

## A group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations

Stathi, Afroditi; Withall, Janet; Greaves, Colin J; Thompson, Janice L; Taylor, Gordon; Medina-Lara, Antonieta; Green, Colin; Snowsill, Tristan; Johansen-Berg, Heidi; Bilzon, James; Gray, Selena; Cross, Rosina; Western, Max J; Koning, Jolanthe L de; Ladlow, Peter; Bollen, Jessica C; Moorlock, Sarah J; Guralnik, Jack M; Rejeski, W Jack; Hillsdon, Melvyn

DOI:  
[10.3310/MQBW6832](https://doi.org/10.3310/MQBW6832)

License:  
Creative Commons: Attribution (CC BY)

*Document Version*  
Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*  
Stathi, A, Withall, J, Greaves, CJ, Thompson, JL, Taylor, G, Medina-Lara, A, Green, C, Snowsill, T, Johansen-Berg, H, Bilzon, J, Gray, S, Cross, R, Western, MJ, Koning, JLD, Ladlow, P, Bollen, JC, Moorlock, SJ, Guralnik, JM, Rejeski, WJ, Hillsdon, M & Fox, KR 2022, 'A group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations: the REACT RCT', *Public Health Research*, vol. 10, no. 14. <https://doi.org/10.3310/MQBW6832>

[Link to publication on Research at Birmingham portal](#)

### General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

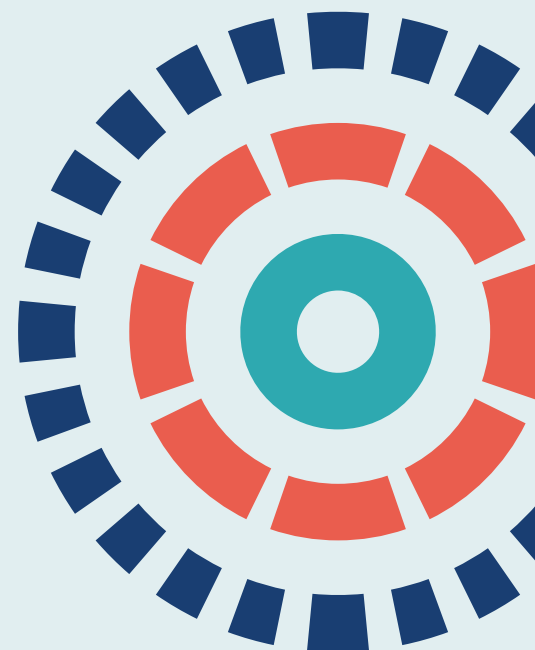
## Public Health Research

Volume 10 • Issue 14 • December 2022

ISSN 2050-4381

# A group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations: the REACT RCT

*Afroditi Stathi, Janet Withall, Colin J Greaves, Janice L Thompson, Gordon Taylor, Antonieta Medina-Lara, Colin Green, Tristan Snowsill, Heidi Johansen-Berg, James Bilzon, Selena Gray, Rosina Cross, Max J Western, Jolanthe L de Koning, Peter Ladlow, Jessica C Bollen, Sarah J Moorlock, Jack M Guralnik, W Jack Rejeski, Melvyn Hillsdon and Kenneth R Fox*





# A group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations: the REACT RCT

Afroditi Stathi<sup>1\*</sup> Janet Withall<sup>4</sup> Colin J Greaves<sup>1</sup>  
Janice L Thompson<sup>1</sup> Gordon Taylor<sup>2</sup>  
Antonieta Medina-Lara<sup>2</sup> Colin Green<sup>2</sup>  
Tristan Snowsill<sup>2</sup> Heidi Johansen-Berg<sup>3</sup>  
James Bilzon<sup>4</sup> Selena Gray<sup>5</sup> Rosina Cross<sup>4</sup>  
Max J Western<sup>4</sup> Jolanthe L de Koning<sup>4</sup>  
Peter Ladlow<sup>6</sup> Jessica C Bollen<sup>2</sup> Sarah J Moorlock<sup>1</sup>  
Jack M Guralnik<sup>7</sup> W Jack Rejeski<sup>8</sup> Melvyn Hillsdon<sup>9</sup>  
and Kenneth R Fox<sup>10</sup>

<sup>1</sup>School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK

<sup>2</sup>University of Exeter Medical School, St Luke's Campus, Exeter, UK

<sup>3</sup>Wellcome Centre for Integrative Neuroimaging, Oxford Centre for Functional MRI of the Brain, Wellcome Centre for Integrative Neuroimaging, John Radcliffe Hospital, University of Oxford, Oxford, UK

<sup>4</sup>Department for Health, University of Bath, Bath, UK

<sup>5</sup>Faculty of Health and Applied Sciences, University of the West of England Bristol, Bristol, UK

<sup>6</sup>Academic Department of Military Rehabilitation, Defence Medical Rehabilitation Centre, Loughborough, UK

<sup>7</sup>Department of Epidemiology and Public Health, University of Maryland, School of Medicine, Baltimore, MD, USA

<sup>8</sup>Department of Health and Exercise Science, Wake Forest University, Worrell Professional Centre, Winston-Salem, NC, USA

<sup>9</sup>College of Life and Environmental Sciences, University of Exeter, Exeter, UK

<sup>10</sup>Centre for Exercise, Nutrition and Health Sciences, School for Policy Studies, University of Bristol, Bristol, UK

\*Corresponding author

**Declared competing interests of authors:** Colin Green was a member of the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) General Committee (2019–20). Colin J Greaves was a member of the NIHR HTA Disease Prevention Panel (2009–13). Antonieta Medina-Lara is a member of the NIHR HTA Committee (2020–present); the South West for Research for Patient Benefit Programme (2021–present); and the NIHR Global Health Units & Groups Research Funding Committee (2021–present). Heidi Johansen-Berg is funded by the Wellcome Trust (110027/Z/15/Z) and the Oxford NIHR Biomedical Research Centre.



Published December 2022

DOI: 10.3310/MQBW6832

This report should be referenced as follows:

Stathi A, Withall J, Greaves CJ, Thompson JL, Taylor G, Medina-Lara A, *et al.* A group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations: the REACT RCT. *Public Health Res* 2022;**10**(14). <https://doi.org/10.3310/MQBW6832>



# Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

*Public Health Research* (PHR) was launched in 2013 and is indexed by Europe PMC, NCBI Bookshelf, DOAJ, INAHTA and Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA).

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@nih.ac.uk](mailto:journals.library@nih.ac.uk)

The full PHR archive is freely available to view online at [www.journalslibrary.nih.ac.uk/phr](http://www.journalslibrary.nih.ac.uk/phr).

## 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.nih.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 13/164/51. The contractual start date was in September 2015. The final report began editorial review in January 2020 and was accepted for publication in November 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 Stathi *et al.* This work was produced by Stathi *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.nih.ac.uk](http://www.journalslibrary.nih.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

---

**Dr Cat Chatfield** Director of Health Services Research UK

## NIHR Journals Library Editors

---

**Professor John Powell** 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** Interim Chair of HTA and EME Editorial Board. Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Ms Tara Lamont** Senior Adviser, School of Healthcare Enterprise and Innovation, 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, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Dr Rob Riemsma** Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, 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 group-based exercise and behavioural maintenance intervention for adults over 65 years with mobility limitations: the REACT RCT

Afroditi Stathi<sup>1\*</sup>, Janet Withall<sup>4</sup>, Colin J Greaves<sup>1</sup>,  
Janice L Thompson<sup>1</sup>, Gordon Taylor<sup>2</sup>, Antonieta Medina-Lara<sup>2</sup>,  
Colin Green<sup>2</sup>, Tristan Snowsill<sup>2</sup>, Heidi Johansen-Berg<sup>3</sup>,  
James Bilzon<sup>4</sup>, Selena Gray<sup>5</sup>, Rosina Cross<sup>4</sup>, Max J Western<sup>4</sup>,  
Jolanthe L de Koning<sup>4</sup>, Peter Ladlow<sup>6</sup>, Jessica C Bollen<sup>2</sup>,  
Sarah J Moorlock<sup>1</sup>, Jack M Guralnik<sup>7</sup>, W Jack Rejeski<sup>8</sup>,  
Melvyn Hillsdon<sup>9</sup> and Kenneth R Fox<sup>10</sup>

<sup>1</sup>School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK

<sup>2</sup>University of Exeter Medical School, St Luke's Campus, Exeter, UK

<sup>3</sup>Wellcome Centre for Integrative Neuroimaging, Oxford Centre for Functional MRI of the Brain, Wellcome Centre for Integrative Neuroimaging, John Radcliffe Hospital, University of Oxford, Oxford, UK

<sup>4</sup>Department for Health, University of Bath, Bath, UK

<sup>5</sup>Faculty of Health and Applied Sciences, University of the West of England Bristol, Bristol, UK

<sup>6</sup>Academic Department of Military Rehabilitation, Defence Medical Rehabilitation Centre, Loughborough, UK

<sup>7</sup>Department of Epidemiology and Public Health, University of Maryland, School of Medicine, Baltimore, MD, USA

<sup>8</sup>Department of Health and Exercise Science, Wake Forest University, Worrell Professional Centre, Winston-Salem, NC, USA

<sup>9</sup>College of Life and Environmental Sciences, University of Exeter, Exeter, UK

<sup>10</sup>Centre for Exercise, Nutrition and Health Sciences, School for Policy Studies, University of Bristol, Bristol, UK

\*Corresponding author [A.Stathi@bham.ac.uk](mailto:A.Stathi@bham.ac.uk)

**Background:** Mobility limitation in older age reduces quality of life, generates substantial health- and social-care costs, and increases mortality.

**Objective:** The RETirement in ACTION (REACT) trial aimed to establish whether or not a community-based active ageing intervention could prevent decline in physical functioning in older adults already at increased risk of mobility limitation.

**Design:** A multicentre, pragmatic, two-arm, parallel-group randomised controlled trial with parallel process and health economic evaluations.

**Setting:** Urban and semi-rural locations across three sites in England.

**Participants:** Physically frail or pre-frail older adults (aged  $\geq 65$  years; Short Physical Performance Battery score of 4–9). Recruitment was primarily via 35 primary care practices.

**Interventions:** Participants were randomly assigned to receive brief advice (three healthy ageing education sessions) or a 12-month, group-based, multimodal exercise and behavioural maintenance programme delivered in fitness and community centres. Randomisation was stratified by site and used a minimisation algorithm to balance age, sex and Short Physical Performance Battery score. Data collection and analyses were blinded.

**Main outcome measures:** The primary outcome was change in lower limb physical function (Short Physical Performance Battery score) at 24 months, analysed using an intention-to-treat analysis. The economic evaluation adopted the NHS and Personal Social Services perspective.

**Results:** Between June 2016 and October 2017, 777 participants (mean age 77.6 years, standard deviation 6.8 years; 66% female; mean Short Physical Performance Battery score 7.37, standard deviation 1.56) were randomised to the intervention arm ( $n = 410$ ) or the control arm ( $n = 367$ ). Data collection was completed in October 2019. Primary outcome data at 24 months were provided by 628 (80.8%) participants. At the 24-month follow-up, the Short Physical Performance Battery score was significantly greater in the intervention arm (mean 8.08, standard deviation 2.87) than in the control arm (mean 7.59, standard deviation 2.61), with an adjusted mean difference of 0.49 (95% confidence interval 0.06 to 0.92). The difference in lower limb function between intervention and control participants was clinically meaningful at both 12 and 24 months. Self-reported physical activity significantly increased in the intervention arm compared with the control arm, but this change was not observed in device-based physical activity data collected during the trial. One adverse event was related to the intervention. Attrition rates were low (19% at 24 months) and adherence was high. Engagement with the REACT intervention was associated with positive changes in exercise competence, relatedness and enjoyment and perceived physical, social and mental well-being benefits. The intervention plus usual care was cost-effective compared with care alone over the 2 years of REACT; the price year was 2019. In the base-case scenario, the intervention saved £103 per participant, with a quality-adjusted life-year gain of 0.04 (95% confidence interval 0.006 to 0.074) within the 2-year trial window. Lifetime horizon modelling estimated that further cost savings and quality-adjusted life-year gains were accrued up to 15 years post randomisation.

**Conclusion:** A relatively low-resource, 1-year multimodal exercise and behavioural maintenance intervention can help older adults to retain physical functioning over a 24-month period. The results indicate that the well-established trajectory of declining physical functioning in older age is modifiable.

**Limitations:** Participants were not blinded to study arm allocation. However, the primary outcome was independently assessed by blinded data collectors. The secondary outcome analyses were exploratory, with no adjustment for multiple testing, and should be interpreted accordingly.

**Future work:** Following refinements guided by the process evaluation findings, the REACT intervention is suitable for large-scale implementation. Further research will optimise implementation of REACT at scale.

**Trial registration:** This trial is registered as ISRCTN45627165.

**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. 14. See the NIHR Journals Library website for further project information.

# Contents

List of tables	xiii
List of figures	xvii
List of supplementary material	xxi
List of abbreviations	xxiii
Plain English summary	xxv
Scientific summary	xxvii
<b>Chapter 1</b> Background	<b>1</b>
Scientific background and explanation of rationale	1
Objectives and hypotheses	2
<i>Primary hypothesis</i>	2
<i>Secondary hypotheses</i>	2
<i>Primary objectives</i>	3
<i>Secondary objectives</i>	3
<b>Chapter 2</b> Trial design and methods	<b>5</b>
Ethics and governance	5
Public involvement and engagement	6
Trial design	6
Changes to the trial design	6
Participants	6
Eligibility screening	7
Consent	8
<i>Verbal consent</i>	8
<i>Written consent</i>	8
Study settings	9
<i>Primary care</i>	9
<i>Third-sector organisations</i>	9
<i>Word of mouth and snowball sampling</i>	10
<i>Local media</i>	10
Study intervention	10
<i>Intervention arm</i>	10
<i>Control arm</i>	11
Intervention delivery	11
Intervention fidelity	12
Health economics	12
Process evaluation	12
Sample size calculation	13
Interim analyses and stopping guidelines	13
Randomisation	13
<i>Randomisation type</i>	13
<i>Randomisation: allocation concealment mechanism</i>	14
<i>Randomisation: implementation</i>	14

## CONTENTS

Blinding	14
Unblinding	15
Statistical analyses	15
Primary and secondary outcome data analysis	15
Adjustment in the models	16
Subgroup analyses	16
Missing data	16
Multiple testing	16
Safety data	16
Amendments to the statistical analysis plan	17
Sensitivity analysis for the primary outcome data analysis	17
Statistical significance	17
Model assumptions	18
Outcomes	18
<i>Assessments</i>	18
<i>Primary hypothesis</i>	18
<i>Secondary hypotheses</i>	18
<i>Primary outcome</i>	18
<i>Secondary outcomes</i>	18
<i>Health economics outcomes</i>	19
<i>Process evaluation outcomes</i>	19
Changes to the study outcomes during the course of the study	19
<b>Chapter 3 Results: randomised controlled trial</b>	<b>21</b>
Participant flow	21
Recruitment	21
Baseline data	21
Main outcome results	25
Numbers analysed	27
Losses and exclusions	27
Programme adherence	27
Sensitivity analysis	27
Health economic results	28
Non-adherence to the protocol	28
Harms	28
Discussion	29
<i>Implications for practice and/or future research</i>	31
<b>Chapter 4 Process evaluation</b>	<b>33</b>
Aims	33
The REACT logic model	33
Study 1: intervention fidelity	34
<i>Methods</i>	34
<i>Results</i>	36
<i>Discussion</i>	36
<i>Conclusions</i>	40
Study 2: qualitative evaluation of intervention processes	40
<i>Design</i>	40
<i>Methods</i>	41
<i>Results</i>	43
<i>Discussion</i>	59

Study 3: quantitative evaluation of intervention mechanisms	63
<i>Aims</i>	63
<i>Methods</i>	63
<i>Results</i>	65
<i>Discussion</i>	75
<b>Chapter 5 Health economic evaluation</b>	<b>79</b>
Within-trial economic evaluation	79
<i>Aims and objectives</i>	79
<i>Methods</i>	79
<i>Results</i>	83
<i>Discussion</i>	88
Model-based economic evaluation	91
<i>Aims and objectives</i>	91
<i>Decision problem</i>	91
<i>Methods</i>	92
<i>Results</i>	96
<i>Discussion</i>	100
<b>Chapter 6 Discussion</b>	<b>103</b>
Key findings and interpretation	103
Strengths, limitations and generalisability	105
Considerations for further research	106
Implications for decision-makers and for practice	107
Conclusion	109
<b>Acknowledgements</b>	<b>111</b>
<b>References</b>	<b>115</b>
<b>Appendix 1 Randomised controlled trial</b>	<b>125</b>
<b>Appendix 2 Process evaluation</b>	<b>137</b>
<b>Appendix 3 Health economic evaluation</b>	<b>147</b>



# List of tables

<b>TABLE 1</b> Eligibility criteria for participation in the REACT study	<b>8</b>
<b>TABLE 2</b> Baseline characteristics	<b>23</b>
<b>TABLE 3</b> Primary and secondary outcomes at 24 months	<b>25</b>
<b>TABLE 4</b> Behaviour change techniques used in each of the sampled sessions	<b>35</b>
<b>TABLE 5</b> Characteristics of the sessions sampled	<b>36</b>
<b>TABLE 6</b> Behaviour change technique and overall intervention fidelity scores	<b>37</b>
<b>TABLE 7</b> Participant characteristics	<b>44</b>
<b>TABLE 8</b> Implications for practice	<b>62</b>
<b>TABLE 9</b> Considerations for research	<b>63</b>
<b>TABLE 10</b> Mechanistic hypotheses and analysis methods used to test them	<b>66</b>
<b>TABLE 11</b> Summary of mediation of intervention effects on SPPB by physical activity variables	<b>72</b>
<b>TABLE 12</b> Summary of moderation of intervention effects from 0 to 12 months and from 0 to 24 months on MSEQ, physical activity (PASE total) and SPPB by demographic variables	<b>73</b>
<b>TABLE 13</b> Unit costs of health and social care resources	<b>81</b>
<b>TABLE 14</b> Breakdown of costs (£) across different categories (excluding intervention costs)	<b>84</b>
<b>TABLE 15</b> Health-related quality of life measured in the REACT study	<b>85</b>
<b>TABLE 16</b> Quality-adjusted life-years in the REACT study	<b>86</b>
<b>TABLE 17</b> Within-trial cost-effectiveness results	<b>88</b>
<b>TABLE 18</b> Cost-effectiveness scenario analyses (within-trial economic evaluation)	<b>90</b>
<b>TABLE 19</b> Scenario analyses for model-based cost-effectiveness analysis	<b>102</b>
<b>TABLE 20</b> The REACT outcome measures and time points	<b>125</b>
<b>TABLE 21</b> The REACT study baseline characteristics compared with the UK population aged over 65 years	<b>127</b>
<b>TABLE 22</b> Study withdrawals	<b>128</b>



<b>TABLE 23</b> Loss to follow-up	<b>128</b>
<b>TABLE 24</b> Details of unexpected SAEs	<b>129</b>
<b>TABLE 25</b> Primary and secondary outcomes at 6 months	<b>132</b>
<b>TABLE 26</b> Primary and secondary outcomes at 12 months	<b>133</b>
<b>TABLE 27</b> Primary and secondary outcomes at 24 months	<b>134</b>
<b>TABLE 28</b> Primary outcome sensitivity analyses using MI without clustering by exercise group and using three population parameters to explore the impact of adherence	<b>135</b>
<b>TABLE 29</b> Analysis of prespecified subgroups: subgroup variables added as covariates and interaction terms with study arm in main model	<b>135</b>
<b>TABLE 30</b> Repeated-measures mixed-model analysis of the primary outcome (SPPB total score) measured at baseline and 6, 12 and 24 months post randomisation (adjusted for site, exercise group, age, sex and baseline SPPB score)	<b>136</b>
<b>TABLE 31</b> Behavioural change techniques and delivery processes included in the REACT intervention fidelity analysis	<b>137</b>
<b>TABLE 32</b> The Dreyfus scoring system for assessing competence in the delivery of clinical consultations and adaptations for scoring REACT delivery fidelity (see also <i>Report Supplementary Material 1</i> )	<b>138</b>
<b>TABLE 33</b> Criteria for summarising scores for each fidelity checklist item	<b>138</b>
<b>TABLE 34</b> Examples of REACT delivery practice	<b>139</b>
<b>TABLE 35</b> Cost of the REACT intervention	<b>152</b>
<b>TABLE 36</b> Sensitivity analyses for intervention cost	<b>153</b>
<b>TABLE 37</b> Baseline resource use and utility values	<b>153</b>
<b>TABLE 38</b> Within-trial cost of health and social care resource usage (excluding intervention costs)	<b>155</b>
<b>TABLE 39</b> Subgroup analyses (within-trial economic evaluation)	<b>155</b>
<b>TABLE 40</b> Predictions from model 1	<b>161</b>
<b>TABLE 41</b> Predictions from model 2	<b>161</b>
<b>TABLE 42</b> Results of ordered logistic regression (model 1) to predict SPPB at 24 months	<b>162</b>
<b>TABLE 43</b> Results of ordered logistic regression (model 2) to predict SPPB at 24 months	<b>163</b>

<b>TABLE 44</b>	Table of input parameters for model-based economic evaluation	<b>164</b>
<b>TABLE 45</b>	Variance-covariance matrix for mobility model 1	<b>169</b>
<b>TABLE 46</b>	Variance-covariance matrix for model for EQ-5D utility by SPPB	<b>170</b>
<b>TABLE 47</b>	Variance-covariance matrix for cost by SPPB	<b>170</b>
<b>TABLE 48</b>	Economic outcomes by Markov model state	<b>170</b>
<b>TABLE 49</b>	Deterministic model-based economic evaluation results	<b>171</b>
<b>TABLE 50</b>	Population average cost-effectiveness results	<b>171</b>



# List of figures

<b>FIGURE 1</b> The REACT logic model and associated data collection	4
<b>FIGURE 2</b> The CONSORT flow diagram of participants through the REACT study	22
<b>FIGURE 3</b> Thematic trees identified from participant experiences: reasons to engage in the REACT study and benefits of participation	45
<b>FIGURE 4</b> Thematic trees identified from participant experiences: barriers to and enablers of participation	46
<b>FIGURE 5</b> Causal diagram for mediators and moderators of intervention effects on physical function	64
<b>FIGURE 6</b> Changes in daily unbouted MVPA from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI)	67
<b>FIGURE 7</b> Changes in muscle-strengthening exercise (MSEQ adherence score) from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI)	68
<b>FIGURE 8</b> Changes in SPPB total score from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI)	68
<b>FIGURE 9</b> Changes in relatedness in relation to MVPA from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI)	70
<b>FIGURE 10</b> Changes in competence in relation to muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI)	71
<b>FIGURE 11</b> Changes in relatedness for muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI)	71
<b>FIGURE 12</b> Changes in enjoyment in relation to muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI)	71
<b>FIGURE 13</b> Changes in perceived benefits of muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI)	72
<b>FIGURE 14</b> Tornado diagram for one-way sensitivity analyses of resource values (NHS/PSS perspective)	87
<b>FIGURE 15</b> Cost-effectiveness scatter plot for within-trial analysis	89
<b>FIGURE 16</b> Cost-effectiveness acceptability curves for within-trial analysis	90
<b>FIGURE 17</b> Expected trial-end (24-month) SPPB score by age, sex and group (error boxes show $\pm 1$ standard error)	93
<b>FIGURE 18</b> Cost by SPPB score range	95

<b>FIGURE 19</b> Predicted EQ-5D utility according to SPPB score	95
<b>FIGURE 20</b> Mean SPPB score and proportion alive in the Markov model	96
<b>FIGURE 21</b> Cumulative incremental discounted QALYs over time	97
<b>FIGURE 22</b> Cumulative incremental discounted costs over time	97
<b>FIGURE 23</b> Heterogeneity analysis, absolute lifetime discounted costs and QALYs	98
<b>FIGURE 24</b> Heterogeneity analysis: incremental QALYs	99
<b>FIGURE 25</b> Heterogeneity analysis: incremental costs	99
<b>FIGURE 26</b> Heterogeneity analysis: INMB	99
<b>FIGURE 27</b> Tornado diagram for one-way sensitivity analyses	101
<b>FIGURE 28</b> The REACT logic model in the context of findings: reasons to engage in REACT	141
<b>FIGURE 29</b> The REACT logic model in the context of findings: benefits of participation in REACT	142
<b>FIGURE 30</b> The REACT logic model in the context of findings: barriers to participation in REACT and daily physical activity	143
<b>FIGURE 31</b> The REACT logic model in the context of findings: enablers of participation in REACT and daily physical activity	144
<b>FIGURE 32</b> Changes in hours spent walking outside from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI)	145
<b>FIGURE 33</b> Changes in adherence to muscle-strengthening exercise (MSEQ adherence score) from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI)	145
<b>FIGURE 34</b> Changes in PASE total score from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI)	145
<b>FIGURE 35</b> Changes in PASE total score from 0 to 12 months for different levels of attendance within the intervention arm (error bars: 95% CI)	146
<b>FIGURE 36</b> Missing data for EQ-5D and NHS/PSS costs	149
<b>FIGURE 37</b> Correlogram of missingness of health-related quality-of-life measures at different time points	149
<b>FIGURE 38</b> Correlogram of missingness of costs at different timepoints from NHS/PSS and societal perspectives	150
<b>FIGURE 39</b> The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram for pragmatic review of economic evaluations	158

<b>FIGURE 40</b> Steady state distribution for Model 2	<b>163</b>
<b>FIGURE 41</b> Probabilistic sensitivity analysis in 65-year-old men: cost-effectiveness plane	<b>171</b>
<b>FIGURE 42</b> Probabilistic sensitivity analysis in 65-year-old men: cost-effectiveness acceptability curve	<b>172</b>



# List of supplementary material

**Report Supplementary Material 1** Intervention fidelity

**Report Supplementary Material 2** Quantitative process evaluation

**Report Supplementary Material 3** Health economic evaluation

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/MQBW6832>).

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

A&E	accident and emergency	MVPA	moderate to vigorous physical activity
ADL	activities of daily living	NCC	National Cost Collection
ANCOVA	analysis of covariance	NHS EED	NHS Economic Evaluation Database
BCT	behaviour change technique	NICE	National Institute for Health and Care Excellence
BMI	body mass index	NIHR	National Institute for Health and Care Research
CCA	complete-case analysis	OR	odds ratio
CI	confidence interval	PASE	Physical Activity Scale for the Elderly
CONSORT	Consolidated Standards of Reporting Trials	PIS	participant information sheet
DMEC	Data Monitoring and Ethics Committee	PSA	probabilistic sensitivity analysis
EQ-5D	EuroQol-5 Dimensions	PSS	Personal Social Services
EQ-5D-3L	EuroQol-5 Dimensions, three-level version	QALY	quality-adjusted life-year
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	QWB-SA	Quality of Well-Being Self-Administered
GP	general practitioner	RA	research assistant
HOT	higher-order theme	RCT	randomised controlled trial
HRG	Healthcare Resource Group	REACT	REtirement in ACTion
ICER	incremental cost-effectiveness ratio	REPS	Register of Exercise Professionals
ID	identification	SAE	serious adverse event
IMD	Index of Multiple Deprivation	SAP	statistical analysis plan
INMB	incremental net monetary benefit	SCT	social cognitive theory
ITT	intention to treat	SDT	self-determination theory
LIFE	Lifestyle Interventions and Independence for Elders	SES	socioeconomic status
MAT-SF	Mobility Assessment Tool-Short Form	SF-36	Short Form questionnaire-36 items
MeSH	medical subject heading	SF-6D	Short Form questionnaire-6 Dimensions
MI	multiple imputation	SPPB	Short Physical Performance Battery
MoCA	Montreal Cognitive Assessment	TMG	Trial Management Group
MRI	magnetic resonance imaging	TSC	Trial Steering Committee
MSEQ	Muscle-Strengthening Exercise Questionnaire	WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index



## Plain English summary

As people get older, they often find that walking, climbing stairs and doing their normal daily activities become more difficult. The REtirement in ACTion (REACT) study tested whether or not a group exercise programme run in local communities could reduce this decline in older adults, and whether or not it provided good value for money. The programme encouraged social interaction and fun, and provided support to find out about and get involved in other physical activities.

General practitioners in Bath, Bristol, Birmingham and Devon invited patients (aged  $\geq 65$  years) to join the REACT study, and 777 patients agreed to participate. In total, 411 participants were randomly chosen to join the REACT sessions twice per week for 12 weeks, then once per week for 40 weeks. The other 366 people attended three 'healthy ageing' education sessions over 1 year. All 777 participants completed physical tests and questionnaires at the start of the study and again at 6, 12 and 24 months. At 24 months, people who had attended the REACT sessions had significantly greater mobility than those who did not attend (8.08 vs. 7.59; scores out of 12). Only 19% of REACT study participants did not complete the study. Participants who attended the REACT group sessions enjoyed the REACT programme, felt more confident and reported better well-being. The REACT programme could help older adults at risk of mobility limitations to stay mobile and independent. REACT provides good value for money: the benefits outweighed the cost, resulting in a saving of £103 in health-care and social care costs per participant. Following some refinements, the REACT programme is ready for large-scale implementation.



# Scientific summary

## Introduction

With increasing age, there is a population-wide decline in physical function. In total, 44% of state pension-age adults in the UK are classified as disabled. The most common form of disability is mobility-related disability (67%). This is a major public health issue significantly reducing the independence and quality of life of older adults, while also contributing to high health-care and social care costs and increased mortality. Reduced gait speed and low levels of physical activity are key markers of frailty, which places an increased pressure on health-care systems worldwide.

The REtirement in ACTion (REACT) study was an effectiveness trial designed to offer real-world, low-cost, tailored exercise and social support to communities of older adults at risk of mobility limitations in the UK.

We hypothesised that, compared with a control arm, participants allocated to the 12-month physical activity and behavioural maintenance intervention would have significantly better lower limb physical function after 24 months' follow-up.

## Methods

### Study design

REACT was a pragmatic, multicentre, two-arm, single-blind, parallel-group randomised controlled trial with an internal pilot phase and comprehensive process and economic evaluations. Participants were recruited from three study sites in England: Bath and Bristol, Birmingham, and Devon. Ethics approval was provided by the NHS South East Coast-Surrey Research Ethics Committee (15/LO/2082). The trial registration number is ISRCTN45627165.

### Participants

Participants were community-dwelling adults, aged  $\geq 65$  years, not in full-time employment and scoring between 4 and 9 (inclusive) on the Short Physical Performance Battery (SPPB). These criteria identify people who have mobility limitations but are still ambulatory, and include people classified as frail (SPPB score of 4–7) and pre-frail (SPPB score of 8–9) by the European Medicines Agency. We excluded people who were unable to walk across a room, were living in residential care, were awaiting hip or knee surgery, were receiving radiation therapy or chemotherapy, had recent heart or spinal surgery or had severe illness that would prevent participation.

We recruited through letters of invitation from general practitioners (GPs) or third-sector/charity organisations and through local media and by word of mouth. We targeted a socioeconomically and ethnically diverse sample that included participants from urban, rural and semirural locations.

### Randomisation and masking

Participants meeting the study inclusion criteria were randomised either to the physical activity and behavioural maintenance intervention or to the control arm using a centralised randomisation website. We used a minimisation algorithm to balance groups by study site, age group (65–74 years,  $\geq 75$  years), sex and baseline functional ability (SPPB score of 4–7 or 8–9). During the internal pilot, the randomisation ratio was 2 : 1 (favouring the intervention) to enable early feasibility testing of intervention engagement. The main trial randomisation was 1 : 1. Throughout the study, 39 couples or pairs of close friends who were both eligible were randomised together to minimise contamination between study arms.

Researchers collecting the primary outcome data (SPPB), statisticians and the senior research team were blinded to group allocation.

### **Procedures**

The intervention arm received a 12-month exercise and behavioural maintenance programme delivered in community centres by qualified exercise professionals. The programme was delivered in three progressive phases: (1) adoption (weeks 1–8), (2) transition (weeks 9–24) and (3) maintenance (weeks 25–52). Established behaviour change techniques (BCTs) were used to enhance motivation, to make realistic plans for sustainable activity, to pre-empt and overcome barriers, to engage social support and to promote self-monitoring and self-regulatory techniques to support the maintenance of behaviour change. The exercise sessions were delivered twice per week for 12 weeks, reducing to once per week for a further 40 weeks, to groups of around 15 participants. The exercise sessions were designed to improve lower limb muscle strength and balance. Exercise prescription and progression methods were based on the functional requirements of each individual:

- **Adoption.** Each participant received a 30-minute individualised, face-to-face introductory session, and participated in two 60-minute group-based exercise sessions per week, which were followed by 20 minutes of social time with refreshments to encourage social interaction and promote session attendance.
- **Transition.** A 45-minute interactive behavioural maintenance session delivered by session leaders was added to one of the two weekly exercise sessions. These sessions were designed to encourage a ‘social club’ atmosphere and long-term maintenance of an active lifestyle, including ongoing exercise classes, home-based exercise, neighbourhood walking and active travel. They incorporated BCTs derived from social cognitive theory, self-determination theory and the skills for maintenance model. After week 12, the exercise session frequency reduced to once per week, with an expectation that participants found an hour per week to exercise at home or in the neighbourhood, or to attend a local exercise session.
- **Maintenance.** The maintenance stage focused further on home- and neighbourhood-based activities while continuing with a weekly centre-based physical activity session followed by a 20-minute social session. The frequency of the 45-minute behavioural maintenance sessions was reduced to once per month, and these sessions focused on enacting participants’ action plans for physical activity outside the REACT programme.

Control arm participants attended three 60- to 90-minute healthy ageing workshops over 24 months. These involved no physical activity content.

Measures were conducted at baseline and 6, 12 and 24 months post randomisation in a group setting at local community centres.

### **Outcomes**

The primary outcome was the SPPB score at 24 months. SPPB measures normal gait speed over 4 metres, time for completing a repeated (×5) rise from a chair, and completion of three standing balance tasks of increasing difficulty, yielding a score from 0 to 12.

Secondary outcomes were (1) measures of physical activity derived from wrist-worn accelerometer [GENEActiv Original (Activinsights, Kimbolton, UK)], including moderate to vigorous intensity physical activity (MVPA); (2) time spent in sedentary activities (time spent below the 40-mg threshold minus accelerometer-estimated sleep time); (3) self-reported physical activity [Physical Activity Scale for the Elderly (PASE) questionnaire]; (4) the Mobility Assessment Tool-Short Form (MAT-SF); (5) self-reported adherence to government guidelines on muscle-strengthening activity [Muscle-Strengthening Exercise Questionnaire (MSEQ)]; (6) grip strength; and questionnaire measures of (7) social well-being; (8) sleep quality; (9) hip, knee and ankle joint pain; (10) health-related quality of life; (11) loneliness; (12) cognitive function [using the UK Biobank Healthy Minds assessment process and the Montreal Cognitive Assessment (MoCA)]; and (13) the Falls Efficacy Scale-International.

A full cost-effectiveness analysis, including an estimation of the intervention costs, within-trial changes in quality-adjusted life-years and a model-based lifetime analysis, as well as detailed sensitivity analyses, were conducted. The analyses used the NHS and Personal Social Services perspective. The price year was 2019.

A mixed-methods process evaluation evaluated the fidelity of delivery of the behavioural maintenance programme using a checklist applied to session recordings, qualitative interviews with REACT participants and session leaders, and quantitative testing of hypotheses about the proposed mechanisms of change derived from the REACT logic model.

A nested substudy, led by the Wellcome Centre for Integrative Neuroimaging, University of Oxford, tested the hypothesis that the REACT intervention slows the rate of brain atrophy and cognitive function decline. This was funded by the National Institute for Health and Care Research (NIHR) Oxford Biomedical Centre, University of Oxford, and will be reported elsewhere.

## Statistical analysis

The power calculation for the primary outcome (SPPB score) at 24 months was based on the published definition of a minimum clinically meaningful change in SPPB score of 0.5 points. To provide 90% power, this required a sample size of 384 participants per arm (768 participants in total).

The primary comparative analysis was carried out using an intention-to-treat (ITT) approach with due emphasis placed on confidence intervals (CIs). The comparison of primary interest was the difference between the intervention and the control arm on SPPB score at 2-year follow-up (24 months after randomisation). Covariates in the model comprised baseline SPPB scores, age, sex and study site. In addition, we adjusted the estimates for clustering by exercise group within the intervention arm, with control arm participants entered as individual groups. Analyses were conducted in Stata® SE version 15.0 (StataCorp LP, College Station, TX, USA).

Secondary outcome analyses were undertaken using the same approach as for the primary analysis (excluding the sensitivity analyses), using the baseline, 6-month, 12-month and 24-month follow-up data, using linear mixed-regression models. Health-related quality of life, as assessed by EuroQol-5 Dimensions (EQ-5D), will be reported elsewhere as part of the health economic evaluation.

As an exploratory analysis, the effects of several predefined factors were further investigated, and these are presented. These included the stratification variables (age categories: 65–74 years and  $\geq 75$  years; sex; and study site: Bath/Bristol, Devon, Birmingham), as well as comorbidity levels at baseline (none or one chronic medical conditions vs. two or more chronic medical conditions), health-related quality of life, socioeconomic subgroups (using education, home ownership and quintiles of area deprivation), history of falls (recorded fall or not during 6 months prior to baseline) and the uptake of any co-interventions during the 24-month study period.

To examine the association between dose and response, we conducted subgroup analyses for participants attending  $\geq 50\%$  and  $\geq 75\%$  of the REACT programme sessions.

Intervention costs were estimated by identifying key resources (programme co-ordination, REACT session leader time and expenses, venue hire, equipment, consumables, and programme-specific training for REACT session leaders) and assigning values to the resources used. The data were collected by the REACT session leaders and trial manager. Costs were estimated per REACT group and then divided by the average group size.



## Results

Between June 2016 and September 2017, 3116 people were telephone screened (of whom 1077 were not eligible) and 1214 attended for baseline screening. Of these, 804 were found to be eligible and 777 were randomised (410 to the intervention arm and 367 to the control arm). The number of participants included in the primary analysis at 24 months was 628 (80.8%), comprising 334 (81.5%) in the intervention arm and 294 (80.1%) control subjects. The mean age of the participants was 77.6 years [standard deviation (SD) 6.8 years], 66.2% of participants were female, 95.1% were white and the mean baseline SPPB score was 7.37 (SD 1.56). Baseline characteristics were similar between the two study arms, although the prevalence of multimorbidity was higher in the intervention arm (47% vs. 39%). The sample was similar to the UK population aged over 65 years in terms of ethnicity and area deprivation, although with some under-representation of men (33.9% vs. 45.6%).

At the 24-month follow-up, the mean SPPB score (adjusted for baseline SPPB score, age, sex, study site and exercise group) was significantly higher in the intervention arm (mean 8.08, SD 2.87) than in the control arm (mean 7.59, SD 2.61), with an adjusted mean difference of 0.49 (95% CI 0.06 to 0.92;  $p = 0.014$ ). Of the 410 participants allocated to the intervention arm, 16.1% did not engage with any of the intervention sessions (non-starters), 19.0% attended < 50% of the sessions offered, 20.2% attended 50–74% of the sessions offered and 44.6% attended  $\geq 75\%$  of the sessions offered. In the case of participants allocated to the intervention arm who started the programme, the mean percentage of sessions attended was 67.7% (95% CI 65.1% to 70.4%). An association between dose and response was observed, with an adjusted mean difference in SPPB score of 0.64 (95% CI 0.23 to 1.05;  $p = 0.002$ ) for those attending  $\geq 50\%$  of intervention sessions and of 0.81 (95% CI 0.38 to 1.23;  $p < 0.001$ ) for those attending  $\geq 75\%$  of sessions.

The SPPB score was significantly higher in the intervention arm than in the control arm at 6 months (adjusted mean difference 0.68, 95% CI 0.39 to 0.96;  $p < 0.001$ ) and 12 months (adjusted mean difference 0.77 points, 95% CI 0.40 to 1.14;  $p < 0.001$ ). Self-reported physical activity was significantly higher in the intervention arm than in the control arm at 6 months (adjusted mean difference in PASE score of 16.3, 95% CI 6.78 to 25.9;  $p = 0.001$ ), 12 months (adjusted mean difference 10.8, 95% CI 3.18 to 18.5;  $p = 0.006$ ) and 24 months (adjusted mean difference 10.7, 95% CI 2.62 to 18.8;  $p = 0.010$ ). Self-reported engagement in muscle-strengthening exercise showed a similar pattern, with highly significant differences ( $p < 0.001$ ) at all three follow-up time points. Accelerometer data indicated a substantial difference favouring the intervention group at 12 months for total MVPA (adjusted mean difference 3.11 minutes per day, 95% CI 0.00 to 6.23;  $p = 0.05$ ) and MVPA accumulated in bouts of at least 10 minutes (1.24 minutes per day, 95% CI 0.22 to 2.26;  $p = 0.018$ ). This equates to a difference of 22 minutes per week of unbouted MVPA. Significant differences favouring the intervention arm were also observed in the Short Form questionnaire-36 items physical component score (at 6 and 12 months), hand-grip strength (at 12 months) and the MAT-SF self-reported lower limb physical functioning scale (at 6, 12 and 24 months). Sensitivity analyses, including imputation of missing values and not adjusting for clustering by exercise group, did not substantially change the pattern or significance of the results.

The mixed findings on secondary outcomes indicate that the effects of the intervention were limited to lower limb physical function and did not extend to substantial increases in physical activity or other domains of physical function (e.g. grip strength). The secondary outcome analyses were exploratory, with no adjustment for multiple testing, and should be interpreted accordingly.

During the study, 95 events were classified as serious adverse events. Only one (a hip fracture following a fall during an exercise session) was related to the study.

The full 12-month REACT programme was estimated to cost £622 per participant. The intervention plus usual care was cost-effective compared with usual care alone over the 2-year time period of REACT. In the base-case scenario analysis, the intervention saved £103 in NHS/Personal Social

Services (PSS) costs per participant with a quality-adjusted life-year (QALY) gain of 0.04 within the 2-year trial window. Lifetime horizon modelling estimated that further cost savings and QALY gains were accrued up to 15 years post randomisation.

In qualitative interviews, participants reported that they enjoyed the programme, and they reported better mental and social well-being, emphasising their higher physical confidence, improved motivation and feeling more outgoing. Improved social connectedness and bonding with REACT groups were key outcomes for the intervention group, who also highlighted improvements in mobility, strength, balance, walking, fitness, sleep and physical independence. Themes identified at 24 months largely mirrored those reported at 6 and 12 months. However, whereas the 6- and 12-month interviews found that social support was a key reason for engaging in REACT, at 24 months individual-level factors, such as perceived benefits, were more prominent themes in explaining physical activity maintenance. Key components of the REACT programme that positively influenced maintenance of physical activity at 24 months were (1) techniques for managing slips/lapses, supporting habit change and resolving sources of tension around increasing physical activity; (2) the person-centred delivery style to build autonomy/intrinsic motivation; and (3) the group-based delivery promoting social connectedness.

The quantitative process evaluation confirmed that, compared with control subjects, participants in the intervention arm reported experiencing greater benefits from exercising in terms of their physical, mental and social well-being. The hypothesis that increased exposure to the intervention will be associated with positive changes in competence, relatedness, enjoyment and perceived benefits (hypothesis 5c) was supported in relation to muscle-strengthening exercise only. Increased exposure to the intervention was associated with positive changes in psychosocial determinants for muscle-strengthening exercise 12 months after baseline, but not with changes in determinants of MVPA.

Taken together, the qualitative and quantitative process evaluations broadly supported the logic model for the REACT intervention. They identified several ways that the intervention and its implementation could be improved. These included possible changes to the logic model (from the qualitative and quantitative studies) and changes to delivery processes (from the intervention fidelity and qualitative studies).

The trial analyses did not show an impact on quality of life. However, the more sophisticated, time-integrated approach used in the health economic analysis revealed a significant difference in EQ-5D (as well as a saving in health-care costs). Indeed, the health economic analysis indicated that the increases in physical function observed were associated with substantial quality-of-life and health economic benefits, both within the 24-month trial window and across a lifetime horizon.

Although the overall results for REACT were positive, the process evaluation indicated, as with most service-based interventions, that there was considerable scope for improvement by session leaders in the facilitation of important self-regulation processes and social/relatedness-building processes during the delivery of the behavioural maintenance programme. To some extent, this may have been mitigated by mutual support among participants and self-delivery of some of the intended processes within the groups during the exercise sessions. However, future implementation of the REACT intervention should aim to improve the training and delivery of the programme accordingly.

## Conclusion

Among older adults with mobility limitations, lower limb physical function after 24 months' follow-up was significantly better among those who participated in the REACT 12-month, community-based group physical activity and behavioural maintenance programme than in the control arm.

Higher intervention effects were associated with increased programme attendance, with once-weekly exercise or more being associated with clinically meaningful benefit. The REACT intervention was cost-saving from an NHS/PSS perspective within a 2-year window, with further cost-savings and QALY benefits estimated in the longer term.

### **Trial registration**

This trial was registered as ISRCTN45627165.

### **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. 14. See the NIHR Journals Library website for further project information.

# Chapter 1 Background

Parts of this chapter have been adapted from Stathi *et al.*<sup>1</sup> Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The text includes minor additions and formatting changes to the original text.

## Scientific background and explanation of rationale

Healthy ageing is defined as ‘the process of developing and maintaining the functional ability that enables wellbeing in older age’.<sup>2</sup> Functional ability comprises the intrinsic capacities, both mental and physical, that people can draw on, relevant environmental characteristics and demands, and how people respond to these demands.

With increasing age, there is a population-wide decline in physical function.<sup>3-5</sup> Frailty and associated comorbidities compromise quality of life for older adults and contribute major societal costs directly to people who live with frailty, to friends and family providing care and losing productivity, and to health-care and social care services.<sup>6,7</sup> The impact of this transition is further heightened by an ever-increasing ageing population in the UK (18.2% in 2017 over the age of 65 years, rising to 20.7% by 2027<sup>8</sup>) that is also reflected worldwide.<sup>9,10</sup>

There is strong evidence that physical activity has a positive impact, slowing or preventing disability in later life.<sup>11-14</sup> A fit and active older person has 36% lower risk of developing functional limitations and 38% lower risk of hip fracture than inactive older people.<sup>15</sup> In the UK-based OPAL PLUS cohort study,<sup>16</sup> older people who undertook at least 25 minutes of moderate to vigorous physical activity (MVPA) per day at baseline received fewer prescriptions and were less likely to be admitted to hospital in an emergency 4–5 years later.<sup>16</sup> Despite these significant benefits, as people age they engage in less physical activity and spend more time being sedentary,<sup>13,17</sup> with 31% of 65- to 74-year-olds reporting < 30 minutes of MVPA per week, rising to 53% of people aged  $\geq 75$  years.<sup>5</sup>

This toxic mix of reduced physical activity leading to compromised physical function and increasing pressure on health-care and social care support services has shifted the focus towards supporting the maintenance of functional capacity in later life, with healthy ageing identified as a key public health priority.<sup>18,19</sup> Clinical trials have provided robust evidence that physical features of frailty, such as reduced muscular strength or endurance, can be reversed or their progression slowed by undertaking an appropriate exercise programme.<sup>14,20,21</sup> According to data from the Health Survey for England, 31% of women and 22% of men aged  $\geq 65$  years report needing help in the last month with one or more activities of daily living (ADLs), such as getting up and down stairs, dressing, getting around indoors or shopping.<sup>5</sup> These people are in transition from independence to frailty and have a great deal to gain if loss of function can be reversed and independence maintained.

Physical inactivity and mobility limitations in older people are more prevalent in socioeconomically deprived sectors of the population.<sup>22,23</sup> Ethnically diverse groups in the USA have a significantly greater risk of developing a range of physical and mental health problems than their white counterparts, and subsequently suffer higher rates of morbidity and premature mortality.<sup>24,25</sup> Self-reported data from the Health Survey for England and the Active Lives survey indicate that older ( $\geq 55$  years) Bangladeshi, Pakistani and Indian adults in England are less likely than their white counterparts to meet physical

activity guidelines.<sup>26-28</sup> Thus, exercise interventions that can successfully engage and retain inactive and ethnically diverse groups of older adults could contribute to the reduction of health and social inequalities.

Lifestyle Interventions and Independence for Elders (LIFE) was a landmark US study in the field of physical activity promotion in older adults.<sup>21</sup> LIFE was a multicentre randomised controlled trial (RCT) comparing the effects of a physical activity programme with a successful ageing educational programme in more than 1600 functionally impaired older persons, over an average follow-up period of 2.6 years. The intervention reduced the risk of developing major mobility disability (defined as the inability to complete a 400-metre walk test within 15 minutes) by 18% relative to the control group [hazard ratio (HR) 0.82, 95% confidence interval (CI) 0.69 to 0.98]. It also reduced the risk of persistent mobility disability by 28% (major mobility disability at consecutive time points) (HR 0.72, 95% CI 0.57 to 0.91). The intervention group maintained an increase of 40 minutes per week (95% CI 29 to 52 minutes;  $p < 0.001$ ) in objectively assessed lifestyle intensity activity ( $\geq 760$  counts/minute compared with the control group at 24 months of follow-up). These estimates are likely to be conservative because the study utilised an active control group that received a substantial health education/lifestyle intervention, including weekly workshops for 6 months and monthly sessions for a further 18 months. Being an efficacy trial, however, LIFE was highly resource intensive, and there was no post-intervention long-term follow-up. The challenge now is to build on this evidence base and develop affordable, scalable interventions that are suitable for delivery in specific contexts and demonstrate maintenance of effect in the long term.

Building on the findings of the LIFE RCT, we designed REtirement in ACTion (REACT), a pragmatic effectiveness trial to test whether or not the LIFE intervention could be adapted into an effective real-life, community programme for older adults at high risk of mobility limitations in the UK. To the best of our knowledge, REACT is the first large-scale UK-based study of its kind. If effective and cost-effective, this programme would offer important health-care and social care benefits, while sustaining health and independence among older adults.

### Objectives and hypotheses

The REACT study aimed to test the effectiveness and cost-effectiveness of a community, group-based physical activity and behaviour maintenance intervention based on the US LIFE programme for reducing or reversing the progression of functional limitations in older people who are at high risk of mobility-related disability.

A nested substudy, led by the Wellcome Centre for Integrative Neuroimaging, University of Oxford, tested the hypothesis that the REACT intervention slows the rate of brain atrophy and cognitive function decline. Measures included a brief battery of tests (paper and pencil and computerised) to assess memory, attention and executive function; structural and functional brain magnetic resonance imaging (MRI) measures; and gait analysis for a subsample of participants. This substudy was funded by the National Institute for Health and Care Research (NIHR) Oxford Biomedical Centre, University of Oxford, and is reported elsewhere.

#### Primary hypothesis

The null hypothesis is that participants allocated to receive the REACT programme will have the same mobility-related limitations, as indicated by Short Physical Performance Battery (SPPB) score, at 24 months of follow-up as the control group.

#### Secondary hypotheses

The null hypothesis is that participants allocated to the REACT intervention will have the same levels of moderate-intensity physical activity, health-related quality of life, cognitive function, ability to perform the ADL, and mental and social well-being at 24 months as the control group.

### Primary objectives

1. To adapt the LIFE intervention from the USA for delivery in UK community settings.
2. To conduct an internal pilot study to evaluate and optimise the feasibility and acceptability of the REACT intervention for older people and intervention providers and of the proposed trial methods across a diverse sample, spanning multiple ethnic groups and geographic areas varying in deprivation index.
3. To conduct a full-scale, pragmatic, multicentre RCT of the REACT intervention, with data collection at baseline and at 6, 12 and 24 months' follow-up.
4. To explore how intervention effectiveness varies with deprivation index and ethnicity (i.e. to explore potential effects on health inequalities).

### Secondary objectives

To compare intervention and control groups in terms of:

- minutes of moderate-intensity physical activity, as measured by accelerometer data
- sedentary time and breaks in sedentary time
- self-reported physical activity
- hand-grip strength of the dominant hand
- performance on a brief test of cognitive function
- the rate of brain atrophy and performance on more detailed tests of cognitive function and gait analysis tests (functional MRI substudy)
- mental and social well-being, energy, sleep quality and pain
- health-related quality of life
- ADL scores.

To conduct a full economic evaluation to estimate the incremental cost-effectiveness of the REACT intervention compared with the control group. To conduct a multimethod process evaluation to evaluate the feasibility of intervention implementation and inform future implementation and possible refinements of the intervention; to evaluate the intervention delivery and inform conclusions about intervention effectiveness; to investigate the proposed mechanisms of change outlined in the REACT logic model (*Figure 1*) and seek alternative explanations if this model is not supported; and to understand the role of context to inform whether or not and how the findings can be generalised.

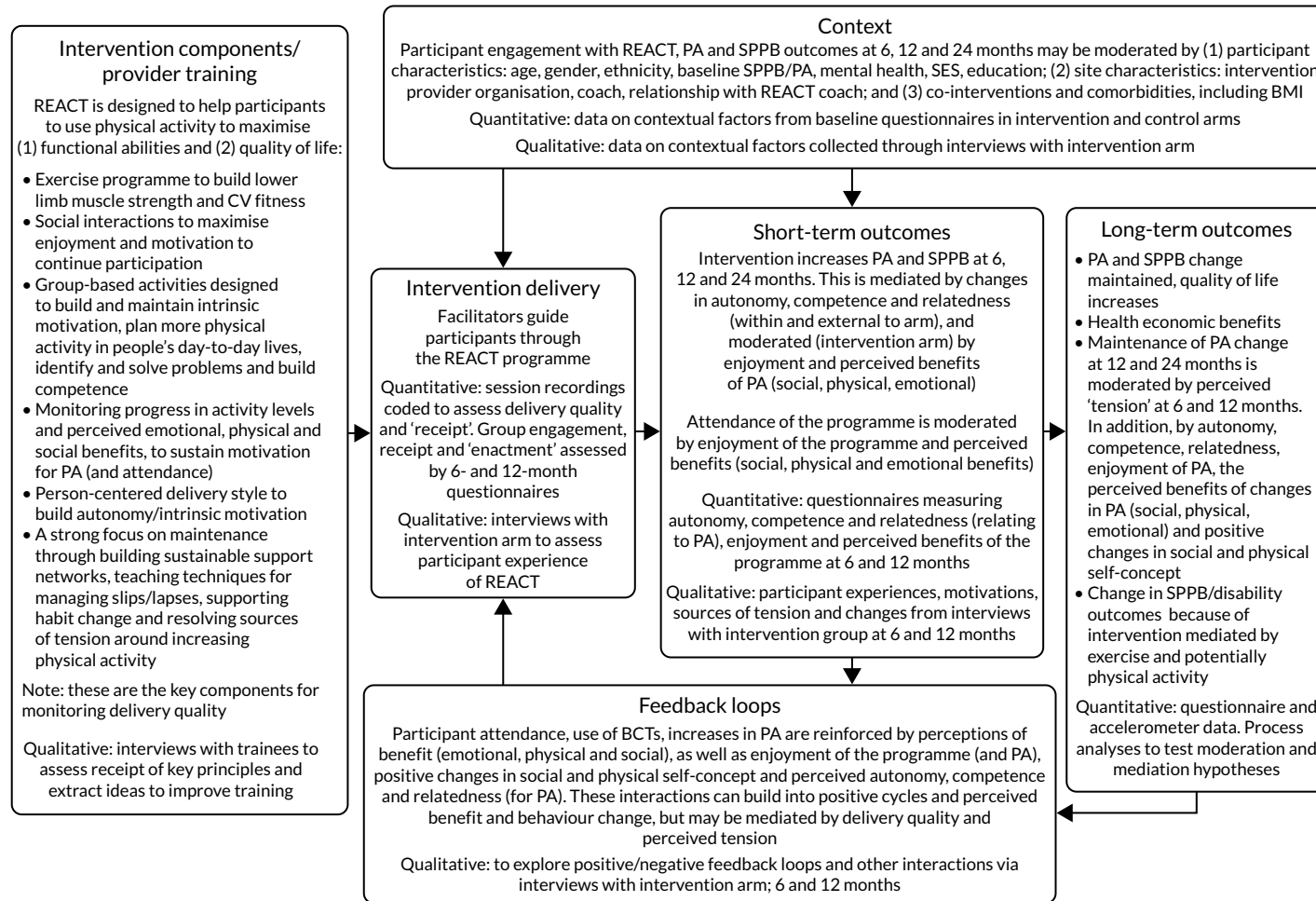


FIGURE 1 The REACT logic model and associated data collection. CV, cardiovascular; PA, physical activity.



## Chapter 2 Trial design and methods

Parts of this chapter have been adapted from Stathi *et al.*<sup>1</sup> Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The text includes minor additions and formatting changes to the original text.

Parts of this chapter have been adapted from Withall *et al.*<sup>29</sup> 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. The text includes minor additions and formatting changes to the original text.

### Ethics and governance

The REACT study was reviewed and approved by the NHS South East Coast – Surrey Research Ethics Committee (15/LO/2082).

A substantial amendment (approved 7 March 2016) was submitted to change the SPPB inclusion criteria from 4–8 to 4–9; to change the randomisation process in the internal pilot from 1 : 1 to 2 : 1; and to streamline the recruitment process.

A second substantial amendment (approved 28 October 2016) was submitted to revert to 1 : 1 randomisation in the main trial without balancing the sample imbalance owing to 2 : 1 randomisation in the pilot; to make a change to the randomisation process to balance sites on key variables; to make a change to the process of informing MRI study participants of incidental findings; to include an additional case report form to be used with a REACT group funded by Bristol Ageing Better (Bristol, UK); to add questions to the Telephone Screening Questionnaire; to randomise couples together; and to make an addition to the consent form regarding audio-recording of REACT sessions. Finally, we refined the recruitment materials and processes after recommendations by participants during the internal pilot study.

A third substantial amendment (approved 2 October 2017) was submitted to add socioeconomic status (SES), marital status and caring-related questions to the Telephone Screening Form; and to enable the requesting of participants' e-Frailty data from general practitioners (GPs) and also summary data from GP mailing databases to allow comparison of responders with non-responders.

A fourth substantial amendment (approved 9 January 2018) was submitted to gain approval for documentation relating to session leader interviews and to exclude planned hospitalisation from serious adverse event (SAE) reporting.

A fifth substantial amendment (approved 16 May 2018) was submitted because three co-applicants changed institution, and a small change was made allowing disclosure of a MRI scan to a GP at the participant's request. No further protocol changes were made. All research and development approvals were sought and obtained for each site and amendment. The study is registered as a current RCT (ISRCTN45627165).



## Public involvement and engagement

The REACT study was built on several years of multidisciplinary work by the study team aimed at understanding influences on the adoption and maintenance of physical activity in community-dwelling older adults. Our Avon Network for the Promotion of Active Ageing in the Community (AVONet) [Medical Research Council (MRC) Lifelong Health and Wellbeing – Collaborative Development Network; ref 90543] used focus groups and workshops with service providers, older people, international experts and service commissioners to assess the needs of older people and their communities for physical activity promotion.<sup>30</sup> The REACT study was considered by our AVONet service user, service provider and commissioner stakeholders to be suitable for delivery across a range of socioeconomic and cultural populations. The REACT protocol was developed based on this input. The Trial Management Group (TMG) was closely involved in the development of the study protocol, and all co-applicants, the trial manager and three people from our service user advisory group (research partners) formed part of that committee. One of the research partners also served on the Trial Steering Committee (TSC). The draft protocol was open to consultation by our service user representatives, our public health expert and members of community organisations. In addition, our Advisory Group (six members) reviewed REACT study materials prior to use.

## Trial design

The REACT study was a multicentre, pragmatic, two-arm, single-blind, parallel-group RCT with an internal pilot phase that incorporated comprehensive process and economic evaluations. Following identification and recruitment, 777 participants who met the study inclusion criteria were randomised to either the intervention arm or the control arm. The internal pilot phase assessed the feasibility of recruitment methods (allowing for some refinement) and confirmed that the prespecified criteria for retention to the intervention sessions and the study were met, prior to progressing with the main trial. A nested substudy, led by the Wellcome Centre for Integrative Neuroimaging at the University of Oxford, employed MRI to test the hypothesis that the REACT intervention slows the rate of brain atrophy and decline in cognitive function. This substudy was funded by the NIHR Oxford Biomedical Centre, University of Oxford, and is reported elsewhere.<sup>31</sup> Outcome data were collected at baseline and 6, 12 and 24 months. The protocol of this study was published as an open access publication in 2018.<sup>1</sup>

## Changes to the trial design

A substantial amendment was accepted (7 March 2016) to change the SPPB inclusion criterion (described below) from scores 4–8 (inclusive) to 4–9 (inclusive) out of 12. This change was deemed to be appropriate to support the recruitment efforts and to adopt the same inclusion criterion as the US LIFE trial.<sup>21</sup>

A further substantial amendment was accepted (28 October 2016) to (1) revert to 1 : 1 randomisation in the main trial without balancing the sample owing to 2 : 1 randomisation in the internal pilot study and (2) clarify the randomisation process regarding balancing sites on key variables [study site, age group, sex and initial functional ability (SPPB)].

These changes were discussed and agreed by the TMG and the TSC.

## Participants

The eligibility criteria were intended to identify sedentary, community-dwelling, older people aged  $\geq 65$  years, who were not in full-time employment and had functional limitations (i.e. who were at

risk of major mobility limitations), but who were still ambulatory (i.e. they could still walk). This was measured using a battery of objective physical function tests, the SPPB, a group of measures that combines the results of the gait speed, chair rise and balance tests:

- gait speed test – a timed walk over 4 metres at normal walking pace (minimum score 0, maximum score 4)
- chair rise test– time taken to stand up and sit down five times from a dining/kitchen-style chair, arms crossed across the chest (minimum score 0, maximum score 4)
- balance test – unsupported balances for 10 seconds in three positions – feet side by side, feet in semi-tandem and feet in full tandem position (minimum score 0, maximum score 4).

Summed, these tests generate a physical function score from 0 to 12. Older adults with scores of 4–9 (inclusive) out of 12 were eligible to take part in REACT. This criterion was based on data showing that older adults with SPPB scores of  $\leq 9$  have substantially higher risk of major mobility disability 3 years later [odds ratio (OR) 8.3, 95% CI 3.3 to 20.67] than those with a score of 12.<sup>21,32</sup> Furthermore, the European Medicines Agency has published specific cut-off points for SPPB to identify three stages of frailty (< 7, frail; 8–9, pre-frail; and 10–12, fit/normal function), which support the REACT study inclusion criterion.<sup>33</sup>

During recruitment, participants were informed that they would receive £15 shopping vouchers for attending each of the 6-, 12- and 24-month assessment sessions.

## Eligibility screening

The eligibility of respondents was assessed in a three-step sequential screening process:

1. Initial self-selection. The participant approach letters and participant information sheet (PIS) clearly stated that we were recruiting people who were still able to carry out, but had some difficulty with, daily activities, such as walking, climbing stairs and getting out of a chair.
2. Phone-based screening. After gaining verbal consent, a preliminary phone screen checked inclusion and exclusion criteria that could be assessed by telephone, including a check on medical exclusion criteria. We excluded people who were unable to walk across a room, who were living in residential care, who were awaiting hip or knee surgery, who were receiving radiation therapy or chemotherapy, who had received recent heart or spinal surgery, or who had severe illness that would prevent participation, including severe arthritis, diagnosed dementia, severe kidney disease, unstable heart disease and severe psychiatric illness. Availability to attend intervention sessions was also checked. Participants who did not meet the eligibility criteria were thanked for their time and provided with the Age UK (London, UK)/NHS guide to healthy ageing and other sources of further advice and information.<sup>34</sup>
3. Face-to-face screening sessions. Potentially eligible participants were invited to a group-based assessment session at a local community centre, at which they could ask questions about the study and were asked to give written informed consent. The SPPB gait speed test was conducted first, and those who failed to complete the 4-metre walk or did not meet the SPPB inclusion criterion (a score of 4–9 inclusive) were thanked for their time and provided with an information pack (as above). Participants who met the eligibility criteria were invited to take part in the remainder of the baseline assessments.

Full inclusion and exclusion criteria are summarised in *Table 1*.

TABLE 1 Eligibility criteria for participation in the REACT study

Criteria	Details
Inclusion criteria	<ul style="list-style-type: none"> <li>• Aged <math>\geq 65</math> years and not in full-time employment</li> <li>• Planning to reside in the target area (Bath/Bristol, Devon, Birmingham) for at least 24 months</li> <li>• Score between 4 and 9 (inclusive) on the SPPB</li> </ul>
Exclusion criteria	<ul style="list-style-type: none"> <li>• Self-reported inability to walk across a room without a walker or the help of another person</li> <li>• Existing major mobility limitation (defined as SPPB score of <math>\leq 3</math> or unable to complete the 4-metre walk component of SPPB)</li> <li>• Living in residential or nursing care</li> <li>• Inability to attend the REACT physical activity sessions as scheduled</li> <li>• A documented or patient-reported medical condition that would preclude participation, including: <ul style="list-style-type: none"> <li>○ arthritis so severe that it would prevent participation in physical activity</li> <li>○ Parkinson's disease or diagnosed dementia</li> <li>○ any terminal illness</li> <li>○ lung disease requiring use of oral corticosteroids or supplemental oxygen</li> <li>○ severe kidney disease requiring dialysis</li> <li>○ severe heart disease that would prevent participation in physical activity (e.g. chest pain when walking 100 or 200 yards or up a flight of stairs)</li> <li>○ implanted cardiac defibrillator</li> <li>○ cardiac arrest that required resuscitation</li> <li>○ severe uncontrolled psychiatric illness</li> <li>○ currently receiving radiation therapy or chemotherapy treatment for cancer</li> <li>○ awaiting knee or hip surgery</li> <li>○ major heart surgery (including valve replacement or bypass surgery) in the last 6 months</li> <li>○ unstable heart condition (e.g. uncontrolled arrhythmia, angina, heart failure or hypertension)</li> <li>○ spinal surgery in the last 6 months</li> <li>○ any other clinical condition that the person's GP or clinician considers would make them unsuitable for participation in a physical activity rehabilitation programme to prevent decline of lower limb functioning</li> </ul> </li> </ul>
Temporary exclusion criteria	<ul style="list-style-type: none"> <li>• Heart attack (or myocardial infarction), stroke, spinal surgery, hip fracture, or hip or knee replacement within the previous 6 months</li> <li>• Currently receiving physical therapy on legs or enrolled in another physical activity research or intervention study</li> </ul>

## Consent

Older adults who were willing to take part in REACT were asked to provide verbal informed consent at the beginning of the telephone screening call and written informed consent prior to commencement of the face-to-face screening sessions.

### Verbal consent

At the beginning of the telephone screening interview, the researcher checked that the participant had read the PIS that had been mailed to them, summarised the study verbally and then offered the participant the chance to ask questions. Verbal consent to the telephone screening was then requested. If the participant failed to give consent, the telephone screen was not conducted. If a participant provided verbal consent, the assignment of a study identification (ID) number was taken as positive evidence that consent was obtained.

### Written consent

#### The environment for consent

The setting in which written consent was obtained at the face-to-face screening session was as private as possible, so that participants could ask questions freely and without embarrassment. To avoid pressuring the participant, only one person associated with the study was present when the participant reviewed the consent form.

## The consent process

The consent process involved a full explanation of the study given by the person taking consent [research assistant (RA) or other authorised researcher], with reference to the PIS that had been mailed to each participant prior to any of the face-to-face screening processes commencing. Potential participants were informed that they may, at any time, withdraw their consent to participate without giving a reason and without it affecting their relationship with their GP, the referring organisation and/or their future treatment and care. The PIS provided details of a contact point at which participants could obtain further information about the study. Following these discussions, people who were willing to participate were asked to complete, sign and date the study consent form, which was also signed and dated by the person obtaining consent.

## Capacity to consent

To be eligible for participation, participants had to have the capacity to give their own informed consent. If a member of the research team considered a participant to be incapable of understanding what was expected of them as a subject in the study, it was not permissible for informed consent to be obtained from a guardian. The study required daily responsibilities that could not be easily assumed by other people. In line with Health Research Authority guidance ([www.hra-decisiontools.org.uk/consent/](http://www.hra-decisiontools.org.uk/consent/); accessed 13 October 2022), consent to participate in the REACT study was presumed to remain legally valid after loss of capacity (provided that the REACT protocol did not change significantly). In all cases, we took note of any signs of objection or distress from participants and consulted closely with them.

## Study settings

The study was conducted in Bath, Bristol, Birmingham and Devon, UK, allowing recruitment of a socioeconomically and ethnically diverse sample, comprising participants from urban, rural and semirural locations.

A range of recruitment strategies to identify suitable participants was employed.

### Primary care

General practices were invited to participate through their local Clinical Research Network and through the research teams' existing networks in primary care. Where possible, we selected practices to maximise diversity in terms of ethnicity, SES and, in Devon, rurality.

Practice staff searched their electronic patient databases for potentially eligible patients using the trial entry criteria that were routinely coded in the database. Lists generated from the searches were further screened for suitability by a GP. GPs focused on screening for items that were not included/partially covered by the electronic searches (e.g. recent bereavement, awaiting knee or hip surgery). A recruitment pack consisting of a participant approach letter printed on the practice headed notepaper, a reply form and the PIS was sent to suitable patients, with a reply-paid envelope addressed to the local research team. GPs and practice nurses also offered the recruitment pack in surgery to patients they considered may be eligible.

### Third-sector organisations

The principal investigators and RAs at each trial site engaged with third-sector and community-based organisations who worked with adults over the age of 65 years. Professionals in these services were asked to approach potentially eligible service users, provide a brief summary of the study and offer recruitment packs. Researchers also attended relevant community groups to present the study and distribute recruitment packs. Publicity materials were also made available through libraries, supermarkets, post offices and general practices.

### ***Word of mouth and snowball sampling***

To enhance recruitment, we used word of mouth and snowball sampling techniques and employed the assistance of bilingual community champions. This approach specifically focused on increasing engagement with ethnically diverse groups.

### ***Local media***

Recruitment was supported by a public relations campaign targeting local newspapers, magazines, radio and community events. This was supported by the public relations team at the University of Bath at no cost to the study.

Each REACT trial site tracked recruitment methods to determine the most successful recruitment strategy, in particular for recruiting from diverse ethnic groups.

## **Study intervention**

### ***Intervention arm***

The intervention arm received a manualised 12-month exercise and behavioural maintenance programme designed for delivery in leisure/community centres by qualified exercise professionals. Sessions were organised as group activities, with up to 15 participants per group. A comprehensive guidance manual outlining the content and structure of the types of exercise to be delivered, methods for progression, safety considerations, methods for tailoring exercises and progression to individual capabilities and the behavioural maintenance sessions was distributed to the session leaders prior to their training event dedicated to the delivery of the REACT intervention. REACT session leaders were qualified to at least Register of Exercise Professionals (REPS) Level 3 (Exercise Referral Diploma or equivalent) and were experienced in delivering safe and effective exercise sessions to older adults. The 1-hour exercise sessions were delivered twice weekly for 12 weeks, reducing to once weekly for a further 40 weeks (64 sessions in total over 12 months) to groups of around 15 participants. Despite being delivered in a group setting, exercise programmes were personalised to each participant, based on their functional status and goals. Towards the end of each session, games-based activities of 15–20 minutes' duration were delivered at 'light to moderate' intensities (points 8 to 13 according to the Rate of Perceived Exertion scale).<sup>35</sup> This individualised approach to exercise prescription enabled each participant to progress at their own pace.

The exercise sessions were each followed by 20 minutes of refreshments and socialising to promote session attendance and contribute to participants' social well-being. After 9 weeks, the behavioural maintenance programme commenced as a 45-minute session delivered once per week (usually immediately following the exercise class). The maintenance sessions were designed to encourage a 'social club' atmosphere. They provided physical activity and health information and emphasised long-term maintenance of an active lifestyle, including the promotion of ongoing engagement in exercise classes, home-based exercise, neighbourhood walking and active travel. They incorporated behaviour change techniques (BCTs) derived from social cognitive theory (SCT), self-determination theory (SDT)<sup>22,23</sup> and the skills for maintenance model.<sup>24</sup> These techniques included building intrinsic motivation; making realistic plans for sustainable activity; pre-empting and overcoming barriers; maximising enjoyment, social interaction and group identity; engaging external social support; and using self-monitoring and self-regulatory techniques to support the maintenance of behaviour change. From week 25 of the intervention, the behaviour maintenance programme reduced to one meeting per month for the remainder of the programme (six further meetings in total). If participants missed two consecutive sessions, REACT session leaders were asked to contact them by telephone to problem solve ways for them to re-engage with the programme.

REACT was delivered in three progressive phases (adoption, transition and maintenance), and established BCTs were used to enhance motivation, to make realistic plans for sustainable activity, to pre-empt and

overcome barriers, to engage social support and to use self-monitoring and self-regulatory techniques to support the maintenance of behaviour change.<sup>36</sup>

### **Start-up (adoption: weeks 1–8)**

The purpose of this phase was to stimulate initial increases in physical activity and fitness, to reduce any anxieties or concerns about exercise, and to build confidence and a sense of attachment to the programme. Each participant received a 30-minute individualised, face-to-face introductory session with the session leader. This was used to personalise the programme for starting levels and progression. Two 60-minute group-based physical activity sessions per week, plus 15–20 minutes of social time, were then delivered by the session leader.

### **Build-up (transition: weeks 9–24)**

A 45-minute interactive behavioural maintenance session delivered by session leaders was added to one of the two weekly sessions. These sessions used evidence-based, person-centred behaviour change strategies to build intrinsic motivation and self-efficacy. Sessions were designed to maximise enjoyment, social interaction and group identity.<sup>37</sup> They incorporated BCTs to address theoretical determinants of behaviour change relating to SCT, SDT<sup>38,39</sup> and the skills for maintenance model:<sup>40,41</sup> for example, to build intrinsic motivation, to make realistic plans, to pre-empt and overcome barriers, to maximise enjoyment, social interaction and group identity, to engage external social support and to use self-monitoring and self-regulatory techniques to support maintenance of behaviour change. A key focus was on exploring and planning transition to more daily lifestyle activities. After week 12, the exercise session frequency reduced to once per week, with an expectation that participants find an hour per week to exercise at home or in the neighbourhood or to attend a local physical activity session.

### **Taking charge (maintenance: weeks 25–52)**

The maintenance stage focused further on home-based and neighbourhood-based activities, while continuing with a weekly centre-based physical activity session followed by a 20-minute social session. The 45-minute behavioural maintenance sessions reduced to once per month and focused on enacting participants' action plans for physical activity outside the REACT programme. We encouraged groups to self-organise and to consider doing activities together beyond the scope of the study. Participants were informed about other local opportunities for physical activity in their community. They were also introduced to the REACT ambassador programme during the maintenance stage. This was a novel element whereby participants could contribute to the group as a co-ordinator. The REACT ambassadors' main role was to sustain activities after the end of the intervention at 12 months by supporting ongoing group meetings and activities without adding to intervention costs, thereby supporting sustainability.

### **Control arm**

After completion of baseline assessments, participants allocated to the control arm were given information regarding healthy ageing. They were invited to three 60- to 90-minute group sessions over the 2 years of the study. These consisted of presentations and discussion groups on various aspects of healthy ageing, such as healthy eating, living with dementia and volunteering. There was no physical activity content.

## **Intervention delivery**

REACT session leaders were qualified to at least REPS Level 3 (Exercise Referral Diploma or equivalent) and were experienced in delivering safe and effective exercise sessions to older adults. Behavioural maintenance sessions were usually delivered by the same session leader but were occasionally delivered by other staff from the same organisation. The REACT co-applicants provided a 2-day intervention delivery training to session leaders, including detailed session plans to ensure consistency in and fidelity to programme delivery based on a written programme of materials and manuals.



## Intervention fidelity

We included a range of the strategies outlined by the National Institutes of Health Behaviour Change Consortium to reinforce intervention fidelity.<sup>42</sup> We (1) ensured 'design fidelity' by building our intervention around a clear logic model (see *Figure 1*); (2) recruited REACT session leaders with appropriate skills and experience; (3) developed an accessible, standardised intervention manual; (4) implemented the standardised REACT session leader training programme; (5) trained more REACT session leaders than needed to accommodate illness or withdrawal; (6) monitored delivery fidelity by recording consultation meetings for a sample of three or four sessions per intervention provider and by applying a fidelity checklist; and (7) checked for intervention 'receipt' and 'enactment' of appropriate levels of physical activity outside the REACT sessions by checking participant understanding of the correct performance of exercises and regularly reviewing progress in the behavioural maintenance sessions. In addition, fidelity was enhanced by incorporating a gradual transition to daily activity within the structure of the REACT intervention (i.e. withdrawal of one session per week after 12 weeks, reduction of behavioural maintenance sessions to once per month after the first 6 months, with targeted planning of ongoing daily lifestyle physical activities around each transition).

## Health economics

Full details of the REACT economic evaluation are given in *Chapter 5*, but, in summary, we used data collected during the trial to estimate the resource use and costs associated with the delivery of the intervention and the wider NHS, social care and participant-level resource use and costs over a 24-month follow-up. The primary economic outcome measure is the quality-adjusted life-year (QALY) gain, derived from participant-reported EuroQol-5 Dimensions, five-level version (EQ-5D-5L), data.

Given the long-term nature of the potential benefits from the REACT intervention, we conducted decision-analytic modelling to assess the longer-term (lifelong) consequences of the intervention compared with the control, including consequences in terms of health-care and social care costs.

## Process evaluation

The REACT study follows the principles of the UK MRC guidance on process evaluation.<sup>43</sup> For full details of the process evaluation see *Chapter 4*; however, the aims and methods are briefly summarised below. The purpose of the process evaluation was to:

- evaluate the feasibility of implementation, including barriers to and facilitators of implementation, to inform future implementation and possible refinements of the intervention
- evaluate the quality and quantity of intervention delivery to inform conclusions about intervention effectiveness
- investigate the proposed mechanisms of change outlined in the REACT logic model (see *Figure 1*) and seek alternative explanations if this model is not supported
- understand the role of context to inform whether or not and how the findings can be generalised.

Methods used in the process evaluation included:

1. a mixed-methods assessment of intervention fidelity (quality of intervention delivery) using a checklist applied to audio-recordings of the REACT social education sessions
2. a qualitative analysis of semistructured interviews conducted with participants and staff providing the intervention
3. quantitative testing of hypotheses derived from the logic model, using study data on demographics and outcomes, as well as a set of questionnaires to measure processes of behaviour change proposed by the REACT logic model.

## Sample size calculation

Power calculations for the primary outcome (SPPB score) at 24 months were based on the published definition of a clinically minimum meaningful change in SPPB score of 0.5 points,<sup>21,44</sup> an expected SD for change in SPPB scores from baseline to 2 years of 2.2,<sup>21</sup> a two-sided significance level of 0.05 and an expected cumulative loss to follow-up of 12.5% per year. To provide 90% power, this required a sample size of 384 participants per arm (768 in total).

To detect a change of 0.5 points with a SD of 2.0, assuming that loss to follow-up accumulates at 12.5% per year throughout REACT's 2-year follow-up period, the required sample size was 384 participants per arm for 85% power using two-sided 5% significance. The REACT study, therefore, looked to recruit a total sample of 768 participants. This sample size also provides 90% power to detect a difference in moderate-intensity physical activity of 50 minutes per week [standard deviation (SD) 185 minutes per week] with 5% significance. A 2 : 1 randomisation process was applied in the internal pilot phase. A 1 : 1 randomisation process without rebalancing was applied in the main trial, resulting in an allocation ratio of 1.11 : 1. On the assumption that the effect size, the dropout rate and the significance level that we were interested in remained unchanged, our power reduced to 84.9% (from 85% power using two-sided 5% significance if the sample was rebalanced).

## Interim analyses and stopping guidelines

The TSC, with advice from the Data Monitoring and Ethics Committee (DMEC), assessed the feasibility of the trial during the internal pilot phase, taking into account findings on the acceptability of trial procedures, intervention adherence and recruitment rates. After 6 months, recruitment data were reviewed by the TSC and, as outlined in the study protocol,<sup>1</sup> when recruitment rates were found to be less than predicted,<sup>29</sup> the research team took actions to increase them (increasing the number of people approached and adapting recruitment procedures). The TSC was happy with the impact on recruitment rates and recommended that we proceed to the main trial. Retention rates (proportion of people providing follow-up data) were also checked at 6 months. Receipt of strong negative feedback from the majority of either participants or intervention providers about the intervention or trial methods would have been considered a stopping criterion. No such negative feedback was received.

## Randomisation

### Randomisation type

When 30 eligible participants had been recruited (enough to form one study group), they were randomised to one of the two arms (intervention arm or control arm) using a secure, centralised web-based randomisation website built by Peninsula Clinical Trials Unit. Randomisation was carried out using a minimisation algorithm<sup>45,46</sup> to balance arms in terms of study site (Bath/Bristol, Birmingham, Devon), age group (65–74 years or  $\geq 75$  years), sex and baseline functional ability (SPPB score of 4–7 or 8–9). This algorithm was built by the Peninsula Clinical Trials Unit and uses the method proposed by Taves<sup>45</sup> and extended by Pocock and Simon.<sup>46</sup> It maintains a stochastic element by computing probabilities proportional to the existing imbalance at the point of randomisation.

Participants were allocated to a specific exercise group within 1 or 2 days of randomisation. This was carried out via a telephone call from a researcher and followed up with a confirmation letter. Session leaders were then provided with contact details for their group members and contacted them directly to arrange a one-to-one meeting prior to the sessions starting.

During the pilot phase, the randomisation ratio was 2 : 1 (intervention to control) to enable feasibility testing of intervention engagement and retention as early as possible. The main trial randomisation



ratio was 1 : 1. The arms were not re-balanced following the pilot. Couples presenting together at the screening, with both people eligible and willing to be involved in the study, were randomised together to reduce contamination between study arms.

### ***Randomisation: allocation concealment mechanism***

To carry out randomisation, an authorised member of the research team accessed the randomisation website using unique username and password log-in details. The website required entry of patient's initials, date of birth and stratification variables [study site, age group (65–74 years or  $\geq 75$  years), sex and baseline functional ability (SPPB score of 4–7 or 8–9)]. The randomisation website generated a unique study ID number for the participant when they were randomised.

Each REACT group consisted of 30 participants. In the event that more than 30 individuals (15 intervention, 15 controls) were recruited within a study site, a small waiting list was maintained. If an individual dropped out of the study (intervention or control arm) during the first 2 weeks, they were replaced by a randomised member of the waiting list, although the original participant was still followed up and included in the ITT analysis. This ensured that the group nature of the intervention was maintained. If a new group was subsequently started at the site, then the members of the waiting list were randomised into the new group.

### ***Randomisation: implementation***

Confirmation that randomisation had been carried out was communicated in a blinded fashion to local site staff and key members of the research team via e-mails automatically generated by the randomisation website.

The clinical trials unit sent the study ID numbers of intervention and control participants to the RA at the local site. The RA, who was not involved in collecting primary outcome data, telephoned participants to inform them of their allocation and sent them a confirmation letter using the contact details collected at the baseline assessment.

Letters to participants in the intervention arm advised participants of the date and venue of their REACT sessions.

## **Blinding**

It is not possible to blind study participants to treatment allocation in behavioural intervention studies, and this is not a problem in pragmatic trial designs, which aim to estimate the benefits of the intervention over and above usual or standardised care. However, we took steps to ensure that data collectors, statisticians and the research team remained blinded to allocation, excluding one RA at each site. The chief investigator was unblinded when needed to allow assessment of SAEs. At follow-up data collection visits, participants were asked not to reveal which arm they were in.

Data were coded so that those undertaking the statistical and economic analyses were blinded. Given the study design, we did not anticipate a substantial risk of contamination (i.e. exposure of the control arm participants to the REACT intervention). However, as part of their briefing on entry to the study, participants in the intervention arm were asked not to share or discuss the content of the intervention sessions with any control arm participants whom they may be in touch with for the duration of the study. The possibility of contamination of control arm participants by REACT session leaders was minimised by giving them clear instructions not to provide intervention materials or information to any participants not assigned to the intervention arm. Attrition bias was minimised by having robust trial procedures to prevent data loss and also by analysing the data using an intention-to-treat (ITT) approach.

Case report forms did not contain any data that would enable the identification of the participant, so staff entering data remained blinded.

## Unblinding

The DMEC undertook regular safety data reviews after recruitment began, and all SAEs were reported to it. The DMEC was responsible for identifying any need for unblinding and periodically reviewed unblinded safety data to determine patterns and trends of events, or to identify safety issues, which would not be apparent on an individual case basis.

## Statistical analyses

The REACT analysis populations were as follows:

- Population 0. The primary analysis was performed on an ITT basis including all participants consented and randomised.
- Population 1. Analysis of all consented participants who completed  $\geq 50\%$  of the intervention (minimum required dose analysis).
- Population 2. Analysis of all consented participants who completed  $\geq 75\%$  of the intervention (high adherers analysis).

There is no published consensus on what constitutes a minimum dose of attendance for behaviour change interventions. Our choice of 50% as a minimum/sufficient dose and 75% as a high dose was based on (1) attendance rates for other successful behaviour change interventions promoting physical activity range from 36%<sup>47</sup> to 61%,<sup>21</sup> and (2) expert opinion within the TMG.

## Primary and secondary outcome data analysis

Analyses were prespecified in the published protocol. Primary outcome analysis was undertaken blinded to group allocation. Primary comparative analysis was undertaken using the principles of ITT with due emphasis placed on CIs. Using appropriate descriptive statistics, we assessed any imbalance between the trial arms at baseline and described the characteristics of participants. The comparison of primary interest was the difference between the intervention and the control arms in SPPB score at the 2-year follow-up (24 months after randomisation). Factors in the model comprised baseline age, sex and study site (Bath/Bristol, Devon or Birmingham). Baseline SPPB scores were added as a covariate in the model and, in addition, we adjusted the estimates for clustering by exercise group within the intervention arm, with control arm participants entered as individual groups following the method cited by Flight *et al.*<sup>48</sup> Analyses were conducted in Stata<sup>®</sup> SE version 15.0 (StataCorp LP, College Station, TX, USA). The randomisation stratification factors (site, age and sex) were entered into the analysis model as categorical variables.

Secondary outcome analyses were undertaken using the same approach as for the primary analysis (excluding the sensitivity analyses), using the baseline, 6-month, 12-month and 24-month follow-up data and linear mixed-regression models. Health-related quality of life, as assessed by EuroQol-5 Dimensions (EQ-5D), will be reported elsewhere as part of the health economic evaluation.

As an exploratory analysis, the effect of several predefined factors was further investigated and is presented. These included the stratification variables [age categories (65–74 years and  $\geq 75$  years), sex and study site (Bath/Bristol, Devon, Birmingham)], as well as comorbidity levels at baseline (none or one chronic medical conditions vs. two or more chronic medical conditions), socioeconomic subgroups (using education, home ownership and quintiles of area deprivation), history of falls (recorded fall or not during 6 months prior to baseline) and the uptake of any co-interventions during the 24-month study period. Health-related quality of life, as assessed by EQ-5D, is reported in *Chapter 10*.

Intervention costs were estimated by identifying key resources (programme co-ordination, session leader time and expenses, venue hire, equipment, consumables, and programme-specific training) and assigning values to the resources used (see *Chapter 5*). The data were collected by the REACT session leaders and trial manager.

### Adjustment in the models

The intention was to adjust the models for the four stratification factors and the intervention group (clustered owing to the group nature of the intervention) and the intervention arm as follows: age group and sex were included as fixed effects, site was included as a random effect, baseline functional ability was included as a covariate, the intervention group was included as a random effect and the intervention arm (intervention or control) was included as a fixed factor. The intervention group effect is the effect due to the intervention being delivered to groups of individuals. In the control arm, the clusters are specified by the individual.<sup>48</sup>

### Subgroup analyses

To examine the association between dose and response, we conducted subgroup analyses comparing participants attending  $\geq 50\%$  and  $\geq 75\%$  of sessions with (all) controls (the three defined analysis populations above). Mirroring the primary analysis, variables representing age group, sex, baseline SPPB and exercise group (within the intervention arm) were entered into the model.

### Missing data

We expected a relatively low level of dropout and missing data. Therefore, the primary analysis was undertaken without imputation of missing values. However, a comparison of baseline covariates between completers and non-completers was undertaken to assess the impact of dropouts on the results.

### Multiple testing

The primary analysis (SPPB at 24 months in the ITT population) is a single test, and, therefore, adjustment for multiple testing was not considered to be appropriate.

The analyses of the secondary outcome measures and the primary outcome at other time points were considered as exploratory analyses, and, therefore, there was no adjustment to account for multiple testing. The results of these analyses were interpreted in the light of the potential for an increased risk of making a type 1 error.

### Safety data

The safety population was all trial participants (see *Chapter 3, Main outcome results*, for the analysis population for the primary outcome).

All SAEs were reported regardless of relatedness to the trial. Non-SAEs (regardless of relatedness) were not reported. All reportable events were followed up until resolution, where possible, or until the end of the data collection period.

No formal comparisons between groups were made because the numbers of events were relatively small. Full details of SAEs are presented in *Chapter 8*.

## Amendments to the statistical analysis plan

Following the provision of baseline scores, but prior to the primary analysis, concerns were raised by members of the TMG about the inclusion of the adjustment for clustering in the primary analysis. Given the group nature of the intervention, there were concerns that adjusting for the clustering (given the lack of clustering in the control arm) would reduce the apparent significance of the intervention effect. Two teleconference meetings were held between the TSC and the TMG: one on 8 January 2020 and one on 27 January 2020. This addendum reflects additions to the analysis following those discussions.

This addendum is an addition to the main REACT statistical analysis plan v5. The statistical analysis plan outlines the primary analysis, including (but not limited to) adjusting the primary analysis for the group nature of the delivery of the intervention in the intervention arm.

In the submitted proposal<sup>30</sup> and the protocol paper,<sup>1</sup> no consideration was given to the clustering in the intervention arm. Therefore, by adjusting for the clustering in the intervention arm in the analysis, there would be a potential reduction in the power of the study. It was also agreed that the clustering in the intervention arm is a potentially important aspect to the structure of the data and, therefore, not including this may result in an increased risk of making a type 1 error. The impact of this, potentially, would be to marginally increase the size of the standard errors associated with the treatment effect estimates. Flight *et al.*<sup>48</sup> provided three case studies suggesting that for moderate levels of intracluster correlation these are unlikely to affect the conclusions. However, this may be the case where a result is borderline significant or where there is a strong clustering effect.<sup>48</sup>

## Sensitivity analysis for the primary outcome data analysis

The structure of the study was such that participants in the intervention arm met in groups whereas participants in the control arm had minimal interaction. Therefore, in the primary analysis, the primary outcome (SPPB at 24 months) was modelled using the approach recommended by Flight *et al.*<sup>48</sup> for a partially clustered design; an analysis of covariance (ANCOVA) was undertaken, adjusting for the site as a fixed effect (Birmingham, Bath/Bristol, Devon) and group as a random effect (the groups within the intervention arm that met for exercise). Following this primary analysis, an additional sensitivity analysis was undertaken without the inclusion of the treatment groupings in the model. This analysis is in line with the analysis proposed in the protocol paper.<sup>1</sup>

The sensitivity analysis was further enhanced by an investigation of the intracluster correlation at each data collection period once other factors (primarily site) had been accounted for. Investigating how the amount of clustering (represented by the intracluster correlation) changes from baseline to 6 months (peak intervention intensity) to 12 and then 24 months may help to indicate the extent to which the intervention itself was associated with any clustering effects (e.g. owing to treatment groups forming close activity-supporting bonds or having strong leadership that may cause the outcomes to cluster by intervention delivery group). This could aid in our interpretation of any discrepancies between effectiveness estimates produced by the clustered and the un-clustered analysis.

## Statistical significance

For the hypothesis test, two-tailed  $p$ -values of  $\leq 0.05$  are considered statistically significant. Given that the additional sensitivity analyses outcomes here are exploratory in nature, there was no adjustment for multiple testing.

## Model assumptions

For all methods outlined, underlying assumptions were checked using standard methods (e.g. residual plots). If assumptions were violated, alternative methods of analysis were sought. In particular, the underlying assumption of the linear impact of baseline covariates was assessed to ensure that baseline covariates were categorised and fitted as factors in the model. The assumption of the consistent clustering effect between the intervention and the control arm was evaluated.

## Outcomes

### Assessments

A full list of measures and time points at which outcome data were collected is presented in *Appendix 1, Table 20*. The person conducting the assessments checked for completion of questionnaires before participants left the assessment premises and made every effort to ensure that missed or spoiled questions were addressed.

### Primary hypothesis

The primary hypothesis was that, compared with the participants in the control arm, participants allocated to receive the REACT programme would have significantly reduced mobility-related limitations, as indicated by SPPB score, at 24 months.

### Secondary hypotheses

The secondary hypotheses were that, compared with the control arm, participants allocated to receive the REACT programme would have significantly increased their levels of moderate-intensity physical activity, health-related quality of life, cognitive function, ability to perform the ADL, and mental and social well-being at 24 months.

### Primary outcome

The primary outcome was the SPPB score at 24 months.

### Secondary outcomes

- Change in minutes of moderate-intensity physical activity, as measured by wrist-worn accelerometers.
- Sedentary time and breaks in sedentary time per day, as measured by wrist-worn accelerometers.
- Self-reported physical activity [Physical Activity Scale for the Elderly (PASE questionnaire)].<sup>49</sup>
- Self-reported adherence to government guidelines on muscle-strengthening activity, assessed by the Muscle-Strengthening Exercise Questionnaire.
- Hand-grip strength of the dominant hand, as measured by a digital dynamometer.
- Ageing Well Profile Social scale (six items) score.<sup>50</sup>
- ADL [EQ-5D, Short Form questionnaire-36 items (SF-36) and Mobility Assessment Tool-Short Form (MAT-SF)] score.<sup>51-53</sup>
- Simple processing speed, episodic memory, fluid intelligence, working memory, visual attention and complex processing speed as determined using the UK Biobank Healthy Minds Questionnaire.<sup>54,55</sup>
- Cognitive impairment, assessed by the Montreal Cognitive Assessment (MoCA).<sup>56</sup>
- The incremental cost-effectiveness of the REACT intervention (EQ-5D, SF-36, Health and Social Service Usage).<sup>51,52</sup>
- Pain [assessed using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)].<sup>57</sup>
- Sleep quality, assessed by the Sleep Condition Indicator.<sup>58</sup>
- Medical history.
- Falls Inventory.<sup>59</sup>

### Health economics outcomes

- Intervention costs. Each component of resource use was estimated at aggregated and at site level. A mean cost per participant of delivering the intervention in REACT was estimated.
- Health-care, social care and other resource use. Resource use is presented for baseline and the 6-, 12- and 24-month follow-up periods and for the adoption, transition and maintenance phases.
- Effectiveness/health-related quality-of-life outcomes. SPPB score at 24 months and QALY data were derived from trial data on EQ-5D-5L, using the UK algorithms/tariffs. Derived health state utility values were used to estimate QALYs through application of standard area-under-the-curve methods using baseline and the 6-, 12- and 24-month assessments.
- Discounting. Costs and health outcomes were discounted using a rate of 3.5%, as recommended by the National Institute for Health and Care Excellence (NICE).
- For full details of the REACT health economic evaluation, see *Chapter 5*.

### Process evaluation outcomes

- Evaluation of the feasibility of implementation, including barriers to and facilitators of it, to inform future implementation and possible refinements of the intervention.
- Evaluation of the quality and quantity of intervention delivery to inform conclusions about intervention effectiveness.
- Investigation of the proposed mechanisms of change outlined in the REACT logic model and identification of alternative explanations if this model is not supported.
- Evaluation of the role of context to inform whether or not and how the findings can be generalised.

For full details of the REACT process evaluation, see *Chapter 4*.

### Changes to the study outcomes during the course of the study

There were no changes to the study outcomes during the course of the study.



## Chapter 3 Results: randomised controlled trial

Parts of this chapter have been adapted from Stathi *et al.*<sup>60</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

Parts of this chapter have been adapted from Withall *et al.*<sup>29</sup> 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. The text includes minor additions and formatting changes to the original text.

In this chapter, we report our findings for the REACT RCT, including recruitment and adherence data, participant baseline characteristics and quantitative study outcomes.

In line with the Consolidated Standards of Reporting Trials (CONSORT) guidelines for reporting parallel-group randomised trials,<sup>61</sup> we outline the flow of participants through the study and report baseline and outcome data for each trial group (*Figure 2*; see *Table 3*).

### Participant flow

Between June 2016 and September 2017, 3116 people were telephone screened (of whom 1077 were not eligible and 825 declined to participate further) and 1214 attended for baseline screening. Of these, 804 were found to be eligible and 777 were randomised (intervention,  $n = 410$ ; control,  $n = 367$ ). Throughout the study, 39 couples or pairs of close friends who were both eligible were randomised together to minimise contamination between study arms. The number of participants included in the primary analysis at 24 months was 628 (80.8%): 334 (81.5%) in the intervention arm and 294 (80.1%) in the control arm. *Figure 2* shows the flow of participants through the study.

### Recruitment

The trial was successful in recruiting. In fact, slightly more participants were recruited than had been planned (target,  $n = 768$ ; actual,  $n = 777$ ). Between February 2016 and September 2017, we contacted 25,559 people (via 35 GP practices, community partners and a PR campaign). A total of 3116 people responded and were telephone screened by the local site RA, 1214 people were screened face to face at local community centres and 777 were randomised (intervention,  $n = 410$ ; control,  $n = 367$ ), slightly exceeding the recruitment goal of 768 participants.<sup>29</sup>

### Baseline data

Baseline characteristics were similar between the two study arms (*Table 2*). The mean age of the participants was 77.6 years (SD 6.8 years) and 66.2% were female. The majority of participants (95.1%) were Caucasian, 1.2% were Asian, 3.0% were African/Caribbean and 0.8% were of other/mixed ethnicity. The mean SPPB score was 7.37 (SD 1.56) and the mean body mass index (BMI) was 29.25 kg/m<sup>2</sup> (SD 5.71 kg/m<sup>2</sup>). Just over half of participants were educated beyond secondary school level ( $n = 417$ ; 53.7%) and the majority were overweight/obese ( $n = 588$ ; 76.5%). The study aimed to broadly represent the diversity of deprivation and ethnicity for individuals over 65 years of age within the UK population. Comparisons of the REACT cohort and the population aged over 65 years in England and Wales are shown in *Appendix 1, Table 21*.<sup>62-64</sup>



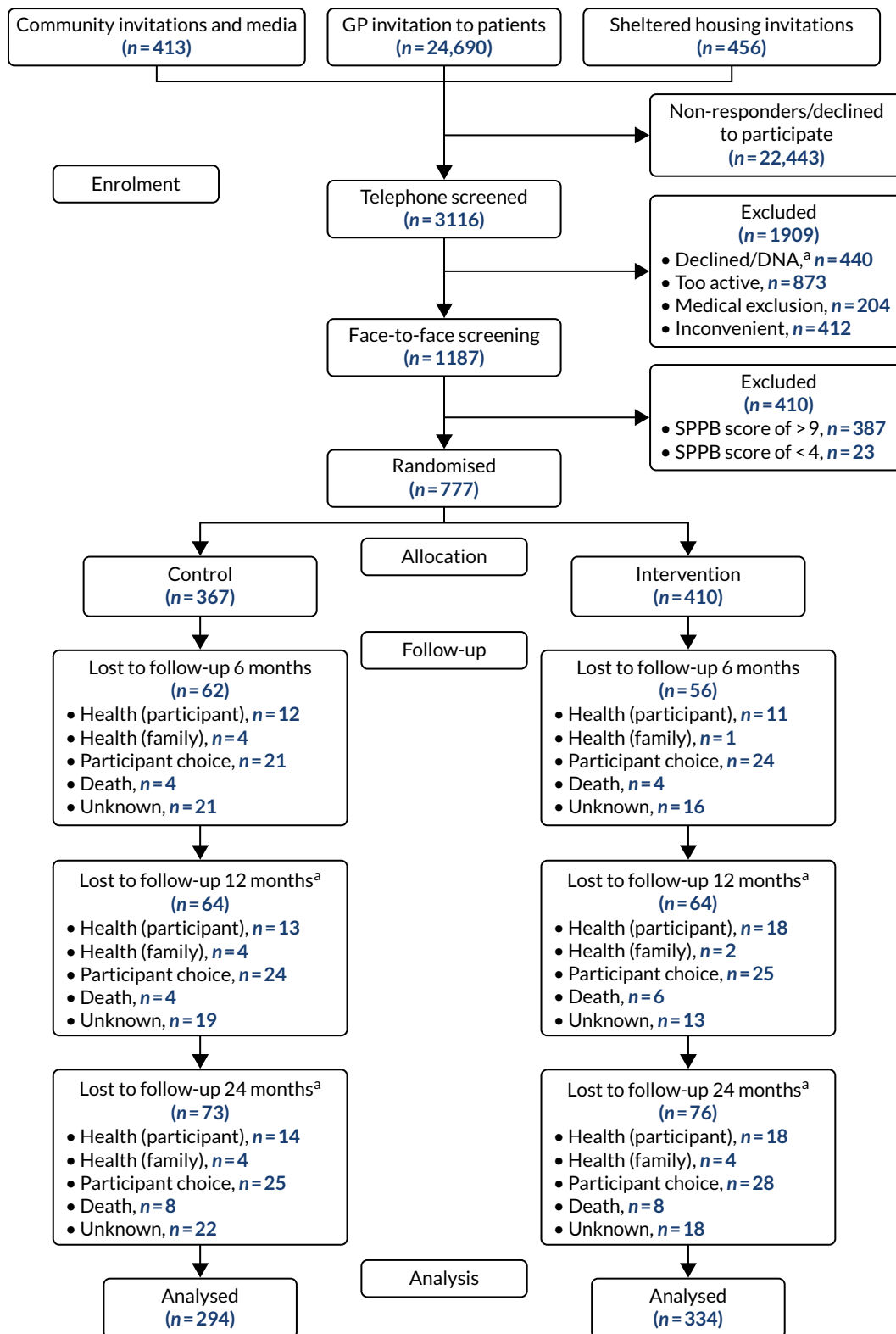


FIGURE 2 The CONSORT flow diagram of participants through the REACT study. a, Loss to follow-up data are cumulative.

TABLE 2 Baseline characteristics

Characteristic	Study arm	
	Control (N = 367)	Intervention (N = 410)
<b>Demographic characteristics</b>		
Age (years), mean (SD); n	77.3 (6.64); 367	77.8 (6.93); 410
<b>Sex, n (%)</b>		
Female (n = 514)	240 (65.6)	274 (66.7)
Male (n = 263)	126 (34.4)	137 (33.3)
<b>Ethnicity, n (%)</b>		
Caucasian/white (n = 739)	352 (96.17)	387 (94.16)
African/Caribbean (n = 23)	9 (2.46)	14 (3.41)
Asian (n = 9)	3 (0.32)	6 (1.46)
Other/mixed (n = 6)	2 (0.55)	4 (0.97)
BMI (kg/m <sup>2</sup> ), mean (SD); n	29.34 (5.51); 363	29.20 (5.67); 404
Cognitive impairment (MoCA), mean (SD); n	24.29 (3.62); 354	24.45 (3.70); 399
<b>Highest education level, n (%)</b>		
Less than secondary school (n = 64)	32 (8.74)	32 (7.79)
Completed secondary school (n = 295)	154 (42.08)	141 (34.31)
Some college/vocational training (n = 206)	89 (24.32)	117 (28.47)
College or university degree (n = 162)	72 (19.67)	90 (21.90)
Graduate degree, or higher (n = 49)	18 (4.92)	31 (7.54)
<b>IMD, n (%)</b>		
Quintile 1 (n = 86)	43 (11.75)	43 (10.46)
Quintile 2 (n = 157)	73 (19.95)	84 (20.44)
Quintile 3 (n = 159)	70 (19.13)	89 (21.65)
Quintile 4 (n = 156)	73 (19.95)	83 (20.19)
Quintile 5 (n = 219)	107 (29.23)	112 (27.25)
<b>Caring responsibilities, n (%)</b>		
Yes (n = 86)	37 (11.9)	49 (14.4)
No (n = 564)	273 (88.1)	291 (85.6)
<b>Marital status, n (%)</b>		
Married or living with partner (n = 334)	158 (50.5)	176 (51.8)
Widowed (n = 200)	90 (28.8)	110 (32.4)
Divorced/separated (n = 79)	48 (15.3)	31 (9.1)
Single and never married (n = 39)	17 (5.4)	22 (6.5)
Other (n = 1)		1 (0.3)
<b>Home ownership, n (%)</b>		
Own home (n = 553)	259 (83.0)	294 (86.5)
Renting/other (n = 99)	53 (17.0)	46 (13.5)

continued

TABLE 2 Baseline characteristics (continued)

Characteristic	Study arm	
	Control (N = 367)	Intervention (N = 410)
Number of chronic illnesses, n (%)		
None (n = 173)	90 (25.0)	83 (20.5)
One (n = 260)	129 (35.8)	131 (32.4)
Two or more (n = 331)	141 (39.2)	190 (47.0)
<b>Outcomes</b>		
SPPB total score, mean (SD); n	7.36 (1.54); 367	7.38 (1.58); 410
Accelerometry		
MVPA (minutes/day): time spent at > 100 milligravitational units in at least 10-minute bouts, mean (SD); n	5.80 (8.62); 330	5.94 (8.91); 374
Unbouted MVPA (minutes/day): all time spent at > 100 milligravitational units, mean (SD); n	58.82 (32.18); 330	55.10 (29.86); 374
Very low PA/sedentary time, excluding sleep (minutes/day), mean (SD); n	804 (91.66); 318	804 (91.52); 362
Breaks in sedentary time (n/day), mean (SD); n	43.23 (13.40); 328	43.53 (13.36); 375
Subjective PA (PASE), mean (SD); n	119.90 (57.61); 359	112.33 (58.13); 400
Muscle-strengthening exercise (MSEQ), mean (SD); n	3.18 (2.12); 338	2.90 (2.01); 388
Hand-grip strength (kg), mean (SD); n	24.92 (8.66); 361	24.68 (8.49); 404
Ageing Well Profile Social Wellbeing subscale, mean (SD); n	23.92 (7.30); 347	23.91 (6.74); 387
Sleep Condition Indicator, mean score (SD); n	21.95 (7.90); 333	22.53 (7.55); 342
Pain (WOMAC), mean score (SD); n	10.12 (3.77); 351	9.73 (3.94); 399
Loneliness, n/N (%)	135/361 (37.2)	135/403 (33.7)
SF-36, mean score (SD); n		
Physical component	30.01 (10.61); 392	29.70 (10.96); 353
Mental component	53.77 (8.66); 392	54.55 (8.33); 353
MAT-SF, mean score (SD); n	49.89 (8.88); 357	49.06 (9.75); 403
UK Biobank Healthy Minds Questionnaire, mean score (SD); n		
Simple processing speed	866.92 (277.42); 337	865.70 (282.35); 383
Fluid intelligence	3.60 (1.70); 332	3.75 (1.59); 377
Executive function	59,849.78 (31,733.12); 254	61,269.89 (38,594.83); 283
Working memory 1	4.29 (1.44); 335	4.37 (1.46); 382
Working memory 2	14.09 (6.43); 336	13.73 (6.14); 383
Episodic memory	5.94 (4.80); 333	6.08 (4.29); 377
Falls inventory		
Number of falls in last 6 months, mean (SD); n	0.72 (1.15); 359	0.69 (1.08); 401
Fall-related injury in last 6 months, n/N (%)	45/355 (12.5)	56/399 (14.2)
MSEQ, Muscle-Strengthening Exercise Questionnaire; PA, physical activity.		

In total, 11.1% of REACT participants fell within quintile 1 (most deprived) of the Index of Multiple Deprivation (IMD), compared with 14.3% of the general UK population aged over 65 years. In quintile 2, these figures were 20.2% and 17.6%, respectively. In terms of ethnicity, REACT under-recruited Asian participants (2.6% in the UK population aged over 65 years, 1.2% in the study) but over-recruited African/Caribbean participants (1.3% in the UK population aged over 65 years, 3.0% in the study).

The proportion of Caucasian/white participants was slightly lower than in the general population (95.1% vs. 95.5%, respectively) while the proportions of other/mixed ethnicities were very similar (0.8% vs. 0.7%, respectively). In terms of sex, 45.6% of the over-65 years population of England and Wales are male, compared with the 33.85% of REACT participants. However, compared with the over-65 years population of England and Wales, the REACT cohort was skewed towards the older age ranges, where the proportion of females increases.<sup>29</sup>

## Main outcome results

At the 24-month follow-up, the mean SPPB score (adjusted for baseline SPPB, age, sex, study site and exercise group) was significantly higher in the intervention arm (mean 8.08, SD 2.87) than in the control arm (mean 7.59, SD 2.61), with an adjusted mean difference of 0.49 (95% CI 0.06 to 0.92;  $p = 0.014$ ). Only one instance of unblinding was reported during the collection of data at 24 months.

The primary and secondary outcomes at 24 months are presented in *Table 3*, and all outcomes at 6 and 12 months are reported in *Appendix 1, Tables 25 and 26*. The SPPB score was significantly higher in the intervention arm than in the control arm at 6 months (adjusted mean difference 0.68 points, 95% CI 0.39 to 0.96 points;  $p < 0.001$ ) and 12 months (adjusted mean difference 0.77 points, 95% CI 0.40 to 1.14 points;  $p < 0.001$ ). Self-reported physical activity was significantly higher in the intervention arm than in the control arm at 6 months (adjusted mean difference in PASE score of 16.3 points, 95% CI 6.78 to 25.9 points;  $p = 0.001$ ), 12 months (adjusted mean difference 10.8 points, 95% CI 3.18 to 18.5 points;  $p = 0.006$ ) and 24 months (adjusted mean difference 10.7 points, 95% CI 2.62 to 18.8 points;  $p = 0.010$ ). Self-reported engagement in muscle-strengthening exercise showed a similar pattern, with highly significant differences ( $p < 0.001$ ) at all three follow-up time points.

TABLE 3 Primary and secondary outcomes at 24 months

	Study arm		Estimated mean difference (95% CI)	<i>p</i> -value <sup>a</sup>
	Control	Intervention		
<b>Primary outcome</b>				
SPPB total score, mean (SD); <i>n</i>	7.59 (2.61); 294	8.08 (2.87); 334	0.49 (0.06 to 0.92)	0.014
<b>Secondary outcomes</b>				
MVPA (minutes/day): time spent at > 100 milligravitational units in at least 10-minute bouts, mean (SD); <i>n</i>	4.50 (6.61); 250	5.15 (5.99); 290	0.65 (-0.48 to 1.78)	0.255
MVPA (minutes/day): all time spent at > 100 milligravitational units, mean (SD); <i>n</i>	48.76 (19.48); 250	51.22 (17.20); 290	2.46 (-0.52 to 5.44)	0.105

continued

TABLE 3 Primary and secondary outcomes at 24 months (continued)

	Study arm		Estimated mean difference (95% CI)	p-value <sup>a</sup>
	Control	Intervention		
Sedentary time, excluding sleep (minutes/day), mean (SD); n	798 (65.80); 249	804 (64.04); 287	6.43 (-4.81 to 17.67)	0.259
Breaks in sedentary time (n/day), mean (SD); n	42.33 (13.54); 248	40.76 (13.21); 287	-1.57 (-3.89 to 0.75)	0.184
Subjective PA (PASE), mean score (SD); n	113.17 (52.10); 301	123.90 (49.79); 328	10.73 (2.62 to 18.84)	0.010
Muscle-strengthening exercise (MSEQ), mean (SD); n	3.18 (1.88); 276	3.86 (2.30); 307	0.68 (0.33 to 1.02)	< 0.001
Hand-grip strength (kg), mean (SD); n	23.43 (4.08); 291	23.74 (3.86); 328	0.31 (-0.33 to 0.94)	0.343
Ageing Well Profile Social Well-being subscale, mean score (SD); n	24.68 (5.85); 295	24.88 (7.07); 306	0.20 (-0.84 to 1.24)	0.700
Sleep Condition Indicator, mean score (SD); n	21.97 (6.10); 285	22.50 (6.65); 311	0.53 (-0.49 to 1.54)	0.306
Pain (WOMAC), mean score (SD); n	10.20 (3.28); 290	9.63 (3.95); 324	-0.57 (-1.15 to 0.00)	0.052
Loneliness, n/N (%)	107/300 (35.7)	110/330 (33.3)	0.037 (-0.064 to 0.074) <sup>b</sup>	0.914
SF-36, mean score (SD); n				
Physical component	29.38 (9.39); 295	30.84 (10.04); 306	1.46 (-0.09 to 3.01)	0.065
Mental component	54.73 (7.64); 295	54.33 (9.18); 306	-0.40 (-1.78 to 0.98)	0.563
EQ-5D, mean (SD); n	0.67 (0.16); 302	0.69 (0.16); 330	0.02 (-0.01 to 0.04)	0.220
MAT-SF, mean (SD); n	47.96 (8.13); 289	49.99 (8.96); 319	2.03 (0.66 to 3.40)	0.004
UK Biobank Healthy Minds Questionnaire, mean score (SD); n				
Simple processing speed	811.28 (240.15); 264	801.67 (246.72); 286	-9.61 (-52.47 to 33.24)	0.657
Fluid intelligence	4.03 (1.41); 262	4.19 (1.61); 282	0.16 (-0.11 to 0.43)	0.234
Executive function	64,770.62 (38,677.48); 210	58,515.77 (35,648.79); 236	-6254.85 (-13,498.22 to 988.52)	0.090
Working memory 1	4.59 (1.29); 263	4.46 (1.22); 282	-0.13 (-0.35 to 0.06)	0.260
Working memory 2	14.27 (5.24); 264	14.62 (5.15); 285	0.36 (-0.56 to 1.28)	0.439
Episodic memory	5.84 (4.19); 263	5.36 (6.85); 286	-0.48 (-1.49 to 0.53)	0.347
Falls inventory				
Number of falls in last 6 months, mean (SD); n	0.73 (1.05); 300	0.70 (1.05); 330	-0.02 (-0.19 to 0.14)	0.772
Fall-related injury in last 6 months, n/N (%)	51/297 (17.2)	57/326 (17.5)	0.3 (-5.92 to 6.46) <sup>b</sup>	0.809

MSEQ, Muscle-Strengthening Exercise Questionnaire; PA, physical activity.

<sup>a</sup> Adjusted for site, exercise group (within the intervention arm), age group, sex and baseline SPPB.

<sup>b</sup> Adjusted estimate and 95% CI for the between-group percentage difference.

Accelerometer data indicated a significant difference favouring the intervention group at 12 months for total MVPA (adjusted mean difference 3.11 minutes per day, 95% CI 0.00 to 6.23 minutes;  $p = 0.05$ ) and MVPA accumulated in bouts of at least 10 minutes (1.24 minutes per day, 95% CI 0.22 to 2.26;  $p = 0.018$ ). This equates to a difference of 22 minutes per week of MVPA lasting  $\leq 10$  minutes. Significant differences favouring the intervention arm were also observed in the SF-36 physical component score (at 6 and 12 months), hand-grip strength (at 12 months) and the MAT-SF self-reported lower limb physical functioning scale (at 6, 12 and 24 months).

## Numbers analysed

All data were analysed based on the participants' originally assigned groups (ITT). The number of participants included in the analysis of each outcome measure is shown in the outcome tables (see *Table 3*; see *Appendix 1, Tables 25–26*).

## Losses and exclusions

Over the 2-year measurement period, we had a relatively low level of withdrawals from the study. Only 14.3% ( $n = 111$ ) of participants withdrew from the study: 59 from the intervention arm and 52 from the control arm. The reasons were personal choice (30.6%), health issues (25.6%), unknown reasons (18.9%), death (14.4%), family health issues (7.2%) and being excluded by the intervention delivery organisation (3.6%) (see *Appendix 1, Table 22*). Final 24-month data collection was completed in October 2019. Given that the number of missing data did not differ substantially between the intervention and the control arm, the primary analysis was undertaken without imputation of missing values.

Data for physical function (SPPB score) at 24 months (primary outcome) were available for 628 participants (80.8%): 334 (81.5%) participants in the intervention arm and 294 (80.1%) participants in the control arm. Compared with the predicted loss to follow-up of 12.5% per year (25% cumulative for two years), the actual loss to follow-up over the 2 years of the study was 19.2% (see *Appendix 1, Table 23*).

## Programme adherence

Of the 410 participants allocated to the intervention arm, 16.1% did not engage with any of the intervention sessions (non-starters), 19.0% attended  $< 50\%$  of the sessions offered, 20.2% attended 50–74% of the sessions offered and 44.6% attended  $\geq 75\%$  of the sessions offered. Among all participants allocated to the intervention arm (including the non-starters), the mean percentage of sessions attended was 56.8% (95% CI 53.6% to 60.1%). Among participants allocated to the intervention arm who engaged with the programme (starters only), this figure was 67.7% (95% CI 65.1% to 70.4%). An association between dose and response was observed (see *Appendix 1, Table 28*), with an adjusted mean SPPB difference of 0.64 (95% CI 0.23 to 1.05;  $p = 0.002$ ) for those attending  $\geq 50\%$  of intervention sessions and 0.81 (95% CI 0.38 to 1.23;  $p < 0.001$ ) for those attending  $\geq 75\%$  of intervention sessions.

## Sensitivity analysis

Sensitivity analyses, including imputation of missing values and not adjusting for clustering by exercise group, did not significantly change the above results (see *Appendix 1, Table 29*). The intraclass correlation coefficient for SPPB scores relating to clustering by exercise group within the intervention arm was 0.02 (95% CI 0.0085 to 0.129). Subgroup analyses for age, education levels and SES (key inequality populations) or other characteristics found no significant interactions with study arm, indicating that the intervention worked equally well for all of the pre-identified subgroups (see *Appendix 1, Table 29*).

## Health economic results

The full 12-month REACT programme, as delivered in the trial, was estimated to cost £9466 per group, an average of £622 per participant. For more details on health economic analysis, see *Chapter 5*.

## Non-adherence to the protocol

As determined in the REACT statistical analysis plan, the following protocol violations were considered:

1. Enrolment protocol violations were considered to occur if a member of the research team failed to appropriately apply the study's eligibility criteria, resulting in the enrolment of an inappropriate patient into the trial.
2. A randomisation protocol violation was defined as a technical or human error leading to the violation of the intended randomisation sequence or any attempts to subvert allocation concealment.
3. A study intervention protocol violation was defined as a delivery error (incorrect number of sessions delivered) in the study intervention attributable to members of the research team. The research team included members of the study co-ordinating centre, site investigators and research co-ordinators.
4. Data collection protocol violations encompassed errors in which the research team failed to comply with specific trial guidelines for data collection and/or outcome evaluation for avoidable reasons.

No type 3 or 4 violations occurred. One type 1 error occurred when a participant scored 3 on the SPPB (inclusion criterion: SPPB score of 4–9) but was mistakenly included in the trial and randomised to the control arm. Data for this participant were included in the analysis. Three type 2 errors occurred. Prior to a protocol change to allow close friends to be randomised together (to avoid contamination), two pairs of friends were randomised separately and allocated to different arms. After discussions with the TMG and the TSC, it was agreed to submit an ethics amendment for the protocol change and both pairs of friends were allowed to attend the intervention programme. There was also one erroneous allocation letter sent to a participant informing them that they were allocated to the control arm when, in fact, their allocation was to the intervention arm. In all cases, analysis was conducted on the basis of original allocation (ITT), and the flow diagram (see *Figure 2*) reflects this.

## Harms

In the 56 months of the study, 93 events were classified as SAEs: 59 from the intervention arm and 34 from the control arm (see *Appendix 1, Table 24*). SAE data were collected when inviting participants to attend assessment sessions, at assessment sessions (6, 12 and 24 months), when inviting control arm participants to attend one of their three social and education sessions and via session leaders reporting SAEs among intervention participants. The somewhat larger number of SAEs reported for intervention arm participants is likely to be a result of the higher frequency of contact with this group by session leaders, providing increased opportunity for reporting illness, injury and hospitalisation.

In the early phases of the study, all hospitalisations were reported as SAEs, including hospitalisation for planned procedures. On 9 January 2018, a substantial ethics amendment was approved to exclude planned hospitalisation from SAE reporting. From that date, any adverse event or adverse reaction was regarded as serious if it:

- resulted in death
- was life-threatening



- required non-elective hospitalisation, prolongation of existing hospitalisation or elective hospitalisation that may be related to taking part in the study
- resulted in persistent or significant disability or incapacity.

In line with the REACT protocol, SAEs were reported within 24 hours to the chief investigator, the Data Management and Ethics Committee chairperson for consultation on the relatedness to the study, and the sponsor of the trial. Ninety of these SAEs were classified as events unrelated to the study and, therefore, were not reported for further consultation with the TSC. One case was related to the study, and two cases considered to possibly have been related to the study were investigated and deemed not to be related to the study. These were reported to the appropriate regulatory body, the sponsor of the REACT trial and the DMEC, as per trial protocol.

## Discussion

In support of the primary hypothesis, older adults with mobility limitations who received the 12-month REACT programme experienced statistically and clinically significant improvements in lower limb physical function compared with control arm participants at the 24-month follow-up (12 months after the end of the intervention) and also at 6 and 12 months, indicating a sustained benefit over time. Higher intervention effects were associated with increased programme attendance.

The baseline SPPB scores were almost identical to the LIFE study population,<sup>21</sup> enabling comparison. The observed difference in SPPB score of 0.49 at 24 months was three times larger than the between-group difference reported in the LIFE trial. In the LIFE trial, this smaller difference in SPPB was sufficient to reduce the subsequent risk of major mobility disability (defined as the objectively assessed inability to walk 400 metres) by 18% and the risk of persistent mobility disability (defined as two consecutive major mobility disability assessments or assessment of major mobility disability followed by death) by 28%. The ability to walk a distance of 400 metres strongly relates to maintenance of independent living and reduced risk of mortality.<sup>65</sup> These meaningful impacts from lower levels of SPPB change in the LIFE study suggest that the minimum clinically meaningful difference in SPPB may be considerably lower than the difference of 0.50 used in calculating our sample size. Indeed, other evidence suggests that changes in SPPB score of  $\geq 0.28$  are meaningful in frail or pre-frail older adults (SPPB score of 4–9).<sup>66</sup>

At the completion of the intervention (12 months post baseline), significant differences were observed in SF-36 physical component score, MAT-SF, MVPA, self-reported physical activity and adherence to muscle-strengthening exercise, and hand-grip strength. These are consistent with the idea that the intervention increased engagement in muscle-strengthening, balance and endurance exercise that mediated the observed effects on physical functioning.

The strong association between session attendance and increased lower limb physical function (see *Appendix 1, Table 28*) also suggests that the intervention worked through engagement in the REACT programme. However, causality cannot be implied from this association and there may be other explanations; for example, older adults who became frailer owing to injuries or life events, or who did not feel they were benefiting from the intervention, may have been less likely to attend the intervention.

It is worth noting the methodological issue of how to choose cut-off points representing a 'sufficient' or minimum dose for behavioural interventions, and this study provides data that may be informative for future studies of group-based exercise interventions. Our data show that attending  $\geq 50\%$  of sessions was associated with a clinically meaningful effect on SPPB (0.64 points) and that a minimum of 75% attendance was associated with a considerably stronger effect (0.81 points).



The increase in MVPA (9 minutes per week bouts or 22 minutes per week unbouted) was small, although it is worth noting that any increase in MVPA has effects on health<sup>67</sup> and these increases must be set against very low levels of initial MVPA in this sample of frail or pre-frail older people (the baseline level of bouts MVPA was just 41 minutes per week). The 2.6-point change in SF-36 physical component score was small (a clinically meaningful difference being cited as around 4 points),<sup>68,69</sup> as was the change in hand-grip strength (0.8 kg compared with a clinically important difference of 5.0 kg).<sup>70</sup> This demonstrates that functional benefits of the REACT programme were specific to the focus of the exercise intervention programme (lower limb mobility) and did not generalise to upper body physical functioning.

At 24 months, of the secondary outcomes, only changes in self-reported physical activity, muscle-strengthening exercise and MAT-SF were sustained. Other differences found at 12 months were reduced (17 minutes per week of unbouted MVPA, 1.5 points in SF-36 physical component score) but fell below the level of significance. This may reflect deterioration of the effects on exercise behaviours over time, as well as a lack of statistical power to detect smaller differences.

To the best of our knowledge, the REACT trial is the first trial targeting physical function with a long-term (12-month) intervention and a 24-month follow-up in this population. It was a well-powered definitive study, being the largest of its kind conducted in the UK. It was a robustly designed and conducted study, with low attrition rates (19% at 24 months) and high adherence.

Although only 3% of those invited to take part were recruited, it should be noted that the REACT study invited everyone in our study area aged  $\geq 65$  years and then applied a two-stage screening process. Based on the screening data from the study, we estimate that over 80% of those invited were likely to be ineligible because their SPPB score was outside the target range or because they met other exclusion criteria. On this basis, the response rate among the eligible population was 17%. Although this still leaves some uncertainty around the generalisability of the findings, it is reassuring that the recruited sample was representative of the UK population aged over 65 years in terms of deprivation and ethnicity, except for an under-representation of South Asian older adults.<sup>29,65</sup>

The main limitation was that, as with other studies of behavioural interventions, blinding of participants to the study arm was not possible, introducing the possibility of social desirability bias in patient-reported measures. However, the primary outcome here consisted of a battery of physical tests assessed by independent observers, with the data collectors blinded to study arm allocation. Given that the secondary outcome analyses were exploratory, there was no adjustment for multiple testing, and the significance of the analyses needs to be interpreted accordingly. A further limitation is that it was not possible to explore variation in outcomes by ethnicity owing to insufficient numbers of ethnic minority participants.

Measurement of physical activity using accelerometers is held to be a gold standard for large-scale field trials and superior to self-report questionnaires. However, in REACT we observed that, despite improved physical function, Muscle-Strengthening Exercise Questionnaire (MSEQ) and self-reported PASE scores, there were no meaningful improvements in accelerometer-measured physical activity over the course of the study. The lack of agreement between objectively measured and self-reported measures could be explained by (1) wrist-mounted accelerometers lacking the sensitivity to detect muscle-strengthening and balance exercise and (2) the fact that the MSEQ is designed to capture muscle-strengthening activity and the PASE includes an item that captures muscle-strengthening exercises and weights it heavily in its scoring algorithm. Furthermore, a focus on MVPA using absolute cut-off points derived from healthier populations may not be appropriate for a pre-frail older adult population whose capacity for aerobic physical activity is likely to be lower than that of healthier, younger adults. Further work is, therefore, needed to understand how valid measures of physical activity and of engagement in muscle-strengthening exercise/activities can be derived from accelerometer data, particularly in older adults.

The subgroup analyses suggested consistency of intervention effects across different subgroups of the population, including both sexes and people of different ages, education levels and SES (as assessed by area deprivation). This indicates that, if rolled out, the intervention has potential to reduce health inequalities, if targeted towards underserved populations.

### *Implications for practice and/or future research*

Programmes such as REACT could help to sustain the health and independence of vulnerable older adults at risk of mobility limitations. The observed dose–response relationship supports the importance of group, multimodal physical activity at least once per week in the initial stage and then at least once per fortnight in the maintenance stage for achieving sustained clinically meaningful changes in physical function. This is a strong clinical and public health message to give to older adults in terms of defining the relatively low level of commitment required to maintain their lower limb physical function.

The REACT exercise intervention provides important evidence supporting the World Health Organization, US and UK physical activity recommendations for multimodal exercise in adults aged over 65 years.

Given that recruitment to this trial was challenging, further research is needed to identify a simple, sensitive and specific assessment process to identify older adults who are likely to benefit from this type of intervention (i.e. likely to have a SPPB score of 4–9). This will be useful both for future research and for implementation of the intervention in this population. Indeed, we hope to generate such a measure from further analyses of the REACT data set.

Future studies will need to focus specifically on people of black, Asian and other non-white ethnic backgrounds to examine the effectiveness of the REACT intervention in these populations, and to identify and address any barriers that might deter them from engaging with this programme.

Finally, research is needed to optimise the implementation of REACT at scale and further evaluate/extend its reach, effectiveness and cost-effectiveness. It may be possible and synergistic, for instance, to integrate the REACT intervention with existing mobility-related prevention and rehabilitation services.

In conclusion, among older adults at risk of mobility limitations, the REACT intervention prevented decline in physical function over a 24-month period. Contrary to the belief that older age comes with an inevitable decline in physical functioning, the REACT study shows that this decline can be slowed or even prevented with modest lifestyle changes.



## Chapter 4 Process evaluation

Parts of this chapter have been reproduced from Stathi *et al.*<sup>1</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

Parts of this chapter have been reproduced from Cross *et al.*<sup>71</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

In this chapter, we report our mixed-methods process evaluation of the REACT RCT, based on MRC guidance on process evaluation<sup>43</sup> and the REACT logic model (see *Figure 1*). This chapter reports the experiences of receiving and delivering the REACT intervention, aiming to identify modifications for future implementation; it explores:

- whether or not the components of the intervention were delivered as intended
- the mechanisms which influenced (1) engagement with the programme and (2) changes in lower limb physical functioning (SPPB).

### Aims

The purposes of the process evaluation in the REACT trial were to:

1. explore participants' and facilitators' experiences of the intervention to identify possible refinements of the intervention for future implementation
2. evaluate the quality and quantity of intervention delivery to inform conclusions about intervention effectiveness
3. investigate proposed mechanisms of change outlined in the REACT logic model and seek alternative explanations if this model is not supported
4. understand the role of context (e.g. setting, area deprivation, ethnicity and sex) to inform whether or not and how the findings can be generalised.

The dose of the intervention delivered and the associations of outcomes with dose are reported in *Chapter 3*. Subgroup analyses exploring differential impacts of the intervention in different population subgroups are also reported in *Chapter 3*, but are combined with the data presented in this chapter to help to inform refinements of the REACT logic model.

### The REACT logic model

The logic model for REACT is shown in *Figure 1*. It identifies:

- the REACT intervention components and how they were intended to be delivered to participants
- hypothesised mechanisms of action of the REACT intervention (the model incorporates a number of causal assumptions about the process by which the intervention was intended to effect change in health behaviours and outcomes)
- hypothesised contextual variables that might moderate or mediate mechanisms of change in motivations and behaviour
- hypothesised interactions between participation in the intervention, delivery quality, motivation, behaviour and outcomes.

## Study 1: intervention fidelity

Parts of this section have been reproduced from Cross *et al.*<sup>71</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

### Methods

#### Design

We conducted a mixed-methods assessment of intervention fidelity (quantity and quality of intervention delivery). This included collecting observational data on intervention exposure (dose) and participant characteristics, as described in *Chapter 3*. We also applied an intervention fidelity checklist to in vivo audio-recordings of a purposive sample of the REACT health behavioural maintenance programme (social education) sessions to generate a descriptive summary of delivery quality scores. Examples of good practice in the delivery of the health behavioural maintenance programme were also extracted from the audio data to inform recommendations for future programme delivery and training. Qualitative data (see *Study 2: qualitative evaluation of intervention processes*) added further context to these data and possible explanations for any deviations from the intended delivery processes.

We did not formally evaluate intervention fidelity with respect to delivery of the group exercise component of REACT. However, we did send observers to one exercise group per site to check that exercise delivery (1) included the intended range of exercises, (2) was individually tailored and (3) promoted progression of exercises. The observers reported that the exercise delivery was close or very close to what the exercise protocol (developed by Dr Ladlow, co-applicant and specialist in strength and conditioning) had specified in all cases.

#### Participants' sampling

Audio-recordings of the REACT health behaviour maintenance group sessions were purposively sampled to include a diverse sample of sessions based on (1) intended inclusion of key BCTs in the session plan, (2) intervention provider (organisations responsible for delivering the intervention groups) and (3) key transition points in the intervention (where the intervention changed in frequency or focus). Key transition points and BCTs in the selected sessions are shown in *Table 4*. Following the internal pilot stage of the study, we asked all providers to record the nine selected sessions for each exercise group that they delivered the REACT intervention to (a total of 54 sessions). Participants in the recorded sessions were REACT intervention arm participants and REACT session leaders.

### Measures

#### Fidelity scoring

We applied a 'fidelity checklist' to assess the delivery fidelity of the session audio-recordings. The 11-item checklist was developed by REACT intervention developers (CG and AS) and a postgraduate researcher (RC) to assess key components of the REACT logic model (see *Figure 1*) and to measure the extent to which REACT intervention facilitators delivered the intervention BCTs and delivery processes as intended (see *Appendix 2, Table 31*).

We used a 6-point Dreyfus scale<sup>72</sup> to measure the session leaders' adherence to the use of intended intervention BCTs and delivery processes, as well as the skill with which they were implemented. The response scale ranges from 0 (very poor/no delivery) to 5 (near-perfect/expert delivery) (see *Appendix 2, Table 32*).

TABLE 4 Behaviour change techniques used in each of the sampled sessions

Intervention BCT	Intervention weeks sampled								
	9 <sup>a</sup>	12	13 <sup>b</sup>	16	20	24 <sup>c</sup>	28	48 <sup>d</sup>	52 <sup>e</sup>
Person-centred delivery	X	X	X	X	X	X	X	X	X
Facilitating enjoyment	X	X	X	X	X	X	X	X	X
Monitoring progress	X	X	X	X	X	X	X	X	X
Self-monitoring	X	X	X	X	X				
Managing setbacks and problem-solving	X	X	X	X	X	X	X	X	X
Action-planning and goal-setting	X	X	X	X	X	X	X	X	X
Modelling	X	X							
Promoting autonomy	X	X	X	X	X	X	X	X	X
Supporting competence and self-efficacy	X	X	X	X	X	X	X	X	X
Supporting relatedness	X	X	X	X	X	X	X	X	X

a Health Behaviour Maintenance sessions start.

b Exercise sessions drop from twice to once per week.

c End of the adoption phase/start of maintenance phase.

d Social education sessions drop to once per month.

e Last REACT session.

Detailed scoring instructions can be found in *Report Supplementary Material 1*. As suggested by other studies,<sup>73,74</sup> the coding was anchored to the key heuristic that a score of 3 was considered to represent 'competent delivery', that is delivery that was considered to be sufficient to deliver the intended behaviour change processes.

In an effort to reduce subjectivity in scoring, two coders – an expert coder (CG) and a postgraduate researcher (RC) – independently coded a sample of 10 sessions. If discrepancies in scoring between coders exceeded more than 1 point on the Likert scale, the sessions were discussed to produce a consensus about how to apply the scoring system. The remaining sessions were coded by the postgraduate researcher (RC).

### Examples of good practice

When coding for intervention fidelity, the researchers (RC and CG) noted the time stamps in the recordings of examples of theorised and non-theorised intervention processes in practice, as well as examples of good and poor intervention delivery. This enabled extraction of examples of good or poor practice and of the intended delivery of specific intervention processes for future intervention training courses.

## Analysis

### Intervention dose and delivery quality

For analysis of the association of dose with the primary outcome, see *Chapter 3*.

### Intervention fidelity

Fidelity checklist scores were summarised by calculating either a mean or a maximum score for each item across all coded sessions. Items representing delivery processes or change techniques that were intended to be delivered in every session (e.g. person-centred delivery, managing setbacks and problem-solving) were summarised with a mean score, whereas items representing processes or techniques that were intended to be delivered in only some of the sessions (e.g. self-monitoring and modelling) were summarised with a maximum score (see *Appendix 2, Table 33*).

A mean item score was then calculated for (1) each BCT, (2) each exercise group and (3) overall delivery fidelity (the mean of all checklist item scores for all exercise groups).

### **Examples of good and suboptimal practice**

Examples of good and poor practice for each checklist item were transcribed and tabulated.

## **Results**

### **Characteristics of the recorded sessions**

From a sample of 54 requested audio-recordings, 25 (46%) were suitable for analysis. Missing data were attributable to equipment failure ( $n = 10$ ), session leaders failing to record the session ( $n = 12$ ), communication issues between researchers and session leaders (resulting in session leaders not recording sessions) ( $n = 5$ ) or sound problems that led to poor quality audio-files ( $n = 2$ ). *Table 5* reports the characteristics of the sessions analysed.

The audio-recordings revealed that the mean session length was 24.6 minutes (SD 16.7 minutes) rather than the intended 45 minutes.

### **Intervention delivery fidelity**

The overall delivery fidelity for the intervention (the mean of the scores for each checklist item taken across all exercise groups) was 2.5 points (SD 0.45 points), indicating that, overall, intervention delivery fidelity was suboptimal (see *Table 6*). The mean fidelity scores broken down by group ranged from 2.4 to 2.9. However, one group (group 4) had consistently lower delivery fidelity scores, with a mean overall fidelity score of 1.7. The fidelity scores for each checklist item (aggregated across all delivery groups) are outlined in *Table 6*.

### **Examples of good and suboptimal practice**

Some examples of both good practice and practices requiring improvement in delivery of each BCT are provided in *Appendix 2, Table 34*.

## **Discussion**

### **Summary of findings**

The delivery of the exercise component of REACT was considered to be good or very good across all sites. However, the overall score for intervention delivery fidelity (2.5, SD 0.45) indicated that, on average, across the sample the delivery of the behaviour maintenance sessions was suboptimal. There were numerous examples of good practice, but also numerous inconsistencies and examples of practice that contradicted the intended delivery model. There was considerable variation in delivery

TABLE 5 Characteristics of the sessions sampled

Intervention group	Intervention site	Intervention provider	Facilitator	Participants (n)	Sessions sampled (n)	Sessions suitable for analysis, n (%)
1	1	1	F1	13	9	6 (67)
2	1	2	F2	15	9	4 (44)
3	1	3	F1	16	9	7 (78)
4	2	4	F3	15	9	2 (22)
5	1	2	F4	14	9	5 (56)
6	3	5	F5	3	9	1 (11)

TABLE 6 Behaviour change technique and overall intervention fidelity scores

Group	Facilitator ID	Item number											Overall score
		1	2	3	4	5	6	7	8	9	10	11	
		Person-centred delivery	Facilitating enjoyment	Monitoring progress: acknowledge and review	Monitoring progress: eliciting benefits of PA	Self-monitoring	Managing setbacks and problem-solving	Action-planning and goal-setting	Modelling	Promoting autonomy	Supporting competency and self-efficacy	Supporting relatedness	
1	F1	3.3	3.5	2.2	1.8	3.0	1.9	1.8	3.0	3.0	2.2	2.2	2.5
2	F2	2.6	2.8	2.6	1.5	3.0	1.0	2.6	3.0	2.4	1.9	2.9	2.4
3	F1	3.4	3.5	3.3	2.0	4.0	2.6	2.5	3.0	3.0	2.7	2.4	2.9
4	F3	2.5	2.5	2.0	2.0	1.0	1.0	0	2.0	2.5	2.0	1.0	1.7
5	F4	3.2	3.2	2.6	2.2	3.0	1.9	3.0	3.0	3.2	2.9	2.5	2.8
6	F5	4.0	3.0	3.5	2.5	3.0	3.0	3.0	0	4.0	2.5	2.0	2.8
Mean item score		3.2	3.1	2.7	2.0	2.8	1.9	2.2	2.3	3.0	2.4	2.2	2.5
Standard deviation		0.55	0.39	0.59	0.34	0.98	0.81	1.14	1.21	0.57	0.39	0.65	0.46
PA, physical activity. Orange shading represents scope for improvement. Light-purple shading represents low fidelity. Aqua shading represents adequate.													



fidelity between intervention components (BCTs and delivery processes), between session leaders and between intervention groups. Key areas identified for future improvement were monitoring progress/eliciting benefits of physical activity (mean score of 2.0), action-planning and goal-setting (mean score of 2.2), modelling (mean score of 2.3), supporting competence and self-efficacy (mean score of 2.2), supporting relatedness (mean score of 2.3) and managing setbacks and problem-solving (mean score of 1.9).

### Relation to other literature and possible explanations

The current study adds to an emerging body of work on intervention fidelity.<sup>73-76</sup> It is consistent with this evidence in finding that the quality of delivery of complex interventions varies considerably between session leaders and from group to group. This illustrates the importance of ensuring consistency of delivery in group-based interventions, as poor facilitation in one group or centre could undermine effectiveness for multiple participants (15 per group in the case of REACT). High-quality training and quality assurance processes may, therefore, be crucial to ensuring that the effectiveness of the intervention is maintained in transitioning from the context of a research study to wider-scale community-based implementation. This might, for example, involve rating of delivery fidelity for each trainee post training (by independent observation or self-rating), performance monitoring or other methods for identifying further training needs.

The fact that the REACT intervention generated clinically meaningful changes in SPPB at both 12 and 24 months (see *Chapter 3*) suggests that, despite scope for improvements in delivery, the intervention still worked. This apparent discrepancy may reflect one or more of several underlying phenomena:

- (a) The approach to rating delivery fidelity may have been overly conservative (i.e. the delivery was actually 'adequate' and so merited a mean score of at least 3.0). However, the authors who reviewed the session delivery are confident that there is clear scope for improvement in delivery of the behaviour maintenance sessions.
- (b) The intervention worked by processes beyond those that were specified in the REACT logic model. However, although some additional processes were identified in the qualitative process evaluation (see *Study 2: qualitative evaluation of intervention processes*), the logic model was broadly supported.
- (c) It may be the case that participants who were more ready to engage and had more intrinsic capacity to overcome the shortcomings of the programme (i.e. to succeed despite suboptimal delivery) did particularly well, lifting the group mean effect to the level of significance. The qualitative data suggest that this may be a possibility and that more individual tailoring could help to engage and benefit a larger proportion of the participants.
- (d) The groups of participants may have had sufficient mutual resources within the group to self-generate some of the intended intervention processes, such as forming a strong sense of relatedness or social identity around the idea of doing exercise together, mutually supporting motivation and helping each other to problem-solve.
- (e) It may be that the shortcomings in the delivery of the health behaviour maintenance sessions have more effect on longer-term outcomes (i.e. they affect maintenance more than short-term changes in physical functioning). This might help to explain the decline in intervention effectiveness from 12 to 24 months (see *Appendix 1, Tables 26-28*). Given that some of the lower-scoring checklist items related to self-regulatory processes, such as managing setbacks, problem-solving, reinforcing benefits and planning future actions, this idea seems plausible. This implies a need to focus more on implementing these and other maintenance-focused behaviour change processes in future implementations of REACT.
- (f) The health behaviour maintenance sessions may not be an important part of the REACT intervention, compared with the structured exercise classes. However, this seems unlikely as the evidence base for the long-term effects on physical functioning of structured exercise programmes delivered without behavioural support is poor.

In relation to item (e) above, a systematic review and meta-regression of physical interventions in older adults suggested that some self-regulation techniques may not be acceptable to older adults.<sup>76</sup> This may be because older people are less likely to be concerned with attaining a particular level of physical activity and more concerned with the enjoyment and social connectedness that can be derived.<sup>77-80</sup> Hence, the poor delivery of self-regulation techniques may to some extent reflect resistance to such techniques from the participants, which the session leaders responded to by downplaying these elements of the intervention. Participant 'pushback' has been cited as a factor in lower delivery fidelity for physical activity promotion in at least one other behavioural intervention.<sup>74</sup>

Scores were low not only for self-regulation, but also for the social processes of supporting relatedness and modelling. Hence, important elements of the intervention's underlying theories (SDT and SCT) were not proactively delivered by session leaders. Despite this, it may be the case that participants gained significant motivation from social interactions which developed spontaneously as a result of the group setting [as suggested in item (d) above]. This process is confirmed as being present in several groups by the qualitative process evaluation (see *Study 2: qualitative evaluation of intervention processes*). Encouraging the development of positive intra-group dynamics is also suggested as a key intervention process in other literature<sup>81-83</sup> and a recent framework of processes for the delivery of group-based intervention.<sup>84</sup>

The discrepancy between planned session length (45 minutes) and mean session length determined from the audio-recordings (24.6 minutes, SD 16.74 minutes) is also an issue of concern, suggesting that workload pressure on the provider staff may have prevented delivery of the sessions as intended. A further possible explanation for the variations between session leaders in fidelity scores may be differences in interest or 'buy-in' between session leaders.

### Strengths and limitations

Assessing intervention fidelity using coding of audio-recorded intervention delivery sessions is considered a gold-standard method.<sup>42</sup> Although time-consuming and labour intensive, this method allowed direct observation of intervention delivery and an assessment that was specifically tailored to the REACT intervention and its associated logic model. The fidelity scores along with the examples of good practice extracted from the session recordings provide direct feedback to the intervention designers about the ways in which the intervention and its training course can be improved in the future. Furthermore, scoring was based on a validated response scale designed for coding the acquisition of skills and reliability.<sup>72</sup> This was enhanced by using independent coders for the first 10 sessions to calibrate the coding and minimise subjective bias, as well as sampling of recordings across a diverse range of intervention components, intervention sites, session leaders and intervention providers.

However, several limitations need to be acknowledged. The data came from a relatively small sample of participants (around 90 of the 410 intervention arm participants). There is a strong potential for sampling bias, given that we were able to score fidelity for only 25 of our intended 54 sessions. This may have led to overestimation of intervention fidelity, assuming that recordings were more likely to be missing at sites at which performance was low. Furthermore, the rating approach used was unavoidably subjective, so there is no definitive way to ensure that a score of 3 represents 'adequate' delivery. Despite this, the raters were confident that there was clear scope for improvement in the delivery of health behaviour maintenance programme.

### Implications for practice/future research directions

The variations observed have implications for intervention design, training and implementation, as well as for the specific refinement of the REACT intervention:

- Both participants and session leaders should be involved in the refinement/adaptation of the REACT health behaviour maintenance programme, the REACT training course, and strategies for translation of theoretical constructs and BCTs into deliverable and acceptable intervention components.

- Refinement of the intervention should focus on improving the delivery of intervention techniques and processes identified as having lower fidelity in this study (i.e. monitoring progress/eliciting benefits of physical activity, action-planning and goal-setting, modelling, supporting competence and self-efficacy, supporting relatedness and managing setbacks/problem-solving).
- Refinement of the intervention should include ideas on promoting intragroup support for the delivery of intended intervention processes.
- The examples of good and suboptimal practice may be helpful in constructing future REACT training materials.
- For implementation, future training of REACT session leaders should invite the session leaders who have been identified (using the checklist developed here or an adapted version) as delivering with competent or high delivery fidelity.
- For implementation, it may be helpful to consider what organisation-level governance or performance monitoring systems might support high-quality delivery of the REACT intervention. This might include incentives for, or monitoring of, intervention sessions to ensure that sufficient time is given to the social education sessions.

However, there is debate concerning whether perfect fidelity is feasible or even desirable.<sup>85</sup> Strict adherence to protocol does not always account for the individual needs of the participants and the context in which the delivery takes place. The intended intervention components and delivery processes need to be considered with reflection on the need for individual tailoring, the receptiveness of participants and session leaders to the intended techniques, and the feasibility of delivering those components in the implementation setting, which may be resource poor in terms of time and facilitator training.

Future research might seek to explore why some BCTs and intervention delivery processes were not delivered as well as others, particularly self-regulation techniques and techniques to build social identity and/or relatedness.

### **Conclusions**

Delivery of the health behaviour maintenance part of the REACT intervention was suboptimal, with considerable scope for improvement in the delivery of both self-regulation processes and social/relatedness-building processes. However, this may have been mitigated by mutual support and self-delivery of some of the intended processes within the groups during the exercise part of the REACT intervention. Two encouraging conclusions that can be drawn are that (1) the REACT intervention can work even with lower than intended delivery quality and (2) there is scope for improvement of delivery quality, which may lead to increased effectiveness and better maintenance of effects.

This study also highlighted the importance of assessing fidelity in the evaluation of complex behavioural interventions and the value that this adds in (1) helping to explain research outcomes, (2) understanding the theoretical underpinnings of health behaviour change interventions and (3) identifying ways to refine interventions and their training courses for future implementation (and further evaluation).

## **Study 2: qualitative evaluation of intervention processes**

### **Design**

We conducted a qualitative study to evaluate the effectiveness of intervention processes and how the REACT programme was received by participants and service providers. This integrated data from:

1. 51 semistructured, longitudinal individual interviews at 6, 12 and 24 months with a purposive sample of participants in the intervention arm (designed to sample diversity in terms of SPPB scores, session attendance, ethnicity, age, sex, service provider and site location)
2. interviews with a sample of session leaders after intervention completion.

For sampling purposes, SPPB scores of 8 and 9 were classified as pre-frail (high SPPB score) and scores of 4–7 were classified as frail (low SPPB score). Attendance was classified as high if participants attended  $\geq 50\%$  of the intervention sessions and low if they attended  $< 50\%$  of the intervention sessions.

## Research questions

- Research question 1: why did older adults engage with the REACT intervention and what were the perceived benefits of this?
- Research question 2: what were the factors (barriers and enablers) associated with adherence to REACT intervention sessions and to daily physical activity outside the REACT sessions?
- Research question 3: what was the apparent mechanism of physical activity behaviour change for older adults participating in REACT, as observed through participants', session leaders' and organisation providers' experiences of receiving/delivering the intervention, and how did this compare to the mechanisms proposed in the REACT logic model?
- Research question 4: was the intervention delivered as planned?
- Research question 5: what changes to the programme were recommended for future delivery of REACT?

## Methods

### Sampling: participants

Using stratified purposive sampling strategy, we selected a diverse sample of participants with a wide range of experience. This provided rich data to explore factors associated with adherence to REACT, the mechanisms of change proposed by the logic model and factors that may be associated with variation in intervention outcomes.<sup>86</sup> Participants were stratified and assigned to one of four groups creating a  $2 \times 2$  matrix of SPPB status and 3-month programme attendance per intervention group.

### Sampling: session leaders

All session leaders and organisation providers involved in the delivery of the REACT programme were invited to participate in focus groups after the completion of the 12-month intervention for the groups that they were leading.

### Data collection: participant interviews

A semistructured topic guide was developed and piloted with the REACT service user advisory group prior to use. The topic guide was divided into four sections, for 6-, 12- and 24-month interviews:

1. factors associated with REACT intervention effectiveness, attendance and adherence to a daily physical activity
2. participant experiences of barriers to and enablers of participation in the REACT intervention and daily physical activity
3. impact of baseline factors, such as physical activity levels, physical function (SPPB), physical activity motivations and past physical activity behaviours, on attendance of the REACT intervention and/or adherence to daily physical activity
4. observed mechanisms of behaviour change for older adults participating in REACT compared with the mechanisms proposed in the REACT logic model.

### Data collection: session leader interviews

A semistructured topic guide was developed and divided into four sections:

1. reasons for and personal benefits from participating in REACT
2. evaluation of the training programme, the structure and content of exercise and behavioural maintenance sessions and suggestions for future delivery
3. impact of REACT on physical function and well-being of participants
4. strengths and weaknesses of the programme.

**Data collection: procedures and informed consent**

Individual interviews were conducted with participants, and a combination of individual interviews and focus groups were conducted with session leaders and providers. REACT participants were recruited following purposive sampling to achieve diversity in ethnicity, sex, variation in SPPB score and adherence at 3 and 6 months (for 6 and 12 month interviews, respectively). Interviews with participants were also conducted at 24 months to examine reasons for maintenance or non-maintenance of physical activity and physical function, with purposive sampling based on the 24-month SPPB scores.

At 6 months, 12 semistructured face-to-face interviews were conducted in participants' homes and five were conducted in community centres. At 12 months, 10 interviews were conducted at home and five were conducted at community centres. At 24 months, 19 interviews were conducted at participants' homes. Twelve session leaders agreed to participate in interviews and group discussions (Bath/Bristol,  $n = 4$ ; Birmingham,  $n = 4$ ; Devon,  $n = 4$ ), while three had left their organisations and were uncontactable. One research team member who had delivered social and education sessions at three groups and two members of staff from one provider organisation were also interviewed. Interviews were recorded using an Olympus VN-741PC digital recorder (Olympus Vietnam Ltd, Long Thành, Vietnam). The average interview lasted approximately 48 minutes (range 22–89 minutes). Informed consent for interview was obtained from participants when they were recruited to the study. Verbal consent was also obtained and recorded at the beginning of each interview. Participants were reminded that their interview would be confidential and that they could withdraw from the interview at any time without any repercussions. No participants declined to participate at that stage.

**Data analysis: participant interviews**

The MRC process evaluation framework was used to explore and understand how the REACT programme was implemented, the causal mechanisms involved and the contextual factors at play that link intervention theory to intervention outcomes.<sup>43</sup> Qualitative data from 6-, 12- and 24-month interviews were used to explore (1) participant interactions with the REACT intervention (analysed using framework analysis<sup>87</sup>) and (2) the accuracy and validity of the REACT logic model. The analysis aimed to capture individual narratives/within-person processes of change, as well as to draw out common themes. Emergent themes were compared and contrasted with the theorised processes of change specified in the logic model. Processes of engagement with the intervention were also explored. Using data collected at 12 and 24 months, factors influencing the maintenance of physical activity/exercise were assessed and linked to participant responses at 6 months. This allowed a qualitative description of participants' experiences, potential pathways and barriers to adoption, and maintenance of active lifestyle to be evaluated at three time points. Techniques to enhance the trustworthiness of the analysis included cross-tabulation, negative case analysis and hypothesis testing.<sup>88</sup>

**Data analysis: session leader interviews and focus groups**

Data from both the individual interviews and the focus group audio-recordings were transcribed verbatim. Data analyses were conducted by members of the research team. Themes were predetermined based on the process evaluation research protocol guide (<https://www.fundingawards.nihr.ac.uk/award/13/164/51>; accessed 17 October 2022).

**Data trustworthiness**

To assure rigour and reduce bias, we employed the following strategies:

- Credibility – via prolonged engagement with the participants. Repeated encounters gave the researchers who were collecting data the context needed to understand the participant experiences and how these changed over time.
- Credibility – via researcher triangulation. Two researchers (Dr Rosina Cross and Professor Afroditi Stathi, chief investigator) conducted coding on the first three transcripts, discussing codes and reaching a consensus before developing the coding framework. The coding framework was discussed with the

process evaluation research team (including Professor Colin J Greaves and Dr Janet Withall) to ensure that there was a consensus in the way in which it had been applied. This framework was then applied to the remaining transcripts by Dr Cross, with regular discussion with Professor Stathi.

- Reflexivity – via regular discussions within the research team, including members with diverse scientific backgrounds and experiences of conducting qualitative research. These diverse voices allowed a better interpretation of data and refinement of coding framework.
- Transferability – via mapping the findings of the qualitative study onto the logic model of the REACT intervention and highlighting similarities and differences among groups with high and low levels of (1) SPPB and (2) adherence to the intervention. This process enables other researchers to transfer the findings of this study to similar contexts and/or settings.
- Auditability – via keeping records of the raw data, field notes, transcripts, analytical decisions made and the way data were interpreted, which created a clear audit trail.

## Results

### Six-, 12- and 24-month semistructured interview findings

#### Participant characteristics

At the 6-month interviews, the study sample comprised 12 women (71%) and five men (29%). The baseline age of the participants ranged from 68 to 88 years. The SPPB score was relatively balanced across the sample (47% frail, 53% pre-frail), as was attendance (47% low attendance, 53% high attendance).

At the 12-month interviews, as it was not possible to contact four women and one man from the 6-month cohort, three additional participants were recruited and the sample comprised 11 women (73%) and four men (27%). The age of the participants ranged from 68 to 89 years. At the 12-month interviews, most participants were classified as pre-frail (73%). Programme attendance was relatively balanced across the sample (40% low attendance, 60% high attendance).

At the 24-month interviews, the study sample comprised 13 women (68%) and six men (32%). The age of the sample ranged from 68 to 89 years. At the 24-month interviews, the majority of the participants were classified as being pre-frail (79%) and most (63%) were classified as having high attendance. *Table 7* summarises participant characteristics and reports baseline and 24-month SPPB scores and programme attendance.

#### Participant characteristics: session leaders

Following completion of the study, all ( $n = 15$ ) session leaders and organisation providers were approached and asked to participate in either an interview or a focus group regarding their experiences of REACT. Twelve agreed to participate (Bath/Bristol,  $n = 4$ ; Birmingham,  $n = 4$ ; Devon,  $n = 4$ ) but three had left their organisations and were uncontactable. In addition, one research team member who had delivered some of the health behaviour maintenance sessions with three groups (Devon) and two members of staff from one provider organisation (Bath/Bristol) were also interviewed. The total number of interviewees was 15 (male,  $n = 7$ ; female,  $n = 8$ ).

#### Thematic tree at 6, 12 and 24 months

The four predetermined central themes were supported by participant descriptions of their experiences of the REACT study and participation in daily physical activity. This section presents the four central themes, higher-order themes (HOTs) and subthemes (*Figures 3 and 4*) and the comparison of responses among participants classified as either pre-frail or frail and participants with low and high attendance, and presents these findings in the context of the REACT logic model.

#### Reasons to engage in REACT

Where relevant, and in an effort to further understand themes generated by participant response, we triangulate these data with relevant data provided by the session leaders.



TABLE 7 Participant characteristics

Pseudonym	Group	Sex	Age (years) (baseline)	SPPB score			Attendance	Interview		
				Baseline: frail	Baseline: pre-frail	24 months		6 months	12 months	24 months
Dorothy	1	Female	68	7		12	72% <sup>a</sup>	X	X	X
Cordelia	1	Female	88	4		5	27%	X		
Etta	1	Female	70		8	10	98% <sup>a</sup>	X	X	X
Anandi	1	Female	76		9	9	49%	X	X	X
Darsha	1	Female	77		9	8	81% <sup>a</sup>		X	X
Iris	2	Female	89		9	8	79% <sup>a</sup>	X	X	X
Mary	2	Female	76	4		4	35%	X		
Cecil	2	Male	69	4		12	65%	X		
Frederick	2	Male	87		8	9	90% <sup>a</sup>	X	X	X
Geraldine	3	Female	78		8	10	61%	X	X	X
Arman	3	Male	71		8	9	95% <sup>a</sup>	X	X	X
Valerie	3	Female	86		8	9	61%		X	X
Rita	3	Female	76		8	10	75% <sup>a</sup>		X	X
Beatrice	4	Female	80		9	9	96% <sup>a</sup>	X	X	
Alvita	4	Female	76	5		8	91% <sup>a</sup>	X	X	
Arthur	4	Male	76		8	9	65%	X	X	
Roger	4	Male	84	5		2	87% <sup>a</sup>	X	X	
Ann	5	Female	74	7		6	78% <sup>a</sup>	X	X	
Evelyn	5	Female	69	7		5	35%	X		
Flora	5	Female	79		8	5	30%	X		
John	4	Male	83	6		1	56%			X
Sam	2	Male	88	5		2	51%			X
Betty	4	Female	88	7		7	98% <sup>a</sup>			X
Timothy	6	Male	70		9	7	5%			X
Angelina	6	Female	73		9	9	65%			X
Mark	7	Male	73		9	9	63%			X
Jane	2	Female	84		8	10	35%			X
Rachel	4	Female	83		6	11	82% <sup>a</sup>			X
Eleanor	6	Female	68		8	12	30%			X

a Attendance of sessions is  $\geq 70\%$  of planned sessions ( $n = 64$ ).

The central theme, 'reasons to engage in REACT', comprised six HOTs: seeking better health, seeking social connectedness, appeal of research, appeal of physical activity, being influenced by others and having no reason not to take part in the 6-month interviews. Four HOTs persisted at the 12-month interviews: seeking better health, seeking social connectedness, appeal of research and appeal of REACT physical activity. This central-order theme was made up of 15 subthemes at the 6-month interviews and seven subthemes at the 12-month interviews. All 12-month subthemes remained from the 6-month interviews, with no new subthemes identified. A total of 19 subthemes remained at the 24-month interviews, with one new subtheme identified.

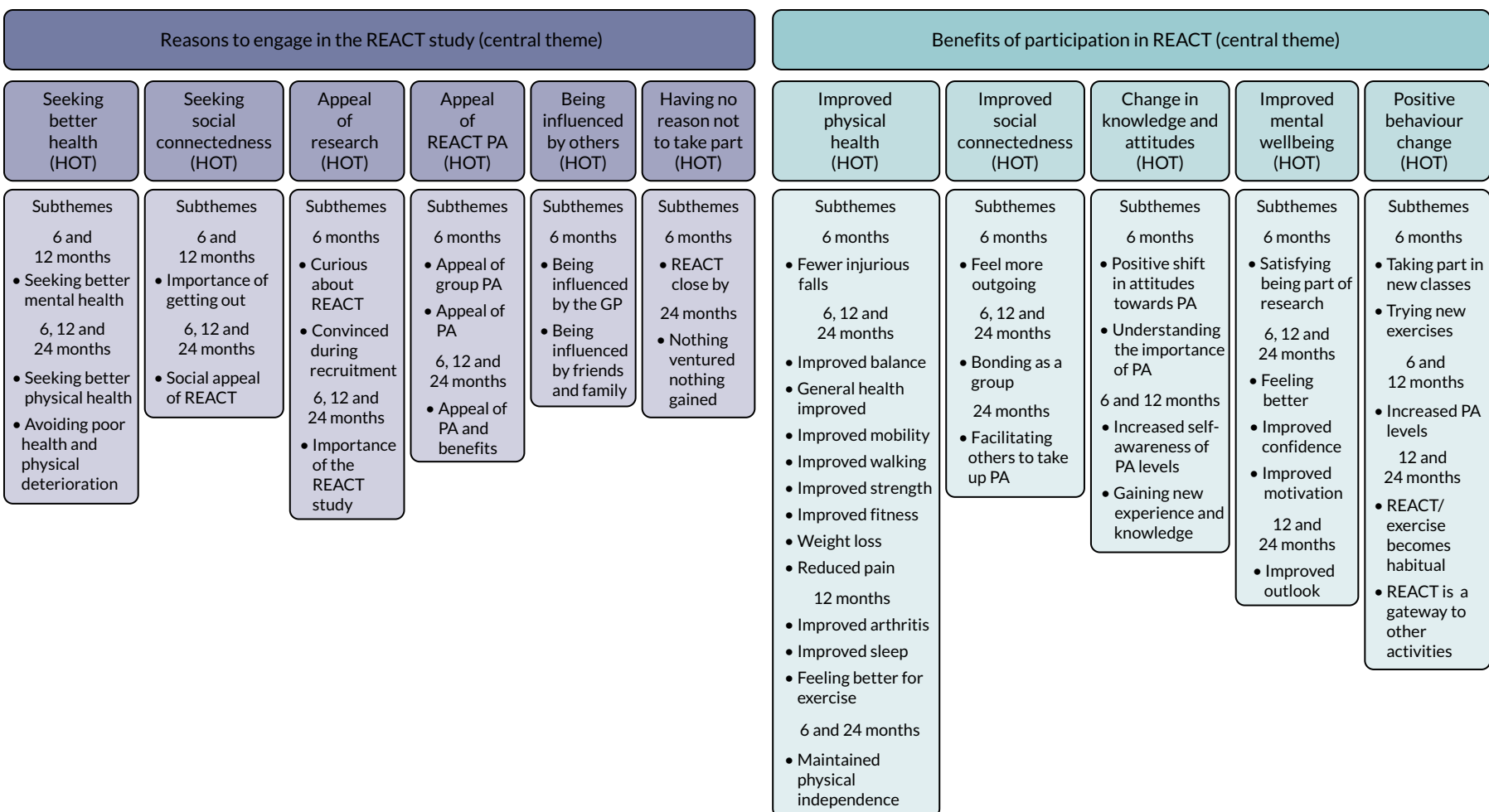


FIGURE 3 Thematic trees identified from participant experiences: reasons to engage in the REACT study and benefits of participation.



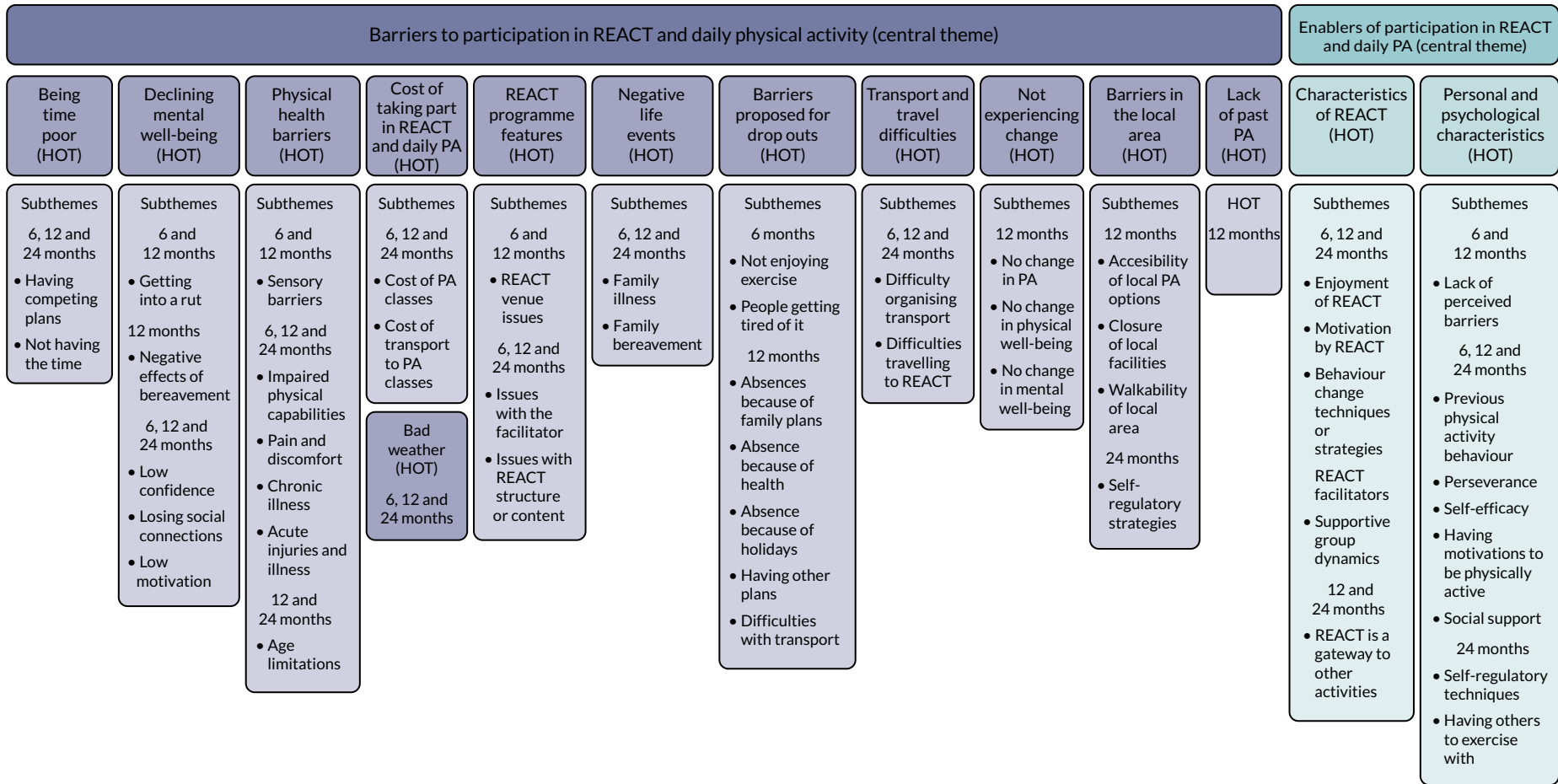


FIGURE 4 Thematic trees identified from participant experiences: barriers to and enablers of participation.

### **Reasons to engage in REACT for people classified as frail and pre-frail**

Reasons for engagement in REACT were very similar between people classified as frail and people classified as pre-frail. At the 6-month interviews, health was the predominant reason to engage in REACT for those classified as frail:

*And then I thought about it and I thought 'well give it a try', anything that will help me to walk better.  
Cordelia, 6 months, SPPB score of 4, attendance 72%*

By contrast, people classified as pre-frail based on their SPPB scores were more concerned with the social aspect of engaging in REACT and the appeal of taking part in research:

*I thought it was interesting that they were studying older people to find out if they do have enough exercise 'cos I don't think they do ... I've done these sort of things before [participate in research] ...  
Iris, 6 months, SPPB score of 9, attendance 79%*

At the 12-month interviews, participants classified as frail or pre-frail were concerned with losing the physical benefits that they experienced during the REACT programme:

*I just wanted to carry on with what we'd been doing, because like you know I could see the benefits.  
Beatrice, 12 months, SPPB score of 9, attendance 96%*

### **Reasons to engage in REACT for people with low and high attendance**

At the 6-month interviews, the key difference between people with low and people with high attendance was that those with high attendance reported that increasing their activity levels was a key reason to participate in REACT:

*... just hoping to get a bit of exercise out it ... yes, I think I've always wanted to be more active really ...  
Beatrice, 12 months, SPPB score of 9, attendance 96%*

This was less commonly reported by those with low attendance, who more often reported improving health and avoiding age-related decline as key reasons for participation. At the 12-month interviews, those with high attendance had multiple reasons to attend, whereas those with low attendance tended to report fewer reasons to continue attending, citing social connectedness as one of the most important reasons.

### **Findings in the context of the REACT logic model**

Participants reported concerns about their poor physical health affecting their independence and social connectedness. Consequently, they sought to participate in exercise that they perceived to be potentially beneficial. The group-based delivery was particularly appealing. Social interactions are proposed by the REACT logic model as a means of maximising enjoyment. At the 6-month interviews, participants confirmed the importance of social interaction and reported that it could be beneficial to their well-being and a key reason to engage in REACT. This was further confirmed at the 12-month interviews, at which participants highlighted the supportive role of this enhanced social connectedness in forming social support networks. This was reported as a motivator of continued participation and confirmed the logic model, which anticipated that changes in relatedness would mediate short-term REACT attendance. The appeal of physical activity reported at 6 months was strengthened with the perceived benefits of participation in REACT reported at 12 months, which motivated participants to maintain these benefits in the long term. This supports the logic model, which highlights that perceived benefits from participation in the REACT sessions would influence attendance of REACT sessions and levels of daily physical activity (see *Figure 2*). One finding, which was not identified in the logic model, was the appeal of participating in research. Initial curiosity in participation translated to appreciation of the value of the research beyond individual gain and a sense of obligation to continue to participate in REACT. These findings in the context of the logic model are shown in a modified REACT logic model (see *Appendix 2, Figure 28*).

**Benefits of participation in REACT**

Under the central theme 'benefits to participation in REACT', participants reported five main benefits: improved physical health, improved social connectedness, changes in knowledge and attitudes, better mental well-being and positive behaviour change. These benefits (HOTs) were reported at both the 6-month and the 12-month interviews. They were made up of 23 subthemes at the 6-month interviews, with 19 subthemes identified at the 12-month interviews: four were new subthemes and 15 remained from the 6-month interviews. A total of 16 subthemes remained at the 24-month interviews, with the emergence of one new subtheme (see *Figure 3*).

At the 6-month interviews, those classified as pre-frail reported experiencing more physical health benefits, whereas those classified as frail seemed to experience more benefits related to mental well-being and social connectedness. Similarly, at the 12-month interviews, participants classified as pre-frail reported more physical health improvements than participants classified as frail:

*Yes it was . . . very interesting, sometimes challenging uh, it was interesting to see what your body could do and how you improved, you know we all improved balance wise.*

*Beatrice, 12 months, SPPB score of 9, attendance 96%*

At the 24-month interviews, participants identified four main categories of physical activity participation:

1. walking (for leisure or active travel)
2. everyday activity (housework, gardening)
3. structured exercise (group or gym based)
4. independent exercises (prescribed by a health-care professional, learned through the REACT programme).

Although not described as physical activity, social activities reported by participants included meeting friends, caring for family and friends, being church members and being members of community groups.

At 24 months, physical activity maintenance was related to specific benefits, including increased mobility and ability to carry out ADLs. Experiencing physical gains from physical activity promoted feelings of competence, which further facilitated physical activity maintenance:

*My health has got better, I'm more motivated and I can get around a lot better. I don't use my stick now whereas I always had my stick. I just got better and better [during REACT], now I walk all round the cycle track.*

*Geraldine, 24 months, SPPB score of 10, attendance 64%*

Four participants reported increased physical activity levels post REACT. Three attributed improved motivation to be active directly to REACT. They described multiple benefits gained physically, developing new friendships and engaging in new activities (e.g. flower arranging and bowling). Two individuals wanted to capitalise on the benefits gained:

*I think because I enjoyed REACT so much, and because before I used to do a tiny bit of exercise, but I enjoyed it so much and I just thought it's a waste if I just leave it.*

*Dorothy, 24 months, SPPB score of 12, attendance 72%*

People with both high and low SPPB scores reported experiencing improved self-efficacy for physical activity, improved social connectedness and increased physical activity levels:

*I think what we did at REACT programme did give me bit of more confidence in walking.*

*Darsha, 12 months, SPPB score of 9, attendance 81%*

Being supported to reach achievable goals enhanced participants' confidence. One participant was spurred on to do more vigorous physical activity (running):

*I was really getting better. So, I better do something with my good health . . . They [REACT] have opened my mind on what I can do. I cannot think of not going to exercise or not go for a walk or anything. I just have to go and do it and I am happy to do it.*

*Angelina, 24 months, SPPB score of 9, attendance 65%*

REACT prompted participants to self-monitor and seek opportunities for exercise outside the REACT programme, in which support from session leaders was important to maintain self-efficacy:

*It's giving me an incentive to be more aware of what I need to do . . . I don't exercise but I do stretching exercises . . . I never did it before . . . this basically has to do with REACT encouraging me. In a way you know, otherwise I wouldn't be doing it because I never did.*

*Eleanor, 24 months, SPPB score of 12, attendance 30%*

*I said [to session leader] that I hadn't really enjoyed it so I wanted to do something much more active, so he said, come to the gym, and then I found out I was in the gym on my own and I didn't like that. So he then said well do aqua aerobics on a Friday.*

*Jane, 24 months, SPPB score of 10, attendance 35%*

### **Benefits of participation in REACT findings and the REACT logic model**

Participants experienced physical, social, emotional and behavioural benefits that supported the mediating role of perceived benefits on REACT attendance and physical activity outcomes, as stated in the REACT logic model. Physical activity outcomes or positive behaviour change in the form of increased physical activity levels and trying new exercise classes were mediated by perceptions of improved self-efficacy for physical activity and improved motivation for physical activity, with participants feeling more capable and motivated to exert autonomy over their choice to participate in daily physical activity and new exercise classes. The logic model proposed that relatedness would act as a mediator of physical activity outcomes. Participants indicated that improved social connectedness not only was a benefit of participation in REACT, but also enhanced enjoyment of REACT and acted as a driver for continued participation. The findings suggested that connectedness, shared identity and sense of relatedness were moderated by the specific intervention groups that participants were members of. The logic model currently identifies site characteristics that could moderate REACT attendance and physical activity outcomes; however, the moderating role of the specific intervention group that participants join on social connectedness and relatedness was not specifically identified. Group dynamics and coherence may differ among different groups and their members, and these processes need more attention and targeting. These findings are shown in the context of the logic model in a modified REACT logic model (see *Appendix 2, Figure 29*).

### **Barriers to participation in REACT and daily physical activity**

Under the central theme 'barriers to participation in REACT and daily physical activity', participants identified nine main barriers (HOTs): being time poor, declining mental well-being, physical health barriers, REACT programme features, difficulty accessing transport, cost of daily physical activity, negative life experiences, bad weather and proposed reasons for dropouts at the 6-month interviews. These HOTs comprise 24 subthemes at 6 months. At 12 months, all nine HOTs identified at 6 months remained, in addition to three new HOTs: not perceiving benefits, barriers in the local area and lack of past physical activity. The 12 HOTs at 12 months comprised 32 subthemes, nine newly identified at 12 months, while 23 remained from the 6-month interviews. At the 24-month interviews, 12 subthemes remained, with the emergence of one new subtheme. Although all HOTs and subthemes within this central theme are presented in *Figure 4*, only the key themes in terms of addressing the research questions of this evaluation are discussed below. Furthermore, we distinguish between barriers that influenced attendance at REACT sessions and those that affected daily physical activity participation.

**Barriers to participation in REACT and daily physical activity and the REACT logic model**

Although people classified as either frail or pre-frail experienced physical health barriers to REACT and daily physical activity, those classified as frail experienced several health problems simultaneously:

*My muscles, because I have this muscle weakness . . . because like sitting down I'm fine but getting up I struggle . . . I used to fall over quite often . . . I'd lose my balance like . . . And then I've seems to got no muscles to hold me back . . . I'm just blop! [laughs] . . . that's my big problem.*

*Alvita, 12 months, SPPB score of 5, attendance 91%*

In addition to physical health barriers, those classified as frail also reported experiencing other barriers to participation, such as having other commitments and difficulty accessing transport. A difference in physical health barriers experienced between people classified as frail and people classified as pre-frail was less evident in the 12-month interviews, with both groups reporting similar physical health barriers. Although the barriers to physical activity for participants classified as frail were dominated by physical health barriers in the 12-month interviews, participants classified as pre-frail reported barriers that related to the REACT programme features, including issues with some venues, some REACT session leaders, the delivery of the health behaviour maintenance programme, some faulty pedometers and moving from two sessions per week to one session per week from week 12 onwards:

*I'm a bit disappointed in we had it [health behaviour maintenance programme] a couple of times when it first started but most of the people don't seem to want to stay behind you know and have a coffee and a chat . . . They seem to have other things they want to do.*

*Cecil, 6 months, SPPB score of 4, attendance 65%*

One participant recounted there being no formal discussion of goal-setting:

*. . . we have talked about the step counter [pedometer] uh but we've sort I don't think we've never talked about the actual, about any actual goals you know.*

*Beatrice, 6 months, SPPB score of 9, attendance 96%*

The barriers that were related to REACT programme features were confirmed in session leader interviews, in which it was reported that, although the use of pedometers was popular in some groups, a number of groups faced difficulties with poor-quality pedometers that did not accurately record steps.

Participants classified as frail reported experiencing numerous and chronic health problems that remained throughout the REACT programme. By contrast, those classified as pre-frail often reported experiencing either acute illness or injuries that were relatively short-lived or health concerns that were stable or they were accustomed to dealing with, such as diabetes. At 12 months, both participants classified as frail and those considered pre-frail shared their experiences of the REACT programme and identified where they faced challenges. These experiences also varied depending on the site of the intervention group and how some session leaders delivered different components of the programme. Reports from interviews with session leaders also noted the challenges involved when working with older adults experiencing numerous physical health barriers, which often included adapting exercise:

*When you're working with 15 people with multiple health conditions you need to be flexible with your delivery.*

At the 24-month interviews, physical health barriers were the most cited reasons for lapses in physical activity participation, together with leading busy lives and having competing commitments. Participants who articulated strategies to overcome 'tensions' to physical activity maintenance had both (1) lived active lives and (2) been engaged in a wide array of activities prior to REACT. Self-regulatory strategies to overcome challenges, such as declining health, competing commitments, motivational barriers and adverse weather, included adjusting expectations or actual physical activity, self-monitoring and making plans. These are discussed in more detail in *Enablers of participation in REACT and daily physical activity*.

### **Barriers to participation in REACT and daily physical activity for people with low and high attendance**

Although the responses between the two groups were similar in the 6-month interviews, in the 12-month interviews there was a distinction between participants with high attendance scores, who tended to experience a single barrier to participation (predominantly physical health), and participants with low attendance scores, who reported multiple barriers to participation (predominantly physical health barriers and having other commitments). When considering participants who were in both the low SPPB score and the low attendance strata, it is important to note that two of the three participants interviewed at 6 months dropped out shortly before the 6-month interview, having experienced multiple barriers to participation relating to physical health, time commitments and travel difficulties. The third participant dropped out a few months after the 6-month interview, having experienced significant physical health barriers and reporting having other commitments. None of the three was available for follow-up at the 12-month interviews.

Being time poor and its associated subthemes – having competing commitments and not having the time – were expected to be a barrier to physical activity, especially in the short term; however, BCTs were incorporated into the REACT programme to help participants to identify and resolve sources of tension around increasing physical activity. Reports that time was a barrier to participation remained throughout the 12 months, indicating that participants were not able to successfully utilise BCTs, the BCTs were not suitable for resolving these tensions or the BCTs were not delivered as planned. The intervention fidelity study, which assessed delivery fidelity of health behaviour maintenance sessions (not exercise sessions), triangulates this finding, as it demonstrated that BCTs, such as managing setbacks and problem-solving, were delivered with low fidelity. Furthermore, reports from session leader interviews at follow-up indicated that reasons for poor delivery fidelity could include (1) not feeling comfortable delivering the content or (2) participants not being engaged with the health behaviour maintenance sessions. This, together with real life issues (e.g. lack of access to transport or competing interests) that participants may have faced, may explain why issues with time availability and tensions with having competing commitments remained at 12 months:

*I mean I have been playing indoor bowls quite recently up and till about a month ago but I've had to give up my car so I can't get around quite as much so it looks like I am not able to drive. And also I've got to find someone to take me.*

*BT2013, 24-month interview*

*I didn't pursue any others. I should have done, I think, but it may be difficult to explain this to you but despite my age I'm very busy, I don't have much spare time. I just deferred starting my last book due to lack of time, I'm going to write it some time. I've written 16 altogether but . . . I've got one more I want to write but certainly things get in the way, but I'll get there. I've got to make sure I live long enough to do it!*

*AT2624, 24-month interviews*

Although REACT programme features such as issues with the REACT session leader and issues with REACT structure and content were not reported to have affected REACT attendance, REACT venue issues were prohibitive of full engagement in REACT sessions. This shows how context in the form of intervention provider and site can, by moderating the way in which interventions are implemented, affect the way that participants experience an intervention and engage with it, as illustrated in the logic model. For example, in the case of issues with REACT structure and content, reports suggest that in some cases BCTs (managing setbacks and problem-solving/goal-setting and action-planning) were not discussed widely in some groups, suggesting a lack of competent delivery of health behaviour maintenance sessions on behalf of the REACT session leader, which confirms findings from the intervention fidelity study and reports from session leaders that some could not remember content from health behaviour maintenance training sessions. The intervention fidelity study showed that the delivery of action-planning and goal-setting was rated as needing improvement (mean score 2.2, SD 1.14), while managing setbacks and problem-solving were delivered with low fidelity (mean score 1.9, SD 0.81).



Participants did not report the delivery fidelity of these BCTs during health behaviour maintenance sessions as being prohibitive of either daily physical activity or REACT attendance; however, intervention outcomes may indicate whether or not this influenced long-term physical activity outcomes as the logic model anticipates it could. Participants with low motivation for physical activity or low self-efficacy for physical activity did not report this as a barrier to REACT attendance, which suggests that participants had the motivation for specifically engaging with REACT and felt a sense of self-efficacy and competence in the REACT context. It is expected that this would not necessarily translate to physical activity in other contexts. This supports the logic model, which anticipated that changes in intrinsic motivation and competence/self-efficacy would mediate changes in physical activity outcomes, in either a negative or a positive feedback loop. Participants reporting physical health barriers seemed to overcome them during REACT classes. This is perhaps because feelings of low motivation for physical activity or low self-efficacy for physical activity were reported more in relation to daily physical activity rather than in REACT itself (i.e. physical health barriers were not further compounded by low self-efficacy for physical activity and low motivation for physical activity). Many participants reported that having the autonomy to manage their own exercise throughout the class enabled them to overcome physical health barriers. Furthermore, the fact that session leaders were able to adapt exercises for those struggling with physical health barriers supported the mediating effect that participants' autonomy over their route of progression and ability to carry out an adapted exercise had on REACT attendance.

Participants and session leaders reported difficulty in accessing transport as an important contextual variable affecting REACT participants, who used a range of strategies to overcome this barrier, such as carpooling and accessing community transport.

The organisation of carpooling is an example of participants building support networks to enable REACT attendance. The logic model anticipated that this could be a determinant of physical activity outcomes and could be mediated by feelings of relatedness among REACT peers. Findings from 'barriers to participation in REACT and daily physical activity' in the context of the logic model are shown in a modified REACT logic model (see *Appendix 2, Figure 30*).

### ***Enablers of participation in REACT and daily physical activity***

Within the subtheme 'enablers of participation in REACT and daily physical activity' there was a clear distinction between pre-existing motivation and components of the REACT project targeting participant motivation. This was reflected in the two HOTs, characteristics of REACT and personal and psychosocial characteristics of participants, including 10 subthemes. At the 12-month interviews, all 10 subthemes remained, with the addition of a new subtheme. At the 24-month interviews, 11 subthemes remained, with the emergence of two new subthemes (see *Appendix 2, Figure 31*).

### ***Enablers of participation in REACT and daily physical activity for people classified as frail and pre-frail***

At the 6-month interviews, participants classified as either frail or pre-frail reported similar enablers of participation to REACT. However, both groups demonstrated a shift in importance of enablers from the 6-month to the 12-month interviews. At 6 months, the predominant enabling factor was the support received from REACT session leaders; however, although this is also discussed at the 12-month interviews, the emphasis was on enjoyment of REACT and supportive group dynamics:

*Just being with everybody and seeing how we could all do, and everybody said 'oh well done'... we all encouraged each other I think you know.*

*Beatrice, 12 months, SPPB score of 9, attendance 96%*

This suggests a transition from REACT session leader-supported participation to a group network of support for physical activity. When focusing on daily physical activity, participants reported that their REACT session leaders' encouragement and support of other activities within the community

were important factors in their continuation of activity independent of REACT. Furthermore, having motivations to return to previous physical activity behaviours seemed to be enabling of daily physical activity for both frail and pre-frail participants. Previous physical activity behaviours were also important at the 24-month interviews, with participants who overcame barriers to physical activity participation being those who had been historically active and engaged in a variety of activities before enrolling in REACT.

In contrast to reports at the 6-month and 12-month interviews, in which self-regulatory techniques were not commonly cited, at the 24-month interviews self-regulatory techniques were reportedly used in an effort to overcome challenges such as declining health, competing commitments, motivational barriers and adverse weather. These self-regulatory strategies included:

- Adjusting expectations of or actual physical activity in accordance with their health limitations and other commitments –

*If I can't [attend exercise class] then I shall do what I do now which is to do light exercises for my back before I get out of bed in the morning and walk and [partner's] corridor is about the same length as my path so we walk up and down that several times . . .*

*Valerie, 24 months, SPPB score of 9, attendance 61%*

- Walking indoors was a coping strategy for staying active when the weather was poor –

*I make a habit of a daily walk you know, of course if the weather is very very bad. Like it was so windy I try to go but then I just walk up and down the corridor.*

*Darsha, 24 months, SPPB score of 8, attendance 81%*

- Doing 'something rather than nothing' was another coping strategy –

*. . . now instead of just going and getting my paper and coming straight back, I could do a circuit if it's a nice day, you know? I'm not counting them [steps] now but it's still reasonable.*

*Frederick, 24 months, SPPB score of 9, attendance 90%*

*You have to make yourself sometimes, you know I think to myself, 'come on get up get up' and I do I reckon every day I reckon to go out somewhere, even if I only walk up the village and back down.*

*Rachel, 24 months, SPPB score of 11, attendance 82%*

*I'm pretty active in silly ways. It sounds nothing, but I don't use the lift . . . it is a small thing . . . so not to use the lift, I have to walk to the far end of the building . . . Go down the stairs and walk back half-way through the building. It's far enough to be an activity.*

*Frederick, 24 months, SPPB score of 9, attendance 90%*

- Repeating the same walking route enabled participants to self-monitor their progress –

*That's a favourite walk because we can sort of tell where we were last year and last week.*

*Mark, 24 months, SPPB score of 9, attendance 63%*

- 'Habit', 'routine' and 'making plans' were key enablers for physical activity, whereas participants did stress that not being as able as before was evident in generally being slow with all activities, which led to frustration –

*I make a plan to walk from home to (city centre) and back . . . and then at home I walk, because I have a lot of stairs, I walk up and down, up and down until I get tired.*

*Dorothy, 24 months, SPPB score of 12, attendance 72%*



Post REACT at the 24-month interviews, individuals who reported frequent physical activity engagement described gaining pleasure and enjoyment from physical activity:

*I love seeing things, I love being out in the fresh air I love noticing things. I don't get up and say right I've gotta do a couple of miles this morning, no I go when I want to enjoy myself.*

*Jane, 24 months, SPPB score of 10, attendance 35%*

Another participant described how physical activity was an enabler to participation in meaningful activities:

*I've got no real interest in ... just exercise as exercise. My interest was always in exercise really incidental to some proper activity, and that would be a social activity among people I liked.*

*Frederick, 24 months, SPPB score of 9, attendance 90%*

Despite the importance of a supportive group network as an enabling factor during the REACT programme, post REACT, social support was a less prominent theme in explaining physical activity maintenance. None of the interviewed participants had utilised support from community agencies (e.g. health visitors, health friend scheme):

*I personally did not need it, because you know I am quite fulfilled with the activities that I am doing, and then I go on holiday and visit my daughters and granddaughters.*

*Arman, 24 months, SPPB score of 8, attendance 95%*

For some participants, friendships developed during REACT and motivational support from friends and spouses facilitated continued physical activity engagement:

*Well [friend] bullied me ... she went and it [attending REACT] was a good way of meeting up and perhaps getting a coffee after ... and then it developed from there ...*

*Valerie, 24 months, SPPB score of 9, attendance 61%*

*I'm not very good at self-motivation ... if you said, oh, you know, go for a run every morning I wouldn't, but if the whole group of us were meeting and I said, 'oh we're going to go for a short walk meet at 10 o'clock at [location]' then I'd go.*

*Rita, 24 months, SPPB score of 10, attendance 75%*

Illness also brought about changes in some friendships and had an impact on people's social activities:

*It was quite hard for me because she [friend] used to come and ring me up every day and we'd meet for an evening meal nearly every day, and all of a sudden it was gone!*

*Rita, 24 months, SPPB score of 10, attendance 75%*

Social isolation and lack of support can have detrimental effects, as vividly reported by one participant:

*I'm less active than I ever was in the past and it's getting worse because I have no friends, no relatives. I don't have children ... They've all died or they've got hip or knee problems ... or they're involved with family, you know, so I have to create things to do to keep myself occupied but in [city] it's a big fat zero. I have nothing, no one, and I don't want to go walking by myself because it can be dangerous. If I fall there's no one you see?*

*Eleanor, 24 months, SPPB score of 12, attendance 30%*

### **Enablers of participation in REACT and daily physical activity of people with low and high attendance**

It was apparent from both the 6-month and the 12-month interviews that people with high attendance were more motivated by REACT than people with low attendance. This indicates that motivation could

act to mediate REACT attendance but barriers could moderate the positive loop between motivation and REACT attendance, as we observed that people in the low-attendance group experienced more physical health barriers. The role of supportive group dynamics was reported more often in the 12-month interviews by people in the high-attendance group. Three participants in the group with low attendance and low SPPB score who reported supportive group dynamics at the 6-month interviews still dropped out, showing that, although supportive group dynamics may be a strong enabling factor, it might not be sufficient to overcome the multiple barriers (physical health, having competing commitments and difficulties accessing transport) that these participants may have faced before they dropped out. People with high attendance reported more participation in previous physical activity behaviours, suggesting a moderating role of previous physical activity behaviours in the attendance of REACT sessions in the short term; however, at the 12-month interviews, this relationship was no longer evident, suggesting that previous experiences may not be a moderator of long-term REACT attendance.

### ***Enablers of participation in REACT and daily physical activity findings in the context of the REACT logic model***

The logic model anticipated that enjoyment of REACT would mediate REACT attendance and physical activity outcomes. Participant reports suggest that enjoyment was a key reason to participate in REACT and start new physical activity classes beyond REACT. Participants reported being motivated by REACT, deriving motivation from their REACT session leaders who supported participants' competence and sense of autonomy, and by watching their REACT peers improve their fitness levels. Participants' perceived competence mediated both short-term and long-term physical activity outcomes. High levels of competence and self-efficacy at baseline, mainly among participants who had previously been active, may influence outcome expectations and need to be accounted for by the REACT logic model (see *Appendix 2, Figure 31*).

### ***Key REACT programme components***

The key programme components responsible for influencing participants' lifestyles at 24 months are summarised below.

#### ***Behaviour change techniques***

Techniques for managing slips/lapses, supporting habit change and resolving sources of tension around increasing physical activity were utilised by participants to support physical activity at 24 months. Participants were able to overcome 'tensions' to physical activity by employing self-regulatory strategies (e.g. incorporating more physical activity into daily life, finding alternative physical activity and monitoring techniques). Routine was an enabler for maintaining PA.

#### ***REACT facilitator delivery style***

REACT session leaders used a person-centred delivery style to build autonomy/intrinsic motivation. Individuals with formerly low physical activity levels expressed increased motivation to be active through mastery and improved self-efficacy. This was sustained post REACT.

#### ***Group-based physical activity***

The group-based format of the REACT programme promoted social connectedness among participants. The social benefits of physical activity were recognised by most participants. Individuals with established social networks reported more physical activity. The success of REACT in supporting sustainable social networks varied between groups. Where this was most successful, social contact continued post intervention.

#### ***Themes identified from session leader and organisation provider interviews***

The REACT programme session leaders and organisation providers were interviewed at follow-up (12 months) to explore whether or not the intervention was delivered as planned and what, if any, changes could be made to the REACT programme for future delivery (research questions 4 and 5).

The analysis of session leader interview data identified 10 themes, which address whether or not the programme was delivered as planned and propose further improvements to the programme components. The programme components presented are manual and training, one-to-one sessions and group exercise classes, health behaviour maintenance sessions, group size/common start time, equipment/recording attendance, ambassadors programme, non-REACT activities, continuation session, transport and cost/funding.

**Manual and training** The REACT manual was a substantial document and was given to session leaders at the training session. It was designed to be used as a reference document throughout the 12-month programme. The manual could appear daunting at first sight, but the training session explained/eased its use and session leaders were very positive about using it:

*... it scares a lot of people ... when you've done a course you realise it's actually quite simple.*

*Exercise specialist, AT9001, Male*

Another session leader recounted:

*I do refer to the training manual, I have it with me at all times.*

*Exercise specialist, AT9002, Male*

Feedback on the REACT 2-day training course delivered by three or four members of the research team was largely positive, with session leaders stressing that it informed them of what they needed to deliver. Some session leaders thought that the sessions focusing on the exercise component of the programme might have been too detailed given that the leaders were well qualified. However, others thought that these sessions were a good refresher:

*I found it very, in depth and possibly ... I didn't need it to be that in depth. But I didn't mind sometimes you need reminders.*

*Exercise specialist, AT9002, Male*

Suggestions for improvements included a practical demonstration of how an entire class would run, and provision of online materials for use during the programme delivery:

*... maybe a demonstration of classes ... this is what session one would look like.*

*Exercise specialist, AT9004, Female*

Another suggestion for improvement was the provision of 'an online resource; an online clip of an hour session can you just watch it when you got time to kind of format and layout' (exercise specialist, AT9001, male).

Feedback was less detailed about the health behaviour change training. Some leaders did not recall the training session, and some recalled only some aspects:

*I'm not quite sure what you mean. I've got to be honest I don't.*

*Exercise specialist, BT001, Female*

### **One-to-one sessions and group exercise classes**

Feedback about these sessions, at which participants met with their session leader on a one-to-one basis, was overwhelmingly positive. One session leader reported that they had taken their participants on a familiarisation tour of the facility as part of the one-to-one session.

There was positive feedback about the structure of the classes, especially the warm-up and games, although 'when you're working with 15 people with multiple health conditions you need to be flexible

with your delivery' (exercise specialist, AT9003, male). Some sessions leaders reported that they had to make some adjustments, including a lengthier warm-up.

Reducing the exercise sessions from two sessions per week to one was perceived as challenging by some, who reported that participants did not want this reduction, as they were engaged with the programme and were making good progress. Some session leaders agreed that moving to one session reduced the rate of progression of the participants.

A change in session leader was challenging for the incoming leader, who reported that the groups were attached to their session leaders and were not happy with changes:

*I took on the first group probably about 4/5 months into it, and I think it was quite hard to engage with them as an outside instructor coming into an already established group.*

*Exercise specialist, AT9003, Male*

### **Health behaviour maintenance sessions**

Feedback about the health behaviour maintenance sessions was mixed, with some session leaders reporting that they did not deliver these sessions as planned because they were uncomfortable with either the content or the format of the sessions:

*I ended up just incorporating what needed to be covered within the classes 'cause I found it really hard to structure specific sessions for that.*

*Exercise specialist, AT9002, Male*

Alternatively, session leaders reported that they did not deliver sessions as planned because their participants did not want to engage with them:

*I've got to say a lot of the time they just wanted to have a coffee and a chat as a group, it was quite difficult to, you know, get all their attention.*

*Exercise specialist, ET9001, Male*

However, some session leaders reported that, although it was challenging to deliver these sessions, on the whole participants enjoyed them:

*... they had a few problems with them but on the whole they enjoyed that.*

*Exercise specialist, BT9004, Female*

### **Group size/common start time**

The fact that all REACT participants joined the group at the same time for the 12-month duration was viewed as both a positive ('being in a group starting off starting off them as a group to get them to know one another which is fantastic'; exercise specialist, BT9003, male) and a negative ('if a few people drop out ... if you don't have a flow of people to keep fresh blood, it's just not sustainable'; exercise specialist, AT9003, male).

### **Equipment/recording attendance**

Everyone was happy using Therabands (Akron, Ohio, USA), but getting ankle weights on and off participants was an issue: 'they took quite a long time to get them on an off so we were kind of losing a bit of time'. Using pedometers as part of the health behaviour maintenance programme was popular, but some groups experienced issues with poor-quality pedometers that did not accurately record steps:

*They really enjoyed having that kind of visual tool [map] to do it and they got really on board with counting their steps and I think that was good motivator.*

*Exercise specialist, AT9003, Male*

There were issues around recording attendance at sessions. Some session leaders did not have access to online systems in the exercise room, so needed to keep paper-based registers. Several session leaders suggested the provision of an online register in future REACT delivery.

### **Ambassadors' programme**

Most session leaders found it difficult to engage participants in a formal ambassador programme because participants 'didn't want that full-on responsibility'. However, in many cases, some elements of the ambassador roles were taken up by participants in a more informal manner:

*There wasn't a specific ambassador . . . they just naturally started to pick their own roles.*

*Exercise specialist, BT9002, Male*

*I think it's a really good idea because it just gives them that more sort of ownership over the programme. But it just didn't work.*

*Exercise specialist, ET9001, Male*

### **Non-REACT activities**

A main focus of the REACT programme was to encourage participants to engage with other community activities. This was particularly the case when the frequency of sessions changed from twice per week to once per week. Session leaders reported some level of success with some groups, but some participants 'were not easily persuaded and were very wedded to their group and session' (exercise specialist, AT9002, male). However, 'one or two of them jumped on quite quickly were very eager to do it' (exercise specialist, AT9001, male). Despite some having concerns about new activities, 'a lot of them were a bit worried about going out and trying other things' (exercise specialist, AT9001, male).

### **Continuation sessions**

Many groups did not want the sessions to end after the 12-month programme:

*When I had to leave the group, it's been a bit traumatic. They were really really upset. So the majority have continued (in some form) with participants paying to attend.*

*Exercise specialist, AT9002, Male*

As a result, additional sessions were created by trusts for the REACT participants:

*There was so much demand for the trust to continue with some form of exercise class that we created the [name] weekly exercise class [REACT continuation].*

*Exercise specialist, AT9008, Male*

### **Transport**

Despite the fact that during the REACT screening process potential participants were given specific information about the location of the REACT classes and asked if they would be able to attend sessions, some still faced transport issues. These were usually solved by community transport or, more commonly, through lift-sharing.

### **Cost/funding**

Some session leaders reported that the programme was a large investment for a small number of people, but they acknowledged the importance of free sessions to improve engagement with exercise sessions:

*But I think you need to hook them with the free yeah just to get them settled.*

*Exercise specialist, ET9001, Male*

## Discussion

### Summary of findings

The longitudinal qualitative evaluation identified the perceived benefits, barriers to and enablers of participation in REACT and daily physical activity at 6, 12, and 24 months, while exploring the perspectives of the session leaders on what worked well and necessary changes for a community roll-out. Exploring the theorised mechanisms proposed by the REACT logic model helped us to understand the way that psychological processes mediate and moderate the programme's impact on physical activity behaviour change. In addition, the interviews with participants and providers helped us to better understand the variation in delivery fidelity that was illustrated by the quantitative fidelity study and whether or not this variation affected participant experiences and outcomes. Exploration of the logic model in this way highlighted where the logic model can be refined to inform changes and adaptations to the REACT programme for a community roll-out.

The findings demonstrate that older adults engaged with the REACT programme, seeking better health, both physical and mental, as well as better social connectedness. During the 12-month programme, they experienced improvements in physical and mental health, improved social connectedness and positive behaviour change. Most benefits described at the 6-month interviews were also reported at 12 months; however, the way in which participants related to these benefits changed. For instance, at 6 months, participants described the physical health benefits that they experienced (e.g. improved balance); however, at 12 months, they described how these benefits translated to better day-to-day quality of life (walking without a stick). Although competence and self-efficacy at the 6-month interviews was derived mainly from the REACT session leader and REACT peers, at the 12-month interviews, participants described deriving self-efficacy and competence from reflecting on their own achievements. This enabled them to explore other exercise opportunities in their community. At 24 months, there was a strong focus on maintenance of functional ability improvements, with participants reporting strategies for long-term maintenance.

There was a distinct transition in how participants experienced enabling factors during the REACT programme. The chief enabling factor at 6 months was the support received from the REACT session leaders; however, the emphasis shifted at 12 months to the support derived from the REACT groups, indicating the creation of a network of group support and a shared identity around the REACT programme.

Personal and psychosocial enablers of REACT attendance and daily physical activity included having good levels of competence, self-efficacy and motivation to be physically active, alongside a history of being physically active. Despite these enablers, daily physical activity was impacted to a greater extent by barriers than REACT attendance. Barriers experienced by participants usually persisted throughout the programme; these were seemingly overcome during REACT exercise sessions but seemed to continue to affect daily physical activity. For example, physical and mental well-being barriers persisted throughout the REACT programme but were overcome with the help of REACT session leaders or REACT peers. Not having enough time affected daily physical activity but not REACT, as did low self-efficacy and competence or low motivation. Furthermore, physical health barriers, such as mobility issues, pain and discomfort and tiredness, were more prohibitive of daily physical activity than REACT attendance. These barriers were seemingly overcome either with the support of the REACT session leader or peers or, in the case of lack of time as a barrier, by prioritising attendance of REACT sessions. However, when participants were outside the REACT programme, the environment was not conducive to overcoming these barriers to the same extent. With the removal of the REACT session leader and a supportive group of peers, these barriers were less likely to be overcome. Planning for the community delivery of REACT needs to focus on a tailored, personalised approach, while at the same time a whole-systems approach needs to be implemented to target barriers operating outside the personal level.

At 24 months, participants' accounts mostly aligned with the hypothesised mechanisms of change depicted in the REACT logic model, although some aspects were more effective than others in



promoting long-term behaviour change. There was mixed evidence for the programme's effectiveness in improving relatedness. Different ability levels among group members interfered with the creation of group relatedness and connectedness. Participants often described having supportive social networks independent of the REACT programme. This may suggest that participants' social support needs were fulfilled elsewhere. However, this was not universally true, and people with limited social networks reported that diminished social interactions influenced their health and well-being and their opportunities for daily activity.

Themes identified at 24 months largely mirror those reported at 6 and 12 months. However, whereas the 6- and 12-month interviews found that social support was a key reason for engaging in REACT, at 24 months individual-level factors were more prominent themes in explaining physical activity maintenance. Self-regulatory strategies were also discussed, which were not identified in the earlier interviews. Key components of the REACT programme that positively influenced maintenance of physical activity at 24 months were (1) techniques for managing slips/lapses, supporting habit change and resolving sources of tension around increasing physical activity; (2) person-centred delivery style to build autonomy/intrinsic motivation; and (3) group-based delivery promoting social connectedness.

The findings from interviews with session leaders and participants triangulated the findings of the fidelity of intervention delivery study. The health behaviour change programme received mixed evaluation. Participants and session leaders highlighted the importance of the programme and its focus on self-regulation, strategies to resolve tensions and competing interests, and increasing knowledge and understanding of the importance of physical activity and ways to incorporate physical activity into daily lives. However, for a REACT community delivery, this programme requires fine-tuning and a good communication and support system for REACT session leaders to enable them to deliver the programme as planned.

### **Considerations for refining the REACT logic model**

One of the main aims of the qualitative process evaluation was to refine the logic model further with new knowledge based on participants' and providers' experiences and evaluation of the REACT programme. The interview findings confirm many of the processes and mechanisms of impact anticipated by the logic model. There are some aspects of the logic model that this study was not designed to address, such as the objectively measured physical activity and SPPB outcomes. The analysis of the qualitative data has highlighted six key modifications for further refining the REACT logic model:

1. Social interactions, formation of shared identity, support networks and feelings of relatedness are moderated by intervention group as a contextual variable. This should be identified in the logic model to enable the assessment of this contextual variable on intervention processes (relatedness) and outcomes (REACT attendance and physical activity outcomes).
2. The responses of participants in the qualitative study stressed that perceived benefits (physical, social and emotional) may be influenced by baseline SPPB score. Consequently, baseline SPPB score needs to be recognised in the logic model as having the potential to impact the mediating effect of perceived benefits on REACT attendance and physical activity outcomes.
3. The delivery of BCTs (monitoring progress, goal-setting and action-planning, and managing setbacks and problem-solving) was moderated by intervention group and session leaders. Session leaders are recognised in the REACT logic model as a key contextual variable and, similarly, intervention group should be identified as a contextual variable potentially moderating intervention delivery. This will highlight the need to monitor and assess delivery fidelity, how this changes from one context to another and any impact on perceived benefits, ability to overcome barriers, REACT attendance and physical activity outcomes.
4. Baseline self-efficacy and competence were identified in the interviews as key enablers of REACT attendance and daily physical activity. Therefore, participant baseline self-efficacy and competence should be identified in the logic model alongside baseline SPPB and physical activity as a key contextual variable in the moderation of REACT attendance and physical activity outcomes.

5. Access to transport was identified as an important environmental barrier to daily physical activity and, in some cases, REACT attendance. Subsequently, it should be recognised within the logic model as an important contextual variable with the potential to moderate both REACT attendance and physical activity outcomes.
6. Competence is recognised in the logic model for its role in mediating REACT attendance and physical activity outcomes; however, interview findings show that the source of the competence needs to be recognised. The source of competence transitioned from external sources (REACT peers and session leaders) to internal sources (drawn from own capabilities).

### Strengths and limitations

A key strength of this qualitative process evaluation was the longitudinal study design, with the collection of data from study participants at multiple time points, including interviews at 6 months (during), 12 months (post intervention) and 24 months (12 months post intervention).<sup>89</sup> This approach provided a robust, in-depth account of participant experiences and the psychological processes responsible for behaviour change, while identifying the interactions that these processes had with context and time over the 24-month period. As previously described, there are few longitudinal qualitative studies of older adults' experiences of physical activity interventions, and this study adds to this limited body of research by highlighting the dynamic processes involved in older adults' engagement with a physical activity intervention, and the uptake and maintenance of daily physical activity over a 24-month period.<sup>90-93</sup>

Having a logic model that clearly defines an intervention helps to identify where evidence agrees with underlying theory and anticipated pathways or if there is limited agreement and the intervention needs to be refined.<sup>43,94,95</sup> The analysis of findings in the context of the REACT logic model enabled the evaluation of the programme and highlighted successful areas of delivery, the mechanisms by which the core elements of the programme work and how they are affected by contextual variables to produce short-term and long-term changes. The integration of the REACT logic model with the qualitative and the quantitative findings will inform the refinement and scale-up of the REACT programme and help to explain variation in outcomes, for example why some participants respond to the programme more successfully than others.

The purposive sampling strategy adopted allowed for sampling of REACT participants across a wide range of characteristics (SPPB, attendance, age, sex, ethnicity, intervention group, intervention session leader and provider). This strategy maximised the diversity in participant experience and provided rich data grounded in the real experiences of community-dwelling older adults. In addition, this sampling strategy facilitated a subgroup analysis of themes based on SPPB and REACT attendance,<sup>96</sup> which highlighted key differences in the way that low (frail) and high (pre-frail) SPPB participants experienced both benefits of REACT participation and barriers to REACT participation and daily physical activity. These differences have important implications for the way that physical activity interventions are designed and delivered.

The purposive sampling ensured diversity of participant characteristics and that thematic saturation was reached during the analysis. However, we cannot be sure that the addition of data from the participants who declined a follow-up interview would not have affected the findings, and we accept this as a limitation of longitudinal qualitative studies. Consequently, we need to be careful when we generalise findings to other populations, especially those less represented in the study sample, for example frail older adults.

A common limitation associated with interviews is the generation of responses from participants that are socially desirable, undermining the credibility of the results. Furthermore, researcher bias when analysing interview data is also possible. Several strategies (prolonged engagement, reflexivity in approach to analysis, involving discussions with other researchers, paraphrasing participant responses to ensure correct interpretations) were employed to increase credibility and minimise bias in the study.



This could have been further limited by using multiple interviewers as a form of triangulation; however, it was deemed that rapport and an awareness of previous interviews was an important factor in facilitating in-depth narratives.<sup>89</sup>

### Implications for practice

Moving forwards to a potential community roll-out, the findings from the interviews with the session leaders highlighted some key actions to fine-tune the REACT programme while providing some key implications for the design of active ageing programmes (*Table 8*).

### Considerations for research

Key considerations for research, based on the findings of this study, are presented in *Table 9*.

TABLE 8 Implications for practice

Key actions	Details
Promote social interaction and enjoyment	REACT attendance and physical activity outcomes were moderated by enjoyment and mediated by relatedness. This was a successful REACT component and a future roll-out should continue targeting enjoyment and social interaction. Active ageing programmes should promote exercise that is enjoyable and facilitate a sense of belonging and social interaction
Health behaviour maintenance sessions	Revise the content and delivery of these sessions and provide ongoing support to session leaders to deliver the content of these sessions with confidence. Ensure that content of these sessions is engaging and interactive and avoid the 'classroom' effect. Capitalise on opportunities to reinforce key messages during the group exercise sessions and during tea and coffee breaks
Transition to non-REACT activities	Evidence of successful transition arrangements during REACT reinforce the importance of focusing on incorporating such arrangements in future REACT roll-out. Provision of information and support to identify meaningful and enjoyable activities and facilitation of taster sessions during the REACT delivery would enable physical activity maintenance post REACT. Furthermore, participants should be given the opportunity to continue with the REACT programme if they wish, so no official programme finish is recommended
Session leaders	The role of the session leader is key in the delivery of a group-based programme. REACT session leaders use social interactions to maximise enjoyment, encourage social bonding and foster a sense of belonging. Their delivery is positive and person centred, and they constantly seek to create opportunities for social interaction, such as exercises in pairs/small groups, water breaks and tea and coffee time afterwards. Focusing on individual progress, adapting exercise to suit individual needs and ensuring that each participant exercises at the right level for them is challenging in a group delivery but achievable. That seems to be particularly important for people with frailty and low SPPB scores, who may face a range of complex barriers, and they may require a more holistic approach to ensure their continuous involvement with the REACT programme
Recruitment	Being able to contribute to research that may benefit the community was a key source of motivation for older adults engaging in REACT. In future roll-out of REACT, the contribution of participants in the group programme, their commitment to their group and ways to facilitate the contribution older adults in the delivery of the programme, especially the health behaviour maintenance programme, could increase the value attached to their engagement. Responsibility to the group, sense of belonging and sense of sharing similar challenges were important motivators for REACT participants

TABLE 9 Considerations for research

Key actions	Details
Understanding the experiences of frail older people	More targeted qualitative research is needed to understand how frail older adults self-evaluate or relate to this, how this affects their experience of physical activity interventions and how they can be supported in the face of numerous barriers to physical activity and fewer perceived benefits. This knowledge can be used in the design of physical activity interventions targeting older adults and the training of intervention session leaders to successfully support them
Utilising longitudinal methodology	Qualitative researchers exploring the processes involved in the long-term behaviour change of older adults need to consider longitudinal study design. This study design can provide in-depth data and a robust narrative over time, from which to understand older adults' experiences of physical activity interventions and the processes involved in long-term physical activity behaviour change

### Study 3: quantitative evaluation of intervention mechanisms

#### Aims

- To investigate the mediating and moderating effects of changes in physical activity on the primary outcome (change in physical function) and other outcomes of interest (cognitive function, sleep quality, well-being, depression).
- To assess changes in proposed psychosocial determinants of physical activity behaviour suggested by SCT and SDT (the theoretical underpinnings of the intervention model).
- To explore the role of contextual variables (e.g. age, sex and ethnicity) in moderating the effects of the REACT intervention.

#### Methods

##### Design

A detailed set of hypotheses (outlined in the following section) were derived from the logic model and quantitatively tested using a range of process measures, including questionnaires and accelerometer data collected at baseline and 6 (short term), 12 (medium term) and 24 months (long term). The hypotheses were stated and published 'a priori', prior to unlocking or any analysis of the REACT data set.<sup>1</sup> Some minor edits were made prior to analysis (e.g. owing to non-availability of some data), and these are indicated in footnotes to the hypotheses below.

##### Hypotheses for quantitative testing of the REACT logic model

The hypotheses were derived from the stated and implied assumptions in the REACT logic model. The time from baseline to 12 months represents the intervention period (0–6 months was the 'initial change phase' and 6–12 months was the 'supported maintenance phase') and the 12- to 24-month time frame represents the post-intervention (unsupported maintenance) phase.

Hypothesis 1 was that being in the intervention arm would lead to changes in physical activity (MVPA, steps and sedentary time) and engagement in muscle-strengthening exercise from 0 to 6 months.

Hypothesis 2 was that increased exposure to the intervention would correlate with increased change in physical activity (MVPA, steps and sedentary time) and engagement in muscle-strengthening exercise from 0 to 6 months.

Hypothesis 3 was that increased exposure to the intervention would correlate with increased change in physical activity (MVPA, steps and sedentary time), engagement in muscle-strengthening exercise and SPPB score from 0 to 12 months.

Hypothesis 4 was that increased exposure to the intervention would lead to increased maintenance of physical activity and engagement in exercise. Therefore, within the intervention arm, intervention dose would correlate negatively with decreases in physical activity, engagement in muscle-strengthening exercise and SPPB score from 6 to 12 months (during the supported maintenance period) and from 12 to 24 months (the unsupported maintenance period).

Hypothesis 5 was that exposure to the intervention would lead to changes in key psychosocial determinants of physical activity and exercise from baseline to 6 and 12 months:

- Compared with the control arm, the intervention arm would experience increases in autonomy, competence (self-efficacy), relatedness for MVPA and exercise, perceived intrinsic benefits of exercise (social, physical and emotional) and enjoyment of physical activity and exercise from 0 to 6 months. An originally intended measure of physical activity-related self-concept was removed from the questionnaire pack to reduce patient burden following review of the process evaluation measures by the TMG and patient and public involvement members. Hence the words ‘self-concept’ have been removed from the original wording of this hypothesis.
- Compared with the control arm, the intervention arm would experience increases in physical activity-related autonomy, competence (self-efficacy), relatedness for physical activity and exercise, perceived intrinsic benefits of physical activity and exercise (social, physical and emotional) and enjoyment of physical activity and exercise from 0 to 12 months. As with the hypothesis above, the words ‘self-concept’ have been removed from the original wording of this hypothesis.
- Increased exposure to the intervention (total contact time from baseline to the relevant time point) would correlate with increased change in the above determinants (and in the expected direction).

Hypothesis 6 was that the intervention effect on SPPB (intervention vs. control) would be mediated by changes in muscle-strengthening exercise and by changes in MVPA, changes in lower intensity physical activity or walking activity (Figure 5). The amount of variance in SPPB explained by the different types of activity/exercise will be of interest. The words ‘changes in balance and co-ordination exercise’ have been removed from the original wording of this hypothesis, as this could not be distinguished from muscle-strengthening exercise by the measure used.

Hypothesis 7 was that the intervention effect on SPPB score and on physical activity and exercise from 0 to 6 and 0 to 24 months may be moderated by a number of potential moderating variables, including age, sex, ethnicity, baseline physical activity and SPPB, co-interventions, comorbidities, BMI, mental health, SES, education level (see Figure 5, and Appendix 2, Figure 30). The words ‘and on PA

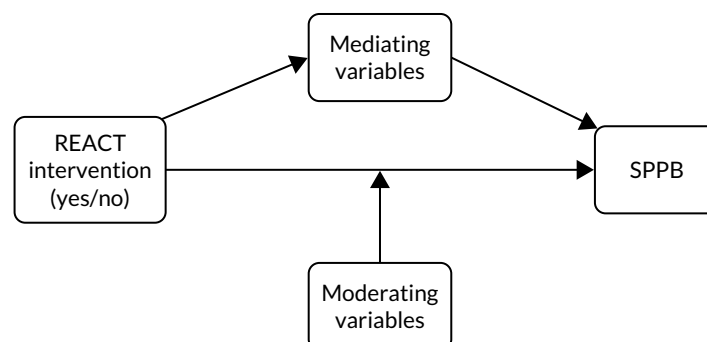


FIGURE 5 Causal diagram for mediators and moderators of intervention effects on physical function.

and exercise' have been added to this hypothesis, as the moderating influence of contextual variables on changes in physical activity behaviours (as well as physical functioning) is of key scientific interest. The time point 0–12 months was removed to reduce the number of analyses. The moderation of short-term (0–6 months) and long-term (0–24 months) effects was considered to be of the most interest from the three time points available.

## Participants

We applied analyses to the whole sample where data were available; however, some hypotheses applied only to the intervention arm.

## Measures

Brief questionnaires assessing the following mechanisms of change suggested by SCT and SDT (the main theoretical underpinnings of the intervention model) were administered at baseline and 6, 12 and 24 months (see *Table 2* for full details):

- autonomy in relation to physical activity
- competence for physical activity
- relatedness for physical activity
- enjoyment of physical activity
- perceived intrinsic benefits of physical activity (social, physical and emotional)
- autonomy for strength-building exercise
- competence for strength-building exercise
- relatedness for strength-building exercise
- enjoyment of strength-building exercise
- perceived intrinsic benefits of strength-building exercise (social, physical and emotional).

To test the hypotheses, we also used measures of intervention engagement (session attendance) and intervention effectiveness [SPPB scores, physical activity/time doing MVPA (by accelerometer and PASE questionnaire), walking (PASE walking subscale), sedentary time (accelerometer), self-reported engagement in muscle strength exercise (MSEQ adherence score)] and secondary outcomes of interest (cognitive function, falls, pain, sleep quality, well-being and grip strength).

Variables used to define individual context were age, sex, ethnicity, baseline physical activity, baseline SPPB, co-interventions, physical comorbidities, BMI, mental health status (any diagnosed mental health conditions in the 6 months prior to baseline), IMD (from postcode) and education level. All of the measures are described in detail in *Chapter 2*.

## Analysis

The above hypotheses were tested using a variety of statistical analyses, as described in *Table 10*. To maximise sensitivity to change and avoid difficulties in constructing the mediation analyses (which require specialist software),<sup>97</sup> the analyses testing hypotheses 1 to 19 have been constructed without accounting for clustering by exercise group within the intervention arm.

## Results

### Hypothesis 1 to hypothesis 5: effects of the intervention on mediators of lower limb physical function

Full data tables are available in *Report Supplementary Material 2*. Key statistics are provided in the following text for significant findings only.

#### Hypothesis 1

Hypothesis 1 was that being in the intervention arm would lead to changes in physical activity (MVPA, steps, sedentary time) and engagement in muscle-strengthening exercise from 0 to 6 months.

TABLE 10 Mechanistic hypotheses and analysis methods used to test them

Hypothesis	Analysis
Hypothesis 1: the intervention will increase physical activity and muscle-strengthening exercise from 0 to 6 months	Comparison of change scores between intervention and control arms, using univariate ANCOVA models with the baseline value of the outcome, age (continuous variable), sex (male, female) and study site (Devon, Bath/Bristol, Birmingham) entered as covariates
Hypothesis 2: intervention exposure will correlate with change in physical activity and muscle-strengthening exercise from 0 to 6 months	Bivariate correlation between the outcome variable and intervention exposure (% of sessions attended) within the intervention arm only. Given the skewed distribution for intervention exposure (46/409 intervention participants did not attend any sessions), we categorised attendance as high ( $\geq 75\%$ ), medium (51–74%) or low (0–50%) and used ANCOVA analysis to look for a linear trend in the mean outcome scores for these three groups. These models also controlled for age, sex and study site (the covariates used in the primary trial analysis)
Hypothesis 3: intervention exposure will correlate with change in physical activity, muscle-strengthening exercise and SPPB from 0 to 12 months	The analyses were the same as for testing hypothesis 2, except for changes to the dependent variable (changes from 0 to 12 months rather than from 0 to 6 months)
Hypothesis 4: intervention exposure will correlate with maintenance of physical activity and engagement in exercise from 6 to 12 and from 12 to 24 months	The analyses were the same as for testing hypothesis 2, except for changes to the dependent variable
Hypothesis 5: intervention will change key psychosocial determinants of physical activity and exercise from baseline to 6 and 12 months. Such increases will correlate with intervention exposure within the intervention arm	Comparison of change scores between intervention and control arms, using univariate ANCOVA models with the baseline value of the outcome, age (continuous variable), sex (male, female) and study site (Devon, Bath/Bristol, Birmingham) entered as covariates
Hypothesis 6: the intervention effect on SPPB (and on key secondary outcomes) will be mediated by changes in muscle-strengthening exercise and in physical activity	Mediation analyses were conducted using linear regression models set up in SPSS v26, using the PROCESS v3.4 macro. <sup>97</sup> Model 4 was selected, which uses a single regression model to estimate coefficients for (1) the direct effect of the independent variable (intervention arm) on the dependent variable (e.g. SPPB) and (2) the indirect effect of the independent on the dependent variable when mediated by a third variable (e.g. physical activity). Bootstrapping with 5000 samples was used to estimate CIs. Age, sex and study site were entered into the models as covariates
Hypothesis 7: the intervention effect on physical activity and exercise from 0 to 6 and from 0 to 24 months may be moderated by a number of contextual variables (11 variables specified)	Linear regression models were used to estimate the interaction effects of individual moderators with trial arm. We followed procedures for moderation analysis recommended by Hayes. <sup>97</sup> Any significant interactions were probed by calculating point estimates at the 16th, 50th and 84th centiles for continuous variables (equivalent to the mean plus or minus one SD for normally distributed data) or at distinct category values where categorical moderators were included. Any significant interactions were visualised to aid interpretation and checked for floor or ceiling effects. The analyses were conducted in SPSS V26 using the PROCESS macro. <sup>97</sup> The <i>p</i> -value representing significance was not adjusted for multiple testing as these analyses are considered to be exploratory (as indicated by the words 'may be' in the hypothesis). It is worth noting that adjusting for the 11 tests conducted at each time point/for each outcome would yield a <i>p</i> -value of 0.0045. Any findings not meeting this criterion are, therefore, potentially vulnerable to analytic bias owing to multiple testing. Note that, for these analyses, ethnic groups were collapsed into two groups owing to numbers being too small to allow finer-grained analysis: (1) black, Asian, Chinese and other ethnic groups and (2) white

There was mixed evidence in support of hypothesis 1, with measures of self-reported physical activity, walking and muscle-strengthening exercise appearing to be more sensitive to the intervention than accelerometer measures. Unbouted MVPA seemed to be more sensitive than bouts MVPA to the intervention for the accelerometer measures. Hence, unbouted MVPA was used to assess changes in physical activity in testing the following hypotheses (hypotheses 2–7). Based on accelerometer data, there were no significant differences between the intervention arm and the control arm in bouts MVPA (in bouts of at least 10 minutes) or sedentary time at 6 months, controlled for baseline MVPA, site, sex and age. However, there was a close to significant difference of 20.2 minutes of unbouted MVPA per week (2.88 minutes per day, 95% CI 5.82 to  $-0.06$  minutes per day;  $p = 0.054$ ) favouring the intervention arm.

Using questionnaire data, significant differences favouring the intervention arm were observed in self-reported physical activity (PASE total score) (mean difference 16.54 points, 95% CI 7.89 to 25.20 points;  $p < 0.001$ ), self-reported walking (PASE walking outside items: days per week  $\times$  hours per day) (mean difference 0.727 hours per week, 95% CI 0.293 to 1.162 hours per week;  $p = 0.001$ ) and self-reported muscle-strengthening exercise (MSEQ adherence score) (mean difference 0.73 scale points, 95% CI 0.42 to 1.04 scale points;  $p < 0.001$ ).

### Hypothesis 2

Hypothesis 2 was that increased exposure to the intervention would correlate with increased change in physical activity (MVPA, steps, sedentary time) and engagement in muscle-strengthening exercise from 0 to 6 months.

Hypothesis 2 was broadly supported, with a consistent pattern of data across all measures (albeit with two out of the five measures achieving only marginal significance in the linear trend analysis). There were significant, weak correlations between intervention attendance and changes from 0 to 6 months in unbouted MVPA ( $R = 0.147$ ,  $p = 0.012$ ,  $n = 292$ ), PASE total score ( $R = 0.136$ ,  $p = 0.012$ ,  $n = 339$ ) (see Appendix 2, Figure 34), hours per week spent walking outside (from PASE walking items) ( $R = 0.165$ ,  $p = 0.002$ ,  $n = 332$ ) and change in MSEQ adherence score ( $R = 0.122$ ,  $p = 0.029$ ,  $n = 320$ ), but not in sedentary time ( $R = -0.063$ ,  $p = 0.290$ ,  $n = 281$ ).

Significant linear trends, whereby positive changes in outcomes were associated with increased session attendance, were identified for unbouted MVPA (Figure 6; difference between low and high attenders 9.8 minutes per day,  $p = 0.002$  for the linear trend,  $n = 265$ ), PASE walking outside (see Appendix 2, Figure 32) (difference between low and high attenders: 1.24 hours per week,  $p = 0.011$ ,  $n = 302$ ) and MSEQ adherence score (see Appendix 2, Figure 33) (difference between low and high attenders 0.82

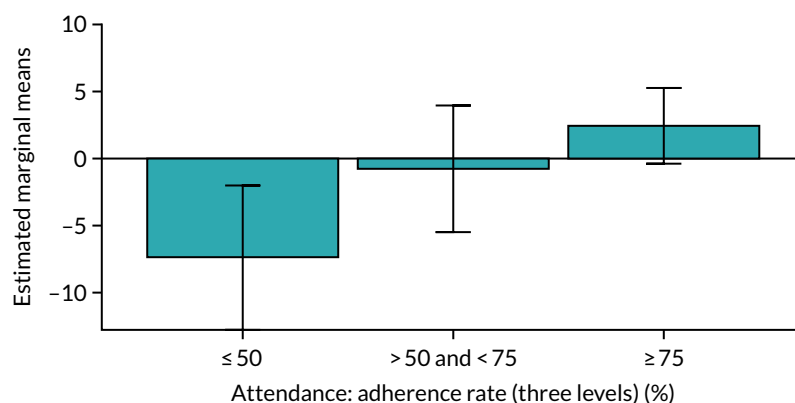


FIGURE 6 Changes in daily unbouted MVPA from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI).

points,  $p = 0.049$ ,  $n = 293$ ). Linear trends that were in the predicted direction were observed for 0- to 6-month changes in sedentary time (difference between low and high attenders 28.7 minutes per day,  $p = 0.056$ ,  $n = 255$ ) and PASE total score (difference between low and high attenders 17.8 points,  $p = 0.051$ ,  $n = 308$ ).

### Hypothesis 3

Hypothesis 3 was that increased exposure to the intervention would correlate with increased change in physical activity (MVPA, steps, sedentary time), engagement in muscle-strengthening exercise and SPPB score from 0 to 12 months.

Hypothesis 3 was broadly supported, although not with respect to accelerometer-assessed sedentary time and with two out of the five measures achieving only marginal significance in the linear trend analysis. There were significant, weak correlations between intervention attendance and changes from 0 to 12 months in unbouted MVPA ( $R = 0.174$ ,  $p = 0.004$ ,  $n = 278$ ), PASE total score ( $R = 0.134$ ,  $p = 0.015$ ,  $n = 329$ ), PASE walking score ( $R = 0.158$ ,  $p = 0.004$ ,  $n = 321$ ), MSEQ adherence score ( $R = 0.252$ ,  $p < 0.001$ ,  $n = 315$ ) and SPPB score ( $R = 0.175$ ,  $p = 0.001$ ,  $n = 346$ ), but not in sedentary time ( $R = -0.076$ ,  $p = 0.215$ ,  $n = 268$ ).

Significant linear trends, whereby positive changes in outcomes were associated with increased session attendance, were identified for PASE total score (see *Appendix 2, Figure 35*) (difference between low and high attenders 17.9 points,  $p = 0.043$  for the linear trend,  $n = 301$ ), MSEQ adherence score (*Figure 7*) (difference between low and high attenders 1.28 points,  $p = 0.001$ ,  $n = 292$ ) and SPPB (*Figure 8*) (difference between low and high attenders 0.89 points,  $p = 0.004$ ,  $n = 316$ ). A linear trend that was

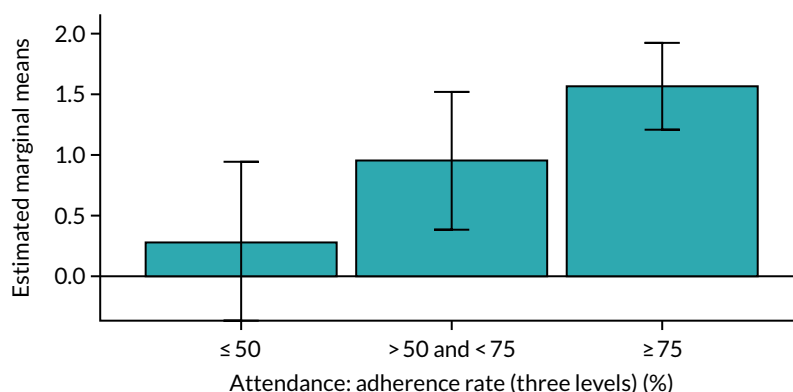


FIGURE 7 Changes in muscle-strengthening exercise (MSEQ adherence score) from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI).

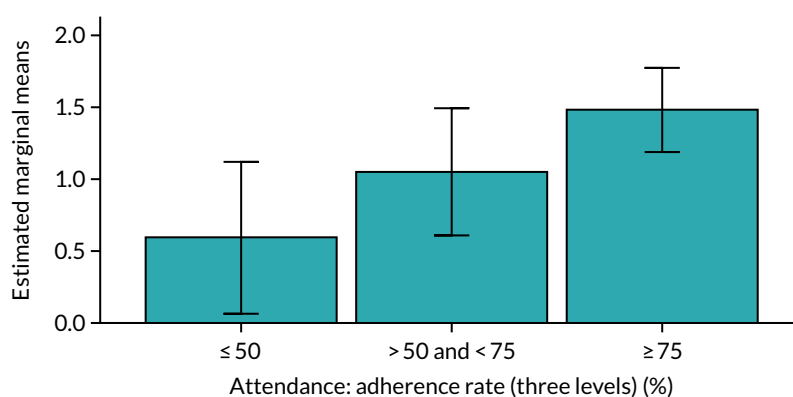


FIGURE 8 Changes in SPPB total score from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI).



in the predicted direction was observed for 0- to 12-month changes in unbouted MVPA (difference between low and high attenders 5.1 minutes per day,  $p = 0.051$ ,  $n = 256$ ) and for the PASE walking outside score (difference between low and high attenders 0.805 hours per week,  $p = 0.083$ ,  $n = 295$ ). No significant trend was observed for sedentary time.

#### **Hypothesis 4**

Hypothesis 4 was that increased exposure to the intervention would lead to increased maintenance of physical activity and engagement in exercise. Therefore, within the intervention arm, intervention dose would correlate negatively with decreases in physical activity, engagement in muscle-strengthening exercise and SPPB score from 6 to 12 months (during the supported maintenance period) and from 12 to 24 months (the unsupported maintenance period).

Hypothesis 4 was not supported by the data (in 13 out of 14 analyses). These analyses were applied in the intervention arm only. No significant correlations were found between percentage session attendance and changes from 6 to 12 months in bouts MVPA, unbouted MVPA, sedentary time, PASE total score, PASE walking score, MSEQ adherence score or SPPB score. There was a weak negative correlation between percentage session attendance and changes in bouts MVPA from 12 to 24 months ( $R = -0.125$ ,  $p = 0.047$ ,  $n = 252$ ). No other changes in outcomes from 12 to 24 months were significantly correlated with session attendance. No significant linear relationships were found for any of the outcomes across the three attendance groups, controlling for age, sex, study site and baseline SPPB score.

#### **Hypothesis 5**

Hypothesis 5 was that exposure to the intervention would lead to changes in key psychosocial determinants of physical activity and exercise from baseline to 6 and 12 months.

**Hypothesis 5a** Hypothesis 5a was that, compared with controls, the intervention arm will experience increases in autonomy and competence (self-efficacy); relatedness for MVPA, exercise and perceived intrinsic benefits of exercise (social, physical and emotional); and enjoyment of physical activity and exercise from 0 to 6 months.

Hypothesis 5a was broadly supported: we saw an intervention effect on most psychosocial determinants specified by the REACT logic model at 6 months. However, no signal was detected in relation to autonomy (for either muscle-strengthening or MVPA) or for enjoyment of muscle-strengthening exercise.

In relation to MVPA, at the 6-month follow-up, there was a significantly greater increase in the intervention arm compared with the control arm for perceived competence (mean difference 0.343 points, 95% CI 0.148 to 0.539 points,  $p = 0.001$ ,  $n = 627$ ), relatedness (mean difference 0.278 points, 95% CI 0.138 to 0.419 points,  $p < 0.001$ ,  $n = 610$ ) and enjoyment of MVPA (mean difference 0.149 points, 95% CI 0.015 to 0.282,  $p = 0.030$ ,  $n = 627$ ). However, there was no significant difference in autonomy related to MVPA (mean difference 0.065 points, 95% CI -0.040 to 0.170 points,  $p = 0.223$ ,  $n = 625$ ).

In relation to muscle-strengthening exercise, at the 6-month follow-up, there was a significant increase in the intervention arm compared with the control arm for perceived competence (mean difference 0.705 points, 95% CI 0.490 to 0.921 points,  $p < 0.001$ ,  $n = 625$ ) and relatedness (mean difference 0.283 points, 95% CI 0.040 to 0.525 points,  $p = 0.023$ ,  $n = 254$ ). The perceived intrinsic benefits of exercise were higher in the intervention arm than in the control arm at 6 months (mean difference 0.862, 95% CI 0.672 to 1.053,  $p < 0.001$ ,  $n = 609$ ). However, there were no significant differences in enjoyment (mean difference 0.185 points, 95% CI -0.037 to 0.407 points,  $p = 0.103$ ,  $n = 261$ ) or in autonomy related to muscle strengthening (mean difference -0.042 points, 95% CI -0.199 to 0.115 points,  $p = 0.597$ ,  $n = 257$ ).

**Hypothesis 5b** Hypothesis 5b was that, compared with control subjects, the intervention arm will experience increases in physical activity-related autonomy and competence (self-efficacy); relatedness for physical activity, exercise and perceived intrinsic benefits of exercise (social, physical and emotional); and enjoyment of physical activity and exercise from 0 to 12 months.



Hypothesis 5b was broadly supported: we saw an intervention effect on most psychosocial determinants specified by the REACT logic model at 12 months. However, no signal was detected in relation to autonomy (for either muscle-strengthening or MVPA) or enjoyment of MVPA.

In relation to MVPA, at the 12-month follow-up, there was a significant increase in the intervention arm compared with the control arm for perceived competence (mean difference 0.279 points, 95% CI 0.073 to 0.486 points,  $p = 0.008$ ,  $n = 615$ ) and relatedness (mean difference 0.211 points, 95% CI 0.064 to 0.357 points,  $p = 0.005$ ,  $n = 594$ ). However, there was no significant difference in autonomy related to MVPA (mean difference  $-0.052$  points, 95% CI  $-0.064$  to 0.168 points,  $p = 0.378$ ,  $n = 613$ ) or in enjoyment of MVPA (mean difference 0.083 points, 95% CI  $-0.051$  to 0.218 points,  $p = 0.225$ ,  $n = 612$ ).

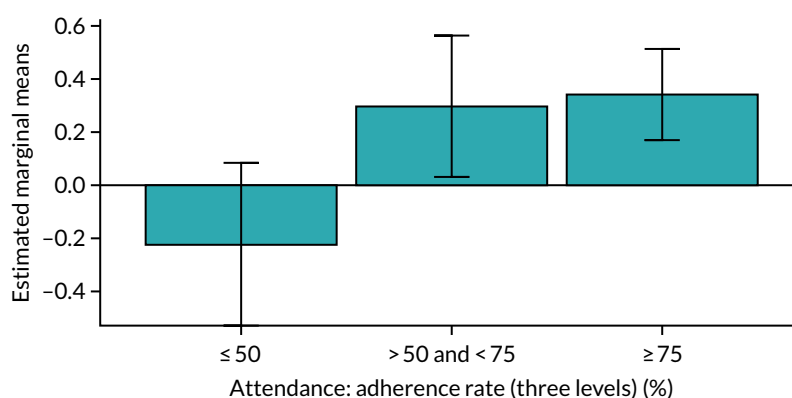
In relation to muscle-strengthening exercise, at the 12-month follow-up, there was a significant increase in the intervention arm compared with the control arm for perceived competence (mean difference 0.625 points, 95% CI 0.403 to 0.847 points,  $p < 0.001$ ,  $n = 609$ ), relatedness (mean difference 0.438 points, 95% CI 0.196 to 0.679 points,  $p < 0.001$ ,  $n = 265$ ) and enjoyment (mean difference 0.250 points, 95% CI 0.023 to 0.476 points,  $p = 0.031$ ,  $n = 268$ ). The perceived intrinsic benefits of exercise were also higher in the intervention arm than in the control arm at 12 months (mean difference 0.757, 95% CI 0.565 to 0.950,  $p < 0.001$ ,  $n = 598$ ). However, there were no significant differences for autonomy related to muscle-strengthening (mean difference  $-0.023$  points, 95% CI  $-0.174$  to 0.128 points,  $p = 0.769$ ,  $n = 268$ ).

**Hypothesis 5c** Hypothesis 5c was that increased exposure to the intervention (total contact time from baseline to the relevant time point) will correlate with increased change in the above determinants (and in the expected direction).

Hypothesis 5c was largely supported in relation to muscle-strengthening exercise but not in relation to MVPA.

In relation to MVPA, at the 12-month follow-up, we identified one significant linear trend, whereby an increase in MVPA relatedness was associated with increased session attendance (*Figure 9*) (difference between low and high attenders 0.564 points in the predicted direction,  $p = 0.002$  for the linear trend,  $n = 295$ ). No significant trend was observed for MVPA autonomy, MVPA competence or MVPA enjoyment.

In relation to muscle-strengthening exercise, at the 12-month follow-up, we identified significant linear trends whereby positive changes in psychosocial determinants were associated with increased session attendance for muscle-strengthening competence (*Figure 10*) (difference between low and high attenders 0.633 points in the predicted direction,  $p = 0.021$ ,  $n = 298$ ), relatedness (*Figure 11*) (difference between



**FIGURE 9** Changes in relatedness in relation to MVPA from 0 to 12 months for different levels of attendance in the intervention arm (error bars: 95% CI).

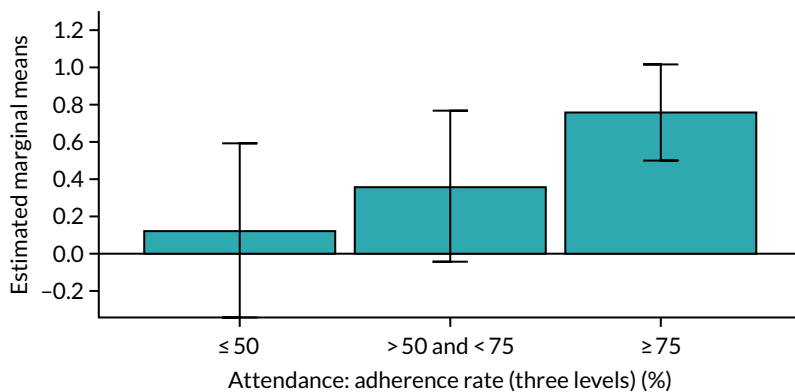


FIGURE 10 Changes in competence in relation to muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI).

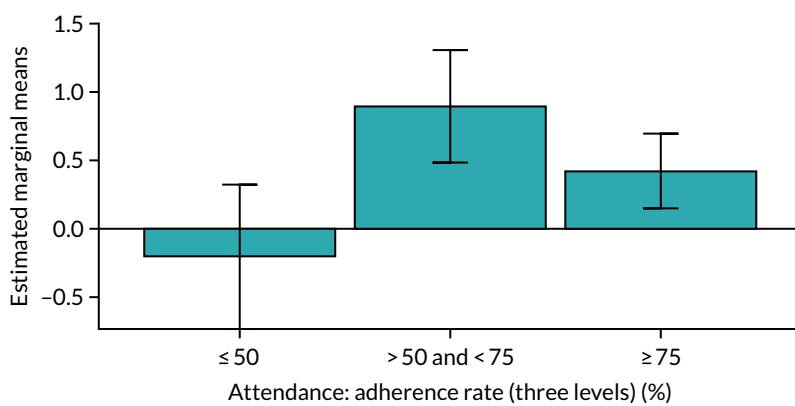


FIGURE 11 Changes in relatedness for muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI).

low and high attenders 0.619 points in the predicted direction,  $p = 0.041$  for the linear trend,  $n = 149$ ), enjoyment (Figure 12) (difference between low and high attenders 0.745 points in the predicted direction,  $p = 0.014$ ,  $n = 153$ ) and perceived benefits of muscle-strengthening exercise (Figure 13) (difference between low and high attenders 1.142 points in the predicted direction,  $p < 0.001$ ,  $n = 291$ ). No significant trend was observed for autonomy for muscle-strengthening exercise.

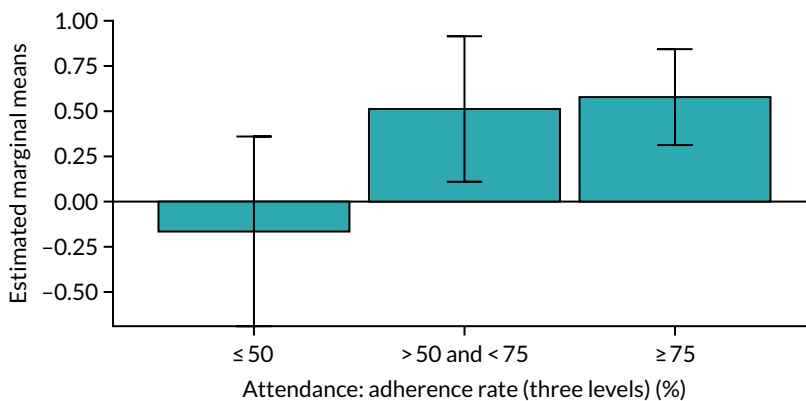


FIGURE 12 Changes in enjoyment in relation to muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI).

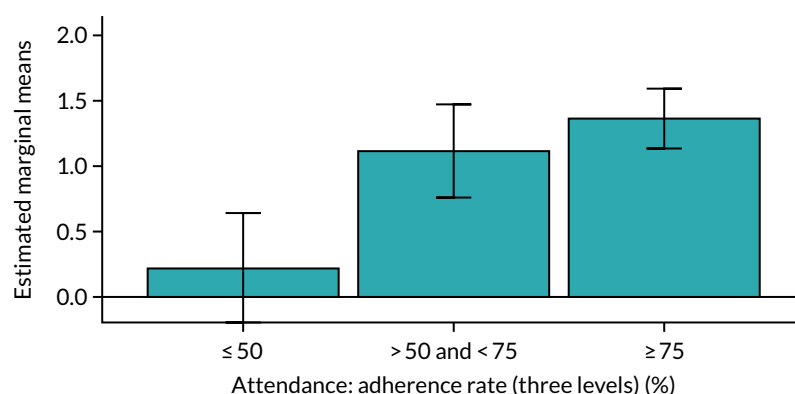


FIGURE 13 Changes in perceived benefits of muscle-strengthening exercise from 0 to 12 months by level of attendance in the intervention arm (error bars: 95% CI).

### Hypotheses 6 and 7 mediation and moderation of intervention effects on lower limb physical function

#### Hypothesis 6

Hypothesis 6 was that the intervention effect on SPPB (intervention vs. control) will be mediated by changes in muscle-strengthening exercise and by changes in MVPA, lower-intensity physical activity or walking activity.

Lower-intensity physical activity was assessed using our accelerometer measures of sedentary time (also called 'very low physical activity').

Hypothesis 6 was largely supported by the data, with the proviso that accelerometer-based measures did not follow the same pattern as the self-reported measures. The intervention effect on SPPB score was partly or completely mediated by changes in self-reported muscle-strengthening exercise and self-reported changes in physical activity (both total activity and walking) between 0 and 12 months and between 0 and 24 months. No accelerometer measures of physical activity significantly mediated the intervention effect on SPPB score at any time point. The standardised mediation effects for muscle-strengthening exercise were generally higher than those for physical activity, and walking seemed to account for the majority of the mediating effect identified by the PASE questionnaire (Table 11).

TABLE 11 Summary of mediation of intervention effects on SPPB by physical activity variables

Mediating variable	Standardised coefficient (95% CI); n		
	6 months	12 months	24 months
Muscle-strengthening (MSEQ)	-	0.051 (0.018 to 0.092); 579	0.053 (0.020 to 0.094); 569
Total activity (PASE)	-	0.025 (0.004 to 0.054); 616	0.039 (0.008 to 0.075); 598
Walking (PASE)	-	0.017 (0.000 to 0.041) 598	0.029 (0.004 to 0.059); 590
MVPA (accelerometer)	-	-	-
VLPA (accelerometer)	-	-	-

VLPA, very low physical activity.

**Note**  
 Orange shading represents no significant mediation. Light-purple shading represents partial mediation. Aqua shading represents total mediation.

### Hypothesis 7

Hypothesis 7 was that the intervention effect on SPPB score, physical activity and exercise from 0 to 6 months and from 0 to 24 months may be moderated by a number of variables, including age, sex, ethnicity, baseline physical activity and SPPB score, co-interventions, comorbidities, BMI, mental health, SES and education level.

In the main trial analysis (see *Chapter 3* and *Appendix 2, Figure 34*), intervention effects were found for PASE total score and muscle-strengthening exercise (MSEQ adherence score), as well as the primary outcome (SPPB score) at both 6 and 24 months. *Report Supplementary Material 2* shows the moderation effects of all of the variables specified in the hypothesis of these three variables. The text below gives an overview of these data, highlighting any significant associations.

Intervention effects on SPPB, physical activity and engagement in muscle-strengthening exercise (MSEQ) were not significantly moderated by age, sex, baseline SPPB score, co-morbidity (number of chronic physical conditions at baseline) or BMI. However, the following significant interactions were observed (*Table 12*).

**TABLE 12** Summary of moderation of intervention effects from 0 to 12 months and from 0 to 24 months on MSEQ, physical activity (PASE total) and SPPB by demographic variables

Moderating Variable	SPPB		Muscle-strengthening (MSEQ)		Total activity (PASE)	
	6 months	24 months	6 months	24 months	6 months	24 months
Age						
Sex						
Ethnicity	-1.426, SE 0.714; $p = 0.048$				-69.519, SE 24.500; $p = 0.005$	
SES (deprivation quintile)	$R^2$ change 0.022, $F = 3.818$ (df = 4, 649); $p = 0.0045$		$R^2$ change 0.017, $F = 2.729$ (df = 4, 587); $p = 0.029$		$R^2$ change 0.017, $F = 2.715$ (df = 4, 604); $p = 0.029$	
Baseline physical activity					0.188, SE 0.079, $p = 0.018$	
Baseline SPPB						
Co-interventions					26.741, SE 11.664, $p = 0.022$	
Comorbidities						
BMI						
Mental health			-2.335, SE 0.972; $p = 0.017$	-2.450, SE 1.057; $p = 0.021$		
Education level				$R^2$ change 0.016, $F = 2.447$ (df = 4, 572); $p = 0.045$		

SE, standard error.

Figures are standardised coefficients for binary variables and  $R^2$  change for adding the trial arm  $\times$  interaction term for variables with more than two categories. Blank cells represent no significant moderation.

At 6 months, the estimated effects of the intervention on SPPB score and the PASE total score were higher for ethnic minority participants than for white participants [1.43 points higher for SPPB score (beta 1.43;  $p = 0.048$ ); 69.52 points higher for PASE total score (beta 69.52;  $p = 0.005$ )]. The interactions are visualised in *Report Supplementary Material 2*. These results need to be treated with considerable caution owing to (1) multiple testing (the Bonferroni adjusted  $p$ -value for multiple testing would be 0.0045) and (2) unbalanced group sizes (31 ethnic minority participants vs. 628 white participants for SPPB score, and 25 ethnic minority participants vs. 625 white participants for PASE total score).

Socioeconomic status (deprivation quintile) significantly moderated the effects of the intervention on SPPB score ( $p = 0.0045$ ) and MSEQ score at 6 months ( $p = 0.029$ ), and on PASE total score at 24 months ( $p = 0.029$ ) (see *Table 12*). The interactions are visualised in *Report Supplementary Material 2*. For SPPB score at 6 months, the intervention worked best in the highest deprivation group (estimated mean difference 1.78 points, 95% CI 0.88 to 2.67 points,  $n = 70$ ) and worst in the lowest deprivation group (estimated mean difference -0.13 points, 95% CI -0.67 to 0.41,  $n = 191$ ), with a general pattern of increasing effect size as area deprivation decreased. However, there was no clear pattern of effect for MSEQ score or physical activity. These results need to be treated with considerable caution owing to (1) multiple testing (the Bonferroni adjusted  $p$ -value for multiple testing would be 0.0045) and (2) unbalanced group sizes (ranging from 63 to 171).

Physical activity level at baseline significantly moderated the effects of the intervention on physical activity at 6 months (beta 0.188;  $p = 0.018$ ), with intervention effects being greater in those who were more physically active at baseline. The interaction is visualised in *Report Supplementary Material 2*. This result needs to be treated with caution owing to multiple testing (the adjusted  $p$ -value would be 0.0045).

The intervention effect from 0 to 24 months on self-reported physical activity (PASE total score) was significantly moderated by engagement in co-interventions during the 24 months of the study (beta 26.741;  $p = 0.022$ ). The intervention effect on physical activity for people who engaged in additional physical activity interventions during the course of the study was 26.741 PASE points higher than for the subgroup who did not engage in such interventions. This result needs to be treated with caution owing to multiple testing (the adjusted  $p$ -value would be 0.0045).

Mental health status at baseline significantly moderated the effects of the intervention on MSEQ score at both 6 months (beta -2.335;  $p = 0.017$ ) and 24 months (beta -2.450;  $p = 0.021$ ), with intervention effects being greater in those who had experienced no diagnosed mental health condition in the last 6 months. The interactions are visualised in *Report Supplementary Material 2*, which shows that the effects in both cases are largely driven by large increases over time in MSEQ adherence for people with mental health conditions in the control arm. This may reflect a regression to the mean effect in participants who had a transient mental health condition at baseline (they may have been less engaged in exercise at baseline), or it may reflect some kind of responding bias (i.e. a tendency for people with a mental health condition to under-report activity at baseline). This result needs to be treated with considerable caution owing to (1) multiple testing (the adjusted  $p$ -value would be 0.0045) and (2) unbalanced group sizes (29 participants with a recent mental health condition at baseline vs. 567 participants with none).

The intervention effect from 0 to 24 months on MSEQ adherence score was also significantly moderated by education level. The explanatory value of adding the arm  $\times$  education interaction term to the regression model ( $R^2$  change) was 0.016 ( $p = 0.045$ ). The interaction was largely driven by an unexpected increase in MSEQ score from 0 to 24 months in the least educated control arm participants. This result needs to be treated with caution owing to multiple testing (adjusting for this would require a  $p$ -value of 0.0045 to denote significance).

## Discussion

### Summary of findings

Testing of the hypotheses supported many elements of the logic model for the REACT intervention (see *Figure 1*). Changes in physical activity in terms of muscle-strengthening exercise, walking and overall physical activity were observed during the initial intervention period (0–6 months). In the intervention arm, these changes, as well as changes in physical activity and SPPB score from 0 to 1 month, were linearly related to the number of intervention sessions attended. No significant changes in sedentary time were observed, however, and maintenance of physical activity from 6 to 12 months and from 12 to 24 months was not associated with intervention exposure.

The intervention generated significant changes in most of the proposed mechanisms of behaviour change at both 6 and 12 months, including (1) changes in perceived competence, relatedness and enjoyment (at 6 months only) relating to MVPA; and (2) changes in perceived competence, relatedness, enjoyment (at 12 months only) and intrinsic benefits relating to muscle-strengthening exercise. However, no signal was detected in relation to perceived autonomy for either muscle-strengthening or MVPA. Changes in the psychosocial determinants of muscle-strengthening exercise were linearly related to session attendance (for competence, relatedness, enjoyment and perceived benefits, but not autonomy). Changes in the psychosocial determinants of MVPA were not related to intervention exposure, except for relatedness.

Mediation analyses showed that the intervention effects on SPPB at 12 and 24 months were mediated by self-reported engagement in muscle-strengthening exercise (MSEQ adherence score). These effects were also mediated by self-reported measures of physical activity (PASE total and PASE walking), but not by accelerometer-assessed physical activity.

Moderation analyses found that the short-term (0–6 months) intervention effects on SPPB were moderated by ethnicity and area deprivation (with ethnic minorities and more-deprived population groups benefiting more from the intervention than white and less-deprived populations). Longer-term (0–24 months) intervention effects on SPPB were not moderated by any of the variables tested, indicating that the intervention worked equally well for participants irrespective of their age, sex, ethnicity, baseline fitness (as indicated by baseline physical activity) and physical functioning, co-interventions, physical and psychological comorbidities, BMI, SES and education level. All of the above are consistent with the idea that the intervention worked through increases in resistance/muscle-strengthening exercise, which in turn was associated with session attendance and changes in the theorised psychosocial variables, as specified in the REACT logic model (with the exception of perceived autonomy).

### Relation to other literature and interpretation

These findings add to the body of literature on the processes of behaviour change for physical activity interventions, particularly in relation to older people. A recent systematic review of the literature on process evaluation in trials of physical activity interventions, conducted as part of the PhD associated with the REACT project,<sup>71</sup> found that, since publication of the MRC guidance on process evaluation in 2014,<sup>66</sup> only 21% of trials (24 papers) quantitatively evaluated mechanisms of behaviour change and only 15% of trials (17 papers) evaluated the moderating effects of context on outcomes. Studies of such processes in older people and in relation to longer-term intervention effects (beyond 12 months) are even scarcer.

The findings on the psychosocial process of behaviour change are consistent with, and help to validate, the REACT logic model and its underpinning theories, in particular SDT<sup>41</sup> and SCT.<sup>98</sup> However, it is worth noting that the SDT process of increasing autonomy did not seem to be associated with increases in either muscle-strengthening exercise or overall physical activity. This may reflect the structured, centre-based nature of the intervention, including formal exercise sessions (the only element of choice being whether or not to attend), although we did try to promote autonomy and

choice in the social education sessions around making plans for physical activity and for continuation of exercise outside the REACT context (and our intervention fidelity analysis indicated that this element was adequately delivered). Nevertheless, the data suggest that promoting autonomy/choice may not be necessary for promoting physical activity, and particularly muscle-strengthening exercise, in older people. This may reflect a preference among some older people for more structured/guided interventions that are enjoyable and socially engaging (and that deliver noticeable physical, social and mental benefits), rather than for individual choice-based activities. In theory, promoting autonomy should be important for longer-term maintenance of exercise behaviours, particularly when the REACT intervention comes to an end. However, this did not seem to be substantiated by the data in this case.

Alternatively, the data on autonomy processes may reflect a measurement error: the items that we used to assess autonomy in relation to physical activity and exercise (i.e. 'I feel free to make my own decisions about muscle-strengthening exercises'; 'I feel like I am in charge of how often I do muscle-strengthening exercises') were taken from a validated scale and had good internal reliability (Cronbach's  $\alpha = 0.71$  for the two-item physical activity autonomy scale and  $\alpha = 0.76$  for the two-item muscle-strengthening exercise autonomy scale). However, the scores were skewed towards the high end of the scale (mean scores 8.1 and 7.7, respectively, on a 0 to 10 scale), suggesting a possible ceiling effect for this measure. The items selected were also part of a longer scale and, therefore, may not function well in isolation.

Importantly, the moderation analyses found that the intervention appeared to perform well in relation to health inequalities. People from ethnic minority backgrounds and areas of high deprivation responded as well to the intervention (and seemingly better at 6 months) as white participants and people from areas of low deprivation. Furthermore, no differences in intervention effect were noted by sex, baseline physical functioning or other demographic characteristics. It is worth noting that the lack of any quantitative difference between people with high or people with low SPPB score at baseline contrasts with the qualitative data reported in *Chapter 4, Study 2: qualitative evaluation of intervention processes*. This may reflect non-generalisation of the reported experiences from the relatively small qualitative sample, or it may reflect a difference between subjectively perceived and objectively measured changes in physical functioning. In particular, the perceptions of benefit for those with lower SPPB score at baseline may have been diminished by comparison with their peers who ended the intervention having achieved (and also started with) better physical functioning. Negative effects of social comparisons involving 'upwards contrasts' on health perceptions have been noted in other literature.<sup>99</sup>

In contrast to the increasingly dominant paradigm favouring objective measurement of physical activity, the findings here suggest that self-reported measures of physical activity may be more valid indicators of engagement in the physical activities targeted by exercise interventions than accelerometer measures.<sup>100,101</sup> In this study, we found that overall physical activity and walking activity, as measured by the PASE questionnaire, and muscle-strengthening activity, as measured by the two-item MSEQ adherence score that we developed specifically for this study, were sensitive to change (producing highly significant differences between groups) and linearly associated with increasing intervention attendance. They, therefore, seem to be valid indicators of participant engagement in the activities targeted by the REACT intervention. In this study, we are assured that physical activity, especially muscle-strengthening exercise, changed as a result of the intervention because this is the only way that physical functioning (as assessed using the SPPB) could have been modified at 12 and 24 months. However, the self-reported measures were better able than the accelerometers to detect these changes. Other literature acknowledges the potentially complementary roles of questionnaires and accelerometers for assessing physical activity in older people.<sup>102</sup>

### Strengths and limitations

The REACT process evaluation benefited from having access to a large data set with good levels of retention and completion for most measures. This offered good statistical power for most analyses. However, some of the moderation analyses (e.g. the analyses including ethnicity and mental health



conditions) need to be treated with caution owing to multiple testing (11 analyses per time point and per variable), small/unbalanced group sizes (e.g. only around 30 participants in the ethnic minority group) and the inability to distinguish between different ethnic minority subgroups (owing to the small numbers available for analysis).

### **Implications for practice/future research directions**

The process evaluation identified ways in which both the theoretical basis and the delivery of the intervention could be further improved.

The data from the quantitative process evaluation broadly support the theoretical basis for behaviour change in the REACT logic model. However, they also imply that some intended processes (notably promoting autonomy) were not necessary for change to occur. Further research is needed to explore the role of autonomy in the REACT intervention and whether or not this construct should be retained in the logic model.

The suggestions of individual-level contextual influences on intervention effectiveness in the logic model (relating to age, sex, etc.) were not supported. However, further research is needed to examine the associations of systemic influences and facilitator-level factors with intervention effectiveness.

Hypotheses about theoretical processes associated with the maintenance of behaviour change (e.g. autonomy, relatedness and enjoyment of physical activity) were not supported and need further investigation. This might involve analyses focused on maintenance of changes in those participants who made initial changes (context-specific analyses), rather than looking for such associations in the whole sample. Further research is also needed to explore the role of perceived tension following changes in physical activity in relation to maintenance of behaviour changes. This element of the logic model has not yet been tested owing to time and resource limitations.





## Chapter 5 Health economic evaluation

Parts of this chapter have been adapted from Stathi *et al.*<sup>1</sup> Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The text includes minor additions and formatting changes to the original text.

Parts of this chapter have been adapted from Snowsill *et al.*<sup>103</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

The economic evaluation assesses whether or not the community-based physical activity and behavioural maintenance intervention (REACT) is cost-effective for reducing the progression of functional limitation in retired or semi-retired men and women aged  $\geq 65$  years who are at risk of mobility-related disability compared with a minimal intervention control group. The perspective used was the NHS and Personal Social Services (PSS), which was complemented with out-of-pocket costs expenditure. All health-care resources used by participants from baseline to 24 months after randomisation were included. The additional (incremental) costs associated with delivering the REACT intervention were estimated using the resource data collected within the trial. Incremental cost-effectiveness ratios (ICERs) were estimated against the primary outcome measure (SPPB) and also for cost per QALYs gained. In addition, the cost-effectiveness of the REACT programme beyond the trial-end time horizon (24 months) up to death was estimated.

### Within-trial economic evaluation

#### *Aims and objectives*

To assess if the REACT programme (plus usual care) was cost-effective compared with usual care alone in retired or semi-retired adults at risk of mobility-related disability from the perspective of the NHS and PSS.

#### *Methods*

##### **Target population and subgroups**

Participants were those eligible for the REACT study, that is adults aged  $\geq 65$  years, retired or semi-retired, with a SPPB score between 4 and 9 (inclusive).

Subgroup analyses were conducted as follows:

- baseline SPPB score – 4–7 compared with 8–9
- SES by beyond secondary school compared with not beyond secondary school and by IMD decile – divided to achieve a split of the trial participants as close as possible to 50 : 50
- age – 65–74 compared with  $\geq 75$  years
- known medical conditions – 0–1 compared with  $\geq 2$
- sex – women compared with men
- site – Devon compared with Bath/Bristol compared with Birmingham
- number of falls in the 6 months before randomisation – 0 compared with  $\geq 1$ .

**Setting and location**

The REACT intervention was delivered in Bath, Bristol, Devon and Birmingham.

**Perspective**

The base-case analysis adopts an NHS and a PSS perspective, in line with the NICE reference case.<sup>104</sup> A scenario analysis considers the costs from a societal perspective, in which the voluntary and out-of-pocket expenditure of long-term care is included, along with an estimate of the value of informal care provided by friends and relatives, and an estimate of private expenditure on health care.

**Intervention and comparators**

The REACT programme is a group-based physical activity and behavioural maintenance intervention that aims to prevent decline in those at risk of mobility-related disability. The programme runs for 12 months, with sessions initially twice weekly and then once weekly, for a total of 64 sessions. Sessions are led by a fitness instructor. REACT participants also receive support as usual from NHS services.

**Time horizon**

The time horizon for the within-trial economic evaluation is 2 years.

**Measurement of effectiveness**

The SPPB score was directly assessed by study investigators at 24 months.

**Measurement and valuation of preference-based outcomes**

Health-related quality of life was measured at baseline and at 6, 12 and 24 months using the EQ-5D-5L and the SF-36 (self-completed in face-to-face sessions).

The EQ-5D-5L covers five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) measured at five different levels.<sup>51</sup> EQ-5D-5L health states were valued through mapping from EQ-5D-5L to EuroQol-5 Dimensions, three-level version (EQ-5D-3L), utilities using the crosswalk method.<sup>105-107</sup> SF-36 responses were converted to Short Form questionnaire-6 Dimensions (SF-6D) health states. The SF-6D has six dimensions (physical functioning, role limitation, social functioning, bodily pain, mental health and vitality), with four to six levels per dimension.<sup>108</sup> Within-trial QALYs were estimated from EQ-5D utilities<sup>109</sup> for the base-case and SF-6D utilities in a scenario analysis.

**Estimation of costs**

There are two cost components:

1. costs of delivering the REACT programme
2. costs associated with the use of health-care and social care services.

Data on the use of NHS and PSS resources were collected at the face-to-face sessions using a resource use questionnaire, which asked participants to recall their usage over the previous 6 months at baseline and at 6, 12 and 24 months.

**Valuation of resources**

Valuation of resources used the Personal Social Services Research Unit (PSSRU) *Unit Costs of Health and Social Care*<sup>110</sup> and NHS National Cost Collection (NCC)<sup>111</sup> (Table 13). Further details are presented in Appendix 3.

Total within-trial costs were estimated by interpolation, assuming that utilisation between months 12 and 18 was midway between utilisation from months 6 to 12 and utilisation from months 18 to 24.

**Currency, price date and conversion**

Great British pounds (2018/19) were used, and purchasing power parity exchange rates were used when converting currencies, as well as appropriate inflation rates for adjusting to the price year 2018/19.

TABLE 13 Unit costs of health and social care resources

Item	Unit cost <sup>a</sup>	Source
<b>Primary and/or community-based services (per contact or consultation)</b>		
GP at surgery/health centre	39.19	PSSRU 2019 <sup>110</sup> (surgery consultation lasting 9.22 minutes, including direct care staff costs, including training)
GP via telephone	15.52	PSSRU 2019 <sup>110</sup>
GP home visit	78.92	PSSRU 2014 <sup>112</sup> and PSSRU 2019 <sup>110</sup> [£156 per hour of GMS activity, 11.4-minute visit, 1 : 0.61 direct-to-indirect time ratio (7.0 minutes indirect), 12 minutes of travel]
Practice nurse at surgery/health centre	12.43	PSSRU 2014 <sup>112</sup> and PSSRU 2019 <sup>110</sup> (£37 per hour, 15.5 minutes per contact, 1 : 0.30 direct-to-indirect time ratio)
Practice nurse via telephone	7.80	PSSRU 2019 <sup>110</sup>
Practice nurse at home	39.68	NCC 2018/19 <sup>111</sup> (N02AF: district nurse, face to face)
Physiotherapist	62.90	NCC 2018/19 <sup>111</sup> (A08A1)
Occupational therapist	83.17	NCC 2018/19 <sup>111</sup> (A06A1)
Dietitian	89.90	NCC 2018/19 <sup>111</sup> (A03)
Podiatrist	42.51	NCC 2018/19 <sup>111</sup> (A09 A)
Counsellor	45.00	PSSRU 2019 <sup>110</sup> (Agenda for Change band 6 for 1 hour)
Walk-in centre	21.00	PSSRU 2019 <sup>110</sup> (assume 15 minutes of Band 6 nurse time)
<b>Day care services</b>		
Per day care centre	40.20 (60.00)	PSSRU 2019 <sup>110</sup>
<b>Overnight hospital stays (per attendance)</b>		
General medical ward or long-stay ward	385.16 plus 503.96 per day	NHS reference costs 2017/18 <sup>113</sup>
Intensive care unit	3532.09	NCC 2018/19 <sup>111</sup>
<b>Other hospital usage</b>		
Outpatient appointment	126.85	NCC 2018/19 <sup>111</sup> (weighted average of all consultant-led and non-consultant-led outpatient attendances)
Day case treatment	751.90	NCC 2018/19 <sup>111</sup> (weighted average of all day-case episodes)
A&E attendance	166.05	NCC 2018/19 <sup>111</sup> (weighted average of all A&E episodes)
<b>Convalescent or nursing home</b>		
Convalescent or nursing home	79.73 (119.00)	PSSRU 2019 <sup>110</sup>
<b>Support from relatives and/or friends</b>		
Help with tasks at home (per hour)	0.00 (11.06)	van den Berg and Ferrer-I-Carbonell (2007); <sup>114</sup> OECD; <sup>115</sup> and Office for National Statistics <sup>116</sup> [€9.65 per hour (Netherlands, 2002) at 0.690 : 0.901 conversion (OECD PPP 2002) and inflated according to nominal earnings growth 2002 (390.9) to 2019 (584.9)]
Stayed off work to help (per day)	0.00 (99.53)	Office for National Statistics <sup>116</sup>

continued

TABLE 13 Unit costs of health and social care resources (continued)

Item	Unit cost <sup>a</sup>	Source
<b>Other (provided as free text)</b>		
Chiropractor	0.00 (40.06)	Newell <i>et al.</i> (2016) <sup>117</sup>
Osteopath	0.00 (44.95)	General Osteopathic Council <sup>118</sup>
Paid care at home (per hour)	18.95 (28.29)	PSSRU 2019 <sup>110</sup>
A&E, accident and emergency.		
a Where a different value applies in the broader (societal) perspective, this is shown in parentheses.		

## Discounting

The discount rate for costs and QALYs is 3.5% per year.<sup>104</sup>

## Analytical methods

All analyses were conducted using the ITT principle.

## Aggregating costs of NHS/Personal Social Services resource use

Costs were aggregated for each time point using the following aggregate categories:

- primary care – GP (at surgery/health centre, via telephone, at home), practice nurse (at surgery/health centre, via telephone, at home), physiotherapist (at surgery/health centre, at home), occupational therapist (at surgery/health centre, at home), nutritionist, chiropractor, counsellor and walk-in centre
- day care – day care attendances
- hospital overnight stays – general medical ward, long-stay ward and intensive care unit
- other hospital attendances – outpatient attendances, day case procedures, and accident and emergency attendances
- convalescent/nursing home – days spent in convalescent/nursing home
- informal care – hours of help from friends/relatives around the home and days friends/relatives took off work (not included in base-case perspective)
- other (from free text) – chiropractor and osteopath (not included in base-case perspective), paid care at home.

## Statistical analysis

A generalised linear model with gamma family and log link was used to model cost data.<sup>119</sup> Within-trial QALYs could not exceed 2 (perfect health for 24 months), so a generalised linear model with gamma family and log link was used to model the difference of 2 minus QALYs and estimates then transformed to QALYs.

Costs and QALYs were adjusted for baseline differences by including the following as covariates:

- age at randomisation (grouped in 5-year age bands)
- sex
- site (Bath/Bristol, Birmingham, Devon)
- baseline SPPB score
- baseline health-care resource use
- baseline utility values<sup>120</sup>
- unadjusted analyses are used in a sensitivity analysis.

Primary analysis was carried out with imputation. Patterns of missingness among resource use and utility value data were explored graphically and through examining the correlation of indicator variables for missingness across follow-up time points. Logistic regression was used to explore the influence of key predictors individually (randomised allocation, sex, age, site and baseline SPPB), with *p*-values calculated from the likelihood ratio test. For details of the analysis of missing data, see *Appendix 3, Figure 36*.

Complete-case analysis (CCA) was used in a sensitivity analysis. Vital status was not recorded in the analysis data set, so any participants who died during the trial would appear as having missing data after they died. They would, therefore, have costs and QALYs imputed just as any participant missing data for other reasons.

Given that REACT is a group-based intervention, analyses producing cluster-robust standard errors were used throughout. In the control arm, each participant was treated as an individual cluster.

Uncertainty in results due to sampling variation was estimated either through cluster-robust Huber-White standard errors [multiple imputation (MI) analyses] or via bootstrapping (1000 iterations for CCAs, 100 iterations for MI analyses where used).

The sensitivity of the results to changing the values assigned to resources by  $\pm 20\%$  was estimated through one-way sensitivity analysis deviating from the CCA, presented in the form of a tornado diagram. Subgroup analyses were conducted by refitting models with interaction terms between the intervention allocation and the subgroup covariate, and producing predictive means for costs and QALYs. These were based on MI, but as these are exploratory we do not report standard errors or CIs. ICER 95% CIs are produced using the bootstrap percentile method.<sup>121</sup>

## Results

### Intervention costs

The cost of the REACT intervention was estimated to be £9465.69 per arm, or £621.83 per participant (see *Appendix 3, Table 35*). The cost was somewhat sensitive to underlying assumptions, with the cost per participant ranging from £554.54 (-11%) to £775.13 (+25%) across the sensitivity analyses (see *Appendix 3, Table 36*). The methods for estimating these costs are described in detail in *Appendix 3*.

### Resource use and utility values

A total of 367 and 410 participants were randomised to the control arm and the intervention arm, respectively. Resource use at baseline was generally well balanced, but there was some evidence of higher baseline resource use in the intervention arm (see *Appendix 3, Table 37*).

The average total value of NHS and PSS resource utilisation in the 6 months prior to randomisation was £653 and £797 in the control arm and the intervention arm, respectively. After multiple imputation and adjusting for covariates, the difference in pre-randomisation resource utilisation was £107 (95% CI -£50 to £263). Over the course of the REACT study, we estimate that NHS and PSS resource use was £724.74 lower in the intervention arm (95% CI £25.79 to £1423.69) (see *Appendix 3, Table 38*). There was generally a trend towards greater resource use over time, and the difference in costs was greatest in the first 6 months. Expanding to a societal perspective moved the expected cost difference towards zero (£645.65), meaning that the cost difference was no longer statistically significant (95% CI -£63.73 to £1355.02).

*Table 14* explores which cost categories lead to savings between the intervention arm and the control arm. In available case analyses and MI analyses, the category 'other hospital' (which includes day-case procedures, outpatients and A&E attendances) was the primary category in which intervention arm participants had lower costs. The difference between MI and CCAs appears to be driven by the difference in hospital overnight stay costs: the unadjusted difference in mean costs is £72.80 among available cases compared with -£78.62 with MI, and this difference may be exacerbated by adjustment.

TABLE 14 Breakdown of costs (£) across different categories (excluding intervention costs)

Cost category	Mean value (SD); n		Difference	
	Intervention arm	Control arm	Unadjusted	Adjusted (95% CI)
<b>Available case analysis</b>				
Primary care	820.49 (680.47); 299	838.12 (704.30); 260	-17.63	-14.40 (-129.17 to 100.37)
Hospital overnight stay	1104.98 (3670.8); 289	1032.18 (3282.4); 246	72.80	254.25 (-358.29 to 866.80)
Other hospital	1372.61 (1799.1); 299	1856.01 (3281.2); 263	-483.40	-484.43 (-931.70 to -37.16)
Day care	1.78 (13.53); 276	7.71 (95.31); 232	-5.93	-8.74 (-31.88 to 14.40)
Convalescent	27.18 (199.77); 299	24.66 (312.97); 263	2.52	26.74 (-7.47 to 60.95)
Other	26.16 (332.18); 410	10.66 (187.38); 366	15.49	<sup>b</sup>
Day care <sup>a</sup>	2.65 (20.20); 276	11.51 (142.26); 232	-8.85	-13.04 (-47.58 to 21.49)
Convalescent <sup>a</sup>	40.56 (298.17); 299	36.80 (467.12); 263	3.76	39.91 (-11.14 to 90.96)
Informal care <sup>a</sup>	169.93 (590.27); 253	128.99 (364.44); 217	40.94	46.69 (-40.82 to 134.19)
Other <sup>a</sup>	46.56 (501.27); 410	22.86 (285.68); 366	23.70	9.96 (-39.66 to 59.58)
<b>MI analysis</b>				
Primary care	848.02 (685.56); 410	853.80 (717.95); 367	-5.79	-25.36 (-127.81 to 77.08)
Hospital overnight stay	1064.98 (3440.2); 410	1143.60 (3492.6); 367	-78.62	-107.76 (-628.11 to 412.60)
Other hospital	1441.98 (2039.3); 410	1836.89 (3255.0); 367	-394.91	-462.03 (-861.98 to -62.07)
Day care	3.72 (37.73); 410	6.64 (82.25); 367	-2.92	-4.79 (-21.69 to 12.12)
Convalescent	69.34 (674.84); 410	35.57 (354.83); 367	33.77	16.02 (-87.99 to 120.04)
Other	26.16 (332.18); 410	10.64 (187.13); 367	15.52	<sup>b</sup>
Day care <sup>a</sup>	5.56 (56.32); 410	9.91 (122.76); 367	-4.35	-7.15 (-32.38 to 18.09)
Convalescent <sup>a</sup>	103.50 (1007.2); 410	53.09 (529.60); 367	50.40	23.92 (-131.33 to 179.17)
Informal care <sup>a</sup>	190.51 (589.62); 410	163.07 (451.56); 367	27.44	23.79 (-54.63 to 102.21)
Other <sup>a</sup>	46.56 (501.27); 410	22.80 (285.29); 367	23.76	9.30 (-37.78 to 56.37)
<p>a Societal perspective.  b Cannot be estimated owing to convergence issues.</p> <p><b>Note</b>  Costs are from NHS/PSS perspective unless otherwise specified. Model for estimating adjusted difference includes only baseline resource use as a covariate owing to convergence issues.</p>				

Participants randomised to receive the intervention had marginally better EQ-5D utility values than those randomised to the control arm at baseline (0.689 vs. 0.677, respectively) and this difference diminished after MI and adjustment (0.007, 95% CI -0.016 to 0.029). However, for SF-6D utility values, there was no difference between the arms at baseline (0.000, 95% CI -0.014 to 0.014) (Table 15).

The resulting predicted mean QALYs (calculated from the EQ-5D and SF-6D) and incremental QALYs are shown in Table 16. The intervention is associated with increased utility values at all follow-up time points, and this difference is statistically significant at 6 months using EQ-5D and at 12 months using SF-6D. The within-trial QALYs estimated using EQ-5D were 1.354 for the intervention arm and 1.314 for the control arm, a statistically significant difference of 0.040 (95% CI 0.006 to 0.074).

TABLE 15 Health-related quality of life measured in the REACT study

HRQoL	Mean value (SD); n								
	Intervention arm				Control arm				
	Baseline	6 months	12 months	24 months	Baseline	6 months	12 months	24 months	
<b>EQ-5D value</b>									
Crosswalk to EQ-5D-3L	0.689 (0.158); 397	0.708 (0.167); 346	0.705 (0.170); 337	0.686 (0.200); 330	0.677 (0.165); 352	0.670 (0.181); 299	0.680 (0.183); 293	0.661 (0.195); 302	
EQ-5D-VAS	70.6 (17.3); 399	72.6 (16.7); 349	71.4 (17.8); 338	70.4 (18.4); 329	72.1 (16.9); 362	72.0 (17.3); 298	70.6 (17.0); 294	69.4 (19.2); 301	
EQ-5D-5L value set	0.789 (0.149); 397	0.805 (0.160); 346	0.801 (0.158); 337	0.782 (0.180); 330	0.781 (0.152); 352	0.770 (0.177); 299	0.785 (0.162); 293	0.767 (0.174); 302	
<b>SF-36</b>									
PCS	29.7 (11.0); 393	32.8 (11.5); 342	31.9 (11.5); 334	30.9 (12.0); 326	30.0 (10.6); 352	30.6 (10.9); 293	29.8 (10.9); 293	29.2 (10.8); 295	
MCS	54.6 (8.3); 393	54.4 (8.6); 342	54.0 (8.8); 334	54.3 (8.6); 326	53.8 (8.7); 352	54.0 (9.1); 293	54.3 (9.4); 293	54.5 (9.1); 295	
SF-6D	0.622 (0.095); 365	0.637 (0.103); 323	0.637 (0.098); 313	0.630 (0.105); 312	0.623 (0.089); 326	0.621 (0.103); 280	0.622 (0.100); 280	0.619 (0.096); 283	
HRQoL, health-related quality of life; MCS, mental component score; PCS, physical component score.									



TABLE 16 Quality-adjusted life-years in the REACT study

Timepoint	EQ-5D utility (crosswalk)				SF-6D utility			
	Intervention arm	Control arm	Difference	95% CI	Intervention arm	Control arm	Difference	95% CI
Baseline	0.685	0.678	0.007	-0.016 to 0.029	0.622	0.622	0.000	-0.014 to 0.014
6 months	0.703	0.668	0.035	0.012 to 0.058	0.634	0.621	0.012	-0.003 to 0.027
12 months	0.692	0.676	0.016	-0.007 to 0.039	0.635	0.617	0.018	0.004 to 0.031
24 months	0.676	0.657	0.019	-0.007 to 0.046	0.627	0.616	0.011	-0.002 to 0.024
Within-trial QALYs	1.354	1.314	0.040	0.006 to 0.074	1.241	1.216	0.025	0.006 to 0.044

### Sensitivity analyses

Varying the unit costs for NHS resources each by  $\pm 20\%$  did not result in a large impact on total costs (Figure 14). Four NHS resources had the potential to affect incremental total costs by more than £20:

1. Increasing the cost of a day-case procedure (base-case value £752) by 20% increased the costs for the intervention relative to the control arm by £39.
2. Increasing the cost of a hospital inpatient bed-day (base-case value £504) by 20% decreased the costs for the intervention relative to the control arm by £32.
3. Increasing the cost of an intensive care unit stay (base-case value £3532) by 20% increased the costs for the intervention relative to the control arm by £25.
4. Increasing the cost of an outpatient appointment (base-case value £127) by 20% increased the costs for the intervention relative to the control arm by £24.

For costs affecting the societal perspective only, one-way sensitivity analyses showed very limited sensitivity to unit costs. Only the cost of a friend/relative taking a day off work and the cost of a chiropractor affected relative costs by more than £5 when varied by  $\pm 20\%$ .

The sensitivity analysis for the plausible extreme values of the societal value of an hour of help around the house revealed some sensitivity in costs from a societal perspective. When the value was reduced to £0.35 per hour, the total incremental costs for the intervention compared with usual care became more negative (to -£541, from a base case of -£458), while when the value was increased to £335 per hour the difference in costs narrowed (to -£347).

### Cost-effectiveness

Cost-effectiveness estimates include intervention costs (£621.83 per participant), as well as the impact of REACT on within-trial resource use, SPPB score and QALYs, as described above. Considering cost per 1-point increase in SPPB score at 24 months, the intervention is dominant in the base case (saves £103 and results in an increase in mean SPPB of 0.49 points).

When using QALYs as the measure of health benefit, as shown in Table 17, the intervention is cost-effective, either being dominant or having ICER below the £20,000 per QALY cost-effectiveness threshold. In the base case, the 95% upper confidence limit is below £20,000 per QALY, suggesting reasonable confidence that the intervention is cost-effective within the trial time horizon. However, uncertain results are found when there is no adjustment for covariates in the MI analysis or when only participants with complete data are included in the analysis adjusting for baseline covariates, because the upper 95% confidence limit for the ICER extends beyond the £20,000 per QALY cost-effectiveness threshold in these analyses.

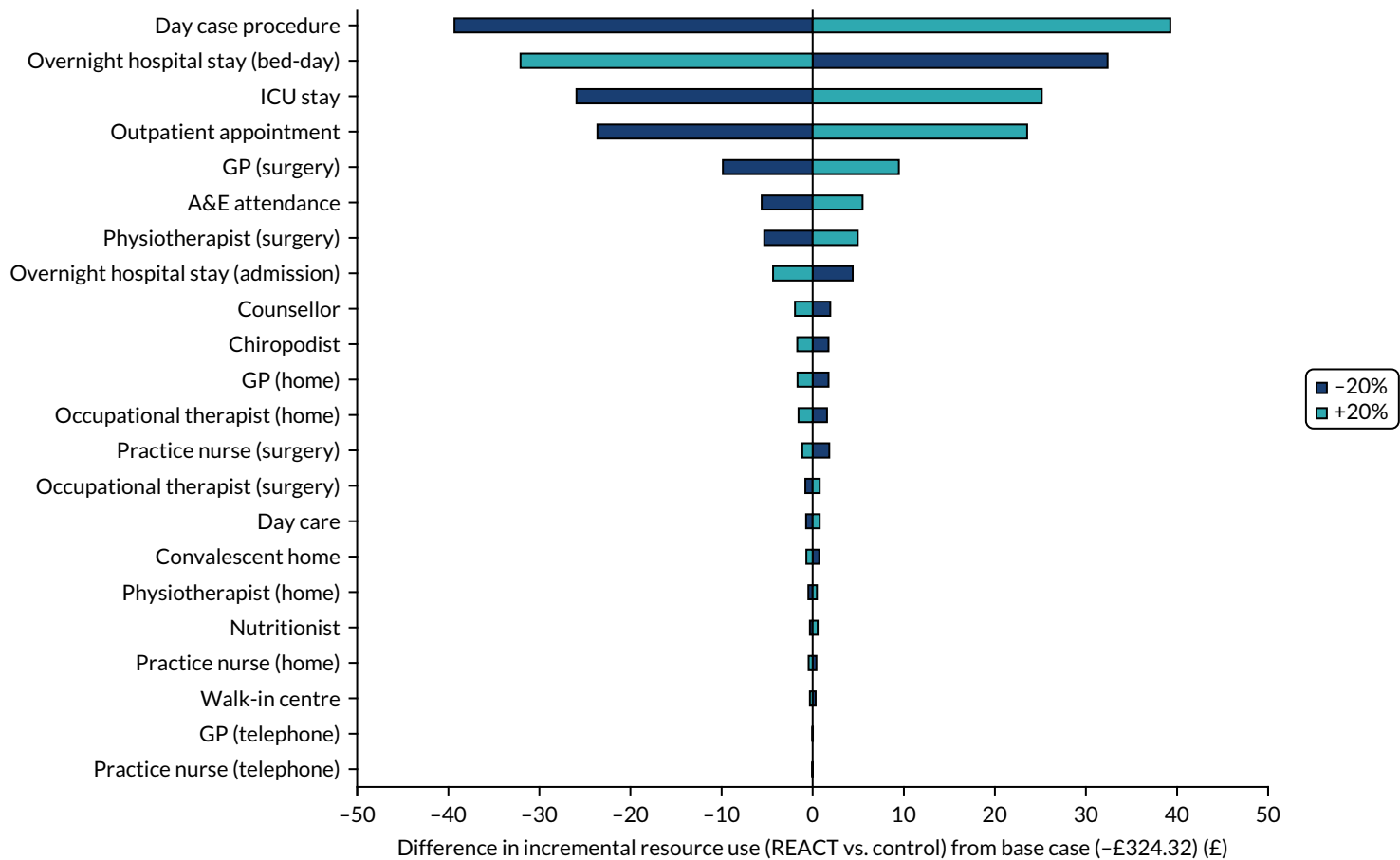


FIGURE 14 Tornado diagram for one-way sensitivity analyses of resource values (NHS/PSS perspective).

TABLE 17 Within-trial cost-effectiveness results

Trial arm	Costs (£)	QALYs	Including costs (£)		Including QALYs		ICER		
			Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	
<b>Base case: MI adjusted</b>									
Control	4046	1.314							
Intervention	3943	1.354	-103	-695 to 489	0.040	0.009 to 0.071	Dominant	Dominant	16,950
<b>Scenario analysis</b>									
<i>MI unadjusted</i>									
Control	3887	1.319							
Intervention	4076	1.364	189	-401 to 778	0.045	-0.001 to 0.091	4200	Dominant	Dominated
<i>CCA adjusted</i>									
Control	3573	1.323							
Intervention	3871	1.372	297	-421 to 1016	0.049	0.010 to 0.089	6000	Dominant	71,250
<i>CCA unadjusted</i>									
Control	3609	1.325							
Intervention	3856	1.389	247	-498 to 992	0.063	0.005 to 0.122	3900	Dominant	18,000
<b>Note</b>									
ICERs reported to nearest £50.									

The uncertainty in cost-effectiveness is demonstrated in cost-effectiveness scatter plots (Figure 15) and cost-effectiveness acceptability curves (Figure 16). The proportion of estimates cost-effective at £20,000 per QALY is 98% in the base case, 84% in the MI unadjusted analysis, 89% in the CCA adjusted analysis, and 98% in the CCA unadjusted analysis.

### Scenario analyses

In Table 18, a number of scenarios were explored. Removing the resource use impact observed in the REACT study makes the intervention less cost-effective but is not enough by itself to make the intervention not cost-effective at a £20,000 per QALY threshold. If the resource use impact is removed and SF-6D utilities are used to estimate QALYs (instead of EQ-5D), then the ICER for the intervention increases to £25,050 per QALY, which would not be considered cost-effective at a £20,000 per QALY threshold but would be considered cost-effective at a £30,000 per QALY threshold.

### Subgroup analyses

Subgroup analyses were conducted by refitting models for within-trial costs and QALYs, including interaction terms for the subgroup of interest with the group allocation variable. The REACT study appears to be cost-effective across all subgroups, except when the trial population is divided by baseline SPPB score. Participants with a baseline SPPB score of 4–7 do appear to obtain some health benefit in terms of QALYs, but there is no beneficial effect of the intervention on NHS/PSS resource use (see Appendix 3, Table 39).

### Discussion

The results indicate that the intervention plus usual care is cost-effective compared with usual care alone over the 2-year time horizon of the REACT study within-trial analysis. We estimated that the REACT intervention would cost £622 per participant to deliver, although this was sensitive to the costs of major inputs (REACT session leader, venue hire) and the group size. We also estimated that the

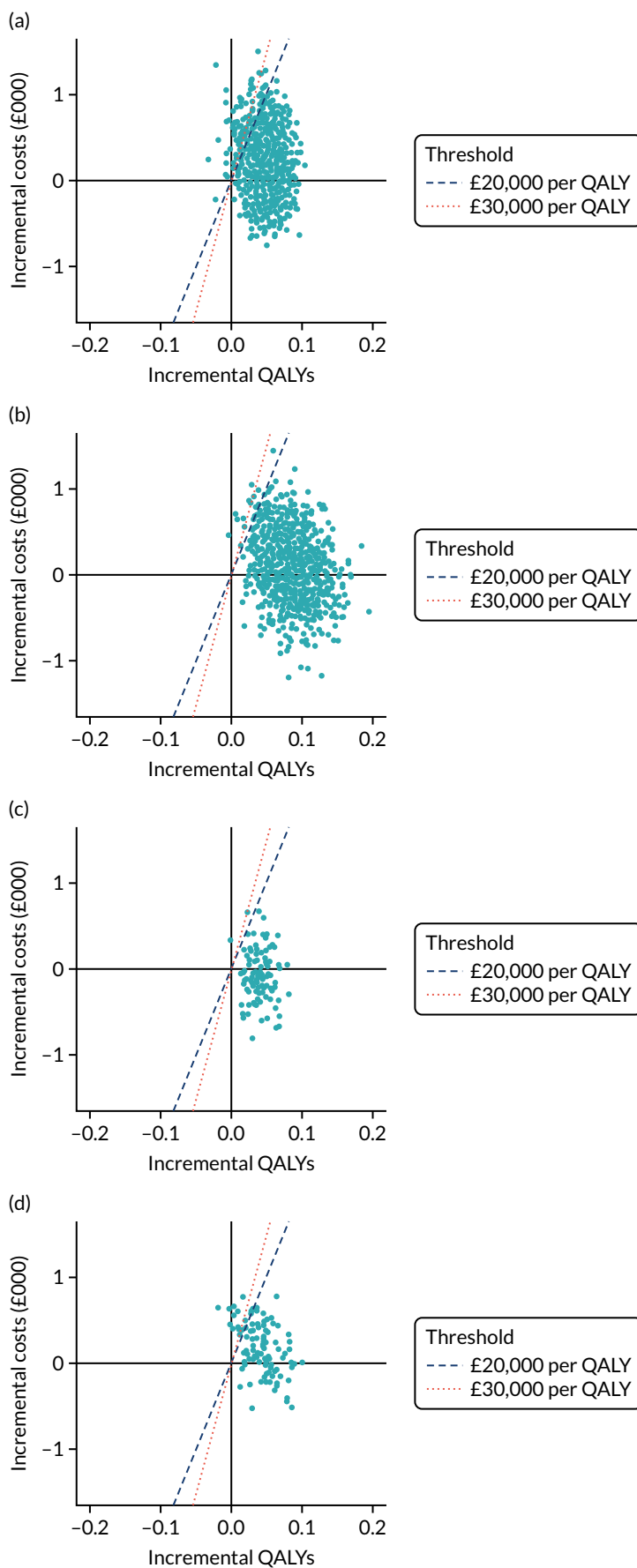


FIGURE 15 Cost-effectiveness scatter plot for within-trial analysis. a, CCA adjusted; b, CCA unadjusted; c, MI adjusted; d, MI unadjusted.

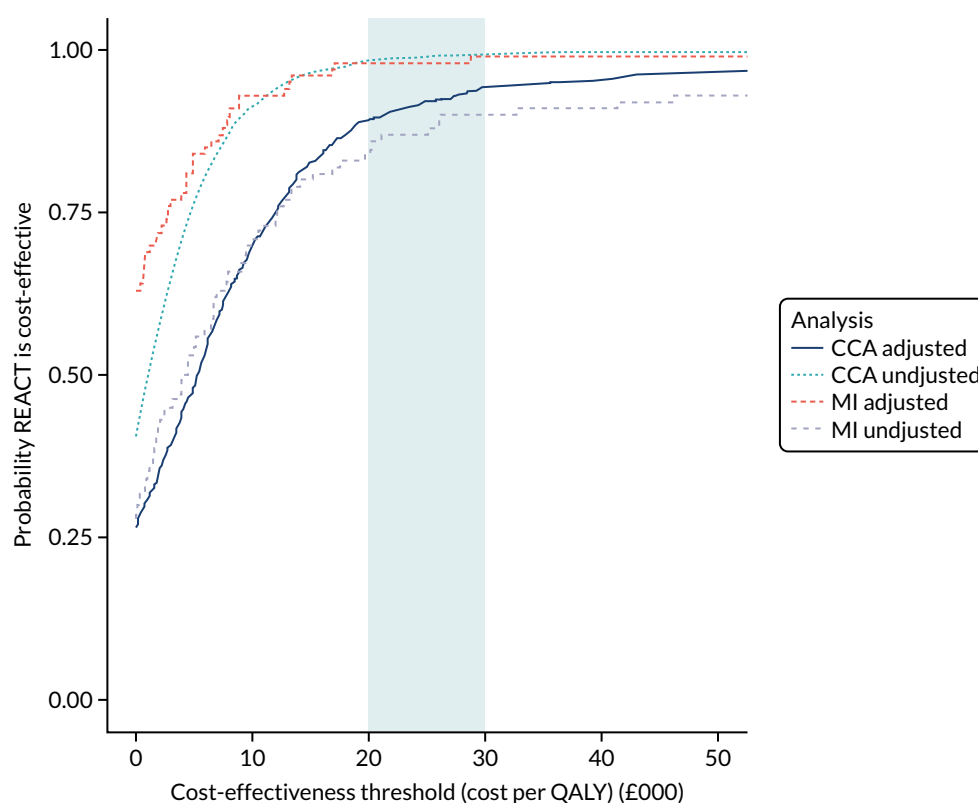


FIGURE 16 Cost-effectiveness acceptability curves for within-trial analysis.

TABLE 18 Cost-effectiveness scenario analyses (within-trial economic evaluation)

Scenario	Cost (£)			QALYs			ICER
	Control arm	Intervention arm	Difference	Control arm	Intervention arm	Difference	
Base case	4046	3943	-103	1.314	1.354	0.040	Dominant
No resource use impact	0	622	622	Unchanged from base case			15,650
Societal perspective	4240	4216	-24	Unchanged from base case			Dominant
SF-6D QALYs	Unchanged from base case			1.216	1.241	0.025	4150
No resource use impact and SF-6D QALYs	0	622	622	1.216	1.241	0.025	25,050
Group size 12	4046	4097	50	Unchanged from base case			1250

ICERs reported to nearest £50.

intervention would lead to a reduction in NHS/PSS resource use, corresponding to savings of £725, although the magnitude of cost savings was subject to uncertainty owing to sampling uncertainty and methodological choices (handling of missing data and baseline imbalances). We also estimated that the intervention would lead to an increase in discounted QALYs and alongside estimated cost saving the intervention is cost-effective (dominant), as it reduces costs to the NHS and society while increasing patient health benefits terms of mobility-related disability compared with usual care alone.

The REACT study had a similar gain in within-trial QALYs as the LIFE study (0.044 to 0.063 over 2 years in REACT, depending on analysis methods, vs. 0.047 over 2.6 years in LIFE),<sup>21</sup> although it should be noted that different preference-based instruments were used [EQ-5D-5L vs. Quality of Well-Being Self-Administered (QWB-SA)]. The cost of the intervention is cheaper than the cost of the LIFE intervention [£622 (2018/19 Great British pounds) vs. £1770 (US\$2301, 2013 US dollars<sup>122</sup>)], and has demonstrated a reduction in health-care utilisation of £324–724 compared with an increase of £1220 [US\$1583 (2013 US dollars<sup>122</sup>)].

This economic evaluation shows that the effect of the intervention on maintaining mobility, as demonstrated through SPPB score, the primary outcome of our study, also led to the expected improvement in self-reported health-related quality of life and health-care utilisation. It included a number of sensitivity and scenario analyses to explore the uncertainty in the results associated with missing data and imbalance in baseline characteristics between trial arms.

The principal limitation of the within-trial analysis is that it covers only 2 years and it is likely (given the persistence of differences in SPPB between the study arms) that further differences in costs and QALYs would occur beyond this time horizon. This particular limitation will be addressed in the next subsection. We have conducted regressions using only the gamma family and log link; we have not explored alternative parameterisations. There were missing data in the study, and it is not possible to rule out the possibility that data were missing not at random. It is estimated that 33% of long-term care spending comes from voluntary and out-of-pocket expenditure,<sup>123</sup> and this is excluded from the base-case perspective.

## Model-based economic evaluation

A pragmatic review identified two economic evaluations of physical activity interventions for older adults at risk of frailty (see *Appendix 3, Figure 39*). In both cases, the authors concluded that the intervention was cost-effective; however, both studies have certain limitations. The study by Alhambra-Borrás *et al.*<sup>124</sup> is at high risk of treatment selection bias, and the authors may have introduced further bias into their modelling by estimating costs and utility values separately for the intervention and control arms. Their modelling approach was based on dividing participants into two states based on the risk of falls/frailty, but this was poorly reported and was, therefore, of little value for informing the development of an economic model for REACT.

In the cost-effectiveness analysis study by Groessl *et al.*,<sup>125</sup> uncertainty is introduced on the impact of health-care resource utilisation, which in turn makes the cost-effectiveness of the LIFE study uncertain; the authors also did not attempt to estimate costs and QALYs beyond mean follow-up from the trial.

Although there is some evidence that group-based physical activity interventions can be cost-effective within their own settings, these findings are unlikely to be applicable to the REACT intervention owing to differences in the intervention designs (e.g. lasting 12 months compared with 36 months in the LIFE study).

### Aims and objectives

To estimate the cost-effectiveness of the REACT programme beyond the trial end time horizon (24 months) up to death.

### Decision problem

What is the cost-effectiveness of the REACT programme evaluated in line with the NICE reference case, specifically adopting:

- a lifetime time horizon
- NHS and PSS perspective (a societal perspective is adopted in a scenario analysis)
- benefit measured in QALYs
- 3.5% discount rate for costs and QALYs?

## Methods

A new decision-analytic model was developed in R (version 4.0.0) using the heemod package<sup>126</sup> (version 0.12.0.9000).

### Model structure and methodology

The decision analytic model focuses on projecting mobility over time and the consequences of mobility on utility values and costs for the NHS and PSS. The cohort of REACT participants enter the model at the end of the trial follow-up, distributed across a set of mutually exclusive mobility states defined by SPPB score. The following year, the patients face the probability of remaining in their end of trial mobility state, of moving to a worse or better health state, or of dying. This annual cycle process is repeated separately for intervention and control arms of the trial, until all cohort members are dead. At the end of the lifetime simulation, the costs and QALYs are aggregated and compared between trial arms to assess the differences in expected costs and QALYs and cost per QALY gained by receiving the intervention over usual care.

In this Markov cohort simulation model, the transition probabilities are allowed to vary with time but the model retains the Markov memoryless property,<sup>127</sup> with 13 health states reflecting different levels of mobility (SPPB score of 0 to SPPB score of 12) and a death state. SPPB score has been shown to be a prognostic for loss of ability to walk 400 metres<sup>32</sup> and development of mobility disability and disability in ADL.<sup>128</sup> We have further found it to be correlated with utility values and NHS/PSS resource use within the REACT study.

The utility values and NHS/PSS costs are related to the SPPB score based on statistical models fitted to the control arm of the REACT study. Transition probabilities between the mobility states are related to the current SPPB score, age and sex, estimated using statistical models fitted to the control arm of the REACT study. It conservatively assumed that mobility does not affect the mortality rate, which means that the model predicts only differences in QALYs based on utility values. Studies have shown an association between SPPB and mortality;<sup>129</sup> however, it was considered plausible that long-term health conditions could affect both mobility and mortality, and that mobility has not been established sufficiently as a surrogate for mortality risk.<sup>130</sup>

The model has a cycle length of 1 year and runs for 35 cycles, at which point the cohort is aged 100–130 years, depending on the starting age, which ranges from 65 to 95 years in 5-year intervals. The model is run separately for men and women. In the first two annual cycles, the costs and QALYs are taken directly from the REACT study, after which the distribution of SPPB values in the cohort is established based on predicted values from the REACT study according to age and sex.

### Model inputs

#### Population

In the base-case analysis, we assume a population of women aged 75 years given that women were the majority of participants (66%) in REACT and the age 75 years is close to the modal age in the study of 77 years for men and women. We also analyse cost-effectiveness for men and women aged 65–95 years in 5-year intervals.

The distribution of SPPB values (after 24 months) for each combination of age, sex and group (intervention or control) is estimated from the REACT study using ordered logistic regression, as shown in *Figure 17*. Specifically, an ordered logistic regression was fitted to trial-end SPPB score, with intervention allocation, age, sex, SPPB score at screening into the study and study site as predictors, and predicted probabilities of each trial-end SPPB score category were computed based on age, sex and intervention allocation using the margins command in Stata.

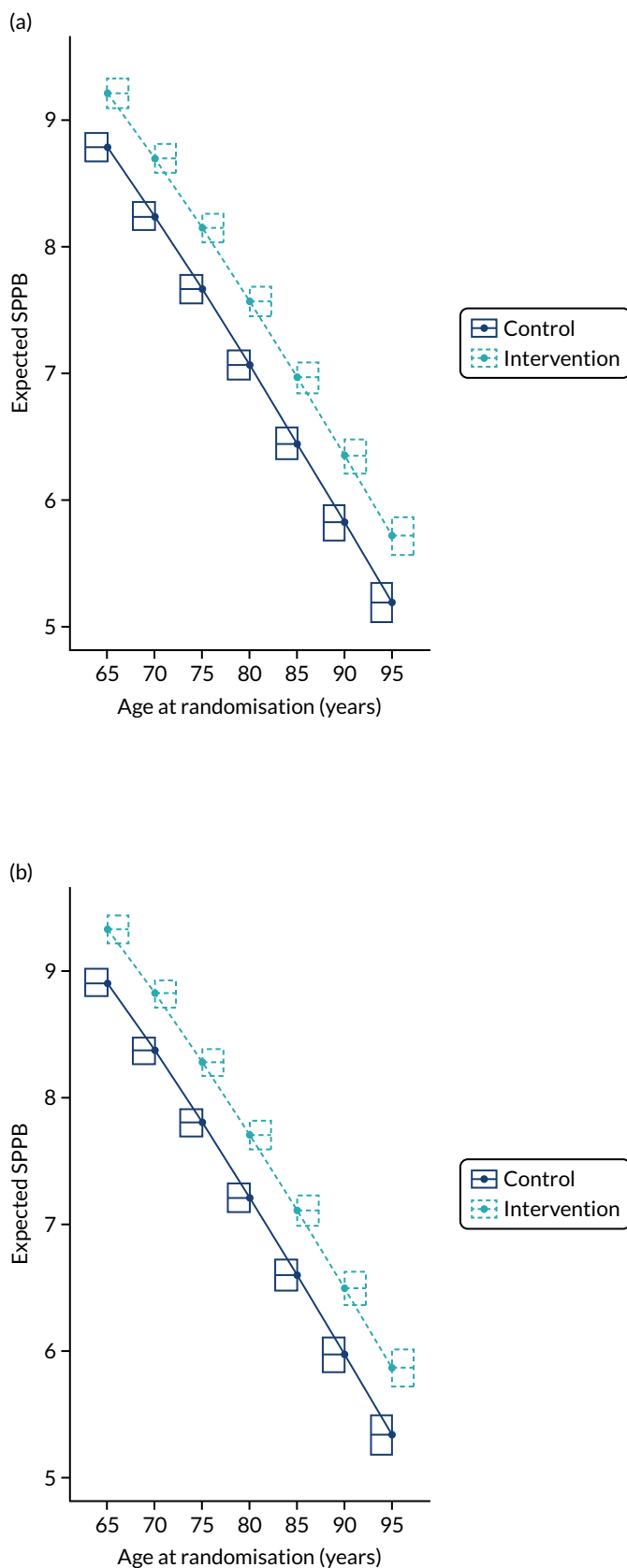


FIGURE 17 Expected trial-end (24-month) SPPB score by age, sex and group (error boxes show  $\pm 1$  standard error). a, male; b, female.



To estimate the average population costs and benefits of REACT, we use two reference populations:

1. the REACT study participants
2. UK population tables.<sup>131,132</sup>

In each case, we estimate weights based on the number of individuals of a given sex whose nearest age is 65, 70, 75, ... and 95 years.

### **Intervention effectiveness**

The effectiveness of the intervention is incorporated in two components:

- the directly observed effect of the intervention on costs (excluding the intervention cost) and QALYs within the REACT study (i.e. up to 24 months after randomisation)
- the SPPB score distribution realised at the end of the trial period.

The first of these components is taken from the analysis in which MI is used to account for missing data and bootstrapping is used to characterise uncertainty. This analysis (see *Results*) produces cost estimates of £4043 [standard error (SE) £254] in the control arm and £3319 (SE £173) in the intervention arm (excluding the cost of delivering the intervention), and produces QALY estimates of 1.348 (SE 0.012) and 1.388 (SE 0.014) in the control arm and the intervention arm, respectively.

The second of these has been described above (see *Figure 17*).

Extrapolation of the intervention effectiveness on health outcomes and costs is modelled as a function of the primary outcome in the trial, SPPB, under variable scenarios about the rate at which the effect of REACT on this outcome wears off.

### **Natural history**

There are two natural history components in the model:

1. the evolution of mobility with age
2. the changing mortality rate with age.

These components do not interact, that is there is no differential mortality rate according to mobility level. Both components apply equally to the intervention and control arms. This means, for example, that there is no assumption that participants in the intervention arm are better able to maintain their mobility after the 24-month trial period. The natural history model is fully described in *Appendix 3*.

### **Costs**

NHS and PSS (annual) costs were estimated as a function of mobility (SPPB score, categorised as 0–3, 4–7, 8–9 or 10–12) from the control arm of the REACT study. For each time point, we estimated the current cost accrual rate based on the cost of recalled health-care resource usage in the 6 months leading up to the time point, which we denote  $x_m$  for the  $m$ -month follow-up time point. At baseline, we estimated costs to be accrued at a rate of:

$$r_0 \approx x_0 + x_6. \quad (1)$$

We further estimate:

$$r_6 \approx x_6 + x_{12} \quad r_{12} \approx \frac{x_6 + 2x_{12} + x_{24}}{2} \quad r_{24} \approx 2x_{24}. \quad (2)$$

A linear regression with dummy variables for the SPPB categories was fitted, adjusting for clustering within participants, as shown in *Figure 18*.

The model shows that costs generally increase with decreasing mobility, but there is considerable uncertainty about costs for those with very poor mobility (SPPB score of  $< 4$ ); this is partly because the REACT study selected participants with a SPPB score of  $\geq 4$ , so there are fewer data available ( $n = 48$  observations with SPPB score of  $< 4$  out of 1269 SPPB observations in the control arm).

### Utility values

Utility values were estimated from SPPB using a linear regression with a restricted cubic regression spline for SPPB.<sup>133</sup> Observations from control arm participants were included, and the analysis was adjusted for clustering within participants (*Figure 19*).

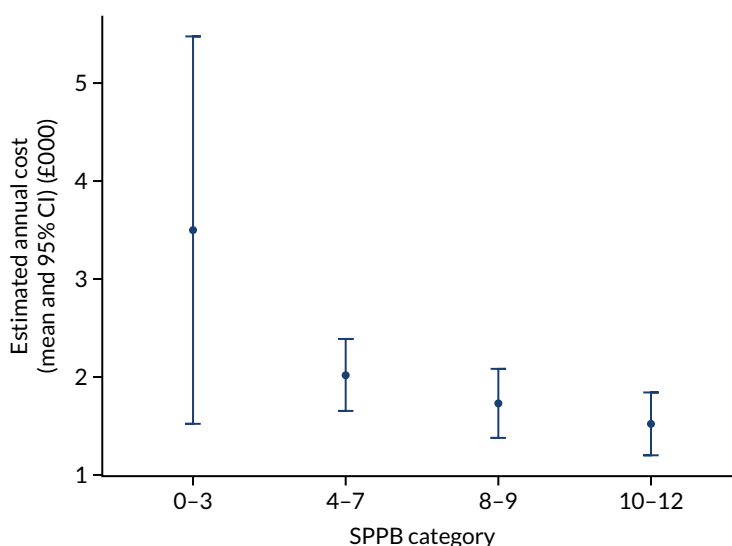


FIGURE 18 Cost by SPPB score range.

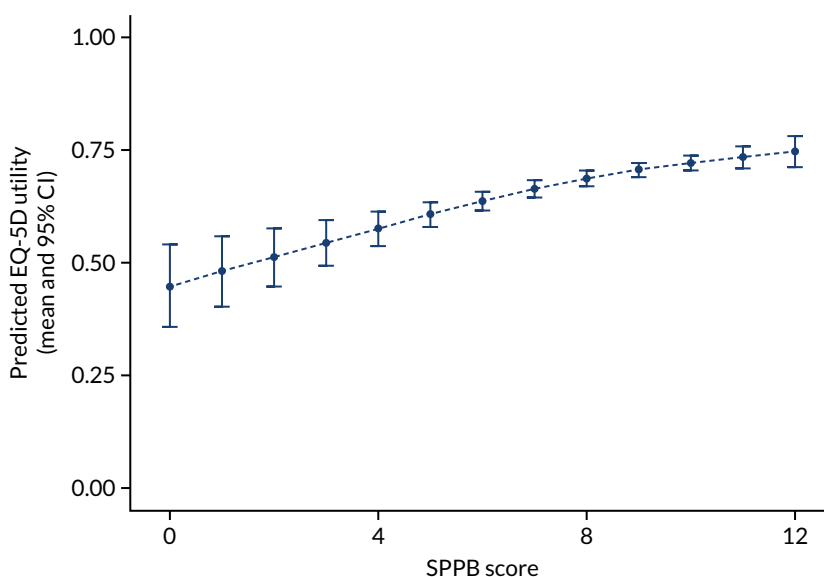


FIGURE 19 Predicted EQ-5D utility according to SPPB score.

Given that SPPB score declines with age (see *Appendix 3*), this means that utility values also decline with age. In the REACT study, a Pearson's chi-squared test showed significant associations of baseline SPPB score with baseline EQ-5D self-care ( $p < 0.001$ ), usual activities ( $p < 0.001$ ) and pain/discomfort ( $p < 0.001$ ) scores. In addition, age was not negatively associated with EQ-5D utility before or after adjusting for baseline SPPB score.

### **Analysis of uncertainty**

Deterministic one-way sensitivity analyses of model parameters and probabilistic sensitivity analysis (PSA) were considered. Correlations between model parameters in the PSA are incorporated through sampling from suitable multivariate normal distributions or sampling from bootstrap estimates. We assume that the cost of the intervention follows a gamma distribution with coefficient of variation of 20%.

We also consider a number of scenario analyses:

- all model inputs re-estimated using CCA rather than MI
- societal costs instead of NHS/PSS costs
- SF-6D utilities instead of EQ-5D
- four alternative models for the evolution of SPPB (as described in *Appendix 3*).

A summary of all model parameters is provided in *Appendix 3, Tables 45–48*.

## **Results**

### **Base case**

The model predicts a decline in SPPB over time for a woman aged 75 years with no difference in mortality between the two arms, as shown in *Figure 20*.

In the lifetime simulation, the cohort receiving the intervention spends more time in higher SPPB states (8–12) but less time in lower SPPB states (<8), and there is no difference in total life years lived. Spending more time in higher SPPB states results in accruing more QALYs and incurring lower costs (see *Appendix 3, Table 48*).

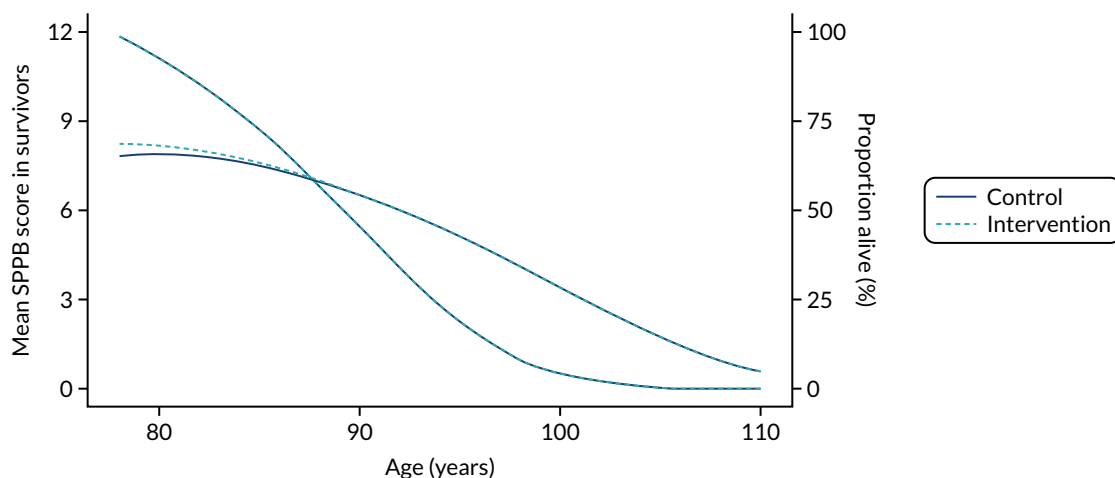


FIGURE 20 Mean SPPB score and proportion alive in the Markov model.

In our base-case analysis, a woman aged 75 years receiving the intervention is expected to gain 0.072 discounted QALYs over her lifetime and to reduce lifetime discounted costs by £290 (see Appendix 3, Table 49). This compares with the trial-end incremental costs of -£103 and 0.040 QALYs. The intervention is both more effective and less costly than usual care and, therefore, cost-effective regardless of the threshold at which the NHS is willing to pay for a QALY.

Figure 21 shows the cumulative incremental discounted QALYs over time. More than half of the lifetime incremental discounted QALYs come in the first 2 years. After 10 years, there are very limited further incremental QALYs, which reflects the fact that SPPB score distributions have largely converged by this point and an increasing proportion of the population is dying and, hence, unable to accrue further QALYs.

Figure 22 shows the cumulative incremental discounted costs over time. In contrast to QALYs, the majority of the lifetime cost savings do not come in the first 2 years, although, similar to QALYs, there is very limited change after 10 years. The intervention is predicted to lead to QALY gains and cost savings and, therefore, be cost-effective at any time horizon beyond 2 years.

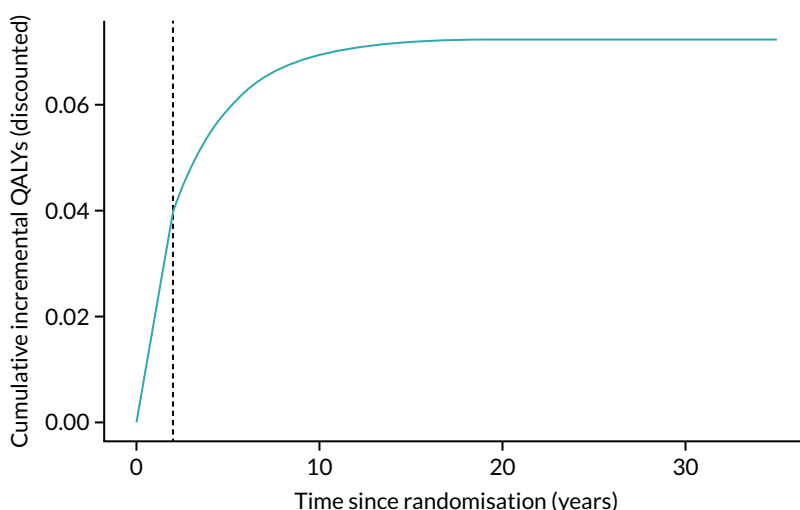


FIGURE 21 Cumulative incremental discounted QALYs over time.

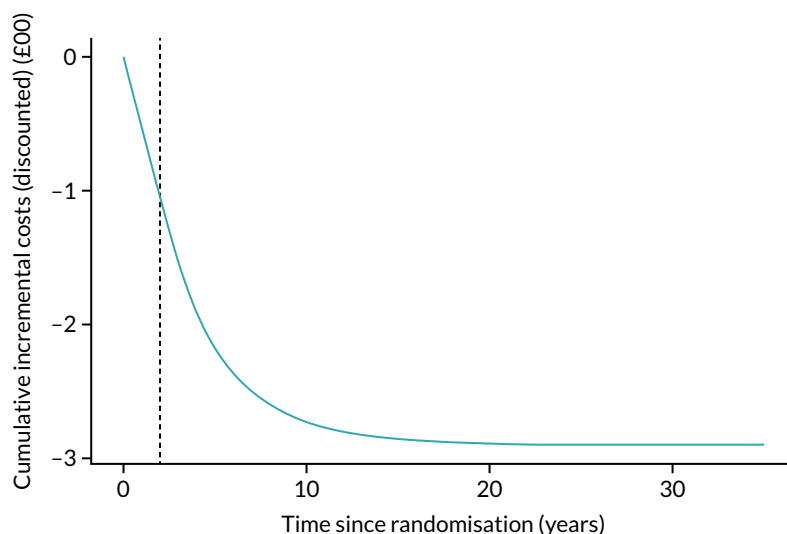


FIGURE 22 Cumulative incremental discounted costs over time.

### Heterogeneity analysis

As shown in *Figure 23*, it is expected that discounted lifetime costs and QALYs are higher for younger participants than for older participants, and are higher for women than for men. This is consistent with the fact that life expectancy is greater for women than for men.

Analysis of the incremental costs and QALYs reveals that the REACT programme is most beneficial to those aged 80–85 years and is more beneficial for women than for men (*Figure 24*). The REACT programme is expected to reduce lifetime costs for participants of any age and sex (*Figure 25*), which means that it is dominant across the population. The incremental net monetary benefit (INMB) at a willingness-to-pay threshold of £20,000 per QALY is shown in *Figure 26*, which demonstrates that the cost of the intervention could increase by £1350 (to around £1970) and still be marginally cost-effective for 65-year-old men and cost-effective for men of other ages and women across the age range.

The base case produces similar incremental costs and QALYs to the REACT study participants, on average. When the reference population was the UK general population aged 65–100 years, REACT was still dominant but produced slightly smaller cost savings and incremental QALYs (see *Appendix 3, Figure 41*).

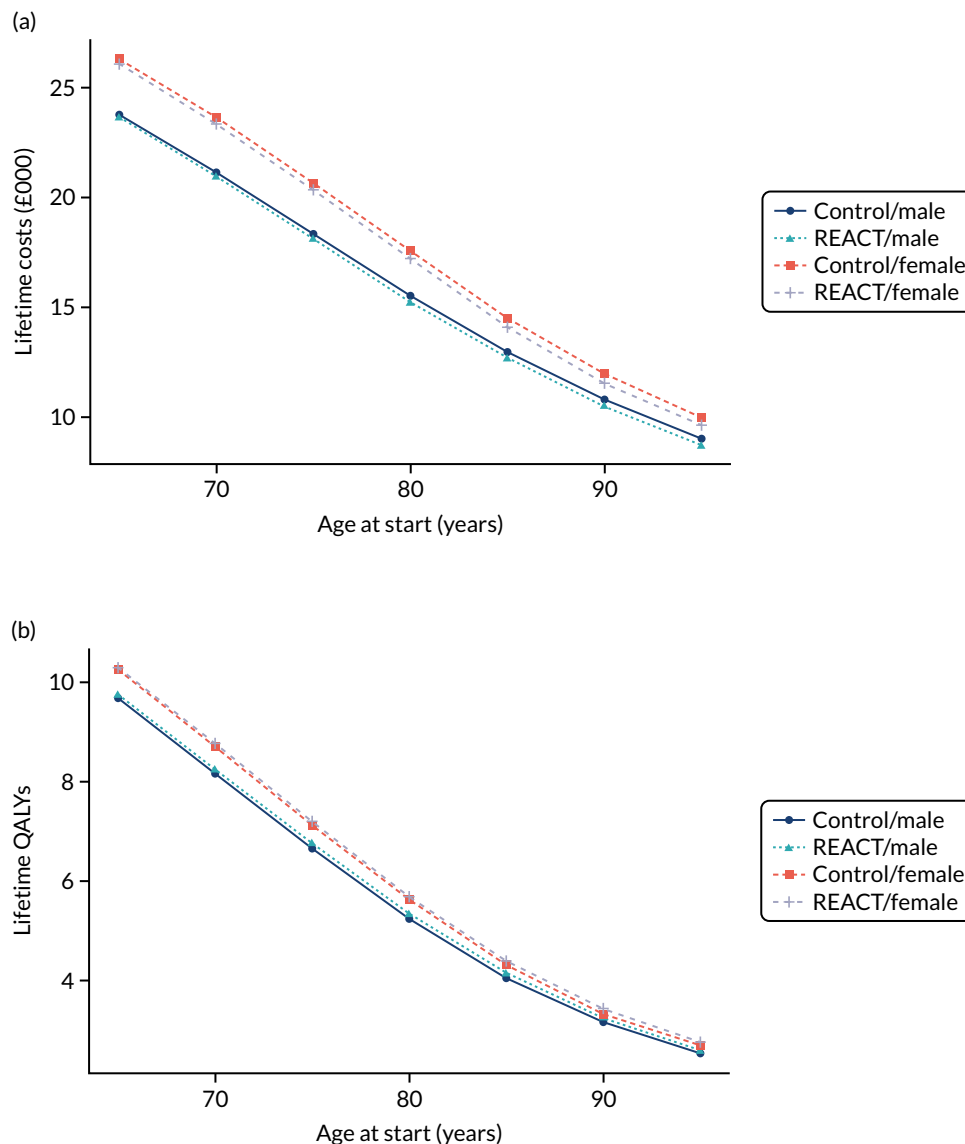


FIGURE 23 Heterogeneity analysis, absolute lifetime discounted costs and QALYs. a, Lifetime costs; b, lifetime QALYs.

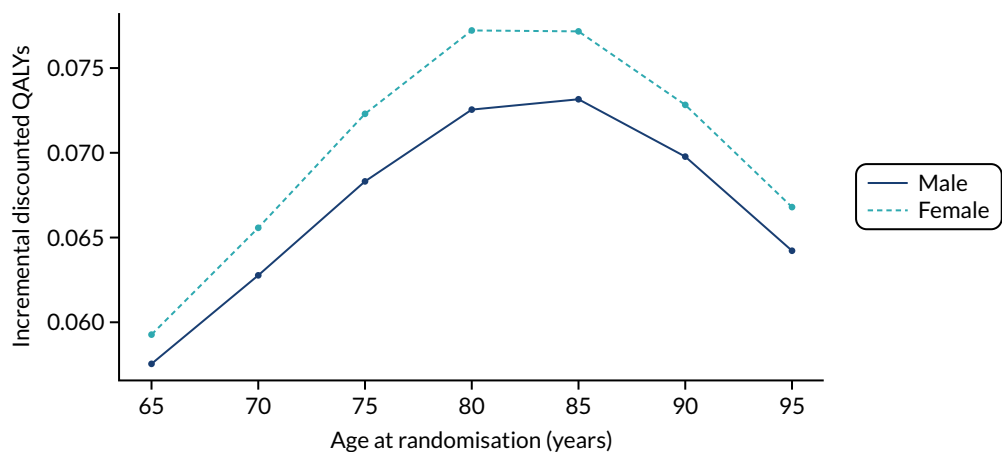


FIGURE 24 Heterogeneity analysis: incremental QALYs.

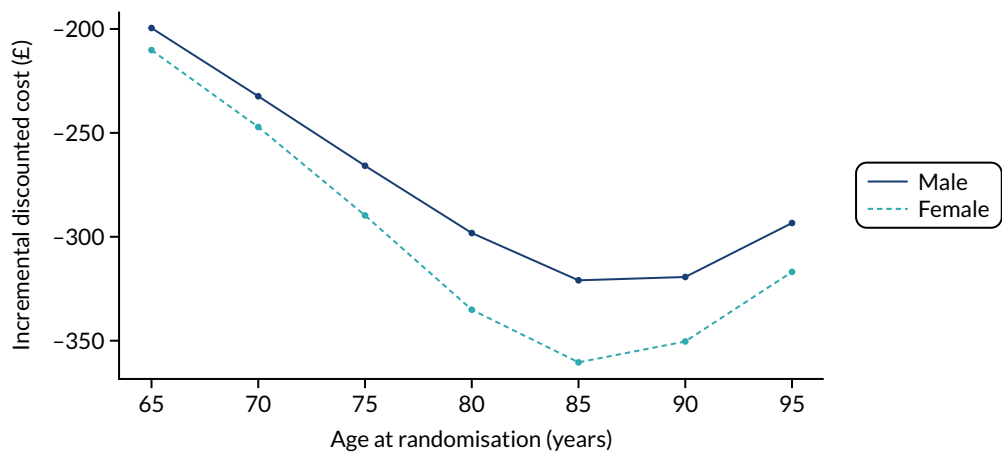


FIGURE 25 Heterogeneity analysis: incremental costs.

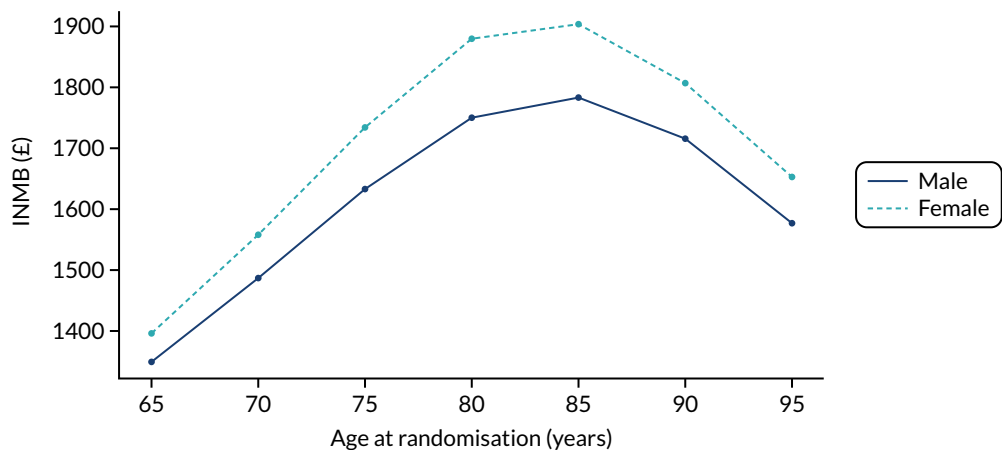


FIGURE 26 Heterogeneity analysis: INMB.

### Sensitivity analyses

A PSA was conducted with 500 iterations. The 95% credible interval for INMB for the base case was £659 to £3550 at a willingness-to-pay threshold of £20,000 per QALY. The 95% credible interval for the ICER was -£14,626 (dominant) to £5537 per QALY. Further heterogeneity analysis is presented in *Appendix 3, Figure 42*.

### One-way sensitivity analyses

One-way sensitivity analyses were conducted for all model parameters (*Figure 27*). The results show that the within-trial QALY and cost estimates strongly influence cost-effectiveness overall, as do the effects of age on utility values and SPPB evolution. There are no parameters that independently cause the intervention to have an ICER above £20,000.

### Scenario analyses

The results of the scenario analyses with a population of women aged 75 years (base-case analysis) are shown in *Table 19*.

All model inputs were re-estimated using CCAs rather than MI analyses. This was the only scenario in which lifetime costs in the intervention arm exceeded those in the control arm (although costs for both arms were lower than in the base-case analysis). The intervention is predicted to be cost-effective in this scenario because the ICER is well below £20,000 per QALY.

Societal perspective on costs had a limited effect on cost-effectiveness. QALYs were unchanged and costs in both arms increased, although incremental costs were marginally smaller.

The SF-6D utilities had a moderate effect on QALYs, reducing lifetime QALYs in both arms and reducing the difference in QALYs between the arms. Costs were unchanged and REACT was still the dominant strategy.

The finding that the intervention produced cost savings and more QALYs than usual care remained in all scenarios analysis of structural uncertainty associated with the model extrapolation.

### Discussion

The results of the model-based economic evaluation reinforce the findings that the intervention is cost-effective compared with usual care alone. Further cost savings and QALY gains are expected to be accrued up to around 15 years from randomisation, which are a result of the intervention reducing and/or preventing frailty, which is then associated with lower costs to the NHS and PSS, as well as better health-related quality of life. Adopting a broader societal perspective had little effect on the cost-effectiveness of the intervention.

The main strengths of the model-based economic evaluation are that it extends the trial-based economic evaluation to a lifetime horizon and assumes no lasting effect of the intervention beyond the end of the trial for maintaining mobility, and no effect on mortality or chronic conditions from REACT. Making these assumptions means that it is less likely that the effects of the intervention will be exaggerated, but also risks underestimating the total benefits of the REACT programme.

Although we model transitions across 13 mobility health states, our granularity came at the cost of statistical model simplicity, given that our ordinal logit regression model relies on the strong untested assumption that the effects of the variables conform to the proportional odds pattern, which is unlikely to apply across all predictors in the model, SPPB at 12 months and covariates in the base-case analysis.

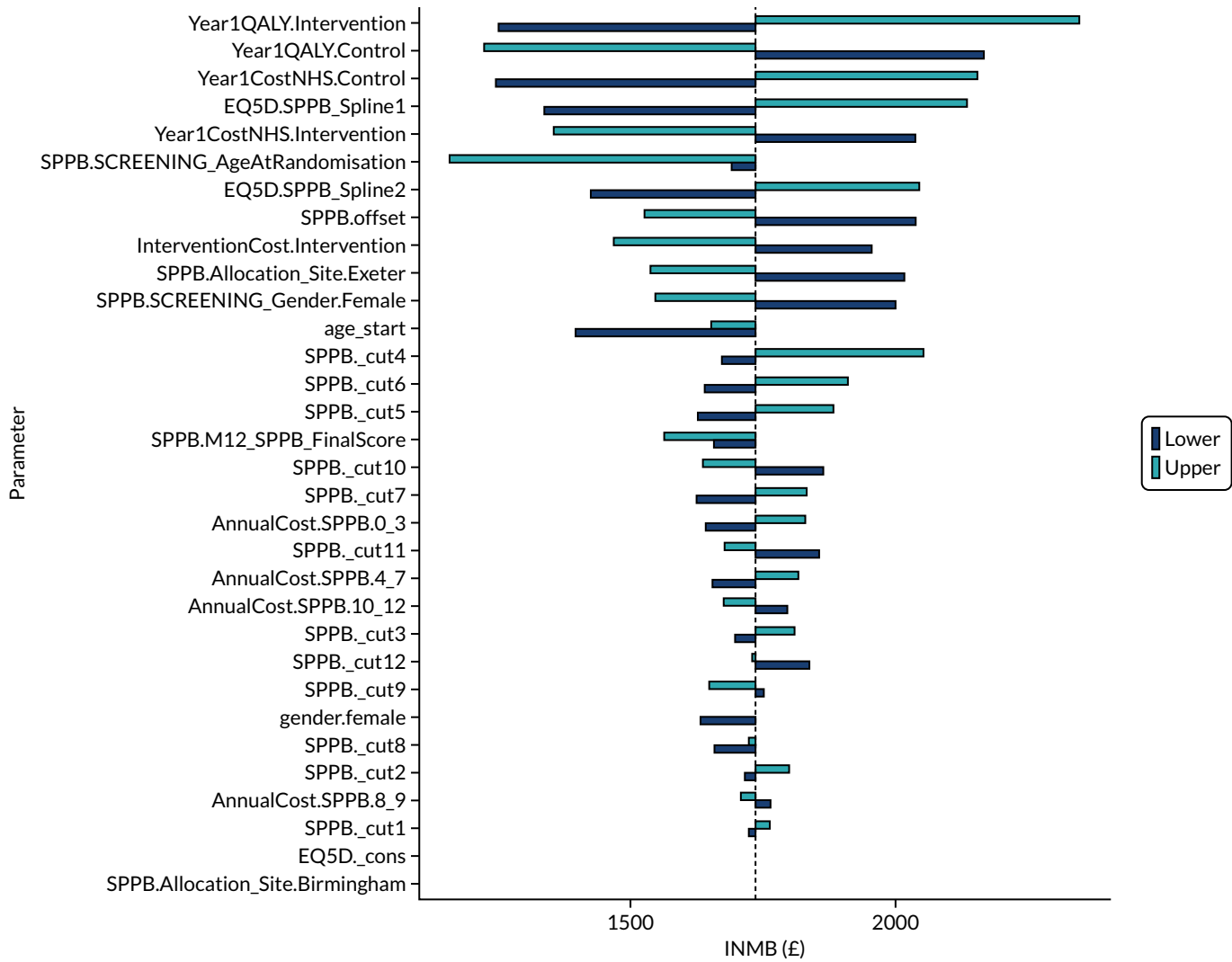


FIGURE 27 Tornado diagram for one-way sensitivity analyses.



TABLE 19 Scenario analyses for model-based cost-effectiveness analysis

Scenario	Cost (£)		QALY		INMB (£)	ICER (£/QALY)
	Control arm	Intervention arm	Control arm	Intervention arm		
Base case	20,627	20,338	7.111	7.183	1735	Dominant
Inputs from CCAs	19,580	19,690	7.177	7.259	1528	1345
Societal perspective	21,168	20,891	7.111	7.183	1723	Dominant
SF-6D utilities	20,627	20,338	6.579	6.623	1165	Dominant
Mobility model 2 (simplified ordered logistic)	21,057	20,690	7.044	7.126	2008	Dominant
Mobility model 3 (static mobility distributions)	20,502	19,838	7.171	7.302	3281	Dominant
Mobility model 4 (alternative method for inclusion of REACT SPPB changes)	23,843	23,268	6.749	6.855	2697	Dominant
Mobility model 5 (artificial transition matrix with 38% annual convergence)	23,192	22,932	6.834	6.896	1501	Dominant

A limitation of our model analysis is that the estimation of costs and utilities fully relied on the REACT trial data, which enrolled relatively mobile participants with SPPB levels of  $\geq 4$ . The predicted model of costs and utility values may, therefore, be biased because it relied on very limited numbers of participants at the lowest mobility levels and may not have accounted for the full spectrum of severity and its impact on health status and resource use. Further research will be required to validate the costs and utility values in representative samples, particularly those at increasing risk of permanent disability.

## Chapter 6 Discussion

This chapter discusses the key findings, process and economic evaluations of the REACT RCT and their implications for future research. Recommendations for scaling up the REACT programme are also provided.

### Key findings and interpretation

The primary REACT hypothesis was supported. Older adults with mobility limitations who received the 12-month, community-based group physical activity and behavioural maintenance programme had significantly better lower limb function at the 24-month follow-up than participants who attended a socioeducation-only control programme (adjusted mean difference 0.49 SPPB points, 95% CI 0.06 to 0.92 SPPB points). The difference in lower limb function between participants in the intervention and control arms was clinically meaningful at 6, 12 and 24 months. The intervention was also found to have significantly increased self-reported physical activity and performance of muscle-strengthening activity, and decreased perceived joint pain, after 24 months' follow-up. However, there was no significant effect on accelerometer-measured physical activity or other secondary outcomes, and some differences found at 12 months were reduced at 24 months. This may reflect deterioration of the effects of exercise behaviours over time.

In qualitative interviews, participants reported that they enjoyed the programme and being part of a research study, and they reported better mental and social well-being, especially higher physical confidence, improved motivation and feeling more outgoing. Improved social connectedness and bonding with REACT groups were key outcomes for participants in the intervention arm, who also cited numerous physical health benefits, including improvements in mobility, strength, balance, walking, fitness, sleep and physical independence. Themes identified at 24 months largely mirrored those reported at 6 and 12 months. However, whereas the 6- and 12-month interviews found that social support was a key reason for engaging in REACT, at 24 months individual-level factors, such as perceived benefits, were more prominent themes in explaining physical activity maintenance. Key components of the REACT programme that seemed to positively influence maintenance of physical activity at 24 months were (1) techniques for managing slips/lapses, supporting habit change and resolving sources of tension around increasing physical activity; (2) the person-centred delivery style to build autonomy/intrinsic motivation; and (3) the group-based delivery promoting social connectedness.

The quantitative process evaluation confirmed that, compared with the control arm, the intervention arm reported experiencing greater benefits from exercising in terms of their physical, mental and social well-being (see *Chapter 4, Results*). In particular, the hypothesis that increased exposure to the intervention will be associated with positive changes in competence, relatedness, enjoyment and perceived benefits (hypothesis 5c) was largely supported in relation to muscle-strengthening exercise only. Increased exposure to the intervention was associated with positive changes in psychosocial determinants for muscle-strengthening exercise 12 months after baseline, but not with changes in determinants of MVPA. Moderation analyses (see *Chapter 4, Results*) and subgroup interactions (see *Appendix 2, Table 33*) showed that REACT appeared to work equally well with men and women, and the positive effects were consistently seen irrespective of age, SES (assessed by education, home ownership and quintiles of area deprivation), baseline physical function, comorbidities and study site. However, the findings in relation to ethnicity and people with mental health problems need to be treated with caution owing to small group sizes.

The intervention fidelity analysis confirmed good delivery of the exercise programme. However, it identified considerable scope for improvement in the delivery of the behaviour maintenance sessions,

particularly in terms of support for monitoring progress/eliciting benefits of physical activity, action-planning/goal-setting, modelling of behaviour, supporting competence/self-efficacy, supporting relatedness and managing setbacks/problem-solving.

Taken together, the qualitative and quantitative process evaluations broadly supported the logic model for the REACT intervention. They identified several ways that the intervention and its implementation could be improved. This included possible changes to the logic model (from the qualitative and quantitative studies) and changes to delivery processes (from the intervention fidelity and qualitative studies). Recommendations for refinement/implementation of the REACT intervention are listed in the discussion sections of the relevant chapters.

Our health economic analyses indicated that REACT offers good value for money. The intervention plus usual care was cost-effective compared with usual care alone over the 2-year time period of the REACT intervention. In the base-case scenario, the intervention saved £103 in NHS/PSS costs per participant, with a QALY gain of 0.04 within the 2-year trial window. Lifetime horizon modelling estimated that further cost savings and QALY gains were accrued up to 15 years post randomisation. This is primarily because the intervention reduced frailty and its associated costs to the NHS and PSS, as well as improving health-related quality of life.

Stronger intervention effects were associated with a higher frequency of session attendance. An association between dose and response was observed, with an adjusted mean SPPB score difference of 0.64 points (95% CI 0.23 to 1.05 points;  $p = 0.002$ ) for those attending  $\geq 50\%$  of intervention sessions and of 0.81 points (95% CI 0.38 to 1.23 points;  $p < 0.001$ ) for those attending  $\geq 75\%$  of sessions. This strong dose-response relationship could simply reflect the effectiveness of the intervention or it could be explained in terms of participants being less likely to attend if their physical function deteriorated over the course of the study (e.g. owing to injury or other life events).

To situate our findings in the context of existing evidence, we searched Google Scholar (Google Inc., Mountain View, CA, USA), the Cochrane Database of Systematic Reviews, the NIHR Journals Library and PubMed, focusing on recent systematic reviews and meta-analyses from inception to 2014, using terms representing 'older people', 'physical activity', 'physical function' and 'randomised controlled trials'.

Based on these searches, the REACT study findings are consistent with a robust existing body of evidence demonstrating the positive impact of physical activity on physical functioning,<sup>11-14</sup> independent living, mobility-related disability, falls, hospital admissions and mortality.<sup>16,134,135</sup>

For example, the largest efficacy trial worldwide in this population, the US LIFE efficacy trial, has already shown that an exercise and behavioural support intervention can reduce the risk of developing major mobility disability by 18%. However, prior to the publication of the current findings, there has been a lack of evidence regarding the effectiveness and cost-effectiveness of long-term ( $\geq 1$  year) exercise programmes administered at a community level in a real-world setting and tailored for a UK population of older adults at risk of mobility disability. Furthermore, our database searches found no long-term, community-based interventions reporting effects for longer than 12 months post intervention.

The REACT study, therefore, provides robust evidence that a relatively low-resource, 1-year exercise intervention can improve physical functioning in real-world community settings in the UK, with clinically meaningful benefits that are sustained over at least 24 months.

The baseline SPPB scores were almost identical with the baseline score of the LIFE study, which enables comparison. The observed difference in SPPB score of 0.49 points at 24 months was three times larger than the between-group difference reported in the LIFE trial.<sup>21</sup> The REACT intervention cost an estimated £622 per participant, which also compares favourably with the cost of the LIFE intervention (£1770 per participant).

## Strengths, limitations and generalisability

To the best of our knowledge, the REACT trial is the first trial in the UK to target physical function with a long-term exercise-based intervention in a population of older adults who have restricted mobility. It was a robustly designed and conducted study with low attrition rates (19% at 24 months) and a comprehensive health economic evaluation. The comprehensive process evaluation increases our understanding of the mechanisms of engagement and behaviour change underpinning the REACT intervention. This will inform further refinement and implementation of the intervention and its training, as well as further development of the logic model and behaviour change theory. As well as informing any future implementation of the REACT intervention, this may help to design other interventions promoting active ageing.

Attendance compared favourably with that reported in other exercise programmes. Of the 83.9% who started the programme, 67.7% (95% CI 65.1% to 70.4%) of participants attended  $\geq 50\%$  of the sessions. However, it is notable that 16.1% of participants allocated to the intervention arm did not attend any intervention sessions, so there is scope to improve initial engagement in the programme. The intervention fidelity analysis also identified scope for improvement in several aspects of delivery of the health behaviour maintenance sessions. Taken together, these intervention delivery deficits may have led to underestimation of intervention effects.

There are some limitations that need to be acknowledged. As with other studies of behavioural interventions, blinding of participants to the study arm was not possible, which introduces the possibility of social desirability bias in patient-reported measures.<sup>136</sup> However, the primary outcome was assessed by independent observers who were blinded to study arm allocation, so this is unlikely to affect the primary outcome.

The generalisability of the findings to a wider UK population is also a key issue. The REACT study aimed to test a programme that was acceptable to as wide a range of the older adult population as possible and to recruit a diverse population in terms of sex, area deprivation and ethnic background. Our attempt to recruit a representative cohort for REACT in terms of deprivation and ethnicity was reasonably successful. Across the IMD quintiles, we recruited 3.2% less than the general population of those most deprived (quintile 1), but 2.6% more than those in quintile 2 (see *Table 2*). Overall, we avoided a substantial skewing towards affluence, which has proved challenging with this age group in previous research.<sup>23</sup> The sample recruited was also representative of the older adult UK population in terms of deprivation and ethnicity, except for an under-representation of South Asian older adults.<sup>29</sup>

Although only 3% of those invited to take part were recruited, the response rate among the eligible population was 17% (see *Chapter 3*). Given this, and the small numbers of ethnic minority participants recruited, the generalisability of the findings (particularly) to ethnic minority populations should be explored further in future studies or as part of a REACT implementation evaluation study.

The mixed findings on the secondary outcomes indicate that the effects of the intervention were limited to lower limb physical function and did not extend to substantial increases in physical activity or other domains of physical function (e.g. grip strength). The secondary outcome analyses were exploratory, with no adjustment for multiple testing, and should be interpreted accordingly.

Furthermore, the trial analyses did not show an impact on quality of life. However, the more sophisticated, time-integrated approach used in the health economic analysis revealed a significant difference in EQ-5D (as well as a saving in health-care costs). Indeed, the health economic analyses indicated that the increases in physical function observed were associated with substantial quality-of-life and health economic benefits, both within the 24-month trial window and across a lifetime horizon.

Although the overall results for REACT were positive, the process evaluation indicated, as with most service-based interventions, that there was considerable scope for improvement by session leaders in the facilitation of important self-regulation processes and social/relatedness-building processes during the delivery of the behavioural maintenance programme. To some extent, this may have been mitigated by mutual support among participants and self-delivery of some of the intended processes within the groups during the exercise sessions. However, future implementations of the REACT intervention should aim to improve the training and delivery of the programme accordingly.

### Considerations for further research

A key consideration for scaling up REACT and for future research in this population is how to identify people who would benefit from interventions (those with SPPB scores of 4–9). Although we met our recruitment target, from invitation to randomisation we had a success rate of only 3%. This recruitment yield is comparable to that of other behavioural interventions targeting older populations.<sup>29</sup> However, finding ways to make recruitment more efficient and more inclusive would be of benefit to future research in this field.

One means by which this might be achieved is via the UK's electronic Frailty Index (which is recorded in GP databases) or similar schemes being tested in Europe.<sup>33</sup> However, evidence of the electronic Frailty Index's ability to identify people with mobility limitations is required to reveal whether or not its use would be beneficial in recruitment for studies targeting improved mobility. Another option might be to develop an easy-to-use, parsimonious self-report measure that has strong predictive value for lower limb physical function. One candidate measure is the MAT-SF measure (a video-scenario-based self-assessment tool designed to assess lower limb physical function) that we used as a secondary outcome of REACT.<sup>53</sup> Further analyses of the REACT (and other) data will reveal whether or not this might prove to be a useful screening tool for use in practice and in future studies. A third option would involve relationship-building with multiple charities, local GPs and social prescribers, and community groups to help identify suitable candidates for intervention.

More research is needed in ethnic minority communities, particularly South Asian communities, to identify strategies for greater inclusivity and any necessary programme modifications to optimise the effectiveness of the REACT programme for ethnic minority populations. Further evaluation of the effectiveness of the programme for people from diverse ethnic backgrounds is also warranted.

The exclusion criteria for REACT meant that some groups of older adults or those with particular conditions were not able to participate, despite the fact that their needs suggest that they may have benefited. Examples include people with diagnosed dementia or people living in a residential care home. Future consideration should be given to strategies and conditions for their inclusion and for assessing the effectiveness of the REACT intervention in these populations.

Further research is also needed to optimise the implementation of REACT at scale and further evaluate and extend its reach, effectiveness and cost-effectiveness.

Using comprehensive participant and public involvement techniques to engage with session leaders, older adults eligible for the REACT programme, together with the research team, will help to fine-tune the behavioural maintenance programme (in terms of both content and training) and its delivery to ensure better fidelity of delivery when REACT is implemented at scale. The examples of good practice identified in the intervention fidelity analysis and the feedback from the qualitative process evaluation will be helpful in refining the REACT training materials.

It may be possible and synergistic to integrate the REACT intervention with existing mobility-related prevention and rehabilitation services, and provide support for participants who do not meet the criteria for specific programmes targeting falls. Given their cost-effectiveness, research into the cost-effectiveness of ongoing/continuous exercise programmes may also be fruitful; this may further improve the maintenance profile of the intervention, thereby increasing the available cost savings. There is also scope for longer-term health outcomes of the REACT programme to be evaluated. For example, maintained or increased mobility in older adults may influence longer-term mental well-being and delay onset of or deterioration through other diseases and conditions, such as dementia, diabetes and heart disease. Finally, the estimated effects of the programme on long-term health service costs, such as medications, primary care visits and emergency secondary care usage, should be directly measured to verify the conclusions of the health economic modelling, particularly to identify more definitive estimates of the cost consequences of delivering the REACT intervention at scale.

The REACT study provides robust evidence that the trajectory of declining physical functioning with age is modifiable and may even be reversible for many adults. It is particularly timely, allowing for some optimism post the COVID-19 pandemic, during which deconditioning was one of the key health challenges that older people experienced.

Future research will need to examine the feasibility of new ways to increase access to the programme, including virtual attendance, which might be useful in future lockdown/pandemic situations for people who are housebound, who may live more rurally or who prefer to engage in home-based interventions in later life. Similar to other programmes,<sup>137</sup> REACT has the potential to be delivered online in its entirety, but we need to evaluate the reach/accessibility of this delivery style and its impact on physical and psychological well-being.

## Implications for decision-makers and for practice

The REACT study has provided robust evidence that a group-based exercise programme with long duration (12 months) that encourages participants to build strong social networks and the habit of exercising in a safe and fun environment can improve their prognosis for maintaining mobility and save costs in both the short and the long term. This adds to the existing literature indicating that physical activity is crucial for older people's physical functioning and health. It should strengthen the case for consideration by funders and policy-makers to support exercise programmes for the critical at-risk population of older people with lower limb frailty or pre-frailty (as defined by a SPPB score of 4–7).

The intervention delivered positive results on mobility across different settings in three different geographical areas of the UK, using a range of session leaders who were already active in the field of exercise programme delivery. For this reason, the REACT intervention is suitable for delivery by existing service providers, including city council providers, voluntary sector organisations, commercial organisations, the leisure industry and the private sector; therefore, the results should generalise well to other real-world settings. However, attention needs to be paid to maintaining (or improving) delivery fidelity when rolling out complex interventions.<sup>43</sup> Therefore, it may be helpful to consider what organisation-level governance or performance-monitoring systems might support high-quality delivery of the REACT intervention.

Given that the intervention appeared to work equally well irrespective of age, gender, SES (assessed by education, home ownership and quintiles of area deprivation) and baseline physical function, REACT may have the potential to reduce health inequalities if delivery is targeted towards underserved populations.



The REACT exercise intervention also provides important evidence supporting both the US and the UK physical activity recommendations for multimodal exercise programmes.<sup>138,139</sup> In particular, it demonstrates the beneficial effect of engaging in at least one exercise session per week. This is a strong public health message for older adults in terms of the level of commitment that they need to demonstrate to achieve maintenance of their lower limb physical function. Contrary to the widespread belief that ageing causes an inevitable decline in physical functioning, the REACT study has proved that this decline can be delayed or reversed with relatively modest lifestyle changes.

The health economic analysis indicates that REACT provides a feasible and cost-effective means of helping older people stay mobile, and this will enhance their chances of remaining independent for longer. Based on the moderation and subgroup analyses presented above, it also has potential to help reduce health inequalities among the elderly. Given this and the cost-saving nature of the intervention, REACT should be given serious consideration by health and public health commissioners.

Although recruitment of older participants to a research study that includes a RCT is much more of a commitment than recruiting to an exercise programme offered by leisure or community services, many lessons have been learned that may help to inform real-world implementation. These lessons are available through our recruitment publication.<sup>29</sup> Furthermore, given that the NIHR Public Health Research programme does not fund intervention costs, the researchers needed to work closely with multiple partners to deliver the REACT intervention programme, which was an excellent test of programme pragmatism and sustainability. The delivery of the REACT study provides an important template for this partnership model.

Two further implications that may be useful for future practice and/or decision-making are that (1) the REACT exercise intervention seems to work even with attendance as low as 50% and a lower-than-intended quality of delivery of the behavioural maintenance programme, and (2) there is scope for improvement of delivery quality through refinements to the behavioural maintenance programme and its training course. This may lead to increased effectiveness and better maintenance of effects in future implementations of the programme.

Finally, it is important to position the REACT programme within the COVID-19 pandemic context and the potential for REACT delivery in the post-COVID-19 era. The COVID-19 pandemic has resulted in the target population of the REACT intervention following stay-at-home guidance during lockdowns and, in some cases, shielding outside lockdowns. During these periods, it would not have been possible to deliver the REACT intervention as originally designed and evaluated. It is unclear whether or not a digital alternative to REACT, delivered in people's homes or online, would be feasible, safe and effective.

As the UK exits from COVID-19 restrictions and a large proportion of the target population have been vaccinated, it is likely that the REACT intervention will once again be deliverable. There may be some concerns around travel to the REACT venues (if the target population are less comfortable taking public transport than previously), which could affect take-up. On the other hand, the demand for appropriate venues may have dropped owing to an increasing number of people working from home, which could lead to a reduction in the cost of delivering REACT (in the short to medium term).

Considering the possibility of future pandemics (e.g. further novel coronaviruses or influenza pandemics), the REACT intervention may be expected to allow more people to maintain their mobility and continue to live in private homes, rather than entering residential or nursing homes because of frailty. A significant burden of COVID-19 in the UK resulted from SARS-CoV-2 spreading through care homes and leading to the hospitalisation and/or death of residents; this may give an additional incentive for local governments and the NHS to fund interventions that stave off frailty.

## Conclusion

The REACT study provides robust evidence that a relatively low-resource, 1-year exercise intervention delivered alongside a behaviour maintenance programme helped UK older adults to retain their lower limb physical function over a 24-month period. The REACT intervention is cost saving from an NHS/PSS perspective within a 2-year window, with further cost savings and QALY benefits estimated in the longer term. The results indicate that the well-established trajectory of declining physical functioning in older age is modifiable and in some cases reversible. Following further refinements guided by the findings of the process evaluation, the REACT intervention is suitable for large-scale implementation. The findings also suggest that there is potential for improvement in the design and delivery of the REACT intervention, and implementation should be evaluated to confirm the generalisability of these findings to the wider population, especially in ethnic minority populations.

Some text in this chapter has been reproduced from Stathi *et al.*,<sup>1</sup> Withall *et al.*<sup>29</sup> and Cross *et al.*<sup>71</sup> These are Open Access articles distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.





# Acknowledgements

## Non-author contributors

We wish to express our thanks to the entire REACT research team, the Clinical Research Networks at each REACT site and all of the general practices and community organisations that supported REACT recruitment. Delivery of the REACT programme was possible only because of the support of our partners Bath and North East Somerset Council; Exeter and Solihull City Councils; Westbank Charity, Exminster; St Monica Trust, Bristol; Bristol Ageing Better; St John's Hospital, Bath; Age UK, Birmingham; Agewell, West Midlands; Sandwell and West Birmingham Hospitals NHS Trust; the Portway Lifestyle Centre, Oldbury; and Solihull Borough Council, Birmingham.

Our Trial Steering Committee, chaired by Professor Yoav Ben-Shlomo, Data Management and Ethics Committee, chaired by Professor Dawn Skelton, and independent statistician, Professor Peter Thomas, provided hugely valuable expert guidance and support.

Dr Andrew F Hayes of The Ohio State University Department of Psychology provided valuable advice on the conduct and interpretation of the moderation analyses.

Finally, we would like to thank the 777 REACT participants without whom this research would not have been possible.

## Contribution of authors

**Afroditi Stathi** (<https://orcid.org.uk/0000-0003-2162-777X>) (Professor of Physical Activity and Community Health) obtained funding for the research, acted as chief investigator and principal investigator at the Bath site, led the design of the study and drafted the report manuscript.

**Janet Withall** (<https://orcid.org.uk/0000-0003-0196-5309>) (Trial Manager, Active Ageing) led study co-ordination, contributed to the design of the study, was involved in the collection and processing of data and drafted the report manuscript.

**Colin J Greaves** (<https://orcid.org.uk/0000-0003-4425-2691>) (Professor of Psychology Applied to Health, Health Psychology) obtained funding for the research, acted as principal investigator at the Devon site, contributed to the design of the study, led the process evaluation and drafted the report manuscript.

**Janice L Thompson** (<https://orcid.org.uk/0000-0002-7312-3226>) (Professor of Public Health Nutrition and Exercise) obtained funding for the research, acted as principal investigator at the Birmingham site, contributed to the design of the study and reviewed the report manuscript.

**Gordon Taylor** (<https://orcid.org.uk/0000-0003-1715-0913>) (Professor of Medical Statistics) obtained funding for the research, contributed to the design of the study, developed the statistical analysis plan and undertook the analyses, and reviewed the report manuscript.

**Antonieta Medina-Lara** (<https://orcid.org.uk/0000-0001-7325-8246>) (Associate Professor in Health Economics) obtained funding for the research, contributed to the design of the health economic study, estimated the cost of the intervention and reviewed the report manuscript.

## ACKNOWLEDGEMENTS

**Colin Green** (<https://orcid.org.uk/0000-0001-6140-1287>) (Honorary Professor of Health Economics) obtained funding for the research, contributed to the design of the study and reviewed the report manuscript.

**Tristan Snowsill** (<https://orcid.org.uk/0000-0001-7406-2819>) (Senior Research Fellow, Health Economics) was involved in the collection and processing of data, estimated the cost of the intervention and reviewed the report manuscript.

**Heidi Johansen-Berg** (<https://orcid.org.uk/0000-0002-4134-9730>) (Professor of Cognitive Neuroscience) obtained funding for the research, contributed to the design of the study (MRI substudy) and reviewed the report manuscript.

**James Bilzon** (<https://orcid.org.uk/0000-0002-6701-7603>) (Professor of Human and Applied Physiology, Rehabilitation Medicine) obtained funding for the research, acted as principal investigator at the Bath site, contributed to the design of the study and reviewed the report manuscript.

**Selena Gray** (<https://orcid.org.uk/0000-0001-6257-4024>) (Professor of Public Health) obtained funding for the research, contributed to the design of the study and reviewed the report manuscript.

**Rosina Cross** (<https://orcid.org.uk/0000-0002-2548-2398>) (Postdoctoral Research Fellow, Primary Care) was involved in the collection and processing of data and contributed to the process evaluation report manuscript.

**Max J Western** (<https://orcid.org.uk/0000-0003-1107-8498>) (Lecturer, Behavioural Science) was involved in the collection and processing of data and reviewed the report manuscript.

**Jolanthe L de Koning** (<https://orcid.org.uk/0000-0001-6811-5726>) (Research Assistant, Active Ageing) was involved in the collection and processing of data and reviewed the report manuscript.

**Peter Ladlow** (<https://orcid.org.uk/0000-0002-9891-9714>) (Higher Scientific Officer, Strength and Reconditioning/Rehabilitation) contributed to the design of the study, led the design of the exercise intervention and reviewed the report manuscript.

**Jessica C Bollen** (<https://orcid.org.uk/0000-0002-6437-5923>) (Research Assistant, Active Ageing) was involved in the collection and processing of data and reviewed the report manuscript.

**Sarah J Moorlock** (<https://orcid.org.uk/0000-0003-4473-7226>) (Research Assistant, Active Ageing) was involved in the collection and processing of data and reviewed the report manuscript.

**Jack M Guralnik** (<https://orcid.org.uk/0000-0001-5108-7263>) (Professor in Epidemiology and Public Health, Epidemiology and Ageing) obtained funding for the research, contributed to the design of the study and reviewed the report manuscript.

**W Jack Rejeski** (<https://orcid.org.uk/0000-0003-2281-4649>) (Professor of Health and Exercise Science, Behaviour Change and Disability in Ageing) obtained funding for the research, contributed to the design of the study and reviewed the report manuscript.

**Melvyn Hillsdon** (<https://orcid.org.uk/0000-0003-2818-3278>) (Associate Professor in Physical Activity Epidemiology and Public Health, Physical Activity and Health) contributed to the processing and analysis of accelerometer data and reviewed the report manuscript.

**Kenneth R Fox** (<https://orcid.org.uk/0000-0003-2216-4287>) (Emeritus Professor of Exercise and Health Sciences, Physical Activity and Health) obtained funding for the research, contributed to the design of the study and reviewed the report manuscript.

## Publications

Stathi A, Withall J, Greaves CJ, Thompson JL, Taylor G, Medina-Lara A, *et al.* A community-based physical activity intervention to prevent mobility-related disability for retired older people (REtirement in ACTION (REACT)): study protocol for a randomised controlled trial. *Trials* 2018;**19**:228.

Withall J, Greaves CJ, Thompson JL, de Koning JL, Bollen JC, Moorlock SJ, *et al.* The tribulations of trials: lessons learnt recruiting 777 older adults into REtirement in ACTION (REACT), a trial of a community, group-based active aging intervention targeting mobility disability. *J Gerontol* 2020;**75**:2387–95.

Cross R, Greaves CJ, Withall J, Rejeski WJ, Stathi A. Delivery fidelity of the REACT (REtirement in ACTION) physical activity and behaviour maintenance intervention for community dwelling older people with mobility limitations. *BMC Public Health* 2022;**22**:1112.

Demnitz N, Stathi A, Withall J, Stainer C, Seager P, de Koning J, *et al.* Hippocampal maintenance after a 12-month physical activity intervention in older adults: the REACT MRI study. *Neuroimage Clin* 2022;**35**:102762.

Snowsill T, Stathi A, Green C, Withall J, Greaves C, Thompson J, *et al.* Cost-effectiveness of a community-based physical activity and behaviour maintenance intervention for preventing decline in physical functioning in older people: an economic evaluation of the REACT (Retirement in Action) intervention. *Lancet Public Health* 2022;**7**:E327–34.

Stathi A, Greaves C, Thompson J, Withall J, Ladlow P, Taylor G, *et al.* Effect of a physical activity and behaviour maintenance programme on functional mobility decline in older adults: the REACT (REtirement in ACTION) randomised controlled trial. *Lancet Public Health* 2022;**7**:E316–26.

## Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review.



## References

1. Stathi A, Withall J, Greaves CJ, Thompson JL, Taylor G, Medina-Lara A, *et al.* A community-based physical activity intervention to prevent mobility-related disability for retired older people (REtirement in ACTion (REACT)): study protocol for a randomised controlled trial. *Trials* 2018;**19**:228. <https://doi.org/10.1186/s13063-018-2603-x>
2. World Health Organisation. *Healthy Ageing and Functional Ability*. 2018. URL: [www.who.int/news-room/questions-and-answers/item/healthy-ageing-and-functional-ability](http://www.who.int/news-room/questions-and-answers/item/healthy-ageing-and-functional-ability) (accessed February 2021).
3. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *Lancet* 2013;**381**:752–62. [https://doi.org/10.1016/S0140-6736\(12\)62167-9](https://doi.org/10.1016/S0140-6736(12)62167-9)
4. Nguyen QD, Wu C, Odden MC, Kim DH. Multimorbidity patterns, frailty, and survival in community-dwelling older adults. *J Gerontol A Biol Sci Med Sci* 2019;**74**:1265–70. <https://doi.org/10.1093/gerona/gly205>
5. NHS Digital. *Adults Health-related Behaviours Health Survey for England 2018*. 2019. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2018> (accessed February 2021).
6. Health and Social Care Information Centre. *Health Survey for England 2014: Health, Social Care and Lifestyles*. London: Health and Social Care Information Centre; 2014. URL: <http://healthsurvey.hscic.gov.uk/media/33473/HSE-2014-summary-report.pdf> (accessed February 2021).
7. Kojima G. Increased healthcare costs associated with frailty among community-dwelling older people: a systematic review and meta-analysis. *Arch Gerontol Geriatr* 2019;**84**:103898. <https://doi.org/10.1016/j.archger.2019.06.003>
8. Office for National Statistics. *Overview of the UK Population: November 2018*. 2018. URL: [www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/november2018](http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/november2018) (accessed February 2021).
9. World Health Organisation. *World Population Ageing*. 2017. URL: [www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017\\_Highlights.pdf](http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf) (accessed February 2021).
10. United States Census Bureau. *Older Population and Aging*. 2019. URL: [www.census.gov/topics/population/older-aging.html](http://www.census.gov/topics/population/older-aging.html) (accessed January 2021).
11. Dipietro L, Campbell WW, Buchner DM, Erickson KI, Powell KE, Bloodgood B, *et al.* Physical activity, injurious falls, and physical function in aging: an umbrella review. *Med Sci Sports Exerc* 2019;**51**:1303–13. <https://doi.org/10.1249/MSS.0000000000001942>
12. Englund DA, Price LL, Grosicki GJ, Iwai M, Kashiwa M, Liu C, *et al.* Progressive resistance training improves torque capacity and strength in mobility-limited older adults. *J Gerontol A Biol Sci Med Sci* 2019;**74**:1316–21. <https://doi.org/10.1093/gerona/gly199>
13. Stathi A, Western M, De Koning J, Perkin O, Withall J. Implementing physical activity programmes for community-dwelling older people with early signs of physical frailty. In Nyman S, editor. *The Palgrave Handbook of Ageing and Physical Activity Promotion*. London: Palgrave Macmillan; 2018. pp. 401–22. [https://doi.org/10.1007/978-3-319-71291-8\\_20](https://doi.org/10.1007/978-3-319-71291-8_20)
14. García-Hermoso A, Ramirez-Vélez R, Sáez de Asteasu ML, Martínez-Velilla N, Zambom-Ferraresi F, Valenzuela PL, *et al.* Safety and effectiveness of long-term exercise interventions in older adults: a systematic review and meta-analysis of randomized controlled trials. *Sports Med* 2020;**50**:1095–106. <https://doi.org/10.1007/s40279-020-01259-y>

15. Cooper R, Kuh D, Cooper C, Gale CR, Lawlor DA, Matthews F, *et al.* Objective measures of physical capability and subsequent health: a systematic review. *Age Ageing* 2011;**40**:14–23. <https://doi.org/10.1093/ageing/afq117>
16. Simmonds B, Fox K, Davis M, Ku PW, Gray S, Hillsdon M, *et al.* Objectively assessed physical activity and subsequent health service use of UK adults aged 70 and over: a four to five year follow up study. *PLOS ONE* 2014;**9**:e97676. <https://doi.org/10.1371/journal.pone.0097676>
17. Jacob ME, Yee LM, Diehr PH, Arnold AM, Thielke SM, Chaves PHM, *et al.* Can a healthy lifestyle compress the disabled period in older adults? *J Am Geriatr Soc* 2016;**64**:1952–61. <https://doi.org/10.1111/jgs.14314>
18. Department of Health and Social Care. *Ageing Well and Supporting People Living with Frailty*. 2018. URL: [www.england.nhs.uk/ourwork/clinical-policy/older-people/frailty/](http://www.england.nhs.uk/ourwork/clinical-policy/older-people/frailty/) (accessed March 2021).
19. World Health Organization. *More Active People for a Healthier World. The Global Action Plan on Physical Activity 2018–2030*. 2018. URL: [www.who.int/news-room/initiatives/gappa](http://www.who.int/news-room/initiatives/gappa) (accessed March 2021).
20. Giné-Garriga M, Roqué-Fíguls M, Coll-Planas L, Sitjà-Rabert M, Salvà A. Physical exercise interventions for improving performance-based measures of physical function in community-dwelling, frail older adults: a systematic review and meta-analysis. *Arch Phys Med Rehabil* 2014;**95**:753–69.e3. <https://doi.org/10.1016/j.apmr.2013.11.007>
21. Pahor M, Guralnik JM, Ambrosius WT, Blair S, Bonds DE, Church TS, *et al.* Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. *JAMA* 2014;**311**:2387–96. <https://doi.org/10.1001/jama.2014.5616>
22. The Marmot Review. *Fair Society, Healthy Lives: Strategic Review of Health Inequalities in England Post-2010*. London: The Marmot Review; 2010.
23. Shaw RJ, Čukić I, Deary IJ, Gale CR, Chastin SF, Dall PM, *et al.* Relationships between socioeconomic position and objectively measured sedentary behaviour in older adults in three prospective cohorts. *BMJ Open* 2017;**7**:e016436. <https://doi.org/10.1136/bmjopen-2017-016436>
24. Centers for Disease Control and Prevention. *CDC Health Disparities and Inequalities Report. Morbidity and Mortality Weekly Report*. 2011. URL: [www.cdc.gov/mmwr/pdf/other/su6001.pdf](http://www.cdc.gov/mmwr/pdf/other/su6001.pdf) (accessed February 2021)
25. Bor J, Cohen GH, Galea S. Population health in an era of rising income inequality: USA, 1980–2015. *Lancet* 2017;**389**:1475–90. [https://doi.org/10.1016/S0140-6736\(17\)30571-8](https://doi.org/10.1016/S0140-6736(17)30571-8)
26. NHS Digital. *Health Survey for England – 2010, Trend Tables*. 2011. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2010-trend-tables> (accessed January 2021).
27. Sport England. *Ethnicity Facts and Figures*. 2019. URL: [www.ethnicity-facts-figures.service.gov.uk/health/diet-and-exercise/physical-activity/latest#data-sources](http://www.ethnicity-facts-figures.service.gov.uk/health/diet-and-exercise/physical-activity/latest#data-sources) (accessed January 2021).
28. GOV.UK. *Ethnicity Facts and Figures*. 2020. URL: [www.ethnicity-facts-figures.service.gov.uk/health/diet-and-exercise/physical-activity/latest](http://www.ethnicity-facts-figures.service.gov.uk/health/diet-and-exercise/physical-activity/latest) (accessed February 2021).
29. Withall J, Greaves CJ, Thompson JL, de Koning JL, Bollen JC, Moorlock SJ, *et al.* The tribulations of trials: lessons learnt recruiting 777 older adults into REtirement in ACTion (REACT), a trial of a community, group-based active aging intervention targeting mobility disability. *J Gerontol A Biol Sci Med Sci* 2020;**75**:2387–95. <https://doi.org/10.1093/gerona/glaa051>
30. Stathi A, Fox KR, Withall J, Bentley G, Thompson JL. *Promoting Physical Activity in Older Adults: A Guide for Local Decision Makers*. 2014. [https://purehost.bath.ac.uk/ws/portalfiles/portal/32338941/AVON\\_strategies\\_report\\_2014\\_Feb.pdf](https://purehost.bath.ac.uk/ws/portalfiles/portal/32338941/AVON_strategies_report_2014_Feb.pdf) (accessed 17 October 2022).



31. Demnitz N, Stathi A, Withall J, Stainer C, Seager P, de Koning J, *et al.* Hippocampal maintenance after a 12-month physical activity intervention in older adults: the REACT MRI study. *Neuroimage Clin* 2021;**35**:102762. <https://doi.org/10.1016/j.nicl.2021.102762>
32. Vasunilashorn S, Coppin AK, Patel KV, Lauretani F, Ferrucci L, Bandinelli S, Guralnik JM. Use of the Short Physical Performance Battery Score to predict loss of ability to walk 400 meters: analysis from the InCHIANTI study. *J Gerontol A Biol Sci Med Sci* 2009;**64**:223–9. <https://doi.org/10.1093/gerona/gln022>
33. European Medicines Agency. *Reflection Paper on Physical Frailty: Instruments for Baseline Characterisation of Older Populations in Clinical Trials*. 2018. URL: [www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-physical-frailty-instruments-baseline-characterisation-older-populations-clinical\\_en.pdf](http://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-physical-frailty-instruments-baseline-characterisation-older-populations-clinical_en.pdf) (accessed January 2021).
34. NHS England. *A Practical Guide to Healthy Ageing*. London: NHS England; 2015. URL: [www.england.nhs.uk/publication/practical-guide-to-healthy-ageing/](http://www.england.nhs.uk/publication/practical-guide-to-healthy-ageing/) (accessed January 2021).
35. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;**14**:377–81. <https://doi.org/10.1249/00005768-198205000-00012>
36. Ntoumanis N, Ng JYY, Prestwich A, Quested E, Hancox JE, Thøgersen-Ntoumani C, *et al.* A meta-analysis of self-determination theory-informed intervention studies in the health domain: effects on motivation, health behavior, physical, and psychological health. *Health Psychol Rev* 2021;**15**:214–44. <https://doi.org/10.1080/17437199.2020.1718529>
37. Ryan RM, Deci EL. A self-determination theory approach to psychotherapy: the motivational basis for effective change. *Canadian Psychology* 2008;**49**:186–93. <https://doi.org/10.1037/a0012753>
38. Deci EL, Ryan RM. *Handbook of Self-determination Research*. Rochester, NY: The University of Rochester Press; 2002.
39. Bandura A. Health promotion from the perspective of social cognitive theory. *Psychology & Health* 1998;**13**:623–49. <https://doi.org/10.1080/08870449808407422>
40. Greaves C, Poltawski L, Garside R, Briscoe S. Understanding the challenge of weight loss maintenance: a systematic review and synthesis of qualitative research on weight loss maintenance. *Health Psychol Rev* 2017;**11**:145–63. <https://doi.org/10.1080/17437199.2017.1299583>
41. Poltawski L, van Beurden SB, Morgan-Trimmer S, Greaves C. The dynamics of decision-making in weight loss and maintenance: a qualitative enquiry. *BMC Public Health* 2020;**20**:573. <https://doi.org/10.1186/s12889-020-08664-y>
42. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, *et al.* Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol* 2004;**23**:443–51. <https://doi.org/10.1037/0278-6133.23.5.443>
43. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, *et al.* Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258. <https://doi.org/10.1136/bmj.h1258>
44. Perera S, Studenski S, Chandler JM, Guralnik JM. Magnitude and patterns of decline in health and function in 1 year affect subsequent 5-year survival. *J Gerontol A Biol Sci Med Sci* 2005;**60**:894–900. <https://doi.org/10.1093/gerona/60.7.894>
45. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. *Clin Pharmacol Ther* 1974;**15**:443–53. <https://doi.org/10.1002/cpt1974155443>



46. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 1975;**31**:103–15. <https://doi.org/10.2307/2529712>
47. Sampson M, Clark A, Bachmann M, Garner N, Irvine L, Howe A, *et al.* Lifestyle intervention with or without lay volunteers to prevent type 2 diabetes in people with impaired fasting glucose and/or nondiabetic hyperglycemia: a randomized clinical trial. *JAMA Internal Med* 2020;**181**:168–78. <https://doi.org/10.1001/jamainternmed.2020.5938>
48. Flight L, Allison A, Dimairo M, Lee E, Mandefield L, Walters SJ. Recommendations for the analysis of individually randomised controlled trials with clustering in one arm – a case of continuous outcomes. *BMC Med Res Methodol* 2016;**16**:165. <https://doi.org/10.1186/s12874-016-0249-5>
49. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;**46**:153–62. [https://doi.org/10.1016/0895-4356\(93\)90053-4](https://doi.org/10.1016/0895-4356(93)90053-4)
50. Stathi A, Fox K. The dimensions of a well-being scale designed for older adults: the Ageing-Well Profile. *J Aging Phys Act* 2004;**12**:300.
51. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**20**:1727–36. <https://doi.org/10.1007/s11136-011-9903-x>
52. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**:271–92. [https://doi.org/10.1016/S0167-6296\(01\)00130-8](https://doi.org/10.1016/S0167-6296(01)00130-8)
53. Rejeski WJ, Rushing J, Guralnik JM, Ip EH, King AC, Manini TM, *et al.* The MAT-sf: identifying risk for major mobility disability. *J Gerontol A Biol Sci Med Sci J* 2015;**70**:641–6. <https://doi.org/10.1093/gerona/glv003>
54. Lyall DM, Cullen B, Allerhand M, Smith DJ, Mackay D, Evans J, *et al.* Cognitive Test Scores in UK biobank: data reduction in 480,416 participants and longitudinal stability in 20,346 participants. *PLOS ONE* 2016;**11**:e0154222. <https://doi.org/10.1371/journal.pone.0154222>
55. Cornelis MC, Wang Y, Holland T, Agarwal P, Weintraub S, Morris MC. Age and cognitive decline in the UK Biobank. *PLOS ONE* 2019;**14**:e0213948. <https://doi.org/10.1371/journal.pone.0213948>
56. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, *et al.* The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;**53**:695–9. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
57. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;**15**:1833–40.
58. Espie CA, Kyle SD, Hames P, Gardani M, Fleming L, Cape J. The Sleep Condition Indicator: a clinical screening tool to evaluate insomnia disorder. *BMJ Open* 2014;**4**:e004183. <https://doi.org/10.1136/bmjopen-2013-004183>
59. Kempen GI, Todd CJ, Van Haastregt JC, Zijlstra GA, Beyer N, Freiberger E, *et al.* Cross-cultural validation of the Falls Efficacy Scale International (FES-I) in older people: results from Germany, the Netherlands and the UK were satisfactory. *Disabil Rehabil* 2007;**29**:155–62. <https://doi.org/10.1080/09638280600747637>
60. Stathi A, Greaves C, Thompson J, Withall J, Ladlow P, Taylor G, *et al.* Effect of a physical activity and behaviour maintenance programme on functional mobility decline in older adults: the REACT (REtirement in ACTion) randomised controlled trial. *Lancet Public Health* 2022;**7**:E316–26. [https://doi.org/10.1016/S2468-2667\(22\)00004-4](https://doi.org/10.1016/S2468-2667(22)00004-4)

61. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;**8**:18. <https://doi.org/10.1186/1741-7015-8-18>
62. Office for National Statistics. *Ethnicity Facts and Figures*. 2018. URL: [www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/age-groups/latest#age-profile-by-ethnicity](http://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/age-groups/latest#age-profile-by-ethnicity) (accessed January 2021).
63. Office for National Statistics. *Population Estimates for the UK, England and Wales, Scotland and Northern Ireland: Mid-2018*. 2019. URL: [www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2018](http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2018) (accessed January 2021).
64. Office for National Statistics. *Populations by Sex, Age Group and Index of Multiple Deprivation (IMD) Quintile, England, 2001 to 2017*. 2018. URL: [www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/009283populationsbysexagegroupandimdquintileengland2001to2017](http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/009283populationsbysexagegroupandimdquintileengland2001to2017) (accessed January 2021).
65. Newman AB, Simonsick EM, Naydeck BL, Boudreau RM, Kritchevsky SB, Nevitt MC, *et al*. Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. *JAMA* 2006;**295**:2018–26. <https://doi.org/10.1001/jama.295.17.2018>
66. Perera S, Studenski S, Newman A, Simonsick E, Harris T, Schwartz A, Visser M, for the Health ABC Study. Are estimates of meaningful decline in mobility performance consistent among clinically important subgroups? (Health ABC study). *J Gerontol A Biol Sci Med Sci* 2014;**69**:1260–8. <https://doi.org/10.1093/gerona/glu033>
67. Jefferis BJ, Parsons TJ, Sartini C, Ash S, Lennon LT, Papacosta O, *et al*. Objectively measured physical activity, sedentary behaviour and all-cause mortality in older men: does volume of activity matter more than pattern of accumulation? *Br J Sports Med* 2019;**53**:1013–20. <https://doi.org/10.1136/bjsports-2017-098733>
68. Badhiwala JH, Witiw CD, Nassiri F, Akbar MA, Jaja B, Wilson JR, Fehlings MG. Minimum clinically important difference in SF-36 scores for use in degenerative cervical myelopathy. *Spine* 2018;**43**:E1260–E1266. <https://doi.org/10.1097/BRS.0000000000002684>
69. Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. *Spine J* 2010;**10**:469–74. <https://doi.org/10.1016/j.spinee.2010.02.007>
70. Bohannon RW. Minimal clinically important difference for grip strength: a systematic review. *J Phys Ther Sci* 2019;**31**:75–8. <https://doi.org/10.1589/jpts.31.75>
71. Cross R, Greaves CJ, Withall J, Rejeski JW, Stathi A. Delivery fidelity of the REACT (REtirement in ACTion) physical activity and behaviour maintenance intervention for community dwelling older people with mobility limitations. *BMC Public Health* 2022;**22**:1112. <https://doi.org/10.1186/s12889-022-13496-z>
72. Dreyfus SE. The five-stage model of adult skill acquisition. *Bull Sci Technol Soc* 2004;**24**:177–82. <https://doi.org/10.1177/0270467604264992>
73. Frost J, Wingham J, Britten N, Greaves C, Abraham C, Warren FC, *et al*. Home-based rehabilitation for heart failure with reduced ejection fraction: mixed methods process evaluation of the REACH-HF multicentre randomised controlled trial. *BMJ Open* 2019;**9**:e026039. <https://doi.org/10.1136/bmjopen-2018-026039>

74. Thompson TP, Lambert JD, Greaves CJ, Taylor AH. Intervention delivery fidelity assessment of a counselling based intervention for promoting smoking reduction and increasing physical activity. *Health Psychol* 2018;**37**:627–37. <https://doi.org/10.1037/hea0000613>
75. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol* 2008;**41**:327–50. <https://doi.org/10.1007/s10464-008-9165-0>
76. Lambert JD, Greaves CJ, Farrand P, Cross R, Haase AM, Taylor AH. Assessment of fidelity in individual level behaviour change interventions promoting physical activity among adults: a systematic review. *BMC Public Health* 2017;**17**:765. <https://doi.org/10.1186/s12889-017-4778-6>
77. Devereux-Fitzgerald A, Powell R, Dewhurst A, French DP. The acceptability of physical activity interventions to older adults: A systematic review and meta-synthesis. *Soc Sci Med* 2016;**158**:14–23. <https://doi.org/10.1016/j.socscimed.2016.04.006>
78. Zubala A, MacGillivray S, Frost H, Kroll T, Skelton DA, Gavine A, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. *PLOS ONE* 2017;**12**:e0180902. <https://doi.org/10.1371/journal.pone.0180902>
79. French DP, Olander EK, Chisholm A, Mc Sharry J. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. *Ann Behav Med* 2014;**48**:225–34. <https://doi.org/10.1007/s12160-014-9593-z>
80. Ashford S, Edmunds J, French DP. What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. *Br J Health Psychol* 2010;**15**:265–88. <https://doi.org/10.1348/135910709X461752>
81. Borek AJ, Abraham C. How do small groups promote behaviour change? An integrative conceptual review of explanatory mechanisms. *Appl Psychol Health Well Being* 2018;**10**:30–61. <https://doi.org/10.1111/aphw.12120>
82. Farrow C, Tarrant M, Khan S. Using social identity to promote health: the impact of group memberships on health in the context of obesity. In Buckingham S, Best D, editors. *Addiction, Behavioral Change and Social Identity: The Path to Resilience and Recovery*. New York, NY: Routledge; 2016. pp. 52–70.
83. Khan SS, Tarrant M, Kos K, Daly M, Gimbuta C, Farrow CV. Making connections: social identification with new treatment groups for lifestyle management of severe obesity. *Clin Psychol Psychother* 2020;**27**:686–96. <https://doi.org/10.1002/cpp.2454>
84. Borek AJ, Smith JR, Greaves CJ, Gillison F, Tarrant M, Morgan-Trimmer S, et al. Developing and applying a framework to understand mechanisms of action in group-based, behaviour change interventions: the MAGI mixed-methods study. *Efficacy and Mech Eval* 2019;**6**(3). <https://doi.org/10.3310/eme06030>
85. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011;**71**(Suppl. 1):52–63. <https://doi.org/10.1111/j.1752-7325.2011.00233.x>
86. Patton MQ. *Qualitative Evaluation and Research Methods*. 2nd edn. London: SAGE Publications Ltd; 1990.
87. Ritchie J, Lewis J. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd; 2003.
88. Lincoln YS, Guba EG. *Naturalistic Inquiry*. London: SAGE Publications Ltd; 1985. [https://doi.org/10.1016/0147-1767\(85\)90062-8](https://doi.org/10.1016/0147-1767(85)90062-8)
89. Nevedal AL, Ayalon L, Briller SH. A qualitative evidence synthesis review of longitudinal qualitative research in gerontology. *Gerontologist* 2019;**59**:e791–e801. <https://doi.org/10.1093/geront/gny134>

90. Maula A, LaFond N, Orton E, Iliffe S, Audsley S, Vedhara K, Kendrick D. Use it or lose it: a qualitative study of the maintenance of physical activity in older adults. *BMC Geriatr* 2019;**19**:349. <https://doi.org/10.1186/s12877-019-1366-x>
91. Moore AJ, Holden MA, Foster NE, Jinks C. Therapeutic alliance facilitates adherence to physiotherapy-led exercise and physical activity for older adults with knee pain: a longitudinal qualitative study. *J Physiother* 2020;**66**:45–53. <https://doi.org/10.1016/j.jphys.2019.11.004>
92. Sims-Gould J, Miran-Khan K, Haggis C, Liu-Ambrose T. Timing, experience, benefits, and barriers: older women's uptake and adherence to an exercise program. *Act Adapt Aging* 2012;**36**:280–96. <https://doi.org/10.1080/01924788.2012.729188>
93. Walker KC, Valentiner LS, Langberg H. Motivational factors for initiating, implementing, and maintaining physical activity behavior following a rehabilitation program for patients with type 2 diabetes: a longitudinal, qualitative, interview study. *Patient Prefer Adherence* 2018;**12**:145–52. <https://doi.org/10.2147/PPA.S150008>
94. Curnan S, LaCava L, Langenburg D, Lelle M, Reece M. *WK Kellogg Foundation Evaluation Handbook*. 1998. URL: [www.wkkf.org/pubs/Tools/Evaluation/Pub770.pdf](http://www.wkkf.org/pubs/Tools/Evaluation/Pub770.pdf) (accessed January 2021).
95. W.K. Kellogg Foundation. *Logic Model Development Guide*. Battle Creek, MI: W.K Kellogg Foundation; 2004. URL: [www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide](http://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide) (accessed January 2021).
96. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013;**13**:117. <https://doi.org/10.1186/1471-2288-13-117>
97. Hayes AF. *Introduction to Mediation, Moderation and Conditional Process Analysis: A Regression Based Approach*. New York, NY: Guildford Press; 2018.
98. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall; 1986.
99. Graham Beaumont J. and Kenealy P. Quality of life perceptions and social comparisons in healthy old age. *Ageing Soc* 2004;**24**:755–69. <https://doi.org/10.1017/S0144686X04002399>
100. Silfee VJ, Haughton CF, Jake-Schoffman DE, Lopez-Cepero A, May CN, Sreedhara M, et al. Objective measurement of physical activity outcomes in lifestyle interventions among adults: a systematic review. *Prev Med Rep* 2018;**11**:74–80. <https://doi.org/10.1016/j.pmedr.2018.05.003>
101. Guo W, Key TJ, Reeves GK. Accelerometer compared with questionnaire measures of physical activity in relation to body size and composition: a large cross-sectional analysis of UK Biobank. *BMJ Open* 2019;**9**:e024206. <https://doi.org/10.1136/bmjopen-2018-024206>
102. Harris TJ, Owen CG, Victor CR, Adams R, Ekelund U, Cook DG. A comparison of questionnaire, accelerometer, and pedometer: measures in older people. *Med Sci Sports Exerc* 2009;**41**:1392–402. <https://doi.org/10.1249/MSS.0b013e31819b3533>
103. Snowsill T, Stathi A, Green C, Withall J, Greaves C, Thompson J, et al. Cost-effectiveness of a community-based physical activity and behaviour maintenance intervention for preventing decline in physical functioning in older people: an economic evaluation of the REACT (Retirement in Action) intervention. *Lancet Public Health* 2022;**7**:E327–34.
104. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal* 2013. London: NICE; 2013. URL: [www.nice.org.uk/process/pmg9/chapter/the-reference-case](http://www.nice.org.uk/process/pmg9/chapter/the-reference-case) (accessed February 2021).
105. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. <https://doi.org/10.1097/00005650-199711000-00002>

## REFERENCES

106. National Institute for Health and Care Excellence (NICE). *Position Statement on Use of the EQ-5D-5L Value Set for England (Updated October 2019)*. London: NICE; 2019. URL: [www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l](http://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l) (accessed 17 October 2022).
107. 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>
108. Kharroubi SA, Brazier JE, Roberts J, O'Hagan A. Modelling SF-6D health state preference data using a nonparametric Bayesian method. *J Health Econ* 2007;**26**:597–612. <https://doi.org/10.1016/j.jhealeco.2006.09.002>
109. Brazier J, Ratcliffe J, Salomon J, Tsuchiya A. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford: Oxford University Press; 2007.
110. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: PSSRU, University of Kent; 2018.
111. NHS Improvement. *National Cost Collection for the NHS*. 2020. <https://improvement.nhs.uk/resources/national-cost-collection/> (accessed January 2021).
112. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: PSSRU, University of Kent; 2014.
113. NHS Improvement. *Reference Costs, 2018*. <https://improvement.nhs.uk/resources/reference-costs/> (accessed January 2021).
114. van den Berg B, Ferrer-i-Carbonell A. Monetary valuation of informal care: the well-being valuation method. *Health Econ* 2007;**16**:1227–44. <https://doi.org/10.1002/hec.1224>
115. Organisation for Economic Cooperation and Development (OECD). *Purchasing Power Parities (PPP)*. 2017. URL: [www.oecd-ilibrary.org/content/data/1290ee5a-en](http://www.oecd-ilibrary.org/content/data/1290ee5a-en) (accessed January 2021).
116. Office for National Statistics. *Employee Earnings in the UK: 2019*. 2019. URL: [www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2019](http://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2019) (accessed January 2021).
117. Newell D, Diment E, Bolton JE. An electronic patient-reported outcome measures system in UK chiropractic practices: a feasibility study of routine collection of outcomes and costs. *J Manipulative Physiol Ther* 2016;**39**:31–41. <https://doi.org/10.1016/j.jmpt.2015.12.001>
118. General Osteopathic Council. *Visiting an Osteopath: What to Expect*. 2020. URL: [www.osteopathy.org.uk/visiting-an-osteopath/what-to-expect/](http://www.osteopathy.org.uk/visiting-an-osteopath/what-to-expect/) (accessed January 2021).
119. Nixon RM, Thompson SG. Parametric modelling of cost data in medical studies. *Stat Med* 2004;**23**:1311–31. <https://doi.org/10.1002/sim.1744>
120. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. <https://doi.org/10.1002/hec.944>
121. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. Oxford: Oxford University Press; 2014. <https://doi.org/10.1093/med/9780199685028.001.0001>
122. Campbell Collaboration, ePPI-Centre. *CCEMG – EPPI-Centre Cost Converter*. 2020. URL: <https://eppi.ioe.ac.uk/costconversion/default.aspx> (accessed January 2021).
123. Incisive Health. *An International Comparison of Long-Term Care Funding and Outcomes: Insights for the Social Care Green Paper*. 2018. URL: [www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/reports-and-briefings/care-support/rb\\_aug18\\_international\\_comparison\\_of\\_social\\_care\\_funding\\_and\\_outcomes.pdf](http://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/reports-and-briefings/care-support/rb_aug18_international_comparison_of_social_care_funding_and_outcomes.pdf) (accessed January 2021).



124. Alhambra-Borrás T, Durá-Ferrandis E, Ferrando-García M. Effectiveness and estimation of cost-effectiveness of a group-based multicomponent physical exercise programme on risk of falling and frailty in community-dwelling older adults. *Int J Environ Res Public Health* 2019;**16**:E2086. <https://doi.org/10.3390/ijerph16122086>
125. Groessl EJ, Kaplan RM, Castro Sweet CM, Church T, Espeland MA, Gill TM, *et al.* Cost-effectiveness of the LIFE physical activity intervention for older adults at increased risk for mobility disability. *J Gerontol A Biol Sci Med Sci* 2016;**71**:656–62. <https://doi.org/10.1093/gerona/glw001>
126. Filipovic-Pierucci A, Zarca K, Durand-Zaleski I. *Markov Models for Health Economic Evaluation: The R Package heemod*. 2017. URL: <https://arxiv.org/pdf/1702.03252.pdf>. <https://doi.org/10.1016/j.jval.2016.09.133> (accessed January 2021).
127. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;**13**:322–38. <https://doi.org/10.1177/0272989X9301300409>
128. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV, *et al.* Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci* 2000;**55**:M221–31. <https://doi.org/10.1093/gerona/55.4.m221>
129. Pavasini R, Guralnik J, Brown JC, di Bari M, Cesari M, Landi F, *et al.* Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis. *BMC Med* 2016;**14**:215. <https://doi.org/10.1186/s12916-016-0763-7>
130. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to review the role of surrogate end points in health policy: state of the art and the way forward. *Value Health* 2017;**20**:487–95. <https://doi.org/10.1016/j.jval.2016.10.011>
131. Office for National Statistics. *Estimates of the Population for the UK, England and Wales, Scotland and Northern Ireland*. 2019. URL: [www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesScotlandandNorthernIreland](http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesScotlandandNorthernIreland) (accessed February 2021).
132. Office for National Statistics. *Mid-year Population Estimates of the Very Old, Including Centenarians: UK*. 2019. URL: [www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/datasets/midyearpopulationestimatesoftheveryoldincludingcentenariansunitedkingdom](http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/datasets/midyearpopulationestimatesoftheveryoldincludingcentenariansunitedkingdom) (accessed January 2021).
133. Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med* 1989;**8**:551–61. <https://doi.org/10.1002/sim.4780080504>
134. Callahan KE, Lovato L, Miller ME, Marsh AP, Fielding RA, Gill TM, *et al.* Self-reported physical function as a predictor of hospitalization in the lifestyle interventions and independence for elders study. *J Am Geriatr Soc* 2018;**66**:1927–33. <https://doi.org/10.1111/jgs.15468>
135. Penninx BW, Ferrucci L, Leveille SG, Rantanen T, Pahor M, Guralnik JM. Lower extremity performance in nondisabled older persons as a predictor of subsequent hospitalization. *J Gerontol A Biol Sci Med Sci* 2000;**55**:M691–7. <https://doi.org/10.1093/gerona/55.11.m691>
136. Hebert JR, Hurley TG, Peterson KE, Resnicow K, Thompson FE, Yaroch AL, *et al.* Social desirability trait influences on self-reported dietary measures among diverse participants in a multicenter multiple risk factor trial. *J Nutr* 2008;**138**:226S–234S. <https://doi.org/10.1093/jn/138.1.226S>
137. Maddison R, Rawstorn JC, Stewart RAH, Benatar J, Whittaker R, Rolleston A, *et al.* Effects and costs of real-time cardiac telerehabilitation: randomised controlled non-inferiority trial. *Heart* 2019;**105**:122–9. <https://doi.org/10.1136/heartjnl-2018-313189>

## REFERENCES

138. US Department of Health & Human Services. *Physical Activity Guidelines for America*. 2018. URL: [https://health.gov/sites/default/files/2019-09/Physical\\_Activity\\_Guidelines\\_2nd\\_edition.pdf](https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf) (accessed January 2021).
139. Department for Health and Social Care. *UK Chief Medical Officers' Physical Activity Guidelines*. 2019. URL: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/832868/uk-chief-medical-officers-physical-activity-guidelines.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/832868/uk-chief-medical-officers-physical-activity-guidelines.pdf) (accessed January 2021).
140. Department of Health and Social Care. *NHS Reference Costs 2013 to 2014*. 2015. URL: [www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014](http://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014) (accessed January 2021).
141. Mentzakis E, Ryan M, McNamee P. Using discrete choice experiments to value informal care tasks: exploring preference heterogeneity. *Health Econ* 2011;**20**:930–44. <https://doi.org/10.1002/hec.1656>
142. Mentzakis E, McNamee P, Ryan M, Sutton M. Valuing informal care experience: does choice of measure matter? *Soc Indic Res* 2012;**108**:169–84. <https://doi.org/10.1007/s11205-011-9873-y>
143. 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>
144. Office for National Statistics. *EMP04: Employment by Occupation (April to June 2018)*. 2018. URL: [www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/employmentbyoccupationemp04](http://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/employmentbyoccupationemp04) (accessed February 2021).
145. Callaghan L, Thompson TP, Creanor S, Quinn C, Senior J, Green C, *et al*. Individual health trainers to support health and well-being for people under community supervision in the criminal justice system: the STRENGTHEN pilot RCT. *Public Health Res* 2019;**7**(20). <https://doi.org/10.3310/phr07200>
146. Department for Communities and Local Government. *English Indices of Deprivation 2019 – LSOA Level*. 2019. URL: <http://opendatacommunities.org/data/societal-wellbeing/imd2019/indices> (accessed January 2021).
147. Glanville J, Kaunelis D, Mensinkai S. How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. *Int J Technol Assess Health Care* 2009;**25**:522–9. <https://doi.org/10.1017/S0266462309990523>
148. Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being (QWB-SA) questionnaire among older adults. *Med Care* 1998;**36**:1349–60. <https://doi.org/10.1097/00005650-199809000-00007>
149. Cabrero-García J, Muñoz-Mendoza CL, Cabañero-Martínez MJ, González-Llopis L, Ramos-Pichardo JD, Reig-Ferrer A. [Short physical performance battery reference values for patients 70 years-old and over in primary health care.] *Aten Primaria* 2012;**44**:540–8. <https://doi.org/10.1016/j.aprim.2012.02.007>
150. Bergland A, Strand BH. Norwegian reference values for the Short Physical Performance Battery (SPPB): the Tromsø Study. *BMC Geriatr* 2019;**19**:216. <https://doi.org/10.1186/s12877-019-1234-8>
151. Ramírez-Vélez R, Pérez-Sousa MA, Venegas-Sanabria LC, Cano-Gutierrez CA, Hernández-Quinonez PA, Rincón-Pabón D, *et al*. Normative values for the Short Physical Performance Battery (SPPB) and their association with anthropometric variables in older Colombian adults. The SABE study, 2015. *Front Med* 2020;**7**:52. <https://doi.org/10.3389/fmed.2020.00052>
152. Office for National Statistics. *Mortality Rates (qx), Principal Projection, UK: 2018-based Edition of this Dataset*. 2019. URL: [www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/mortalityratesqxprincipalprojectionunitedkingdom](http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/mortalityratesqxprincipalprojectionunitedkingdom) (accessed January 2021).

## Appendix 1 Randomised controlled trial

Parts of this appendix have been adapted from Stathi *et al.*<sup>1</sup> Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The text includes minor additions and formatting changes to the original text.

Parts of this appendix have been adapted from Withall *et al.*<sup>29</sup> 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. The text includes minor additions and formatting changes to the original text.

Parts of this appendix have been adapted from Stathi *et al.*<sup>60</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

TABLE 20 The REACT outcome measures and time points

Outcome measure	Visit type				
	Screening	F06	F12	F24	
Visit code		SV1	F06	F12	F24
Visit number		1	2	3	4
Telephone call	1				
Activity/assessment month	-0.5	0	6	12	24
Form name					
Verbal consent	X				
Telephone screening	X				
Written informed consent		X			
Contact information update	X	X	X	X	X
Demographic data					
Age	X				
Sex	X				
Ethnicity	X				
BMI (weight only at 12 months)		X		X	X
MoCA		X	X	X	X
Highest education level	X				
IMD	X				
Caring responsibilities		X			
Marital status		X	X	X	X

continued



TABLE 20 The REACT outcome measures and time points (continued)

Outcome measure	Visit type			
	Screening	Follow-up		
Home ownership	X			
Number of chronic illnesses	X	X	X	X
SPPB	X	X	X	X
Accelerometry	X	X	X	X
Subjective Physical Activity (PASE questionnaire-10 item)	X	X	X	X
Muscle-Strengthening Exercise Adherence Questionnaire	X	X	X	X
Dynamometer (hand-grip strength)	X	X	X	X
Ageing Well profile (6 items social well-being scale, only at Bath/Bristol site)	X		X	X
Sleep Condition Indicator	X	X	X	X
Pain (WOMAC)	X			X
Loneliness	X	X	X	X
Health-related quality of life (EQ-5D, SF-36)	X	X	X	X
MAT-SF	X	X	X	X
Cognitive function (UK Biobank Healthy Minds Questionnaire)	X	X	X	X
Falls Inventory	X	X	X	X
Health and Social Service Usage	X	X	X	X
Process evaluation				
Session attendance (intervention group only)		X	X	
Total contact time for each participant		X	X	
Physical activity-related self-concept	X	X	X	
Perceived tension of maintaining current physical activity		X	X	X
Perceived tension of maintaining current exercise		X	X	X
Autonomy in relation to physical activity	X	X	X	X
Competence for physical activity	X	X	X	X
Relatedness for physical activity	X	X	X	X
Enjoyment of physical activity	X	X	X	X
Perceived intrinsic benefits of physical activity (social, physical and emotional)	X	X	X	X
Autonomy for strength-building exercise	X	X	X	X
Competence for strength-building exercise	X	X	X	X
Relatedness for strength-building exercise	X	X	X	X
Enjoyment of strength-building exercise	X	X	X	X
Perceived intrinsic benefits of strength-building exercise (social, physical and emotional)	X	X	X	X
Enjoyment of the REACT programme (intervention group)		X	X	
Functional MRI imaging sub-study				
MRI scan, detailed cognitive assessment and gait analysis	X	X	X	

F06, follow-up at 6 months; F12, follow-up at 12 months; F24, follow-up at 24 months; SV1, screening visit 1.

TABLE 21 The REACT study baseline characteristics compared with the UK population aged over 65 years

Characteristic	REACT (N = 777), n (%)	UK general population aged over 65 years (N = 9,223,073), n (%)
Age (years)		
65–69	95 (12.2)	2,674,161 (29.0)
70–74	191 (24.6)	2,178,672 (23.6)
75–79	190 (24.5)	1,777,547 (19.3)
80–84	160 (20.6)	1,338,005 (14.5)
≥ 85	141 (18.1)	1,254,688 (13.6)
Sex		
Female	514 (66.2)	6,617,318 (54.4)
Male	263 (33.9)	5,548,239 (45.6)
Race/ethnicity		
Caucasian/white <sup>a</sup>	739 (95.1)	8,806,190 (95.5)
African/Caribbean	23 (3.0)	115,288 (1.3)
Asian	9 (1.2)	238,878 (2.6)
Other/mixed	6 (0.8)	60,872 (0.7)
IMD <sup>b</sup>		
Quintile 1	86 (11.1)	1,321,666 (14.3)
Quintile 2	157 (20.2)	1,618,649 (17.6)
Quintile 3	159 (20.5)	1,975,582 (21.4)
Quintile 4	156 (20.1)	2,127,763 (23.7)
Quintile 5	219 (28.2)	2,180,334 (23.6)

a Total of all the white categories (white British, white Irish, other white background).

b IMD Q1 is most deprived.

**Note**

Some of the percentages do not total 100% because of rounding.

TABLE 22 Study withdrawals

Group	No. of participants (%)		
	Total	Control arm	Intervention arm
Control	52 (46.85)		
Intervention	59 (53.15)		
Male	45 (40.54)	25 (48.08)	25 (42.37)
Female	66 (59.46)	27 (51.92)	39 (66.10)
Bath/Bristol	47 (42.34)	24 (46.15)	23 (38.98)
Birmingham	18 (16.22)	11 (21.15)	7 (11.86)
Devon	46 (41.44)	17 (32.69)	29 (49.15)
65–74 years	24 (21.62)	14 (29.92)	10 (16.95)
≥ 75 years	87 (78.38)	38 (73.08)	49 (83.05)
Reasons			
1: Excluded	4 (3.60)	0 (0.00)	4 (6.78)
2: Choice	34 (30.63)	17 (32.69)	17 (28.81)
3: Health issues (participant)	28 (25.23)	13 (25.00)	15 (25.42)
4: Health issues (family)	8 (7.21)	4 (7.69)	4 (6.78)
5: Death	16 (14.41)	8 (15.38)	8 (13.56)
6: Other	21 (18.92)	10 (19.23)	11 (18.64)
Total	111 (100)	52 (100)	59 (100)

TABLE 23 Loss to follow-up

Group	No. of participants enrolled	6 months		12 months		24 months	
		No. with valid SPPB score	No. (%) with missing SPPB score	No. with valid SPPB score	No. (%) with missing SPPB score	No. with valid SPPB score	No. (%) with missing SPPB score
Bath/Bristol	335	283	52 (15.5)	280	55 (16.4)	275	60 (17.9)
Devon	268	222	46 (17.2)	215	53 (19.8)	209	59 (22.0)
Birmingham	174	154	20 (11.5)	154	20 (11.5)	144	30 (17.2)
Overall	777	659	118 (15.2)	649	128 (16.5)	628	149 (19.2)

TABLE 24 Details of unexpected SAEs

Study number	Allocation	Type	Related to intervention
AT1464	Intervention	Deceased: unknown cause	Not related
AT1564	Intervention	Stroke	Not related
AT1767	Intervention	Scheduled hip operation	Not related
AT1832	Intervention	Heart attack	Not related
AT1905	Control	Deceased: unknown cause	Not related
AT1948	Intervention	Deceased: unknown cause	Not related
AT2171	Control	Deceased: unknown cause	Not related
AT2187	Intervention	Stomach ulcer	Not related
AT2275	Intervention	Mini stroke	Possible
AT2305	Intervention	Fall	Not related
AT2341	Intervention	Deceased: unknown cause	Not related
AT2414	Intervention	Deceased: stroke	Not related
AT2486	Intervention	Deceased: unknown cause	Not related
AT2536	Intervention	Fall: fractured hip	Related
AT2630	Control	Deceased: unknown cause	Not related
AT2648	Intervention	Deceased: unknown cause	Not related
AT2675	Intervention	Deceased: frailty in old age	Not related
BT1231	Intervention	Deceased: Alzheimer's disease	Not related
BT1348	Intervention	Cardiac event (non-fatal)	Not related
BT1585	Control	Cancer diagnosis	Not related
BT1636	Intervention	Severe nose bleeds	Not related
BT1704	Control	Deceased: cardiac event	Not related
BT1753	Control	Deceased: suspected cardiac arrest	Not related
BT1770	Intervention	Cancer diagnosis	Not related
BT1785	Intervention	Heart attack	Not related
BT1979	Control	Fall: broken shoulder	Not related
BT2002	Control	Bowel cancer diagnosis	Not related
ET1036	Control	Fall: injured arm and head (hospitalised). Diagnosed with pneumonia and kidney issues	Not related
ET1044	Control	Stroke	Not related
ET1073	Intervention	Diarrhoea, dehydration and rheumatoid arthritis	Not related
ET1295	Control	Stomach pains, a hernia and memory deterioration	Not related
ET1392	Intervention	Sciatica which aggravated MS	Not related
ET1392	Intervention	Spinal surgery, fractured sacrum owing to surgery	Not related
ET1393	Control	Fall: fractured head of femur, led to half-hip replacement	Not related
ET1425	Intervention	Non-Hodgkin's lymphoma	Not related
ET1439	Intervention	Spine compression	Not related

continued

TABLE 24 Details of unexpected SAEs (continued)

Study number	Allocation	Type	Related to intervention
ET1506	Intervention	Elective surgery resulting in sepsis and pneumonia	Not related
ET1565	Intervention	Fall: socket of femur break. Hip replacement needed	Not related
ET1603	Control	Bad reaction to flu injection necessitating cardioversion	Not related
ET1689	Control	Sepsis	Not related
ET1717	Intervention	Pneumonia	Not related
ET1752	Control	Deceased: pneumonia, emphysema and lung cancer	Not related
ET1754	Intervention	Terminal cancer diagnosis	Not related
ET1804	Control	Asthma attack (hospitalised), followed by fall resulting in brain bleed	Not related
ET1844	Control	Collapsed at home: suspected bleeding ulcer	Not related
ET1866	Intervention	Deceased: colon cancer	Not related
ET1953	Intervention	Knee injury	Not related
ET1959	Control	Fall: double fracture and dislocation of left ankle	Not related
ET1969	Intervention	Fall: cracked ribs	Not related
ET1977	Intervention	Fall: broken hip	Not related
ET1999	Intervention	Elective surgery: swollen testicles	Not related
ET2017	Control	TIA (hospitalised)	Not related
ET2024	Control	Fall: broken shoulder	Not related
ET2024	Control	Dementia diagnosis	Not related
ET2035	Control	Two stents fitted and aortic valve replaced	Not related
ET2036	Intervention	Hernia (surgery performed)	Not related
ET2045	Intervention	Chest pains (high troponin levels), contracted flu in hospital	Not related
ET2067	Intervention	Pneumonia (hospitalised). Minor heart attack while hospitalised	Not related
ET2067	Intervention	Pneumonia, second bout (hospitalised)	Not related
ET2069	Control	Fall: fractured humerus	Not related
ET2069	Control	Planned knee replacement	Not related
ET2081	Intervention	TIA	Not related
ET2092	Control	Sepsis	Not related
ET2117	Intervention	Bleeding (details unknown)	Not related
ET2128	Control	Deceased: unknown cause	Not related
ET1281	Intervention	Cardiac issues	Not related
ET2191	Intervention	Triple fusion of ankle joint	Not related
ET2205	Intervention	Cellulitis	Not related
ET2255	Intervention	Stroke	Not related
ET2255	Intervention	Fall (hospitalised)	Not related
ET2290	Intervention	Stroke	Not related
ET2309	Intervention	Lung cancer	Not related
ET2374	Control	Accident: fractured hip resulting in replacement	Not related

TABLE 24 Details of unexpected SAEs (continued)

Study number	Allocation	Type	Related to intervention
ET2377	Control	Stroke	Not related
ET2402	Control	Fall: haematoma on leg	Not related
ET2401	Intervention	Pneumonia (hospitalised). Treatment caused AF and bleed to abdomen	Not related
ET2416	Intervention	Planned knee replacement	Not related
ET2416	Intervention	Problems with knee replacement	Not related
ET2425	Intervention	Terminal bone cancer	Not related
ET2449	Intervention	Deceased: heart attack and possible appendicitis	Not related
ET2455	Intervention	Heart attack: three stents fitted	Not related
ET2472	Control	Chest pains: stents fitted	Not related
ET2481	Control	TIA	Not related
ET2511	Intervention	Big toe removed	Not related
ET2528	Control	Cardiac 'episode' (hospitalised)	Not related
ET2357	Intervention	Hospitalisation (pneumonia)	Not related
ET2588	Intervention	COPD (hospitalised). On palliative care	Not related
ET2709	Intervention	Fall: broke left femur	Not related
ET2710	Intervention	Sepsis, pneumonia and pleurisy	Not related
ET2710	Intervention	Deceased: ischaemic heart disease, COPD and diabetes	Not related
ET2716	Control	Gall stone in bowel	Not related
ET2733	Control	Fall (hospitalised): pacemaker fitted	Not related
ET2733	Intervention	Hematoma (hospitalised)	Not related

COPD, chronic obstructive pulmonary disease; TIA, transient ischaemic attack.

TABLE 25 Primary and secondary outcomes at 6 months

Outcome	n (control, intervention)	Study arm		Estimated mean difference (95% CI)	p-value <sup>b</sup>
		Control <sup>a</sup>	Intervention <sup>a</sup>		
<b>Primary outcome</b>					
SPPB (total score)	305, 354	8.07 (1.91)	8.74 (1.87)	0.68 (0.39 to 0.96)	< 0.001
<b>Secondary outcomes</b>					
<i>Accelerometry</i>					
MVPA (minutes per day)					
Time spent at > 100 milligravitational units in at least 10-minute bouts	270, 314	5.68 (7.36)	6.56 (7.09)	0.86 (-0.37 to 2.12)	0.165
All time spent at > 100 milligravitational units	270, 314	53.74 (18.17)	56.59 (22.42)	2.85 (-0.67 to 6.37)	0.11
Sedentary time, excluding sleep (minutes per day)	261, 304	810 (78.84)	809 (95.02)	-0.91 (-15.52 to 13.71)	0.902
Breaks in sedentary time (n per day)	269, 312	41.48 (9.51)	42.91 (9.71)	1.42 (-0.17 to 3.02)	0.080
Subjective physical activity score (PASE)	296, 346	115.49 (58.21)	131.82 (61.76)	16.33 (6.78 to 25.89)	0.001
Muscle-strengthening exercise score (MSEQ)	277, 320	3.22 (1.90)	3.93 (1.90)	0.70 (0.39 to 1.01)	< 0.001
Hand-grip strength (kg)	302, 350	24.78 (3.93)	25.34 (4.09)	0.55 (-0.08 to 1.18)	0.085
Sleep Condition Indicator score	255, 288	24.34 (5.74)	24.23 (5.45)	-0.16 (-1.05 to 0.81)	0.806
Loneliness	301, 345	111 (36.8%)	115 (33.3%)	-0.009 (-0.070 to 0.052)	0.778
SF-36 score					
Physical component	293, 342	30.64 (8.68)	32.75 (8.42)	2.11 (0.73 to 3.48)	0.003
Mental component	293, 342	54.19 (7.46)	54.41 (6.88)	0.23 (-0.93 to 1.38)	0.699
EQ-5D	299, 346	0.70 (0.15)	0.68 (0.15)	0.026 (0.003 to 0.05)	0.028
MAT-SF	297, 345	50.00 (7.05)	51.40 (6.53)	1.40 (0.32 to 2.48)	0.012
<i>UK Biobank Healthy Minds Questionnaire score</i>					
Simple processing speed	287, 333	811.79 (243.82)	821.25 (259.59)	9.45 (-32.48 to 51.39)	0.655
Fluid intelligence	286, 327	3.87 (1.42)	3.89 (1.32)	0.02 (-0.21 to 0.25)	0.854
Executive function	240, 270	62,291.45 (39,262.35)	60,343.81 (39,565.45)	-1947.64 (-8835.47 to 4940.18)	0.576
Working memory 1	286, 331	4.34 (1.21)	4.42 (1.13)	0.08 (-0.12 to 0.29)	0.423
Working memory 2	286, 330	13.59 (5.53)	13.83 (5.77)	0.24 (-0.70 to 1.19)	0.609
Episodic memory	287, 331	5.97 (4.98)	5.96 (4.59)	-0.00 (-0.80 to 0.79)	0.994
<i>Falls Inventory</i>					
No. of falls in last 6 months	295, 335	0.61 (0.93)	0.58 (1.02)	-0.03 (-0.18 to 0.12)	0.712
Fall-related injury in last 6 months, n (%)	291, 328	40 (13.7)	43 (13.1)	0.64 (-4.96 to 6.35) <sup>c</sup>	0.728
a Mean (SD) unless otherwise stated.					
b Adjusted for site, exercise group (within the intervention arm), age group, sex and baseline SPPB.					
c Adjusted estimate and 95% CI for the between-group percentage difference.					

TABLE 26 Primary and secondary outcomes at 12 months

Outcome	n (control, intervention)	Control <sup>a</sup>	Intervention <sup>a</sup>	Estimated mean difference (95% CI)	p-value <sup>b</sup>
<b>Primary outcome</b>					
SPPB (total score)	303, 346	7.85 (2.05)	8.62 (2.58)	0.77 (0.40 to 1.14)	< 0.001
<b>Secondary outcomes</b>					
<b>Accelerometry</b>					
MVPA (minutes per day)					
Time spent at > 100 milligravitational units in at least 10-minute bouts	277, 299	4.98 (5.51)	6.22 (6.20)	1.24 (0.22 to 2.26)	0.018
All time spent at > 100 milligravitational units	277, 299	52.04 (15.92)	55.15 (19.67)	3.11 (-0.00 to 6.23)	0.05
Sedentary time, excluding sleep (minutes per day)	274, 300	790 (67.54)	789 (66.51)	-0.92 (-12.11 to 10.27)	0.871
Breaks in sedentary time (number per day)	275, 301	42.92 (9.62)	43.49 (10.93)	0.57 (-1.15 to 2.28)	0.512
Subjective physical activity (PASE)	296, 306	120.22 (49.34)	131.05 (46.23)	10.84 (3.18 to 18.50)	0.006
Muscle-strengthening exercise (MSEQ)	266, 315	3.32 (1.84)	3.98 (2.22)	0.67 (0.33 to 1.00)	< 0.001
Hand-grip strength (kg)	297, 340	24.42 (3.64)	25.23 (4.05)	0.81 (0.20 to 1.43)	0.010
Ageing Well Profile Social Wellbeing subscale	284, 322	24.34 (5.74)	24.23 (5.44)	-0.12 (-1.05 to 0.81)	0.805
Sleep Condition Indicator	266, 303	22.44 (5.28)	23.30 (4.89)	0.86 (-0.05 to 1.77)	0.064
Loneliness	296, 340	104 (35.1%)	111 (32.6)	-0.003 (-0.069 to 0.064)	0.935
<b>SF-36</b>					
Physical component	293, 334	29.66 (8.53)	32.25 (8.23)	2.59 (1.22 to 3.95)	< 0.001
Mental component	293, 334	54.52 (7.52)	53.95 (7.81)	-0.57 (-1.82 to 0.67)	0.362
EQ-5D	293, 337	0.70 (0.14)	0.69 (0.14)	0.018 (0.004 to 0.041)	0.112
MAT-SF	292, 328	49.25 (7.51)	51.47 (7.79)	2.21 (1.00 to 3.43)	0.001
<b>UK Biobank Healthy Minds Questionnaire</b>					
Simple processing speed	257, 297	829.66 (241.06)	842.97 (253.15)	13.31 (-30.01 to 56.63)	0.543
Fluid intelligence	254, 295	4.18 (1.47)	4.22 (1.54)	0.04 (-0.23 to 0.31)	0.768
Executive function	204, 244	66,552.53 (44,421.06)	62,642.94 (44,187.88)	-3909.59 (-12,564.60 to 4745.42)	0.371
Working memory 1	257, 294	4.48 (1.35)	4.36 (1.52)	-0.13 (-0.38 to 0.13)	0.322
Working memory 2	259, 296	14.22 (5.68)	13.74 (5.83)	-0.49 (-1.50 to 0.53)	0.343
Episodic memory	258, 297	5.91 (4.87)	5.86 (4.45)	-0.05 (-0.88 to 0.77)	0.896
<b>Falls inventory</b>					
Number of falls in last 6 months	300, 330	0.73 (1.05)	0.70 (1.05)	-0.02 (-0.19 to 0.14)	0.772
Fall-related injury in last 6 months, n (%)	297, 326	51 (17.2)	57 (17.5)	0.3 (-5.92 to 6.46) <sup>c</sup>	0.809

a Mean (SD) unless otherwise stated.

b Adjusted for site, exercise group (within the intervention arm), age group, sex and baseline SPPB.

c Adjusted estimate and 95% CI for the between-group percentage difference.



TABLE 27 Primary and secondary outcomes at 24 months

Outcome	n (control, intervention)	Control arm <sup>a</sup>	Intervention arm <sup>a</sup>	Estimated mean difference (95% CI)	p-value <sup>b</sup>
<b>Primary outcome</b>					
SPPB total score	294, 334	7.59 (2.61)	8.08 (2.87)	0.49 (0.06 to 0.92)	0.014
<b>Secondary outcomes</b>					
MVPA (minutes per day)					
Time spent at > 100 milligravitational units in at least 10-minute bouts	250, 290	4.50 (6.61)	5.15 (5.99)	0.65 (-0.48 to 1.78)	0.255
All time spent at > 100 milligravitational units	250, 290	48.76 (19.48)	51.22 (17.20)	2.46 (-0.52 to 5.44)	0.105
Sedentary time, excluding sleep (minutes per day)	249, 287	798 (65.80)	804 (64.04)	6.43 (-4.81 to 17.67)	0.259
Breaks in sedentary time (n/day)	248, 287	42.33 (13.54)	40.76 (13.21)	-1.57 (-3.89 to 0.75)	0.184
Subjective physical activity (PASE)	301, 328	113.17 (52.10)	123.90 (49.79)	10.73 (2.62 to 18.84)	0.010
Muscle-strengthening exercise (MSEQ)	276, 307	3.18 (1.88)	3.86 (2.30)	0.68 (0.33 to 1.02)	< 0.001
Hand-grip strength (kg)	291, 328	23.43 (4.08)	23.74 (3.86)	0.31 (-0.33 to 0.94)	0.343
Ageing Well Profile Social Wellbeing subscale	295, 306	24.68 (5.85)	24.88 (7.07)	0.20 (-0.84 to 1.24)	0.700
Sleep Condition Indicator	285, 311	21.97 (6.10)	22.50 (6.65)	0.53 (-0.49 to 1.54)	0.306
Pain (WOMAC)	290, 324	10.20 (3.28)	9.63 (3.95)	-0.57 (-1.15 to 0.00)	0.052
Loneliness, n (%)	300, 330	107 (35.7)	110 (33.3)	0.037 (-0.064 to 0.074) <sup>c</sup>	0.914
SF-36					
Physical component	295, 306	29.38 (9.39)	30.84 (10.04)	1.46 (-0.09 to 3.01)	0.065
Mental component	295, 306	54.73 (7.64)	54.33 (9.18)	-0.40 (-1.78 to 0.98)	0.563
EQ-5D	302, 330	0.67 (0.16)	0.69 (0.16)	0.02 (-0.01 to 0.04)	0.220
MAT-SF	289, 319	47.96 (8.13)	49.99 (8.96)	2.03 (0.66 to 3.40)	0.004
<b>UK Biobank Healthy Minds Questionnaire</b>					
Simple processing speed	264, 286	811.28 (240.15)	801.67 (246.72)	-9.61 (-52.47 to 33.24)	0.657
Fluid intelligence	262, 282	4.03 (1.41)	4.19 (1.61)	0.16 (-0.11 to 0.43)	0.234
Executive function	210, 236	64,770.62 (38,677.48)	58,515.77 (35,648.79)	-6254.85 (-13,498.22 to 988.52)	0.090
Working memory 1	263, 282	4.59 (1.29)	4.46 (1.22)	-0.13 (-0.35 to 0.06)	0.260
Working memory 2	264, 285	14.27 (5.24)	14.62 (5.15)	0.36 (-0.56 to 1.28)	0.439
Episodic memory	263, 286	5.84 (4.19)	5.36 (6.85)	-0.48 (-1.49 to 0.53)	0.347
<b>Falls inventory</b>					
Number of falls in last 6 months	300, 330	0.73 (1.05)	0.70 (1.05)	-0.02 (-0.19 to 0.14)	0.772
Fall-related injury in last 6 months, n (%)	297, 326	51 (17.2)	57 (17.5)	0.3 (-5.92 to 6.46) <sup>c</sup>	0.809
<p>a Mean (SD) unless otherwise stated.</p> <p>b Adjusted for site, exercise group (within the intervention arm), age group, sex and baseline SPPB.</p> <p>c Adjusted estimate and 95% CI for the between-group percentage difference.</p>					

TABLE 28 Primary outcome sensitivity analyses using MI without clustering by exercise group and using three population parameters to explore the impact of adherence

Sensitivity analysis	n; mean SPPB total score (SD)		Estimated mean difference (95% CI)	p-value <sup>a</sup>
	Control	Intervention		
MI (n = 777)	367; 7.50 (2.44)	410; 7.87 (2.93)	0.38 (0.02 to 0.73)	0.040
With no clustering by exercise group <sup>b</sup> (n = 628)	294; 7.59 (2.61)	334; 8.7 (2.21)	0.49 (0.14 to 0.84)	0.006
Population 0 (whole sample) (n = 628)	294; 7.59 (2.61)	334; 8.08 (2.87)	0.49 (0.06 to 0.92)	0.014
Population 1 ( $\geq$ 50% adherence) (n = 546)	294; 7.59 (2.20)	252; 8.23 (2.49)	0.64 (0.23 to 1.05)	0.002
Population 2 ( $\geq$ 75% adherence) (n = 471)	294; 7.64 (2.12)	177; 8.45 (2.28)	0.81 (0.38 to 1.23)	< 0.001

a Adjusted for site, exercise group (within the intervention arm), age group, sex and baseline SPPB.

b The intraclass correlation coefficient for SPPB scores relating to clustering by exercise group within the intervention arm was 0.02 (95% CI 0.0085 to 0.129).

TABLE 29 Analysis of prespecified subgroups: subgroup variables added as covariates and interaction terms with study arm in main model

Factor/covariate	Coefficient (SE); p	Coefficient for interaction with study arm (SE); p <sup>a</sup>	Control arm, mean (SE)	Intervention arm, mean (SE)
Baseline SPPB				
4–7	0		6.20 (0.20)	6.71 (0.22)
8–9	2.35 (0.19); < 0.001	0.094 (0.361); 0.794	8.59 (0.17)	9.03 (0.18)
Education level				
1	0	0	7.41 (0.44)	7.79 (0.45)
2	0.28 (0.35); 0.423	0.229 (0.690); 0.740	7.57 (0.20)	8.18 (0.23)
3	0.39 (0.36); 0.278	-0.317 (0.718); 0.659	8.00 (0.27)	8.06 (0.24)
4	0.15 (0.37); 0.696	0.303 (0.738); 0.681	7.39 (0.29)	8.08 (0.27)
5	-0.06 (0.46); 0.901	0.760 (0.926); 0.412	6.89 (0.54)	8.03 (0.42)
Deprivation level				
1	0	0	7.14 (0.36)	7.68 (0.42)
2	0.03 (0.34); 0.935	0.344 (0.670); 0.607	6.98 (0.30)	7.86 (0.29)
3	0.28 (0.34); 0.410	-0.343 (0.670); 0.609	7.60 (0.29)	7.80 (0.27)
4	0.38 (0.35); 0.280	0.040 (0.668); 0.952	7.48 (0.29)	8.06 (0.27)
5	1.04 (0.34); 0.002	-0.211 (0.645); 0.743	8.28 (0.24)	8.61 (0.25)
Age (years)				
65–74	0		8.35 (0.21)	8.78 (0.21)
$\geq$ 75	-1.20 (0.19); < 0.001	-0.016 (0.368); 0.965	7.16 (0.17)	7.57 (0.18)
Comorbidity				
No	0		7.86 (0.17)	8.17 (0.18)
Yes	-0.41 (0.18); 0.022	-0.283 (0.476); 0.553	7.19 (0.20)	7.97 (0.20)

continued

TABLE 29 Analysis of prespecified subgroups: subgroup variables added as covariates and interaction terms with study arm in main model (continued)

Factor/covariate	Coefficient (SE); <i>p</i>	Coefficient for interaction with study arm (SE); <i>p</i> <sup>a</sup>	Control arm, mean (SE)	Intervention arm, mean (SE)
Sex				
Male	0		7.54 (0.23)	7.85 (0.22)
Female	0.22 (0.19); 0.238	0.273 (0.375); 0.466	7.62 (0.16)	8.20 (0.17)
Site				
Bath/Bristol	0		7.18 (0.19)	7.95 (0.22)
Birmingham	0.70 (0.25); 0.007	-0.193 (0.497); 0.698	7.96 (0.27)	8.54 (0.29)
Devon	0.39 (0.23); 0.090	-0.711 (0.574); 0.116	7.88 (0.23)	7.94 (0.25)
Fallers				
Yes	0		7.48 (0.22)	7.76 (0.21)
No	0.39 (0.18); 0.033	0.331 (0.365); 0.366	7.69 (0.16)	8.30 (0.17)
Housing type				
Home owner	0		7.73 (0.14)	8.19 (0.14)
Other	-0.935 (0.26); 0.007	0.012 (0.509); 0.981	6.79 (0.32)	7.27 (0.36)
Other interventions: 24 months <sup>b</sup>				
No	0		7.73 (0.25)	8.43 (0.31)
Yes	0.08 (0.30); 0.777	-0.291 (0.634); 0.513	7.82 (0.16)	8.22 (0.15)

a Unstandardised coefficient for each factor in the regression model.

b Whether or not additional (non-REACT) physical activity interventions were undertaken at any point in the 24-month study period.

TABLE 30 Repeated-measures mixed-model analysis of the primary outcome (SPPB total score) measured at baseline and 6, 12 and 24 months post randomisation (adjusted for site, exercise group, age, sex and baseline SPPB score)

Time point	Control arm, mean (SE)	Intervention arm, mean (SE)	Estimated mean difference <sup>a</sup> (95% CI)	<i>p</i> -value
Baseline	7.37 (0.12)	7.45 (0.28)		-
6 months	8.02 (0.13)	8.75 (0.28)	0.65 (0.25 to 1.03)	0.001
12 months	7.83 (0.13)	8.65 (0.28)	0.74 (0.34 to 1.13)	< 0.001
24 months	7.54 (0.13)	8.11 (0.28)	0.48 (0.08 to 0.88)	0.018

a Estimated group × time interaction effect on SPPB score.

## Appendix 2 Process evaluation

Parts of this appendix have been reproduced from Cross *et al.*<sup>71</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

### Intervention fidelity

TABLE 31 Behavioural change techniques and delivery processes included in the REACT intervention fidelity analysis

Checklist item	Intended delivery strategies
1. Person-centred delivery style communication should be participant focused, maximising participant autonomy (delivery style)	<p>Use of open-ended questions</p> <p>Affirmations for positive behaviours, recognising efforts to change, as well as participants' autonomy</p> <p>Reflective listening (actively engage with participant, empathise, reflect emotional state)</p> <p>Summaries of the discussion, drawing together the ideas discussed (e.g. around motivation, overcoming barriers)</p> <p>Using the ask-tell-discuss technique to exchange/deliver key information</p>
2. Facilitating enjoyment (delivery style)	Using the techniques associated with person-centred delivery (as above), session leaders should encourage and reinforce enjoyment of social interactions within the group by making the social interactions positive, supportive and enjoyable
3. Monitoring progress (acknowledging and reviewing) (self-regulatory technique)	Using the techniques associated with person-centred delivery (as above), session leaders should regularly acknowledge and review the progress of group members in terms of their physical activity levels
4. Monitoring progress (eliciting and reinforcing the benefits of physical activity) (self-regulatory technique)	Using the techniques associated with person-centred delivery (as above), session leaders should encourage discussion on the emotional, social and physical benefits of physical activity
5. Self-monitoring (self-regulatory technique)	Using techniques associated with person-centred delivery (as above), session leaders should encourage participant self-monitoring or acknowledge participant attempts to self-monitor
6. Managing setbacks and problem-solving (self-regulatory technique)	Using techniques associated with person-centred delivery (as above), session leaders should encourage discussion on setbacks participants have experienced and encourage problem-solving. This should include reframing and normalising setbacks. Problems should be broken down, and the sustainability of coping plans and the support others can provide should also be considered
7. Action-planning and goal-setting (self-regulatory technique)	Using techniques associated with person-centred delivery (as above), session leaders should work with the participants to agree on action plans, including goal-setting and identifying any barriers that may arise
8. Modelling (SCT)	Using techniques associated with person-centred delivery (as above), session leaders should give participants the opportunity to observe others engaging appropriately with the programme
9. Promoting autonomy (SDT)	Using techniques associated with person-centred delivery (as above), session leaders should encourage proactive involvement in the classes and discussion and create opportunities for participant input, while acknowledging participant perspectives, encouraging participants to be the driver of change and develop a sense of control

continued

**TABLE 31** Behavioural change techniques and delivery processes included in the REACT intervention fidelity analysis (*continued*)

Checklist item	Intended delivery strategies
10. Supporting competence and self-efficacy (SDT and SCT)	Using techniques associated with person-centred delivery (as above), session leaders should encourage participants to identify and break down barriers to change, set achievable goals/encourage gradual progress and encourage problem-solving
11. Supporting relatedness (SDT)	Using techniques associated with person-centred delivery (as above), session leaders should fulfil participants' needs for relatedness (social engagement/acceptance, approval of one's behaviour and giving support to others). This can be promoted by encouraging engagement in physical activity, where there are opportunities for positive social interactions, as well as highlighting physical activity as a social opportunity

**TABLE 32** The Dreyfus scoring system for assessing competence in the delivery of clinical consultations and adaptations for scoring REACT delivery fidelity (see also *Report Supplementary Material 1*)

Competence level <sup>72</sup>	Scoring	Examples	REACT delivery fidelity categories
Absence	0	Absence of feature and/or highly inappropriate performance	Low fidelity
Novice	1	Minimal use of feature and/or inappropriate performance	Low fidelity
Advanced beginner	2	Evidence of competence, but numerous problems or inconsistencies	Scope for improvement
Competent	3	Competent, good features but some minor problems or inconsistencies	Competent
Proficient	4	Very good features, but minimal problems or inconsistencies	Proficient
Expert	5	Excellent features, no problems or inconsistencies	Expert

**TABLE 33** Criteria for summarising scores for each fidelity checklist item

Checklist item	Criterion for summarising scores across multiple sessions
Person-centred delivery style	Mean
Facilitating enjoyment	Mean
Monitoring progress	Mean
Self-monitoring	Maximum
Managing setbacks and problem-solving	Mean
Action-planning and goal-setting	Mean
Modelling	Maximum
Promoting autonomy	Mean
Supporting competence and self-efficacy	Mean
Supporting relatedness	Mean

TABLE 34 Examples of REACT delivery practice

Checklist item	Examples of good practice	Examples of practice requiring improvement
Person-centred delivery	Group 5, week 12: the facilitator starts a discussion with open-ended questions, reflects on the responses of participants, responding where appropriate. The facilitator gives options throughout the discussion, but lets the group direct the conversation. The facilitator highlights the group social aspect. The conversation is natural, and the facilitator confirms and summarises details at the end before praising them	Group 4, week 9: the facilitator's communication is not participant focused. For example, the facilitator asks a question but does not reflect on answers, quickly moving on to ask another question. The facilitator frequently talks over participants and directs conversation back to themselves
Facilitating enjoyment	Group 3, week 9: the facilitator introduces a name game as a means of getting to know each other. The whole group is involved. The facilitator supports the idea by going first, dispelling awkwardness and encourages 'banter' as she goes along. Throughout the game, the facilitator praises the group and reinforces positive comments that are made	Group 3, week 12: the session leader is slow to react when groups are not interacting or the mood of the group has dropped
Monitoring progress: acknowledge and review	Group 6, week 16: the facilitator asks how the group got on with their 'activity snacks'. Group discussion ensues with a bit of joking, a positive light-hearted environment in which to discuss progress. One participant reflects on a failure to progress, but the facilitator reframes it positively as an improvement (since starting), asking how the lady felt about it. The facilitator and participant move on to discuss a specific goal and action plan for keeping it up	Group 4, week 12: the facilitator jokes that a participant is finding it too easy and needs more weight but fails to use it as an opportunity to formally praise them and highlight the progress made
Monitoring progress: eliciting benefits of physical activity	Group 6, week 16: the facilitator praises the group for the progress made in increasing activity levels and opens a discussion about the benefits. The facilitator asks participants about ways they feel they have benefited. The session leaders' approach to the discussion is participant focused	Group 6, week 16: the facilitator affirms that the group is more active and learning different movements, but fails to use the opportunity to explore the benefits of this with the group
Self-monitoring	Group 3, week 9: the facilitator describes how to use the pedometers to monitor steps, constantly checking for understanding and clarifying where needed. The facilitator praises those who ask questions about the device and highlights barriers they may face and how they might go about overcoming them. The facilitator summarises and encourages participants to do their best to improve their steps and not to compare their efforts with those of others	Group 4, week 28: a participant reminds the facilitator to discuss and record their pedometer steps. Session leader records them and asks for them to reset their pedometers to zero for the following week. The facilitator offers no praise or feedback on participants' efforts to self-monitor steps
Managing setbacks and problem-solving	Group 3, week 13: the facilitator asks an open question about a participant and the participant reports she has not done well. The facilitator is positive and reframes the comment, highlighting the fact that the participant has shown up to REACT and goes on to offer her praise	Group 6, week 16: the group brings up the fact that they are struggling to exercise in the hot weather; the facilitator agrees but does not encourage them or offer any ways to overcome this as a barrier  Group 6, week 16: a participant starts a discussion about managing their knee injury and the effect that this has had on their

continued

TABLE 34 Examples of REACT delivery practice (continued)

Checklist item	Examples of good practice	Examples of practice requiring improvement
	Group 3, week 48: the facilitator asks for progress on a participant's leg injury. Despite the participant's negative outlook on it, the facilitator is positive, reassuring her that her plan of rehabilitation is a good one. The facilitator then focuses on the fact that organising a physiotherapy appointment is a positive step	activity; however, the facilitator misses the opportunity to offer any reassurance, support or problem-solving strategies
Action-planning and goal-setting	Group 6, week 16: the facilitator discusses an action plan for keeping up recent progress and encourages participant input, making positive affirmations about the action plan that participants have devised	Group 5, week 48: the facilitator asks about participants' plans for exercise post REACT. The facilitator does not take the opportunity to discuss the specificity of goals or plans, how participants will execute them or discuss barriers or problems that may arise
Modelling	Group 1, week 13: the facilitator asks a participant to share their experiences of starting a new activity with the rest of the group. The facilitator highlights their success and praises the participant for attending	Group 4, week 28: the facilitator highlights a participant's correct execution of a particular exercise but does not explain why it is correct or congratulate the participant for successful engagement
Promoting autonomy	Group 5, week 12: the facilitator supports participant choice during a discussion about new activities that they are thinking of joining. The facilitator offers options for a planned group activity but insists it is 'their choice' and asks open questions about how they feel about this. The facilitator lets them discuss and offers help in summarising the discussion, asking again how they feel about the particular plan that they have formulated	Group 5, week 13: the facilitator consistently points out that participants are able to choose the way in which they complete exercises, but does not facilitate a discussion about this or encourage participant input
Supporting competence and self-efficacy	Group 6, week 16: after discussing an action plan with a participant, the facilitator asks the participant if she feels that she could manage it. The facilitator goes on to affirm that she believes that the participant can	Group 5, week 13: the facilitator works well at identifying barriers to change but then does not continue the discussion adequately to include breaking down the barriers, and seeking the input of participants was limited in this instance
	Group 3, week 28: while action-planning, the facilitator asks participants for barriers that they face to joining a new class. The facilitator breaks each down, offering up solutions that the group goes on to discuss. The interaction is very participant focused	
Supporting relatedness	Group 5, week 12: the facilitator supports discussion about organising a group walk, encouraging participants to choose their own pace but support each other through it. The facilitator continues to encourage a social aspect by organising coffee afterwards. The facilitator supports a participant-focused discussion, during which most of the interaction occurs between the participants	Group 6 week 16: a participant discusses problems that they experience with their knees and that they do not want to drop out of REACT. Another participant offers to work with the first, as both have limited mobility, but the facilitator fails to encourage this pairing and does not use it as an opportunity to discuss the social benefits of physical activity or how participants may support each other in the future



## Qualitative evaluation of intervention processes

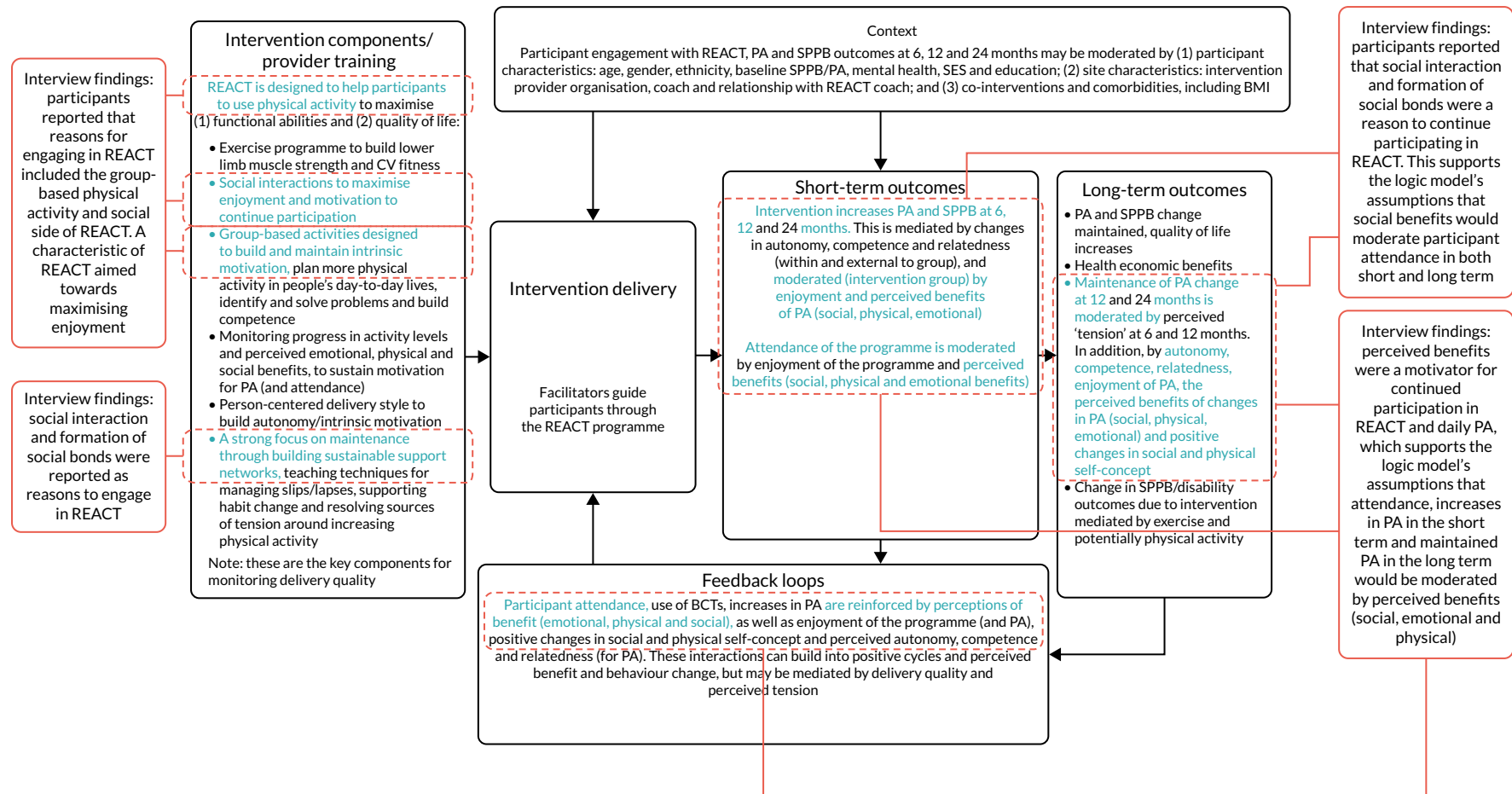


FIGURE 28 The REACT logic model in the context of findings: reasons to engage in REACT. CV, cardiovascular; PA, physical activity.



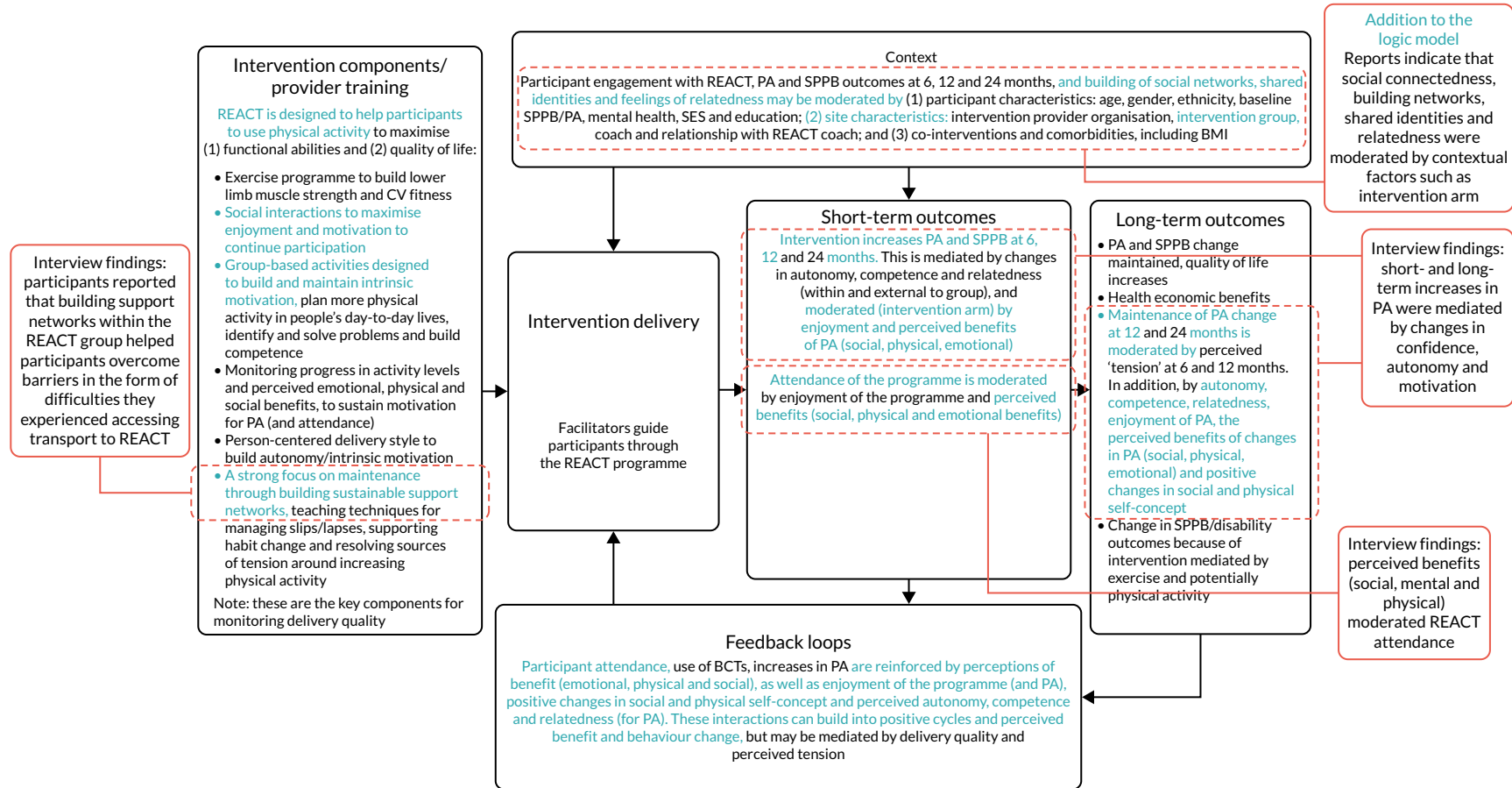


FIGURE 29 The REACT logic model in the context of findings: benefits of participation in REACT. CV, cardiovascular; PA, physical activity.

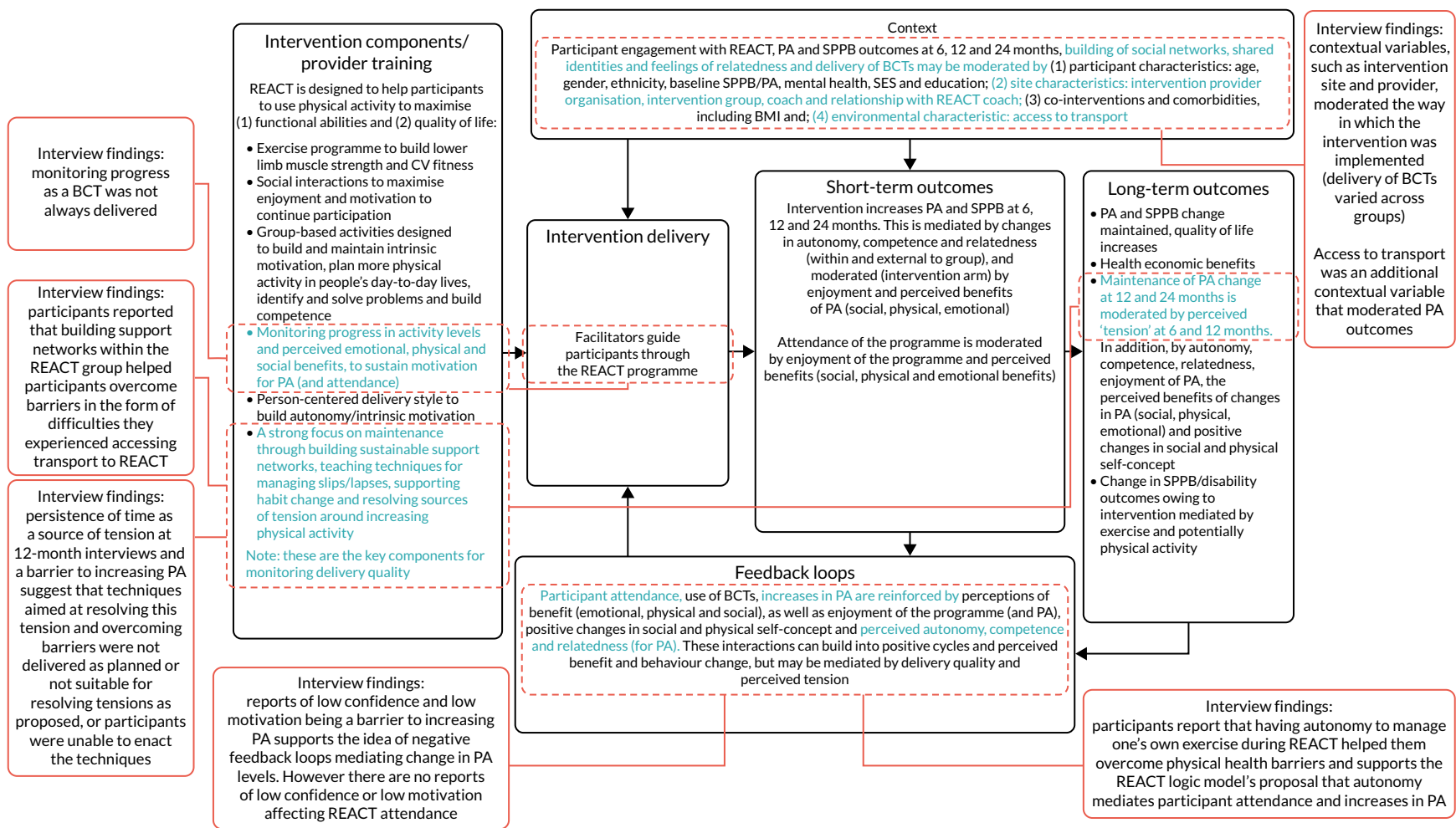


FIGURE 30 The REACT logic model in the context of findings: barriers to participation in REACT and daily physical activity. CV, cardiovascular; PA, physical activity.

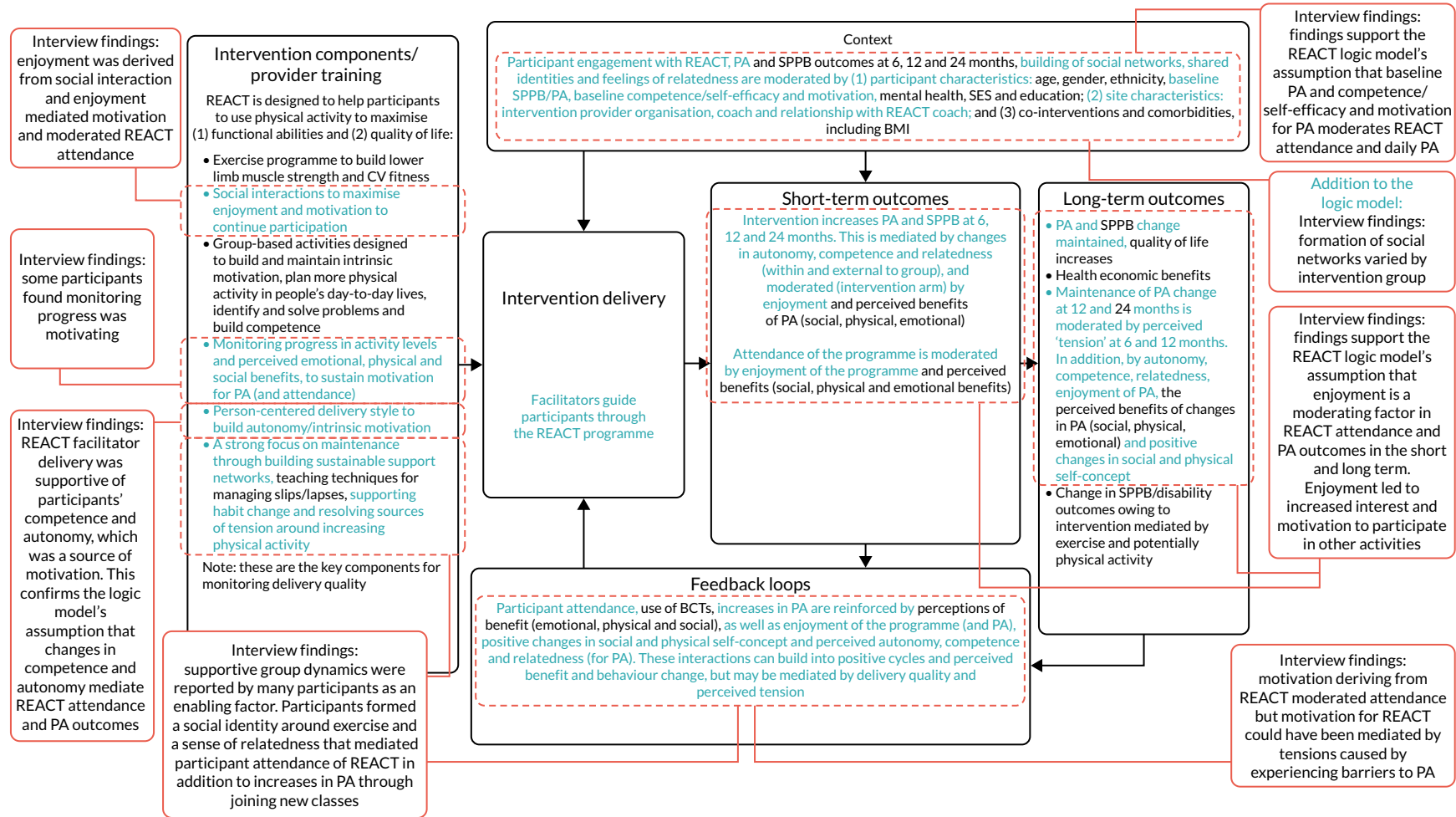


FIGURE 31 The REACT logic model in the context of findings: enablers of participation in REACT and daily physical activity. CV, cardiovascular; PA, physical activity.

## Quantitative process evaluation

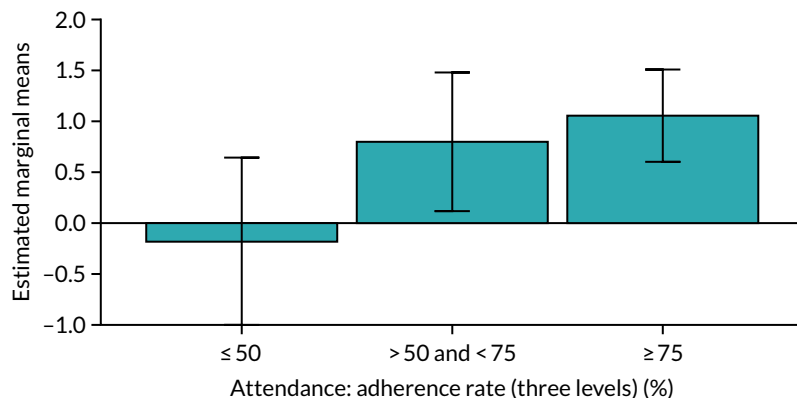


FIGURE 32 Changes in hours spent walking outside from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI).

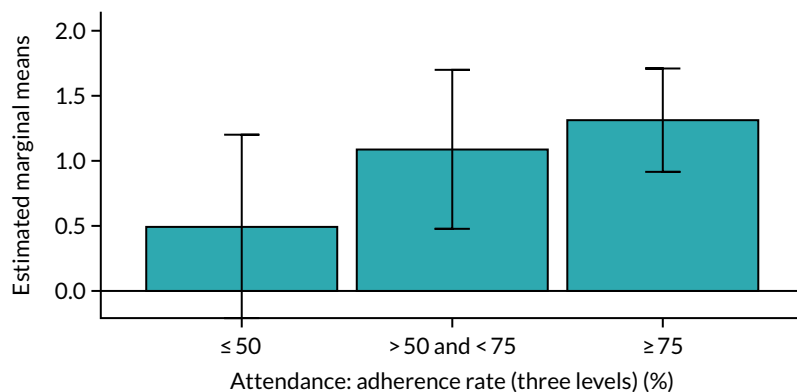


FIGURE 33 Changes in adherence to muscle-strengthening exercise (MSEQ adherence score) from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI).

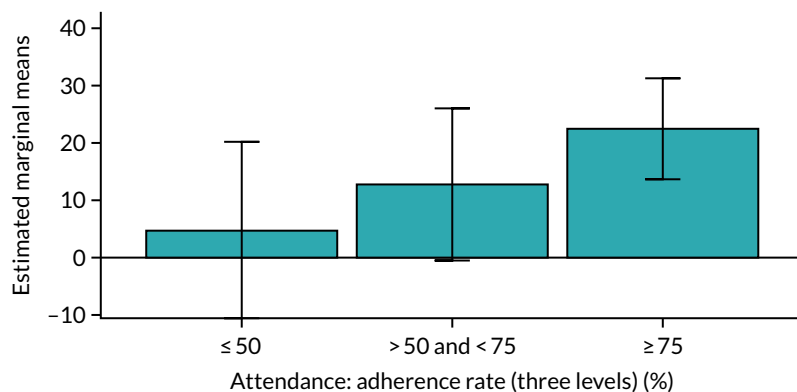


FIGURE 34 Changes in PASE total score from 0 to 6 months for different levels of attendance within the intervention arm (error bars: 95% CI).

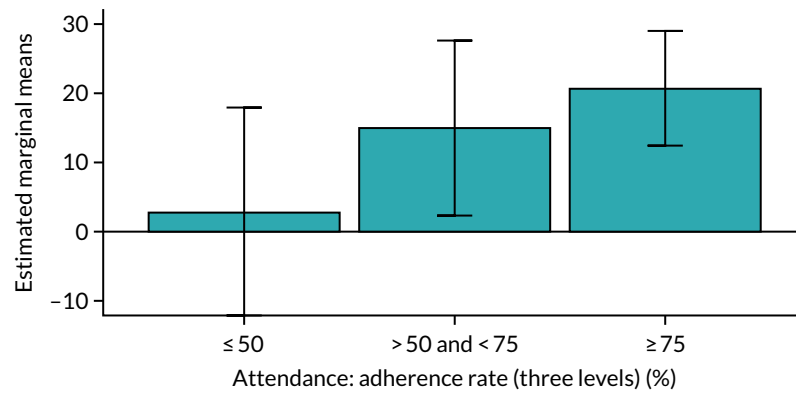


FIGURE 35 Changes in PASE total score from 0 to 12 months for different levels of attendance within the intervention arm (error bars: 95% CI).

## Appendix 3 Health economic evaluation

Parts of this appendix have been adapted from Snowsill *et al.*<sup>103</sup> This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by/4.0/>. The text includes minor additions and formatting changes to the original text.

### Valuation of resources

The costs of hospital overnight stays were estimated from the 2017/18 NHS reference costs<sup>113</sup> and the 2018/19 NCC for the NHS schedule of NHS costs.<sup>111</sup> Unlike previous editions, the current edition does not present costs for excess bed-days (or the number of bed-days contributing to estimated costs), so it is challenging to estimate costs per bed-day.

We examined the number of admissions and number of bed-days blinded to treatment allocation. A total of 186 general medical ward admissions were reported, accumulating 810 bed-days (mean 4.4 days per admission). In the case of long-stay wards, there were eight admissions totalling 101 bed-days (mean 12.6 days per admission), and there were 20 admissions to intensive care units, totalling 108 bed-days (mean 5.4 days per admission).

The cost for general medical ward and long-stay ward admissions were estimated using the 2017/18 NHS reference costs<sup>113</sup> as follows:

- All data for elective, non-elective long stay and non-elective short stay finished consultant episodes were collated.
- Any Healthcare Resource Groups (HRGs) clearly not applicable to the population (e.g. paediatric or relating to maternity, or specifically giving an age range for children and young people) were filtered out.
- The total number of bed-days for each HRG was calculated from the 'inlier' and excess bed-days.
- A linear regression (weighted for activity) was used to relate the mean cost for a HRG to the mean bed-days for that HRG.

As a result, it was assumed that the cost for an admission would be £376.47 plus £492.58 per bed-day, and these figures were then inflated by 2.31% to 2018/19 prices.<sup>110</sup> Estimating a cost for spending time in an intensive care unit is challenging because critical care reference costs are 'unbundled', meaning that they refer only to the costs of critical care specifically and not to the costs of treating the underlying disease that caused admission, which are reported and reimbursed separately. In addition, current reference costs refer to a period spent in critical care rather than bed-days. The most recent reference costs using bed-days as activity are the 2013/14 reference costs,<sup>140</sup> which suggest an average cost of £1190 per bed-day after excluding paediatric and neonatal critical care (£1274 in 2018/19 prices). For comparison, the most recent NCC estimates are an average cost of £1428 per critical care period (again, after excluding paediatric and neonatal critical care).

We assumed that a period of critical care would cost £1428.09 (specifically for critical care) plus £2104.00 (the average cost for admissions after exclusions as above), for a total cost of £3532.09.

The societal cost of individuals paying for a chiropractor or osteopath was estimated from market prices. For a chiropractor, we used estimates from a study by Newell *et al.*,<sup>117</sup> who estimated, based on data from 33 clinics relating to 1895 patients, a mean of 5.4 chiropractic consultations per patient over a 90-day period, with associated mean per-patient costs of £202.15. The resulting mean cost per

appointment (£37.44) was then inflated to £40.06 in 2018/19 prices. The cost of an appointment with an osteopath was estimated from the public reporting of a private survey by the General Osteopathic Council that the average follow-up session cost £42,<sup>118</sup> which was inflated to £44.95 in 2018/19 prices. The costs of care at home were estimated from the PSSRU 2019 costs for a home care worker (£28 per hour on weekdays, £29 per hour at the weekend).

The societal value of an hour of help around the home from friends or relatives was estimated based on a compensating variation study among informal caregivers in the Netherlands.<sup>114</sup> This provided an estimated value of €9.65 per hour, which was converted to £7.39 in 2002 prices using purchasing power parity rates,<sup>115</sup> and then inflated to £11.06 in 2019 prices according to nominal weekly earnings growth.<sup>116</sup> Sensitivity analyses for plausible extreme values for an hour of help around the home came from two separate British studies.<sup>141,142</sup> First, a discrete choice experiment with 270 respondents from Scotland was analysed with a latent class model,<sup>141</sup> which suggested three classes with average class probabilities of 0.528, 0.192 and 0.204. The value at which respondents were willing to accept an hour of household tasks in each class was £0.25, £0.58 and £0.53, respectively, giving a weighted average willingness-to-accept value of £0.35 per hour of help with household tasks. Second, a compensating income variation analysis of the British Household Panel Survey<sup>142</sup> suggested that a compensation of £5020 per week would be required to offset the loss to well-being for an individual providing 10–19 hours per week of informal care, which is approximately £335 per hour (assuming 15 hours per week). The societal value of a day taken off work by a friend or relative was estimated from median daily earnings (£99.53).<sup>116</sup>

## Missing data

Multiple imputation was used to account for missing data on the assumption that cost and QALY data were missing at random (i.e. their missingness can be predicted from observed data). Fifty imputation sets were generated using MI chained equations with predictive mean matching (five nearest neighbours). The number of imputation sets was chosen based on the rule of thumb that the number of imputations should be similar to (or exceed) the percentage of cases that are incomplete.<sup>143</sup>

The variables used in the predictive mean matching were:

- random allocation variable (intervention vs. control)
- adjustment variables (sex, age, baseline SPPB score and site)
- SPPB scores at each follow-up time point
- EQ-5D and SF-6D utilities at each follow-up time point
- disaggregated costs at each follow-up time point.

Of the 777 trial participants, 451 (58%) provided complete data for cost-effectiveness analysis (i.e. resource use and EQ-5D-5L at all time points, including baseline). A further 25 participants provided complete data for an unadjusted cost-effectiveness analysis, that is they provided all data specified above apart from baseline resource use.

*Figure 36* shows that, although most participants provided baseline data, a fair proportion of the participants failed to provide one or more measurement of EQ-5D-5L or NHS/PSS resource use, meaning that QALYs and/or total costs could not be estimated.

Patterns of missing data for EQ-5D and SF-6D were examined. At each time point, EQ-5D was more likely to be present than SF-6D, but the missingness of EQ-5D and SF-6D was highly correlated. In total, 359 (46%) participants had at least one utility value missing for at least one time point: 230 (30%) were missing EQ-5D for at least one time point and 346 (46%) were missing SF-6D for at least one time point (*Figure 37*). Participants who had an instrument missing for one of the follow-up time

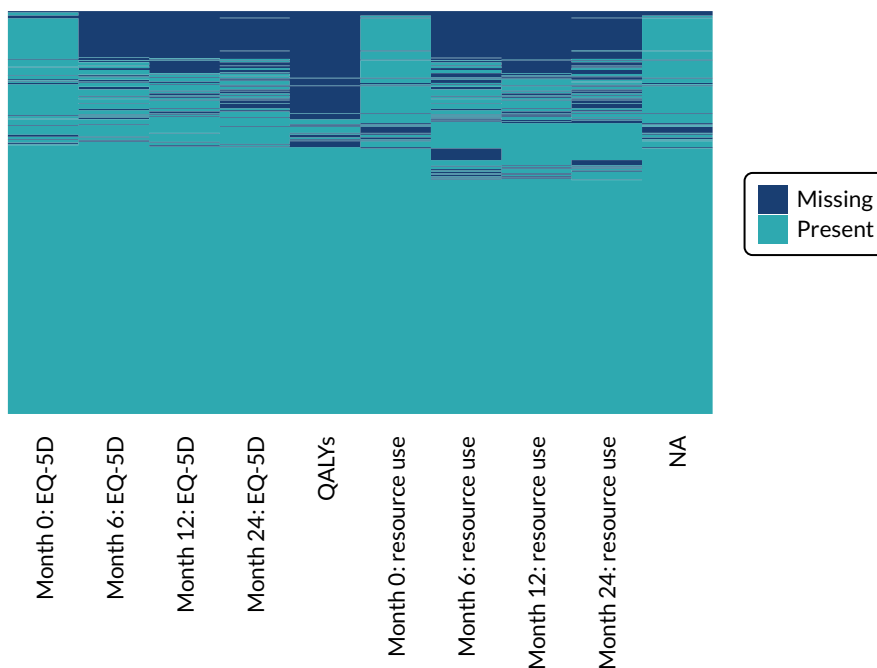


FIGURE 36 Missing data for EQ-5D and NHS/PSS costs.

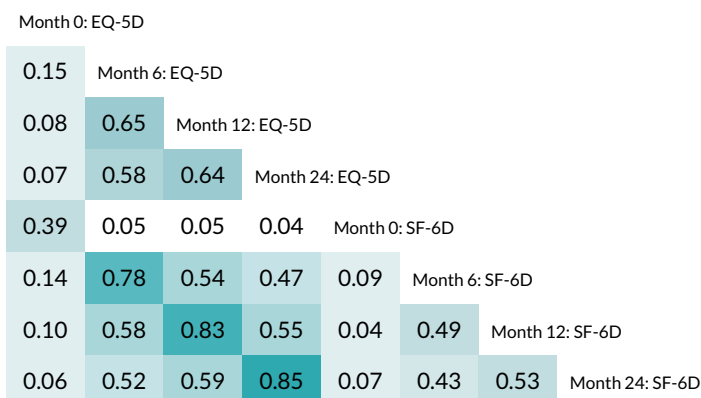


FIGURE 37 Correlogram of missingness of health-related quality-of-life measures at different time points.

points were more likely to have that instrument missing for the other follow-up time points. Logistic regression on an indicator variable for missingness revealed that treatment allocation did not appear to influence missingness (EQ-5D,  $p = 0.95$ ; SF-6D,  $p = 0.61$ ). Age at randomisation appeared to have an effect on missingness, with a fairly consistent pattern that older participants were more likely than younger participants to have missing EQ-5D and SF-6D (EQ-5D,  $p < 0.001$ ; SF-6D,  $p < 0.001$ ) data. Study site appeared to influence missingness; however, when age was also included in the logistic model, study site was no longer statistically significant (EQ-5D,  $p = 0.12$ ; SF-6D,  $p = 0.063$ ). Sex did not appear to influence missingness (EQ-5D,  $p = 0.49$ ; SF-6D,  $p = 0.22$ ).

For resource use, there was naturally a relationship between the missingness of societal costs and the missingness of NHS/PSS costs, given that all resources included in societal costs are also included in NHS/PSS costs (although sometimes with different values). Having missing resource use at one follow-up time point increased the probability of having missing resource use at subsequent time points (Figure 38). Treatment allocation and sex did not appear to influence missingness (treatment allocation,  $p = 0.2$ ; sex,  $p = 0.74$ ). Increased age was positively associated with the likelihood of costs



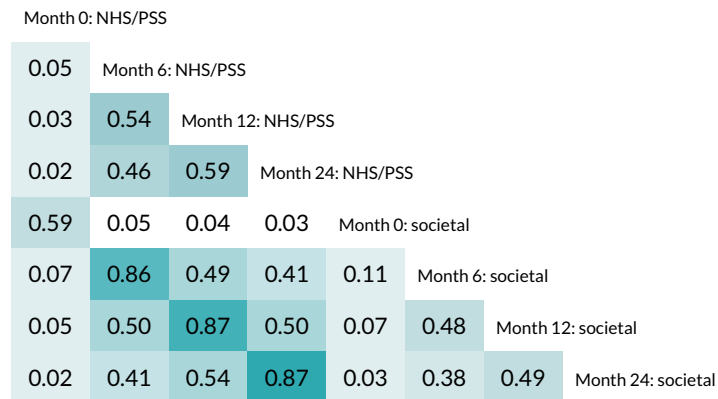


FIGURE 38 Correlogram of missingness of costs at different timepoints from NHS/PSS and societal perspectives.

being missing ( $p < 0.001$ ). Site was also associated with missingness of costs, with participants in Devon most likely and participants in Birmingham least likely to be missing costs ( $p < 0.001$ ). This relationship was preserved when age was also included in the logistic model.

## Intervention costs

### Referral and co-ordination

Our base-case assumption is that there would be no incremental GP time cost associated with referral, as the time taken to make a referral is not expected to be significant or notably different from inputs in the usual care arm, such as GP advice to increase physical activity, for which no costs have been included in the analysis. We included 30 minutes of time for a social prescribing link worker per patient, whose job is to ensure that the patient is appropriate for referral, briefly describe the REACT intervention and refer the patient to a REACT co-ordinator (see below). REACT study investigators estimated that it took 30 minutes to screen an individual for the REACT study.<sup>29</sup> As a sensitivity analysis, we replace these costs with the estimated recruitment costs for the REACT study, which included screening GP databases to identify suitable individuals, mailing invitations, a PR campaign and other recruitment methods, at an estimated total cost of £81,597 to recruit 777 participants (£105.02 per participant recruited).<sup>29</sup>

The REACT co-ordinator would perform activities such as:

- identifying appropriate venues and making bookings
- recruiting session leaders
- ensuring that session leaders are trained appropriately to deliver REACT and are intermittently supervised for intervention fidelity
- maintaining a waiting list
- administering related payments (to venues, session leaders and other expenses)
- sending schedules and reminders to participants.

We assume that these activities will require 36 hours of co-ordinator time (approximately 1 working week) per REACT group.

### REACT session leader

In general, the NHS does not employ sport and fitness instructors. The mean hourly salary for an employed fitness instructor was estimated to be £10.63.<sup>134</sup> However, over half of fitness instructors are self-employed<sup>144</sup> and the hourly charge for these is not well estimated by national statistics. In the REACT study, all session leaders were qualified to at least level 3 (whereas many fitness instructors may

be qualified only to level 2). One REACT study provider reported a base salary cost of £15.50 per hour and one freelance trainer charged £26 per hour. Using the Agenda for Change band 4 cost information from the PSSRU, a recent economic evaluation of individual health trainers<sup>145</sup> estimated the cost of a health trainer to be £28 per hour (revised to £29.43 in the most recent version of PSSRU unit costs).<sup>110</sup> The base annual salary for band 4 community-based scientific and professional staff is £22,256, which, spread over 1618 working hours per year (as assumed for band 4 workers by PSSRU), corresponds to an hourly gross rate of £13.76. We assume that the cost of REACT session leaders will be £29.43 per hour, and this should not be highly sensitive to whether REACT session leaders are employed or freelance. We, therefore, assume that this includes the cost of any supervision required, a suitable location to meet REACT participants and perform any administrative duties, and any travel expenses. In sensitivity analyses we vary the cost by  $\pm 20\%$ .

### Venue hire

The NHS does not (in general) have facilities to accommodate an intervention such as REACT, which requires a hall or large room suitable for group exercise with up to 15 participants. If the REACT intervention were to be rolled out, it is likely that a venue would need to be hired from a private sector or charity/non-profit sector organisation or from a public body (e.g. a local authority community centre or school). The costs of hiring such a venue will probably vary significantly depending on geographical location, type of space and facilities. The REACT intervention would operate during off-peak hours for health centres. The cost of venue hire was reported by the investigators as £15–25 per hour for some sites and £32.50 per session for a multisite provider. In the case of some sites, an overall cost was reported, including costs for staff and refreshments, so no specific venue cost could be identified.

To attempt to identify a representative sample of venue hire costs from outside the REACT trial, we randomly sampled 10 lower-layer super output areas in England, one from each decile of the IMD.<sup>146</sup> The retrieved prices ranged from £12 to £25 per hour, with a mean of £17.32 per hour, which is used as the base-case cost per hour of venue hire. In sensitivity analyses, we vary the cost from £15 per hour to £25 per hour.

### One-to-one introductory sessions

We assume that the 45-minute introductory sessions are attended by all individuals allocated to the intervention. As this may be an overestimate (some may not attend the introductory session and it is unlikely that the opportunity cost of a non-attendance is equal to the opportunity cost of a session), we do not include other preparation or travel costs for these sessions. In a sensitivity analysis, we include an additional 30 minutes of REACT trainer time for travel and/or preparation.

### Cost per session

The cost per session of REACT is estimated from the REACT session leader time (preparation, travel and delivery), REACT session leader travel expenses, venue hire and consumables. The nominal session lengths are 80 minutes (41 sessions) and 105 minutes (23 sessions), giving a mean nominal session length of 89 minutes. Session leaders were requested to note the start and end times of each session, which provides empirical data on the length of sessions (719 sessions provided these data out of a maximum of 1728). The data suggest that a significant proportion of 80-minute sessions were shortened to 75 minutes or extended to 90 minutes. The mean session length was 86.3 minutes. We estimate the cost per session by assuming that each session requires 90 minutes of the session leader time and venue hire. In a sensitivity analysis, we vary this from 80 minutes to 105 minutes.

Session leaders were heterogeneous in their recording of preparation time. We assumed 90 minutes per session and 45 minutes for 15 participants of one-to-one introductory sessions: this leaves  $\approx 51$  hours ( $\approx 48$  minutes per session) for preparation. In the base case, we assume 30 minutes of preparation per session. In sensitivity analyses, we vary this from 10 minutes to 90 minutes.

Session leader travel time was recorded in only 139 out of 737 (18.9%) session records, with a mean travel time of 22.2 minutes in those sessions. In a sensitivity analysis, we remove travel time to represent the cost if staff are already on site.

Consumables per session were assumed to be limited to refreshments and to cost £1 per attendee per session. We assumed that refreshments would be ordered for the full class for the first session (mean 14.63) and thereafter would be ordered for the number attending the first session (mean 9.63).

### Printed materials

We assume a cost per participant of £2, suggested by the investigators.

### Group size

In total, there were 410 individuals randomised, who were allocated to 27 groups; therefore, the average group size was 15.2 individuals. In sensitivity analyses we vary the group size from 12 to 17 individuals.

### Overall costs

The breakdown of the costs is presented in *Table 35*.

The cost was somewhat sensitive to underlying assumptions, as shown in *Table 36*. The cost per participant ranged from £554.54 (-11%) to £775.13 (25%) across the sensitivity analyses.

TABLE 35 Cost of the REACT intervention

Item	Resource use per group	Unit cost (£)	Cost per group (£)	Cost per participant (£)
Link worker	30 minutes per participant	33.83	257.48	16.92
Co-ordinator	36 hours per group	33.83	1217.88	80.01
Introductory sessions	45 minutes per participant	29.43	335.99	22.07
Equipment				
Pedometers	One per participant	9.95	151.46	9.95
Other	Ankle weights and Therabands, assumes reuse and sharing between groups	29.16	29.16	1.92
REACT session leader				
Preparation	30 minutes per session	29.43	941.76	61.87
Travel time	30 minutes per session	29.43	941.76	61.87
Delivery	90 minutes per session	29.43	2825.28	185.60
Consumables				
Refreshments	Approximately one per participant per session	1.00	621.32	40.82
Printed materials	One set per participant	2.00	30.44	2.00
Venue hire	90 minutes per session	17.32	1662.72	109.23
Training session leaders				
Training leaders	One leader to four trainees	33.83	22.20	1.46
REACT session leaders time	1.5 days for every four programmes delivered	29.43	77.25	5.08
Training venue	As above	29.43	77.25	5.07
Training manual	One per REACT trainer	12.00	3.00	0.20
Total			9465.69	621.83

TABLE 36 Sensitivity analyses for intervention cost

Sensitivity analysis	Cost per group (£)	Cost per participant (£)
REACT study recruitment costs (£105.02 per recruited participant)	10,806.84	709.94
REACT session leader		
20% less per hour	8441.28	554.54
20% more per hour	10,490.10	689.13
Additional 30 minutes for introductory one-to-one sessions	9689.68	636.55
Venue cost		
£15 per hour	9242.97	607.20
£25 per hour	10,202.97	670.27
Session duration		
80 minutes	8967.02	589.07
105 minutes	10,213.69	670.97
REACT session leader		
10 minutes preparation per session	8837.85	580.59
90 minutes preparation per session	11,349.21	745.57
Zero travel time	8523.93	559.97
Group size		
n = 12	9301.56	775.13
n = 17	9556.24	562.13

## Baseline resource use and utility values and within-trial costs analysis

TABLE 37 Baseline resource use and utility values

Characteristic	Unit	Control arm (N = 367)	Intervention arm (N = 410)
<i>Primary care resource use</i>			
Primary care usage reported, n (%)	Participants	359 (98)	404 (99)
GP, mean (SD)	Visits	2.91 (4.03)	2.75 (2.99)
Practice nurse, mean (SD)	Visits	1.40 (2.41)	1.54 (2.68)
Physiotherapist, mean (SD)	Visits	0.28 (1.05)	0.67 (2.54)
Occupational therapist, mean (SD)	Visits	0.03 (0.26)	0.04 (0.42)
Nutritionist, mean (SD)	Visits	0.04 (0.56)	0.04 (0.33)
Chiropodist, mean (SD)	Visits	0.84 (1.63)	0.87 (1.66)
Counsellor, mean (SD)	Visits	0.09 (1.29)	0.03 (0.37)
Walk-in centre, mean (SD)	Visits	0.03 (0.22)	0.03 (0.21)

continued

TABLE 37 Baseline resource use and utility values (continued)

Characteristic	Unit	Control arm (N = 367)	Intervention arm (N = 410)
<b>Hospital overnight stay resource use</b>			
Overnight stay usage reported, n (%)	Participants	350 (95)	392 (96)
General medical ward, mean (SD)	Visits	0.04 (0.22)	0.06 (0.25)
Long-stay ward, mean (SD)	Visits	0.00 (0.00)	0.00 (0.05)
Intensive care unit, mean (SD)	Visits	0.00 (0.00)	0.01 (0.07)
<b>Other hospital resource use</b>			
Other hospital usage reported, n (%)	Participants	362 (99)	406 (99)
Outpatient appointment, mean (SD)	Visits	1.19 (2.44)	1.11 (1.86)
Day case, mean (SD)	Visits	0.30 (0.92)	0.31 (0.88)
A&E attendance, mean (SD)	Visits	0.09 (0.34)	0.12 (0.37)
<b>Social care resource use, mean (SD)</b>			
Day care	Visits	0.01 (0.16)	0.07 (1.26)
Convalescent home	Visits	0.00 (0.00)	0.03 (0.70)
Relatives/friends help	Hours	1.60 (6.80)	2.20 (7.90)
Relatives/friends off work	Days	0.07 (0.83)	0.10 (1.34)
<b>Outside NHS/PSS perspective resource use, mean (SD)</b>			
Chiropractor	Visits	0.03 (0.34)	0.02 (0.30)
Osteopath	Visits	0.06 (0.67)	0.02 (0.23)
Care at home	Hours	0.00 (0.00)	0.00 (0.00)
<b>Summary costs (NHS/PSS perspective)</b>			
Complete cases, n (%)	Participants	347 (95)	387 (94)
Primary care, mean (SD)	£	185 (218)	212 (247)
Overnight stay, mean (SD)	£	68 (491)	188.16 (1461.92)
Other hospital, mean (SD)	£	394 (825)	390 (762)
Other care (excluding informal), mean (SD)	£	0 (6)	6 (75)
Total costs from NHS/PSS perspective, mean (SD)	£	653 (1006)	797 (1765)
<b>Summary costs (societal perspective)</b>			
Complete cases, n (%)	Participants	306 (83)	358 (87)
Other care (excluding informal), mean (SD)	£	4 (35)	10 (114)
Informal care, mean (SD)	£	25 (116)	36 (204)
Total societal costs, mean (SD)	£	659 (1025)	841 (1890)
<b>Baseline health-related quality of life</b>			
Complete cases, n (%)	Participants	320 (87)	364 (89)
EQ-5D, mean (SD)	Utility	0.68 (0.17)	0.69 (0.16)
SF-6D, mean (SD)	Utility	0.62 (0.09)	0.62 (0.10)

TABLE 38 Within-trial cost of health and social care resource usage (excluding intervention costs)

Analysis	n	Intervention cost (£), mean (SE)		Control cost (£), mean (SE)		Difference in cost (£), mean (95% CI)
<b>MI, NHS/PSS perspective</b>						
0–6 months	777	634.56 (51.35)		877.47 (104.39)		<b>-242.91 (-464.94 to -20.89)</b>
6–12 months	777	751.24 (68.15)		908.26 (97.83)		-157.01 (-366.61 to 52.59)
18–24 months	777	1107.04 (106.83)		1224.89 (131.04)		-117.84 (-426.60 to 190.91)
Total within-trial costs	777	3321.39 (203.36)		4046.13 (310.05)		<b>-724.74 (-1423.69 to -25.79)</b>
<b>MI, societal perspective</b>						
0–6 months	777	681.59 (53.43)		915.78 (106.40)		<b>-234.19 (-459.73 to 8.64)</b>
6–12 months	777	804.36 (73.77)		964.86 (102.52)		-160.50 (-378.93 to 57.93)
18–24 months	777	1201.00 (113.23)		1280.22 (134.01)		-79.23 (-398.96 to 240.51)
Total within-trial costs	777	3594.56 (212.52)		4240.21 (317.77)		-645.65 (-1355.02 to 63.73)
<b>Complete case analysis, NHS/PSS perspective</b>						
0–6 months	578	623.34 (53.96)		876.30 (118.93)		-252.96 (-508.74 to 2.82)
6–12 months	584	724.98 (73.49)		793.90 (83.25)		-68.91 (-266.53 to 128.71)
18–24 months	579	1129.23 (123.52)		1217.96 (147.71)		-88.72 (-450.07 to 272.63)
Total within-trial costs	477	3248.86 (257.30)		3573.37 (306.69)		-324.51 (-1086.45 to 437.43)
<b>Complete case analysis, societal perspective</b>						
0–6 months	496	653.83	64.22	782.55	85.71	-128.73 (-352.09 to 94.64)
6–12 months	501	669.51	67.25	766.94	81.81	-97.43 (-295.12 to 100.26)
18–24 months	492	1164.12	148.57	1297.73	173.41	-133.61 (-561.72 to 294.50)
Total within-trial costs	377	3114.50	241.16	3572.72	330.51	-458.23 (-1254.94 to 338.49)
Total within-trial costs include linear interpolation to estimate costs between 12 and 18 months and discounting of costs from 12 to 24 months at 3.5%. Bold indicates that the difference is statistically significant ( $p < 0.05$ ).						

## Subgroup analyses

TABLE 39 Subgroup analyses (within-trial economic evaluation)

Subgroup	Costs (£)		QALYs		ICER (£)
	Intervention arm	Control arm	Intervention arm	Control arm	
Baseline SPPB score					
4–7	4842.42	4156.20	1.315	1.288	25,008
8–9	3291.05	3771.44	1.394	1.343	Dominant
Age (years)					
65–74	3394.97	3589.01	1.390	1.333	Dominant
≥ 75	4230.99	4257.29	1.333	1.305	Dominant
Sex					
Female	3822.75	3796.69	1.347	1.313	784
Male	4154.65	4540.45	1.373	1.321	Dominant
continued					

TABLE 39 Subgroup analyses (within-trial economic evaluation) (continued)

Subgroup	Costs (£)		QALYs		ICER (£)
	Intervention arm	Control arm	Intervention arm	Control arm	
Known medical conditions					
0–1	3791.61	3367.29	1.376	1.340	11,881
≥ 2	4028.48	4867.41	1.339	1.286	Dominant
SES (by education)					
Not beyond secondary school	4361.20	4239.89	1.350	1.324	4725
Beyond secondary school	3614.64	3799.94	1.359	1.307	Dominant
SES (by IMD decile)					
1–6	3871.42	3989.77	1.343	1.292	Dominant
7–10	4004.43	4089.17	1.371	1.341	Dominant
Site					
Bath/Bristol	3833.29	3603.76	1.364	1.308	4075
Birmingham	4301.44	4235.21	1.360	1.306	1235
Devon	3841.29	4509.19	1.342	1.333	Dominant
Falls in previous 6 months					
None	4016.40	3815.74	1.377	1.330	4266
One or more	3857.19	4273.30	1.326	1.293	Dominant

Costs include the intervention cost of £621.83, which does not vary between subgroups.

## Review of economic evaluations of interventions to prevent or reduce mobility-related disability

We have conducted a pragmatic review to identify existing economic evaluations of interventions similar to REACT, performed a within-trial economic evaluation and developed a new decision analytic model to extrapolate any long-term costs and benefits that may arise from the REACT intervention.

### Introduction

The aim of this review was to identify any economic evaluations of interventions with similar aims to REACT, that is group-based physical activity interventions for older adults with the intention to maintain mobility or delay frailty. The purpose of identifying any such evaluations was to explore modelling approaches that may be appropriate and to be able to put the results of the economic evaluation of REACT in the wider evidence base for any existing competing interventions in the same patient population.

The broader set of economic evaluations of physical activity interventions (individual and group based) was not considered in this review because these studies often focus on the impact of physical activity on long-term health conditions, such as cardiovascular disease, which is not considered to be a primary aim of REACT.

### Methods

A pragmatic review of economic evaluations of interventions to prevent or reduce mobility-related disability in older people was undertaken. To be eligible for inclusion, a study needed to be an economic evaluation comparing at least two options, with at least one including a physical activity component.

Studies must incorporate costs and a health outcome measure. The study population must be older adults (mostly or all aged  $\geq 65$  years) and must not be selected according to the presence and/or history of specific health conditions. The primary purpose of the intervention must be to maintain mobility or delay the incidence of frailty, and its effect may be evaluated in terms of maintaining or improving general mobility or balance or reducing falls.

MEDLINE was searched via Ovid from inception to 25 March 2020 by using medical subject heading (MeSH) terms and the study design filter used by the NHS Economic Evaluation Database (NHS EED) to restrict records of full economic evaluations:

1. exp Frailty/or exp Aging/or exp Mobility Limitation/ [n = 247,743]
2. exp Exercise/or exp Exercise Therapy/ [n = 221,376]
3. 1 and 2 [n = 8623]
4. (NHS EED search filter begins)

...

29. (NHS EED search filter ends) [n = 827,917]
30. 3 and 29 [n = 296].

The MeSH terms were identified by the reviewer (TS) as relevant to the study question by considering the MeSH terms associated with REACT protocol<sup>1</sup> and the LIFE study,<sup>21</sup> and also searching for relevant MeSH terms. Free-text terms were not included because this was a pragmatic review. The NHS EED study design filter is one that has been used for many systematic reviews of economic evaluations.<sup>147</sup>

## Results

One reviewer (TS) screened the titles and abstracts from the 296 retrieved records and selected 21 to retrieve as full texts for assessment of eligibility. One citation could not be retrieved as full text, so 20 full texts were assessed for eligibility (see *Figure 39*). Forward citation chasing from protocol papers was also conducted using Google Scholar to find published results.

A significant proportion of the full texts was protocols for studies with no associated economic evaluation published, and, in addition to this, several studies were excluded because they did not include physical activity as a core intervention component (e.g. it was an optional component in a multicomponent intervention or the intervention was designed to facilitate physical activity) or because they did not complete a full economic evaluation. For a list of excluded studies at the full-text screening stage, see *Report Supplementary Material 3*.

Two studies<sup>124,125</sup> were ultimately judged to be eligible, and these are described briefly below.

### *Alhambra-Borrás et al. 2019*

Alhambra-Borrás *et al.*<sup>124</sup> investigated the effectiveness and cost-effectiveness of a group-based exercise programme for community-dwelling adults aged  $\geq 65$  years in Valencia, Spain. Their study was quasi-experimental, with participants from two different health centres being recruited concurrently and all participants from one centre being offered the intervention, and all participants from the other centre receiving usual care. However, an estimated 100 out of 258 participants in the 'intervention centre' did not agree to participate in the physical exercise programme and were, therefore, excluded. This introduces a substantial risk of bias into all of the study analyses in addition to the normal caveats associated with the lack of randomisation in allocating patients to treatments.

Health-care resource utilisation was estimated by asking participants to recall doctors' visits and hospitalisations for the previous 12 months (at baseline) or 9 months (at the 9-month follow-up). The SF-12 was used to assess health-related quality of life (at baseline and at the 9-month follow-up).



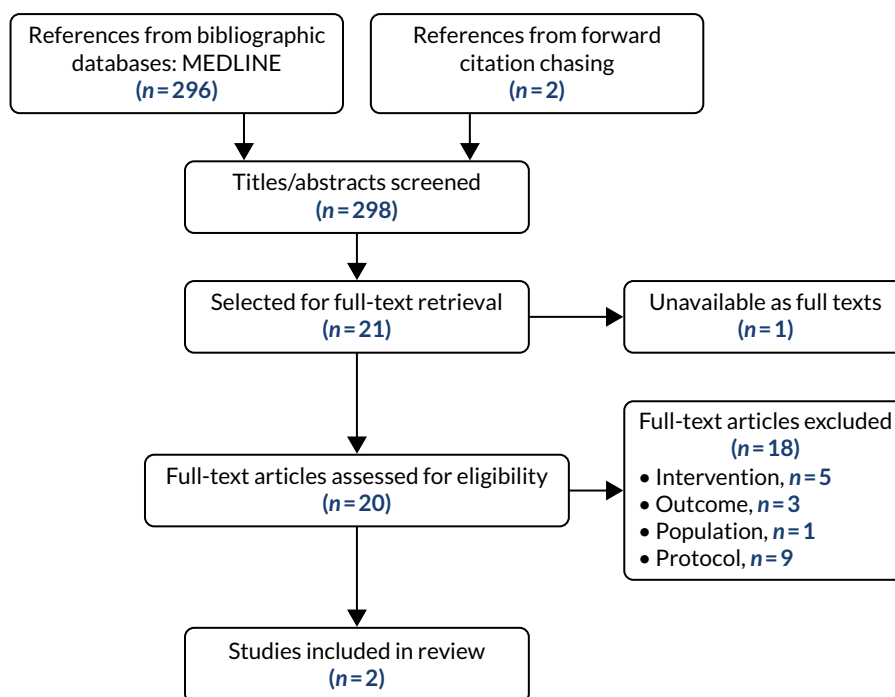


FIGURE 39 The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram for pragmatic review of economic evaluations.

The cost of the intervention was estimated solely from the time of the physiotherapist delivering the intervention, suggesting that no cost was used for the venue or any equipment. It is not clear whether the cost per hour for the physiotherapist is salary cost only or if it includes overhead costs.

A within-trial cost-effectiveness analysis was not conducted. Instead, trial data were used to estimate the associated health-care costs and quality of life (SF-12 mapped to EQ-5D-3L utility) of participants in the 'baseline state' (improvement in risk of falls/frailty after 9 months) and 'deteriorated state' (no improvement or worsening of risk of falls/frailty after 9 months). These costs and utility values were estimated separately for the two centres (intervention and control). This may be problematic because differences in costs and health-related quality of life could arise from pre-existing differences between the centres rather than owing to the effect of the intervention. The authors do not report what proportion of participants was expected to be in each health state or how future transitions between these health states were accounted for in the analysis.

The authors conclude that the intervention, if offered continuously, would lead to reductions in overall costs (from a health-care perspective) and increases in QALYs, meaning that the intervention would be dominant. For all age groups (up to 92 years), the intervention is predicted to be dominant, with younger participants experiencing greater QALY gains and realising greater cost savings. The analysis predicted average cost savings of €44,833 (approximately £48,500, converted using 2019 PPP<sup>122</sup>) and QALY gains of 0.513 (both discounted at 3.5% per year).

The authors did not report the uncertainty in incremental costs and QALYs, but they did conduct a sensitivity analysis, which suggested that participants who were older at the start of the intervention would achieve smaller cost savings and gain fewer QALYs.

### Groessl et al. 2016

Groessl *et al.*<sup>125</sup> investigated the cost-effectiveness of the LIFE physical activity programme. Adults aged 70–89 years at risk of developing major mobility disability were recruited and randomised to

either the LIFE physical activity programme (including group sessions and home-based activity plans) or a health education programme as a control. Every 6 months, participants reported their health-care resource utilisation using a published resource use questionnaire and reported their health-related quality of life using the QWB-SA tool.<sup>148</sup>

Intervention costs were estimated from personnel costs and adjusted for overheads, such as facilities. The longest follow-up time point was at 36 months after randomisation and mean follow-up was 2.6 years, which served as the time horizon for the analysis.

The physical activity intervention was estimated to cost \$2300 more than the health education programme per participant and to lead to 0.047 additional QALYs, for an ICER of \$49,000 per QALY in 2013 prices. Health-care resource utilisation costs were not included in the base-case analysis, with the justification that the difference did not reach statistical significance; however, when these were included in a sensitivity analysis, the ICER was \$83,000 per QALY. Limited sensitivity analyses were conducted, and showed that, if overheads, wages or fringe benefits were higher, the physical activity intervention would be less cost-effective.

No attempt was made to extrapolate with decision analytic modelling to estimate lifetime costs and QALYs. It is not clear whether or not extrapolation would have made a material difference to the conclusions given that differences between the arms in resource use and QWB-SA score were not statistically significant at 36 months, but there was some evidence that fewer participants in the physical activity arm had persistent mobility disability at 36 months.

## Natural history model

There are two natural history components in the model:

1. the evolution of mobility with age
2. the changing mortality rate with age.

### *The evolution of mobility with age*

It has been demonstrated in a number of population studies from countries other than the UK that SPPB score is negatively correlated with age, with the mean SPPB score typically decreasing by around 1 point per decade of life beyond 60 years of age, and that women typically have lower SPPB scores.<sup>149–151</sup> However, it should be noted that these are cross-sectional studies rather than longitudinal studies; therefore, these may not correspond to a typical pattern of SPPB scores over an individual's life course. For example, it may be that those with high SPPB scores are able to maintain their mobility, whereas those with some impairment are at high risk of becoming more impaired. It is also possible that surgical management (e.g. hip replacement) could result in an increase in SPPB score. Mortality may also be associated with mobility impairment (although not necessarily directly caused by poor mobility),<sup>129</sup> which could lead to a less pronounced decline of SPPB with increasing age when assessed in a cross-section of the population compared to if SPPB were measured over a typical life course. Population studies have generally shown that very few older people fall in the lowest SPPB category (0–3),<sup>149,150</sup> although it is possible that such individuals were selected out (or selected themselves out) of the studies by not completing the SPPB, as well as being less likely to survive into old age.

There is a trend to convergence in the predicted mean SPPB score observed over the course of the REACT study, as described in *Chapter 3*. In particular, there is an increase in the benefit of the intervention from 6 to 12 months, which is followed by a decline from the 12- to 24-month follow-up points, which amounts to yearly reduction of 36%.

The extrapolation of mobility from the end of the trial follow-up to lifetime was derived from an ordered logistic regression of SPPB score at 24 months as a function of SPPB score at 12 months post randomisation, age at randomisation, sex and allocation site (Bath/Bristol, Birmingham, Devon) in the base-case analysis (model 1) and a modified model that included only SPPB score at 12 months post randomisation (model 2) in sensitivity (scenario) analysis. The model may be described by the following equation:

$$\widehat{SPPB}_{24i} = \alpha SPPB_{12i} + \beta'X_i + \varepsilon_i, \quad (1)$$

where  $\widehat{SPPB}_{24i}$  is the value of a latent (unobserved) continuous mobility status variable for individual  $i$ , a trial participant allocated to the control arm at 24 months, which is determined by that person's own observed  $SPPB_{12i}$  value at 12 months with coefficient  $\alpha$  plus a linear combination of the product of variables and their coefficients,  $\beta'X_i$ , which in the case of model 1 is equal to:

$$\beta_1 \text{Age at randomisation}_i + \beta_2 \text{gender}_i + \beta_3 \text{DExeter}_i + \beta_4 \text{DBirmingham}_i, \quad (2)$$

where Devon takes the value of 1 if the individual is from the Devon trial site and is zero otherwise, while  $\text{DBirmingham}$  is similarly defined (with Bath and Bristol as the reference);  $\beta'X_i$  is omitted from model 2. The residual term,  $\widehat{SPPB}_{24i}$ , is a random unobserved variable following a logistic distribution. The continuous dependent variable, [inline], is not observed; instead the discrete variable,  $SPPB_{24i}$ , is observed, taking the value of 0, 1, 2, ..., 11, 12, with probabilities determined by the cumulative logistic distribution centred on the mean value of  $\alpha SPPB_{12i} + \beta'X_i$  relative to a set of fixed thresholds defining the boundaries of SPPB response categories.

To assess convergence, we used the appropriate 24-month SPPB distributions in the control and intervention arms, as described above, and then used the model to evolve the SPPB distribution over time in the absence of mortality. The mean SPPB at 24 months, 36 months (3 years) and 120 months (10 years) was calculated. Convergence is given by the reduction in the difference in mean SPPB score between intervention and control from 24 to 36 months.

The two models resulted in a convergence (i.e. a decline in the benefit of the intervention over usual care) rate for the first year of extrapolation ranging from 30% at age 65 years to 15% at age 95 years in women and from 31% to 15% in men in the base-case analysis (see *Table 40*); the corresponding predicted convergence rates in scenario analysis were 22% to 16% in women and 22% to 14% in men (see *Table 42*). Details of the models are presented below. We also consider three other models for the evolution of mobility. Model 3 is a static distribution, that is all individuals maintain their 24-month end-of-trial SPPB status until death, so there is no convergence of SPPB distributions between the intervention group and the control group. Model 4 uses the observed frequency of changes in SPPB between month 12 and month 24 in the REACT control group (annual convergence 5.2–7.3%, depending on age and sex). Model 5 assumes a normal distribution for changes in SPPB that does not achieve long-term decline of SPPB but is matched to the 36% annual convergence between the intervention group and the control group in the last year of the trial.

Predictions from model 1 are shown in *Table 40* (predictions have been made assuming that the site is Devon). Notable features are that the convergence is not as fast as observed between month 12 and month 24 in the REACT study (36%), that the mean SPPB is predicted to increase for younger individuals (e.g. women aged 65 years, men aged 75 years), and is predicted to become very low on average for older individuals (especially women).

We use model 1 in the base-case and model 2 (see *Table 41*) in a scenario analysis. We also consider three other models for the evolution of mobility. Model 3 is a static distribution, that is all individuals maintain their 24-month SPPB status until death. This implies no convergence of the SPPB distributions between the REACT group and the control group. Model 4 is based on the frequency of observed

TABLE 40 Predictions from model 1

Sex	Age (years)	Mean SPPB score after 2 years	Mean SPPB score after 10 years	Convergence
Female	65	8.91	9.26	30%
Female	75	7.81	7.41	23%
Female	85	6.60	4.77	19%
Female	95	5.34	1.81	15%
Male	65	8.79	9.80	31%
Male	75	7.67	8.17	24%
Male	85	6.45	5.78	19%
Male	95	5.19	2.74	15%

SPPB values are given for the control arm. Convergence is the annual percentage reduction in the difference between SPPB in the intervention arm and SPPB in the control arm for the first year of extrapolation.

TABLE 41 Predictions from model 2

Sex	Age (years)	Mean SPPB score after 2 years	Mean SPPB score after 10 years	Convergence
Female	65	8.91	7.06	22%
Female	75	7.81	6.81	19%
Female	85	6.60	6.49	16%
Female	95	5.34	6.09	14%
Male	65	8.79	7.04	22%
Male	75	7.67	6.77	19%
Male	85	6.45	6.44	16%
Male	95	5.19	6.22	13%

SPPB values are given for the control arm. Convergence is the annual percentage reduction in the difference between SPPB in the intervention arm and SPPB in the control arm for the first year of extrapolation.

changes in SPPB between month 12 and month 24 in the REACT control group (annual convergence 5.2–7.3% depending on age and sex). Model 5 is an artificial model (based on a normal distribution for changes in SPPB) that does not aim to achieve long-term decline of SPPB but does achieve 38% annual convergence between the REACT and the control groups.

### Details of mobility models

The evolution of mobility was initially estimated from the REACT study using ordered logistic regression (model 1) with mobility at 24 months post randomisation in the control arm as the dependent variable, with estimated coefficients for mobility at 12 months post randomisation, age at randomisation, sex and allocation site (Bath/Bristol vs. Birmingham vs. Devon) as independent variables (Table 42).

As shown in Table 43, the SPPB score at 12 months was positively associated with SPPB score at 24 months. Furthermore, age had a negative independent effect; for example, the predicted SPPB score distribution at 24 months for a 75-year-old with a SPPB score of 8 at 12 months would be worse than the predicted SPPB score distribution at 24 months for a 65-year-old also with a SPPB score of 8.

TABLE 42 Results of ordered logistic regression (model 1) to predict SPPB at 24 months

Term	Estimate	95% CI
SPPB score at 12 months	0.787	0.660 to 0.915
Age at randomisation	-0.052	-0.088 to -0.016
Sex: female	-0.208	-0.646 to 0.230
Site		
Birmingham	0.050	-0.542 to 0.642
Devon	0.572	0.107 to 1.037
Threshold SPPB score		
0/1	-3.988	-7.070 to -0.906
1/2	-3.291	-6.379 to -0.203
2/3	-2.476	-5.592 to 0.640
3/4	-1.830	-4.931 to 1.270
4/5	-0.617	-3.755 to 2.521
5/6	0.119	-3.028 to 3.265
6/7	1.060	-2.079 to 4.200
7/8	1.972	-1.177 to 5.120
8/9	3.115	-0.024 to 6.255
9/10	4.181	1.029 to 7.332
10/11	5.188	2.038 to 8.337
11/12	6.877	3.616 to 10.14

$n = 367$ . Model  $F$ -test for joint covariate significance:  $F_{5,5129} = 35.61$ ;  $p < 0.0001$ .

There was no evidence to suggest that the decline in SPPB score was faster for women (because the coefficient CI included the null). Participants in Devon had a lower predicted probability than those from other sites of finding themselves in a low-mobility health state at 24 months. The mean distance between thresholds was 0.988, meaning that the SPPB distribution predicted at any one time will be concentrated in three to four SPPB categories (i.e. unlikely to have large jumps in SPPB). The thresholds appear uncertain owing to wide CIs, but the thresholds are highly positively correlated, so the distance between them remains relatively stable.

A more parsimonious model was fitted, removing the terms for sex, site and age, leaving only SPPB score at 12 months to predict SPPB score at 24 months (model 2). *Table 43* shows the terms for this model.

The predicted transitions across states using this model will (in the absence of death) converge to a steady state distribution with mean SPPB score of 6.53 (*Figure 40*).

### The changing mortality rate with age

The mortality rate for individuals was estimated from UK life tables for men and women aged 65–100 years.<sup>152</sup> For men and women aged > 100 years, a Gompertz model was fitted to the data for men and women aged 65–100 years and was used to extrapolate to age 135 years. It is assumed that the population that would participate in a REACT programme is well approximated by the general population. There are two reasons why this may not be an ideal approximation, but these operate in opposite directions: (1) to participate in a REACT programme an individual will need to have a minimum level of mobility and be able to attend (which is likely to exclude those in residential or nursing homes or with severe health conditions) and (2) to participate in a REACT programme an individual must have some level of mobility impairment (which would exclude those in very good health).

TABLE 43 Results of ordered logistic regression (model 2) to predict SPPB at 24 months

Term	Estimate	95% CI
SPPB score at 12 months	0.799	0.676 to 0.922
Threshold SPPB score		
0/1	0.196	-0.888 to 1.280
1/2	0.868	-0.091 to 1.827
2/3	1.681	0.813 to 2.549
3/4	2.320	1.456 to 3.183
4/5	3.512	2.625 to 4.399
5/6	4.227	3.332 to 5.122
6/7	5.135	4.161 to 6.109
7/8	6.014	4.978 to 7.050
8/9	7.121	6.019 to 8.222
9/10	8.146	6.973 to 9.318
10/11	9.129	7.897 to 10.36
11/12	10.796	9.404 to 12.19

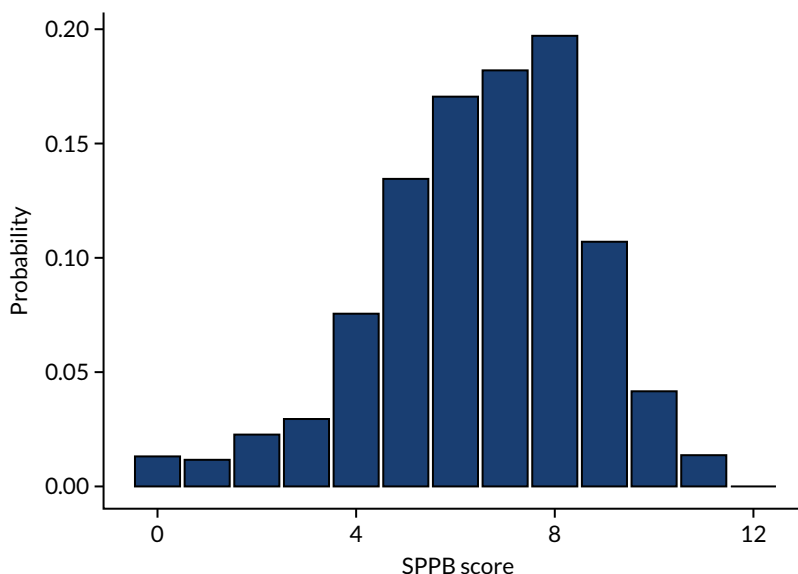


FIGURE 40 Steady state distribution for Model 2.

The scenario analyses considered four alternative models for the evolution of SPPB (Table 44):

- model 2 – ordered logistic model fitted to REACT study control arm with no terms for age, sex or site
- model 3 – no transitions between SPPB states (i.e. remain in the same SPPB state until death)
- model 4 – transitions based on distribution of changes in SPPB in the REACT study control arm between month 12 and month 24
- model 5 – artificial transition matrix designed to achieve fast convergence between REACT and control arms (difference in mean SPPB cut by 36% each year).

TABLE 44 Table of input parameters for model-based economic evaluation

Parameter name	Description	Base-case value	One-way sensitivity analysis		PSA		
			Lower	Upper	Distribution	Parameter 1	Parameter 2
Dr	Discount rate for costs and QALYs	3.5%					
age_start	Age at the start of the model (years)	75	65	95			
sex.female	Population is female?	1	0	1			
TrialEnd_SPPB.Control.0	Probability of having given SPPB (0 to 12) in control arm at 24 months	0.008			Multinomial	5	
TrialEnd_SPPB.Control.1		0.015				8.9	
TrialEnd_SPPB.Control.2		0.015				8.9	
TrialEnd_SPPB.Control.3		0.028				16.4	
TrialEnd_SPPB.Control.4		0.051				30.1	
TrialEnd_SPPB.Control.5		0.069				41	
TrialEnd_SPPB.Control.6		0.089				52.8	
TrialEnd_SPPB.Control.7		0.134				79.5	
TrialEnd_SPPB.Control.8		0.149				88.1	
TrialEnd_SPPB.Control.9		0.164				97.3	
TrialEnd_SPPB.Control.10		0.121				71.6	
TrialEnd_SPPB.Control.11		0.104				61.6	
TrialEnd_SPPB.Control.12		0.053				31.3	

Parameter name	Description	Base-case value	One-way sensitivity analysis		PSA		
			Lower	Upper	Distribution	Parameter 1	Parameter 2
TrialEnd_SPPB.Intervention.0	Probability of having given SPPB (0 to 12) in REACT arm at 24 months	0.006			Multinomial	3.1	
TrialEnd_SPPB.Intervention.1		0.010				5.7	
TrialEnd_SPPB.Intervention.2		0.011				5.9	
TrialEnd_SPPB.Intervention.3		0.020				11.2	
TrialEnd_SPPB.Intervention.4		0.039				21.4	
TrialEnd_SPPB.Intervention.5		0.056				30.6	
TrialEnd_SPPB.Intervention.6		0.075				41.3	
TrialEnd_SPPB.Intervention.7		0.119				65.5	
TrialEnd_SPPB.Intervention.8		0.142				77.8	
TrialEnd_SPPB.Intervention.9		0.171				94.2	
TrialEnd_SPPB.Intervention.10		0.141				77.3	
TrialEnd_SPPB.Intervention.11		0.135				74.2	
TrialEnd_SPPB.Intervention.12		0.075				41.4	
site.BathBristol	Input to mobility model 1	0			Multinomial	335	
site.Birmingham	Input to mobility model 1	0				174	
site.Devon	Input to mobility model 1	1				268	

continued



TABLE 44 Table of input parameters for model-based economic evaluation (continued)

Parameter name	Description	Base-case value	One-way sensitivity analysis		PSA		
			Lower	Upper	Distribution	Parameter 1	Parameter 2
SPPB.M12_SPPB_FinalScore	Coefficient for SPPB at 12 months in mobility model 1	0.787	0.660	0.915	Multivariate normal (parameter 1 = mean vector, parameter 2 = variance-covariance matrix)	0.787	See Appendix 3, Table 45
SPPB.Allocation_Site.Birmingham	Coefficient for site = Birmingham in mobility model 1	0.050	-0.542	0.642		0.050	
SPPB.Allocation_Site.Devon	Coefficient for site = Devon in mobility model 1	0.572	0.107	1.037		0.572	
SPPB.SCREENING_AgeAtRandomisation	Coefficient for age at randomisation in mobility model 1	-0.052	-0.088	-0.016		-0.052	
SPPB.SCREENING_Sex.Female	Coefficient for sex = female in mobility model 1	-0.208	-0.646	0.230		-0.208	
SPPB_cut1	Mobility model 1 ordinal logistic regression cut points	-3.988	-7.070	-3.291		-3.988	
SPPB_cut2		-3.291	-3.988	-2.476		-3.291	
SPPB_cut3		-2.476	-3.291	-1.830		-2.476	
SPPB_cut4		-1.830	-2.476	-0.617		-1.830	
SPPB_cut5		-0.617	-1.830	0.119		-0.617	
SPPB_cut6		0.119	-0.617	1.060		0.119	
SPPB_cut7		1.060	0.119	1.972		1.060	
SPPB_cut8		1.972	1.060	3.115		1.972	
SPPB_cut9		3.115	1.972	4.181		3.115	
SPPB_cut10		4.181	3.115	5.188		4.181	
SPPB_cut11		5.188	4.181	6.877		5.188	
SPPB_cut12		6.877	5.188	10.139		6.877	

Parameter name	Description	Base-case value	One-way sensitivity analysis		PSA		
			Lower	Upper	Distribution	Parameter 1	Parameter 2
SPPB.offset	Additional offset parameter (drives mobility up or down)	0	-0.5	0.5	Normal	0	0.25
EQ5D.SPPB_Spline1	Coefficient for SPPB in EQ-5D utility function	0.032	0.017	0.046	Multivariate normal	0.032	See Appendix 3, Table 46
EQ5D.SPPB_Spline2	Coefficient for SPPB (spline) in EQ-5D utility function	-0.011	-0.025	0.003		-0.011	
EQ5D._cons	Intercept in EQ-5D utility function	0.449	0.357	0.540		0.449	
SPPB_Spline2.SPPB.0	SPPB spline	0					
SPPB_Spline2.SPPB.1		0					
SPPB_Spline2.SPPB.2		0					
SPPB_Spline2.SPPB.3		0					
SPPB_Spline2.SPPB.4		0					
SPPB_Spline2.SPPB.5		0.020					
SPPB_Spline2.SPPB.6		0.163					
SPPB_Spline2.SPPB.7		0.551					
SPPB_Spline2.SPPB.8		1.306					
SPPB_Spline2.SPPB.9		2.503					
SPPB_Spline2.SPPB.10		4.027					
SPPB_Spline2.SPPB.11		5.714					
SPPB_Spline2.SPPB.12		7.429					

continued

TABLE 44 Table of input parameters for model-based economic evaluation (continued)

Parameter name	Description	Base-case value	One-way sensitivity analysis		PSA		
			Lower	Upper	Distribution	Parameter 1	Parameter 2
AnnualCost.SPPB.0_3	Annual cost with SPPB 0–3	3494.62	1516.90	5472.34	Multivariate normal	3494.62	See Appendix 3, Table 47
AnnualCost.SPPB.4_7	Annual cost with SPPB 4–7	2019.66	1652.31	2387.00		2019.66	
AnnualCost.SPPB.8_9	Annual cost with SPPB 8–9	1730.03	1376.41	2083.64		1730.03	
AnnualCost.SPPB.10_12	Annual cost with SPPB 10–12	1520.00	1200.29	1839.72		1520.00	
Year1CostNHS.Control	Within-trial costs for control arm	4043.00	3553.79	4462.12	Bootstrap samples (n = 100)		
Year1CostNHS.Intervention	Within-trial costs for REACT arm (excluding intervention)	3319.00	3018.50	3699.12			
Year1QALY.Control	Within-trial QALYs for control arm	1.348	1.327	1.374			
Year1QALY.Intervention	Within-trial QALYs for REACT arm	1.388	1.364	1.419			
InterventionCost.Control	Intervention costs (control arm)	0					
InterventionCost.Intervention	Intervention costs (REACT arm)	621.83	402.42	888.22	Gamma	Shape 25	Scale 24.87
p_survive	Probability of surviving 1 year (given at age 75; varies with age)	0.978					

TABLE 45 Variance-covariance matrix for mobility model 1

Parameter	Variance-covariance matrix																
SPPB.M12_SPPB_FinalScore	0.004	0.000	0.000	-0.002	0.000	0.030	0.031	0.033	0.035	0.038	0.039	0.043	0.045	0.047	0.049	0.051	0.053
SPPB.Allocation_Site.Birmingham	0.000	0.000	0.000	0.000	0.000	0.027	0.027	0.028	0.028	0.028	0.028	0.028	0.028	0.027	0.027	0.027	0.027
SPPB.Allocation_Site.Devon	0.000	0.000	0.050	0.004	-0.004	0.059	0.061	0.058	0.054	0.058	0.061	0.060	0.060	0.060	0.057	0.056	0.054
SPPB.SCREENING_AgeAtRandomisation	-0.002	0.000	0.004	0.091	0.022	0.012	0.015	0.016	0.014	0.016	0.018	0.015	0.012	0.013	0.013	0.013	0.011
SPPB.SCREENING_Sex.Female	0.000	0.000	-0.004	0.022	0.056	0.018	0.013	0.014	0.015	0.019	0.021	0.023	0.023	0.022	0.021	0.023	0.027
SPPB_cut1	0.030	0.027	0.059	0.012	0.018	2.472	2.402	2.402	2.379	2.402	2.409	2.394	2.391	2.376	2.374	2.358	2.410
SPPB_cut2	0.031	0.027	0.061	0.015	0.013	2.402	2.482	2.456	2.424	2.437	2.443	2.426	2.425	2.407	2.406	2.389	2.436
SPPB_cut3	0.033	0.028	0.058	0.016	0.014	2.402	2.456	2.528	2.489	2.499	2.502	2.487	2.486	2.470	2.469	2.456	2.503
SPPB_cut4	0.035	0.028	0.054	0.014	0.015	2.379	2.424	2.489	2.502	2.507	2.506	2.491	2.492	2.476	2.476	2.462	2.511
SPPB_cut5	0.038	0.028	0.058	0.016	0.019	2.402	2.437	2.499	2.507	2.563	2.557	2.541	2.542	2.527	2.526	2.514	2.563
SPPB_cut6	0.039	0.028	0.061	0.018	0.021	2.409	2.443	2.502	2.506	2.557	2.576	2.559	2.558	2.543	2.543	2.530	2.581
SPPB_cut7	0.043	0.028	0.060	0.015	0.023	2.394	2.426	2.487	2.491	2.541	2.559	2.566	2.564	2.548	2.548	2.537	2.588
SPPB_cut8	0.045	0.028	0.060	0.012	0.023	2.391	2.425	2.486	2.492	2.542	2.558	2.564	2.581	2.563	2.563	2.552	2.603
SPPB_cut9	0.047	0.027	0.060	0.013	0.022	2.376	2.407	2.470	2.476	2.527	2.543	2.548	2.563	2.566	2.565	2.553	2.606
SPPB_cut10	0.049	0.027	0.057	0.013	0.021	2.374	2.406	2.469	2.476	2.526	2.543	2.548	2.563	2.565	2.586	2.571	2.621
SPPB_cut11	0.051	0.027	0.056	0.013	0.023	2.358	2.389	2.456	2.462	2.514	2.530	2.537	2.552	2.553	2.571	2.582	2.627
SPPB_cut12	0.053	0.027	0.054	0.011	0.027	2.410	2.436	2.503	2.511	2.563	2.581	2.588	2.603	2.606	2.621	2.627	2.769

TABLE 46 Variance-covariance matrix for model for EQ-5D utility by SPPB

Parameter	Variance-covariance matrix		
EQ5D.SPPB_Spline1	5.24E-05	-4.6E-05	-0.00033
EQ5D.SPPB_Spline2	-4.6E-05	5.05E-05	0.000267
EQ5D_cons	-0.00033	0.000267	0.002181

TABLE 47 Variance-covariance matrix for cost by SPPB

Parameter	Variance-covariance matrix			
AnnualCost.SPPB.0_3	1,018,202	12,910	3949	-7816
AnnualCost.SPPB.4_7	12,910	35,128	7045	5855
AnnualCost.SPPB.8_9	3949	7045	32,551	5919
AnnualCost.SPPB.10_12	-7816	5855	5919	26,609

## Model-based economic evaluation results

Results are presented in *Tables 48–50*.

TABLE 48 Economic outcomes by Markov model state

State	Control arm			Intervention arm			Difference		
	Life-years	QALYs	Costs (£)	Life-years	QALYs	Costs (£)	Life-years	QALYs	Costs (£)
<i>Within-trial</i>									
	2.00	1.31	4046	2.00	1.35	3943	= 0.000	+ 0.040	-103
<i>Markov model extrapolation</i>									
0	0.09	0.02	188	0.09	0.02	167	-0.007	-0.003	-21
1	0.09	0.03	197	0.08	0.02	169	-0.010	-0.004	-28
2	0.18	0.06	394	0.16	0.05	345	-0.017	-0.007	-49
3	0.25	0.09	564	0.22	0.08	497	-0.023	-0.010	-66
4	0.85	0.33	1148	0.78	0.29	1035	-0.068	-0.032	-113
5	0.86	0.36	1211	0.80	0.33	1110	-0.060	-0.030	-101
6	1.55	0.70	2233	1.47	0.66	2100	-0.078	-0.042	-133
7	1.96	0.96	2925	1.91	0.93	2829	-0.054	-0.032	-97
8	2.57	1.35	3390	2.58	1.35	3398	0.011	0.003	8
9	1.81	1.01	2480	1.90	1.06	2607	0.089	0.052	127
10	0.94	0.55	1169	1.04	0.62	1305	0.105	0.065	136
11	0.47	0.29	607	0.56	0.35	731	0.093	0.060	124
12	0.05	0.04	75	0.07	0.05	101	0.019	0.013	27
Total	13.68	7.11	20,627	13.68	7.18	20,338	0.000	0.072	-290

Costs and QALYs discounted at 3.5% per year; life-years are not discounted.

TABLE 49 Deterministic model-based economic evaluation results

Group	Discounted cost (£)	Discounted QALYs	Net monetary benefit (£)	ICER (£/QALY)
Control	20,627	7.111	121,594	
Intervention	20,338	7.183	123,330	
Difference	-290	0.072	1735	Dominant

TABLE 50 Population average cost-effectiveness results

Population	Incremental costs	Incremental QALYs	INMB	ICER
Base case (75-year-old woman)	-290	0.072	1735	Dominant
REACT study participants	-296	0.071	1718	Dominant
UK general population aged 65-100 years	-273	0.068	1631	Dominant

### Heterogeneity analysis

As demonstrated in heterogeneity analysis, the intervention is least cost-effective in 65-year-old men, so we repeated the PSA in 65-year-old men. All other parameters were left unchanged from the original PSA.

The 95% credible intervals for INMB and the ICER were £387 to £2811 and -£14,133 (dominant) to £10,147 per QALY, respectively.

Figures 41 and 42 show the cost-effectiveness scatter plot and cost-effectiveness acceptability curve.

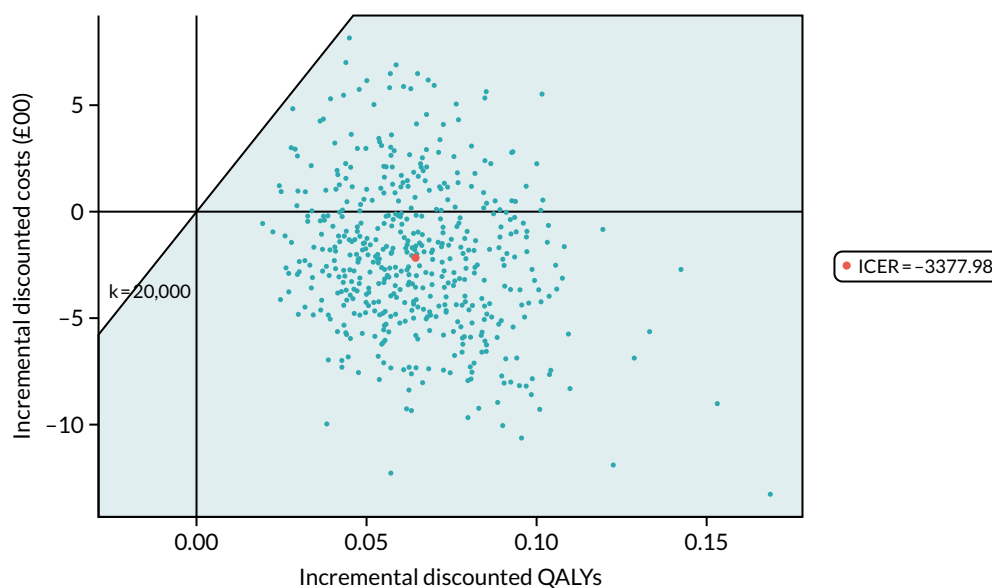


FIGURE 41 Probabilistic sensitivity analysis in 65-year-old men: cost-effectiveness plane.

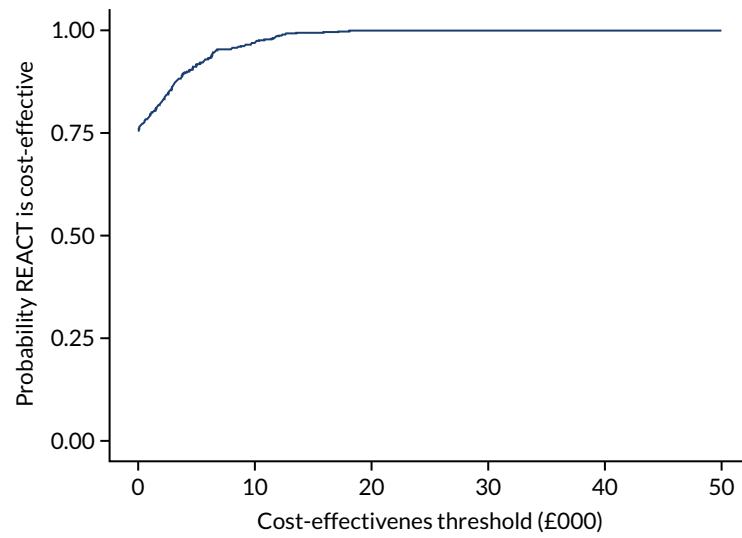


FIGURE 42 Probabilistic sensitivity analysis in 65-year-old men: cost-effectiveness acceptability curve.





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***