

## Public Health Research

Volume 10 • Issue 8 • June 2022

ISSN 2050-4381

# A proportionate, universal parenting programme to enhance social-emotional well-being in infants and toddlers in England: the E-SEE Steps RCT

*Tracey Bywater, Vashti Berry, Sarah Blower, Matthew Bursnall, Edward Cox, Amanda Mason-Jones, Sinéad McGilloway, Kirsty McKendrick, Siobhan Mitchell, Kate Pickett, Gerry Richardson, Kiera Solaiman, M Dawn Teare, Simon Walker and Karen Whittaker*





# A proportionate, universal parenting programme to enhance social-emotional well-being in infants and toddlers in England: the E-SEE Steps RCT

Tracey Bywater<sup>1\*</sup>, Vashti Berry<sup>2</sup>, Sarah Blower<sup>1</sup>,  
Matthew Bursnall<sup>3</sup>, Edward Cox<sup>4</sup>,  
Amanda Mason-Jones<sup>1</sup>, Sinéad McGilloway<sup>5</sup>,  
Kirsty McKendrick<sup>3</sup>, Siobhan Mitchell<sup>2</sup>,  
Kate Pickett<sup>1</sup>, Gerry Richardson<sup>4</sup>, Kiera Solaiman<sup>3</sup>,  
M Dawn Teare<sup>3,6</sup>, Simon Walker<sup>4</sup>  
and Karen Whittaker<sup>7</sup>

<sup>1</sup>Department of Health Sciences, University of York, York, UK

<sup>2</sup>College of Medicine and Health, University of Exeter, Exeter, UK

<sup>3</sup>Sheffield Clinical Trials Research Unit, Sheffield, UK

<sup>4</sup>Centre for Health Economics, University of York, York, UK

<sup>5</sup>Centre for Mental Health and Community Research, Maynooth University, Maynooth, Ireland

<sup>6</sup>Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

<sup>7</sup>School of Nursing, University of Central Lancashire, Preston, UK

\*Corresponding author

**Declared competing interests of authors:** Tracey Bywater is a trustee for Children's Early Intervention Trust (CEIT) (Bangor, Wales). Early Intervention Wales Training (EIWT) (Bangor, Wales) is owned by CEIT and offers training courses, including Incredible Years®. Trustees do not benefit financially from trainings or CEIT/EIWT activities. Enhancing Social and Emotional health in the Early years study trainings were arranged via the Incredible Years developer in the USA. In addition, Tracey Bywater is supported by the National Institute for Health and Care Research (NIHR) Yorkshire and Humber Applied Research Collaboration [(ARC-YH) reference NIHR200166]. Vashti Berry is supported by the South West Peninsula Applied Research Collaboration [(PenARC) reference NIHR200167]. Sarah Blower is supported by the NIHR ARC-YH (reference NIHR200166). Matthew Bursnall has received funding from the NIHR Public Health Research programme (reference 15/49/32). Kate Pickett is supported by the NIHR ARC-YH (reference NIHR200166). Gerry Richardson is supported by the NIHR ARC-YH (reference NIHR200166). M Dawn Teare has received funding from the NIHR Public Health Research programme (reference 15/49/32). Karen Whittaker reports grants from the Community Practitioners' and Health Visitors' Association Education and Development Trust (London, UK) to support the fathers and co-parents substudy during the conduct of the study.

Published June 2022

DOI: 10.3310/BCFV2964



This report should be referenced as follows:

Bywater T, Berry V, Blower S, Bursnall M, Cox E, Mason-Jones A, *et al.* A proportionate, universal parenting programme to enhance social-emotional well-being in infants and toddlers in England: the E-SEE Steps RCT. *Public Health Res* 2022;**10**(8). <https://doi.org/10.3310/BCFV2964>



# Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

The full PHR archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/phr](http://www.journalslibrary.nihr.ac.uk/phr). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

## Criteria for inclusion in the *Public Health Research* journal

Reports are published in *Public Health Research* (PHR) if (1) they have resulted from work for the PHR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Public Health Research* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## PHR programme

The Public Health Research (PHR) programme, part of the National Institute for Health and Care Research (NIHR), is the leading UK funder of public health research, evaluating public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health.

For more information about the PHR programme please visit the website: <https://www.nihr.ac.uk/explore-nihr/funding-programmes/public-health-research.htm>

## This report

The research reported in this issue of the journal was funded by the PHR programme as project number 13/93/10. The contractual start date was in February 2015. The final report began editorial review in October 2020 and was accepted for publication in October 2021. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PHR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the PHR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the PHR programme or the Department of Health and Social Care.

Copyright © 2022 Bywater *et al.* This work was produced by Bywater *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source - NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## NIHR Journals Library Editor-in-Chief

---

**Professor Ken Stein** Professor of Public Health, University of Exeter Medical School, UK

## NIHR Journals Library Editors

---

**Professor John Powell** Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Professor of Digital Health Care, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

**Professor Andrée Le May** Chair of NIHR Journals Library Editorial Group (HSDR, PGfAR, PHR journals) and Editor-in-Chief of HSDR, PGfAR, PHR journals

**Professor Matthias Beck** Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Eugenia Cronin** Consultant in Public Health, Delta Public Health Consulting Ltd, UK

**Dr Peter Davidson** Consultant Advisor, Wessex Institute, University of Southampton, UK

**Ms Tara Lamont** Senior Adviser, Wessex Institute, University of Southampton, UK

**Dr Catriona McDaid** Reader in Trials, Department of Health Sciences, University of York, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Emeritus Professor of Wellbeing Research, University of Winchester, UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professor of Child Health Research, Child and Adolescent Mental Health, Palliative Care and Paediatrics Unit, Population Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, London, UK

**Professor Jonathan Ross** Professor of Sexual Health and HIV, University Hospital Birmingham, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

**Professor Ken Stein** Professor of Public Health, University of Exeter Medical School, UK

**Professor Jim Thornton** Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of editors: [www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)



# Abstract

## A proportionate, universal parenting programme to enhance social-emotional well-being in infants and toddlers in England: the E-SEE Steps RCT

Tracey Bywater <sup>1\*</sup> Vashti Berry <sup>2</sup> Sarah Blower <sup>1</sup>  
Matthew Bursnall <sup>3</sup> Edward Cox <sup>4</sup> Amanda Mason-Jones <sup>1</sup>  
Sinéad McGilloway <sup>5</sup> Kirsty McKendrick <sup>3</sup> Siobhan Mitchell <sup>2</sup>  
Kate Pickett <sup>1</sup> Gerry Richardson <sup>4</sup> Kiera Solaiman <sup>3</sup>  
M Dawn Teare <sup>3,6</sup> Simon Walker <sup>4</sup> and Karen Whittaker <sup>7</sup>

<sup>1</sup>Department of Health Sciences, University of York, York, UK

<sup>2</sup>College of Medicine and Health, University of Exeter, Exeter, UK

<sup>3</sup>Sheffield Clinical Trials Research Unit, Sheffield, UK

<sup>4</sup>Centre for Health Economics, University of York, York, UK

<sup>5</sup>Centre for Mental Health and Community Research, Maynooth University, Maynooth, Ireland

<sup>6</sup>Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

<sup>7</sup>School of Nursing, University of Central Lancashire, Preston, UK

\*Corresponding author [tracey.bywater@york.ac.uk](mailto:tracey.bywater@york.ac.uk)

**Background:** Mental disorders have become a public health crisis. Early prevention is key. Parenting programmes are effective for children aged  $\geq 3$  years; however, there is a lack of evidence of their effectiveness for children aged  $\leq 2$  years.

**Objectives:** To establish if the model named Enhancing Social–Emotional Health and Well-being in the Early Years (E-SEE) Steps can (1) enhance child social emotional well-being and establish whether or not it is cost-effective at 20 months of age when compared with services as usual; and (2) be delivered as a proportionate universal model with fidelity.

**Design:** A pragmatic two-arm randomised controlled trial and economic appraisal, with an embedded process evaluation to examine the outcomes, implementation and cost-effectiveness of the intervention, and intervention uptake, compared with services as usual. The study had an external pilot phase (which was originally planned as an internal pilot).

**Setting:** The intervention was delivered in community settings by early years children’s services and/or public health staff in four sites.

**Participants:** A total of 341 parents of infants aged  $\leq 8$  weeks were randomised in a ratio of 5 : 1 (intervention,  $n = 285$ ; control,  $n = 56$ ). The target sample was 606 parents.

**Intervention:** Two Incredible Years<sup>®</sup> parenting programmes (i.e. infant and toddler) delivered in a proportionate universal model with three levels [one universal (book) and two targeted group-based parenting programmes].

**Main outcome measures:** Child social and emotional well-being (primary outcome) was assessed using the Ages and Stages Questionnaire: Social and Emotional, 2nd edition, at 2, 9 and 18 months after randomisation. Parent depression (secondary key outcome) was assessed using the Patient Health Questionnaire-9 items. Both questionnaires were eligibility screeners for targeted groups.

**Results:** The primary outcome analysis provided no evidence that the E-SEE Steps model was effective in enhancing child social and emotional well-being. The adjusted mean difference was 3.02 on the original Ages and Stages Questionnaire: Social and Emotional, 2nd edition, in favour of the control [95% confidence interval -0.03 to 6.08;  $p = 0.052$ ;  $N = 321$  (intervention,  $n = 268$ ; control,  $n = 53$ )]. Analysis of the key secondary outcome (i.e. parent depression levels as assessed by the Patient Health Questionnaire-9 items) provided weak evidence on the Patient Health Questionnaire-9 items in favour of the intervention (adjusted mean difference -0.61, 95% confidence interval -1.34 to 0.12;  $p = 0.1$ ). Other secondary outcomes did not differ between arms. The economic analysis showed that the E-SEE Steps model was associated with higher costs and was marginally more effective (0.031 quality-adjusted life-years gained from E-SEE Steps compared with SAU, 95% confidence interval -0.008 to 0.071) than services as usual, resulting in an incremental cost-effectiveness ratio of approximately £20,062 per quality-adjusted life-year compared with services as usual. Overall take-up of the targeted parenting programmes was low. Sites, although enthusiastic, identified barriers to delivering the intervention.

**Limitations:** The target sample size was not met and the study was not powered to explore the effectiveness of each level of intervention. Most parents in the sample were well educated and, therefore, the results are unlikely to be generalisable, particularly to those at greatest risk of poor social and emotional well-being.

**Conclusions:** The E-SEE Steps proportionate universal model did not enhance child social and emotional well-being, but generated non-significant improvements in parent health outcomes, resulting in considerable uncertainty around the cost-effectiveness of the intervention. The primary and key secondary outcome gave inconsistent signals. Although, with system changes, increased resources and adaptations to the intervention, the model could be implemented, evidence for positive outcomes from the E-SEE Steps model is poor.

**Future work:** The universal-level E-SEE Step data (i.e. the *Incredible Years* book) from the external pilot will be pooled with the main trial data for further exploration up to follow-up 1, which is the time point at which most change was seen.

**Trial registration:** This trial is registered as ISRCTN11079129.

**Funding:** This project was funded by the National Institute for Health and Care Research (NIHR) Public Health Research programme and will be published in full in *Public Health Research*; Vol. 10, No. 8. See the NIHR Journals Library website for further project information.

# Contents

List of tables	xiii
List of figures	xv
List of supplementary material	xvii
List of abbreviations	xix
Plain English summary	xxi
Scientific summary	xxiii
<b>Chapter 1 Introduction</b>	<b>1</b>
Social and emotional well-being	1
Parenting interventions to promote infant social and emotional well-being	1
The case of proportionate universal interventions	2
Lack of proportionate universal interventions and evidence base	2
The Incredible Years parenting programmes	2
Introduction to the E-SEE Steps model and its components	3
Explanation of rationale	3
Specific objectives	4
<b>Chapter 2 Methods</b>	<b>5</b>
Pilot phase	5
Main trial design	7
Participants and eligibility criteria	8
Settings and locations where data were collected	8
Intervention	10
Incredible Babies <i>book</i>	10
<i>Incredible Years infant programme</i>	10
<i>Incredible Years toddler programme</i>	10
Universal proportionate delivery of the intervention	10
Setting for intervention delivery	11
Control treatment	11
Outcomes and measures	11
<i>Child primary outcome</i>	12
<i>Key secondary parent outcome</i>	12
<i>Other secondary outcomes for primary and co-parents</i>	12
<i>Child secondary outcomes</i>	12
<i>Parent-child dyad secondary outcome measures</i>	13
<i>Other measures</i>	13
<i>Changes to trial outcomes after the trial commenced</i>	13
Data collection	13
Sample size	13
Additional information about the ASQ:SE-2	14
Randomisation	14
<i>Blinding</i>	15

## CONTENTS

Statistical methods	15
<i>Treatment effectiveness</i>	15
<i>Inclusion of covariates</i>	15
<i>Missing data</i>	15
<i>Analysis populations</i>	16
<i>Additional analyses</i>	16
Ethics arrangements	16
Changes to protocol	16
<i>Version 5: 23 February 2017</i>	17
<i>Version 6: 8 May 2017</i>	17
<i>Version 7: 16 May 2017</i>	17
<i>Version 8: 7 November 2017</i>	17
<i>Version 9: 8 November 2018</i>	17
<i>Version 10: 17 June 2019</i>	17
Patient and public involvement	18
<i>Pre-funding preparation</i>	18
<i>Post-funding preparatory work</i>	18
<i>Throughout the pilot and main trial</i>	18
<i>Dissemination</i>	19
<i>Reflections on what worked well and not so well</i>	19
<b>Chapter 3 Outcome evaluation results</b>	<b>21</b>
Main outcome measures	21
Data completeness	21
Baseline covariates	21
Primary analysis of primary outcome: ASQ:SE-2	25
Sensitivity analysis of the primary outcome	26
Exploratory analysis of the primary outcome	29
Key secondary outcome: depression (PHQ-9) score of primary caregiver	30
Other secondary outcomes	32
Summary of intervention take-up	34
Characteristics of compliers and non-compliers	35
Overall summary of outcomes	41
<b>Chapter 4 Process evaluation</b>	<b>43</b>
Introduction	43
Method	43
<i>Sampling</i>	43
<i>Data collection</i>	44
<i>Quantitative measures</i>	45
<i>Qualitative methods</i>	45
<i>Data analysis</i>	46
Findings	46
<i>Aim 1</i>	46
<i>Aim 2</i>	54
Conclusions	73
<b>Chapter 5 Economic evaluation</b>	<b>75</b>
Introduction	75
Methodology	75
<i>Resource use and costs</i>	75
<i>Outcomes</i>	80
<i>Analysis</i>	80

Results	81
<i>Resource use and costs</i>	81
<i>Health outcomes</i>	84
<i>Cost-consequence analysis</i>	88
<i>Cost-effectiveness analysis</i>	88
<i>Scenario analyses</i>	90
<i>Sensitivity analyses</i>	92
Discussion	94
<i>Principal findings</i>	94
<i>Cost-effectiveness and longer-term impacts</i>	94
<b>Chapter 6 Discussion</b>	<b>95</b>
Was the E-SEE Steps programme effective?	95
Considerations of why the E-SEE programme did not demonstrate effectiveness	97
<i>Services as usual: was the parenting programme offer and uptake 'better' in the control arm?</i>	97
<i>Low take-up of IY-I and IY-T</i>	97
<i>Programme uptake according to screener measures</i>	97
<i>Robust measures and assumptions?</i>	98
Can Incredible Years be delivered as a proportionate universal model, and what are the organisational, or systems-level, barriers to and facilitators of delivering in this way, with fidelity?	98
<i>Challenges for the specific intervention: the E-SEE Steps programme</i>	99
<i>Challenges inherent to the public services system within which the intervention is delivered</i>	100
<i>Co-parents</i>	100
Is Incredible Years, and the proposed delivery model, cost-effective in enhancing child social and emotional well-being at 20 months when compared with services as usual?	101
<i>Costs</i>	101
<i>Outcomes</i>	102
<i>Cost-effectiveness</i>	102
Limitations	103
Strengths	104
Future research	105
Conclusions	107
<i>Implications for practice or local service delivery</i>	107
<b>Acknowledgements</b>	<b>109</b>
<b>References</b>	<b>113</b>
<b>Appendix 1</b> Logic model	<b>123</b>
<b>Appendix 2</b> Normative data for outcome measures	<b>125</b>
<b>Appendix 3</b> Baseline differences in outcomes	<b>127</b>
<b>Appendix 4</b> Difference in ASQ:SE-2 scores between arms	<b>129</b>
<b>Appendix 5</b> Summary of ASQ:SE-2 missing data	<b>133</b>
<b>Appendix 6</b> Primary outcome by site	<b>135</b>

## CONTENTS

<b>Appendix 7</b> Indicative analysis of ASQ:SE-2	<b>137</b>
<b>Appendix 8</b> Analysis of Patient Health Questionnaire-9 items data	<b>139</b>
<b>Appendix 9</b> Transitions in Infant CARE-Index scores	<b>143</b>
<b>Appendix 10</b> Planned substudies	<b>145</b>

# List of tables

<b>TABLE 1</b> Pilot progression criteria and assessment and design implications for a definitive RCT	5
<b>TABLE 2</b> Thresholds for concern: ASQ:SE-2	14
<b>TABLE 3</b> Baseline demographic characteristics of child, primary caregiver and co-parent	21
<b>TABLE 4</b> Summary statistics and difference between arms for ASQ:SE-2	27
<b>TABLE 5</b> Summary of primary and sensitivity analysis models	28
<b>TABLE 6</b> Primary outcome subgroup analysis	29
<b>TABLE 7</b> Transition between ASQ:SE-2 categories	30
<b>TABLE 8</b> Difference between arms for the key secondary outcome: PHQ-9	31
<b>TABLE 9</b> Summary of effect estimates from all analysis models for key secondary outcome: PHQ-9	32
<b>TABLE 10</b> Transition between PHQ-9 categories	33
<b>TABLE 11</b> Difference between arms for the Infant CARE-Index	33
<b>TABLE 12</b> Difference between arms for the PSOC survey	34
<b>TABLE 13</b> Difference between arms for the EQ-5D-5L	35
<b>TABLE 14</b> Difference between arms for outcomes at 18 months (FU3) only	35
<b>TABLE 15</b> Comparing characteristics of compliers and non-compliers: IY-I and IY-T	38
<b>TABLE 16</b> Summary of questions and data sources for the process evaluation	44
<b>TABLE 17</b> Service design decisions	47
<b>TABLE 18</b> Distribution of parenting support programmes by age group across sites	48
<b>TABLE 19</b> Eligibility and take-up for IY-I and IY-T groups	49
<b>TABLE 20</b> Attendance at IY-I groups	50
<b>TABLE 21</b> Attendance at IY-T groups	51
<b>TABLE 22</b> Incredible Years infant programme and IY-T group adherence based on IY checklists	52
<b>TABLE 23</b> Agreed PPIC scores for IY-I and IY-T	53

## LIST OF TABLES

<b>TABLE 24</b> Use of child care during IY-I and IY-T sessions	<b>63</b>
<b>TABLE 25</b> Unit costs	<b>76</b>
<b>TABLE 26</b> Average E-SEE Steps programme costs by resource category (£ 2018/19)	<b>81</b>
<b>TABLE 27</b> Average IY-I and IY-T group session costs by resource category (£ 2018/19)	<b>82</b>
<b>TABLE 28</b> Adult- and child-related costs by resource category (£ 2018/19)	<b>83</b>
<b>TABLE 29</b> Adult- and child-related outcomes by treatment arm	<b>85</b>
<b>TABLE 30</b> Within-trial cost-consequence analysis	<b>88</b>
<b>TABLE 31</b> Within-trial cost-effectiveness analysis	<b>89</b>
<b>TABLE 32</b> Cost-effectiveness scenario analysis	<b>91</b>
<b>TABLE 33</b> Normative data for ASQ:SE-2	<b>125</b>
<b>TABLE 34</b> Normative data from women for their PHQ-9	<b>125</b>
<b>TABLE 35</b> Differences between arms in outcome measures at baseline	<b>127</b>
<b>TABLE 36</b> Difference in ASQ:SE-2 score between arms using square root transformation	<b>129</b>
<b>TABLE 37</b> Difference in ASQ:SE-2 score between arms using primary model assuming unstructured correlation	<b>129</b>
<b>TABLE 38</b> Missing ASQ:SE-2 data summary	<b>133</b>
<b>TABLE 39</b> Indicative analysis of ASQ:SE-2 for each stage using eligible subgroups	<b>137</b>
<b>TABLE 40</b> Transitions between Infant CARE-Index risk categories	<b>143</b>
<b>TABLE 41</b> Mother and child characteristics by outcome status	<b>148</b>
<b>TABLE 42</b> Summary of hospital admissions and ED attendances	<b>152</b>
<b>TABLE 43</b> Reasons for hospital admission (self-reported)	<b>152</b>
<b>TABLE 44</b> Significant predictors of outcomes related to the child	<b>153</b>
<b>TABLE 45</b> Significant predictors of outcomes related to the mother	<b>155</b>



# List of figures

<b>FIGURE 1</b> Flow of participants through the main trial	9
<b>FIGURE 2</b> Summary of overall ASQ:SE-2 scores by follow-up point	26
<b>FIGURE 3</b> Forest plot for primary analysis models on original scale	28
<b>FIGURE 4</b> Summary of PHQ-9 by treatment arm and follow-up point	31
<b>FIGURE 5</b> Eligibility, compliance and session attendance: IY-I	36
<b>FIGURE 6</b> Eligibility, compliance and session attendance: IY-T	37
<b>FIGURE 7</b> Comparison of PPIC and checklist adherence for IY-I and IY-T	54
<b>FIGURE 8</b> Participant responses to the EQ-5D-5L questionnaire by follow-up period and treatment arm	86
<b>FIGURE 9</b> Sensitivity of IY programme costs and measures of cost-effectiveness with respect to the number of future programmes delivered	93
<b>FIGURE 10</b> Sensitivity of IY programme costs and measures of cost-effectiveness with respect to the number of IY group session attendees	93
<b>FIGURE 11</b> Logic model for the IY-I and IY-T	123
<b>FIGURE 12</b> Residuals from the primary analysis model	129
<b>FIGURE 13</b> ASQ:SE-2 scores at each follow-up point	130
<b>FIGURE 14</b> Residuals from the primary analysis model of ASQ:SE-2 after square root transformation	131
<b>FIGURE 15</b> Residuals from the primary analysis model of ASQ:SE-2 after calculating difference from baseline	131
<b>FIGURE 16</b> Residuals against fitted values for the ASQ:SE-2 primary analysis model	131
<b>FIGURE 17</b> Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 3	135
<b>FIGURE 18</b> Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 2	135
<b>FIGURE 19</b> Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 4	135
<b>FIGURE 20</b> Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 1	136

## LIST OF FIGURES

<b>FIGURE 21</b> Histogram of PHQ-9 data	<b>139</b>
<b>FIGURE 22</b> Distribution of PHQ-9 difference from baseline	<b>140</b>
<b>FIGURE 23</b> Residuals against fitted values for PHQ-9	<b>141</b>
<b>FIGURE 24</b> Residuals from planned PHQ-9 analysis	<b>141</b>
<b>FIGURE 25</b> Residuals from unplanned analysis of PHQ-9 after calculating difference from baseline	<b>141</b>

# List of supplementary material

**Report Supplementary Material 1** Process evaluation

**Report Supplementary Material 2** Overview of missing data at questionnaire and item level

**Report Supplementary Material 3** Breastfeeding outcomes

**Report Supplementary Material 4** Service use as reported by participants split by trial arm

**Report Supplementary Material 5** Health economic evaluation

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/BCFV2964>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.



## List of abbreviations

ASQ:SE-2	Ages and Stages Questionnaire: Social and Emotional, 2nd edition	IY	Incredible Years
BFI	Baby Friendly Initiative	IY-B	<i>Incredible Babies</i> book
CHU-9D	Child Health Utility	IY-I	Incredible Years infant programme
CI	confidence interval	IY-T	Incredible Years toddler programme
CSRI	Client Service Receipt Inventory	LA	local authority
DMEC	Data Monitoring and Ethics Committee	MCID	minimum clinically important difference
DVD	digital versatile disc	MI	multiple imputation
ED	emergency department	MPAS	Maternal Postnatal Attachment Scale
ENRICH	EvaluatioN of wRaparound in Ireland for CHildren and families	NICE	National Institute for Health and Care Excellence
EQ-5D	EuroQol-5 Dimensions	OR	odds ratio
EQ-5D-3L	EuroQol-5 Dimensions, three-level version	PAC	Parent Advisory Committee
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	PedsQL	Pediatric Quality of Life Inventory
E-SEE	Enhancing Social-Emotional Health and Well-being in the Early Years	PHQ-9	Patient Health Questionnaire-9 items
FFCPin	Fitting Fathers and Co-parents In	PPAS	Paternal Postnatal Attachment Scale
FU1	follow-up 1 (2 months post baseline)	PPIC	parent programme implementation checklist
FU2	follow-up 2 (9 months post baseline)	PSOC	Parent Sense of Competence
FU3	follow-up 3 (18 months post baseline)	PSSRU	Personal Social Services Research Unit
GBP	Great British pound	QALY	quality-adjusted life-year
GP	general practitioner	RCT	randomised controlled trial
HENRY	Health Exercise and Nutrition for the Really Young	SAP	statistical analysis plan
HRQoL	health-related quality of life	SAU	services as usual
ICC	intracluster coefficient	SD	standard deviation
ICER	incremental cost-effectiveness ratio	SDQ	Strengths and Difficulties Questionnaire
INHB	incremental net health benefit	TMG	Trial Management Group
IRR	incidence rate ratio	Triple P	Positive Parenting Program
		TSC	Trial Steering Committee



# Plain English summary

## What was the question?

We wanted to help parents to support their child's mental health between the ages of 2 and 20 months. We know that group-based parenting programmes (in which parents meet to discuss issues and learn from each other and group leaders) can work for children aged  $\geq 3$  years; however, we do not know if they work well for younger children.

## What did we do?

We picked a programme to support young children called Incredible Years<sup>®</sup> and compared parents who tried it with those who did not. We asked all parents questions about their health and their child's health at four different time points over 18 months.

Parents who were offered the Incredible Years programme received an Incredible Years baby book to support them with their parenting. When babies were aged 4 months, some parents were offered the Incredible Years infant programme. At 12 months, some parents were offered the Incredible Years toddler programme. An offer to attend these programmes depended on how parents rated their own or their child's mental health.

When children were aged 20 months we compared data from parents who followed the Incredible Years programme with data from those who had not been offered the programme.

## What did we find?

We found that the Incredible Years programme did not improve child well-being or significantly improve parent mental health.

Services that delivered the groups, and parents who attended, enjoyed the experience and felt some benefits. However, staff struggled to find the time and capacity to deliver the programme because of high workloads and competing priorities. Not all parents who were offered the programme accepted or turned up.

## What does this mean?

Organisations and people who help parents with young children, such as children centre staff or health visitors, need additional support and time to engage parents and to deliver programmes like Incredible Years. The findings of this study suggest that the programme would not improve child well-being.





# Scientific summary

## Background

Behavioural and mental disorders have become a public health crisis. Early prevention is key to prevent mental health issues in childhood and to mitigate the personal, familial and societal costs of related later negative outcomes. Evidence-based group parenting programmes are effective for parents of children aged  $\geq 3$  years; however, there is a lack of evidence for programmes for parents of children aged  $\leq 2$  years. A proportionate universal approach has been advocated to reduce the overall gradient of health inequality by offering support/services commensurate with individual/family level of need.

Two parenting programmes that aim to enhance child well-being and development are the Incredible Years<sup>®</sup> infant programme (IY-I) and the Incredible Years toddler programme (IY-T). These programmes were delivered and evaluated in a proportionate universal intervention model called Enhancing Social-Emotional Health and Well-being in the Early Years (E-SEE) Steps.

## Objectives

- Does the E-SEE Steps model enhance child social and emotional well-being at 20 months of age when compared with services as usual (SAU)?
- Is the E-SEE Steps model cost-effective in enhancing child social and emotional well-being at 20 months when compared with SAU?
- Can the E-SEE Steps model be delivered as a proportionate universal model, and what are the organisational or systems-level barriers to and facilitators of delivering in this way, with fidelity?

## Methods

### Design

The design was informed by a large randomised pilot study, which involved two research sites, with over 200 families, and parent advisory committees. The trial was a pragmatic two-arm randomised controlled trial and economic appraisal, with an embedded process evaluation to examine the outcomes, implementation and cost-effectiveness of the intervention, and intervention uptake, plus three additional substudies.

### Setting

The intervention was delivered in community settings by early years children's services and/or public health staff in four local authorities. Research sites had adequate birth rates to support the trial, were not offering the intervention as SAU and were willing to train staff to deliver the intervention.

### Blinding

Data collectors, referrers, the chief investigator, the statistician (until final analysis), the Trial Steering Committee and the Trial Management Group were blind to participant allocation. Participants, Incredible Years leaders and some study team members, such as trial co-ordinators, were not blind.

### Sample size calculation

Sample size was calculated on the child primary outcome of social and emotional well-being using the Ages and Stages Questionnaire: Social and Emotional, 2nd edition (ASQ:SE-2). We defined the clinically important difference at follow-up 3 (18 months post baseline) (FU3) to be 5 units on the ASQ:SE-2 for

the intervention arm when compared with SAU. Assuming a standard deviation of 18 units on the ASQ:SE-2 at FU3, the correlation between baseline and FU3 scores is 0.26, and between pairs of measurements after baseline is 0.40. For a design effect of 1.25 for the intervention arm, two-sided 5% significance level and 90% power, we would require the study to have retained 441 intervention participants and 92 control participants. Allowing for 12% overall attrition, the target was 606 randomised parents, with an allocation ratio of 5 : 1 to ensure that sufficient numbers of eligible parents were able to attend the group parenting programmes.

### **Participants**

#### **Inclusion criteria**

Parents were eligible for inclusion if they consented to participate, had a child aged  $\leq 8$  weeks, were willing to be randomised and, if allocated to intervention, were able to receive the Incredible Years services offered.

#### **Exclusion criteria**

Parents whose child had obvious, or diagnosed, organic developmental difficulties or who were enrolled in another group parenting programme at sign-up were not eligible to participate.

#### **Recruitment**

Health visitors and family support workers invited families to hear more about the study. Those parents who consented were contacted by the research team. Researchers recruited parents and obtained informed consent during a home visit. Parents could also self-refer to the study and co-parents could participate in the study if the 'primary' parent invited them. Families received shopping vouchers of modest value, increasing at each data collection point.

#### **Randomisation and allocation**

Randomisation occurred following baseline data collection, using a web-based randomisation system. Parents were randomly allocated to the intervention or control arm in a 5 : 1 ratio stratified according to level of need at baseline based on the parent Patient Health Questionnaire-9 items (PHQ-9) score or child ASQ:SE-2 score, sex of child and parent, and research site.

#### **Intervention**

The E-SEE Steps model comprised a proportionate universal intervention model with three levels [one universal level (i.e. *The Incredible Years* baby book) and two targeted levels (i.e. IY-I and IY-T, which were 10 and 12 weeks long, respectively, with one 2-hour group session per week)]. Parents were offered the groups if they rated themselves as at least mildly depressed on the PHQ-9 or if they rated their child in the monitoring zone or above on the ASQ:SE-2. Within the E-SEE Steps model, four intervention 'doses' were possible for each family: (1) the book only, (2) the book plus IY-I, (3) the book plus IY-I plus IY-T and (4) the book plus IY-T.

#### **Main outcome measures**

Data were collected in the home by a researcher at four time points [baseline, follow-up 1 (2 months post baseline), follow-up 2 (9 months post baseline) and FU3].

#### **Primary outcome**

Child social and emotional well-being assessed by the ASQ:SE-2.

#### **Secondary key outcomes**

Parent depression assessed by the PHQ-9.

### **Child secondary outcomes**

Parent-child interaction was independently observed and was assessed using the Infant CARE-Index. Cognitive development and health (quality of life) was assessed using the Pediatric Quality of Life Inventory (PedsQL™) at final follow-up only. Child behaviour was assessed by the Strengths and Difficulties Questionnaire 2-4 version at final follow-up only.

### **Parent secondary outcomes**

Maternal/paternal-child attachment/interaction was assessed by the Maternal Postnatal Attachment Scale and Paternal Postnatal Attachment Scale at final follow-up only. Parenting skill was assessed using the Parent Sense of Competence questionnaire. Health (quality of life) was assessed using the EuroQol-5 Dimensions, five-level version (EQ-5D-5L).

### **Other outcomes**

Demographic information was captured via a bespoke structured interview form and included age, ethnicity, religion, income, marital status, parent/co-parent education, housing, family composition, infant feeding and prematurity. Further economic evaluation outcomes to examine resource use and costs based on access to health, social and educational services by parents and children were reported by parents using a modified Client Service Receipt Inventory. Costs of intervention delivery were gathered via implementation staff and existing data sources. The process evaluation included group attendance, leader contact rates, leader self-rated adherence checklists, a researcher-rated parenting programme implementation checklist and (modified) Incredible Years parent satisfaction questionnaires.

### **Analysis**

Primary analysis applied intention to treat.

The marginal model for overall effectiveness was fitted using general estimating equations with a Gaussian family, identity link and autoregressive covariance structure of order 1 AR(1). AR(1) means that each observation in the time series is directly related to the observation that preceded it. Model estimates with standard errors that are robust to the non-normality and non-independence of observations were computed. Statistical analyses used Stata®/MP 16.0 (StataCorp LP, College Station, TX, USA). Item non-response was imputed using questionnaire developer rules. Robustness to outcome and baseline non-response was explored using multiple imputation (MI) methods. Covariates included in the model were baseline PHQ-9 and ASQ:SE-2 scores, whether or not the parent had a degree, whether or not the parent was in a relationship, ethnicity, child's sex, follow-up time and delivery site. We explored the extent to which intervention effectiveness differed between subpopulations by testing the significance of the interaction between randomised treatment group and subgroup. The impact of individual E-SEE Step components was investigated using non-randomised observational analysis where participants in the control arm with outcome scores above the eligibility threshold were used as a pseudo-control group. To assess the robustness of the outcome analysis, the primary analysis was repeated with alternative specifications of the primary outcome measure using MI. Per-protocol and complier-average causal effect analyses were not conducted, as there is no satisfactory way of defining compliers without biasing the estimated impact of IY-I and IY-T on compliers because of the conditional design. Descriptive analysis of the characteristics associated with compliance was undertaken.

### **Process evaluation/treatment processes: method and analysis**

A multimethod approach assessed fidelity of delivery and explored parents', leaders' and service managers' experiences of the E-SEE Steps model, as well as the organisational, team and individual factors that facilitate or hinder its implementation. Quantitative monitoring data were collected for all parent groups. Pre- and post-training questionnaires assessed leaders' qualifications, existing experience of parenting groups and working with families, perceived competence to deliver the programme, perceived organisational support, and experiences of delivering Incredible Years. All quantitative data were reported descriptively. Qualitative data were gathered during focus groups and semistructured interviews with key stakeholders. Thematic analysis was applied to the qualitative data.

### **Economic evaluation: method and analysis**

Cost-effectiveness and cost-consequence analyses were conducted. Costs in both trial arms were estimated from alternative perspectives, including a NHS and Personal Social Services perspective, a wider public sector perspective and a societal perspective, which includes costs to participants. A micro-costing of Incredible Years group delivery established delivery costs.

Initial analysis presents incremental results for the primary/key outcome measures for children (i.e. ASQ:SE-2) and adults (i.e. PHQ-9) separately. These results were compared with the incremental costs measured from the alternative perspectives. Secondary outcomes in terms of quality-adjusted life-years (QALYs) (using the PedsQL for children and EQ-5D-5L for adults) were also considered.

Alternative methods for combining primary and secondary outcomes across children and adults and across outcomes were explored for a full assessment of the benefits, and compared with costs. Links between trial outcome measures and longer-term outcomes were explored. Probabilistic sensitivity analyses were conducted to reflect the uncertainty around the adoption decision. Sensitivity analyses determined the robustness of the results to altering certain assumptions.

## **Results**

The target sample size was not reached. A total of 341 parents were randomised, with a retention rate of 94%. There were no baseline differences between arms. The E-SEE Steps model was not effective in enhancing child social and emotional well-being compared with SAU. No significant differences were found between the E-SEE Steps and control arms. All secondary outcomes, including the key parent outcome (i.e. depression) also did not show any difference between arms. Sensitivity analyses confirmed the findings.

Overall, parent take-up of the targeted parenting programmes was low, which could have had an impact on results. Sites, although enthusiastic, identified barriers to delivering the intervention, such as lack of capacity, resource and time. Adaptations to the intervention were also suggested (e.g. to ensure compatibility with UK guidance and context-specific video content).

The E-SEE Steps model had higher costs and more QALYs [0.031 QALY gain, 95% confidence interval (CI) -0.008 to 0.071 QALY gain] than SAU, resulting in an incremental cost-effectiveness ratio of approximately £20,062 per QALY compared with SAU. The mean incremental public sector cost of the E-SEE Steps model was £621 per family compared with SAU (95% CI -£103 to £1288). There was a small gain in mean QALYs, with positive increments in adults exceeding minor decrements reported in child outcomes over the trial period. Findings were sensitive to changes in the key assumptions used in the analysis. All scenarios found the E-SEE Steps model cost-effective at the maximum recommended threshold of £30,000 per QALY.

The trial also found that child emergency department (ED) attendance was predicted by younger gestational age, older age at recruitment to trial, mothers' poorer mental health and younger age, mother attending ED and study site. For mothers, ED attendance was predicted by mixed ethnic origin, having a boy, having poorer quality of life at baseline and having been hospitalised during the trial. Mothers' admission to hospital was predicted by ED attendance for themselves and being anxious or depressed.

## **Limitations**

The study was not powered to establish the effectiveness of each of the intervention's three individual levels, only the effectiveness of the overall E-SEE Steps model. Planned secondary analysis to explore each level of intervention could not be conducted because of low attendance rates in the group programme (and the very small number of parents in the control arm). *The Incredible Years* baby book will be explored further by combining pilot and main trial data; however, it will not be possible to combine data for the group levels because of design changes made within the pilot.

Owing to low co-parent numbers ( $n = 68$ ), we could not provide insights into the role of co-parents in shaping children's social and emotional development, but our co-parent substudy explored further co-parent perceptions of parenting programmes and engagement into such programmes.

We had to resort to using parent-reported attendance at EDs and hospital admissions, as Hospital Episode Statistics data were unobtainable within the study time frame because of various barriers to, and during, the NHS Digital application process.

We are unsure how representative our sample is of the population in each site, or nationally. We have documented the challenges of exploring trial sample 'representativeness'; however, our main trial sample is predominantly well educated, with only a small proportion (11%) of parents identifying themselves as single and not in a live-in relationship. Therefore, it is unlikely that our sample is representative of families that are experiencing the greatest threats to their social and emotional well-being.

## Conclusions

- The E-SEE Steps proportionate universal delivery model did not enhance child social and emotional well-being.
- The E-SEE Steps model can be implemented in community settings and delivered by health and/or family and children services; however, system changes need to occur, and resources and capacity increased, to embed any such model successfully. Intervention adaptations are also required.
- Although the E-SEE Steps model demonstrates higher costs over the trial duration, this may not be the case longer term. Programme delivery costs will reduce over time (e.g. training costs may no longer be needed in the future).
- More work is needed to ensure better engagement of parents. We suggest that a pre-intervention component or a set of implementation strategies devoted to identification and engagement (and retention) of parents and also co-parents be developed and adhered to consistently, while being fully resourced.
- Maternal mental health predicted child ED attendance and mothers' own hospital admissions. This finding highlights the importance of tackling/preventing maternal depression or anxiety during pregnancy and within the perinatal period through enhanced evidence-based service provision.

## Trial registration

This trial is registered as ISRCTN11079129.

## Funding

This project was funded by the National Institute for Health and Care Research (NIHR) Public Health Research programme and will be published in full in *Public Health Research*; Vol. 10, No. 8. See the NIHR Journals Library website for further project information.



# Chapter 1 Introduction

## Social and emotional well-being

Social and emotional well-being in childhood refers to the ability of a child to express and manage their emotions in socially and culturally appropriate ways and to develop positive relationships with their parents, other adult caregivers and peers.<sup>1,2</sup> The development of social and emotional well-being requires many foundational competencies, skills and characteristics, including, for example, inhibitory control, emotion regulation, self-confidence, perspective-taking and attachment.<sup>1,3</sup> Throughout infancy and toddlerhood, the typically developing child experiences rapid cognitive and maturational changes that support the accomplishment of increasingly complex developmental tasks in social and emotional domains of development.<sup>2</sup> By age 20 months, infants can respond to their own name, recognise themselves, display interest in other people (e.g. adults and peers) and are able to engage in co-ordinated interaction with them. In addition, infants at this age can form attachments with their caregivers and develop internal working models of relationships, are capable of expressing basic emotions and can also regulate their emotions through self-soothing with support from adults.<sup>3</sup> These social and emotional competencies form the foundations of healthy development and school readiness.

There are strong continuities between infant social and emotional well-being and later life outcomes. Impaired social and emotional well-being in the early years increases the risks of negative outcomes throughout childhood and into adulthood. For example, longitudinal studies document increased risks of poor mental health, as well as increased risks of antisocial behaviour and criminality, and poor educational and employment outcomes.<sup>4-6</sup> Conversely, positive social and emotional well-being in the early years is associated with good health and development outcomes throughout the life course and provides a basis for adaptive resilience to future adversities.<sup>4,7,8</sup> Evidence suggests that promoting social and emotional well-being in the early years is more effective, and less costly, than interventions delivered later in life once difficulties are more entrenched.<sup>9</sup>

Parent-child relationships play a critical role in the development of social and emotional well-being. Parenting practices, styles, skills and knowledge, as well as parental mental health, can have an impact on the quality of parent-child relationships and attachment bonds. Unresponsive parenting (e.g. when a parent is under stress or experiencing depression) can lead to ineffective parenting strategies and (inadvertent) emotional neglect.<sup>2</sup> Although the majority of research has been conducted with mothers, there is a growing field also documenting the influential role of fathers and other co-parents, such as grandparents, in the development of children's social and emotional well-being.<sup>10,11</sup>

## Parenting interventions to promote infant social and emotional well-being

Parenting programmes are effective in reducing behavioural, social and emotional difficulties in school-aged children,<sup>12,13</sup> as well as improving parent psychosocial health.<sup>14</sup> However, as highlighted by the National Institute for Health and Care Excellence (NICE) in 2012, there is a gap in our understanding of the extent to which parenting and family-focused interventions are effective in enhancing social and emotional well-being specifically in infants.<sup>15</sup> An earlier systematic review found some preliminary support for the use of group-based parenting programmes to improve the emotional and behavioural adjustment of children aged < 3 years.<sup>16</sup> However, there was insufficient evidence to support any firm conclusions and further research on parenting programmes for younger age groups was recommended.<sup>16</sup> Although there is significant policy interest and increasing research in the area of early intervention and prevention [with the subsequent establishment of the Early Intervention Foundation (London, UK) and the 1001 Critical Days Manifesto movement], the evidence gap identified by NICE still exists.



## The case of proportionate universal interventions

There has been a particular call for research on interventions provided to families within a framework of 'proportionate universalism'. Proportionate universalism was first proposed in the Marmot review (2010)<sup>17</sup> and refers to services designed to be delivered universally, but with a scale and intensity that is proportionate to the level of need or disadvantage. Proportionate universalism is a response to an overwhelming body of evidence highlighting social gradients of health, whereby people at the top have less threat to health than the deprived people at the lower end of the gradient.<sup>29</sup> A proportionate universalism approach argues that services will not reduce health inequalities if they are focused solely on the most disadvantaged in society; however, if we solely concentrate on the most disadvantaged then the health gradient will not decrease and will only tackle a small part of the problem.<sup>17</sup>

The Marmot review's call to action included strategic objectives concerning the provision of proportionate universal parenting services.<sup>17</sup> Although there are numerous examples of parenting programmes for young children in the UK (see the Early Intervention Foundation's *Guidebook*<sup>18</sup>), many such programmes are targeted at high-risk families or children who are already experiencing difficulties. Targeted approaches, although potentially effective, have been criticised for causing stigma towards parents experiencing disadvantage and, in turn, affecting the uptake of services. In addition, it has also been argued that targeted approaches often do not identify all children and families in need of additional support.<sup>8</sup> Proportionate universal approaches that embed interventions within existing universal services may improve uptake and interaction with, and access to, care.<sup>8,16</sup>

## Lack of proportionate universal interventions and evidence base

Despite clear potential, there is a large gap in the evidence base for proportionate universal interventions. A recent systematic review of universal parenting interventions for 0- to 2-year-olds reported a lack of evidence for effectiveness of interventions in the postpartum to 24-month period.<sup>8</sup> Although the quality of evidence was low, previous controlled studies reported no differences between intervention families and controls. The authors<sup>8</sup> speculate that one reason for the lack of intervention effects may be a lack of behaviour change content in the programme theory for included interventions, despite their reliance on parental behaviour change as a mechanism for improving social and emotional well-being in infants. Hurt *et al.*<sup>8</sup> concluded that 'there is an urgent need for robust evaluation of existing interventions, and to develop and evaluate novel interventions to enhance the offer to all families'.

## The Incredible Years® parenting programmes

The Incredible Years® (IY) parenting programmes [URL: [www.incredibleyears.com](http://www.incredibleyears.com) (accessed 6 January 2022)] are parent interventions that are underpinned by social learning theory and are designed to enhance the social and emotional well-being of children (aged 0–12 years). The programmes encourage rewards for behaviours in children parents want to see more of, while ignoring behaviours parents want to see less of. The IY programmes are manualised and delivered by trained leaders to groups of 10–12 parents for 2 hours a week for 10–14 weeks. The Incredible Years infant programme (IY-I) and Incredible Years toddler programme (IY-T) versions for 0- to 1-year-olds and 1- to 3-year-olds, respectively, build on decades of research evidence that demonstrate the effectiveness of the IY programmes for parents of children aged  $\geq 3$  years. Previously, the older-age IY programmes have demonstrated effectiveness, cost-effectiveness and transportability in independent trials across several countries/contexts.<sup>12</sup> Meta-analyses suggest that IY may be beneficial to younger children and their parents,<sup>19</sup> and that families with severe depression and severe conduct problems demonstrate co-occurring changes for the better.<sup>20</sup> An analysis of data pooled from several IY trials suggested a large moderating effect for depression and that IY was more beneficial for children where parents were more depressed.<sup>21</sup>



Evidence for the IY-I is slowly growing.<sup>22</sup> A small ( $n = 80$ ) randomised non-targeted study in Wales, UK, showed that control mothers were significantly less sensitive during play with their baby at a 6-month follow-up.<sup>23</sup> In addition, a small trial in Denmark, delivered universally, found differential outcomes for the lowest and highest functioning families, suggesting that IY-I should be targeted (as originally designed).<sup>24</sup> Results from 12 IY-I groups ( $n = 79$  group participants) showed parental benefits of improved mental health and parenting confidence post course (pre-post only, no comparator) and influenced a rural county in East Wales to scale delivery of IY-I.<sup>25</sup>

Two IY-T trials, one in the UK<sup>26</sup> and one in the USA,<sup>27</sup> were inconclusive. The UK trial<sup>26</sup> was a small community-based trial in Wales. The trial<sup>26</sup> relied on geographical targeting to disadvantaged 'Flying Start' areas and did not always reach families that needed most support. The US trial<sup>27</sup> delivered IY-T through primary care (i.e. paediatric practices), rather than community settings.

Both the IY-I and IY-T require further evaluation to establish their clinical effectiveness and cost-effectiveness and to inform the evidence base. Neither IY or similar parenting programmes [e.g. the Positive Parenting Program (Triple P)<sup>28</sup>] have, to date, been delivered or tested using a targeted proportionate universal approach delivered in community settings.

## Introduction to the E-SEE Steps model and its components

The Enhancing Social-Emotional Health and Well-being in the Early Years (E-SEE) Steps model is a unique multilayer intervention delivery model that combines three elements of existing IY programmes in a proportionate universal approach. The universal dose is the *Incredible Babies* book (IY-B), with two targeted group-based programmes (i.e. IY-I and IY-T) providing support at a greater intensity for those with greater need at different points in their child's development.

The IY age-appropriate programmes lend themselves well to a proportionate universal delivery model, as they can be delivered in cumulative 'doses' according to need.<sup>29</sup> The E-SEE Steps model has the potential to provide robust evidence and inform NICE guidance. Recent research<sup>30</sup> on child outcomes using longitudinal data has demonstrated the usefulness of using 'stacked' early interventions across the early years of children's lives to maximise impacts on child outcomes. In addition, a recent systematic review<sup>31</sup> suggested that a proportionate universal approach, although underutilised, is useful for mental health interventions and, given previous evidence for older age IY programmes, this intervention may be useful when delivered in this way.

## Explanation of rationale

Although there is significant policy interest, there is a lack of robust UK evidence for promoting social and emotional well-being and for programmes specifically designed to prevent later mental health issues developing in children aged  $\leq 2$  years. The early years are a critical period of development for children, during which empathic and responsive parenting promotes positive outcomes. However, the majority of parenting programmes are designed for older children for whom social, emotional and behavioural difficulties are identifiable. Recent UK policy and guidance has placed emphasis on a whole family approach (i.e. including fathers and grandparents in an integrated proportionate approach).

The proposed study will evaluate a preventative approach for parents of very young children at a time when the child may show no obvious signs of mental health or behavioural difficulties (or at least these signs are difficult to detect), although risk factors, such as parent or co-parent depression, may be present. The IY basic programme (for parents of children aged  $\geq 3$  years) has demonstrated substantial post-intervention improvements in a variety of parent and child outcomes and has a robust evidence base in the UK.<sup>32</sup> However, although IY-I and IY-T have been developed with the same successful

format and support infrastructure as the basic programme, they have not yet been evaluated in a proportionate delivery model within a community-based trial.

### Specific objectives

The E-SEE study comprised two phases: (1) a pilot trial and (2) a definitive randomised controlled trial (RCT). The pilot phase informed the main trial design and trial procedures. The main trial was designed to (1) establish the effectiveness of the E-SEE Steps model on clinical outcomes, (2) evaluate the processes around service delivery and (3) assess cost-effectiveness.

The primary research questions were:

- Does the E-SEE Steps model enhance child social and emotional well-being at 20 months of age when compared with services as usual (SAU)?
- Can IY be delivered as a proportionate universal model, and what are the organisational, or systems-level, barriers to and facilitators of delivering in this way, with fidelity?
- Is IY, and the proposed delivery model, cost-effective in enhancing child social and emotional well-being at 20 months when compared with SAU?

Alongside the outcome, process and economic studies designed to answer these questions (see *Chapters 3–5*, respectively), a series of substudies were planned (see *Appendix 10*), including (1) a study exploring the experiences of co-parents, (2) a study exploring hospital visits and admissions and (3) a study exploring comparisons with a similar trial conducted in Ireland.

## Chapter 2 Methods

Parts of this chapter have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This report is concordant with the Consolidated Standards of Reporting Trials (CONSORT) statement.<sup>34</sup>

### Pilot phase

The study was funded as a two-phase RCT, comprising an internal 18-month pilot conducted in two local authority (LA) areas and followed by a 30-month pragmatic two-arm RCT conducted in four LAs.<sup>35</sup> The study met stop/go criteria to progress to a main trial phase; however, the pilot identified design changes needed to ensure viability [e.g. changes to programme materials, the addition of another screener and more flexibility for the delivery of the model by one organisation (i.e. health or LA) as opposed to the originally required two organisations (i.e. health and LA)].

Table 1 summarises the progression criteria, associated measures and assessment of the pilot, and consequent changes to study design and/or implementation.

A recalculation of sample size for the main trial was also undertaken using data collected during the pilot phase.

Although the overarching design remained the same, the pilot was reclassified as an external pilot. A full account of the pilot phase of the study and how the learning informed the main trial design is provided in Blower *et al.*<sup>36</sup>

TABLE 1 Pilot progression criteria and assessment and design implications for a definitive RCT

Progression criterion	End point/measure	Assessment of criterion	Implications for design of definitive RCT
Recruitment: recruit and randomise at least 144 participants at each site ( $n = 288$ ) to achieve 192 participants at final follow-up	Sample size (recruited and randomised) achieved at baseline	Not achieved: $n = 205$ , representing 71% of the target baseline sample of 288. Although recruitment was below target, high rates of retention mitigated this (at FU3 $n = 181$ , i.e. 94% of original FU3 target of $n = 192$ )	Develop site recruitment targets/expectations and document in site-level agreements  Create awareness-raising materials to promote the study among referring practitioners  Utilise QuinteT approach  Recruitment challenges suggest an alternative design, consisting of a single research question 'Do the scores of children in the IY arm, on average, stay below those scores for children in SAU over the three follow-up measures?', which reduces the sample size required in a definitive trial and increases statistical power

continued

TABLE 1 Pilot progression criteria and assessment and design implications for a definitive RCT (continued)

Progression criterion	End point/measure	Assessment of criterion	Implications for design of definitive RCT
Retention: maximum 12% loss at each data collection (follow-up) time point (equivalent to 32% overall loss)	Sample size at the three follow-up data collection time points (i.e. FU1, FU2 and FU3)	Achieved: 181 participants were retained by the end of FU3 (12% overall loss). There was an average of 4% loss at each data collection point. This was much lower than the 12% loss at each time point and the 32% overall loss that was anticipated	<p>Replicate piloted strategies for retaining parents</p> <p>Change of address procedures</p> <p>Branded tokens with study contact details on</p> <p>PAC input to data collector training</p> <p>Consistent data collector across time points</p> <p>Appointment by letter procedures</p> <p>Re-attempt contact with those lost to follow-up at next time point</p> <p>Consider SWAT to explore differential effectiveness</p>
Intervention delivery: books received by parents and ability for local sites to successfully deliver the required number of groups (to include identification and training of group leaders and suitable venues for groups)	<p>Book receipt monitored via track and trace postage</p> <p>Programme monitoring data collated by the study team on IY training attendance, group venue and number of groups delivered in each site</p>	<p>Partially achieved: all parents in the intervention arm received an IY-B. Sufficient practitioners identified, trained in IY-I and IY-T and attended supervision in each site</p> <p>Suitable venues used (all groups delivered in children's centres)</p> <p>Viable group size was impacted by initial identification and recruitment issues and staff capacity to run groups at times convenient for parents and, therefore, was not achieved in all groups</p>	<p>Replicate postage and delivery procedures</p> <p>Ensure minimum of 10 practitioners trained in each programme in each site</p> <p>Allow flexibility in group delivery (could be health visitors only or LA only or combination)</p>
IY-I planned: four groups per site	Parent contact logs completed by group leaders		Ensure that referring practitioners offer an application to the study to all parents of children aged $\leq 8$ weeks
IY-T planned: two groups per site			
Intervention groups consist of a viable number (minimum eight parents invited, with five parents in attendance > 50% of sessions)		<p>IY-I: two groups ran in site 1 (<math>n = 6</math>). Two groups started in site 2, which were combined to one group (<math>n = 3</math>)</p> <p>IY-T: two groups ran in site 1 (<math>n = 2</math> in both groups). Two groups ran in site 2 (<math>n = 3</math> and <math>n = 7</math>)</p> <p>Spaces offered to eight or more eligible participants in most groups</p> <p>IY-I site 1: <math>n = 15</math> eligible, <math>n = 6</math> accepted and attended</p>	<p>Supplement the PHQ-9 (scores <math>\geq 5</math>) with ASQ:SE-2 scores (in the monitoring or refer zones) when determining eligibility for targeted IY-I and IY-T</p> <p>Emphasise need for crèche, transport, flexible timing and wrap-around support in service design processes for IY-I and IY-T</p> <p>Utilise matched non-research participants to supplement groups where needed to ensure minimum group sizes</p>

TABLE 1 Pilot progression criteria and assessment and design implications for a definitive RCT (continued)

Progression criterion	End point/measure	Assessment of criterion	Implications for design of definitive RCT
		IY-I site 2: $n = 12$ eligible, $n = 4$ accepted and attended	Revise sample calculation from 3 : 1 (intervention to control) to 5 : 1 to facilitate the identification of eligible parents and viable group size
		IY-T site 1: $n = 24$ eligible, $n = 8$ accepted, $n = 4$ attended	
		IY-T site 2: $n = 31$ eligible, $n = 17$ accepted, $n = 10$ attended	
Intervention acceptability: IY retention levels to reach 70% at IY-I and IY-T end	Parent contact logs completed by group leaders	Achieved: achieved when viewed as the percentage of parents retained when parents attended at least one session. Retention lower when calculated as percentage of anyone initially accepting a place (this included 'no shows')	Efforts should be focused on increasing uptake of groups among eligible parents (see above)
Parent satisfaction (supplements data on retention levels)	Satisfaction forms completed weekly after each group session and at the end of the programme	IY-I: average weekly attendance as a percentage total of those parents attending at least one session = 73%	Some suggestions from parents included increasing group sizes, ensuring locations are accessible/ providing transport and timing the groups so that they start when children are younger
		IY-T: average weekly attendance as a percentage total of those parents attending at least one session = 87%	Additional feedback from parents in relation to increasing the length of the programme and each session need further exploration
		Parent satisfaction forms completed by parents indicate a positive and high level of parent satisfaction	
Intervention fidelity: adherence and quality of delivery assessment of 80% in each LA across co-leaders	Group leader weekly self-report checklist (adherence) and PPIC completed independently by the study team (quality of delivery)	Achieved: overall adherence to key components of IY-I and IY-T ranged from 80% to 96% according to self-report checklists completed by group leaders	Retain independent measure of intervention fidelity alongside standard self-report checklists
		The PPIC tool shows acceptable threshold of 80% was exceeded for quality of delivery for IY-I and IY-T. Overall PPIC scores for fidelity were good for IY-I (73%) and IY-T (78%)	Provide accredited IY training and fortnightly supervision to group leaders and promote delivery 'dry runs' through service design processes
ASQ:SE-2, Ages and Stages Questionnaire: Social and Emotional, 2nd edition; FU1, follow-up 1 (2 months post baseline); FU2, follow-up 2 (9 months post baseline); FU3, follow-up 3 (18 months post baseline); PAC, Parent Advisory Committee; PHQ-9, Patient Health Questionnaire-9 items; PPIC, parent programme implementation checklist; QuinteT, Qualitative Research Integrated within Trials; SWAT, study within a trial.			

## Main trial design

The trial was designed as a pragmatic two-arm RCT and economic appraisal, with an embedded process evaluation to examine the outcomes, implementation and cost-effectiveness of the intervention, and intervention uptake.

## Participants and eligibility criteria

Eligible trial participants were parents (or primary caregivers) with the main parenting responsibility for the index child (aged  $\leq 8$  weeks) and their co-parent (if appropriate).

Primary caregiver is used here as an umbrella term to describe any person (e.g. biological parents, step-parents, foster parents, grandparents or legal guardians) who has the primary parental responsibilities of a child. English law states that if the child lives with their mother, then the mother is recognised as the primary carer.

Co-parent is a term used to describe any individual who may or may not be the 'biological' parent of the child, but who is involved in the upbringing of the child alongside the child's primary caregiver (i.e. the father, or a parent by partnership or marriage to one of the child's biological parents, such as stepmother). One co-parent could be recruited for each recruited primary carer.

Parents with a child aged  $\leq 8$  weeks were approached by health visitors or family and child service staff/services to see if they wished to hear more about the E-SEE study. For those who were interested, a form was completed, giving permission for the research team to receive their contact details, assess their eligibility and to contact the parent to make an appointment to visit them in their home (or an alternative venue to suit the parent) to discuss the trial further. Parents then decided if they wished to participate in the trial or not. Recruitment, therefore, was conducted within the home by the research team, with written informed consent being given by parents and co-parents wishing to participate. Parents could also self-refer if they had heard about the study via other channels (e.g. community groups and forums).

Non-eligible parents were provided with information about how to access local children centres and health service provision. *Figure 1* illustrates the flow of participants throughout the trial.

Parents were eligible if they:

- had the main parental responsibility for a child aged  $\leq 8$  weeks at initial engagement
- were willing to participate in the research
- were willing to be randomised and, if allocated to intervention, were able to receive IV services offered
- were not enrolled in another group parenting programme at consent stage
- were fully competent to give consent.

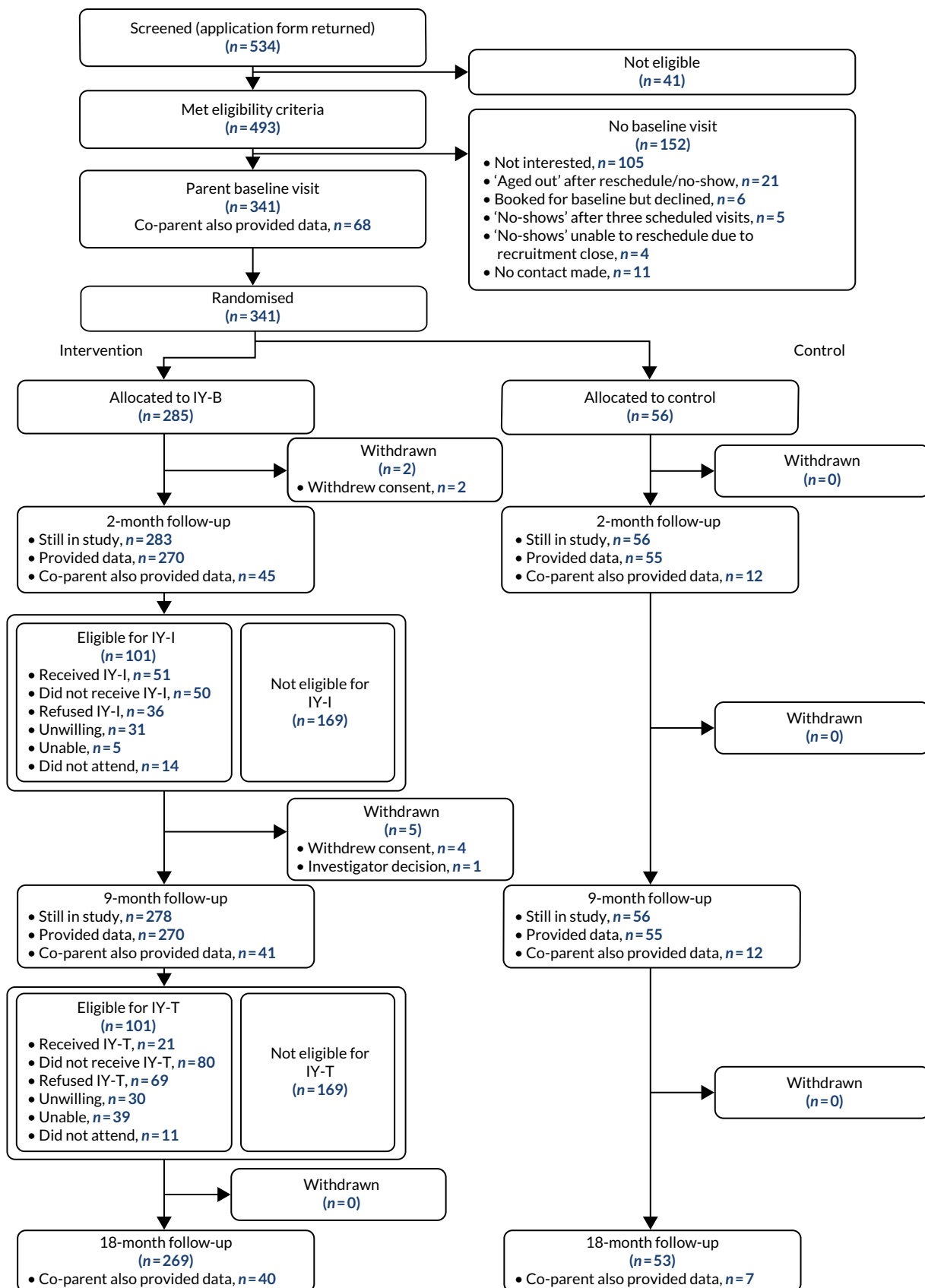
Co-parents were eligible if the primary carer agreed to their involvement and if they lived with the index child or looked after them for at least three evenings each week.

The exclusion criteria were the opposite of the above and, in addition, parents were excluded if:

- the child had obvious organic or developmental difficulties or had been diagnosed with the same.

## Settings and locations where data were collected

Expressions of interest, via completed proformas, were received from 16 potential sites in England. The proformas and initial discussions with sites were used to establish (1) levels of local deprivation, (2) sufficient live birth rates per year to allow recruitment of eligible and interested families, and to achieve an adequate randomisation sample with viable numbers for group delivery, and (3) willingness to support staff and intervention delivery costs and time. A number of interested sites were unable to participate because of limits on health visitor capacity/contracts or because sites were already



**FIGURE 1** Flow of participants through the main trial. Note that 'unwilling' is defined as not wishing to attend the intervention and 'unable' is defined as unable to attend (e.g. because of lack of child care, returned to work or time the intervention was delivered).



delivering the IY-I or IY-T or were committed to running other parenting programmes. Four trial sites were selected that met the requirements. Two sites were in the north of England, both with council and NHS IY delivery, one site was in the Midlands of England, involving LA-only delivery (although health visitors also approached potential participants), and one site from the south of England, involving an NHS-only provider.

### Intervention

The E-SEE Steps model adopts the IY series [URL: [www.incredibleyears.com](http://www.incredibleyears.com) (accessed 7 January 2022)] of programmes for parents, children and teachers to enhance social and emotional well-being in children aged 0–12 years as content for the different levels of intervention. Two new programmes have been developed for parents of children aged 0–2 years: the IY-I and the IY-T (see *Appendix 1, Figure 11*, for the IY-I and IY-T logic model). Each programme is accompanied by a parenting book (i.e. the IY-B), reflecting the content of the programmes delivered in group sessions, with activity and journal pages. In addition, the E-SEE Steps model used the IY-B as a universal educational resource to increase parental knowledge and understanding about babies and toddlers' social and emotional development. All intervention parents received this book immediately following randomisation to intervention and postal tracking numbers were used to establish receipt of the book. There is evidence<sup>37,38</sup> to support a link between parents' knowledge and their practice/parenting behaviour. Within the universal proportionate framework, three levels of the IY parenting programme were investigated in this RCT.

#### Incredible Babies book

The IY-B is guide and journal of a baby's first year. The IY-B provides information on how to promote and understand a baby's physical, social, emotional and language development, and includes safety alerts, developmental principles and a journal section to record progress.

#### Incredible Years infant programme

In the infant group-based programme, parents learn how to help their babies feel loved, safe and secure. Parents learn how to encourage their babies' physical and language development. The programme involves parents attending a 2-hour session with their babies once per week for 10 weeks in groups of 8–10 parents (the standard IY-I is 8 weeks; however, developer guidance suggests that the content should be delivered over 10 weeks for 'high-need' families). The programme uses video clips of real-life situations to support the training and there are opportunities for group discussions and practice exercises for parents to do with their babies. Key content covered in the IY-I includes 'getting to know your baby', 'babies as intelligent learners', 'providing physical, tactile and visual stimulation', 'parents learning to read babies' minds' and 'gaining support'. The programme is delivered by two trained IY group leaders.

#### Incredible Years toddler programme

In the toddler group-based programme, parents learn how to help their toddlers feel loved and secure, how to encourage their toddler's language, social and emotional development, how to establish clear and predictable routines, how to handle separations and reunions and how to use positive discipline to manage misbehaviour. The programme (delivered by two trained group leaders) involves parents attending a 2-hour session once a week for 12 weeks in groups of 10–14 parents. A crèche may be provided during each session. Key IY-T content includes 'playing with your child', 'supporting your child's social, emotional and language development', 'using praise to encourage positive child behaviour', 'reinforcing positive behaviour', 'setting limits', 'handling separations' and 'managing unwanted behaviour'.

### Universal proportionate delivery of the intervention

Following baseline data collection, the universal-level intervention dose (i.e. the IY-B) was posted to all intervention families to read and use at home. Postal tracking was used to establish receipt of the book.



An active placebo (e.g. another type of book) was not specifically given to the control arm, as the control arm received SAU.

The IY-I and IY-T were targeted using two screener measures. Eligibility was based on the primary parent's level of depression, as measured by the Patient Health Questionnaire-9 items (PHQ-9), and/or the child's social and emotional well-being score, as measured by the Ages and Stages Questionnaire: Social and Emotional, 2nd edition (ASQ:SE-2). Eligibility for the IY-I and IY-T were assessed at follow-up 1 (2 months post baseline) (FU1) and follow-up 2 (9 months post baseline) (FU2), respectively, during home data collection visits. Eligible parents (and their co-parent if applicable) were invited to attend the IY-I and/or IY-T. The IY developer recommends home visits to eligible parents before the group sessions begin to better understand the family's needs and to build rapport and engagement.

In an attempt to ensure viable group size (i.e. a minimum of five parents), non-research parents could be invited to the groups. These parents were invited by service staff based on professional judgement/assessment. We did not collect data for these parents, as they attended as they would for any other parenting intervention delivered at that site. This is accepted practice in types of research interventions that require viable group size. The intervention families could also access all SAU.

## Setting for intervention delivery

Both IY-I and IY-T were delivered in convenient community venues and locations to reduce participant travel time, burden and drop-out. Travel and crèche facilities were provided by sites, where possible, and provision monitored. Each IY group was delivered by two IY-trained staff from health/LA staff, such as health visitors, infant mental health practitioners, children centre staff and nursery nurses. Group leaders attended separate 3-day training workshops for IY-I and IY-T that were delivered by accredited UK-based IY trainers. Sites were advised to deliver a 'dry run' practice of an IY-I or IY-T group prior to delivering the research groups. In addition, the IY developer recommended regular peer review and/or delivery supervision for facilitators for effective implementation and encouraged agencies to support accreditation of group leaders. Further details on training and supervision are provided in *Report Supplementary Material 1*.

## Control treatment

Control parents/co-parents had access to SAU. IY-I and IY-T did not form SAU in participating sites, although other parenting programmes may have been available. We documented the nature of SAU in each locality and collected data on which health (and social) services families had accessed via completion of an adapted Client Service Receipt Inventory (CSRI). A waiting list control design was not feasible, as children of control group parents would exceed the IY-I and IY-T age range after intervention group completion. Implementation partners were asked to consider offering the IY-T book or IY basic programme to control parents at the trial end.

## Outcomes and measures

Questionnaire packs were pre-tested with non-research parent representatives of the ethnic and socioeconomic profiles of the regions in the study [facilitated by the E-SEE Parent Advisory Committee (PAC)] to assess user-friendliness and comprehension of the questionnaire materials and length of time for completing them. Measure selection was also informed by a suite of systematic reviews of the measurement properties of outcome measures frequently used in RCTs of parenting programmes for children in the early years.<sup>39,40</sup> The child and parent measures presented were administered at all time points [i.e. baseline, FU1, FU2 and follow-up 3 (18 months post baseline) (FU3)] unless otherwise stated.

Provenance, properties and rationale for all measures are presented in the trial protocol.<sup>41</sup> All measures were completed by parents, with the exception of the Infant CARE-Index observational measure and, in addition, researchers completed the CSRI and demographic form based on parent responses.

### *Child primary outcome*

#### **Social and emotional well-being**

The parent-reported ASQ:SE-2<sup>42</sup> was used to establish effectiveness of the overall proportionate delivery of the E-SEE Steps model. This questionnaire was completed by the primary parent only. Population norms for ASQ:SE-2 are provided in *Appendix 2, Table 33*.

### *Key secondary parent outcome*

#### **Depression**

The parent-reported PHQ-9<sup>43</sup> was used to establish effectiveness of the proportionate delivery of the overall E-SEE Steps model. The PHQ-9 was completed by primary and co-parents. Population norms for the PHQ-9 are provided in *Appendix 2, Table 34*, along with a copy of the measure.

### *Other secondary outcomes for primary and co-parents*

#### **Parent-child attachment/interaction**

Parent-child attachment/interaction was measured at FU3 using the primary and co-parent-completed Maternal Postnatal Attachment Scale (MPAS) and/or Paternal Postnatal Attachment Scale (PPAS).<sup>44,45</sup>

#### **Parenting skill**

Parenting skill was assessed by the primary and co-parent-completed Parent Sense of Competence (PSOC) scale.<sup>46</sup>

#### **Health-related quality of life**

Health-related quality of life (HRQoL) was measured by the primary and co-parent-completed EuroQol-5 Dimensions, five-level version (EQ-5D-5L).<sup>47</sup>

#### **Service use**

Service use was assessed using an adapted CSRI<sup>48</sup> for the primary parent only.

### *Child secondary outcomes*

#### **Behaviour**

Behaviour was measured at FU3 using the primary and co-parent-completed Strengths and Difficulties Questionnaire (SDQ).<sup>49</sup>

#### **Cognitive development**

Cognitive development was measured at FU3 using the primary and co-parent-completed Pediatric Quality of Life Inventory (PedsQL™) infant scale.<sup>50</sup>

#### **Health (quality of life)**

Health (quality of life) was measured at FU3 using the primary and co-parent completed PedsQL Infant Scale.<sup>50</sup>

#### **Service use**

Service use was captured by the primary parent-completed CSRI.<sup>48</sup>

## Parent-child dyad secondary outcome measures

### Dyadic synchrony

Dyadic synchrony was assessed using the Infant CARE-Index (PM Crittenden, Family Relations Institute, 2010) observational report solely conducted with the primary parent-child dyad. Researchers filmed interactions in the home and videos were coded by researchers trained in the Infant CARE-Index.

### Other measures

#### Demographics

A bespoke parent/co-parent demographics report form was developed specifically for the study, capturing key information on age, ethnicity, religion, income, marital status and parent/co-parent education. The form also contained questions on breastfeeding and other feeding practices.

The co-parent and follow-up demographics forms were a shorter version of the baseline form.

#### Quality of relationships

The quality of relationships between parents (if applicable) was assessed via parent self-report using a very brief form developed specifically for the study.

Further measures for the process and economic evaluation are described in *Chapters 4 and 5*, respectively.

### Changes to trial outcomes after the trial commenced

Use of the Eyberg Child Behaviours Inventory was provided as an example child secondary measure to be used in the FU3 visits with participants. However, following work on a systematic review of measures in relevant domains/constructs as part of E-SEE study,<sup>40</sup> the SDQ was found to be more appropriate for the study. The SDQ is available in multiple languages, has a subscale of emotional well-being, is free of charge and is more widely available to services. In protocol version 10, the PHQ-9 is described as the 'parent and co-parent primary outcome'. That description is technically incorrect, as the trial is powered on ASQ:SE-2 alone. PHQ-9 is now described as the key secondary outcome. In addition, in protocol version 10, the Infant CARE-Index is described as a child attachment measure and listed as a child secondary outcome. Child attachment is just one part of the index and we will use the overall dyadic synchrony score, which is a parent-child dyad measure.

## Data collection

Data collection took place in the participant's own home. All data collectors were educated to postgraduate level (or with equivalent experience) and had prior experience working with children or families. All data collectors attended a study-specific training delivered by the E-SEE trial co-ordinator and E-SEE York trial manager prior to data collection. Data collectors also undertook safeguarding training and good clinical practice training. Further detail on training and data collection procedures can be found in the study protocol version 10.<sup>41</sup>

## Sample size

Sample size was calculated on the child primary outcome of social and emotional well-being using the ASQ:SE-2. The study was designed to evaluate the effectiveness of the whole E-SEE Steps model (i.e. 'Do the scores of children in the E-SEE Steps arm, on average, stay below those scores for children in the SAU arm over the three follow-up measures?'). Parents were eligible to be offered the proportionate intervention if they scored  $\geq 5$  on the PHQ-9 or their child scored within the monitoring

or cause for concern range on the ASQ:SE-2. We defined the minimal clinically important difference at FU3 to be 5 units on the ASQ:SE-2 in the E-SEE Steps arm compared with the SAU arm. We expected this effect to be consistently seen over the three follow-up points. Assuming a standard deviation (SD) of 18 units on the ASQ:SE-2 at FU3, the correlation between baseline and FU3 scores is 0.26, and between pairs of measurements after baseline is 0.40. For a design effect of 1.25 for the intervention arm, two-sided 5% significance level and 90% power, we would require the study to have retained 441 intervention participants and 92 control participants. Allowing for overall attrition of 12%, this would require 606 parents to be randomised with an allocation ratio of 5 : 1. The high allocation ratio was required to ensure that a sufficient number of parents were eligible and able to attend IY groups. Assuming the attrition rate of 12%, we expected 151 parents to be eligible for IY-I, with 50 parents attending, and 147 parents to be eligible for IY-T, with 48 parents attending, based on our pilot data.

## Additional information about the ASQ:SE-2

The ASQ:SE-2 thresholds for concern are shown in *Table 2*. For interventions aimed at improving the mental health of parents, a minimal(ly) clinically important difference of 5 has been suggested; however, given that this is an intervention aimed at child behaviour, it is not appropriate here but is provided for reference.

## Randomisation

Randomisation was at the individual level, using a web-based randomisation system developed by the Sheffield Clinical Trials Research Unit (Sheffield, UK) in collaboration with epiGenesys (Sheffield, UK) and using a randomisation sequence prepared by the trial statistician. E-SEE Step participants were randomised in a 5 : 1 ratio (intervention to control arms) stratified by sex of child, sex of carer, recruitment site and PHQ-9 baseline score or ASQ:SE-2 baseline score (with the last two variables representing level of need).

Randomisation occurred after eligibility had been established, informed consent obtained and baseline measures collected from parents to reduce initial attrition. The allocation schedule was concealed and the intervention arm was confirmed only once eligibility and consent were confirmed by researchers. A member of the research team inputted participant information to the online system to enable randomisation, with allocation results returned immediately. The trial co-ordinator informed families of allocation to condition.

Parent-child dyads were randomly allocated to intervention or control arms and stratified according to level of need based on the depression score of the parent with the main parenting responsibility or the child's social and emotional well-being (assessed using ASQ:SE-2), sex of the child and primary parent, and recruitment site. The co-parent was automatically assigned the same allocation as the parent.

TABLE 2 Thresholds for concern: ASQ:SE-2

Threshold for concern	Age interval (version of ASQ:SE-2)			
	2 months	6 months	12 months	18 months
Monitor: it is close to the cut-off point (review behaviours of concern and monitor)	25-34	30-44	40-49	50-64
Refer: it is above the cut-off point (further assessment with a professional may be needed)	35+	45+	50+	65+

If the parent had missing depression data, measured using the PHQ-9, then randomisation was not possible. We did not randomise participants who had missed three or more of the main questions on the PHQ-9 measure at baseline.

### Blinding

Participants and group leaders delivering the intervention were not blind. Data managers and research staff involved in recruitment, initial assessments and fidelity/process assessment were also not blind to allocation. Production of ongoing Trial Management Group (TMG), Trial Steering Committee (TSC) and Data Monitoring and Ethics Committee (DMEC) reports was undertaken by unblinded programmers. The trial statistician remained blind while the study was in progress, but was unblinded when conducting the final analysis. All data collectors were blind to the research questions and condition allocation. Unblinding of blinded members of the E-SEE team was made known to study management via the protocol non-compliance procedure and logged. Further information on steps to ensure the maintenance of blinding can be found in our full protocol.<sup>41</sup>

## Statistical methods

Information on the calculation of the sample size for the main trial and the clinical meaning of intervention effects has been provided in *Sample size*. Information on the analysis for the ancillary substudies is included in the study protocol, which is available along with the E-SEE main trial statistical analysis plan (SAP) and pilot SAP.<sup>41</sup>

Effectiveness was assessed using intention-to-treat analysis, accounting for the repeated measurements of individuals at FU1, FU2 and FU3 using general estimating equations. Membership of the IY-I and IY-T groups is confounded with outcome at FU1 and FU2 and so programme group clustering was unaccounted for because it is not possible to do this appropriately without biasing the treatment effect estimate in the methods we have explored. This is also current practice in the wider community.<sup>31</sup> Simulations conducted during SAP development suggested that estimates from our chosen model are robust to intracluster coefficients (ICCs) < 0.2.

### Treatment effectiveness

The overall effectiveness of the proportionate delivery of IY was assessed by intention-to-treat analysis using a marginal model fitted using general estimating equations with a Gaussian family, identity link and autoregressive covariance structure of order 1 AR(1). AR(1) means that each observation in the time series is directly related to the observation that preceded it. Model estimates with standard errors that are robust to the non-normality and non-independence of observations were computed. Statistical analyses used Stata<sup>®</sup>/MP 16.0 (StataCorp LP, College Station, TX, USA).

### Inclusion of covariates

There were no major differences between the groups at baseline. Covariates included were baseline PHQ-9 and ASQ:SE-2 scores, whether or not the parent had a degree, whether or not the parent was in a relationship, the primary parent's ethnicity, the child's sex, follow-up time and delivery site. Site fixed effects were included to minimise unexplained variance in site-specific effects and facilitate generalisability by capturing factors that explain why effects vary across sites (e.g. differences in parental recruitment and retention, and implementation fidelity).

### Missing data

Full case and survey item non-response were explored. Item non-response was imputed using questionnaire developer rules. Sensitivity to missing outcome data and missing information on whether or not the primary parent held a degree was explored using chained equations to impute 25 multiply imputed data sets using a logit model and a multivariate normal mode for the repeated-measure

outcome variable. Covariates were the same as those in the primary analysis model plus randomisation group and whether or not the dyad met the criteria for inclusion in the IY-I and IY-T groups.

### *Analysis populations*

#### **Inequalities**

We conducted a simple moderator analysis to explore the extent to which intervention effectiveness differed between distinct subpopulations by testing the significance of the interaction between randomised treatment group and subgroup (e.g. parent education level, first child, sex of child, site).

#### **Intervention components**

The impact of individual components (i.e. IY-B, IY-I and IY-T) was investigated using non-randomised observational analysis where participants in the control arm with outcome scores above the eligibility threshold were used as a pseudo-control group.

### *Additional analyses*

#### **Sensitivity analysis**

To assess the robustness of the outcome analysis, the primary analysis was repeated with alternative specifications of the primary outcome measure and using multiple imputation (MI). Per-protocol and complier-average causal effect analyses were not conducted, as there is no satisfactory way of defining compliers without biasing the estimated impact of IY-I and IY-T on compliers because of the conditional design whereby eligible participants have already scored highly on the outcome measure. A descriptive analysis of the characteristics associated with compliance was undertaken.

#### **Study oversight and management**

The University of York (York, UK) was the trial sponsor. A TSC comprised an independent chairperson, a member with early years expertise, an independent statistician and lay representatives (including a member of the PAC when available). The TMG comprised the chief investigator, trial managers, a trial statistician, a data manager, a health economist and other co-investigators. The DMEC included an independent chairperson, a professor of psychology and an independent trial statistician. These committees functioned in accordance with Sheffield Clinical Trials Research Unit standard operating procedures.

### **Ethics arrangements**

This study was approved by NHS Wales Research Ethics Committee 5 (Bangor, UK) on 22 May 2015 (reference 15/WA/0178 and IRAS 173946). Health Research Authority approval, which was issued on the basis of an existing assessment of regulatory compliance, was received on 13 September 2016 (i.e. a letter of Health Research Authority approval for a study processed through pre-Health Research Authority approval systems). The chief investigator's departmental ethics committee at the University of York required submission of project documentation and approval was given on 10 August 2015 (reference FC15/03).

### **Changes to protocol**

Version 10 is the final version of the protocol. Please refer to the protocol for a full list of amendments.<sup>41</sup> Significant amendments (i.e. those not related to simple language and terminology) to the protocol in the main trial phase (i.e. versions 5–10) are detailed below. All amendments were approved by the Research Ethics Committee, E-SEE TSC and DMEC, and were accepted by the funder.

### Version 5: 23 February 2017

- Section 4 (Gantt chart and key milestones): removal of study Gantt chart to allow flexible timing of group programme delivery. Milestones table updated to reflect access of health data at end of study only.
- Section 5 (ancillary substudies): change made to access health records at a single time point at the end of the study, rather than during pilot and at the end of study, in the light of NHS Digital costs.
- Section 6 (selection, recruitment and withdrawal of participants): parents may be given/sent a reminder card about the focus groups/interviews.
- Section 8 (intervention): change made under 'setting and delivery' to allow a more flexible approach to intervention delivery, which will be guided by each site/service provider.
- Section 9 (assessments and procedures): the Eyberg Child Behaviour Inventory was replaced by the SDQ as the child secondary outcome measure.

### Version 6: 8 May 2017

Trial design, sample size, randomisation and statistical sections of the protocol were amended because of changes to the main trial, as informed by the pilot phase:

- The pilot was external with references to internal pilot removed.
- Inclusion of the ASQ:SE-2 as an eligibility screener (protocol version 4) meant that ASQ:SE-2 would be used as a stratification variable in the main phase. Clarification added that recruitment site is a stratification variable (pilot and main phase).
- Strategy of lowering the PHQ-9 eligibility threshold in the pilot would not be used in the main trial.
- Change to sample size calculations for the main trial. Calculations for the pilot and main phase are presented separately in section 6.
- Allocation ratio of 4.8 : 1 applied in the main trial.
- Adaptation of the statistical design to look at effectiveness of the programme overall, rather than at each stage of the IY programme.
- Removal of the requirement to calculate ICC for the pilot (protocol version 4) meant that an independent statistician is no longer required at that stage.

### Version 7: 16 May 2017

- Section 6 (selection, recruitment and withdrawal of participants): parents and co-parents who have taken part in the pilot phase will not be eligible to take part in the main phase of the trial.

### Version 8: 7 November 2017

- Study summary, section 6 (sample size calculation) and section 7 (randomisation): adjustment of the allocation ratio from 4.8 : 1 to 5 : 1.

### Version 9: 8 November 2018

- Section 6 (selection, recruitment and withdrawal of participants): minor amendment to allow inclusion of non-research participants to the group intervention based on professional judgement/assessment made by of service staff.

### Version 10: 17 June 2019

- Section 9 (assessment and procedures): clarification that the effectiveness of the proportionate delivery of the IY E-SEE Steps model overall (i.e. three levels of IY-B, IY-I and IY-T) will be determined through use of the ASQ:SE-2 and PHQ-9. Footnote 4 added to the parent and co-parent primary outcome. Infant CARE-Index videos will be coded by trained members of the research team.



In addition to the amendments above, three changes were made to the analysis model in the protocol. These changes were justified and agreed by all stakeholders (i.e. the National Institute for Health Research, DMEC and TSC) prior to database close and unblinding. First, we replaced a multilevel mixed model with treatment group and participants as random effects, with a marginal model fitted using general estimating equations.

Second, we no longer accounted for treatment group clustering because the offer of IY-I and IY-T was conditional on outcome at FU1 and FU2 and, therefore, clustering was confounded with treatment effect, leading to biased estimation of the latter. We used a marginal model fitted using general estimating equations because accounting for repeated measures using a mixed model inflates the type I error (random intercept-only model) or gives a biased estimate of the treatment effect (random intercept and slope model). Simulations conducted during SAP development suggested that estimates from this alternative model are robust to ICCs < 0.2.

Third, cluster-level analysis using summary measures was no longer included because participants can get IY-I alone, IY-T alone or both IY-I and IY-T and, therefore, there was no way of grouping participants into clusters that remain stable throughout the intervention. The sex of primary caregiver covariate was not used because findings from the pilot showed too few male primary caregivers for the associated model parameter to be estimated.

### **Patient and public involvement**

The involvement of parents and caregivers of very young children has been integral to the E-SEE trial. In this section we outline the different ways that patient and public involvement has contributed to the E-SEE model, from the planning stages through to the dissemination of trial findings. Reflections on what worked well and not so well are also provided.

#### ***Pre-funding preparation***

We conducted focus groups with family and children service staff and fathers to establish how to be 'father inclusive' when inviting fathers to groups, and how best to retain them. We also asked for fathers' opinions on the developing study design and the importance of the research questions. The meetings were informative and a brief report was produced to summarise the key learning.<sup>51</sup>

#### ***Post-funding preparatory work***

A PAC was established with members from two study sites. Members included mothers, fathers, step-parents and grandparents with children of similar ages to the trial index children. The group was ethnically diverse. The main roles of the PAC were to advise and support researchers on recruitment to the trial and advise on, and assist with, training in the measures to be used (e.g. the PAC supported the training of E-SEE data collectors by role playing data collection visits with the researchers). The PAC also advised on trial retention strategies, and on publicity and dissemination. PAC members received training for their committee role.

#### ***Throughout the pilot and main trial***

The PAC raised awareness of the trial through community events in their local communities, sharing flyers with local groups, and also kept the research team informed of local events and groups targeted at the trial population of interest. In addition, the PAC contributed to the E-SEE trial website design and had their own E-SEE Facebook (Facebook, Inc., Menlo Park, CA, USA) page to enhance engagement with each other and the research team. One PAC member became a lay member on the TSC for a period of time.



### **Dissemination**

Pre COVID-19, PAC members were scheduled to participate in and support site dissemination events. PAC members contributed to pilot site awareness events and parent and staff end-of-pilot events. Infographics for participating families have been reviewed by the PAC for levels of understanding and acceptability. PAC members reviewed the *Plain English summary* of this report.

### **Reflections on what worked well and not so well**

Ensuring representation of parents from all our study sites was not possible, despite several attempts to recruit PAC members via social media adverts, e-mails to existing networks and groups, and posters in local children's centres. However, we were successful in recruiting PAC members from two of our sites through local voluntary organisations working with families. PAC members' local knowledge of where the best places/groups to raise awareness were was very useful. It was particularly useful to get the PAC members' views on measure choice (e.g. if a measure looked simple and nicely laid out then it was preferred to a shorter measure that looked less professional or attractive). The wording of questions was also scrutinised for understanding and informed our choice. Having PAC members support E-SEE data collectors in role play was invaluable, as it gave 'real-life' experience of trying to complete measures with parents while children demanded their parents' attention.

Although we were well informed by father groups pre funding and offered some father inclusivity training, few fathers were recruited to the trial and fewer still attended groups when their partner was eligible to attend. We write more about this in our substudy in *Appendix 10*.



## Chapter 3 Outcome evaluation results

Parts of this chapter have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Details of the sample recruited and retention of participants across time points are available in *Figure 1*.

### Main outcome measures

The primary outcome was child social and emotional well-being using the ASQ:SE-2. Higher values of ASQ:SE-2 indicate a higher degree of risk in terms of the social and emotional well-being of children. The key secondary outcome was the PHQ-9 depression module for the primary carer, subsequently referred to as PHQ-9. Higher values on the PHQ-9 indicate higher levels of depression for the parent. Therefore, for both the primary and key secondary outcome, we were looking for the difference in average scores between the treatment and control group to be negative for the results to indicate efficacy. We were powered on the primary outcome only and so present findings in terms of significance for ASQ:SE-2 and trends or confidence intervals (CIs) for PHQ-9.

There are age-appropriate versions of the ASQ:SE-2. Three versions were used in this trial.

### Data completeness

The trial had very high levels of response to all surveys at all follow-up points both in terms of the number of surveys returned and the number of items completed on each survey (see *Report Supplementary Material 2*). Therefore, very little imputation was required.

### Baseline covariates

No major imbalance between arms at baseline was found, both in terms of covariates (*Table 3*) and baseline outcome scores (see *Appendix 3, Table 35*). Mean ASQ:SE-2 score at baseline was moderately lower in the treatment group than in the control group, but this was adjusted for in the primary analysis, as baseline ASQ:SE-2 was a planned covariate.

TABLE 3 Baseline demographic characteristics of child, primary caregiver and co-parent

Variable	Treatment arm		All
	Treatment	Control	
<b>Child</b>	<b>N = 285</b>	<b>N = 56</b>	<b>N = 341</b>
<i>Categorical variable, n (%)</i>			
Sex of child			
Male	145 (51)	29 (52)	174 (51)
Female	140 (49)	27 (48)	167 (49)
Child's ethnicity			
English/Welsh/Scottish/Northern Irish/British	215 (75)	37 (66)	252 (74)
Any other white background	9 (3)	4 (7)	13 (4)
			continued

OUTCOME EVALUATION RESULTS

TABLE 3 Baseline demographic characteristics of child, primary caregiver and co-parent (continued)

Variable	Treatment arm		
	Treatment	Control	All
White and black Caribbean	2 (1)	1 (2)	3 (1)
White and black African	3 (1)	1 (2)	4 (1)
White and Asian	5 (2)	0 (0)	5 (1)
Any other mixed/multiple ethnic group	6 (2)	0 (0)	6 (2)
Indian	14 (5)	6 (11)	20 (6)
Pakistani	19 (7)	5 (9)	24 (7)
Bangladeshi	4 (1)	0 (0)	4 (1)
Any other Asian background	2 (1)	0 (0)	2 (1)
African	5 (2)	1 (2)	6 (2)
Any other ethnic group	1 (0)	1 (2)	2 (1)
Premature			
No	274 (96)	53 (95)	327 (96)
Yes	9 (3)	3 (5)	12 (4)
Missing data	2 (1)	0 (0)	2 (1)
Difficulties at birth			
No	132 (46)	26 (46)	158 (46)
Yes	153 (54)	30 (54)	183 (54)
<i>Continuous variable</i>			
Child's age (weeks)			
n (%)	285 (100)	56 (100)	341 (100)
Mean (SD)	6.1 (2.1)	5.9 (2.2)	6.0 (2.1)
Median (IQR)	6 (4–8)	6 (4–8)	6 (4–8)
Minimum, maximum	2, 11	2, 10	2, 11
<b>Primary caregiver</b>			
<i>Categorical variable, n (%)</i>			
Parent's age group (years)			
18–21	9 (3)	2 (4)	11 (3)
22–25	36 (13)	7 (13)	43 (13)
26–30	88 (31)	15 (27)	103 (30)
31–35	95 (33)	21 (38)	116 (34)
≥ 36	57 (20)	11 (20)	68 (20)
Sex			
Female	285 (100)	56 (100)	341 (100)
Ethnicity			
English/Welsh/Scottish/Northern Irish/British	218 (76)	38 (68)	256 (75)
Irish	2 (1)	0 (0)	2 (1)
Any other white background	12 (4)	4 (7)	16 (5)

TABLE 3 Baseline demographic characteristics of child, primary caregiver and co-parent (continued)

Variable	Treatment arm		All
	Treatment	Control	
White and black Caribbean	1 (0)	0 (0)	1 (0)
White and black African	0 (0)	1 (2)	1 (0)
White and Asian	0 (0)	1 (2)	1 (0)
Any other mixed/multiple ethnic group	5 (2)	0 (0)	5 (1)
Indian	15 (5)	7 (13)	22 (6)
Pakistani	18 (6)	3 (5)	21 (6)
Bangladeshi	4 (1)	0 (0)	4 (1)
Any other Asian background	3 (1)	0 (0)	3 (1)
African	7 (2)	1 (2)	8 (2)
Any other ethnic group	0 (0)	1 (2)	1 (0)
Highest qualification previously achieved			
Post-doctorate qualification	8 (3)	0 (0)	8 (2)
Master's degree	28 (10)	8 (14)	36 (11)
Undergraduate degree (e.g. BA or BSc)	96 (34)	14 (25)	110 (32)
A certificate or diploma in higher education	33 (12)	5 (9)	38 (11)
A, AS or S Levels	19 (7)	7 (13)	26 (8)
O Levels or GCSE: five or more	15 (5)	6 (11)	21 (6)
O Levels or GCSE: four or fewer	9 (3)	3 (5)	12 (4)
Overseas qualifications	10 (4)	2 (4)	12 (4)
Vocational qualifications	53 (19)	8 (14)	61 (18)
None of these qualifications	14 (5)	1 (2)	15 (4)
Missing data	0 (0)	2 (4)	2 (1)
Relationship status			
Married and living together	184 (65)	38 (68)	222 (65)
Cohabiting/living together	70 (25)	12 (21)	82 (24)
Living together part of the time	4 (1)	3 (5)	7 (2)
Separated	4 (1)	0 (0)	4 (1)
A couple but not living together	13 (5)	0 (0)	13 (4)
Dating	1 (0)	1 (2)	2 (1)
Not in a relationship	9 (3)	2 (4)	11 (3)
<i>Continuous variable</i>			
Age (years)			
<i>n</i> (%)	285 (100)	56 (100)	341 (100)
Mean (SD)	30.9 (5.1)	31.1 (5.0)	30.9 (5.0)
Median (IQR)	31 (28–35)	32 (27–34)	31 (28–34)
Minimum, maximum	18, 43	20, 40	18, 43

continued

## OUTCOME EVALUATION RESULTS

TABLE 3 Baseline demographic characteristics of child, primary caregiver and co-parent (continued)

Variable	Treatment arm		
	Treatment	Control	All
Baseline weekly income (£)			
n (%)	226 (79)	43 (77)	269 (79)
Mean (SD)	733.1 (470.7)	766.9 (454.4)	738.5 (467.5)
Median (IQR)	603 (400–950)	710 (400–1000)	630 (400–973)
Minimum, maximum	0, 2500	151, 1850	0, 2500
<b>Co-parent</b>	<b>N = 53</b>	<b>N = 15</b>	<b>N = 68</b>
<i>Categorical variable, n (%)</i>			
Parent's age group (years)			
18–21	2 (4)	0 (0)	2 (3)
22–25	1 (2)	2 (13)	3 (4)
26–30	13 (25)	2 (13)	15 (22)
31–35	13 (25)	5 (33)	18 (26)
≥ 36	22 (42)	5 (33)	27 (40)
Missing data	2 (4)	1 (7)	3 (4)
Sex			
Male	51 (96)	15 (10)	66 (97)
Female	1 (2)	0 (0)	1 (1)
Missing data	1 (2)	0 (0)	1 (1)
Relationship status			
Married and living together	39 (74)	13 (87)	52 (76)
Cohabiting/living together	12 (23)	1 (7)	13 (19)
Living together part of the time	0 (0)	1 (7)	1 (1)
A couple but not living together	1 (2)	0 (0)	1 (1)
Missing data	1 (2)	0 (0)	1 (1)
Highest qualification previously achieved			
Post-doctorate qualification	1 (2)	1 (7)	2 (3)
Master's degree	7 (13)	0 (0)	7 (10)
Undergraduate degree (e.g. BA or BSc)	16 (30)	3 (20)	19 (28)
A certificate or diploma in higher education	6 (11)	4 (27)	10 (15)
A, AS or S Levels	1 (2)	2 (13)	3 (4)
O Levels or GCSE: five or more	6 (11)	1 (7)	7 (10)
O Levels or GCSE: four or fewer	4 (8)	1 (7)	5 (7)
Overseas qualifications	1 (2)	1 (7)	2 (3)
Vocational qualifications	8 (15)	2 (13)	10 (15)
None of these qualifications	1 (2)	0 (0)	1 (1)
Missing data	2 (4)	0 (0)	2 (3)

TABLE 3 Baseline demographic characteristics of child, primary caregiver and co-parent (continued)

Variable	Treatment arm		
	Treatment	Control	All
<i>Continuous variable</i>			
Age (years)			
n (%)	51 (96)	14 (93)	65 (96)
Mean (SD)	34.5 (8.3)	35.1 (8.2)	34.6 (8.2)
Median (IQR)	33 (29–39)	35 (30–40)	34 (29–39)
Minimum, maximum	18, 68	24, 55	18, 68
Baseline weekly income (£)			
n (%)	44 (83)	11 (73)	55 (81)
Mean (SD)	915.8 (895.9)	690.4 (483.4)	870.7 (831.1)
Median (IQR)	760 (460–1079)	600 (250–1000)	700 (450–1058)
Minimum, maximum	93, 6000	115, 1645	93, 6000

A Level, Advanced Level; AS Level, Advanced Subsidiary Level; BA, Bachelor of Arts; BSc, Bachelor of Science; GCSE, General Certificate of Secondary Education; IQR, interquartile range; O Level, Ordinary Level; S Level, Scholarship Level.

Although we assessed imbalance between the arms, we did not assess representativeness of the sample to local or national population data. What was noticeable here was the large number of E-SEE parents in live-in relationships (90% and 89% of intervention and control parents, respectively, with 65% and 68% married, respectively). According to the Office for National Statistics,<sup>52</sup> marriage prevalence is 50% and 60% of the population (aged  $\geq 16$  years) were living as a couple in 2019, the majority of whom were married. Only 3% and 4% of parents (intervention and control, respectively) said that they were not in a relationship, which is below the national average of 25% of families being lone-parent families. The numbers were too small to conduct subgroup analyses.

### Primary analysis of primary outcome: ASQ:SE-2

The mean ASQ:SE-2 score was higher in the treatment group than in the control group at all follow-up points apart from FU3. As higher scores relate to poorer outcomes, this provides no evidence in the direction of efficacy (Figure 2).

There was a borderline statistically significant positive difference in ASQ:SE-2 scores between arms after controlling for baseline covariates and stratification variables, suggesting that the intervention may have a detrimental effect (3.02, 95% CI -0.03 to 6.08) (Table 4).

Residuals from the primary analysis model were not perfectly normal, as they were symmetric around zero, with no outlying or overtly influential values (see Appendix 4, Figure 12). The validity of the primary analysis was, nevertheless, investigated using a sensitivity analysis because the ASQ:SE-2 itself has a positively skewed distribution and, together with the extreme imbalance in allocation the type I error, may be increased because there is a heightened risk that, by chance, we observe more extreme values in the larger treatment arm.

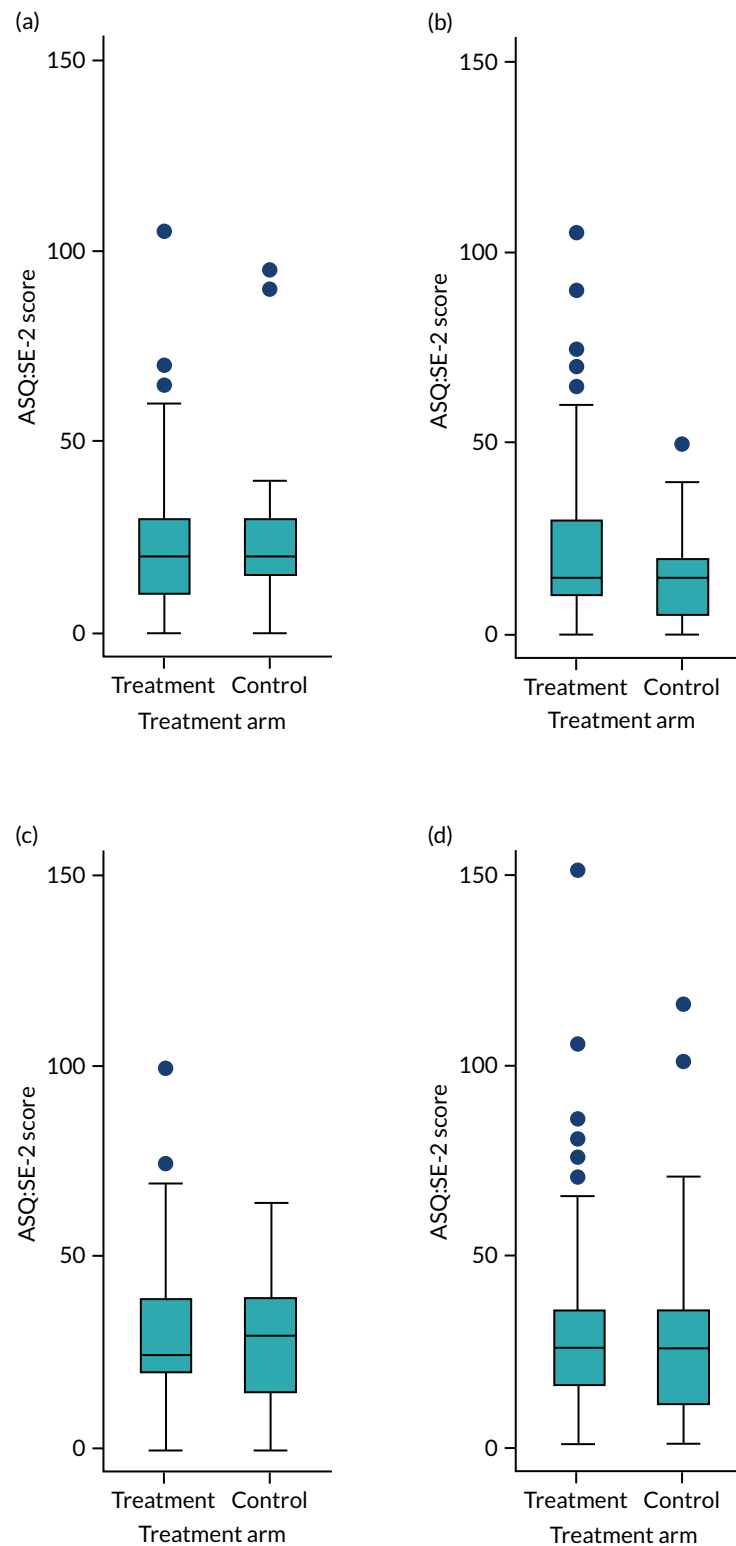


FIGURE 2 Summary of overall ASQ:SE-2 scores by follow-up point. (a) Baseline; (b) FU1; (c) FU2; and (d) FU3.

### Sensitivity analysis of the primary outcome

In addition to the pre-planned sensitivity analysis outlined in the SAP [see URL: [www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/#tab-3](http://www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/#tab-3) (accessed 10 January 2022)], unplanned additional sensitivity analyses were undertaken because of the two issues above.



TABLE 4 Summary statistics and difference between arms for ASQ:SE-2

Time point	Treatment arm		All (n = 341)	Mean difference (95% CI)	Adjusted mean difference (95% CI); p-value
	Treatment (n = 285)	Control (n = 56)			
Baseline					
n (%)	285 (100)	56 (100)	341 (100)		
Mean (SD)	22.8 (15.1)	23.8 (16.6)	23.0 (15.4)	-0.91 (-5.32 to 3.50)	
Median (IQR)	20 (10-30)	20 (15-30)	20 (15-30)		
Minimum, maximum	0, 105	0, 95	0, 105		
FU1					
n (%)	270 (95)	55 (98)	325 (95)		
Mean (SD)	20.5 (15.7)	16.5 (10.9)	19.8 (15.1)	4.06 (-0.29 to 8.41)	
Median (IQR)	15 (10-30)	15 (5-20)	15 (10-25)		
Minimum, maximum	0, 105	0, 50	0, 105		
FU2					
n (%)	269 (94)	55 (98)	324 (95)		
Mean (SD)	29.0 (16.2)	26.8 (14.7)	28.6 (16.0)	2.14 (-2.50 to 6.78)	
Median (IQR)	25 (20-40)	30 (15-40)	25 (20-40)		
Minimum, maximum	0, 100	0, 65	0, 100		
FU3					
n (%)	268 (94)	53 (95)	321 (94)		
Mean (SD)	26.8 (19.5)	28.1 (23.6)	27.0 (20.2)	-1.30 (-7.26 to 4.66)	
Median (IQR)	25 (15-35)	25 (10-35)	25 (15-35)		
Minimum, maximum	0, 150	0, 115	0, 150		
Overall					3.02 (-0.03 to 6.08); 0.052

IQR, interquartile range.

The first planned sensitivity analysis refitted the primary analysis model after transforming ASQ:SE-2 to z-scores and percentage scores (where percentage score is the difference from minimum score expressed as a percentage of the difference between the minimum and maximum scores possible). This transformation, to some extent, also addressed the issue that three different (age-specific) ASQ:SE-2 versions were used over the follow-up periods. The second planned sensitivity analysis was MI. However, because of the skewed distribution of the outcome and concerns over inflated type I error rates due to unbalanced allocation, we felt that further sensitivity analyses were justified.

Four types of unplanned sensitivity analyses were conducted. First, we re-ran the primary analysis on square root-transformed primary outcome scores. *Appendix 4, Table 36*, shows that the residuals for this model are less skewed. Second, we re-ran the primary analysis model, assuming an unstructured rather than auto-correlated structure (see *Appendix 4, Table 37*). This allowed more of these data to be used because Stata omitted participants who do not have data for at least two adjacent follow-up points. Third, owing to the skewed raw data and imbalance in the size of each arm, a permutation test was carried out. We permuted the allocation to arm 1000 times, keeping all other variables unchanged, and calculated the proportion of times we observed an effect estimate greater than or equal to the actual effect estimate to calculate a permutation-based p-value.

Finally, MI was undertaken using the difference between each participant’s ASQ:SE-2 follow-up score and baseline ASQ:SE-2 score. MI on this ‘differenced’ data removed much of the skew in these data because participants with extreme ASQ:SE-2 values at follow-up typically had similarly extreme values for ASQ:SE-2 at baseline and so the differences were more normally distributed. The MI was conducted using chained equations and 25 replications, with qualification imputed using a logit model and ASQ:SE-2 imputed using a multivariate normal model by treating each time point as a separate variable in the chain. Complete data were available for all other variables in the primary analysis model.

Table 5 summarises the outputs from all models (primary and sensitivity) and Figure 3 shows a forest plot for all analyses, with results transformed into original ASQ:SE-2 units. The sensitivity analyses suggest that the results were not particularly sensitive to the imbalance between arms and the skewed distribution of the primary outcome. However, the strength of the signal diminishes when methods that allow for missing and incomplete data are employed. To understand why the additional observations had this effect, we explored some features of the missing data by comparing baseline ASQ:SE-2 scores and treatment allocation between the missing and non-missing observations (see Appendix 5, Table 38). It is difficult to draw any conclusion from this comparison because of the complex relationship between missing data, time point, allocation and effect.

TABLE 5 Summary of primary and sensitivity analysis models

Analysis	Adjusted mean difference (95% CI)	p-value
Primary analysis model	3.02 (-0.03 to 6.08)	0.052
Sensitivity analyses		
z-score transformation	0.19 (0.01 to 0.36)	0.036
Percentage transformation	0.81 (0.06 to 1.56)	0.035
Primary with unstructured correlation	2.56 (-0.69 to 5.80)	0.122
Permutation test		0.069
Difference from baseline: MI	1.97 (-1.22 to 5.17)	0.23

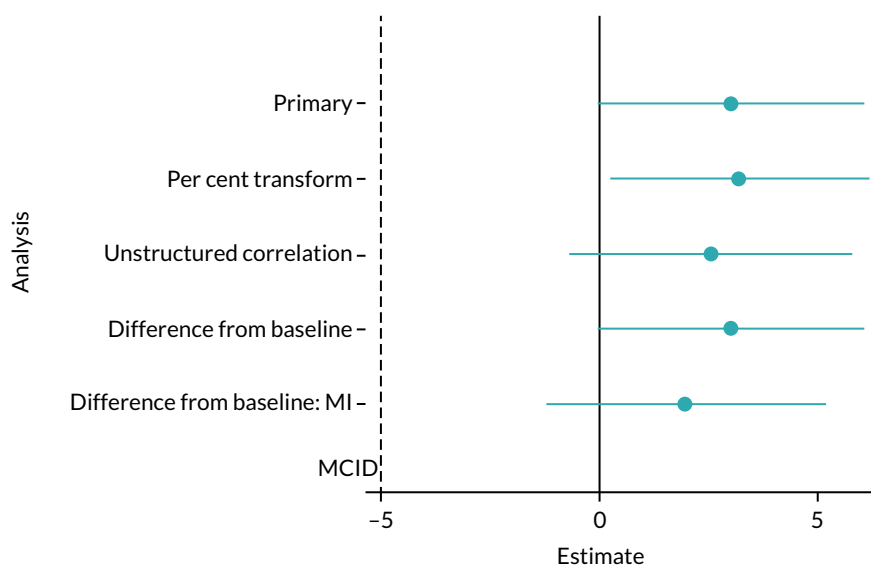


FIGURE 3 Forrest plot for primary analysis models on original scale. MCID, minimum clinically important difference.

## Exploratory analysis of the primary outcome

It is interesting to note that comparing the unadjusted differences between arms at each follow-up point suggested a stronger early effect in favour of SAU, which tended to diminish over time (see *Table 4*). Most of the weight of the signal in the overall effect was due to the larger difference seen at FU1, which then diminished over time. This might suggest that the book alone increased the primary outcome in the treatment group compared with the control group. This pattern was more pronounced after subtracting the baseline score from the score at each follow-up point and looking at unadjusted differences between arms. In this adjusted analysis, the size of the effect decreased from 4.6 at FU1 to 2.9 at FU2 and -0.9 at FU3.

Two planned and one unplanned exploratory analysis were conducted. First, the subgroup analysis was conducted for social and economic background (using 'primary parent educated to degree level' as a proxy), whether or not the index child was the first child, sex of the child and site by adding an interaction between treatment group and subgroup into the primary analysis model. For 'first child', the main effect was additionally included because it is not a covariate. There was weakly significant evidence of an interaction with treatment for site ( $p = 0.06$ ), but no evidence of a significant interaction for the other three subgroups (*Table 6*). The site-level control groups were too small for additional statistical analysis (*Figures 17–20 in Appendix 6* do not suggest better scores in the treatment groups compared with control for any of the sites).

Second, planned indicative analysis of whether or not the baby and toddler groups had an impact were undertaken. This analysis is indicative only because allocation to group programme occurs after randomisation. Eligibility for group programme is determined by ASQ:SE-2 and PHQ-9 scores at FU1 for IY-I and at FU2 for IY-T. The indicative analysis compared the eligible participants with the subgroups of control participants that had ASQ:SE-2 and PHQ-9 scores in the eligible range for the groups (referred to as pseudo-controls), using the same model as used for the primary outcome. The analysis found no evidence of a difference between group programme participants and the pseudo-controls (see *Appendix 7, Table 39*).

Finally, we conducted an unplanned analysis by converting ASQ:SE-2 score into a categorical variable indicating different risk categories in line with how it is commonly used in practice. This analysis was undertaken to explore the possibility that, even though the intervention was not effective at the aggregate level, dyads in certain risk categories may have responded in a clinically important way. We compared the proportion of children in each risk category at each follow-up as a proportion of the number in that risk category at baseline and found no difference between arms in how participants

TABLE 6 Primary outcome subgroup analysis

Subgroup	Interaction coefficient (95% CI)	Interaction $p$ -value
Below degree level	-2.88 (-9.19 to 3.43)	0.37
Male (child)	-0.23 (-6.28 to 5.81)	0.94
Not first child	1.95 (-4.15 to 8.06)	0.53
Site		
1	-9.99 (-17.25 to -2.74)	
2	-3.49 (-9.9 to 2.93)	
3	-6.80 (-14.43 to 0.83)	
4	Omitted	
Overall site interaction		0.06

transitioned between risk categories following the intervention (Table 7). Similar analyses are provided for the PHQ-9 (see Table 10) and the Infant CARE-Index (see Table 11) and these analyses, again, showed no difference.

### Key secondary outcome: depression (PHQ-9) score of primary caregiver

Table 8 and Figure 4 show that the key secondary outcome (i.e. parent depression score) was lower (i.e. better) in the treatment group at FU1 after the IY-B universal dose than at baseline.

The PHQ-9 items scores also had a skewed distribution (see Appendix 8, Figures 21 and 22) and so an unplanned sensitivity analysis was undertaken following the same argument as applied to the sensitivity analysis of the primary outcome. We applied three of the unplanned sensitivity analyses used for the primary outcome. It is of interest that, conversely to the primary outcome, the direction of the effect for PHQ-9 consistently favours the intervention, although the effect is not statistically significant (Table 9). The magnitude of the signal shifts in favour of the treatment as we move from the planned model to the two sensitivity analyses (see Appendix 8, Figure 23). It is difficult to interpret the meaning of an effect of this size (i.e. < 1 unit) for the PHQ-9 in the context of this trial. In depression studies, a minimum clinically important difference (MCID) of 5 units is typical. However, in the public health context, a more modest effect would be meaningful because a trial would typically not recruit severely depressed parents.

TABLE 7 Transition between ASQ:SE-2 categories

ASQ:SE-2 category	Treatment arm	N	ASQ:SE-2 category at baseline, n (%)		
			Low or no risk	Monitoring zone	Refer zone
FU1					
Low or no risk	Treatment	153	134 (88)	15 (10)	4 (3)
	Control	30	27 (90)	3 (10)	0 (0)
Monitoring zone	Treatment	61	42 (69)	11 (18)	8 (13)
	Control	14	10 (71)	3 (21)	1 (7)
Refer zone	Treatment	56	19 (34)	20 (36)	17 (30)
	Control	11	8 (73)	1 (9)	2 (18)
FU2					
Low or no risk	Treatment	154	125 (81)	18 (12)	11 (7)
	Control	29	22 (76)	6 (21)	1 (3)
Monitoring zone	Treatment	57	43 (75)	5 (9)	9 (16)
	Control	15	12 (80)	2 (13)	1 (7)
Refer zone	Treatment	58	29 (50)	11 (19)	18 (31)
	Control	11	7 (64)	3 (27)	1 (9)
FU3					
Low or no risk	Treatment	150	141 (94)	5 (3)	4 (3)
	Control	29	26 (90)	1 (3)	2 (7)
Monitoring zone	Treatment	61	54 (89)	3 (5)	4 (7)
	Control	14	12 (86)	1 (7)	1 (7)
Refer zone	Treatment	57	41 (72)	9 (16)	7 (12)
	Control	10	8 (80)	1 (10)	1 (10)

TABLE 8 Difference between arms for the key secondary outcome: PHQ-9

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)	p-value
	Treatment			Control					
	n (%)	Mean	SD	n (%)	Mean	SD			
Primary caregiver									
Baseline	285 (100)	3.06	3.49	56 (100)	2.84	3.18	0.22 (-0.77 to 1.21)		
FU1	270 (95)	2.37	2.78	55 (98)	3.29	4.27	-0.92 (-1.81 to -0.03)		
FU2	270 (95)	2.44	2.97	55 (98)	2.80	3.60	-0.36 (-1.25 to 0.53)		
FU3	269 (94)	2.90	3.50	53 (95)	2.89	3.66	0.02 (-1.02 to 1.06)		
Overall								-0.61 (-1.34 to 0.12)	0.10
Co-parent									
Baseline	53 (100)	3.17	3.64	15 (100)	3.20	3.84	-0.03 (-2.14 to 2.08)		
FU1	45 (85)	3.53	3.86	12 (80)	2.75	3.82	0.78 (-1.67 to 3.23)		
FU2	41 (77)	3.83	4.89	12 (80)	1.83	1.95	2.00 (-0.84 to 4.84)		
FU3	40 (75)	3.25	3.51	7 (47)	2.14	1.77	1.11 (-1.56 to 3.78)		
Overall								0.15 (-1.09 to 1.40)	0.81

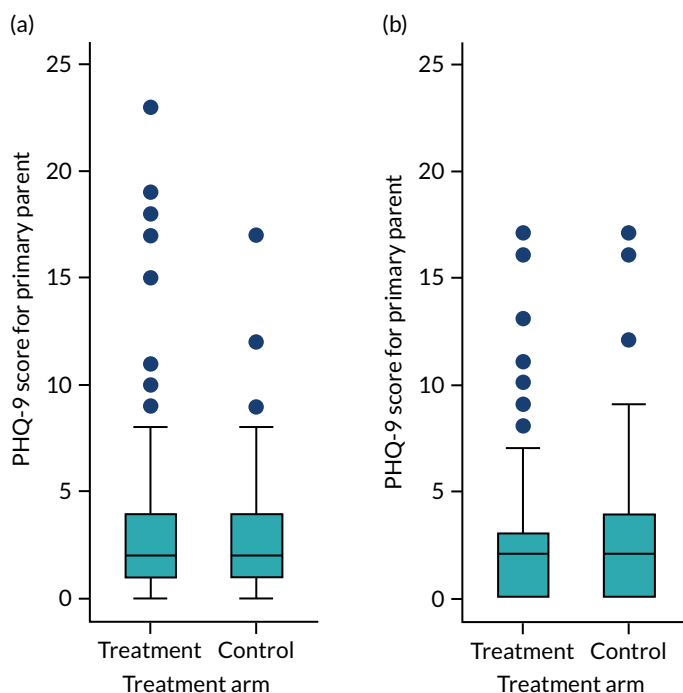


FIGURE 4 Summary of PHQ-9 by treatment arm and follow-up point. (a) Baseline; (b) FU1; (c) FU2; and (d) FU3. (continued)

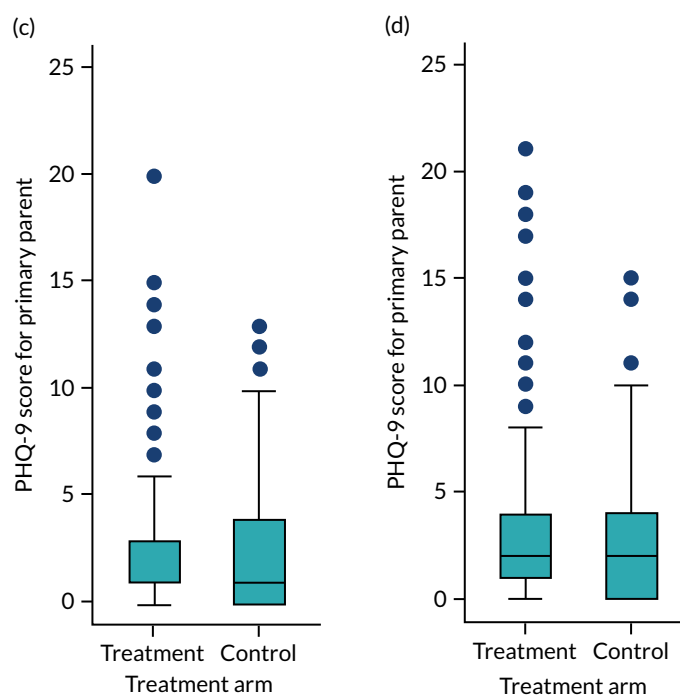


FIGURE 4 Summary of PHQ-9 by treatment arm and follow-up point. (a) Baseline; (b) FU1; (c) FU2; and (d) FU3.

TABLE 9 Summary of effect estimates from all analysis models for key secondary outcome: PHQ-9

Analysis model	Adjusted mean difference (95% CI)	p-value
Primary analysis model	-0.61 (-1.34 to 0.12)	0.100
Primary with unstructured correlation	-0.64 (-1.35 to 0.07)	0.077
Difference from baseline: MI	-0.68 (-1.37 to 0.12)	0.054

Similarly, as for the primary outcome, we undertook an additional unplanned analysis of how parents transitioned between PHQ-9 categories during the intervention. PHQ-9 can allocate parents to five levels of depression. Very few parents were moderately or severely depressed and so the transition analysis was based on three categories: (1) no depression, (2) mild depression and (3) moderate to severe depression. *Table 10* shows the proportion of people in each depression category at follow-up as a proportion of the people in that category at baseline. Consistent with the weakly significant finding in favour of the treatment for PHQ-9, *Table 10* suggests that people in the treatment group were slightly more likely to be categorised as having less extreme depression at follow-up than people in the control group.

### Other secondary outcomes

No secondary outcome measures provided any signal to suggest a difference between arms. First, we measured dyadic synchrony for the primary parent-child dyad at all follow-up points using the Infant CARE-Index (*Table 11*) (for which a higher score indicates higher dyadic synchrony). The number of people available with complete data for the Infant CARE-Index across all four time points was a relatively small subset of the study (i.e. 119 of the 269 participants who consented to video observation).

TABLE 10 Transition between PHQ-9 categories

PHQ-9 category	Treatment arm	N	PHQ-9 category at baseline, n (%)		
			No depression	Mild depression	Moderate or severe depression
FU1					
No depression	Treatment	52	32 (62)	19 (37)	1 (2)
	Control	13	7 (54)	6 (46)	0 (0)
Mild depression	Treatment	156	36 (23)	106 (68)	14 (9)
	Control	29	8 (28)	17 (59)	4 (14)
Moderate or severe depression	Treatment	62	3 (5)	26 (42)	33 (53)
	Control	13	0 (0)	5 (38)	8 (62)
FU2					
No depression	Treatment	56	27 (48)	28 (50)	1 (2)
	Control	13	9 (69)	4 (31)	0 (0)
Mild depression	Treatment	151	35 (23)	95 (63)	21 (14)
	Control	29	9 (31)	17 (59)	3 (10)
Moderate or severe depression	Treatment	63	5 (8)	32 (51)	26 (41)
	Control	13	1 (8)	3 (23)	9 (69)
FU3					
No depression	Treatment	51	25 (49)	24 (47)	2 (4)
	Control	13	8 (62)	5 (38)	0 (0)
Mild depression	Treatment	156	36 (23)	92 (59)	28 (18)
	Control	28	7 (25)	16 (57)	5 (18)
Moderate or severe depression	Treatment	62	5 (8)	30 (48)	27 (44)
	Control	12	2 (17)	4 (33)	6 (50)

TABLE 11 Difference between arms for the Infant CARE-Index

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)
	Treatment			Control				
	n (%)	Mean	SD	n (%)	Mean	SD		
Baseline	105 (100)	5.87	2.48	14 (100)	5.57	2.14	0.30 (-1.06 to 1.66)	
FU1	105 (100)	6.10	2.62	14 (100)	7.79	3.07	-1.69 (-3.18 to -0.20)	
FU2	105 (100)	8.18	2.79	14 (100)	8.07	2.70	0.11 (-1.44 to 1.66)	
FU3	105 (100)	8.33	2.59	14 (100)	7.93	2.50	0.40 (-1.04 to 1.84)	
Overall								-0.25 (-1.09 to 0.59)

Although the results are not statistically significant, with a modest signal in favour of the control overall (-0.25), a similar slight trend can be seen, that is the magnitude of the signal at FU1 (-1.69) is larger than at FU2 (0.11) and FU3 (0.44). Similarly, as for the ASQ:SE-2 and PHQ-9, we undertook an additional unplanned analysis of how parents transitioned between the Infant CARE-Index dyadic synchrony categories during the intervention. The Infant CARE-Index has four levels of dyadic synchrony: (1) sensitive, (2) adequate, (3) inept and (4) risk. *Appendix 9, Table 40*, shows the proportion of people in each category at follow-up as a proportion of the people in that category at baseline. *Appendix 9, Table 40*, suggests that there is no difference between arms with regard to transition through the Infant CARE-Index categories.

Third, two measures were collected at all follow-up points for both parents: (1) parenting skills using the PSOC survey (*Table 12*) and HRQoL using the EQ-5D-5L (*Table 13*).

Fourth, three outcomes were collected from primary and co-parents at FU3 only: (1) child behaviour using the SDQ, (2) cognitive development using the PedsQL and (3) parent-child interaction using the MPAS and PPAS, as appropriate (*Table 14*). Finally, data on patterns in how children were fed were collected to see if the intervention had an impact on the rates and length of breastfeeding, but no obvious differences between arms over time were apparent (see *Report Supplementary Material 3*).

### Summary of intervention take-up

With some exceptions, the IY parenting groups had relatively low take-up, particularly for the IY-T groups, but once a participant attended one session then they typically attended at least half of the remaining sessions, suggesting a very high level of satisfaction once engaged (*Figures 5 and 6* and see *Report Supplementary Material 1*). This is explored more formally in the process evaluation (see *Chapter 4*). There was a relatively large difference between sites in levels of take-up for IY-I, with sites 2 and 3 both having > 50% of eligible parents participating and sites 1 and 4 both having < 50% of eligible parents participating (see *Figure 6* and *Report Supplementary Material 1*). There was less variation between sites in the uptake of toddler groups, with sites 2 and 4 having 30% of eligible parents participating and sites 1

TABLE 12 Difference between arms for the PSOC survey

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)
	Treatment			Control				
	n (%)	Mean	SD	n (%)	Mean	SD		
Primary caregiver								
Baseline	285 (100)	80.94	9.56	56 (100)	80.21	9.76	0.72 (-2.03 to 3.47)	
FU1	270 (95)	82.81	9.66	55 (98)	81.84	10.63	0.97 (-1.88 to 3.82)	
FU2	270 (95)	83.84	8.90	55 (98)	82.69	10.00	1.15 (-1.49 to 3.79)	
FU3	269 (94)	81.88	9.41	53 (95)	82.43	9.70	-0.56 (-3.35 to 2.23)	
Overall								0.07 (-1.74 to 1.87)
Co-parent								
Baseline	53 (100)	79.19	10.24	15 (100)	79.87	11.21	-0.68 (-6.67 to 5.31)	
FU1	45 (85)	78.38	9.28	12 (80)	83.17	10.81	-4.79 (-10.91 to 1.33)	
FU2	41 (77)	79.68	8.74	12 (80)	77.00	10.64	2.68 (-3.23 to 8.59)	
FU3	40 (75)	78.40	9.95	7 (47)	81.43	14.50	-3.03 (-11.60 to 5.54)	
Overall								-1.26 (-5.16 to 2.63)



TABLE 13 Difference between arms for the EQ-5D-5L

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)
	Treatment			Control				
	n (%)	Mean	SD	n (%)	Mean	SD		
Primary caregiver								
Baseline	285 (100)	0.94	0.08	56 (100)	0.92	0.10	0.02 (-0.00 to 0.04)	
FU1	270 (95)	0.95	0.08	55 (98)	0.93	0.10	0.03 (0.01 to 0.05)	
FU2	270 (95)	0.95	0.09	55 (98)	0.92	0.12	0.03 (0.00 to 0.06)	
FU3	269 (94)	0.94	0.10	53 (95)	0.93	0.10	0.01 (-0.02 to 0.04)	
Overall								0.02 (0.00 to 0.04)
Co-parent								
Baseline	52 (98)	0.95	0.08	15 (100)	0.95	0.11	0.00 (-0.05 to 0.05)	
FU1	45 (85)	0.93	0.12	12 (80)	0.96	0.07	-0.03 (-0.10 to 0.04)	
FU2	41 (77)	0.93	0.11	12 (80)	0.96	0.08	-0.03 (-0.10 to 0.04)	
FU3	40 (75)	0.94	0.10	7 (47)	0.98	0.03	-0.04 (-0.11 to 0.03)	
Overall								-0.01 (-0.05 to 0.02)

TABLE 14 Difference between arms for outcomes at 18 months (FU3) only

Outcome	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)
	Treatment			Control				
	n (%)	Mean	SD	n (%)	Mean	SD		
Primary caregiver								
SDQ	266 (99)	9.67	4.27	53 (100)	9.15	4.53	0.52 (-0.75 to 1.79)	0.64 (-0.64 to 1.91)
PedsQL	268 (100)	86.70	9.61	53 (100)	87.31	9.15	-0.61 (-3.42 to 2.20)	-0.60 (-3.22 to 2.01)
MPAS	257 (96)	84.61	5.82	52 (98)	84.27	7.15	0.34 (-1.47 to 2.15)	0.94 (-0.76 to 2.64)
Co-parent								
SDQ	39 (98)	9.26	4.29	7 (100)	8.71	3.35	0.54 (-2.82 to 3.90)	0.10 (-3.56 to 3.76)
PedsQL	40 (100)	83.63	11.21	7 (100)	91.43	6.28	-7.80 (-16.38 to 0.78)	-6.28 (-15.98 to 3.41)
PPAS	38 (95)	78.15	6.67	7 (100)	78.81	9.42	-0.67 (-6.41 to 5.07)	1.84 (-3.85 to 7.54)

and 3 having 10% of eligible parents participating. This echoes the finding of a significant interaction between treatment and site allocation. Therefore, it may be appropriate to undertake further analysis of outcomes for the sites with higher uptake using causal methods applicable to smaller sample sizes.

### Characteristics of compliers and non-compliers

We expected 151 participants to be eligible for IY-I, with 50 attending, and 147 participants to be eligible for IY-T, with 48 attending. We achieved 51 participants for IY-I and 21 participants for IY-T.

OUTCOME EVALUATION RESULTS

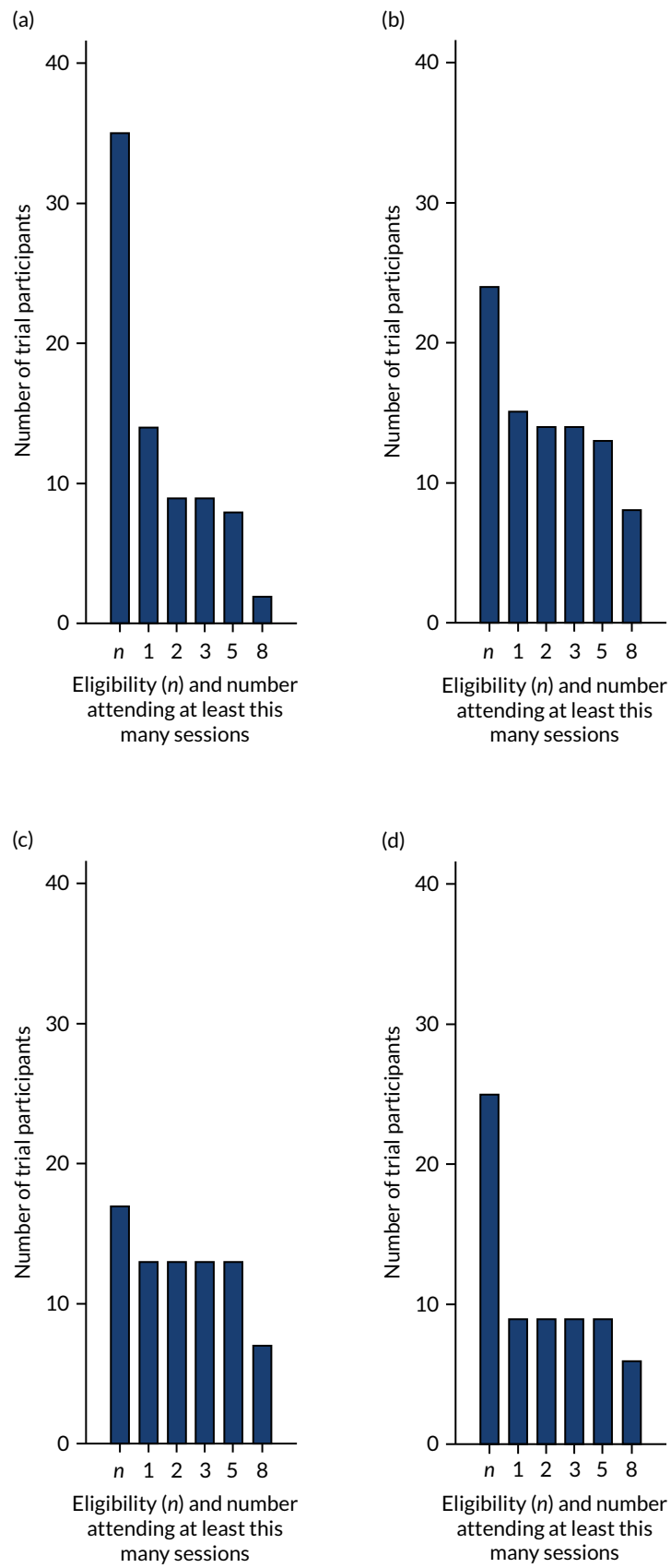


FIGURE 5 Eligibility, compliance and session attendance: IY-I. (a) Site 1; (b) site 2; (c) site 3; and (d) site 4.

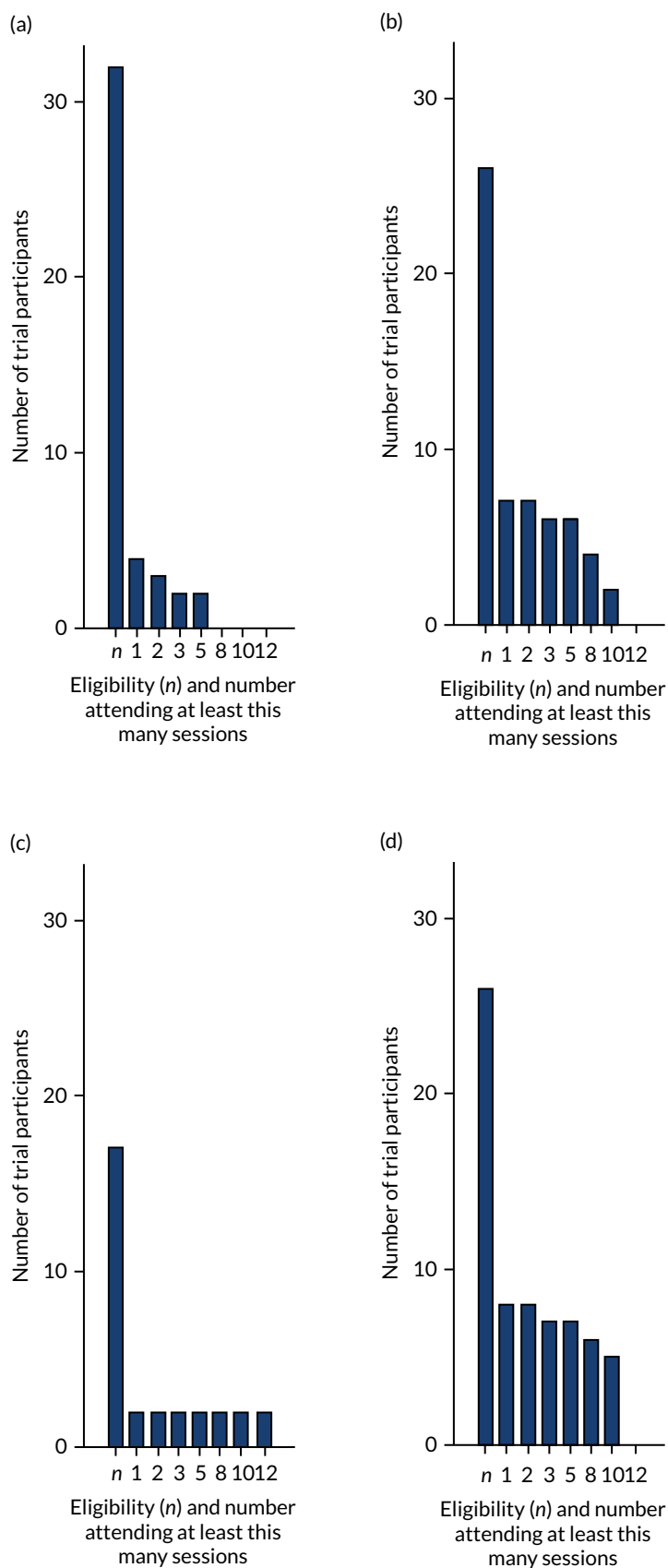


FIGURE 6 Eligibility, compliance and session attendance: IY-T. (a) Site 1; (b) site 2; (c) site 3; and (d) site 4.

To explore the relatively low take-up further, we compared the characteristics of compliers with non-compliers, defining compliance as ‘attended at least one session’ and non-compliance as being ‘eligible but attending no sessions’. This was a relatively low threshold; however, in the absence of any empirical or theoretically based threshold for complete service receipt, compliance was treated only as an indicator of initial intervention engagement and, as mentioned above (see *Summary of intervention take-up*), participants who attended one session were likely to attend at least half of the subsequent sessions. This analysis is descriptive to avoid inflating the type I error by multiple testing. *Table 15* compares compliers with non-compliers in terms of baseline characteristics, including ASQ:SE-2 and PHQ-9 scores (using FU1 for IY-I and FU2 for IY-T). There appears to be little difference between compliers and non-compliers overall, but there are a few exceptions. Parents educated to higher level and in the higher-income bands were marginally more likely to take up the intervention. Dyads with ASQ:SE-2 child development scores below the monitoring zone were more likely to take up IY-T sessions. This result should be viewed with caution because of the small numbers of participants who took up IY-T overall. Dyads with lower PHQ-9 parental depression scores and ASQ:SE-2 child development scores below the monitoring zone were more likely to take up IY-I sessions. The lower take-up by people with more severe baseline scores is of particular interest, as it suggests that the intervention may be less appealing to people who are most likely to benefit.

TABLE 15 Comparing characteristics of compliers and non-compliers: IY-I and IY-T

Compliance	Non-complier, n (%)	Complier, n (%)
<b>IY-I</b>		
ASQ:SE-2 (development) FU1		
Below monitoring zone (n = 22)	12 (55)	10 (45)
In monitoring zone (n = 17)	7 (41)	10 (59)
In refer zone (n = 58)	28 (48)	30 (52)
PHQ-9 (depression) FU1		
None (n = 53)	23 (43)	30 (57)
Mild (n = 37)	22 (59)	15 (41)
Moderate and moderately severe (n = 11)	5 (45)	6 (55)
Whether or not qualified to a degree level		
Below degree level (n = 37)	22 (59)	15 (41)
Degree level or higher (n = 64)	28 (44)	36 (56)
Quartile of IY-I eligible participants' weekly income distribution		
Lowest quartile (< £360) (n = 20)	11 (55)	9 (45)
Second quartile (between £360 and £600) (n = 25)	14 (56)	11 (44)
Third quartile (between £600 and £900) (n = 16)	8 (50)	8 (50)
Highest quartile (> £900) (n = 19)	6 (32)	13 (68)
Income not given (n = 21)	11 (52)	10 (48)
Whether or not first child		
Not first child (n = 56)	27 (48)	29 (52)
First child (n = 45)	23 (51)	22 (49)

TABLE 15 Comparing characteristics of compliers and non-compliers: IY-I and IY-T (continued)

Compliance	Non-complier, n (%)	Complier, n (%)
Sex of child		
Male (n = 55)	30 (55)	25 (45)
Female (n = 46)	20 (43)	26 (57)
Site		
3 (n = 17)	4 (24)	13 (76)
2 (n = 24)	9 (38)	15 (63)
4 (n = 25)	16 (64)	9 (36)
1 (n = 35)	21 (60)	14 (40)
Relationship status		
Married and living together (n = 65)	31 (48)	34 (52)
Cohabiting/living together (n = 20)	11 (55)	9 (45)
Other type of relationship (n = 8)	4 (50)	4 (50)
Not in a relationship or separated (n = 8)	4 (50)	4 (50)
Parent age group (years)		
18–21 (n = 6)	4 (67)	2 (33)
22–25 (n = 17)	10 (59)	7 (41)
26–30 (n = 23)	11 (48)	12 (52)
31–35 (n = 35)	20 (57)	15 (43)
≥ 36 (n = 20)	5 (25)	15 (75)
Ethnicity		
English/Welsh/Scottish/Northern Irish/British/Irish (n = 73)	32 (44)	41 (56)
Any other white background (n = 4)	3 (75)	1 (25)
Indian (n = 9)	7 (78)	2 (22)
Pakistani (n = 6)	3 (50)	3 (50)
Any other ethnic group (n = 9)	5 (56)	4 (44)
<b>IY-T</b>		
ASQ:SE-2 (development) FU2		
Below monitoring zone (n = 29)	20 (69)	9 (31)
In monitoring zone (n = 34)	30 (88)	4 (12)
In refer zone (n = 38)	30 (79)	8 (21)
PHQ-9 (depression) FU2		
None (n = 53)	45 (85)	8 (15)
Mild (n = 39)	30 (77)	9 (23)
Moderate and moderately severe (n = 8)	5 (63)	3 (38)
Severe (n = 1)	0 (0)	1 (100)
Whether or not qualified to a degree level		
Below degree level (n = 38)	28 (74)	10 (26)
Degree level or higher (n = 63)	52 (83)	11 (17)
continued		

## OUTCOME EVALUATION RESULTS

TABLE 15 Comparing characteristics of compliers and non-compliers: IY-I and IY-T (continued)

Compliance	Non-complier, n (%)	Complier, n (%)
Quartile of IY-T eligible participants' weekly income distribution		
Lowest quartile (< £450) (n = 20)	14 (70)	6 (30)
Second quartile (between £450 and £610) (n = 20)	16 (80)	4 (20)
Third quartile (between £610 and £1000) (n = 21)	19 (90)	2 (10)
Highest quartile (> £1000) (n = 18)	14 (78)	4 (22)
Income not given (n = 22)	17 (77)	5 (23)
Whether or not first child		
Not first child (n = 57)	44 (77)	13 (23)
First child (n = 44)	36 (82)	8 (18)
Sex of child		
Male (n = 59)	45 (76)	14 (24)
Female (n = 42)	35 (83)	7 (17)
Site		
3 (n = 17)	15 (88)	2 (12)
2 (n = 26)	19 (73)	7 (27)
4 (n = 26)	18 (69)	8 (31)
1 (n = 32)	28 (88)	4 (13)
Relationship status		
Married and living together (n = 65)	53 (82)	12 (18)
Cohabiting/living together (n = 22)	17 (77)	5 (23)
Other type of relationship (n = 7)	6 (86)	1 (14)
Not in a relationship or separated (n = 7)	4 (57)	3 (43)
Parent's age group (years)		
18–21 (n = 4)	2 (50)	2 (50)
22–25 (n = 15)	11 (73)	4 (27)
26–30 (n = 26)	23 (88)	3 (12)
31–35 (n = 32)	26 (81)	6 (19)
≥ 36 (n = 24)	18 (75)	6 (25)
Ethnicity		
English/Welsh/Scottish/Northern Irish/British/Irish (n = 81)	64 (79)	17 (21)
Any other white background (n = 0)	0	0
Indian (n = 7)	6 (86)	1 (14)
Pakistani (n = 6)	5 (83)	1 (17)
Any other ethnic group (n = 7)	5 (71)	2 (29)

## Overall summary of outcomes

The primary analysis of the primary outcome (i.e. the social and emotional well-being of the child) was borderline statistically significant in favour of the control group. Sensitivity analyses of the primary outcome analysis diminished the strength of the signal. The MCID of the primary outcome stated in the sample size calculation was 5 units (i.e. -5 units, as we expected those receiving the E-SEE Steps model to have lower scores). Therefore, we have strong evidence that the proportionate intervention E-SEE Steps is not clinically effective, as -5 was well outside of the 95% CIs for both the primary and sensitivity analyses. The primary analysis was a repeated-measure analysis over the three follow-up time points, adjusted for baseline. When the three time points were analysed separately (with each adjusted for baseline), the strongest signal was seen at FU1 in favour of SAU, which was the measurement following the universal dose of the IY-B.

By contrast, the analysis of the key secondary outcome (i.e. depression levels among primary parent) provided a moderate signal in favour of the treatment. Considering each time point separately for depression of the primary parent, the strongest signal was, again, seen at FU1.

All of the 95% CIs for the primary analyses and associated sensitivity analyses excluded -5. Therefore, although we did not achieve the target sample size, the high retention rate to the trial resulted in robust evidence that the E-SEE Steps model is not clinically effective in enhancing social and emotional well-being. However, it is worth noting that take-up of IY-T was substantially lower than expected and this will have had an impact on its ability to have an effect overall.





## Chapter 4 Process evaluation

Parts of this chapter have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Introduction

The overarching aim of the process evaluation was to contextualise the main trial's effectiveness and cost-effectiveness findings in terms of why and how the intervention (i.e. the E-SEE Steps model) may or may not have worked. The study was focused on understanding the context and processes supporting the model as a whole, rather than the individual programme components/levels (i.e. IY-I and IY-T). For the pilot trial findings, see Blower *et al.*<sup>36</sup>

The process evaluation had three main aims:

1. assess the uptake and fidelity of the intervention when delivered by public sector services in the UK
2. identify the factors associated with successful implementation of the intervention in the UK
3. assess the acceptability and feasibility of the intervention for delivery in public services in the UK.

These aims were addressed via the following research questions:

- Can a multiagency service deliver the E-SEE Steps model (i.e. IY) in a proportionate universal model?
- What level of fidelity is achieved at each level of the intervention?
- What are the organisational- or systems-level barriers to and facilitators of delivering in this way?
- How acceptable and feasible is delivery of the E-SEE Steps model for key intervention stakeholders?

### Method

We used a multimethod approach to explore these process evaluation questions, embedded within the main processes of the trial. In addition, a qualitative substudy (see *Appendix 10*) explored the views of fathers as co-parents, which supplemented findings from the process evaluation.

#### Sampling

##### Group leaders

All staff who were selected by their site to be trained in and/or deliver the E-SEE Steps intervention were invited to take part in the process evaluation. Staff were provided with an information sheet and asked to sign a consent form. All staff selected by sites as potential group leaders ( $n = 94$ ) were asked to complete a pre-training online questionnaire. Group leaders who went on to deliver the groups ( $n = 29$ ) were also asked to complete a post-delivery questionnaire and were invited to attend a focus group, which explored their views on the set-up, delivery and success of the programme in their site. Nineteen group leaders participated in these focus groups.

##### Service managers

Two service managers in each site were selected from the health and children's services partner organisations that had been involved in designing and delivering the E-SEE Steps model to participate in the process evaluation. Service managers were provided with an information sheet and asked to sign a consent form.

## Parents

All parents who were eligible and attended at least one of the group-based intervention sessions were invited to take part in a focus group on completion of the group. We also used purposive sampling to select a group of parents for telephone interviews. These parents were eligible for the groups, but declined the offer or did not attend. From the 98 eligible parents, 12 (i.e. three from each site) were sampled purposively.

## Data collection

Data collection for the process evaluation comprised both quantitative and qualitative elements to assess fidelity, feasibility and acceptability. Research questions and data sources are summarised in *Table 16*.

TABLE 16 Summary of questions and data sources for the process evaluation

Research question	Dimension	Data source/measure
Can a multiagency service deliver the E-SEE Steps model (i.e. IY) in a proportionate universal model?	Service design	Site-level decisions about service design
	Selection of group leaders	Demographics of group leaders: pre- and post-delivery questionnaires
What level of fidelity is achieved at each level of the intervention?	Eligibility	Proportion of intervention and control families who meet the criteria for IY-I and IY-T groups, respectively
	Uptake/reach	Proportion of intervention families who accept a group place offer
	Fidelity of delivery (adherence and quality)	Group leader-completed adherence checklists
		Observations of delivery by PPIC (e.g. adherence, dose/exposure, quality of delivery, participant responsiveness)
	Dose/exposure	Session length
		Group attendance registers
Participant responsiveness	Weekly participant contact sheets	
	Parent/co-parent satisfaction: weekly and end-of-programme evaluations	
What are the organisational- or systems-level barriers to and facilitators of delivering in this way?	Barriers to/facilitators of delivery	Group leader focus groups
		Parent focus groups
		Service manager interviews
		Interviews with non-attending parents
		Interviews with IY mentors
		Group leader post-delivery questionnaire
How acceptable and feasible is delivery of the E-SEE Steps model for key intervention stakeholders?	Acceptability of E-SEE Steps model delivery	Group leader focus groups
		Parent focus groups
	Feasibility of the E-SEE Steps model for future implementation and scale-up	Service manager interviews
		Interviews with non-attending parents
		Interviews with IY mentors
		End-of-programme parent/co-parent evaluations

PPIC, parent programme implementation checklist.

### Quantitative measures

Attendance by parents and co-parents was recorded by group leaders in a weekly attendance log, which included space to record reasons for non-attendance, whether or not a catch-up session had been delivered and how many crèche places and interpreters were required. Contact with parents outside group sessions was recorded by group leaders on a contact sheet, weekly and by each type of contact made (e.g. home visit, telephone call, text and letter).

Group leaders' self-reported adherence to core components was assessed using the standard weekly-completed IY checklists that correspond with the components set out in the respective programme manuals. Adherence to an average of 80% of the content is generally considered acceptable fidelity. The quality of delivery was assessed using the parenting programme implementation checklist (PPIC),<sup>53</sup> an 18-item tool developed to appraise aspects of programme implementation fidelity by independent observation. The checklist has indices for adherence, dose/exposure, quality of delivery and participant responsiveness. All IY group sessions were video-recorded and a random subset of both IY-I and IY-T programmes were used as observational data. One video from each group was coded, 14 in total (IY-I,  $n = 8$ ; IY-T,  $n = 6$ ). Where possible, the same session was coded for each group. Exceptions to this included where the video was not available because of technical issues during delivery or where there were too few participants, in which case the session with greatest attendance was chosen for coding. For IY-I, session 4 was coded. For IY-T, session 6 was coded. All videos were coded by two researchers.

Parent satisfaction was assessed using standard IY satisfaction questionnaires, adapted for use with parents with low literacy and for UK English language (originally US English language). The weekly questionnaire was completed after each session and at the end of the 10- and 12-week programme, respectively. Group leader data and feedback were obtained using online questionnaires before training and again immediately after completing delivery of the IY programme. The pre-training questionnaire gathered information on the group leaders' demographics, qualifications and existing experiences of working with children and families, whereas the post-delivery questionnaire was used to supplement the qualitative data on group leaders' views of delivering IY-I and/or IY-T.

### Qualitative methods

To avoid influencing the impact of the intervention, focus groups and interviews were undertaken once intervention delivery was complete at each site. A total of eight focus groups were conducted (four focus groups with group leaders from each site and four focus groups with a selection of parents and co-parents across the sites). In site 1, an interview was conducted instead of a focus group because of only one participant attending. The focus groups/interviews explored:

- the acceptability and usefulness of the IY-B as a universal intervention
- the acceptability of a proportionate model with stepped intervention
- the processes for identification, screening and recruitment
- the strategies/approaches for engaging fathers and extended carers
- barriers to and facilitators of attendance
- experiences of participation in the groups.

A total of seven semistructured interviews with health and children's services managers were conducted. In addition, interviews were held with three IY trainers and/or mentors. All interviews were approximately 30 minutes in duration and conducted largely via telephone. The interviews explored managers' and mentors' views on the accommodations required or adaptations made to the service to enable delivery of the IY intervention, and the particular system-, organisational- and team-level barriers to and facilitators of delivery in their locality.

Eleven telephone interviews were conducted with parents who were eligible for the groups, but who declined the offer or did not attend. Interviews explored why parents did not engage or discontinued

engagement and will help to inform engagement practices in future parenting interventions. All interviews were approximately 15–20 minutes in duration. Interviews were audio-recorded and transcribed. All data were anonymised and any comments or observations relating to specific individuals that could allow them to be recognised were removed.

### *Data analysis*

#### **Quantitative data analysis**

The attendance log and contact sheet and weekly and end-of-programme parent satisfaction forms were analysed in SPSS 25 (IBM Corporation, Armonk, NY, USA) and reported descriptively using means or medians, percentages and measures of variability. Descriptive data were also reported for the pre-training and post-delivery questionnaires completed by group leaders. Observational data from the sample of video-recorded sessions were observed and coded, and average scores for indices were collated for each group. Inter-rater reliability was assessed using percentage agreement. Data from the session adherence checklists were reported using descriptive statistics, as above.

#### **Qualitative data analysis**

Data were analysed by thematic analysis in NVivo 11 (QSR International, Warrington, UK), using a deductive/theoretical approach.<sup>54</sup> The process evaluation adopted a multilevel framework to synthesise the interview and focus group data gathered about the factors that impede or facilitate successful delivery of the programme in different contexts (research question 1). The aim of the analysis was twofold: (1) to understand (individual-, organisational- or systems-level) barriers to and facilitators of fidelity within a multiagency delivery of IY in a proportionate universalism model; and (2) to understand the acceptability and feasibility of delivering the E-SEE Steps model for key intervention stakeholders.

Interviews were read and re-read, and data were coded using a semantic approach.<sup>54</sup> Data were organised to explore patterns in semantic content within and across data sources in relation to the research questions. Coding was progressively refined, with the researcher moving between codes, accounts, thematic maps and analytic memos and notes, which were written throughout the process to explore emergent thoughts, ideas, concepts and themes. In the final stage, the analytic process progressed from description to interpretation and culminated in a defined set of themes relating to each research question. Themes were reviewed and defined, and the significance of the patterns and their broader meanings and implications theorised in relation to the research questions.<sup>55</sup>

## **Findings**

### *Aim 1*

#### **Research question 1: designing and delivering a multiagency intervention**

##### ***Service design meetings***

The E-SEE research team provided service design support to sites in setting up the targeted group-based programmes in their areas. Senior managers and key delivery stakeholders were invited to attend service design meetings. Sites differed in their ability to provide the support known to increase engagement, access and retention to parenting interventions, including the provision of child care, transport support and preparation to ensure leaders were proactive and confident in the materials (Table 17). Suggested leader preparation included undertaking dry runs of delivery to increase familiarity with the materials and processes. Only one site managed to achieve this under the direction of a colleague who was experienced in IY.

TABLE 17 Service design decisions

Site	Programme	Number of groups	Number of practitioners trained	Co-delivery?	Lead where not co-delivery	Transport and/or crèche provided	Home visits before groups started?
1	IY-I	2	13	Yes		Crèche and transport	Yes (n = 4)
2	IY-I	2	16	No	Health	Crèche	Yes (n = 6)
3	IY-I	2	17	Yes		Crèche	Yes (n = 1)
4	IY-I	2	10	Yes		Crèche and transport	No
1	IY-T	2	9	Yes		Crèche and transport	Yes (n = 7)
2	IY-T	2	10	No	Health	Crèche	No
3	IY-T	1	11	No	Children's services	Crèche	No
4	IY-T	1	8	Yes		Crèche and transport	No

Five groups across two sites were delivered by a single agency in the main trial because of the need for flexibility on delivery in context, informed by the pilot findings. Health professionals' time for training and delivery was paid out of NHS excess treatment costs, whereas no budget support was provided to children's services staff. The number of groups delivered in each site was smaller than the number intended.

### Selecting group leaders

**Demographics and experience of site practitioners** Ninety-four practitioners were identified by sites to be trained as prospective group leaders, based on their role/remit for supporting parents of children under 2 years of age, experience in parenting programmes and group facilitation, interest in being involved in the E-SEE trial, and capacity to prepare and deliver groups over a full term. Twenty-nine (31%) of these practitioners went on to deliver the IY-I and IY-T groups. Of the practitioners who consented to completing a pre-training research questionnaire (n = 50), all were female, 86% (n = 43) were white British and 14% (n = 7) identified as British 'other'. Forty per cent of practitioners were aged ≥ 46 years (36% were aged between 31 and 45 years and 24% were aged between 18 and 30 years). Fifty-six per cent (n = 28) of practitioners worked full time and 44% worked part-time, with hours ranging between 15 and 32 per week. All practitioners who completed a questionnaire held a qualification [i.e. an undergraduate degree (n = 21, 42%), a level 3 qualification (e.g. Qualifications and Credit Framework, National Vocational Qualification, National Nursery Examination Board or Advanced Level) (n = 13, 26%), a Business and Technology Education Council qualification (n = 6, 12%), a pre-graduate diploma (n = 5, 10%), a post-graduate diploma (n = 4, 8%) or a master's qualification (n = 1, 2%)].

Practitioners (n = 26) selected by sites to deliver the E-SEE Steps groups were more likely to work full time and to have previously delivered parenting interventions, including the IY programme. Perhaps unsurprisingly, 42% of those practitioners who went on to deliver the trial groups also reported feeling 'confident' prior to being trained in setting up and running a parenting group compared with the same proportion (42%) of practitioners who reported feeling only 'somewhat confident' and also went on to deliver the trial groups. *Report Supplementary Material 1* presents differences between practitioners who were trained and those who went on to deliver the trial groups.

**Training and supervision** Training for IY-I and IY-T was delivered to each site, on average, 2–3 (range 1–11) months before delivering the groups. Ninety-four staff were trained in 10 separate training sessions. Of this total, 29 went on to deliver at least one session of the trial groups. Supervision sessions were carried out remotely via teleconference, with the aim to review goals, questions and video clips. Organisation firewalls prevented the sharing of video footage of delivery for review by the mentor for most of the supervisory groups, thereby relying on leaders' self-perceptions of their strengths and weaknesses. *Report Supplementary Material 1* provides site-level details of the training and supervision provided. Feasibility of the supervision model for future implementation is explored in the themes below.

**Site-level services as usual overview** Services as usual were collated for each site for parenting support programmes spanning behaviour management, healthy weight and nutrition, early learning and development, postnatal support and support for children with additional needs. This is important, given the outcome findings of no difference between the E-SEE Step model and SAU.

Across sites, there were a larger number of services and support targeted at children in the early years (i.e. 0–5 years), although there was a proportion of programmes available to all ages. In terms of more targeted provision, this was for ages 0–18 months and 0–5 years across all sites. Sites 2 and 4 had the lowest proportion of programmes focusing on the 0- to 18-month period, and site 3 had the highest proportion of services on offer for this age group, with fewer general/any-age programmes. *Table 18* shows the distribution of programmes across age groups.

In site 1, the majority of parenting support was focused on early learning and development. Evidence-based programmes focusing on behaviour management (e.g. the Strengthening Families programme) and healthy weight and nutrition [e.g. the Health Exercise and Nutrition for the Really Young (HENRY) programme] were available to parents. Early learning and development services included messy play and story sessions, but no evidence-based programmes were offered.

In site 2, SAU involved evidence-based programmes, including the Solihull Approach, Triple P, Stepping Stones and IY-I, with IY-I as the only universally offered programme. Alongside these services, social toddler workshops and play and sensory sessions were delivered.

Services as usual in site 3 were the widest ranging of the four sites, with a total of 61 programmes on offer. A large proportion of these programmes were focused on the 0- to 18-month stage, including Amazing Babies, Mums in Mind, Baby and Me and IY-I. Both IY-I and IY-T were part of targeted SAU. However, IY was not delivered widely and an agreement was made between the research team and site that the research would be conducted in areas of their site where IY-I and IY-T were not currently offered to reduce any risk of contamination.

TABLE 18 Distribution of parenting support programmes by age group across sites

Site	Age group (%)		
	0–18 months	0–5 years	General/all ages
1	26.9	30.8	42.3
2	15	45	40
3	37.7	50.8	11.5
4	13.3	26.7	60

Site 4 had the smallest number of SAU focusing on 0- to 18-month stage, with a larger number of programmes available for a broader age range. SAU comprised several evidence-based programmes, including Triple P, Stepping Stones, Family Transitions and Solihull Approach, and these were offered as part of targeted services. Other services included baby massage, sensory groups and play groups.

Chapter 5 explores SAU further as part of an economic evaluation. According to parent reports on the CSRI, only four control families and three intervention families participated in additional parenting programmes listed on the CSRI. The attended programmes were Triple P, Solihull Approach and Home Start (see *Report Supplementary Material 4* for a detailed breakdown of SAU accessed by study participants).

## Research question 2: fidelity of delivery

### Universal level: the Incredible Babies book

A minimum of 82 per cent (82.1%) of books were received by families according to postal tracking numbers. A mailroom error led to missing tracking numbers for the remaining participants across three sites. This error was rectified by asking parents during a scheduled call if they had received a book.

### Targeted level: group programmes

**Eligibility and attendance** The percentage of parents at each site accepting an offer of a place on the group intervention varied across programmes and between sites (range 29–88%). *Table 19* illustrates the comparison between eligibility and take-up across sites. Those parents marked as ‘declined’ were sometimes uncontactable. On average, three attempts were made by group leaders to contact parents via telephone, with an average of two answerphone messages left by group leaders. If answerphone messages were not responded to, then one follow-up text was sent. Letters and home visits were occasionally used when group leaders could not contact the parent by other means.

Reasons provided by parents for not taking up a place in the group differed slightly for IY-I and IY-T groups. For IY-I, reasons included not wanting to interrupt the baby’s sleep pattern and not wanting to attend a group programme. For IY-T, key reasons were work commitments, perceiving no need of the course and having other groups to attend. For both IY-I and IY-T, length and duration of sessions and child care concerns were raised.

TABLE 19 Eligibility and take-up for IY-I and IY-T groups

Site	Programme	Number eligible (% of intervention arm)	Number accepted (% of those eligible)	Number 'declined'	Reasons for decline, n	
					Uncontactable	Other
1	IY-I	35/105 (33)	20 (57)	15	11	4
2	IY-I	24/71 (34)	21 (87.5)	3	1	2
3	IY-I	17/50 (34)	13 (76.5)	4	2	2
4	IY-I	25/59 (42)	11 (44)	14	10	4
1	IY-T	32/105 (30)	12 (37.5)	20	5	15
2	IY-T	26/71 (37)	15 (57.7)	11	2	9
3	IY-T	17/50 (34)	5 (29.4)	12	4	8
4	IY-T	26/59 (44)	16 (61.5)	10	3	7
Total		202	113 (55.9)	89	38	51



**Group attendance** The IY-I groups are designed as eight manualised sessions to be delivered over 10 weeks (the extra 2 weeks allow more time for higher-need vulnerable parents); however, none of the trial IY-I groups utilised the tenth week. Table 20 provides IY-I attendance data. Site 1b finished in seven sessions, as only two parents were available for the majority of sessions. Average individual attendance ranged from 3.3 sessions in site 1b to 8.1 sessions in site 4a. Overall, 15 parents who had accepted a place on an IY-I group did not attend the first session. This drop out is attributable to two sites (sites 1 and 2) in particular. Attendance declined by one-fifth after the first week. However, after this initial drop, rates were largely stable.

Table 21 provides IY-T attendance data. Less than half of the parents who accepted a place attended the first session, and this was particularly noticeable for site 1b where none of the research parents attended any of the sessions. Of those parents who did attend week 1, there was relatively stable retention for the remaining weeks of the intervention, with a small drop appearing in week 3 and very few parents attending the final week of delivery. The average individual attendance of research parents and co-parents ranged from just two sessions in site 1b to more than nine sessions in site 3.

### **Adherence to process and content**

Most groups delivered between 75% and 100% of the intended sessions and six out of eight IY-I groups reported group sizes in line with study requirements. By comparison, only one out of six IY-T groups was a viable size. Weekly parent contact was adhered to across both programmes and the IY-I groups reported strong adherence to the checklists, where six of the eight groups delivered met the 80% threshold. Many IY group leaders ( $n = 18$ , 46.2%) reported regularly utilising and reviewing weekly fidelity data, six IY group leaders reported using these data occasionally and the rest ( $n = 15$ , 38%) said that they seldom or never reviewed the weekly fidelity data.

TABLE 20 Attendance at IY-I groups

Site	Number accepted	Week number (n)										Average individual attendance (n)
		1	2	3	4	5	6	7	8	9	10	
1a	13 (1)	6 (1)	5 (1)	4	5 (1)	5 (1)	3 (1)	5	4	6		5.3
1b	7	5	4	3	2	<sup>a</sup>	2	2	2			3.3
2a	21	10 <sup>b</sup>	4	7 <sup>b</sup>	10 <sup>b</sup>	9 <sup>b</sup>	11 <sup>b</sup>	10 <sup>b</sup>	9 <sup>b</sup>	10 <sup>b</sup>		5.8
2b		5	4	3	5	5	5	4	5	5		6.8
3a	13 (2)	6 (1)	6 (1)	5 (1)	5 (1)	4 (1)	5 (1)	5 (1)	6 (1)	3 (1)		7.7
3b		7 (1)	6	5 (1)	6	5 (1)	7 (1)	5 (1)	7	7 (1)		7.6
4a	11 (5)	9 <sup>b</sup> (1)	7 <sup>b</sup> (1)	9 <sup>b</sup> (1)	9 <sup>b</sup> (1)	6 <sup>b</sup> (1)	9 <sup>b</sup> (1)	7 <sup>b</sup> (1)	7 <sup>b</sup> (1)	9 <sup>b</sup> (1)		8.1
4b		2 (4)	2 (4)	0 (3)	2 (3)	2 (4)	1 (3)	0 (2)	2 (4)	2 (4)		7.3
Total	65 (8)	50 (8)	38 (7)	36 (6)	44 (6)	36 (8)	43 (7)	38 (5)	42 (6)	42 (7)	0	

<sup>a</sup> Only one parent could attend and so the session was cancelled.

<sup>b</sup> Includes research co-parents: two in site 2a and in site 4a.

#### **Notes**

Non-research parents included in brackets.

Number of parents who attended at least one session: site 1a,  $n = 9$ ; site 1b,  $n = 6$ ; site 2a,  $n = 12$  (includes two co-parents); site 2b,  $n = 6$ ; site 3a,  $n = 7$ ; site 3b,  $n = 8$ ; site 4a,  $n = 10$  (includes two co-parents); site 4b,  $n = 6$ . Site 1b is a second group in site 1, but was part of cohort 2 (whereas site 1a was the same site but part of cohort 1). In all other instances where a/b are used, it refers to the second group for that site (i.e. all other site groups ran within the same cohorts).



TABLE 21 Attendance at IY-T groups

Site	Number accepted	Week number (n)												Average individual attendance (n)
		1	2	3	4	5	6	7	8	9	10	11	12	
1a	6	4 <sup>a</sup>	3 <sup>a</sup>	3 <sup>a</sup>	1 <sup>a</sup>	2	1	2	0	1				4.3
1b	6 (1)	0 (1)	1 (1)	0 (1)	0									2
2a	15 (3)	4	6 <sup>a</sup>	4	4 <sup>a</sup>	2	4 <sup>a</sup>	2	5 <sup>a</sup>	3	3	4	4	7.5
2b		2 (1)	1 (3)	1 (1)	0	2 (2)	0 (1)	2 (2)	2 (2)	2 (2)	2 (1)	0		6.6
3	5 (3)	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	2 (2)	2 (1)	2 (2)	2 (2)	2 (2)	2 (2)	9.4
4	16 (1)	9 <sup>a</sup> (1)	7 <sup>a</sup> (1)	6 <sup>a</sup> (1)	7 (1)	8 <sup>a</sup> (1)	5 (1)	7 <sup>a</sup> (1)	6	7 <sup>a</sup>	5 (1)	7 (1)		8.3
Total	48 (8)	21 (5)	20 (7)	16 (5)	14 (3)	16 (5)	12 (4)	15 (5)	15 (3)	15 (4)	12 (4)	13 (3)	6 (2)	

a Includes research co-parents: one in site 1a, one in site 2a and two in site 4.

#### Note

Non-research parents included in brackets.

Only a single non-research parent attended the site 1b group on weeks 1 and 3, with no research parents in attendance. The group was stopped after week 4 because of limited research parent attendance. Number of parents who attended at least one session (includes non-research): site 1a,  $n = 4$  (includes one co-parent); site 1b,  $n = 2$ ; site 2a,  $n = 6$  (includes one co-parent); site 2b,  $n = 5$ ; site 3,  $n = 5$ ; site 4,  $n = 10$  (includes two co-parents). Site 1b is a second group in that site but was part of cohort 2 (whereas site 1a was in the same site but part of cohort 1) and this is why this one site is dealt with differently. In all other instances where a/b are used, it refers to the second group for that site (i.e. all other site groups ran within the same cohorts).

**Group leader contact with parents** IY delivery guidance states that group leaders should contact parents weekly, between each group session, to check in on homework tasks and to increase motivation, by telephone call, home visit, individual session, text, letter or a combination of these methods. The E-SEE Step group leaders recorded their weekly contacts and mainly adhered to IY contact guidance (see *Report Supplementary Material 1*). For IY-I retention, contacts ranged from zero to nine contacts per parent, with a mean of 4.5 contacts per parent across all groups. IY-T retention contact ranged from 3 to 12 contacts per parent, with a mean of 7.2 contacts per parent across all groups.

**Group leader checklists** One measure of programme fidelity was the group leaders' adherence to the recommended vignettes or the number of vignettes in each session and the actions specified in the 'Did I' section of the weekly programme self-complete checklists (*Table 22*). For IY-I, the checklists indicate which of the programme's vignettes are recommended as core by the developer. The IY-T checklists do not specify core vignettes, but, instead, give leaders flexibility to choose the most appropriate ones for parents, with a recommended minimum of six vignettes per session. The checklists 'Did I' section requests a 'yes' or 'no' response for core actions, for example 'Did I... write the agenda on the board'.

TABLE 22 Incredible Years infant programme and IY-T group adherence based on IY checklists

Group	Recommended vignettes (IY-I, N = 43; IY-T, N = 117), n (%)	Mean number of recommended vignettes shown per session	Total number of additional vignettes shown	'Did I' statements (IY-I, N = 73; IY-T, N = 120), n (%)
IY-I				
1a	27 (62.8)	3.4	14	72 (98.6)
1b	30 (69.8)	4.2	3	61 (83.6)
2a	42 (97.7)	5.3	24	65 (89)
2b	36 (83.7)	4.5	19	66 (90.4)
3a	38 (88.4)	4.6	4	64 (87.7)
3b	37 (86)	4.6	19	62 (84.9)
4a	37 (86)	4.6	1	65 (89)
4b	28 (65.1)	3.9	10	64 (87.7)
Average	34.4 (79.9)	4.4	11.8	65.1 (88.9)
IY-T				
1a	52 (44.4)	4.3	0	88 (73.3)
1b <sup>a</sup>	15 (53.6)	5.0	0	29 (90.1)
2a	98 (83.8)	8.2	0	113 (94.2)
2b	94 (80.3)	7.8	3	110 (91.7)
3	71 (60.7)	6.5	1	86 (71.7)
4	58 (49.6)	4.9	1	92 (76.7)
Average	64.7 (55.3)	6.1	0.83	86.3 (71.9)

a Site 1b completed only three sessions and so the average presented is across three sessions only.

**Note**

For IY-I, the total for 'Did I' statements differs for sites 2 and 1, as the sheet for session 7 has eight items for site 2 and 10 for site 1. For IY-T, site 1b group delivery was stopped after session 3 and adherence presented from sessions delivered is as follows: recommended vignettes, 15/28 (53.6%); and 'Did I' statements, 29/32 (90.6%).

The majority (six out of eight) of IY-I groups met the criteria of 80% on fidelity of adherence based on self-completed checklists, with the exception of sites 1b and 4b. For IY-T, only the site 2 groups met the criteria on fidelity of adherence based on the checklists overall, the 'Did I' statements and vignette use. IY-I groups, on average, showed 4.4 vignettes per session, whereas IY-T groups showed an average of 6.1 vignettes per session, in line with the recommended minimum.

### Quality of delivery

**Parent programme implementation checklist** Average scores for both IY-I and IY-T show over 50% fidelity (63.7% and 73.6%, respectively) using the PPIC assessment (Table 23). Scores indicate that adherence was the weakest subscale for both IY-I and IY-T. Eight groups achieved a fidelity score of over 80% for certain components (e.g. site 2 quality of delivery and site 3a participant responsiveness), but only three groups achieved this level overall.

TABLE 23 Agreed PPIC scores for IY-I and IY-T

Group	Adherence (% for subscale)	Quality (% for subscale)	Participant responsiveness (% for subscale)	Total score (% total)
IY-I				
1a	29 (72.5)	26 (74.3)	11 (73.3)	66 (73.0)
1b	17 (42.5)	27 (77.1)	9 (60.0)	53 (58.0)
2a	31 (77.5)	31 (88.6)	10 (66.6)	72 (80.0)
2b	23 (57.5)	29 (82.9)	8 (53.3)	60 (66.0)
3a	21 (52.5)	24 (68.6)	13 (86.7)	58 (64.0)
3b	13 (32.5)	23 (65.7)	7 (46.7)	43 (47.0)
4a	28 (70.0)	31 (88.6)	11 (73.3)	70 (77.0)
4b	12 (30.0)	17 (58.6)	7 (46.7)	36 (40.0)
Mean scores	21.8 (54.5)	26 (74.3)	9.5 (63.3)	57.3 (63.7)
IY-T				
1a	25 (62.5)	25 (71.4)	9 <sup>a</sup> (60)	59 (65.0)
1b	29 (72.4)	28 (80.0)	13 (86.7)	70 (77.0)
2a	18 (45.0)	20 (57.1)	7 (46.7)	45 (50.0)
2b	36 (90.0)	31 (88.6)	13 (86.7)	80 (88.0)
3	30 (75.0)	27 (77.1)	14 (93.3)	71 (78.0)
4	29 (72.4)	29 (82.9)	14 (93.3)	72 (80.0)
Mean scores	27.8 (69.5)	26.7 (76.3)	11.7 (78.0)	66.2 (73.6)

a Item 4 ('Does the group leader encourage all parents to participate?') was coded as not applicable for site 1a, as the item is about attending to all members of the group and there was only one parent in attendance.

#### Notes

The minimum total score is 18 and the maximum total score is 90.

For individual subscales: adherence – eight items, scores range from 8 to 40; quality of programme delivery – seven items, scores range from 7 to 35; and participant responsiveness – three items, scores from 3 to 15.

**Session length** Session length was assessed to understand intervention dose. Sessions should run for 120 minutes when including breaks and 90 minutes without breaks. IY-I sessions ranged from 60 to 112 minutes, with a mean time of 69 minutes and, therefore, the majority of sessions were not within 10% of the expected time. IY-T sessions ranged from 92 to 116 minutes, with a mean time of 102 minutes.

Figure 7 compares adherence scores derived from independent observation (i.e. PPIC) with adherence scores from self-completed checklists. In almost all cases, self-rated adherence was higher than observed adherence, although concurrence was closer for IY-T than for IY-I.

**Participant responsiveness**

**Parent satisfaction: weekly forms** Brief parent satisfaction questionnaires were given to parents at the end of each IY session. Out of 374 weekly attendees, 303 evaluation forms were completed for IY-I (i.e. a 81% response rate). Out of 179 weekly attendees, 158 responses were completed for IY-T (i.e. a 88.3% response rate). A high level of parent satisfaction with the programme content and delivery was found (see *Report Supplementary Material 1*).

Mean satisfaction scores for IY-T are slightly higher than for IY-I. Parents indicated that they were likely to use the parenting strategies discussed in each session and felt able to participate in the group sessions (IY-I, 98.2%; IY-T, 96.7%). Of those parents who did not feel able to participate fully, reasons included unsettled children, lack of sleep, tending to their child and not contributing to the group discussion. Across both IY-I and IY-T groups, parents commonly found the following most useful: group discussion, advice from other mums, ‘knowing you’re not alone’, ‘confidence that I’m doing things right’ and group-sharing experiences. High parent satisfaction was also reported in the final end-of-programme questionnaire (see *Report Supplementary Material 1*).

**Aim 2**

**Research question 3: barriers to and facilitators of delivering the E-SEE Steps model**

During interviews and focus groups, stakeholders highlighted several barriers to and facilitators of adherence and quality. Themes were categorised under service engagement processes, the IY programme, organisational support, supervision and mentorship, and connection and relevance. Despite rich data, we have limited the number quotations illustrating each point because of space constraints.

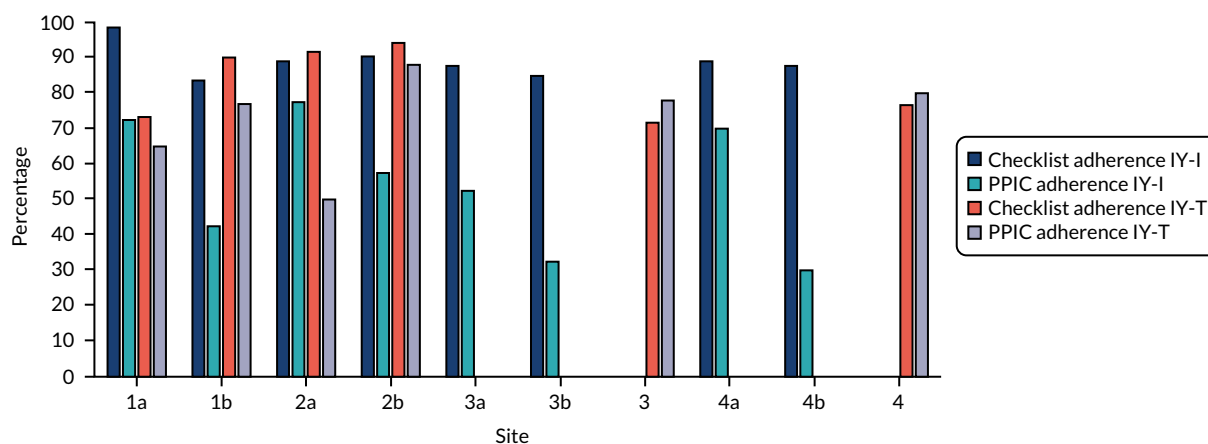


FIGURE 7 Comparison of PPIC and checklist adherence for IY-I and IY-T. Note that there was only one IY-T group for sites 3 and 4 (labelled accordingly). Sites 3 and 4 had two IY-I groups split as 3a and 3b, and 4a and 4b, respectively.

## Service parent engagement processes

**Timing and relevance of the offer of an Incredible Years infant and toddler programme group place** Group leaders and service managers highlighted barriers to engaging parents, particularly around timing or relevance of the offer, and how parents viewed the opportunity:

*That was the only . . . the biggest challenge and, you know getting hold of people when they're back at work, particularly for the toddler programme. You know, we found we left lots of messages and struggled to kind of get in touch with people, whereas it was a lot easier when it was the baby programme to get in touch and find out what people could do and what was best.*

Site 3 group leader

Incredible Years mentors and service managers described the importance of timing:

*I think it's always about pathways. I think it's about getting the right parents at the right time.*

IY mentor 3

*It's really about maintaining that contact and identifying when is the right time because for a lot of parents you don't walk past your front door for quite some time when you've had a newborn baby. You're getting your routine in place . . .*

Site 1 service manager

Across groups, a large number of parents accepted the offer initially, but did not attend the group sessions. Service managers highlighted that a quicker time frame was needed (than was possible in the trial) from participant identification to group access, maintaining contact after initial identification and not losing momentum, and working in partnership to do this effectively:

*We're working on quick timescales so if we say a family is identified for a group we would be hoping that at the start of the next group we'll be forecasting for planning stuff when one is looking towards the middle, we're already starting our next group, that then families will be accessing straight away so they haven't lost that momentum. They know what it is, they're not waiting a long time.*

Site 1 service manager 2

Parents said that by the time they were invited to the group they had already started to sign up for other classes, confirming the importance of early contact with parents:

*. . . I'd started to sign up for other classes by that point.*

Non-attending parent P7

Interviews with non-attending parents also highlighted the importance of maintaining contact after the initial offer, suggesting that follow-up telephone calls and/or home visits by group leaders would have helped them to feel more comfortable attending the group. A proactive welcome from the group leader would act as a bridge or a link for that parent, given the challenge of a group intervention when confidence and personal well-being is low. Short notice was also described as a barrier to attendance by these parents:

*I don't maybe, maybe a visit before. Maybe send a letter and then if possible . . . I know it's difficult, but go out and visit and see if they've got any issues about going to the group and they can . . . That would be really good.*

Non-attending parent P2

**'Selling' the offer** Lack of clarity in the offer and unconfirmed crèche availability impeded group leaders in 'selling' the programme to parents. Group leaders cited a number of issues, including parents not receiving adequate information from the research team prior to leader contact, parents receiving

inconsistent information about groups (e.g. a different time or venue) and leaders being unable to provide a clear offer (e.g. unconfirmed crèche availability):

*The kind of thing that came up when I was contacting parents for toddler group was that we didn't have the crèche and we weren't sure if we were going to have a crèche . . . So that was kind of like yeah, I can come if that's there but if not, definitely not.*

Site 4 group leader

Some sites felt the initial 'sell' at the point of identification and recruitment to the study by midwives and health visitors was not as strong as it needed to be:

*. . . it needed to be the early identification from the visits and from the midwives and from the health visitors trying to sell the . . . the project really . . . And we didn't feel that was happening as much as it could be doing. If they were the ones with the information . . . we . . . we don't have knowledge of which parents and parents to be that are out there, so we need to rely on our partners to be able to do that. So, I . . . I would just encourage the buy-in from other partners really.*

Site 3 service manager 3

### **The Incredible Years programme**

Many barriers to adherence stemmed from the IY programme itself. Group leaders found processes, such as role play, challenging and the collaborative (non-didactic) nature of the programme delivery was different from usual practice for many of them.

**Group leader resistance to Incredible Years processes** Perceived resistance was specific to one site in particular; however, it presented a significant barrier from both the perspective of leaders in that site and their IY mentor:

*A couple of them were resistant, very resistant, to role play anyway. They were resistant in the training and they, that moved on and although we spent, I spent more time than anything else I think on trying to problem solve that with them.*

IY mentor 1

Group leaders suggested that role play was less challenging in the context of IY-I, as parents had their babies with them. However, group leaders did not think it would be feasible to do this with toddlers. Other mentors also identified that difficulty with core IY elements could be a barrier and related this back to the individual qualities and experience of group leaders:

*. . . I think that the usual thing that you see on training is some very strong participants who just seem to get it, and their experience lends themselves to . . . to kind of understanding the programme quickly and being comfortable with role-plays and that kind of thing. And then you get participants who really didn't, or don't quite understand what was expected of them, and it's just not a good fit . . .*

IY mentor 3

**Grasping the collaborative approach and 'trusting the process'** Some leaders found the IY collaborative (non-didactic) approach difficult to employ. IY mentors linked this to a departure from the requirements of their day-to-day job where often they give direct answers to parent questions:

*I think a common theme would be getting them to not give the parents the direct answers. So where a parent would say . . . 'Well what if I do this? Or should I do this?' And in their normal roles it would be very appropriate for them obviously to give a direct response. So it was encouraging . . . a collaborative approach and get the parents to come up with the ideas, come up with those sorts of type of things . . . And we just kept saying, 'Just trust that process. Just trust it. Give it a little longer, open it up to the group'. And I think the more they did that, I think they saw that the parents were much more engaged.*

IY mentor 2

Leaders noted that because they were focusing so much on delivering the content, how they were delivering it often got lost:

*... when we were delivering the baby group we were very content focused and we kind of slipped on doing some of the more sort of ethos type, like giving out the stickers and ... We were so focused on the content ... And delivering that content and making sure we got the content right that actually all the other stuff that kind of went around it ... And the way it was supposed to be delivered I think we kind of slipped up a little bit there just because it ... there was just so much to think of.*

*Site 3 group leader*

**Programme tailoring** Several groups had to negotiate delivery of the E-SEE Steps model to very small groups of parents. IY group leaders described the challenges of adhering to the programme while tailoring content for their groups.

In addition, owing to trial recruitment delays, some IY-I sessions were delivered to babies of different ages, some of whom were older than the early programme content was designed for:

*And the particular challenge was for us, with ... the babies that we had, and I can't remember why this happened now, but they were 4 months old when it started, not [8 weeks], and that was quite a challenge, trying to tailor ... still deliver the programme how it should have been delivered.*

*Site 2 group leader*

**The need for a training booster or preparation session** The majority of group leaders ( $n = 35$ , 89.7%) were not able to deliver a recommended 'dry run' group delivery prior to trial group delivery, with over one-third ( $n = 14$ , 35.9%) feeling only 'somewhat prepared' to deliver groups. Group leaders described the challenges of having a large gap between training and delivery of the programme, with some group leaders suggesting a need for more training in the practice of delivery and the facilitation of groups:

*Site 2 group leaders F1: I think it's also the timing, like when you actually do your training and like the gap but then you've got to up and run it and you're like, oh my God ... Just like a refresher ... And like but just like we're talking about that, literally, what are you going to do?*

*Site 2 group leaders F2: Yeah. Like a mock session.*

This suggestion aligns with IY mentor feedback that a further session would be beneficial:

*I think it's a real challenge ... But then I think that there needs to definitely be either a fourth day that's a follow-up day once they've gone away and processed some of the stuff.*

*IY mentor 3*

### **Organisational support**

Stakeholders identified a number of ways in which organisational support could facilitate or hamper fidelity, including capacity for planning and preparation, and concern around recruitment.

**Capacity for planning and preparation** The majority of group leaders reported that their organisations had been 'supportive' (48.7%) or 'very supportive' (30.8%) during delivery. Group leaders and service managers described issues of capacity and a lack of protected time needed to deliver the programme as intended. For some group leaders, this was due to uncertainty in the service and changing roles, and the responsibilities within their own service or within the partnership service during the trial. Others group leaders reported having underestimated the time and capacity needed in their staffing:

*I tried to do this model, this time, but ended up having to pull [name] off it because we just didn't have the capacity for that ... we tried to do gold standard again, but we couldn't quite get there. That's our problem, things can change from one week to the next and one day to the next.*

*Site 3 service manager 3*



A lack of awareness at commissioner level resulted in a mismatch between the demands of the delivery and the capacity available to deliver it:

*I think the decision was actually taken at commission level . . . We were virtually instructed we had to do it. So there has been a slight mismatch in the capacity to deliver it . . . We would like to do it, but there is very little leeway in the service . . .*

*Site 3 service manager 1*

Although some sites struggled with capacity, when adequate planning and preparation time were provided to group leaders it was really valued:

*We have had that time allotted, because they know how important it is . . . Because it has come from the top, so they have passed it down to pass down, to pass down, to pass down, and they're the ones running it, so they know how important it is. And we need that time, and it is protected, and we've . . . I don't think we've had to change our planning sessions, anything like that.*

*Site 1 group leader*

Although many services had trouble with group leader capacity, those who invested additional capacity described a benefit in their group delivery:

*Because we had [name] as well, didn't we, shadowing when we did it, and she was sort of very much in there with the parents wasn't she? And picking up on bits that we might have missed . . .*

*Site 1 group leader*

**Flexibility and understanding of managers** Flexibility and understanding of managers was perceived as beneficial. Managers who attended the IY training found it easier to understand the demands and were able to support their staff. For example, understanding that group leaders needed time to recap learning and IY processes from training to prepare for the sessions.

Managers who had not been on the training or did not have a good understanding of the IY programme struggled more:

*Site 2 group leaders F2: I I think with the baby and having been on the other side now, with the baby we sat, and we felt that [manager's name] . . . it would have been useful if [manager's name] had done the training . . .*

*Site 2 group leaders F4: Yes.*

*Site 2 group leaders F2: To understand our level of sort of . . . panic, frustration, overthinking.*

*Site 2 group leaders F4: Yeah, that's true.*

*Site 2 group leaders F2: Like I don't think she understood, I think she kind of felt well, you've done the training, what's wrong.*

Although the post-delivery survey suggested that most leaders felt well supported by their organisation, focus group feedback suggested that this issue is complex. Despite, in some instances, caseloads being reduced, a lack of awareness in the wider team often meant that tasks and e-mails continued to go to those delivering IY, which led to difficulties in managing time to prepare for sessions:

*[Manager's name] did say that she had reduced the tasks sent to me but . . . it just didn't feel like it.*

*Site 2 group leader*



**Motivation, engagement and investment** An expressed commitment to the programme from leaders and service managers contrasted with organisational-level difficulties, resulting in examples of leaders going 'above and beyond' to keep the programme running:

*I think we're all absolutely invested in it. We've all been trained on it. We see the potential of the outcomes for families and when it's run and we've got that attainment from the group, that we've seen some wonderful outcomes from it.*

Site 3 service manager 2

Leader motivation and investment was acknowledged by mentors as an asset that enabled them to be resilient in negotiating the challenges of delivery:

*I remember one of the early baby programmes, I mean, they had dropouts which had nothing to do with them before the group started, and all sorts of things. But they . . . talk about resilient, they just kept going, because they were so passionate about, 'Let's get this going, we can do it'. And that I think was really lovely to see.*

IY mentor 2

### **Supervision and mentorship**

Ninety-two per cent of leaders attended the supervision sessions and the majority reported that the supervision provided was 'useful' (n = 10, 41.7%) or 'very useful' (n = 10, 41.7%). Supervision and mentorship were identified as central to successful delivery; however, there were challenges negotiating the logistics of remote supervision.

**Incredible Years mentor guidance** When leader-mentor relationships worked well it could boost adherence. Group leaders described how guidance from their IY mentor enabled them to maintain fidelity in their delivery, in addition to providing confidence and reassurance:

*Site 1 group leaders F4: Even when we had sessions that we thought, 'Oh God, that has gone really, really bad', she would still be able to pick out the positives, and, 'Oh actually there were positives for me'.*

*Site 1 group leaders F3: Yeah but very good at actually encouraging you, but also honing in on the needs of the people you are working with, and getting that bit where . . . because it can be quite prescriptive. It is that bit about getting it to be individualised and personalised to the people that were in the group. They were very good at that, and all the wealth of experience comes out, doesn't it, when you are talking to them. So you pick up so much, even with the limitations of not having supervision as it could have been. I think it was really, really rewarding and so it was really good.*

**Logistics and mode of supervision** Supervision sessions were intended to assist leaders in enhancing the quality of programme delivery; however, across the groups, logistical issues disrupted this process:

*. . . we only had a certain amount of time together as a three of us to be able to meet to prep for the sessions because we all had different sessions running and obviously we had [name] from health and us from support workers' side, the supervision we found very tricky to find the time when we all were together at the same time . . .*

Site 3 group leader

The remote nature of supervision was problematic, both in terms of not being face to face and not being able to review videos from the sessions:

*Site 1 group leaders F2: Because [of] the Skype [Microsoft Corporation, Redmond, WA, USA] supervision, you were relying on phone connections, everyone else's internet, it didn't work, and you couldn't unpick your videos, and I think that's what you need for the Incredible Years.*

Site 1 group leaders F3: *And that's the way it should be, we tried to figure it out, but with governance we couldn't really do it. And I think it is a shame because that was really, really, really useful.*

Site 1 group leaders F2: *And that's for me, the live supervision, how it should be, and I benefited from that because it just helps you keep to that true fidelity, doesn't it? It helps you stick to what you're meant to.*

*... the barrier throughout was I wasn't there ... I wasn't able to give the video like I'd have liked to have done ... So you are doing phone supervision, because that was another thing, Skype just didn't seem to work anywhere in ... you see? I mean we tried but it tended to end up mainly being on the phone, and I mean I have to say they were prepared, and I think we sort of set that out quite early on. So they knew what sort of questions, and the goal of the supervision would be ... Because when you are on supervision on the phone, you can only respond to a certain point, on what they are giving you.*

*IY mentor 2*

The supervision process was enhanced when groups were able to find a way around viewing the video either in person or virtually:

*I was able to sit with them, with their video, show them the video, and straight away visually looking at it and just picking up things that they hadn't actually thought of, just by looking at the video, like the seating position and a few other little bits. So I felt actually that the supervision has a richness. What I can see on the video is some of the fantastic stuff they were doing, and then you were able to give some positive feedback.*

*IY mentor 2*

### **Connection and relevance**

The group leader–parent relationship prior to and during group delivery was highlighted as crucial, particularly in terms of initial engagement, retention and quality of delivery.

**Importance of building rapport** Group leaders described the importance of building rapport with parents early on, often pre-delivery, and how this was made more challenging by capacity and the short period of time allotted for engagement. Rapport was particularly important in cases where the parent was anxious or there were other challenges going on:

Site 4 group leaders F1: *... I seemed to speak to them a lot before we started ... Backwards and forwards so I felt that we all really knew each other quite well, really ... But that was quite nice because I think they needed that support to ... Come join the group, so ...*

Site 4 group leaders F3: *Yeah, especially if we're asking them to travel as well ...*

Site 4 group leaders F1: *Yeah ... That definitely helped with ... I think the two that came from that area were really vulnerable and ... in for completely different ways and actually had [parent's name] not had that connection before easily we could have lost them.*

**Knowing your group** Understanding the individual goals and needs of parents enabled more sensitive facilitation and personalisation of the manual and better management of group dynamics:

Site 1 group leaders F3: *And the baby one where they hone into that development, that stage of development, what they actually need to, as it building up, and if the baby is not sort of like a newborn, and it's like, 'Well, they are past that stage', so they are not that particularly interested, but when it comes onto this stage of development where the babies were at, it was like ... you could see that they actually were more interested and taking on board what ... and because they could relate it to their own children, 'Oh, mine is doing that'.*

Site 1 group leaders F2: *And I think that is the art of IY actually.*

Site 1 group leaders F1: *... it is that real personalisation ...*

Contact with parents during the groups was particularly important in continuing this person-centred approach. Although most leaders described difficulty in making direct contact with parents over the telephone, they adapted their practices to using texts or other methods and felt that this was a critical part of engaging with and valuing parents:

*The phone call was well received ... you keep the engagement, and you do that check in, and you actually say to that parent, 'How are you? How have you got on?' And they are valued as much as the baby is valued, so we didn't have anybody kind of say, 'We don't like it'. And I think there is something in that, it is kind of a reminder. So even though they are not answering the phone maybe they might see the message.*

*Site 1 group leader*

Service managers also felt that this was a helpful practice and something that they would take forward:

*I think the model, the sort of where we're phoning parents, sort of, to check in with them and see how things are, I think that's something that's really good and I think that, you know, encourages parents. You can pick up easily if there's anything, if a parent's struggling, or wasn't happy with a session and things. I think that's a really good practice that we'll probably look at doing moving forward.*

*Site 3 service manager 2*

#### **Research question 4: acceptability and feasibility of the E-SEE Steps model delivery**

In focus groups and interviews, parents, group leaders and service managers reflected on long-term or future implementation of the programmes, both in terms of stakeholder acceptability and feasibility of future implementation.

##### **Acceptability**

Several barriers to and facilitators of acceptability were identified, including a lack of engagement with the IY-Bs, the nature of the offer and the demands of attendance and programme processes.

**Utility of the *Incredible Babies* book** Interviewees reported struggling to engage with the IY-Bs and they had more utility for those who did not attend (i.e. those who were not invited to IY-I or IY-T) than for those who did. The language and text-heavy nature of the book was off-putting. Parents said that on its own this may be a barrier to using it; however, when using it alongside the course it was different:

*A lot of it's intellectual. So it's really hard to understand ... So it's like, what's that word? Google [Google Inc., Mountain View, CA, USA] it.*

*Site 2 parent focus group P1*

*As it's anonymous, I'm going to admit that I didn't read the book.*

*Site 4 parent focus group (parent who attended IY-I and IY-T)*

Some parents suggested an audio or electronic book may have been more accessible. However, participants were asked if they would like an audio or hard copy when they were randomised to

intervention, but may not have remembered this. The audiobook was considered potentially beneficial for low literacy families, but it may have also been useful in other scenarios:

*It would have been a lot easier if that book was electronic . . . when you're trying to feed and it's 4.00 a.m. and you can see the book on the other side of the room and you think, 'I would really like to read that right now'.*

*Site 4 parent focus group*

**Nature of the offer** Key stakeholders perceived familiarity and trust as facilitators of acceptability, with the relational nature of the offer highlighted as critical. Familiarity and trust were seen as important for attendance and as key factors in relation to the retention and engagement of parents.

**Familiarity** Familiarity, both in terms of knowing other parents attending and in terms of familiarity with IY processes, were contributors to acceptability:

*I felt . . . because I came in the second time and they'd already pretty much gone through it and then with myself and two other parents came along I felt that they'd already settled in doing it then, they knew exactly what to expect and that made it quite easy, it was, you know, we didn't have to go through everything and, like you say, repeat things, they kind of were in the groove . . .*

*Site 4 group leader*

For non-attending parents, familiarity was important, and this included the group taking place locally and attending with a co-parent or friend:

*Yeah, I think that would have helped, knowing at least one other person that was going. Obviously you don't know anyone. That probably would have been more encouraging.*

*Non-attending parent P3*

**Contact and information provided** How contact is made, who makes the contact and the quality of contact prior to groups starting was highlighted as pivotal in parents' decision to attend.

Although many parents had positive first impressions of the person who made the offer of the parenting group, some did not recall receiving any contact at all from group leaders despite being eligible for a place:

*No, I haven't had anything. That's why I thought they probably haven't chosen me because am I right in thinking that not everyone got picked?*

*Non-attending parent P5*

For others, there was a lack of information about the offer. For some, this was not a barrier, but in some cases it contributed to a decision not to attend:

*I didn't have any [information]. All mine was, was when the lady [a trial researcher] come to my house she kept saying I don't want to hear if you're going or not and I just got invited to it but I didn't have a clue what it was. Until I actually got there . . .*

*Site 2 parent focus group*

Most parents were happy to receive a letter, but agreed that this should be followed up with a home visit or telephone call, ideally by a group leader who could provide them with all the group details. Some parents recalled only receiving a letter or text to tell them the location and time of the group. For parents struggling with low mood, it was felt that more encouragement and information would be needed to get them to attend.

**Location of groups** IY group-based programmes are designed to be delivered in community settings, such as in community centres, church halls or similar. The number and location of groups in each site were planned around the spread of the trial's intervention participants; however, low take-up rates meant that some groups were not viable and had to be closed/merged. This meant that many intervention parents did not have a local group they could attend and, therefore, the proximity of the groups was, perhaps unsurprisingly, reported as a key barrier:

*And like it needs to be within an hour's walking distance . . . The group wasn't.*

*Non-attending parent P11*

*I think I only went to one of the groups and then particularly where the centre was positioned, I had to pay for parking if I had to park, stressed out with two children in the car so, rather than participate in something, I wanted to enjoy it, I was looking at it as a bit of a chore . . .*

*Non-attending parent P1*

Although group location was reported as convenient for 86% of IY-I attendees and 92% of IY-T attendees, these data are biased towards those who attended most of the sessions and who completed and returned the end-of-programme questionnaire.

**Crèche and child care** Weekly evaluation forms asked parents whether or not they made child care arrangements for their other children while attending the sessions. Twenty-seven per cent ( $n = 74$ ) of IY-I parents and 60.7% ( $n = 91$ ) of IY-T parents made additional child care arrangements (Table 24). End-of-programme questionnaire data confirms that child care was necessary to attend the group for 20% ( $n = 8$ ) of IY-I attendees and 71% ( $n = 10$ ) of IY-T attendees, supporting the need for child care support. All IY-T parents who completed the end-of-programme evaluation used the crèche provided for the programme at some point during the trial, compared with 15% of IY-I attendees.

Although child care support was important, leaving children in an associated crèche divided opinion among parents. Those who attended the groups described how much they enjoyed some time to themselves during the sessions, with the crèche providing some respite, and how much their child enjoyed the crèche:

*Site 2 parent focus group interviewer: Anything else that you really liked about the group?*

*Site 2 parent focus group P3: The kids playing downstairs!*

TABLE 24 Use of child care during IY-I and IY-T sessions

Programme	Child care provider	Number	Percentage
IY-I	Crèche at same children's centre	19	25.7
	Relative or friend	16	21.6
	Paid nursery	22	29.7
	Child worker	2	2.7
	Other	15	20.3
	Total	74	
IY-T	Crèche at same children's centre	76	83.5
	Relative or friend	5	5.5
	Paid nursery	10	11.0
	Total	91	

Site 2 parent focus group P1: *The group workers, the leaders were good they were lovely all of them 'cos it was like all of them on the first one then there was two the same but there was only two then another two but yeah they were all nice and a bit of respite too.*

Site 2 parent focus group P3: *Yeah. The second one he [co-parent] was working full time and that would be my only break, it was awful.*

Site 2 parent focus group P2: *... I liked comparing [name] to other babies 'cos they were all the same age, and they were all starting to sit up and some of them started to and some of them weren't, some of them were rolling, some of them weren't it was just like quite nice to see what who was doing what and they were all so different really um and [name] really enjoyed the crèche ...*

For some who attended, being separated from their child for the first time was challenging, but was something they accomplished during the course of the programme. For others it was a barrier to attending altogether:

*Yes. They're toddlers, but obviously nobody's ... I'm not on maternity leave any more, it meant giving up quite a chunk of time with my kids on the only day I have off with them ... So it would have meant even longer without my child ... And once I'm back at work, that was just too long.*

*Non-attending parent P6*

**Co-parent attendance** Co-parent involvement was a substudy of the E-SEE trial, which focused on fathers' involvement (see Appendix 10), as well as a theme in the process evaluation, which explored all types of co-parent support. The substudy revealed that the timing of the group was important for fathers' participation. Fathers were often excluded if there were no evening or out-of-hours groups. In addition, the substudy identified the importance of ensuring that the parenting offer felt relevant to fathers, with many fathers considering parenting programmes that are offered to new parents as primarily 'for mothers'.

The process evaluation found that involvement of co-parents was important in relation to parent attendance, yet the offer of co-parent attendance was inconsistent across sites. For parents who struggled with low mood or anxiety, co-parent availability to attend was critical to feeling able to attend the group. Although work was cited as the biggest barrier to co-parent attendance, in some cases it was not clear to parents that co-parents were allowed to attend:

*I don't think it was something that was made clear but, even if it was, he works so he wouldn't have been able. I'm self-employed but on maternity leave at the time, my partner, he wouldn't have been able to get time off. He's an engineer so he's pretty much always on call.*

*Non-attending parent P1*

Site 3 parent focus group P2: *Better if you could bring somebody with you ... So not ... I know you're a single mum but if you could bring like your mum with you ...*

Site 3 parent focus group P1: *Or my mum, or my husband, or whoever ...*

Site 3 parent focus group P2: *Just to be ... a, to listen to what's been said and sort of engage with it too because they're your main caregivers for your child ... But also to have the kid in the room if ... That's what you need.*

For those parents who attended groups with dads or other co-parents in them, they felt that this brought a positive dynamic to the group. Other parents described the positive effects they noticed



since their co-parent attended the sessions. However, it also seemed important to some dads that there would be others attending too:

*I've got two kids and I can tell you, he is an amazing dad because of this group and he will say the same . . . He just gets it because of this course and I thank you so much because, I'll tell you what, I probably wouldn't be married to him now if it wasn't for this course and I'm not even joking . . . He only came because there was another dad here. If there wasn't another dad here, he wouldn't have . . . yeah, and today, he was like, 'Should I come?' I was like, 'Yeah, come on', and he's like, 'Will there be any other dads?'*

*Site 4 parent focus group*

**Demands of attendance and programme processes** There was mixed acceptability of some of the IY processes with parents, particularly for those who described struggling with low mood/depression.

**Incredible Years processes: role play** Many of the parents reflected that they struggled to engage with role play:

*Site 4 parent focus group P2: I don't like the role play.*

*Site 4 parent focus group P3: I don't like it either but I found what I learnt from it worth doing it, absolutely . . .*

*Site 4 parent focus group P5: I think it's just not a natural thing for most people to do, is it?*

*Site 4 parent focus group P2: No, especially when there was no children in the room.*

*Site 4 parent focus group P3: It was good when [group leaders] would take the reins as the parents and sort of show us how to handle situations.*

Some parents identified the role play as a particularly Americanised aspect of the programme, something they did not feel translated well for UK parents:

*Site 2 parent focus group P3: We kept getting the leaders to do it 'cos . . . They wanted us to practise giving praise and um we wouldn't do it so we had the leaders and one of them was putting toys away and the other one was praising them and you could just see she'd gone bright red.*

*Site 2 parent focus group P1: Yeah.*

*Site 2 parent focus group P2: Americans are quite cheesy aren't they, as well they're quite, like, quite expressive with their emotions and things and we're rubbish at that.*

**Low mood and demands of the Incredible Years programme** Group leaders questioned the fit between the demands of the IY programme and parents with low mood and depression. Group leaders identified low mood and additional mental health issues as a barrier to group attendance. Group leaders felt that there was a bias in those that attended, whereby those who were able to come were not suffering with low mood and were more proactive in general:

*Interviewer: Did low mood prevent any parents from actually really participating and engaging?*

*Site 2 group leader F1: Definitely.*

*Site 2 group leader F3: One of ours, she came to a little bit towards the end, didn't she?*

Site 2 group leader F1: *That's why we had to be careful about role play, wasn't it, that sort of woman who shutdown. Also, in the baby group there was a couple that we kept trying to get but they just couldn't. Because being in a group, you know the group for some people's really difficult as well.*

Site 2 group leader F3: *I think one of them dropped out, didn't she, she wasn't able to do it.*

Interviews with eligible but non-attending parents also reflected anxiety and low mood as barriers to attendance:

*Yeah, and I didn't go only because I was suffering quite bad with anxiety at the time and I didn't feel comfortable going without my partner but, because it was only set times, my partner was at work for most of those days . . .*

*Non-attending parent P3*

However, parents who attended the sessions reflected that it had helped them with their low mood and anxiety:

Site 2 parent focus group P3: *. . . postnatal depression especially when we found out about having [name] because he was a massive surprise, um, it helped a lot being able to talk to people that have children close together and, um, that when their partner works and thing it's very helpful. It gets you out the house as well so it's never the same when you go to meetings.*

Site 2 parent focus group P1: *Yeah. It is good support that way it is, when you're meeting new mums 'cos you get some mums in [a] similar situation as you isn't there. Some of them with depression and some with anxiety as well you know.*

Group leaders felt that active engagement – core to the IY model – with parents between sessions (e.g. telephone calls and follow-up) was central to ensuring that low mood did not continue to be a barrier to attendance:

*. . . so in the baby one, the only person we've had that's done both baby and toddler with us, the second week she didn't come to baby and we gave her a phone call and were chatting to her and she said yeah, I really struggled, I just couldn't make . . . I couldn't . . . couldn't get out the house this week, I just couldn't do it. So we had a good chat with her . . . And actually, she then did come out of her shell quite a bit more and she's done other things with us at the children's centre.*

*Site 3 group leader*

One of the core IY processes – role play of scenarios – was identified as a particular problem for these parents. Although group leaders found different ways around this, such as using another 'buzz' instead of the role play or demonstrating the role play themselves, they were unsure whether or not it was feasible to deliver the programme successfully to parents struggling with low mood. A 'buzz' is where parents are asked to pair up to share thoughts about or discuss a particular topic or issue. For example, pairing up parents to share how they can get extra sleep and who might provide them with backup support if needed. In addition, group leaders were uncomfortable with the use of labelled praise when working with parents with low mood:

*I think you've got to be very careful that you're making sure everyone does get a sticker and that's where you get unstuck because then if you . . . if one person perhaps is maybe not partaking in it as well, perhaps there's other issues or whatever, like low mood and so they're not doing what . . . you know participating then they never get a sticker so . . . and then that's just going to make [them] feel even more crap.*

*Site 2 group leader F3*



## Nature of a parenting group

**Being part of the group** Being part of the group was described as the most valued outcome for parents attending the sessions. The opportunity to get out of the house and meet other mums was a big draw:

*Well, because it was my first, I have no friends with babies. In my office no-one had kids. They're all my age so for me it was also like another chance but I'm not a massive fan of baby groups. I do them but I don't particularly socialise. I'm like off you go, they'll play and I'll just sit. So for me it was a chance to meet other mums as well. That was a big draw.*

*Site 4 parent focus group*

**A parent-led approach** From legitimising play to self-care and 'seeing it work', a parent-led approach, relaxed environment, use of vignettes and group discussion facilitated parents in feeling a sense of empowerment.

A parent-led approach enabled parents to have ownership over the sessions and to drive the direction of the content:

*Site 4 parent focus group P3: They were really nice, let us take our own direction, more of a 'I need to get this across to you but you can do it your way'.*

*Site 4 parent focus group P2: Also here for moral support if you have no sleep tonight, we can get through today, yeah.*

*Site 4 parent focus group P3: We then needed to like, I don't know, say feeding, she was talking about the feeding then we would say how ours was going and it would change to like how it would help us instead of just carrying on with the course content, if that makes sense.*

## Feasibility

Complex interventions, such as the E-SEE Steps model, can be difficult to accommodate in existing systems. We investigated the extent to which different stakeholders felt that it was possible to embed and implement the model successfully. From the group leaders and service managers' perspective, this relates to the viability of the E-SEE Steps model within their existing service system(s) and the ability of staff to deliver it. From the parents' perspectives, this relates to their ability to access and make use of the E-SEE Steps support offered, given their circumstances and resources. Feasibility is discussed in terms of programme content and delivery, flexibility of service and system fit.

## Programme content and delivery

**Baby versus toddler programmes** Stakeholders reflected on the practicality of delivering IY-I and IY-T, considering factors such as difficulty recruiting and engaging parents and ease of implementation. Group leaders reflected that recruitment and retention of parents was easier for IY-I than for IY-T. This reflection is supported by the intervention take-up (IY-I, 64%; IY-T, 48%), attrition between acceptance and completion rates for the programmes (IY-I, 35%; IY-T, 73%) and in the number of sessions attended by parents across IY-I and IY-T (see *Research question 2: fidelity of delivery*):

*Site 2 group leader F4: I think life is just easier in that first year with a baby ...*

*Site 2 group leader F1: Yeah.*

*Site 2 group leader F4: That's when they bring baby with them ... So they're just there ... As well ...*

Site 2 group leader F2: *When they're sleeping, aren't they?*

Site 2 group leader F1: *Yeah.*

Site 2 group leader F4: *I think it's just easier, I guess . . .*

Site 2 group leader F1: *And they don't have lots of other plans, do they? Than with a toddler in tow. Life is just different, isn't it?*

Although some services initially thought that IY-T would be easier to implement, issues such as crèche provision, difficulty recruiting and subsequent smaller group sizes meant that it was a trickier programme to implement. Although services were more confident that they would continue to deliver IY-I in the future, there was uncertainty around the practicalities of delivering IY-T:

*So a discussion was with one of the advisors for the Children's Centres and how they can be using that in their offer, so, you know, I think pretty keen to be running Incredible Babies. I suppose the . . . the only reservation we currently have is with Incredible Toddlers and just the practical considerations and the financial considerations about providing child care, which . . . I think it's become apparent that it's really necessary to do, but, you know . . . and it's . . . they showed us the benefits of doing it, but how much of that we will be able to . . . to manage from a practical point of view . . .*

*Site 4 service manager*

For non-attending parents, there was also a distinction between the two offers (IY-I and IY-T). Working patterns were the greatest barrier to attendance for IY-T and leaving their child in the crèche was an additional source of anxiety that many parents found off-putting:

*The only other thing would have been if there was like flexible time sort of thing, the only thing I can think of more than just one session the same time, same day every week, if there were different ones, work it around your kids and your work if you're back at work and things like that.*

*Non-attending parent P3*

**Compatibility of programme content with the Baby Friendly Initiative** One concern raised by many group leaders was in relation to credibility. This concern was directly associated with what group leaders considered to be outdated videos and advice and some issues with regard to how IY-I content, in particular, aligned with guidelines from UNICEF's Baby Friendly Initiative (BFI):

*Yeah, this was quite a bone of contention for the health colleagues in that they were very concerned about the bottles in some of the vignettes.*

*Site 3 group leader*

*When you were actually going through some of the vignettes and stuff like that it was there, we'd be look at some of vignettes and going, ooh, can't really say that can you, it's like that doesn't fit in with baby friendly particularly, I think it was the baby one wasn't it? It was . . . Yeah. Use of bottles or something.*

*Site 2 group leader F2*

There were several things highlighted in vignettes that group leaders felt would be important to update, including use of electronics, weaning safety and general safety practices. However, the programme gave an opportunity to raise these issues with parents:

Site 4 group leaders F2: *I think even in the vignettes I remember the current practice around safety was . . . I remember you updating the parents on that . . .*

Site 4 group leaders F1: *Oh, yeah.*

Site 4 group leaders F4: *Yeah.*

Site 4 group leaders F2: *Because there's a few things ...*

Site 4 group leaders F1: *That's the only problem because they are very dated ...*

Site 4 group leaders F2: *Yeah.*

Site 4 group leaders F1: *And there's a lot ... Of things have changed. So [name], you were very good at updating the parents because they did ...*

The need to update videos for a UK audience was also highlighted by parents:

Site 2 parent focus group P3: *A change they could do is the videos and have them in English please! Not American slang English, have English ... English that would be a bit more relatable.*

Site 2 parent focus group P1: *And how the technology is now, some toddlers are sitting there with tablets now because you can get the learning apps [applications] ...*

### Flexibility of service

**The E-SEE Steps model as a development opportunity** The E-SEE Steps model was seen as a development opportunity for staff, from training opportunities in parent support and group facilitation to being involved in research and accessing supervision:

*... it wouldn't take all of our health resource, we could do it multi-agency, it would benefit both of our services, and the children of [site], to be able to offer them this intervention that we knew was evidence based, you know, recommended as part of the Healthy Child programme, but we'd never been able to really get the funding to do that. So it was a real incentive to think we can train ... get some of our workforce trained, and we can also deliver this to the population, and can this ... is this something we can harness going forward so that we carry on using those skills that we've been taught, and be able to deliver future groups? So sort of multi elements for doing it.*

*Site 2 service manager 2*

Post-delivery questionnaire data showed an increase in leader confidence post-delivery, with group leaders reporting feeling confident or very confident ( $n = 23$ , 59%) in setting up and running a parenting group, an increase from the 42.3% prior to delivery. In addition, for some leaders, involvement in the E-SEE Steps project resulted in changes to their wider practice:

*I've got one of my family support team leaders who was trained in IY Baby and she was fortunate enough to be videoed when she was delivering, so got that real raw feedback which was needed and she will say it has changed her practice, it really has, where before she would have stepped in, intervened and offered the advice and guidance, she now knows how to facilitate that conversation to enable parents to reach that understanding and that decision making and empowering them back to be able to do that ...*

*Site 3, service manager 2*

**Infrastructure: staff roles and levels** Staff level and role was raised as a challenge for long-term feasibility. The IY developer suggests that staff have extensive (i.e. 2 years minimum) experience with young children, experience with parenting skills and family interactions, at least one course in child development and involvement with group activities/awareness of group dynamics. In addition, training in social learning theory is recommended. The developer also recommends that at least one of the two leaders conducting a group should have a master's degree or a higher qualification, or a comparable educational background, with the therapeutic skills needed to assess and manage parents

and children with mental health problems. The majority of those staff who delivered the programme satisfied these criteria, having extensive experience with young children. Many of the staff who delivered the programme also had previous experience in leading group activities/programmes; however, only one group leader held a master's degree.

Some services suggested that IY delivery would be more feasible in the UK context if it could be delivered by a lower band of staff:

*Site 2 service manager 1: So I think all I'm trying to say is, it wouldn't [be feasible] ... we are the 0-5 service. We're an expensive service ... that is ... where the issues would lie. So I think we could do it, if we did it with our staff that ...*

*Interviewer: The band 5?*

*Site 2 service manager 1: Yeah, where their time isn't so pressured ... and costly ... our band 5 staff would have the flexibility and ... to be able to commit to those kinds of programmes more than our band 6s.*

Incredible Years mentors also challenged the suitability of particular individuals to facilitate. Mentors suggested that background experience and ability to embrace the collaborative aspects may be more important than band or level of qualification per se:

*So there may be something these days in how health visitors are trained that actually don't see day-to-day management of a child's tempers or a umm I'm, I'm not sure that they ... something like an IAPTs [Improving Access to Psychological Therapies] training which now children centres workers go on and possibly gives them a far better underpinning umm health visitors knowledge ... but that maybe a thought that they are, that they are working at the very high child protection level umm and therefore they may not see ... see, see this model as something that, that, that they feel they have time for or is their role.*

*IY mentor 1*

**Incredible Years accreditation process** The accreditation process requires submission of session or lesson protocols, evaluations, digital versatile disc (DVD)-recorded sessions for review and paperwork to support the application. The DVD review of the group leader's programme sessions is utilised to ensure that the quality of programme delivery is measured by observations of performance. There is an accreditation certification application fee of US\$600. Accreditation is considered to maximise the quality of group leader performance and programme delivery fidelity. It is believed that accredited leaders implementing the full programme will achieve results similar to those in the published literature. Accreditation grants eligibility for an invitation to continue training to become an IY-certified peer coach and/or mentor of group leaders in selected programmes.

Although more than half ( $n = 21$ , 53.8%) of IY group leaders reported that they were planning on running more IY groups in the future, only seven (17.9%) said that they were working towards becoming accredited. IY mentors described a need for more group leaders to work towards accreditation to have capacity to support any future roll out of the E-SEE Steps model. A service with accredited leaders is more likely to deliver the programme(s) with quality and fidelity to the model, and such leaders are then eligible to train to become IY peer coaches and mentors, providing the infrastructure in the system for monitoring and supervising others' delivery. Although many group leaders described being keen to work towards IY accreditation, they described the process as problematic with regard to financial and logistical support from their organisations:

*Site 4 group leaders F1: I really, really would like to ...*

Site 4 group leaders F2: *Yeah.*

Site 4 group leaders F1: *But again it comes down to budget and if I'm going to be here in 6 months.*

Site 4 group leaders F3: *We did put in that request for the funding for that and that got politely declined for site 4 group leaders right now, which is fine, I get that . . .*

F1: *Yeah.*

Site 4 group leaders F3: *Because we've got so much going on so I can't.*

Many issues stemmed from a lack of capacity in the system, or uncertainty in the future capacity of staff in the service to run another three groups to have the video footage evidence of group delivery to submit for accreditation:

Site 3 group leaders F1: *I think I was looking into getting accredited but I think because it . . . you know, with our organisation, I just don't think . . . and I know my manager had looked into it as well to see whether it would be funded by our organisation but I think at the moment, because we're just so . . . the uncertainty is . . . you know, we just don't know what's going to be happening moving forward that I just don't think for me it's an option at the moment.*

Site 3 group leaders F2: *No, I certainly can't afford to go through that process so at the moment that, yeah, it's not . . . that's not something that I'm thinking about.*

Site 3 group leaders F3: *Yeah, I'd say very similar because we're not going to have capacity to run each group three times in the next few years.*

## System fit

**Organisational structure and provision to support delivery of the E-SEE Steps model** Practitioners and service managers perceived a good fit between the programme and their service needs, but service managers described a mismatch between the support they would like to offer their communities and their commissioning arrangements, meaning that they felt it unlikely they would have the organisational structure to support delivery of the E-SEE Steps model going forward:

*The problem is as an organisation we would probably like to do it and support it, but we're very, very constrained by the commissioning arrangements. So for [name] trust I think it would fit very well, but because we are so constrained by the commissioning arrangements, we can't agree to do things for which we are not funded.*

*Site 3 service manager 1*

Several of the service managers acknowledged that they had underestimated the demand on their service. Service managers also described structural changes to services that were already stretching capacity, or likely to stretch capacity further, making the time and staff commitment required to deliver the E-SEE Steps model problematic:

*I think the difficulty probably was a slight underestimation of the demand on our service, but in conjunction with that, we were going through a huge period of change, probably the biggest the service has ever seen, slightly after agreeing to participate, so . . . and I don't think at that time, we could have had the insight to know how big an impact those changes would have on our staff and service . . . It's all our universal service having to have a big shift to absorb a lot of the work.*

*Site 2 service manager 1*

Incredible Years mentors acknowledged that this was central to supporting programme fidelity in any future roll out:

*... I don't think IY structure is the issue, I think it's the individual services' issue in making sure that they have got their structure to be able to support a parenting programme to the proper protocols ... I think they need to be thinking, 'OK, it is not just a 3-day training or a 2-day training'. They've got to think, 'Have I got enough resources to start with?' Have we costed out if we are doing so many groups over the next 2 or 3 years, what they will cost? Have we costed out the time for the staff, and do we put in the accreditation process, so that staff go to supervision, do the videos, do look at their practice? Because I think that will keep it at a quality level that will keep the programme going.*

*IY mentor 2*

Group leaders and service managers acknowledged that a lack of support and funding at a commissioner level was problematic in the feasibility of delivery going forward:

*I just think what's difficult moving forward is what they are thinking at a strategic level in terms of what parenting programmes they are going to deliver in the future ... And that still seems to be out for, I don't know, discussion, decision, whatever. So we are sort of slightly tied in terms of what will happen now in the future with it and whether the Incredible Years section of it will be included in it or not.*

*Site 3 group leaders F1*

**Alignment with current service offer** How well IY programmes aligned with the current service offer and how well they served to address gaps in the current offer was an important consideration for service managers in terms of future implementation:

*I think it is supporting, isn't it? But really supporting the Incredible Years and they [wider service] have done it in a programme change, haven't they? Where they have stopped doing other things and they have focused on that, which is brilliant. So we look at my area as perinatal mental health child mental health is part of that. There is nothing like that in our area at all, so I try and kind of talk to anyone I can about referring in because it is such a good opportunity for families, and I am really, really happy that it is continuing.*

*Site 1 group leader*

For some services, the evidence base of IY was a draw alongside providing families with help at the earliest opportunity through IY-I. In addition, the potential for continued contact across time through delivering IY-I and IY-T addressed a gap for some services:

*It's within our structure. There's no additional funding. It's just part of our delivery so we've committed and we will continue to commit to this offer remaining here in [name of site] and the reshaping of our resource will enable us to continue to do that ... if we know that a programme will create the most impact then why would we not do that? Why would we not? You can offer a hundred universal activities as lovely fantastic play but you may not be reaching the people that are most in need, the families that are most in need, and I just think that, especially for IY Baby, the earliest the opportunity the better.*

*Site 1 service manager 1*

**Partnership working** Working in partnership was described both as a strength and as a challenge to feasibility. Services and group leaders reflected that sharing skills and knowledge was beneficial and that sharing time and resources to deliver the groups was also a strength of that collaboration:

*Site 1 group leaders F3: I've co-delivered with health visitors initially, and their team leaders were part of the research meetings, they were part of the discussions around what we do. They were also tasked with, 'Have we got anybody else we can kind of boost it with?' And then this is your time ... I think it worked well having health as a core facilitator, definitely.*



Site 1 group leaders F2: *Yeah, you can see the benefits of you, can't you?*

Site 1 group leaders F2: *I think it does help from observing the session with [name], I learned more with what [name] does with your mental health side of it, and it can . . . it helped a lot because we have got our child care, nursery development side, but then [name] has got . . . That parental side of it, which links.*

Site 1 group leaders F1: *There is so much back the other way as well . . . So it makes sense, if you have got wide skill sets you can bring that into the group. I think it is a shame maybe health visitors didn't continue as well in the delivery, because obviously I am a mental health practitioner, I am not a health visitor, so I think that would have been really, really beneficial, especially for some of the baby . . .*

Although working in partnership was feasible and brought benefits for some group leaders, challenges stemming from a lack of capacity and a disparity in investment of time, effort and resource between partners was noted by others:

*The biggest grievance, if you like, that people had about it was how it was going to work in partnership with Health because immediately we got told about it Health were saying, 'Oh, we can't do that. We can't do that'. There was always we seemed to be the one having to try and push it.*

*Site 3 service manager 3*

## Conclusions

The E-SEE Steps intervention is a complex intervention that combines versions of an evidence-based intervention (i.e. IY) into a tiered, proportionate universal model. The process evaluation has explored the feasibility and potential of the components as they relate to the model as a whole and, as such, does not extrapolate its findings to any stand-alone delivery of either of the group-based programmes. Despite expressed commitment by services to providing good-quality and proportionate parenting support, acceptable levels of fidelity and high levels of satisfaction for those parents who attended sessions, the process evaluation data indicate a number of barriers and challenges in the potential scale-up delivery of the E-SEE Steps intervention. Some of these barriers and challenges are specific to the intervention, which, with adaptation, could be addressed. Other barriers and challenges are inherent to the public services system within which the intervention is situated and are arguably more difficult to resolve or, at least, require more complex system accommodation. Indications of barriers to delivery existed within the pilot, and some were addressed. However other challenges, such as IY programme adaptation, may also be required, which would need buy-in from the developer. The developer was, however, amenable and flexible in revising and republishing the IY-B in the light of BFI concerns.





## Chapter 5 Economic evaluation

Parts of this chapter have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This chapter seeks to address the research question of whether or not the E-SEE Steps programme represents value for money compared with SAU. It also provides a brief summary of a planned substudy exploring predictors of hospital attendance and admissions.

### Introduction

The aim of this economic analysis was to assess the cost-consequences and the cost-effectiveness of the E-SEE Steps programme using evidence from the E-SEE trial.

Health outcomes were considered from parental and child perspectives and were expressed as quality-adjusted life-years (QALYs) and as other available measures of parental and child social and emotional well-being. Costs were measured in Great British pounds (GBPs) (2018–19) from a public sector, family and broader (i.e. both public sector and family) perspective. Costs and QALYs were discounted at 3.5% per annum in line with UK guidelines.<sup>56</sup> Cost-effectiveness results are presented for primary caregivers, children and jointly using incremental cost-effectiveness ratios (ICERs) and incremental net health benefits (INHBs).<sup>57</sup> Econometric modelling was used to control for patient co-variables and uncertainty was explored using alternative assumptions through scenario and sensitivity analyses.

### Methodology

The health economic analysis plan is available at URL: [www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/#tab-3](http://www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/#tab-3) (accessed 13 January 2022).

#### Resource use and costs

Resource use data for the E-SEE Steps programme were collected at all four time points, with primary caregivers using a modified service use questionnaire (i.e. the CSRI). Baseline forms recorded resource use from 2 months prior to trial entry. A micro-costing framework calculated a cost for each primary caregiver and their enrolled child according to reported resource use, site-specific factors and group-based compositions. The costs of providing the E-SEE Steps programme comprised four elements: (1) providing the IY-B to all participants, (2) training (e.g. staff training, training materials and operational elements required to train group leaders in the IY-I and IY-T), (3) group costs (e.g. staff time and the materials required to practically run the IY-I and IY-T group-sessions) and (4) contact (e.g. the letters, texts, telephone calls and home visits necessary to provide support between group sessions). Operational costs for training included venue hire charges, catering expenses, programme trainer fees and telephone supervision costs. Training material costs included workshop manuals, IY-Bs, leader manuals and instructional DVD sets. Group costs included the materials used in the sessions, any catering provided and the provision of crèche facilities for participants.

Costs reflected GBP at 2018–19 prices and were grouped into E-SEE Steps programme-related costs, primary care costs, secondary care costs, mental health-care costs, community service costs, social services costs, child care costs and absent workday costs. Costs were estimated separately for parent and child, with E-SEE Steps programme-related costs and absent workday costs allocated to parent-related costs and child care costs allocated to the child-related costs. E-SEE Steps programme-related

unit costs were predominantly informed by financial costs recorded as part of the trial. Health care-, social care- and community-level unit costs were based on NHS reference costs and Personal Social Services Research Unit (PSSRU) costs.<sup>58,59</sup> Child care costs were sourced from relevant surveys,<sup>60,61</sup> and absent workday costs (productivity losses) were calculated using the human capital approach.<sup>62</sup> Unit costs were inflated to 2018–19 prices using the PSSRU pay and prices index where necessary (Table 25). E-SEE Steps programme-related costs were divided into within-programme fixed costs (i.e. those costs that are independent of the number of course attendees, e.g. group session costs), between-programme fixed costs (i.e. those costs that are independent of the number of programmes provided, e.g. staff

TABLE 25 Unit costs

E-SEE Steps programme resource	Unit cost (£)			Source
	Average	Minimum <sup>a</sup>	Maximum <sup>a</sup>	
<b>IY-I</b>				
<i>Between-programme fixed costs</i>				
IY trainer fees and expenses	2842.50	1555.71	4211.18	Recorded trial cost
Venue hire with catering	602.35	172.80	1330.01	Recorded trial cost
Other equipment/materials	100.00			Recorded trial cost
IY-B (for group leaders)	20.00			Recorded trial cost
Workshop manuals	20.00			Recorded trial cost
Programme trainee fees	16.20			Recorded trial cost
Leader manual and DVD set	600.00			Recorded trial cost
Telephone supervision of team leaders	763.81	330.00	1294.24	Recorded trial cost
<i>Within-programme fixed costs</i>				
IY staff delivery costs (per hour)				
Mental health practitioners	45.00			PSSRU 2019 <sup>59</sup>
Early years worker/parenting co-ordinator	62.53			PSSRU 2019 <sup>59</sup>
Family support worker	56.84			PSSRU 2019 <sup>59</sup>
Child health worker	61.50			PSSRU 2019 <sup>59</sup>
Health visitor	72.56			PSSRU 2019 <sup>59</sup>
Community health nurse	71.46			PSSRU 2019 <sup>59</sup>
Clinical team lead	87.00			PSSRU 2019 <sup>59</sup>
Equipment/materials (for up to 12 participants)	200.00			Recorded trial cost
Catering (each session for up to 12 participants)	57.32			Recorded trial cost
<i>Programme variable costs</i>				
IY-B	20.00			Recorded trial cost
Child centre/crèche venue costs (per hour)	15.00			Recorded trial cost
Other forms of contact				
Contact by telephone	6.00			PSSRU 2019 <sup>59</sup>
Contact by home visit	42.88			PSSRU 2010 <sup>63</sup>
Contact by letter	2.39			Gidlow <i>et al.</i> <sup>64</sup>
Contact by text	0.00			Assumed cost free as texts are already being sent

TABLE 25 Unit costs (continued)

E-SEE Steps programme resource	Unit cost (£)			Source
	Average	Minimum <sup>a</sup>	Maximum <sup>a</sup>	
<b>IY-T</b>				
<i>Between-programme fixed costs</i>				
Practitioner training costs				
IY trainer fees and expenses	3014.88	2146.00	4797.50	Recorded trial cost
Venue hire with catering	491.19	250.00	588.00	Recorded trial cost
Other equipment/materials	100.00			Recorded trial cost
IY-B	16.20			Recorded trial cost
Workshop manuals	20.00			Recorded trial cost
Programme trainee fees	16.20			Recorded trial cost
Leader manual and DVD set	750.00			Recorded trial cost
Telephone supervision of team leaders	1312.38	552.50	2687.00	Recorded trial cost
<i>Within-programme fixed costs</i>				
IY staff delivery costs (per hour)				
Mental health practitioners	45.00			PSSRU 2019 <sup>59</sup>
Early years worker/parenting co-ordinator	62.53			PSSRU 2018 <sup>65</sup>
Family support worker	56.84			PSSRU 2017 <sup>66</sup>
Child health worker	61.50			PSSRU 2018 <sup>65</sup>
Health visitor	72.56			PSSRU 2015 <sup>67</sup>
Community health nurse	71.46			PSSRU 2015 <sup>67</sup>
Clinical team lead	87.00			PSSRU 2019 <sup>59</sup>
Equipment/materials (for up to 12 participants)	200.00			Recorded trial cost
Catering (each session for up to 12 participants)	57.32			Recorded trial cost
<i>Programme variable costs</i>				
IY-B	20.00			Recorded trial cost
Child centre/crèche venue costs (per hour)	15.00			Recorded trial cost
Contact by telephone	6.00			PSSRU 2019 <sup>59</sup>
Contact by home visit	42.88			PSSRU 2010 <sup>63</sup>
Contact by letter	2.39			Gidlow <i>et al.</i> <sup>64</sup>
Contact by text	0.00			Assumed cost free as texts are already being sent
Child care				
Playgroup	2.50			Average child care costs: Money Advice Service <sup>59</sup>
Childminder	4.57			<i>Childcare Survey 2019</i> <sup>58</sup>
Family friend	4.57			Assumed equal to childminder cost
Day nursery	5.16			<i>Childcare Survey 2019</i> <sup>58</sup>
Parenting classes				
All parenting programmes	1501.00			PSSRU 2016 <sup>68</sup>

continued

TABLE 25 Unit costs (continued)

E-SEE Steps programme resource	Unit cost (£)			Source
	Average	Minimum <sup>a</sup>	Maximum <sup>a</sup>	
Health and social care				
GP surgery visit	33.00			PSSRU 2019 <sup>59</sup>
GP home visit	107.07			PSSRU 2010 <sup>63</sup>
GP doctor telephone call	15.32			PSSRU 2019 <sup>59</sup>
GP nurse surgery visit	5.69			PSSRU 2019 <sup>59</sup>
GP nurse home visit	31.44			PSSRU 2010 <sup>63</sup>
GP nurse telephone call	6.00			PSSRU 2019 <sup>59</sup>
Health visitor	93.35			NHS reference cost 2017/18 <sup>58</sup>
District nurse	39.42			NHS reference cost 2017/18 <sup>58</sup> (N02AF)
Other doctor appointment	81.08			NHS reference cost 2017/18 <sup>58</sup> (N29AF)
Psychiatrist (adult)	189.72			NHS reference cost 2017/18 <sup>58</sup> (724 – Perinatal Psychiatry)
Psychiatrist (child)	283.66			NHS reference cost 2017/18 <sup>58</sup> (MHSTOTHPLC)
Psychologist	142.82			NHS reference cost 2017/18 <sup>58</sup> (713 – Psychotherapy)
Other counsellor/therapist (adult)	190.57			NHS reference cost 2017/18 <sup>58</sup> (MHSTOTHPLA)
Other counsellor/therapist (child)	283.66			NHS reference cost 2017/18 <sup>58</sup> (MHSTOTHPLC)
Mental health nurse	92.00			PSSRU 2010 <sup>63</sup>
Accident and emergency visit	112.41			PSSRU 2010 <sup>63</sup>
Outpatient appointment	143.73			NHS Reference cost 2017/18 <sup>58</sup> (Outpatient – General Surgery)
Social worker (office visit)	20.25			PSSRU 2019 <sup>59</sup>
Social worker (home visit)	29.25			PSSRU 2019 <sup>59</sup>
Midwifery	81.61			NHS reference cost 2017/18 <sup>58</sup> (Outpatient – 560)
111 telephone call	13.87			Turner <i>et al.</i> <sup>69</sup>
Dentist/orthodontist	123.38			PSSRU 2019 <sup>59</sup>
Walk-in centre visit	47.15			NICE guideline 94 <sup>66</sup>
Dietitian	88.15			NHS reference cost 2017/18 <sup>58</sup> (A03)
General surgery	143.74			NHS reference cost 2017/18 <sup>58</sup> (A03)
Ambulance	190.53			National Audit Office 2017 <sup>70</sup>
Orthopaedic appointments	127.04			NHS reference cost 2017/18 <sup>58</sup> (Outpatient – 110)
Smoking cessation support	208.57			NHS reference cost 2017/18 <sup>58</sup> (Day case – DZ58Z)
Hearing assessment	83.77			NHS reference cost 2017/18 <sup>58</sup> (CA37A)

TABLE 25 Unit costs (continued)

E-SEE Steps programme resource	Unit cost (£)			Source
	Average	Minimum <sup>a</sup>	Maximum <sup>a</sup>	
Hospital admissions/procedures				
Hospital day neonatal	405.00			NHS reference cost 2017/18 <sup>58</sup> (PB06J–PB06M)
Hospital day paediatrician	314.97			NHS reference cost 2017/18 <sup>58</sup> (PX57A–PX57C)
Birth	2026.30			NHS reference cost 2017/18 <sup>58</sup> (NZ30A–C)
C-section birth	3809.72			NHS reference cost 2017/18 <sup>58</sup> (NZ50A–C)
Miscarriage	1955.37			NHS reference cost 2017/18 <sup>58</sup> (MB08A)
Knee fracture surgery	4575.33			NHS reference cost 2017/18 <sup>58</sup> (HE21B–HE21BD)
General renal disorders	3919.51			NHS reference cost 2017/18 <sup>58</sup> (LA09L)
Cleft lip and palate surgery	607.04			NHS reference cost 2017/18 <sup>58</sup> (CA66B)
Heart surgery	4336.82			NHS reference cost 2017/18 <sup>58</sup> (ED31A–ED31C)
Gall bladder surgery	3640.81			NHS reference cost 2017/18 <sup>58</sup> (GA10G)

GP, general practitioner.

<sup>a</sup> The minimum/maximum IY costs across participating sites (in cases where costs varied across sites).

training costs) and variable costs (i.e. those costs that are dependent on the number of recipients, e.g. IY-Bs) (see *Table 25*). Between-programme staff training costs were site specific and distributed across primary caregivers who attended at least one group session at each site. Weekly within-programme group session costs were apportioned across the number of attendees at each session. It was assumed that no costs were incurred when a session had zero attendance.

Costs were estimated from a public sector perspective, a family perspective (i.e. financial impacts on families) and a broader perspective that encompasses both public sector and family costs. It was assumed that child care and absent workday costs were family-related costs, whereas E-SEE Steps programme, primary care, secondary care, mental health, community service and social service costs fell on the public sector. Adult, child and overall costs were aggregated within and across data collection periods. The base-case cost-effectiveness analysis was conducted from a public sector perspective using aggregated adult and child costs over the trial's 18-month time horizon. Scenario analyses explored the impacts of alternative costing perspectives, applying the minimum, maximum and site-specific E-SEE Steps programme-related costs recorded at trial sites and controlling for baseline costs as a means of correcting for potential imbalances in baseline care requirements between arms.<sup>71</sup> In addition, sensitivity analyses considered the extent to which average programme costs and measures of cost-effectiveness vary with respect to the number of programmes a trained practitioner may be expected to deliver and the number of primary caregivers who may be expected to attend group sessions.

*Table 25* is unpacked by arm and more details regarding SAU parenting classes are provided in *Report Supplementary Material 4*.

### Outcomes

A range of HRQoL outcome measures for parents and children were collected throughout the trial and were each presented. Generic outcome measures were used for cost-consequence and cost-effectiveness analyses. A descriptive summary for each child and parental HRQoL outcome measure was conducted. To examine what factors mediated treatment effects, a comparative assessment of item-specific responses (e.g. PHQ-9: feeling down, depressed or hopeless) was conducted for outcomes that notably deviated between arms. The primary outcome used in the cost-consequence and cost-effectiveness frameworks was QALYs, a generic measure of health capturing survival and HRQoL. Cost-consequence analyses also compared PedsQL-adjusted life-years, a measure weighting survival according to mean PedsQL scores. An area under the curve approach using linear interpolation between time points was used to estimate QALYs and PedsQL-adjusted life-years.

The HRQoL of primary caregivers was measured at baseline, FU1, FU2 and FU3 using the EQ-5D-5L questionnaire, which asks individuals to rate their health (i.e. no problems, some problems, moderate problems, severe problems or extreme problems) across five dimensions (mobility, self-care, usual activity, pain/discomfort and anxiety/depression). Each combined rating defines a health state that has a preference-based weighting derived from a representative sample of the UK population.<sup>72</sup> Consistent with UK guidance for economic evaluations, HRQoL weights were estimated from a published mapping of EQ-5D-5L responses onto HRQoL values based on a UK population survey using the EuroQol-5 Dimensions, three-level version (EQ-5D-3L) instrument.<sup>73,74</sup> An alternative scenario calculated adult QALYs using the associated preference values for the EQ-5D-5L questionnaire. The PHQ-9 questionnaire was also assessed to better compare depression symptoms and monitor severity across nine known mental health disorders (*Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition criteria). Higher PHQ-9 scores are associated with decreased functional status and increased symptom-related difficulties.

Child HRQoL was estimated using the SDQ, a measure for assessing emotional and behavioural problems for children aged 2–4 years. The SDQ was completed by the primary caregiver at FU3. Preference-based HRQoL weights were derived from a published mapping of the SDQ onto the Child Health Utility (CHU-9D) questionnaire, a generic preference-based measure of paediatric HRQoL, and used to estimate child QALYs.<sup>74</sup> As an alternative assessment of child HRQoL, the PedsQL infant (ages 13–24 months) questionnaire was used, which asks primary caregivers to report infant-related problems with physical functioning, physical symptoms, emotional functioning, social functioning and cognitive functioning. The PedsQL was completed by the primary caregiver at FU3. PedsQL mean values range from 0 (reflecting constant problems) to 100 (perfect health status). Generic preference-based HRQoL weights were not calculated using the PedsQL because of fundamental differences between the PedsQL infant version used in the trial and the children and adolescent versions used in published mapping algorithms.<sup>75,76</sup> Instead, PedsQL scores themselves were used as weights. As the SDQ and PedsQL were not completed at baseline, it was assumed baseline values for both arms were equal to the mean SAU values at FU3.

### Analysis

The within-trial economic analysis assessed adult- and child-related resource use, costs by category, total costs, QALYs and PedsQL-adjusted life-years over the trial's 18-month time horizon. The cost-consequence analysis was conducted using observed differentials in costs and outcomes between trial arms. Cost-effectiveness results were obtained from an adjusted analysis, controlling for a set of relevant participant co-variables using regression methods. In line with UK guidelines, costs and QALYs were discounted at 3.5% per annum.<sup>72</sup>

All regression analyses controlled for treatment allocation, child age, adult age, child sex, highest qualification and parental relationship status. QALY regression analyses also controlled for baseline EuroQol-5 Dimensions (EQ-5D) (adults), baseline PHQ-9 and baseline ASQ:SE-2 to control for potential imbalances in mean utility values between arms.<sup>77</sup> Generalised linear models were used for the cost analysis, accounting for the non-normality and skewed nature of the cost data with a log-link

and gamma distribution. Ordinary least square regressions were applied for the QALY analysis. Missing cost, outcome and income data were populated via MI by chained equations using predicted mean matching, with Rubin's rules used for the subsequent analysis of multiple data sets.<sup>78,79</sup>

Cost-effectiveness results set out the adjusted mean cost and QALY estimates for adults, children and the aggregate (overall). The cost-effectiveness of the E-SEE Steps programme was assessed according to the ICERs and INHBs offered by the intervention compared with SAU. ICERs represent the cost per additional measure of benefit (QALY) of a strategy compared with the next best alternative.<sup>80</sup> INHBs capture the health gain from the strategy less the health that would have otherwise been generated elsewhere had the same resources been allocated for alternative purposes (e.g. the health generated by alternative social programmes). Positive INHB values represent health gains that exceed the health equivalent of the additional health-care costs (as defined by the cost-effectiveness threshold =  $k$ ).

The uncertainty in the regression estimates and overall decision uncertainty were estimated using Monte Carlo simulation, assuming multivariate normality of the regression coefficients.<sup>81</sup> Uncertainty estimates include 95% credible intervals around mean values (e.g. costs, QALYs and INHB) and the probability of the E-SEE Steps programme being the most costly, effective and cost-effective compared with SAU. The associated INHB of the E-SEE Steps programme, and the associated probability of being cost-effective, is presented for three measures of health opportunity cost (i.e. the health that would have been generated elsewhere using the same resources): £15,000 per QALY (i.e. the Department for Health and Social Care's chosen threshold)<sup>82</sup> and £20,000 and £30,000 per QALY (i.e. the range used by NICE).<sup>56</sup> Cost-effectiveness thresholds represent the marginal productivity of the health-care system, where interventions with ICERs below the threshold are deemed cost-effective when compared with the next best relevant comparator.

## Results

Income data had an elevated degree of missingness (22%). The characteristics of the participants who informed the economic analysis can be found in *Table 3*.

### Resource use and costs

#### Intervention-level costs

*Table 26* presents the overall average costs of delivering the programme, alongside the average cost for each component. The average costs of the E-SEE Steps programme (% total average costs) were as follows: £20 (4.3%) for the IY-B, £233.19 (50.8%) for training, £196.50 (42.9%) for group costs and £8.81 (1.9%) for IY contact, amounting to a total average cost of £458.50 per primary carer.

TABLE 26 Average E-SEE Steps programme costs by resource category (£ 2018/19)

Resource category	Average IY cost (£)	
	Mean	SD
IY-B	20.00	0.00
IY-I group costs	97.08	242.75
IY-T group costs	99.42	393.38
IY-I staff training	158.56	379.18
IY-T staff training	74.63	334.26
IY contact	8.81	21.98
Total	458.50	1021.23



Table 27 presents the average group-related costs for those primary caregivers who attended at least one session. The IY-I and IY-T group sessions had an average cost of £512 and £1232 per participant, respectively. The IY-I group session costs were largely staff related (69%), whereas IY-T group session costs were predominantly staffing (44%) and child care related (45%) (see Table 27). Average group costs were sensitive to the number of attendees, with E-SEE Step programme costs potentially inflated as a result of relatively low levels of participation. Training and group session costs were highly variable across sites. The total IY-I (IY-T) training-related costs at each site ranged between £2082 (£2765) and £5257 (£5514) because of a variety of site-specific contextual factors, including venue costs, catering and the accredited 2-day IY training course fees and expenses. Group costs also varied across sites as a result of differences in attendance, child care requirements and job profiles of those delivering the course(s).

### Trial resource use and costs

Table 28 reports the average costs by trial arm for adults, children and overall, in total, and for each resource category. Report Supplementary Material 5 provides a complete-case analysis breakdown of resource use by treatment arm, individual (i.e. adult/child) and resource category. Report Supplementary Material 5 also reports all estimated costings by follow-up period.

Intervention-related resource use followed protocol, such that 100% of E-SEE Step participants received the IY-B. Approximately 20% ( $n = 54$ ) of those participants randomised to the E-SEE Steps programme went on to attend an IY-I group session, whereas only 8% ( $n = 23$ ) attended an IY-T group session. Contact with participants throughout the IY-I and IY-T was primarily conducted by telephone and text, with an average of four telephone calls and 3.2 texts made per family. Contact by letter or home visit was relatively rare for IY attendees, with an average of 0.4 letters sent and 0.2 visits made per family. Primary care was utilised more frequently by children than caregivers during the trial. General practitioner (GP) and nurse contacts were relatively balanced between trial arms, albeit with children in the IY arm having 1.06 fewer GP visits, on average, than children in the SAU arm. Secondary care resource use was highest for the E-SEE Steps programme, with average adult and child inpatient nights (adult: E-SEE Steps programme 0.34 vs. SAU 0.04; child: E-SEE Steps programme 0.41 vs. SAU 0.25) and other hospital-based services (adult: E-SEE Steps programme 0.84 vs. SAU 0.57; child: E-SEE Steps programme 0.62 vs. SAU 0.33) above those observed in SAU families. Mental health-care resource use was rare for adults and children. However, adults in the SAU arm had 0.48 more contacts with other counselling or therapy (i.e. non-psychiatrist/non-psychologist) than those in the E-SEE Steps programme (E-SEE Steps programme 0.44 vs. SAU 0.92). Community service uptake was generally higher for those in the E-SEE Steps programme, where adult and child district nurse visits, midwife sessions, parenting courses, other community services and child health visitors were more common than with SAU. Participants in both arms rarely reported resource use associated with social services (adult: E-SEE Steps programme 0.00 vs. SAU 0.08; child: E-SEE Steps programme 0.05 vs. SAU 0.08). Child care hours were lower in the E-SEE Steps arm than in the SAU arm, with childminder

TABLE 27 Average IY-I and IY-T group session costs by resource category (£ 2018/19)

Average attendee cost	Group costs (£)			
	IY-I ( $n = 54$ )		IY-T ( $n = 23$ )	
	Mean	SD	Mean	SD
Total average	512.36	313.13	1231.96	724.56
Staff costs	354.10	213.33	543.68	460.78
Catering	36.09	50.43	95.00	113.77
Handouts	38.83	17.43	42.50	28.99
Child centre costs	83.33	185.54	550.78	378.01



TABLE 28 Adult- and child-related costs by resource category (£ 2018/19)

Resource category	IY						SAU					
	Adult		Child		Overall		Adult		Child		Overall	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Public sector perspective												
IY	458.50	1021.23	0.00	0.00	458.50	1021.23	0.00	0.00	0.00	0.00	0.00	0.00
Primary care	135.97	173.63	225.12	225.48	361.09	324.01	132.17	163.57	248.75	220.36	380.92	300.64
Secondary care	458.66	1359.10	417.77	830.04	876.43	1664.99	434.64	1602.20	387.14	793.32	821.78	1844.70
Mental health care	205.58	1131.00	23.97	281.52	229.55	1199.39	365.54	1603.58	0.00	0.00	365.54	1603.58
Community-based services	323.02	727.41	548.37	803.07	871.39	1268.97	251.34	699.68	413.63	430.47	664.96	915.96
Social services	15.65	104.95	12.77	101.01	28.42	203.23	13.14	93.15	16.58	101.57	29.71	193.43
Total cost	1597.37	2886.56	1228.00	1543.07	2825.37	3709.42	1196.82	3290.44	1066.09	975.85	2262.92	3650.21
Family perspective												
Child care	0.00	0.00	1137.19	1054.22	1137.19	1054.22	0.00	0.00	1292.30	1435.30	1292.30	1435.30
Absent days from work	517.93	4433.31	0.00	0.00	517.93	4433.31	399.61	951.63	0.00	0.00	399.61	951.63
Total cost	517.93	4433.31	1137.19	1054.22	1655.12	4621.61	399.61	951.63	1292.30	1435.30	1691.91	2066.91
Broader perspective												
Total cost	2115.30	5096.83	2365.19	1965.73	4480.49	5869.73	1596.43	3395.34	2358.40	1674.23	3954.83	4031.70

hours (E-SEE Steps programme 39 vs. SAU 59), friend and family hours (E-SEE Steps programme 96 vs. SAU 107) and day nursery hours (E-SEE Steps programme 98 vs. SAU 106) all lower in the E-SEE Steps arm over the trial's 18-month follow-up.

The average total costs from a public sector perspective were found to be higher in the E-SEE Steps arm than in the SAU arm for adults (E-SEE Steps programme £1598 vs. SAU £1197), children (E-SEE Steps programme £1228 vs. SAU £1067) and overall (E-SEE Steps programme £2826 vs. SAU £2263), with incremental costs of £401 (95% CI -£434 to £1236), £161 (95% CI -£244 to £568) and £563 (95% CI -£470 to £1595), respectively. Differentials in average public sector mental health-related costs (£136, 95% CI -£491 to £219; E-SEE Steps programme £230 vs. SAU £366), primary care-related costs (-£20, 95% CI -£111 to £71; E-SEE Steps programme £362 vs. SAU £381) and social services-related costs (-£2, 95% CI -£66 to £64; E-SEE Steps programme £29 vs. SAU £30) were higher for SAU than for the E-SEE Steps programme, whereas intervention-related costs [E-SEE Steps programme £459 (95% CI £190 to £728) vs. SAU £0 (95% CI £0 to £0)], community services-related costs (£207, 95% CI -£143 to £555; E-SEE Steps programme £872 vs. SAU £665) and secondary care-related costs (£55, 95% CI -£415 to £525; E-SEE Steps programme £877 vs. SAU £822) were higher for those in the E-SEE Steps programme.

Primary care- and community-based services were predominantly child related, with 64% and 63% of costs associated with child admissions and child services, respectively. Our planned substudy exploring predictors of emergency department (ED) attendance and hospital admissions (see *Appendix 10*) highlighted that study site, younger gestational age, older age at recruitment to trial and mothers' poorer mental health, younger age and ED attendance predicted child ED attendance (see *Appendix 10, Tables 44 and 45*). Health problems or injury since birth predicted hospital admission. For mothers, ED attendance was predicted by mixed ethnic origin, having a boy, having poorer quality of life at baseline and having been hospitalised during the trial. Mothers' admission to hospital was predicted by ED attendance for themselves and being anxious or depressed. Mental health service costs were mainly adult related (96%), whereas secondary care costs (adult, 53%; child, 47%), social services costs (adult, 50%; child, 50%) and overall public costs (adult, 55%; child, 45%) were relatively balanced between adult- and child-related costs.

Average total costs from a family perspective were comparable between arms (E-SEE Steps programme £1655 vs. SAU £1692), with the E-SEE Steps programme associated with an average cost saving of £37 (95% CI -£956 to £1029). Child care costs were reduced by £155 (95% CI -£165 to £475), on average, in the E-SEE Steps arm (E-SEE Steps programme £1138 vs. SAU £1,293) compared with SAU. Alternatively, average absent workday costs were £118 [95% CI -£790 to £1027] higher for the intervention (E-SEE Steps programme £518 vs. SAU £400) than for SAU.

Combined health and social care, child care and absent workday costs were relatively similar between the E-SEE Steps and SAU arms, with non-intervention E-SEE Steps programme costs for adults (E-SEE Steps programme £1657 vs. SAU £1597) and children (E-SEE Steps programme £2365 vs. SAU £2358) comparable to SAU over 18 months. Taking a broader costing perspective (relating to both families and the public sector), costs totalled £4481 for those in the E-SEE Steps arm (adult, £2116; child, £2366) and £3,955 for SAU (adult, £1597; child, £2359), making for an average total cost differential of £526 (95% CI -£939 to £1990) per family. The primary factor in the E-SEE Steps programme's higher average cost was the cost of the intervention itself, with cost savings in some resource categories (e.g. mental health costs) counterbalanced with additional costs in others (e.g. community-based services).

### Health outcomes

Table 29 presents mean values for each health outcome for adults and children at each available follow-up period. PedsQL-adjusted life-years and QALY values reported at each follow-up period represent those accrued since the previous time point. Figure 8 provides a detailed breakdown of adult responses to each of the PHQ-9 and EQ-5D-5L questionnaire domains.

TABLE 29 Adult- and child-related outcomes by treatment arm

Outcome	Treatment arm								Difference			
	E-SEE Steps programme				SAU				Baseline	FU1	FU2	FU3
	Baseline	FU1	FU2	FU3	Baseline	FU1	FU2	FU3				
Adult HRQoL												
PHQ-9	3.06316	2.39439	2.48351	2.96632	2.83929	3.34286	2.78571	3.0500	0.22387	-0.94847	-0.30221	-0.08368
EQ-5D-5L	0.93585	0.94729	0.94758	0.93457	0.91457	0.91970	0.91534	0.92878	0.02127	0.02759	0.03224	0.00579
(Mapped) EQ-5D-3L	0.89386	0.91285	0.91859	0.89645	0.86446	0.87181	0.87721	0.89115	0.02940	0.04105	0.04137	0.00530
QALYs <sup>a</sup>		0.15047	0.53392	0.68113		0.14475	0.51001	0.66229		0.00572	0.02390	0.01884
Child HRQoL												
PedsQL: overall				86.54722				87.16596				-0.61874
PedsQL: psychosocial				83.86168				84.37054				-0.50886
PedsQL: physical				90.22215				90.95386				-0.73171
SDQ score				9.67293				9.15094				0.52199
(Mapped) CHU-9D				0.85324				0.85933				-0.00609
PedsQL-adjusted life-years <sup>b</sup>		0.14482	0.50596	0.65026		0.14528	0.50847	0.65374		-0.00046	-0.00251	-0.00348
QALYs <sup>b</sup>		0.14265	0.49928	0.64194		0.14311	0.50090	0.64402		-0.00046	-0.00162	-0.00208

a Accrued from previous time point and calculated from EQ-5D-3L scores mapped from EQ-5D-5L responses.

b Accrued from previous time point, assuming that child baseline HRQoL in both arms is equal to that reported in the SAU arm at the 18-month follow-up.

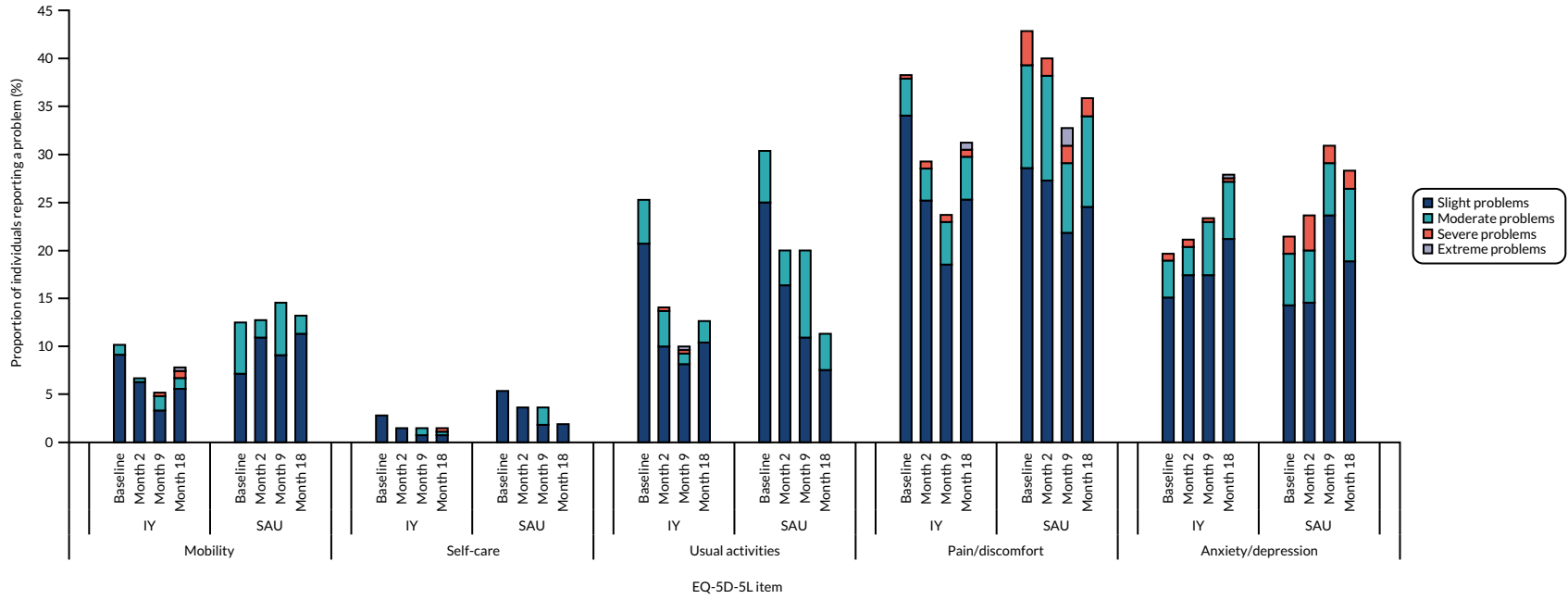


FIGURE 8 Participant responses to the EQ-5D-5L questionnaire by follow-up period and treatment arm.

Primary caregivers randomised to the E-SEE Steps arm reported higher baseline PHQ-9 scores (3.06) than the SAU arm (2.84), with an average difference of 0.22 points between arms. Baseline PHQ-9 responses identified tiredness (E-SEE Steps programme 71.2% vs. SAU 69.6% reporting any problems), eating issues (E-SEE Steps programme 33.3% vs. SAU 35.7% reporting any problems) and sleeping difficulties (E-SEE Steps programme 27.7% vs. SAU 19.6% reporting any problems) as the three most common problems experienced by primary caregivers (see *Report Supplementary Material 5*). All subsequent follow-up periods found PHQ-9 scores to be higher for SAU participants than for E-SEE Step participants, with average differences of 0.95, 0.30 and 0.08 points at FU1 (E-SEE Steps programme 2.39 vs. SAU 3.34), FU2 (E-SEE Steps programme 2.48 vs. SAU 2.79) and FU3 (E-SEE Steps programme 2.97 vs. SAU 3.05), respectively. E-SEE Steps programme-associated reductions in depression and severity scores were concentrated at FU1 and FU2, with modest reductions in the proportion of participants reporting problems with apathy, depression and eating-related issues (see *Figure 8*). Average PHQ-9 scores were similar between arms by FU3, with E-SEE Steps participants seeing modest reductions in the proportions reporting problems with apathy (E-SEE Steps programme 18.8% vs. SAU 24.5%) and eating (E-SEE Steps programme 28.6% vs. SAU 35.8%), balanced with modest elevations in the proportions reporting issues with tiredness (E-SEE Steps programme 66.4% vs. SAU 58.5%) and concentration (E-SEE Steps programme 20.1% vs. SAU 13.2%).

Baseline EQ-5D-5L scores for primary caregivers were higher in the E-SEE Steps arm than in the SAU arm, with an average HRQoL score of 0.936 in the E-SEE Steps arm compared with 0.915 in the SAU arm. Increased EQ-5D-5L scores were observed at FU1 (0.947) and FU2 (0.948) compared with baseline levels (0.936) for those randomised to the E-SEE Steps programme. Scores did, however, return below baseline values at FU3 (0.935). For those participants randomised to SAU, EQ-5D-5L scores fluctuated between time points, rising at FU1 (0.920), falling at FU2 (0.915) and finally rising again at FU3 (0.929). The differential in EQ-5D-5L scores between the E-SEE Steps programme and SAU rose from a 0.021 (95% CI -0.004 to 0.046) increment at baseline to 0.027 (95% CI 0.002 to 0.052) and 0.032 (95% CI 0.001 to 0.063) at FU1 and FU2, respectively. EQ-5D-5L scores converged at FU3, with a differential score of 0.006 (95% CI -0.026 to 0.038). Results using mapped EQ-5D-3L were aligned with those observed using the EQ-5D-5L, albeit with notably lower values and increased differentials between arms, reflecting differences between the EQ-5D-3L and EQ-5D-5L value sets. Mapped EQ-5D-3L scores were higher for the E-SEE Steps programme than for SAU, with differential scores of 0.029 (95% CI -0.006 to 0.064) (E-SEE Steps programme 0.894 vs. SAU 0.864), 0.041 (95% CI 0.006 to 0.076) (E-SEE Steps programme 0.913 vs. SAU: 0.872), 0.041 (95% CI 0.001 to 0.081) (E-SEE Steps programme 0.919 to SAU 0.877) and 0.005 (95% CI -0.040 to 0.050) (E-SEE Steps programme 0.896 vs. SAU 0.891) at baseline, FU1, FU2 and FU3, respectively. E-SEE Steps programme-associated improvements in EQ-5D-5L- and mapped EQ-5D-3L-defined HRQoL were concentrated at FU1 and FU2. Improvements stemmed from reductions in the proportion of primary caregivers reporting problems with mobility and usual activities, and for reductions in the rates and severity of pain/discomfort and anxiety/depression problems reported. Given that no mortality was observed during trial follow-up, differences in mapped EQ-5D-3L HRQoL scores explain the QALY differentials observed. Those participants randomised to the E-SEE Steps programme accrued 0.006 (E-SEE Steps programme 0.150 vs. SAU 0.145), 0.024 (E-SEE Steps programme 0.534 vs. SAU 0.510) and 0.019 (E-SEE Steps programme 0.681 vs. SAU 0.662) more QALYs than SAU between baseline and FU1, FU1 and FU2 and FU2 and FU3, respectively. E-SEE Steps programme-associated QALY gain was highest between FU1 and FU2 (i.e. the periods with the highest increments in mapped EQ-5D-3L HRQoL scores compared with SAU). We found that the E-SEE Steps programme was associated with moderate gains in HRQoL, primarily at FU1 and FU2. Differences between arms were multidimensional, happening in and across a variety of health domains defined by the EQ-5D-5L and PHQ-9.

Children of primary caregivers randomised to the E-SEE Steps arm had lower PedsQL scores overall and across psychosocial and physical domains than children of primary caregivers randomised to SAU. Compared with SAU, PedsQL responses for E-SEE Steps participants had 0.619 (E-SEE Steps programme 86.5 vs. SAU 87.2), 0.509 (E-SEE Steps programme 83.9 vs. SAU 84.4) and 0.732 (E-SEE Steps programme

90.2 vs. SAU 91.0) decrements in overall, psychosocial and physical summary scores, respectively. This represented a modest reduction in parent-reported paediatric psychosocial and physical HRQoL. SDQ scores were also found to show modest decrements in child outcomes, with E-SEE Steps primary caregivers reporting a 0.52 increase (E-SEE Steps programme 9.67 vs. SAU 9.15) in child SDQ distress and impairment scores compared with SAU. At FU3, mapped CHU-9D scores were lower in the E-SEE Steps arm than in the SAU arm, with an associated 0.006 (95% CI -0.005 to 0.018) (E-SEE Steps programme 0.853 vs. SAU 0.859) decrement in HRQoL. Resultantly, those randomised to E-SEE Steps arm accrued 0.0005 (E-SEE Steps programme 0.1427 vs. SAU 0.1431), 0.002 (E-SEE Steps programme 0.499 vs. SAU 0.510) and 0.002 (E-SEE Steps programme 0.642 vs. SAU 0.644) fewer QALYs than SAU between baseline and FU1, FU1 and FU2 and FU2 and FU3, respectively. PedsQL-adjusted life-years were higher than mapped CHU-9D QALY estimates and had larger differences between arms. Children of participants randomised to the E-SEE Steps programme accrued 0.0005 (E-SEE Steps programme 0.1448 vs. SAU 0.1453), 0.003 (E-SEE Steps programme 0.506 vs. SAU 0.508) and 0.004 (E-SEE Steps programme 0.650 vs. SAU 0.654) fewer PedsQL-adjusted life-years than children of participants randomised to SAU between baseline and FU1, FU1 and FU2 and FU2 and FU3, respectively. We found no evidence of the E-SEE Steps programme enhancing child HRQoL compared with SAU.

### Cost-consequence analysis

Table 30 presents the mean incremental adult- and child-related public sector costs of the E-SEE Steps programme and the mean differences (i.e. consequences) in adult and child health outcomes over the trial's 18-month time horizon. The table also includes the associated ICER for each outcome measure.

Adult-related public sector costs were, on average, £401 higher in the E-SEE Steps arm than in the SAU arm. Adult QALYs were 0.047 higher in the E-SEE Steps arm than in the SAU arm when using mapped EQ-5D-3L scores and 0.035 higher when using EQ-5D-5L preference weights. The E-SEE programme had corresponding ICERs of approximately £8449 (EQ-5D-3L score) and £11,300 (EQ-5D-5L score).

Average child-related public sector costs were £162 higher in the E-SEE Steps arm than in the SAU arm. Children randomised to the E-SEE Steps arm were estimated to have 0.004 (E-SEE Steps programme 1.284 vs. SAU 1.288) fewer QALYs and 0.006 (E-SEE Steps programme 1.301 vs. SAU: 1.307) fewer PedsQL-adjusted life-years over the trial's 18-month time horizon. Consequently, the E-SEE Steps programme was found to be a dominated strategy for children, given additional child-related costs and worse health outcomes compared with SAU.

### Cost-effectiveness analysis

Table 31 reports the adjusted base-case mean costs, QALYs and cost-effectiveness results for the E-SEE Steps programme for adults, children and overall (i.e. adults and children) over the time horizon of the trial. *Report Supplementary Material 5* presents all regression results used to inform the cost-effectiveness analysis.

TABLE 30 Within-trial cost-consequence analysis

Outcome	Incremental costs (£)	Consequence	ICER (£)
Adult			
QALYs: mapped EQ-5D-3L	400.55	0.047	8448.39
QALYs: EQ-5D-5L	400.55	0.035	11,298.70
Child			
PedsQL-weighted life-years	161.90	-0.00645	Dominated
QALYs: mapped CHU-9D	161.90	-0.00416	Dominated

TABLE 31 Within-trial cost-effectiveness analysis

Cost-effectiveness	Costs (£) (95% CI) [probability of most costly]	QALYs (95% CI) [probability of most effective]	ICER (£)	INHB (95% CI) [probability of being cost-effective]		
				Cost-effectiveness threshold = £15,000	Cost-effectiveness threshold = £20,000	Cost-effectiveness threshold = £30,000
Overall						
SAU	1988.61 (1465.79 to 2615.43) [0.037]	2.58680 (2.54927 to 2.62129) [0.06]		[0.640]	[0.512]	[0.332]
E-SEE Steps programme	2609.46 (2312.07 to 2951.04) [0.963]	2.61775 (2.60252 to 2.6342) [0.94]	20,061.02	-0.01044 (-0.07207 to 0.05246) [0.360]	-0.00009 (-0.07207 to 0.05246) [0.488]	0.01025 (-0.036 to 0.0574) [0.668]
Adult						
SAU	942.44 (604.51 to 1461.11) [0.052]	1.31392 (1.27465 to 1.35166) [0.044]		[0.450]	[0.316]	[0.192]
E-SEE Steps programme	1388.26 (1142.31 to 1639.19) [0.948]	1.34818 (1.33322 to 1.36373) [0.956]	13,010.68	0.004544 (-0.04636 to 0.06414) [0.550]	0.01197 (-0.04636 to 0.06414) [0.684]	0.01940 (-0.024 to 0.06746) [0.808]
Child						
SAU	1000.28 (746.07 to 1322.25) [0.143]	1.27420 (1.26722 to 1.28191) [0.868]		[0.913]	[0.929]	[0.940]
E-SEE Steps programme	1177.33 (1034.65 to 1340.39) [0.856]	1.26957 (1.26629 to 1.27261) [0.132]	Dominated	-0.01644 (-0.03847 to 0.0087) [0.087]	-0.01349 (-0.03847 to 0.0087) [0.071]	-0.01053 (-0.03847 to 0.0087) [0.060]



Overall, the E-SEE Steps programme was found to have higher costs and to be more effective than SAU. Total costs were estimated at £2610 and £1989 for the E-SEE Steps programme and SAU, respectively, giving the E-SEE Steps arm an incremental cost of approximately £621 per family. The E-SEE Steps programme was found to have an incremental QALY gain of 0.031 (E-SEE Steps programme 2.618 vs. SAU 2.587) compared with SAU, and was associated with a 96% probability of being the costliest strategy and a 94% probability of being the most effective strategy. The incremental cost and QALY estimates for the E-SEE Steps programme resulted in an ICER of approximately £20,062 per QALY compared with SAU. INHBs were only positive at a cost-effectiveness threshold of £30,000, with the probability of being cost-effective estimated to be 36%, 48.8% and 66.8% for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively. INHB 95% credible intervals fell below zero for all cost-effectiveness thresholds considered.

For primary caregiving adults, the E-SEE Steps programme was found to be the costliest and the most effective strategy. Adult-related costs were estimated at approximately £1389 for the E-SEE Steps programme and £943 for SAU, giving the E-SEE Steps arm an incremental cost of approximately £446. Estimated costs were lower and cost differential between arms higher than those observed when not controlling for relevant covariables. The QALY gain associated with the E-SEE Steps programme was estimated at approximately 0.034 (E-SEE Steps programme 1.348 vs. SAU 1.314) compared with SAU, notably lower than that calculated without controlling for relevant covariables (0.047). The E-SEE Steps programme was calculated as having a 95% probability of being the costliest strategy and a 96% probability of being the most effective strategy. The associated E-SEE Steps ICER was calculated as £13,011 per QALY in adults. INHBs were positive at all cost-effectiveness thresholds considered, with the probability of being cost-effective estimated to be 55%, 68.4% and 80.8% for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively. INHB 95% credible intervals fell below zero for all cost-effectiveness thresholds considered.

For children, the E-SEE Steps programme was found to have higher costs and to be less effective than SAU. Total child-related costs were estimated at £1178 and £1001 for the E-SEE Steps programme and SAU, respectively, making for an incremental cost of approximately £178 per child. As for adults, estimated costs were lower and cost differential between arms higher than those observed when not controlling for relevant covariables. The E-SEE Steps programme was estimated as having an incremental QALY decrement of approximately 0.005 (E-SEE Steps programme 1.270 vs. SAU 1.274), an 85.6% probability of being the costliest strategy and a 13.2% probability of being the most effective strategy. SAU was less costly and more effective than the E-SEE Steps programme and was, therefore, deemed a dominant strategy in children. INHBs for the E-SEE Steps programme were negative at all cost-effectiveness thresholds considered, with the probability of being cost-effective at 8.7%, 7.1% and 6.0% for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively.

The 95% credible intervals for adult, child and overall costs and QALYs had overlapped considerably between the E-SEE Steps programme and SAU, suggesting a significant level of uncertainty in the incremental costs and outcomes the E-SEE Steps programme may be expected to deliver.

### Scenario analyses

Table 32 presents the overall cost, QALY and cost-effectiveness estimates for a set of scenario analyses, along with base-case results for comparison. *Report Supplementary Material 5* presents a more detailed assessment for each scenario.

When using EQ-5D-5L preference weights to calculate adult QALYs made, QALY estimates were notably higher, raising overall values from 2.618 to 2.668 for the E-SEE Steps programme and from 2.587 to 2.645 for SAU. The incremental QALY gain associated with the E-SEE Steps programme was reduced from 0.031 to 0.023, which, in turn, increased the ICER to £27,068 per QALY. Using the EQ-5D-5L value set gave the E-SEE Steps programme a 23.7%, 34.2% and 53.9% probability of being cost-effective for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds considered, respectively.



TABLE 32 Cost-effectiveness scenario analyses

Scenario	Cost (£)	QALYs	ICER (£)
Base case			
SAU	1988.61	2.58680	
IY	2609.46	2.61775	20,061
Controlling for baseline costs			
SAU	1775.42	2.58680	
IY	2589.71	2.61775	26,312
12 participants per IY-I and IY-T group			
SAU	1988.61	2.58680	
IY	2297.85	2.61775	9992
EQ-5D-5L			
SAU	1988.61	2.64511	
IY	2609.46	2.66805	27,068
Broader perspective			
SAU	3727.13	2.58680	
IY	4180.62	2.61775	14,653
Site 1 site costs			
SAU	1988.61	2.58680	
IY	2595.88	2.61775	19,622
Site 2 site costs			
SAU	1988.61	2.58680	
IY	2564.65	2.61775	18,613
Site 3 site costs			
SAU	1988.61	2.58680	
IY	2681.61	2.61775	22,392
Site 4 site costs			
SAU	1988.61	2.58680	
IY	2575.83	2.61775	18,974
Minimum site costs			
SAU	1988.61	2.58680	
IY	2515.23	2.61775	17,016

In the 2 months leading up to trial baseline, the SAU arm accumulated an average public sector cost of approximately £1355 per family (i.e. £168 above the average accrued in the E-SEE Steps arm during the same period). Controlling for baseline cost, imbalances can capture unobservable factors that impact trial costs, but are unrelated to the intervention under investigation (e.g. not conflating the impacts of differential baseline care needs with intervention allocation). Average SAU cost estimates fell from a base-case cost of £1989 to approximately £1776 when controlling for baseline values. The estimated E-SEE Step programme costs remained largely unchanged, meaning that incremental costs of the E-SEE Steps programme rose from £621 to £815 compared with SAU. Controlling for baseline costs gave the E-SEE Steps programme a 99.7% probability of being the costliest strategy and an ICER of £26,312 per QALY. In addition, the E-SEE Steps programme had a 20.4%, 35.0% and 57.0% probability of being cost-effective for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively.

Considering a broader perspective (i.e. both public sector costs and family-related costs) increased the estimated mean costs by approximately £1572 and £1739 in the E-SEE Steps arm (total £4181) and SAU arm (total £3728), respectively. The E-SEE Steps programme remained the costliest strategy, albeit with incremental costs reduced relative to the base-case analysis (from £621 to £432) due to E-SEE Steps programme-related savings in child care costs. Taking a broader costing perspective found the E-SEE Steps programme to have an ICER of £14,654, positive net-health benefits at all cost-effectiveness thresholds considered, an 85% probability of being the costliest strategy and a 49.6%, 61.1% and 73% probability of being cost-effective at the £15,000, £20,000 and £30,000 cost-effectiveness thresholds considered, respectively.

Participation in the group elements of the IY-I and IY-T programme was highly variable both within and between sites. As participation was relatively low for both IY-I and IY-T programmes, we explored an alternative best-case scenario where the E-SEE Steps programme would maintain the same proportion of participants attending sessions (to align with trial efficacy), but those participants who do attend are assumed to be in groups at full capacity (i.e. 12 families). In this scenario, within-programme costs (e.g. staff time running group sessions) and between-programme costs (e.g. staff training) would be distributed across the largest group possible for any given programme. When assuming full capacity group sizes, the E-SEE Steps programme was associated with a reduced overall cost (£2298) and an incremental cost of £312 compared with SAU, approximately half base-case values (£621). The E-SEE Steps programme had an 82.5% probability of being the costliest strategy, an ICER of £9993 and reported positive INHBs at each cost-effectiveness threshold considered. The probabilities of being cost-effective were 62.2%, 71.8% and 81.9% for the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively.

The cost of providing the IY-I and IY-T courses varied between sites, as training expenses and practitioner pay bands differed across sites. Taking the minimum observed site-specific costs reduced the E-SEE Steps programme's estimated mean cost from £2610 to £2516 and reduced incremental costs from £621 to £527. Resultantly, the E-SEE Steps programme had a 94.4% probability of being the costliest strategy, an ICER of £17,017 and a 44.1%, 56.5% and 71.5% probability of being cost-effective at the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively. Alternatively, taking the maximum training and practitioner costs raised the E-SEE Steps programme's estimated costs by £117 (from £2610 to £2726) and incremental costs from £621 to £738. When using the highest observed site costs, the E-SEE Steps programme was found to have a 98.2% probability of being most costly, an ICER of £23,820, positive INHB at only the highest £30,000 per QALY cost-effectiveness threshold and had a probability of 26.2%, 39%, and 60% probability of being cost-effective at the £15,000, £20,000 and £30,000 cost-effectiveness thresholds, respectively. Site-specific results were bounded between the minimum and maximum costing scenarios. Using site-specific programme costs generated ICERs of £19,623, £18,614, £18,975 and £22,393 for sites 1–4, respectively.

### **Sensitivity analyses**

Figures 9 and 10 examine the average programme costs and measures of cost-effectiveness with respect to the number of programmes trained practitioners may be expected to deliver (between programme), and the number of attendees group sessions may be expected to reach (within programme), respectively. As the trial ran only a single E-SEE Steps programme, between-programme fixed costs (training costs, materials costs, etc.) were distributed over a single programme only. In addition, group session within-programme fixed costs and between-programme fixed costs were distributed over a relatively low and highly variable IY-I and IY-I course attendance. Average within-trial programme costs may, therefore, not accurately represent those borne if the E-SEE Steps programme was to be introduced nationwide, where more programmes may be run by trained practitioners for alternative group sizes.

Figure 9 shows that average programme costs, and corresponding ICERs, fall at a decreasing rate with respect to the number of programmes delivered. In this costing, average programme costs were sensitive to initial expansions in programme delivery, as within-trial values (£486 per family) reduced

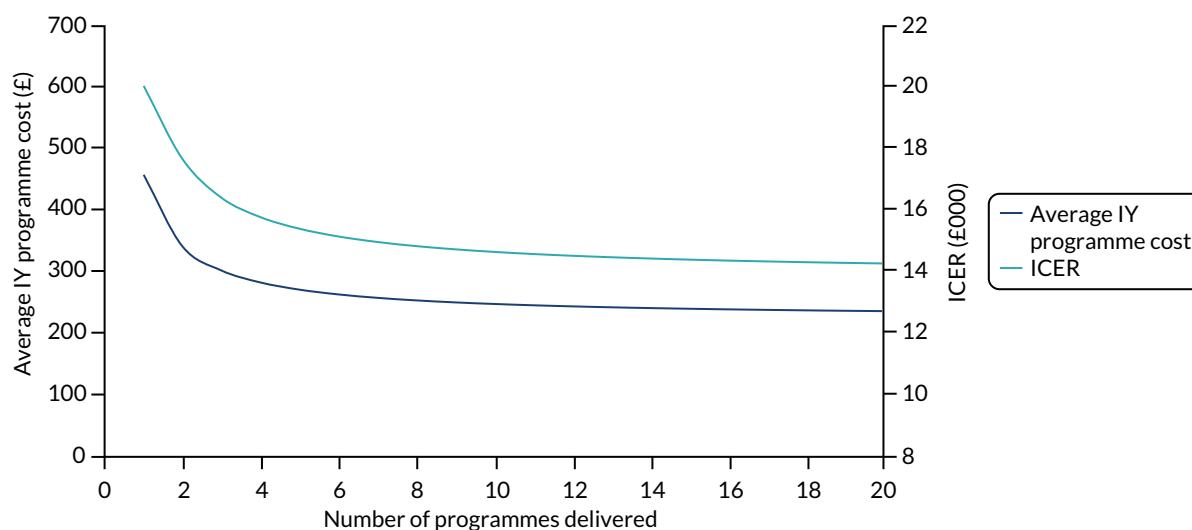


FIGURE 9 Sensitivity of IY programme costs and measures of cost-effectiveness with respect to the number of future programmes delivered.

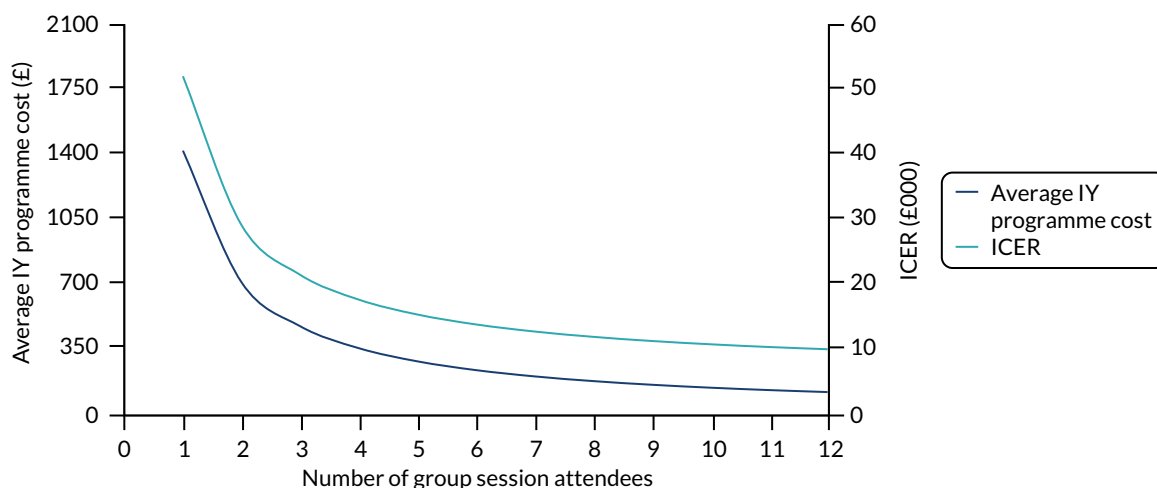


FIGURE 10 Sensitivity of IY programme costs and measures of cost-effectiveness with respect to the number of IY group session attendees.

to £342, £304 and £284 when two, three and four programmes were delivered, respectively. Cost-effectiveness was also particularly sensitive to initial expansions in the number of programmes delivered, with corresponding ICERs of approximately £17,680, £16,424 and £15,797, respectively. Average intervention costs plateaued within £235–250 and measures of cost-effectiveness within £14,300–14,700 for between 10 and 20 deliverable programmes.

Figure 10 shows positive but diminishing reductions in average programme costs and measures of cost-effectiveness with respect to the number of IY-I and IY-T group session attendees. Results were particularly sensitive to changes in attendances, with average programme-specific costs reducing from £1423 per family when provided on an individual basis to £711, £475 and £356 per family when two, three and four caregivers attend group sessions. Corresponding cost-effectiveness estimates were an ICER of £52,110 for individual course attendance and ICERs of £29,137, £21,479, £17,650 for two, three and four attendees, respectively. Average intervention costs plateaued around £118–178 and ICERs around £9993–11,906 for attendances of between 8 and 12 participants.

## Discussion

### *Principal findings*

Our results show considerable uncertainty around the cost-effectiveness of the E-SEE Steps intervention when the time period of the trial is considered. The mean incremental public sector cost of the E-SEE Steps programme was found to be £621 per family compared with SAU. There was a small gain in mean QALYs, with positive increments in adults exceeding minor decrements reported in child outcomes over the trial period. The E-SEE programme had an ICER of £20,062, which is above the minimum threshold recommended by NICE of £20,000 per QALY, resulting in the INHB of introducing the E-SEE Steps programme being negative in the base-case scenario, with the probability of the programme being cost-effective at approximately 48.8%. In contrast, at the maximum threshold (i.e. £30,000/QALY), the INHB of introducing the E-SEE Steps programme was positive (0.01 QALYs), with the probability of the programme being cost-effective at approximately 66.8%. Findings were sensitive to changes in the key assumptions used in the analysis. All scenarios found the E-SEE Steps programme to be cost-effective at the maximum threshold recommended by NICE of £30,000 per QALY.

### *Cost-effectiveness and longer-term impacts*

The broader impacts and the longer-term implications of the E-SEE Steps programme may be important in determining the overall cost-effectiveness of the programme. However, results from the trial suggest that any impact on future child conduct disorder would favour SAU, as childhood behavioural outcomes were worse in the E-SEE Steps group. Extrapolating these results into the future may increase costs and worsen outcomes associated with the E-SEE Steps programme, rendering it unlikely to be cost-effective compared with SAU. In addition, the positive incremental parental outcomes associated with the E-SEE Steps programme converged between arms at final follow-up, suggesting that the programme offered only shorter-term parental benefits. The differentials observed in parental HRQoL were relatively modest, short in duration and informed by a trial not powered to detect differences in HRQoL. Conclusions regarding the cost-effectiveness of E-SEE Steps are largely contingent on such differentials in parental HRQoL estimates and must, therefore, be interpreted with caution.

## Chapter 6 Discussion

This RCT of an intervention delivered in a proportionate universal approach was designed to (1) evaluate whether or not the 'E-SEE Steps' programme (i.e. the IY programmes delivered in a proportionate universal approach) had an impact on clinical outcomes; (2) evaluate the processes around service delivery; and (3) assess cost-effectiveness.

We tested the effectiveness of a proportionate universal model (E-SEE Steps) with three components. All intervention families received the IY-B as a universal element; we then applied a needs-based approach using the ASQ:SE-2 and the PHQ-9 as screeners at FU1 and FU2, offering parents the IY-I and IY-T group-based programmes, respectively. We did not explore the effectiveness of each individual level of 'dose', as the trial was not powered to do so (because of design revisions and compromises following our pilot study; see Blower *et al.*<sup>37</sup>).

Our target sample size was 606 parents with a child aged  $\leq 8$  weeks (allowing for 12% overall attrition). We recruited 341 parents, and 285 and 56 parents were randomised to intervention and control, respectively, with an allocation ratio of 5 : 1 to ensure a total of 50 and 48 parents attending IY-I and IY-T, respectively, across four research sites in England.

We considered any implications of the (unusual) 5 : 1 ratio. We recognise that this ratio is not as 'power-efficient' as balanced randomisation, but it was necessary to ensure that enough intervention participants would be eligible for the group interventions in each site.

The ASQ:SE-2 and PHQ-9 had skewed distributions and, given the extreme imbalance in the size of each arm, the type I error may be increased because there is a heightened risk that, by chance, we observed more extreme values in the larger treatment arm. However, sensitivity analyses on both the ASQ:SE-2 and PHQ-9 showed that our conclusions were not sensitive to this concern. In fact, all 95% CIs for the primary outcome (including sensitivity analyses) excluded the MCID. The high retention to the trial (i.e. 94%, which mitigated not reaching the target sample size) meant that we had enough power to be confident in our conclusions and assured that the allocation ratio did not negatively impact on the results.

Stratification variables included child social and emotional development score (ASQ:SE-2), depression scores (PHQ-9), sex of child and carer. There were no observable differences between arms at baseline, suggesting that our stratification and randomisation process was successful in ensuring comparable arms.

The average baseline age of primary caregivers and their enrolled children were 31 years and 6.05 weeks, respectively. We are not able to determine whether or not the sample were representative of the population of families with 8-week-old infants in each area, or nationally. This is a complex issue, which needs further methodological consideration (we briefly touch on these issues in *Appendix 10*). However, the E-SEE Steps sample differs from the Office for National Statistics' estimates on cohabitation, as indicated in *Chapter 3*.

We will now consider each aim in turn.

### Was the E-SEE Steps programme effective?

The E-SEE Steps programme did not enhance child social and emotional well-being at 20 months of age when compared with the control arm that received SAU. Although the finding was borderline statistically significant, the analysis of the primary outcome (i.e. social and emotional well-being of the

child assessed on the ASQ:SE-2) favoured the control group. Sensitivity analyses of the primary outcome diminished the strength of the signal, although results remained in favour of the control group (see *Table 5*). What is concerning, however, is that both arms showed declining child well-being by our final time point (FU3) on the ASQ:SE-2, with the intervention arm declining more, on average, than the control arm.

The results did not differ depending on how educated parents were, whether or not their child was their firstborn or whether their child was a boy or girl. However, subgroup analysis showed weakly significant evidence of an interaction with treatment for site ( $p = 0.06$ ). The site-level control groups were too small for additional statistical analysis, but do not suggest better scores in the treatment groups than in the control groups for any of the sites.

The analysis of the key secondary outcome (i.e. parent depression levels as assessed on the PHQ-9) demonstrated the opposite (non-significant) trend (i.e. a moderate non-statistically significant signal in favour of the intervention, with sensitivity analyses slightly strengthening the signal).

Follow-up 1 (2 months post baseline) is when the main difference in arms arises for the ASQ:SE-2 and PHQ-9. The children in the intervention arm demonstrated poorer well-being on the ASQ:SE-2 than children in the control arm at this point. FU1 occurs after the universal dose of the IY-B. We considered whether or not the IY-B led to intervention parents questioning their perceptions and how they interpreted their child's social and emotional well-being when completing the ASQ:SE-2. Some parents, when interviewed, said that they found the book too 'difficult' to understand, which may have left them feeling 'inadequate' in their abilities to parent, whereas others may have been more sensitised to their child's well-being and, therefore, more critical of their developmental progress than they may have been if they had not received the book. We do not have the evidence to support these hypotheses, and we cannot be sure that all intervention families read the book.

Follow-up 1 (2 months post baseline) is also when the largest difference between arms occurred for the PHQ-9, with parents in the intervention arm demonstrating improved depression levels compared with parents in the control arm. This runs counter to the pattern expected, as we would expect the ASQ:SE-2 and PHQ-9 to demonstrate similar change patterns based on existing literature.<sup>83,84</sup> We cannot explain this inverse patterning; however, one possibility lies in the psychometric properties of the ASQ:SE-2, which we intend to explore further (the version used was only just developed at the time the trial started and has not been validated in the UK). A previous systematic review<sup>85</sup> of the worldwide evidence on ASQ:SE-2 highlighted that more research is needed to determine its measurement properties in an English sample. There is a lack of research on the reliability and validity of the specific version and associated age intervals administered in the E-SEE trial. The cut-off scores and norms are derived from data on infants and toddlers in the USA.

There was no evidence of a significant difference between trial arms on any of the other secondary parent-reported measures, that is the PSCO survey, HRQoL (using the EQ-5D-5L) and co-parent levels of depression (using the PHQ-9), all of which were collected at all four time points. The same pattern was true for measures of child behaviour (using the SDQ), child cognitive development (using the PedsQL) and parent-child interaction (using the MPAS and PPAS), which were collected at final time point only. We also did not find evidence of differences across arms with regard to how children were fed (e.g. breastfed, bottle fed or mixed feeding, and duration of feeding).

Similar to parent-reported data, the independent observation of dyadic synchrony for the primary parent-child dyad, using the Infant CARE-Index at each time point, demonstrated no difference between study arms. Infant CARE-Index data are a subset of these data from 119 families who consented to be video-recorded and had complete data across all four time points.



## Considerations of why the E-SEE programme did not demonstrate effectiveness

### *Services as usual: was the parenting programme offer and uptake 'better' in the control arm?*

Given that no evidence of a positive impact of the E-SEE Steps programme was found, we considered whether or not control families had taken up local offers of parenting programmes to possibly enhance child well-being. We thought that perhaps uptake would be higher following allocation to control, and lack of the IY offer, possibly leading to control arm bias. However, take-up of alternative programmes was low and only four control and three intervention parents accessed alternative parenting programmes across the entire trial, as reported on the CSRI. In addition, the pattern of change in outcomes suggests a worsening of development in the control arm over time (albeit at a lower rate than the intervention arm), ruling out explanations that they were gaining added value elsewhere. Owing to such small numbers, we can conclude that it was not uptake in these programmes that was driving the null outcome. The programmes were Triple P, the Solihull Approach and Home Start.

As part of the process evaluation we also collated the support programmes (e.g. behaviour management, healthy weight and nutrition, early learning and development, postnatal support and support for children with additional needs) on offer in each site. Sites differed on type and number of programmes offered. SAU in site 3 were the widest ranging of the four sites, with a total of 61 programmes on offer. Across the sites, evidence-based group-based programmes included Strengthening Families,<sup>86</sup> HENRY,<sup>87</sup> the Solihull Approach,<sup>88</sup> Triple P<sup>89</sup> and Family Transitions<sup>90</sup> (and IY in some non-research areas of sites). We did not ascertain what local take-up rates were on these programmes in each area, but for E-SEE Step parents, as stated above, it was low (see *Report Supplementary Material 4*). It may be useful in future studies to gauge local engagement, recruitment and retention strategies for ongoing programme delivery to inform trial site selection.

### *Low take-up of IY-I and IY-T*

Our null findings may have been influenced by the overall low uptake and retention in the IY-I and IY-T groups (for further information on uptake see the following section *Programme uptake according to screener measures and Challenges inherent to the public services system within which the intervention is delivered*). Our imbalanced allocation ratio of 5 : 1 was to ensure a total of 50 and 48 parents attending IY-I and IY-T, respectively, across research sites. At FU1, we assumed that 32% of parents in the E-SEE Steps arm would be eligible for IY-I and expected that 33% of these parents would actually attend. Therefore, at each stage, approximately 10% of parents in the E-SEE Step arm would attend a group [ $n = 51$  (18%) for IY-I, which was better than the expected 10%, and  $n = 21$  (7%) for IY-T, which was lower than the expected 10%]. The level of uptake was disappointing and may have had an impact on the results.

Some attendees were not retained throughout the programme. For both programmes, retention reduced after five sessions. At five sessions, > 80% of participants were still attending in both programmes, but this reduced to 45% for IY-I at session eight (of nine sessions) and 43% for IY-T at session 10 (of 12 sessions). Although numbers were small, and we cannot draw any conclusions, we noted that compliers (i.e. those who attended at least one session) and non-compliers did not differ on outcomes.

In terms of baseline characteristics, parents educated to a higher level and in the higher income bands were marginally more likely to take up the group-based programmes, which reflects findings from the extant literature, demonstrating impact and benefits for socially advantaged families compared with those disadvantaged by income or education.<sup>91-93</sup> However, a recent individual patient data analysis of IY programmes found no evidence for differential effects by social disadvantage, at least in relation to children's conduct problems.<sup>94</sup>

### *Programme uptake according to screener measures*

We considered whether or not take-up was associated with differences in screener scores or arms.

Eligibility for IY-I and IY-T was determined by ASQ:SE-2 and PHQ-9 scores at FU1 and FU2, respectively. The indicative analysis compared the participants eligible for IY-I and IY-T with the subgroups of control participants that had ASQ:SE-2 and PHQ-9 scores in the eligible range for the groups (referred to as pseudo-controls), using the same model as used for the primary outcome. No differences were found between arms.

However, we found that parents with lower PHQ-9 depression scores were more likely to take up IY-I, whereas there was no difference in intervention acceptance by ASQ:SE-2 score. This suggests that services need to work hard to engage with parents who may be suffering low mood or depression to facilitate their attendance at groups.

### ***Robust measures and assumptions?***

We chose our outcome measures carefully for reliability, validity and utility across the age range, as well as for being acceptable and understandable, short and readily available. Our PAC also advised on the acceptability and user-friendliness of the measures.

During the trial set-up stage, we began three systematic reviews of measures frequently used in (three or more) RCTs of parenting programmes for children aged  $\leq 5$  years of age to include (1) parent-report parenting measures,<sup>39</sup> (2) parent-report child measures<sup>40</sup> and (3) independent observational measures.<sup>95</sup> The intention was to inform measure choice and produce a core set of measures. The reviews found a concerning lack of quality and robustness of the measures included therein, when assessed against COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklists.<sup>96</sup>

We considered if the measures were robust enough, given the opposite pattern in ASQ:SE-2 and PHQ-9, but the fact remains that all measures were consistent in finding no significant difference across arms. However, we plan to explore the psychometric properties of ASQ:SE-2, the Infant CARE-Index and the PSOC survey in future work. The Ages and Stages Questionnaire and ASQ:SE-2 are routinely used in the UK for the 24-month child development assessment, with a test-retest reliability of 89%, internal consistency of 84%, sensitivity of 81% and specificity of 84%, which gave us confidence in using the ASQ:SE-2. However, the ASQ:SE-2 has not been validated in the UK.

We defined the clinically important difference at FU3 to be an improvement of 5 units on the ASQ:SE-2 in the IY group when compared with SAU. Although this seems small in a range of 0–240, most people score 0 for most questions, and the ‘monitor’ zone – indicating cause for concern – starts low, in the score range 25–50 (depending on time point and age-appropriate version). We reflected back to our original sample size calculation for reassurance. Given the lack of evidence for what would constitute a clinically significant change on this measure in the original sample size, we used data from the ASQ:SE-2 technical report.<sup>95</sup> We used a difference of 11 points when comparing the at-risk subgroups, which then translated to an effect of 5.5 points at the population level if only half of participants receive the intervention. We, therefore, conclude that the MCID was sound.

Having considered reasons as to why the E-SEE Step programme was not effective, we will now consider the implementation of the E-SEE Steps programme and the process evaluation, as our second aim.

## **Can Incredible Years be delivered as a proportionate universal model, and what are the organisational, or systems-level, barriers to and facilitators of delivering in this way, with fidelity?**

We found a number of barriers and challenges to the delivery of the E-SEE Steps programme through our discussions with key service stakeholders and parents, which directly relate to poor take-up, and retention, of parents to the group programmes, and also inform potential future delivery of the E-SEE Steps programme. Two overarching areas will be discussed, comprising challenges for the actual intervention and challenges for the system that delivers it.



## Challenges for the specific intervention: the E-SEE Steps programme

### Materials

Some challenges for the intervention are amenable to change and may increase wider uptake in the UK. For instance, IY programme materials and video resources could be adapted to be UK specific and temporally relevant (e.g. featuring technology widely used by parents in daily interactions with their children). The IY-B would also benefit by being more easily accessible, for example by being available online or as an electronic book (it is already available as an audiobook on compact disc), and suitable for a wider audience, not just those who struggle with literacy or language. We had low take-up of the offer of the audiobook compact disc in our study.

As part of the pilot phase, a revised version of the IY-B book was produced by the developer to incorporate feedback from the sites to comply more with standards set by UNICEF's BFI. This was unplanned work and took a lot of time, resource and energy to complete within a short time frame so that it could be used within the main trial. Similar amendments have been suggested for the video resources to ensure compatibility with the UK's (and other countries') increasing desire to be BFI accredited. However, this would require additional investment and agreement from the developer to do so, which would mean making amendments, such as removing all pictures and videos of bottles (in case it encourages parents to bottle feed instead of breast feed).

Vignettes showing bottles are currently there to be inclusive of those who cannot, or wish not to, breastfeed, and can also serve as a discussion point around breastfeeding benefits. It is a leader skill to be inclusive within their group and raise awareness of context-specific guidance.

During parent interviews, some parents found the IY-B too 'academic' or difficult to understand as a universal dose when given in isolation of the IY-I programme (it usually is given as part of the group programme).

### Training and delivery

A structured booster training day or delivery preparation support for practitioners after the core 3-day training to embed practice of the IY collaborative process would be beneficial, as would a 'dry run' of delivery prior to delivering to research participants; however, only one site managed to perform a brief dry run. Going straight into delivering IY for the first time as part of a trial may have had a negative impact on parental engagement and retention. The home-coaching training element of IY could have been introduced and implemented to support service staff in engaging with parents prior to, and during, IY delivery, but this would have taken more time, commitment and resource for already financially stretched services.

### Supervision

Owing to the geographical location of sites, availability of IY mentors across the UK and costs involved in face-to-face supervision, a remote supervision method was employed, with independent UK-based IY mentors offering supervision virtually. However, this method, although seemingly simple in concept, proved extraordinarily difficult to implement. Supervision was plagued by difficulties due to site data protection systems that prevented the sharing of video files between agencies that were co-delivering and with the accredited external IY mentors for their analysis.

For future delivery of remote supervision, significant investment is needed in UK systems to upskill the workforce to participate in critical supervision to ensure fidelity and facilitator support. These issues may have an impact on UK staff achieving IY accreditation status, which takes hours of delivery and supervision to reach. Given the current COVID-19 pandemic, it is crucial that these issues are rectified to support remote supervision during remote delivery of IY to support families.

### ***Challenges inherent to the public services system within which the intervention is delivered***

Other system features require consideration before contemplating future intervention delivery. These features include the alignment between nationally and regionally commissioned service offers, and the capacity of providers to deliver intensive parenting programmes via a specific service or agency. If the E-SEE Steps programme is to be delivered via a multiagency partnership, working arrangements and roles for the different agencies and services need to be transparent and agreed prior to delivery.

In 2015, at the start of the pilot trial, responsibility for health visiting passed from the NHS to the LAs, resulting in lengthy large service transformation programmes that affected staff job roles and capacity in the system to engage with research. These changes also had a direct impact on health visitors' involvement in the delivery of parenting support, affecting their capacity to deliver the E-SEE Steps programme. Moreover, although sites with established multiagency partnerships were significantly quicker to set up and establish the E-SEE Steps model, this did not guarantee successful recruitment or retention of parents in the group programmes.

Take-up of the intervention was low, and greater investment in building rapport prior to groups starting could mitigate this; however, the time and capacity of staff presented some challenges to undertaking this engagement work in trial sites. This lack of rapport may have had an impact on parents who struggled with low mood and anxiety, as these parents often need more support to enable their attendance at sessions. Therefore, parents with less/no depression were more likely to take up the offer of IY-I. Parents educated to a higher level and in the higher income bands were marginally more likely to take up the intervention, possibly because they have the social and economic resources to do so, such as a flexible jobs, owning a car and having available child care where needed. Further consideration is needed on how to encourage participation of those who may struggle to get to a group, be it through low mood or accessibility. Not all sites were able to provide crèches or travel costs, which, again, discriminates against those arguably most in need, and may widen the inequalities gap. Owing to costs and other restrictions, crèches and travel costs were difficult to initiate in some sites and may not become an embedded element of delivery; however, reducing these barriers has consistently been shown to be important in parent choice and uptake of services.<sup>97</sup>

The trial design and processes became barriers at times. Both the external pilot and main trial interviews with service managers revealed that trial design/processes further contributed to participant low take-up (e.g. by delaying timely service provision to accommodate data collection points). There were site differences in parent take-up of IY-I, which may be related to differing engagement strategies in sites and across delivery organisations. Future delivery of the E-SEE Steps programme would be free from such trial design restrictions.

It was encouraging to find that parents who attended IY-I and/or IY-T reported high satisfaction and were likely to continue with the group if they managed to attend one session. This is important, as other studies<sup>90</sup> have demonstrated that word of mouth and peer perceptions are key influencers within a community context to encourage others to take up an offer.

Newly trained E-SEE Steps IY leaders were enthusiastic about IY and saw additional benefits to their usual professional practice. For some, IY provided a shift away from being the 'informers or educators' in a didactic approach to working more collaboratively and empowering parents to figure out problems together as a group.

### ***Co-parents***

The qualitative substudy (see *Appendix 10*) explored co-parents' perceptions of the benefits of parenting programmes to inform organisations and group leaders on how to engage with, and retain, fathers and other co-parents within programmes and services. Timing and relevance were the key themes to emerge from interviewing 21 fathers. Only five fathers recollected receiving invitations to IY-I or IY-T and only one father had attended IY-I, with none attending IY-T. The timing theme demonstrated that fathers felt excluded because groups were not offered in the evenings and weekends, suggesting that services, unwittingly, discriminate against working fathers/parents.

In the relevance theme, fathers held the belief that parent groups were mainly for mothers and that they could go along to 'support' the mother rather than learn anything useful for themselves. Fathers did, however, say that they would like to learn more about child development and be more educated, but they failed to recognise that a parenting programme serves these purposes. These points reflect the process evaluation, in which there seems to be some misunderstanding of 'the offer' of attending a parenting programme and how it may be able to support families. Engagement is key to retention and, therefore, to outcomes (if delivered with fidelity). Service design should pay specific attention to these elements.

Despite the fact that only 11 primary parents (3%) said that they were not in a relationship at baseline, only 68 co-parents participated in the trial (i.e. 20% of primary carers had a co-parent that participated in the trial). This figure is typically 10–20% in parenting trials.<sup>92</sup> We originally planned to look at co-parents' impact on child mental health; however, owing to (1) low uptake by co-parents to the trial, with very small numbers in the control arm (intervention,  $n = 53$ ; control,  $n = 15$ ), and (2) low uptake to the intervention ( $n = 8$ ), we were unable to perform the planned analyses.

Intervention co-parents, on average, presented with slightly higher rates of depression than both control co-parents and intervention primary carers. Fathers/co-parents, as well as mothers/primary carers, may find the transition to parenthood challenging and may need additional mental health and well-being support. In 2015, the National Childbirth Trust surveyed 296 fathers, and 38% reported that they were concerned about their mental health.<sup>98</sup> It is important that we consider the mental health of co-parents, and the impact on their children, and partners, for instance, the presence of a non-depressed father is associated with lower rates of behavioural problems among children of depressed mothers.<sup>99</sup> Services and parenting programmes need to be more inclusive, as mentioned above.

Having discussed clinical outcomes and the process evaluation we turn now to the final main aim of our evaluation.

### **Is Incredible Years, and the proposed delivery model, cost-effective in enhancing child social and emotional well-being at 20 months when compared with services as usual?**

The aim of the economic analysis was to assess the cost-consequences and the cost-effectiveness of the E-SEE Steps model (see *Chapter 5*).

#### **Costs**

The average costs of the E-SEE Steps programme amounted to £458.50 per primary carer. The E-SEE Steps programme's proportionate framework was associated with lower costs than the IY interventions in previous IY trials in which the observed average costs of IY interventions ranged between £1496 and £1792 per family.<sup>21</sup> Training and group session costs varied across sites because of differences in attendance, child care requirements and the job profiles of those delivering IY (reflecting differences in salary levels of health and LA staff).

Although staff training costs were high, over time the average intervention costs per primary carer would be reduced if trained staff were retained to deliver future groups. The initial set-up costs would become recurrent delivery costs, as demonstrated in earlier work (e.g. Edwards *et al.*<sup>100</sup>). Our analysis suggests that the average cost of delivering the E-SEE Steps programme falls by approximately 48% (i.e. £236.97) if a trained practitioner delivers 20 programmes. All IY leaders in the E-SEE Steps programme were trained specifically for the trial, as most sites did not have trained or accredited IY facilitators. This was purposeful, as we wanted contamination of IY delivery in SAU to be low or non-existent at the research sites.

The average public-related cost for those allocated to SAU was £1197 and £2263 for primary caregivers and families, respectively. Relative to the E-SEE Steps programme, SAU was found to have an average total public cost saving of £401 and £526 per primary caregiver and per family, respectively. The primary factor for the E-SEE Steps programme's higher average cost was the cost of the intervention itself, with cost savings in some resource categories counterbalanced by additional costs in others. For example, participants in the SAU arm had 0.48 more contacts with counselling or therapy for adults (non-psychiatrist/non-psychologist) for mental health issues than those in the E-SEE Steps programme. However, community service uptake was generally higher for those in the E-SEE Steps arm, for whom adult and child district nurse visits, midwife sessions, parenting courses, other community services and child health visitors were more common than for those in the SAU arm. Other IY research has demonstrated a similar pattern (i.e. more service use in the intervention arm than in the control arm during trial conditions, e.g. O'Neill *et al.*<sup>101</sup>).

### Outcomes

The EQ-5D-5L scores for primary carers were highest in the E-SEE Steps arm, rising from a 0.021 increment at baseline to 0.028 and 0.032 at FU1 and FU2, respectively. The E-SEE Steps programme was also associated with reductions in PHQ-9 depression and severity scores at FU1 and FU2. EQ-5D-5L and PHQ-9 scores did, however, converge at FU3. Although these differences between arms were modest in magnitude and not statistically significant, the results favouring the intervention arm may have important implications for the children, as living with a depressed parent has been shown to lead to negative outcomes for children. For example, living with an adult with poor mental health is associated with higher infant attendance at EDs<sup>102</sup> and, in addition, that parents' depressive symptoms are related to adverse perinatal outcomes.<sup>103</sup> Our results reflected this. In *Appendix 10* we explore the predictors for ED attendance and hospital admission for E-SEE Steps children and their mothers. ED attendance of the child and mother was predicted by the mothers' mental health status among other factors, suggesting a potential mechanism for improvements in parental mental health to avert adverse outcomes and health-related resource use. It should be noted, however, that the magnitude of the reductions in adult depression scores observed in the E-SEE Steps arm was modest and did not translate into improved child outcomes or reductions in secondary care resource utilisation over the trial period. Specifically, children in the E-SEE Steps arm had £31-higher average secondary care costs, an average decrement of 0.619 in PedsQL scores and an average increment of 0.522 in SDQ difficulty scores compared with children in the SAU arm.

### Cost-effectiveness

The E-SEE Steps proportionate universal model did not improve child well-being, but generated improvements in parent health outcomes, resulting in considerable uncertainty around the cost-effectiveness of the intervention. The mean incremental public sector cost and QALY gain of the E-SEE Steps programme was found to be £621 and 0.031 per family compared with SAU, respectively. The positive increments in adult QALYs (0.034) exceeded the minor decrements reported in child outcomes (-0.005), leading to a modest overall gain in mean QALYs. This could possibly be a reflection that a parenting programme generally impacts initially on parents, as parents are the direct programme recipients.

The E-SEE Step programme had an ICER of £20,062 per QALY, which falls within the £20,000–30,000 per QALY threshold range used by NICE. At the minimum threshold, the INHB of introducing the E-SEE Steps programme was negative in the base-case scenario, with the probability of the programme being cost-effective falling below that equivalent to a coin toss (48.8%). At the maximum threshold (i.e. £30,000 per QALY), the INHB associated with the E-SEE Steps programme was positive (0.01 QALYs), with the probability of the programme being cost-effective at approximately 66.8%. Findings were predicated on modest differences in HRQoL values observed in adults and were sensitive to changes in the key assumptions used in the analysis.

The broader impacts and the longer-term implications of the E-SEE Steps programme on caregivers and their children may be important in determining the overall cost-effectiveness of the programme. Our findings showed no evidence of improvements in child-related outcomes during the E-SEE Steps trial, whereas parental outcomes converged between arms at final follow-up, suggesting that the

programme offered only shorter-term parental benefits. Going forward, the future implications of the E-SEE Steps programme on children, primary caregivers and health and social service providers are highly uncertain. The modest decrements in well-being and heightened difficulties observed in children in the E-SEE Steps arm suggest that any longer-term extrapolation of trial findings for children is likely to increase costs and worsen outcomes associated with the E-SEE Steps programme, rendering it unlikely to be cost-effective. In addition, there is little evidence to suggest that the extrapolation of trial findings for adults would be in favour of the E-SEE Steps programme, given that the modest short-term differentials observed in parental HRQoL were informed by a trial not powered to detect differences in HRQoL and converged at final follow-up. The cost-effectiveness of the E-SEE Steps programme may, therefore, be contingent on perspective. Any trade-offs between recipients must be considered with respect to the ultimate goal(s) of the intervention (i.e. family, parental or child outcomes).

## Limitations

Several limitations to the trial and data have been acknowledged above and are, in summary, listed below:

- Our internal pilot became external because of design/intervention changes. We had to rework our budget and timeline while maintaining feasible research questions. We recalculated the sample size and had to forgo our planned analysis of effectiveness of each level of dose within the main trial in favour of exploring whether or not the whole E-SEE Steps model was effective.
- We failed to meet the trial recruitment target of 606 participants; however, owing to the high participant retention rate to the trial (94%), low levels of missing data, range of 95% CIs and sensitivity analyses, we are confident in our conclusions that E-SEE model is not effective in enhancing child social and emotional well-being, but there was low take-up of the group-based programmes.
- Implementation issues, specifically service engagement processes, may have resulted in too few participants receiving the intervention. This could be improved in future delivery, but would require additional, or potentially different, service resource/training. Owing to the low take-up, per-protocol and complier-average causal effect analysis were rejected, as the pseudo-control groups were too small. The 5 : 1 random allocation ratio also had an impact on this decision.
- We do not know if our sample is representative at local or national levels; however, as we only had 11 (3%) participants who reported not being in a relationship at baseline this suggests not. Our previous work with E-SEE Step pilot data, and the similar ENRICH (Evaluation of wRapararound in Ireland for CHildren and families) study in Ireland,<sup>104</sup> demonstrates the many difficulties in assessing representativeness of community samples (see *Appendix 10* for substudy information).
- IY has been delivered in many countries and is, typically, delivered with interpreters where needed (as in E-SEE Steps trial and our other IY trials<sup>26,105-107</sup>). However, within the E-SEE Steps trial, we do not have a large enough, or diverse enough, sample to explore cultural variation or relevance of the intervention in detail.
- The trial timeline had an impact on timings of some group session deliveries and thereby services had reduced time to engage with families.
- We do not know the extent of families' engagement with the IY-B after receipt. Some families may not have read any of the IY-B, whereas others may have digested every page. Qualitative evidence suggests that families found it difficult to access the content without support/facilitation.
- Service/contextual shifts across the duration of the trial, including commissioning changes for health, significant reductions in budgets and high staff turnover, meant that the recommended 'dry run' of group delivery to iron out any teething issues and to give facilitators an opportunity to practise new skills prior to research group delivery did not go ahead in three out of the four sites. In the one site that did have a dry run this was reduced to four summary sessions, as opposed to the whole programme. Therefore, leaders were not as experienced in engaging families in IY and leading IY-I and IY-T groups as they could have been.



- Supervision during delivery enhances performance, skills and confidence; however, supervision was conducted remotely and many access issues were experienced because of an inability of organisational IT systems to communicate with each other (e.g. firewalls prevented communication), or a lack of appropriate IT equipment.
- Co-parent engagement was limited because primary carers were the initial gatekeepers (i.e. co-parents could only be invited via the primary parent). In addition, services were not able to deliver flexible groups to mainly working co-parents (e.g. during the evening or weekend). For these reasons we were unable to look at the relationship between co-parent and child outcomes, such as depression and well-being, respectively.
- There were over the robustness of some measures, including the primary outcome measure (due to lack of UK validation), and the limited age range on other measures, with reliance on older age versions for psychometric information.
- Only a subset of data for the Infant CARE-Index analyses was used, as not all families consented to be video-recorded and, owing to research staff capacity and resource issues, we could include only data from families with complete data at four time points.
- Parent and child service use is reliant only on parent report because of several issues, over a prolonged period of time, when requesting access to Health Episodes Statistics data. In December 2015, an initial application to use health service data for the cohort was submitted to the Health and Social Care Information Centre (now NHS Digital). During this time, the Health and Social Care Information Centre migrated its data access request online and despite contact at least twice a month there was no response to this application. Eventually, it emerged that the application had been lost in the changeover to an online system and to NHS Digital and a new application had to be submitted in 2017. We were allocated a case officer and a number of changes were requested for the consent forms and website. Further changes were then required because of the introduction of General Data Protection Regulation and further requested in autumn 2019 prior to submission to the Independent Group Advising on the Release of Data. By this time, costs had increased substantially to well beyond the budget allocated and there was limited staff support and time left for the ongoing work required. It was decided, therefore, to use self-reported data at this point and explore further funding at a later date to explore the routinely collected data.
- A limitation of all trials is potential control arm bias. The very fact that control families were randomised and told that they had been allocated to the SAU arm could have led to bias, for example through control families taking up more/less services or answering/responding to study assessments differently. Alternative designs, such as the trials within cohorts,<sup>108</sup> using random selection to the intervention rather than random allocation may be useful in the future.
- The economic analysis of child outcomes relied on data from FU3 only. By assuming that baseline child HRQoL in both arms were aligned to the SAU control arm at FU3, the analysis failed to account for potential baseline differences in HRQoL or the underlying dynamics of the intervention's impact on infants over time (as conducted with adult HRQoL). Estimated differentials in child QALYs were, therefore, susceptible to bias.

## Strengths

This trial has several strengths as alluded to throughout the discussion. These are summarised below:

- The proportionate universal trial design arguably reflects real-world intervention delivery where services may adopt different approaches or levels of intervention relative to the stages of child development and/or for differing levels of need [e.g. the Health Visiting model – Universal, Universal Plus and Universal Partnership Plus (URL: <https://www.gov.uk/government/publications/healthy-child-programme-0-to-19-health-visitor-and-school-nurse-commissioning>; accessed 22 March 2022)]. A RCT of one standalone intervention does not consider doses of one or more intervention for families with differing needs at different times.

- We included an in-depth process and economic evaluation.
- The trial was flexible with regard to organisational/site-specific delivery models (i.e. delivery by health or council staff, or both, was acceptable). The areas were also new to IY and so there was no risk of contamination.
- High retention to trial (94%) mitigated the low recruitment figures and enabled us to retain power to answer the main research question with confidence. We also had very few missing data.
- The trial included both parent report and independent observational measures.
- Although there may be some issues with the robustness of some parent-report measures, we are confident, because of our systematic review outcomes and previous research,<sup>39,40,109</sup> that those administered were the most appropriate, given the poor choices available in some outcome domains. Furthermore, we included an independent observation measure in addition to parent report. Our reviews have highlighted weaknesses in certain measures, which we will seek to explore or address.
- Our PAC helped in measure selection and were very helpful in various aspects of the trial (see *Chapter 2, Patient and public involvement*).
- There was no major imbalance between arms at baseline, in terms of either covariates and baseline outcome scores, and we conducted sensitivity analyses to explore these data where there were uncertainties.
- We included substudies that broadened the utility of the research with regard to process, representativeness of sample, predictors of ED and hospital admissions, and father inclusion.
- We have very rich data, some of which will be accessible to other researchers.
- We have consent to follow up parents in the future and will be able to compare arms because control SAU families did not receive the intervention.

## Future research

We have identified the following areas for future research:

- The E-SEE Step trial was funded in response to a National Institute for Health Research Public Health Research call to explore the effectiveness of parenting programmes to enhance child social and emotional well-being in children aged < 2 years. This evidence gap still exists and needs exploring further. Other potential alternatives to parenting programmes in early prevention are now gathering an evidence base, but definitive RCTs are needed. For example, video interactive guidance<sup>110</sup> demonstrated promise in improving parental sensitivity to their preterm infants in a pilot RCT.<sup>111</sup> Video interactive guidance is used in the UK's Children and Young People's Improving Access to Psychological Therapies<sup>112</sup> for behavioural and emotional problems in the 0- to 5-year age range. Video interactive guidance is delivered in a targeted way on a 1-to-1 basis, thereby it is potentially more costly than a group-based parenting programme. Continuity of care, whereby women have a named midwife throughout pregnancy, birth and beyond, is building an evidence base and *The NHS Long Term Plan*<sup>113</sup> sets specific targets to ensure that 75% of women from black and minority ethnic groups and those living in social deprivation receive continuity of care from a known midwife by 2024. However, although a RCT in Australia<sup>114</sup> demonstrated that, for women at low risk of medical complications, caseload midwifery increased women's satisfaction with antenatal, intrapartum and postpartum care, there remains relatively little information about the effects of midwife-led continuity models of care on mothers' and babies' health and well-being in the longer postpartum period.<sup>115</sup>
- Future studies should aim to develop and test a 'pre-intervention' component of implementation strategies devoted to identification and engagement of parents and co-parents, particularly those who may have low mood, as this influences (in this case, low) take-up to the group-based programme, which is in opposition to the high retention rate of parents to the trial. This component should pay particular attention to engaging co-parents too.

- It is clear from this study and our own and others' reviews of measurement tools that there is a paucity of appropriate and robust measures to explore domains of interest across infancy and toddlerhood. This is particularly true for 'attachment' in 0- to 12-month-olds, as highlighted by NICE.<sup>116</sup> We aim to explore some measures, and possibly their psychometric properties or relationships with other measures further, as follows:
  - Although the primary outcome ASQ:SE-2 is widely used in the UK, it has not yet been validated in the UK context. It did not perform as expected with the PHQ-9 (which has been validated in the UK). The Infant CARE-Index has also not yet been validated in the UK. The SDQ and PSOC survey have been used, but both have been validated on older samples. We will explore psychometric properties of the above measures.
  - We will complete the coding all videos with the Infant CARE-Index, for example all pilot data and remaining main trial data (current analyses included parents who consented and with complete data at all four timepoints only). This would also enable more in-depth exploration of the relationship between ASQ:SE2 and Infant CARE-Index scores, which we would expect to correlate, plus subgroup analyses.
  - During the trial we identified a gap for a UK short parent-child 'bonding' measure to be used as a screener owing to the shortage of appropriate measures. A measure has been co-produced with service managers and health visitors, with parent input, and piloted. Results are promising, with some additional validation work planned.
- We will explore the E-SEE Steps sample with regard to 'representativeness', which will build on our pilot work<sup>117</sup> that looks at our UK sample and an Irish sample, and consider how representativeness can be assessed. Preliminary findings suggest that it is extremely difficult to assess at both a local and national level, which casts doubt on trial results and generalisability of findings.
- We will compare any differences in process, particularly around pre-intervention engagement strategies by services, in the E-SEE Steps programme and the similar Irish ENRICH study.<sup>104</sup>
- We will perform a meta-analysis combining the results from the pilot study (pre-design changes) with the results from the full trial, limiting the outcome to FU1. This will enable us to explore the universal IY-B outcome with increased power.
- Our service use data are based on parent report only because of issues with accessing routine data (see *Limitations*). It is important to check that parent report is accurate and we will re-apply to access Health Episode Statistics data to assess this.
- Longer-term follow-up is needed when index children enter school to assess any differences between arms on the primary outcome (i.e. social and emotional well-being), including child behaviour, possibly through school records. In addition, parental mental health could be assessed via a postal survey. We have consent for future contact, and participants in our SAU arm did not receive the trial intervention and would not be eligible in any future delivery in the site (because of age appropriateness), which gives an opportunity for longer follow-up comparisons. Parenting programmes typically have an impact on parenting behaviour (proximal) before child outcomes (distal). This makes sense, as it is the parent behaviour that needs to change to influence child outcomes. The IY BASIC programme for parents of children aged  $\geq 3$  years has successfully demonstrated parent, and child, outcome change post intervention in 'stand-alone' intervention RCTs.<sup>105-107</sup> As the E-SEE Steps intervention is a long-term intervention, that is, longer than a single/stand-alone parent programme (with one to four possible doses over an 18-month period dependent on the level of family need), and with more time for behaviour change to become embedded, we would have expected some changes, or trends, in parenting and subsequent child well-being by the end of the trial.



## Conclusions

### *Implications for practice or local service delivery*

The E-SEE Steps model is a proportionate universal delivery model, whereby successful implementation relies on the whole delivery system working cohesively; however, there were issues in reaching the target families. The model was shown not to be effective when delivering IY-B, IY-I and IY-T to the same cohort of parents (i.e. the intervention produced no benefit to children and insignificant benefits to parental mental health). We cannot categorically state that each level, or individual element, of the E-SEE Steps model (i.e. IY-B, IY-I or IY-T) is not effective, as the study was not powered to establish this. Therefore, it is fair to say that lack of evidence for standalone IY-I and IY-T programmes remains.

The success, or otherwise, of E-SEE Steps programme delivery is contingent on both the faithful delivery of the intervention component programmes and a supportive, flexible and resilient organisational system to host and deliver them as a whole proportionate approach.

These data highlight real difficulties for individual practitioners and service organisations in ensuring that the resources and capacity are available to identify eligible parents, properly engage and retain parents and co-parents, and to prepare sufficiently to deliver with fidelity, with appropriate supervision and oversight and robust service infrastructure to support parents to attend group-based programmes.

A pre-intervention component, or a set of implementation strategies devoted to the identification and engagement (and retention) of parents and also co-parents, may be useful and could be co-developed, applied and evaluated to assess impact on programme uptake.

There are structural and/or organisational challenges in accessing remote supervision for group leaders, which should be addressed.<sup>118</sup> Given the current pandemic, programme delivery and supervision has been moving to an online, remote model to provide continued support to families; however, some difficulties remain (e.g. firewalls can make it difficult to access remote supervision and share video-recorded group delivery on some organisations' laptops, such as the NHS). Evaluations are needed to explore the effectiveness of online delivery and supervision of parenting programmes, as this mode may be preferable for some families (e.g. those that work, partners/parents/carers that live separately or for rural families or those that cannot travel to a group). Remote delivery will also reduce costs of delivery.

The process evaluation suggests that the system struggles to deliver such intensive programmes because of budget cuts, high staff turnover and capacity issues, despite practitioners', service managers' and local commissioners' good will and commitment. However, leaders who were trained were enthusiastic about the interventions and saw additional benefits to their usual professional practice, and families who attended one session were likely to continue with the group and they demonstrated high satisfaction levels.

The proportionate universal approach delivered to the same cohort of parents is the right approach to deliver and assess the effectiveness of parenting programmes (with sufficient power), such as IY, which have several age-related versions. Age as an additional axis also provides the possibility to offer support at different stages of development when the challenges may be different so that parenting support is not a one-size fits all. It makes financial sense for geographical areas to invest in a programme and train leaders to deliver all programmes for continuity. Some countries (e.g. Norway) and UK counties (e.g. Powys) invest heavily in not just the IY parenting programmes, but also in the evidence-based IY child and teacher programmes.

Within the timeline of the trial, there was considerable uncertainty around the cost-effectiveness of the E-SEE Steps programme. However, in general, a proportionate universal approach to delivery may offer potential financial savings by building and consolidating a local skills base with a consistent approach (e.g. various elements of the same, or different, interventions) to reduce health inequalities, as Marmot *et al.*<sup>17</sup> purports. However, more work is needed to support parents with symptoms of

depression to attend a parenting programme, and to increase mental health provision during pregnancy (see NICE guidance<sup>15</sup>), and in the perinatal period. Evidence shows that some parenting programmes reduce parent depression.<sup>20</sup> Some families may, therefore, require 1 : 1 provision/delivery rather than group-based delivery. IY offers that option in the form of home coaching versions of the intervention. Services with a suite of options for the form of delivery of a parenting programme are likely to be able to be more responsive to the needs of their clients (e.g. enhancing relationship skills in those with reduced sensitivity due to low mood<sup>20</sup>) and do better when it comes to engagement and retention in services.

We found that parent anxiety or depression in mothers was a predictor for their admission to hospital and attendance at EDs for their child. Evidence-based parenting programmes could supply a solution in some areas that lack mental health provision. However, careful attention would need to be paid to engaging with families with low mood, depression or anxiety, and a pre-intervention component may need to be developed as part of service design, as mentioned earlier in this section.

# Acknowledgements

The authors would like to thank site staff for participating in and supporting the study, as well as for their identification of potential participants and delivery of groups. We are grateful to the developer of IY and all accredited IY trainers and mentors who trained and supervised study group leaders. Our thanks to all members, past and present, of our PAC. We would also like to thank the independent members of our TSC, Michael Robling (chairperson), Lee Shepstone, Catherine Coucill and lay person (parent of study-age children), and our DMEC, Stavros Petrou (chairperson), Jacqueline Barnes and Lucy Bradshaw, for all their support, wisdom and guidance. We would like to thank the wider research project team for all their efforts and hard work, including research assistants, trial co-ordinators, data collectors, data managers and previous co-applicants/team members. We would also like to thank the anonymous reviewers who gave their time to review this report and subsequently improve its quality. We thank the NIHR for funding the trial, Public Health England for additional support for excess treatment costs and the study sponsor (i.e. the University of York).

Our most grateful thanks are extended to all of the families who participated in the study and who gave up their time to share their experiences with us during a very busy period of family life.

## Contributions of authors

**Tracey Bywater** (<https://orcid.org/0000-0001-7207-8753>) (chief investigator) was involved in the following: conceptualisation, methodology, validation, investigation, writing (original draft, review and editing), supervision, project administration and funding acquisition.

**Vashti Berry** (<https://orcid.org/0000-0001-6438-3731>) (process evaluation lead and co-applicant) was involved in the following: conceptualisation, methodology, validation, formal analysis, investigation, writing (original draft, review and editing), supervision and funding acquisition.

**Sarah Blower** (<https://orcid.org/0000-0002-9168-9995>) (trial manager) was involved in the following: conceptualisation, methodology, validation, investigation, data curation, resources, writing (original draft, review and editing), supervision and project administration.

**Matthew Bursnall** (<https://orcid.org/0000-0002-6519-3558>) (trial statistician) was involved in the following: methodology, validation, formal analysis, data curation, writing (original draft and review) and editing.

**Edward Cox** (<https://orcid.org/0000-0001-8981-0699>) (health economist) was involved in the following: validation, formal analysis, data curation and writing (original draft, review and editing).

**Amanda Mason-Jones** (<https://orcid.org/0000-0002-4292-3183>) (substudy 1 lead and co-applicant) was involved in the following: conceptualisation, methodology, validation, investigation, writing (original draft, review and editing) and funding acquisition.

**Sinéad McGilloway** (<https://orcid.org/0000-0003-0988-3201>) (substudy 3 lead and co-applicant) was involved in the following: conceptualisation, methodology, validation, writing (original draft, review and editing) and funding acquisition.

**Kirsty McKendrick** (<https://orcid.org/0000-0003-0497-3582>) (trial manager) was involved in the following: methodology, validation, investigation, data curation, writing (original draft, review and editing), supervision and project administration.

## ACKNOWLEDGEMENTS

**Siobhan Mitchell** (<https://orcid.org/0000-0002-4085-3898>) (research associate, process evaluation) was involved in the following: methodology, validation, investigation, formal analysis, data curation, writing (original draft, review and editing) and project administration.

**Kate Pickett** (<https://orcid.org/0000-0002-8066-8507>) (co-applicant) was involved in the following: conceptualisation, methodology, writing (review and editing) and funding acquisition.

**Gerry Richardson** (<https://orcid.org/0000-0002-2360-4566>) (health economic lead and co-applicant) was involved in the following: conceptualisation, methodology, validation, writing (review and editing), supervision and funding acquisition.

**Kiera Solaiman** (<https://orcid.org/0000-0003-2244-5345>) (research assistant) was involved in the following: investigation, data curation, writing (review and editing) and project administration.

**M Dawn Teare** (<https://orcid.org/0000-0003-3994-0051>) (lead statistician and co-applicant) was involved in the following: conceptualisation, methodology, validation, formal analysis, writing (original draft, review and editing) and funding acquisition.

**Simon Walker** (<https://orcid.org/0000-0002-5750-3691>) (health economist and co-applicant) was involved in the following: conceptualisation, methodology, validation, writing (review and editing), supervision and funding acquisition.

**Karen Whittaker** (<https://orcid.org/0000-0002-3493-9396>) (substudy 2 lead and co-applicant) was involved in the following: conceptualisation, methodology, formal analysis, investigation, writing (original draft, review and editing), supervision, project administration and funding acquisition.

All authors participated in the interpretation of the findings, contributed ideas and were involved in critically revising this report for important intellectual content. All authors read and agreed the final report. Author contribution statements are provided in line with the CRediT statement guidelines.

## Publications

Bywater T, Berry V, Blower SL, Cohen J, Gridley N, Kiernan K, *et al.* Enhancing Social-Emotional Health and Wellbeing in the Early Years (E-SEE): a study protocol of a community-based randomised controlled trial with process and economic evaluations of the incredible years infant and toddler parenting programmes, delivered in a proportionate universal model. *BMJ Open* 2018;**8**:e026906.

Blower SL, Gridley N, Dunn A, Bywater T, Hindson Z, Bryant M. Psychometric properties of parent outcome measures used in RCTs of antenatal and early years parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:367–87.

Bywater T, Gridley N, Berry V, Blower S, Tobin K. The Parent Programme Implementation Checklist (PPIC): the development and testing of an objective measure of skills and fidelity for the delivery of parent programmes. *Child Care Pract* 2019;**25**:281–309.

Candlish J, Teare MD, Cohen J, Bywater T. Statistical design and analysis in trials of proportionate interventions: a systematic review. *Trials* 2019;**20**:151.

Gridley N, Blower S, Dunn A, Bywater T, Bryant M. Psychometric properties of child (0-5 Years) outcome measures as used in randomized controlled trials of parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:388–405.

Gridley N, Blower S, Dunn A, Bywater T, Whittaker K, Bryant M. Psychometric properties of parent-child (0-5 years) interaction outcome measures as used in randomized controlled trials of parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:253–71.

Whittaker K. Public Health and the Life Course: Supporting Parents and Families. In Cowley S, Whittaker K. editors. *Community Public Health in Policy and Practice*. London: Elsevier; 2020. pp. 159–80.

Blower SL, Berry VL, Bursnall MC, Cohen J, Gridley N, Loban A, *et al*. Enhancing Social-Emotional Outcomes in Early Years (E-SEE): randomised pilot study of Incredible Years infant and toddler programs. *J Child Fam Stud* 2021;**12**:1–7.

Gridley N, Hickey G, Bywater T, McGilloway S, Blower S, Whittaker K, Berry V. The challenges of assessing sample representativeness within community-based evaluations of parenting programmes delivered in England and Ireland [published online ahead of print August 24 2021]. *Health Soc Care Community* 2021.

Bywater T, Barry V, Blower S, Bursnall, Cox E, *et al*. A randomized controlled trial of a proportionate universal parenting program delivery model (E-SEE Steps) to enhance child social-emotional wellbeing. *PLoS ONE* 2022;**17**:e0265200.

### Protocol

See the NIHR Journals Library [URL: [www.journalslibrary.nihr.ac.uk/programmes/phr/139310/#/](http://www.journalslibrary.nihr.ac.uk/programmes/phr/139310/#/) (accessed 14 January 2022)] for the full protocol (version 10).

A brief version of protocol is available here via URL: <https://bmjopen.bmj.com/content/8/12/e026906> (accessed 14 January 2022).

### Data-sharing statement

All data requests should be submitted to the corresponding author (TB) and/or third author (SB) for consideration from September 2021. Access to available anonymised data may be granted following completion of a proforma for review by the data access team (comprising members of the TMG).



## References

1. Darling-Churchill KE, Lippman L. Early childhood social and emotional development: advancing the field of measurement. *J Appl Dev Psychol* 2016;**45**:1–7. <https://doi.org/10.1016/j.appdev.2016.02.002>
2. Rosenblum KL, Dayton CJ, Muzik M. Infant social and emotional development: emerging competence in a relational context. In Zeanah CH, Bakermans-Kranenburg MJ, Barton M, Berens AE, Bernard K, editors. *Handbook of Infant Mental Health*. 3rd edn. New York, NY: The Guilford Press; 2009. pp. 80–103.
3. Denham SA, Wyatt TM, Bassett HH, Echeverria D, Knox SS. Assessing social–emotional development in children from a longitudinal perspective. *J Epidemiol Community Health* 2009;**63**(Suppl. 1):i37–52. <https://doi.org/10.1136/jech.2007.070797>
4. Shonkoff JP, Phillips DA, editors. *From Neurons to Neighbourhoods: The Science of Early Childhood Development*. Washington, DC: National Academies Press; 2000.
5. Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, et al. Nurturing care: promoting early childhood development. *Lancet* 2017;**389**:91–102. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3)
6. Lorber MF, Egeland B. Parenting and infant difficulty: testing a mutual exacerbation hypothesis to predict early onset conduct problems. *Child Dev* 2011;**82**:2006–20. <https://doi.org/10.1111/j.1467-8624.2011.01652.x>
7. Waylen A, Stallard N, Stewart-Brown S. Parenting and health in mid-childhood: a longitudinal study. *Eur J Public Health* 2008;**18**:300–5. <https://doi.org/10.1093/eurpub/ckm131>
8. Hurt L, Paranjothy S, Lucas PJ, Watson D, Mann M, Griffiths LJ, et al. Interventions that enhance health services for parents and infants to improve child development and social and emotional well-being in high-income countries: a systematic review. *BMJ Open* 2018;**8**:e014899. <https://doi.org/10.1136/bmjopen-2016-014899>
9. Allen G. *Early Intervention: Smart Investment, Massive Savings – The Second Independent Report to Her Majesty’s Government*. London: The Stationery Office; 2011.
10. Cowan CP, Cowan PA. Enhancing parenting effectiveness, fathers’ involvement, couple relationship quality, and children’s development: breaking down silos in family policy-making and service delivery. *J Fam Theory Rev* 2019;**11**:92–111. <https://doi.org/10.1111/jftr.12301>
11. Kirby JN, Sanders MR. A randomized controlled trial evaluating a parenting program designed specifically for grandparents. *Behav Res Ther* 2014;**52**:35–44. <https://doi.org/10.1016/j.brat.2013.11.002>
12. Furlong M, McGilloway S, Bywater T, Hutchings J, Smith SM, Donnelly M. Behavioural and cognitive–behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database Syst Rev* 2012;**2**:CD008225. <https://doi.org/10.1002/14651858.CD008225.pub2>
13. Zwi M, Jones H, Thorgaard C, York A, Dennis JA. Parent training interventions for attention deficit hyperactivity disorder (ADHD) in children aged 5 to 18 years. *Cochrane Database Syst Rev* 2011;**12**:CD003018. <https://doi.org/10.1002/14651858.CD003018.pub3>
14. Barlow J, Smailagic N, Ferriter M, Bennett C, Jones H. Group-based parent-training programmes for improving emotional and behavioural adjustment in children from birth to three years old. *Cochrane Database Syst Rev* 2010;**3**:CD003680. <https://doi.org/10.1002/14651858.CD003680.pub2>



15. National Institute for Health and Care Excellence (NICE). *Social and Emotional Wellbeing: Early Years Public Health Guideline*. Public Health Guideline [PH40]. London: NICE; 2012. URL: [www.nice.org.uk/guidance/ph40](http://www.nice.org.uk/guidance/ph40) (accessed 6 July 2021).
16. Barlow J, Parsons J, Stewart-Brown S. Preventing emotional and behavioural problems: the effectiveness of parenting programmes with children less than 3 years of age. *Child Care Health Dev* 2005;**31**:33–42. <https://doi.org/10.1111/j.1365-2214.2005.00447.x>
17. Marmot M, Allen J, Goldblatt P, Boyce T, McNeish D, Grady M, et al. *Fair Society, Healthy Lives (The Marmot Review)*. 2010. URL: [www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review](http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review) (accessed 6 July 2021).
18. Early Intervention Foundation. *Guidebook*. URL: <https://guidebook.eif.org.uk> (accessed 6 January 2022).
19. Menting ATA, Orobio de Castro B, Matthys W. Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behaviour: a meta-analytic review. *Clin Psychol Rev* 2013;**33**:901–13. <https://doi.org/10.1016/j.cpr.2013.07.006>
20. Leijten P, Gardner F, Melendez-Torres GJ, Weeland J, Hutchings J, Landau S, et al. Co-occurring change in children's conduct problems and maternal depression: latent class individual participant data meta-analysis of the Incredible Years parenting program. *Dev Psychopathol* 2019;**31**:1851–62. <https://doi.org/10.1017/S0954579419001068>
21. Gardner F, Leijten P, Mann J, Landau S, Harris V, Beecham J, et al. Could scale-up of parenting programmes improve child disruptive behaviour and reduce social inequalities? Using individual participant data meta-analysis to establish for whom programmes are effective and cost-effective. *Public Health Res* 2017;**5**(10). <https://doi.org/10.3310/phr05100>
22. Leckey Y, Hickey G, Stokes A, McGilloway S. Parent and facilitator experiences of an intensive parent and infant programme delivered in routine community settings. *Prim Health Care Res Dev* 2019;**20**:e74. <https://doi.org/10.1017/S146342361900029X>
23. Jones CH, Erjavec M, Viktor S, Hutchings J. Outcomes of a comparison study into a group-based infant parenting programme. *J Child Fam Stud* 2016;**25**:3309–21. <https://doi.org/10.1007/s10826-016-0489-3>
24. Pontoppidan M, Klest SK, Sandoy TM. The Incredible Years parents and babies program: a pilot randomised controlled trial. *PLOS ONE* 2016;**11**:e0167592. <https://doi.org/10.1371/journal.pone.0167592>
25. Evans S, Davies S, Williams M, Hutchings J. Short-term benefits from the Incredible Years parents and babies programme in Powys. *Community Pract* 2015;**88**:46–8.
26. Hutchings J, Griffith N, Bywater T, Elen Williams M, Baker-Henningham H. Targeted vs. universal provision of support in high-risk communities: comparison of characteristics in two populations recruited to parenting interventions. *J Child Serv* 2013;**8**:169–82. <https://doi.org/10.1108/jcs-03-2013-0009>
27. Perrin EC, Sheldrick RC, McMenamy JM, Henson BS, Carter AS. Improving parenting skills for families of young children in paediatric settings: a randomised clinical trial. *JAMA Pediatr* 2014;**168**:16–24. <https://doi.org/10.1001/jamapediatrics.2013.2919>
28. Sanders MR. Triple P – Positive Parenting Program: towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clin Child Fam Psychol Rev* 1999;**2**:71–90. <https://doi.org/10.1023/a:1021843613840>



29. Michael M, Jessica A, Tammy B, Peter G, Joana M. *Health Equity in England: The Marmot Review 10 Years On*. URL: [www.health.org.uk/publications/reports/the-marmot-review-10-years-on](http://www.health.org.uk/publications/reports/the-marmot-review-10-years-on) (accessed 6 July 2021).
30. Molloy C, O'Connor M, Guo S, Lin C, Harrop C, Perini N, Goldfeld S. Potential of 'stacking' early childhood interventions to reduce inequities in learning outcomes. *J Epidemiol Community Health* 2019;**73**:1078–86. <https://doi.org/10.1136/jech-2019-212282>
31. Candlish J, Teare MD, Cohen J, Bywater T. Statistical design and analysis in trials of proportionate interventions: a systematic review. *Trials* 2019;**20**:151. <https://doi.org/10.1186/s13063-019-3206-x>
32. Webster-Stratton C. Preventing conduct problems in Head Start children: strengthening parenting competencies. *J Consult Clin Psychol* 1998;**66**:715–30. <https://doi.org/10.1037/0022-006x.66.5.715>
33. Bywater T, Barry V, Blower S, Bursnall, Cox E, *et al*. A randomized controlled trial of a proportionate universal parenting program delivery model (E-SEE Steps) to enhance child social-emotional wellbeing. *PLoS ONE* 2022;**17**:e0265200.
34. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c332. <https://doi.org/10.1136/bmj.c332>
35. Bywater T, Berry V, Blower SL, Cohen J, Gridley N, Kiernan K, *et al*. Enhancing Social-Emotional Health and Wellbeing in the Early Years (E-SEE): a study protocol of a community-based randomised controlled trial with process and economic evaluations of the incredible years infant and toddler parenting programmes, delivered in a proportionate universal model. *BMJ Open* 2018;**8**:e026906. <https://doi.org/10.1136/bmjopen-2018-026906>
36. Blower SL, Berry VL, Bursnall MC, Cohen J, Gridley N, Loban A, *et al*. Enhancing Social-Emotional Outcomes in Early Years (E-SEE): randomised pilot study of Incredible Years infant and toddler programs. *J Child Fam Stud* 2021;**12**:1–7. <https://doi.org/10.1007/s10826-021-01991-7>
37. Bartlett JD, Guzman L, Ramos-Olazagasti MA. *Parenting Knowledge Among First-Time Parents of Young Children*. URL: [www.childtrends.org/publications/parenting-knowledge-among-first-time-parents-of-young-children-a-research-to-practice-brief](http://www.childtrends.org/publications/parenting-knowledge-among-first-time-parents-of-young-children-a-research-to-practice-brief) (accessed 6 July 2021).
38. Breiner H, Ford M, Gadsden VL, National Academies of Sciences, Engineering, and Medicine. Parenting knowledge, attitudes, and practices. In Breiner H, Ford M, Gadsden VL, editors. *Parenting Matters: Supporting Parents of Children Ages 0–8*. Washington, DC: National Academies Press; 2016. pp. 45–81.
39. Blower SL, Gridley N, Dunn A, Bywater T, Hindson Z, Bryant M. Psychometric properties of parent outcome measures used in RCTs of antenatal and early years parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:367–87. <https://doi.org/10.1007/s10567-019-00276-2>
40. Gridley N, Blower S, Dunn A, Bywater T, Bryant M. Psychometric properties of child (0–5 years) outcome measures as used in randomized controlled trials of parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:388–405. <https://doi.org/10.1007/s10567-019-00277-1>
41. Bywater T, Blower S. *Enhancing Social-Emotional Health and Wellbeing in the Early Years: A Community-Based Randomised Controlled Trial (and Economic) Evaluation of the Incredible Years Infant and Toddler (0–2) Parenting Programmes*. URL: [www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/](http://www.york.ac.uk/healthsciences/research/public-health/projects/e-see-trial/) (accessed 7 January 2022).

42. Squires J, Bricker D, Potter L. Revision of a parent-completed development screening tool: Ages and Stages Questionnaires. *J Pediatr Psychol* 1997;**22**:313–28. <https://doi.org/10.1093/jpepsy/22.3.313>
43. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;**16**:606–13. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
44. Condon JT, Corkindale CJ. The assessment of parent-to-infant attachment: development of a self-report questionnaire instrument. *J Reprod Infant Psychol* 1998;**16**:57–76. <https://doi.org/10.1080/02646839808404558>
45. Condon JT, Corkindale CJ, Boyce P. Assessment of postnatal paternal–infant attachment: development of a questionnaire instrument. *J Reprod Infant Psychol* 2008;**26**:195–210. <https://doi.org/10.1080/02646830701691335>
46. Johnston C, Mash EJ. A measure of parenting satisfaction and efficacy. *J Clin Child Psychol* 1989;**18**:167–75. [https://doi.org/10.1207/s15374424jccp1802\\_8](https://doi.org/10.1207/s15374424jccp1802_8)
47. Essink-Bot ML, Krabbe PF, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures: the Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Med Care* 1997;**35**:522–37. <https://doi.org/10.1097/00005650-199705000-00008>
48. Beecham J, Knapp M. Costing psychiatric interventions. In Thornicroft G, Brewin CR, Wing J, editors. *Measuring Mental Health Needs*. 2nd edn. London: Gaskell/Royal College of Psychiatrists; 1992. pp. 163–83.
49. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 1997;**38**:581–6. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
50. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the paediatric quality of life inventory. *Med Care* 1999;**37**:126–39. <https://doi.org/10.1097/00005650-199902000-00003>
51. Tracey L, Bywater TJ, Blower SL, Berry V, Cohen J, Kiernan K, et al. *Public Involvement in Research: Social and Emotional Well-being in Early Years*. York: Institute for Effective Education; 2014.
52. Office for National Statistics (ONS). *Population Estimates by Marital Status and Living Arrangements, England and Wales 2019*. London: ONS; 2020.
53. Bywater T, Gridley N, Berry V, Blower S, Tobin K. The Parent Programme Implementation Checklist (PPIC): the development and testing of an objective measure of skills and fidelity for the delivery of parent programmes. *Child Care Pract* 2019;**25**:281–309. <https://doi.org/10.1080/13575279.2017.1414031>
54. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. <https://doi.org/10.1191/1478088706qp063oa>
55. Patton MQ. *Qualitative Evaluation and Research Methods*. 2nd edn. Thousand Oaks, CA: SAGE Publications Ltd; 1990.
56. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. Process and Methods [PMG9]. London: NICE; 2013. URL: [www.nice.org.uk/process/pmg9/chapter/foreword](http://www.nice.org.uk/process/pmg9/chapter/foreword) (accessed 6 July 2021).
57. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. 4th edn. Oxford: Oxford University Press; 2015.
58. NHS Improvement. *Reference Costs 2017–18*. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed 6 July 2021).

59. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2019*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/) (accessed 6 July 2021).
60. Coleman L, Cottell J. *Childcare Survey 2019*. URL: [www.familyandchildcaretrust.org/childcare-survey-2019](http://www.familyandchildcaretrust.org/childcare-survey-2019) (accessed 6 July 2021).
61. Money Advice Service. *Average Childcare Costs*. URL: [www.moneyadviceservice.org.uk/en/articles/childcare-costs](http://www.moneyadviceservice.org.uk/en/articles/childcare-costs) (accessed 6 July 2021).
62. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. *PharmacoEconomics* 1996;**10**:460–6. <https://doi.org/10.2165/00019053-199610050-00003>
63. Curtis LA. *Unit Costs of Health and Social Care 2010*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2010/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2010/) (accessed 13 January 2022).
64. Gidlow CJ, Ellis NJ, Riley V, Chadborn T, Bunten A, Iqbal Z, *et al*. Randomised controlled trial comparing uptake of NHS Health Check in response to standard letters, risk-personalised letters and telephone invitations. *BMC Public Health* 2019;**19**:224. <https://doi.org/10.1186/s12889-019-6540-8>
65. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2018*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/) (accessed 13 January 2022).
66. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2017*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/) (accessed 13 January 2022).
67. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2015*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2015/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2015/) (accessed 13 January 2022).
68. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2016*. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2016/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2016/) (accessed 13 January 2022).
69. Turner J, O’Cathain A, Knowles E, Nicholl J, Tosh J, Sampson F, *et al*. *Evaluation of NHS 111 Pilot Sites. Final Report*. Sheffield: University of Sheffield; 2012.
70. National Audit Office. *NHS Ambulance Services*. URL: [www.nao.org.uk/report/nhs-ambulance-services/](http://www.nao.org.uk/report/nhs-ambulance-services/) (accessed 21 January 2022).
71. van Asselt AD, van Mastrigt GA, Dirksen CD, Arntz A, Severens JL, Kessels AG. How to deal with cost differences at baseline. *PharmacoEconomics* 2009;**27**:519–28. <https://doi.org/10.2165/00019053-200927060-00007>
72. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ* 2018;**27**:7–22. <https://doi.org/10.1002/hec.3564>
73. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, *et al*. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;**15**:708–15. <https://doi.org/10.1016/j.jval.2012.02.008>
74. Furber G, Segal L, Leach M, Cocks J. Mapping scores from the Strengths and Difficulties Questionnaire (SDQ) to preference-based utility values. *Qual Life Res* 2014;**23**:403–11. <https://doi.org/10.1007/s11136-013-0494-6>
75. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL™ generic core scales. *PharmacoEconomics* 2014;**32**:693–706. <https://doi.org/10.1007/s40273-014-0153-y>
76. Lambe T, Frew E, Ives NJ, Woolley RL, Cummins C, Brettell EA, *et al*. Mapping the Paediatric Quality of Life Inventory (PedsQL™) generic core scales onto the Child Health Utility Index-9 Dimension (CHU-9D) score for economic evaluation in children. *PharmacoEconomics* 2018;**36**:451–65. <https://doi.org/10.1007/s40273-017-0600-7>

77. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**:779–87. <https://doi.org/10.1002/hec.635>
78. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York, NY: John Wiley & Sons; 2002. <https://doi.org/10.1002/9781119013563>
79. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–99. <https://doi.org/10.1002/sim.4067>
80. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2015.
81. Briggs AH, Claxton K, Sculpher MJ. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.
82. Hernandez-Villafuerte K, Zamora B, Towse A. *Issues Surrounding the Estimation of the Opportunity Cost of Adopting a New Health Care Technology: Areas for Further Research*. 2018. URL: <https://ideas.repec.org/p/ohe/respap/002044.html> (accessed online 6 July 2021).
83. Bornstein MH, Suwalsky JT, Breakstone DA. Emotional relationships between mothers and infants: knowns, unknowns, and unknown unknowns. *Dev Psychopathol* 2012;**24**:113–23. <https://doi.org/10.1017/S0954579411000708>
84. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal depression and child psychopathology: a meta-analytic review. *Clin Child Fam Psychol Rev* 2011;**14**:1–27. <https://doi.org/10.1007/s10567-010-0080-1>
85. Velikonja T, Edbrooke-Childs J, Calderon A, Slead M, Brown A, Deighton J. The psychometric properties of the Ages & Stages Questionnaires for ages 2–2.5: a systematic review. *Child Care Health Dev* 2017;**43**:1–17. <https://doi.org/10.1111/cch.12397>
86. Kumpfer KL, Molgaard V, Spoth R. The Strengthening Families Program for the prevention of delinquency and drug use. In Peters RD, McMahon RJ, editors. *Preventing Childhood Disorders, Substance Abuse, and Delinquency*. Thousand Oaks, CA: SAGE Publications Ltd; 1996. pp. 241–67. <https://doi.org/10.4135/9781483327679.n11>
87. Rudolf MC, Hunt C, George J, Hajibagheri K, Blair M. HENRY: development, pilot and long-term evaluation of a programme to help practitioners work more effectively with parents of babies and pre-school children to prevent childhood obesity. *Child Care Health Dev* 2010;**36**:850–7. <https://doi.org/10.1111/j.1365-2214.2010.01116.x>
88. Douglas H. The Solihull Approach: helping health visitors to help families with young children. *Young Minds* 1999;**40**:19–20.
89. Sanders MR, Mazzucchelli TG, Studman LJ. Stepping Stones Triple P: the theoretical basis and development of an evidence-based positive parenting program for families with a child who has a disability. *J Intellect Dev Disabil* 2004;**29**:265–83. <https://doi.org/10.1080/13668250412331285127>
90. Stallman HM, Sanders MR. Family Transitions Triple P: the theoretical basis and development of a program for parents going through divorce. *J Divorce Remarriage* 2007;**47**:133–53. [https://doi.org/10.1300/J087v47n03\\_07](https://doi.org/10.1300/J087v47n03_07)
91. 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**:1476. <https://doi.org/10.1136/bmj.38853.451748.2F>
92. Lundahl B, Risser HJ, Lovejoy MC. A meta-analysis of parent training: moderators and follow-up effects. *Clin Psychol Rev* 2006;**26**:86–104. <https://doi.org/10.1016/j.cpr.2005.07.004>

93. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalising behaviour problems – a meta-analytic review. *J Child Psychol Psychiatry* 2006;**47**:99–111. <https://doi.org/10.1111/j.1469-7610.2005.01544.x>
94. Gardner F, Leijten P, Harris V, Mann J, Hutchings J, Beecham J, *et al.* Equity effects of parenting interventions for child conduct problems: a pan-European individual participant data meta-analysis. *Lancet Psychiatry* 2019;**6**:518–27. [https://doi.org/10.1016/S2215-0366\(19\)30162-2](https://doi.org/10.1016/S2215-0366(19)30162-2)
95. Squires J, Bricker D, Twombly E. *Ages & Stages Questionnaires®: Social-Emotional, Second Edition (ASQ®:SE-2). A Parent-Completed Child Monitoring System For Social-Emotional Behaviours.* London: Brookes Publishing; 2015.
96. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, *et al.* The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010;**10**:22. <https://doi.org/10.1186/1471-2288-10-22>
97. Axford N, Lehtonen M, Kaoukji D, Tobin K, Berry V. Engaging parents in parenting programs: lessons from research and practice. *Child Youth Serv Rev* 2012;**34**:2061–71. <https://doi.org/10.1016/j.childyouth.2012.06.011>
98. Public Health England. *Early Years High Impact Area 2: Supporting Maternal and Family Mental Health.* URL: [www.gov.uk/government/publications/commissioning-of-public-health-services-for-children/early-years-high-impact-area-2-supporting-maternal-and-family-mental-health](http://www.gov.uk/government/publications/commissioning-of-public-health-services-for-children/early-years-high-impact-area-2-supporting-maternal-and-family-mental-health) (accessed 6 July 2021).
99. Dietz LJ, Jennings KD, Kelley SA, Marshal M. Maternal depression, paternal psychopathology, and toddlers' behaviour problems. *J Clin Child Adolesc Psychol* 2009;**38**:48–94. <https://doi.org/10.1080/15374410802575362>
100. Edwards RT, Jones C, Berry V, Charles J, Linck P, Bywater T, *et al.* Incredible Years parenting programme: cost-effectiveness and implementation. *J Child Serv* 2016;**11**:54–72. <https://doi.org/10.1108/JCS-02-2015-0005>
101. O'Neill D, McGilloway S, Donnelly M, Bywater T, Kelly P. A cost-effectiveness analysis of the Incredible Years parenting programme in reducing childhood health inequalities. *Eur J Health Econ* 2013;**14**:85–94. <https://doi.org/10.1007/s10198-011-0342-y>
102. Paranjothy S, Evans A, Bandyopadhyay A, Fone D, Schofield B, John A, *et al.* Risk of emergency hospital admission in children associated with mental disorders and alcohol misuse in the household: an electronic birth cohort study. *Lancet Public Health* 2018;**3**:e279–e288. [https://doi.org/10.1016/S2468-2667\(18\)30069-0](https://doi.org/10.1016/S2468-2667(18)30069-0)
103. Alvarez SL, Meltzer-Brody S, Mandel M, Beeber L. Maternal depression and early intervention: a call for an integration of services. *Infants Young Child* 2015;**28**:72–87. <https://doi.org/10.1097/IYC.0000000000000024>
104. Hickey G, McGilloway S, Leckey Y, Leavy S, Stokes A, O'Connor S, *et al.* Exploring the potential utility and impact of a universal, multi-component early parenting intervention through a community-based, controlled trial. *Child Youth Serv Rev* 2020;**118**:1054–58. <https://doi.org/10.1016/j.childyouth.2020.105458>
105. Hutchings J, Gardner F, Bywater T, Daley D, Whitaker C, Jones K, *et al.* Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ* 2007;**334**:678. <https://doi.org/10.1136/bmj.39126.620799.55>



## REFERENCES

106. McGilloway S, Mhaille GN, Bywater T, Furlong M, Leckey Y, Kelly P, *et al*. A parenting intervention for childhood behavioral problems: a randomized controlled trial in disadvantaged community-based settings. *J Consult Clin Psychol* 2012;**80**:116–27. <https://doi.org/10.1037/a0026304>
107. Morpeth L, Blower S, Tobin K, Taylor RS, Bywater T, Edwards RT, *et al*. The effectiveness of the Incredible Years pre-school parenting programme in the United Kingdom: a pragmatic randomised controlled trial. *Child Care Pract* 2017;**23**:141–61. <https://doi.org/10.1080/13575279.2016.1264366>
108. Kim SY, Flory J, Relton C. Ethics and practice of trials within cohorts: an emerging pragmatic trial design. *Clin Trials* 2018;**15**:9–16. <https://doi.org/10.1177/1740774517746620>
109. Gridley N, Blower S, Dunn A, Bywater T, Whittaker K, Bryant M. Psychometric properties of parent-child (0–5 years) interaction outcome measures as used in randomized controlled trials of parent programs: a systematic review. *Clin Child Fam Psychol Rev* 2019;**22**:253–71. <https://doi.org/10.1007/s10567-019-00275-3>
110. O’Hara L, Smith ER, Barlow J, Livingstone N, Herath NI, Wei Y, *et al*. Video feedback for parental sensitivity and attachment security in children under five years. *Cochrane Database Syst Rev* 2019;**11**:CD012348. <https://doi.org/10.1002/14651858.CD012348.pub2>
111. Barlow J, Sembi S, Underdown A. Pilot RCT of the use of video interactive guidance with preterm babies. *J Reprod Infant Psychol* 2016;**34**:511–24. <https://doi.org/10.1080/02646838.2016.1217404>
112. Ludlow C, Hurn R, Lansdell S. A current review of the Children and Young People’s Improving Access to Psychological Therapies (CYP IAPT) program: perspectives on developing an accessible workforce. *Adolesc Health Med Ther* 2020;**11**:21–8. <https://doi.org/10.2147/AHMT.S196492>
113. NHS England. *The NHS Long Term Plan*. London: NHS England; 2019. URL: <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf> (accessed 22 February 2022).
114. Forster DA, McLachlan HL, Davey MA, Biro MA, Farrell T, Gold L, *et al*. Continuity of care by a primary midwife (caseload midwifery) increases women’s satisfaction with antenatal, intrapartum and postpartum care: results from the COSMOS randomised controlled trial. *BMC Pregnancy Childbirth* 2016;**16**:28. <https://doi.org/10.1186/s12884-016-0798-y>
115. Sandall J, Soltani H, Gates S, Shennan A, Devane D. Midwife-led continuity models versus other models of care for childbearing women. *Cochrane Database Syst Rev* 2016;**4**:CD004667. <https://doi.org/10.1002/14651858.CD004667.pub5>
116. National Institute for Health and Care Excellence. *Social and Emotional Wellbeing for Children and Young People Overview*. URL: <https://pathways.nice.org.uk/pathways/social-and-emotional-wellbeing-for-children-and-young-people> (accessed 6 July 2021).
117. Gridley N, Hickey G, Bywater TJ, McGilloway S, Blower SL, Whittaker K, Berry V. The challenges of assessing sample representativeness within community-based evaluations of parenting programmes delivered in England and Ireland [published online ahead of print August 24 2021]. *Health Soc Care Community* 2021. <https://doi.org/10.1111/hsc.13535>
118. Martin P, Kumar S, Lizarondo L. Effective use of technology in clinical supervision. *Internet Interv* 2017;**8**:35–9. <https://doi.org/10.1016/j.invent.2017.03.001>
119. Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2013;**35**:551–5. <https://doi.org/10.1016/j.genhosppsy.2013.04.006>

120. Dreyer K, Williamson RAP, Hargreaves DS, Rosen R, Deeny SR. Associations between parental mental health and other family factors and healthcare utilisation among children and young people: a retrospective, cross-sectional study of linked healthcare data. *BMJ Paediatr Open* 2018;**2**:e000266. <https://doi.org/10.1136/bmjpo-2018-000266>
121. Walker SP, Wachs TD, Grantham-McGregor S, Black MM, Nelson CA, Huffman SL, *et al*. Inequality in early childhood: risk and protective factors for early child development. *Lancet* 2011;**378**:1325–38. [https://doi.org/10.1016/S0140-6736\(11\)60555-2](https://doi.org/10.1016/S0140-6736(11)60555-2)
122. Barlow J, Bergman H, Kornør H, Wei Y, Bennett C. Group-based parent training programmes for improving emotional and behavioural adjustment in young children. *Cochrane Database Syst Rev* 2016;**8**:CD003680. <https://doi.org/10.1002/14651858.CD003680.pub3>
123. Panter-Brick C, Burgess A, Eggerman M, McAllister F, Pruett K, Leckman JF. Practitioner review: engaging fathers – recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *J Child Psychol Psychiatry* 2014;**55**:1187–212. <https://doi.org/10.1111/jcpp.12280>
124. Amodia-Bidakowska A, Laverty C, Ramchandani PG. Father–child play: a systematic review of its frequency, characteristics and potential impact on children’s development. *Dev Rev* 2020;**57**:100924. <https://doi.org/10.1016/j.dr.2020.100924>
125. Alexander CP, Zhu J, Paul IM, Kjerulff KH. Fathers make a difference: positive relationships with mother and baby in relation to infant colic. *Child Care Health Dev* 2017;**43**:687–96. <https://doi.org/10.1111/cch.12445>
126. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In Bryman A, Burgess RG, editors. *Analysing Qualitative Data*. London: Routledge; 1994. pp. 173–94. [https://doi.org/10.4324/9780203413081\\_chapter\\_9](https://doi.org/10.4324/9780203413081_chapter_9)
127. Pope C, Ziebland S, Mays N. Qualitative research in health care. *Analysing qualitative data*. *BMJ* 2000;**320**:114–6. <https://doi.org/10.1136/bmj.320.7227.114>





# Appendix 1 Logic model

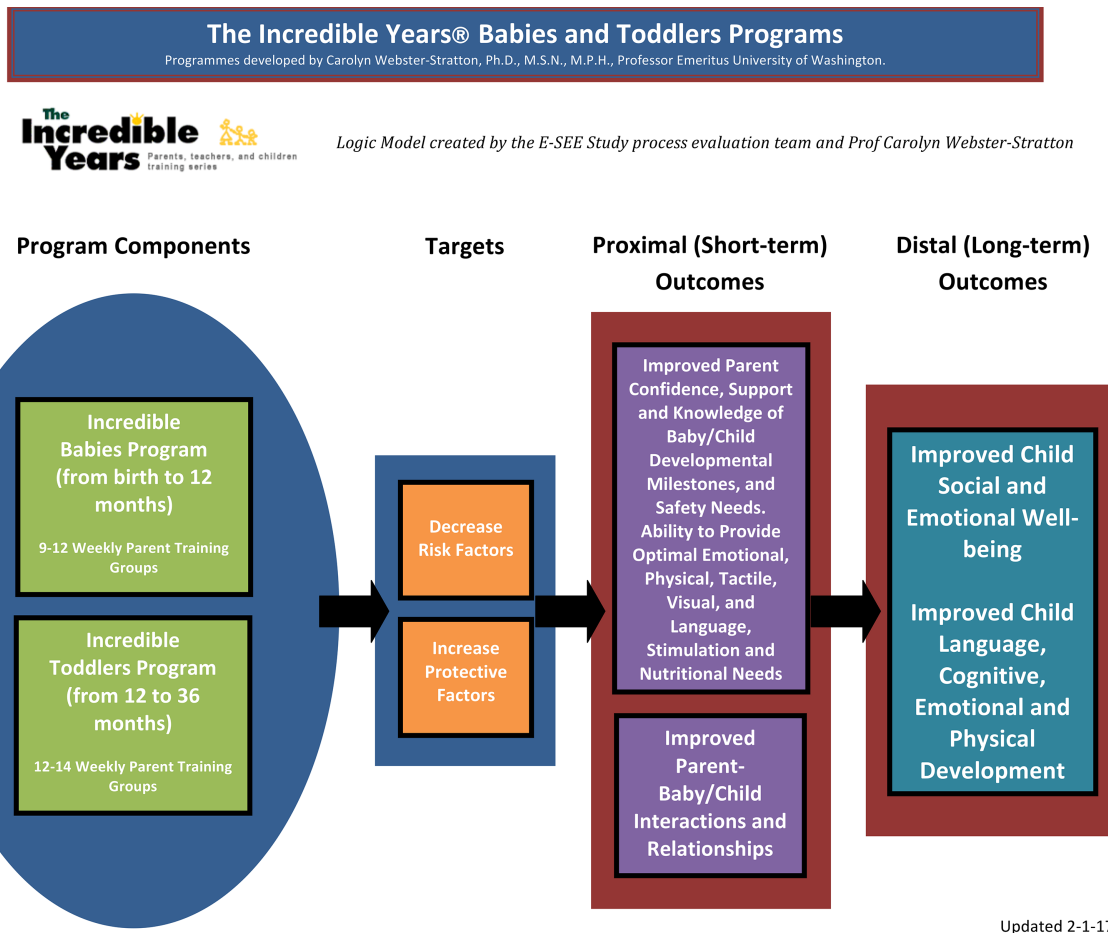


FIGURE 11 Logic model for the IY-I and IY-T. (continued)

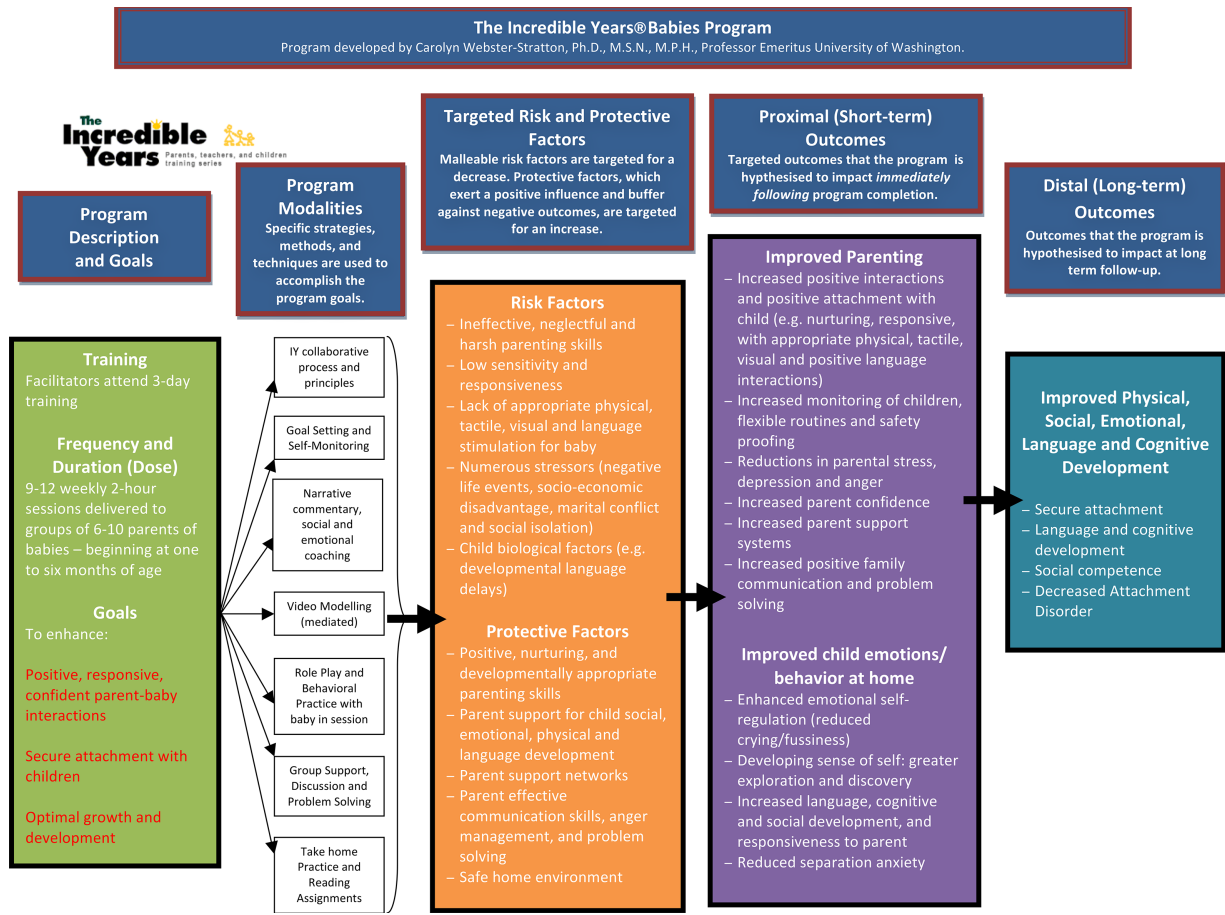


FIGURE 11 Logic model for the IY-I and IY-T.

## Appendix 2 Normative data for outcome measures

TABLE 33 Normative data for ASQ:SE-2

ASQ:SE-2 age interval	Male			Female		
	<i>n</i>	Mean	Median	<i>n</i>	Mean	Median
2 months	148	26.0	20	139	22.8	20
6 months	1066	28.1	21	976	26.6	20
12 months	1278	42.8	35	996	34.6	30
18 months	1239	51.4	40	885	44.1	30

More information on normative data can be found at URL: <https://brookespublishing.com/resource-library/asqse-2-technical-appendix/> (accessed 17 January 2022).

TABLE 34 Normative data from women for their PHQ-9

Normative data	Age (years)			
	14–24	25–34	35–44	45–54
Mean	2.53	2.51	2.81	2.69
SD	3.38	3.01	3.60	3.29

Data from Kocalevent *et al.*<sup>119</sup>

Copy of the Patient Health Questionnaire-9 items



**PATIENT HEALTH QUESTIONNAIRE  
(PHQ-9)**

S   /

Over the last 2 weeks, how often have you been bothered by any of the following problems? (Use "✓" to indicate your answer)	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself - or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

FOR OFFICE USE ONLY    0 +    \_\_\_ +    \_\_\_ +    \_\_\_  
= Total Score:    \_\_\_

10. If you checked off <u>any</u> problems, how <u>difficult</u> have these problems made it for you to do your work, take care of things at home, or get along with other people?			
Not difficult at all <input type="checkbox"/>	Somewhat difficult <input type="checkbox"/>	Very difficult <input type="checkbox"/>	Extremely difficult <input type="checkbox"/>

## Appendix 3 Baseline differences in outcomes

Parts of this appendix have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

TABLE 35 Differences between arms in outcome measures at baseline

Outcome measure	Treatment			Control			Mean difference (95% CI)
	n (%)	Mean	SD	n (%)	Mean	SD	
Child	(n = 285)			(n = 56)			
ASQ:SE-2	285 (100)	22.8	15.1	56 (100)	23.8	16.6	-0.9 (-5.3 to 3.5)
Primary caregiver	(n = 285)			(n = 56)			
PHQ-9	285 (100)	3.1	3.5	56 (100)	2.8	3.2	0.2 (-0.8 to 1.2)
PSOC	285 (100)	80.9	9.6	56 (100)	80.2	9.8	0.7 (-2.0 to 3.4)
EQ-5D	285 (100)	0.9	0.1	56 (100)	0.9	0.1	0.0 (-0.0 to 0.0)
Co-parent	(n = 53)			(n = 15)			
PHQ-9	53 (100)	3.2	3.6	15 (100)	3.2	3.8	-0.0 (-2.1 to 2.1)
PSOC	53 (100)	79.2	10.2	15 (100)	79.9	11.2	-0.7 (-6.7 to 5.3)
EQ-5D	52 (98)	1.0	0.1	15 (100)	0.9	0.1	0.0 (-0.1 to 0.1)



## Appendix 4 Difference in ASQ:SE-2 scores between arms

Parts of this appendix have been adapted from Bywater *et al.*<sup>33</sup> © 2022 Bywater *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

TABLE 36 Difference in ASQ:SE-2 score between arms using square root transformation

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)	p-value
	Treatment			Control					
	n (%)	Mean	SD	n (%)	Mean	SD			
Baseline	285 (100)	4.47	1.68	56 (100)	4.60	1.64	-0.12 (-0.60 to 0.36)		
FU1	270 (95)	4.15	1.82	55 (98)	3.82	1.38	0.33 (-0.18 to 0.84)		
FU2	269 (9)	5.13	1.64	55 (98)	4.87	1.77	0.25 (-0.23 to 0.73)		
FU3	268 (94)	4.80	1.95	53 (95)	4.78	2.32	0.02 (-0.57 to 0.61)		
Overall								0.33 (-0.01 to 0.67)	0.057

TABLE 37 Difference in ASQ:SE-2 score between arms using primary model assuming unstructured correlation

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)	p-value
	Treatment			Control					
	n (%)	Mean	SD	n (%)	Mean	SD			
Baseline	285 (100)	22.84	15.13	56 (100)	23.75	16.60	-0.91 (-5.32 to 3.50)		
FU1	270 (95)	20.52	15.72	55 (98)	16.45	10.87	4.06 (-0.29 to 8.41)		
FU2	269 (94)	28.96	16.23	55 (98)	26.82	14.73	2.14 (-2.50 to 6.78)		
FU3	268 (94)	26.81	19.50	53 (95)	28.11	23.60	-1.30 (-7.26 to 4.66)		
Overall								2.56 (-0.69 to 5.80)	0.122

### Residual plots for primary outcome

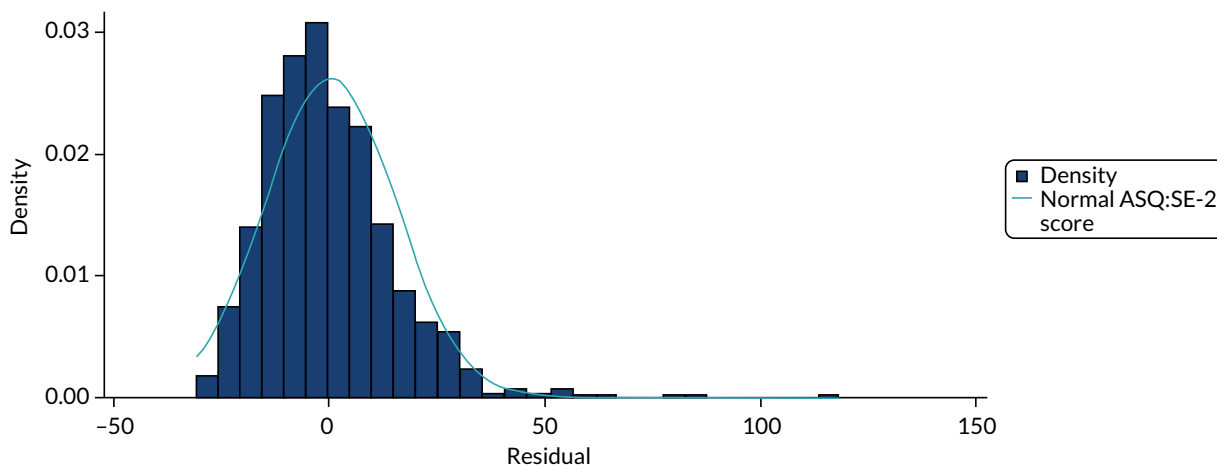


FIGURE 12 Residuals from the primary analysis model.

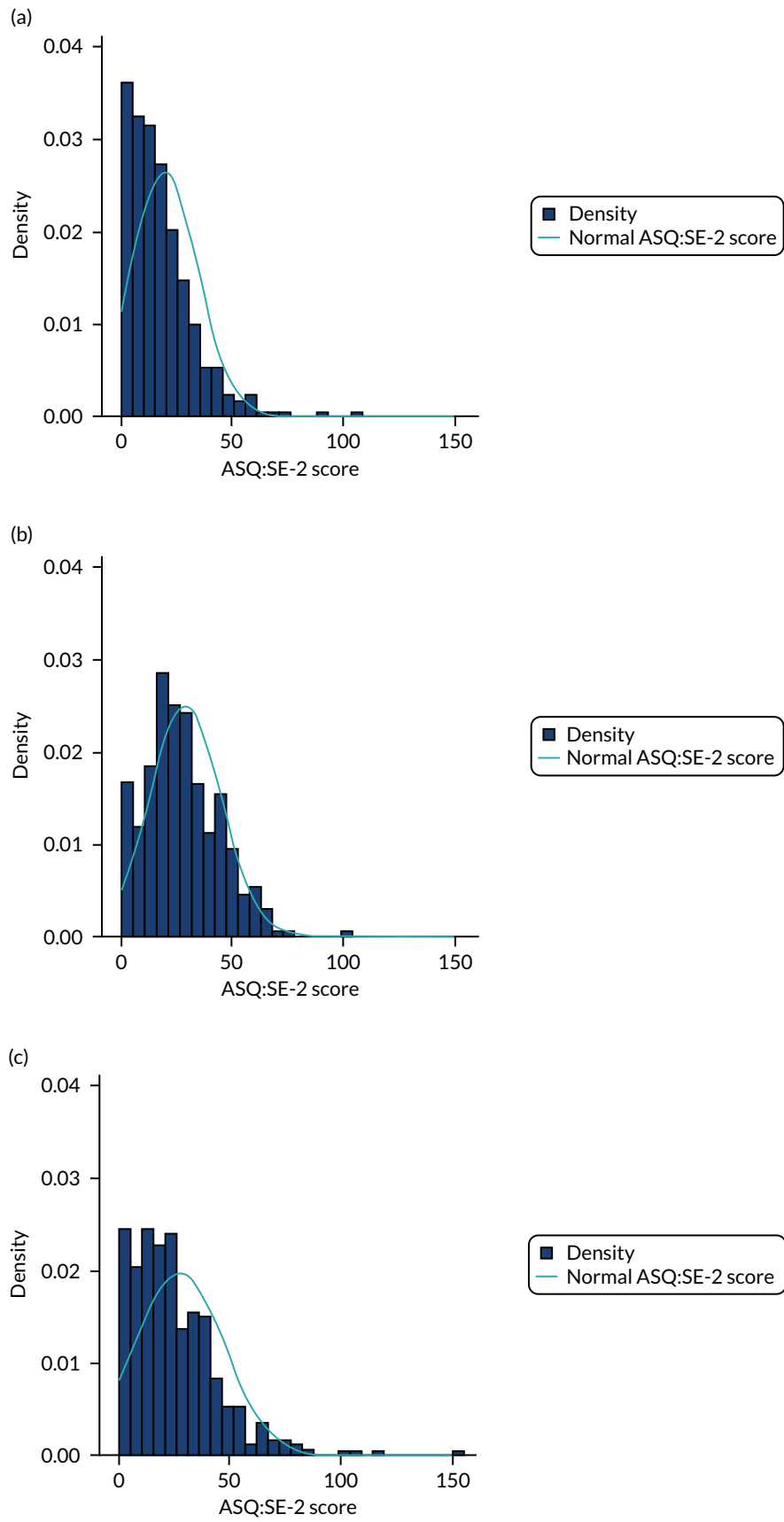


FIGURE 13 ASQ:SE-2 scores at each follow-up point. (a) FU1; (b) FU2; and (c) FU3.



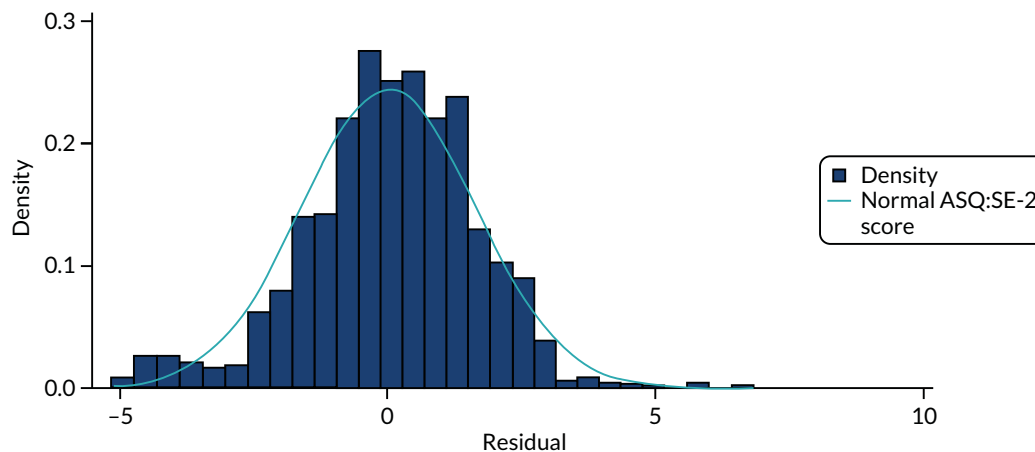


FIGURE 14 Residuals from the primary analysis model of ASQ:SE-2 after square root transformation.

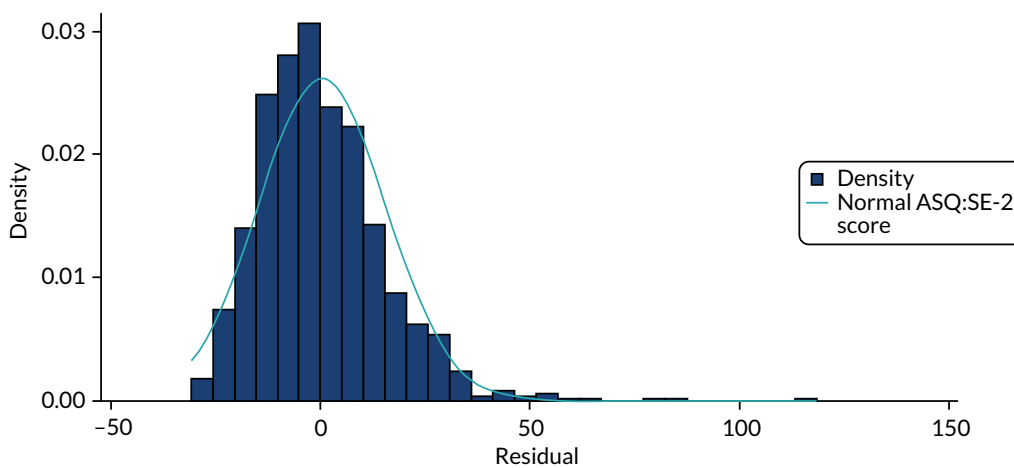


FIGURE 15 Residuals from the primary analysis model of ASQ:SE-2 after calculating difference from baseline.

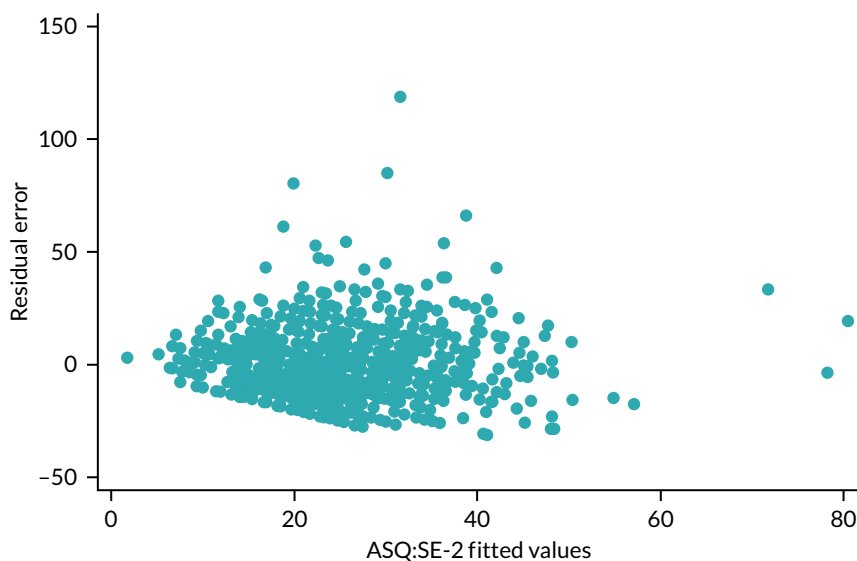


FIGURE 16 Residuals against fitted values for the ASQ:SE-2 primary analysis model.



## Appendix 5 Summary of ASQ:SE-2 missing data

TABLE 38 Missing ASQ:SE-2 data summary

Treatment arm	ASQ:SE-2 provided (N = 970)	ASQ:SE-2 missing (N = 53)
Intervention (N = 855), n (%)	807 (94)	48 (6)
Control (N = 168), n (%)	163 (97)	5 (3)
Baseline ASQ:SE-2		
Mean (SD)	23.3 (15.5)	17.5 (10.3)
Median (IQR)	20 (15–30)	15 (10–25)
Minimum, maximum	0, 105	0, 45



## Appendix 6 Primary outcome by site

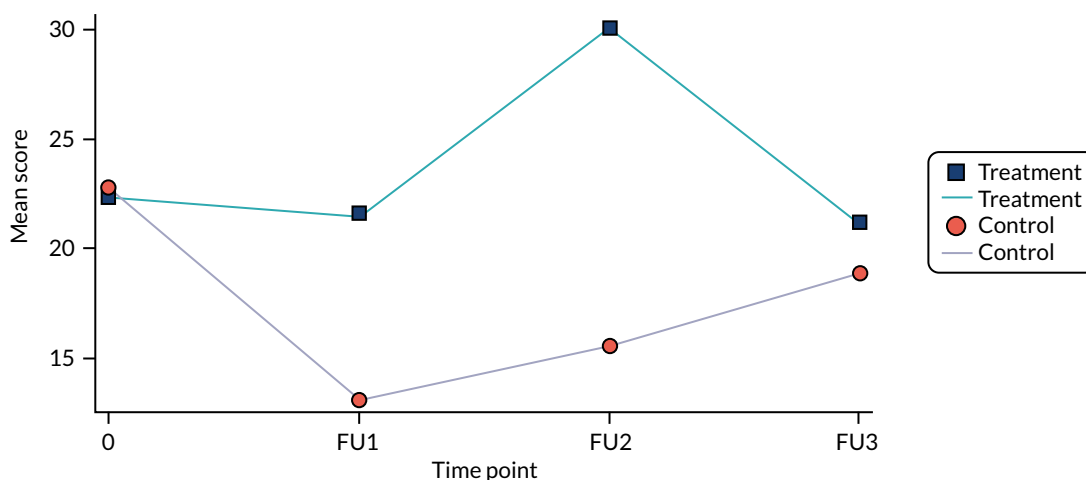


FIGURE 17 Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 3.

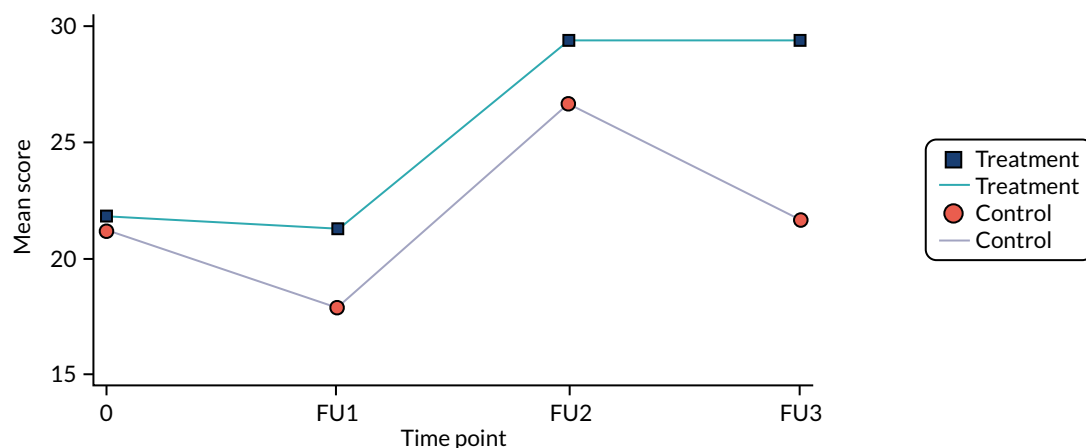


FIGURE 18 Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 2.

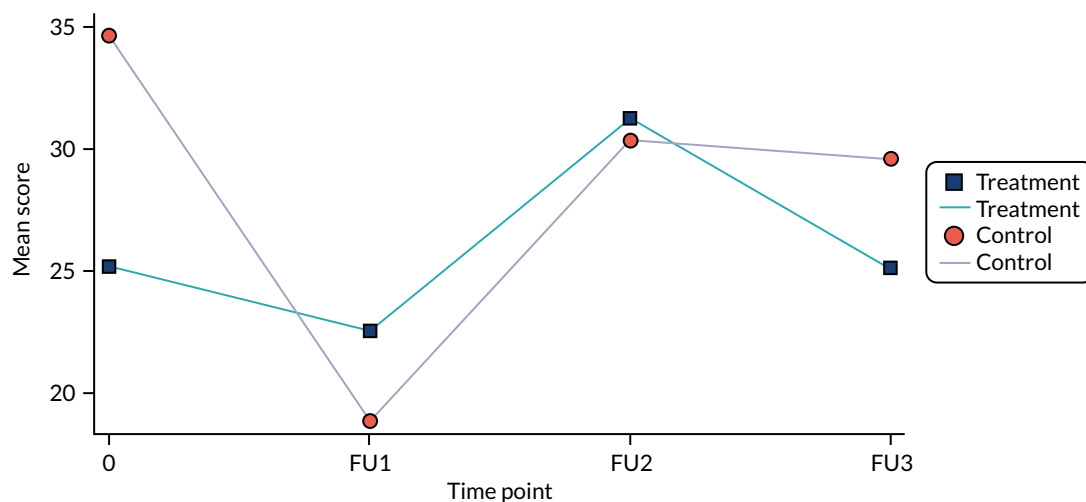


FIGURE 19 Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 4.

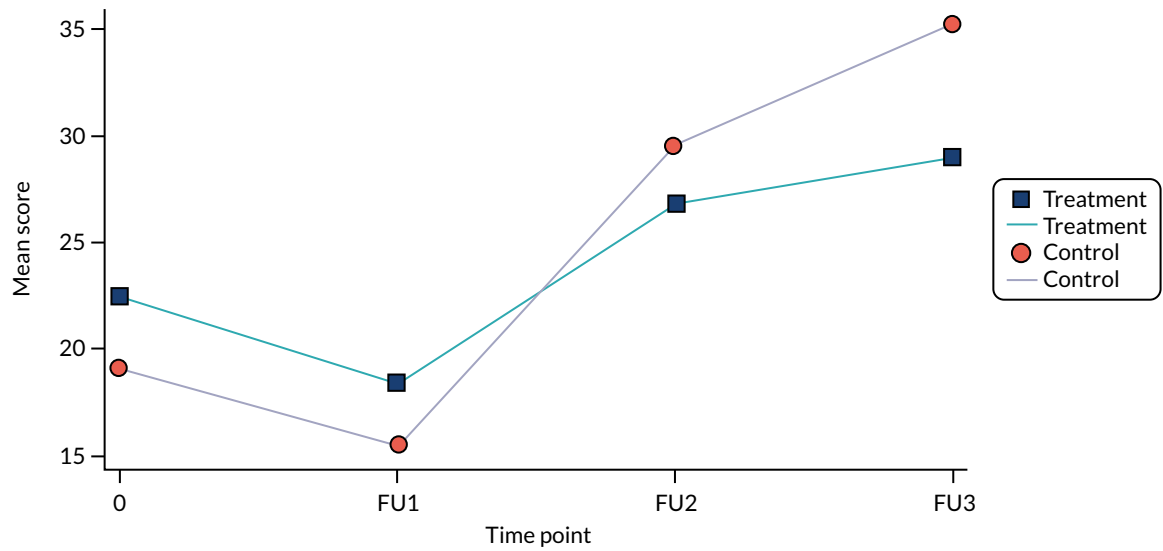


FIGURE 20 Primary outcome (ASQ:SE-2) mean score by follow-up time point and arm for site 1.

## Appendix 7 Indicative analysis of ASQ:SE-2

TABLE 39 Indicative analysis of ASQ:SE-2 for each stage using eligible subgroups

Time point	Treatment arm						Mean difference (95% CI)	Adjusted mean difference (95% CI)	p-value
	Treatment			Pseudo-control					
	n (%)	Mean	SD	n (%)	Mean	SD			
Baseline	101 (100)	32.62	17.60	19 (100)	33.16	22.93	-0.53 (-9.60 to 8.54)		
FU2	97 (100)	36.86	17.17	19 (100)	35.00	12.58	1.86 (-6.27 to 9.99)		
Overall								1.90 (-5.93 to 9.73)	0.63
Baseline	101 (100)	28.51	17.51	22 (100)	27.95	18.62	0.56 (-7.61 to 8.73)		
FU3	99 (100)	35.40	22.23	21 (100)	37.86	24.63	-2.45 (-13.12 to 8.22)		
Overall								-0.85 (-12.09 to 10.39)	0.88





## Appendix 8 Analysis of Patient Health Questionnaire-9 items data

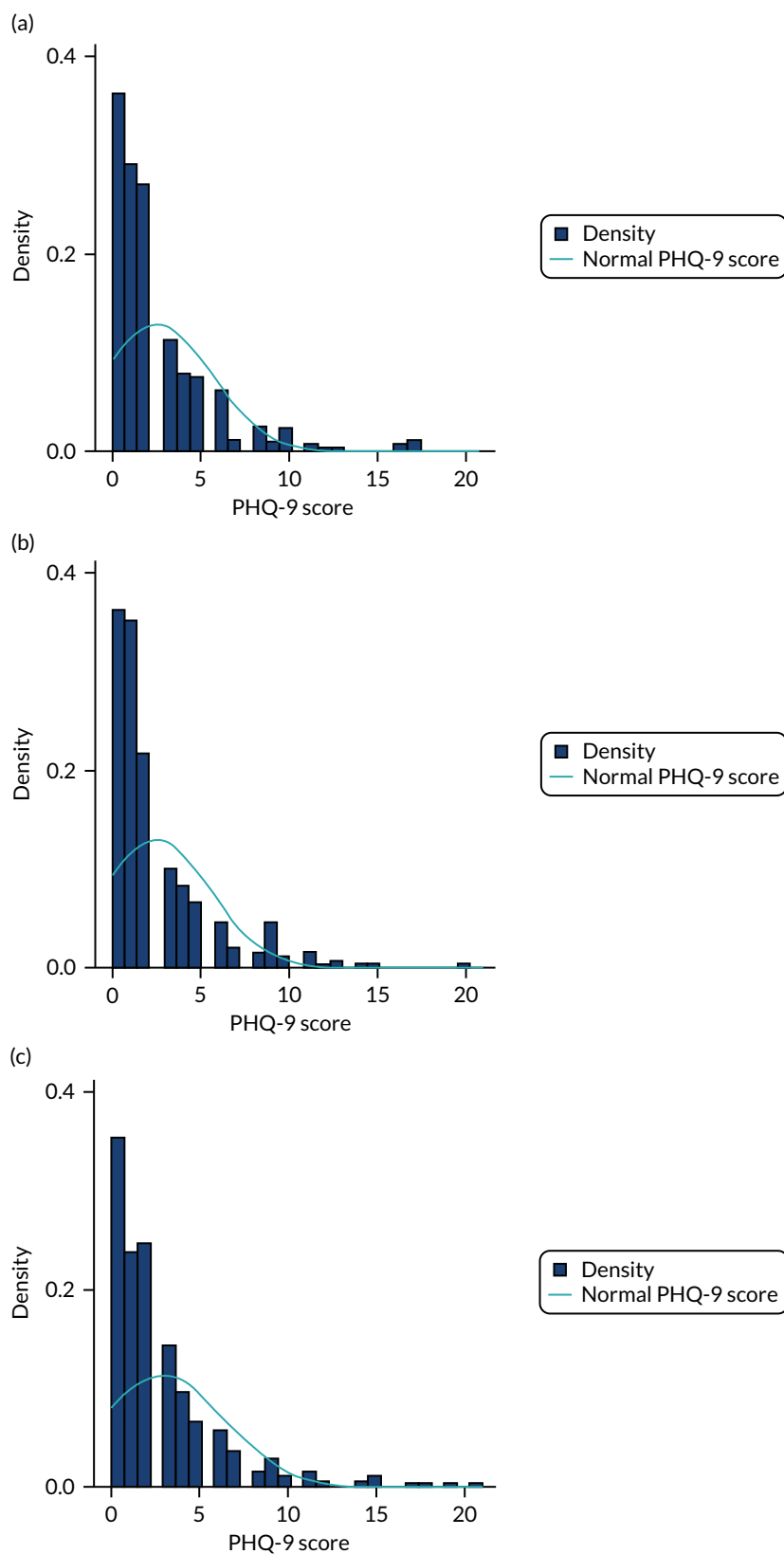


FIGURE 21 Histogram of PHQ-9 data. (a) FU1; (b) FU2; and (c) FU3.

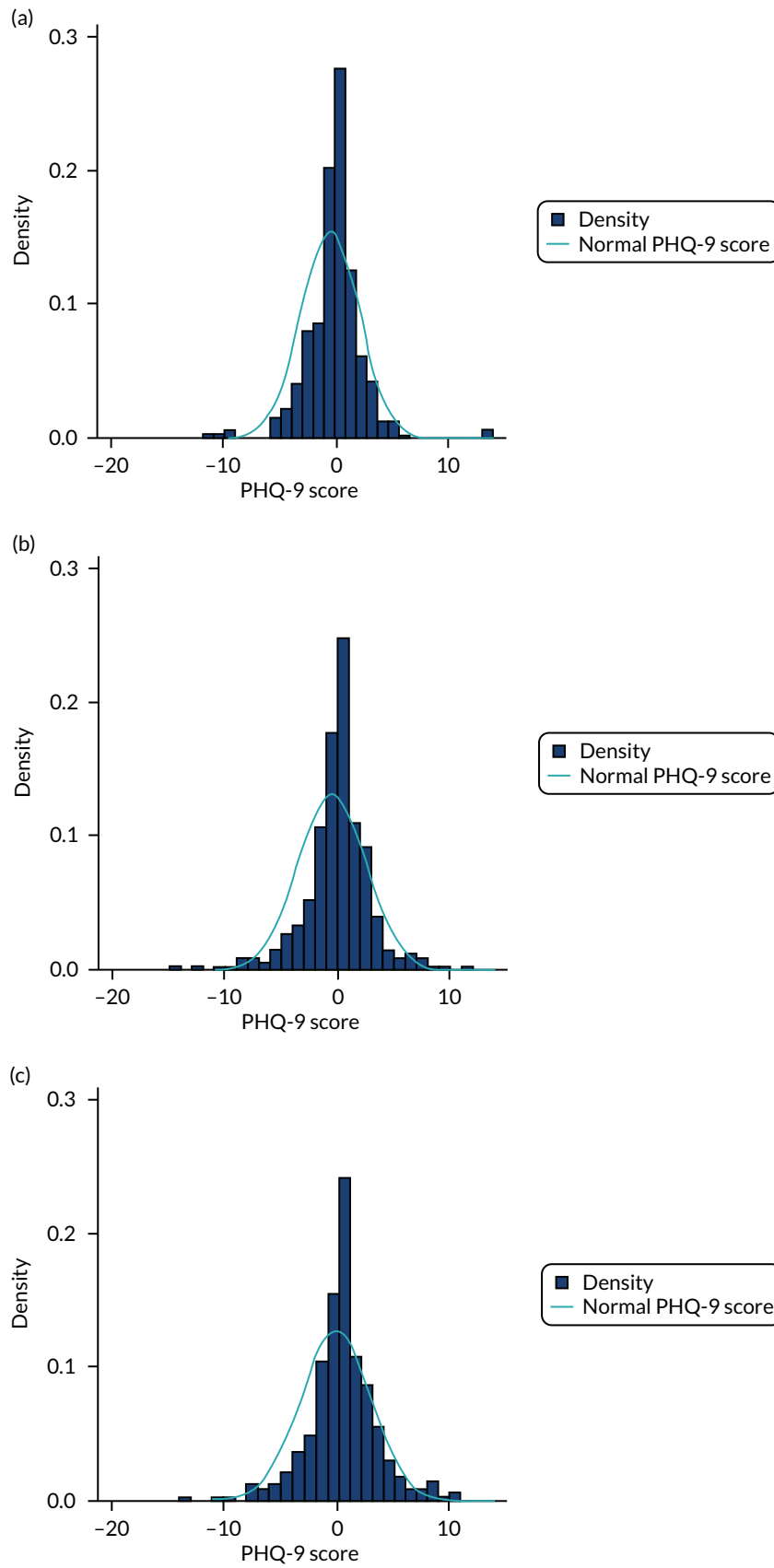


FIGURE 22 Distribution of PHQ-9 difference from baseline. (a) FU1; (b) FU2; and (c) FU3.

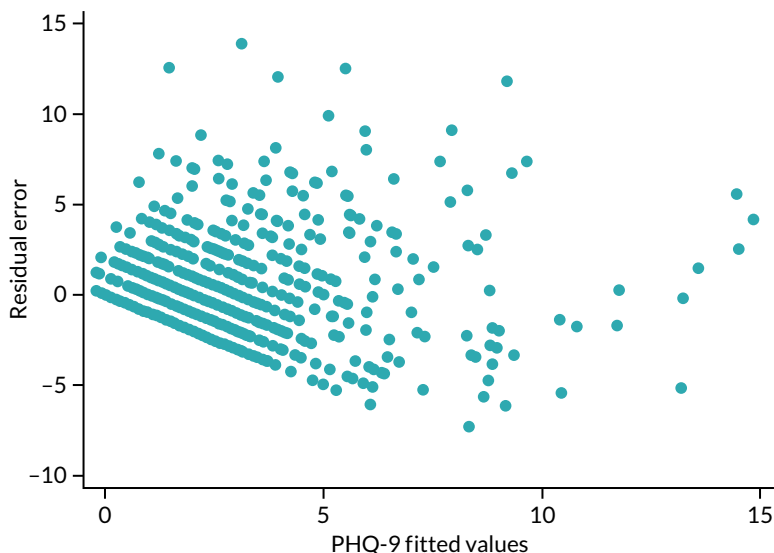


FIGURE 23 Residuals against fitted values for PHQ-9.

### Residual plots for primary parent Patient Health Questionnaire-9 items

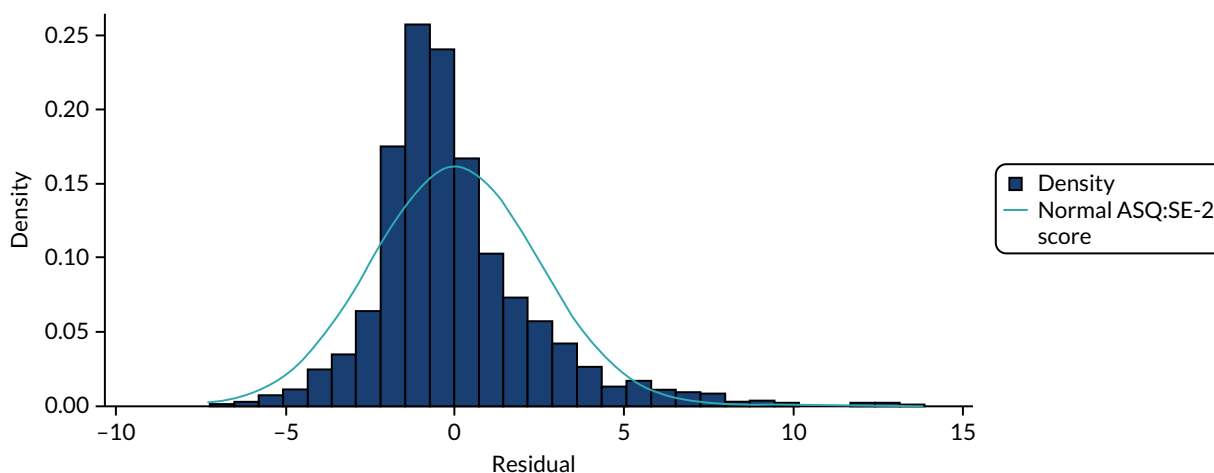


FIGURE 24 Residuals from planned PHQ-9 analysis.

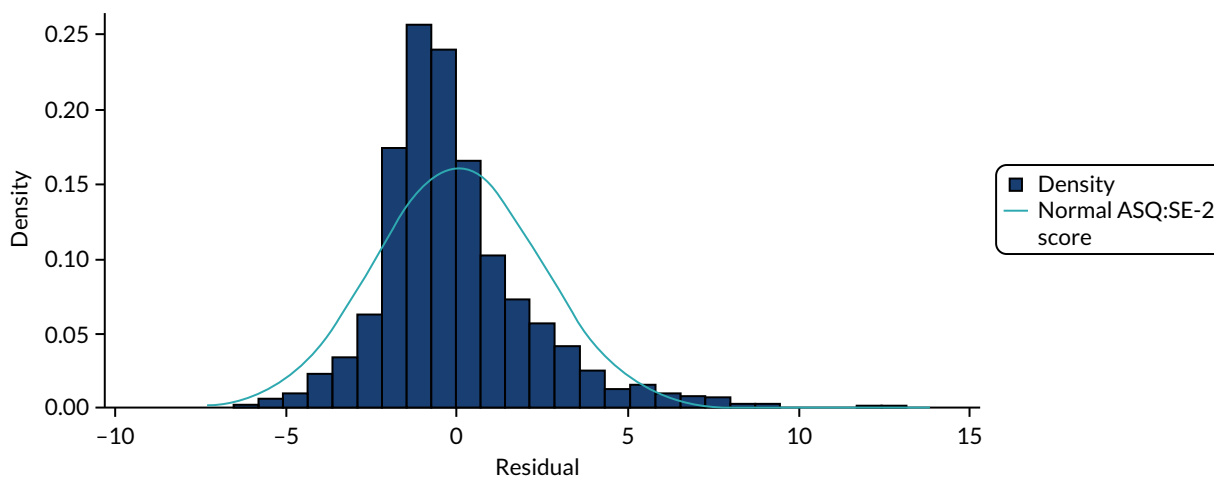


FIGURE 25 Residuals from unplanned analysis of PHQ-9 after calculating difference from baseline.



## Appendix 9 Transitions in Infant CARE-Index scores

TABLE 40 Transitions between Infant CARE-Index risk categories

Infant CARE-Index rating at baseline	Treatment arm	N	Infant CARE-Index rating, n (%)			
			Risk	Inept	Adequate	Sensitive
<b>FU1</b>						
Risk	Treatment	38	14 (37)	8 (21)	13 (34)	3 (8)
	Control	5	1 (20)	1 (20)	1 (20)	2 (40)
Inept	Treatment	36	12 (33)	13 (36)	8 (22)	3 (8)
	Control	6	1 (17)	2 (33)	3 (50)	0 (0)
Adequate	Treatment	24	6 (25)	8 (33)	7 (29)	3 (13)
	Control	3	0 (0)	1 (33)	1 (33)	1 (33)
Sensitive	Treatment	7	2 (29)	1 (14)	4 (57)	0 (0)
	Control	0	0 (0)	0 (0)	0 (0)	0 (0)
<b>FU2</b>						
Risk	Treatment	38	3 (8)	8 (21)	19 (50)	8 (21)
	Control	5	0 (0)	1 (20)	4 (80)	0 (0)
Inept	Treatment	36	6 (17)	8 (22)	17 (47)	5 (14)
	Control	6	2 (33)	0 (0)	3 (50)	1 (17)
Adequate	Treatment	24	4 (17)	2 (8)	12 (50)	6 (25)
	Control	3	0 (0)	1 (33)	1 (33)	1 (33)
Sensitive	Treatment	7	0 (0)	0 (0)	3 (43)	4 (57)
	Control	0	0 (0)	0 (0)	0 (0)	0 (0)
<b>FU3</b>						
Risk	Treatment	38	2 (5)	11 (29)	17 (45)	8 (21)
	Control	5	0 (0)	1 (20)	2 (40)	2 (40)
Inept	Treatment	36	2 (6)	7 (19)	19 (53)	8 (22)
	Control	6	0 (0)	2 (33)	4 (67)	0 (0)
Adequate	Treatment	24	0 (0)	5 (21)	12 (50)	7 (29)
	Control	3	0 (0)	2 (67)	1 (33)	0 (0)
Sensitive	Treatment	7	0 (0)	1 (14)	5 (71)	1 (14)
	Control	0	0 (0)	0 (0)	0 (0)	0 (0)



## Appendix 10 Planned substudies

This appendix provides an overview and findings of three planned E-SEE Step substudies. These substudies comprise (1) a study of predictors of ED and hospital admissions, (2) a study qualitatively exploring father engagement with parenting support and (3) a study bringing together the E-SEE Step trial and a similar trial in Ireland<sup>104</sup> to explore sample representativeness. The last substudy is ongoing and we plan to compare patterns of process and implementation across the trials.

### Substudy 1: predictors of emergency department attendance and hospital admission for children and their mothers

#### Background

Parental mental health has been proposed as a factor that increases utilisation of emergency services and other services by children and young people.<sup>120</sup> Both maternal and paternal depression are associated with an increase in child mortality and suboptimal cognitive and developmental trajectories during both childhood and adolescence.<sup>121</sup> Children of parents with mental health problems also show an increase in recurrent health problems themselves<sup>122</sup> and are more susceptible to infections and injuries, particularly when mothers have experienced mental health problems before birth or up to 1 year after the birth.<sup>123</sup>

#### Aim

The aim was to find out more about predictors of ED utilisation and subsequent admission to hospital of participant dyads (i.e. mothers and their child) enrolled in the E-SEE Step trial.

An initial application to use health service data for the cohort was submitted to the Health and Social Care Information Centre (now NHS Digital) in December 2015. During this time, the Health and Social Care Information Centre migrated its data access request online and, despite contact at least twice a month, there was no response to this application. Eventually, it emerged that the application had been lost in the changeover to an online system and to NHS Digital and a new application had to be submitted in 2017. We were allocated a case officer and a number of changes were requested for the consent forms and website. Further changes were then required because of the introduction of General Data Protection Regulation and further requested in autumn 2019 prior to submission to the Independent Group Advising on the Release of Data. By this time, costs had increased substantially to > £8000, which was well beyond the budget allocated, and there was limited staff support and time left for the ongoing work required. It was decided, therefore, to use self-reported data at this point and explore further funding at a later date to explore the routinely collected data.

#### Methods

##### Data sources

A secondary analysis was undertaken based on data available from the E-SEE Step trial, specifically information from demographic and other baseline data, as well as follow-up data from the four previously described time points, including self-reported adverse events and hospital admissions from the CSRI, information from the trial's relationship questionnaire, EuroQol quality-of-life inventory (EQ-5D-5L) and PSOC survey data.

##### Outcomes

The following self-reported data from the CSRI were used as substudy outcomes:

- child hospital admissions
- mother hospital admission
- child ED use
- mother ED use.

All outcomes were recorded as binary incidence (i.e. if any event as above was recorded) and continuous frequency (i.e. number of separate occurrences) per caregiver–child dyad. Duplicate records were identified by date and event details, and were consolidated into unique events. Study data contained reasons for hospital admission (which were reported thematically using the first reported reason if there were more than one). No reasons for ED use were available.

### Statistical analysis

Baseline characteristics were summarised descriptively for the total study sample and mothers and children for whom one or more study outcomes were recorded (hospital admissions or ED attendances).

To identify predictors of the study outcomes, a series of univariate and multivariate regression analyses were conducted. The incidence of hospitalisations or ED attendances was analysed using mixed-effect logistic regression, and the frequency of hospitalisations and ED attendances was analysed using negative binomial or Poisson regression, depending on the model fit based on overdispersion of the outcome, and assessed separately for each model.

Initially, each potential predictor (as detailed below) was analysed separately. Each model included an interaction term of the predictor with allocation to the E-SEE intervention or control as a fixed effect to account for possible confounding by the intervention. Where such an interaction term resulted in non-convergence of the regression model, allocation was included as a main effect only. Recruitment site (i.e. 1, 2, 3 or 4) was included as a random effect in the mixed-effects logistic regression models and explored as a separate fixed effect in the Poisson and negative binomial models.

Statistically significant predictors (i.e. a  $p$ -value  $< 0.05$ ) from the univariate regressions were then included together with allocation in a combined multivariate model for each outcome. Where significant predictors used the same information (e.g. reporting of partner violence at baseline and at any time during the study), only the overall variable was included in the combined mode. Odds ratios (ORs) or incidence rate ratios (IRRs) together with 95% CIs and  $p$ -values were reported for each predictor. All data were analysed on a complete-case basis, with cases with any missing data for a given analysis excluded. No adjustments were made for multiple testing. All analyses were conducted using Stata.

The following potential predictors were included in the initial analyses. Based on the distribution of categorical variables in these data, some categories were combined (indicated below in *italics*) to allow viable and meaningful analyses.

#### **Demographics: mother**

- Age (in years).
- Sex (female).
- Ethnicity (white, Asian/Asian British, black/black British, mixed/multiple, other).
- Highest educational qualification achieved [Ordinary Level or lower, Advanced/Advanced Subsidiary/Scholarship Level, vocational qualification, higher education (*certificate or diploma, bachelor's degree, master's degree, post-doctorate qualification*), other (*none of the above, overseas qualifications*)].
- Employment status [paid employed (*employed, maternity leave*), unemployed (*unemployed, housewife*), other (*student, retired, other*)].
- Weekly income (£).
- Number of biological children (including current child).
- Age when her first child was born (in years).



### **Relationship information: mother**

- Relationship status [married and living together, in a relationship (couple not living together, cohabiting, living together part of the time, dating), not in a relationship (not in a relationship, separated, divorced)].
- Quality of the relationship between parents [high (excellent, good), low (average, poor, very poor)].
- Current physical violence or emotional abuse towards from partner (yes, no) recorded at baseline and any incidents over all collected time points (yes, no).
- History of emotional or physical abuse by partner or someone else important (yes, no) recorded at baseline and any incidents over all collected time points (yes, no).

### **Self-reported outcome measures: mother**

- PHQ-9 depression severity (score range 0–27, with a higher score indicating greater depressive symptoms).
- EQ-5D-5L quality of life (score range < 0 to 1, where 1 is a state of perfect health and 0 is death).
- EQ-5D-5L anxiety/depression [anxious or depressed (*slightly/moderately/severely/extremely anxious or depressed*), not anxious or depressed].

### **Pregnancy and parenting information**

- Being a first-time parent (yes/no).
- Attendance at parenting classes (yes/no).
- Number of parenting class sessions attended, at baseline and during the time of the trial.
- Problems during pregnancy (yes/no).
- Problems or difficulties at the time of child's birth (yes/no).
- Breastfeeding reported at baseline (yes/no).
- PSOC survey (score range 17–102, with higher scores indicating higher confidence).

### **Child characteristics**

- Gestational age (in weeks).
- Sex (male, female).
- Premature birth (yes/no).
- Age at baseline (in weeks).
- Health problems or serious injuries since birth (yes/no).

### **Service use during the trial**

- Mother hospitalisation (one or more, none), included only in analyses where this was not the outcome.
- Mother ED attendance (one or more, none), included only in analyses where this was not the outcome.

## **Results**

Demographic information on the 341 mother–baby dyads included in the analysis is presented in *Chapter 3*. Approximately 5% of babies had been born prematurely and approximately one-third had experienced health problems or serious injury since birth (*Table 41* provides full baseline characteristics by outcome status).

TABLE 41 Mother and child characteristics by outcome status

Characteristic	Total (N = 341)	Child admitted to hospital at least once (N = 65)	Child attended ED at least once (N = 145)	Mother admitted to hospital at least once (N = 36)	Mother attended ED at least once (N = 43)
<b>E-SEE Step trial allocation, n (%)</b>					
Intervention	285 (84)	57 (88)	124 (86)	32 (89)	35 (81)
Control	56 (16)	8 (12)	21 (15)	4 (11)	8 (19)
<b>Mother demographics</b>					
Age (in years)					
n (%)	341 (100)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	30.9 (5.0)	30.0 (5.5)	30.3 (5.0)	28.8 (4.9)	30.5 (5.9)
Median (IQR)	31 (28–34)	30 (26–33)	30 (26–33)	29 (25.5–32)	30 (27–35)
Ethnicity, n (%)					
White/white British	274 (80)	54 (84)	126 (87)	34 (94)	32 (74)
Asian/Asian British	50 (15)	8 (12)	15 (10)	2 (6)	7 (16)
Black/black British	8 (2)	1 (2)	2 (1)	0 (0)	0 (0)
Mixed/multiple ethnic groups	8 (2)	2 (3)	2 (1)	0 (0)	4 (9)
Other ethnic group	1 (0.3)	0 (0)	0 (0)	0 (0)	0 (0)
Education (highest attained), n (%)					
O Level or lower	33 (10)	8 (12)	13 (9)	6 (17)	4 (9)
A, AS or S Levels	26 (8)	2 (3)	11 (8)	6 (17)	7 (16)
Vocational qualification	61 (18)	12 (19)	24 (17)	8 (22)	7 (16)
University or a degree higher than A Level	192 (56)	37 (57)	88 (61)	13 (36)	21 (49)
Other	27 (8)	6 (9)	8 (6)	3 (8)	4 (9)
Missing	2 (1)	0 (0)	1 (1)	0 (0)	0 (0)
Employment, n (%)					
Paid employed	250 (73)	46 (71)	104 (72)	20 (56)	25 (58)
Unemployed	84 (25)	17 (26)	37 (26)	14 (39)	16 (37)
Other	6 (2)	2 (3)	4 (3)	2 (6)	2 (5)
Missing	1 (0.3)	0 (0)	0 (0)	0 (0)	0 (0)
Weekly income (£)					
n (%)	269 (79)	52 (80)	113 (78)	32 (89)	36 (84)
Mean (SD)	738.50 (467.50)	751.80 (527.50)	776.20 (522.20)	539.10 (400.30)	609.60 (394.80)
Median (IQR)	630 (400–973)	675 (365–1000)	650 (400–1000)	500 (250–633)	545.50 (372.50–730)

TABLE 41 Mother and child characteristics by outcome status (continued)

Characteristic	Total (N = 341)	Child admitted to hospital at least once (N = 65)	Child attended ED at least once (N = 145)	Mother admitted to hospital at least once (N = 36)	Mother attended ED at least once (N = 43)
Number of biological children (including this child), n (%)					
1	147 (43)	28 (43)	72 (50)	17 (47)	20 (47)
2	123 (36)	23 (35)	51 (35)	10 (28)	11 (26)
3	44 (13)	10 (15)	15 (10)	5 (14)	8 (19)
4–6	24 (7)	4 (6)	5 (4)	4 (11)	4 (9)
Missing	3 (01)	0 (0)	2 (1)	0 (0)	0 (0)
Age at first child					
n (%)	340 (99.7)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	27.5 (5.4)	26.3 (5.5)	27.6 (5.7)	25.0 (5.5)	26.9 (5.8)
Median (IQR)	28 (23–31)	27 (23–30)	28 (24–32)	25 (21–30)	27 (23–31)
<b>Mother: relationship information</b>					
Relationship status, n (%)					
Married and living together	222 (65)	37 (57)	89 (61)	19 (53)	26 (61)
In a relationship or living together	104 (31)	21 (32)	47 (32)	13 (36)	13 (30)
Not in a relationship	15 (4)	7 (11)	9 (6)	4 (11)	4 (9)
Quality of relationship, n (%)					
Excellent	162 (48)	26 (40)	69 (48)	8 (22)	17 (40)
Good	94 (28)	17 (26)	37 (26)	12 (33)	15 (35)
Average	10 (3)	3 (5)	5 (3)	2 (6)	1 (2)
Poor	4 (1)	1 (2)	1 (1)	0 (0)	0 (0)
Very poor	1 (0.3)	0 (0)	0 (0)	0 (0)	1 (2)
Missing	55 (16)	11 (17)	24 (17)	10 (28)	6 (14)
Not in a relationship	15 (4)	7 (11)	9 (6)	4 (11)	4 (9)
Current partner physically violent or emotionally abusive, n (%)					
Yes	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
No	271 (79)	47 (72)	112 (77)	22 (61)	33 (77)
Prefer not to answer	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Missing	55 (16)	11 (17)	24 (16)	10 (28)	6 (14)
Not in a relationship	15 (4)	7 (11)	9 (6)	4 (11)	4 (9)
History of emotional or physical abuse, n (%)					
Yes	40 (12)	11 (17)	22 (15)	7 (19)	8 (19)
No	231 (68)	38 (59)	92 (64)	15 (42)	26 (61)

continued

TABLE 41 Mother and child characteristics by outcome status (continued)

Characteristic	Total (N = 341)	Child admitted to hospital at least once (N = 65)	Child attended ED at least once (N = 145)	Mother admitted to hospital at least once (N = 36)	Mother attended ED at least once (N = 43)
Prefer not to answer	3 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Missing	67 (20)	16 (25)	31 (21)	14 (39)	9 (21)
<b>Mother: self-reported outcome measures</b>					
EQ-5D-5L quality-of-life index					
n (%)	341 (100)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	0.94 (0.08)	0.92 (0.10)	0.93 (0.08)	0.92 (0.08)	0.92 (0.07)
Median (IQR)	1.00 (1–1)	0.90 (0.87–1)	0.95 (0.87–1)	0.93 (0.86–1)	0.94 (0.87–1)
EQ-5D-5L anxiety/depression, n (%)					
Not anxious or depressed	273 (80)	46 (71)	106 (73)	23 (64)	35 (81)
Slightly anxious or depressed	51 (15)	13 (20)	29 (20)	10 (28)	5 (12)
Moderately or severely anxious or depressed	17 (5)	6 (9)	10 (7)	3 (8)	3 (7)
PHQ-9 depression score					
n (%)	341 (100)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	3.0 (3.4)	3.7 (4.5)	3.2 (3.7)	4.9 (4.1)	3.7 (2.8)
Median (IQR)	2 (1–4)	2 (1–5)	2 (1–5)	4 (2.5–6)	3 (1–6)
PHQ-9 depression classification, n (%)					
No depression	241 (71)	41 (63)	98 (68)	16 (44)	25 (58)
Mild depression	83 (24)	18 (28)	39 (27)	16 (44)	17 (40)
Moderate or severe depression	17 (5)	6 (9)	8 (6)	4 (11)	1 (2)
<b>Mother: pregnancy and parenting information</b>					
First-time parent, n (%)					
Yes	145 (43)	28 (43)	72 (50)	16 (44)	20 (47)
No	196 (58)	37 (57)	73 (50)	20 (56)	23 (54)
Attended parenting classes, n (%)					
Yes	125 (37)	27 (42)	59 (41)	11 (31)	15 (35)
No	216 (63)	38 (59)	86 (59)	25 (69)	28 (65)
Number of parenting classes attended (not IY)					
n (%)	341 (100)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	6.2 (13)	0.8 (2.4)	0.5 (1.7)	0.3 (0.7)	0.3 (1.1)
Median (IQR)	0 (0–5)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)

TABLE 41 Mother and child characteristics by outcome status (continued)

Characteristic	Total (N = 341)	Child admitted to hospital at least once (N = 65)	Child attended ED at least once (N = 145)	Mother admitted to hospital at least once (N = 36)	Mother attended ED at least once (N = 43)
Problems during pregnancy, n (%)					
Yes	148 (43)	37 (57)	71 (49)	20 (56)	26 (60)
No	193 (57)	28 (43)	74 (51)	16 (44)	17 (40)
Problems/difficulties at the time of birth, n (%)					
Yes	183 (54)	43 (66)	83 (57)	29 (81)	26 (60)
No	158 (46)	22 (34)	62 (43)	7 (19)	17 (40)
Breastfeeding, n (%)					
Yes	299 (88)	58 (89)	125 (86)	30 (83)	35 (81)
No	42 (12)	7 (11)	20 (14)	6 (17)	8 (19)
PSOC					
n (%)	341 (100)	65 (100)	145 (100)	36 (100)	43 (100)
Mean (SD)	80.8 (9.6)	79.2 (10.7)	79.4 (9.6)	79.0 (10.2)	81.8 (7.6)
Median (IQR)	82.0 (74–88)	81.0 (72–87)	81.0 (74–86)	80.5 (71–85)	81.0 (77–87)
<b>Child</b>					
Age at baseline (weeks)					
n	341	65	145	36	43
Mean (SD)	6.0 (2.1)	6.2 (1.9)	6.4 (2.0)	6.1 (2.0)	5.9 (2.3)
Median (IQR)	6.0 (4.0–8.0)	6.0 (5.0–8.0)	7.0 (5.0–8.0)	6.0 (4.5–8.0)	5.0 (4.0–8.0)
Sex, n (%)					
Male	174 (51)	33 (51)	80 (55)	22 (61)	29 (67)
Female	167 (49)	32 (49)	65 (45)	14 (39)	14 (33)
Gestational age (in weeks)					
n (%)	338 (99)	65 (100)	145 (100)	36 (100)	42 (98)
Mean (SD)	39.6 (1.5)	38.8 (1.9)	39.3 (1.6)	39.5 (1.7)	39.5 (1.4)
Median (IQR)	39.7 (39.0–40.6)	39.0 (37.9–40.0)	39.3 (38.4–40.4)	40.0 (38.2–40.6)	40.0 (38.6–40.4)
Born prematurely, n (%)					
Yes	12 (4)	7 (11)	8 (6)	1 (3)	1 (2)
No	327 (96)	57 (88)	136 (94)	35 (97)	42 (98)
Missing	2 (1)	1 (2)	1 (1)	0 (0)	0 (0)
Health problems or serious injury since birth, n (%)					
Yes	103 (30)	36 (55)	51 (35)	11 (31)	10 (23)
No	238 (70)	29 (45)	94 (65)	25 (69)	33 (77)
A Level, Advanced Level; AS Level, Advanced Subsidiary Level; IQR, interquartile range; O Level, Ordinary Level; S Level, Scholarship Level.					

### Child outcomes

For the children, there was a total of 95 hospital admissions recorded across all time points (for 65 children, with up to five admissions per child) and 263 ED attendances (for 145 children, with up to 14 per child) (Table 42). Admissions were related predominantly to respiratory conditions (40%) and other infectious diseases (21%) (Table 43). Table 44 illustrates parent and child characteristics that were significant predictors of child hospital admissions and ED attendances from the univariate and multivariate regression models.

TABLE 42 Summary of hospital admissions and ED attendances

Number of events and length of stay	Mother		Child	
	Hospital admission	ED attendance	Hospital admission	ED attendance
Number of events/individuals				
Number of events	46	67	95	263
Number of individuals with one or more events (% total)	36 (10.6)	43 (12.6)	65 (19.1)	145 (42.5)
Number of events per individual				
<i>n</i>	36	43	65	145
Mean (SD)	1.3 (0.7)	1.6 (1.3)	1.5 (0.8)	1.8 (1.5)
Median (IQR)	1 (1-1)	1 (1-2)	1 (1-2)	1 (1-2)
Minimum, maximum	1, 4	1, 8	1, 5	1, 14
Length of stay (overnights)				
<i>n</i>	36		65	
Mean (SD)	6.2 (14.5)		3.2 (3.2)	
Median (IQR)	2 (1-5)		2 (1-5)	
Minimum, maximum	1, 87		1, 19	

IQR, interquartile range.

TABLE 43 Reasons for hospital admission (self-reported)

Reason	Mother (N = 46 admissions), n (%)	Child (N = 95 admissions), n (%)
Injury	0 (0)	3 (3.2)
Digestive conditions	1 (2.2)	8 (8.4)
Monitoring	2 (4.4)	5 (5.3)
Neonatal conditions		6 (6.3)
Other infectious disease	5 (10.9)	20 (21.1)
Reproductive health	28 (60.8)	4 (4.1)
Respiratory conditions	1 (2.2)	38 (40.0)
Surgery	6 (13.0)	5 (5.3)
Other	3 (6.5)	6 (6.3)

TABLE 44 Significant predictors of outcomes related to the child

Predictor	Univariate			Multivariate		
		95% CI	p-value		95% CI	p-value
	OR			OR		
Outcome: at least one hospital admission (child)						
Mother not in a relationship (vs. married and living together)	3.92	1.24 to 12.43	<b>0.020</b>	2.38	0.69 to 8.22	0.172
Problems during pregnancy (vs. no problems)	1.86	1.04 to 3.34	<b>0.037</b>	1.13	0.59 to 2.16	0.711
Child born prematurely (vs. born full term)	5.47	1.42 to 21.07	<b>0.014</b>	1.70	0.34 to 8.39	0.516
Child health problems or serious injuries since birth (vs. no health problems)	3.70	2.01 to 6.81	<b>&lt; 0.001</b>	3.25	1.76 to 6.01	<b>&lt; 0.001</b>
Number of parenting classes attended	1.19	1.01 to 1.41	<b>0.041</b>	1.07	0.86 to 1.34	0.546
Gestational age (in weeks)	0.72	0.60 to 0.88	<b>0.001</b>	0.78	0.61 to 1.00	0.050
Age of mother at birth of first child (in years)	0.94	0.89 to 0.99	<b>0.021</b>	0.97	0.92 to 1.03	0.376
Mother EQ-5D-5L score at baseline	0.02	0.00 to 0.53	<b>0.020</b>	0.14	0.00 to 4.73	0.270
	IRR			IRR		
Outcome: frequency of hospital admissions (child)						
Mother not in a relationship (vs. married and living together)	3.54	1.28 to 9.81	<b>0.015</b>	0.82	0.22 to 3.09	0.775
Problems during pregnancy (vs. no problems)	1.80	1.04 to 3.11	<b>0.037</b>	1.21	0.68 to 2.15	0.510
Child born prematurely (vs. born full term)	3.81	1.18 to 12.31	<b>0.026</b>	0.99	0.27 to 3.61	0.985
Difficulties at the time of child's birth (vs. no problems)	2.00	1.12 to 3.56	<b>0.018</b>	1.71	0.97 to 3.01	0.062
Child health problems and serious injuries since birth (vs. no health problems)	2.83	1.64 to 4.87	<b>&lt; 0.001</b>	2.75	1.62 to 4.68	<b>&lt; 0.001</b>
History of emotional or physical abuse by partner or someone important: baseline (vs. no history of abuse)	0.41	0.19 to 0.86	<b>0.018</b>			
History of emotional or physical abuse by partner or someone important: over all follow-ups (vs. no history of abuse)	1.89	1.01 to 3.54	<b>0.048</b>	1.17	0.62 to 2.22	0.625
Gestational age	0.76	0.64 to 0.89	<b>0.001</b>	0.82	0.66 to 1.00	0.052
Age of mother at birth of first child (in years)	0.92	0.87 to 0.97	<b>0.001</b>	0.96	0.91 to 1.02	0.168
	OR			OR		
Outcome: at least one ED attendance (child)						
Mother not in a relationship (vs. married and living together)	3.77	1.09 to 13.08	<b>0.037</b>	1.40	0.42 to 4.70	0.586
First-time parent (vs. existing parent)	1.80	1.11 to 2.93	<b>0.018</b>	1.06	0.49 to 2.28	0.878
Mother attending ED at least once during trial	2.10	1.00 to 4.39	<b>0.049</b>	2.34	1.13 to 4.84	<b>0.022</b>
Mother slightly, moderately or severely anxious or depressed based on baseline EQ-5D-5L subscale score (vs. not depressed)	1.85	1.01 to 3.38	<b>0.046</b>	2.40	1.30 to 4.41	<b>0.005</b>
Age of mother at baseline (years)	0.92	0.87 to 0.97	<b>0.001</b>	0.98	0.93 to 1.03	0.440
Gestational age (in weeks)	0.74	0.62 to 0.88	<b>0.001</b>	0.73	0.61 to 0.88	<b>0.001</b>
Child's age at baseline (in weeks)	1.21	1.07 to 1.36	<b>0.002</b>	1.19	1.06 to 1.35	<b>0.004</b>
Number of biological children	0.72	0.56 to 0.94	<b>0.016</b>	0.69	0.45 to 1.06	0.090

continued

TABLE 44 Significant predictors of outcomes related to the child (continued)

Predictor	Univariate			Multivariate		
		95% CI	p-value		95% CI	p-value
	IRR			IRR		
Outcome: frequency of ED attendances (child)						
Site 2 (vs. site 3)	0.54	0.31 to 0.93	<b>0.026</b>	0.37	0.21 to 0.62	< <b>0.001</b>
Site 1 (vs. site 3)	0.61	0.37 to 0.99	<b>0.047</b>	0.46	0.29 to 0.73	<b>0.001</b>
Mother not in a relationship (vs. married and living together)	3.17	1.57 to 6.41	<b>0.001</b>	2.18	0.91 to 5.20	0.079
First-time parent (vs. existing parent)	1.59	1.10 to 2.29	<b>0.013</b>	1.25	0.88 to 1.79	0.213
Problems during pregnancy (vs. no problems)	1.45	1.01 to 2.09	<b>0.044</b>	1.31	0.92 to 1.88	0.134
Difficulties at the time of child's birth (vs. no problems)	1.60	1.10 to 2.33	<b>0.013</b>	1.34	0.94 to 1.91	0.105
Child health problems and serious injuries since birth (vs. no health problems)	1.55	1.05 to 2.27	<b>0.026</b>	1.38	0.98 to 1.94	0.066
Mother slightly, moderately or severely anxious or depressed based on baseline EQ-5D-5L subscale score (vs. not anxious or depressed)	1.57	1.02 to 2.42	<b>0.041</b>	1.24	0.84 to 1.84	0.274
History of emotional or physical abuse by a partner or someone important: baseline (vs. no history of abuse)	0.56	0.36 to 0.88	<b>0.012</b>			
History of emotional or physical abuse by a partner or someone important: over all follow-ups (vs. no history of abuse)	1.55	1.02 to 2.34	<b>0.038</b>	1.12	0.75 to 1.66	0.582
Age of mother at baseline (years)	0.95	0.91 to 0.98	<b>0.003</b>	0.96	0.92 to 1.00	<b>0.028</b>
Gestational age (in weeks)	0.85	0.76 to 0.96	<b>0.009</b>	0.90	0.80 to 1.01	0.074
Child's age at baseline (in weeks)	1.10	1.00 to 1.20	<b>0.044</b>	1.10	1.02 to 1.20	<b>0.016</b>
<b>Note</b>						
Bold text indicates <i>p</i> -values of < 0.5.						

In univariate regressions predicting hospital admissions, significant predictors included physical problems during pregnancy, being born prematurely, child health problems since birth and mothers being of younger age and not currently in a relationship. In a combined regression using all significant univariate predictors, only child health problems or serious injury since birth remained statistically significant (OR 3.25, 95% CI 1.76 to 6.01;  $p < 0.001$ ). A similar pattern was observed for the number of admissions, with only serious child health problems remaining statistically significant in the combined regression model (IRR 2.75, 95% CI 1.62 to 4.68;  $p < 0.001$ ).

Based on univariate regressions, ED attendances were more likely among children who had a lower gestational age, were older by the time of entering the trial and whose mothers were first-time parents, had fewer existing children, were younger, not in a relationship, attended ED at least once themselves during the course of the trial and who had rated themselves as moderately or severely anxious or depressed (on the EQ-5D dimension) at baseline. In a combined regression model, factors that remained statistically significant were gestational age (OR 0.73, 95% CI 0.61 to 0.88;  $p = 0.001$ ), age of child when entering the trial (OR 1.19, 95% CI 1.06 to 1.35;  $p = 0.004$ ), mothers' ED attendance (OR 2.34, 95% CI 1.13 to 4.84;  $p = 0.022$ ) and EQ-5D anxiety or depression (OR 2.40, 95% CI 1.30 to 4.41;  $p = 0.005$ ). The frequency of ED attendances was associated with a large number of univariate predictors, including higher frequencies at sites 3 and 4, for children of lower gestational age or who had serious health problems since birth, and mothers who were younger, first-time parents, not in a



relationship, had problems during pregnancy or delivery, had a history of emotional or physical abuse and rated themselves as moderately anxious or depressed. The mother's age (IRR 0.96, 95% CI 0.92 to 1.00;  $p = 0.028$ ) and child's age (IRR 1.10, 95% CI 1.02 to 1.20;  $p = 0.016$ ) at baseline and geographical site [summary effect of site  $\chi^2_{(df=3)} = 16.79$ ;  $p < 0.001$ ] remained statistically significant in a multivariate regression.

### Mother outcomes

For the mothers, there were a total of 46 hospital admissions recorded (for 36 mothers, up to four admissions per person) and 67 ED attendances (for 43 mothers, up to eight attendances per person). Admissions were related predominantly to reproductive health (61%), surgery (13%) and other infectious diseases (11%). Table 45 illustrates parent and child characteristics that were significant predictors of mother hospital admissions and ED attendances from the univariate and multivariate regression models.

TABLE 45 Significant predictors of outcomes related to the mother

Outcome	Univariate			Multivariate		
	OR	95% CI	p-value	OR	95% CI	p-value
At least one hospital admission (mother)						
Unemployed (vs. paid employed)	2.40	1.09 to 5.28	<b>0.030</b>	1.08	0.37 to 3.17	0.885
University or a degree higher than A Level (vs. O Level or lower)	0.30	0.10 to 0.87	<b>0.027</b>	0.50	0.11 to 2.22	0.364
Problems or difficulties at the time of child's birth (vs. no problems or difficulties)	4.07	1.73 to 9.59	<b>0.001</b>	2.18	0.82 to 5.83	0.119
Mother having attended the ED department at least once during the trial (vs. having not attended)	4.15	1.76 to 9.74	<b>0.001</b>	2.36	0.84 to 6.62	0.104
Mother slightly, moderately or severely anxious or depressed based on baseline EQ-5D-5L subscale score (vs. not anxious or depressed)	3.34	1.53 to 7.28	<b>0.002</b>	2.31	0.80 to 6.65	0.120
History of emotional or physical abuse by a partner or someone important: baseline (vs. no history of abuse)	0.35	0.13 to 0.92	<b>0.034</b>			
History of emotional or physical abuse by a partner or someone important: over all follow-ups (vs. no history of abuse)	2.78	1.27 to 6.09	<b>0.011</b>	1.92	0.72 to 5.12	0.190
Weekly income (£)	1.00	1.00 to 1.00	<b>0.023</b>	1.00	1.00 to 1.00	0.425
Age of mother at baseline (years)	0.91	0.85 to 0.98	<b>0.017</b>	0.96	0.87 to 1.06	0.415
Age of mother when first child was born (years)	0.93	0.86 to 0.99	<b>0.027</b>	0.99	0.88 to 1.10	0.803
PHQ-9 score at baseline	1.13	1.04 to 1.23	<b>0.004</b>	1.08	0.94 to 1.23	0.273
	<b>IRR</b>			<b>IRR</b>		
Frequency of hospital admissions (mother)						
Unemployed (vs. paid employed)	2.89	1.34 to 6.24	<b>0.007</b>	1.16	0.53 to 2.52	0.715
Problems or difficulties at the time of child's birth (vs. no problems or difficulties)	3.45	1.47 to 8.08	<b>0.004</b>	2.02	0.88 to 4.61	0.097

continued

TABLE 45 Significant predictors of outcomes related to the mother (continued)

Outcome	Univariate			Multivariate		
		95% CI	p-value		95% CI	p-value
Mother having attended the ED department at least once during the trial (vs. having not attended)	4.29	1.90 to 9.69	< 0.001	3.39	1.66 to 6.93	<b>0.001</b>
Mother slightly, moderately or severely anxious or depressed based on baseline EQ-5D-5L subscale score (vs. not anxious or depressed)	3.70	1.74 to 7.85	<b>0.001</b>	3.10	1.14 to 8.45	<b>0.027</b>
History of emotional or physical abuse by a partner or someone important: over all follow-ups (vs. no history of abuse)	2.57	1.17 to 5.62	<b>0.019</b>	1.51	0.73 to 3.14	0.268
Weekly income (£)	1.00	1.00 to 1.00	<b>0.012</b>	1.00	1.00 to 1.00	0.465
Age of mother at baseline (years)	0.89	0.83 to 0.95	<b>0.001</b>	0.98	0.91 to 1.05	0.516
Age of mother when first child was born (years)	0.91	0.85 to 0.97	<b>0.007</b>	0.97	0.90 to 1.06	0.547
PHQ-9 score at baseline	1.15	1.05 to 1.27	<b>0.003</b>	1.08	0.99 to 1.18	0.072
	<b>OR</b>			<b>OR</b>		
At least one ED attendance (mother)						
Mixed or multiple ethnic group (vs. white ethnicity)	7.47	1.78 to 31.42	<b>0.006</b>	9.62	2.19 to 42.27	<b>0.003</b>
Having a male child (vs. female child)	2.34	1.10 to 4.99	<b>0.028</b>	2.08	1.03 to 4.20	<b>0.042</b>
Mother being hospitalised at least once during the trial (vs. not being hospitalised)	4.15	1.76 to 9.78	<b>0.001</b>	4.15	1.81 to 9.56	<b>0.001</b>
	<b>IRR</b>			<b>IRR</b>		
Frequency of ED attendances (mother)						
EQ-5D-5L index score at baseline	0.01	0.00 to 0.96	<b>0.048</b>			
A Level, Advanced Level; O Level, Ordinary Level.						
<b>Note</b>						
Bold text indicates <i>p</i> -values < 0.5.						

In univariate regressions, hospital admission was more likely for mothers who were younger, less educated, unemployed, had a lower weekly income, had attended ED at least once during the trial, had a history of emotional or physical abuse, were more anxious or depressed based on the EQ-5D-5L and PHQ-9 and had problems during delivery. However, none of these predictors was statistically significant in the combined multivariate model. The frequency of hospital admissions was associated with similar factors univariately, and having attended ED at least once during the trial (IRR 3.39, 95% CI 1.66 to 6.93;  $p = 0.001$ ) and being slightly, moderately or severely depressed based on the EQ-5D-5L (IRR 3.10, 95% CI 1.14 to 8.45;  $p = 0.027$ ) remained statistically significant in the multivariate model.

Significant predictors of ED attendances were the same in the univariate and multivariate models. ED attendance was more likely for mothers of mixed ethnicity (OR 9.62, 95% CI 2.19 to 42.27;  $p = 0.003$ ), mothers whose child in the E-SEE Steps study was a boy (OR 2.08, 95% CI 1.03 to 4.20;  $p = 0.042$ ) or who had been hospitalised at least once during the trial (OR 4.15, 95% CI 1.81 to 9.56;  $p = 0.001$ ).

Higher frequency of these attendances was significantly associated with lower EQ-5D-5L quality-of-life scores at baseline (IRR 0.01, 95% CI 0.00 to 0.96;  $p = 0.048$ ), but no other predictors.

### Discussion

This study explored the predictors for ED attendance and hospital admission for children and their mothers enrolled in the E-SEE Steps trial. We found that ED attendance of the child was predicted by their younger gestational age, older age at recruitment to the trial, mothers' poorer mental health and younger age, the mother attending ED herself and the study site (i.e. sites 3 and 4). Health problems or injury since birth predicted a child's hospital admission and children were most often admitted for respiratory or other infectious diseases. For mothers, ED attendance was predicted by mixed ethnic origin, having a boy, having poorer quality of life at baseline and having been hospitalised during the trial. The mothers' admission to hospital was predicted by having attended the ED for themselves at least once and being anxious or depressed. Mothers were most commonly admitted to hospital for reproductive health issues.

The study benefits from a comprehensive data set, with few missing data. However, the limited sample size meant that we had reduced power in the multivariate analyses, although the main trial was sufficiently powered for outcome analyses in *Chapter 3*. Furthermore, these data were self-reported and the report of relationship abuse was likely to be under-reported.

A complex multilevel model of ED attendance and admission has emerged. It is undeniable that poorer mental health in mothers has a major impact on ED utilisation for both them and their babies. More exploration is needed to understand the complex pathway for this and to find ways in which to support new mothers at the most appropriate time during their journey to parenthood, and to understand how local ED accessibility affects the utilisation of health services.

## Substudy 2: Fitting Fathers and Co-parents In – an exploration of fathers' and co-parents' engagement with parenting support

### Background

Much of the evidence for parenting support programmes is based on mothers' experiences and less is understood about the appropriateness of parenting programmes for fathers.<sup>122</sup> Fathers play a vital role in the social, emotional and cognitive development of young children,<sup>124</sup> and the well-being of their child's mother.<sup>125</sup> However, the needs of fathers and co-parents are poorly understood and can be marginalised by parenting support services.<sup>123</sup>

### Aim

The aim of the substudy was to examine the E-SEE Steps fathers'/co-parents' reasoning and resources for engagement or disengagement with parenting programmes. These data are used to help understand how organisations and group leaders can engage with, and retain, fathers and other co-parents within programmes and services.

### Study design and methods

This substudy was approved by the University of Central Lancashire e-Ethics Committee for Science, Technology, Engineering, Medicine and Health. An interpretive approach guided the study design and data collection activities.

### Sample

Our sample comprised fathers and co-parents recruited as participants to either arm of the E-SEE Steps trial ( $n = 68$ ) who had indicated an interest in ongoing research ( $n = 51$ ) and were eligible for inclusion. Twenty-one fathers responded to an invitation to join the Fitting Fathers and Co-parents In (FFCPin) substudy.

## Data collection

Individual telephone interviews were used for data collection purposes. Participating fathers indicated consent for interviews to be audio-recorded and anonymised information to be used as substudy data. Interviewers (KW and Lisa Kane) used a topic guide to sequence questions depending on E-SEE Steps course or other group attendance.

## Data analysis

Audio-recorded interviews were verbatim transcribed and anonymised for analysis. The framework approach<sup>126</sup> was used to analyse all data, broadly following the five steps outlined by Pope *et al.*<sup>127</sup> This approach enabled the analysis to be 'grounded' and inductive, while also maintaining a specific focus on topics, such as programme engagement.

## Results

All 21 fathers lived with their index partner and child (age range 9–18 months; median 10 months) and were recruited from E-SEE Steps trial sites (site 1,  $n = 4$ ; site 2,  $n = 4$ ; site 3,  $n = 7$ ; site 4,  $n = 6$ ). Two fathers identified as being of South Asian heritage and the remaining identified as white British. The fathers were aged between 27 and 48 years. All except one father were employed (full time,  $n = 17$ ; part time,  $n = 2$ ) or self-employed ( $n = 2$ ) in occupations including taxi driver, armed forces, hover/ferry services, psychologist, photographer, human resources, teaching, postal service and publican. All FFCPin fathers had accessed courses or groups designed to support parents, including local antenatal classes ( $n = 11$ ), community groups (e.g. baby swim) ( $n = 10$ ), music, yoga, sensory classes, National Childbirth Trust's new parent groups and 'dads'-specific groups ( $n = 2$ ). Five fathers identified receiving invitations to E-SEE Steps IY-I or IY-T groups, although only one had attended (IY-I) sessions.

Analysis of the full data set revealed two core themes (i.e. 'time' and 'relevance') and associated subthemes. To preserve participant anonymity, all names have been changed when presenting fathers' quotations.

## Time

The theme of time concerned having 'enough time' and the 'timing' of sessions.

### Enough time

Limited time was raised by most of the fathers who explained that a busy life often meant that they had competing demands, with little space in their week for joining a group or course. As Ed put it:

*... generally I'm busy or working and Bob specified that an 8-week course, for me personally that would be a big commitment, I would maybe do a week or two a day or two-day courses at find extremely difficult to commit to longer.*

Ed

Those fathers invited but not attending the E-SEE Steps programme gave reasons of work demands and session timings for non-attendance:

*I haven't attended any of them, my wife has been to some of them [ ... ] The biggest one for me is the timing with work, from 8 o'clock to 5 o'clock I'm not available to do stuff which limits you to weekends or taking time off to attend.*

Ryan

### Timing

Difficulty in attending the E-SEE Steps group could also be due to poor timing with respect to the parents' current state of health and thereby perceived ability to manage group attendance:

*No we didn't [attend E-SEE Steps], my partner suffers with severe anxiety and depression and for the first couple of sessions that we were due to go [ ... ] I was at work and I couldn't get it off because of obviously*

*work reasons short-staffed et cetera and would have meant my partner going on her own where she [ . . . ] feels severe anxiety she didn't want to do it on her own she wanted me with her.*

Liam

A general problem, as noted above by Ryan, was the daytime scheduling of groups for parents, which meant that many saw them as inaccessible for fathers:

*. . . groups that I've been, that I've looked into, and you know they run at times that are completely inconvenient for somebody that works full time.*

Ishmal

Daytime scheduling was also assumed to mean that they were meant for mothers who, when their babies were younger, were less likely to be working:

*They're all designed, so they[re] based on women and mothers and if you're a working father there's nothing on an evening or a weekend.*

Andy

This suggests that the timing of groups can also indirectly communicate messages concerning the relevance of groups to mothers or fathers.

When working fathers had managed to attend a group, they had been proactive in making arrangements to make their own attendance possible. Examples included Damien, who set up 'a forum for new dads' in the workplace, and Greg, who 'made an agreement with work which allowed me to have 3 days off a week' and during this time he attended a parenting course.

## Relevance

The theme relevance was connected to notions of the applicability of parenting groups to fathers, owing to the fact that they were generally dominated by women and thereby 'mother-heavy'. Relevance to fathers was also dependant on their own 'purpose' of attending.

## Mother-heavy

The fathers interviewed spoke of parent groups as places where conversations were 'just naturally sort of . . . mother-heavy, [laughs] you know what I mean rather than parent' (Finley).

For Ed, the father who attended the E-SEE Steps IY-I group, the three sessions he joined were noted as 'fairly friendly', although when he described his experience he stated that 'you are largely spectating and involved on a superficial level', which he believed was, in part, due to his own feeling that it was 'still targeted at mums', as opposed to the actual content or behaviours of group leaders. Ed noted that when there were mainly women in the room the talk was mainly by 'mothers, and they're talking to each other as mothers. Rather than as parents', limiting the relevance for men and reinforcing a sense of being a spectator.

Liam voiced that 'everything is for mums, baby yoga baby massage although dads can come, if a dad does go all the other mums kind of look [at] them like that's kind of weird'. In these environments, fathers referred to being 'an outsider and it can be really difficult to "break in" if you like, to the chat that the ladies are having' (Ozzie). Being the only dad could be quite 'nerve-racking' (Mike) and their outside status may be compounded by the relationships women have with each other outside the parent group:

*. . . it was quite difficult because the interaction wasn't really there. [ . . . ] they did a lot more together after the clubs which I wasn't really invited to but they, they tended to do like baby massage together and they did loads of things together whereas I only went to this thing.*

Greg

*It's kind of what you expect anyway, I don't feel anything, health visitors are there for the mum and baby to make sure they're healthy.*

Andy

*... whatever you do it's not going to be right for your partner, it's got to be right for the mothers.*

Ozzie

### **Purpose**

The reason to go would be to support the best actions for their child and partner:

*But I'm there for... I'm not there for me and how I feel, I'm there for my daughter to try and bring her up.*

Neil

*Whereas when I joined the groups I was just very much there to do it as support for my wife and imagine the other guys were as well. You know we didn't go there to like get new, you know, we were just there to support our wives, and if they made some new friends, brilliant! [laughter].*

Ed

For those who had not attended a parenting (education) course, there was uncertainty about what the sessions would cover. Calum presumed they'd 'encourage, I guess bonding and things like that I'm not 100% sure what goes on there', whilst Bob voiced 'I don't even know what the parenting groups are to be honest!'

Many of the fathers expressed that they needed a clear purpose for attending a parenting course and would want it to be a learning experience, 'gaining information' (Finley) 'from the experts... with expert knowledge' (Calum). Therefore, attendance at a course 'would depend on what the subject was and if I felt I needed it at the time' (Andy). Otherwise, as Harry sums up 'I think if I went to a group and didn't feel that it was educational, I would withdraw from it':

*The subject areas would have to appeal to me and be something that I felt I was getting something from it, so it would have to be around kind of you know behaviour, you know development, that type of thing.*

Ishmal

New parenthood was identified as a relevant stage for attending parenting classes, when motivation was high and the purpose was to 'know what to expect' (Mike) and be prepared for 'what we were gonna be going through' (Ishmal):

*So the motivation for that was it was our first child, obviously, we haven't been through anything like this before and it would be nice to have some kind of advice on what to expect.*

John

### **Conclusion**

Fathers interviewed as part of the FFCPin substudy provided evidence of a continued perception of parenting groups and services being oriented towards mothers, especially in the early days. It was clear that the fathers in this study were interested in being prepared for parenthood and learning about how to support the health and development of their child. However, the female-dominated conversations and being in the minority as a man made engagement with sessions difficult. Service providers could assist with the engagement of fathers in parenting groups/courses by ensuring that information about the purpose and content of courses is clearly communicated. Course group leaders also have an important role in managing the focus of sessions to reduce sex bias and, instead, support a parent focus to which fathers, co-parents and mothers can contribute.

## Substudy 3: an E-SEE Steps programme and ENRICH comparison

### Background

Our final substudy involved a comparative project between the E-SEE Steps programme and the 6-year ENRICH research programme funded by the Health Research Board (Dublin, Ireland).

One of the two models that has been evaluated as part of the ENRICH programme is called the parent and infant model, a central pillar of which includes the IY-I and IY-T programmes.

### Aim

The original overarching aim of this substudy was to assess the effectiveness of IY-I and IY-T in enhancing children's (and, where applicable, parents') social and emotional well-being via a pooled analysis of data drawn from the ENRICH<sup>104</sup> and E-SEE Steps trials. Unfortunately, however, it transpired that this was not possible because of differences in the ways that these data were collected and the lack of scope for meaningful comparisons across jurisdictions.

Therefore, the specific objectives of the study shifted to using both data sets ( $n = 395$  parent–infant dyads) to (1) investigate the feasibility of assessing sample representativeness using publicly available data in two countries and to subsequently identify the challenges involved in considering the generalisability of study findings, and (2) undertake a comparative case study analysis to gain an understanding of the processes and factors that influence the exploration, installation and initial implementation of group-based parenting programmes in real-world settings.

### Results

A paper reporting findings relating to objective 1 has been published.<sup>117</sup> In brief, we were unable to source open/available data sets that could be manipulated to provide information pertaining only to mothers/families of very young infants. Our paper describes three key challenges associated with establishing the representativeness of trial samples (i.e. how to define representativeness, availability of data sets and quality of data) and we make recommendations for the conduct of future research in this area.

A second paper pertaining to objective 2 is currently in preparation.







EME  
HSDR  
HTA  
PGfAR  
**PHR**

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).  
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
Department of Health and Social Care*

***Published by the NIHR Journals Library***