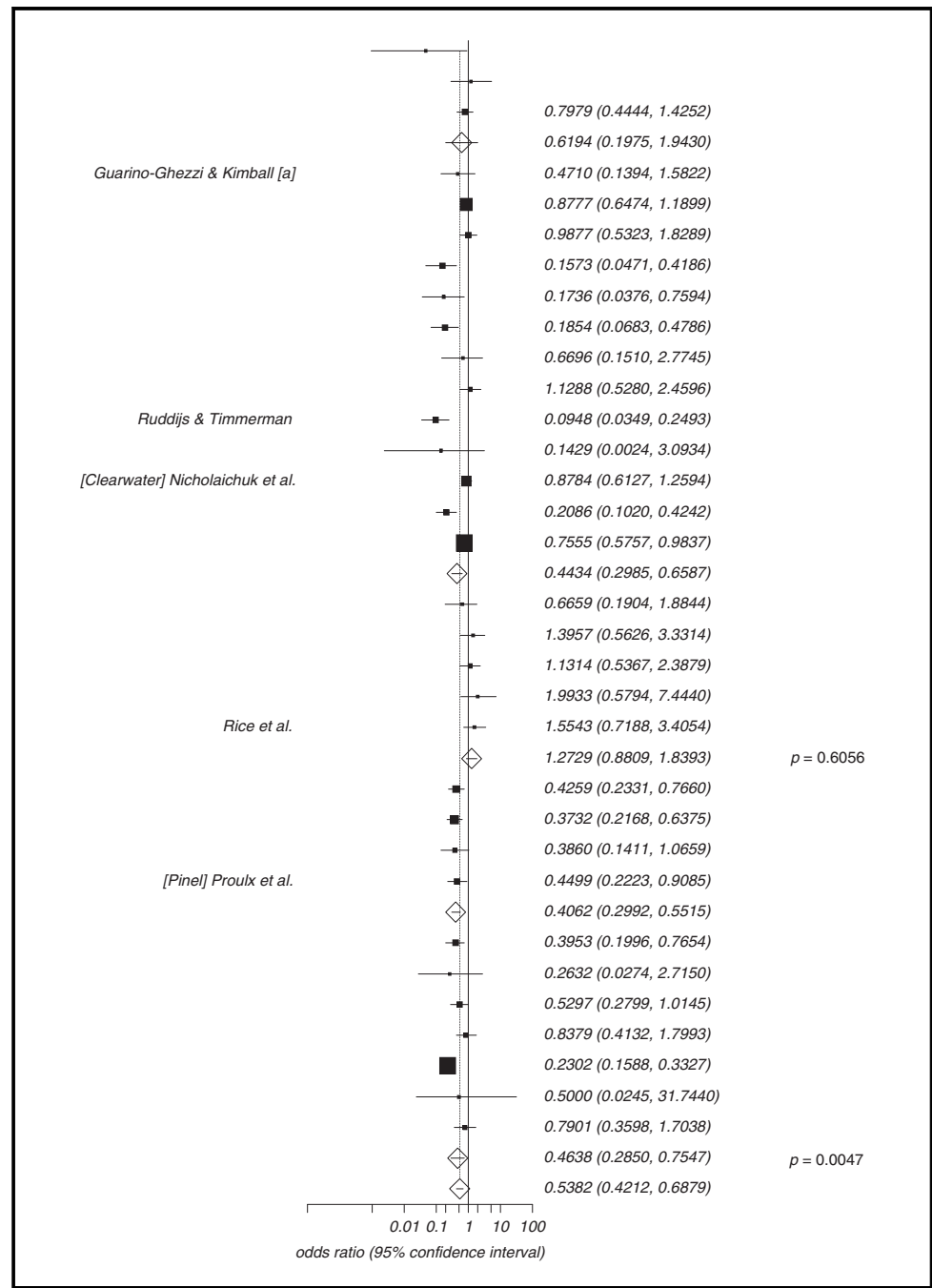
It can be seen from Figure 2 that although numerically in line with the overall results, no effect of treatment was identified in the three RCT designs, while between study variability was at the limits of what would be expected by chance alone (OR = 0.62, 95%, CI 0.201.94, $p = 0.41$). For the 13 studies employing an incidental cohort design, the results here indentified significant reductions in general recidivism for treated vs untreated offenders (OR = 0.44, 95%, CI 0.300.66, $p < 0.0001$). Incidental studies did however demonstrate more variability than would be expected by chance alone ($p < 0.0001$). In the five studies, where assignment was based on need, those referred to treatment demonstrated higher rates of general recidivism compared to offenders deemed not to need treatment although this estimate is not statistically significant (OR = 1.27, 95%, CI 0.881.84, $p = 0.20$). For the four studies examining those who completed treatment, vs those who dropped out of treatment, general recidivism was shown to be significantly lower in those completing treatment (OR = 0.41, 95%, CI 0.300.55, $p < 0.0001$). A significant effect of treatment was identified in the seven studies examining the general recidivism rates of treatment volunteers vs treatment refusers (OR = 0.46, 95%, CI 0.280.75, $p = 0.002$).

Meta-regression analysis

A full random effects meta-regression analysis closely replicated the findings of the conventional meta-analysis on the sexual recidivism data reported in Figure 1. The results of the meta-regression are reported in Table II.

It can be seen from Table II that including a classification variable in the model denoting randomized vs non-randomized designs, indicated that RCTs were associated with a similar treatment effect to non-randomised designs. However, the wide confidence intervals indicate the substantial uncertainty of these findings. However, we would note that the effects of treatment remained strongly statistically significant and were largely unaffected by the inclusion of the design covariate term in the model.

Discussion

Significant reductions were found for both sexual (OR = 0.58, 95%, CI 0.450.74), and general (OR = 0.46, 95%, CI 0.280.75) recidivism in treated vs control groups. These results are broadly in line with rates reported by Lösel and Schmucker (2005; OR = 0.59, CI = 0.47-0.74), and stronger than those reported by Hanson *et al.* (2002; OR = 0.81, CI = 0.71-0.94), or Hanson *et al.* (2009; OR = 0.77, CI = 0.65-0.91). The results for general recidivism are again broadly in line with Lösel and Schmucker (OR = 0.60, CI = 0.48-0.75), and Hanson *et al.* (2002; OR = 0.56, CI = 0.50-0.64), and are better than the results reported by Hanson *et al.* (2009; OR = 0.92, CI 0.78-1.77). Importantly, the results of the study indicate that the design of individual studies impacts upon the estimated effectiveness of treatment. Specifically, the overall results of the RCTs in the study did not identify any significant effect of treatment. But this finding must be qualified in that the confidence intervals in the analysis lacked precision and indicate that this design could potentially either show very good, or poor, outcomes. Positive effects of treatment (in both sexual and general recidivism) were identified in studies employing incidental treatment designs, and data using treatment dropouts as a control group. However, it should be noted that designs, using treatment dropouts as controls are open to substantial systematic bias, as there is a higher recidivism rate in those that dropped out of therapy. Unsurprisingly, assigning offenders to treatment based on need indicated a negative effect of treatment, in that higher risk individuals received treatment, while lower risk cases did not.

Table II	Treatment effectiveness for reducing sexual recidivism by design: results of meta-regression analysis	
Treatment type	*Odds ratio*	*95% CI*
Overall	0.55	0.40-0.73
Randomized trials analysis		
Treatment	0.55	0.40-0.76
Trial (yes/no)	0.96	0.29-2.91

The full random effects meta-regression analysis, based upon numerical simulation techniques, found no relationship between study design and treatment effect. There are several potential explanations for this finding. First, the small numbers of RCTs (the only design to deal adequately with bias) were simply underpowered to demonstrate a convincingly smaller "true" treatment effect. Second, trials may be poorly designed and conducted, leading to biased estimates of treatment effect, which are not so different from those identified in the designs more open to systematic bias. Third, there may be no difference in estimates derived from different study designs.

Positive results were also found in designs that examined treatment completers vs treatment dropouts, attendance vs treatment refusal, and dropouts vs treatment refusal. There may be a number of reasons for these results. For example, evidence suggests that those dropping out of treatment programmes may be likely to possess personality characteristics associated with an increased risk of recidivism (e.g. Marques *et al.*, 1994; McGonaghy, 1999), such as a general uncooperative, antisocial lifestyle (Hanson and Morton-Bourgon, 2004). In short, the use of both treatment refusals and treatment dropouts as a basis for comparison may increase the possibility of between group differences in recidivism occurring for reasons other than the presence, or absence, of treatment. Therefore, it is not possible to draw reliable conclusions with the use of these data.

Evidence also suggests that similar caution should be applied to designs that compare the recidivism rates of offenders assigned to treatment based on need compared with those deemed not to need treatment. The initial difference in risk level is not accounted for when assessing data from such designs, and in fact it is highly likely that those being offered treatment will be those at higher risk (as based on the "What Works" principles described above). Therefore, data from these studies may distort results by indicating the non-effectiveness of treatment in such study designs.

Conclusions

This study reports studies identified up until 2009, and examines both published, and unpublished research originating from a variety of samples employing a random effects model. Consequently, it can be argued that the results are reflective not only of identified studies, but are also representative of a random set of observations drawn from the common population distribution (Fleiss, 1993). The results of the study suggest that what is required in future research is methodological rigour, and consistency, in the way in which researchers measure the effectiveness of sexual offender treatment. The CODC (2007) have suggested guidelines for those designing new studies or evaluating current programmes, which will hopefully result in research in which the possible range of plausible interpretations is minimized, and enable researchers to present results of treatment outcome with increased confidence.

References

Studies included in the meta-analysis are marked with an asterisk.

Alexander, M. (1999), "Sexual offender treatment efficacy revisited", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 11, pp. 101-16. doi: 10.1007/BF02658841.

*Allam, J. (1998), *Community-Based Treatment for Sex Offenders: An Evaluation*, University of Birmingham, Birmingham.

*Allam, J. (1999), *Effective Practice in Work with Sex Offenders: A Re-Conviction Study Comparing Treated and Untreated Sex Offenders*, West Midlands Probation Service Sex Offender Unit, Birmingham.

Andrews, D.A. and Bonta, J. (2010), *The Psychology of Criminal Conduct*, 5th ed., Matthew Bender & Company, New Providence, NJ.

*Bakker, L., Hudson, S., Wales, D. and Riley, D. (1999), "And there was light", *An Evaluation of the Kia Marama Treatment Programme for New Zealand Sex Offenders Against Children*, Psychological Service, Department of Corrections New Zealand, Christchurch.

*Barnes, J.M. and Peterson, K.D. (1997), *The Kentucky Sex Offender Treatment Program*, Commonwealth of Kentucky, Justice Cabinet, Department of Corrections, Division of Mental Health, Kentucky State Reformatory, La Grange, KY.

*Berlin, F.S., Hunt, W.P., Malin, M.H., Dyer, A., Lehne, G.K. and Dean, S. (1991), "A five-year plus follow-up survey of criminal recidivism within a treated cohort of 406 pedophiles, 111 exhibitionists and 109 sexual aggressives: issues and outcomes", *American Journal of Forensic Psychiatry*, Vol. 12, pp. 5-28.

Beyko, M.J. and Wong, S.C.P. (2005), "Predictors of treatment attrition as indicators for program improvement not offender shortcomings: a study of sex offender treatment attrition", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 17, pp. 375-89. doi: 10.1177/107906320501700403.

*Borduin, C.M., Henggeler, S.W., Blaske, D.M. and Stein, R.J. (1990), "Multisystemic treatment of adolescent sexual offenders", *International Journal of Offender Therapy and Comparative Criminology*, Vol. 34, pp. 105-13. doi: 10.1177/0306624×9003400204.

*Borduin, C.M., Schaeffer, C.M. and Heilblum, N. (2009), "A randomized clinical trial of multisystemic therapy with juvenile offenders: effects on youth social ecology and criminal activity", *Journal of Consulting and Clinical Psychology*, Vol. 77, pp. 26-37. doi: 10.1037/a0013035.

Bracken, M.B. (2001), "On stratification, minimization and protection against types 1 and 2 error", *Journal of Clinical Epidemiology*, Vol. 54, pp. 104-5. doi: 10.1016/S0895-4356(00)00286-9.

*Bremer, J.F. (1992), "Serious juvenile sex offenders: treatment and long-term follow-up", *Psychiatric Annals*, Vol. 22, pp. 326-32. doi: 10.1080/14789940600819341.

Brooks-Gordon, B., Bilby, C. and Wells, H. (2006), "A systematic review of psychological interventions for sexual offenders i: randomised control trials", *The Journal of Forensic Psychiatry and Psychology*, Vol. 17, pp. 442-66. doi: 10.1080/14789940600819341.

*Clelland, S.R., Studer, L.H. and Reddon, J.R. (1998), "Follow-up report of rapists treated in a forensic psychiatric hospital", *Violence and Victims*, Vol. 13, pp. 79-86, (Alberta Hospital).

Collaborative Outcome Data Committee (2007), *Sexual Offender Treatment Outcome Research: CODC Guidelines for Evaluation*, Public Safety Canada Department of Corrections, Ottawa, available at: www. publicsafety.gc.ca/res/cor/rep/codc-eng.aspx

*Cooper, H.M. (2000), "Long-term follow-up of a community-based treatment program for adolescent offenders", *Master Abstracts International*, Vol. 45 No. 03, UMI No. MR21542, p. 172.

*Craissati, J. and McClurg, G. (1997), "The challenge project: a treatment program evaluation for perpetrators of child sexual abuse", *Child Abuse and Neglect*, Vol. 21, pp. 637-48. doi: 10.1016/S0145-2134(97)00028-8.

*CS/RESORS (1991), *An Evaluation of Community Sex Offender Programs in the Pacific Region (No. 21803-0-A602/01-XSB)*, Author, Vancouver.

DerSimonian, R. and Laird, N. (1986), "Meta-analysis in clinical trials", *Controlled Clinical Trials*, Vol. 7, pp. 177-86. doi: 10.1016/0197-2456(86)90046-2.

*Dwyer, S.M. (1997), "Treatment outcome study: seventeen years after sexual offender treatment", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 9, pp. 149-60. doi: 10.1007/BF02674865.

Egger, M., Smith, G.D. and Altman, D.G. (2005), *Systematic Reviews in Health Care: Meta-Analysis in Context*, 5th ed., BMJ Publishing Group, London.

*England-Aytes, K., Olsen, S.S., Zakrajsek, T., Murray, P. and Ireson, R. (2001), "Cognitive behavioral treatment for sexual offenders: an examination of recidivism", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 13, pp. 223-31. doi: 10.1177/107906320101300401.

Fisher, R.A. (1935), *The Design of Experiments*, Hafner Press, New York, NY.

Fleiss, J.L. (1993), "The statistical basis of meta-analysis", *Statistical Methods in Medical Research*, Vol. 2, pp. 121-45. doi: 10.1177/096228029300200202.

Fleiss, J.L. (1994), "Measures of effect size for categorical data", in Cooper, H. and Hedges, L.V. (Eds), *The Handbook of Research Synthesis*, Sage, New York, NY, pp. 245-60.

Fleiss, J.L. and Gross, A.J. (1991), "Meta-analysis in epidemiology, with special reference to studies of the association to environmental tobacco smoke and lung cancer: a critique", *Journal of Clinical Epidemiology*, Vol. 44, pp. 127-39. doi: 10.1016/0895-4356(91)90261-7.

*Friendship, C., Mann, R.E. and Beech, A.R. (2003), "Evaluation of a national prison-based treatment program for sexual offenders in England and Wales", *Journal of Interpersonal Violence*, Vol. 18, pp. 744-59. doi: 10.1177/0886260503018007004.

Gallagher, C.A., Wilson, D.B., Hirschfield, P., Coggeshall, M.B. and MacKenzie, D.L. (1999), "A quantitative review of the effects of sex offender treatment of sexual reoffending", *Corrections Management Quarterly*, Vol. 3, pp. 19-29.

*Gordon, A. and Packard, R. (1998), The impact of community maintenance treatment on sex offender recidivism", paper presented at the 17th annual meeting of the Association for the Treatment of Sexual Abusers, Vancouver, BC (Twin Rivers), October 16.

Greenland, S. and O'Rourke, K. (2001), "On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions", *Biostatistics*, Vol. 2, pp. 463-71. doi: 10.1093/biostatistics/2.4.463.

*Guarino-Ghezzi, S. and Kimball, L.M. (1998), "Juvenile sex offenders in treatment", *Corrections Management Quarterly*, Vol. 2, pp. 45-54.

Hall, G.C.N. (1995a), "Sexual offender recidivism revisited: a meta-analysis of recent treatment studies", *Journal of Consulting and Clinical Psychology*, Vol. 63, pp. 802-9. doi: 10.1037//0022-006X.63.5.802.

*Hall, G.C.N. (1995b), "The preliminary development of theory-based community treatment for sexual offenders", *Professional Psychology: Research and Practice*, Vol. 26, pp. 478-83. doi: 10.1037//0735-7028.26.5.478.

Hanson, R.K. (1997), "How to know what works with sexual offenders", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 9, pp. 129-45. doi: 10.1007/BF02674863.

*Hanson, R.K., Broom, I. and Stephenson, M. (2004), "Evaluating community sex offender treatment programs: a 12-year follow-up of 724 offenders", *Canadian Journal of Behavioural Science*, Vol. 36, pp. 85-94. doi: 10.1037/h0087220.

Hanson, R.K., Bourgon, G., Helmus, L. and Hodgson, S. (2009), "The principles of effective correctional treatment also apply to sexual offenders: a meta-analysis", *Criminal Justice and Behavior*, Vol. 36, pp. 865-91. doi: 10.1177/0093854809338545.

Hanson, R.K. and Bussière, M.T. (1998), "Predicting relapse: a meta-analysis of sexual offender recidivism studies", *Journal of Consulting and Clinical Psychology*, Vol. 66, pp. 348-62. doi: 10.1037//0022-006X.66.2.348.

Hanson, R.K., Gordon, A., Harris, A.J.R., Marques, J.K., Murphy, W., Quinsey, V.L. and Seto, M.C. (2002), "First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 14, pp. 169-97. doi: 10.1177/107906320201400207.

Hanson, R.K. and Morton-Bourgon, K. (2004), *Predictors of Sexual Recidivism: An Updated Meta-Analysis. Corrections Research*, Public Safety and Emergency Preparedness Canada, Ottawa, available at: www.psepcsppcc.gc.ca/publications/corrections/pdf/200402_e.pdf (accessed August 12, 2013).

*Hanson, R.K., Steffy, R.A. and Gauthier, R. (1992), Long-term follow-up of child molesters: risk prediction and treatment and treatment outcome", (User Report No. 1992-02, Corrections Branch, Ministry of the Solicitor General of Canada, Ottawa.

*Hanson, R.K., Steffy, R.A. and Gauthier, R. (1993), "Long-term recidivism of child molesters", *Journal of Consulting and Clinical Psychology*, Vol. 61, pp. 646-52. doi: 10.1037//0022-006X.61.4.646.

Harkins, L. and Beech, A.R. (2007), "Measurement of the effectiveness of sexual offender treatment", *Aggression and Violent Behavior*, Vol. 12, pp. 36-44. doi: 10.1016/j.avb.2006.03.002.

Harris, G.T., Rice, M.E. and Quinsey, V.L. (1998), "Appraisal and management of risk in sexual aggression: implications for criminal justice policy", *Psychology, Public Policy and Law*, Vol. 4, pp. 73-115. doi: 10.1037//1076-8971.4.1-2.73.

*Hedderman, C. and Sugg, D. (1996), *Does Treating Sex Offenders Reduce Re-Offending? Home Office Research and Statistics Directorate: Research Findings, 45*, Home Office, London, available at: www.homeoffice.gov.uk/rds/pdfs/r45.pdf

Hedges, L.V. and Olkin, I. (1985), *Statistical Methods for Meta-Analysis*, Academic Press, New York, NY.

*Huot, S. (1999), Minnesota study of sex offender recidivism", unpublished data.

Jones, N., Pelissier, B. and Klein-Saffran, J. (2006), "Predicting sex offender treatment entry among individuals convicted of sexual offense crimes", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 18, pp. 83-98. doi: 10.1177/107906320601800106.

Kenworthy, T., Adams, C.E., Bilby, C., Brooks-Gordon, B. and Fenton, M. (2004), *Psychological Interventions for those who have Sexually Offended or are at Risk of Offending (Cochrane Review). The Cochrane Library, Issue 3*, Wiley, Chichester.

*Kramer, S.P. (1985), "Sex offender treatment and tracking: problems, perspectives and outcomes in the Utah criminal justice system", unpublished manuscript.

*Lab, S.P., Shields, G. and Schondel, C. (1993), "Research note: an evaluation of juvenile sex offender treatment", *Crime and Delinquency*, Vol. 39, pp. 543-53. doi: 10.1177/0011128793039004008.

*Lindsay, W.R., Marshall, I. and Neilson, C. (1998), "The treatment of men with a learning disability convicted of exhibitionism", *Research in Developmental Disabilities*, Vol. 19, pp. 295-316. doi: 10.1016/S0891-4222 (98)00010-9.

Lipsey, M.W. and Wilson, D. (2001), *Practical Meta-Analysis*, Sage Publications, Newbury Park, CA.

*Looman, J., Abracen, J. and Nicholaichuk, T.P. (2000), "Recidivism among treated sexual offenders and matched controls: data from the regional treatment centre (ontario)", *Journal of Interpersonal Violence*, Vol. 15, pp. 279-90, (RTC Ontario), doi:10.1177/088626000015003004.

Lösel, F. and Schmucker, M. (2005), "The effectiveness of treatment for sexual offenders: a comprehensive meta-analysis", *Journal of Experimental Criminology*, Vol. 1, pp. 117-46. doi: 10.1007/s11292-004-6466-7.

*McGrath, R.J., Cumming, G., Livingston, J.A. and Hoke, S.E. (2003), "Outcome of a treatment program for adult sex offenders: from prison to community", *Journal of Interpersonal Violence*, Vol. 18, pp. 3-17. doi: 10.1177/0886260502238537.

McGonaghy, N. (1999), "Methodological issues concerning evaluation of treatment for sexual offenders: randomization, treatment dropouts, untreated controls, and within-treatment studies", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 11, pp. 183-93.

*McGuire, T.J. (2000), "Correctional institution based sex offender treatment: a lapse behavior study", *Behavioral Sciences and the Law*, Vol. 18, pp. 57-71. doi: 10.1002/(SICI)1099-0798(200001/02)18:1 < 57:: AID-BSL377 > 3.3.CO;2-K.

Marques, J.K., Day, D.M., Nelson, C. and West, M.A. (1994), "Effects of cognitive- behavioral treatment on sex offenders recidivism: preliminary results of a longitudinal study", *Criminal Justice and Behavior*, Vol. 21, pp. 28-54. doi: 10.1177/0093854894021001004.

Marques, J.K. and Murphy, W. (2004), "Random assignment designs", *paper presented at the 23nd Annual Conference of the Association for the Treatment of Sexual Abusers*, *Alberquerque, NM*, October.

*Marques, J.K., Weideranders, M., Day, D.M., Nelson, C. and van Ommeren, A. (2005), "Effects of a relapse prevention programme on sexual recidivism. Final results from California's sex offender treatment and evaluation project (SOTEP)", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 17, pp. 79-107. doi: 10.1177/107906320501700108.

*Marshall, W.L. and Barbaree, H.E. (1988), "The long-term evaluation of a behavioral program for child molesters", *Behaviour Research and Therapy*, Vol. 26, pp. 499-511. doi: 10.1016/0005-7967(88)90146-5.

*Marshall, W.L., Eccles, A. and Barbaree, H.E. (1991), "The treatment of exhibitionists: a focus on sexual deviance versus cognitive and relationship features", *Behaviour Research and Therapy*, Vol. 29, pp. 129-35. doi: 10.1016/0005-7967(91)90041-Z.

Marshall, W.L. and Marshall, L.E. (2007), "The utility of the random controlled trial for evaluating sexual offender treatment: the gold standard or an inappropriate strategy?", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 19, pp. 175-91. doi: 10.1177/107906320701900207.

*Martin, I. (1998), Efficacite d'un programme cognitive-behavioral institutionnel pour delinquents sexuels", unpublished doctoral dissertation, Department of psychology, University of Montreal (La Macaza).

Minnesota Department of Corrections (2007), Sex offender recidivism in Minnesota", April, Author, St. Paul, MN, available at: www.doc.state.mn.us/publications/documents/04-07SexOffenderReport-Recidivism.pdf (accessed November 26, 2010).

Mishra, S. and Lalumière, M. (2009), "Is the crime drop of the 1990s in Canada and the USA associated with a general decline in risk and health-related behavior?", *Social Science and Medicine*, Vol. 68, pp. 39-48. doi: 10.1016/j.socscimed.2008.09.060.

*Nicholaichuk, T.P. (1996), "Sex offender treatment priority: an illustration of the risk/need principle", *Forum on Corrections Research*, Vol. 8, pp. 30-2 (Saskatchewan).

*Nicholaichuk, T., Gordon, A., Gu, D. and Wong, S. (2000), "Outcome of an institutional sexual offender treatment program: a comparison between treated and matched untreated offenders", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 12, pp. 139-53 (Clearwater), doi:10.1177/107906320001200205.

Pelissier, B. (2007), "Treatment retention in a prison-based residential sex offender treatment program", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 19, pp. 333-46. doi: 10.1177/107906320701900401.

*Perkins, D. (1987), "A psychological treatment programme for sexual offenders in McGurk, B.J., Thornton, D. M. and Williams, M. (Eds), *Applying Psychology to Treatment: Theory and Practice*, Her Majesty's Stationary Office, London, pp. 192-217.

*Proctor, E. (1996), "A five-year outcome evaluation of a community-based treatment program for convicted sexual offenders run by the probation service", *Journal of Sexual Aggression*, Vol. 2, pp. 3-16. doi: 10.1080/13552609608413253.

*Proulx, J., Ouimet, M., Pellerin, B., Paradis, Y., Mckibben, A. and Aubut, J. (1998), *Post-Treatment Recidivism Rates in Sexual Aggressors: A Comparison Between Dropout and Non-Dropout Subjects*, (*No. 98-22*), Les cahiers de l'école de criminology, University of Montreal, Montreal, (Pinel).

Quinsey, V.L., Harris, G.T., Rice, M.E. and Lalumière, M.L. (1993), "Assessing the treatment efficacy in outcome studies of sex offenders", *Journal of Interpersonal Violence*, Vol. 13, pp. 621-44. doi: 10.1177/088626093008004006.

*Rattenbury, F.R. (1986), The outcomes of hospitalized and incarcerated sex offenders: a study of offender types, recidivism rates, and identifying characteristics of the repeat offender", unpublished doctoral dissertation, School of Loyola University of Chicago, Chicago, IL.

Rice, M.E. and Harris, G.T. (1997), "The treatment of adult offenders", in Stoff, D.M., Breiling, J. and Master, J.D. (Eds)*Handbook of Antisocial Behavior*, Wiley, Toronto, pp. 425-35.

Rice, M.E. and Harris, G.T. (2003), "The size and signs of treatment effects in sex offender therapy", in Prentky, R.A., Janus, E.S. and Seto, M.C. (Eds), *Sexually Coercive Behavior: Understanding and Management. Annals of the New York Academy of Sciences*, Academic of Sciences, New York, NY, Vol. 989, pp. 428-40.

*Rice, M.E., Quinsey, V.L. and Harris, G.T. (1991), "Sexual recidivism among child molesters released from a maximum security psychiatric institution", *Journal of Consulting and Clinical Psychology*, Vol. 59, pp. 381-6. doi: 10.1037//0022-006X.59.3.381.

*Ruddijs, F. and Timmerman, H. (2000), "The stichting ambulante preventie projecten method: a comparative study of recidivism in first offenders in a Dutch outpatient setting", *International Journal of Offender Therapy and Comparative Criminology*, Vol. 44, pp. 725-39. doi: 10.1177/0306624×00446008.

*Scalora, M.J. and Garbin, C. (2003), "A multivariate analysis of sex offender recidivism", *International Journal of Offender Therapy and Comparative Criminology*, Vol. 47, pp. 309-23. doi: 10.1177/0306624×03047003005.

Schmidt, F.L., Oh, I.-S. and Hayes, T.L. (2009), "Fixed- versus random-effects models in meta-analysis: model properties and an empirical comparison of differences in results", *British Journal of Mathematical and Statistical Psychology*, Vol. 62, pp. 97-128. doi: 10.1348/000711007×255327.

*Schweitzer, R. and Dwyer, J. (2003), "Sex crime recidivism: Evaluation of a sexual offender treatment program", *Journal of Interpersonal Violence*, Vol. 18, pp. 1292-310. doi: 10.1177/0886260503256658.

*Seager, J.A., Jellicoe, D. and Dhaliwal, G.K. (2004), "Refusers, dropouts and completers: measuring sex offender treatment efficacy", *International Journal of Offender Therapy and Comparative Criminology*, Vol. 48, pp. 600-12. doi: 10.1177/0306624×04263885.

*Seto, M.C. and Barbaree, H.E. (1999), "Psychopathy, treatment behavior, and sex offender recidivism", *Journal of Interpersonal Violence*, Vol. 14, pp. 1235-48. doi: 10.1177/088626099014012001 (Warkworth).

*Smiley, W.C. and Mulloy, R. (1995), Treatment of sex offenders: does intervention work?", unpublished manuscript (RHC Pacific).

Smith, T.C., Spiegelhalter, D.J. and Thomas, A. (1995), "Bayesian approaches to randomeffects meta-analysis: a comparative study", *Statistics in Medicine*, Vol. 14, pp. 2685-99. doi: 10.1002/sim.4780142408.

*Taylor, R. (2000), "A seven-year reconviction study of HMP grendon therapeutic community". Home Office Research, Development and Statistics Directorate, Research Findings, No. 115, Home Office, London, available at: www.homeoffice.gov.uk/rds/pdfs/r115.pdf (accessed August 12, 2013).

*The Missouri Sexual Offender Program (1989), *Updated Figures for Missouri Sex Offender Study*, Department of Corrections, Division of Classification and Treatment, Jefferson City, MO (Missouri).

*Washington (1998), *Sex offenses in Washington State, 1998 Update*, Washington State Institute for Public Policy, Olympia, WA.

*Worling, J.R. and Curwen, T. (2000), "Adolescent sexual offender recidivism: success of specialized treatment and implications for risk prediction", *Child Abuse and Neglect*, Vol. 4, pp. 965-82. doi: 10.1016/S0145-2134(00)00147.

Further reading

Abracen, J. and Looman, J. (2004), "Issues in the treatment of sexual offenders: recent developments and directions for future research", *Aggression and Violent Behavior*, Vol. 9, pp. 229-46. doi: 10.1016/S1359-1789(01)00074-X.

*Florida (1984), "Report on treatment programs for sex offenders: summary and recommendations", In conjunction with the Governors Task Force on sex offenders and their victims, Florida Mental Health Institute, Tampa, FL.

*Lindsay, W.R. and Smith, A.H.W. (1998), "Response to treatment for sex offenders with intellectual disability: a comparison of men with one and two-year probation sentences", *Journal of Intellectual Disability Research*, Vol. 42, pp. 346-53. doi: 10.1046/j.1365-2788.1998.00147.x.

*RHC (Rural Health Care) Pacific (1998), *Sex Offender Programs: Program Manuals, Descriptions, Evaluations, and Research*, Regional Health Centre, Washington, DC.

*Studer, L.H. and Reddon, J.R. (1998), "Treatment may change risk prediction for sexual offenders", *Sexual Abuse: A Journal of Research and Treatment*, Vol. 10, pp. 175-81. doi: 10.1177/107906329801000302 (Alberta Hospital).

Weirzbicki, M. and Pekaririk, G. (1993), "A meta-analysis of psychotherapy drop-out", *Professional Psychology: Research and Practice*, Vol. 24, pp. 190-95. doi: 10.1037/0735-7028.24.2.190.

Appendix.

Search terms, Databases, Journals, Web sites Searched

Search terms

The following lists contain those used within the literature search:

sex(ual) offender, rapist, child molester, pedophile, pedophilia, exhibitionist, exhibitionism, sexual assault, incest, voyeur, frotteur, indecent exposure, sexual deviant, paraphilia(c), and, treatment, outcome, recidivism, recidivist, outcome, recidivate, reoffend, reoffense, relapse, failure.

Additional search terms used

Randomized control(led) trial, Randomized clinical trial, SOTP, Intervention, STEP, SOTEP.

Databases

PsychLit; PsychInfo; PsychNet UK; Bath Information Data Service; Biological Abstracts; The Cochrane Library; The Cochrane Register of Trials; Medline; Sociofile; National Criminal Justice Reference System; Science Direct; Ingenta; Web of Science.

Journals

Aggression and Violent Behavior; *American Journal of Forensic Psychiatry*; *American Psychologist*; *Behavior Research and Therapy*; *Behavioral Sciences and the Law*; *British Journal of Clinical Psychology*; *British Journal of Criminology*; *Canadian Journal of Behavioral Science*; *Child Abuse and Neglect*; *Corrections Management Quarterly*; *Crime and Delinquency*;

Criminal Justice and Behavior; *Forensic Update*; *Forum on Corrections Research*; *International Journal of Forensic Mental Health*; *International Journal of Offender Therapy and Comparative Criminology*; *Journal of Consulting and Clinical Psychology*; *Journal of Experimental Criminology*; *Journal of Forensic Psychiatry and Psychology*; *Journal of Intellectual Disability Research*; *Journal of Interpersonal Violence*; *Journal of Sexual Aggression*; *Legal and Criminological Psychology*; *Professional Psychology: Research and Practice*; *Psychology, Crime and Law*; *Research in Developmental Disabilities*; *Sexual Abuse: A Journal of Research and Treatment.*

Web sites

www.sexual-offender-treatment.org (*Journal of the International Association for the Treatment of Sexual Offenders*; *IATSO*).

www.csom.org (Centre for Sexual offender Management – U.S Department of Justice). www.csc-scc.gc.ca (Correctional Service Canada).

www.publicsafety.gc.ca (Public Safety Canada). www.homeoffice.gov.uk (Home Office).

www.hmprisonservice.gov.uk (HM Prison Service).

www.cochrane.org (The Cochrane Collaboration).

About the authors

Professor Anthony Beech is Head of the Centre for Forensic and Criminological Psychology at the University of Birmingham, UK. He has authored over 170 peer-reviewed articles, 40 book chapters and seven books in the area of forensic science/criminal justice. His particular areas of research interests are: risk assessment; the neurobiological bases of offending; reducing online exploitation of children; and increasing psychotherapeutic effectiveness of the treatment given to offenders. Professor Anthony Beech is the corresponding author and can be contacted at: a.r.beech@bham.ac.uk

Professor Nick Freemantle is Chair of Clinical Epidemiology and Biostatistics at the University College London, UK. Throughout his career Nick has worked in clinical trials and he has substantial experience of trials of pharmaceuticals, medical devices and of complex behaviour and educational interventions. He has a particular interest in strategies to ensure the quality of randomised trials and in designing trials to maximise their scientific integrity.

Dr Caroline Power is a Forensic Psychologist working for Partnerships in Care, UK. She was involved in the research while completing a ForenPsyD at the University of Birmingham.

Dr Dawn Fisher is a Chartered and Registered Forensic and Clinical Psychologist and has been involved in sex offender work for over 30 years, and she is a founder member of NOTA (The UK National Organisation for the Treatment of Abusers). She has authored over 60 peer-reviewed articles, book chapters and a book over mainly in this area.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com

This article has been cited by:

1. William L. Marshall, Liam E. MarshallThe Treatment of Adult Male Sex Offenders 1227-1243. [CrossRef]
2. Ana Martínez-Catena, Santiago Redondo. 2016. Etiología, prevención y tratamiento de la delincuencia sexual. *Anuario de Psicología Jurídica* **26**:1, 19-29. [CrossRef]
3. Brandy L. BlaskoOverview of Sexual Offender Typologies, Recidivism, and Treatment 11-29. [CrossRef]