Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

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

W1: Description of Incredible Years Parenting programme

This suite of parenting programmes was developed in Seattle by Professor Carolyn Webster-Stratton, University of Washington. Like many similar evidence-based programmes, it is based on social learning theory, and was originally aimed at treatment of children aged 3-8 with identified conduct problems. It was subsequently employed as a preventive programme for improving parenting and reducing risk of conduct problems, in families at risk due to social disadvantage. A distinctive feature of the IY programmes is their collaborative group-based model, enabling parents to recognise their skills, and empowering them to identify effective strategies to achieve their goals, in ways that fit with their own family context and values. Access to groups is facilitated by providing food, child care and sometimes transport. Parenting behaviour change is aided by practice-based methods including problem solving, discussion of videos illustrating different parenting strategies, role-play practice and homework assignments. Parents work together to find solutions that will work for their family circumstances, meeting weekly in groups of 10-14, generally for 12-14 weeks with two trained group leaders. Leaders are often psychologists or social workers, but this is not a requirement; IY is often delivered, including in many of the European trials by teachers, day care workers, health visitors, nurses and others. The programmes are manualised.

Most of the evidence for IY (and hence, most of the trials included in our IPD meta-analysis) is based on RCTs of the IY ‘BASIC’ 12-14 week parenting programme (Webster-Stratton & Reid, 2010) compared most commonly to waiting-list control groups. More recently, newer versions are available for specific age groups, including babies, toddlers and school age, some with 18 or more sessions, as described in Webster-Stratton (2016), but most of these do not yet have an evidence base. The basic parent programme encourages warm relationships between parents and children, by emphasising the importance of responsive play with children, and encouraging positive behaviours through limit-setting and constructive discipline strategies. These help parents learn strategies that will serve as alternatives to physical punishment or other harsh disciplining of children.

IY has a particularly strong emphasis on fidelity to the programme, through a high-quality, rigorous training system, involving regular supervision, and submission of videotaped group material, for feedback and eventual certification. The certification process supports the development of locally based trainers and mentors in implementing countries. Manuals and materials are not available freely, but are supplied along with training by a private for-profit company based in Seattle. The process is exacting and costly, but ensures the programme is delivered with fidelity, with the aim that it should achieve similar results in the ‘real world’ to those obtained in randomised trials.
References and further reading:


W3: Search strategy

Eligibility criteria (PRISMA-IPD #6)
We sought to include all completed randomised controlled trials of the IY parenting programme in Europe for children aged 1–12 years. Non-randomised controlled trials were excluded because no causal inference about the effects of the IY programme can be drawn from non-randomised designs. No restrictions were placed on the years trials were conducted, required minimum follow-up, or included outcome measures. Within each randomised controlled trial we included individuals who had received the IY (or a combination of IY and a reading intervention that focused on similar parenting behaviour) and individuals in control conditions. We excluded trials with additional non-parenting programmes, as well as excluding individuals who had received additional treatment for disruptive behaviour such as the IY child programme, because the focus of this project was to examine the effect of the parenting programme as a sole intervention. We excluded programmes that were much more minimal than the standard IY programme of 12–14 sessions, for example, highly abbreviated non-standard versions. We also excluded individuals that received a reading intervention only.

Identifying studies—information sources (PRISMA-IPD #7)
Studies were identified through (1) a systematic literature search in the following databases: CINAHL, Embase, Global Health, MEDLINE and PsycINFO, (2) the Incredible Years website, which provides information on trials evaluating IY, (3) the European IY mentors’ network, and (4) asking experts. Searches in January 2015 revealed no further completed trials.

Identifying studies—search (PRISMA-IPD #8)
Embase, Global Health, MEDLINE (<1946 to Present) and PsycINFO were searched via OVID using the following search terms: 1. incredible year$.mp; 2. webster-stratton.mp; 3. 1 or 2. CINAHL was searched via EBSCO using the phrase “incredible years”.

Study selection processes (PRISMA-IPD #9)
Eligibility was assessed by the first author and double checked by four additional authors (SS, JH, PL; JM). There were no differences of opinion.

W4. Details of harmonisation procedures
Two different data harmonisation strategies were used for the present analyses:

1. Combining similar classification systems to harmonise data for different socioeconomic status indicators.
Indicators of socioeconomic status were screened for comparability. Fortunately, most trials used similar indicators that were operationalised in similar ways. For example, of the 15 included trials, 14 trials supplied data, and among these, whether a family had low income was defined by i) receiving financial benefits (10 trials), receiving financial benefits and having below-median income (1 trial), ii) scoring below the low
socioeconomic status threshold on the Hollingshead Index (1 trial), or iii) living in social housing or living in the household of family/friends (2 trials). Indicators of families’ (risk for) low income were dichotomised. Educational level categories used across trials were compared with the UNESCO Institute for Statistics ISCED-11. Although some categories had to be combined (e.g. less than primary education and primary education) because these had already been combined in some trial datasets, five main categories (primary education or less, lower secondary education, upper secondary education, post-secondary education, and university degree) were present in data from all trials and used in the final pooled dataset.

2. **Using norm deviation scores to harmonize scores on child conduct problems.** A primary measure (i.e., the most frequently used measure) was selected for each construct. If data on the primary measure was unavailable, data from similar measures were converted into scores on the primary measure using norm deviation scores (i.e., number of standard deviations the individual scores are above or below the population mean). This approach assumes that both instruments measure the same construct with the same measurement error on different instruments and thus the scores can be converted using known population characteristics.

Norm deviation scores are similar to z-scores, in that values are standardised by taking off a mean dividing by an SD. However, they differ in that we used external information from published norms²⁷,²⁸ to supply the population mean and standard deviation, while standard z-scoring estimates these from the sample at hand.

Advantages of using norm deviation scores are (1) that absolute scores can easily be interpreted, because they are on the original scale of the primary measure. This allows for interpretation of clinical significance of intervention effects; (2) that the scores of individuals remain the same after adding data from new trials, because harmonisation is done on an individual family level.

In contrast, integrative data analysis using latent variables based on the different measures included across trials, although strong in its use of multiple measures within trials, has the disadvantage of scores that are hard to interpret on an absolute level, as well as scores that depend on the model tested and thus the trials or families included in the model. Norm deviation scores were used to harmonise data from the following constructs: Parent-reported disruptive child behaviour (reported here), children’s co-morbid ADHD symptoms, children’s co-morbid emotional problems, and parental depression.

**W5. Risk of bias across studies**