Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): 5-year follow-up of a pragmatic, randomised controlled, superiority trial

Peter Fonagy[1]*, Stephen Butler[1], David Cottrell[2], Stephen Scott[3], Stephen Pilling[1], Ivan Eisler[3], Peter Fuggle[4], Abdullah Kraam[5], Sarah Byford[3], James Wason[6,7], Jonathan A. Smith[9], Alisa Anokhina[1], Rachel Ellison[1], Elizabeth Simes[1], Poushali Ganguli[3], Elizabeth Allison[1], Ian M Goodyer[8]

[1]Research Department of Clinical, Educational and Health Psychology, University College London, London, UK

[2]Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

[3]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

[4]Anna Freud National Centre for Children and Families, London, UK

[5]University of Leeds and Rotherham Doncaster and South Humber NHS Foundation Trust, UK

[6]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

[7]MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

[8]Department of Psychiatry, University of Cambridge, Cambridge UK

[9]Department of Psychological Sciences, Birkbeck College, University of London, UK

*Corresponding Author: p.fonagy@ucl.ac.uk

SUMMARY

Background

Multisystemic therapy is a manualised treatment programme for young people who exhibit antisocial behaviour. The Systemic Therapy for At Risk Teens trial is the first large-scale randomised controlled trial of multisystemic therapy in the UK. Previous findings reported to 18 months post-baseline did not indicate superiority of multisystemic therapy compared with management as usual. Here, we report outcomes of the trial to 60 months.

Methods

Young people aged 11–17 years with moderate-to-severe antisocial behaviour were randomly allocated to management as usual (n=342) or 3–5 months of multisystemic therapy followed by management as usual (n=342). The primary outcome was proportion of offences with convictions in the groups**.**

Findings

By 60 months' follow-up, 55% of the multisystemic therapy group had at least one offence with a criminal conviction, compared with 53% of the management-as-usual group (odds ratio 1·13, 95% CI: 0·82, 1·56; p=0·44).

Interpretation

The results of the 5-year follow-up show no evidence of longer-term superiority for multisystemic therapy compared with management as usual.

Funding

National Institute for Health Research Health Services and Delivery Research programme.

Introduction

Antisocial behaviour is a key component of conduct disorder, a clinical syndrome that can have significant personal, interpersonal and societal costs.[1] Young people with conduct disorder are at risk of developing persistent long-term psychological and behavioural problems.[2] There were an estimated 1,240,000 cases of conduct disorder in England in 2014[3]. The lifetime savings associated with preventing conduct disorder by early intervention are estimated to be £260,000 per severe case.[4] The burden of antisocial behaviour is long-term for both young people and the communities in which they live; thus, it is essential to understand the long-term efficacy of interventions in order to appraise their value beyond limiting the strain on current services.

Multisystemic therapy (MST) was initially designed in the USA as an intervention for families with young people who engage in antisocial behaviour and are at risk of becoming young offenders. It is an intensive, family-focused programme that helps young people manage their behaviour in various contexts, including at home, at school, and in the community.[5] High-quality, quantitative systematic reviews showed that MST helps to reduce adolescent antisocial and offending behaviour and improves individual and family problems, but the majority of studies with positive results are from the USA, and replications in other countries have had mixed outcomes[6]. All longer-term follow-up studies are based on US samples and carried out by the developers of the treatment; these typically report short follow-up periods for secondary outcomes other than criminal behaviour.[7,8] In essence, the longer-term outcomes of MST are not known.

The Systemic Therapy for At Risk Teens (START) trial was a pragmatic randomised controlled trial evaluating the effectiveness of MST, implemented to international treatment fidelity standards, compared with management as usual (MAU) at nine pilot sites across

England. MAU involved standard local care offered by a range of services, including mental health, juvenile justice, social care and education, in varying combinations across settings. Overall, 684 families took part in the trial, with 50% (n=342) assigned to MST and 50% (n=342) assigned to MAU. Full demographic characteristics at baseline are shown in Table 1.

The outcomes of the START trial to 18 months were mixed.[9] The primary outcome selected by commissioners of the trial was the proportion of out-of-home placement because of its association with high costs and poor long-term outcomes for young people. The results indicated that out-of-home placements were infrequent (approximately 20%) in both groups, with no significant difference between the two groups. Similarly, although offending decreased over time in both groups, there was no evidence to suggest that MST was superior to MAU in reducing criminal behaviour, and the mean number of recorded offences was slightly higher in the MST group than the MAU group at 18 months. Secondary outcomes were more promising, consistently suggesting that parents in the MST group perceived greater improvements to young people's antisocial behaviour compared with those in the MAU group. Parents also saw improvements in the young people's mental health, mood, and family functioning, and reported positive changes to their own parenting strategies. These findings, however, were not sustained at 18 months, and the overall results of the trial did not show any clear evidence of superiority of MST over MAU.

The purpose of this follow-up study is to evaluate the medium- to long-term effectiveness of MST compared with MAU. The follow-up period was 60 months for the primary outcome (the proportion of young people with any criminal conviction) and 48 months for secondary outcomes (psychiatric problems and areas in which conduct disorder is likely to result in poorer outcomes, including educational and work attainment, social relationships, pregnancy and physical health).

Methods

Study design and participants

The START study was a pragmatic trial to determine the superiority of brief MST followed by MAU compared with MAU alone in addressing antisocial behaviour in young people. The study design and procedures are fully described in the published trial protocol (https://www.isrctn.com/ISRCTN77132214) and results up to 18 months have been reported.[9] The study protocol was approved by the London South-East Research Ethics Committee for data collection to 18 months (09/H1102/55) and extended data collection to 60 months (09/H1102/55). All participants met at least one of the following indicators of antisocial behaviour (violent and aggressive behaviour; at least one conviction plus three additional warnings, reprimands, or convictions; diagnosis of treatment-resistant conduct disorder; school exclusion; risk of harm) across several settings and minimal exclusion criteria (appendix p.1).

Randomisation and masking

The investigators and objective assessors used a secure randomisation protocol and were strictly masked to treatment allocation. Masking was maintained throughout the follow-up period, with clinical and research staff located separately to avoid leakage of information. All coding, data entry and data cleaning were done by individuals masked to allocation. Data were warehoused separately from the research teams. A random sample (25%) of data was double-entered to check for entry errors. Allocation data were kept physically inaccessible to investigators and research assistants. The evaluation of treatment fidelity (see below) was undertaken by a geographically separate research group without access to the effectiveness data.

Procedures

MST[10] is an intensive family- and home-based intervention for young people with serious antisocial behaviour. The MST therapist worked with the young person's family to improve parenting, enhance relationships, garner support from social networks, address communication problems, encourage adaptive behaviour (eg, school attendance) and reduce maladaptive habits (eg, association with delinquent peers). Therapists saw only a few families at a time. Therapists saw each family three times a week for 3–5 months, and were available for crisis calls 24 h a day, 7 days a week. Programme fidelity was maintained by expert supervision and the use of a well-developed quality assurance system implemented at each clinical site (details can be found in the appendix to the first START paper[9]). The Therapist Adherence Measure-Revised (TAM-R), based on interviews conducted with the parents or carers as the treatment was in progress, indicated that therapists at all sites delivered MST that was adherent to the criteria specified by the treatment developers.[9] Following MST, families received MAU (described below).

MAU was based on the best available local services for each young person, in line with current community practice informed by treatment guidelines, and was offered on an as-needed basis (eg, support to re-engage with education, anger management, or victim awareness programmes) without formulation, an overarching plan, standardisation or supervision. The MAU interventions were supported by Youth Offending Teams (YOTs), Child and Adolescent Mental Health Services (CAMHS), and social and education services (see appendix p.7 for details).

Outcomes

The primary outcome was the proportion of young people who are known to have committed one or more criminal offence that led to a conviction in each arm of the trial at 60 months

post-randomisation. To ensure comparability with other trials of MST, the total number of recorded offences was taken as the outcome variable, based on official records from the Police National Computer and Young Offender Information System (a centralised database that records all offending within the UK, including date and type of offence, whether the offence resulted in a conviction (as not all offences do), and whether the conviction resulted in a custodial sentence, community sentence, or a caution). Offending data were collected over the course of the trial from 8th August 2011 to 4th September 2017**.** For the purposes of the analyses, offences were categorised as violent, non-violent, or breach of probation. Time to first offence was based on the date of the first recorded offence.

We also report a wide range of secondary outcomes prespecified in our research protocol consistent with the ecological focus of MST in addressing youth antisocial behaviour with its impact on criminality, education, and mental health. All outcomes were initially assessed at baseline and at 6, 12, and 18 months after randomisation. For the current follow-up period, secondary measures were collected at 24, 36, and 48 months post-baseline, and primary measures were collected up to 60 months. Families were sent an opt-out letter inviting them to decide whether they wanted to continue taking part in the trial. Families that did not opt out were contacted to confirm whether they wished to participate, to explain to them what participation would entail, and to ask them to complete a new consent form. Data collection was overseen by an independently chaired Trial Steering Committee and a Data Monitoring and Ethics Committee.

Secondary outcomes were self-report measures, completed by the young person and their parent or carer with the assistance of a research assistant, typically within the family home. The questionnaire pack took approximately 2 hours to complete, and families received £25 in

remuneration. Questionnaires were collected between 8th August 2011 and 4th September 2016**.** (For brief descriptions and associated hypotheses see the appendix pp.8-12).

Self-reported outcomes collected from baseline included measures of antisocial behaviour and attitudes that were targets of MST, completed by either the young person (Y) or the parent (A) or both (YA) (Strengths and Difficulties Questionnaire [SDQ-YA],[11] Inventory of Callous-Unemotional Traits [ICU-YA],[12] Self-Report Delinquency Measure [SRD-Y],[13] Antisocial Beliefs and Attitudes Scale [ABAS-Y],[14] Youth/Adult Materialism Scale [YMS-Y],[15] and the ADHD and Learning & Language subscales of the Conners Comprehensive Behaviour Rating Scales [CBRS-Y][16]). Some adult-specific measures were used from 24-month follow-up for participants who turned 18 during this period. These included the Adult Behaviour Checklist (ABC-Y),[17] the SDQ for Siblings, and the Adult Self Report (ASR-Y).[17]

Intermediate outcome measures of parenting skills and family functioning, which are assumed by the developers to account for treatment effects, included the Alabama Parenting Questionnaire (APQ-YA),[18] the Loeber Caregiver Questionnaire (LCQ-A),[19] the Family Adaptability and Cohesion Evaluation Scale (FACES-IV-A),[20] the Level of Expressed Emotion Questionnaire (LEE-Y),[21] and the Conflict Tactics Scale (CTS2-A).[22] As antisocial behaviour impacts on parental wellbeing, questionnaire measures concerning parental and young people's wellbeing and adjustment included the Short Mood and Feelings Questionnaire (SMFQ-Y),[23] the SDQ, and the General Health Questionnaire (GHQ-A).[24] Some measures were administered only from 24-month follow-up. These included the SF-36[25] (a quality of life measure), the Coddington Life Events Scale (CLES-A),[26] and the Adolescent Resilience Questionnaire (ARQ-Y).[27]

Economic outcomes included the use of health, social care, education, and criminal justice sector services. The Child and Adolescent-Service Use Schedule, designed specifically for

the trial, enabled us to monitor service use. Young people's quality of life was assessed using the EQ-5D-3L.[28]

The outcome measures administered and their schedule are provided in the appendix.

Changes to trial outcomes include the lack of data from the National Pupil Database, as it was not possible to reliably link the data to trial follow-up points. A characterisation of MST services was also planned during this follow-up period, including characteristics of the service, team operations, and the range of interventions available. Unfortunately, it was not possible to do this due to services being discontinued.

Statistical analysis

All analyses, except where noted, were prespecified in a statistical analysis plan agreed with the data monitoring committee.

The primary outcome was analysed using a mixed-effects logistic regression model with a fixed effect for treatment arm allocation. The model was adjusted for number of criminal offences which led to a conviction before randomisation, sex, age at onset of criminal behaviour (early or late) as fixed effects, and site as a random effect. The model was fitted in R using the library lme4.

The primary analysis estimated the odds ratio between MAU and MST, with confidence interval and p-value.

The time-to-event outcome (time to first offence) was analysed using a Cox proportional hazards model adjusted for the same fixed effects as the primary outcome. Count data outcomes (number and types of delinquent acts) were analysed using Poisson mixed models. The continuous outcomes (all questionnaire outcomes) were analysed with linear mixed

effects. All models included covariates from the primary outcome model as well as the respective baseline measure.

All analyses were performed with statistical methods that handle outcome data that are missing at random. As an additional analysis, as recommended by the Data Monitoring Committee, we used multiple imputation on a dataset with all baseline and outcome variables.

The primary outcome (number of recorded offences with convictions) and time to first offence outcome were tested for subgroup effects, with the following moderators considered: age, sex, baseline diagnoses of conduct disorder, depression, anxiety or ADHD, early-onset conduct disorder, offences before the trial, referral path as categorical variables, and baseline callous-unemotional traits, antisocial beliefs and attitudes, and peer delinquency scores as continuous variables.

A health economic evaluation was also carried out as part of the study. Unfortunately, due to an administrative error there was a large amount of missing data for the EQ-5D-5L measure. Consequently, we report the health economic methods and findings in the appendix (p. 37)

Role of the funding source

Representatives of the funders and MST-UK were invited to and were present at all Trial Steering Committee meetings but had no input into the design, data collection, analysis, or interpretation of the study findings. The corresponding author had full access to all the study data and had final responsibility for the decision to submit for publication.

Results

Patient flow is illustrated in Figure 1. In total, 684 young people were randomised, and 609 (89%) were available for collecting the primary outcome measure at 60-month follow-up.

Service use data were available for 607 (88·7%) of the participants, excluding only those who withdrew from the study. Attrition in relation to interview and self-report measures was greater, with 491 (77·9%), 478 (69·9%), 433 (63·3%), and 349 (51·0%) of participants contributing at 18, 24, 36, and 48 months, respectively. Excluding the 15% loss of participants for the 6-month visit, (19·4% attrition in MAU arm), the average loss of data was approximately 7% between each 6-monthly visit to the 48-month point (1% per month). Loss to follow-up could not be predicted from baseline characteristics (appendix pp.4-6).

For the primary outcome, there were no significant differences between interventions in the proportion of young people with a criminal offence with a conviction at 60-month follow-up: 188 (55%) of 342 young people in the MST group had at least one criminal offence compared with 180 (53%) of 341 in the MAU group (odds ratio 1·13, 95% CI: 0·82, 1·56; p=0·44).

There were no significant differences in median time to first offence, which was 36 months post-baseline for the MST group and 48 months for the MAU group (hazard ratio 1·03, 95% CI: 0·84, 1·26; p=0·78). Analyses of the number of recorded offences for both groups are shown in Table 2, and breakdown by violent and non-violent offences is provided in the appendix (p.18). Young people in the MST group had significantly (after adjustment) more recorded offences than those in the MAU group at 24 months (0·75 and 0·41, respectively; adjusted mean difference 0·35, 95% CI: 0·03, 0·67, p=0·031) and at 48 months (0·39 and 0·39; adjusted mean difference 0·35, 95% CI: 0·00, 0·69; p=0·049). When violent and non-violent offences were analysed separately, no significant difference between the groups was found at any timepoint.

We report on secondary outcomes between 18 months and 48 months aggregated, and at 24, 36, and 48 months. After correcting for multiple testing with the Benjamini–Hochberg method, no group differences were found on measures of antisocial behavioural problems

(SDQ), callous-unemotional traits (ICU), conduct problems (SRD), antisocial beliefs and attitudes (ABAS), attention, hyperactivity, learning, or language problems (CBRS), behavioural and emotional problems (ABC and ASR), or the Materialism Scale.

Parents in the MAU group reported higher levels of inconsistent discipline on the APQ compared with those in the MST group at 24-month follow-up (difference=–0·64, (95% CI –1·05 to –0·24; p=0·0023). Other measures of parenting skills, including family functioning (Loeber Caregiver Questionnaire), family adaptability (FACES-IV), expressed emotions (LEE), and family conflict (CTS2), did not identify any significant between-group differences.

No between-group differences were observed on measures of wellbeing and adjustment (SMFQ), young people's and parental wellbeing (SDQ), general psychiatric wellbeing (GHQ), quality of life (SF-36), or significant life events (CLES-A).

No significant differences were found in the percentage of young people in employment or education by 48-month follow-up; 70% of participants in the MST group were in education or employment, compared with 82% of those in the MAU group (odds ratio 0·53; 95% CI: 0·27, 1·03; p=0·062). There were no significant differences in the likelihood of participants in either group experiencing or causing pregnancy; the reported rates by 48-month follow-up were 17% in the MST group and 22% in the MAU group (odds ratio 0·72; 95% CI: 0·46, 1·13; p=0·16).

The results of prespecified interaction tests performed are given in the appendix (p.19). Only peer delinquency score had evidence of significantly moderating treatment outcome (Interaction OR per one-unit increase 0.91, CI 0.85 to 0.98, p=0.012): MST had significantly higher benefit as peer delinquency increased. There was a marginal moderation from sex, with female adolescents being slightly more likely to do better in the MST group compared

with those in the MAU group. Contrary to expectation, none of the indicators of severity we measured (early onset of conduct disorder, high callous-unemotional behaviour, previous offences) provided a selective indication for MST.

A similar pattern of results was observed for the analysis of data on time to first offence. Participants with higher baseline peer delinquency scores on the SRD were more likely to have a delay in offending in the MST group compared with the MAU group, but this effect was reversed for those with lower baseline peer delinquency scores (p=0·015). Figure 2 illustrates the proportion of non-offenders over the follow-up period who had high or low peer delinquency scores at baseline. Other prespecified moderators had no effect on the capacity of MST to delay time to first offence (appendix, p.19).

Discussion

In our previous study following participants up to 18 months post-randomisation, the START trial found no evidence of superiority of MST in reducing out-of-home placements or criminal offending, but in the same period there was parental and self-reported evidence of greater, and particularly more rapid, changes in antisocial behaviours and some superiority on measures of mood and wellbeing.[9] The 5-year follow-up reported here found no significant difference in overall recorded offending rates with convictions in young people in the MST group compared with those in MAU. The analyses of mean offending at different timepoints, however, suggested that outcomes were somewhat better among those receiving standard intervention from regular services (ie, MAU), as at two timepoints the MAU group showed significantly lower rates of offending. It should be noted that the number of recorded offences was not large and was always less than one offence per participant. However, consistent with the higher number of recorded offences in the MST group, there was a tendency for more young people from the MAU group to be in education or employment.

Analyses of key secondary outcomes did not indicate that MST was superior to MAU. The more rapid improvement in parent-rated antisocial behaviour observed within 1 year of the intervention[9] was not generally sustained over 48-month follow-up. The data showed that MST was superior in continuing to reduce inconsistent discipline at 24 months, but there were no other indications of lasting benefit from MST on the measures associated with antisocial behaviour and attitudes, parenting skills and family functioning, and young people's and parental wellbeing and adjustment.

These findings are not consistent with results from MST studies in the USA,[6] which found that MST was superior to MAU on measures of criminal offending. There are key differences in the social context and study designs that might account for disparities between the long-term effectiveness data from the USA and the UK. First, the criminal justice system in the USA has a stronger emphasis on punishment than that in the UK, where the emphasis is on rehabilitation.[30] Participants in studies in the USA may have been more motivated to engage in MST if the other options open to them were less desirable.

Because referrals were made by a variety of services rather than just coming from the criminal justice system, the START trial probably captured a group of young people with more diverse needs compared with previous studies of MST and may have comprised less severe cases than the bulk of trials in the USA. Importantly, the principles underpinning MAU services in England overlap with the evidence-based, social learning theory-inspired principles for parenting that underlie MST, making MAU in the UK more similar to the MST approach than MAU services in the USA might be. Further, effective systemic approaches for antisocial behaviour and criminality have been shown to have numerous common working elements.[31]

Despite the contextual differences in relation to international comparisons, the findings of the trial may be generalisable to the UK population. The trial is the largest evaluation of the long-term effects of MST to date. Families were recruited into the trial using carefully selected criteria to cater for the multi-agency approach the UK takes to young people with antisocial behaviour problems, and 87% of the eligible families agreed to take part. We were able to obtain offending and service use data on almost all participants who continued to consent to data collection, and the proportion who actively wished to discontinue participation was relatively small (14%).

To attempt to gain information on problems for which MST may be particularly helpful, we pre-planned a significant number of moderator analyses based on a range of indicators with the potential to identify distinct subgroups of young people with antisocial problems. The results of these analyses were disappointing. However, the results suggest that MST may be particularly helpful for young people with more delinquent peer social connections, who thus have more opportunity to engage in crime; higher numbers of deviant peers reported by participants were associated with elevation in virtually all symptoms of conduct disorder, including truancy, substance use, and perpetration of aggression. Our findings suggest that this group may specifically benefit from MST while those with fewer delinquent contacts are less well served by this approach.

The study has several limitations. It was not possible to deliver on some aspects of the design, including incorporating data on school absence from the National Pupil Database (which we were unable to link to reliably), and there was substantial attrition by 48 months on secondary outcomes. The description of the characteristics of the services delivered to both arms of the trial across the 5-year follow-up period is partial, as most of the MST sites that took part in the trial had shut down by the 36-month follow-up point (when the evaluation was planned to

be carried out) and the researchers could not contact the clinicians who had delivered the intervention. Furthermore, an administrative error resulted in gaps in the baseline EQ-5D-3L data. Efforts were made to impute the information, but it is possible that the measure was not sensitive enough to detect broader changes in these young people's quality of life (see appendix p.46)**.** It is possible that because services made proactive bids to participate in the trial, MAU interventions may have been of a higher standard than MAU services across the country on average, although we have no evidence to support such a claim. If this was indeed the case then we can conclude only that MST shows no superiority over MAU delivered to a high standard.

Although these limitations should be taken into account, the findings evidently do not suggest that MST is more beneficial than MAU in the long term for young people with antisocial behaviour problems. There is also no evidence to suggest that MST saves resources or that investment in MST is cost-effective. Treatment effects associated with MST and MAU appear stable and relatively invariant in their trajectory after 18 months. As MAU was delivered in the context of an RCT, we can assume that a relatively rigorous approach was taken in both arms of the intervention in this trial and to case management. While this may have contributed to our finding no significant benefit associated with MST, the findings suggest that unexpectedly good outcomes can be achieved in this clinical population which is generally considered to respond poorly to interventions when sufficient focused clinical effort is made.

Contributors

PFo, SBu, DC, JS, SS, SP, IE, PFu, SBy, and IMG were responsible for the original proposal, for securing funding for the trial, and for drafting the original protocol with assistance from EA. PFo, as chief investigator, had overall responsibility for the management of the study, with support from SBu as the clinical research lead. IMG had responsibility for the East Anglia site; DC and AK for the northern site; and IE, SS, and PFu for the southeast sites. SP, PFu, and SBu were responsible for development of the measure of management-as-usual interventions, and JS for qualitative measures. RE and ES were project managers throughout different phases of the trial and coordinated the randomisation and minimisation protocol. AA was the project manager who prepared data for analysis and coordinated trial closure and manuscript preparation. RE, with the support of PFo and SBu, set up and coordinated the database, with all data held in a single repository managed by the Mental Health Research Network at the East Anglia site. JW and SBy wrote the statistical analysis plan. JW, PG, SBy, and PG did the statistical analyses. PFo wrote the initial draft of the manuscript with support from EA. All authors contributed to and approved the final manuscript.

Declaration of interests

PF, SB, DC, SS, SP, IE, PFu, AK, SBy, JW, JS, AA, RE, ES, PG, EA and IMG report grants from the Department for Children, Schools and Families, the Department of Health and Social Care, and the National Institute for Health Research Clinical Research Network, and non-financial support from the Youth Justice Board and the National Institute for Health Research. DC reports his role as the co-chair of the National Institute for Health Research Advanced Fellowship Panel.

Acknowledgements

This study was funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research programme. The first phase of the study (to 18 months) was funded by the Department for Children, Schools and Families in conjunction with the Department of Health. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. We thank the research assistants involved in acquiring the data across the nine sites; Nicole Hickey (Imperial College London) who prepared, coded, and analysed the youth offending data from the PNC records; Eric Taylor and the members of the Trial Steering Committee he chaired, and Philip Graham and the members of the Data Management and Ethics Committee he chaired, for their counsel; Kathryn Harney (Associate Director of Research, Greater Manchester Mental Health NHS Foundation Trust) for her support, co-operation and guidance.

Proposals may be submitted up to 36 months following article publication. After 36 months the data will be available in our university's data warehouse but without investigator support other than deposited metadata. Information regarding submitting proposals and accessing data may be found at [LINK]

Trial registration

ISRCTN 77132214

References

1 Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. BMJ 2001; **323:** 191.
2 Blair RJR, Leibenluft E, Pine DS. Conduct disorder and callous–unemotional traits in youth. N Engl J Med 2014; **371:** 2207–16.

3 NICE. Antisocial behaviour and conduct disorders in children and young people: recognition and management (CG158). 2017. https://www.nice.org.uk/guidance/cg158/chapter/introduction (accessed 12 December 2012).
4 Parsonage M, Khan L, Saunders A. Building a better future: the lifetime costs of childhood behavioural problems and the benefits of early intervention. London, UK: Centre for Mental Health, 2014.
5 Henggeler SW. Multisystemic therapy. Resid Treat Child Youth 2001; **18:** 75–85.
6 Markham A. A review following systematic principles of multisystemic therapy for antisocial behavior in adolescents aged 10–17 years. Adolesc Res Rev 2018; **3:** 67–93.
7 Sawyer AM, Borduin CM. Effects of multisystemic therapy through midlife: a 21.9-year follow-up to a randomized clinical trial with serious and violent juvenile offenders. J Consult Clin Psychol 2011; **79:** 643–52.
8 Schaeffer CM, Borduin CM. Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. J Consult Clin Psychol 2005; **73:** 445–53.
9 Fonagy P, Butler S, Cottrell D, et al. Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. Lancet Psychiatry 2018; **5:** 119–33.
10 Henggeler SW, Borduin CM. Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents. Pacific Grove, CA: Brooks/Cole; 1990.
11 Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. J Am Acad Child Adolesc Psychiatry 2001; **40:** 1337–45.
12 Frick PJ, Cornell AH, Bodin SD, Dane HE, Barry CT, Loney BR. Callous-unemotional traits and developmental pathways to severe conduct problems. Dev Psychol 2003; **39**: 246-60.
13 Elliott DS, Canter RJ, Knowles BA. The epidemiology of delinquent behavior and drug use among American adolescents, 1976-1978. A report of the National Youth Survey. Boulder, CO: Behavioral Research Institute; 1981.
14 Butler SM, Leschied AW, Fearon P. Antisocial beliefs and attitudes in pre-adolescent and adolescent youth: the development of the Antisocial beliefs and Attitudes Scales (ABAS). J Youth Adolesc 2007; **36:** 1058–71.
15 Goldberg ME, Gorn GJ, Peracchio LA, Bamossy G. Understanding materialism among youth. J Consum Psychol 2003; **13:** 278–88.
16 Conners CK, Sitarenios G, Parker JD, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. J Abnorm Child Psychol 1998; **26:** 257–68.
17 Achenbach TM, Rescorla LA. The Achenbach system of empirically based assessment (ASEBA) for ages 1.5 to 18 years. In: Maruish ME, ed. The use of psychological testing for treatment planning and outcomes assessment: instruments for children and adolescents. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2004: 179–213.
18 Essau CA, Sasagawa S, Frick PJ. Psychometric properties of the Alabama Parenting Questionnaire. J Child Fam Stud 2006; **15:** 595–614.
19 Loeber R, Farrington DP. Serious & violent juvenile offenders: risk factors and successful interventions. Thousand Oaks, CA: Sage; 1999.
20 Olson DH, Bell R, Portner J. Family Adaptability and Cohesion Evaluation Scales. FACES-II. Minneapolis, MN: Family Inventories Manual Life Innovations; 1992.

21 Gerlsma C, Hale WW, 3rd. Predictive power and construct validity of the Level of Expressed Emotion (LEE) scale. Depressed out-patients and couples from the general community. Br J Psychiatry 1997; **170:** 520–25.
22 Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The revised Conflict Tactics Scales (CTS2): development and preliminary psychometric data. J Fam Issues 1996; **17:** 283–316.
23 Sharp C, Goodyer IM, Croudace TJ. The Short Mood and Feelings Questionnaire (SMFQ): a unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. J Abnorm Child Psychol 2006; **34:** 379–91.
24 Goldberg D. Manual of the General Health Questionnaire. Windsor, UK: NFER-Nelson; 1978.
25 Ware JE, Jr. SF-36 health survey update. Spine 2000; **25:** 3130–9.
26 Coddington RD. The signifance of life events as etiologic factors in the diseases of children. I. A survey of professional workers. J Psychosom Res 1972; **16:** 7–18.
27 Gartland D, Bond L, Olsson CA, Buzwell S, Sawyer SM. Development of a multi-dimensional measure of resilience in adolescents: the Adolescent Resilience Questionnaire. BMC Med Res Methodol 2011; **11:** 134.
28 Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011; **20:** 1727–36.
29 NHS England. Model specification for transitions from child and adolescent mental health services. 2015. https://www.england.nhs.uk/wp-content/uploads/2015/01/mod-transt-camhs-spec.pdf (accessed 4 April 2019).
30 Timmons-Mitchell J, Bender MB, Kishna MA, Mitchell CC. An independent effectiveness trial of multisystemic therapy with juvenile justice youth. J Clin Child Adolesc Psychol 2006; **35:** 227–36.
31 van der Pol TM, van Domburgh L, van Widenfelt BM, Hurlburt MS, Garland AF, Vermeiren RRJM. Common elements of evidence-based systemic treatments for adolescents with disruptive behaviour problems. Lancet Psychiatry 2019; **6:** 862–8.

RESEARCH IN CONTEXT

Evidence before this study

We did a systematic review to identify randomised controlled trials and systematic reviews of multisystemic therapy (MST) for adolescent antisocial behaviour for our previous report on the outcomes of this trial to 18 months. The review encompassed studies to Dec 31, 2016, and found 22 primary randomised trials of MST. Previous reviews (eg, for NICE) identified MST as a promising intervention for young people with conduct problems. However, the outcomes of the trials were mixed, with generally good outcomes in studies based in the USA, but some reports of non-USA studies suggesting that MST was no more effective at reducing antisocial behaviour than usual services.

We did an additional search of Embase, MEDLINE and PsycINFO for any relevant randomised trials of MST that had been published from Jan 1, 2017 to Dec 1, 2019, using the terms "multisystemic therapy" or "MST". We found one additional trial published in 2018. In total, 13 studies were carried out in the USA and 10 in Europe. The median follow-up period was 18 months, ranging from 6 months to 22 years. With the exception of two USA-based studies that followed up a 1995 trial of MST 13 and 22 years later, no trial evaluated MST outcomes for more than 48 months.

We have previously published the outcomes of the START trial to 18 months. At the time, it was the only independent large-sample trial evaluating the medium-term superiority and cost-effectiveness of MST in the UK. We found no additional benefit of MST compared with management as usual in terms of out-of-home placements, but MST did have a positive effect

for parent-reported offending behaviour at 18 months and was associated with a faster rate of positive behavioural changes.

Added value of this study

The present study follows on from the initial 18-month follow up, reporting criminal convictions to 60 months post-baseline and secondary (self-reported) outcomes, including an economic analysis, to 48 months. Most evaluations of the effectiveness of MST from the USA to date have indicated that it is more effective than other interventions. Evidence from European countries has been more equivocal. To our knowledge, this is the longest follow-up for a randomised controlled evaluation of MST in the UK to date. The outcomes do not support the long-term superiority of MST, despite the results of the initial 18-month follow-up where parents report suggested more rapid change associated with MST. These outcomes include the proportion of young people who offended at 60-month follow-up and secondary evaluations of young people's behaviour and adjustment, psychosocial functioning, family functioning, and quality of life, as well as the cost-effectiveness of MST compared with management as usual.

Implications of all the available evidence

Our results do not support the superiority of MST in the UK for young people with conduct problems, and we found no evidence of MST being more cost-effective in the long term. It is possible that MST is more beneficial in the context of the social services and criminal justice system in the USA, but that in the UK the needs of young people with conduct problems are met equally well by the usual services currently offered to them.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-5
	2b	Specific objectives or hypotheses	5, 7-9
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	4-6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	10
Participants	4a	Eligibility criteria for participants	6

	4b	Settings and locations where the data were collected	4-5 (and appendix)
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Supplementary material to 2018 paper
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	7-9/Appendix
	6b	Any changes to trial outcomes after the trial commenced, with reasons	9
Sample size	7a	How sample size was determined	Supplementary material to 2018 paper
	7b	When applicable, explanation of any interim analyses and stopping guidelines	N/A
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	6 & supplementary material to 2018 paper
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6 & supplementary material to 2018 paper

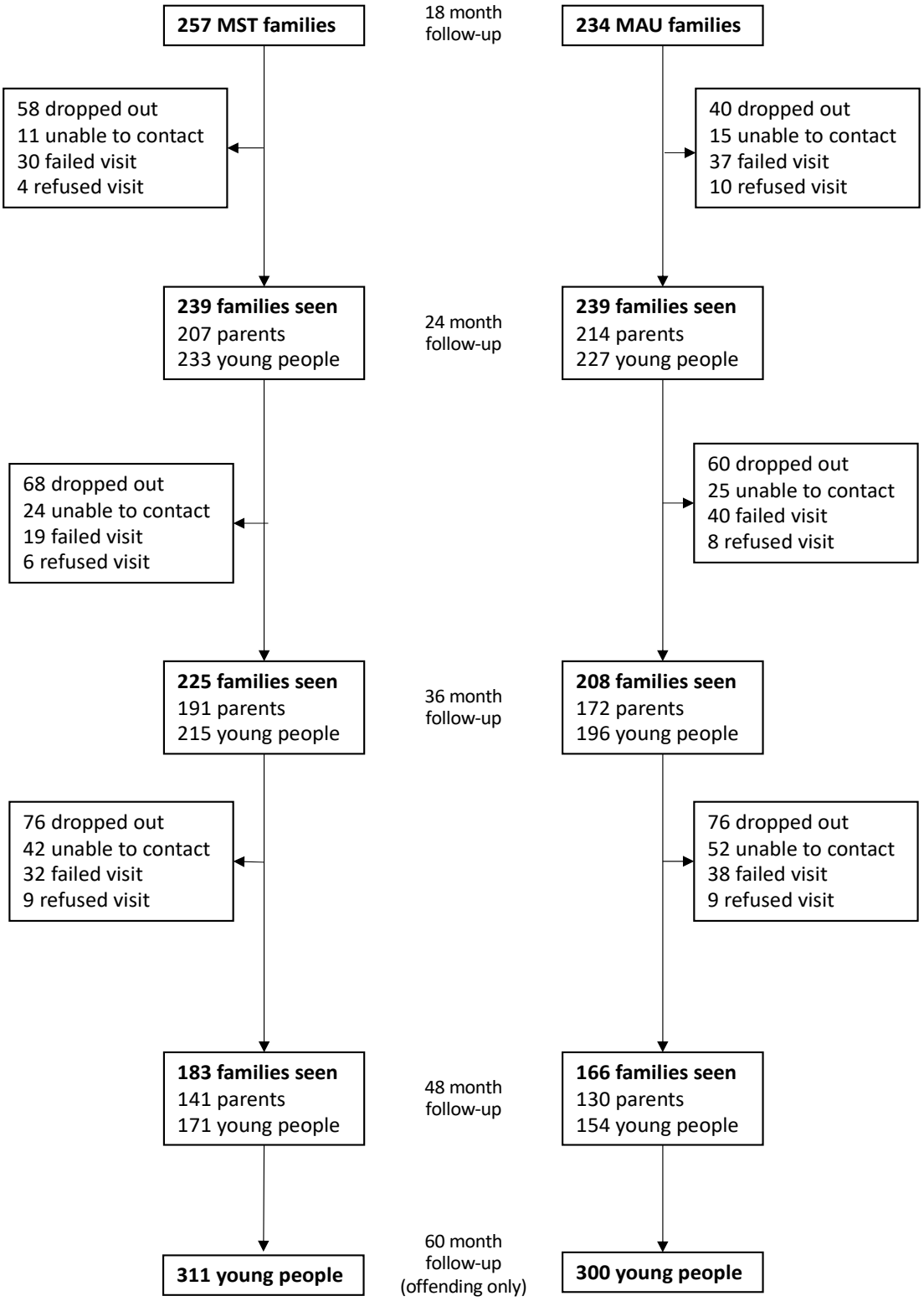
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6 & supplementary material to 2018 paper
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6 & supplementary material to 2018 paper
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	6 & supplementary material to 2018 paper
	11b	If relevant, description of the similarity of interventions	6 & supplementary material to 2018 paper
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	9-11
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	10
Results	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	5 & Figure 1(a&b)

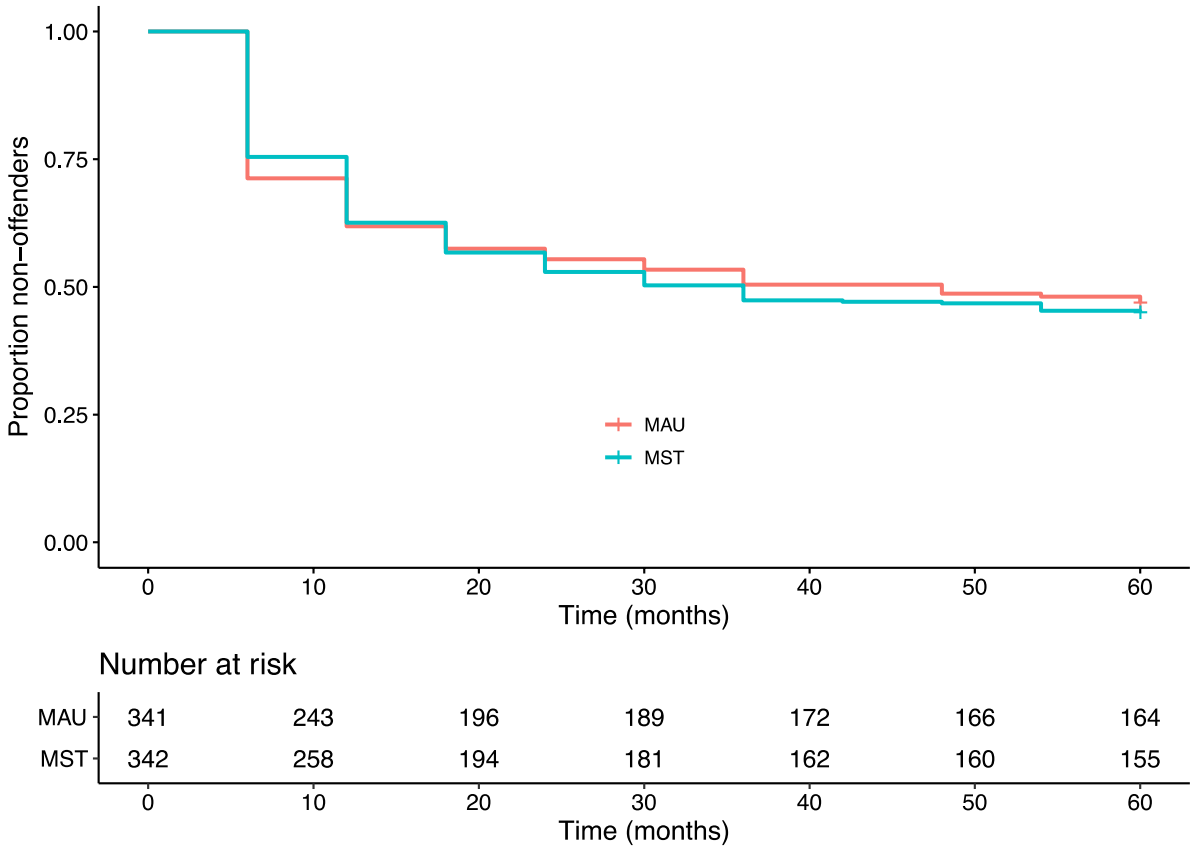
Participant flow (a diagram is strongly recommended)	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1(a&b)
Recruitment	14a	Dates defining the periods of recruitment and follow-up	7-8
	14b	Why the trial ended or was stopped	N/A
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Figure 1
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	12-15 & Appendix
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	12-15 & Appendix
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	14
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	N/A

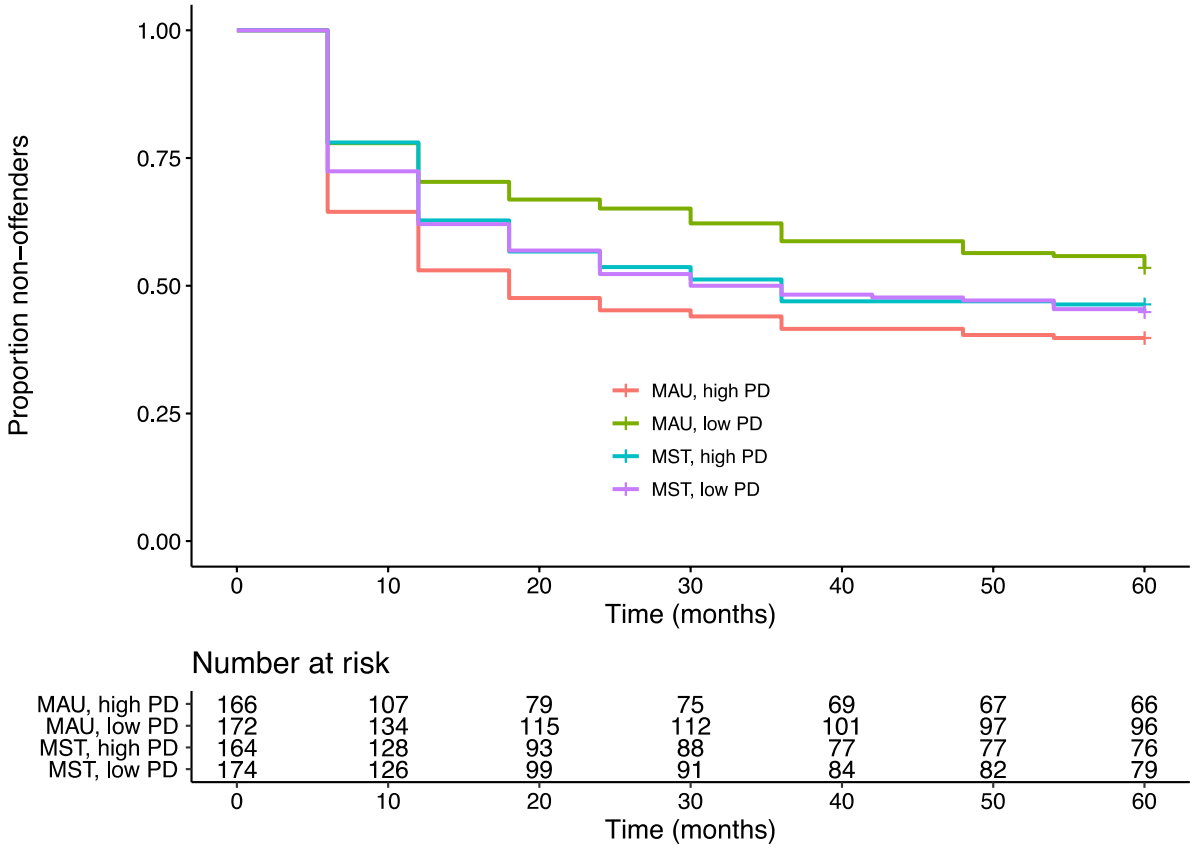
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	17-18
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	17
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	15-18
Other information			
Registration	23	Registration number and name of trial registry	20
Protocol	24	Where the full trial protocol can be accessed, if available	6
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	3, 19

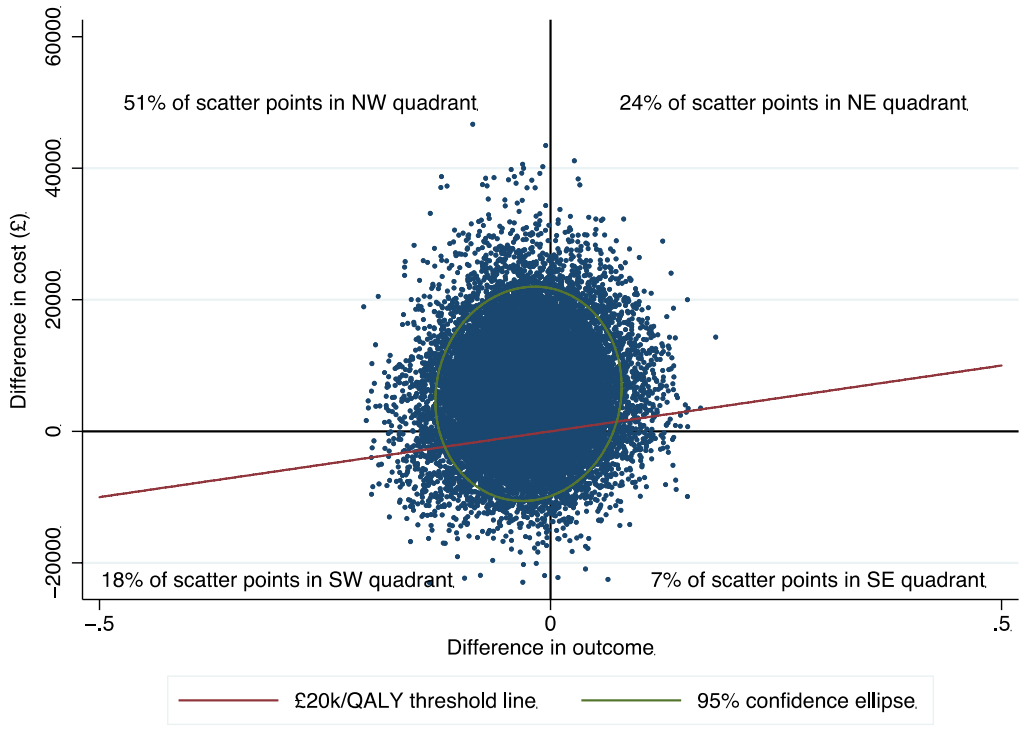
*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Figure 1b


257 MST families
18 month follow-up
234 MAU families
58 dropped out
11 unable to contact
30 failed visit
4 refused visit
40 dropped out
15 unable to contact
37 failed visit
10 refused visit
239 families seen
207 parents
233 young people
24 month follow-up
239 families seen
214 parents
227 young people
68 dropped out
24 unable to contact
19 failed visit
6 refused visit
60 dropped out
25 unable to contact
40 failed visit
8 refused visit
225 families seen
191 parents
215 young people
36 month follow-up
208 families seen
172 parents
196 young people
76 dropped out
42 unable to contact
32 failed visit
9 refused visit
76 dropped out
52 unable to contact
38 failed visit
9 refused visit
183 families seen
141 parents
171 young people
48 month follow-up
166 families seen
130 parents
154 young people
60 month follow-up (offending only)
311 young people
300 young people


1.00
0.75
0.50
0.25
0.00
Proportion non−offenders
MAU
MST
0
10
20
30
40
50
60
Time (months)
Number at risk
MAU 341 243 196 189 172 166 164
MST 342 258 194 181 162 160 155
0
10
20
30
40
50
60
Time (months)


1.00
0.75
0.50
0.25
0.00
Proportion non−offenders
MAU, high PD
MAU, low PD
MST, high PD
MST, low PD
0
10
20
30
40
50
60
Time (months)
Number at risk
MAU, high PD 166 107 79 75 69 67 66
MAU, low PD 172 134 115 112 101 97 96
MST, high PD 164 128 93 88 77 77 76
MST, low PD 174 126 99 91 84 82 79
0
10
20
30
40
50
60
Time (months)


51% of scatter points in NW quadrant
24% of scatter points in NE quadrant
18% of scatter points in SW quadrant
7% of scatter points in SE quadrant
Difference in cost (£)
60000
40000
20000
0
−20000
−.5
0
.5
Difference in outcome
£20k/QALY threshold line
95% confidence ellipse