Equity effects of parenting interventions for child conduct problems: a pan-European individual participant data meta-analysis

Frances Gardner, DPhil (University of Oxford, corresponding author)

Patty Leijten, PhD (University of Oxford, University of Amsterdam, Netherlands)

Victoria Harris, PhD (King's College London, UK)

Joanna Mann, PhD (University of Oxford, UK)

Judy Hutchings, DClinPsy (Bangor University, UK)

Jennifer Beecham, PhD & Eva-Maria Bonin, PhD (London School of Economics, UK)

Vashti Berry, PhD (University of Exeter, UK)

Sinead McGilloway, PhD (Maynooth University, Ireland)

Maria Gaspar, PhD, Maria João Seabra-Santos, PhD (University of Coimbra, Portugal)

Bram Orobio de Castro, PhD, Ankie Menting PhD (Utrecht University, Netherlands)

Margiad Williams, PhD (Bangor University, UK)

Ulf Axberg, PhD (University of Gothenburg, Sweden)

Willy-Tore Morch PhD (UIT Tromso, Norway)

Stephen Scott, FRCPsych (King's College London, UK)*

Sabine Landau, PhD (King's College London, UK)*

*Joint last authors

Abstract

Background: Childhood conduct problems are a costly public health problem, five times more common in socially disadvantaged groups. Untreated, they have a poor prognosis, with increasing gaps between socio-economic groups, and high rates of subsequent criminality. The Incredible Years (IY) is a high-quality parenting programme as recommended by NICE for reducing conduct problems, and is widely disseminated in Europe. Many trials show IY to be effective, but the potential effects on social inequality of parenting interventions are unknown. This matters since some behavioural interventions (e.g. smoking cessation programmes), while beneficial overall, can widen inequality gaps. Since single trials and aggregate-level meta-analysis are ill-equipped for examining differential intervention (moderator) effects, we pooled individual-level trial data.

Method: Individual participant data (IPD) meta-analysis of a near-complete set of randomised trials of European IY parenting programmes (N=1696; 15 trials eligible; 7% [1/15], data unavailable; 7% [1/15] lacked primary outcome). Children were aged 2-10 years (*M* 5.1; 30% [492/1651] ethnic minority; 58% [931/1614] low-income). Primary outcome was child conduct problems, using Eyberg Child Behavior Inventory (ECBI-I). Moderators were analysed using multilevel modelling with multiple imputation.

Findings: IY led to an overall reduction in child conduct problems (13.5 points on ECBI-I, 95% CI 10.9 to 16.1)**.** There was no evidence for differential effects by family disadvantage (poverty, lone/teen parenthood, joblessness; low education), or ethnic minority status.

Interpretation: This world-first IPD meta-analysis of parenting trials, the largest pooled data set to date, found no evidence for differential effects by social disadvantage, suggesting IY is

unlikely to widen socioeconomic inequalities in conduct problems. Furthermore, the programme may be an important tool for reducing social disparities and improving poor long-term outcomes in disadvantaged families, since follow-up studies indicate that benefits persist. Clinicians and commissioners can be reassured the programme is similarly effective for families from different backgrounds.

Funding: NIHR, #12-3070-04

Research in context

Evidence before this study:

Childhood conduct problems (disruptive behaviour problems) are a costly public health problem that is five times more common in socially disadvantaged groups. They are a major social issue as they precede criminal behaviour in the majority of youth. The Incredible Years (IY) is a high-quality parenting programme as recommended by NICE for reducing child conduct problems, and is widely disseminated across Europe. Many trials and systematic reviews show IY to be effective, but the potential effects on social inequality of parenting interventions are unknown. This matters, since some behavioural interventions, while beneficial on average, can often widen health gaps between rich and poor, e.g. smoking cessation programmes.

We systematically searched five databases (including PsychINFO, CINAHL, MEDLINE) for evaluations of parenting interventions published from inception to 2014, with no language restriction, using search terms “parenting”, “children” “conduct problems” and synonyms. We found two 2006 meta-analyses addressing our question about moderation of parenting intervention effects by social disadvantage, based on aggregate-level data from randomised

and non-randomised trials. Both concluded that socially disadvantaged families benefited less from parenting interventions, compared to more advantaged families. However, aggregate-level meta-analysis is a poor quality method for testing intervention moderation effects, due to problems of low power, greater risk of bias, and lack of information about variability in outcome between individuals. To overcome these limitations, Individual Participant Data (IPD) meta-analysis is recommended. Our searches found no IPD meta-analyses of parenting intervention trials in children. Hence, we conducted the first IPD meta-analysis of a parenting intervention, harmonising individual-level data from virtually all trials across Europe of IY.

Added value of this study:

Our study is the first to test whether parenting programmes widen social inequalities in children's conduct problems, in a large, well-powered study, by pooling data across trials using IPD meta-analysis. We obtained data from a near-complete set (93%, [14/15]) of eligible randomised trials of IY parenting programmes in Europe, and analysed data from 13/14 trials (1696 children, aged 2-10), as one toddler trial lacked data on conduct problems, due to the young age of the children. A wide range of social and ethnic backgrounds, and community-based service contexts, were represented in the pooled sample (30% [492/1651] minority; 58% [931/1614] low-income). We found that IY was effective for reducing child conduct problems, and that there was no evidence to suggest that its effectiveness varied with different levels of social disadvantage, including families on low income, lone or teen parents, jobless households, or those with low education, as well as for ethnic minority families. Updated searches in March 2019 revealed two further eligible trials; both found no variation in effectiveness by social disadvantage.

Implications of all the available evidence:

Our data suggest that this parenting programme, unlike many behavioural interventions for conditions that are considerably more common in disadvantaged groups, is unlikely to widen inequalities in conduct problems, based on socioeconomic status, or ethnicity. Our evidence updates and substantially alters the picture of inequality effects found in influential prior aggregate-level meta-analyses. Furthermore, if the effects are maintained over time (initial studies suggest they are), then the programme may reduce the increasing social gap in untreated conduct problems, with the social difficulties and criminality that otherwise ensue. If it does so, this could be an important tool for promoting equity. In the meantime, clinicians, social care providers, and commissioners can be reassured that the programme is similarly effective for families from different backgrounds.

Introduction

Conduct problems in childhood are common, costly and persistent over time, foreshadowing a range of adverse outcomes, including school failure, criminal offending, unemployment, and poor physical and mental health.[1,2] All carry substantial economic and personal burden[2], and are of major societal concern. Parenting interventions are a common and effective strategy[3,4] for preventing and treating conduct problems, which also reduce harsh and abusive parenting. Recent years have seen policy directives encouraging their implementation at scale in many European countries, and globally (e.g. NICE, WHO[5,6]).

However, in going to scale, it is important not only to establish that interventions are effective, but also that they are beneficial for the most socially disadvantaged, who typically have the highest levels of ill-health. Poor parenting and child conduct problems are highly patterned by income and social class, and possibly by ethnicity.[7,8] A recent systematic review[9] suggested some types of behavioural interventions, for example for smoking cessation or obesity, whilst effective overall, may actually increase health inequalities,

especially those that involve individual behaviour change, compared to those employing external incentives or regulatory strategies.[9] Therefore, it is vital to assess whether behaviour change interventions, including parenting programmes, have the unintended consequence of widening social inequalities in health, especially given their widespread use. This is a key question for commissioners, and practitioners who refer families to interventions, or deliver them to disadvantaged or minority client groups.

Equity effects by social disadvantage

Despite substantial evidence that parenting interventions are effective overall in preventing and treating child conduct problems,[3,4] there are many reasons why disadvantaged families might benefit less. For example, economic stressors present barriers to attending sessions and implementing new strategies; moreover, there may be a mismatch between parents' values and expectations, and those of the intervention. Furthermore, conflicting findings from individual trials and systematic reviews are equivocal about differential effects for socially disadvantaged or ethnic minority families[10-13]. For example, two systematic reviews addressing moderator effects conclude that disadvantaged children benefit less than those from more advantaged families.[12,13] If this were so, parenting interventions may fail to help the most vulnerable families, whilst also inadvertently widening social inequalities. Such unintended effects were seen in Sure Start community services in England, whereby the most socially disadvantaged families had worse outcomes compared to more advantaged families.[14]

Equity effects by ethnicity

Parenting styles, practices and norms vary across cultures, but most parenting interventions are developed in Western countries. Thus, they might be a poor fit, and produce diminished effects for parents from different cultural backgrounds. Identifying effective cross-cultural parenting interventions is of huge importance to governments seeking to

enhance health outcomes for minority and new immigrant families. Furthermore, worldwide policies promoting early intervention have led to considerable efforts to identify effective parenting interventions across cultures.[5,15] Thus, understanding the effects of parenting interventions for minority families in Western countries has important policy and practice implications, whilst also contributing useful knowledge on the generalisability of interventions developed (and often delivered) by professionals from one ethnic background.

Most work on ethnicity and parenting interventions has been conducted in the USA, and is equivocal about whether outcomes differ by ethnicity.[16,17] Despite this, there is a predominant view that parenting interventions should be adapted for different ethnic groups.[18] However, such approaches imply delivering parenting groups separately by ethnicity, which may be neither a practical or desirable service model for multi-ethnic European cities. Moreover, data are much needed to guide policy in Europe, where immigration has increased, but where individual trials are not powered to test intervention effects by ethnicity.

IPD meta-analysis for testing equity effects

Traditionally, differential intervention effects are tested using subgroup analyses within trials, or aggregate-level meta-analysis across trials. However, the validity of both traditional methods is limited; subgroup analyses in individual trials have well-documented problems of inadequate power, multiple testing, and selective reporting bias, linked to a failure to pre-specify secondary analyses.[19,20] Aggregate-level meta-analyses fail to address these problems because moderators are analysed only at trial level (e.g., mean poverty level of sample), leading to substantial loss of power and meaningful information, and creating high risk of confounding moderators (e.g., where poverty is correlated with problem severity).[19,20]

Individual Participant Data (IPD) meta-analysis overcomes these problems. First, it fully exploits within-trial variation in sociodemographic characteristics, rather than lacking this information as in traditional meta-regression.[20] Second, pooling across multiple trials allows for inclusion of data on all outcomes, and consistent analysis across trials. This maximises transparency and minimises bias, particularly selective outcome reporting and publication bias. However, bias-reducing benefits of IPD meta-analysis only apply if a large proportion of triallists share data and measure similar outcomes. Thus, IPD offers substantial advantages for addressing questions about equity effects, provided sufficient data can be included.[20] Equity questions are particularly pressing for problems that are highly patterned by social disadvantage, as with conduct problems, whose sequelae also confer further disadvantage.[2]

The present study

The aim of this study was to assess, using IPD meta-analysis, the equity effects — in a European context — of the Incredible Years (IY) programme (Web-extra, W1), a well-established group-based, evidence-based parenting intervention for reducing child conduct problems.[21] Specifically, we examined whether social disadvantage and ethnicity moderated the effects of IY intervention on conduct problems, capitalizing on individual-level data (IPD) across 13 European randomised trials.

We focused on the IY programme because:(1) it is a manualised intervention with a substantial evidence base[21] as recommended by NICE and other policymakers; (2) it has been widely disseminated in many European countries; and (3) there are active European research networks for IY, raising the probability of obtaining data from a near-complete set of trials for IPD meta-analysis. We focused on Europe for the following additional reasons: (1) European countries where IY is implemented tend to have relatively similar health and social care systems (unlike, for example, the USA); (2) most European trials have been conducted

independently of the programme developer – (developer involvement is associated with stronger intervention effects, and may represent a source of bias); (3) little is known about ethnicity effects in Europe, so pooling IPD uniquely, allowed for the inclusion of substantial numbers of families from a range of minority backgrounds.

Methods

Reporting, registration and ethics

PRISMA IPD reporting guidelines were followed (Web-extra, W2). Protocol and analysis plan are at www.spi.ox.ac.uk/parentingIPD. Ethical approval was granted by Oxford University, Social Policy and Intervention DREC.

Eligibility criteria

We sought to include all data from all completed randomised trials of the IY parenting intervention in Europe, for children aged 1–12 years, including unpublished trials, without restriction on publication year, or outcome measures. We included both prevention (selective or universal) and treatment/ indicated prevention trials (for children diagnosed or above clinical cut-off on conduct problems). We excluded trials, or conditions within trials that: (1) were not randomised; (2) included additional non-parenting material (e.g., child-focused interventions); and (3) were abbreviated, non-standard versions of the usual weekly 12–14 session IY intervention.

Identifying and selecting trials (Figure 1; Web-extra, W3)

Trials were identified through: (1) systematic searches in five databases (CINAHL, Embase, Global Health, MEDLINE, PsycINFO), in January 2015; (2) searching IY website library; and (3) consultation with experts including European IY mentors' network. Search terms via OVID were: 1. incredible years.mp; 2. webster-stratton.mp; 3. 1 or 2. Search strings were adapted for other databases. Eligibility was assessed by the first author and double-

checked by four additional authors, with no disagreements. Searches were updated in March 2019 in order to assess any relevant new trials.

Data collection and data integrity

All available fully anonymised data were requested for the 15 randomised trials of IY parenting intervention (Table 1) identified in the 2015 searches. Trial PIs signed data-sharing agreements specifying ethical and ownership issues. Raw, individual item-level data were supplied and checked for missing items, scale validity and scores, internal consistency, baseline imbalance, and consistency with trial protocols and reports. Copies of original questionnaires were supplied to check for consistent use across trials. Queries were resolved in collaboration with trial investigators. No substantial integrity issues arose. Trial quality was assessed using Cochrane risk-of-bias tool.

Measures and harmonisation (Web-extra, W4 harmonisation details)

Child conduct problems. The pre-specified primary outcome was the Eyberg Child Behavior Inventory, Intensity Scale (ECBI-I)[22] as it is the most frequently used across trials (n=11). This 36-item psychometrically robust scale assesses parent-reported frequency of child conduct problems on a 7-point Likert scale[22] Two trials[23,24] used a different measure of conduct problems (Parental Account of Children's Symptoms, PACS)[25] and, in both cases, data were converted to scores on the ECBI-I, using norm deviation scores. This is similar to using z-scores, but utilises means and standard deviations on each measure, from published norms[26,27]. PACS and ECBI-I scores were strongly positively correlated (*r*=.71), based on IPD from four trials[28-31] that included both measures. Internal consistency at baseline was high (ECBI-I α= .94; PACS, α= .82). Data from primary caregivers (98% mothers) were used because few trials included data from both parents. Only limited data (k=3) were available from other informants (e.g. teachers) and were excluded.

Social disadvantage moderator variables (five binary indicators): Given that social disadvantage is multifaceted, we included a range of indicators. Due to variation in assessment of social disadvantage, it was necessary to harmonise indicators based on similarities in operationalisation, as follows:

Low income: Indicators were defined as receiving income-related financial benefits (10 trials), scoring below Hollingshead Index's low-SES threshold (1 trial),[32] or living in social housing (2 trials).[33,34]

Educational level: Highest educational level of parent was dichotomised using UNESCO ISCED-11 categories,[35] where 'low'=primary / lower secondary, and 'high'=upper secondary/ degree-level education.

Lone parenthood: Primary parent lives without partner/ spouse.

Teen parenthood: Parent <20 years at target child's birth.

Unemployment: No parent in household employed.

Ethnic minority: Primary parent's ethnic background, any category other than 'white' (ONS classification).

Statistical methods

Power calculations for an anticipated sample size of N=1400 gave 96% and 80% power to detect a small interaction effect between two binary variables (Cohen's d=0.20, 0.15 respectively) using ANOVA F-test at 5% significance level.

The purpose of the analyses was to assess whether any of the six binary participant-level measures of social disadvantage moderated the effect of IY on ECBI-I post-intervention. Three statistical issues needed addressing: (i) the pooled data had a hierarchical structure where families (Level 1) in the intervention arm were nested within parent-groups (Level 2), and parent-groups were nested within trials (Level 3); (ii) there was some variation

in design of the trials that needed accounting for (e.g. stratified randomisation; changes in allocation ratios during the trial); and (iii) missing data biases needed minimising. We addressed these using a one-stage model, which tackles relevant moderation questions, all in one-step.[20]

Putative moderators were assessed one at a time. The resulting multilevel/mixed effects modelling used post-intervention ECBI-I as the dependent variable and contained fixed effects for trial arm, trial-level moderator summaries (between-trial variables, e.g. percent ethnic minority) and participant-level deviations from trial summaries (within-trial variables) and respective interaction terms. Tests of the effects of interaction terms then provided an assessment of the trial-level and participant-level moderating effects. Importantly, this allowed us to assess empirically whether these two moderating effects differed. If such a difference was significant at a liberal 10% test level, then two separate moderating effects were allowed; if not, a more powerful model with a single interaction term was fitted. The sizes of any moderation effects were described by an effect moderation index that expressed the difference in IY effect (active-control arm) between "presence" and "absence" levels of the binary moderator on the ECBI scale.

The hierarchical structure of the data was modelled by random intercepts that varied with trial (Level 3) and a further random intercept that varied with parent-training group within the active trial arm (Level 2). Trial design features were accommodated by including relevant fixed effects (e.g. for randomisation stratifiers) or random intercepts that varied with cluster in a cluster-randomised trial. Known predictors of post ECBI-I (baseline ECBI-I, child gender, age) were also included as fixed effects, as was the possible confounder - prevention vs. treatment trial - and its interaction with trial arm, in order to adjust moderation effects. Finally, to allow for further treatment effect heterogeneity (e.g. due to service contexts or

composition of the trial population) a trial-varying random coefficient of trial arm was included in the model.

IPD had missing values in moderator and outcome variables. We used multiple imputation by chained equations (MICE)[36] to produce valid estimates of moderation effects under missing-at-random (MAR) assumptions.[36] Analyses were conducted in Stata 14; significance level is 5% unless specified.

Role of funder: The funder played no role in data analysis or interpretation.

Results

Study characteristics

Fifteen trials met inclusion criteria (Table 1), conducted in England (*k*=7),[23,28-31,37,38] Wales (*k*=2),[39,40] Netherlands (*k*=2),[33,34] with one each in Ireland,[41] Norway,[42] Portugal[24] and Sweden.[32] Thirteen trials (N=1696) were included in the analyses, with two UK trials (13%, [2/15]) excluded, one where data were no longer available,[38] and one where IPD was supplied but which lacked data on our primary outcome,[40] because of the young age of the children. Due to uneven (2:1) randomisation ratios in some trials, there were 1046 families in the intervention, and 650 in the control arms. For all trials, we included data for baseline and first post-intervention assessment, which was normally 4-6 months later; in most studies this was the primary endpoint. All trials were conducted independently of the US-based developer (Webster-Stratton). Risk of bias within studies was assessed as low on most items (Web-extra, W5).

Of the 13 trials included in IPD analyses, ten were treatment (referred for clinical-level conduct problems, n=5) or indicated prevention (screened for high levels of conduct problems; n=5). Three were selective prevention trials (targeting high-risk families such as disadvantaged families or mothers released from prison).[30,33,34] Some targeted low-income

areas (Wales Sure Start)[39] or schools in low-income wards[28] prior to screening for conduct problems. Overall, most trials (10/13) included families who were predominantly socially disadvantaged due to low income or lone parenthood.[23,28-31,33,34,37,39,41] Six trials in England and Netherlands accounted for over 90% of the families from ethnic minorities (range 19–78% per trial).[28,30,31,33,34,37] In nine trials, the control condition was a waiting-list; in four trials, there was a minimal or no intervention. Most sites (9/13) delivered IY in community settings (e.g. schools, family centres, NGOs), rather than health services (3/13).

Over half the families (58%, [931/1614]) had low income, 35% were lone parents [57/1606], or jobless [45/1303]; 30% [492/1651] were from ethnic minorities (Table 2). In one Dutch trial, families identified as mainly Middle Eastern and North African; in one Dutch and three London trials, mainly African-Caribbean, alongside multiple other ethnic groups. In Birmingham, families identified as belonging to 18 different ethnic groups. The mean child age was five (63 months, SD 17.8); 23% [326/1393] of parents reported clinical levels of depressive symptoms. Families in the trial with no IPD available[38] (n=116) were moderately comparable to those in other trials. Children aged 2-8 from three English General Practices were screened for above-average levels of conduct problems. Families were less disadvantaged than the pool average, and 9% [11/115] were from an ethnic minority. Updated 2019 searches found two further eligible trials, not included in the IPD, in Sweden[43] and Netherlands[44]. Both were conducted in community settings, and were aimed at children showing elevated levels of conduct problems. Levels of social disadvantage and ethnic minority status were somewhat lower than the pool average (Table 1).

Main effect of the intervention.

There was a significant overall effect of IY intervention on child conduct problems ($z=10.08$, $p<0.001$), estimated to be a reduction of 13.5 points on ECBI-I (95% CI 10.9 to 16.1). Our sample mean ECBI-I score at baseline was 137 (SD 37) with an intervention group post-test mean of 116 (SD 35), and control group post-test mean of 125; possible range, 35-252. Conduct problem clinical cut-off is 127, thus intervention group children on average moved from 10 points above, to 10 points below the cut-off; control children moved to near the cut-off at post-test. Table 2 summarizes ECBI-I pre-post data by trial arm. Figure 2 shows trial-specific and overall intervention effects. Most trials found that IY reduced conduct problems. Standardised group differences varied from very small (-0.10, equivalent to 3.9 ECBI-I point reduction based on baseline ECBI-I SD=37 points) to moderate/large (-0.65; 23.9 points). Our overall effect was small-to-moderate sized (-0.37, CI -0.44 to -0.29, 13.5 points decrease). As found in aggregate-level reviews,[3] indicated prevention or treatment trials had larger effect sizes,[32,39,42] than selective prevention trials.[30,33,34] Between-trial heterogeneity in intervention effects was moderate (I^2 =42.5%).

Equity analyses: effect moderation by disadvantage and ethnicity?

Five binary variables - low income and education, joblessness, lone and teenage parent status - were used to index aspects of social disadvantage. These displayed moderate-to-large positive associations in our sample (*r*'s 0.16 to 0.54; the strongest correlation was between unemployment and low income). Membership of ethnic minority was not consistently associated with these indices (*r*'s -0.12 to 0.07).

Analyses of interaction effects of participant-level variables showed that, overall, there were no significant moderation effects by any social disadvantage indicator, or by ethnicity (Table 3). Moderation effects at participant level did not usually differ from those at trial level (all p-values > 0.27), and therefore we conducted moderation assessment across trials and participants. Exceptions were teenage parent status and ethnic minority for which

moderation effects differed between- and within- trials at the 10% level (p=0.051 and 0.042 respectively). Therefore, we report the within-trial results for these variables. None of the relevant moderation tests were significant (all p-values > 0.1; Table 3). Figure 3 shows raw data means by trial arm and disadvantage indicator, low income. Table 3 shows estimated effect modification indices for each moderator. Even the largest estimated index (7.3 points) would separate the overall effect (d=-0.37) into two "moderator present" and "moderator absent" effects in the same direction (d=-0.47; d=-0.27). Thus, as well as these moderation indices being statistically insignificant, estimates of their magnitudes are mostly small and unlikely to be clinically important. Notably, both trials retrieved from updated searches reported within-trial moderator analyses[43, 44], finding no differential effects by social disadvantage indicators, including lone parent, low education, income, and immigrant status. In summary, there was no evidence to suggest that the intervention benefits were diminished for families who were disadvantaged by low income or education, joblessness, lone or teenage parent status, or ethnic minority status.

Discussion

This study is the first in the world to use the power of IPD meta-analysis to assess the equity effects of one of the highest quality parenting programmes for reducing conduct problems, a common and disabling childhood mental health condition. Conduct problems are several times more common in disadvantaged groups,[7] and untreated lead to a high prevalence of criminality, mental ill-health and poor social functioning, such as poverty and joblessness[2]. It would be a matter of grave concern if parenting interventions further widened the differences between socioeconomic groups, thereby contributing to an inequitable society. It is plausible that disadvantaged families would derive less benefit than advantaged families, since parenting interventions require social circumstances that allow sufficient time and

organisation to implement required behaviour change. Such interventions have been suggested to be particularly susceptible to adverse equity effects.[9] Our careful synthesis of data from almost all trials of the IY parenting intervention in Europe provided a uniquely large and diverse sample that allowed us to address equity effects using the most stringent and well-powered tests to date.

The findings do not provide any evidence to suggest that families with social disadvantage, and those from ethnic minorities, benefit less from IY than more advantaged families. For most variables, we were powered to detect even a small moderation effect (Cohen's d=0.2 translating into a true moderation index of 7.4 ECBI-I points). Thus, it is highly unlikely that the intervention increases existing social inequalities with respect to the amelioration of child conduct problems. Moreover, there is a reasonable possibility that the programme may reduce social inequity over the long term, if the effects are maintained over the years. To date the only long-term follow-up of European IY trials supported this, with a halving in the rate of Oppositional Defiant Disorder 8-10 years after initial treatment.[45]

IY has features that may enhance its effectiveness across social groups. While its content and principles are similar to other evidence-based parenting programmes derived from social learning theory (e.g. Triple P; PMTO, Parent Management Training-Oregon) it has a particular focus on a collaborative delivery model. Thus, parenting goals and strategies are tailored to families' needs, whether these arise from child characteristics, culture, social disadvantage or family values.[46] This careful attention to individual needs may help enhance intervention effectiveness across a range of families, offsetting differential effects that might otherwise result from educational and behaviour change-based programmes.

This study is notable in providing a sophisticated analysis of equity effects by ethnicity in Europe. These are potentially generalisable findings, based on the inclusion of substantial numbers of families from diverse ethnic groups, with very different cultural and

immigration histories, albeit mainly from six trials in two countries, UK and Netherlands. Due to small numbers in most categories of ethnicity, we were unable to look further at specific groups. The intervention was delivered in groups with mixed ethnicities, and rather than being specifically culturally adapted, is collaborative and flexible in its approach. This suggests that despite variation in parenting styles and values across ethnicities, IY does not need adaptation for separate ethnic groups. This is important in view of the additional practical and financial demands of ethnic-specific service adaptation and delivery. Moreover, separating delivery systems for different ethnic groups is not beneficial for community integration. The lack of any differential effect by ethnicity may be a feature of IY with its built-in flexibility and cultural sensitivity.[46] However, further assessment is needed to test if other parenting programmes show comparable equity effects by SES or ethnicity.[10,47] Additionally, the effects of key delivery components relevant to equity (including IY's collaborative approach, which may be costly to train) require further investigation in dismantling or factorial designs in order to optimise effectiveness.

The study has several further strengths that increase our confidence in the findings. First, our unique pooled data set yielded a large sample and results, which, due to diversity of samples and contexts, are potentially generalisable across countries, service settings and level of child problems. Second, compared to other IPD studies, we obtained an unusually complete set of data from 93% [14/15] of eligible trials, thereby minimising the possibility of availability bias. One toddler trial[40] supplied IPD, but was excluded as it lacked conduct problem outcomes. Although our original searches were conducted in 2015, we updated the searches in 2019, finding two trials whose results were consistent with our IPD analyses. Both showed no moderation by social disadvantage[43, 44], suggesting their inclusion would be unlikely to alter the conclusions. Thirdly, the plausibility of the findings is enhanced by our robust analytic strategies; we included only randomised trials, and accounted for all relevant

trial design features. Finally, we accounted for missing data using multiple imputation, which requires less restrictive assumptions than using completely observed cases.[36]

The study has several limitations. First, we focused on only one parenting programme, albeit one of the most established in Europe, and whose content (if not delivery) is in many ways similar to other evidence-based parenting programmes. Second, we focused only on Europe, in order to better understand the effects of a commonly-imported US programme in the region, and because of availability of a near-complete sample of trials conducted independently of the commercial developer. Third, several assumptions had to be made during data harmonisation, notably that two different instruments measured the same construct of conduct problems, and that indices of low SES (e.g. receipt of benefits) were broadly comparable across countries. Fourth, data were available only on variation in intervention effectiveness; trials lacked data on variation in access to parenting services, another potential source of inequalities. Nevertheless, most trials appeared successful in accessing disadvantaged families, as low-SES and lone-parent families were over-represented, compared to national norms.[23,28-31,37,39,40] Fifth, insufficient data were available from other informants (e.g. fathers, teachers); hence, we could only include data from one parent, usually mother, as primary outcome. Parent-reported conduct problems (typically mother) is the usual primary outcome for trials in this age group. As with any pooled data study, we were able to include only those outcomes that were measured in individual trials. Direct observational measures of conduct problems provide one solution for validating self-reported outcomes,[21] however, these were highly heterogeneous and available only in a subset of trials. Future trials should involve multiple data sources, using common instruments across the field.

To conclude, our pooled IPD meta-analysis, the largest on any conduct problem or parenting intervention, suggests that IY does not increase socioeconomic inequalities in child

conduct problems. This contrasts with earlier findings, based on aggregate-level analyses and weaker trial designs,[12,13] which may have been underpowered, and subject to ecological fallacy. Families from both disadvantaged and advantaged backgrounds, as well as those from ethnic minorities, are all likely to benefit from the intervention. This conclusion likely applies to other parenting interventions with similar content and collaborative delivery, although we cannot be certain without further IPD data. For example, a small Norwegian study combined two trials of another US programme, PMTO, and found no adverse equity effects.[47] Our findings are important for policymakers and commissioners needing to identify, fund and target interventions toward those most at risk, as well as practitioners recommending or delivering parenting programmes. The IY intervention is clearly an effective strategy to reduce conduct problems in families facing social disadvantage. However, to avoid adverse equity effects, it is vital services continue to ensure that the most disadvantaged families can access the programme.

Declaration of interests: FG, PL, JH, SS, VB, SM, MG, MJSS, BOdC, AM, UA, WM led trials that were included in IPD set. JH, MG, MJSS have received occasional payments for training leaders in the Incredible Years parent programme. FG, JH are co-developers of a non-profit parenting programme with WHO, 'Parenting for Lifelong Health'. The authors declare no other conflicts of interest.

Author contributions: FG was Principal Investigator and drafted the paper. FG, JH, SS conceived the study. FG, SS, SL, JH, JB designed the study, wrote the grant and interpreted the findings. SL designed the analyses; SL, VH conducted the analyses. PL managed the study and data, contributed to the design and writing, and interpreted the findings, assisted by

JM. FG, PL, JH, SS, VB, SM, MG, MJSS, BOdC, AM, UA, WM led included trials and contributed trial data and interpretation of findings. MW and EMB contributed to interpretation of findings. All authors commented upon and approved the final manuscript.

Acknowledgements: NIHR Public Health Research funded the study, grant 12-3070-04, PI FG. SL received salary support from the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. VB's time is supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South West Peninsula (NIHR CLAHRC South West Peninsula). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Erskine HE, Norman RE, Ferrari AJ, et al. Long-Term Outcomes of Attention-Deficit/Hyperactivity Disorder and Conduct Disorder: A Systematic Review and Meta-Analysis. *J Am Acad Child Adolesc Psychiatry* 2016; **55**(10): 841-50.
2. Rivenbark JG, Odgers CL, Caspi A, et al. The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *J Child Psychol Psychiatry* 2018; **59**(6): 703-10.
3. Leijten P, Gardner F, Melendez-Torres GJ, et al. What to Teach Parents to Reduce Disruptive Child Behavior: Two Meta-Analyses of Parenting Program Componnts. *J Am Acad Child Adolesc Psychiatry* 2019. doi.org/10.1016/j.jaac.2018.07.900
4. van Aar J, Leijten P, Orobio de Castro B, Overbeek G. Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clin Psychol Rev* 2017; **51**: 153-63.
5. WHO. Inspire: Seven strategies for ending violence against children. Geneva: WHO, 2016.
6. Scott S. A national approach to improving child and adolescent mental health care: The children and young people's improving access to psychological therapies (IAPT) programme in England. In: Weisz JR, Kazdin AE, eds. *Evidence-based psychotherapies for children and adolescents*. NY: Guilford Press 2017.
7. Piotrowska PJ, Stride CB, Maughan B, Goodman R, McCaw L, Rowe R. Income gradients within child and adolescent antisocial behaviours. *Br J Psychiatry* 2015; **207**(5): 385-91.
8. Kouider EB, Koglin U, Petermann F. Emotional and behavioral problems in migrant children and adolescents in Europe: a systematic review. *Eur Child Adolesc Psychiatry* 2014; **23**(6): 373-91.
9. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health* 2013; **67**(2): 190-3.
10. Gardner F, Connell A, Trentacosta CJ, Shaw DS, Dishion TJ, Wilson MN. Moderators of Outcome in a Brief Family-Centered Intervention for Preventing Early Problem Behavior. *J Consult Clin Psychol* 2009; **77**(3): 543-53.
11. Leijten P, Raaijmakers MAJ, de Castro BO, Matthys W. Does Socioeconomic Status Matter? A Meta-Analysis on Parent Training Effectiveness for Disruptive Child Behavior. *J Clin Child Adolesc Psychol* 2013; **42**(3): 384-92.
12. Lundahl B, Risser HJ, Lovejoy MC. A meta-analysis of parent training: Moderators and follow-up effects. *Clin Psychol Rev* 2006; **26**(1): 86-104.
13. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *J Child Psychol Psychiatry* 2006; **47**(1): 99-111.
14. Belsky J, Melhuish E, Barnes J, Leyland AH, Romaniuk H. Effects of Sure Start local programmes on children and families: early findings from a quasi-experimental, cross sectional study. *BMJ* 2006; **332**(7556): 1476.
15. Gardner F, Montgomery P, Knerr W. Transporting Evidence-Based Parenting Programs for Child Problem Behavior (Age 3-10) Between Countries: Systematic Review and Meta-Analysis. *J Clin Child Adolesc Psychol* 2016; **45**(6): 749-62.
16. Reid M, Webster-Stratton C, Beauchaine TP. Parent Training in Head Start: A Comparison of Program Response Among African American, Asian American, Caucasian, and Hispanic Mothers. *Prev Sci* 2001; **2**(4): 209-27.
17. van Mourik K, Crone MR, de Wolff MS, Reis R. Parent Training Programs for Ethnic Minorities: a Meta-analysis of Adaptations and Effect. *Prev Sci* 2017; **18**(1): 95-105.

18. Castro FG, Barrera M, Holleran Steiker LK. Issues and Challenges in the Design of Culturally Adapted Evidence-Based Interventions. *Annu Rev Clin Psychol* 2010; **6**: 213-39.
19. Petticrew M, Tugwell P, Kristjansson E, Oliver S, Ueffing E, Welch V. Damned if you do, damned if you don't: subgroup analysis and equity. *J Epidemiol Community Health* 2012; **66**(1): 95-8.
20. Fisher DJ, Carpenter JR, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017; **356**.
21. Menting AT, Orobio de Castro B, Matthys W. Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: a meta-analytic review. *Clin Psychol Rev* 2013; **33**(8): 901-13.
22. Eyberg SM, Ross AW. Assessment of child behavior problems: The validation of a new inventory. *J Clin Child Psychol* 1978; **7**(2): 113-6.
23. Scott S, Spender Q, Doolan M, Jacobs B, Aspland H. Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *BMJ* 2001; **323**(7306): 194-8.
24. Seabra-Santos MJ, Gaspar MF, Azevedo AF, et al. Incredible Years parent training: What changes, for whom, how, for how long? *J Appl Dev Psychol* 2016; **44**: 93-104.
25. Taylor E, Schachar R, Thorley G, Wieselberg M. Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *Br J Psychiatry* 1986; **149**(6): 760-7.
26. Taylor E, Sandberg S, Thorley G. The epidemiology of childhood hyperactivity (Vol. 33). Oxford: Oxford University Press; 1991.
27. Burns GL, Patterson DR. Normative Data on the Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory: Parent and Teacher Rating Scales of Disruptive Behavior Problems in Children and Adolescents. *Child Fam Behav Ther* 2001; **23**(1): 15-28.
28. Scott S, Sylva K, Doolan M, et al. Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: the SPOKES project. *J Child Psychol Psychiatry* 2010b; **51**(1): 48-57.
29. Gardner F, Burton J, Klimes I. Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *J Child Psychol Psychiatry* 2006; **47**(11): 1123-32.
30. Scott S, O'Connor TG, Futh A, Matias C, Price J, Doolan M. Impact of a parenting program in a high-risk, multi-ethnic community: the PALS trial. *J Child Psychol Psychiatry* 2010a; **51**(12): 1331-41.
31. Scott S, Sylva K, Kallitsoglou A, Ford T. Which Type of Parenting Programme Best Improves Child Behaviour and Reading? Follow-Up of the Helping Children Achieve Trial. London: Nuffield Foundation; 2014.
32. Axberg U, Broberg AG. The Incredible Years: Evaluation of the Webster Stratton parent management training programme in Sweden. *Scand J Psychol* 2012; **53**: 224-32.
33. Menting ATA, de Castro BO, Wijngaards-de Meij LDNV, Matthys W. A trial of parent training for mothers being released From incarceration and their children. *J Clin Child Adolesc Psychol* 2014; **43**(3): 381-96.
34. Leijten P, Raaijmakers MA, Orobio de Castro B, van den Ban E, Matthys W. Effectiveness of the Incredible Years Parenting Program for Families with Socioeconomically Disadvantaged and Ethnic Minority Backgrounds. *J Clin Child Adolesc Psychol* 2017; **46**(1): 59-73.
35. UNESCO. International Standard Classification of Education 2011. Montreal: UNESCO Institute for Statistics, 2011.

36. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011; **30**(4): 377-99.
37. Morpeth L, Blower S, Tobin K, et al. The effectiveness of the Incredible Years pre-school parenting programme in the United Kingdom: a pragmatic randomised controlled trial. *Child Care in Practice* 2017; **23**(2): 141-61.
38. Patterson J, Barlow J, Mockford C, Klimes I, Pyper C, Stewart-Brown S. Improving mental health through parenting programmes: block randomised controlled trial. *Arch Dis Child* 2002; **87**(6): 472-7.
39. Hutchings J, Bywater T, Daley D, et al. Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ* 2007; **334**(7595): 678.
40. Hutchings J, Griffith N, Bywater T, Williams M. Evaluating the Incredible Years Toddler Parenting Programme with parents of toddlers in disadvantaged (Flying Start) areas of Wales. *Child Care Health Dev* 2017; **43**(1): 104-13.
41. McGilloway S, Ni Mhaille G, Bywater T, et al. A parenting intervention for childhood behavioral problems: a randomized controlled trial in disadvantaged community-based settings. *J Consult Clin Psychol* 2012; **80**(1): 116-27.
42. Larsson B, Fossum S, Clifford G, Drugli MB, Handegard BH, Morch WT. Treatment of oppositional defiant and conduct problems in young Norwegian children: results of a randomized controlled trial. *Eur Child Adolesc Psychiatry* 2009; **18**(1): 42-52.
43. Stattin H, Enebrink P, Özdemir M, Giannotta F. A national evaluation of parenting programs in Sweden: the short-term effects using an RCT effectiveness design. *J Consult Clin Psychol* 2015; **83**(6): 1069-84.
44. Weeland J, Chhangur RR, van der Giessen D, Matthys W, de Castro BO, Overbeek G. Intervention effectiveness of the Incredible Years: New insights into sociodemographic and intervention-based moderators. *Behav Ther* 2017; **48**(1): 1-18. See also Corrigendum, *Behav Ther* 2018; **49**(2): 308.
45. Scott S, Briskman J, O'Connor TG. Early prevention of antisocial personality: long-term follow-up of two randomized controlled trials comparing indicated and selective approaches. *Am J Psychiatry* 2014; **171**(6): 649-57.
46. Webster-Stratton C. Affirming diversity: Multi-cultural collaboration to deliver the incredible years parent programs. *Int J Child Health Hum Dev* 2009; **2**(1): 17-32.
47. Tømmeraas T. Social gradients and participant characteristics in child behavior problem interventions. *Child Youth Serv Rev* 2016; **70**: 57-64.

Table 1. Characteristics of trials that met inclusion criteria

Trial	Lead author (year)	Country	Setting	Screened for conduct problems	*N*	Child age (*M*)	% low income	% ethnic minority
#1	Larsson (2009)[42]	Norway	Outpatient psychiatric clinics	Yes	75	3–8 (6.58)	25	1
#2	Axberg (2012)[32]	Sweden	Outpatient psychiatric clinics	Yes	62	3–8 (5.97)	41	0
#3	Seabra-Santos (2016)[24]	Portugal	University clinics	Yes	124	3–6 (4.66)	0	0
#4	McGilloway (2012)[41]	Ireland	Community services	Yes	149	2–7 (4.84)	47	6
#5	Menting (2014)[33]	Nether-lands	Community services	No	99	1–11 (6.30)	93	78
#6	Leijten (2017)[34]	Nether-lands	Outpatient psychiatric clinics & schools	Yes & No	156	2–8 (5.59)	74	65
#7	Hutchings (2007)[39]	Wales	Community services	Yes	153	3–4 (3.84)	79	1
#8*	Hutchings (2017)[40]	Wales	Community services	No	103	0–2 (1.85)	56	0
#9	Morpeth (2017)[37]	England	Community services	Yes	161	2–4 (3.68)	63	52
#10	Scott (2010b)[28]	England	Schools	Yes	112	4–6 (5.21)	44	40
#11	Scott (2010a)[30]	England	Schools	No	174	4–6 (5.50)	44	75
#12	Scott (2014)[31]	England	Schools	Yes	214	3–7 (6.07)	80	19
#13	Gardner (2006)[29]	England	Community services	Yes	76	2–9 (5.93)	64	2
#14	Scott (2001)[23]	England	Outpatient psychiatric clinics	Yes	141	2–10 (5.67)	58	15
#15*	Patterson 2002[38]	England	General practice	Yes	116	2-8	25	9
Updated searches, 2019, IPD not included								
U1	Stattin 2015[43]	Sweden	Community services	No	277	3-12 (6.8)	6	18
U2	Weeland 2017[44]	Nether-lands	Community services	Yes	387	4-8 (6.2)	^21	14

* #15, IPD not available; #8, excluded from analyses, as no data on primary outcome
^ based on parent low education, as no data collected on low income

Table 2. Summaries for demographics and clinical outcome by randomised group, pooled sample, 13 trials.

	Total N, max 1696	# trials info available	Control (max N, 650)		Incredible Years (max N, 1046)	
Variable (categorical)	N	K	N	Percent	N	Percent
Child gender (male)	1696	13	650	63.8	1046	63.1
Low income	1614	13	615	57.9	999	57.6
Low education	1561	13	581	39.8	980	43.3
Lone parent	1606	13	606	33.0	1000	36.8
Teen parent	1609	12	605	12.6	1004	11.7
Unemployed	1303	11	522	30.3	781	37.5
Ethnic minority	1651	13	629	30.0	1022	30.9
Continuous variables:			Mean (SD)			Mean (SD)
Child conduct problem score ECBI-I- Baseline	1622	13	611	135.5 (37.0)	1011	139.4 (37.0)
ECBI-I Post intervention	1445	13	567	125.5 (37.9)	878	116.2 (34.7)
Child age (months)	1682	13	643	64.2 (16.9)	1039	62.4 (18.3)

Table 3. Results of formal assessments of within-trial moderators.

Effects are expressed on the ECBI-I scale.

Social deprivation variable	% with modifier	Estimated** intervention effect, ECBI-I	Estimated moderation Index	95% C.I.	p-value
Low income	58	-11.52	1.91	–4.77 to 8.59	0.58
Not low income	42	-13.43			
Low education	39	-9.88	4.37	–2.17 to 10.90	0.29
Not low education	61	-14.25			
Lone parent	35	-12.57	0.5	-6.1 to 7.1	0.88
Not lone parent	65	-13.07			
Teen parent	12	-6.36	7.32*	–2.24 to 16.87*	0.13*
Not teen parent	88	-13.68			
Unemployed	34	-8.17	4.88	–2.67 to 12.42	0.21
Employed	66	-13.05			
Ethnic minority	30	-14.96	-1.37*	–9.81 to 7.08*	0.75*
Not ethnic minority	70	-13.59			

*Separate interaction terms fitted for between- and within-trial effects. Reported inferences refer to within-trial moderation effects.

**Moderators were modelled one at a time and thus adjustments for missing data biases can vary between models.

Figure 1. Flow chart for included studies

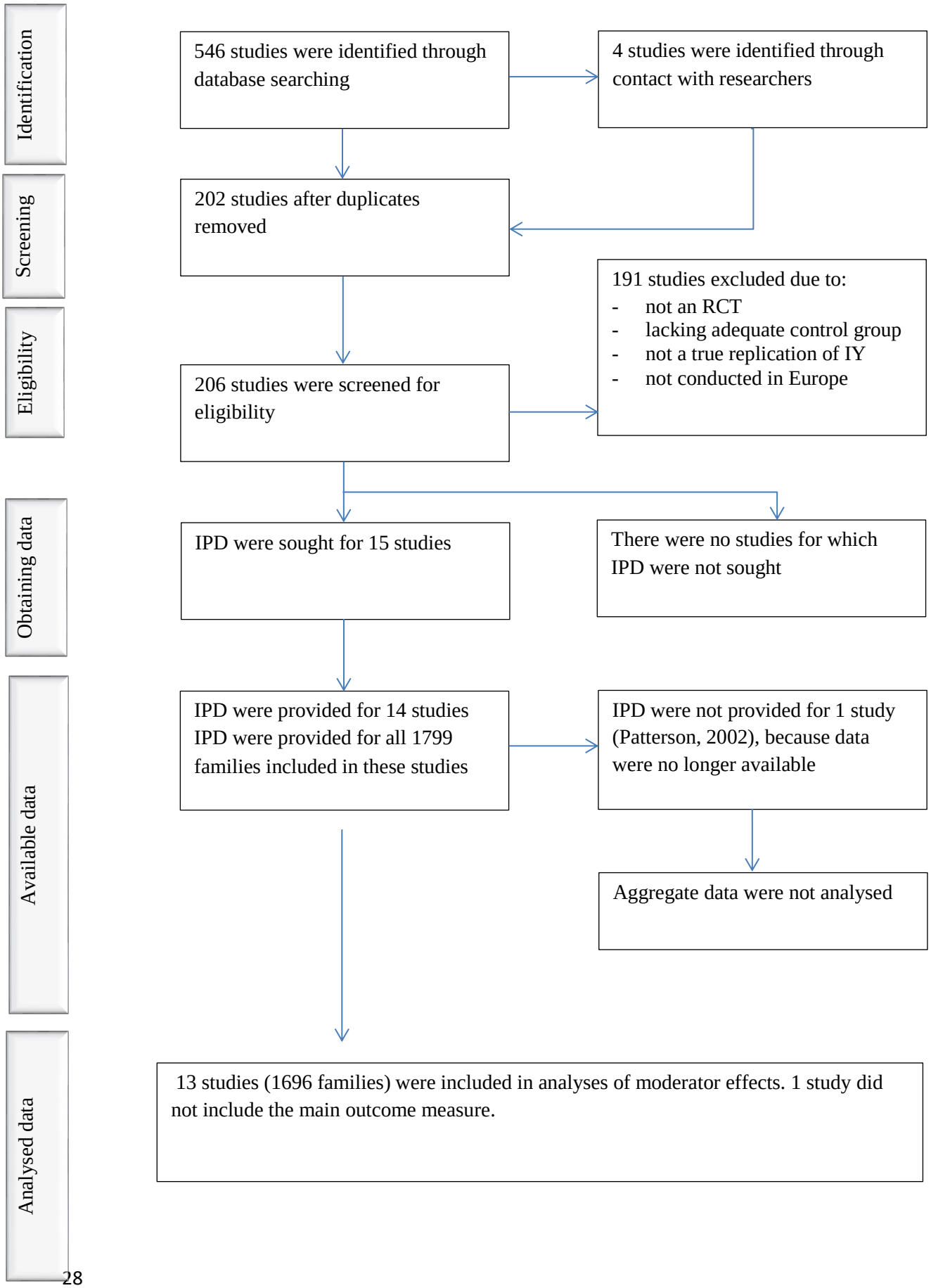


Figure 2. Forest plot of estimated effects of IY parenting programme (difference IY – Control on ECBI-I scale) on child conduct problems.

Symptom reductions indicate a benefit of IY. Effects were estimated based on the IPD by fitting the (simplified) analysis model to each trial separately.

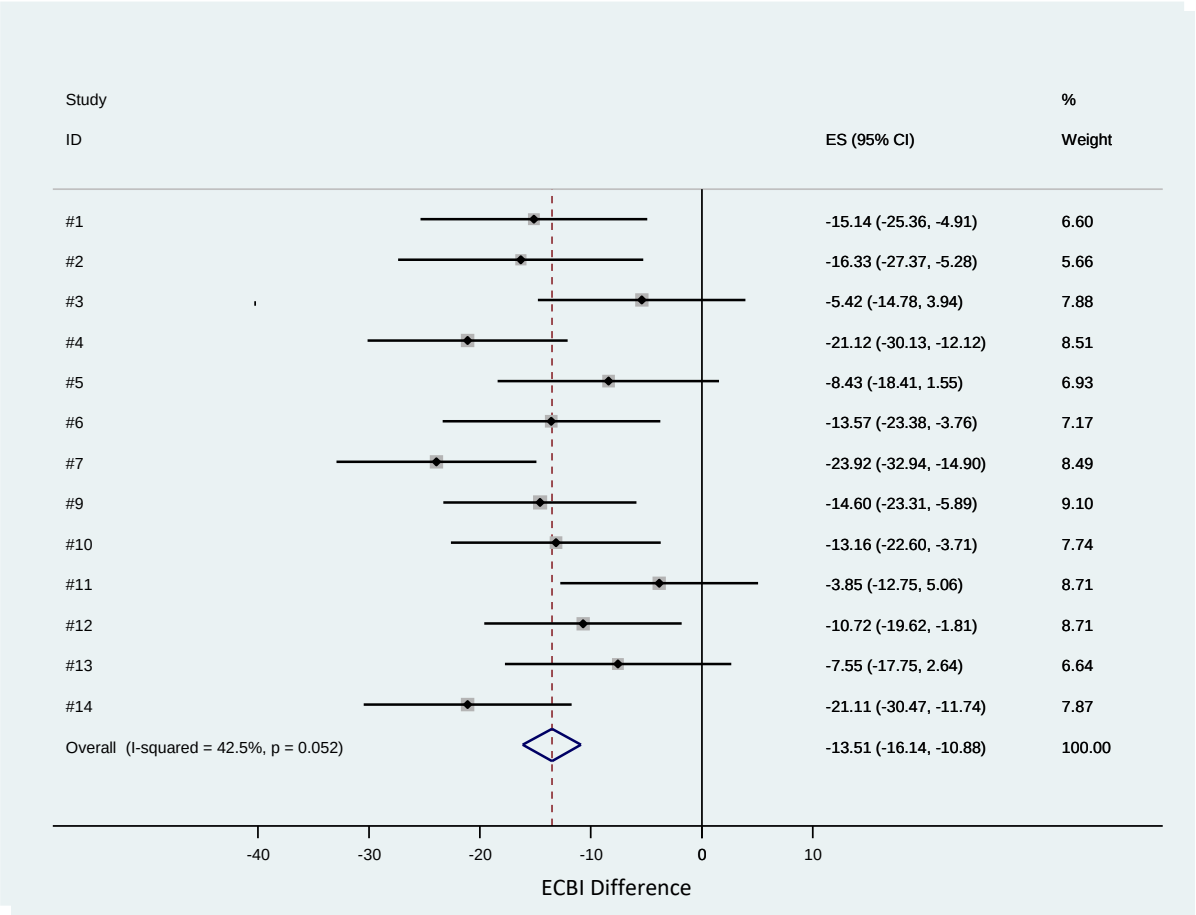
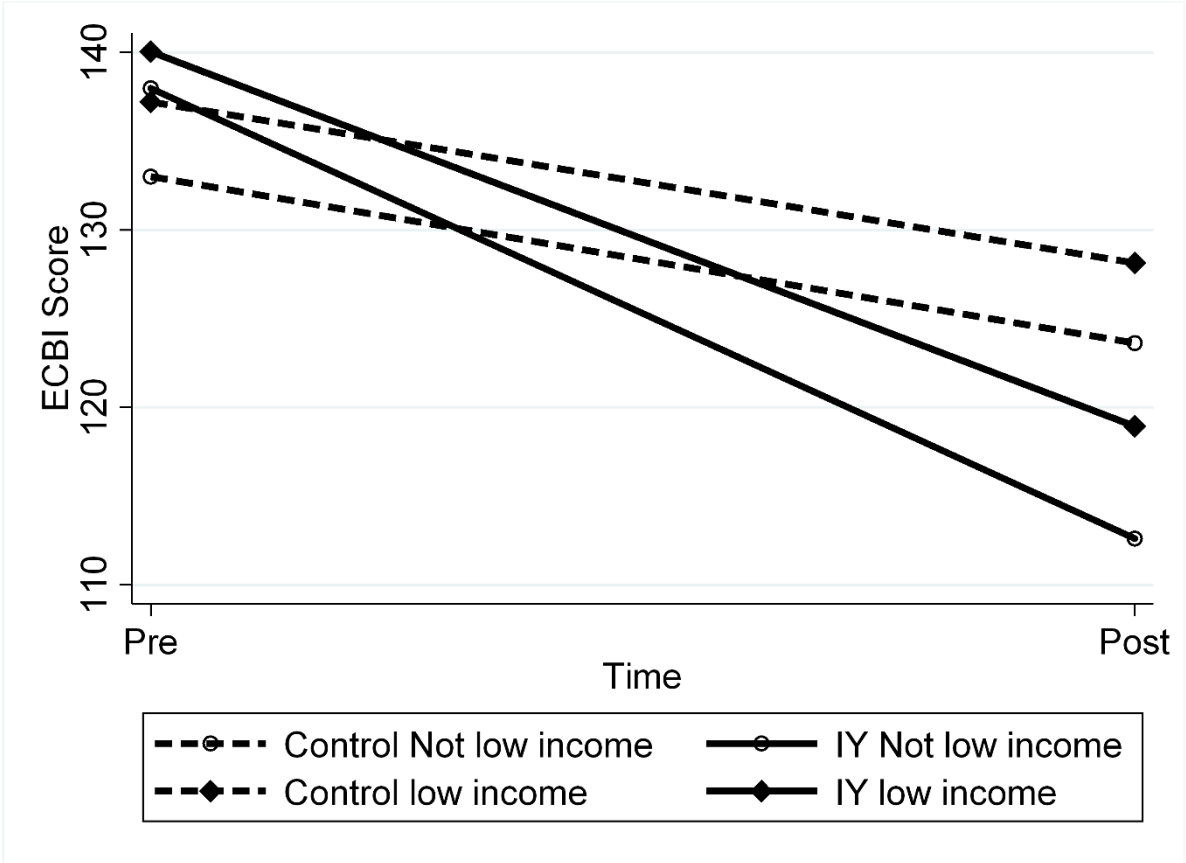


Figure 3. Moderation of intervention effects on child conduct problems (unadjusted raw mean ECBI-I scores) by low income; *N*=1614



Web-extra (supplementary) materials

W1:
Description of Incredible Years Parenting programme

W2
Checklist PRISMA IPD reporting guideline (Stewart et al., 2015), attached separately.

W3:
Search strategy

W4
Details of harmonisation procedures

W5:
Risk of bias table, across studies