

A Review Following Systematic Principles of Multisystemic Therapy for Antisocial Behavior in Adolescents Aged 10–17 Years

Anna Markham¹

Received: 17 September 2017 / Accepted: 19 September 2017 / Published online: 11 October 2017
© The Author(s) 2017. This article is an open access publication

Abstract One intervention that has shown promise in reducing serious adolescent antisocial behavior is Multisystemic Therapy. This approach is ecologically driven and considers those social systems within which adolescents are embedded. The aims of this systematic review were firstly to investigate whether Multisystemic Therapy for adolescents aged 10–17 years reduces antisocial behavior and out-of-home placement and, secondly whether improvements in other domains, such as, substance use, adolescent emotional and behavioral difficulties, family functioning, peer relations and school are observed. An initial scoping exercise undertaken to explore the available literature found a systematic review undertaken over 10 years ago. Since then, Multisystemic Therapy has been the subject of a number of randomized control trials across the world. Subsequently an updated review following systematic principles was undertaken utilizing inclusion criteria and quality control measures. This resulted in 11 studies, published from 2006 to 2014 conducted in and outside of America, assessed in quality as ranging from weak to strong. Results indicated that the outcomes for Multisystemic Therapy continue to be mixed across studies. Comparisons between studies were challenging and the review highlighted the need for increased consistency in reporting about “usual services,” deeper consideration about cultural differences in the international transportation of Multisystemic Therapy, adequate sample sizes and improved documenting of aftercare services.

Keywords Multisystemic Therapy (MST) · Juvenile delinquency · Treatment outcome · Effectiveness

Introduction

Considerable research has been conducted over the past 20 years to understand the causes and correlates of adolescent offending, including several longitudinal studies following children well into adulthood (Koops and Orobio de Castro 2004). These have allowed for a detailed and invaluable insight into associations between both risk and protective factors across several domains including individual, family, peer, school and community factors with general offending (see reviews by Lipsey 1995; Loeber and Farrington 1998; Shader 2002; Rennie and Dolan 2010). Despite this, adolescent offending continues to be a complex and persistent societal problem with significant consequences for individuals, families and communities (Blackburn 2003).

Preventative efforts are based on the assumption that the life course trajectories of adolescents can be changed by actively reducing those risk factors associated with antisocial behavior and building on the strengths and protective factors that support desistance. Youth justice and social care agencies are committed to empirically supported interventions that reduce persistent patterns of adolescent antisocial behavior. Some promising results have been obtained with cognitive or behavioral approaches, parent management training, pharmacological approaches and multimodal therapies (Walton 2012). Very few interventions have adopted a structured multimodal approach despite research indicating that serious and violent adolescent offending is multi-determined (Shader 2002). This systematic review focuses on one such intervention, namely, Multisystemic Therapy

✉ Anna Markham
A.Markham.1@bham.ac.uk

¹ School of Psychology, University of Birmingham,
Birmingham B15 2TT, UK

(Henggeler and Borduin 1990) (for a critique of the model, see Markham 2016).

Multisystemic Therapy is quite possibly one of the most empirically investigated interventions for antisocial behavior by adolescents and is currently delivered by more than 500 teams across 16 countries worldwide. Given the rapid widespread implementation, it might be supposed that the effectiveness of Multisystemic Therapy has been consistently empirically demonstrated. A preliminary search of the Cochrane Library, Campbell Collaboration and Google Scholar undertaken in December 2013 yielded a systematic review of the effectiveness of Multisystemic Therapy undertaken by Littell et al. (2005); a meta-analysis (Curtis et al. 2004) and several narrative literature reviews (e.g., LaFavor and Randall 2013). While narrative reviews can provide an informative overview, they are frequently vulnerable to sources of error and bias and little attempt is made to sample all of the available literature and critically consider study design and quality (Petticrew and Roberts 2006). By contrast, systematic reviews adopt a more rigorous, comprehensive and transparent process to synthesize large bodies of information (Petticrew and Roberts 2006).

The systematic review undertaken by Littell et al. (2005) examined 21 randomized control trials of Multisystemic Therapy from eight independent samples (total participants = 1230) including non-published studies (see Table 3 in “Appendix” for studies). A range of different outcome measures from official records of antisocial behavior to caregiver report of child problem behavior and self-report involvement in delinquency were identified across included studies. The authors concluded that whilst there was no evidence suggesting that Multisystemic Therapy had harmful effects; it remained unclear whether the treatment program had clinically significant advantages compared with other interventions. One critique of the previous research trials by Littell and colleagues was the involvement of the program developers in all but one of the included studies. The possible effect of developers-as-evaluators has been investigated by Petrosino and Soydan (2005) in a meta-analysis of 300 randomized control trials of interventions targeting recidivism and a review of 12 meta-analyses of offender treatment. In both cases larger mean effect sizes were found when evaluators were influential in the design and delivery of treatment. While highly involved researchers may be unduly influenced at various stages, it is also possible this finding is explained by an increased attention to integrity and delivery. The only independent study reviewed by Littell et al. (2005) did not find Multisystemic Therapy to be superior to usual services in reducing adolescent antisocial behavior (Canada, Leschied and Cunningham 2002) highlighting the need for further independent investigation of the effectiveness of Multisystemic Therapy.

Additional critiques by Littell et al. (2005) involved poor descriptions of “usual services”, incomplete information

about randomization procedures and unexplained attrition in at least three studies in the number of participants who had agreed to be assessed, who were then randomly assigned and reported in the results. Follow-up periods varied considerably within studies meaning that one participant could be followed up for twice as long as another participant thus making between group comparisons difficult.

The finding by Littell et al. (2005) that the effectiveness of Multisystemic Therapy was not convincingly demonstrated contradicted a meta-analysis by the program developers undertaken at a similar time and including six of the eight same studies (total participants = 708) (Curtis et al. 2004). An overall average effect size of $d = 0.55$ for criminal behavior (based on official records) was reported and Multisystemic Therapy participants and their families were found to be functioning better than 70% of participants treated by usual methods. The different conclusions reached by these two reviews sparked debate about quality assessment, inclusion criteria, allegiance effects and the estimation of effect sizes (see Henggeler et al. 2006; Littell 2006).

It is worth noting that the meta-analysis found a clear difference between efficacy and effectiveness trials. The former is typically conducted under optimal circumstances, i.e., closely supervised by developers, often university based with therapists who are graduate students. Larger effect sizes were observed under these conditions ($d = 0.81$, CI 95% = ± 0.33) than effectiveness studies involving treatment delivered by therapists in community settings ($d = 0.26$, CI 95% = ± 0.06). The significance of the study condition variable highlights possible challenges in the dissemination of Multisystemic Therapy to the real world.

Current Study

The systematic review undertaken by Littell et al. (2005) made an important contribution in bringing together previous findings and highlighting methodological limitations. However, it was based on research literature from 1985 up to January 2003 and therefore can be considered somewhat anachronistic. This is particularly relevant given the increased body of both international and independent research in line with the global spread of Multisystemic Therapy. Littell et al. (2005) were only able to include studies across three countries (America, Canada and Norway) and gaps in knowledge about international transportation remain. It was reported that there were 13 “ongoing” possibly randomized studies and it may be that the sufficient number of new primary studies may alter previous conclusions.

Given the mixed findings, continued efforts are needed to determine the effectiveness of Multisystemic Therapy. This current review aims to determine whether Multisystemic Therapy is more effective than usual services or no

treatment for adolescents who are at risk of serious antisocial behavior and/or out-of-home placement. The primary focus is further offending as measured by official data which, despite its problems, remains the most significant test of any intervention designed to reduce antisocial behavior. The review followed the methodology outlined in the Cochrane Handbook for Systematic reviews of Interventions (Higgins and Green 2011) with the exception of the inclusion of non-English studies due to time and resource constraints.

The first objective was to identify the highest quality experimental studies that have measured the effectiveness of Multisystemic Therapy since the search undertaken by Littell et al. (2005) and provide a detailed description of their methods. The second objective was to determine whether Multisystemic Therapy (including adaptations for problem sexual behavior and contingency management/substance abuse) was more effective than treatment as usual/no treatment in addressing outcomes (primary outcomes: antisocial behavior and out-of-home placement; secondary outcomes: substance use, adolescent functioning; family functioning, peer relations and school performance) in adolescents aged 10–17 years.

Methods

Search Strategy

The current study searched several data sources (current as of 21 June 2014) including seven online bibliographic general reference databases [Cochrane; PsychINFO; Medline; EMBASE; Applied Social Sciences Index and Abstracts (ASSIA); National Criminal Justice Reference Service (NCJRS) and Web of Science]; two dissertation and thesis portals (ProQuest Dissertations and Theses Global and British Library ETHos); Government policy sources (United States Office of Juvenile Justice and Delinquency Prevention; United States Department of Health and Human Services; United Kingdom Ministry of Justice; United Kingdom Department of Health and NHS evidence) and four Multisystemic Therapy related websites (<http://msti.org>, <http://mstuk.org>, <http://www.mstservices.com>; Family Services Research Centre of the Medical University of South Carolina).

The following search terms were modified where appropriate to meet the search requirements of each database “Multisystemic Therapy OR Multisystemic OR Multi-systemic” AND “therap* or treat* or interven* or program*” AND “Outcome* OR evaluat* OR effect* OR experiment* OR trial OR compare* OR impact OR consequen* OR recidiv* OR reoffen* OR relapse OR reconvict* OR research” AND “youth* or adolesc* or young* or teen* or juvenile* or child* or minors* or boys* or girls*.”

The search was restricted to studies published after January 2002 in accordance with guidelines by Petticrew and

Roberts (2006) to allow for an overlap of approximately 1 year before the end date for review updates. To broaden the search, reference lists of shortlisted articles as well as relevant book chapters were hand searched to identify potentially eligible studies. Additionally, a number of prominent authors in the field were contacted directly to identify unpublished or on-going research that might meet the inclusion criteria (Dr Littell, Dr Fonagy, Professor Henggeler, Professor Borduin, Professor Leschied and Professor Ogden). All but one responded within the time frame of the study.

Study Selection

Inclusion/Exclusion Criteria

A population, intervention, comparators and outcome (PICO) framework was used to support a robust search strategy and identify potential studies to be included. The inclusion criteria are laid out in Table 1.

The aim was to evaluate the effectiveness of Multisystemic Therapy as an intervention for reducing antisocial behavior and therefore adapted versions for problem sexual behavior and contingency management/substance abuse were included given the highly elevated rates of alcohol and drug use among those involved in the youth justice system (Tripodi and Bender 2011). Only programs licensed by Multisystemic Therapy Services Inc. were considered due to the stringent training and ongoing supervision/consultation processes. Where possible the methods employed in a review update should mimic those of the original review (Higgins and Green 2011). One of the studies included in the review by Littell et al. (2005) involved an adaptation for psychiatric emergencies. The validity of including this adaptation when examining the treatment effect for offending is questionable given the potential differences in clinical presentation. This was further indicated from the substantially different mean treatment lengths reported between the psychiatric emergency adaptation and standard Multisystemic Therapy (90 h versus 23–40 h, respectively).

Eligible study designs involved the random assignment of participants to treatment and comparison/control groups. Although it is recognized that there are challenges with random allocation, the randomized control trial is recognized as the optimal design for minimizing possible pre-existing differences between treatment and comparison groups as well as risk of inadvertent researcher bias. As highlighted by Peto, Collins and Gray (1995) “There is simply no serious scientific alternative to the generation of large-scale randomized evidence.... [randomized control trials] have a central role to play in the development of rational criteria for the planning of health care throughout the world” (p. 39).

The initial systematic search of reference databases provided a total of 2400 potentially relevant hits. An additional 184 articles were identified from other sources including hand searching

Table 1 The PICO guide to identify relevant literature

	Inclusion
Population	Adolescents aged 10–17 years at risk of serious antisocial behavior/out of home placement/foster care/residential setting/incarceration; males and females; different nationalities; different ethnicities
Intervention	Licensed Multisystemic Therapy programs including the adapted versions for Problem Sexual Behavior and Contingency Management/Substance Abuse
Comparator	Treatment as usual in youth justice or social care system/other treatments (e.g. individual therapy) or no treatment
Outcome	Primary outcomes: antisocial behavior (arrest/criminal conviction); family living arrangements (at home or placement) Secondary outcomes: alcohol and drug use; adolescent functioning; family functioning; peer relations; school attendance and performance
Study type	Experimental where participants are randomly allocated to treatment and comparison/control groups Studies were included if they followed up a sample of adolescents who had engaged in Multisystemic Therapy over time. Prospective and retrospective studies were included. Studies covering the same population were included only where each study examined unique factors. Where similar factors were measured the more up to date study, i.e. the one with the longer follow up period, was included and the other excluded
Language	English language only
Year of publication	From January 2002 until search date (21st June 2014)
Exclusion	Specifically where the primary presenting problem is related to physical health; or as an alternative to hospitalization for psychiatric care ^a ; or Multisystemic Therapy for abuse and neglect Non-licensed Multisystemic Therapy programs Narrative reviews, editorials, commentaries, single case studies, opinion papers or other group designs Any publications prior to January 2002 Studies with children out of the age range Studies without follow-up Studies reported in Littell et al. (2005) systematic review; unless updates were available

^aThis differs from the systematic review undertaken by Littell et al. (2005)

reference lists, dissertation portals or government policy sources. The majority of hits were excluded during the initial screening of titles and abstracts only by the first author, due to clear irrelevance to the current review, duplication or non-English language. Of the 172 remaining studies; attempts were made to retrieve the full copies of the article via the University of Birmingham e library, on site library, interlibrary loans or direct contact with authors. Five articles could not be located.

A predefined form was used to assess study eligibility of those articles deemed potentially relevant during initial sifting. The search process is presented in Fig. 1. This resulted in 18 articles including three updates to two studies included in the systematic review by Littell et al. (2005) and 15 cross referenced publications of nine new trials (see Table 4 in “Appendix” for studies).

Quality Assessment

The 11 studies meeting inclusion criteria were each rigorously quality assured using a pre-defined checklist. Information was cross referenced across publications as recommended by Littell, Corcoran and Pillai (2008). A checklist developed by the United Kingdom’s National Institute of Clinical Evidence (2012) of the most common and well documented biases relevant to randomized control trials was adapted for this purpose. There are no summary scores or specific numerical algorithms within the tool; the use of which is questionable given the lack

of standard techniques establishing reliability and validity of quality assessment scales (Juni et al. 1999). Furthermore, there is an explicit emphasis on considering the likely magnitude and direction of any possible bias supporting the critical evaluation of the implications for interpreting findings (Centre for Review and Dissemination 2009). A structured judgment process was used to combine the overall appraisal of bias and confidence in the findings into four possible categories: strong, good, weak and, rejected. All of the included studies were critically appraised and a random selection, 36.4% (4 papers) were quality checked by a second independent reviewer to ensure reliable ratings. The consensus ratings, where applicable, are presented in Table 5 in “Appendix”.

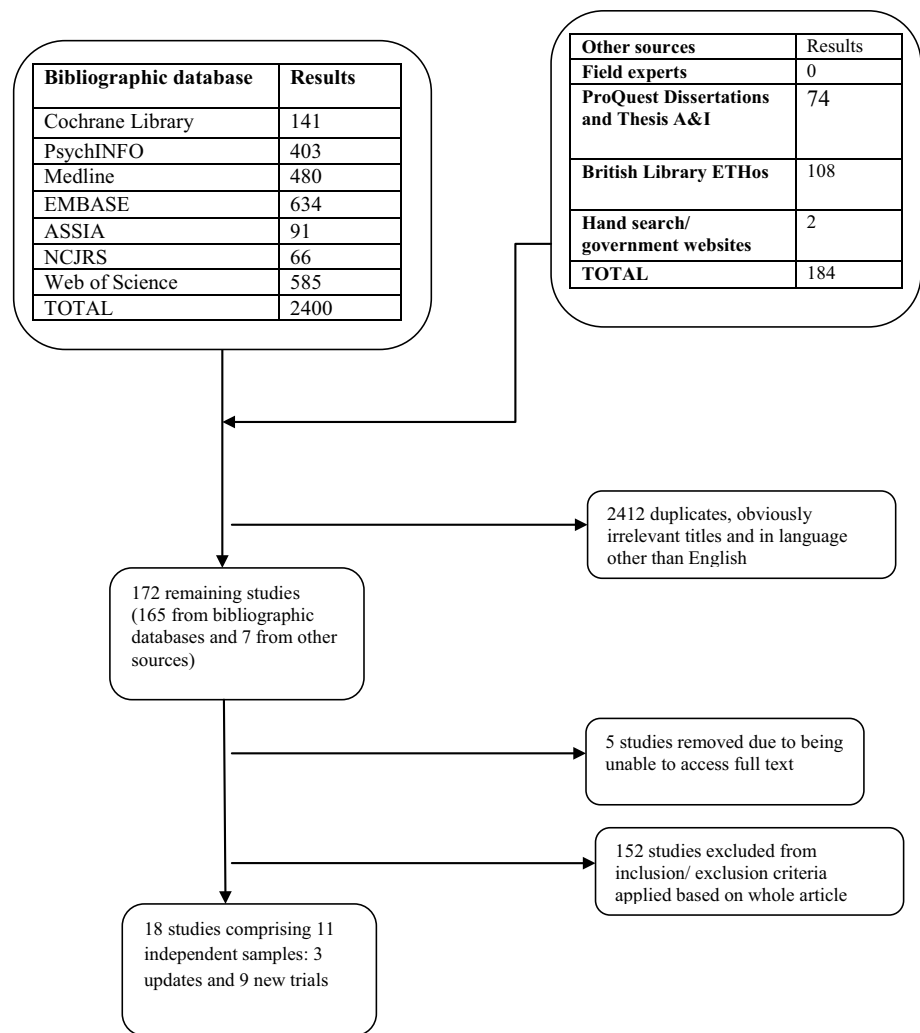
Data Extraction

Data were extracted from the included studies by the primary author using a predetermined form and any absent or unclear information was marked next to the relevant item.

Results

Quality of Included Studies in Review

The included studies ranged in quality (see Table 5 in “Appendix”). Four of these were rated “strong” (Asscher

Fig. 1 Search results and study selection

et al. 2013; Butler et al. 2011; Sawyer and Borduin 2011; Sundell et al. 2008); four others were rated as “good” (Borduin et al. 2009; Letourneau et al. 2009; Ogden and Hagen 2006; Weiss et al. 2013) and three were rated “weak” (Glisson et al. 2010; Henggeler et al. 2006; Timmons-Mitchell et al. 2006).

Referral pathways were frequently poorly described in that a discrepancy existed between eligible participants and those included for randomization. For example, in the United Kingdom, about a quarter of those referred could not be contacted or refused to consent to assessment. This was a similar proportion in the referrals of both Weiss et al. (2013) and Letourneau et al. (2009). It is possible that this proportion may represent those families with more chaotic lives who are less willing to cooperate with services and whose profiles could be substantially different from those who do agree. Furthermore, other staff frequently carried out the initial screening (e.g., social workers, Sundell et al. 2008; probation staff; Henggeler et al. 2006). Various local agencies may have made different judgments on possible

eligibility introducing a level of selection bias among and across sites or referrers. It is therefore not known how representative the samples within the trials may be of the general target population of Multisystemic Therapy.

Appropriate randomization was undertaken in some studies, although there was variety in the method, the point at which this occurred and concealment of allocation. The method of randomization was not reported in three studies (Henggeler et al. 2006; Ogden and Hagen 2006; Weiss et al. 2013). A further two studies used the coin toss (Sawyer and Borduin 2011; Timmons-Mitchell et al. 2006); the validity of which is questionable (Clark and Westerberg 2009). Almost half of the studies did not explicitly state who undertook randomization and how this remained concealed (e.g., Weiss et al. 2013).

It was positive that almost all of the studies explored group differences at baseline on demographic, criminal histories and/or psychosocial characteristics. No statistically significant differences were found for most studies indicating that randomization had been successful (e.g., Asscher et al.

2013). Exceptions to this included Henggeler et al. (2006) and Borduin et al. (2009) where the likely direction of the effect of the bias was assessed as favoring the comparison group. Furthermore, in the Norway randomized control trial (Ogden and Hagen 2006), differences in baseline scores could have overinflated the estimates of treatment effects.

Sample sizes were relatively small with three studies having less than 100 participants (Borduin et al. 2009; Ogden and Hagen 2006; Timmons-Mitchell et al. 2006). Few studies specifically reported on having undertaken power calculations (only Butler et al. 2011; Glisson et al. 2010) making it difficult to assess whether the sample size was adequate to detect a true effect.

Almost half of the included studies provided inadequate descriptions of the comparison condition making it difficult to know what Multisystemic Therapy was being compared with. In almost all of the studies; it was unclear whether the groups had received the same care apart from the intervention being studied.

Some studies reported low rates of missing data (e.g., at 2 year follow-up 94% of participants completed assessment measures, Letourneau et al. 2009); whereas other studies had relatively high rates of non-completion (e.g., in Glisson et al. 2010, just over half of participants (57.0%) completed the 18 month assessment measures). Only one study provided no information on drop outs (Timmons-Mitchell et al. 2006). There was variation across studies in whether analyses had been undertaken to examine the missing and completed data and any possible group differences. This is potentially problematic as participants who drop out may have more significant difficulties and as a consequence treatment effects could be overinflated.

It was positive that all of the studies had follow-up periods of over 1 year. Recidivism rates increase with time indicating the need for long term observation strategies. However, the degree of difference between studies in follow-up periods is potentially problematic. It may be that, for the studies with shorter follow-up periods, recidivism rates would increase as time since assessment/discharge increases. For the two problem sexual behavior studies; the low base rate of sexual recidivism is an inherent difficulty in outcome research. Letourneau et al. (2009) reported only four sexual offences across the sample at the 2 year follow-up.

Most of the studies considered a wide range of outcomes, used reliable measures for their assessment and multiple sources of information. The exception to this was Timmons-Mitchell et al. (2006) where research assistants completed the Child and Adolescent Functional Assessment Scale solely from court records. No inter rater reliability information was reported for the coding process and it is not clear how comprehensive the file information was. Four studies relied on caregivers to report out-of-home placement (Glisson et al. 2010; Henggeler et al. 2006; Letourneau et al.

2009; Ogden and Hagen 2006) which is potentially less reliable as parents may be motivated in various different ways. In America, the vast majority of the studies only examined official data on antisocial behavior within the state (e.g., Borduin et al. 2009) potentially missing a proportion of those adolescents who moved to another state during the follow-up period.

It is positive that seven studies used other sources of information for involvement in antisocial behavior given that many offences are not reported; and even those that are brought to the attention of police may not be officially recorded. For the majority, this involved the Self-Reported Delinquency Scale (Elliott et al. 1985), which has demonstrable reliability and validity across a range of settings. However, this tool focuses on serious criminal behaviors and it is possible that adolescents may incorrectly label certain behaviors thus overstating seriousness. As with all self-report measures; there are issues related to social desirability. This may be particularly relevant given that most of the populations under examination were involved with the justice system and adolescents may be reluctant to provide accurate accounts for fear of further consequences.

In some studies the treatment and comparison groups were assessed on the same schedule (e.g., Weiss et al. 2013); however, in other trials, data was collected at different points (e.g., Timmons-Mitchell et al. 2006). This can be problematic because the outcome data for cases may not be comparable due to the differences in the length of observation periods.

In some studies, there were good descriptions of the care taken to ensure that those collecting the data were blind to the condition (e.g., teachers were informed that questionnaires were for a study on teen socialization in Borduin et al. 2009). However, in other studies, blinding was unclear (e.g., Timmons-Mitchell et al. 2006) or it was stated that those collecting the data were aware of the assignment (e.g., Letourneau et al. 2009). This potentially introduces an element of bias due to people's preconceived beliefs about Multisystemic Therapy and how these may consciously or unconsciously influence their behavior.

Just over half of the studies were assessed as independent of the program developers and their associates (Asscher et al. 2013; Butler et al. 2011; Ogden and Hagen 2006; Sundell et al. 2008; Timmons-Mitchell et al. 2006; Weiss et al. 2013).

Characteristics of Included Studies

Population

Table 2 provides an overview of the characteristics of the 11 included studies. The majority were undertaken in America (Borduin et al. 2009; Glisson et al. 2010; Henggeler et al. 2006; Letourneau et al. 2009; Sawyer and Borduin 2011;

Table 2 Characteristics of included studies

Authors ^a ; date; country; method of randomization	Population N ^a ; source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual services; length	Follow up period	Main findings ^a
Asscher et al. (2014) Netherlands Computerized program	256 (MST = 147; US = 109) Severe antisocial behavior	16.02 years (SD = 1.3) 73% male 55% Dutch; 34% Moroccan; 32% Surinamese	Length not reported Mean TAM score M = 4.36 (SD = 0.51)	Individual treatment (21%); family based (53%); combination (7%); detention (4%); no treatment (14%) Length not reported	Reoffending: 36.7 months (SD = 15.8) Measures: post Tx 5.7 months (SD = 1.90) and follow up 13.0 months (SD = 3.0)	No significant differences in percentage arrested at least once, number of arrests, time to first arrest or type of recidivism MST significantly more effective in reducing externalizing behavior (parent and adolescent report); ODD and CD symptoms; self-reported property offences but not violent offences MST parents reported significantly greater increase in sense of competence; significantly more improvement for positive discipline (caregiver, adolescent and observer report); relationship quality (caregiver and observer report) and inept discipline (observer rated)

Table 2 (continued)

Authors ^a , date; country; method of randomization	Population N ^a , source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual services; length	Follow up period	Main findings ^a
Borduin et al. (2009) America Random number table	48 (MST = 24; US = 24) Juvenile court	14.0 years (SD = 1.0) 95.8% male 72.9% white	30.8 weeks (range 14.3–63.7) Fidelity data not reported	CBT group (90 min 2x week) and individual Tx (1x week) 30.1 weeks (range 17.0–89.9)	Reoffending: 8.9 years (range 7.31–10.64) Measures: within 1 week post Tx	MST had significantly fewer arrests for sexual and other offences, fewer days incarcerated; less self-reported person and property offences Significant group effect on psychiatric symptoms, behavioral difficulties (caregiver and adolescent) and family cohesion and adaptability favoring MST Significant increases in emotional bonding and social maturity; decrease in youth aggression towards peers; increase in youth grades for MST group over US (caregiver and teacher)
Butler et al. (2011) UK Stochastic minimization balanced for offence type gender and ethnicity	108 (MST = 56; US = 52) Court referred	15.2 years (SD = 1.2) 82.8% male 37.3% White British/ European; 33.3% Black African; 24.7% other	20.4 weeks (range 11–30) TAMs not reported	Youth offending services: individually tailored to prevent reoffending Mean contacts 20.88 (SD = 12.88)	Reoffending: 18 months Measures: 4 weeks post Tx	Significant between group differences in mean number of offences only at 18 months Custodial sentences increased in both groups Significantly greater reduction in MST group for adolescent reported offending behaviors Significant increase in favor of MST for positive parenting (parent report only) Significant decrease in antisocial processes in MST group (parent report only) YSR/ABAS/DP/SFIT/CBCL: no significant interactions

Table 2 (continued)

Authors ^a , date; country; method of randomization	Population N ^a , source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual services; length	Follow up period	Main findings ^a
Glisson et al. (2010) America Sequence numbers	674 (MST = 349; US = 325) Juvenile court	14.9 years (SD = 1.59) 69% male 91% White	15.0 weeks (SD = 6.3) TAMs not reported; weekly logs; coding of audiotaped sessions	Inpatient (24%); outpatient (90%); family/partner focused Tx (50%) 186.6 days (SD = 138.3)	Out of home placement: 18 months Measures: 6, 12 and 18 months	Out-of-home placements significantly lower for MST group CBCL: at 18 months no significant between group differences No significant between groups differences in mean number of arrests or days in placement SRD: 3 DC conditions significantly fewer status offences and crimes against the person (not general theft) than FC (adolescent report) CBCL: no significant between groups difference Urine screens: significant difference between groups. DC/MST and DC/MST-CM had significantly lower positive % than DC Form 90: DC/MST and DC/MST-CM reported significantly less alcohol, marijuana and polydrug use than FC
Henggeler et al. (2006) America Not reported	161 (DC and MST = 38; DC and MST-CM = 43; DC = 38; FC = 42) Juvenile court	15.2 years (SD = 1.1) 83% male 67% African American; 31% White	MST 66 h (SD = 32) MST-CM 57 h (SD = 30) CM: voucher system; functional analysis and management planning	Family court (yearly) and group and individual Tx Drug court (weekly) and outpatient Tx Length not reported	Reoffending: 12 month Measures: 4 and 12 months	

Table 2 (continued)

Authors ^a , date; country; method of randomization	Population N ^a , source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual ser- vices; length	Follow up period	Main findings ^a
Letourneau et al. (2009) America Stratified permuted blocks based on index victim age	131 (MST=68; US =63) Court ordered	14.6 years (SD=1.7) 97.6% male 54% Black 44% White	30.9 weeks (SD=12.2) Mean TAM score 3.99 (SD=0.68)	Weekly 60 min group cognitive behavioral Tx Standardized therapeutic activities 54.1 weeks (SD=43.1)	Reoffending and out of home placement: 24 months Measures: 6 and 12 months	No between group dif- ference for nonsexual arrests MST caregivers sig- nificantly less likely to report out-of-home placement SRD: MST youth reported significantly less involve- ment in delinquent behavior No between group differ- ences for self-reported substance use ABSI: Significantly greater reduction in problem sexual behavior over time for MST youth (adoles- cent report only) CBCL: MST group showed significantly greater reduction in externalizing and internalizing symptoms (caregiver and adolescent report) PYS: Significant MST treatment effects on Lax Discipline (youth report) and Bad Friends only

Table 2 (continued)

Authors ^a , date; country; method of randomization	Population N ^a , source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual services; length	Follow up period	Main findings ^a
Ogden and Hagen (2006) Norway Not reported	75 (MST=46; US=29)	15.02 years (SD=1.36) 64.0% male 98.7% Norwegian	Length not reported TAMs not reported	Not reported. Original study: institution (36.9%); crisis (13.2%); social worker care (15.8%); home Tx (18.4%); No Tx (15.8%)	Out of home placement and measures; 24 months	MST less likely to be living out of the family home either at or in the 6 months prior to follow up MST group showed significantly greater reduction in total score and internalizing (not externalizing) symptoms (caregiver report not adolescent). Teacher report for scores (internalizing, externalizing and total) significantly differed favoring MST MST group scored significantly lower on self-reported involvement in delinquency than US
Sawyer and Borduin (2011) America Coin toss	176 (MST=92; IT=84)	14.5 years (SD=1.4) 69.3% male 76.1% White 22.2% African American	20.7 h (SD=7.4) Summary of systems addressed	Individual therapy: psychodynamic, client center or behavioral approaches	Reoffending: 21.9 years (range 18.3–23.8 years)	For all categories of arrests (violent or nonviolent felony) excluding misdemeanors, rates were significantly lower for MST group than IT MST participants were significantly less likely to have been sentenced to prison
Sundell et al. (2008) Sweden Computer generated	156 (MST=79; US=77) Conduct disorder	15.0 years (SD=1.45) 61% male 47% not Swedish: Asia (30); Europe outside of Scandinavia (25) and Africa (14)	20.8 weeks (SD=7.4) Mean TAM score 4.00 (SD=0.61; range 1.73–5.00)	Individual Tx (26.0%); Family Tx (20.8%) Mentor (15.6%) Placed (10.4%) ART (5.2%); Addiction Tx (2.6%); Special education (2.6%); No Tx (16.9%)	Report on police contact and measures 7 months (SD=1.0) and 24 months (SD=0.6)	No significant between group differences on social care and caregiver reports of police involvement, out-of-home placement, self-reported delinquency, CBCL, YSR, SOC, AUDIT and DUDIT, social competence, parenting skill, school attendance or mothers mental health

Table 2 (continued)

Authors ^a , date; country; method of randomization	Population N ^a , source	Mean age; % male; ethnicity	MST: length; fidelity	Comparison: usual services; length	Follow up period	Main findings ^a
Timmons-Mitchell et al. (2006) America Coin toss by court staff	93 (MST=48; US=45) County family court	15.1 years (SD=1.25) 78% Male 77.5% European American 15.5% African American	20.7 weeks (SD=8.7, range 6.1 to 62.6) Mean TAM score 4.2 (SD=0.38)	Not tracked. Probation officers reported that referrals made to drug and alcohol Tx, anger management, individual and family Tx	Reoffending: 24 months Measures: post Tx and 6 month	Recidivism significantly lower for MST group and arraigned for significantly fewer new offenses No significant difference in average time to first arrest or percentage of felonies versus misdemeanors CAFAS: MST scores significantly greater improvement on four of the six subscales: school/work, home, community, moods and emotions No significant difference between groups for felony arrest or time to being arrested CBCL and YSR: rate of reduction in problem behaviors greater for MST group No significant difference for teacher reported problem behavior, SRD, drug use, reintegration to mainstream; grades, number of days absent or suspended, family adaptability or cohesion PAQ: significant curvilinear treatment effect on permissive subscale indicated a greater rate of decrease for MST group
Weiss et al. (2013) America Not reported	164 (MST=84; No Tx=80) Conduct problems in behavior class	14.6 years (SD1.3) 83% male 60% African American 40% European American	Length not reported Mean TAM score=4.41 (SD=0.51) “moderately high to high” Coded sessions	Behavior focused classroom management; usual school and community services	Reoffending: 30 months Reoffending from official sources: charges	

^aPrimary study in bold (see Table 4 in “Appendix” for associated studies); MST Multisystemic Therapy, US Usual services, IT Individual therapy, ART Aggression Replacement Training, Tx Treatment

Timmons-Mitchell et al. 2006; Weiss et al. 2013) and Europe including the Netherlands (Asscher et al. 2013); Sweden (Sundell et al. 2008); the United Kingdom (Butler et al. 2011) and an update to the Norway randomized control trial (Ogden and Hagen 2006).

The participants' characteristics were well described in most studies and it was possible to quantitatively synthesize some demographic information. The sample size ranged from 48 (Borduin et al. 2009) to 674 (Glisson et al. 2010) with over a 100 participants in eight of the studies. The size of the sample across all studies was 2042 adolescents (mean 185.6, SD = 171.5). The mean average age for the whole sample was 14.9 years (SD = 0.5) with Asscher et al. (2013) reporting the highest mean age (16.0 years) and Borduin et al. (2009) reporting the lowest (14.0 years). The samples were predominantly male; percentages of females ranged from 2.4% (Letourneau et al. 2009) to 39% (Sundell et al. 2008). The ethnicity of participants varied and the use of ethnic categories was not consistent across studies.

Samples were recruited from various sources; the majority from youth justice (Borduin et al. 2009; Butler et al. 2011; Glisson et al. 2010; Henggeler et al. 2006; Letourneau et al. 2009; Sawyer and Borduin 2011; Timmons-Mitchell et al. 2006); social care (Ogden and Hagen 2006; Sundell et al. 2008); alternative education (Weiss et al. 2013) or a combination (Asscher et al. 2013). Participant offending history, where reported, varied. The offending histories of participants were typically more severe in the American studies compared with Europe (e.g., Timmons-Mitchell et al. 2006 reported that the average number of all pre-treatment offences was 6.87, SD = 4.4). The exception to this was the problem sexual behavior sample in Letourneau et al. (2009) which appeared less delinquent (62% had no prior general offences).

Intervention

The majority (8) of the included studies examined standard Multisystemic Therapy; a further two studies examined problem sexual behavior and one for contingency management/substance abuse. The average length of treatment was provided by eight studies (excluding Asscher et al. 2013; Ogden and Hagen 2006; Weiss et al. 2013) and this was converted to weeks where possible. Glisson et al. (2010) had the lowest mean treatment length (15.0 weeks, SD = 6.3) and Multisystemic Therapy-problem sexual behavior had the longest (approximately 31 weeks) (Borduin et al. 2009; Letourneau et al. 2009). Only two studies provided information on actual therapy contact. Henggeler et al. (2006) reported that approximately half of therapist contacts were with family; 13% school, 5% peer; 14% youth and 19% community. By contrast, Weiss et al. (2013) reported increased involvement with adolescents (95% individual sessions) and school (94%).

The proportion of participants who successfully completed Multisystemic Therapy was frequently unreported (e.g., Butler et al. 2011; Glisson et al. 2010). Only Borduin et al. (2009) reported 100% treatment completion. Few studies provided clear information about how many of those randomly allocated actually started treatment, were discharged on the mutual agreement with caregivers or prematurely terminated (e.g., only Letourneau et al. 2009; Sundell et al. 2008 gave clear descriptions).

With regard to the therapists, most studies provided information about professional background and demographic characteristics. This ranged from qualification details (e.g., 86% had Masters degrees in Asscher et al. 2013) to specific information about professional background, additional training and years of clinical experience (e.g., Sundell et al. 2008; Butler et al. 2011). Given that this review only included licensed Multisystemic Therapy programs, as standard all therapists would have received the 5 day orientation training; weekly group supervision/consultation and quarterly boosters (Schoenwald 2008).

Every study in this review attempted to measure treatment integrity, most frequently using the Therapist Adherence Measure which is a 26 item Likert scale developed through Multisystemic Therapy expert consensus and validated in two studies (Henggeler et al. 1997, 1999). The measure is collected regularly with each family and is said to measure those therapist features which are specific to the Multisystemic Therapy principles. However, not all of the studies reported on the actual mean adherence score (e.g., Borduin et al. 2009; Butler et al. 2011) and those which did provided little referential information about what the score signified. Only Weiss et al. (2013) reported the mean therapist score to be in the moderately high to high range for adherence to treatment. In many studies; the Therapist Adherence Measure was not available from all families for reasons unreported (e.g., Sundell et al. 2008; Timmons-Mitchell et al. 2006).

Other measures of treatment integrity included independent coding of adherence to the nine Multisystemic Therapy principles from audio taped sessions by consultants (Glisson et al. 2010; Weiss et al. 2013). However, it was not clear whether those involved in the coding were aware of the ongoing research trial. Given that there continues to be a lack of clarity as to how the principles reflect the theory of change and are operationalized in clinical practice, it remains unclear as to whether such codings provide a reliable measure of fidelity to Multisystemic Therapy specifically. Glisson and colleagues (2010) used therapist logs to detail time spent addressing sub-systems (e.g., individual child, family with primary caregiver, school with caregiver). While this may provide a good indication of the multi-modal approach, it does not provide a measure of treatment integrity per se.

Comparison

All of the studies compared Multisystemic Therapy with another approach apart from Weiss et al. (2013). In this trial, the participants were all recruited from a self-contained intervention classroom and so were involved with a behavioral change program at some level.

There was variation in the information provided about comparison groups. The most comprehensive reports were provided by the two problem sexual behavior studies which included the theoretical orientation and content, format, supervision processes and therapist characteristics (Borduin et al. 2009; Letourneau et al. 2009). Other studies provided a general overview (e.g., in the Netherlands usual services consisted of counseling, family based treatments, detention or no treatment, Asscher et al. 2013). Some studies also reported relatively high proportions of comparison group participants to be placed out-of-home (e.g., about a quarter, Glisson et al. 2010; Ogden and Hagen 2006).

Data Synthesis

The majority of the studies examined outcomes across a range of domains including antisocial behavior, drug and alcohol use, adolescent functioning, family functioning, peer relations and school.

Antisocial Behavior

A standard operationalization of antisocial behavior is lacking and outcomes were reported in different ways including official data: arrests/charges (yes or no; rates, time to first arrest and type) and out-of-home placement (yes/no; number of days/years; combining various types such as prison, treatment setting, foster parents, institution or supervised living facility; official and caregiver report). Information was typically obtained from a range of official sources including correctional, probation and police services. In America, only Letourneau et al. (2009) accessed national sources as well as within state data. The follow-up periods for the official recidivism data ranged widely between 1 year (Henggeler et al. 2006) and 21.9 years (Sawyer and Borduin 2011) making it difficult to directly compare rates across studies.

About half of the studies found no significant differences on official measures of antisocial behavior (e.g., Asscher et al. 2014; Henggeler et al. 2006; Letourneau et al. 2013; Löfholm et al. 2009; Weiss et al. 2013) (see Table 6 in “Appendix” for findings from official sources). The study with the longest follow-up period (21.9 years) found significant differences on arrests for both violent and non-violent felonies, years incarcerated but not misdemeanor offences (Sawyer and Borduin 2011). In the United Kingdom, Butler et al. (2011) found significant between group differences in

favor of Multisystemic Therapy for violent and non-violent offences at 18 months follow-up. The charges for which adolescents were formally arraigned were also significantly fewer for the Multisystemic Therapy group in Timmons-Mitchell et al. (2006). Survival analyses indicated that Multisystemic Therapy participants survived for longer without any type of arrest than the comparison group [$\chi^2(1) = 6.06$, $p = 0.01$]. Lastly, Borduin et al. (2009) found that the Multisystemic Therapy-problem sexual behavior group had significantly fewer arrests for sexual crimes, non-sexual crimes and fewer days in custody (effect sizes reported between 0.086 and 0.155 which can be considered small). Survival analysis examining the proportion of participants who were not arrested in each group by time favored treatment efficacy [$\chi^2(1, n = 48) = 8.17$, $p < 0.01$]. Three studies reported a significant difference in out-of-home placement in favor of Multisystemic Therapy based on caregiver report (Glisson et al. 2010; Letourneau et al. 2013; Ogden and Hagen 2006).

Seven studies used the Self-Report Delinquency scale (Elliott et al. 1985). Various subscales were used and reported upon differently (e.g., raw score/t score) thus making comparisons difficult. One trial could not be reported upon as analysis was not undertaken between the Multisystemic Therapy and comparison condition (Henggeler et al. 2006). Of the remaining six, differences between group scores were significant for four studies (Asscher et al. 2014 for property but not violent offences; Borduin et al. 2009; Letourneau et al. 2013; Ogden and Hagen 2006). Both Asscher et al. (2014) and Ogden and Hagen (2006) also reported effect sizes which were 0.37 and 0.26, respectively and could be considered small.

Alcohol and Substance Use

For the study investigating Multisystemic Therapy-contingency management, a significant group and time interaction effect on the self-report Form 90 for alcohol consumption was found ($p = 0.049$) (Henggeler et al. 2006). This indicated that adolescents under drug court for both the Multisystemic Therapy and Multisystemic Therapy-contingency management conditions reported less alcohol use at follow-up (12 months) controlling for baseline scores. This was also indicated for heavy alcohol, marijuana and polydrug use. There was also a significant difference between urine screens (drug court only: 45% positive; drug court and Multisystemic Therapy: 7% and drug court and Multisystemic Therapy-contingency management: 17%, $p < 0.001$) in favor of the two Multisystemic Therapy conditions.

Very few other trials used measures to assess substance use directly. The exception to this was Löfholm et al. (2009) who found that drug related problems reduced over time whilst risky alcohol increased for both groups. At the longest follow-up (2 years), Letourneau et al. (2009) also found

no significant between group differences on self-reported substance use as measured by a subscale of the Personal Experience Inventory.

Adolescent Functioning

Adolescent outcomes were typically gathered from multiple informants including caregivers, adolescents and school staff. A range of measures were used; the most common being the Child Behavior Checklist (Achenbach 1991). Psychiatric symptomology (e.g., Borduin 2009), antisocial beliefs (e.g., Butler et al. 2011), self-esteem (Asscher et al. 2013) and psychopathy (e.g., Asscher et al. 2013) were also examined. Only one study included a measure to assess a protective factor (Sense of Coherence Scale, Asscher et al. 2013).

The caregiver reported Child Behavior Checklist was used by eight studies and significant between group differences were reported for half of these in favor of Multisystemic Therapy (Asscher et al. 2014 (externalizing); Butler et al. 2011 (aggression and delinquency subscales); Ogden and Hagen, 2006 (total) and Weiss et al. 2013 (externalizing)). Findings could not be directly compared because different scales were used; subscales and reporting of scores (raw or t scores). Four studies reported on the effect size which ranged from 0.35 (Weiss et al. 2013) to 0.53 (Asscher et al. 2014), which is in the small to medium range.

The adolescent report version of the Child Behavior Checklist (Youth Self-Report, Achenbach 1991) was used by seven studies. For the Multisystemic Therapy-contingency management trial, the scores were not reported in the “results” section as the t scores for the total were almost the same as the mean for the normative sample (Henggeler et al. 2006). Three of the six studies found a significant positive treatment effect; all on the externalizing scale (Asscher et al. 2014; Letourneau et al. 2009; Weiss et al. 2013). Two studies provided effect sizes that could be considered small (0.39 in Asscher et al. 2014; 0.26 in; Weiss et al. 2013).

Family Functioning

Six of the included studies examined the family domain and this was where the range in measures used varied the most. Areas of parenting (e.g., Butler et al. 2011), quality of parent–youth relationships (e.g., Asscher et al. 2013) and parental mental health (e.g., Borduin et al. 2009) were all explored. Some researchers used adapted scales of various parent and family assessments and combined them together (e.g., Asscher et al. 2013). One study also included an observer assessment of the family examining positive parenting, inept discipline and relationship quality (Asscher et al. 2013).

There were some mixed findings within and between studies and the variety in the measures used makes it challenging to draw any generalizations. For example, Butler

et al. (2011) found that positive parenting increased in the Multisystemic Therapy group and decreased in the control group [ES time(multisystemic therapy) = 0.29, 95% CI = −0.12, 0.72], although this was only for caregiver and not adolescent report. Löfholm et al. (2009) in Sweden found no treatment effect on parenting skills and across both conditions; parenting skill increased significantly over time as reported by caregivers but not by adolescents. In contrast, positive findings were found in Henggeler et al. (2009b) for youth report of Lax Discipline involving an increased follow through on adolescent misbehavior by caregivers in the Multisystemic Therapy group ($p < 0.05$). Only Asscher et al. (2014) found significant improvement for positive discipline for the Multisystemic Therapy group as rated by caregivers, adolescents and observers.

Peers

Peer relationships, most typically association with procriminal peers, were assessed in five of the studies mostly from youth and/or caregiver reports. Some studies found no significant differences in the association with delinquent peers between Multisystemic Therapy and comparison groups (e.g., Butler et al. 2011; Sundell et al. 2008) while others reported favorable treatment effects. This included Henggeler et al. (2009b) who reported that the scores for “Bad Friends” for Multisystemic Therapy participants decreased significantly more over time than for usual services ($p < 0.05$). Borduin et al. (2009) also reported favorable treatment effects for both the parent and teacher composite measure and adolescent report of emotional bonding and social maturity for the Multisystemic Therapy group. Asscher et al. (2014) found significant between group differences on increased contact with prosocial peers in favor of Multisystemic Therapy but not decreased affiliation with deviant peers.

School

Education was the area least considered and was assessed by only three studies involving information about grades, suspensions and attendance (Borduin et al. 2009; Sundell et al. 2008; Weiss et al. 2013). In the alternative education sample, no treatment effects were found on the Teacher externalizing scale, school grades or days suspended (Weiss et al. 2013). Borduin et al. (2009) reported a significant effect for parent and teacher report of grades in favor of Multisystemic Therapy. Lastly, Löfholm et al. (2009) reported no treatment effect for school attendance.

After Multisystemic Therapy

The vast majority of the studies provided no information regarding aftercare services. Only Sundell et al. (2008)

reported that at the 7 month follow-up; two-thirds (66%) of Multisystemic Therapy adolescents were still receiving services and 39% had been re-referred for investigations resulting in new services. This was similar to the comparison group. At the 24 month follow-up, a third of the Multisystemic Therapy group was still receiving services (Löfholm et al. 2009).

Discussion

The current study aimed to investigate the evidence for Multisystemic Therapy, which is an intensive family and community based intervention for young people aged 10–17 years old who are at high risk of offending behavior and/or out-of-home placement. Multisystemic Therapy is currently delivered worldwide across 16 countries and is a recommended intervention for the prevention of serious adolescent offending by a number of governmental departments including the Office of Juvenile Justice and Delinquency Prevention, United States Department of Justice and the National Institute of Clinical Excellence, Department of Health in the United Kingdom. It might be supposed, therefore, that the effectiveness of Multisystemic Therapy has been consistently empirically demonstrated. However, the most recent independent systematic review found inconclusive evidence for the program's effectiveness in comparison to other services (Littell et al. 2005).

The systematic review undertaken by Littell et al. (2005) included trials conducted over 10 years ago, almost all of which had involvement from the program developers in the United States. An updated review of the evidence was therefore considered pertinent, particularly in light of the probable increase in both international and independent research. Developing understanding of interventions that effectively reduce persistent patterns of antisocial behavior among adolescents continues to be an ongoing priority in delinquency prevention policy. The current study, therefore, aimed to identify the highest quality and most recently conducted experimental studies to determine whether Multisystemic Therapy was more effective than treatment as usual/no treatment in addressing both primary and secondary outcomes. The findings of the 11 independent samples (3 updates and 9 new trials) will be discussed in the context of previous research and arranged by program type.

Multisystemic Therapy: Standard

Six new randomized control trials and two updates to previous trials have been conducted to evaluate the effectiveness

of Multisystemic Therapy adding to the previous five randomized control trials (four published, one unpublished) reported by Littell et al. (2005). The six most recent studies were all conducted in real world settings with community practitioners, by research teams not associated with the original developers (excluding Glisson et al. 2010 where the second author is on the Board of Directors). Multisystemic Therapy has now been evaluated in six countries (America, Canada, Netherlands, Norway, Sweden and the United Kingdom) potentially increasing the external validity of findings.

International research is not, however, without challenges. There continues to be debate about the level of adaptation needed to transport Multisystemic Therapy to local contexts and systems (Ogden et al. 2009; Kiddy 2014). Furthermore, the interpretation of findings is complicated by the social, cultural and ethnic factors that are unique to a particular country and influence comparative “usual services” for managing adolescents with antisocial behavior (Epping-Jordan 2004). Aggregating findings can mask real differences in contributing contextual factors. In both the Netherlands and Sweden; adolescents (up to 20 years of age) with antisocial behavior are primarily managed through the child welfare system. Services frequently adopt an in-home and family orientated therapeutic approach. By contrast, out-of-home placements are often the primary intervention in America where young offenders are managed through the justice system. Whilst the legal system in the United Kingdom is potentially more comparable to the United States, placement in secure settings is limited to those most serious young offenders and youth offending services deliver a comprehensive package of support to target those individual and system factors that put an adolescent at risk of (re)offending and build upon protective factors.

Norway was the first European country to implement Multisystemic Therapy and the most recent report at 24 month follow-up showed some positive treatment effects on out-of-home placement, self-reported delinquency and parent/teacher reported adolescent difficult behaviors (Ogden and Hagen 2006). However, one of the four original sites was removed from analyses leading to a substantial reduction in the sample size in the update due to the lack of Therapist Adherence Measure data, for reasons unreported, thereby preventing assessment of treatment integrity.

By contrast, the Swedish trial found that usual services performed equally as well as Multisystemic Therapy. Decreases in adolescent problem behaviors, improvements within family and social skills were observed for both groups (Sundell et al. 2008). The researchers suggest that these findings may be explained by the difference in implementation (Löfholm et al. 2009). In Sweden, the program was guided by local initiatives; whereas in Norway there was a national strategy thus potentially increasing the support given to teams, demonstrating commitment, increasing the

acceptability of the intervention to practitioners and their level of accountability for outcomes.

An alternative explanation is that, in Norway, about 50% of participants in the comparison condition were in residential settings compared with 18% in Sweden. The proportion of participants out-of-home is a complicating factor given that increased contact with other adolescents with risky behaviors may well increase the chance of iatrogenic effects (Dishion et al. 1999), thereby potentially disfavoring the Norwegian comparison group. A further possibility is that usual services in Sweden may be of a higher quality and, therefore, the comparison group experienced more positive outcomes. It is also possible that Multisystemic Therapy was not delivered with adequate fidelity in Sweden where lower mean Therapist Adherence Measure scores by 1 SD were found compared with American studies. Previous research undertaken by the program developers has linked Therapist Adherence Measure scores with positive outcomes (Schoenwald et al. 2005), although this correlation has not been a consistent finding (e.g., studies in the United Kingdom, Sweden and Canada have found no such associations).

Whether the Therapist Adherence Measure itself actually provides a measure of adherence remains contentious (Littell 2006). Sample items such as (“My family and the therapist worked together effectively”) may be related to constructs such as client satisfaction and therapeutic alliance rather than adherence to Multisystemic Therapy principles per se. Furthermore, the Therapist Adherence Measure is a family rated measure; which arguably provides little independent assessment of adherence. Families who experience positive outcomes may quite simply give better feedback. Treatment integrity can be defined as the extent to which an intervention is implemented as intended involving therapist adherence, competence and treatment differentiation (Perepletchikova et al. 2007). In this respect, the Therapist Adherence Measure itself provides little indication about the level of competence with which therapists deliver the multiple components of Multisystemic Therapy. This remains unexplored within the research literature but is important given that therapists are required to be expert in several different therapeutic approaches. It is not clear how this is achieved in clinical practice, especially given that therapists come from a broad range of backgrounds including applied psychology, social work, youth justice, family therapy and nursing (Fox and Ashmore 2014).

Other countries where no significant differences on a range of primary outcomes between Multisystemic Therapy and usual services were found include convictions or out-of-home placement in Canada (Leschied and Cunningham 2002) and for frequency, timing and type of rearrest in the Netherlands (Asscher et al. 2014). In the latter case, small positive treatment effects (ranging from 0.25 to 0.36)

were found for parent and adolescent reported externalizing behavior.

In this review, the three studies with favorable treatment effects on official data were the United Kingdom study (Butler et al. 2011), an American family court (Timmons-Mitchell et al. 2006) and the 21.9 year follow-up by the program developers (Sawyer and Borduin 2011).

In the United Kingdom, Multisystemic Therapy was investigated in an ethnically diverse urban sample and compared with existing youth offending protocols. While in both conditions there were reduced re-offenses and out-of-home placements, there was a significant between group difference in the number of non-violent offences at 18 month follow-up. Consistent with this finding, post treatment adolescent and caregiver reported externalizing behavioral problems showed significantly greater reduction in the Multisystemic Therapy group. No group differences immediately post treatment were found for any of the secondary outcomes (e.g., parental supervision or association with deviant peers). The researchers suggest that changes within these domains may occur later, as with the official data where between-group differences only emerged at the 18 month follow-up. However, given that the assessment measures were only completed at intervention end, it is difficult to draw any conclusions about a possible delay in the effects of treatment.

Butler and colleagues concluded that Multisystemic Therapy adds value to current services in the United Kingdom; however, these findings do need to be interpreted with some caution. The sample size was relatively small ($N = 108$) and underpowered to be able to explore any mechanisms of change contributing to outcomes. The trial was also conducted in two North London boroughs limiting the external validity to other parts of the United Kingdom.

A positive treatment effect on official antisocial behavior was also found by Timmons-Mitchell et al. (2006), which was the first independent replication in America with serious juvenile offenders. Worryingly, however, two-thirds of the adolescents in the Multisystemic Therapy condition still went on to be arrested within the 18 month follow-up period. There were some substantial methodological limitations with the study including the randomization method (coin toss by court personnel), poor description of usual services, collection of data from a single secondary source (court records) and limited examination of the 11% drop outs. Treatment effects are likely to be overinflated when drop outs are not used in analyses as these cases tend to have more negative outcomes; thus, the direction of bias would likely have been in favor of Multisystemic Therapy.

Lastly, in the longest follow-up trial by Sawyer and Borduin (2011), significant differences in arrests for both violent and non-violent felonies, years incarcerated but not misdemeanor offences in favor of Multisystemic Therapy were

found. It is positive that these participants continue to be followed up and such lengthy periods of observation are rare in interventions aimed at reducing antisocial behavior. However, Multisystemic Therapy was compared with individual therapies underpinned by psychodynamic, client centered or behavioral approaches which are unlikely to represent current practices among health, youth justice and social care agencies.

Multisystemic Therapy: Problem Sexual Behavior

This review found two new randomized control trials for Multisystemic Therapy-problem sexual behavior adding to one previously reported upon in Littell et al. (2005). Multisystemic Therapy is quite possibly one of the only programs for adolescents with sexually harmful behavior which has been investigated using a randomized control trial research design (Langstrom et al. 2013). The worthwhile efforts by researchers to overcome some of the significant logistical, legal, and ethical challenges in the pursuit of conducting randomized control trials with this specific population should be noted.

Multisystemic Therapy-problem sexual behavior has only been investigated in America, with oversight by the developers either as the main researchers (e.g., Borduin et al. 1990, 2009) or expert consultants (Letourneau et al. (2009)), thereby limiting the external validity of findings. The study reported in Littell et al. (2005) had a very small sample size ($n = 16$; 8 multisystemic therapy and 8 individual therapy with no sex offender treatment component) (Borduin et al. 1990). Recidivism rates (arrest data from court and police records) at the 3 year (range 21–49 months) follow-up were considerably lower for Multisystemic Therapy adolescents than the comparison group (sexual offences: 12.5 vs. 75%; non-sexual offences 25 vs. 50%).

In this updated review; one of the few studies that found a significant treatment effect of Multisystemic Therapy on antisocial behavior from both official data and self-report was for problem sexual behavior (Borduin et al. 2009). Findings from the assessment measures pre and post treatment indicated that Multisystemic Therapy was more effective in decreasing problem behaviors in youth, improving family relations (cohesion and adaptability), peer relations (emotional bonding and social maturity) and academic performance (improved grades). There was a large observation period (average 8.9 years), which is particularly important given the relatively low base rate for sexual recidivism although the sample size was relatively small ($N = 48$). It is worth noting that the treatment length was the longest compared with the 11 included studies in this review (approximately 31 weeks) and almost twice the average treatment length as stated on the Multisystemic Therapy Services website (approximately 17.4 weeks). Any possible interaction between treatment dose on outcomes is unknown given the highly individualized approach in Multisystemic Therapy.

The second included study investigating Multisystemic Therapy for problem sexual behavior did not report any treatment effect on officially recorded offending (general arrests) but did on self-reported delinquency and out-of-home placement (Letourneau et al. 2013). The 2 year follow-up period is not considered long enough for investigating sex offender treatment (Collaborative Outcome Data Committee 2007). However, this was a larger clinical trial ($N = 128$) and the only one to use community practitioners. Multisystemic Therapy was also reported to have been more effective in decreasing problematic sexual behaviors and externalizing behaviors. It is worth noting that participants were substantially less “delinquent” than in other trials. At baseline, scores on the Child Behavior Check List were in the normal range and various measures were dichotomized due to low incidence, for example, the Self-Report Delinquency Scale and Adolescent Sexual Behavior Inventory. How generalizable these findings may be to the chronic and versatile juvenile offenders which Multisystemic Therapy purports to target is therefore questionable.

Multisystemic Therapy: Contingency Management/ Substance Abuse

This updated review found one randomized control trial conducted specifically for substance abusing adolescents (Henggeler et al. 2006), which adds to one previous study in the review by Littell et al. (2005). Previous research by Henggeler, Pickrel and Brondino (1999) involved a sample of adolescents randomly assigned to Multisystemic Therapy or a community program ($N = 118$). Six months following completion, Multisystemic Therapy participants reported less use of alcohol, marijuana and other drugs than those accessing usual services. Fewer out-of-home placements were also observed. A smaller proportion of the participants ($n = 80$) were followed up for an average of 4 years. Multisystemic Therapy adolescents had fewer convictions and higher levels of abstinence from marijuana as indicated from self-report and urine analysis (abstinence: 55% for multisystemic therapy vs. 28% treatment as usual) (Henggeler et al. 2002).

The randomized control trial in this current review was a relatively complex trial including four conditions, three of which involved drug court (and usual services or and Multisystemic Therapy or and Multisystemic Therapy-contingency management) or family court (Henggeler et al. 2006). The three drug court conditions all appeared to be more effective in decreasing substance use and criminal behavior than family court. Data from urine screens and self-report indicated that adding Multisystemic Therapy or Multisystemic Therapy-contingency management further improved the substance abuse rated outcomes for adolescents but not the criminal or placement outcomes.

The Multisystemic Therapy-contingency management study was part efficacy and effectiveness in that all of the

therapists were employed by the research center and had supervision from the program developers (Henggeler et al. 2006). It remains to be seen how these findings may be replicated in real world settings. Furthermore, it is highly concerning that about three quarters of adolescents were placed out-of-home in the drug court conditions given that “at home” is one of the key measures of program success (drug court 87%; drug court and multisystemic therapy 71%; drug court and multisystemic therapy-contingency management 74%). The researchers argue that the high level of supervision and weekly review involved with the drug court conditions contributed to this finding. The full resource implications of this intensive approach need to be considered and a 5 year follow-up is planned which will help in understanding the sustainability of the findings.

Limitations of Existing Research

The current study identified a number of new trials (nine) and three updates to previous trials investigating the effectiveness of Multisystemic Therapy since the systematic review undertaken by Littell et al. (2005). All of the studies used comparison groups, therefore, any treatment effects must be considered relative to those rather than absolute (Löfholm et al. 2013). What has emerged from the findings is the complexity involved when synthesizing data and drawing conclusions across international contexts. Usual services are heavily influenced by social, legal and political systems; they consist of changing and active approaches influenced by new theory and methods (Löfholm et al. 2013). Relative effects of treatment may also vary over time as community agencies adopt key features of Multisystemic Therapy, most likely to involve an increased emphasis on systemic and community approaches within services for adolescent offenders.

Most studies continue to use relatively small sample sizes; the largest randomized control trial was undertaken in the Appalachian counties ($N = 674$, Glisson et al. 2010). Sample size is important for sufficient statistical power to more reliably detect any possible group differences and begin to explore some of the possible moderating factors, such as gender, age and ethnicity. This would help with identifying those who may benefit the most from Multisystemic Therapy. Further to this, the majority of studies have been undertaken with predominantly male samples. Female delinquents have consistently been found to have specific difficulties including greater levels of sexual abuse victimization and mental health problems (Emeka and Sorensen 2009). It is important for future trials to allow for establishing the effectiveness of Multisystemic Therapy with females, exploring their experience of the program and making adaptations as necessary. Age as a moderating factor also requires further examination given that Multisystemic Therapy views the primary caregiver as the main conduit of change (Henggeler et al. 2009a) and that adolescence is a critical stage

of development where the influence of peers and school factors may become stronger. There is a large developmental gap in the target age range for Multisystemic Therapy (10–17 years) and it is unlikely that models can be uniformly applied to this group as a whole. Greater use of split samples on the basis of participant age may support in identifying discrepancies between adolescents of differing ages.

The outcomes for the proportion of adolescents and their families who drop out of treatment does not appear to have been investigated within Multisystemic Therapy research. The average completion rate has been reported to be 74% (as cited in Sundell et al. 2008), which is comparable to other offending behavior treatment programs (Olver et al. 2011). A consistent finding is that those who do not complete treatment tend to fare worse, therefore, highlighting the value of exploring key predictors of attrition.

On the Multisystemic Therapy Services website, it states that the average length of treatment is up to 60 h of contact provided during a 4-month period (approximately 17.4 weeks). Average treatment length where stated in the included studies for the most part was generally higher than this reported figure. Multisystemic Therapy adopts an individualized approach to meet the needs of young people and families and as such there is no set treatment manual. Treatment length within and across studies varied considerably and the multicomponent nature of Multisystemic Therapy makes it difficult to know what exactly is being delivered, evaluated and how replicable this is in practice. It is difficult to know if there is any interaction between treatment dose on outcomes; and whether a particular number of sessions over a period of time may be associated with successful outcomes. The most widely cited measure to examine treatment integrity was the Therapist Adherence Measure, which is arguably a poor indicator given that it is completed by family members, neglects to consider the competence with which interventions are delivered and has not been established to measure any knowledge or clinical skill specific to Multisystemic Therapy.

It was positive that many of the studies used a number of outcome measures across domains and from multiple sources (caregiver, official, adolescent, teacher, social worker) given that Multisystemic Therapy adopts a social ecological approach. Furthermore, examining the depth and breadth of treatment effects can help to form better decisions. A general critique of the way in which outcomes are assessed in Multisystemic Therapy is that the specific referral problem behaviors (e.g., nonattendance at school; family aggression), treatment goals for each individual case and progress towards these is not reported. This adds to the challenge of how “success” is defined and how changes in the various systems contribute to treatment progress. Given that inherent within the program is building upon individual and system strengths, it was surprising that only Sundell et al. (2008) used a measure to explore a protective factor.

Furthermore, research would benefit from wider consideration of education given its relevance to desistance (Lösel and Bender 2003; Payne et al. 2003); the value of school in measuring adolescent progress and that “in work or school” is one of the routine outcomes gathered by Multisystemic Therapy Services Inc. (<http://www.mstservices.com>).

The lack of carefully defining and tracking other services accessed by participants in the conditions was evident across studies. Given that most samples were referred from justice and/or social care agencies, there are likely to be contacts with other services (e.g., probation officers) but what this involved was often left unreported. Only Weiss et al. (2013) reported that three quarters of participants had received some form of mental health service outside of the project. While this was similar between conditions and assessed as being not significantly related to the primary outcome, access to other services is an important area to explore because Multisystemic Therapy is an intensive, costly resource and efforts should be made to reduce any possible duplication. Furthermore, whether and what contribution other services may make to the outcomes achieved is left unexamined.

Surprisingly, a complete absence of data about the arrangements for aftercare services at treatment end was found across the studies. In Sweden, around one-third of Multisystemic Therapy participants were still receiving services at the 2 year follow-up. The chronicity of conduct disorder is well recognized as are relapse rates for substance misuse. Despite the assertion by Multisystemic Therapy Services that most cases need minimal formal after care services (Multisystemic Therapy Services 2008, p. 1), from the authors own clinical experience it is highly unlikely that adolescents and their families who often have intergenerational dysfunction, trauma and abuse histories and long standing contact with social care or justice services will not need some form of formal aftercare package.

The advantages of Multisystemic Therapy do need recognition. It is one of the most widely evaluated and internationally transported interventions. Over and above effectiveness, the program addresses those known risk factors for reoffending among multiple domains within a structured framework. Multisystemic Therapy is delivered within the adolescents’ natural ecology thus potentially reducing barriers to accessing services and increasing the generalisability of the skills taught. Assessing and promoting treatment fidelity as part of the outcome literature as well as focusing on clinician accountability are all highly valued features.

Study’s Limitations and Strengths

This review has a number of strengths and limitations. Published outcome research has been well documented by Multisystemic Therapy Services (e.g., Multisystemic therapy: research at a glance, 2015) and it is unlikely that

any published studies would have been missed. Comprehensive search terms were used and substantial attempts to find unpublished research from several sources (including dissertation and thesis portals, government websites, searching relevant reference lists and contact with experts). All of the experts who responded were not aware of any further research which might be relevant.

It is recognized that the search process is not without bias. Firstly, due to time constraints and resources, only research in English was included. With a movement toward publishing research in English, the risk of language bias likely presents as less of a potential issue (Higgins and Green 2011). During the search, reference lists were scanned and it is acknowledged that the sole use of titles to identify articles of potential relevance involved some level of subjectivity. The expansion of the search in this way relied on the reference lists of the shortlisted articles, regardless of how the article itself had been identified.

Bias is also evident in the use of pre-defined criteria to establish which studies to include within this review. It could be argued that high criteria were set for inclusion (i.e., randomized control trials), which may be to the detriment of considering equally valuable studies. This includes a number of quasi experimental designs investigating Multisystemic Therapy (e.g., Painter 2009) as well as benchmarking studies (e.g., Curtis et al. 2009). Furthermore there are a few published and unpublished qualitative studies that make important contributions to the knowledge base, for example, investigating client or therapist experience of Multisystemic Therapy (e.g., Tighe et al. 2012; Markham 2016).

This review focused on randomized control trials due to various factors that can affect nonrandom allocation potentially predisposing the treatment group to better or worse outcomes. One good example of this is a trial conducted in Washington State where allocation to Multisystemic Therapy was left to the discretion of court personnel or inappropriately based on case numbers (Barnoski 2004, sample size $N = 145$). The review found no treatment effect for recidivism data at 18 month follow-up, however, on examination of the data, participants who had been allocated to Multisystemic Therapy scored significantly higher on the risk assessment tool used at baseline. The conclusions were that the validity of the trial had been compromised and a re-evaluation was recommended. The stringent inclusion criteria, therefore, served to minimize the risk of any inadvertent bias in allocation and increase the chances that included studies were appropriate and measured similar concepts.

Conclusion

The current study provided an update to a previous systematic review by Littell et al. (2005) of the most recently

conducted randomized control trials for the effectiveness of Multisystemic Therapy. Despite the rapid international expansion of Multisystemic Therapy, it would seem that the evidence base continues to present conflicting evidence. Some studies found that Multisystemic Therapy had a positive treatment effect on official measures of antisocial behavior, self-reported involvement in delinquency, caregiver report of externalizing behavior problems and affiliation with antisocial peers. However, this review found that the findings were neither consistent across studies or within studies on the various measures used to assess outcomes.

Empirical investigations of Multisystemic Therapy using randomized control trials have now been undertaken in six countries. All of the studies used comparison groups, therefore, any treatment effects must be considered relative to those rather than absolute (Löfholm et al. 2013). This review has highlighted that the impact of the cultural, legal and political differences across countries upon the usual treatment condition and outcome measures needs to be taken into account when synthesizing data and drawing generalised conclusions. This is especially relevant for the Scandinavian countries where adolescents are typically managed through the child welfare system.

A further confounding factor is that samples sizes continue to be relatively small (8 of the 11 included studies had just over 100 participants and this exceeded 200 in only 2 studies) resulting in a lack of power in findings in the literature. Four studies made some attempt to examine subgroups of participants (e.g., age, gender and ethnicity) to explore who may benefit the most from Multisystemic Therapy; but subgroup sample sizes were very small. The majority of studies have been undertaken with predominantly male samples and are often too small to be able to fully explore any interactions between participant characteristics and outcome. This would help with more specifically answering the question of

who may benefit the most which is essential to support the prioritization of this relatively expensive and intensive program. Greater consistency in the descriptions of samples, the demographic factors and specific cultural information from which they are drawn is recommended and would enable comparisons to be more easily made.

In conclusion, there is a considerable amount of research literature available on Multisystemic Therapy, which perhaps leads some to erroneously presume that the empirical support has been consistently demonstrated and that those families for which the program may be more successful is well known. Multisystemic Therapy research has been driven by examining effectiveness and the current study indicates that greater effort needs to be made to investigate both moderating factors and the mechanisms of change. This review has demonstrated the complexity of comparing randomized control trials across international contexts and identified that there is much work to be done in terms of understanding why Multisystemic Therapy might work and under which conditions it may be most successful.

Compliance with Ethical Standards

Conflict of interest The author reports no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Tables 3, 4, 5 and 6.

Table 3 Included studies in the systematic review undertaken by Littell et al. (2005)

Author(s)	Title
Borduin et al. (1990)	Multisystemic treatment of adolescent sexual offenders
Borduin et al. (1995)	Multisystemic treatment of serious juvenile offenders: long-term prevention of criminality and violence
Henggeler et al. (1992)	Family preservation using multisystemic therapy: an effective alternative to incarcerating serious juvenile offenders
Henggeler et al. (1997)	Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination
Henggeler et al. (1999)	Multisystemic treatment of substance-abusing and -dependent delinquents: outcomes, treatment fidelity, and transportability
Leschied and Cunningham (2002)	Seeking effective interventions for young offenders: Interim results of a 4-year randomized study of multisystemic therapy in Ontario
Ogden and Halliday-Boykins (2004)	Multisystemic treatment of antisocial adolescents in Norway: replication of clinical outcomes outside of the US

Henggeler et al. (1999) has been not been included in this table as multisystemic therapy was offered as an alternative to the hospitalization of adolescents in psychiatric crisis

Table 4 Quality checking of included studies

Author(s)	Appropriate randomization	Adequate concealment	Comparable groups at baseline	Adequate sample size	Well described conditions	Same care	Missing data appropriately handled	Equal follow up or adjusted	Available outcome data	Intent to treat analysis	Follow up at least a year	All outcomes considered	Reliable measures	Researchers blind	Author(s) independent	Overall rating
Asscher et al. (2013)	Y	N	Y	U	N	Y	Y	Y	Y	Y	Y	Y	U	U	Y	Strong
Borduin et al. (2009)	Y	Y	N	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Good
Butler et al. (2011)	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	U	Y	Strong
Glisson et al. (2010)	Y	Y	U	Y	N	U	U	Y	U	U	Y	N	N	N	N	Weak
Henggeler et al. (2006)	Y	U	N	N	Y	U	Y	Y	Y	Y	Y	Y	Y	U	N	Weak
Letourneau et al. (2009)	Y	Y	Y	U	Y	N	U	Y	Y	Y	Y	Y	Y	N	N	Good
Ogden and Hagen (2006)	U	Y	Y	U	N	U	N	Y	Y	Y	Y	Y	Y	U	Y	Good
Sawyer and Borduin (2011)	Y	U	Y	U	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	N	Strong
Sundell et al. (2008)	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Strong
Timmons-Mitchell et al. (2006)	U	U	Y	U	N	U	U	Y	U	N	Y	Y	N	U	Y	Weak
Weiss et al. (2013)	U	U	Y	U	N	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Good

Henggeler et al. (1999) has been not included in this table as multisystemic therapy was offered as an alternative to the hospitalization of adolescents in psychiatric crisis

Table 5 Articles accessed in full and meeting inclusion criteria

Author(s)	Title	Additional notes
Asscher et al. (2013)	A randomized controlled trial of the effectiveness of multisystemic therapy in the Netherlands: post-treatment changes and moderator effects	Netherlands randomized control trial with four associated publications
Asscher et al. (2014)	Sustainability of the effects of multisystemic therapy for juvenile delinquents in the Netherlands: effects on delinquency and recidivism	
Manders et al. (2013)	Psychopathy as predictor and moderator of multisystemic therapy outcomes among adolescents treated for antisocial behavior	
Dekovic et al. (2012)	Within-intervention change: mediators of intervention effects during multisystemic therapy	
Borduin et al. (2009)	A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: effects on youth social ecology and criminal activity	Multisystemic therapy-problem sexual behavior
Butler et al. (2011)	A randomized controlled trial of multisystemic therapy and a statutory therapeutic intervention for young offenders	United Kingdom randomized control trial
Glisson et al. (2010)	Randomized trial of multisystemic therapy and ARC in a two-level evidence-based treatment implementation strategy	Rural Appalachian counties
Henggeler et al. (2006)	Juvenile drug court: enhancing outcomes by integrating evidence-based treatments	Multisystemic therapy-contingency management
Letourneau et al. (2009)	Multisystemic therapy for juvenile sexual offenders: 1-year results from a randomized effectiveness trial	Multisystemic therapy-problem sexual behavior with three associated publications
Letourneau et al. (2013)	2-year follow-up of a randomized effectiveness trial evaluating multisystemic therapy for juveniles who sexually offend	
Henggeler et al. (2009b)	Mediators of change for multisystemic therapy with juvenile sexual offenders	
Sundell et al. (2008)	The transportability of multisystemic therapy to Sweden: short-term results from a randomized trial of conduct-disordered youths	Sweden randomized control trial with two associated publications
Löfholm et al. (2009)	Multisystemic therapy with conduct disordered youth: stability of treatment outcomes 2 years after intake	
Timmons-Mitchell et al. (2006)	An independent effectiveness trial of multisystemic therapy with juvenile justice youth	First American independent
Weiss et al. (2013)	An independent randomized clinical trial of multisystemic therapy with non-court-referred adolescents with serious conduct problems	Alternative education setting
Ogden and Hagen (2006)	Multisystemic treatment of serious behavior problems in youth: sustainability of effectiveness 2 years after intake	Update to Norway randomized control trial
Schaeffer and Borduin (2005)	Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders	Two updates to original study by Borduin et al. (1995)
Sawyer and Borduin (2011)	Effects of multisystemic therapy through midlife: a 21.9-year follow-up to a randomized clinical trial with serious and violent juvenile offenders	

Table 6 Findings from official records of antisocial behavior and out-of-home placement

Authors (mean follow up)	Primary outcome				Sig: effect size where reported		
	Measurement	Multisystemic Therapy		Comparison			
Asscher et al. (2014) ^a (36.7 months)	Percentage arrested at least once (%)	58%		49%	NS*		
	Violent recidivism	50%		41%	NS		
	Number of arrests	1.12		1.22	NS		
	Time until first arrest	8.28 (6.24)		8.16 (8.69)	NS		
Borduin et al. (2009) (8.9 years)	Arrests: sexual offence (%)	8.3%		45.8%	p < .001		
	Sexual crimes: mean count (SD)	0.13 (0.34)		0.79 (1.02)	p < .001 Effect size 0.155		
	Arrests: non sexual offence (%)	29.2%		58.3%	p < .001		
	Other crimes: mean count (SD)	1.46 (3.27)		4.88 (8.25)	p < .001 Effect size 0.037		
	Incarcerated: mean number of days (SD)	393.42 (1221.11)		1942.50 (3121.04)	p < .001 Effect size 0.086		
	Butler et al. (2011) (18 months)	Recidivism rate (%): all offences (violent and nonviolent)	8%		36%	p = .001	
Recidivism rate (%): nonviolent		8%		34%	p = .005		
Percentage that committed no offence		90%		63%	p < .001		
Last 6 months of the study only	Average count all offences mean (SD)	0.10 (0.36)		0.51 (0.78)	p = .02		
	Youths in custody	10%		17%	NS		
Glisson et al. (2010) (18 months)	Out-of-home placement	Not reported: 23% of overall sample			p = .019		
Henggeler et al. (2006) (12 months)	Four Tx conditions	Drug court		Only DC	FC		
		Multisystemic therapy		Multisystemic therapy-CM			
	Arrests: mean count (SD)	1.40 (1.52)		1.28 (1.44)	1.45 (1.35)	1.00 (1.15)	NS
	Days in placement	67 (127)		60 (85)	95 (114)	76 (123)	NS
	Out-of-home placement (%)	71%		74%	87%	55%	NS
Letourneau et al. (2013) ^a (24 months)	Arrests for nonsexual offences mean (SD)	2.36 (3.56)		2.59 (4.41)		NS	
	Out-of-home placement mean days (SD)	309.46 (89.64)		323.18 (80.82)		p < .05	
Löfholm et al. (2009) ^a (24 months)	Arrest 6 months prior to follow up	33%		23%		NS	
	Out-of-home placement	40%		38%		NS	
	Out-of-home placement mean days (SD) over	112.9 (21.4)		125.4 (25.4)		NS	
Ogden and Hagen (2006) ^a (24 months)	Out-of-home at follow up	26%		45%		p = .08	
	Out-of-home previous 6 months to follow up	28%		48%		p = .07	

Table 6 (continued)

Authors (mean follow up)	Primary outcome			Sig: effect size where reported
	Measurement	Multisystemic Therapy	Comparison	
Sawyer and Borduin (2011) ^a (21.9 years)	Arrest for at least one felony	34.8%	54.8%	p = .01 Effect size 0.62
	Arrest for at least one violent felony	4.3%	15.5%	p = .02 Effect size 1.32
	Arrest for at least one nonviolent felony	34.8%	51.2%	p = .03 Effect size 0.51
	Arrest for at least one misdemeanor	60.9%	65.5%	NS
	Felony offences mean (SD)	1.53 (2.94)	1.99 (2.61)	p < .05
	Misdemeanor offences mean (SD)	2.39 (3.15)	3.58 (6.16)	NS
	Incarceration (years)	5.25 (12.91)	7.88 (13.78)	p < .05
	Probation (years)	1.72 (2.64)	1.85 (2.72)	
Timmons-Mitchell et al. (2006) (24 months)	Charges for which youth formally arraigned (%)	66.7%	86.7%	p < .05
	Felony charges (%)	12.5%	10.3%	NS
	Time to first arrest mean days	135	117	N
	Mean new offences	M = 1.44 (SD = 1.5)	M = 2.29 (SD = 1.5)	p < .01
Weiss et al. (2013) (2.5 years)	Criminal charges (felony or misdemeanor) (%)	45%	45%	NS
	Time to arrest	Not reported		NS

^aFindings from most recent publication with sample

NS difference not significant

References

* Primary studies and associated publications used in the review

- Achenbach, T. (1991). *Manual for the Child Behavior Checklist/4–18*. Burlington: University of Vermont Department of Psychiatry.
- *Asscher, J. J., Dekovic, M., Manders, W., van der Laan, P. H., Prins, P. J. M., & van Arum, S. (2014). Sustainability of the effects of multisystemic therapy for juvenile delinquents in the Netherlands: effects on delinquency and recidivism. *Journal of Experimental Criminology, 10*, 227–243.
- *Asscher, J. J., Dekovic, M., Manders, W. A., van der Laan, P. H., & Prins, P. J. M. (2013). A randomized controlled trial of the effectiveness of multisystemic therapy in the Netherlands: post-treatment changes and moderator effects. *Journal of Experimental Criminology, 9*, 169–187.
- Barnoski, R. (2004). Providing evidence-based programs with fidelity in Washington State juvenile courts: cost analysis. Washington State Institute for public policy. Retrieved June 23, 2014, from <http://www.wsipp.wa.gov/ReportFile/1058>.
- Blackburn, R. (2003). *The psychology of criminal conduct: theory, research and practice*. Chichester: Wiley.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34*, 105–114.
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology, 63*, 569–578.
- *Borduin, C. M., Schaeffer, C. M., & Heiblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: effects on youth social ecology and criminal activity. *Journal of Consulting and Clinical Psychology, 77*, 26–37.
- *Butler, S., Baruch, G., Hickley, N., & Fonagy, P. (2011). A randomized controlled trial of multisystemic therapy and a statutory therapeutic intervention for young offenders. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*, 1220–1235.
- Centre for Reviews and Dissemination (2009) Systematic reviews: CRD's guidance for undertaking reviews in health care. University of York, York.
- Clark, M. P. A., & Westerberg, B. D. (2009). How random is the toss of a coin? *Canadian Medical Association Journal, 181*, E306–E308.
- Collaborative Outcome Data Committee. (2007). *The Collaborative Outcome Data Committee's guidelines for the evaluation of sexual offender treatment outcome research Part 2: CODC guidelines (corrections research user report no. 2007-03)*. Ottawa: Public Safety Canada.

- Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic treatment: a meta-analysis of outcome studies. *Journal of Family Psychology, 18*, 411–419.
- Curtis, N. M., Ronan, K. R., Heiblum, N., & Crellin, K. (2009). Dissemination and effectiveness of Multisystemic Treatment in New Zealand: a benchmarking study. *Journal of Family Psychology, 23*(2), 119–129.
- *Dekovic, M., Asscher, J. J., Manders, W. A., Prins, P. J. M., & van der Laan, P. (2012). Within-intervention change: mediators of intervention effects during multisystemic therapy. *Journal of Consulting and Clinical Psychology, 80*, 574–587.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm—peer groups and problem behavior. *American Psychologist, 54*, 755–764.
- Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Self-report delinquency scale in explaining delinquency and drug use*. Beverly Hills: Sage.
- Emeka, T., & Sorensen, J. (2009). Female juvenile risk: Is there a need for gendered assessment instruments? *Youth Violence and Juvenile Justice, 7*, 313–330.
- Epping-Jordan, J. E. (2004). Research to practice: international dissemination of evidence-based behavioral medicine. *Annals of Behavioral Medicine, 28*(2), 81–87.
- Fox, S., & Ashmore, Z. (2014). Multisystemic therapy as an intervention for young people on the edge of care. *The British Journal of Social Work, 45*(7), 1–17.
- *Glisson, C., Schoenwald, S. K., Hemmelgarn, A. L., Green, P. D., Dukes, D., Armstrong, K. S., & Chapman, J. E. (2010). Randomized trial of multisystemic therapy and ARC in a two-level evidence-based treatment implementation strategy. *Journal of Consulting and Clinical Psychology, 78*(4), 537–550.
- Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: a multisystemic approach to treating the behavior problems of children and adolescents*. Pacific Grove: Brooks/Cole.
- Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow up of multisystemic therapy with substance abusing and dependent juvenile offenders. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*, 868–874.
- *Henggeler, S. W., Halliday-Boykins, C. A., Cunningham, P. B., Randall, J., Shapiro, S. B., & Chapman, J. E. (2006). Juvenile drug court: enhancing outcomes by integrating evidence-based treatments. *Journal of Consulting and Clinical Psychology, 74*(1), 42–54.
- *Henggeler, S. W., Letourneau, E. J., Chapman, J. E., Borduin, C. M., Schewe, P. A., & McCart, M. R. (2009b). Mediators of change for multisystemic therapy with juvenile sexual offenders. *Journal of Consulting and Clinical Psychology, 77*, 451–462.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65*, 821–833.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: an effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60*, 953–961.
- Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance abusing and dependent delinquents: outcomes, treatment fidelity, and transportability. *Mental Health Services Research, 1*, 171–184.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009a). *Multisystemic therapy for antisocial behavior in children and adolescents*. New York: Guildford.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., & Swenson, C. C. (2006). The littell paper: methodological critique and meta-analysis as Trojan horse. *Children and Youth Services Review, 28*(4), 447–457.
- Higgins, J. P., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions*. Retrieved May 26, 2014, from <http://handbook.cochrane.org/>.
- Juni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *The Journal of the American Medical Association, 282*, 1054–1060.
- Kiddy, C. (2014). Exploring the tension between adherence and cultural fit when delivering multisystemic therapy in England. Unpublished doctoral thesis, Royal Holloway, University of London.
- Koops, W., & Orobio de Castro, B. (2004). The development of aggression and its linkages with violence and youth delinquency. *European Journal of Developmental Psychology, 1*, 241–269.
- LaFavor, T., & Randall, J. (2013). Family therapy review: Contrasting contemporary models. In A. Rambo, C. West, A. Schooley, & T. Boyd (Eds.), *Multisystemic family therapy* (pp. 100–103). New York: Routledge/Taylor Francis Group.
- Langstrom, N., Enebrink, P., Lauren, E.-M., Lindblom, J., Werko, S., & Hanson, R. K. (2013). Preventing sexual abusers of children from reoffending: systematic review of medical and psychological interventions. *British Medical Journal, 347*, 4630.
- Leschied, A. W., & Cunningham, A. (2002). *Seeking effective interventions for young offenders: interim results of a four year randomized study of multisystemic therapy in Ontario, Canada*. London: Centre for Children and Families in the Justice System.
- *Letourneau, E. J., Henggeler, S. W., Borduin, C. M., Schewe, P. A., McCart, M. R., Chapman, J. E., & Saldana, L. (2009). Multisystemic therapy for juvenile sexual offenders: 1-year results from a randomized effectiveness trial. *Journal of Family Psychology, 23*, 89–102.
- *Letourneau, E. J., Henggeler, S. W., McCart, M. R., Borduin, C. M., Schewe, P. A., & Armstrong, K. S. (2013). Two-year follow-up of a randomized effectiveness trial evaluating multisystemic therapy for juveniles who sexually offend. *Journal of Family Psychology, 27*, 978–985.
- Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works: reducing reoffending—guidelines from research and practice* (pp. 63–78). Chichester: Wiley.
- Littell, J. H. (2006). The case for multisystemic therapy: evidence or orthodoxy? *Children and Youth Services Review, 28*(4), 458–472.
- Littell, J. H., Campbell, M., Green, S., & Toews, B. (2005). *Multisystemic therapy for social, emotional, and behavioral problems in youth aged 10–17*. Cochrane Database of Systematic Reviews.
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis. Pocket guides to social work research methods*. New York: Oxford University Press.
- Loeber, R., & Farrington, D. P. (1998). *Serious and violent juvenile offenders: risk factors and successful interventions*. Thousand Oaks: Sage.
- Löfholm, C. A., Brännström, L., Olsson, M., & Hansson, K. (2013). Treatment-as-usual in effectiveness studies: what is it and does it matter? *International Journal of Social Welfare, 22*, 25–34.
- Löfholm, C. A., Olsson, M., Sundell, K., & Hansson, K. (2009). MST with conduct disordered young people: Stability of treatment outcomes two years after intake. *Evidence & Policy: A Journal of Research, Debate and Practice, 5*(4), 373–397.
- Lösel, F., & Bender, D. (2003). Protective factors and resilience. In D. P. Farrington & J. W. Coid (Eds.), *Early prevention of adult antisocial behavior* (pp. 130–204). Cambridge: Cambridge University Press.
- *Manders, W. A., Dekovic, M., Asscher, J. J., van der Laan, P. H., & Prins, P. J. M. (2013). Psychopathy as predictor and moderator of multisystemic therapy outcomes among adolescents treated

- for antisocial behavior. *Journal of Abnormal Child Psychology*, 41, 1121–1132.
- Markham, A. (2016). Multisystemic therapy: therapist experience of programme delivery, processes and outcomes. Unpublished doctoral thesis, University of Birmingham.
- Multisystemic Therapy Services. (2008). Required and recommended program practices and characteristics: program review form and rationale statements. Retrieved July 28, 2014, from <http://www.mstservices.com/Requirement%20Rationale.doc> 17/05/2014.
- Multisystemic Therapy Services (2015). Multisystemic therapy: research at a Glance. Retrieved July 29, 2014, from <http://mst-services.com/files/outcomestudies>.
- National Institute for Health and Care Excellence. (2012). Process and methods guides. The guidelines manual: appendices B–I. Retrieved May 26, 2014, from <http://publications.nice.org.uk/pmg6b>.
- Ogden, T., Hagen, Askeland, & Christensen (2009). Implementing and evaluating evidence-based treatments of conduct problems in children and youth in Norway. *Research on Social Work Practice*, 19(5), 582–591.
- *Ogden, T., & Hagen, K. A. (2006). Multisystemic therapy of serious behavior problems in youth: Sustainability of therapy effectiveness two years after intake. *Journal of Child and Adolescent Mental Health*, 11(3), 142–149.
- Ogden, T., & Halliday-Boykins, C. A. (2004). Multisystemic treatment of antisocial adolescents in Norway: replication of clinical outcomes outside of the US. *Child and Adolescent Mental Health*, 9, 77–83.
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology*, 79, 6–21.
- Painter, K. (2009). Multisystemic therapy as community based treatment for youth with severe emotional disturbance. *Research on Social Work Practice*, 19, 314–324.
- Payne, A., Gottfredson, D., & Gottfredson, G. (2003). Schools as communities: The relationships among communal school organisation, student bonding, and social disorder. *Criminology*, 41, 749–777.
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829–841.
- Peto, R., Collins, R., & Gray, R. (1995). Large-scale randomized evidence: large, simple trials and overviews of trials. *Journal of Clinical Epidemiology*, 48, 23–40.
- Petrosino, A., & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1(4), 435–450.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: a practical guide*. Oxford: Blackwell
- Rennie, L. C., & Dolan, M. C. (2010). The significance of protective factors in the assessment of risk. *Criminal Behavior and Mental Health*, 20, 8–22.
- *Sawyer, A. M., & Borduin, C. M. (2011). Effects of multisystemic therapy through midlife: a 21.9-year follow-up to a randomized clinical trial with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, 79, 643–652.
- *Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, 73 (3), 445–453.
- Schoenwald, S. K. (2008). Toward evidence-based transport of evidence based treatments: multisystemic therapy as an example. *Journal of Child & Adolescent Substance Abuse*, 17(3), 69–91.
- Schoenwald, S. K., Letourneau, E. J., & Halliday-Boykins, C. (2005). Predicting therapist adherence to a transported family-based treatment for youth. *Journal of Clinical Child and Adolescent Psychology*, 34, 658–670.
- Shader, M. (2002). *Risk factors for delinquency: an overview*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- *Sundell, K., Hansson, K., Löfholm, C. A., Olsson, T., Gustle, L. H., & Kadesjo, C. (2008). The transportability of multisystemic therapy to Sweden: short-term results from a randomized trial of conduct disordered youth. *Journal of Family Psychology*, 22, 550–560.
- Tighe, A., Pistrang, N., Casdagli, L., Baruch, G., & Butler, S. (2012). Multisystemic therapy for young offenders: families' experiences of therapeutic processes and outcomes. *Journal of Family Psychology*, 26, 187–197.
- *Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. *Journal of Clinical Child and Adolescent Psychology*, 35(2), 227–236.
- Tripodi, S., & Bender, K. (2011). Substance abuse treatment for juvenile offenders: a review of quasi-experimental and experimental research. *Journal of Criminal Justice*, 39(3), 246–252.
- Walton, A. (2012). Antisocial personality disorder: a systematic review of indicated preventative interventions in adolescence. Peterborough Youth Offending Service.
- *Weiss, B., Han, S., Harris, V., Catron, T., Ngo, V. K., Caron, A., Gallop, R., & Guth, C. (2013). An independent randomized clinical trial of multisystemic therapy with non-court-referred adolescents with serious conduct problems. *Journal of Consulting and Clinical Psychology*, 81, 1027–1039.