Attitudes toward treatment for individuals who have committed sexual offenses have varied between two extremes and all points in between. Early sexual psychopath laws enacted in the 1930s and 1940s allowed for civil commitment in lieu of incarceration based on the belief that some individuals who committed sexual offenses required and were amenable to treatment (e.g., LaFond, 2005). As belief in the effectiveness of treatment waned, punishment was favored over rehabilitation and many such laws were repealed or fell into disuse (e.g., Janus & Walbek, 2000). In addition to the broader “nothing works” conclusions being reached in criminology (e.g., Lipton, Martinson, & Wilks, 1975), pessimism about sex offender treatment specifically emerged in the 80s (e.g., Furby, Weinrott, & Blackshaw, 1989). Nonetheless, the imperative to respond to sexual violence drove the rapid increase in treatment programs over the subsequent decades (e.g., from 643 respondents in the Safer Society Survey in 1986 to 1380 in 1996; McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010). Though clinicians have sought to provide effective services based on available literature, there remains a lack of clearly identified and supported evidence-based practices (Deming & Jennings, 2019).

Defining Treatment Effectiveness

One of the primary issues is how to measure treatment effectiveness and thereby define what “works.” As the primary goal is to improve public safety by reducing offending, the field has often looked toward recidivism rates to measure success. Officially detected sexual crime is known to underestimate true rates; yet sexual crimes by known offenders may be less likely to go undetected given greater surveillance and a lower bar for charging individuals on a sexual offender registry (Letourneau, Bandyopadhyay, Sinha, & Armstrong, 2009). Additionally, any underestimate associated with using officially detected recidivism as the outcome would apply equally to all groups. Unfortunately, there are several unresolvable concerns regarding the use of sexual recidivism despite its obvious primacy as a treatment goal. First, investigators must choose between using arrest rates, which likely include false positives, or convictions, which likely excludes true recidivists when prosecutors cannot meet the burden of proof or a defendant enters a plea bargain. Second, low rates of officially detected sexual recidivism create study design challenges, requiring long follow-up periods and large samples for sufficient power to detect differences across groups. Third, the use of recidivism as a binary outcome requires a more rigorous test of effectiveness than other treatments face (Levenson & Prescott, 2014). That is, treatments for mental health disorders are often evaluated based on symptom reduction or remission over a period of several years; in contrast, using recidivism as the outcome means “sex offender treatment must result in lifelong remission after one dose” (Levenson & Prescott,
While certainly an appropriate goal, this is clearly a hefty task.

Challenges to Internal & External Validity

Beyond problematic outcome measurement, there are numerous challenges to both external and internal validity in the sex offender treatment literature. Careful study design is needed to maximize external validity by addressing possible sampling bias (i.e., individuals who volunteer for and complete treatment likely differ substantially from those who do not volunteer and/or do not complete treatment). Using volunteer controls and intent-to-treat analyses can manage these potential biases, but treatment outcome studies have not uniformly used such approaches. Additionally, given the many different motivations for participating in treatment and the associated variable treatment engagement, it may be appropriate to parse out those who truly engaged in a program to examine treatment effectiveness in this (admittedly ideal) subgroup. For example, Marques et al. (2005) found differences at the time of treatment completion indicative of engagement and responsiveness between those who did not sexually reoffend and those who did over the eight-year follow-up. Specifically, non-reoffenders demonstrated lower deviant arousal, fewer cognitive distortions, better understanding of the consequences of offending, and better relapse prevention plans than those who went on to sexually recidivate, particularly if they were higher risk offenders at baseline. The ultimate goal is to reduce recidivism in all participants, not just a selected subsample. Nonetheless, it is important to recognize most non-forensic mental health treatments do not face the unique challenge of trying to create change in individuals, a subset of whom may be participating for some secondary benefit (e.g., reduced sentence) and whose engagement may be disingenuous.

There are also several threats to internal validity present in the majority of relevant outcome research. First, most studies don’t assess treatment fidelity, which has been shown to impact treatment effectiveness in related fields (e.g., treatment of serious juvenile offenders; Huey, Henggeler, Brondino, & Pickrel, 2000). Additionally, therapist qualities such as warmth and empathy impact intermediate treatment variables in research with individuals who have committed sexual offenses (e.g., decreases in denial and improved relationships; Marshall, 2005). Although therapist qualities have not been shown to impact the key outcome of sexual recidivism, they are associated with treatment outcomes for other conditions (e.g., CBT outcome studies, Keijers, Schaap, & Hoogduin, 2000). This may require researchers to evaluate clinician-related variability in outcomes; otherwise the degree to which treatment effects (or lack thereof) are attributable to such qualities remains unknown.

Employing Meta-Analysis

Unsurprisingly, based on the issues noted above, research on treatment effectiveness with adults convicted of sexual crimes is quite mixed when examined on a study-by-study basis. Statistical methods for combining results across studies may increase confidence in research findings. This can be seen in the Collaborative Outcome Data Project’s meta-analysis which found significantly lower sexual recidivism rates in treatment than comparison groups (12.3% vs. 16.8%), with larger declines for “newer” (at the time) programs: CBT and MST (Hanson et al., 2002). Hanson and colleagues (2002) acknowledged certain “treatment effects” reported in original studies were likely related to subject selection and condition assignment (e.g.,

---

1 Subsequent articles will specifically look at research with youth offenders.
treatment dropouts in the comparison sample), but a significant treatment effect remained when only including studies that completed intent-to-treat analyses. The authors concluded, “We believe that the balance of evidence suggests that current treatments reduce recidivism, but that firm conclusions await more and better research” (Hanson et al., 2002, p. 186).

Subsequent meta-analytic reviews have reported overall significant treatment effects across studies yet no effect in better-designed studies (Beech, Freemantle, Power, & Fisher, 2015; Långström et al., 2013; Soldino & Carbonell-Vaya, 2017). More than a decade after the Hanson et al. (2002) meta-analysis, Schmucker & Lösel (2015) reached a slightly more optimistic conclusion. Specifically, in their meta-analysis of 29 studies likely to have equivalence between treatment and control groups, there was a significant effect of treatment overall. Despite highlighting numerous methodological problems with the existing literature base, the authors concluded, “Sexual offender treatment has made progress towards an evidence-oriented crime policy” (Schmucker & Lösel, 2015, p. 623). In contrast, a meta-analysis examining outcomes specific to offenders with child victims found no effect of treatment when analyses were limited to those rated as good or weak in methodological quality; no studies were classified as strong and those with the poorest study quality showed a treatment effect (Grønnerød, Grønnerød, & Grøndahl, 2015). Two recent studies using propensity score matching to create equivalence also found no differences between treatment and control groups in sexual recidivism (Grady, Edwards, & Pettus-Davis, 2017; Mews, DiBella, & Purver, 2017).

Where Do We Go From Here?

Despite these lackluster findings, certain treatment approaches may be effective and/or treatment may be effective for particular subgroups. For example, Hanson and colleagues (2009) demonstrated higher treatment effectiveness when programs adhere to more Risk Need Responsivity Principles. Thus, treatment may be more effective for certain offenders (high-risk) when it targets criminogenic needs (dynamic risk factors for offending) in certain contexts (when tailored to offender abilities and characteristics).

Calls for examining attribute-by-treatment interactions (i.e., what works for whom; e.g., Schmucker & Lösel, 2015; Soldino & Carbonell-Vaya, 2017) and the influence of various content and process variables (Levenson & Prescott, 2014) have gone unanswered but may shed important light on the “does treatment work” conundrum. Reviewers also consistently call for more methodologically rigorous study designs, such as randomized controlled trials (e.g., Hanson et al., 2002; Soldino & Carbonell-Vaya, 2017) and measures of effectiveness beyond binary outcomes (recidivism). The field would benefit from expanding the definition of treatment outcome to include behavioral change and harm reduction, using continuous measures such as time to re-offense, offense frequency, and offense severity (Schmucker & Lösel, 2015). As noted by Levenson & Prescott (2014), “It is time to consider that “does treatment work?” may not be the only—or even the best—question to ask” (p. 263).
References


