

This is a repository copy of *Chair-based yoga programme for older adults with multimorbidity:RCT with embedded economic and process evaluations*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/217882/>

Version: Published Version

---

**Article:**

Tew, Garry, Wiley, Laura [orcid.org/0000-0001-9619-4484](https://orcid.org/0000-0001-9619-4484), Ward, Lesley et al. (14 more authors) (2024) Chair-based yoga programme for older adults with multimorbidity:RCT with embedded economic and process evaluations. *Health technology assessment*. pp. 1-186. ISSN 2046-4924

<https://doi.org/10.3310/KPGN4216>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



## Health Technology Assessment

Volume 28 • Issue 53 • September 2024

ISSN 2046-4924

# Chair-based yoga programme for older adults with multimorbidity: RCT with embedded economic and process evaluations

*Garry Alan Tew, Laura Wiley, Lesley Ward, Jessica Grace Hugill-Jones, Camila Sofia Maturana, Caroline Marie Fairhurst, Kerry Jane Bell, Laura Bissell, Alison Booth, Jenny Howsam, Valerie Mount, Tim Rapley, Sarah Jane Ronaldson, Fiona Rose, David John Torgerson, David Yates and Catherine Elizabeth Hewitt*





# Chair-based yoga programme for older adults with multimorbidity: RCT with embedded economic and process evaluations

Garry Alan Tew<sup>1,2,3\*</sup>, Laura Wiley<sup>2</sup>, Lesley Ward<sup>2,3</sup>,  
Jessica Grace Hugill-Jones<sup>2</sup>, Camila Sofia Maturana<sup>2</sup>,  
Caroline Marie Fairhurst<sup>2</sup>, Kerry Jane Bell<sup>2</sup>,  
Laura Bissell<sup>4</sup>, Alison Booth<sup>2</sup>, Jenny Howsam<sup>4</sup>,  
Valerie Mount<sup>5</sup>, Tim Rapley<sup>6</sup>, Sarah Jane Ronaldson<sup>2</sup>,  
Fiona Rose<sup>2</sup>, David John Torgerson<sup>2</sup>, David Yates<sup>7</sup> and  
Catherine Elizabeth Hewitt<sup>2</sup>

<sup>1</sup>Institute for Health and Care Improvement, York St John University, York, UK

<sup>2</sup>York Trials Unit, Department of Health Sciences, University of York, York, UK

<sup>3</sup>Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle-upon-Tyne, UK

<sup>4</sup>British Wheel of Yoga Qualifications (BWYQ), Sleaford, Lincs, UK

<sup>5</sup>Member of the Public, York, UK

<sup>6</sup>Department of Social Work, Education and Community Well-being, Northumbria University, Newcastle-upon-Tyne, UK

<sup>7</sup>Department of Anaesthesia, York Hospitals NHS Foundation Trust, York, UK

\*Corresponding author

Published September 2024

DOI: 10.3310/KPGN4216

This report should be referenced as follows:

Tew GA, Wiley L, Ward L, Hugill-Jones JG, Maturana CS, Fairhurst CM, *et al.* Chair-based yoga programme for older adults with multimorbidity: RCT with embedded economic and process evaluations. *Health Technol Assess* 2024;**28**(53). <https://doi.org/10.3310/KPGN4216>





# Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, 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@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hta](http://www.journalslibrary.nihr.ac.uk/hta).

## Criteria for inclusion in the *Health Technology Assessment* journal

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

Reviews in *Health Technology Assessment* 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.

## HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

## This article

The research reported in this issue of the journal was funded by the HTA programme as award number 17/94/36. The contractual start date was in January 2019. The draft manuscript began editorial review in January 2023 and was accepted for publication in August 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article 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 HTA 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 HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Tew *et al.* This work was produced by Tew *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Newgen Digitalworks Pvt Ltd, Chennai, India ([www.newgen.co](http://www.newgen.co)).



# Abstract

## Chair-based yoga programme for older adults with multimorbidity: RCT with embedded economic and process evaluations

Garry Alan Tew<sup>1,2,3\*</sup> Laura Wiley<sup>2</sup> Lesley Ward<sup>2,3</sup>  
Jessica Grace Hugill-Jones<sup>2</sup> Camila Sofia Maturana<sup>2</sup>  
Caroline Marie Fairhurst<sup>2</sup> Kerry Jane Bell<sup>2</sup> Laura Bissell<sup>4</sup>  
Alison Booth<sup>2</sup> Jenny Howsam<sup>4</sup> Valerie Mount<sup>5</sup> Tim Rapley<sup>6</sup>  
Sarah Jane Ronaldson<sup>2</sup> Fiona Rose<sup>2</sup> David John Torgerson<sup>2</sup>  
David Yates<sup>7</sup> and Catherine Elizabeth Hewitt<sup>2</sup>

<sup>1</sup>Institute for Health and Care Improvement, York St John University, York, UK

<sup>2</sup>York Trials Unit, Department of Health Sciences, University of York, York, UK

<sup>3</sup>Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle-upon-Tyne, UK

<sup>4</sup>British Wheel of Yoga Qualifications (BWYQ), Sleaford, Lincs, UK

<sup>5</sup>Member of the Public, York, UK

<sup>6</sup>Department of Social Work, Education and Community Well-being, Northumbria University, Newcastle-upon-Tyne, UK

<sup>7</sup>Department of Anaesthesia, York Hospitals NHS Foundation Trust, York, UK

\*Corresponding author [g.tew@yorks.ac.uk](mailto:g.tew@yorks.ac.uk)

**Background:** Older adults with multimorbidity experience impaired health-related quality of life and treatment burden. Yoga has the potential to improve several aspects of health and well-being. The British Wheel of Yoga's Gentle Years Yoga<sup>®</sup> programme was developed specifically for older adults, including those with chronic conditions. A pilot trial demonstrated feasibility of using Gentle Years Yoga in this population, but there was limited evidence of its effectiveness and cost-effectiveness.

**Objective:** To determine the effectiveness and cost-effectiveness of the Gentle Years Yoga programme in addition to usual care versus usual care alone in older adults with multimorbidity.

**Design:** Pragmatic, multisite, individually randomised controlled trial with embedded economic and process evaluations.

**Setting:** Participants were recruited from 15 general practices in England and Wales from July 2019 with final follow-up in October 2022.

**Participants:** Community-dwelling adults aged 65 years and over with multimorbidity, defined as two or more chronic health conditions from a predefined list.

**Interventions:** All participants continued with any usual care provided by primary, secondary, community and social services. The intervention group was offered a 12-week programme of Gentle Years Yoga.

**Main outcome measures:** The primary outcome and end point were health-related quality of life measured using the EuroQol-5 Dimensions, five-level version utility index score over 12 months. Secondary outcomes were health-related quality of life, depression, anxiety, loneliness, incidence of falls, adverse events and healthcare resource use.

**Results:** The mean age of the 454 randomised participants was 73.5 years; 60.6% were female, and participants had a median of three chronic conditions. The primary analysis included 422 participants (intervention,  $n = 227$  of 240, 94.6%; usual care,  $n = 195$  of 214, 91.1%). There was no statistically or clinically significant difference in the EuroQol-5 Dimensions, five-level version utility index score over 12 months: the predicted mean score for the intervention group was 0.729 (95% confidence interval 0.712 to 0.747) and for usual care it was 0.710 [95% confidence interval (CI) 0.691 to 0.729], with an adjusted mean difference of 0.020 favouring intervention (95% CI  $-0.006$  to 0.045,  $p = 0.14$ ). No statistically significant differences were observed in secondary outcomes, except for the pain items of the Patient-Reported Outcomes Measurement Information System-29. No serious, related adverse events were reported. The intervention cost £80.85 more per participant (95% CI £76.73 to £84.97) than usual care, generated an additional 0.0178 quality-adjusted life-years per participant (95% CI 0.0175 to 0.0180) and had a 79% probability of being cost-effective at the National Institute for Health and Care Excellence threshold of £20,000 per quality-adjusted life-year gained. The intervention was acceptable to participants, with seven courses delivered face to face and 12 online.

**Limitations:** Self-reported outcome data raise the potential for bias in an unblinded trial. The COVID-19 pandemic affected recruitment, follow-up and the mode of intervention delivery.

**Conclusions:** Although the Gentle Years Yoga programme was not associated with any statistically significant benefits in terms of health-related quality of life, mental health, loneliness or falls, the intervention was safe, acceptable to most participants and highly valued by some. The economic evaluation suggests that the intervention could be cost-effective.

**Future work:** Longer-term cost-effectiveness modelling and identifying subgroups of people who are most likely to benefit from this type of intervention.

**Trial registration:** This trial is registered as ISRCTN13567538.

**Funding:** This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 17/94/36) and is published in full in *Health Technology Assessment*; Vol. 28, No. 53. See the NIHR Funding and Awards website for further award information.

# Contents

List of tables	xiii
List of figures	xvii
List of supplementary material	xix
List of abbreviations	xxi
Plain language summary	xxiii
Scientific summary	xxv
<b>Chapter 1 Introduction</b>	<b>1</b>
Burden of multimorbidity in older people	1
Interventions for improving outcomes for people with multimorbidity	1
Yoga as an intervention for improving health and well-being	1
Research aims and objectives	3
<b>Chapter 2 Methods</b>	<b>5</b>
Trial design	5
Setting	5
Eligibility and recruitment	5
<i>Eligibility criteria for participants</i>	5
<i>Recruitment of participants</i>	7
<i>Identification of trial groups</i>	7
<i>Consenting participants</i>	7
<i>Eligibility criteria for primary care general practices</i>	9
<i>Recruitment of primary care general practices</i>	9
<i>Eligibility criteria for yoga teachers</i>	9
<i>Recruitment of yoga teachers</i>	9
Intervention and comparator conditions	9
<i>Comparator description</i>	9
<i>Intervention description</i>	10
<i>Criteria for discontinuing or modifying allocated interventions</i>	14
<i>Strategies to improve adherence to interventions</i>	14
<i>Relevant concomitant care permitted or prohibited during the trial</i>	14
<i>Provisions for post-trial care</i>	14
Outcomes	14
<i>Primary outcome measure</i>	14
<i>Secondary outcomes measures</i>	15
<i>Process evaluation</i>	16
<i>Intervention fidelity</i>	16
<i>Other data collected</i>	17
Sample size	17
<i>Original</i>	17
<i>Revised</i>	17
Randomisation	18
Blinding	18

## CONTENTS

Internal pilot phase	19
Statistical methods	20
<i>Recruitment</i>	20
<i>Baseline characteristics of randomised and analysed participants</i>	20
<i>Withdrawals and follow-up</i>	20
<i>Interim pilot phase</i>	20
<i>Primary outcome analysis</i>	20
<i>Sensitivity analyses</i>	21
<i>Subgroup analysis</i>	21
<i>Secondary outcome analyses</i>	22
<i>Adverse events</i>	22
Data collection and management	22
<i>Design and processing of participant questionnaires</i>	24
<i>Plans to promote participant retention and complete follow-up</i>	24
<i>Participant withdrawal</i>	25
<i>Trial completion and exit</i>	25
Adverse event reporting and harms	25
Regulatory approvals and monitoring	26
<i>Trial registration</i>	26
<i>Summary of changes to the protocol</i>	26
<i>Trial Management Group</i>	26
<i>Trial Steering Committee</i>	26
<b>Chapter 3 Clinical effectiveness results</b>	<b>27</b>
Participant flow	27
<i>Baseline characteristics of randomised participants</i>	30
<i>Withdrawals and follow-up</i>	30
Internal pilot phase	34
Primary outcome (EQ-5D-5L utility index score) analysis	36
<i>Raw scores</i>	36
<i>Baseline characteristics of participants included in primary analysis</i>	36
<i>Primary end-point analysis</i>	39
Sensitivity analyses	41
<i>Adjusting for other covariates (sensitivity 1)</i>	41
<i>Clustering by yoga teacher (sensitivity 2)</i>	41
<i>Compliance with random allocation and treatment received</i>	41
<i>Yoga practices</i>	43
<i>Intervention fidelity</i>	44
Subgroup analyses	46
<i>Intended mode of delivery of Gentle Years Yoga</i>	46
Secondary analysis	46
<i>EuroQol-5 Dimensions, five-level version utility index scores at the secondary time points</i>	46
<i>EuroQol-5 Dimensions, five-level version visual analogue scale</i>	46
<i>Generalised Anxiety Disorder-7</i>	46
<i>Patient Health Questionnaire-8</i>	46
<i>University of California, Los Angeles-3 loneliness</i>	46
<i>English Longitudinal Study of Ageing single-item direct loneliness question</i>	46
<i>Patient-Reported Outcomes Measurement Information System-29</i>	49
<i>Falls</i>	49
Adverse events	51
<b>Chapter 4 Economic evaluation</b>	<b>53</b>

Introduction and aim	53
Methods	53
<i>Within-trial economic evaluation</i>	53
<i>Economic data collection</i>	53
<i>Health-related quality of life</i>	53
<i>Healthcare resource use</i>	54
Costs	54
<i>Costing the intervention</i>	55
<i>Missing data</i>	56
<i>Incremental analysis</i>	57
<i>Sensitivity analysis</i>	58
<i>Cost-consequences analysis</i>	58
Results	58
<i>Participant population and missing data</i>	58
<i>Health-related quality-of-life utility</i>	59
<i>Healthcare resource use and costs</i>	60
<i>Intervention cost</i>	66
<i>Cost-utility analysis and uncertainty</i>	66
<i>Sensitivity analysis</i>	68
<i>Cost-consequence analysis</i>	69
Discussion	72
<b>Chapter 5 Process evaluation</b>	<b>75</b>
Introduction and aims	75
Methods	75
<i>Study design</i>	75
<i>Sampling strategy</i>	75
<i>Recruitment, consent and data collection</i>	75
<i>Qualitative data management and analysis</i>	76
<i>Relationship between process evaluation and trial</i>	76
Results	76
<i>Enrolment: clear willingness to engage with Gentle Years Yoga</i>	78
<i>Integration: many already felt (relatively) active and healthy</i>	79
<i>Workability: engaged well with online and face-to-face delivery</i>	80
<i>Appraisal: experienced Gentle Years Yoga as gentle exercise (with mindful breathing)</i>	81
<i>Integration: variable engagement with Gentle Years Yoga, social time and home practice</i>	82
<i>Appraisal: embodied impact (and future engagement) varied</i>	83
<i>Reconfiguration: introducing and embedding gentle yoga</i>	85
Discussion	87
<b>Chapter 6 Methodological substudies</b>	<b>89</b>
Substudy 1: What is the concurrent validity of the PROMIS-29 with the EQ-5D-5L?	89
Introduction	89
Objectives	89
Methods	90
Results	91
Discussion	91
Substudy 2: Does including £5 and/or a pen in the recruitment pack enhance recruitment?	98
Background	98
Methods	98



## CONTENTS

<i>Results</i>	98
<i>Conclusion</i>	99
Substudy 3: Does sending a pen with a follow-up questionnaire enhance return rates?	99
<i>Background</i>	99
<i>Methods</i>	99
<i>Results</i>	99
<i>Conclusion</i>	99
Substudy 4: Does offering a free yoga session to control participants after the 12-month follow-up assessment enhance retention and reduce contamination?	100
<i>Introduction</i>	100
<i>Methods</i>	100
<i>Results</i>	101
<i>Discussion</i>	101
<b>Chapter 7 Discussion</b>	<b>105</b>
Summary of main findings	105
Interpretation of findings	105
<i>Intervention fidelity</i>	106
<i>Sensitivity of outcome measures</i>	106
<i>Not clinically effective</i>	107
<i>Comparison with previous studies</i>	107
<i>Cost-effectiveness</i>	108
Strengths and limitations	109
Equality, diversity and inclusion	110
Patient and public involvement	111
Implications for health care	111
<i>Does this trial provide sufficient evidence to justify implementing Gentle Years Yoga courses as a funded NHS service?</i>	111
<i>Should self-funded Gentle Years Yoga classes be recommended to older people with multimorbidity?</i>	112
<i>A word of caution</i>	112
Recommendations for research	112
Conclusion	114
<b>Additional information</b>	<b>115</b>
<b>References</b>	<b>119</b>
<b>Appendix 1</b> Research Ethics Committee approved amendments to the study protocol	<b>127</b>
<b>Appendix 2</b> Software output of model coefficients for the covariates with 95% CIs	<b>129</b>
<b>Appendix 3</b> Number and proportion of participants with complete case data by group	<b>131</b>
<b>Appendix 4</b> Further missing data information	<b>133</b>
<b>Appendix 5</b> Completion and missingness of EQ-5D-5L questionnaires	<b>135</b>
<b>Appendix 6</b> Number of missing dimensions for invalid EQ-5D-5L questionnaires	<b>137</b>
<b>Appendix 7</b> Proportion reporting EQ-5D-5L levels 1–5 by dimension	<b>139</b>

<b>Appendix 8</b> Cost-effectiveness planes for sensitivity analyses 1–6	<b>143</b>
<b>Appendix 9</b> Cost-effectiveness acceptability curves for sensitivity analyses 1–6	<b>147</b>
<b>Appendix 10</b> Expected and actual site-level ethnicity (percentages)	<b>151</b>



# List of tables

<b>TABLE 1</b> Details of props used	12
<b>TABLE 2</b> Schedule of assessments	23
<b>TABLE 3</b> Reason for declining participation in the GYY trial	29
<b>TABLE 4</b> Reason for ineligibility in the GYY trial	29
<b>TABLE 5</b> Baseline characteristics of randomised participants by group	31
<b>TABLE 6</b> Self-reported health conditions at baseline by randomised group	32
<b>TABLE 7</b> Participant-reported expectations and preferences concerning the health care being offered in the GYY trial, collected at baseline, by randomised group	33
<b>TABLE 8</b> Values of outcome measures assessed at baseline by randomised group	33
<b>TABLE 9</b> Return rates for post-randomisation follow-up questionnaires	35
<b>TABLE 10</b> Summaries of raw EQ-5D-5L utility index score by trial arm and time point	36
<b>TABLE 11</b> Baseline characteristics of randomised participants by group for those included in primary analysis	36
<b>TABLE 12</b> Grouped conditions self-reported at baseline by randomised group as analysed	37
<b>TABLE 13</b> Participant-reported expectations and preferences concerning the health care being offered in the GYY trial, collected at baseline, by randomised group as analysed	38
<b>TABLE 14</b> Values of outcome measures assessed at baseline by randomised group as analysed	39
<b>TABLE 15</b> Difference in adjusted mean EQ-5D-5L utility index score over time by randomised group from primary and SA models	40
<b>TABLE 16</b> Summary of GYY class attendance by week and recruitment wave	42
<b>TABLE 17</b> Definitions of adherence to GYY intervention by recruitment wave	43
<b>TABLE 18</b> Self-reported yoga practice at 3-month follow-up	44
<b>TABLE 19</b> Self-reported yoga practice at 6-month follow-up	45
<b>TABLE 20</b> Self-reported yoga practice at 12-month follow-up	45

<b>TABLE 21</b> Summary of raw scores for secondary outcomes	47
<b>TABLE 22</b> Difference in adjusted means over time by randomised group for secondary outcomes	47
<b>TABLE 23</b> Summary of raw scores for PROMIS-29 secondary outcomes	48
<b>TABLE 24</b> Difference in adjusted means over time by randomised group for PROMIS-29 outcomes	50
<b>TABLE 25</b> Summary of non-SAEs	52
<b>TABLE 26</b> Unit costs of healthcare services	55
<b>TABLE 27</b> Summary of EQ-5D-5L utility index scores and unadjusted and adjusted MD (95% CI) at each time point (all available cases)	59
<b>TABLE 28</b> Total QALYs for all available cases by allocation over 12-month trial duration and difference in mean QALYs (95% CI)	60
<b>TABLE 29</b> Mean resource use at each study time point, based on all available cases	60
<b>TABLE 30</b> Total mean costs based on all available cases, up to 12-month follow-up	64
<b>TABLE 31</b> Total mean costs of health service use, medication cost and intervention cost based on multiply imputed data set, up to 12-month follow-up	65
<b>TABLE 32</b> Cost of the GYY intervention	66
<b>TABLE 33</b> Mean costs and QALYs by group (multiply imputed data set)	67
<b>TABLE 34</b> Cost–utility analysis results	67
<b>TABLE 35</b> Cost–consequence balance sheet	70
<b>TABLE 36</b> Summary of all participants' involvement in interviews and observation	77
<b>TABLE 37</b> Characteristics of qualitative interview participants: yoga and usual care arms ( $n = 27$ )	77
<b>TABLE 38</b> Overview of participants' main reasons for joining the trial	79
<b>TABLE 39</b> Examples of reported physical and psychosocial impacts of GYY on participants	84
<b>TABLE 40</b> Examples of recommendations on potential implementation strategies to support any future roll-out of GYY	86
<b>TABLE 41</b> Correlations between the PROMIS-29 and EQ-5D-5L scores	92
<b>TABLE 42</b> Mean change score (paired $t$ -test) and responsiveness estimates of SRM and Cohen's $d$ effect size for each measure and time interval	94

<b>TABLE 43</b> Correlation of change scores of EQ-5D-5L with change scores for PROMIS-29	<b>96</b>
<b>TABLE 44</b> Participants achieving minimum clinically important difference, based on an increase of 0.06 in EQ-5D-5L index value between time points	<b>97</b>
<b>TABLE 45</b> Mean and SD of EQ-5D-5L index value by time point and trial site	<b>130</b>
<b>TABLE 46</b> Number and proportion of participants with complete case data by group	<b>131</b>
<b>TABLE 47</b> Logistic regression for indicators of missing cost and QALY data on baseline variables	<b>133</b>
<b>TABLE 48</b> Completion and missingness of EQ-5D-5L questionnaires by group and time point	<b>135</b>
<b>TABLE 49</b> Number of missing dimensions for invalid EQ-5D-5L questionnaires by group and time point	<b>137</b>
<b>TABLE 50</b> Proportion reporting EQ-5D-5L levels 1–5 by dimension, group and time point	<b>140</b>
<b>TABLE 51</b> Expected and actual site-level ethnicity (percentages)	<b>152</b>



# List of figures

<b>FIGURE 1</b> General practice recruitment flowchart	<b>8</b>
<b>FIGURE 2</b> Chair-based postures that are commonly used in GYY.	<b>11</b>
<b>FIGURE 3</b> Consolidated Standards of Reporting Trials diagram: recruitment to the GYY trial in the pilot phase waves	<b>27</b>
<b>FIGURE 4</b> Consolidated Standards of Reporting Trials diagram: recruitment to the GYY trial in the main phase waves	<b>27</b>
<b>FIGURE 5</b> Consolidated Standards of Reporting Trials diagram: randomisation and follow-up in the GYY trial. (a) Withdrawals and deaths over time are cumulative	<b>28</b>
<b>FIGURE 6</b> Adjusted mean EQ-5D-5L utility index scores (with 95% CIs) for primary analysis over time by randomised group	<b>40</b>
<b>FIGURE 7</b> Number of sessions attended by GYY intervention group participants	<b>43</b>
<b>FIGURE 8</b> Cost-effectiveness plane for the GYY intervention relative to usual care: base-case analysis (MI adjusted for all covariates)	<b>68</b>
<b>FIGURE 9</b> Cost-effectiveness acceptability curve for the GYY intervention relative to usual care: for the base-case analysis (MI adjusted for all covariates)	<b>68</b>
<b>FIGURE 10</b> Receiver operating characteristic curves for the PROMIS-29 physical and mental health scores against the external criterion of an improvement equating to the minimum clinically important difference of 0.06 in the EQ-5D-5L index value between baseline and month 12	<b>97</b>
<b>FIGURE 11</b> Location of GYY-trained yoga teachers	<b>113</b>
<b>FIGURE 12</b> Cost-effectiveness plane for SA1: complete-case analysis	<b>143</b>
<b>FIGURE 13</b> Cost-effectiveness plane for SA2: personal costs	<b>143</b>
<b>FIGURE 14</b> Cost-effectiveness plane for SA3: face-to-face yoga classes only	<b>144</b>
<b>FIGURE 15</b> Cost-effectiveness plane for SA4: online yoga classes only	<b>144</b>
<b>FIGURE 16</b> Cost-effectiveness plane for SA5: medication costs removed	<b>145</b>
<b>FIGURE 17</b> Cost-effectiveness plane for SA6: age and gender removed	<b>145</b>
<b>FIGURE 18</b> Cost-effectiveness acceptability curve for SA1: CCA	<b>147</b>
<b>FIGURE 19</b> Cost-effectiveness acceptability curve for SA2: personal cost	<b>147</b>



## LIST OF FIGURES

<b>FIGURE 20</b> Cost-effectiveness acceptability curve for SA3: face-to-face yoga classes only	<b>148</b>
<b>FIGURE 21</b> Cost-effectiveness acceptability curve for SA4: online yoga classes only	<b>148</b>
<b>FIGURE 22</b> Cost-effectiveness acceptability curve for SA5: medication costs removed	<b>149</b>
<b>FIGURE 23</b> Cost-effectiveness acceptability curve for SA6: age and gender removed	<b>149</b>

# List of supplementary material

**Report Supplementary Material 1** Summary of GP practices that supported recruitment

**Report Supplementary Material 2** Summary of yoga courses delivered in the trial (yoga teachers, delivery mode, location and number of participants allocated)

**Report Supplementary Material 3** Actual time between randomisation and commencement and any gaps in delivery for each yoga course delivered

**Report Supplementary Material 4** Yoga teacher research training manual

**Report Supplementary Material 5** Substudy 2: Does including £5 and/or a pen in the recruitment pack enhance recruitment?

**Report Supplementary Material 6** Substudy 3: Does sending a pen with a follow-up questionnaire enhance return rates?

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/KPGN4216>).

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

AE	adverse event	MD	mean difference
AIC	Akaike information criterion	MI	multiply imputed
AUC	area under the curve	NICE	National Institute for Health and Care Excellence
BWY	British Wheel of Yoga	NIHR	National Institute for Health and Care Research
CACE	complier-average causal effect	PHQ-8	Patient Health Questionnaire-8 item
CCA	complete-case analysis	PIS	participant information sheet
CI	confidence interval	PROMIS-29 v2.1	29-item Patient-Reported Outcomes Measurement Information System
CONSORT	Consolidated Standards of Reporting Trials	QALY	quality-adjusted life-year
ELSA	English Longitudinal Study of Ageing	RCT	randomised controlled trial
EQ-5D-5L	EuroQol-5 Dimensions, five-level version	ROC	receiver operating characteristic
GAD-7	Generalised Anxiety Disorder-7	SA	sensitivity analysis
GP	general practitioner	SAE	serious adverse event
GYG	Gentle Years Yoga	SD	standard deviation
HCRW	Health and Care Research Wales	SMD	standardised mean difference
HR	Hazard Ratio	SRM	standardised response mean
HRA	Health Research Authority	TSC	Trial Steering Committee
HRQoL	health-related quality of life	UCLA	University of California, Los Angeles
HTA	Health Technology Assessment	VAS	visual analogue scale
IMD	Index of Multiple Deprivation	WTP	willingness to pay
ITT	intention to treat	YTU	York Trials Unit
MAR	missing at random		
MCID	minimal clinically important difference		



## Plain language summary

It is common for older adults to have two or more long-term health conditions. These conditions affect quality of life differently, with some people feeling well and others needing healthcare support. The Gentle Years Yoga programme was developed to improve quality of life for older adults, including those with long-term health conditions. We wanted to see how well the programme worked and if it offered good value for money for the NHS.

We tested whether offering a 12-week course of Gentle Years Yoga improved the quality of life and reduced anxiety, depression, loneliness and falls for people aged 65 years and over who had two or more long-term health conditions. We recruited 454 people through general practices across England and Wales, with 240 people selected at random to be invited to take part in the Gentle Years Yoga programme and the other 214 to continue with their usual care and not be offered Gentle Years Yoga. The average age of participants was 74 years, nearly two-thirds were female and the number of long-term health conditions participants had ranged from two to nine (average was three). They completed four questionnaires over a 12-month period. We also interviewed some of the participants and the yoga teachers to find out how the approach worked in practice. The yoga was delivered either face to face or online.

We did not find any significant benefits in terms of quality of life, anxiety, depression, loneliness or falls. At interview, some yoga participants noted no or a modest impact on their health or lifestyle, while others described Gentle Years Yoga as transformative, having substantial impacts and improvements on their physical health and emotional well-being. Because running the yoga classes was relatively inexpensive and some insignificant benefits were seen, the Gentle Years Yoga programme may be good value for money.



# Scientific summary

## Background

Multimorbidity, having two or more chronic health conditions, is a major challenge for older adults and increases healthcare utilisation and associated costs. Multimorbidity is associated with reduced health-related quality of life (HRQoL), impaired functional status, worse physical and mental health and premature death. In 2015, 54% of people aged 65 years and over in England had multimorbidity.

There has been limited exploration of the effectiveness of interventions to improve outcomes for people with multimorbidity. There is some evidence to suggest that yoga may help to prevent and treat various physical and mental illnesses and improve HRQoL. The available data offer support for the beneficial effects of yoga in older adults and for several chronic conditions. However, robust evidence of clinical and cost effectiveness is limited, and little research has specifically focused on older adults with multimorbidity.

The Gentle Years Yoga (GY) programme was developed to cater specifically to the needs of older adults, including those with health conditions common to an older cohort such as osteoarthritis, hypertension and cognitive impairment. A pilot randomised trial of the GY programme ( $n = 52$  adults, mean age 75 years) demonstrated feasibility, with a potential for a clinically important benefit on health status [EuroQoL-5 Dimensions, five-level version (EQ-5D-5L) utility index score] at 3 months after randomisation [mean difference 0.12, 95% confidence interval (CI) 0.03 to 0.21]. Consequently, we conducted this larger trial, as if shown to be clinically and cost-effective, GY could be widely implemented, leading to improved outcomes for this population.

## Objectives

The primary objective was to establish if the offer of a 12-week GY programme in addition to usual care is more effective compared with usual care alone in improving HRQoL (EQ-5D-5L utility index score) over 12 months in people aged 65 years or over with multimorbidity.

Secondary objectives were to:

- explore the effect of the GY programme on HRQoL, depression, anxiety, loneliness and incidence of falls
- explore the safety of the GY programme in terms of the occurrence of adverse events
- assess the cost-effectiveness of the GY programme
- undertake a qualitative process evaluation to explore the acceptability of the intervention and the experience of participants and teachers, explain the determinants of delivery and identify the optimal implementation strategies.

## Methods

### Design

This was a multisite, two-arm, parallel-group, superiority, individually randomised controlled trial comparing an experimental strategy of offering a 12-week GY programme against a control strategy of no offer of GY in community-dwelling people aged 65 years or over who had multimorbidity. Both trial arms continued with any usual care provided by primary, secondary, community and social services.



The study also included cost-effectiveness and qualitative process evaluations.

### **Setting**

Participants were recruited from primary care general practices in the UK. General practices interested in taking part in the trial were identified with help from the NHS Clinical Research Networks and the Health and Care Research Wales Support and Delivery Centre. General practices were selected based on their proximity to yoga class venues and local transport routes.

The yoga courses were delivered either face-to-face in a yoga studio, community hall or leisure centre or online via video conferencing during periods of social distancing restrictions.

The trial's yoga consultants identified yoga teachers who were eligible and potentially interested in taking part in the trial. Yoga teachers needed to have completed the British Wheel of Yoga (BWY) Qualification Level 4 Teaching GYY and have valid BWY membership and insurance. For online courses, teachers also needed to be proficient in remote teaching.

### **Participant recruitment and consent**

Patients were eligible to join the study if they were aged 65 years or older, community-dwelling and had two or more of the predefined chronic health conditions derived from the NHS Quality and Outcomes Framework. Exclusion criteria were: inability to attend one of the GYY courses on offer; yoga practice in the previous 6 months; contraindications to yoga participation; severe mental health problem; learning disability; or being unable to provide consent and/or return the baseline questionnaire. For online classes, ineligibility also included no internet access; inability to use the internet; no suitable device; insufficient space at home; and/or no sturdy chair for use during the classes.

Potential participants were identified by searching general practitioner (GP) electronic patient databases. Participating practices ran a custom-built search, based on predefined read codes, which identified patients with eligible health conditions.

Potentially eligible patients were sent a recruitment pack including an information sheet and, if interested, were invited to sign and return a consent form. After the research team checked their suitability for the trial with their GP, eligible patients were asked to provide baseline data on sociodemographic measures, primary and secondary outcome measures and preferences/beliefs for the treatments on offer in the trial. Participants indicated on the consent form if they also wanted to be considered to take part in the process evaluation interviews.

### **Intervention and comparison**

Gentle Years Yoga is designed for older adults, including those with chronic conditions. Based on standard Hatha Yoga, it incorporates physical postures and transitions as well as breathing, concentration and relaxation activities. The aims of GYY are to improve muscle strength, flexibility, balance, mobility and mental and social well-being. Chairs are used for seated exercise and support when standing. The yoga practices are modified so individuals with varying medical conditions and functional abilities can participate safely. Props are used to modify some of the postures and concentration activities. The physical challenge of each posture can be progressed throughout the course as participants become more able and confident.

Participants randomised to the intervention were invited to take part in a free GYY course. Each course involved 12, 75-minute sessions of group-based yoga, usually delivered 1 week apart, either face to face or online. Each class included: 'housekeeping' activities (5 minutes); an introduction to the theme and practices of the class, basic breathing and focusing activities (5 minutes); an extended warmup/mobilisation and preparatory postures (30–35 minutes); focused postures and restorative activities (10–15 minutes); breathing exercises (5–10 minutes); and relaxation and concentration activities (5–10 minutes), followed by optional after-class social time (15–30 minutes).

Throughout the trial, both trial arms continued with any usual care provided by primary care, secondary care, community and social services.

### Sample size

We proposed to randomise 586 participants in a 1 : 1 ratio to be able to detect a difference of 0.06 in EQ-5D-5L utility index score, assuming a standard deviation (SD) of 0.20, with 90% power, a two-sided alpha of 0.05 and 20% attrition. In October 2021, an interim calculation of the correlation between baseline and 12-month EQ-5D-5L utility index score indicated we would be able to detect a clinically important difference with close to or greater than 90% power with 454 participants, since the primary analysis adjusted for baseline score, which affords gains in power.

### Randomisation

Participants were randomised via a central, computer-based randomisation system designed and managed by York Trials Unit (YTU), University of York. Randomisation was stratified by site using varying block sizes and allocation ratios. Blinding of participants or the yoga teachers was not possible.

### Outcome measures

Outcomes were self-reported by the participant and collected using questionnaires at baseline and 3, 6 and 12 months after randomisation.

The primary outcome and end point was the EQ-5D-5L utility index score over the 12-month follow-up period. Utility index scores were calculated following current National Institute for Health and Care Excellence guidance.

Secondary outcomes were:

- EQ-5D-5L utility index score at 3, 6 and 12 months after randomisation
- EQ-5D-5L visual analogue scale (VAS) score at 3, 6 and 12 months and overall
- HRQoL at 3, 6 and 12 months and overall using PROMIS-29 (v2.1)
- depression severity at 3, 6 and 12 months and overall using the Patient Health Questionnaire-8 (PHQ-8)
- anxiety severity at 3, 6 and 12 months and overall using the Generalised Anxiety Disorder-7 (GAD-7)
- loneliness at 3, 6 and 12 months and overall. Questions used to capture loneliness were taken from the English Longitudinal Study of Ageing (ELSA), based on the University of California, Los Angeles 3-item (UCLA-3) loneliness scale and a direct question about how often the respondent felt lonely
- the incidence of falls, adverse events and healthcare resource use over 12 months.

### Analysis

All analyses were conducted in Stata version 17 (StataCorp LP, College Station, TX, USA) following the principles of intention-to-treat using two-sided statistical tests at the 5% significance level. The primary outcome was analysed using a linear mixed model, including data at all available follow-up time points, adjusting for baseline EQ-5D-5L utility index score, time point, trial arm and an arm-by-time interaction as fixed effects. Participant (to account for the repeated measures) and site were included as random effects.

In sensitivity analyses, age, gender and adapted Bayliss (severity-adjusted count of health conditions) score were added as covariates to the primary analysis model, and the site was substituted for yoga teacher as a random effect.

Complier-average causal effect (CACE) analyses, using a two-stage instrumental variable regression approach with randomised group as the instrumental variable, were implemented to assess the impact of receiving GYY on the primary treatment estimate.

A subgroup analysis was conducted to assess for differential effects of the intervention based on mode of delivery (face to face or online).

The secondary outcomes of EQ-5D VAS, GAD-7, PHQ-8, T-scores from each of the seven subscales of the PROMIS-29 and the physical and mental health component scores and the global item score, UCLA-3 score and ELSA single-item direct loneliness question were analysed using the same methods as for the primary outcome, with baseline EQ-5D-5L utility index score swapped as a covariate for baseline value of the outcome.

The incidence of falls during the 12-month follow-up period was analysed using a mixed-effect negative binomial regression model, adjusting for the number of falls in the 3 months prior to baseline and site as a random effect.

Adverse events are summarised descriptively.

### **Economic analysis**

The economic analysis assessed the relative cost-effectiveness of the GYY programme in addition to continued access to usual care compared with usual care alone. Costs and health outcomes were evaluated from the perspective of the NHS and Personal Social Services using a within-trial economic analysis and a cost-consequence analysis, both over a 12-month time horizon; hence, discounting was not required. Health outcomes were assessed in terms of quality-adjusted life-years (QALYs) using EQ-5D-5L data, with costs collected for healthcare resource use, medications and the intervention. Findings were presented in terms of the incremental cost-effectiveness ratio for the intervention versus usual care and net monetary benefit. The base-case analysis was undertaken on an intention-to-treat basis, with multiple imputation used to deal with missing data and sensitivity analyses conducted to explore uncertainty around the cost-effectiveness findings.

### **Process evaluation**

The process evaluation was informed by qualitative interviews with trial participants, trial decliners, trial yoga teachers and stakeholders, as well as by observations of standardisation training sessions and yoga classes. Intervention fidelity was assessed by class observation and via interviews with teachers as part of the process evaluation.

Interviews were audio-recorded, transcribed verbatim and edited to ensure respondent anonymity. Data analysis was iterative throughout the trial and conducted according to standard procedures of qualitative analysis.

## **Results**

Between July 2019 and August 2021, 13,070 invitation packs were sent to potentially eligible participants, of which 1297 responded. Of these, 454 (35.0%) participants were randomised: 240 to the intervention and 214 to usual care. The mean age of participants was 73.5 years (range 65–99); 60.6% were female, and participants had a median of three chronic conditions.

Among the intervention group, the mean number of GYY sessions attended was 8.8 (SD 3.7, median 10, range 0–12). Two hundred and twenty-two (92.5%) participants attended at least one session, and fifty-three (22.1%) attended all twelve. Eighty per cent ( $n = 192$ ) attended at least three of the first six sessions and at least three other sessions. One participant in the usual care group was invited to attend classes in error; they attended eight sessions, including five of the first six.

The primary analysis included 422 participants with valid EQ-5D-5L data at baseline and at least one post-randomisation time point (intervention  $n = 227$  of 240, 94.6%; usual care  $n = 195$  of 214, 91.1%).

There was no statistically or clinically significant difference in the EQ-5D-5L utility index score over 12 months: the predicted mean score for the intervention group was 0.729 (95% CI 0.712 to 0.747) and for usual care was 0.710 (95% CI 0.691 to 0.729), with an adjusted mean difference of 0.020 favouring intervention (95% CI -0.006 to 0.045;  $p = 0.14$ ). The sensitivity analyses produced very similar results. The CACE analyses, which considered compliance as attending (1)  $\geq 1$  GYY session and (2)  $\geq 6$  including 3 of the first 6, produced slightly greater, but not clinically relevant, treatment estimates (0.025, 95% CI -0.002 to 0.052;  $p = 0.07$ ; and 0.029, 95% CI -0.002 to 0.059;  $p = 0.06$ , respectively).

There was no evidence of an interaction between trial arm and intended mode of delivery (interaction effect 0.007, 95% CI -0.042 to 0.057;  $p = 0.77$ ).

No statistically significant differences were observed in secondary outcomes, except in the T-score for the pain interference subscale of the PROMIS-29 at 3 months (-1.44, 95% CI -2.63 to -0.26;  $p = 0.02$ ) and over the 12 months (-1.14, 95% CI -2.24 to -0.04;  $p = 0.04$ ), and in the global (pain intensity) PROMIS-29 item at 12 months (-0.45, 95% CI -0.83 to -0.08;  $p = 0.02$ ) and over the 12 months (-0.32, 95% CI -0.61 to -0.04;  $p = 0.03$ ), favouring intervention.

No serious, related adverse events were reported.

### **Economic evaluation**

The base-case economic evaluation found that the intervention cost £80.85 more per participant (95% CI £76.73 to £84.97) than usual care, generated an additional 0.0178 QALYs per participant (95% CI 0.0175 to 0.0180) and had a 79% probability of being cost-effective at the willingness-to-pay threshold of £20,000 per QALY gained.

### **Process evaluation**

Participants found both face-to-face and online courses acceptable. Participants were highly motivated; most viewed their health as good and engaged well with the GYY classes over time. The majority viewed GYY as a form of gentle exercise with mindful breathing. Some participants noted no or only a modest impact of GYY on their health or lifestyle, including physical, psychological and self-management benefits, while others described GYY as transformative, having substantial impacts and improvements on their physical health and emotional well-being.

## **Conclusions**

The offer of a 12-week GYY programme was not associated with any statistically significant benefits in terms of HRQoL, mental health, loneliness or falls in older adults with multimorbidity. However, the intervention was safe, acceptable to most participants and highly valued by some. The economic evaluation suggests that the intervention could be cost-effective.

## **Trial registration**

This trial is registered as ISRCTN13567538.

## **Funding**

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 17/94/36) and is published in full in *Health Technology Assessment*; Vol. 28, No. 53. See the NIHR Funding and Awards website for further award information.



# Chapter 1 Introduction

Material throughout this chapter has been reproduced from Tew *et al.*<sup>1</sup>

This is an open access article that is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made. See: <https://creativecommons.org/licenses/by/4.0/>. The text below includes minor additions and formatting changes to the original text.

## Burden of multimorbidity in older people

Multimorbidity, often defined as the co-existence of two or more chronic medical conditions,<sup>2</sup> is a major challenge for health and care systems worldwide and of particular relevance for older adults. In 2015, 54% of people aged 65 years or older in England exhibited multimorbidity; this percentage is projected to increase to 68% by 2035.<sup>3</sup> Multimorbidity is associated with poorer outcomes such as reduced quality of life, impaired functional status, worse physical and mental health and premature death.<sup>4,5</sup> Multimorbidity also increases healthcare utilisation and associated costs.<sup>6,7</sup>

## Interventions for improving outcomes for people with multimorbidity

There has been limited exploration of the effectiveness of interventions to improve outcomes for people with multimorbidity. A 2021 systematic review identified 16 randomised controlled trials (RCTs) with 4753 participants that had evaluated a range of complex interventions for people with multimorbidity in primary care and community settings.<sup>8</sup> Eight studies examined multifaceted interventions that targeted the co-ordination of care and healthcare providers while also providing self-management support for patients. Four studies reported on self-management support interventions that did not have a clear link to the patients' healthcare provision. The other four studies focused primarily on medication management. The results suggested that all intervention types probably make little or no difference to health-related quality of life (HRQoL) or mental health outcomes. Five of the 10 studies with HRQoL outcomes reported EuroQoL 5 Dimensions (EQ-5D) (a generic measure of health utility) scores that could be included in a meta-analysis, with a mean difference (MD) of 0.03 [95% confidence interval (CI) -0.01 to 0.07], consistent with the overall effect suggesting no difference in this outcome. There was also little or no effect on clinical, psychological or medication outcomes or healthcare utilisation. There were mixed effects on function, activity and patient health behaviours, and limited data on costs. There was a low risk of bias overall; however, the evidence for all outcomes was downgraded to low certainty due to serious concerns about inconsistency and imprecision. This review highlighted the need for further research to determine the clinical and cost-effectiveness of interventions that are ideally simple, generalisable and which can address several medical conditions simultaneously. Yoga is a potential candidate intervention.

## Yoga as an intervention for improving health and well-being

Yoga originated thousands of years ago in India as an integrated mind-body practice based on ancient Vedic philosophy. During the 20th century, yoga became increasingly recognised outside India, and over the past decades, it has continued to grow in popularity worldwide as a practice for improving health and well-being. While modern yoga often focuses primarily on physical poses and is sometimes thought

of as a type of exercise, the practice usually incorporates one or more of the mental or mindful elements that are traditionally part of yoga, such as relaxation, concentration or meditation. There are currently many different styles or schools of yoga, each with a variable emphasis and approach to practice. Research evidence suggests that some of these yoga practices may help to prevent and treat various physical and mental illnesses and improve HRQoL.<sup>9,10</sup>

In November 2017, the Cochrane Library published a special collection of 14 systematic reviews that focused on the effectiveness of yoga for improving physical or mental health symptoms and quality of life in a range of health conditions, including musculoskeletal, pulmonary, cancer, cardiovascular, neurological and mental health. A summary of four diverse but pertinent reviews is as follows:

- Yoga for chronic non-specific low back pain:<sup>11</sup> For yoga compared to non-exercise controls (9 trials; 810 participants), there was moderate-certainty evidence that yoga produced small-to-moderate improvements in back-related function [standardised mean difference (SMD) -0.44, 95% CI -0.66 to -0.22] and pain (MD -7.81, 95% CI -13.37 to -2.25) at 6 months. The authors recommended additional high-quality research to improve confidence in estimates of effect and to evaluate long-term outcomes.
- Yoga for asthma:<sup>12</sup> There was some evidence that yoga may improve quality of life (MD in Asthma Quality of Life Questionnaire score per item 0.57 units on a 7-point scale, 95% CI 0.37 to 0.77; five studies;  $n = 375$ ) and symptoms (SMD 0.37, 95% CI 0.09 to 0.65; three studies;  $n = 243$ ) and reduce medication usage (risk ratio 5.35, 95% CI 1.29 to 22.11; two studies) in people with asthma. The authors concluded that large, high-quality trials are needed to confirm the effects of yoga on asthma.
- Yoga for improving HRQoL, mental health and cancer-related symptoms in women diagnosed with breast cancer:<sup>13</sup> Seventeen studies that compared yoga versus no therapy provided moderate-quality evidence showing that yoga improved HRQoL (SMD 0.22, 95% CI 0.04 to 0.40; 10 studies,  $n = 675$ ), reduced fatigue (SMD -0.48, 95% CI -0.75 to -0.20; 11 studies,  $n = 883$ ) and reduced sleep disturbances in the short term (SMD -0.25, 95% CI -0.40 to -0.09; six studies,  $n = 657$ ). No serious adverse events (SAEs) were reported. Additional research was recommended to assess medium- and longer-term effects.
- Yoga for primary prevention of cardiovascular disease:<sup>14</sup> Yoga was found to produce reductions in diastolic blood pressure (MD -2.90 mmHg) and triglycerides (MD -0.27 mmol/l) and increase high-density lipoprotein cholesterol (MD 0.08 mmol/l). There was no clear evidence of a difference between groups for low-density lipoprotein cholesterol, although there was moderate statistical heterogeneity. Adverse events (AEs), occurrence of type 2 diabetes and costs were not reported in any of the studies. No study reported cardiovascular mortality, all-cause mortality or non-fatal events, and most studies were small and short term.

Elsewhere, studies have sought to determine the effects of yoga in older populations. For example, a 2012 systematic review of 16 studies ( $n = 649$ )<sup>15</sup> and a more recent trial of 118 participants<sup>16</sup> demonstrated that yoga may provide greater improvements in physical functioning and self-reported health status than conventional physical activity interventions in older adults. More recently, a systematic review of six trials ( $n = 307$ ) of relatively high methodological quality reported that yoga interventions had a small beneficial effect on balance (SMD 0.40, 95% CI 0.15 to 0.65, six trials) and a medium effect on physical mobility (SMD 0.50, 95% CI 0.06 to 0.95, three trials) in people aged 60 and over.<sup>17</sup>

In summary, these data offer support for the beneficial effects of yoga in older adults and for several chronic conditions. However, many of the previous studies had limitations, including small sample sizes, a single yoga teacher delivering the programme and short-term follow-up. Robust economic evaluations of yoga are also limited, although a recent systematic review concluded that 'medical' yoga is likely to be a cost-effective option for low back pain.<sup>18</sup> Very little research has specifically focused on older people with multimorbidity.



In 2009, the Gentle Years Yoga© (GYG) programme was developed by the Yorkshire Yoga and Therapy Centre to cater specifically to the needs of older adults, including those with conditions common to an older cohort such as osteoarthritis, hypertension and cognitive impairment. As part of the pilot research study conducted at Yorkshire Yoga in 2016, a standardised GYG teacher training programme was manualised with the creation of a quality-assured teacher training course which became the British Wheel of Yoga (BWY) GYG programme that is being delivered by the BWY. British Wheel of Yoga is the National Governing Body of Yoga in Great Britain, with a nationwide network of over 5000 qualified yoga teachers. Gentle Years Yoga is based on standard Hatha Yoga, incorporating traditional physical poses and transitions as well as breathing, concentration and relaxation activities. Adaptations to challenging Hatha Yoga poses have been made so that older adults can participate safely while still obtaining the fitness, health and well-being benefits of yoga. Each programme involves one group-based session per week for 12 weeks (each session includes a 75-minute chair-based yoga class and after-class social time) and promotion of regular self-managed yoga practice at home.

In a pilot trial of the GYG programme,<sup>19</sup> 82 older adults expressed an interest within a 2-month recruitment period, of which 52 (mean age 75 years) were recruited and randomised. Participants had up to six chronic conditions, the most common of which were osteoarthritis, hypertension and depression. Trial yoga courses were delivered across four community venues by four yoga teachers. Two-thirds of participants had an acceptable attendance of  $\geq 80\%$ . The study demonstrated feasibility of evaluating the GYG programme in a fully powered RCT and the potential for a positive clinically important effect on health status [EuroQol-5 Dimensions, five-level version (EQ-5D-5L) utility index score] at 3 months after randomisation (MD 0.12, 95% CI 0.03 to 0.21).

Consequently, we conducted a larger trial – The GYG Trial – to establish the clinical and cost-effectiveness of the GYG programme in older adults with multimorbidity. If this intervention was shown to be clinically effective and cost-effective, it could be implemented more widely, leading to improved outcomes for this population.

## Research aims and objectives

The GYG Trial was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) Programme in response to a themed call on complex health and care needs in older people. The aim of the trial was to establish the clinical and cost effectiveness of the GYG programme in addition to usual care versus usual care alone in community-dwelling older adults with multimorbidity.

The primary objective was to establish if the offer of a 12-week GYG programme in addition to usual care is more effective compared with usual care alone in improving HRQoL (EQ-5D-5L utility index score) over 12 months in people aged 65 years or over with multimorbidity.

Secondary objectives were as follows:

- to explore the effect of the GYG programme on HRQoL, depression, anxiety and loneliness at 3, 6 and 12 months after randomisation
- to explore the effect of the GYG programme on the incidence of falls over 12 months from randomisation
- to explore the safety of the GYG programme relative to control in terms of the occurrence of AEs over 12 months after randomisation
- to assess if the GYG programme is cost-effective, measured using differences in the cost of health resource use between the intervention and usual care groups and the incremental cost-effectiveness



ratios (ICER) using quality-adjusted life-years (QALYs) derived from the EQ-5D-5L measured at 3, 6 and 12 months after randomisation

- to undertake a qualitative process evaluation to describe the experience of the intervention, explain the determinants of delivery (including treatment fidelity) and identify the optimal implementation strategies for embedding and normalising the GYY programme in preparation for a wider roll-out.

## Chapter 2 Methods

Material throughout this chapter has been reproduced from Tew *et al.*<sup>1</sup>

This is an open access article that is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made. See: <https://creativecommons.org/licenses/by/4.0/>. The text below includes minor additions and formatting changes to the original text.

### Trial design

This was a multisite, two-arm, parallel-group, superiority, individual RCT comparing an experimental strategy of offering a 12-week GYY programme against a control strategy of no offer of GYY in community-dwelling people aged 65 years or over who had multimorbidity. Both trial arms continued with any usual care provided by primary, secondary, community and social services.

The study also included an internal pilot, economic analysis of cost-effectiveness (see [Chapter 4](#)), a qualitative process evaluation (see [Chapter 5](#)) and four methodological substudies (see [Chapter 6](#)) that addressed the following questions:

- What is the concurrent validity of the 29-item Patient-Reported Outcomes Measurement Information System® (PROMIS-29) against the EQ-5D-5L?
- Does including £5 and/or a pen in the recruitment pack enhance recruitment?
- Does sending a pen with a follow-up questionnaire enhance return rates?
- Does offering a free yoga session to control participants after the 12-month follow-up assessment enhance retention and reduce contamination?

### Setting

Participants were recruited from primary care general practices serving nine geographical areas: Harrogate, Hull, Wirral, Kent, Bristol, Oxford, Wantage and Banbury in England and Newport in Wales. The 15 general practitioner (GP) practices that supported recruitment are listed in [Report Supplementary Material 1](#). The yoga courses were delivered either face to face in a non-medical community-based facility (e.g. yoga studio, community hall, leisure centre) or online via video conferencing during periods of social distancing restrictions resulting from the COVID-19 pandemic (see [Report Supplementary Material 2](#) for further details).

### Eligibility and recruitment

#### Eligibility criteria for participants

Patients were eligible to join the study if they were aged 65 years or older (both male and female), community-dwelling (including sheltered housing living with support) and had two or more of the following chronic conditions, as derived from a list included in the NHS Quality and Outcomes

## METHODS

Framework (QOF version 31.0) and following discussions with the Trial Management Group clinical oversight and yoga consultants:

- arthritis: including osteoarthritis, rheumatoid arthritis and history of shoulder, hip or knee arthroplasty for arthritis
- asthma or chronic obstructive pulmonary disease
- atrial fibrillation
- bowel problems: including irritable bowel syndrome, diverticulitis and inflammatory bowel disease
- cancer, diagnosed within the last 5 years
- cardiovascular disease: including coronary heart disease (includes angina and history of heart attack, bypass surgery or angioplasty), hypertension, heart failure and peripheral arterial disease
- chronic kidney disease
- dementia (only if patients have the capacity to provide written informed consent)
- depression or anxiety
- diabetes
- epilepsy
- fibromyalgia
- multiple sclerosis
- osteoporosis or osteopenia
- Parkinson's disease
- sensory conditions: including hearing loss, macular degeneration, cataracts and glaucoma
- stroke, within the last 5 years.

Patients were ineligible for the study if they met one or more of the following exclusion criteria:

- inability to attend one of the yoga courses on offer [participants needed to indicate that they would be available to attend at least 9 of the 12 classes on offer for a particular course. In relation to online classes, additional factors that made someone ineligible included: no internet access, unfamiliarity with or inability to use the internet, no suitable device for accessing the online classes (e.g. tablet-size screen or larger; device with camera and microphone), insufficient space at home and no sturdy chair for use in the classes]
- attended yoga classes twice a month or more in the previous 6 months
- contraindications to yoga participation (as identified by the patient's GP)
- severe mental health problem: Schizophrenia, bipolar affective disorder or other psychotic illness (on advice from the Trial Management Group yoga consultants, potential participants with severe mental health problems and learning disabilities were listed as exclusions on the basis that the GYY teacher training programme does not cover how to accommodate these people, who may have specific support needs, and it was thought that this might make the yoga class size of 12–15 difficult to manage)
- learning disability (on advice from the Trial Management Group yoga consultants, potential participants with severe mental health problems and learning disabilities were listed as exclusions on the basis that the GYY teacher training programme does not cover how to accommodate these people, who may have specific support needs, and it was thought that this might make the yoga class size of 12–15 difficult to manage)
- unable to read or speak English (potential participants who were unable to read or speak English were listed as exclusions due to the uncertainty of being able to adapt the course delivery and questionnaires to accommodate different languages)
- unable to provide consent
- unable to complete and return a valid baseline questionnaire
- no more than one trial participant per household [no more than one trial participant per household could take part to avoid contamination effects (i.e. if one was allocated to the intervention group and the other to the usual care group)]
- currently enrolled in another research study for which concurrent participation is deemed inappropriate by their GP or a clinician co-investigator.

All eligibility criteria were assessed by reviewing responses to specific questions posed in a screening questionnaire [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] that was self-reported by potential participants either in written format or over the telephone with a researcher. Patients who did not meet the eligibility criteria were notified in writing or via a phone call that they were ineligible, and no further correspondence was sent. Patients who were deemed eligible were required to complete a baseline questionnaire [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] either in written format or over the telephone with a researcher. Participants were eligible to be randomised once the study team had received their completed consent form and baseline questionnaire. If a completed baseline questionnaire was not received by the end of the recruitment period, the patient was sent a letter to inform them that recruitment to the trial had closed and that they were unable to participate.

### **Recruitment of participants**

Potential participants were identified by searching the electronic patient databases (SystemOne or EMIS) of 15 general practices. Participating practices ran a custom-built search based on pre-defined read codes, which identified patients with health conditions within the eligibility criteria. A GP then reviewed the resulting list to rule out patients who did not meet the eligibility criteria. Where there were more potentially eligible patients identified than required, practices were asked to order the list of patients by NHS number and select the top required number. Docmail (a third-party information handler) then mailed out a recruitment pack to each of the remaining potentially eligible patients, which included a covering invitation letter, a participant information sheet (PIS) which included a link to an audio version on the trial website, a consent form and a screening questionnaire [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] and two prepaid envelopes for returning the completed forms. If the patient was deemed eligible (based on the information provided in the screening questionnaire), a researcher notified the patient's GP and asked them to confirm the patient's suitability for participation. Eligible patients were then sent a baseline questionnaire [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] to complete and return by a specified date. This questionnaire collected data on sociodemographic measures, primary and secondary outcome measures and preferences/beliefs for the treatments on offer in the trial.

Implementation of an alternative process was required during the COVID-19 pandemic as it was difficult to collect study forms by post, particularly when the trial team was required to work from home. Therefore, participants gave consent electronically by completing an online form and provided screening and baseline questionnaire data via a telephone call with a researcher, who entered these data into an online form during the call.

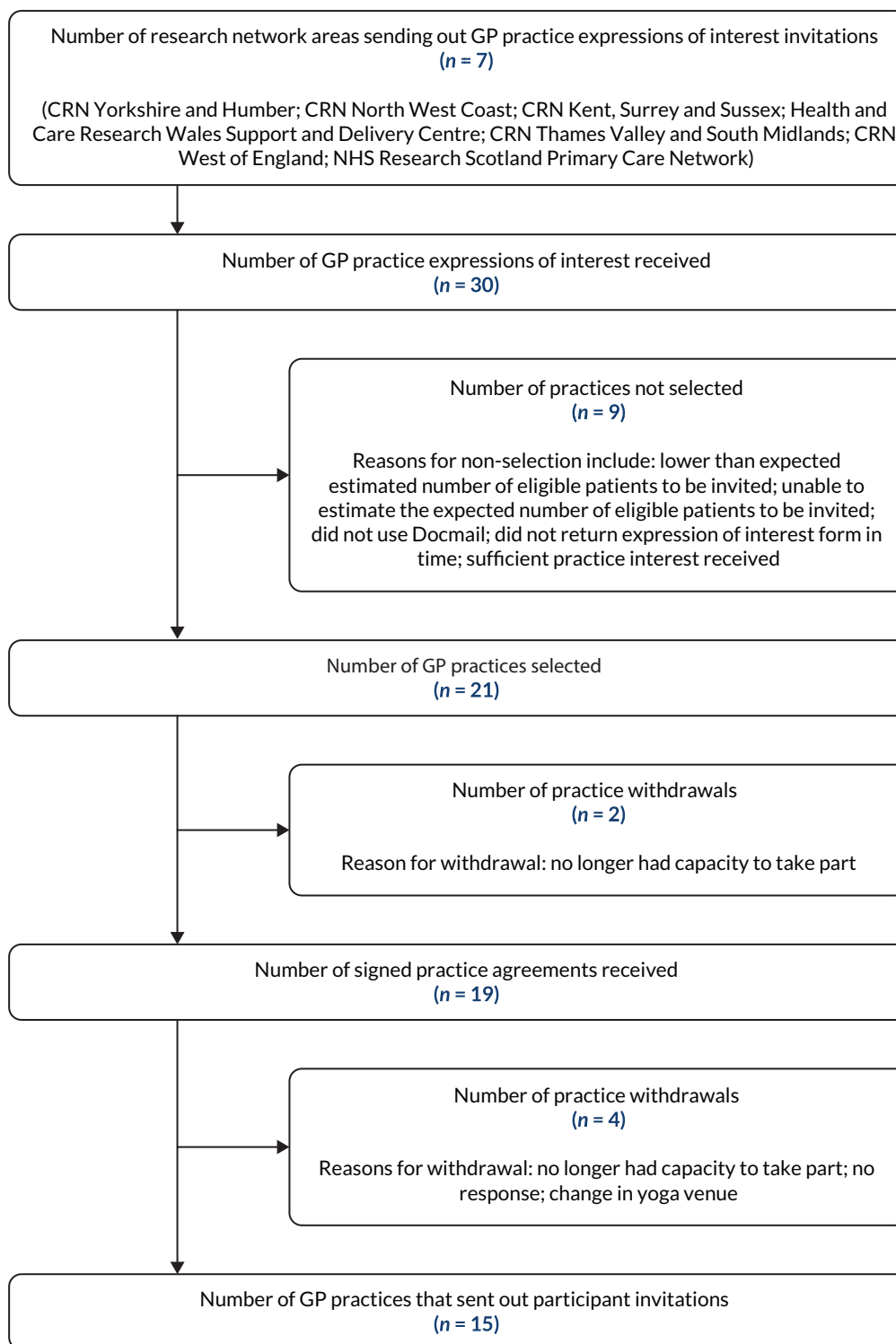
### **Identification of trial groups**

Participants had to be recruited in groups for the yoga courses, and the courses were run over a period of 29 months. To facilitate identification, each recruitment drive was called a 'wave'. There were two waves in the internal pilot phase: pilot wave 1 (PPW1) and pilot wave 2 (PPW2) and two waves in the main phase of the trial: main wave 1 (MPW1) and main wave 2 (MPW2). The number of courses run within a wave varied depending on GP practice capacity to recruit and yoga teacher availability.

### **Consenting participants**

As detailed above, potentially eligible patients were posted a recruitment pack, which included a covering invitation letter, an information sheet, a screening questionnaire and a hard copy of (or electronic link to) the consent form.

The information sheet provided a balanced written account of the purpose and design of the trial and also included details of who to contact to ask any questions and how they could access an audio recording of the information sheet on the trial website. Part way through recruitment, we introduced a simple diagram



**FIGURE 1** General practice recruitment flowchart.

of the trial design to be provided along with the information sheet to facilitate participants’ understanding of the randomisation process and what each group was to receive [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)].

Participants indicated on the consent form if they also wanted to be considered to take part in the process evaluation interviews. If the participant indicated ‘yes’ and was selected by the process

evaluation researcher to take part in the interviews, they were then provided with an information sheet regarding the interviews [see Process evaluation project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. Additional consent was sought, either via hard copy or electronically, in line with alternative arrangements, from those who agreed to take part [see Process evaluation project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. Individuals who declined to participate in the trial were able to indicate their willingness to take part in the interviews as a 'trial decliner', and additional consent was sought from them, as above.

### ***Eligibility criteria for primary care general practices***

The selection of general practices was informed by location (i.e. located close, generally within five miles, to the yoga class venue); local transport routes and teacher recommendations for face-to-face classes; computer system used; patient list size and practice staff availability for conducting recruitment activities.

### ***Recruitment of primary care general practices***

General practices that were potentially interested in taking part in the trial were identified with help from the NHS Clinical Research Networks (CRN) in England and the Health and Care Research Wales (HCRW) Support and Delivery Centre. We also worked with NHS Research Scotland (NRS) Primary Care Research Network but did not receive any expressions of interest from invited general practices via this network. Trial coordinators then liaised with key stakeholders at each practice (e.g. practice managers, GPs) to explain the requirements of the study. If a practice agreed to take part, a practice-level agreement form was signed to confirm capacity and capability before any recruitment activity commenced. This initially was the Health Research Authority (HRA) and HCRW Statement of Activities document [see GP practice project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] which was later replaced by the Integrated Research Application System (IRAS) Organisations Information Document [see GP practice project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. Practice recruitment is represented in [Figure 1](#).

### ***Eligibility criteria for yoga teachers***

To be eligible for inclusion to deliver the intervention programme, yoga teachers needed to have completed the British Wheel of Yoga Qualifications (BWYQ) Level 4 Teaching GYY qualification and have valid BWY membership and insurance. For online courses, teachers also needed to be proficient in remote teaching. The selection of yoga teachers was informed by observations of them leading online non-trial yoga classes, as conducted by the trial's yoga consultants (LB and JH).

### ***Recruitment of yoga teachers***

Yoga teachers who were potentially interested in taking part in the trial were identified by the trial's yoga consultants (LB and JH). The consultants were aware of who had completed the Level 4 GYY teacher training course (because of their work at the awarding organisation BWYQ) and where they were based. The consultants and trial coordinators liaised with potential teachers to explain the requirements of the study. This included both delivery of the courses and provision of backup yoga teachers to cover absence. If a teacher agreed to take part, a contract with the University of York was signed before any trial classes were delivered.

## **Intervention and comparator conditions**

### ***Comparator description***

The comparator was usual care alone. Usual care was defined as 'The wide range of care that is provided in a community, whether it is adequate or not, without a normative judgment'.<sup>20</sup> Throughout the trial, both trial arms continued with any usual care provided by primary care, secondary care, community and social services. This approach reflects the main aim of this pragmatic trial: to determine the clinical effectiveness and cost-effectiveness of offering the GYY programme in addition to usual care.

To characterise and quantify usual care, self-reported healthcare resource use (NHS and private care) was collected at baseline and at each follow-up assessment for all participants in both intervention and usual care groups. Prescription data from general practices were also collected for the period between 3 months prior to baseline and up until 12 months after, for a sub-sample of participants in both groups.

The protocol did not restrict access to yoga or any other intervention during the follow-up period. However, to reduce contamination, the trial yoga teachers were asked to only deliver trial classes to the participants who had been randomly allocated to their courses. In the follow-up questionnaires, all participants were asked to report any trial and non-trial, supervised yoga classes and self-managed yoga practice that they had done since the previous study time point.

### **Intervention description**

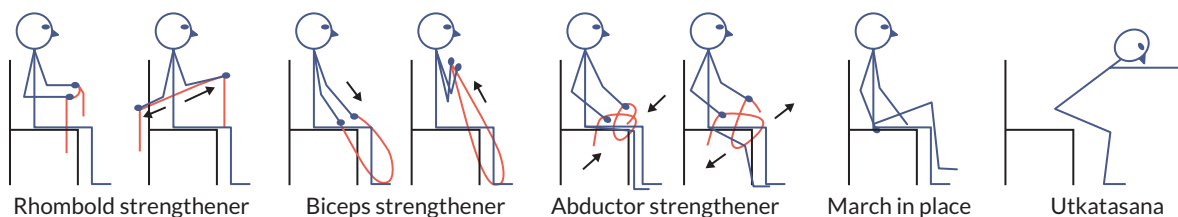
Gentle Years Yoga is a yoga programme for older adults, including those with chronic health conditions. It is based on standard Hatha Yoga and incorporates traditional physical postures and transitions as well as breathing, concentration and relaxation activities. Each class also includes optional, additional time immediately following the class for participants to stay on and socialise. The main aims of the programme are to improve muscle strength, flexibility, balance, mobility and mental and social well-being. Chairs are used for seated exercise and can be used to provide support when standing, although the whole session can be carried out on a chair. [Figure 2](#) shows examples of seated postures that are commonly used. The yoga practices are modified to allow individuals with varying medical conditions and functional abilities to participate safely. Props are also sometimes used to modify some of the postures and concentration activities. A list of props used in the trial is detailed in [Table 1](#). The physical challenge of each posture can be progressed throughout the course as participants become more able and confident. The following summarises how the GYY classes differ from standard Hatha Yoga classes:

- For the most part, participants are seated on chairs, and when standing, they use the chair or other aids for support.
- The classes do not use supine, semi-supine or prone postures; instead, the key elements of traditional supine and prone postures are integrated into seated or standing postures.
- The classes hold static postures for a shorter length of time, especially those that could cause more pronounced acute increases in blood pressure.
- The physical set-up of classes has been adapted to suit people with sensory impairments; specifically, participants being relatively close to the teacher; lighting levels being higher; the colour of equipment being in contrast to that of the walls, the floor and the teacher, and no music played during verbal instructions.
- The pace and overall structure of the class allow greater time for recovery from the more intense activities (e.g. by having a simple breathing practice follow a more-challenging physical posture).
- If there are individuals with mild cognitive impairment in the group, the teacher will use short, single-subject phrases and pace the instructions to allow time for processing each element of the instructions.
- The classes have a longer warm-up period and an overall slower pace, making it safer for older adults and at a level where they can work without feeling 'left behind' or 'too old for yoga' or having their self-confidence eroded.
- Breathing practices avoid retention, as this is contraindicated for individuals with hypertension.
- Mobilisation, postures and concentration activities are incorporated that specifically focus on balance and co-ordination.

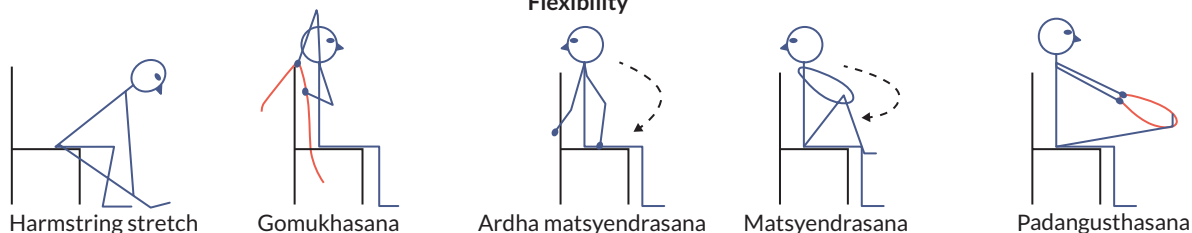
Participants randomised to the intervention were invited to take part in 12 free-of-charge, 75-minute group-based GYY classes. Each class was immediately followed by an optional 15- to 30-minute period of socialising. All courses commenced within 3 weeks of randomisation. Each course involved one class per week for 12 (mostly) consecutive weeks; however, there was allowance for a gap in delivery for

## BWY GENTLE YEARS YOGA® ASANA

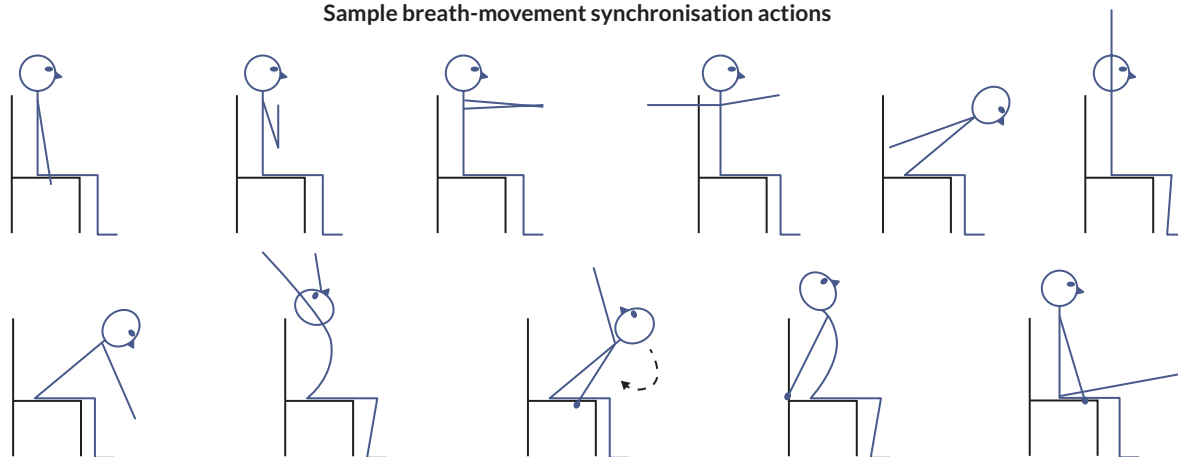
## Muscle strengthening



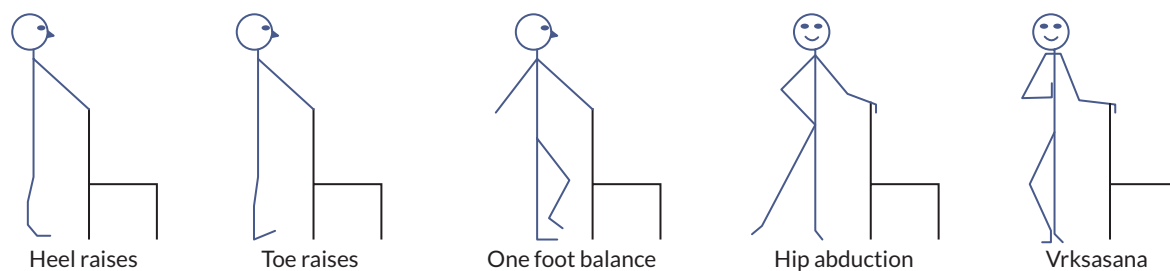
## Flexibility



## Sample breath-movement synchronisation actions



## Steps to balancing



## Asana with support

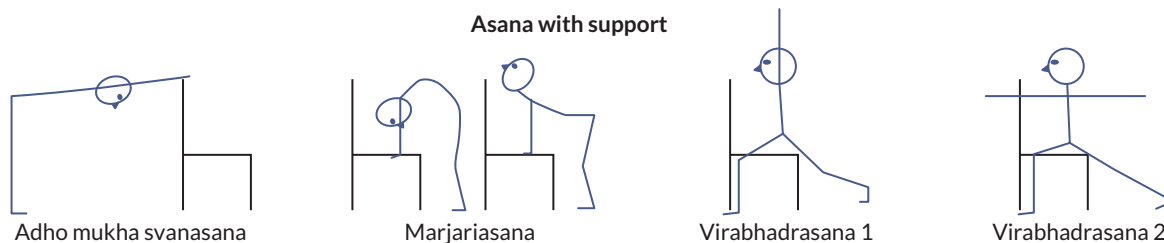


FIGURE 2 Chair-based postures that are commonly used in GYY. Reproduced from Tew *et al.*<sup>19</sup>

public holidays and unforeseen circumstances. A summary of actual time between randomisation and commencement of a yoga course and between each class, indicating any gaps in delivery, is provided



**TABLE 1** Details of props used

Item	Description	Possible uses
Resistance bands	Brand: Meglio Material: thermoplastic elastomer Resistance level: extra light, 1–4 lbs Size: 1.2 m Colour: yellow	Upper and lower body resistance – biceps, triceps, quadriceps; finger/hand mobilisation and dexterity; concentration and focused sensation
Scarfs	Brand: TecUnite <sup>a</sup> /Geebor <sup>b</sup> /QLOUNI <sup>b</sup> Material: silk fabric Size: 24 × 24 inches Colour: assorted	Toss/catch; squeeze hands; finger/hand mobilisation and dexterity; hand-to-hand passing; shoulder mobilisation; breath work, esp. longer exhalation focus; concentration and focused sensation
Beanbag	Brand: Pretext/First-Play Material: nylon Covered/Cotton covered <sup>b</sup> Size: 10 × 8 cm/15 × 10 cm Weight: 255.15 g/110 g Colour: assorted	Toss/catch; squeeze hands; squeeze toes; finger dexterity; foot balance; hand balance; concentration and focused sensation; hand-to-hand/hand-to-foot passing; counting beans meditation
Squishy ball <sup>a</sup>	Brand: MIMIEYES Material: polyurethane Size: 2.5 inches Weight: 170 g Colour: yellow (with smiley faces)	Toss/catch; squeeze hands; squeeze toes; roll hands (limited); finger dexterity; foot balance; hand balance; concentration
Tennis ball <sup>b</sup>	Brand: Wilson Material: felt Size: 2.5 inches Weight: 230 g Colour: yellow/orange	Toss/catch and bounce pulse raisers; roll hands; roll feet; foot balance; hand balance; concentration; focused sensation
Block <sup>b</sup>	Brand: Yoga Studio Material: recycled chip foam Size: 12 × 2 × 8 inches Weight: 500 g Colour: mottled/multicoloured	Aligning body in seated Tadasana; use as a platform for exercising toes, feet, ankles, fingers, hands, wrists

<sup>a</sup> Used in Pilot Phase Only.

<sup>b</sup> Used in Main Phase Only (note: change in equipment due to yoga teacher feedback from process evaluation).

in *Report Supplementary Material 3*. Before the first class, participants were required to complete and submit a BWY Health Questionnaire to their yoga teacher [see Yoga participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. This questionnaire was needed for BWY teaching insurance purposes and so that the teachers were aware of each participant's medical conditions and physical activity status.

Each class included: (1) 'housekeeping' activities, such as completing the class register and discussing any home practice or health issues (5 minutes); (2) an introduction to the theme and practices of the class, basic breathing and focusing activities (5 minutes); (3) an extended warmup/mobilisation and preparatory postures (30–35 minutes); (4) focused postures and restorative activities (10–15 minutes); (5) breathing exercises (5–10 minutes) and (6) relaxation and concentration activities (5–10 minutes), followed by optional after-class social time (15–30 minutes).

The after-class social time provided an opportunity for participant interaction and building of social networks. The yoga teachers encouraged participants to stay on for this component, but it was not mandatory. The focus of conversations was not standardised; however, the teachers were advised that there was a preference for non-yoga-based discussions. The teacher could still provide general yoga advice to participants if directly asked.

The classes also included instruction on yoga activities that the participants could do at home. Yoga teachers provided participants with a home practice sheet that detailed these activities, which were to be completed at home on non-class days. As the class-based activities became more challenging, students were given new home practice sheets to allow progression of their home yoga routine. Each yoga participant received four home practice sheets in total, covering weeks 1–3, 4–6, 7–9 and 10–12, respectively [see Yoga participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. Each sheet included at least five practices, providing an expected practice time of 10–20 minutes per session.

Towards the end of the course (i.e. classes 11 or 12), the yoga teachers provided participants with general verbal advice about continuing yoga practice and a written or electronic handout [see Yoga participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] sign-posting participants to suitable yoga classes (i.e. GYY or similar) in their local community or online, which they could attend on a self-pay basis.

The delivery mode of the trial yoga classes was originally face to face. Online classes were implemented during Pilot Phase Wave 2 (autumn 2020) and addressed the closure of venues during the first national lockdown and, later, the need for social distancing during the COVID-19 pandemic. The trial classes included intervention participants only (i.e. members of the public could not participate), with each 12-week course having 12 participants allocated if delivered online and up to 15 participants allocated if delivered face to face (see [Report Supplementary Material 2](#)). Face-to-face classes were conducted in non-medical community-based facilities (e.g. yoga studio, community hall, leisure centre) following checks for venue suitability by the yoga consultants (LB and JH). Accessibility factors that were considered included close proximity to public transport links, parking facilities and disability access. Online classes were conducted via Zoom, an online livestream platform freely accessible to participants using a computer or tablet. Pre-course contact by the yoga teachers with participants was introduced from the start of Pilot Phase Wave 2. Prior to online classes commencing, the yoga teacher conducted one-to-one Zoom meetings with their participants to optimise the set-up of their equipment and environment and to discuss any health issues or course queries. Prior to face-to-face classes commencing, the yoga teachers conducted one-to-one telephone meetings with their participants to discuss any health issues or course queries.

### Yoga teacher training

Nineteen yoga courses were delivered within the trial by 12 yoga teachers (1 teacher delivered 3 courses, 5 teachers delivered 2 courses each and 6 teachers delivered 1 course each). All teachers had the BWYQ Level 4 Teaching GYY qualification, appropriate insurance and experience of working with older adults. They also received specific training about the trial and its procedures from the research team.

The GYY programme is copyrighted by the BWY, and since 2017 they have provided training in GYY to qualified yoga teachers. Training for the regulated level 4 Teaching GYY qualification takes place over approximately 12 months and covers the National Occupational Standards for understanding the principles of adapting physical activity for older adults and the planning, adaptation and delivery of sessions to meet the requirements of participants with specific needs. This includes information on ageing, barriers and motivators to exercise, ethical and legal responsibilities, the physiology of ageing and common chronic conditions and how to modify yoga for different health states. After distance learning modules and face-to-face instruction, the teachers demonstrate their understanding through worksheets, multiple-choice questions, two case studies, designing a GYY programme and being observed and assessed on their teaching of GYY sessions on two occasions.

To minimise inter-teacher variation and enhance fidelity of intervention delivery, the yoga teachers received standardisation training from the research team via a 1-day interactive workshop. The training included background information about the GYY programme and the trial, clarification of

standardised class content and structure and practical delivery tips, including intervention progression and provision of home practice sheets. It also stressed the importance of only allowing people in the trial's intervention group to access the classes and explained trial processes such as AE reporting and class attendance monitoring. To supplement this training day, teachers also received a research training manual (see [Report Supplementary Material 4](#)) and feedback from a yoga consultant (JH) on a 12-week course plan that each teacher needed to submit at least 2 weeks before their first trial class. [see Yoga teacher project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024) for GYY course plan feedback template.] During the intervention delivery period, the yoga teachers were able to contact the yoga consultants (LB and JH, who developed the GYY teacher training course) for advice about the intervention and the trial coordinators for advice about trial processes.

### ***Criteria for discontinuing or modifying allocated interventions***

There were no specific criteria for discontinuing or modifying allocated interventions. Participants could decide to stop doing the yoga programme at any point and for any reason.

### ***Strategies to improve adherence to interventions***

To optimise and encourage attendance, the teachers were asked to contact participants who missed two consecutive classes without prior notification.

### ***Relevant concomitant care permitted or prohibited during the trial***

Participants continued to receive independent, usual care throughout the trial, and this was not prohibited in any way.

### ***Provisions for post-trial care***

Towards the end of the course (i.e. classes 11 or 12), the yoga teachers provided intervention participants with general verbal advice about continuing yoga practice and a paper or electronic handout [see Yoga participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] sign-posting participants to suitable yoga classes (i.e. GYY or similar) in their local community or online, which they could attend on a self-pay basis. The usual care participants received the same information after completing the final (12-month) follow-up questionnaire. Usual care participants were also randomised into a methodological substudy in which half were offered a GYY class at the end of their 12-month follow-up. They were informed of this offer shortly after randomisation in order to determine whether this improved their retention in the trial. Methods and results of this substudy are detailed in [Chapter 6](#).

## **Outcomes**

### ***Primary outcome measure***

The primary outcome was the EQ-5D-5L utility index score.<sup>21</sup> The EQ-5D-5L was self-reported by the participant and collected using questionnaires at baseline and at 3, 6 and 12 months post randomisation. The primary end point was the overall difference over the 12-month follow-up period. There is currently work ongoing to develop a valuation set for the EQ-5D-5L for England; in the meantime, the UK's National Institute for Health and Care Excellence (NICE) recommends<sup>22</sup> that utility index scores should be calculated using the crosswalk developed by van Hout *et al.*;<sup>23</sup> hence, they were calculated on this basis.

The EQ-5D™ is a widely used self-reported health utility measure that comprises two parts: the classification of five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a visual analogue scale (VAS), which records participants' overall evaluation of their health on a scale from 100 (best imaginable health) to 0 (worst imaginable health). The EQ-5D has been validated in many different patient populations, including those with diabetes, cardiovascular problems, chronic obstructive pulmonary disease, cancer, chronic pain and rheumatoid arthritis.

There are currently two versions of the instrument that can be used for adults: the original EuroQol-5 Dimensions, three-level version (EQ-5D-3L) with five dimensions of health and three response levels of problems and the more recent EQ-5D-5L that has the same five dimensions of health but has five response levels of problems (1 = no problems, 2 = slight problems, 3 = moderate problems, 4 = severe problems and 5 = unable/extreme problems). The EQ-5D-5L helps overcome problems with ceiling effects and has greater sensitivity.<sup>21</sup> It showed evidence of good sensitivity in the pilot trial of the GYY programme<sup>19</sup> and has been the primary outcome measure in other primary care-based multimorbidity trials.<sup>24</sup> Responses to the EQ-5D-5L lead to 3125 unique possible combinations of health states where each health state is mapped to a utility index score (on a scale where negative values correspond to a state worse than death, 0 corresponds to a health state equivalent to being dead and 1 corresponds to perfect health) by making use of a valuation set. Participants who die can be given a score of 0 (for both the utility index score and VAS) for any assessment time point following their date of death.

Besides being used as the primary outcome measure in the analysis, the EQ-5D-5L was also used to estimate QALYs for the economic evaluation (see [Chapter 4](#)).

### Secondary outcomes measures

All secondary outcomes were self-reported by the participant and collected using questionnaires at baseline and during the 12-month follow-up [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. The secondary outcomes were:

- EQ-5D-5L utility index score at 3, 6 and 12 months after randomisation
- EQ-5D-5L VAS score at 3, 6 and 12 months and overall
- HRQoL at 3, 6 and 12 months and overall using the PROMIS-29<sup>25</sup>
- depression severity at 3, 6 and 12 months and overall using the Patient Health Questionnaire-8 (PHQ-8)<sup>26</sup>
- anxiety severity at 3, 6 and 12 months and overall using the Generalised Anxiety Disorder-7 (GAD-7)<sup>27</sup>
- loneliness at 3, 6 and 12 months and overall. Four questions were used to capture different aspects of loneliness. The first three questions were taken from the University of California, Los Angeles 3-item (ULCA-3) loneliness scale.<sup>28</sup> The wording of the UCLA questions and response options was taken from the English Longitudinal Study of Ageing (ELSA).<sup>29</sup> The last was a direct question about how often the respondent feels lonely
- incidence of falls over 12 months
- AEs over 12 months
- healthcare resource use over 12 months (see [Chapter 4](#), Economic Evaluation).

### Patient-Reported Outcomes Measurement Information System-29

The PROMIS-29 v2.1 measurement scale consists of seven subscales and a global item. Both a physical and a mental health component summary score can also be generated.

The seven subscales look at the following aspects of HRQoL: physical function, anxiety, depression, fatigue, sleep disturbance, ability to participate in social roles and activities and pain interference. Each of the items in the seven subscales takes an integer-valued response ranging from 1 to 5, where higher-numbered responses indicate a worse quality of life for the anxiety, depression, fatigue, sleep disturbance and pain interference subscales but better for the physical function and ability to participate in social roles and activities subscales. The PROMIS can be scored in multiple ways. Manual scoring: for each of the seven subscales, a raw score is calculated as the sum of the values of the responses for each item in the subscale, and this raw score is converted to a T-score using the conversion tables specified in the PROMIS Adult Profile Scoring Manual ([www.healthmeasures.net/images/PROMIS/manuals/Scoring\\_Manuals/\\_PROMIS\\_Adult\\_Profile\\_Scoring\\_Manual.pdf](http://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manuals/_PROMIS_Adult_Profile_Scoring_Manual.pdf)). However, this does not deal with missing data as it uses a 'complete-case' approach, where an overall score for a subscale is only calculated if there are no missing responses for any of the items in the subscale. The preferred method

of scoring is to use the HealthMeasures Scoring Service ([www.assessmentcenter.net/ac\\_scoring-service](http://www.assessmentcenter.net/ac_scoring-service)). This method of scoring uses responses to each item for each participant (referred to as 'response pattern scoring'). Response pattern scoring is preferred because it is more accurate than the use of raw score/scale score look-up tables included in the manual. Response pattern scoring is especially useful when there is missing data (i.e. a respondent skipped an item) (text copied and adapted from the PROMIS Profile Scoring Manual). We used the online Assessment Centre scoring in this trial.

The global item asks the participant to rate their pain on an 11-point scale ranging from 0 to 10, where 0 corresponds to 'no pain at all' and 10 is 'worse imaginable pain'.

The mental and physical health component summary scores were calculated as detailed in [https://labs.dgsom.ucla.edu/hays/files/view/docs/programs-utilities/prom29/PROMIS29\\_Scoring\\_08082018.pdf](https://labs.dgsom.ucla.edu/hays/files/view/docs/programs-utilities/prom29/PROMIS29_Scoring_08082018.pdf). A higher score indicates a more favourable outcome.

Each of the PROMIS-29 v2.1 subscales, the physical and mental health component scores (PCS, MCS) and the global item at 3, 6 and 12 months post randomisation, as well as overall, were used as secondary outcomes.

### **Patient Health Questionnaire-8**

The eight-item PHQ-8 instrument is used to measure depression severity and asks the participant to indicate how often in the last 2 weeks they have been bothered by eight problems, each scored on the scale 0 = 'Not at all', 1 = 'Several days', 2 = 'More than half the days', and 3 = 'Nearly every day'. A total score is obtained from summing the eight-item scores. If one item was missing from the score, it was substituted with the average score of the non-missing items (scored pro rata and total score rounded to the nearest integer).<sup>30</sup> Questionnaires with two or more missing values were not scored. A total score of 5–9 represents mild depressive symptoms; 10–14, moderate; 15–19, moderately severe and 20–24, severe.

### **Generalised Anxiety Disorder-7**

The GAD-7 is a seven-item instrument asking the participant to indicate how often over the last 2 weeks they have been bothered by seven problems; each scored 0 = 'not at all', 1 = 'several days', 2 = 'more than half the days' and 3 = 'nearly every day'. A total score is obtained by summing the seven item scores from 0 to 21. If one item was missing from the score, then it was substituted with the average score of the non-missing items (scored pro rata and total score rounded to the nearest integer). Questionnaires with two or more missing values were not scored. Scores of 5, 10 and 15 represent cut-off points for mild, moderate and severe anxiety, respectively.

### **Loneliness**

The UCLA-3 is a three-item instrument asking the participant to indicate how often they feel they lack companionship, feel left out or feel isolated. The scoring of the three items is as follows: 1 = 'hardly ever or never', 2 = 'some of the time' and 3 = 'often'. A summary score is calculated by summing the three item scores, where none of the item responses are missing.

In addition to the UCLA-3 instrument, loneliness was also measured using the ELSA single-item direct loneliness question at each of the follow-up time points, which asks the participant how often they feel lonely with responses: 1 = 'Never', 2 = 'Hardly Ever', 3 = 'Occasionally', 4 = 'Some of the time' and 5 = 'Often or Always'.

### **Falls**

Incidence of falls was measured as the total number of falls experienced by the participant during the 12-month follow-up period. The baseline and 3-, 6- and 12-month follow-up questionnaires asked whether the participant fell in the 3 (or 6 on the 12-month questionnaire) months since the last questionnaire, and if so, how many times.



### Process evaluation

A qualitative process evaluation was informed by qualitative interviews with trial participants, trial decliners, trial yoga teachers and stakeholders, as well as by observations of standardisation training sessions and yoga classes (see [Chapter 5](#), Process Evaluation).

### Intervention fidelity

To ensure the trial yoga courses were delivered in accordance with the GYY teacher training programme and GYY trial guidelines, each yoga teacher underwent an observation of one of their trial classes by one of the yoga consultants (LB and JH). A checklist was completed for each observation [see Yoga teacher project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024) for the yoga teacher fidelity check assessment template.]. Intervention fidelity was also assessed via class observations and interviews with each yoga teacher as part of the process evaluation.

### Other data collected

Sociodemographic measures (age, gender, ethnicity, residential status, employment status, smoking status) and details of health conditions were collected at baseline via the screening and baseline questionnaires. An adapted Bayliss measure of illness burden<sup>31</sup> was calculated by summing the amount that each self-reported health condition limits a participant's daily activities from 1 = 'Not at all' to 5 = 'A lot'. Participant beliefs and preferences for the GYY programme and usual care were assessed in the questionnaires at baseline and 12 months. Adherence by participants to the supervised GYY classes was recorded by the yoga teachers using class attendance registers [see Yoga teacher project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)]. Participants also self-reported in the follow-up questionnaires any other supervised or self-managed yoga practice that they had done since the previous study time point. Adverse events were recorded (see 'Adverse event reporting and harms' below).

## Sample size

### Original

Walters and Brazier, in a review paper of the EQ-5D-3L,<sup>32</sup> found a difference of 0.074 (mean) or 0.081 (median) to be a minimum clinically important difference among a variety of patients, while McClure and colleagues found a minimum clinically important difference of 0.063 (mean) or 0.064 (median) for the EQ-5D-5L using simulated data.<sup>33</sup> To be conservative, we powered the trial to detect the lowest estimate (0.06), assuming a standard deviation (SD) of 0.20.<sup>19</sup> Accounting for loss to follow-up of 20%, we needed to randomise 586 participants to have 90% power with 5% significance (two sided).

Although this was an individually randomised trial, there was the potential for clustering within the intervention arm by yoga class. Assuming an intraclass correlation coefficient of 0.04 and average class size of 12 in the intervention arm, with the proposed sample size of 586, we would have still retained 84% power to detect the same magnitude of effect (*ceteris paribus*). In this calculation, we considered the level of clustering at the yoga class level rather than at the level of the yoga teacher, since we believed this to be the most influential level of clustering. Accounting for potential clustering within the intervention arm only leads to small reductions in power, which could potentially be recovered in the analysis of the repeated measures, adjusting for baseline value, which was not accounted for in the original sample size calculation.

### Revised

Trial recruitment commenced prior to the COVID-19 pandemic. The onset of the pandemic caused challenges to recruitment. Considering these challenges, in October 2021, the sample size calculation was revisited. At this point, 454 participants had been randomised (240 intervention; 214 usual care) and the trial team were considering whether to apply for a funded 7-month extension to give extra time to pursue the original recruitment target of 586.

The original calculation did not account for the correlation of outcome with baseline, as this provides the most conservative target. At the start of the trial, there was little to base the estimate of the correlation on, and assuming a correlation that is larger than is eventually observed can lead to an underpowered trial. However, in practice, it was always planned to include the baseline EQ-5D-5L utility index score as a covariate in the primary analysis model, and this provides gains in power.

Assuming all other parameters remained the same, with 454 participants, we would have had 81% power to detect the same effect size. With a higher rate of follow-up that was consistent with what had been observed up until this time (90%), this would have increased the power to 86%. Accounting for a higher rate of follow-up and also correlation of the outcome with baseline of at least 0.35, then we would have retained 90% power. An interim calculation of the correlation between baseline and 12-month EQ-5D-5L utility index score based on data received and processed up to 15 October 2021 ( $n = 86$  observations) was 0.67 (95% CI 0.54 to 0.77), and at that time we had observed a response rate of 90% at 3 months. Therefore, we were confident that we would be able to detect a clinically important difference with close to or greater than 90% power with 454 participants randomised. Hence, it was agreed to cease recruitment at 454 participants and not apply for an extension.

### Randomisation

Participants were randomised via a central, computer-based randomisation system designed and managed by York Trials Unit (YTU), University of York. The randomisation was stratified and used varying block sizes and allocation ratios as follows.

When enough patients (ideally 20–30) had provided baseline data and confirmed their availability for a specific GYY course, they were randomised collectively as a 'batch' by a member of the research team using the randomisation system. The patients were allocated to either the intervention or usual care group in a ratio that was variable to ensure that each GYY course was full to begin with (12–15 participants randomised to the intervention group for each face-to-face course and 12 participants for each online course, with the remaining participants allocated to usual care). We targeted an overall allocation ratio of 1 : 1.

Each batch of randomisations was then considered as a 'site' (and it is this level of 'site' that is included as a random effect in the statistical analysis as detailed later in this chapter), but each site could contain patients from more than one GP practice. One yoga teacher led each GYY course, but some teachers led more than one course.

Since a group of participants were randomised simultaneously, as opposed to participants being randomised one by one, the allocation sequence could not be predicted in advance.

Participants were notified of their group allocation via telephone, letter or e-mail and sent a participant diary to prospectively record their healthcare resource use [see Trial Participant project documents, <https://fundingawards.nihr.ac.uk/award/17/94/36> (accessed May 2024)] so that they had this to refer to when asked about their healthcare resource use in the follow-up questionnaires. If randomised to the intervention group, they were sent details of the class that they should attend and the name of the yoga teacher.

Following randomisation, the research team at YTU provided the yoga teachers with the names and contact details of those who were going to be attending their classes. They required this information so that they could make appropriate arrangements to support intervention delivery. The randomisation of participants was timed to occur no more than 3 weeks before the start date of the course that had been agreed with the yoga teacher.

For randomisation of participants to the methodological substudies, see [Chapter 6](#).

## Blinding

Due to the nature of the intervention, participants and yoga teachers were not blinded to group allocation. Outcome measures were primarily self-reported. Members of the research team collecting the outcome data over the telephone took reasonable steps to ensure that they remained blinded to group allocation.

Although GPs were informed about their patients' participation in the study, they were not informed about allocation status to reduce the risk of inducing GP behaviour change based on this knowledge. The wider health and social care team were not informed about a participant's study participation or allocation status.

There were no specific emergency unblinding procedures. A participant's group allocation could have been revealed to their GP in response to an AE, but there were no instances where this was necessary.

## Internal pilot phase

The original overall recruitment target was 586 participants across at least 12 sites over a total of 24 months. Mailouts at different GP practices were staggered. The period covering the first eight sites/courses formed the internal pilot phase.

Internal pilot objectives:

- to assess whether the provision and acceptability of the intervention met the pre-defined progression criteria thresholds, via the proportion of participants receiving their first yoga session within 3 weeks of randomisation and retention of intervention participants, respectively
- to assess whether recruitment and 6-month follow-up rates met the pre-defined progression criteria thresholds, measured by recruitment and EQ-5D-5L completion data.

The progression criteria assessed the level of recruitment for each site, follow-up rates and provision and acceptability of the intervention and informed study continuation beyond the internal pilot phase. The progression criteria were assessed using a traffic light system of green (go), amber (review) and red (stop), as follows:

- intervention provision (assessed 4 months after start of pilot intervention period):
  - green: 3–4 sites offering their first group yoga session within 3 weeks of participant randomisation
  - amber: 1–2 sites offering their first group yoga session within 3 weeks of participant randomisation
  - red: 0 sites offering their first group yoga session within 3 weeks of participant randomisation.
- intervention acceptability (assessed 4 months after start of pilot intervention period):
  - green: ≥ 80% retention of intervention participants
  - amber: < 80% but ≥ 65% retention of intervention participants
  - red: < 65% retention of intervention participants.
- recruitment (assessed at 6 months after start of internal pilot recruitment):
  - green: 3–4 sites recruited ≥ 20 patients each within 4 months (based on number of participants needed to allow randomisation and formation of a yoga class)
  - amber: 1–2 sites recruited ≥ 20 patients each within 4 months
  - red: 0 sites recruited ≥ 20 patients each within 4 months.



- six-month follow-up (assessed at 8 months after start of pilot intervention period):
  - green:  $\geq 80\%$  completion of the EQ-5D-5L
  - amber:  $< 80\%$  but  $\geq 65\%$  completion of the EQ-5D-5L
  - red:  $< 65\%$  completion of the EQ-5D-5L.

If any criteria were graded as amber, a rescue plan was to be considered outlining steps to be taken to improve intervention provision, recruitment, retention and/or follow-up (as appropriate) and approved by the Trial Steering Committee (TSC) before submission to the funding body (NIHR HTA). If all the progression criteria were failed (red), then the internal pilot would not have progressed to the main phase of the study. If the TSC deemed that the progression criteria were met to a satisfactory standard by the end of the internal pilot, then the study was to continue, and outcome data from participants in the internal pilot were included in the main study analysis.

### Statistical methods

Analyses were conducted once at the end of the trial using Stata version 17.<sup>34</sup> For all outcomes, the analysis population set included all randomised participants with data available for that outcome, and participants were analysed in the group to which they were randomised, irrespective of deviations based on non-compliance, under the principles of intention to treat (ITT). Statistical tests were two sided at the 5% significance level, and two-tailed 95% CIs and *p*-values were used.

### Recruitment

The flow of participants through the trial is detailed in three CONSORT diagrams depicting (1) recruitment of participants in the two pilot phase waves; (2) recruitment of participants in the two main phase waves; and (3) the retention and follow-up of participants post randomisation. The total number of participants approached and randomised are reported, with reasons for non-participation (ineligibility or non-consenting) provided where available.

### Baseline characteristics of randomised and analysed participants

All participant baseline data are summarised descriptively by the trial arm, both as randomised and as included in the primary analysis. No formal statistical comparisons were undertaken on baseline data between the arms. Continuous measures are reported as means and SDs (and/or median, interquartile range and range), and categorical data are reported as counts and percentages.

### Withdrawals and follow-up

Response rates to the participant questionnaires at 3, 6 and 12 months are presented by the trial arm, and overall, as number expected (i.e. not withdrawn before the time point), they are presented as a percentage of number randomised, number returned (% of expected, % of randomised) and median (interquartile range) time to completion in days from due date. Mode of completion (postal or over the telephone) is summarised for the 6-month time point. Reasons for non-completion are provided where known. Type and timing of withdrawals are presented overall and by the randomised group.

All outcome data are summarised descriptively by the randomised group and overall at each time point.

### Interim pilot phase

Descriptive statistics only were used to evaluate the progression criteria for the internal pilot sites. Data from participants in the internal pilot were included in the main study analysis.

### Primary outcome analysis

The EQ-5D-5L utility index score was modelled using a linear mixed-effects model to incorporate the outcome at all post-randomisation time points. The model adjusted for baseline EQ-5D-5L utility index score, time point, trial arm and trial arm by time interaction as fixed effects. Participant and site

were included as random effects, to account for the repeated measures of scores by participants over time, and the clustering of participants within sites. The different covariance structures for repeated measurements that are available as part of the analysis software were applied to the model. The most appropriate pattern was used for the final model based on diagnostics, including Akaike's information criterion (smaller values are preferred). The treatment effect in the form of the adjusted MD in EQ-5D-5L utility index score was extracted with its 95% CI and *p*-value for each time point and overall. The overall difference between the two groups over the 12 months from randomisation was the primary end point, but differences at each time point were extracted for secondary investigations aimed at determining the potential pattern of improvement. Model coefficients for the covariates with 95% CIs are also presented to aid understanding of the fitted model. Participants were only included in the model if they had full data for the baseline covariates and outcome data for at least one post-randomisation time point; the model assumes any missing data were missing at random (MAR). Model assumptions were checked as follows: the normality of the standardised residuals was checked using Q-Q plot, and homoscedasticity was assessed by means of a scatter plot of the standardised residuals against fitted values.

### **Sensitivity analyses**

#### **Adjusting for other covariates**

The primary analysis was repeated, but age, gender and adapted Bayliss score were additionally adjusted for as fixed effects.

#### **Clustering by yoga teacher**

Some of the yoga teachers taught more than one course of GYY as part of the trial. Analyses to account for possible clustering by yoga teachers were also undertaken by including the intended yoga teacher as a random effect instead of site.

#### **Compliance with random allocation and treatment received**

The number of GYY classes attended by participants is summarised. Attendance at GYY classes is presented as the percentage of participants who attended each week, overall and by recruitment wave. The number of participants who attended all 12 of their sessions, at least 9 sessions and at least 6 (including 3 of the first 6) is reported as measures of intervention adherence. We also report the amount of non-GYY yoga conducted by the two groups.

Complier-average causal effect (CACE) analyses for the primary outcome were undertaken to explore the impact of non-compliance on treatment effect estimates. Three analyses were conducted. The first considered participants who are fully compliant, defined as attendance at least three of the first six sessions and at least three other sessions. The second CACE analysis defined compliance as attendance of one yoga session or more (i.e. any compliance), which included participants who were fully and partially compliant. The final CACE analysis considered the number of sessions attended in its continuous form. Two-stage least squares instrumental variable (IV) regression for the EQ-5D-5L at 12 months was used, with randomised group as the IV and robust standard errors to account for clustering within site. The CACE analysis adjusted for gender (in the first stage) since gender was likely to be associated with attendance.

#### **Missing data**

The extent of missingness for the primary outcome was investigated and reported. The mixed-effects model incorporated data collected at all post-randomisation time points, and any missing outcome data were assumed to be MAR. We explored patterns of missingness and considered undertaking a sensitivity analysis (SA) to assess departures from the MAR assumption using a pattern mixture model; however this was deemed unnecessary due to the low amount of attrition in the primary analysis model (< 10%). The attrition rate here refers to those with only baseline data.

### *Subgroup analysis*

#### **Intended mode of delivery of Gentle Years Yoga**

Participants in the first wave of recruitment were randomised in October 2019, and for those allocated to the yoga arm, the classes took place between October 2019 and February 2020. This was before any restrictions were imposed on our daily lives as a result of the COVID-19 pandemic and, therefore, the classes were delivered face to face, as initially intended. However, recruitment was paused in mid-March 2020, and once recruitment recommenced, it was decided that yoga classes would be delivered online. Therefore, GYY participants who were randomised between September 2020 and June 2021 had their classes delivered online. Beyond this, up to the end of recruitment in October 2021, some of the sites delivered online and the others face to face.

A subgroup analysis was conducted in which the primary analysis was repeated including, as a fixed effect, an indicator for whether participants were randomised in a site intended for online GYY delivery (as opposed to face to face) and an interaction term between trial arm and intended mode of delivery.

#### *Secondary outcome analyses*

Each of the following secondary outcomes was analysed using the same methods as described for the primary outcome, with baseline EQ-5D-5L utility index score swapped as a covariate for baseline value of the outcome:

- EQ-5D VAS
- GAD-7 score
- PHQ-8 score
- T-scores from each of the seven subscales of the PROMIS-29 v2.1, the PCS and MCS and the global item score
- UCLA-3 score
- ELSA single-item direct loneliness question.

The incidence of falls during the 12-month follow-up period was analysed by modelling the number of participant falls over the 12-month follow-up period with a mixed-effect negative binomial regression model, adjusting for the number of falls in the 3 months prior to baseline, and site as a random effect. The model includes an exposure variable for the number of months for which the participant provided falls data. The point estimate for the treatment effect in the form of an incidence rate ratio and its associated 95% CI and *p*-value is provided.

#### *Adverse events*

Adverse events for this trial are summarised descriptively, stratified by whether they are classified as serious or non-serious. Only events deemed at least possibly related to the study are reported. The number, type, outcome, action and relatedness of the events are summarised.

### **Data collection and management**

Participant-completed baseline and follow-up questionnaires were sent from and returned to YTU. A central database at YTU was used to prompt the sending out and return of follow-up questionnaires. Where necessary, participants provided the questionnaire data via a telephone call with a researcher at YTU (e.g. when the research team needed to work from home during the COVID-19 pandemic and thus were unavailable to send and receive postal questionnaires or when participants required assistance). In addition to the data collected at the follow-up time points, some data were collected on an ongoing basis (such as AEs and changes in participant status when a participant requested to be withdrawn). The data collected in this trial, and the timing of its collection, are given in [Table 2](#).

TABLE 2 Schedule of assessments

Assessment	Follow-up time point				
	Screening/ baseline	3 months	6 months	12 months	Over 12 months
Eligibility and consent					
Eligibility	X				
Consent	X				
Background and follow-up data					
Demographic data	X				
Medical history	X				
Patient expectations and preferences	X				
EQ-5D-5L	X	X	X	X	
PROMIS-29 v2.1					
GAD-7	X	X	X	X	
PHQ-8	X	X	X	X	
UCLA	X	X	X	X	
Falls	X	X	X	X	
Healthcare resource use	X	X	X	X	
Prescription use <sup>a</sup>	X				X
Ongoing data collection					
AE reporting					X
Change of participant status					X

<sup>a</sup> Collected for 3 months prior to baseline and for 12 months post baseline for a subset of participants in both groups.

Screening and baseline data were collected via postal questionnaires in pilot phase wave 1; however due to the COVID-19 pandemic, data were collected via a telephone call and entered electronically into a Qualtrics® (online survey software, Qualtrics, Provo, Utah, USA) survey by the researcher in pilot phase wave 2 and main phase waves 1 and 2.

Follow-up data were collected via postal questionnaire in all phases of the trial, except for a short period due to COVID-19 restrictions for pilot phase wave 1 (6-month follow-up), where it was collected via a telephone call and entered electronically into a Qualtrics survey by the researcher. If a follow-up questionnaire was not returned to YTU 14 days after it was posted, a reminder questionnaire was sent to the participant. If a follow-up questionnaire was not returned to YTU 14 days after a reminder questionnaire was posted, a researcher contacted the participant via telephone call if the participant had consented to this means of communication. If during the telephone call the participant advised that they were unable to complete the postal questionnaire but agreed to provide the information over the telephone, primary outcome data collection was prioritised and then any other outcome data that the participant was willing to provide was collected. If a questionnaire was returned to YTU incomplete or with errors, a researcher contacted the participant via telephone for clarification or completion of missing data if the participant had consented to this means of communication. Telephone data in this case were entered onto a paper questionnaire by the researcher.

Essential trial documentation, which individually and collectively permits evaluation of the conduct of a clinical trial and the quality of the data produced, was kept in an electronic Trial Master File on a secure

file server at the University of York with access restricted to named YTU trial staff. This documentation will be retained for a minimum of 20 years in electronic format in accordance with Good Clinical Practice. Data were handled in accordance with the Data Protection Act 2018, General Data Protection Regulation (GDPR) legislation, the latest Directive on Good Clinical Practice and local policy. All paper data records are stored in a secure facility at the University of York in locked filing cabinets in a locked room. The key to the cupboard is held by the data archivist. The paper data records will be archived at an approved off-site location. All electronic data records are stored on a secure file server at the University of York, with access restricted to YTU trial staff. Data entered electronically onto a database will be stored on a private network protected by a firewall at the University of York. Access to the database is restricted to YTU trial staff by login and password. The trial database will be securely archived for a minimum of 10 years on the YTU computer network. Access to the archived data will be restricted to YTU trial staff and named individuals but will be retrievable at the request of the sponsor or investigator. All data will be stored for a minimum of 10 years, in line with the University of York's policy.

### ***Design and processing of participant questionnaires***

The participant postal questionnaires were designed using TeleForm software (version 11.2; Cardiff Software, Cambridge, UK). Specification questionnaires were populated with variable names and appropriate scoring. To maximise data quality, when questionnaires were returned to YTU (by post), key variables required for the statistical analysis were reviewed for completion and accuracy by a data administrator who resolved any queries by contacting the participant via telephone. After this initial check, all questionnaires were prepared for scanning by a data administrator using the TeleForm software. When a form would not scan, the data were manually entered. When a form was scanned, the data were then verified depending on what TeleForm identified as requiring correction. The verified data were then temporarily held in the Download Database and available for second checking. This involved each hard copy of all forms being compared against the entry stored in the Download Database and correcting the electronic data as necessary. All data were scanned, downloaded and second-checked in the Validate Database. The automated data validation was undertaken by the Data Manager who applied predetermined rules to check for agreed variables and whether the data were recorded correctly.

The participant electronic questionnaires were designed using Qualtrics. To maximise data quality, response requirements and validations were used for key variables required for the statistical analysis (e.g. to alert when a question had been missed or entered in the incorrect format). The data from each form submitted was regularly downloaded in batches from Qualtrics and imported into a Structured Query Language (SQL) database. Images of each form submitted were also downloaded in PDF format from Qualtrics. The data verification process was similar to that for the postal questionnaires in that the data were checked against the images and a code was run to apply the data rules and identify any anomalies.

Data that had been validated were made available, upon a formal request, to the trial statistician and health economist.

### ***Plans to promote participant retention and complete follow-up***

The following processes were implemented to promote participant retention and follow-up:

- Yoga teachers were asked to encourage regular class attendance and to contact participants who missed two consecutive classes without prior notification.
- Participants who withdrew from the yoga programme were asked if they would be happy to remain in the trial to provide follow-up data.
- Participants received a £5 shopping voucher or £5 cash with every follow-up questionnaire as a thank you for their continuing participation in the trial.
- A text message was sent to participants who consented to receive this form of communication 1–7 days before each follow-up questionnaire was sent, saying that they would soon receive a

follow-up questionnaire and to complete and return it as soon as possible. The text message also acted as a prompt for the participant to inform the research team if they had moved address.

- A reminder letter was sent if the follow-up questionnaire had not been returned within 14–21 days.
- A researcher called the participant to complete the follow-up questionnaire over the telephone if it had not been returned within 28 days.
- If a questionnaire was returned incomplete or with errors, a researcher called the participant for clarification or completion of missing or invalid data.
- Participants were advised that they were able to call a member of the research team if they required assistance with completing a questionnaire.
- A newsletter containing information about trial progress and any relevant updates was sent out to participants via post or e-mail every 3–6 months.
- After completing the final follow-up questionnaire (12 months), all usual care participants received details of GYY or other suitable yoga classes that they could join on a self-pay basis. That they would do so was specified in the PIS.
- After randomisation to the main trial, participants allocated to the usual care group were randomised again to receive either the offer of a one-off yoga class, which would take place after the final follow-up or no offer (see substudy 4 in [Chapter 6](#): Does offering a free yoga session to control participants after the 12-month follow-up assessment enhance retention and reduce contamination?).
- After randomisation to the main trial, participants allocated to the intervention group were randomised again to receive either a pen with their 3-month follow-up questionnaire or no pen (see substudy 3 in [Chapter 6](#): Does sending a pen with a follow-up questionnaire enhance return rates?).

### **Participant withdrawal**

Participants could withdraw from the trial at any point during the study. If a participant indicated that they wanted to withdraw from the study, they were asked whether they wished to withdraw from the intervention only (i.e. withdraw from the yoga programme if allocated to the intervention group) or withdraw fully from the study. When withdrawal was only from the intervention, follow-up data continued to be collected. Data provided by participants who fully withdrew were retained for analysis up to the point at which they withdrew. A member of the research team completed a change of status form for any participant who changed status (died, withdrew, lost to follow-up) during the trial.

### **Trial completion and exit**

Participants completed the trial once they had returned a complete 12-month follow-up questionnaire. Participants exited the trial prematurely if they had fully withdrawn, were lost to follow-up or died.

## **Adverse event reporting and harms**

Adverse events were defined as any untoward medical occurrence in a trial participant which did not necessarily have a causal relationship with the treatment. SAEs were defined as any untoward medical occurrence that:

- resulted in death
- was life-threatening
- required inpatient hospitalisation or prolongation of existing inpatient hospitalisation
- resulted in persistent or significant disability or incapacity
- was a congenital anomaly or birth defect
- was any other important medical event which, although not included in the above, may have required medical or surgical intervention to prevent one of the outcomes listed. A planned operation or hospitalisation for a pre-existing condition was not considered an SAE.



Adverse events that were possibly related to yoga practice or study participation during the 12 months after randomisation were recorded by trial coordinators or trial support officers, followed by an assessment of seriousness, relatedness and expectedness by the Chief Investigator and a clinician co-investigator (DY). We also recorded all deaths and all falls that resulted in hospitalisation, regardless of causality. Yoga teachers and participants were asked to notify the trial office of any AEs they thought were related to yoga practice or study participation. Researchers reported all falls resulting in hospitalisation via a telephone check with the participant if they indicated that they had experienced a fall on one of the follow-up questionnaires.

All (S)AEs were reported to the Trial Management Group, TSC and Sponsor. SAEs that were related to the research and unexpected were reported to the Research Ethics Committee (REC).

### **Regulatory approvals and monitoring**

Ethics approval for the study was obtained from the North East – York REC (REC reference number 19/NE/0072) on 24 April 2019. The study was also approved by the Health Research Authority on 24 April 2019. Amendments to the protocol and study documentation were submitted to the REC and the Health Research Authority for their approval as required during the study. The trial sponsor was Northumbria University.

#### ***Trial registration***

The study was registered to the ISRCTN Registry on 18 March 2019 (reference number ISRCTN13567538).

#### ***Summary of changes to the protocol***

Amendments to the trial protocol are summarised in [Appendix 1](#).

#### ***Trial Management Group***

The day-to-day management of the trial was overseen by the Trial Management Group, which met every 2 months. This group comprised the Chief Investigator, co-applicant co-investigators, Trial Manager, Trial Coordinators, Trial Support Officers, Data Manager, Data Administrator, Statistician, Health Economist, Process Evaluation Researcher, Clinician, Patient Representative and a BWY Representative (see [Acknowledgements](#) section for member details).

#### ***Trial Steering Committee***

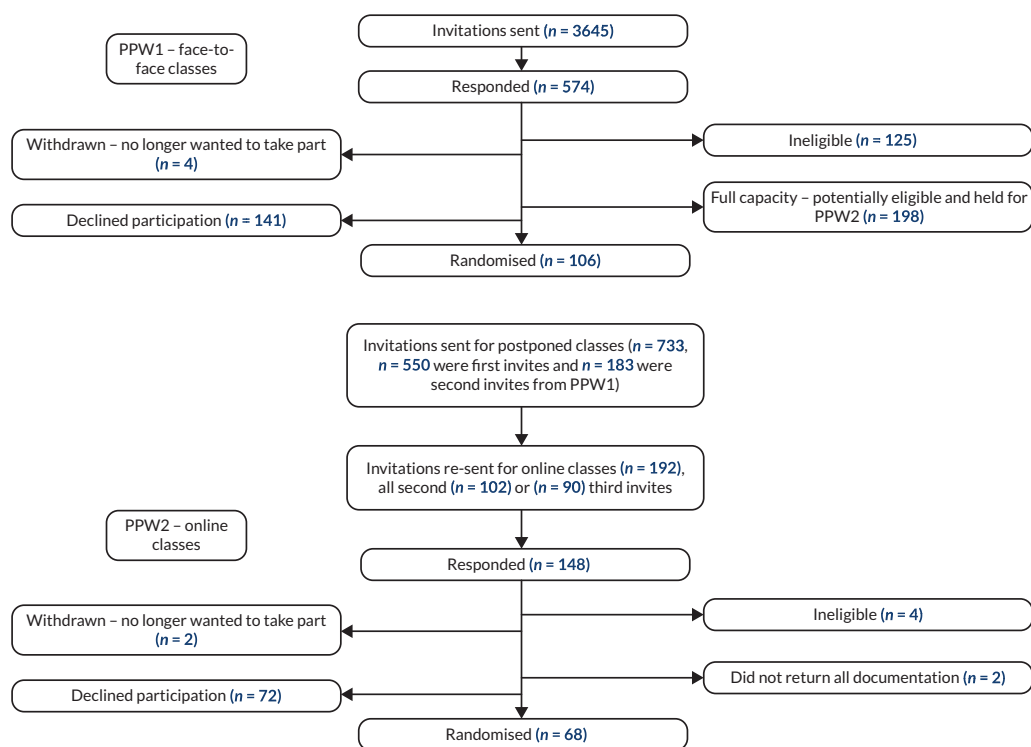
A TSC was appointed by the funding body to provide overall supervision for the trial and to advise on its continuation. It met every 6–9 months and included an independent chairperson and members as well as representatives of the Trial Management Group (the Chief Investigator, a Statistician and a Trial Coordinator). Membership is listed in the [Acknowledgements](#) section.

The TSC agreed to additionally take on the role of a data monitoring committee. A separate Data Monitoring Committee was considered unnecessary due to the low risk and open nature of the trial. Therefore, the TSC reviewed and discussed data pertaining to both participant safety and trial progress.

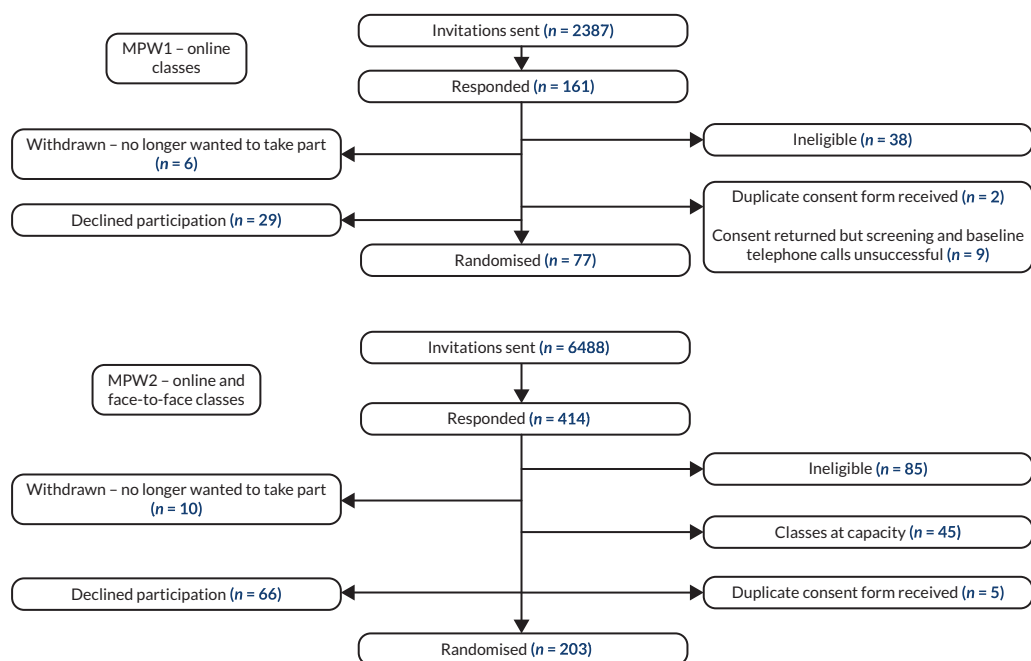
## Chapter 3 Clinical effectiveness results

### Participant flow

The flow of participants is illustrated in CONSORT flow diagrams in [Figures 3–5](#).

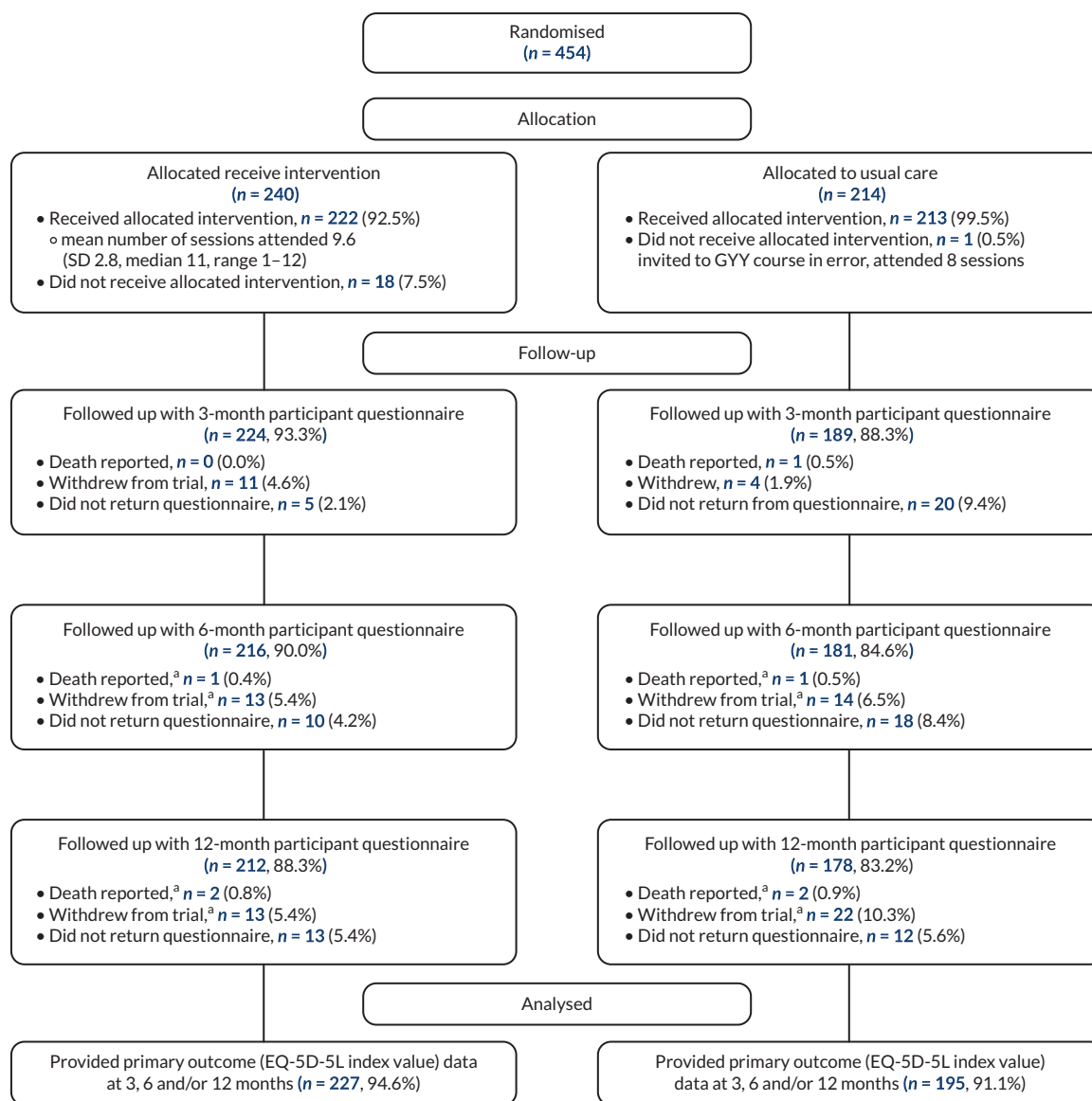


**FIGURE 3** Consolidated Standards of Reporting Trials diagram: recruitment to the GYY trial in the pilot phase waves.



**FIGURE 4** Consolidated Standards of Reporting Trials diagram: recruitment to the GYY trial in the main phase waves.





**FIGURE 5** Consolidated Standards of Reporting Trials diagram: randomisation and follow-up in the GYY trial. <sup>a</sup> Withdrawals and deaths over time are cumulative.

Participants were recruited from 15 GP practices across 6 CRNs: Yorkshire and Humber (2 sites in Harrogate and 1 in Hull); North West Coast (2 sites in Wirral); Kent, Surrey and Sussex (1 site in Kent); Health and Care Research Wales (1 site in Newport); West of England (4 sites in Bristol) and Thames Valley and South Midlands (2 sites in Oxford, 1 in Wantage and 1 in Banbury).

These 15 practices had a total estimated list size of 320,512, of which 13,070 (4.1%) were sent an invitation pack between July 2019 and August 2021. A number of participants invited in the first wave of recruitment, who were potentially eligible but were not randomised in the first wave due to sufficient numbers being reached to fill the GYY courses, were reinvited and rescreened in the second wave of recruitment (n = 285). A response to the invitation pack was received from 1,297 participants (9.9% of 13,070). A quarter declined participation (n = 308, 23.7%) (see [Table 3](#)) or withdrew their interest after initially providing consent (n = 22, 1.7%) and 252 (19.4%) were ineligible (see [Table 4](#)). The remaining 261 (20.1%) were not randomised for other reasons, most commonly that sufficient participants had been recruited to fill the GYY courses (n = 243).

**TABLE 3** Reason for declining participation in the GYY trial

Reason for declining participation (n = 326, participants could provide more than one reason)	n	%
No reason	163	50
Health condition/poor health	27	8.3
Does not want to do online yoga	27	8.3
Does not have the technology to do online yoga	18	5.5
Does not think it will benefit them as already active	18	5.5
No longer wants to take part – no specific reason	16	4.9
Other commitments	13	4
Already takes part in an exercise activity	9	2.8
Already practices yoga	8	2.5
Less than two long-term health conditions	8	2.5
No access to a computer	5	1.5
Feels they are too old	3	0.9
Lack of transport	2	0.6
Unable to travel to face-to-face classes	2	0.6
Unable to complete online consent form	1	0.3
Not computer literate	1	0.3
Not enough space at home	1	0.3
Does not want to do face-to-face yoga	1	0.3
Does not want to complete more forms	1	0.3
Does not like study title	1	0.3
Not active enough for them	1	0.3

**TABLE 4** Reason for ineligibility in the GYY trial

Reason for ineligibility (n = 281, more than one reason could apply)	n	%
No multimorbidity	168	59.8
Do too much yoga already	28	10
Unable to attend	24	8.5
Did not return screening or consent form	22	7.8
Unreturned baseline	19	6.8
No suitable electronic device	6	2.1
Not able to use internet	3	1.1
Insufficient space to practice yoga	3	1.1
Severe mental health problem	2	0.7
Enrolled in an unsuitable study	2	0.7

continued

TABLE 4 Reason for ineligibility in the GYY trial (continued)

Reason for ineligibility (n = 281, more than one reason could apply)	n	%
No access to internet	2	0.7
Belongs to same household as another participant	2	0.7
Not over 65	0	0
Not community dwelling	0	0
Learning disability	0	0
Does not speak English	0	0
No sturdy chair	0	0

In total, between 18 October 2019 and 5 October 2021, 454 eligible and consenting participants were randomised: 240 to the intervention and 214 to usual care. Participants were randomised across 19 sites (mean 23.9 per site, SD 4.9, median 24, range 16–35). Seven sites delivered face-to-face GYY courses, and 12 were online. Twelve participants were randomised to intervention for every online course and either 12 or 15 (median 15) for every face-to-face course (see [Report Supplementary Material 2](#)).

### Baseline characteristics of randomised participants

The mean age of randomised participants was 73.5 years (range 65–99); 60.6% were female, and participants had a median of three chronic conditions (see [Table 5](#)). The most commonly reported conditions were cardiovascular diseases ( $n = 307$  participants, 67.6%), which included participants who reported at least one of coronary heart disease, hypertension, heart failure or peripheral arterial disease (of which hypertension was the most prevalent) and arthritis (osteo or rheumatoid arthritis,  $n = 242$ , 53.3%) (see [Table 6](#)). The intervention and usual care groups were reasonably comparable in terms of baseline characteristics, except that there was a slightly higher proportion of females in the intervention group (64.2%) than usual care (56.5%).

Participants were asked at baseline about their expectations and preferences in relation to the health care offered in the GYY trial (see [Table 7](#)). Half of the respondents ( $n = 235$ , 52.3%) thought that usual care would be fairly or very effective at improving their quality of life, and a slightly higher proportion ( $n = 277$ , 61.0%) thought the GYY programme would be fairly or very effective at improving their quality of life. Given the choice, three-quarters ( $n = 339$ , 74.7%) said they would prefer to be allocated to the intervention group rather than usual care alone. Most of the rest had no preference ( $n = 103$ , 22.7%), and only a small number preferred usual care ( $n = 12$ , 2.6%).

The baseline values of the primary and secondary outcome measures are summarised in [Table 8](#) and are reasonably well-balanced between groups.

### Withdrawals and follow-up

Participant follow-up was completed in October 2022.

In total, we became aware of seven deaths [2 (0.8%) in the intervention group and 5 (2.3%) in the usual care group]. Six of these occurred within the 12 months from randomisation, and one just beyond this. For one death that occurred within 12 months, we only became aware of the event after their 12-month questionnaire was sent out.

A further 36 participants [14 (5.9%) intervention participants and 22 (10.5%) usual care participants] withdrew from follow-up data collection during the trial; 15 before month 3, 12 between months 3 and 6, 8 between months 6 and 12 and 1 just beyond 12 months (participants contacted the research team upon receipt of their 12-month postal questionnaire to say they were unable to complete the questionnaire due to ill health).

TABLE 5 Baseline characteristics of randomised participants by group

	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
Age, years	73.4 (6.0)	73.5 (6.4)	73.5 (6.2)
Gender, n (%)			
Male	86 (35.8)	93 (43.5)	179 (39.4)
Female	154 (64.2)	121 (56.5)	275 (60.6)
Ethnic group, n (%)			
White British	230 (95.8)	205 (95.8)	435 (95.8)
White Irish	2 (0.8)	0 (0.0)	2 (0.4)
White other	5 (2.1)	3 (1.4)	8 (1.8)
Black Caribbean	1 (0.4)	0 (0.0)	1 (0.2)
Asian Indian	1 (0.4)	2 (0.9)	3 (0.7)
Asian Pakistani	0 (0.0)	1 (0.5)	1 (0.2)
White and Asian	0 (0.0)	1 (0.5)	1 (0.2)
Other mixed	1 (0.4)	0 (0.0)	1 (0.2)
Missing	0 (0.0)	2 (0.9)	2 (0.4)
Employment status, n (%)			
Employed	10 (4.2)	8 (3.7)	18 (4.0)
Retired	219 (91.3)	196 (91.6)	415 (91.4)
Other	11 (4.6)	10 (4.7)	21 (4.6)
IMD decile (1 = most deprived to 10 = least deprived)	7.6 (2.6)	7.5 (2.7)	7.5 (2.7)
Smoking status, n (%)			
Yes	5 (2.1)	5 (2.3)	10 (2.2)
No, never smoked	115 (47.9)	109 (50.9)	224 (49.3)
No, used to smoke	120 (50.0)	100 (46.7)	220 (48.5)
No. of conditions, <sup>a</sup> n (%)			
2	86 (35.8)	85 (39.7)	171 (37.7)
3	86 (35.8)	71 (33.2)	157 (34.6)
4	41 (17.1)	34 (15.9)	75 (16.5)
5	16 (6.7)	14 (6.5)	30 (6.6)
6	7 (2.9)	7 (3.3)	14 (3.1)
7	1 (0.4)	3 (1.4)	4 (0.9)
8	2 (0.8)	0 (0.0)	2 (0.4)
9	1 (0.4)	0 (0.0)	1 (0.2)
No. of conditions, median (minimum, maximum)	3.0 (2.0, 9.0)	3.0 (2.0, 7.0)	3.0 (2.0, 9.0)
Bayliss <sup>b</sup>	9.6 (6.5)	9.7 (7.6)	9.7 (7.1)

a Conditions grouped as per the inclusion criteria.

b Higher score indicates worse outcome.

**Note**

Data are mean (SD) unless otherwise stated.

**TABLE 6** Self-reported health conditions at baseline by randomised group

Health condition	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
Cardiovascular disease	162 (67.5)	145 (67.8)	307 (67.6)
<i>Hypertension</i>	132 (55.0)	119 (55.6)	251 (55.3)
<i>Coronary heart disease, including angina, history of heart attack, bypass surgery or angioplasty</i>	32 (13.3)	36 (16.8)	68 (15.0)
<i>Heart failure</i>	15 (6.3)	8 (3.7)	23 (5.1)
<i>Peripheral artery disease</i>	22 (9.2)	19 (8.9)	41 (9.0)
Arthritis	135 (56.3)	107 (50.0)	242 (53.3)
<i>Osteoarthritis of the shoulder, hip or knee</i>	123 (51.2)	99 (46.3)	222 (48.9)
<i>Rheumatoid arthritis of the shoulder, hip or knee</i>	19 (7.9)	16 (7.5)	35 (7.7)
Sensory conditions	90 (37.5)	78 (36.4)	168 (37.0)
<i>Deafness or severe problem with hearing</i>	74 (30.8)	63 (29.4)	137 (30.2)
<i>Blindness or severe problem with vision</i>	30 (12.5)	23 (10.7)	53 (11.7)
Depression or anxiety	63 (26.3)	47 (22.0)	110 (24.2)
<i>Anxiety</i>	45 (18.8)	36 (16.8)	81 (17.8)
<i>Depression</i>	48 (20.0)	31 (14.5)	79 (17.4)
Asthma or COPD	62 (25.8)	47 (22.0)	109 (24.0)
<i>Asthma</i>	47 (19.6)	36 (16.8)	83 (18.3)
<i>COPD</i>	21 (8.8)	15 (7.0)	36 (7.9)
Bowel problems	55 (22.9)	36 (16.8)	91 (20.0)
Osteoporosis or osteopenia	38 (15.8)	41 (19.2)	79 (17.4)
Diabetes	36 (15.0)	37 (17.3)	73 (16.1)
Atrial fibrillation	37 (15.4)	35 (16.4)	72 (15.9)
Cancer (last 5 years)	35 (14.6)	35 (16.4)	70 (15.4)
Chronic kidney disease	14 (5.8)	15 (7.0)	29 (6.4)
Stroke (last 5 years)	6 (2.5)	10 (4.7)	16 (3.5)
Fibromyalgia	8 (3.3)	7 (3.3)	15 (3.3)
Epilepsy	1 (0.4)	5 (2.3)	6 (1.3)
Multiple sclerosis	3 (1.3)	2 (0.9)	5 (1.1)
Parkinson's disease	1 (0.4)	4 (1.9)	5 (1.1)
Dementia	2 (0.8)	1 (0.5)	3 (0.7)

COPD, chronic obstructive pulmonary disease.

**Notes**

Data are n, number (%).

Health conditions are presented as grouped according to the trial eligibility criteria and then broken down (text in italics) by individual condition.

**TABLE 7** Participant-reported expectations and preferences concerning the health care being offered in the GYY trial, collected at baseline, by randomised group

Expectations and preferences	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
<b>How effective do you think that usual care would be in improving your quality of life?</b>			
Very ineffective	17 (7.1)	6 (2.8)	23 (5.1)
Fairly ineffective	20 (8.3)	28 (13.1)	48 (10.6)
Can't decide	71 (29.6)	72 (33.6)	143 (31.5)
Fairly effective	80 (33.3)	65 (30.4)	145 (31.9)
Very effective	49 (20.4)	41 (19.2)	90 (19.8)
Missing	3 (1.3)	2 (0.9)	5 (1.1)
<b>How effective do you think the GYY programme would be in improving your quality of life?</b>			
Very ineffective	5 (2.1)	2 (0.9)	7 (1.5)
Fairly ineffective	7 (2.9)	4 (1.9)	11 (2.4)
Can't decide	81 (33.8)	78 (36.4)	159 (35.0)
Fairly effective	82 (34.2)	72 (33.6)	154 (33.9)
Very effective	65 (27.1)	58 (27.1)	123 (27.1)
<b>Given the choice, which study group would you prefer to be in?</b>			
Yoga and usual care	183 (76.3)	156 (72.9)	339 (74.7)
Usual care	5 (2.1)	7 (3.3)	12 (2.6)
No preference	52 (21.7)	51 (23.8)	103 (22.7)
<b>Note</b> Data are n, number (%).			

**TABLE 8** Values of outcome measures assessed at baseline by randomised group

Outcome scores at baseline	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
EQ-5D-5L utility index score <sup>a</sup>	0.742 (0.176)	0.736 (0.162)	0.739 (0.169)
EQ-5D-5L VAS <sup>a</sup>	75.0 (18.2)	73.4 (17.6)	74.3 (17.9)
PHQ-8 <sup>b</sup>	3.7 (3.9)	3.8 (4.3)	3.8 (4.1)
GAD-7 <sup>b</sup>	2.5 (3.4)	2.7 (3.6)	2.6 (3.5)
PROMIS-29 physical functioning <sup>a</sup>	46.7 (8.5)	46.3 (8.5)	46.5 (8.5)
PROMIS-29 anxiety <sup>b</sup>	46.9 (8.0)	48.1 (8.5)	47.5 (8.2)
PROMIS-29 depression <sup>b</sup>	46.4 (7.6)	46.8 (8.1)	46.6 (7.8)
PROMIS-29 fatigue <sup>b</sup>	47.4 (9.7)	48.7 (9.8)	48.0 (9.8)
PROMIS-29 sleep disturbances <sup>b</sup>	49.1 (9.5)	49.8 (9.6)	49.5 (9.6)
PROMIS-29 social roles <sup>a</sup>	54.7 (9.3)	54.1 (9.9)	54.4 (9.6)
PROMIS-29 pain interference <sup>b</sup>	53.3 (8.7)	53.6 (8.9)	53.5 (8.8)
PROMIS-29 pain intensity <sup>b</sup>	3.1 (2.5)	3.2 (2.4)	3.1 (2.4)

continued

**TABLE 8** Values of outcome measures assessed at baseline by randomised group (*continued*)

Outcome scores at baseline	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
PROMIS-29 mental health summary score <sup>a</sup>	52.9 (8.0)	52.0 (8.5)	52.5 (8.2)
PROMIS-29 physical health summary score <sup>a</sup>	47.6 (8.8)	47.1 (8.8)	47.4 (8.8)
UCLA-3 loneliness <sup>b</sup>	4.2 (1.7)	4.4 (1.9)	4.3 (1.8)
ELSA single-item direct loneliness question, <sup>b</sup> n (%)	2.2 (1.3)	2.3 (1.3)	2.2 (1.3)
<b>Fallen past 3 months, n (%)</b>			
Yes	61 (25.4)	49 (22.9)	110 (24.2)
No	179 (74.6)	164 (76.6)	343 (75.6)
Missing	0 (0.0)	1 (0.5)	1 (0.2)
<b>If yes, number of falls</b>	1.5 (0.8)	1.9 (2.8)	1.7 (1.9)
Median (minimum, maximum)	1 (1, 4)	1 (1, 20)	1 (1, 20)

a Higher score indicates better outcome.

b Lower score indicates better outcome.

**Note**

Data are mean (SD) unless otherwise stated.

The overall follow-up rates for the 454 randomised participants were 91.0% at month 3, 87.4% at month 6 and 85.9% at month 12 (see [Table 9](#)). At all time points, the return rate was slightly higher (by approximately 5 percentage points) in the intervention group than in the usual care group. Median time to completion was 7 days from the due date at both months 3 and 6 and 10 days at month 12 and was similar for the two groups at each time point.

At the 6-month time point, 70 (17.6%) of the questionnaires were completed over the phone with a researcher rather than on paper and returned by post. This was when COVID-19 restrictions prevented researchers from being in the office to facilitate the mailing and return of postal questionnaires.

## Internal pilot phase

The progression criteria for the internal pilot phase were assessed after the last participant recruited in the pilot sites was followed up at 6 months and was graded against the pre-defined traffic light style thresholds:

- Intervention provision

The pilot phase consisted of eight sites; these eight sites held the first intervention session of the GYY course between 13 and 20 days following participant randomisation. Therefore, the 'green' threshold was met for this criterion as at least three sites offered the first group yoga session within 3 weeks of participant recruitment.

- Intervention acceptability

Of the 108 participants randomised to the intervention in the pilot phase, an average of 76 (70.4%) attended each GYY session. Therefore, the 'amber' threshold was met for this criterion, as between 65% and 80% of intervention participants were retained in the programme.

**TABLE 9** Return rates for post-randomisation follow-up questionnaires

		Intervention (n = 240)	Usual care (n = 214)	Total (n = 454)
<b>Month 3</b>	Expected, <sup>a</sup> n (%)	229 (95.4)	209 (97.7)	438 (96.5)
	Death or withdrawal recorded before month 3, n (%)	11 (4.6)	5 (2.3)	16 (3.5)
	Returned, n (% expected, % randomised)	224 (97.8, 93.3)	189 (90.4, 88.3)	413 (94.3, 91.0)
	Days to completion, median (IQR)	8 (4–18)	7 (4–14)	7 (4–15)
<b>Month 6</b>	Expected, <sup>a</sup> n (%)	226 (94.2)	199 (93.0)	425 (93.6)
	Death or withdrawal recorded before month 6, n (%)	14 (5.8)	15 (7.0)	29 (6.4)
	Returned, n (% expected, % randomised)	216 (95.6, 90.0)	181 (91.0, 84.6)	397 (93.4, 87.4)
	Days to completion, median (IQR)	8 (4–17)	7 (4–14)	7 (4–16)
	Completed over phone with researcher, n (% returned) <sup>b</sup>	40 (18.5)	30 (16.5)	70 (17.6)
<b>Month 12</b>	Expected, <sup>a</sup> n (%)	225 (93.8)	190 (88.8)	415 (91.4)
	Death or withdrawal recorded before month 12, n (%)	15 (6.2)	24 (11.2)	39 (8.6)
	Returned, n (% expected, % randomised)	212 (94.2, 88.3)	178 (93.7, 83.2)	390 (94.0, 85.9)
	Days to completion, median (IQR)	9.5 (5–15)	10 (5–16)	10 (5–15)

IQR, interquartile range.

a Not withdrawn before time point.

b Only questionnaires from the 6-month follow-up were due during the time when postal questionnaires could not be sent out due to COVID-19 restrictions; the other time points were postal only.

- Recruitment

The eight sites each recruited between 16 and 28 participants. Therefore, the 'green' threshold was met for this criterion as at least three sites recruited  $\geq 20$  participants.

- Six-month follow-up

Overall, 148 (85.1%) of the 174 participants recruited during the pilot phase provided valid EQ-5D-5L data at the 6-month follow-up. Therefore, the 'green' threshold was met for this criterion as completion rates exceeded 80%.

Although one criterion was graded amber, the rest were green, and so the TSC was satisfied to recommend that the trial continue without the need for any major changes in recruitment or retention processes.



## Primary outcome (EQ-5D-5L utility index score) analysis

The EQ-5D-5L utility index score was assessed at baseline and at 3, 6 and 12 months post randomisation. The EQ-5D-5L utility index score is a value between 0 and 1, where a higher score indicates better health. The trial was powered to detect a difference of 0.06 (assuming a SD of 0.20).

### Raw scores

Summaries of the EQ-5D-5L utility index score by trial arm and time point are presented in [Table 10](#). At each time point, mean scores are slightly higher in the intervention arm than in the usual care arm. Overall, at baseline, the mean score was 0.739 (SD 0.169) and decreased over time to 0.707 (SD 0.214) at 12 months.

The correlation between baseline EQ-5D-5L utility index score and scores at the follow-up time points is: 3 months 0.72 (95% CI 0.67 to 0.77), 6 months 0.63 (95% CI 0.57 to 0.69) and 12 months 0.59 (95% CI 0.52 to 0.65).

### Baseline characteristics of participants included in primary analysis

The primary analysis included participants with a valid EQ-5D-5L utility index score at baseline and at least one post-randomisation time point ( $n = 422$ , 93.0%; intervention  $n = 227$ , 94.6%; usual care  $n = 195$ , 91.1%). The baseline characteristics of these participants are included in [Tables 11–14](#); these are very similar to the randomised population, which indicates that there is little evidence that loss to follow-up has introduced attrition or selection bias.

**TABLE 10** Summaries of raw EQ-5D-5L utility index score by trial arm and time point

Mean (SD), N	Intervention ( $n = 240$ )	Usual care ( $n = 214$ )	Total ( $n = 454$ )
Baseline	0.742 (0.176), 240	0.736 (0.162), 213	0.739 (0.169), 453
3 months	0.749 (0.168), 224	0.723 (0.201), 190	0.737 (0.184), 414
6 months	0.732 (0.207), 218	0.706 (0.219), 180	0.720 (0.212), 398
12 months	0.723 (0.210), 213	0.689 (0.219), 182	0.707 (0.214), 395

**TABLE 11** Baseline characteristics of randomised participants by group for those included in primary analysis

	Intervention ( $n = 227$ )	Usual care ( $n = 195$ )	Overall ( $n = 422$ )
Age, years	73.2 (5.9)	73.4 (6.2)	73.3 (6.0)
Gender, $n$ (%)			
Male	84 (37.0)	90 (46.2)	174 (41.2)
Female	143 (63.0)	105 (53.8)	248 (58.8)
Ethnic group, $n$ (%)			
White British	217 (95.6)	186 (95.4)	403 (95.5)
White Irish	2 (0.9)	0 (0.0)	2 (0.5)
White Other	5 (2.2)	3 (1.5)	8 (1.9)
Black Caribbean	1 (0.4)	0 (0.0)	1 (0.2)
Asian Indian	1 (0.4)	2 (1.0)	3 (0.7)
Asian Pakistani	0 (0.0)	1 (0.5)	1 (0.2)

**TABLE 11** Baseline characteristics of randomised participants by group for those included in primary analysis (continued)

	Intervention (n = 227)	Usual care (n = 195)	Overall (n = 422)
White and Asian	0 (0.0)	1 (0.5)	1 (0.2)
Other mixed	1 (0.4)	0 (0.0)	1 (0.2)
Missing	0 (0.0)	2 (1.0)	2 (0.5)
<b>Employment status, n (%)</b>			
Employed	10 (4.4)	7 (3.6)	17 (4.0)
Retired	208 (91.6)	178 (91.3)	386 (91.5)
Other	9 (4.0)	10 (5.1)	19 (4.5)
IMD decile (1 = most deprived to 10 = least deprived)	7.7 (2.6)	7.4 (2.7)	7.5 (2.7)
<b>Smoking status, n (%)</b>			
Yes	5 (2.2)	5 (2.6)	10 (2.4)
No, never smoked	109 (48.0)	103 (52.8)	212 (50.2)
No, used to smoke	113 (49.8)	87 (44.6)	200 (47.4)
<b>No. of conditions,<sup>a</sup> n (%)</b>			
2	82 (36.1)	76 (39.0)	158 (37.4)
3	84 (37.0)	66 (33.8)	150 (35.5)
4	36 (15.9)	31 (15.9)	67 (15.9)
5	15 (6.6)	13 (6.7)	28 (6.6)
6	7 (3.1)	6 (3.1)	13 (3.1)
7	0 (0.0)	3 (1.5)	3 (0.7)
8	2 (0.9)	0 (0.0)	2 (0.5)
9	1 (0.4)	0 (0.0)	1 (0.2)
<b>No. of conditions, median (minimum, maximum)</b>	3.0 (2.0, 9.0)	3.0 (2.0, 7.0)	3.0 (2.0, 9.0)
<b>Bayliss<sup>b</sup></b>	9.4 (6.4)	9.7 (7.7)	9.6 (7.0)

a Conditions grouped as per the inclusion criteria.

b Higher score indicates worse outcome.

**Note**

Data are mean (SD) unless otherwise stated.

**TABLE 12** Grouped conditions self-reported at baseline by randomised group as analysed

	Intervention (n = 227)	Usual care (n = 195)	Overall (n = 422)
Cardiovascular disease	154 (67.8)	135 (69.2)	289 (68.5)
Arthritis	127 (55.9)	98 (50.3)	225 (53.3)
Sensory conditions	82 (36.1)	75 (38.5)	157 (37.2)
Depression or anxiety	59 (26.0)	42 (21.5)	101 (23.9)
Asthma or COPD	57 (25.1)	40 (20.5)	97 (23.0)
Bowel problems	53 (23.3)	30 (15.4)	83 (19.7)
Osteoporosis or osteopenia	36 (15.9)	37 (19.0)	73 (17.3)

continued

**TABLE 12** Grouped conditions self-reported at baseline by randomised group as analysed (*continued*)

	Intervention (n = 227)	Usual care (n = 195)	Overall (n = 422)
Diabetes	33 (14.5)	34 (17.4)	67 (15.9)
Atrial fibrillation	34 (15.0)	32 (16.4)	66 (15.6)
Cancer (last 5 years)	33 (14.5)	31 (15.9)	64 (15.2)
Chronic kidney disease	14 (6.2)	15 (7.7)	29 (6.9)
Stroke (last 5 years)	5 (2.2)	10 (5.1)	15 (3.6)
Fibromyalgia	8 (3.5)	6 (3.1)	14 (3.3)
Epilepsy	1 (0.4)	5 (2.6)	6 (1.4)
Parkinson's disease	1 (0.4)	4 (2.1)	5 (1.2)
Multiple sclerosis	3 (1.3)	1 (0.5)	4 (0.9)
Dementia	2 (0.9)	1 (0.5)	3 (0.7)

COPD, chronic obstructive pulmonary disease.

**TABLE 13** Participant-reported expectations and preferences concerning the health care being offered in the GYY trial, collected at baseline, by randomised group as analysed

Expectations and preferences	Intervention (n = 227)	Usual care (n = 195)	Overall (n = 422)
<b>How effective do you think that usual care would be in improving your quality of life?</b>			
Very ineffective	15 (6.6)	6 (3.1)	21 (5.0)
Fairly ineffective	18 (7.9)	25 (12.8)	43 (10.2)
Can't decide	68 (30.0)	63 (32.3)	131 (31.0)
Fairly effective	75 (33.0)	61 (31.3)	136 (32.2)
Very effective	48 (21.1)	38 (19.5)	86 (20.4)
Missing	3 (1.3)	2 (1.0)	5 (1.2)
<b>How effective do you think the GYY programme would be in improving your quality of life?</b>			
Very ineffective	4 (1.8)	2 (1.0)	6 (1.4)
Fairly ineffective	7 (3.1)	4 (2.1)	11 (2.6)
Can't decide	77 (33.9)	74 (37.9)	151 (35.8)
Fairly effective	78 (34.4)	65 (33.3)	143 (33.9)
Very effective	61 (26.9)	50 (25.6)	111 (26.3)
<b>Given the choice, which study group would you prefer to be in?</b>			
Yoga and usual care	176 (77.5)	139 (71.3)	315 (74.6)
Usual care	5 (2.2)	6 (3.1)	11 (2.6)
No preference	46 (20.3)	50 (25.6)	96 (22.7)

**TABLE 14** Values of outcome measures assessed at baseline by randomised group as analysed

Outcome scores at baseline	Intervention (n = 227)	Usual care (n = 195)	Overall (n = 422)
EQ-5D-5L utility index score <sup>a</sup>	0.742 (0.175)	0.736 (0.163)	0.739 (0.169)
EQ-5D-5L VAS <sup>a</sup>	75.4 (18.2)	73.9 (17.2)	74.7 (17.7)
PHQ-8 <sup>b</sup>	3.6 (3.8)	3.7 (4.2)	3.7 (4.0)
GAD-7 <sup>b</sup>	2.4 (3.3)	2.6 (3.6)	2.5 (3.4)
PROMIS-29 physical functioning <sup>a</sup>	47.0 (8.4)	46.4 (8.4)	46.7 (8.4)
PROMIS-29 anxiety <sup>b</sup>	46.9 (8.0)	48.0 (8.6)	47.4 (8.3)
PROMIS-29 depression <sup>b</sup>	46.4 (7.5)	46.5 (8.0)	46.4 (7.7)
PROMIS-29 fatigue <sup>b</sup>	47.3 (9.7)	48.4 (9.8)	47.8 (9.7)
PROMIS-29 sleep disturbances <sup>b</sup>	49.1 (9.6)	49.5 (9.5)	49.3 (9.5)
PROMIS-29 social roles <sup>a</sup>	54.8 (9.2)	54.3 (10.0)	54.6 (9.6)
PROMIS-29 pain interference <sup>b</sup>	53.2 (8.7)	53.6 (8.9)	53.4 (8.8)
PROMIS-29 pain intensity <sup>b</sup>	3.1 (2.5)	3.1 (2.4)	3.1 (2.4)
PROMIS-29 mental health summary score <sup>a</sup>	53.0 (7.9)	52.2 (8.4)	52.6 (8.1)
PROMIS-29 physical health summary score <sup>a</sup>	47.9 (8.6)	47.2 (8.7)	47.6 (8.7)
UCLA-3 loneliness <sup>b</sup>	4.2 (1.7)	4.3 (1.8)	4.2 (1.7)
ELSA single-item direct loneliness question, <sup>b</sup> n (%)	2.1 (1.3)	2.2 (1.3)	2.2 (1.3)
Fallen past 3 months, n (%)			
Yes	58 (25.6)	46 (23.6)	104 (24.6)
No	169 (74.4)	148 (75.9)	317 (75.1)
Missing	0 (0.0)	1 (0.5)	1 (0.2)
If yes, number of falls			
Median (minimum, maximum)	1 (1, 4)	1 (1, 20)	1 (1, 20)
a Higher score indicates better outcome.			
b Lower score indicates better outcome.			
<b>Note</b>			
Data are mean (SD) unless otherwise stated.			

### Primary end-point analysis

There was no evidence of a statistically or clinically significant difference in EQ-5D-5L utility index score between the intervention and usual care arms over 12 months, with an adjusted MD of 0.02 in favour of the intervention group (95% CI -0.006 to 0.045,  $p = 0.14$ ). The predicted means and associated 95% CIs over time are presented in [Table 15](#) and displayed in [Figure 6](#), by group.

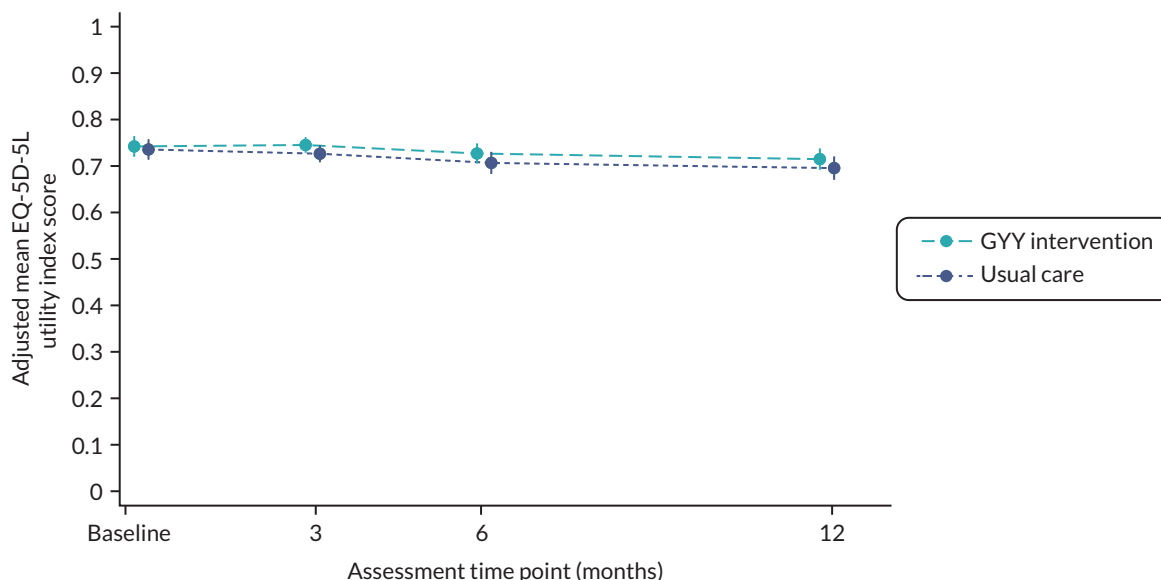
Different covariance structures were applied to the model, and the Akaike information criterions (AICs) were compared. An unstructured pattern that models all variances and covariances separately was used in the final model, as this resulted in the lowest AIC.

Model fit diagnostics indicated that the standardised residuals demonstrated only a minor deviation from normality and were uniform against fitted values; therefore untransformed values were used in analyses.

**TABLE 15** Difference in adjusted mean EQ-5D-5L utility index score over time by randomised group from primary and SA models

Time point, months	Intervention Mean (95% CI)	Usual care Mean (95% CI)	Difference (95% CI)	p-value
<b>Primary</b>				
3	0.745 (0.728 to 0.762)	0.726 (0.708 to 0.745)	0.019 (-0.006 to 0.044)	0.14
6	0.727 (0.705 to 0.749)	0.707 (0.683 to 0.730)	0.020 (-0.012 to 0.053)	0.22
12	0.715 (0.692 to 0.738)	0.696 (0.671 to 0.720)	0.019 (-0.015 to 0.053)	0.26
Overall	0.729 (0.712 to 0.747)	0.710 (0.691 to 0.729)	0.020 (-0.006 to 0.045)	0.14
<b>Sensitivity 1</b>				
3	0.745 (0.728 to 0.762)	0.727 (0.709 to 0.745)	0.018 (-0.007 to 0.042)	0.16
6	0.727 (0.705 to 0.748)	0.708 (0.684 to 0.731)	0.019 (-0.013 to 0.051)	0.24
12	0.715 (0.692 to 0.737)	0.696 (0.672 to 0.721)	0.018 (-0.015 to 0.052)	0.28
Overall	0.729 (0.712 to 0.746)	0.711 (0.692 to 0.729)	0.018 (-0.007 to 0.044)	0.16
<b>Sensitivity 2</b>				
3	0.745 (0.728 to 0.762)	0.726 (0.708 to 0.745)	0.019 (-0.006 to 0.044)	0.14
6	0.727 (0.705 to 0.749)	0.707 (0.683 to 0.730)	0.020 (-0.012 to 0.053)	0.22
12	0.715 (0.692 to 0.738)	0.696 (0.671 to 0.720)	0.019 (-0.015 to 0.053)	0.26
Overall	0.729 (0.712 to 0.747)	0.710 (0.691 to 0.729)	0.020 (-0.006 to 0.045)	0.14

**Note**  
(n = 422; intervention, n = 227; usual care, n = 195).



**FIGURE 6** Adjusted mean EQ-5D-5L utility index scores (with 95% CIs) for primary analysis over time by randomised group.

Model coefficients for the covariates with 95% CIs are provided as software output in [Appendix 2](#) to aid understanding of the fitted model, along with summaries of the EQ-5D-5L index value by trial site and time point to assess variation between sites (see [Table 45](#)).

## Sensitivity analyses

### *Adjusting for other covariates (sensitivity 1)*

Results were very similar when the primary analysis was repeated with age, gender and adapted Bayliss score additionally adjusted for as fixed effects (see [Table 15](#)).

### *Clustering by yoga teacher (sensitivity 2)*

Nineteen yoga courses were delivered within the trial by 12 yoga teachers (1 teacher delivered 3 courses, 5 teachers delivered 2 courses each and 6 teachers delivered 1 course each). Analyses to account for possible clustering by yoga teacher were undertaken by including the intended yoga teacher as a random effect instead of site in the primary analysis model; results were virtually unchanged (see [Table 15](#)).

### *Compliance with random allocation and treatment received*

One participant in the usual care group was invited to attend classes in error; they attended eight sessions, including five of the first six sessions.

A summary of attendance at weekly GYY sessions for intervention participants is presented in [Table 16](#). Among the intervention group, 222 (92.5%) participants attended at least 1 GYY class, while 53 (22.1%) attended all 12 (see [Figure 7](#)). The mean number of sessions attended among all randomised intervention participants was 8.8 (SD 3.7, median 10) and 9.6 (SD 2.8, median 11) among those who attended at least one. Eighty per cent ( $n = 192$ ) of participants attended at least six sessions, including at least three of the first six (see [Table 17](#)).

On average, the first class in a course took place 18.2 days (SD 2.7, median 19, range 13–21) after the participant was randomised, and classes were scheduled a median of 7 days apart (range 7–28; longer intervals tended to be due to the Christmas period) (see [Report Supplementary Material 3](#)).

Three CACE analyses for the primary outcome were undertaken to explore the impact of non-compliance on treatment effect estimates, with compliance defined as:

- Attendance at one yoga session or more ( $n = 222$  intervention participants, 92.5%;  $n = 1$  usual care participant, 0.5%). The CACE estimate of the treatment effect is a difference of 0.025 at 12 months in favour of the intervention group (95% CI  $-0.002$  to  $0.052$ ,  $p = 0.07$ ). This difference is larger than the ITT estimate [The CACE analysis is not directly comparable with the primary ITT analysis as the CACE analysis cannot take account of the repeated measures for the EQ-5D-5L utility index score at 3, 6 and 12 months; it simply considers the difference at 12 months. Therefore, we conducted a linear regression with 12-month EQ-5D-5L utility index score as the outcome variable, adjusting for baseline score and gender with robust standard errors to account for clustering within site], but neither the treatment effect nor the upper 95% CI limit exceeds the clinically meaningful difference of 0.06.
- Attendance of at least three of the first six sessions and at least three other sessions ( $n = 192$  intervention participants, 80.0%;  $n = 1$  usual care participant, 0.5%). The CACE estimate of the treatment effect is a difference of 0.029 at 12 months in favour of the intervention group (95% CI  $-0.002$  to  $0.059$ ,  $p = 0.06$ ). This difference is larger than the ITT estimate [The CACE analysis is not directly comparable with the primary ITT analysis as the CACE analysis cannot take account of the repeated measures for the EQ-5D-5L utility index score at 3, 6 and 12 months; it simply considers the difference at 12 months. Therefore, we conducted a linear regression with 12-month EQ-5D-5L utility index score as the outcome variable, adjusting for baseline score and gender with robust standard errors to account for clustering within site.], but neither

**TABLE 16** Summary of GYY class attendance by week and recruitment wave

		Week number												
		1	2	3	4	5	6	7	8	9	10	11	12	Average
<b>Pilot phase</b>														
<b>Pilot phase wave 1</b> (n = 60)	No. of non-withdrawn patients	56	56	55	52	50	50	50	50	50	50	49	49	51
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	42 (70, 75)	44 (73, 79)	46 (77, 84)	40 (67, 77)	37 (62, 74)	34 (57, 68)	35 (58, 70)	33 (55, 66)	29 (48, 58)	32 (53, 64)	36 (60, 73)	34 (57, 69)	37 (61, 72)
<b>Pilot phase wave 2</b> (n = 48)	No. of non-withdrawn patients	45	45	45	45	45	45	45	45	45	45	45	45	45
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	40 (83, 89)	41 (85, 91)	43 (90, 96)	40 (83, 89)	39 (81, 87)	41 (85, 91)	38 (79, 84)	39 (81, 87)	37 (77, 82)	36 (75, 80)	41 (85, 91)	35 (73, 78)	39 (81, 87)
<b>Pilot phase overall</b> (n = 108)	No. of non-withdrawn patients	101	101	100	97	95	95	95	95	95	95	94	94	96
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	82 (76, 81)	85 (79, 84)	89 (82, 89)	80 (74, 82)	76 (70, 80)	75 (69, 79)	73 (68, 77)	72 (67, 76)	66 (61, 69)	68 (63, 72)	77 (71, 82)	69 (64, 73)	76 (70, 79)
<b>Main phase</b>														
<b>Main phase wave 1</b> (n = 36)	No. of non-withdrawn patients	34	33	33	33	32	31	31	31	31	31	31	31	32
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	32 (89, 94)	28 (78, 85)	32 (89, 97)	31 (86, 94)	29 (81, 91)	28 (78, 90)	27 (75, 87)	28 (78, 90)	25 (69, 81)	26 (72, 84)	24 (67, 77)	27 (75, 87)	28 (78, 88)
<b>Main phase wave 2</b> (n = 96)	No. of non-withdrawn patients	90	90	89	89	88	88	87	87	84	82	80	80	86
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	79 (82, 88)	78 (81, 87)	81 (84, 91)	73 (76, 82)	71 (74, 81)	70 (73, 80)	75 (78, 86)	74 (77, 85)	73 (76, 87)	67 (70, 82)	66 (69, 83)	66 (69, 83)	73 (76, 85)
<b>Main phase overall</b> (n = 132)	No. of non-withdrawn patients	124	123	122	122	120	119	118	118	115	113	111	111	118
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	111 (84, 90)	106 (80, 86)	113 (86, 93)	104 (79, 85)	100 (76, 83)	98 (74, 82)	102 (77, 86)	102 (77, 86)	98 (74, 85)	93 (70, 82)	90 (68, 81)	93 (70, 84)	101 (76, 85)
<b>Overall</b>														
<b>Overall</b> (n = 240)	No. of non-withdrawn patients	225	224	222	219	215	214	213	213	210	208	205	205	214
	No. of participants attending (% of rand pts, % of non-withdrawn pts)	193 (80, 86)	191 (80, 85)	202 (84, 91)	184 (77, 84)	176 (73, 82)	173 (72, 81)	175 (73, 82)	174 (73, 82)	164 (68, 78)	161 (67, 77)	167 (70, 81)	162 (68, 79)	177 (74, 82)

rand pts, randomised participants.

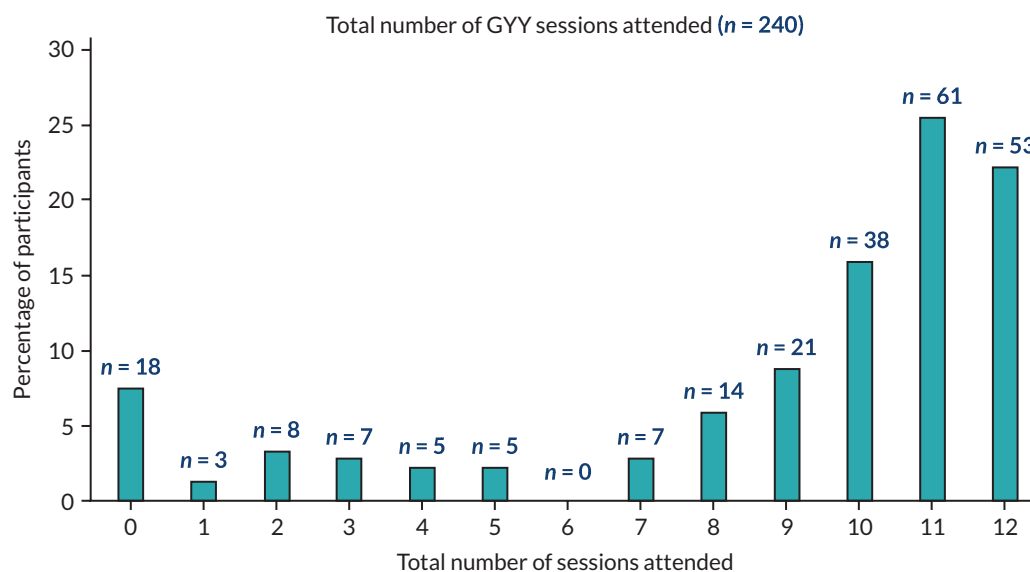


FIGURE 7 Number of sessions attended by GYY intervention group participants.

TABLE 17 Definitions of adherence to GYY intervention by recruitment wave

	Pilot phase wave 1 ( $n = 60$ )	Pilot phase wave 2 ( $n = 48$ )	Overall (pilot) ( $n = 108$ )	Main phase wave 1 ( $n = 36$ )	Main phase wave 2 ( $n = 96$ )	Overall (main) ( $n = 132$ )	Overall ( $n = 240$ )
Attended all 12 sessions, $n$ (%)	10 (17)	13 (27)	23 (21)	9 (25)	21 (22)	30 (23)	53 (22)
Attended $\geq 9$ sessions, $n$ (%)	32 (53)	41 (85)	73 (68)	28 (78)	72 (75)	100 (76)	173 (72)
Attended (at least) 3 of the first 6 sessions and 3 others, $n$ (%)	39 (65)	43 (90)	82 (76)	31 (86)	79 (82)	110 (83)	192 (80)

the treatment effect nor the upper 95% CI limit exceeds the clinically meaningful difference of 0.06.

- Number of sessions attended in its continuous form (intervention: mean 8.8, SD 3.7, one usual care participant attended eight sessions). The CACE estimate was a difference of 0.003 per session (95% CI  $-0.000$  to  $0.005$ ,  $p = 0.07$ ), indicating a very small additional benefit of the intervention for each session attended.

### Yoga practices

Participant self-reported data on attendance at yoga classes and home yoga practice at 3-, 6- and 12-month follow-ups are presented in [Tables 18–20](#), respectively.

The self-reported data relating to attendance at GYY classes at the 3-month follow-up match well with the attendance register data (see [Table 18](#)). All 214 participants in the intervention group who self-reported as having attended a GYY session were recorded as having attended at least one session (8 participants were recorded on the attendance registers but did not return a 3-month questionnaire). The one usual care participant who self-reported as attending was the person we expected. Only a small number of participants in both groups reported attending other (non-trial) yoga sessions. Most of the intervention group reported practising yoga at home, which could include as part of the GYY programme ( $n = 185$ , 82.6%), but only a small number of usual care participants ( $n = 6$ , 3.2%). Where undertaken,



participants in the intervention group did twice as many home yoga sessions as usual care participants (median 4 vs. 2), though these sessions tended to last for a similar length of time (median 15 minutes).

At 6 months, 72 (33.3%) intervention participants reported that they had attended a GYY session in the previous 3 months, all of whom were confirmed to have attended at least one session according to the class registers (see [Table 19](#)). It is likely the other attendees had completed their sessions more than 3 months prior to completing this follow-up questionnaire. Three usual care participants reported having attended a GYY session, though none of these were present on the class registers. GYY classes are available to the public through the BWY, so it is possible these participants had sought out and attended a GYY session not delivered as part of the trial, and therefore this would not be captured as part of our evaluation. The proportion of participants reporting home yoga practices decreased relative to month 3 (in accordance with the cessation of the GYY course) in the intervention group (as did the median number of home yoga sessions from 4 to 3) but increased very slightly in the usual care group.

At 12 months, only a fifth of participants reported having attended GYY classes in the previous 6 months ( $n = 41$ , 19.3%) and two usual care participants (again, not participants who were on a class register) (see [Table 20](#)). As at 6 months, the proportion reporting home yoga practice decreased in the intervention group relative to the previous follow-up time point (to just less than half) but increased in the usual care group (9.6%). Participants in the intervention group reported doing a median of three home yoga sessions a week lasting a median of 15 minutes, while for the usual care group, this was a median of two sessions a week for 10 minutes.

### Intervention fidelity

All yoga teachers submitted a course plan to the yoga consultant for pre-approval ahead of their first class. Each teacher received timely feedback on their plan. The feedback was mostly positive, and all plans met the assessment criteria and were therefore deemed appropriate for delivery.

**TABLE 18** Self-reported yoga practice at 3-month follow-up

Month 3	Intervention ( $n = 224$ )	Usual care ( $n = 189$ )	Overall ( $n = 413$ )
Attended any yoga classes in the <b>past 3 months</b> , $n$ (%)			
No	9 (4.0)	179 (94.7)	188 (45.5)
Yes – GYY classes	214 (95.5)	1 (0.5)	215 (52.1)
Yes – other group-based yoga classes	3 (1.3)	4 (2.1)	7 (1.7)
Yes – one-to-one yoga sessions	0 (0.0)	0 (0.0)	0 (0.0)
If yes, how many sessions:			
Mean (SD), $N$	9.3 (2.6), 210	7.6 (1.5), 5	9.2 (2.5), 215
Median (IQR)	10.0 (8.0–11.0)	8.0 (6.0–9.0)	10.0 (8.0–11.0)
Practising yoga at home in the <b>past 3 months</b> (including as part of the GYY programme)?, $n$ (%)			
	185 (82.6)	6 (3.2)	191 (46.2)
If Yes, number of home sessions done in a typical week:			
Mean (SD), $N$	4.3 (1.9), 184	2.3 (0.5), 4	4.2 (1.9), 188
Median (IQR)	4.0 (3.0–6.0)	2.0 (2.0–2.5)	4.0 (3.0–5.5)
Usual duration of each session, in minutes			
Mean (SD), $N$	18.3 (14.6), 183	25.0 (23.5), 4	18.4 (14.7), 187
Median (IQR)	15.0 (10.0–20.0)	15.0 (12.5–37.5)	15.0 (10.0–20.0)

IQR, interquartile range.

TABLE 19 Self-reported yoga practice at 6-month follow-up

Month 6	Intervention (n = 216)	Usual care (n = 181)	Overall (n = 397)
Attended any yoga classes in the <b>past 3 months</b> , n (%)			
No	129 (59.7)	173 (95.6)	302 (76.1)
Yes – GYY classes	72 (33.3)	2 (1.1)	74 (18.6)
Yes – other group-based yoga classes	16 (7.4)	3 (1.7)	19 (4.8)
Yes – one-to-one yoga sessions	3 (1.4)	0 (0.0)	3 (0.8)
If yes, how many sessions:			
Mean (SD), N	6.6 (4.3), 83	6.8 (3.0), 5	6.6 (4.3), 88
Median (IQR)	6.0 (3.0–10.0)	8.0 (6.0–8.0)	6.0 (3.0–10.0)
Practising yoga at home in the <b>past 3 months</b> (including as part of the GYY programme)?, n (%)	123 (56.9)	11 (6.1)	134 (33.8)
If Yes, number of home sessions done in a typical week:			
Mean (SD), N	3.8 (2.3), 123	3.2 (1.9), 10	3.7 (2.3), 133
Median (IQR)	3.0 (2.0–5.0)	2.5 (2.0–5.0)	3.0 (2.0–5.0)
Usual duration of each session, in minutes			
Mean (SD), N	19.1 (12.8), 123	26.5 (16.2), 10	19.6 (13.1), 133
Median (IQR)	15.0 (10.0–20.0)	30.0 (15.0–30.0)	15.0 (10.0–25.0)
IQR, interquartile range.			

TABLE 20 Self-reported yoga practice at 12-month follow-up

Month 12	Intervention (n = 212)	Usual care (n = 178)	Overall (n = 390)
Attended any yoga classes in the <b>past 6 months</b> , n (%)			
No	157 (74.1)	167 (93.8)	324 (83.1)
Yes – GYY classes	41 (19.3)	2 (1.1)	43 (11.0)
Yes – other group-based yoga classes	14 (6.6)	7 (3.9)	21 (5.4)
Yes – one-to-one yoga sessions	0 (0.0)	0 (0.0)	0 (0.0)
If yes, how many sessions:			
Mean (SD), N	15.4 (8.4), 49	12.0 (7.8), 7	15.0 (8.4), 56
Median (IQR)	16.0 (10.0–20.0)	10.0 (5.0–20.0)	14.0 (9.5–20.0)
Practising yoga at home in the <b>past 6 months</b> (including as part of the GYY programme)?, n (%)	102 (48.1)	17 (9.6)	119 (30.5)
If Yes, number of home sessions done in a typical week			
Mean (SD), N	3.1 (2.0), 100	2.8 (1.9), 17	3.1 (2.0), 117
Median (IQR)	3.0 (2.0–4.0)	2.0 (1.0–4.0)	2.0 (2.0–4.0)
Usual duration of each session, in minutes			
Mean (SD), N	20.3 (13.3), 99	13.6 (7.6), 17	19.3 (12.8), 116
Median (IQR)	15.0 (10.0–30.0)	10.0 (10.0–15.0)	15.0 (10.0–20.0)
IQR, interquartile range.			

Each yoga teacher underwent an observation of one of their trial classes by one of the yoga consultants. A fidelity check assessment form was completed for each observation. All yoga teachers passed all aspects of the fidelity check assessment criteria.

## Subgroup analyses

### *Intended mode of delivery of Gentle Years Yoga*

More participants were randomised in a site intended for online GYY delivery (61.9% across 12 sites; intervention group  $n = 144$ , 60.0%; usual care group  $n = 137$ , 64.0%) than face to face (38.1% across 7 sites; intervention group  $n = 96$ , 40.0%; usual care group  $n = 77$ , 36.0%). A subgroup analysis was conducted in which the primary analysis was repeated, including, as a fixed effect, an indicator for this factor plus an interaction with trial arm. There was no evidence of an interaction between trial arm and intended mode of delivery (interaction effect 0.007, 95% CI  $-0.042$  to  $0.057$ ,  $p = 0.77$ ).

## Secondary analysis

### *EuroQol-5 Dimensions, five-level version utility index scores at the secondary time points*

Adjusted EQ-5D-5L utility index score means and group differences from the primary analysis model are presented in [Table 22](#) and displayed in [Figure 6](#). There was no evidence of a statistically significant difference at 3, 6 or 12 months, and none of the CIs for the differences contained the clinically meaningful difference of 0.06.

### *EuroQol-5 Dimensions, five-level version visual analogue scale*

Raw EQ-5D-5L VAS scores are summarised in [Table 21](#). Adjusted means and group differences are presented in [Table 22](#). The analysis included data from 423 participants (intervention  $n = 227$ , 94.6%; usual care  $n = 196$ , 91.6%). There was no evidence of a statistically significant difference at any time point.

### *Generalised Anxiety Disorder-7*

Raw GAD-7 scores are summarised in [Table 21](#). Adjusted means and group differences are presented in [Table 22](#). The analysis included data from 420 participants (intervention  $n = 227$ , 94.6%; usual care  $n = 193$ , 90.2%). There was no evidence of a statistically significant difference at any time point.

### *Patient Health Questionnaire-8*

Raw PHQ-8 scores are summarised in [Table 21](#). Adjusted means and group differences are presented in [Table 22](#). The analysis included data from 419 participants (intervention  $n = 227$ , 94.6%; usual care  $n = 192$ , 89.7%). There was no evidence of a statistically significant difference at any time point.

### *University of California, Los Angeles-3 loneliness*

Raw UCLA-3 scores are summarised in [Table 21](#). Adjusted means and group differences are presented in [Table 22](#). The analysis included data from 419 participants (intervention  $n = 227$ , 94.6%; usual care  $n = 192$ , 89.7%). There was no evidence of a statistically significant difference at any time point.

### *English Longitudinal Study of Ageing single-item direct loneliness question*

Raw ELSA single-item direct loneliness question scores are summarised in [Table 21](#). Adjusted means and group differences are presented in [Table 22](#). The analysis included data from 421 participants (intervention  $n = 227$ , 94.6%; usual care  $n = 194$ , 90.7%). There was no evidence of a statistically significant difference at any time point.

TABLE 21 Summary of raw scores for secondary outcomes

Time point, months	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
<b>EQ-5D-5L VAS, mean (SD)</b>			
3	76.1 (16.7)	74.1 (16.9)	75.2 (16.8)
6	74.4 (18.1)	71.8 (19.2)	73.2 (18.6)
12	73.8 (18.3)	70.1 (21.5)	72.1 (19.9)
<b>GAD-7, mean (SD)</b>			
3	3.8 (3.9)	4.5 (4.5)	4.1 (4.2)
6	4.0 (4.2)	4.3 (4.1)	4.1 (4.1)
12	4.1 (4.4)	4.6 (4.8)	4.3 (4.6)
<b>PHQ-8, mean (SD)</b>			
3	2.7 (3.4)	3.1 (4.0)	2.9 (3.7)
6	2.8 (3.5)	3.0 (3.9)	2.9 (3.7)
12	2.8 (3.9)	3.1 (4.0)	2.9 (4.0)
<b>UCLA-3 loneliness, mean (SD)</b>			
3	4.4 (1.8)	4.3 (1.7)	4.3 (1.7)
6	4.4 (1.7)	4.4 (1.8)	4.4 (1.8)
12	4.3 (1.7)	4.5 (1.8)	4.4 (1.7)
<b>ELSA single-item direct loneliness question, mean (SD)</b>			
3	2.3 (1.2)	2.3 (1.2)	2.3 (1.2)
6	2.3 (1.3)	2.3 (1.2)	2.3 (1.3)
12	2.2 (1.2)	2.5 (1.3)	2.3 (1.3)

TABLE 22 Difference in adjusted means over time by randomised group for secondary outcomes

Time point, months	Intervention Mean (95% CI)	Usual care Mean (95% CI)	Difference (95% CI)	p-value
<b>EQ-5D-5L VAS</b>				
3	75.6 (73.8 to 77.4)	74.5 (72.6 to 76.5)	1.08 (-1.55 to 3.71)	0.42
6	73.8 (71.7 to 75.9)	72.1 (69.8 to 74.4)	1.74 (-1.39 to 4.86)	0.28
12	73.1 (70.8 to 75.3)	70.9 (68.4 to 73.4)	2.18 (-1.19 to 5.55)	0.20
Overall	74.2 (72.5 to 75.8)	72.5 (70.7 to 74.3)	1.67 (-0.78 to 4.12)	0.18
<b>GAD-7</b>				
3	2.8 (2.4 to 3.2)	3.0 (2.6 to 3.4)	-0.17 (-0.72 to 0.37)	0.53
6	2.9 (2.5 to 3.3)	3.0 (2.5 to 3.4)	-0.10 (-0.70 to 0.50)	0.74
12	3.0 (2.5 to 3.4)	2.9 (2.5 to 3.4)	0.01 (-0.61 to 0.63)	0.98
Overall	2.9 (2.5 to 3.2)	3.0 (2.6 to 3.3)	-0.09 (-0.57 to 0.40)	0.72

continued

**TABLE 22** Difference in adjusted means over time by randomised group for secondary outcomes (*continued*)

Time point, months	Intervention Mean (95% CI)	Usual care Mean (95% CI)	Difference (95% CI)	p-value
<b>PHQ-8</b>				
3	3.9 (3.5 to 4.2)	4.4 (4.0 to 4.8)	-0.53 (-1.12 to 0.05)	0.07
6	4.1 (3.6 to 4.5)	4.4 (3.9 to 4.9)	-0.30 (-0.97 to 0.36)	0.37
12	4.3 (3.8 to 4.7)	4.5 (4.0 to 5.0)	-0.25 (-0.93 to 0.43)	0.48
Overall	4.1 (3.7 to 4.4)	4.4 (4.0 to 4.8)	-0.36 (-0.90 to 0.18)	0.19
<b>UCLA-3 loneliness</b>				
3	4.3 (4.2 to 4.5)	4.3 (4.1 to 4.4)	0.07 (-0.15 to 0.29)	0.54
6	4.4 (4.3 to 4.6)	4.4 (4.2 to 4.6)	0.03 (-0.21 to 0.26)	0.83
12	4.4 (4.3 to 4.6)	4.4 (4.3 to 4.6)	-0.00 (-0.24 to 0.23)	0.97
Overall	4.4 (4.3 to 4.5)	4.4 (4.2 to 4.5)	0.03 (-0.16 to 0.22)	0.75
<b>ELSA loneliness</b>				
3	2.3 (2.2 to 2.4)	2.3 (2.2 to 2.5)	-0.01 (-0.17 to 0.16)	0.94
6	2.4 (2.3 to 2.5)	2.3 (2.2 to 2.4)	0.07 (-0.10 to 0.25)	0.41
12	2.3 (2.2 to 2.4)	2.4 (2.3 to 2.5)	-0.10 (-0.27 to 0.08)	0.28
Overall	2.3 (2.2 to 2.4)	2.3 (2.2 to 2.5)	-0.01 (-0.15 to 0.13)	0.88

**TABLE 23** Summary of raw scores for PROMIS-29 secondary outcomes

Time point, months	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
<b>Physical function, mean (SD)</b>			
3	47.0 (8.5)	45.9 (8.8)	46.5 (8.6)
6	47.1 (8.5)	45.9 (9.1)	46.5 (8.8)
12	46.5 (8.4)	44.9 (8.5)	45.8 (8.5)
<b>Anxiety, mean (SD)</b>			
3	47.8 (8.8)	47.9 (8.9)	47.9 (8.8)
6	47.6 (8.5)	48.7 (9.1)	48.1 (8.8)
12	47.6 (8.8)	48.0 (10.0)	47.7 (9.4)
<b>Depression, mean (SD)</b>			
3	47.3 (8.0)	47.5 (8.3)	47.4 (8.2)
6	47.8 (8.0)	48.0 (8.1)	47.9 (8.0)
12	47.9 (8.2)	47.6 (8.8)	47.7 (8.5)
<b>Fatigue, mean (SD)</b>			
3	47.7 (10.5)	49.0 (10.2)	48.3 (10.4)
6	47.5 (10.3)	49.1 (10.5)	48.2 (10.5)
12	48.5 (9.9)	49.3 (10.9)	48.9 (10.4)

TABLE 23 Summary of raw scores for PROMIS-29 secondary outcomes (continued)

Time point, months	Intervention (n = 240)	Usual care (n = 214)	Overall (n = 454)
<b>Sleep disturbance, mean (SD)</b>			
3	49.8 (9.2)	50.5 (9.4)	50.1 (9.3)
6	49.5 (8.7)	50.6 (9.1)	50.0 (8.9)
12	49.7 (8.5)	50.4 (9.5)	50.1 (9.0)
<b>Social participation, mean (SD)</b>			
3	53.1 (9.5)	51.3 (10.2)	52.3 (9.9)
6	52.4 (10.7)	52.2 (10.3)	52.3 (10.5)
12	52.6 (9.3)	50.8 (10.3)	51.8 (9.8)
<b>Pain interference, mean (SD)</b>			
3	52.8 (8.8)	54.6 (8.6)	53.6 (8.8)
6	52.7 (9.2)	54.0 (9.3)	53.3 (9.3)
12	53.2 (9.4)	54.6 (8.9)	53.8 (9.2)
<b>Physical health, mean (SD)</b>			
3	47.7 (8.9)	46.3 (9.1)	47.0 (9.0)
6	47.5 (8.8)	46.5 (9.4)	47.0 (9.1)
12	47.2 (8.8)	45.4 (8.9)	46.4 (8.9)
<b>Mental health, mean (SD)</b>			
3	52.1 (8.6)	50.8 (8.7)	51.5 (8.6)
6	51.9 (8.5)	50.9 (8.8)	51.5 (8.6)
12	51.8 (8.4)	50.5 (9.1)	51.2 (8.8)
<b>Global (pain intensity), mean (SD)</b>			
3	3.0 (2.4)	3.3 (2.4)	3.1 (2.4)
6	3.1 (2.5)	3.4 (2.5)	3.3 (2.5)
12	3.2 (2.5)	3.8 (2.4)	3.4 (2.4)

### Patient-Reported Outcomes Measurement Information System-29

Raw PROMIS-29 scores are summarised in [Table 23](#). Adjusted means and group differences are presented in [Table 24](#). These analyses included data from between 419 and 421 participants. There was evidence of a statistically significant difference in the T-score for the pain interference subscale of the PROMIS-29 at 3 months ( $-1.44$ , 95% CI  $-2.63$  to  $-0.26$ ;  $p = 0.02$ ) and over the 12 months ( $-1.14$ , 95% CI  $-2.24$  to  $-0.04$ ;  $p = 0.04$ ) and in the global (pain intensity) PROMIS-29 item at 12 months ( $-0.45$ , 95% CI  $-0.83$  to  $-0.08$ ;  $p = 0.02$ ) and over the 12 months ( $-0.32$ , 95% CI  $-0.61$  to  $-0.04$ ;  $p = 0.03$ ). Differences favoured the intervention. Otherwise, no statistically significant differences were observed.

### Falls

A total of 421 participants responded to the question asking whether they had had a fall in the previous 3 or 6 months on at least one of the post-randomisation questionnaires, of which 112 (26.6%) said they had [60/227 (26.4%) in the intervention group and 52/194 (26.8%) in usual care].

**TABLE 24** Difference in adjusted means over time by randomised group for PROMIS-29 outcomes

Time point, months	Intervention Mean (95% CI)	Usual care Mean (95% CI)	Difference (95% CI)	p-value
<b>Physical function</b>				
3	46.9 (46.2 to 47.6)	46.4 (45.6 to 47.1)	0.50 (-0.55 to 1.55)	0.35
6	46.9 (46.1 to 47.7)	46.0 (45.1 to 46.8)	0.90 (-0.27 to 2.07)	0.13
12	46.1 (45.2 to 46.9)	45.3 (44.3 to 46.2)	0.80 (-0.45 to 2.04)	0.21
Overall	46.6 (46.0 to 47.3)	45.9 (45.2 to 46.6)	0.73 (-0.22 to 1.69)	0.13
<b>Anxiety</b>				
3	48.1 (47.2 to 49.0)	47.3 (46.3 to 48.3)	0.80 (-0.53 to 2.13)	0.24
6	48.0 (47.0 to 48.9)	48.3 (47.3 to 49.3)	-0.35 (-1.72 to 1.03)	0.62
12	48.0 (47.0 to 49.1)	47.4 (46.3 to 48.5)	0.67 (-0.85 to 2.19)	0.39
Overall	48.0 (47.2 to 48.9)	47.7 (46.8 to 48.5)	0.37 (-0.82 to 1.56)	0.54
<b>Depression</b>				
3	47.3 (46.4 to 48.2)	47.4 (46.5 to 48.4)	-0.14 (-1.42 to 1.15)	0.83
6	47.9 (47.0 to 48.7)	48.2 (47.3 to 49.1)	-0.33 (-1.56 to 0.91)	0.60
12	48.1 (47.2 to 49.0)	47.4 (46.4 to 48.4)	0.70 (-0.68 to 2.08)	0.32
Overall	47.7 (47.0 to 48.5)	47.7 (46.9 to 48.5)	0.08 (-1.02 to 1.17)	0.89
<b>Fatigue</b>				
3	48.1 (47.1 to 49.1)	48.4 (47.3 to 49.5)	-0.28 (-1.75 to 1.19)	0.71
6	47.8 (46.8 to 48.9)	48.8 (47.7 to 49.9)	-0.97 (-2.51 to 0.57)	0.22
12	49.1 (48.0 to 50.1)	48.6 (47.5 to 49.8)	0.43 (-1.15 to 2.02)	0.59
Overall	48.3 (47.5 to 49.2)	48.6 (47.7 to 49.5)	-0.27 (-1.52 to 0.97)	0.67
<b>Sleep disturbance</b>				
3	50.1 (49.2 to 51.0)	50.2 (49.2 to 51.2)	-0.16 (-1.43 to 1.11)	0.80
6	49.8 (48.9 to 50.7)	50.2 (49.2 to 51.2)	-0.42 (-1.68 to 0.85)	0.52
12	50.0 (49.1 to 50.9)	49.9 (48.9 to 50.9)	0.08 (-1.22 to 1.37)	0.91
Overall	50.0 (49.2 to 50.7)	50.1 (49.3 to 51.0)	-0.17 (-1.19 to 0.85)	0.75
<b>Social participation</b>				
3	53.0 (51.9 to 54.0)	51.6 (50.5 to 52.7)	1.39 (-0.14 to 2.92)	0.08
6	52.3 (51.1 to 53.5)	52.1 (50.8 to 53.4)	0.21 (-1.60 to 2.01)	0.82
12	52.4 (51.3 to 53.5)	51.1 (49.9 to 52.3)	1.28 (-0.37 to 2.92)	0.13
Overall	52.6 (51.6 to 53.5)	51.6 (50.6 to 52.6)	0.96 (-0.40 to 2.32)	0.17
<b>Pain interference</b>				
3	53.0 (52.2 to 53.8)	54.4 (53.5 to 55.3)	-1.44 (-2.63 to -0.26)	0.02
6	52.9 (52.0 to 53.9)	54.0 (52.9 to 55.0)	-1.03 (-2.40 to 0.34)	0.14
12	53.4 (52.4 to 54.4)	54.3 (53.2 to 55.5)	-0.94 (-2.47 to 0.59)	0.23
Overall	53.1 (52.3 to 53.8)	54.2 (53.4 to 55.1)	-1.14 (-2.24 to -0.04)	0.04

**TABLE 24** Difference in adjusted means over time by randomised group for PROMIS-29 outcomes (continued)

Time point, months	Intervention Mean (95% CI)	Usual care Mean (95% CI)	Difference (95% CI)	p-value
<b>Physical health</b>				
3	47.6 (46.9 to 48.3)	46.8 (46.0 to 47.6)	0.76 (−0.30 to 1.82)	0.16
6	47.5 (46.7 to 48.3)	46.6 (45.7 to 47.5)	0.89 (−0.31 to 2.09)	0.14
12	46.8 (45.9 to 47.7)	45.8 (44.9 to 46.8)	0.97 (−0.32 to 2.25)	0.14
Overall	47.3 (46.6 to 48.0)	46.4 (45.7 to 47.2)	0.87 (−0.11 to 1.86)	0.08
<b>Mental health</b>				
3	51.9 (51.2 to 52.6)	51.4 (50.6 to 52.2)	0.48 (−0.61 to 1.56)	0.39
6	51.7 (51.0 to 52.5)	51.2 (50.4 to 52.1)	0.50 (−0.65 to 1.66)	0.39
12	51.3 (50.5 to 52.2)	51.2 (50.3 to 52.1)	0.11 (−1.12 to 1.35)	0.86
Overall	51.7 (51.0 to 52.3)	51.3 (50.6 to 52.0)	0.36 (−0.63 to 1.36)	0.47
<b>Global (pain intensity)</b>				
3	3.0 (2.8 to 3.2)	3.3 (3.0 to 3.5)	−0.26 (−0.58 to 0.06)	0.11
6	3.1 (2.9 to 3.4)	3.4 (3.1 to 3.7)	−0.26 (−0.62 to 0.09)	0.15
12	3.2 (2.9 to 3.5)	3.7 (3.4 to 4.0)	−0.45 (−0.83 to −0.08)	0.02
Overall	3.1 (2.9 to 3.3)	3.4 (3.2 to 3.7)	−0.32 (−0.61 to −0.04)	0.03

A mean of 0.82 (SD 2.0, median 0, range 0–21) falls per person was reported over an average of 10.5 (SD 3.6, median 12) months [intervention 0.91 (SD 2.1, median 0, range 0–21) falls over 10.8 (SD 3.2, median 12) months; usual care 0.71 (SD 1.9, median 0, range 0–15) falls over 10.2 (SD 3.9, median 12) months]. There was no evidence of a statistically significant difference in the rate of falls reported over the 12 months of follow-up (incidence rate ratio 1.38, 95% CI 0.95 to 2.01,  $p = 0.09$ ).

## Adverse events

There were no reported serious and related AEs.

There were seven reported non-SAEs that were deemed to be at least possibly related to the intervention and that were expected (see [Table 25](#)). These were reported for seven participants, all in the intervention group. The events all related to the onset or aggravation of pain during or after the yoga sessions (back pain  $n = 3$ , shoulder  $n = 1$ , knee  $n = 1$ , knee and shoulder  $n = 1$ , thigh  $n = 1$ ), though none required medical attention beyond taking pain killers. Four of the events were recorded as resolved in their initial report. Of the three that were ongoing, two were subsequently followed up, and the events were deemed to be resolved without the need for further medical intervention. Three of the seven participants subsequently withdrew from the intervention due to the pain, including the two participants for whom the event was deemed definitely related.



**TABLE 25** Summary of non-SAEs

	<b>Non-SAEs (n = 7)</b>
<b>Description of event, n (%)</b>	
Pain	7 (100)
<b>Action taken,<sup>a</sup> n (%)</b>	
Study treatment interrupted/halted	5 (71.4)
Therapy prescribed	1 (14.3)
Other	5 (71.4)
<b>Days from randomisation to onset, mean (SD)</b>	39.1 (41.4)
<b>Presence of event, n (%)</b>	
Continuous	4 (57.1)
Intermittent	3 (42.9)
<b>Outcome of event, n (%)</b>	
Resolved	4 (57.1)
Ongoing	3 (42.9)
<b>Participant withdrew from intervention, n (%)</b>	
Yes	3 (42.9)
No	4 (57.1)
<b>Relationship to study treatment, n (%)</b>	
Probably related	5 (71.4)
Definitely related	2 (28.6)
<sup>a</sup> Not mutually exclusive.	

# Chapter 4 Economic evaluation

## Introduction and aim

This chapter presents the methods and results of the economic evaluation that was conducted as part of the GYY trial. The aim was to assess if the GYY programme is cost-effective relative to usual care/ determine the cost-effectiveness of the GYY programme (the intervention) plus usual care compared to usual care alone, in terms of the impact on participants' health-related quality-of-life utility score.

## Methods

The economic evaluation comprised the following:

1. a within-trial economic analysis to assess the cost-effectiveness of the GYY programme compared with usual care based on the time horizon of the trial (i.e. 12 months)
2. a cost-consequences analysis to present disaggregated costs alongside the full range of outcomes, over a 12-month time horizon.

Further details of both analyses are provided below and were summarised in the trial protocol and health economics analysis plan.

### *Within-trial economic evaluation*

A within-trial economic evaluation was undertaken using a cost-utility analysis to evaluate the intervention relative to usual care in terms of the incremental cost per QALY. The perspective of the UK NHS and personal social services (PSS) was taken for the evaluation on an ITT basis. Analyses used a 12-month time horizon; therefore discounting of health benefits and costs was not necessary. Multiple imputation was undertaken in the base-case analysis, with a complete-case analysis (CCA) explored also as a SA. Further sensitivity analyses were undertaken to investigate the effect of varying assumptions made in the analysis in terms of the cost-effectiveness findings. Personal costs paid by participants regarding private health care were also explored as a SA. All analyses were conducted in line with NICE recommendations as far as possible<sup>35</sup> and taking into consideration the reporting requirements of the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) [www.ispor.org/TaskForces/EconomicPubGuidelines.asp](http://www.ispor.org/TaskForces/EconomicPubGuidelines.asp).

### *Economic data collection*

Health outcome and healthcare resource use data were collected during the 12-month study follow-up period. Specifically, healthcare resource use and EQ-5D-5L data were obtained at baseline, 3, 6 and 12 months. Medication prescription data were collected via GP practices for a sample of participants over a period of 3 months prior to randomisation and for the 12 months following randomisation. Information regarding the intervention cost was obtained from the study team (i.e. invoices, etc.).

### *Health-related quality of life*

The health-related quality-of-life utility data for the economic evaluation comprised the EQ-5D-5L, as described previously in [Chapter 2](#), which was used to elicit patient utility index scores at the different time points in order to estimate QALYs. QALYs are able to combine the effect of a treatment on both the quantity and quality of life for an individual, where one QALY is equivalent to a year spent in full health. Participants' responses to the EQ-5D-5L questions were used to derive utilities, with the EQ-5D-5L scored according to the User Guide<sup>36</sup> and the mapping function developed by van Hout *et al.*<sup>23</sup> used to calculate utility index scores, as recommended by NICE.<sup>35</sup> QALYs were estimated for each participant based on utility index scores at the different time points (i.e. at baseline, 3, 6 and 12 months) using the

area under the curve (AUC) approach; this assumes linear interpolation between measurement points, following the trapezium rule.<sup>37</sup> In the primary analysis, the MD in QALYs gained over the 12-month period between the groups was adjusted for baseline utility only. An additional SA was further adjusted for baseline utility, age, gender and study site.

The EQ-5D-5L utility index scores are provided for the two groups, with movements between levels examined for each domain by trial arm. Descriptive statistics of the utilities for both groups and the difference in utilities between the groups at each time point have been presented. The overall difference in EQ-5D-5L utility index scores between the two groups has been examined through regression methods, consistent with the statistical analysis.

### **Healthcare resource use**

Healthcare resource use data were collected for each participant, specifically resource use within primary care and the community (i.e. appointments with a GP, nurse, physiotherapist, mental health services) and the hospital setting [i.e. hospital outpatient attendances, day-case visits, inpatient stays and accident and emergency attendances]. Information on private treatments was also collected (i.e. private GP visits, physiotherapists or mental health service specialists) to feed into a SA. The resource use data were collected during the trial via self-reported participant questionnaires at baseline, 3, 6 and 12 months. Healthcare resource use results are presented for both arms in terms of mean value, SD and MD (with 95% CIs) between the groups.

### **Costs**

Unit costs for the resources used were obtained from established costing sources such as NHS Reference Costs<sup>38</sup> and Personal Social Services Research Unit (PSSRU) Unit Costs of Health and Social Care.<sup>39</sup> No data were collected as to how privately funded health care was paid for, for example, private plan, employer plan or pay as you go. For the purposes of costing, we assumed participants paid out of pocket on a pay-as-you-go basis. Private healthcare costs were sourced from published online sources. The unit cost for all items used in the analysis is summarised in [Table 26](#). All costs were evaluated in Great British pounds (£) for 2020–1. In cases where costs were sourced from previously published data, costs were inflated to the appropriate year figures. The mean cost per participant for each resource use item is provided by trial arm, with costs calculated by multiplying the resources used by their corresponding unit costs. For participants' responses to accident and emergency visits that involved overnight stays, the corresponding cost was calculated by multiplying the number of visits to accident and emergency by the cost of a visit to accident and emergency plus multiplying the number of nights they spent in hospital (as part of this visit) by the cost of an inpatient night in hospital.

Medications data were collected for a sample of participants across three sites using prescription data collected from GP practices. This sample comprised all participants in the first pilot phase who remained within the trial at 12 months post randomisation. These data were used to generate an estimate of the average medication cost per participant; all prescribed medications were recorded for the 3-month period prior to randomisation and for the 12-month period post randomisation. For each participant in the sample, information was collected regarding the medication name, dose, quantity taken, frequency (i.e. once daily, twice weekly, etc.), number of days taken and whether it was a repeat prescription. Unit costs were attached to the medications using costs sourced from the British National Formulary (BNF)<sup>40</sup> and NHS Drug Tariff.<sup>41</sup> Costs were applied using the drug that matched as closely as was available from BNF, and where there was a discrepancy between the information reported in the frequency and quantity, the more conservative option was selected. For participants who had not had medication data collected for them (i.e. they had missing medication data), the average medication cost for the relevant group (i.e. GYY or usual care) was applied.

The total cost for each participant comprised the following: healthcare resource use, medication cost and intervention cost (for those who received the intervention). Additional information regarding the GYY intervention is provided in the 'Costing the intervention' section below.

TABLE 26 Unit costs of healthcare services

Item	Unit cost	Notes	Source
GP visit at GP practice	£39.23	Per consultation of 9.22 minutes duration (including direct care staff costs and qualifications)	PSSRU 2021 <sup>39</sup>
GP visit at home	£100.62	Per home visit, comprising 11.4-minute consultation and 12 minutes travel time (PSSRU 2015); £4.30 per minute of patient contact (PSSRU 2021)	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
GP consultation over the phone	£30.53	Per phone contact of 7.1 minutes duration (PSSRU 2015); at £4.30 per minute of patient contact (PSSRU 2021)	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
Nurse visit at GP practice	£11.37	Per consultation lasting 15.5 minutes (PSSRU 2015); at £44 per hour (PSSRU 2021)	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
Nurse visit at home	£27.13	Per home visit, comprising 25-minute consultation (PSSRU 2014) and 12 minutes travel time (PSSRU 2015); at £44 per hour (PSSRU 2021)	PSSRU 2021, <sup>39</sup> 2015 <sup>42</sup> and 2014 <sup>43</sup>
Nurse consultation over the phone	£4.40	Per phone contact of 6 minutes duration (PSSRU 2015); at £44 per hour (PSSRU 2021)	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
District nurse visit at GP practice	£11.37	In the absence of a recent cost, these were assumed equal to the cost of a standard nurse visit at the practice	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
District nurse visit at home	£27.13	In the absence of a recent cost, these were assumed equal to the cost of a standard nurse visit at home	PSSRU 2021, <sup>39</sup> 2015 <sup>42</sup> and 2014 <sup>43</sup>
District nurse consultation over the phone	£4.40	In the absence of a recent cost, these were assumed equal to the cost of a standard nurse consultation over the phone	PSSRU 2021 <sup>39</sup> and 2015 <sup>42</sup>
Physiotherapist	£58.75	Per appointment, of 1 hour duration. Average cost per working hour for physiotherapists in bands 5–8a	PSSRU 2021 <sup>39</sup>
Mental health services	£285.70	Per appointment	NHS Ref Costs 20/21 <sup>38</sup>
Physiotherapist (hospital)	£56.50	Per appointment, of 1 hour duration. Average cost per working hour for physiotherapists in bands 5–8a	PSSRU 2021 <sup>39</sup>
Other hospital outpatient appointment	£181.54	Per outpatient attendance	NHS Ref Costs 20/21 <sup>38</sup>
Day-case appointment	£1191.96	Per (non-specific) attendance	NHS Ref Costs 20/21 <sup>38</sup>
Accident and emergency visit	£296.87	Per (non-specific) attendance	NHS Ref Costs 20/21 <sup>38</sup>
Inpatient night in hospital	£372.70	Total expenditure on excess bed days (elective and non-elective) divided by total activity, inflated to 2021 prices	NHS Ref Costs 17/18 <sup>44</sup>
Private GP appointment	£79	Per appointment	Bupa <sup>45</sup>
Private physiotherapist appointment	£63.94	Average of £75.07 for first appointment and £52.81 for follow-up appointment	myTribe Health Insurance <sup>46</sup>
Private mental health services specialist visit	£360	Per appointment	Psychiatry-UK <sup>47</sup>

### Costing the intervention

The cost of the intervention comprised (1) the cost of training the yoga teachers to undertake the GYY classes and (2) the cost of running the GYY classes, which included the cost of providing participants with the equipment and home practice sheets provided as part of the intervention. The intervention cost is described further below.

### Cost of training the yoga teachers

The cost of training involved the cost of the time spent by the yoga teachers being trained (i.e. the trainees) and the cost of the trainers' time, plus room hire costs (for face-to-face training). Information regarding the time spent on the training workshops and the costs claimed by trainees was obtained by the trial team. The training sessions lasted for 5 hours, at a cost of £17 per hour for trainees. The cost of trainers' time incorporated the cost of the yoga consultants and university staff who undertook the training, based on university pay scales and invoices submitted (ranging from £19 to £100 per hour); the trainers' costs were based on two yoga consultants and two trial coordinators being present for 5 hours, and the CI and qualitative researcher being present for 2 hours.

The corresponding unit costs were applied for the time spent by trainers and trainees present at the workshops. Although the training workshops comprised a combination of online and face-to-face training workshops, the costs of the trainer and trainee time were the same for both modes of delivery; the only difference in cost for the face-to-face training arose from room hire costs also being included. The total cost of training that took place over the study period was divided by the number of participants in the intervention group.

### Cost of running the Gentle Years Yoga classes

The cost of running the course of GYY classes as part of the intervention was estimated using information obtained from the trial team. Some courses were delivered face to face, while others were delivered online; hence the relevant costs have been applied accordingly. Costs common to both modes of delivery of the classes were a yoga class payment provided to yoga teachers of £90 per class (i.e. £1080 per 12-week course) and a £279 administration cost per 12-week course.

Costs specific to the face-to-face classes only were costs of refreshments (£10 per class, i.e. £120 per 12-week course), travel costs for yoga teachers (£0.45 per mile, up to a maximum of 60 miles per class) and venue costs. The venue costs ranged from £15 per class to £49 per class, generating an average venue cost of £26 per class (i.e. £311 per 12-week course). Costs incurred for online classes only were a video-conferencing subscription cost per yoga teacher running the online classes (£12 per month for Zoom pro) and the cost for liaising with participants prior to the course starting (£135 by phone, or £270 by Zoom, per course).

The following equipment was provided to participants: tennis ball, band, scarf, beanbag and equipment bag (£4.40), plus block (£4) at a total of £8.80 per participant (source: trial team). For participants who undertook the course of GYY classes online, the equipment was delivered by post; hence a postage and packing cost was added to generate a total cost of £14.97 for these participants. The face-to-face classes involved the yoga teacher bringing the equipment to the class; hence no postage cost was incurred. The relevant costs were applied according to whether participants received classes face to face or online, to generate a total equipment cost for the whole intervention group. This cost was then divided by the total number of participants in the intervention group to estimate an average equipment cost per participant. In addition, home practice sheets were also provided to participants as part of the intervention, with a corresponding unit cost of £0.82 for printing the nine-page document (source: trial team).

### Missing data

Missing data are likely to arise in trial-based economic analyses involving participant-level data. This is due to the summative nature of the data; the total cost is comprised of several cost items over different time points, whereby if one item is missing, the total cost is then considered missing. Similarly, QALYs can be calculated only when all EQ-5D-5L questionnaires have been completed fully at all time points. The use of CCA, that is, including only participants with complete data on costs and EQ-5D-5L, can introduce bias unless the data are missing completely at random. An approach used to deal with missing data in economic evaluations in clinical trials is multiple imputation<sup>48</sup> which estimates plausible, unbiased values for the missing data. This is based on the assumption of the data being MAR, that is, the observed

data can be used to predict the missing values, but the missingness of the data is not associated with its value. We explored the validity of this assumption through the exploration of missing economic data at all follow-up points for both study groups and missing data patterns. Logistic regression was used to investigate which factors were associated with the missingness of data. Our findings indicated the MAR assumption to be reasonable; hence, multiple imputation with chained equations was used for the base-case analysis, with predicted mean matching on QALY and cost estimates, thereby generating plausible values from the imputation.

The imputation model included age, gender, study site, utilities (at baseline, 3, 6 and 12 months) and total costs at the resource use level. The base-case analysis undertook 60 imputations due to the finding of 58% missing data and following guidance that the number of imputed data sets should be similar to the proportion of incomplete cases.<sup>49</sup> All sensitivity analyses also followed this guidance. Graphical plots compared the distributions of the observed and imputed data to check whether the imputed data resembled the observed data. The imputed data sets were combined using Rubin's rules<sup>50</sup> to generate mean QALY and cost estimates. Usual imputation methods were followed for participants who died during the trial, where they had any missing data prior to their death. For data that would have been obtained after their death, zero QALYs and costs were assumed.

The base-case analysis used the multiply imputed data set with a CCA (i.e. includes only participants with observed data for all utilities and costs) undertaken as a SA. This enabled an alternative missing data assumption to be explored, as CCA relies on the data being not MAR. An initial exploration of the data was conducted using available case analysis.

### Incremental analysis

Costs and QALYs are presented for the two groups in terms of mean value and SD and MD (with 95% CIs) between the groups. Findings are presented using ICERs, where appropriate, with standard decision rules followed when evaluating the two options. The incremental cost per QALY gained was determined based on the following:

$$\text{ICER} = \frac{\text{Cost (Intervention)} - \text{Cost (Comparator)}}{\text{Effect (Intervention)} - \text{Effect (Comparator)}} = \frac{\Delta C}{\Delta E} \Delta \quad (1)$$

where incremental costs and incremental effects are shown by  $\Delta C$  and  $\Delta E$ , respectively.

The ICER is compared to a willingness-to-pay (WTP) threshold, that is, the amount that a decision-maker is willing to pay for an additional QALY. The threshold of £20,000, following NICE recommendations, has been used for the analyses;<sup>35</sup> where interventions have an ICER below £20,000, they would generally be deemed to be cost-effective. The results are also presented in terms of net monetary benefit (NMB),<sup>51</sup> thereby representing the intervention in monetary terms based on a WTP threshold  $\lambda$ . Where the mean incremental NMB (i.e. the difference in NMB between the intervention and usual care) is positive, the intervention is found to be cost-effective at the given threshold.

$$\text{NMB} = (\lambda \Delta E) - \Delta C$$

The cost-utility analysis produced estimates of the MDs in costs and effects using seemingly unrelated regression equations, with 95% CIs estimated using bias-corrected and accelerated (BCA) bootstrap methods. Bootstrapping was conducted for 10,000 replications, producing 10,000 estimates of incremental costs and incremental effects. The 10,000 point estimates generated from the analyses are plotted as a scatterplot on the cost-effectiveness plane, across four quadrants, as summarised below:

1. North-east quadrant: the intervention is more costly and more effective than usual care; an ICER is estimated to evaluate whether the incremental effects (i.e. QALYs here) are worth the incremental costs.



2. South-east quadrant: the intervention is cost saving and more effective than usual care; it therefore dominates usual care.
3. South-west quadrant: the intervention is cost saving but less effective than usual care; an ICER is estimated to determine the cost per QALY lost.
4. North-west quadrant: the intervention is more costly and less effective than usual care; it is therefore dominated by usual care.

Cost-effectiveness acceptability curves (CEACs) were used to investigate the probability of the intervention being cost-effective<sup>52</sup> compared with usual care at different WTP threshold possibilities faced by decision-makers.

Regression methods have been used to account for differences in stratification or prognostic variables and other sources of heterogeneity. QALY data were adjusted for baseline EQ-5D-5L scores, age, gender and trial site. Differences between the groups were found to be statistically significant if  $p < 0.05$  and are presented alongside CIs around the differences in costs and outcomes. Analyses were undertaken using Stata release 17.<sup>34</sup>

### **Sensitivity analysis**

The uncertainty around the cost-effectiveness findings was explored by sensitivity analyses to investigate the robustness of the findings to assumptions made in the base-case analysis. Specifically, the following sensitivity analyses have been undertaken:

1. CCA, as an alternative to the use of multiple imputation which dealt with missing data
2. inclusion of personal expenses for private health care
3. face-to-face yoga classes only (i.e. no online classes included)
4. online yoga classes only (i.e. no face-to-face classes included)
5. medication costs removed
6. age and gender removed as covariates.

### **Cost-consequences analysis**

A cost-consequences analysis was also undertaken to present disaggregated costs alongside the full range of outcomes. The analysis summarises resource use captured during the trial and the costs associated with these for each of the trial arms, as well as the consequences of the trial. As per the cost-utility analysis, the cost-consequence analysis was conducted on an ITT basis, following the 12-month time horizon for the trial and using Stata release 17.<sup>34</sup>

For each resource use item, the mean resource use was calculated according to the appropriate unit (e.g. mean number of visits/appointments). To calculate the incremental difference for each individual item of resource use between the two trial arms, the values for participants allocated to the usual care group were subtracted from the values of participants allocated to the intervention (GY) group.<sup>53</sup> Mean unit costs were calculated by resource item for each of the trial arms by multiplying the resource by an appropriate unit cost, with resource use costs summed by type, for example, primary care. Multiple imputation was not undertaken for the purposes of the cost-consequence analysis, as the model serves only to highlight the likely resource use and the average costs per resource, alongside the consequences (outcomes) of the trial. The consequences presented in [Table 35](#) reflect the outcomes of the GY trial, and full details as to how these were calculated are presented in [Chapter 2](#).

## **Results**

### **Participant population and missing data**

As previously reported, overall, seven participants died during the study period: 2/240 (1%) in the intervention group and 5/214 (2%) in the usual care group. A CCA was undertaken (SA1) that included participants with complete data on all EQ-5D-5L and cost items over the 12-month study period. In

total, 192 (42.3%) participants had complete economic data: 105 in the intervention group and 87 in the usual care group. Therefore, overall, 262 (57.7%) participants had missing economic data. Details of the completion of data at the data collection points in the trial can be seen in [Appendix 3](#) (see [Table 46](#)). At baseline, 72.9% participants in the intervention group and 77.1% in the usual care group had complete data, which decreased at 3 months and further by 6 months for the two groups. Completion levels increased by 12 months, however, to 70.8% in the intervention group and 67.8% in the usual care group (see [Appendix 6, Table 49](#)). The pattern of missing data for both cost and EQ-5D-5L data was not monotonic, that is, there were some participants who had missing data at 6 months but who had observed data at 12 months. The odds ratios (ORs) from a logistic regression for indicators of missing data, for QALYs and costs, on study group allocation and several baseline variables are presented in [Appendix 4](#) (see [Table 47](#)). Age, site and allocation were found to be significantly associated with both missing cost data at 12 months and missing QALY data at 12 months. In addition, gender was significantly associated with missing QALY data at 12 months. When investigating observed values, observed QALYs at 3 and 6 months were found to be significant predictors of missing QALY data at 12 months and cost data at 12 months. Our findings therefore support the assumption of the data being MAR rather than missing completely at random. There is no way of definitively distinguishing whether the data are MAR rather than missing not at random (MNAR), so the best we can do is assess how *reasonable* the assumption is based on how likely it is that people would not want to provide data depending on what their response was. Data that are particularly sensitive or personal or that reflect an unfavourable opinion may be more likely to be MNAR, but arguably resource use data are not particularly sensitive. Missingness was therefore assumed to depend on baseline covariates (EQ-5D-5L at baseline, age and gender) and observed QALYs and costs. The fit of the imputation model was checked by comparing the distributions of the observed data with the imputed data for costs and utilities.

### Health-related quality-of-life utility

The EQ-5D-5L was very well completed at baseline, with completion rates being 100% for intervention participants and 99.5% for usual care participants (only one participant missed one item) (see [Appendix 5, Table 48](#)). Completion rates decreased over time for the two groups, though they remained high at 12 months, at 87.9% and 83.2% for the intervention and usual care groups, respectively. Further information on the missing dimensions and completion of the EQ-5D-5L questionnaires is provided in [Appendix 6](#) (see [Table 49](#)). A breakdown of the different EQ-5D-5L levels reported by participants over the 12-month follow-up period, according to dimension, group and time point is shown in [Appendix 7](#) (see [Table 50](#)), for all available cases (i.e. not imputed). Mean (SD) VAS scores were similar at baseline, though slightly higher in the intervention group: 75.1 (18.2) for the intervention group versus 73.4 (17.6) in the usual care group. As presented previously (see [Chapter 3, Table 21](#)), mean VAS scores increased at 3 months, then decreased over time for both groups. The mean EQ-5D-5L utility index scores are presented in [Table 27](#) for all time points, based on the available cases, alongside the MD between the

**TABLE 27** Summary of EQ-5D-5L utility index scores and unadjusted and adjusted MD (95% CI) at each time point (all available cases)

Utility	Intervention (n = 240)		Usual care (n = 214)		Unadjusted MD (95% CI) (intervention – usual care)	Adjusted MD (95% CI) <sup>a</sup> (intervention – usual care)
	Follow-up	N	Mean score (SD)	N		
Baseline	240	0.742 (0.176)	213	0.736 (0.162)	0.006 (–0.025 to 0.038) p = 0.684	
3 months	224	0.749 (0.168)	190	0.723 (0.201)	0.026 (–0.010 to 0.062) p = 0.153	0.018 (–0.006 to 0.043) p = 0.145
6 months	218	0.732 (0.207)	180	0.706 (0.219)	0.026 (–0.016 to 0.068) p = 0.227	0.021 (–0.012 to 0.053) p = 0.215
12 months	213	0.723 (0.210)	182	0.689 (0.219)	0.035 (–0.008 to 0.077) p = 0.109	0.024 (–0.010 to 0.059) p = 0.166

a The difference at 3, 6 and 12 months is adjusted for baseline utility.



groups (both unadjusted and adjusted for baseline utility). Utilities were slightly higher at baseline for participants in the intervention group (0.742) compared with those in the usual care group (0.736). The MD was not found to be statistically significant at any time point, irrespective of whether it was adjusted for baseline utility or not. The total QALYs gained over the duration of the trial are shown in [Table 28](#) for all available cases, indicating a small gain in QALYs for the intervention group over the 12-month period. This increase in QALYs was not found to be statistically significant based on the non-imputed (all available cases) data set when adjusting for baseline utility only or when adjusting for baseline utility, age, gender and study site.

### Healthcare resource use and costs

The healthcare resource use and associated costs for all available cases are summarised in [Tables 29](#) and [30](#), respectively, according to item of resource and trial arm.

In terms of resource use outside of the hospital, although participants in the intervention group had, on average, fewer nurse visits at the GP practice, they had, on average, additional GP visits at the GP practice/home/phone, nurse visits by phone/home, physiotherapist visits (community-based) and mental health services (community-based). For hospital-based resource use, intervention group participants had, on average, fewer physiotherapist visits (hospital-based), hospital outpatient visits, inpatient nights in hospital and accident and emergency visits, though they had, on average, more day-case hospital visits and attendances with mental health services (hospital-based) than usual care participants over the study period.

**TABLE 28** Total QALYs for all available cases by allocation over 12-month trial duration and difference in mean QALYs (95% CI)

Trial arm	Total	Mean (SD) Total QALYs	Difference (95% CI) <sup>a</sup> (intervention – usual care)	Difference (95% CI) <sup>b</sup> (intervention – usual care)
Intervention	209	0.742 (0.173)	0.020 (–0.004 to 0.044)	0.021 (–0.004 to 0.045)
Usual care	171	0.710 (0.190)	<i>p</i> = 0.101	<i>p</i> = 0.098

a Adjusted for baseline utility.

b Adjusted for all covariates (baseline utility, gender, age, study site); estimated using ordinary least squares.

**TABLE 29** Mean resource use at each study time point, based on all available cases

Type of resource use (months)	Intervention (N = 240)		Usual care (N = 214)	
	Mean (SD)		Mean (SD)	MD (95% CI) (intervention – usual care)
<b>GP visit at GP practice</b>				
Baseline	0.60 (1.03) N = 236		0.56 (1.00) N = 212	0.045 (–0.144 to 0.235)
3	0.48 (0.97) N = 220		0.48 (0.97) N = 180	0.008 (–0.183 to 0.199)
6	0.41 (0.93) N = 213		0.41 (0.78) N = 176	0.004 (–0.169 to 0.177)
12	0.69 (1.19) N = 208		0.63 (1.07) N = 175	0.053 (–0.175 to 0.282)
<b>GP visit at home</b>				
Baseline	0.017 (0.20) N = 239		0.005 (0.07) N = 212	0.012 (–0.017 to 0.041)
3	0.02 (0.16) N = 222		0 (0) N = 181	0.018 (–0.006 to 0.042)
6	0.005 (0.07) N = 171		0.005 (0.08) N = 208	–0.001 (–0.016 to 0.014)
12	0.005 (0.07) N = 172		0.01 (0.11) N = 208	0.007 (–0.025 to 0.011)

TABLE 29 Mean resource use at 12 months, based on all available cases (continued)

Type of resource use (months)	Intervention (N = 240)	Usual care (N = 214)	MD (95% CI) (intervention – usual care)
	Mean (SD)	Mean (SD)	
<b>GP consultation over the phone</b>			
Baseline	0.73 (1.05) N = 237	0.64 (0.97) N = 211	0.091 (–0.098 to 0.280)
3	0.64 (0.99) N = 214	0.68 (1.08) N = 180	–0.038 (–0.243 to 0.166)
6	0.69 (0.99) N = 204	0.62 (0.88) N = 176	0.072 (–0.119 to 0.262)
12	1.39 (1.63) N = 194	1.13 (1.38) N = 167	0.260 (–0.055 to 0.576)
<b>Nurse visit at GP practice</b>			
Baseline	0.64 (1.22) N = 237	0.71 (1.69) N = 208	–0.074 (–0.347 to 0.199)
3	0.57 (0.97) N = 214	0.80 (1.36) N = 183	–0.229 (–0.460 to 0.003)
6	0.47 (0.88) N = 212	0.57 (0.92) N = 175	–0.099 (–0.279 to 0.081)
12	0.90 (1.40) N = 199	1.01 (1.76) N = 171	–0.107 (–0.431 to 0.216)
<b>Nurse visit at home</b>			
Baseline	0.03 (0.40) N = 239	0.01 (0.15) N = 212	0.019 (–0.038 to 0.076)
3	0.04 (0.31) N = 221	0.01 (0.07) N = 181	0.031 (–0.016 to 0.078)
6	0.01 (0.15) N = 211	0.01 (0.11) N = 173	0.003 (–0.025 to 0.030)
12	0.12 (1.67) N = 207	0.01 (0.08) N = 175	0.110 (–0.138 to 0.358)
<b>Nurse consultation over the phone</b>			
Baseline	0.17 (0.46) N = 179	0.25 (0.63) N = 167	–0.078 (–0.193 to 0.037)
3	0.17 (0.51) N = 167	0.21 (0.69) N = 141	–0.038 (–0.173 to 0.097)
6	0.18 (0.52) N = 163	0.26 (0.69) N = 137	–0.079 (–0.217 to 0.059)
12	0.34 (0.77) N = 200	0.28 (0.63) N = 173	0.058 (–0.087 to 0.203)
<b>District nurse visit at GP practice</b>			
Baseline	0.004 (0.06) N = 238	0.02 (0.21) N = 211	–0.019 (–0.047 to 0.008)
3	0.03 (0.25) N = 222	0 (0) N = 184	0.027 (–0.009 to 0.063)
6	0 (0) N = 212	0 (0) N = 173	0 (0 to 0)
12	0.04 (0.26) N = 207	0.02 (0.18) N = 178	0.016 (–0.029 to 0.062)
<b>District nurse visit at home</b>			
Baseline	0.09 (1.30) N = 239	0.07 (0.96) N = 212	0.017 (–0.196 to 0.231)
3	0.01 (0.09) N = 221	0.06 (0.43) N = 184	–0.051 (–0.110 to 0.008)
6	0.005 (0.07) N = 211	0.02 (0.17) N = 173	–0.013 (–0.038 to 0.013)
12	0.12 (1.67) N = 208	0.07 (0.53) N = 178	0.047 (–0.209 to 0.303)
<b>District nurse consultation over the phone</b>			
Baseline	0.01 (0.11) N = 180	0.006 (0.08) N = 168	0.005 (–0.014 to 0.025)
3	0.02 (0.24) N = 171	0 (0) N = 143	0.023 (–0.016 to 0.063)
6	0 (0) N = 164	0.007 (0.09) N = 138	0.007 (–0.020 to 0.006)
12	0.01 (0.12) N = 208	0.02 (0.13) N = 177	0.003 (–0.028 to 0.022)

continued

TABLE 29 Mean resource use at 12 months, based on all available cases (continued)

Type of resource use (months)	Intervention (N = 240)	Usual care (N = 214)	MD (95% CI) (intervention – usual care)
	Mean (SD)	Mean (SD)	
<b>Physiotherapist visit (community)</b>			
Baseline	0.15 (0.57) N = 239	0.16 (0.72) N = 212	-0.014 (-0.134 to 0.106)
3	0.25 (1.15) N = 223	0.11 (0.56) N = 183	0.136 (-0.046 to 0.318)
6	0.11 (0.89) N = 212	0.13 (0.61) N = 171	-0.021 (-0.178 to 0.136)
12	0.33 (1.67) N = 207	0.30 (1.04) N = 179	0.032 (-0.252 to 0.315)
<b>Mental health services (community)</b>			
Baseline	0.05 (0.31) N = 239	0 (0) N = 212	0.050 (0.008 to 0.093)
3	0.06 (0.42) N = 222	0.02 (0.23) N = 184	0.041 (-0.027 to 0.110)
6	0.06 (0.64) N = 211	0.02 (0.17) N = 173	0.044 (-0.054 to 0.143)
12	0.12 (0.84) N = 208	0.02 (0.17) N = 178	0.099 (-0.027 to 0.224)
<b>Physiotherapist visit (hospital)</b>			
Baseline	0.17 (0.92) N = 238	0.17 (1.11) N = 212	-0.002 (-0.190 to 0.187)
3	0.06 (0.36) N = 221	0.19 (0.94) N = 186	-0.130 (-0.265 to 0.005)
6	0.09 (0.54) N = 215	0.05 (0.31) N = 175	0.042 (-0.049 to 0.132)
12	0.13 (0.67) N = 209	0.17 (0.65) N = 176	-0.036 (-0.169 to 0.096)
<b>Mental health services (hospital)</b>			
Baseline	0.02 (0.16) N = 239	0 (0) N = 213	0.017 (-0.005 to 0.038)
3	0.01 (0.12) N = 221	0 (0) N = 186	0.014 (-0.003 to 0.030)
6	0.02 (0.17) N = 213	0 (0) N = 176	0.019 (-0.006 to 0.044)
12	0.01 (0.12) N = 210	0 (0) N = 176	0.014 (-0.003 to 0.032)
<b>Other hospital outpatient appointment</b>			
Baseline	0.72 (1.40) N = 239	0.88 (1.38) N = 210	-0.157 (-0.415 to 0.101)
3	0.69 (1.26) N = 218	0.88 (1.29) N = 188	-0.195 (-0.445 to 0.055)
6	0.62 (1.18) N = 214	0.65 (0.99) N = 176	-0.022 (-0.241 to 0.198)
12	1.29 (2.00) N = 206	1.28 (1.84) N = 176	0.008 (-0.380 to 0.396)
<b>Day-case appointment</b>			
Baseline	0.09 (0.39) N = 237	0.07 (0.36) N = 212	0.022 (-0.048 to 0.092)
3	0.08 (0.39) N = 218	0.09 (0.37) N = 186	0.013 (-0.089 to 0.062)
6	0.13 (0.51) N = 212	0.08 (0.41) N = 176	0.048 (-0.046 to 0.142)
12	0.17 (0.62) N = 208	0.13 (0.41) N = 176	0.042 (-0.066 to 0.150)
<b>Accident and emergency visit (not overnight)</b>			
Baseline	0.07 (0.27) N = 180	0.08 (0.46) N = 168	-0.017 (-0.095 to 0.062)
3	0.07 (0.56) N = 170	0.08 (0.31) N = 146	-0.005 (-0.108 to 0.098)
6	0.36 (0.22) N = 167	0.04 (0.19) N = 140	0.0002 (-0.046 to 0.046)
12	0.19 (1.10) N = 211	0.09 (0.31) N = 178	0.104 (-0.063 to 0.272)

TABLE 29 Mean resource use at 12 months, based on all available cases (continued)

Type of resource use (months)	Intervention (N = 240)	Usual care (N = 214)	MD (95% CI) (intervention – usual care)
	Mean (SD)	Mean (SD)	
<b>Accident and emergency visit (involving overnight stay)</b>			
Baseline	0.06 (0.29) N = 238	0.04 (0.27) N = 213	0.025 (–0.027 to 0.078)
3	0.05 (0.27) N = 221	0.03 (0.23) N = 186	0.013 (–0.036 to 0.062)
6	0.02 (0.14) N = 212	0.03 (0.17) N = 175	–0.010 (–0.040 to 0.021)
12	0.08 (0.41) N = 211	0.03 (0.17) N = 175	0.047 (–0.017 to 0.112)
<b>Accident and emergency nights spent in hospital</b>			
Baseline	0.08 (0.63) N = 235	0.29 (2.26) N = 213	0.206 (–0.508 to 0.097)
3	0.04 (0.31) N = 219	0.03 (0.33) N = 184	0.008 (–0.054 to 0.071)
6	0.05 (0.48) N = 212	0.20 (1.89) N = 175	–0.148 (–0.413 to 0.117)
12	0.23 (1.59) N = 210	0.16 (1.01) N = 176	0.069 (–0.205 to 0.342)
<b>Inpatient nights in hospital</b>			
Baseline	0.11 (0.69) N = 239	0.01 (0.15) N = 213	0.095 (–0.0002 to 0.1896)
3	0.06 (0.48) N = 223	0.09 (1.05) N = 186	–0.029 (–0.183 to 0.126)
6	0.02 (0.10) N = 211	0.13 (1.11) N = 176	–0.116 (–0.267 to 0.036)
12	0.02 (0.17) N = 209	0.30 (2.44) N = 178	–0.284 (–0.616 to 0.048)
<b>Private GP appointment</b>			
Baseline	0.03 (0.25) N = 239	0.02 (0.14) N = 213	0.011 (–0.027 to 0.048)
3	0.03 (0.21) N = 219	0.05 (0.26) N = 183	–0.022 (–0.068 to 0.025)
6	0.04 (0.26) N = 210	0.07 (0.46) N = 174	–0.037 (–0.109 to 0.036)
12	0.03 (0.22) N = 208	0.08 (0.54) N = 177	–0.050 (–0.130 to 0.030)
<b>Private physiotherapist appointment</b>			
Baseline	0.46 (2.01) N = 239	0.33 (1.40) N = 213	0.127 (–0.197 to 0.451)
3	0.48 (1.68) N = 223	0.26 (1.27) N = 183	0.213 (–0.082 to 0.508)
6	0.83 (1.27) N = 209	0.30 (0.96) N = 173	0.015 (–0.215 to 0.246)
12	0.70 (3.11) N = 209	0.55 (1.66) N = 177	0.145 (–0.367 to 0.657)
<b>Private mental health services specialist visit</b>			
Baseline	(0.13) N = 239	0 (0) N = 213	0.008 (–0.009 to 0.026)
3	0 (0) N = 216	0.02 (0.17) N = 183	–0.016 (–0.038 to 0.006)
6	0.05 (0.54) N = 207	0 (0) N = 173	0.053 (–0.028 to 0.134)
12	0.05 (0.69) N = 208	0 (0) N = 176	0.048 (–0.054 to 0.151)

Hospital-based services tended to be the major cost drivers, along with the cost of medications and the intervention, for participants in the intervention group. Based on the multiply imputed data set (see [Table 31](#)), the largest (mean; 95% CI) reductions in costs were seen for the intervention group versus usual care group for inpatient nights spent in hospital (–£139.55; –£80.22 to –£98.89), accident and emergency visits (–£25.89; –£73.32 to £21.54) and hospital outpatient attendances (–£11.30; –£46.32 to £23.73). In addition, medication costs were lower for the intervention group (–£68.90; –£77.19

to -£60.62). The most notable increases in costs for participants in the intervention group over the 12-month period were found for day-case visits (£85.03; £15.22 to £154.85) and mental health services (£32.88; £19.66 to £46.11).

**TABLE 30** Total mean costs based on all available cases, up to 12-month follow-up

Cost item	Total mean cost £ (SD)		
	Intervention (N = 240)	Usual care (N = 214)	MD (95% CI) (intervention - usual care)
GP visit at GP practice	62.13 (82.93) N = 197	56.41 (75.68) N = 153	5.72 (-11.20 to 22.64)
GP visit at home	3.18 (22.92) N = 190	2.05 (14.28) N = 147	1.12 (-3.12 to 5.37)
GP phone consultation	84.92 (85.72) N = 174	73.23 (73.65) N = 148	11.69 (-6.00 to 29.38)
Nurse visit at GP practice	21.69 (27.05) N = 184	25.49 (34.85) N = 153	-3.80 (-10.44 to 2.84)
Nurse visit at home	4.89 (47.78) N = 194	0.72 (5.40) N = 150	4.17 (-3.54 to 11.89)
Nurse phone consultation	2.97 (5.72) N = 145	2.91 (6.15) N = 118	0.07 (-1.38 to 1.51)
District nurse visit at GP practice	0.70 (4.25) N = 196	0.29 (2.22) N = 155	0.40 (-0.34 to 1.14)
District nurse visit at home	3.88 (46.63) N = 196	4.55 (22.89) N = 155	-0.68 (-8.71 to 7.36)
District nurse phone consultation	0.17 (1.22) N = 153	0.07 (0.56) N = 123	0.10 (-0.13 to 0.34)
Physiotherapist (community)	36.98 (171.08) N = 197	33.41 (102.59) N = 153	3.57 (-27.20 to 34.35)
Mental health services (community)	62.68 (380.33) N = 196	16.59 (109.15) N = 155	46.09 (-15.93 to 108.11)
Physiotherapist (hospital)	16.39 (58.75) N = 200	18.59 (62.30) N = 158	-2.21 (-14.84 to 10.42)
Mental health services (hospital)	14.36 (89.66) N = 199	0 (0) N = 158	14.36 (0.33 to 28.39)
Other hospital outpatient appointment	474.44 (660.91) N = 194	498.39 (556.69) N = 161	-23.95 (-153.08 to 105.18)
Day-case appointment	442.38 (1469.34) N = 194	359.84 (897.16) N = 159	82.54 (-179.35 to 344.43)
Accident and emergency visit (not overnight)	75.16 (227.10) N = 158	63.11 (161.44) N = 127	12.04 (-35.00 to 59.09)
Accident and emergency visits (including overnight hospital stays)	108.62 (607.51) N = 194	122.25 (786.00) N = 156	-13.62 (-160.12 to 132.87)
Inpatient nights in hospital	32.00 (190.63) N = 198	191.01 (1101.16) N = 160	-159.01 (-315.69 to -2.33)

**TABLE 30** Total mean costs based on all available cases, up to 12-month follow-up (continued)

Cost item	Total mean cost £ (SD)		MD (95% CI) (intervention – usual care)
	Intervention (N = 240)	Usual care (N = 214)	
SA2 costs:			
Private GP appointment	7.86 (35.84) N = 191	16.82 (86.81) N = 155	-8.96 (-22.55 to 4.63)
Private physiotherapist appointment	98.34 (360.92) N = 197	66.00 (185.12) N = 155	32.34 (-30.31 to 94.99)
Private mental health services specialist visit	28.88 (394.89) N = 187	7.06 (64.91) N = 153	21.82 (-41.69 to 85.32)

**TABLE 31** Total mean costs of health service use, medication cost and intervention cost based on multiply imputed data set, up to 12-month follow-up

Cost item	Total mean cost £ (SD)		MD (95% CI) (intervention – usual care)
	Intervention (N = 2400)	Usual care (N = 2140)	
GP visit at GP practice/home	66.50 (88.27)	58.92 (73.90)	7.58 (2.82 to 12.35)
GP phone consultation	84.38 (83.78)	71.46 (73.33)	12.92 (8.31 to 17.52)
Nurse/district nurse at GP practice	21.99 (27.43)	24.34 (33.90)	-2.35 (-4.13 to -0.56)
Nurse phone consultation	3.02 (6.23)	2.94 (6.08)	0.07 (-0.29 to 0.43)
Nurse/district nurse home visit	8.21 (85.07)	6.00 (45.94)	2.21 (-1.84 to 6.26)
Physiotherapist visit (community)	36.50 (162.10)	35.99 (118.74)	0.51 (-7.85 to 8.86)
Physiotherapist visit (hospital)	18.46 (61.60)	19.88 (65.79)	-1.42 (-5.13 to 2.28)
Mental health services	51.31 (289.87)	18.42 (122.17)	32.88 (19.66 to 46.11)
Hospital outpatient visit	475.63 (643.27)	486.93 (549.48)	-11.30 (-46.32 to 23.73)
Day-case hospital visit	437.05 (1385.08)	352.02 (944.61)	85.03 (15.22 to 154.85)
Inpatient hospital nights	33.39 (213.30)	172.94 (990.71)	-139.55 (-80.22 to -98.89)
Accident and emergency visits	194.98 (731.14)	220.88 (897.42)	-25.89 (-73.32 to 21.54)
Medication cost	346.06 (143.97)	414.97 (140.13)	-68.90 (-77.19 to -60.62)
Total cost (excluding intervention costs)	<b>1777.48 (2063.12)</b>	<b>1885.69 (2126.77)</b>	<b>108.21 (-13.81 to 230.23)</b>
Intervention cost	187.49 (0)	N/A	N/A
Total cost (including intervention costs)	<b>1964.96 (2063.12)</b>	<b>1885.69 (2126.77)</b>	<b>79.28 (-42.74 to 201.30)</b>

N/A, not applicable.

### Intervention cost

A total of 240 participants were randomised to receive the intervention, with the mean cost of the intervention estimated to be £187.49 per participant. The different cost components of the intervention can be seen in [Table 32](#).

For the training of yoga teachers for intervention delivery, in PPW1, face-to-face training took place (six trainees), with all remaining training taking place online in the subsequent training workshops for PPW2 (six trainees), MPW1 (three trainees) and MPW2 (nine trainees, involving three half-day sessions and one 1-hour additional meeting). The total cost of training that took place over the study period was divided by the number of participants in the intervention group to generate a mean training cost per participant of £31.92.

The cost of running the GYY classes was estimated to be £142.25 for a course of 12 classes; some courses were delivered face to face, and some were delivered online; hence both delivery methods were considered in the cost estimations. Four face-to-face courses took place in PPW1, four online courses in PPW2 and three online courses in MPW1. In MPW2, there was a combination of (three) face to face and (five) online courses. In terms of the cost of liaising with participants prior to the course starting, in PPW2 a cost of £135.00 (i.e. by phone) was applied for four courses, in MPW1 £270.00 (i.e. by Zoom) was applied for three courses, and during MPW2 three phone (£135.00) and five Zoom (£270.00) costs were applied.

The cost of equipment used by participants was estimated to be £12.50 on average per participant, which considered that an additional charge was incurred for the participants who received their equipment via post, for those who undertook classes online (those who undertook face-to-face classes received their equipment at the classes, hence no postage charge). The cost of home practice sheets was £0.82 per participant.

The intervention cost was explored as part of the sensitivity analyses to investigate whether the mode of delivery (i.e. face to face vs. online) affected the cost-effectiveness findings. When only face-to-face classes were considered, the intervention cost became £175.44, and for online classes only, it became £195.52. The online classes had higher costs associated with them due to the costs of liaising beforehand with participants via the additional one-to-one session and also the video-conferencing subscription cost, plus the cost of equipment was slightly higher due to sending out via post; though the face-to-face classes did involve the room hire cost, refreshment expenses and some travel expenses, overall there was an increase in cost for the online classes when compared with face-to-face classes.

### Cost-utility analysis and uncertainty

In the base-case analysis, participants in the intervention group incurred higher costs than the usual care group, at a mean (95% CI) cost of £1964.96 (£1882.38 to £2047.55) per participant in the intervention group and £1885.69 (£1795.53 to £1975.85) per participant in the usual care group (see [Table 33](#)). Additional QALYs were experienced by participants in the intervention group; mean (95% CI) QALYs

**TABLE 32** Cost of the GYY intervention

Cost element	Total mean cost £ per participant
Training of yoga teachers	31.92
Running GYY classes (course of 12 classes)	142.25
Equipment	12.50
Home practice sheets	0.82
<b>Total intervention cost</b>	<b>187.49</b>



were found to be 0.731 (0.724 to 0.738) in the intervention group versus 0.708 (0.700 to 0.716) in the usual care group (see [Table 33](#)).

The incremental analysis found that the intervention was associated with additional costs of £80.85 per participant, and generated an additional 0.0178 QALYs per participant, on average over the 12 months (see [Table 34](#)) in the base case. The resulting ICER was found to be £4546 per QALY, which falls under the WTP threshold used by NICE, of £20,000 per QALY. The intervention is therefore considered to be cost-effective. The incremental NMB, at the £20,000 per QALY threshold, was £274.85 (£268.29 to £281.41), also indicating that the intervention falls in the cost-effective region in comparison to usual care. This is shown by the incremental NMB associated with the intervention being positive; the NMB is positive if the ICER is below the WTP threshold, that is, indicating here that the benefit in QALYs gained exceeds the resources displaced if the intervention were to be adopted.

The point estimates of the bootstrap estimates presented in the cost-effectiveness plane in [Figure 8](#) indicate there is uncertainty associated with the cost-effectiveness findings, as estimates populate all four quadrants of the plane. The estimates are more heavily populated in the north-east and south-east quadrants, that is, showing an improvement in QALYs from the intervention, but indicate quite a variation in findings in terms of the costs, with points below the x-axis showing a reduction in cost for

**TABLE 33** Mean costs and QALYs by group (multiply imputed data set)

	Mean cost (SE) (95% CI)	Mean QALYs (SE) (95% CI)
Intervention	1964.96 (42.11) (1882.38 to 2047.55)	0.731 (0.004) (0.724 to 0.738)
Usual care	1885.69 (45.97) (1795.53 to 1975.85)	0.708 (0.004) (0.700 to 0.716)
Difference (intervention–usual care) <sup>a</sup>	79.28 (62.24) (–42.74 to 201.30) $p = 0.203$	0.023 (0.005) (0.013 to 0.034) $p < 0.001$

SE, standard error.

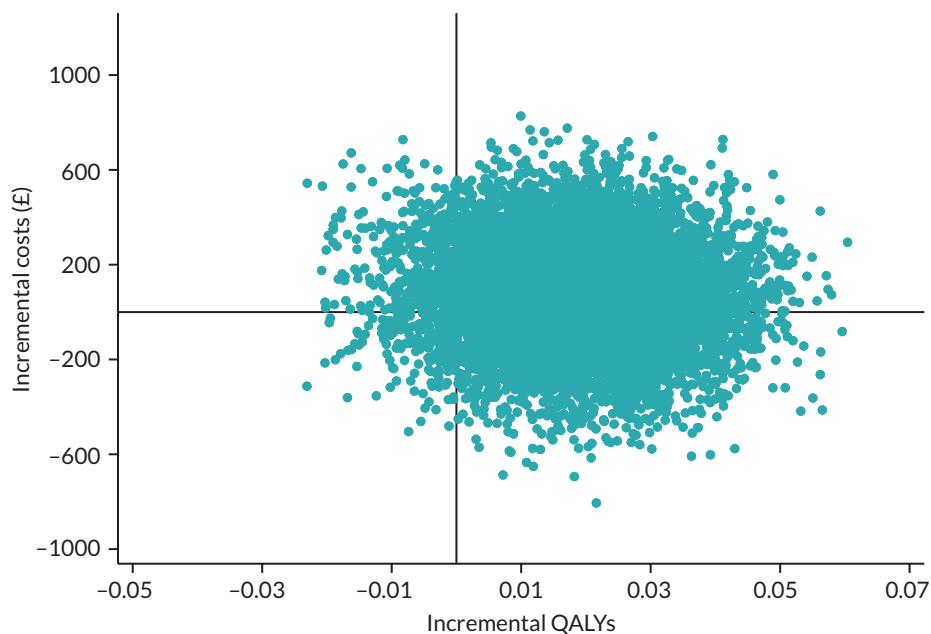
a Adjusted for baseline utility, age, gender and study site.

**TABLE 34** Cost–utility analysis results

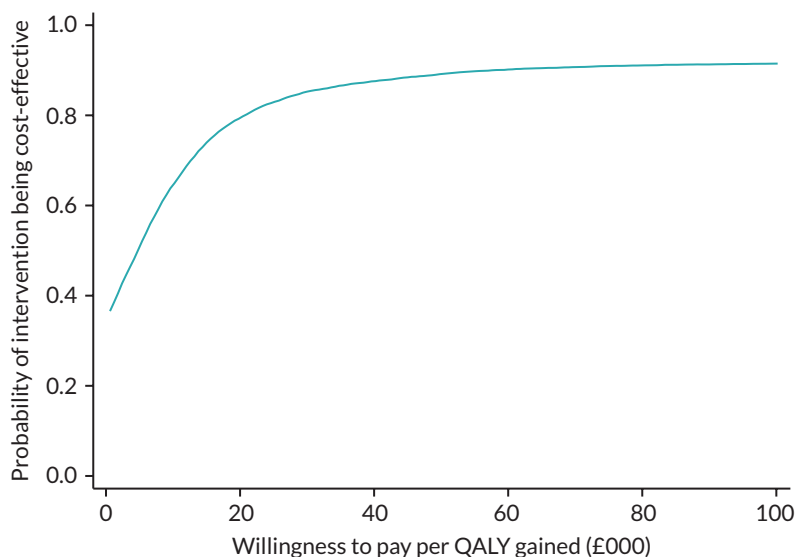
SA	Incremental mean cost (95% CI) <sup>a</sup>	Incremental mean QALYs (95% CI) <sup>a</sup>	ICER (£): cost per QALY	Probability cost-effective, £20,000/QALY (%)
Base case [multiply imputed (MI)], NHS perspective	80.85 (76.73 to 84.97)	0.0178 (0.0175 to 0.0180)	£4546.03	79
SA1: CCA	96.08 (–360.00 to 552.16)	0.0237 (–0.0136 to 0.0611)	£4049.20	77
SA2: personal expenses	116.94 (112.72 to 121.15)	0.0170 (0.0168 to 0.0172)	£6883.00	74
SA3: face-to-face yoga courses only	68.80 (64.69 to 72.92)	0.0178 (0.0175 to 0.0180)	£3868.60	81
SA4: online yoga courses only	88.88 (84.77 to 93.00)	0.0178 (0.0175 to 0.0180)	£4997.65	79
SA5: medication cost excluded	149.23 (145.14 to 153.32)	0.0178 (0.0175 to 0.0180)	£8395.54	73
SA6: removing age and gender	28.31 (24.35 to 32.27)	0.0184 (0.0182 to 0.0186)	£1537.66	85

a Difference between groups (intervention–usual care) with a bivariate model using seemingly unrelated regression used to estimate 95% CIs. All analyses are adjusted for the following covariates: baseline utility, age, gender and study site (with the exception of SA6).





**FIGURE 8** Cost-effectiveness plane for the GYY intervention relative to usual care: base-case analysis (MI adjusted for all covariates).



**FIGURE 9** Cost-effectiveness acceptability curve for the GYY intervention relative to usual care: for the base-case analysis (MI adjusted for all covariates).

the intervention, so a cost saving, and points above showing an increase in cost. The associated CEAC is shown in [Figure 9](#), illustrating the probability of the intervention being cost-effective for different WTP thresholds. The probability of the intervention being cost-effective at a WTP threshold of £20,000 per QALY gained threshold is 79% (base case) and 85% at a £30,000 per QALY gained WTP threshold.

### Sensitivity analysis

The results of the sensitivity analyses are presented in terms of the incremental mean costs, incremental mean QALYs, the cost-effectiveness finding (i.e. the ICER) and the probability of the intervention being cost-effective (based on a WTP threshold of £20,000 per QALY) in [Table 34](#). The QALY findings remained relatively similar across the sensitivity analyses, with most notable differences seen for the

CCA in SA1 for the complete analysis and in SA6 (where gender and age were no longer adjusted for in the model). The intervention group experienced additional QALYs, on average, with the difference found to be statistically significant for all analyses with the exception of SA1. In terms of the cost findings, all sensitivity analyses found there to be higher costs incurred by the intervention group, though these were of varying magnitude ranging between £28.31 (£24.35 to £32.27) for SA6 and £149.23 (£145.14 to £153.32) for SA5 (see [Table 34](#)). The cost-effectiveness results remained robust to the sensitivity analyses that were conducted; the ICER remained below the WTP threshold of £20,000 in all instances. The cost-effectiveness planes and CEACs for all sensitivity analyses can be seen in [Appendix 8](#) (see [Figures 12–17](#)) and [Appendix 9](#) (see [Figures 18–23](#)).

### Cost-consequence analysis

Unit costs of all NHS and private healthcare resource items are presented in [Table 35](#). The cost-consequence balance sheet summarises the results of the cost-consequences analysis (see [Table 35](#)). For each item of cost, the mean per participant resource use for both arms is displayed, alongside the incremental resource use. The mean cost per participant is also shown for each item according to trial arm, along with the incremental cost. In terms of the consequences, the outcomes of the GYY trial, as per [Chapter 3](#), are listed. A descriptive summary of the costs and consequences is provided, and items are discussed in a disaggregated format.

### Costs

In general, the between-group differences were minimal for most items of resource use. In terms of community-based care provided via GP clinics, participants in the intervention arm typically reported, on average, greater levels of resource use than participants receiving usual care except for nursing-based care. Of the service use collected, the most frequent community service used by participants in both arms was phone-based GP consultations, followed by clinic-based nurse and GP visits.

As only small between-group differences in average resource use were observed, only small between-group differences were observed in terms of cost for most items. However, more notable cost differences were observed for GP phone consultations and community-provided mental health services, with participants in the intervention arm costing more on average, £11.69 and £46.09 per person, respectively.

An inverse pattern was observed for care provided through hospitals. Participants allocated to receive the intervention had, on average, few hospital-based physiotherapy visits, outpatient visits, inpatient nights in hospital and accident and emergency visits resulting in an inpatient stay. This translated into lower average service use costs for these services, with participants in the intervention arm costing, on average, £2.21 less for physiotherapy, £23.95 less for outpatient visits, £159.01 less for inpatient stays and £13.62 less for accident and emergency visits, resulting in an inpatient stay during the 12-month course of the trial.

Medication costs were estimated for the trial population based on a subsample. Based on this estimate, it appears that participants who received the intervention had, on average, lower associated medication costs than those who did not.

In addition to NHS-provided care, participants also reported their use of privately funded health care. Data were not collected as to how this was funded, for example, through a private healthcare plan or through one-off payments, so an assumption was made that participants paid on a 'pay-as-you-go' basis. Unlike for NHS-funded care, participants in the intervention arm reported, on average, fewer private GP visits than participants in the usual care arm; however, they did report, on average, greater uptake of privately funded physiotherapy and mental health services. This translated into higher average costs for participants in the intervention arm for private physiotherapy services (additional £32.34 per person) and mental health services (additional £21.82 per person).

TABLE 35 Cost-consequence balance sheet

Cost item	Mean resource use per participant			Mean cost per participant £ (SD)		
	Intervention (N = 240)	Usual care (N = 214)	Incremental difference (intervention - usual care)	Intervention (N = 240)	Usual care (N = 214)	Incremental difference (intervention - usual care)
<b>Community-based health care</b>						
GP visit at GP practice	1.58 (2.11) N = 197	1.44 (1.93)	0.15 (-0.29 to 0.58)	62.13 (82.93) N = 197	56.41 (75.68) N = 153	5.72 (-11.20 to 22.64)
GP visit at home	0.03 (0.23) N = 190	0.02 (0.14) N = 147	0.01 (-0.03 to 0.05)	3.18 (22.92) N = 190	2.05 (14.28) N = 147	1.12 (-3.12 to 5.37)
GP phone consultation	2.78 (2.81) N = 174	2.40 (2.41) N = 148	0.38 (-0.20 to 0.96)	84.92 (85.72) N = 174	73.23 (73.65) N = 148	11.69 (-6.00 to 29.38)
Nurse visit at GP practice	1.91 (2.39) N = 184	2.24 (3.07) N = 153	-0.33 (-0.92 to 0.25)	21.69 (27.05) N = 184	25.49 (34.85) N = 153	-3.80 (-10.44 to 2.84)
Nurse visit at home	0.18 (1.76) N = 194	0.03 (0.20) N = 150	0.15 (-0.13 to 0.44)	4.89 (47.78) N = 194	0.72 (5.40) N = 150	4.17 (-3.54 to 11.89)
Nurse phone consultation	0.68 (1.30) N = 145	0.66 (1.40) N = 118	0.01 (-0.31 to 0.34)	2.97 (5.72) N = 145	2.91 (6.15) N = 118	0.07 (-1.38 to 1.51)
District nurse visit at GP practice	0.06 (0.37) N = 196	0.03 (0.20) N = 155	0.04 (-0.03 to 0.10)	0.70 (4.25) N = 196	0.29 (2.22) N = 155	0.40 (-0.34 to 1.14)
District nurse visit at home	0.14 (1.72) N = 196	0.17 (0.84) N = 155	-0.02 (-0.32 to 0.27)	3.88 (46.63) N = 196	4.55 (22.89) N = 155	-0.68 (-8.71 to 7.36)
District nurse phone consultation	0.04 (0.28) N = 153	0.02 (0.13) N = 123	0.02 (-0.03 to 0.08)	0.17 (1.22) N = 153	0.07 (0.56) N = 123	0.10 (-0.13 to 0.34)
Physiotherapist (community)	0.63 (2.91) N = 197	0.57 (1.75) N = 153	0.06 (-0.46 to 0.58)	36.98 (171.08) N = 197	33.41 (102.59) N = 153	3.57 (-27.20 to 34.35)
Mental health services (community)	0.22 (1.33) N = 196	0.05 (0.38) N = 155	0.16 (-0.06 to 0.38)	62.68 (380.33) N = 196	16.59 (109.15) N = 155	46.09 (-15.93 to 108.11)
<b>Hospital-based health care</b>						
Physiotherapist (hospital)	0.29 (1.04) N = 200	0.33 (1.10) N = 158	-0.04 (-0.26 to 0.18)	16.39 (58.75) N = 200	18.59 (62.30) N = 158	-2.21 (-14.84 to 10.42)
Mental health services (hospital)	0.05 (0.31) N = 199	0 (0) N = 158	0.05 (0.00 to 0.10)	14.36 (89.66) N = 199	0 (0) N = 158	14.36 (0.33 to 28.39)
Other hospital outpatient appointment	2.61 (3.64) N = 194	2.74 (3.07) N = 161	-0.13 (-0.84 to 0.58)	474.44 (660.91) N = 194	498.39 (556.69) N = 161	-23.95 (-153.08 to 105.18)
Day-case appointment	0.37 (1.223) N = 194	0.30 (0.75) N = 159	0.07 (-0.15 to 0.29)	442.38 (1469.34) N = 194	359.84 (897.16) N = 159	82.54 (-179.35 to 344.43)
Accident and emergency visit (not overnight)	0.25 (0.76) N = 158	0.21 (0.54) N = 127	0.04 (-0.12 to 0.20)	75.16 (227.10) N = 158	63.11 (161.44) N = 127	12.04 (-35.00 to 59.09)
Accident and emergency visits (including overnight hospital stays)	0.25 (1.40) N = 194	0.34 (2.11) N = 156	-0.09 (-0.46 to 0.28)	108.62 (607.51) N = 194	122.25 (786.00) N = 156	-13.62 (-160.12 to 132.87)
Inpatient nights in hospital	0.09 (0.51) N = 198	0.51 (2.95) N = 160	-0.43 (-0.85 to -0.01)	32.00 (190.63) N = 198	191.01 (1101.16) N = 160	-159.01 (-315.69 to -2.33)

Cost item	Mean resource use per participant			Mean cost per participant £ (SD)		
	Intervention (N = 240)	Usual care (N = 214)	Incremental difference (intervention - usual care)	Intervention (N = 240)	Usual care (N = 214)	Incremental difference (intervention - usual care)
<b>Medications</b>						
Medications <sup>a</sup>				346.06 (144.24) N = 240	414.97 (140.42) N = 214	-68.90 (-95.22 to -42.58)
<b>Privately funded health care</b>						
Private GP appointment	0.10 (0.45) N = 191	0.21 (1.10) N = 155	-0.11 (-0.29 to 0.06)	7.86 (35.84) N = 191	16.82 (86.81) N = 155	-8.96 (-22.55 to 4.63)
Private physiotherapist appointment	1.54 (5.64) N = 197	1.03 (0.23) N = 155	0.51 (-0.47 to 1.49)	98.34 (360.92) N = 197	66.00 (185.12) N = 155	32.34 (-30.31 to 94.99)
Private mental health services specialist visit	0.80 (1.10) N = 187	0.02 (0.18) N = 153	0.06 (-0.12 to 0.24)	28.88 (394.89) N = 187	7.06 (64.91) N = 153	21.82 (-41.69 to 85.32)
<b>Consequences (primary time point of over 12 months)</b>						
EQ-5D-5L Utility				Mean adjusted difference: 0.02 (95% CI -0.006 to 0.045) <i>p</i> = 0.14		
EQ-5D-5L VAS				Mean adjusted difference: 1.67 (95% CI -0.78 to 4.12) <i>p</i> = 0.18		
GAD-7				Mean adjusted difference: -0.09 (95% CI -0.57 to 0.40) <i>p</i> = 0.72		
PHQ-8				Mean adjusted difference: -0.36 (95% CI -0.90 to 0.18) <i>p</i> = 0.19		
UCLA-3 loneliness				Mean adjusted difference: 0.03 (95% CI -0.16 to 0.22) <i>p</i> = 0.75		
ELSA single-item direct loneliness question				Mean adjusted difference: -0.01 (95% CI -0.15 to 0.13) <i>p</i> = 0.88		
PROMIS-29				Mean adjusted difference: Physical function: 0.73 (95% CI -0.22 to 1.69) <i>p</i> = 0.13 Anxiety: 0.37 (95% CI -0.82 to 1.56) <i>p</i> = 0.54 Depression: 0.08 (95% CI -1.02 to 1.17) <i>p</i> = 0.89 Fatigue: -0.27 (95% CI -1.52 to 0.97) <i>p</i> = 0.67 Sleep disturbance: -0.17 (95% CI -1.19 to 0.85) <i>p</i> = 0.75 Social participation: 0.96 (95% CI -0.40 to 2.32) <i>p</i> = 0.17 Pain Interference: -1.14 (95% CI -2.24 to -0.04) <i>p</i> = 0.04 Physical health: 0.87 (95% CI -0.11 to 1.86) <i>p</i> = 0.08 Mental health: 0.36 (95% CI -0.63 to 1.36) <i>p</i> = 0.47 Global (pain intensity): -0.32 (95% CI -0.61 to -0.04) <i>p</i> = 0.03		
Falls				Incidence rate ratio: 1.38 (95% CI 0.95 to 2.01) <i>p</i> = 0.09		

a Medications costs were estimated based on a subsample of the participants.

## Consequences

The primary trial analysis showed a small difference in favour of the intervention; however, this was not statistically significant ( $p = 0.14$ ). No statistically significant differences were observed in secondary outcomes, except for in the T-score for the pain interference subscale of the PROMIS-29 at 3 months ( $-1.44$ , 95% CI  $-2.63$  to  $-0.26$ ;  $p = 0.02$ ) and over the 12 months ( $-1.14$ , 95% CI  $-2.24$  to  $-0.04$ ;  $p = 0.04$ ) and in the global (pain intensity) PROMIS-29 item at 12 months ( $-0.45$ , 95% CI  $-0.83$  to  $-0.08$ ;  $p = 0.02$ ) and over the 12 months ( $-0.32$ , 95% CI  $-0.61$  to  $-0.04$ ;  $p = 0.03$ ).

The cost–consequence analysis shows some differences between the resource uptake and associated costs between the two trial arms (GY intervention and usual care), particularly in terms of NHS-provided hospital care, where costs, on average, were considerably lower for several of the measured services.

In all, the cost–consequence model summarises the between-group differences in resource use and costs, alongside the primary trial outcomes. As can be seen, although some resources were used on average more frequently by participants in the intervention arm and thus attracted higher costs, hospital-based services were used less frequently, on average, and thus costs were lower in the intervention arm.

Given these costs and the way in which small, albeit not statistically significant, benefits were observed in the trial analysis when considering the trial outcomes, the cost–consequence model provides further support to the within-trial analysis that the GY intervention may represent good value for money.

## Discussion

The within-trial cost–utility analysis presented in this chapter indicated that the intervention group experienced higher costs (£80.85; 95% CI £76.73 to £84.97) and additional QALYs (0.0178 QALYs; 95% CI 0.0175 to 0.0180 QALYs) over the 12-month period versus usual care. The resulting ICER of £4546 indicates that the GY intervention in addition to usual care was found to be cost-effective when compared with usual care alone over a 12-month time horizon, based on a WTP threshold of £20,000 per QALY gained. The probability of the intervention being cost-effective at a WTP threshold of £20,000 per QALY gained was found to be 79%. Uncertainty around the findings was displayed on the cost-effectiveness plane by the point estimates from the base-case analysis featuring in all four quadrants, most predominantly in the north-east and south-east quadrants. The cost-effectiveness findings appear robust across a range of assumptions that were explored using sensitivity analyses, with the finding of cost effectiveness remaining for all sensitivity analyses undertaken. A cost–consequence balance sheet was implemented as a supplementary analysis to the within-trial analysis to provide a clear descriptive summary of the NHS-funded and privately funded costs associated with the GY intervention over a time horizon of 1 year. Results from a cost–consequence analysis highlighted that participants who were in receipt of the intervention, on average, reported lower resource use for several healthcare services, most notably hospital-based care, and this translated into lower associated costs.

The base-case analysis was undertaken from the perspective of the NHS and PSS, and, originally, we planned to undertake a secondary analysis from the wider societal perspective which included the exploration of productivity losses and personal expenses (i.e. private treatments). However, due to the age of the study population, and on consideration of the length of the questionnaires, it was felt that the productivity loss questions were not appropriate to include. The personal expenses were therefore explored as part of a SA. Although baseline completion rates were high for the economic data, missing responses for a large proportion of items at later time points led to a high rate of missing data overall (58%). While this is not uncommon in trial-based economic evaluations involving patient-level data, it is important to highlight this when interpreting the findings from the analysis. The pattern of missing data was found to be non-monotonic, and the analysis of missing data suggested it was reasonable to

assume the data were MAR. Therefore, missing data were accounted for using multiple imputation, a validated approach in RCTs<sup>54</sup> in the base case, with a CCA (SA1) undertaken as an alternative to the use of multiple imputation for dealing with missing data.

Key cost drivers of the analysis comprised hospital-based services, specifically inpatient stays, accident and emergency visits (including overnight stays), day-case visits and outpatient attendances. Also, as expected, the cost of the intervention itself was prominent at £187.49 in the base case, which comprised a combination of face-to-face and online courses. The delivery mode of the GYY classes provided as part of the intervention was explored using SA (SA3 and SA4): online classes were found to be slightly more expensive than face-to-face classes. This was attributable to the costs of liaising with participants beforehand, video-conferencing subscription costs and postage and packing costs for the equipment being sent out to participants for the online classes. Although face-to-face classes incurred costs of venue hire, refreshments and travel expenses, the online classes were still estimated to be slightly higher in cost overall, which is an interesting finding. Medications were also a key cost driver. Although this cost was estimated from a sample of 62 (i.e. 14% of) participants who were considered to be representative, the effect of removing it from the analysis was explored via SA; the intervention remained cost-effective, though incremental costs were found to be higher, with a correspondingly higher ICER of £8395.54; still cost-effective based on a threshold of £20,000 per QALY gained.

The training cost applied as part of the intervention cost can vary depending on the number of attendees and the number of sessions that the training is held over. Some of the yoga teachers attended the training more than once, and there were some who attended but did not actually teach any of the GYY classes (i.e. they provided backup). We investigated an alternative scenario to explore how the cost-effectiveness findings would be affected if the training had instead been undertaken over two sessions (one face to face and one online) each with six attendees, that is, for the 12 yoga teachers who undertook the GYY classes as part of the intervention during the study. The training cost was found to reduce to £17.38 per participant, resulting in the overall intervention cost reducing to £172.95. In terms of the cost-effectiveness findings, costs were found to be £66.31 higher for the intervention group overall, with an ICER of £3406.60; therefore the intervention was found to be more cost-effective (than the base case).

In conclusion, the intervention resulted in more QALYs than usual care, though it was found to be more expensive. Based on the information available from this trial, the intervention has been estimated to be cost-effective when considering the WTP threshold used by NICE of £20,000 per QALY gained.



# Chapter 5 Process evaluation

## Introduction and aims

This chapter presents the methods and results of the process evaluation that was conducted as part of the GYY trial. The aim was to establish the determinants of delivery, trial processes, participant experience, implementation and dissemination aspects of the GYY trial and yoga intervention.

## Methods

### *Study design*

The qualitative process evaluation involved interviews and observations with a diverse range of trial participants and yoga teachers at all sites, as well as additional stakeholders.

### *Sampling strategy*

Our sampling strategy aimed to achieve a balance between spread of data and depth. We were also responsive to the study context, with additional data collection in response to our emerging analysis and study events in order to support meaning saturation.<sup>55</sup> Our sample of people recruited to interview was purposive,<sup>56</sup> using the following criteria:

- Trial participants: a maximum variation sample of people in the yoga arm to include the breadth of demographic data relating to age, gender, ethnicity, Index of Multiple Deprivation (IMD) and number, severity and type of health condition, as well as site and mode of delivery. A convenience sample of people in usual care arm and trial decliners.
- Trial yoga teachers: a total population sample of all yoga teachers in the trial.
- Additional stakeholders: a snowball, emergent and theoretical sample of representatives of healthcare, academic and community-based groups involved in the development, delivery, evaluation, implementation and/or commissioning of yoga and related UK health programmes.

One trial class was observed for all but one of the yoga teachers. Sampling was monitored and discussed regularly at each team meeting.

### *Recruitment, consent and data collection*

All interview participants were sent interview-specific study information. In relation to observations, all the trial participants attending a class and the yoga teacher were sent observation-specific study information. Interviews and observations took place throughout the trial, from January 2020 to April 2022. All process evaluation participants provided written informed consent for the interviews and/or observations, in addition to main trial consent.

The majority of interviews and observations were conducted remotely, with the exception of PPW1, which was undertaken face to face prior to COVID-19 restrictions. Some yoga participants and yoga teachers took part in a follow-up interview.

Interview topic guides were initially developed and agreed within the team and included the following:

- Trial participants: trial processes (e.g. ideas and/or concerns about randomisation and consent); the yoga intervention (e.g. expectations about yoga and/or concerns about impact on health and acceptability, delivery format, the timing, location, content, home practice); general health and well-being.
- Trial yoga teachers: views and experiences of yoga, trial processes, trial delivery and participant engagement.



- Additional stakeholders: views and experiences of yoga research, yoga provision for health and social prescribing.

The content of the interview was flexible to accommodate additional unanticipated areas, the developing analysis and in the case of follow-up interviews, what was known from the prior interview. Interviews and observations were conducted by an experienced qualitative researcher (LW), who, as necessitated by the purposive sampling strategy, was unblinded to allocation.

### ***Qualitative data management and analysis***

Interviews were, with consent, audio-recorded, transcribed verbatim and edited to ensure anonymity of respondents. Contemporaneous field notes from non-participant observation of yoga classes were edited to ensure anonymity of participants. The analysis was conducted according to the standard procedures of rigorous qualitative analysis,<sup>57</sup> including pre-coding,<sup>58</sup> open and focused coding, constant comparison, memoing,<sup>59,60</sup> deviant case analysis<sup>61</sup> and within- and across-participant thematic charting.<sup>62</sup> A sample of data was independently coded and cross-checked, to identify and reflect on similarities and differences. The majority of data were also analysed collectively in weekly 'data clinics' where the research team shared and exchanged interpretations of key issues emerging from the data. Analysis was conducted by LW/TR (both with social science backgrounds) who were joined for some discussions by GT. All names and other potentially identifying details have been changed to ensure the anonymity of participants.

### ***Relationship between process evaluation and trial***

The emergent analysis was regularly discussed at monthly TMGs and, where appropriate, informed changes to study processes. We produced two overviews with large numbers of detailed (minor) recommendations during the PPW1, in relation to the provision of face-to-face classes in March 2020 for initial pilot work, then following the shift to online provision, PPW2, in February 2021. Examples of recommendations ranged from supporting more effective class register practices to clarifying the concept, purpose and meaning of 'usual care' in participant trial information to enhancing class engagement during online class delivery. Nearly all recommendations were acted upon. This chapter will not focus on exploring the trial optimisation and fidelity work; instead it will offer an overview of some core issues that help enable reflection on the effectiveness outcomes of the trial.

## **Results**

Forty interviews were conducted with twenty-five yoga participants (see [Table 36](#)). The majority of initial interviews took place during the 12-week intervention period ( $n = 22$ ), with three in the main phase occurring within 3 months post intervention. Trial yoga participants interviewed had engaged in face-to-face delivery of yoga classes ( $n = 7$ ) and online delivery ( $n = 18$ ) across the pilot ( $n = 13$ ) and main phase ( $n = 12$ ) periods of the study. Follow-up interviews ( $n = 15$ ) took place between 3 and 8 months post intervention. Interviews typically lasted between 60 and 90 minutes. Interview participants' ages ranged from 66 to 91 years, with six participants 80 years of age or over (see [Table 37](#)). Participants were predominately female ( $n = 14$ , 56%) and White British ( $n = 22$ , 88%), with IMD ranging from 1 to 10. Sixteen participants lived with others, eight lived alone and one lived in sheltered housing. Albeit that we did not seek to recruit participants to echo those in the wider trial cohort, the demographics of the process evaluation trial participants are broadly reflective of the wider trial cohort. The wider trial yoga arm cohort is slightly more likely to be female (60.6%), ethnically white (95.8%) and on average from a higher IMD (7.54). However, trial participants in the process evaluation typically had a higher number of conditions, 28% ( $n = 7$ ) with six or more, compared to 4.5% in the yoga arm of the trial. During the pilot phase, two interviews were conducted with usual care trial participants and one with a trial decliner, and these lasted between 30 and 60 minutes.

**TABLE 36** Summary of all participants' involvement in interviews and observation

		Pilot phase	Main phase	Total
Yoga participants interviewed	Initial interview only	3	7	10
	Initial and follow-up interview	10	5	15
	<b>Total</b>			<b>25</b>
Usual care participants interviewed	Initial interview only	2	0	2
	<b>Total</b>			<b>2</b>
Trial decliners interviewed	Initial interview only	1	0	1
	<b>Total</b>			<b>1</b>
Yoga teachers interviewed	Initial interview only	2	6	8
	Initial and follow-up interview	3	0	3
	<b>Total</b>			<b>11</b>
Other stakeholders <sup>a</sup> interviewed	Initial interview only	0	15	15
	<b>Total</b>			<b>15</b>
Yoga classes observed	Face to face	3	0	3
	Online	4	3	7
	<b>Total</b>			<b>10</b>

a Includes trial team members ( $n = 4$ ), yoga teachers/researchers ( $n = 4$ ), health researchers ( $n = 2$ ), NHS, Voluntary and Community Sector and Local Authority affiliated stakeholders ( $n = 5$ ).

**TABLE 37** Characteristics of qualitative interview participants: yoga and usual care arms ( $n = 27$ )

	Yoga interview participants ( $n = 25$ )	Usual care interview participants ( $n = 2$ )
Age, years	73.6 (6.0)	69.5 (3.5)
Gender, $n$ (%)		
Male	11 (44)	1 (50)
Female	14 (56)	1 (50)
Ethnic group, $n$ (%)		
White British	22 (88)	2 (100)
White Irish	1 (4)	0 (0)
Black Caribbean	1 (4)	0 (0)
Asian Indian	1 (4)	0 (0)
IMD decile <sup>a</sup>	6.92 (2.49)	N/A
No. of conditions, <sup>b</sup> $n$ (%)		
2	3 (13)	0 (0)
3	4 (16)	0 (0)

continued

**TABLE 37** Characteristics of qualitative interview participants: yoga and usual care arms ( $n = 27$ ) (continued)

	Yoga interview participants ( $n = 25$ )	Usual care interview participants ( $n = 2$ )
4	7 (28)	0 (0)
5	4 (16)	0 (0)
6 or more	7 (28)	2 (100)
<b>No. of conditions, median (min, max)</b>	4 (2, 8)	7 (7, 7)

a 1 = most deprived to 10 = least deprived.  
b Conditions grouped as per the inclusion criteria.

**Note**  
Data are mean (SD) unless otherwise stated.

Fourteen interviews were conducted with 11 yoga teachers (see [Table 36](#)). All initial interviews took place during or within 3 weeks of the 12-week intervention period ( $n = 12$ ). Follow-up interviews ( $n = 3$ ) took place within 3 months post intervention. Interviews typically lasted between 60 and 100 minutes. Where reported ( $n = 8$ ), yoga teaching experience ranged from 4 to 35 years. Ten of the yoga teachers received their yoga training through BWY, with five currently serving as Diploma course tutors. Five yoga teachers indicated additional yoga qualifications, and five were also qualified exercise or Pilates instructors.

Fifteen interviews were conducted with fifteen additional stakeholders, including trial team members ( $n = 4$ ) and non-trial yoga teachers, three with research experience ( $n = 4$ ) (see [Table 36](#)). They typically lasted between 45 and 60 minutes. Finally, observations during the pilot phase ( $n = 7$ ) were undertaken across weeks 4–11 of the 12-week intervention to evaluate the range of delivery determinants. During the main phase, all observations ( $n = 3$ ) were conducted at week 7 to explore standardisation of delivery at a single time point. Three face-to-face yoga classes were observed, and seven were delivered online.

### **Enrolment: clear willingness to engage with Gentle Years Yoga**

Sixteen of the interviewed trial participants had not tried yoga before. Of these, some were familiar with what yoga involved through family members who practised, while others were unaware of the variety of yoga styles available. Trial participants' preconceptions of yoga, based on media representation, experience or word of mouth, ranged from a physically challenging form of exercise targeting younger adults, women and the financially comfortable, through to a 'soft' form of stretching targeting physical and mental well-being. Several participants noted a generational knowledge of yoga among older adults of the 1960' culture, associating its practice with Eastern spirituality. Common reasons for lack of engagement in yoga prior to the trial included barriers of cost, transport and class availability. Those participants with prior engagement mostly described practising yoga in their 30s or 50s, with most stopping due to the physical content or ability to self-mobilise to and from the floor precluding ongoing engagement with their community- or gym-based classes. Only one participant described a long-term, mat-based yoga practice.

Participants outlined four main reasons for joining the GYY trial: potential benefit to health, invitation-enabled opportunity for exercise, curiosity about yoga and, as with all studies, discussions of altruism (see [Table 38](#)). Several participants noted the encouragement of family members and their GPs to try the yoga classes. Notably, the majority of participant interviewees, those with prior experience of yoga and the yoga naïve, did not know yoga could be adapted either to a seated practice or to older adults prior to receiving their trial invitation. Despite some concerns about the potential physical challenges of yoga,

TABLE 38 Overview of participants' main reasons for joining the trial

Rationale	Description	Example
Health-related	GYG presented a potential age-appropriate form of activity to maintain, improve and/or manage their physical and mental health	Grace joined the trial as she 'wanted to see if it would be any good for me, to help me with my aches and pains, and, as I said, being told that yoga can help with depression, anxiety and stress and all this, and I wanted to see if that would help' (YPart-Int1)
Opportunism	Classes provided extrinsic motivation for a group-based activity. During COVID lockdown/isolation, online provision of GYG classes provided access to exercise and social contact	Gwen noted that 'we were in the middle of a COVID pandemic ... But I, like so many people of my age, was more or less in isolation. Not getting as much exercise, no swimming. ... And I looked at the letter very carefully, but decided I ought to do something to keep myself fit' (YPart-Int1)
Curiosity	They were interested to see what GYG was about	John was 'intrigued' by yoga, without the trial invitation he 'wouldn't have even thought about it. No, not in a million years' (YPart-Int1)
Altruism	They were happy to take part in research that may benefit others. This was especially relevant to the two participants in the usual care arm	Abe, part of the usual care arm, described two reasons 'I mean, my first reaction - I had two things, really. One was, I've always thought about doing yoga anyway, to build up body strength. ... That's probably the primary reason. But the other one was, you know, it'll help others. ... But I think, you know, it'll help others. I'd also previously signed up for that - you know every couple of years they do - it's a national study, it's a database where they basically measure your whole body, everything. You get scanned. You go for the whole day, and every three years they do it, something like that. ... And I joined that, again because I thought it might help others, you know' (UCPart-Int1)

the trial participants considered the GYG intervention would be within their physical capabilities, due to it being chair-based, and delivered by yoga instructors who had undertaken specialist training in yoga for older adults with health issues. As such, there were minimal concerns regarding the physical content of the trial classes. However, several noted social anxiety around meeting new people and potentially feeling embarrassed at learning a new activity in front of others.

Unsurprisingly for individuals joining a yoga trial, many held a belief, or were interested in exploring, the use of positive health behaviours as an effective and proactive means of maintaining independent living and improving health-related well-being. Self-management was regularly mentioned in preference to pharmaceutical management of chronic health conditions - 'I hate taking tablets, I really do ... I would choose an activity, you know, rather than [medication]' (Joy, YPart-Int1) - with participants concerned regarding the side effects and potential dependency issues of certain pharmaceuticals. However, trial participants' advocacy of proactive self-management came with the caveat of ensuring management choices were age-appropriate.

### **Integration: many already felt (relatively) active and healthy**

The number of self-reported chronic health conditions ranged from two to eight per participant (mean 4.5). Fifteen interview participants reported physical health issues, predominantly related to mobility and pain, and eight participants reported both physical and mental health issues. Conditions were often co-related, for example, pain was commonly associated with sleep disturbance, fatigue and low mood. Six of the fifteen longitudinal interview yoga participants reported decreased health status over time, two due to accidents and one related to mental health. However, despite the high level of multimorbidity, most participants viewed their health as good. As Derek notes:

*I've got high blood pressure too. That's well controlled on medication. What else? I've been on antidepressants for about ten years now, and again that's reasonably well controlled. [...] I've got some gastric reflux, which I'm on treatment for as well, but that's not very severe.*

YPart-Int 1

In his follow-up interview, Derek went on to note that 'I'd describe myself as being fairly healthy despite all of the issues, high blood pressure, arthritis, reflux and that' (YPart-Int2). Many did not consider multimorbidity indicative of poor or ill health. Certain health conditions, such as osteoarthritis, were expected among this older adult age group. The presence of multiple conditions was not seen as cumulative, with more conditions not necessarily equating to poorer health. Additionally, the diagnosis of a chronic health issue has often instigated adopting healthier lifestyle behaviours to minimise disease progression and maximise quality of life.

The impact of their multimorbidity was related to acute rather than chronic health status, and pharmaceutical management meant many health conditions had little to no impact on their activities of daily living or general well-being. As such, it was the day-to-day presentation, rather than the presence, of a health condition that determined its impact on the person. For example, when asked, Heather noted that her health is 'not bad really, because, although I've got these conditions, they're controlled' (YPart-Int1). Several chronic health conditions, such as hypertension and coronary heart disease, were experienced as stable and/or asymptomatic conditions, often amendable to pharmaceutical management. These had little impact on participants' daily life, to the point that many participants did not mention them in the interviews when prompted, even though they had indicated them on their screening questionnaires. Similarly, many participants had adjusted to living with chronic but stable pain, viewing it as a protective signal for managing their health condition. However, conditions with variable and/or intense symptom presentation, such as rheumatoid arthritis or clinical anxiety, were highly impactful, requiring adaptive and tailored management of both the acute and chronic phases of the condition. Notably, seven of the participants deemed their health to be above what they considered the level required to qualify for this trial. For example, Anne noted, while filling in the initial trial-related forms 'I thought, "oh, they won't want me, because there's nothing wrong with me" sort of thing' (YPart-Int1). All the trial participants interviewed queried their inclusion in the trial.

### **Workability: engaged well with online and face-to-face delivery**

There was no clear delivery preference among participants. Some participants favoured online delivery, citing pragmatic reasons including the removal of geographical barriers around choosing a yoga teacher, the removal of travel barriers related to transport logistics, adverse weather and acute health status. This in turn enabled consistency of practice, independent of external or health factors. Several participants noted that the ability to look at the yoga teacher without being distracted by other participants enabled them to focus more on their practice. It also provided a level of anonymity and emotional security for those with emotional and mental health issues. Conversely, those preferring face-to-face delivery favoured the physical presence of class peers for social support and interaction and of the yoga teachers for receiving positional information and adaptations. On a pragmatic level, face-to-face classes were also preferred for those with poor internet connections impacting online classes.

Most yoga teachers indicated they adapted quickly to the onset of COVID-19 restrictions in March 2020, quickly moving their community yoga classes online where possible. Albeit that participants' IT skills were variable, ranging from basic computer literacy to highly competent business users, the majority adapted quickly to the computer-based classes. Despite the lifting of restrictions, which allowed a return to face-to-face teaching, many yoga teachers took a pragmatic approach to class delivery and continued to offer online class options. They have continued online delivery of GYY and their general yoga provision in response to a stable client base who prefer the convenience and accessibility of being able to engage in yoga without having to travel outside their home. Centrally, whether delivered online or face to face, the GYY style of yoga was viewed as a suitable and safe form of physical movement for older adults with varying health issues, independent of delivery format.

Despite restricted participant visibility and communication, yoga teachers had no safety concerns for online yoga beyond those inherent to face-to-face-delivered classes. The only potential safety concerns Kirsty had were the times when:

*participants will want to do a little bit more than I am ready for them to do. But would that make much difference between in person and online? Not really, no. Because if they're going to push themselves a bit more, they're going to do it anyway, in person or online. I think that the responsiveness, when it's in-person, it's the two-way verbal. When it's online, it's one-way verbal, and I'm observing so much more.*

YTeach-Int1

Safety reminders are a consistent and integral element of all yoga teaching. Both participants and yoga teachers identified two key safety messages of self-responsibility and non-harm, working within your personal limits, within a safe practice environment. The most common adverse impacts of a GYY class reported were those typically associated with exercise classes in general, including delayed-onset muscle soreness and upper body pain.

Finally, nearly all yoga teachers reported delivering class content which contained elements that differed from their 12-week plan submitted to and authorised by the yoga consultants prior to intervention delivery. The consistent reason given for this was that the class plans, written prior to meeting their participants, needed to be adapted once their health issues and physical abilities were known. Most notably, we also observed a variable focus on home practice content and the timely introduction of home practice sheets. Home practice should have been covered in the first segment (i.e. 'housekeeping') of each yoga class as per yoga teacher research training ([Report Supplementary Material 4](#)). Home practice was, in some cases, however only referred to in passing during social time after class, meaning only those who stayed on for this optional time received any home practice guidance.

#### **Appraisal: experienced Gentle Years Yoga as gentle exercise (with mindful breathing)**

The majority of participant interviewees viewed GYY as a form of gentle exercise. Only two participants found the class content physically challenging, with one further describing the level of physicality as appropriate both for beginners and for online delivery. However, while lacking overt physical challenge, participants commonly mentioned the challenge presented by coordination of the breath with movement. So, Michael, who did not find GYY 'at all physically challenging' did however find 'the complexity of the movements' challenging:

*making certain you replicated what [trial YT] was doing, so there was a lot of thinking: 'hang on, that's out, then up and breathe out and breathe in', you know [laughs].*

YPart-Int1

This was noted as a unique feature differentiating yoga from exercise and provided a mental focus commonly described as inducing a calmness not found in other forms of physical activity. General participant feedback noted a lack of background information on defining what 'yoga' was in terms of basic concepts and the role of the breath within its practice.

Overall, the participant interviewees had mixed views on the role, place and target audience of the GYY form of yoga. Those in favour of the practice viewed GYY as filling a gap in the market for age-appropriate exercise, particularly for those with no previous experience of physical activity or those seeking an alternative to previous aerobic-based activities which had become too physically challenging. For these participants, GYY was viewed as both a physical and mental health tool, with the use of chairs and adaptability of the physical postures key to making the practice inclusive to all regardless of chronic and acute health status and physical ability. However, those with physical health issues found the repetitive movements and upper body focus of some GYY classes induced or increased musculoskeletal pain. Conversely, those with minimal health issues found the GYY classes lacked physical challenge beyond gentle stretching.



*Dave: 'I thought it complemented what I'm doing with my [other diagnosis-related exercise] class really. Although a little different I have to say. I find the yoga gentler; the [other exercise] class is more challenging physically.'*

*Interviewer: 'Which do you prefer?'*

*Dave: 'I prefer the challenging one. A lot of the stuff so far in terms of the yoga we've been sat down. Well I never sit down when I have my exercise class with [Diagnosis-group]. ... No, it's not physically demanding enough. For me.'*

YPart-Int1

They often preferred their pre-existing physical routines to this new form of physical activity. Some questioned its potential to be beneficial as a form of exercise – with Walter noting that: 'but it was very easy, I thought, I didn't think that I was doing anything that would exercise me in any way ... to me, exercise should be getting the heart rate up and sweating' (YPart-Int1).

Several yoga teachers also noted that the GYY style of yoga may not be sufficiently challenging for more functionally able individuals they worked with. As Faye notes:

*But our bodies age, and our bodies change, and it's not a number. And there are some people who are very, very able. In my classes, in this trial, they're in their 70s and 80s but they could be doing so much more. And so I feel sometimes I'm holding them back a little bit.*

YTeach-Int1

Another yoga teacher, Jackie, noted that she felt 'like I've got people that are more able than this course is designed for' (YTeach-Int1). She describes her surprise at how her participants were 'really keen, really able' from the first class onwards. Some yoga teachers thought some participants would be better suited to attending general mat-based yoga classes. Several trial participants also queried the inclusion criteria of the trial, feeling that they should have been recruited based on health status rather than age. Others were uncomfortable being classified as an 'older' adult with multiple health issues. Participants in the younger end of the age range, and those who felt in good health, stated feeling out of place among participants older and in poorer health than them.

### ***Integration: variable engagement with Gentle Years Yoga, social time and home practice***

Participant interviewees held diverse views on the role of social time. Some described it as potentially beneficial, others as non-essential and independent of a GYY class. Irrespective of whether the classes were face to face or online, those in favour cited both social and educational benefits of social time as an opportunity to meet new people, gain peer support for those experiencing isolation and low mood and learn more about the practice of yoga. Notably, the value of social time was tied to the facilitation skill of inclusivity, independent of delivery format. Yoga teachers more adept at facilitation promoted inclusivity through moderating the group dynamic to ensure everyone had the opportunity to contribute to conversation if desired. However, a majority of participant interviewees had no interest in the social aspect of classes, and yoga teachers and participants noted that attendance was low across several classes. These participants commonly stated a lack of need or interest in socialising with a new group of people, noting they had joined the trial due to an interest in yoga not meeting new people. For example, Dot explained that 'I don't go for the social business' (YPart-Int1). Some also noted the communication issues inherent to group conversation via an online platform, outlining that there could be limited opportunity to verbally socialise for those of a quieter nature.

Reported adherence to home practice during the 12-week intervention and into the follow-up period was also variable, ranging from zero through to daily practice. One of the factors in adherence was the level of integration of the home practice content into the group classes by the yoga teachers. Some participants expressed disappointment in the lack of focus on home practice by their yoga

teachers. Some suggested more dynamic and progressive home practice options for those more physically capable, synchronising home practice with the progressive class content, including props and relaxation practices.

The primary facilitator of sustained self-practice was the perceived biopsychosocial benefit gained from practice. Those who continued to practice yoga at home into the follow-up period described a practice which had evolved over time. Initially basing their home practice on the content of the trial-provided home practice sheets, they gained a skillset from the 12-week intervention, enabling them to tailor the exercises, breathing and relaxation practices to their acute health needs, time commitments and/or personal preferences. For several participants, this skillset precluded the need for continued attendance at group-based classes, with a preference for self-directed practice at home. Conversely, the primary reason for a lack of home practice during or following the trial was a noted lack of benefit from, or interest in yoga, or a preference for pre-existing physical exercise routines. So Simon, who was 'a bit lukewarm with the exercises' initially said that he did not find home practice 'motivating' (YPart-Int1) and in the follow-up interview noted that 'Well, I didn't do very much home practice in the trial, to be honest' (YPart-Int2). Many non-adherent participants also noted a lack of extrinsic motivation for self-practice, particularly among those experiencing low mood or mental health challenges. Finally, several participants with work or social demands cited the time cost of trying to fit home practice into an already busy lifestyle.

### **Appraisal: embodied impact (and future engagement) varied**

The reported impact of GYY practice focuses around two main areas of improvements: physical health and emotional well-being (see [Table 39](#)). Tangible improvements in functional activities of daily living, such as dressing and housework, were also noted. Most notably, the impact of the GYY classes ranged from minimal to transformative. Some participants noted no impact of yoga on their health or lifestyle. This was primarily associated with describing a state of good health and physical activity when entering the trial, with the physical yoga content not being at a level capable of providing additional functional or sustained benefits. Irrespective of impact, some noted that taking part in the trial had stimulated an interest in engaging in other forms of yoga or physical activity, such as mat-based yoga, Tai Chi or more physically challenging forms of exercise.

In contrast, some participants described the transformative impact of the intervention had placed yoga as integral to their lives. For example, GYY has impacted Audrey at a biopsychosocial level, in a way she describes as 'lifechanging':

*It is the holistic bit. It's the mental, the physical, the intellectual, the social. [...] It closes down negative thoughts but opens up possibilities, and that very much resonates with me because that's the sort of person I am. It puts me in control. But it's something about letting go of control as well, if that makes sense. It's just okay to be me. And, yes it can be one of many things, but it is – it is lifechanging. That is without a doubt.*

YPart-Int2

Continued engagement was aligned with initial and sustained biopsychosocial improvements in physical health, emotional well-being and independence over the 12-week intervention, with associated decreases in analgesic, opiate-based medication.

Some participants noticed physical and psychosocial improvements from within the first few classes. For example, after one class, Rose felt a notable impact:

*Rose: 'And my first thing that I felt came from it really, was that first night of doing yoga, after I'd done that first class, and it was very gentle, I actually didn't have any pain at all going to bed for the first time in years.'*



**TABLE 39** Examples of reported physical and psychosocial impacts of GYY on participants

Domain	Description	Examples of Impact
<b>Primary physical impacts</b>	<p>Included improved muscle strength, musculoskeletal pain and stiffness, postural awareness, mobility, balance and co-ordination and upper body range of motion.</p> <p>In terms of pain, the practice of yoga was tied to a reduction in pain intensity, an ability to detach from the pain, and consequently, an improvement in pain-related mental health.</p> <p>In relation to posture and balance, participants noted a heightened spatial awareness of their body's interaction with the external environment. This in turn led to decreased episodes of feeling at risk of falling, both within and outside the home environment.</p>	<p><u>None/Minimal</u>: Frank was 'hoping it would help me strengthen my leg muscles which are a bit weak, and help me to walk better'. He thought through doing yoga that his balance has 'improved slightly. But because the legs are a bit weak you have to fight it, especially first thing in the morning'. (YPart-Int1)</p> <p><u>Moderate</u>: Gwen, discussed with the yoga teacher, specific problems she was having. 'So I said, "the thing that I was noticing, when driving the car, it was becoming more difficult to look left and right, with the turning". And she did address it, and because I'm still doing the exercises every day, I can now turn my head'. (YPart-Int1)</p> <p><u>Substantive</u>: Heather noted a reduction in pain from the first class. 'It was, it was amazing. The first class that we did, after we'd done the class, or the exercises, I was thinking "oh, some of this is really hurting", and then we did the relaxation and it went. Didn't stay away but it went, and it went for several hours, I can't remember now exactly how long but it went for a considerable length of time. And that for me was astonishing because I never have no pain. Never have no pain.</p> <p>...</p> <p>When I had the shingles, and I wasn't well, so I hadn't done the yoga for a while, and I went back to doing it and I did it for 2 days on the run and the second day I came down and I said to [Husband] "I'd forgotten how much difference this made", because it just does. ... But because he said to me, "When are you going to start your yoga again?" Because he's like that, [Husband], if he thinks something is doing me good he'll poke me until I do it' (YPart-Int2)</p>
<b>Primary psychosocial impacts</b>	<p>Improved management of sleep, emotional and mental well-being.</p> <p>These improvements were notably tied to a new-found awareness of the breath, with conscious breathing regularly cited as a powerful tool for participants to manage both chronic and acute stress and anxiety.</p>	<p><u>None/Minimal</u> – Anne noted that 'Well, it's doing things that I wouldn't normally do. Or, as I say the breathing, I don't normally do. It's making you think about your breathing. Because I just do things without breathing, and I know I should be, you know (mimics controlled breathing) more' (YPart-Int1)</p> <p><u>Moderate</u> – The impact of yoga for Marie was 'in terms of, I think, thinking about breathing, which I hadn't done before. So, I have actually used some of the techniques I've learnt when, you know, something was bothering me or I was getting in a panic about something. So that, I've never sort of realised was so important and, you know, what a difference it could make' (YPart-Int1)</p> <p><u>Substantive</u>: Mindful breathing was now central to June's everyday life. 'Well, I learned how to relax and that's been very useful. I don't sleep well, and what [trial YT] taught us I was able to learn how to actually, to breathe into it and to find a way of stopping my mind racing. Getting to bed, I go over everything and then I think: "did I do? I should have done. I wonder if?" And I've learnt how to calm all that down. So, it has helped me a great deal with sleeping.</p> <p>...</p> <p>I think it's here to stay. Whether I go back to the classes or I don't it's here to stay. But I know I've got a tool I can use that will calm me down and will help with the pain.<sup>a</sup> So that's a really useful tool to have. ... learning how to breathe properly, and to go into the moment, that was more important – it's given me a tool that I can use, I can use anywhere' (YPart-Int2)</p>

<sup>a</sup> June also found yoga impacted on her rheumatoid arthritis-related pain through reducing pain intensity, enabling detachment from it and improving pain-related mental health. This, in turn, reduced her pharmaceutical analgesic use.

Interviewer: 'After just one class?'

Rose: 'After just one class, I couldn't believe it. I was a little tiny bit stiff, in the more sensitive areas of the knee and things like that, but that wasn't unbearable by any stretch of the imagination. So, I just thought, well, you know, this has got real good benefits. And it was just like flowing in water, really, because I really felt into it. Once I'd started, it was just like, you know, swimming. You just do it [laughter].'

YPart-Int1

For others, the improvements occurred more gradually over the 3-month trial period and/or beyond. Tom's awareness of the impact of yoga on his well-being increased once the classes had stopped:

*It opened up something of which I was unaware. Because I was unaware, I could see no value. I could put no value on it and I could see no value in it beforehand. And now looking back, I can. [...] I can't say there's like this little sort of magic moment of revelation or something like that, but I suppose that glimpse or that awareness of that sort of connection of mind, body, brain, movement, you know, just going from a flash to being in that world for a little bit of time, into being in that world for a slightly longer period of time. That awareness grew during the course.*

Ypart-Int2

Some had engaged with (or intended to) post-trial GYY classes, as well as different types of local yoga classes (some with different yoga teachers) and different styles of yoga, including mat-based for those who found the GYY classes lacking in physical challenge. Others had integrated self-directed yoga practice as part of their home routine, having a set time and place in the home for a daily or multiweekly practice.

Participants who had not continued with yoga post trial described two main reasons: a lack of benefit from the 12-week intervention and a lack of self-motivation to find a new class or engage with self-practice. Tom reflected on his lack of continuation:

*And now it's [GYG classes] gone out of my life and I've been talking to you about how beneficial it's been and things like that, and yet I've not made any efforts to go back to it, or yet if I had an email saying, 'There's a follow-up class,' I'd say, 'Yeah, ding.' That's my experience. That's a very strange combination.*

Ypart-Int2

Decreased self-motivation was often associated with low mood rather than lack of interest in yoga, with these participants suggesting some form of extrinsic motivation may be enough to re-engage them with a regular practice. Additionally, some of the more physically able participants noted that while GYY was not something they were interested in pursuing further at the present time, it presented as a suitable option for when they were physically unable to continue with their current physical activities.

### **Reconfiguration: introducing and embedding gentle yoga**

Yoga teachers, trial participants and other stakeholders were generally supportive of the integration of yoga within the NHS as a means of providing a more holistic approach to health care than the predominant biomedical model. The potential biopsychosocial health benefits of yoga were viewed as aligning with the initial NHS vision of preventative medicine, focusing on maintaining health rather than treating illness, as well as new agendas like social prescribing. They outlined a range of factors that would be central to consider in any future roll-out of GYY in the UK context (see [Table 40](#)).

Many favoured the representation of yoga as a proactive health behaviour aimed at health maintenance, sustained independence and illness prevention independent of age. The impact of a health condition was deemed more relevant than its presence. Several participants noted that the option of a gentle seated form of yoga, such as GYY, would have been useful when many of their chronic health conditions presented in their 30s–50s. Age was not viewed as indicative of health status or ability, with people noting little commonality or homogeneity among 'older adults' as a group beyond their similar ages.

**TABLE 40** Examples of recommendations on potential implementation strategies to support any future roll-out of GYY

Focus of implementation work	Examples of relevant implementation strategies <sup>a</sup>	Examples of activities needed and issues to be understood
Public preconceptions of yoga	<p><i>Use mass media</i></p> <p><i>Develop educational materials</i></p> <p><i>Distribute educational materials</i></p>	<p>Media campaigns needed to connect with those who either don't know about yoga or who don't see it as something that they could do, for example,</p> <ul style="list-style-type: none"> <li>• Need to actively reject the stereotype of yoga being for young, white, slim, middle-class females and portray the broad demographic of people who could potentially benefit from yoga</li> <li>• Older people need to see 'themselves' doing GYY</li> <li>• Address media stereotypes of older adults – negative portrayal of frailty and misleading portrayal of being in the healthy prime of their lives. Need to balance the use of visual imagery to promote a 'realistic version' of the target demographic and what engaging in a GYY class would look like</li> </ul>
Health professionals' preconceptions of yoga	<p><i>Inform local opinion leaders</i></p> <p><i>Obtain and use patients/consumers and family feedback</i></p> <p><i>Make training dynamic</i></p>	<p>Provide different forms of evidence, for example,</p> <ul style="list-style-type: none"> <li>• Outline potential health benefits, safety and reduced service demand shown in GYY trials</li> <li>• Provide direct GYY participant feedback on physical and mental health benefits</li> <li>• Provide taster classes to enable them to understand the potential suitability of GYY for their older patients</li> </ul>
Delivery of GYY classes	<p><i>Promote adaptability</i></p> <p><i>Identify early adopters</i></p> <p><i>Build a collation</i></p>	<p>Align with existing models of provision, for example,</p> <ul style="list-style-type: none"> <li>• 12-week course duration of GYY is disparate with NHS rehabilitation and fitness industry courses, which typically run in 6- to 8-week blocks</li> <li>• Social prescribing, GP referral schemes, annual health checks, specialist nurse clinics have effective information delivery infrastructures and referral pathways amenable to GYY delivery</li> <li>• Commissioners tend to commission organisations rather than individuals</li> </ul>
Inequalities	<p><i>Inform local opinion leaders</i></p> <p><i>Alter patient/consumer fees</i></p> <p><i>Promote adaptability</i></p>	<p>Adjust practices to support more inclusion, for example,</p> <ul style="list-style-type: none"> <li>• Work with someone with direct links to underserved groups – same community, ethnicity, language – to raise awareness about GYY and provide inclusive options (e.g. women-only classes)</li> <li>• Commercial GYY programmes can be used to help support some subsidised or free places and/or classes</li> <li>• Hybrid and online GYY classes can enable accessibility for some who have difficulty accessing community-based classes due to physical or mental health issues or carer responsibilities, overcoming transport or weather barriers to attending community-based classes</li> </ul>

<sup>a</sup> Strategies derived from a refined compilation of implementation strategies that was generated as part of the Expert Recommendations for Implementing Change project.<sup>63</sup>

Notably, several yoga teachers reported not putting age restrictions on their current community-based GYY and related classes. Alison, a yoga teacher, noted that:

*So, in my chair-based classes, I have somebody with rheumatoid arthritis who is in her fifties. I think that the age cut-off is – I wouldn't say it's irrelevant, but I suppose you've got to ... Do we need it? I don't know if we do need it, you know, a chair-based class ... I don't enforce it, in my chair-based class. It's just anybody who would prefer to do something – you know, whatever we do on a mat, we do on a chair. So, if people can't get down onto the floor and up off the floor easily then they come to my chair-based class.*

*Yteach-Int1*

As such, GYY could focus on inclusivity rather than age, promoting GYY as a beginner's form of gentle yoga, suitable for those who are currently non-active or those unable or uninterested to engage with a mat-based class.

Moving the emphasis of GYY from age to inclusivity may facilitate additional engagement of GYY classes among health professionals and patients. The potential target demographic could include those with mobility, mental health and rehabilitation issues. Gentle Years Yoga was seen as suitable for those who are wheelchair dependent and for a younger demographic with health-related mobility restrictions who require a gentler style of yoga. As one participant noted, 'People I think with chronic [illnesses] could easily, you know, keep up with it. ... I think it's open for all, open for all' (Joy, Ypart-Int1). Broadening the GYY remit to include people based on their physical ability would provide accessibility to a younger demographic with chronic health needs, to individuals in transition from injury or surgery and to experienced yoga practitioners no longer able to meet the physical requirements of a general yoga class due to onset or progression of health issues.

## Discussion

Participant engagement with the GYY trial was sustained throughout the life of the trial. Initially, participant interviewees agreed to take part for a range of reasons. Many understood GYY as a potential age-appropriate way to impact on their health and well-being. Some also saw it as an opportunity to socialise with others as well as a way to motivate themselves to exercise. Some were curious to discover more about (age-appropriate) yoga and, as with all trials, people reported (conditional) altruism.<sup>64</sup> In so doing, nearly all that agreed to join the trial, be they allocated to the intervention or usual care groups, demonstrated a clear desire to make some form of change in their health and well-being in and through what they understood as an age-appropriate physical activity. In relation to ideas around the COM-B model of behaviour change,<sup>65</sup> participants felt they had *capability* – notably, some even felt that they exceeded the inclusion criteria but that they were too healthy to be included. The trial itself offered an *opportunity* to make a change, and recruitment aligned with a moment when they had *motivation* to change. For those in the intervention arm, the trial also provided further extrinsic *motivation* through the provision of weekly classes.

Over time, participants actively engaged with the classes. The GYY style of yoga was viewed as a suitable and safe form of physical movement for older adults with varying health issues. Most notably, this was independent of the delivery format. As a recent systematic review of online yoga also shows,<sup>66</sup> GYY, via synchronous online delivery, is clearly feasible, acceptable and satisfactory to participants. However, of the GYY participants who took part in the qualitative work, there was no clear preference for one delivery format – distal/online or proximal/face to face – over the other. While they found online provision satisfactory, especially during the periods of COVID-19 lockdown, some preferred face-to-face contact once restrictions were lifted. Some favoured online for pragmatic reasons, including removal of barriers related to transport, adverse weather and acute health status. Some favoured face to face for physical presence and support of peers and of the yoga teacher. Most yoga teachers also continued to offer some online general and GYY classes since restrictions lifted. As with the Australian Successful AGEing (SAGE) yoga trial for people aged 60 years and older that went online due to COVID-19,<sup>67</sup> the yoga teachers and participants responses in this study demonstrate that online yoga is suitable for an older adult cohort. The participants challenged some stereotypes around the information and communication technology capabilities of older adults<sup>68,69</sup> and demonstrated the potential for online delivery of group-based health-related services for older adults.<sup>70</sup> However, engagement in social time, after the formal movement and meditative aspects of the class, was variable. In part, this tied to the presence or absence of a desire to socialise with others. For some, yoga was the motivation to take part, not additional social interaction with others. Relatedly, engagement with home practice was also variable, with reported adherence ranging from zero through to daily practice. Echoing the SAGE yoga trial,<sup>71</sup> engagement was mediated by the perceived biopsychosocial benefit gained from practice.

The majority of participant interviewees viewed GYY as a form of gentle exercise with mindful breathing. Only two found the class content physically challenging. Many outlined that GYY was an age-appropriate exercise, particularly for those with no previous experience of physical activity or those seeking a less demanding alternative to prior activities. Aligning with the SAGE yoga trial and other studies on yoga with a range of populations,<sup>72-74</sup> participants did value the focus on breathing, mindful movement and the calmness this could provide. Those who saw themselves with minimal health issues found the GYY classes lacked physical challenge beyond gentle stretching and preferred their pre-existing physical routines. Notably, irrespective of the high level of multimorbidity, most participants viewed their health as good. They made it very clear that multimorbidity does not equate with being in poor health, and that it was the presentation, not the presence, of a health condition that determined its impact on them. Echoing many of the participants, an editorial by Valderas outlines multimorbidity should not be understood as 'a health condition in its own right'.<sup>75</sup> Valderas notes that:

*[M]ultimorbidity has come to replace the notion of complexity in clinical care and in policy documents. A number of yet unresolved complexity issues – which have long captured the attention of healthcare managers and policy makers – are being reframed or rather simply relabelled as multimorbidity. However, the root of these problems actually lies elsewhere, most frequently in the severity of the conditions rather than in their mere presence.*

*ibid: 213–214*

Ford and Ford similarly note that multimorbidity is 'doctor- and research-centred rather than patient-centred' as it focuses on 'counting diseases' over people's 'lived experience' (p. 7).<sup>76</sup> Echoing the participants, they note that 'patients with multiple health issues view their health in terms of function, symptoms and well-being' (p. 7). As such, the trial participants may fit a category of 'multimorbidity', but such labelling hid a wide diversity in functional ability, symptoms severity and well-being with the participants interviewed and observed. Participants and yoga teachers noted a good level of functional ability. Participants routinely reported low symptom severity and good well-being.

As such, the process evaluation shows that participants were highly motivated and engaged well with the GYY classes over time. As well as through engaging with GYY classes, relatively healthy participants felt that they had undertaken a form of gentle exercise with mindful breathing. However, the impact of this process varied dramatically. Some participants noted no impact of GYY on their health or lifestyle. Some described a modest impact on aspects of health, including physical, psychological and self-management benefits, as seen in other recent studies.<sup>67,77</sup> Some participants described GYY as transformative, having substantive impacts and improvements on their physical health and emotional well-being. In some senses, the process of engaging in the GYY intervention enabled each participant to take part in their own *N-of-1* trial.<sup>78,79</sup> They were their own chief investigator, their own study participant – the individual unit of observation – collecting real-time data on the efficacy of GYY and other interventions. In and through taking part in the intervention arm of the trial, they discovered what was the right sort of intervention for them. For some, GYY was the optimal intervention; for others, other forms of yoga – some mat based – were seen as optimal; and for others, other forms of new and pre-existing physical and/or medicative activities were optimal. Irrespective of the aggregated trial result, this was a meaningful experience for many. It enabled them to learn what form of activity was right for them, given their physical (and emotional) well-being at a specific point in time.



## Chapter 6 Methodological substudies

This chapter describes four methodological substudies embedded within the GYY host trial, which addressed the following questions:

- What is the concurrent validity of the PROMIS-29 with the EQ-5D-5L?
- Does including £5 and/or a pen in the recruitment pack enhance recruitment?
- Does sending a pen with a follow-up questionnaire enhance return rates?
- Does offering a free yoga session to control participants after the 12-month follow-up assessment enhance retention and reduce contamination?

### Substudy 1: What is the concurrent validity of the PROMIS-29 with the EQ-5D-5L?

#### Introduction

Health-related quality of life has been defined as 'how well a person functions in their life and his or her perceived well-being in physical, mental, and social domains of health'.<sup>80</sup> A common mode of assessing HRQoL is through patient-reported outcome measures (PROMs), via self-administered surveys (e.g. postal) or interviewer administration (e.g. telephone). Measuring HRQoL accurately is important in a range of fields, including research, policy-making and informing clinical care. Two HRQoL PROMs were collected in this trial, the PROMIS-29 and the EQ-5D-5L.

The PROMIS was developed by the US National Institute of Health (NIH) using item response theory to calibrate a large number of questions for each health domain to create an item bank.<sup>81</sup> PROMIS has over 100 health domains with a calibrated 'item bank' of questions that can be administered via standard short forms, custom short forms or computer adaptive testing.<sup>82</sup> PROMIS aims to address the lack of generalisable and universal PROMs to assess many HRQoL domains.<sup>82</sup> It has undergone extensive psychometric testing for validity and reliability in the general adult population, but it has not been widely used in older adults with multimorbidity<sup>83</sup> and has received limited attention in the UK.

The EQ-5D has been validated in many different patient populations including diabetes, cardiovascular problems, chronic obstructive pulmonary disease, cancer, chronic pain and rheumatoid arthritis. The EQ-5D has some limitations arising from the descriptions of health used in the measure, which include (1) large proportions of the respondents scoring at the very top or very bottom of the scale (i.e. ceiling effects in the very healthy or floor effects in the very ill), (2) imprecise measurement for individuals, (3) poorly worded questions such as those that combine concepts (double-barrelled items) and (4) lack of coverage of the full range of health.<sup>84</sup> PROMIS addresses several of these limitations, including: (1) capturing a wider range of each health domain, (2) measuring individual health status with greater precision and (3) using rigorously designed and tested questions.<sup>84</sup>

Although EQ-5D-5L and PROMIS-29 are both concise generic measures that allow the estimation of HRQoL, a recent systematic review of studies comparing EQ-5D and PROMIS-29 instruments has noted the lack of research comparing their measurement properties and utilities.<sup>85</sup> The impact of choosing either instrument for research, policy and clinical care decision-making is therefore unclear, and more research is needed to understand the relationship between the two instruments.

#### Objectives

To assess the concurrent validity of the PROMIS-29, a short-form version of the whole PROMIS system, against the EQ-5D-5L in adults aged 65 years or older with multimorbidity.

## Methods

### Study design

This non-randomised, methodological (non-interventional) study within a trial (SWAT) was embedded into the main study. The SWAT was registered with the Northern Ireland Network for Trials Methodology Research SWAT repository on 1 April 2018 (SWAT95; [www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/SWATSWARInformation/Repositories/SWATStore/](http://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/SWATSWARInformation/Repositories/SWATStore/)). Both the EQ-5D-5L and PROMIS-29 were delivered to all participants at baseline and 3, 6 and 12 months post randomisation in the main study. The concurrent validity and responsiveness to change of the PROMIS-29 were measured relative to the well-established EQ-5D-5L.

### Participants

All of the 454 participants allocated to either the intervention or usual care arm of the GYY trial were eligible for inclusion in this substudy.

### Outcome measures

The primary outcome of this substudy was the concurrent validity of the PROMIS-29 (v2.1) relative to the EQ-5D-5L to see how well it compares to the well-established EuroQoL test. Secondary outcomes were: internal and external responsiveness of the PROMIS-29 MCS and PCS and EQ-5D-5L (index value and VAS) to test the ability of the PROMIS-29 to detect clinically important change.

### Statistical analysis

The PROMIS-29 and EQ-5D-5L scores were calculated as described in [Chapter 2](#).

The concurrent validity of the PROMIS-29 was measured relative to the EQ-5D-5L. At all time points, correlation ( $\rho$ ) was calculated between each of the seven domain scores and the MCS and PCS of the PROMIS-29 and:

- the EQ-5D-5L index value, using Pearson's correlation coefficient
- the EQ-5D-5L VAS score, using Pearson's correlation coefficient
- each domain of the EQ-5D-5L, using Kendall's tau b.

The internal responsiveness of the PROMIS-29 (MCS and PCS) and EQ-5D-5L (index value and VAS) was evaluated by comparing changes in scores using a paired *t*-test. Comparisons were made between the following time points:

- baseline and 3 months
- baseline and 6 months
- baseline and 12 months
- 3 and 6 months
- 3 and 12 months
- 6 and 12 months.

The standardised response mean (SRM; mean change score divided by the SD of the change scores) and Cohen's *d* effect size (mean change score divided by the SD of the earlier score) for each measure and time interval were calculated with bootstrapped 95% CIs (1000 iterations). A common interpretation of the magnitude of effect sizes is that 0.2 is considered small, 0.5 moderate and 0.8 large.<sup>86</sup>

The external responsiveness of the PROMIS-29 was evaluated by correlating change scores of the MCS and PCS with the change scores for EQ-5D-5L index values and VAS score. Participants were categorised according to whether or not their scores changed (increased) by the minimal clinically important difference (MCID) of 0.06 on the EQ-5D-5L index value between each time interval. Receiver operating characteristic (ROC) curves were produced to assess how well the PROMIS-29 MCS and

PCS discriminate between improved and non-improved participants (defined as EQ-5D-5L change score > MCID of 0.06) for the time interval between baseline and 12 months. An AUC value of 0.5 indicates a discriminatory value equivalent to chance. Analyses were conducted in Stata version 17.<sup>34</sup>

## Results

Summary scores for the PROMIS-29 and EQ-5D-5L outcomes at each time point are presented in [Chapter 3](#).

Correlations between the PROMIS-29 and EQ-5D-5L scores are presented in [Table 41](#). For the EQ-5D-5L index value and VAS score, and the PROMIS-29 physical function and social participation subscales and both the PCS and MCS, a higher score indicates a more favourable outcome, whereas for all other scores a lower score indicates a better outcome. Hence, some correlations are negative. At all time points, the strongest correlation was between the EQ-5D-5L index value and the PROMIS-29 PCS ( $\rho \geq 0.74$ ). The EQ-5D-5L utility index score also correlated well with the PROMIS-29 MCS ( $\rho \geq 0.66$ ).

Both the PROMIS-29 MCS and PCS showed statistically significant changes between baseline and all post-randomisation time points and for the PCS between months 3 and 12 and months 6 and 12 (see [Table 42](#)). The EQ-5D-5L index and VAS scores showed statistically significant changes between baseline and month 6, baseline and month 12, months 3–6 and months 3–12. On average, scores for all four measures decreased between baseline and month 12, suggesting a slight deterioration in HRQoL among responders. However, according to Cohen's *d* statistics, changes between time intervals for all measures were minimal to small only (< 0.2), indicating little to no meaningful change over the 12-month follow-up. The Cohen's *d* effect size for the change from baseline to 12 months was similar for the four measures (range -0.19 to -0.16).

Positive and statistically significant, though reasonably weak (maximum 0.57), correlations were observed between the change in EQ-5D-5L index value and VAS and the change in PROMIS-29 MCS and PCS (see [Table 43](#)). Correlations were reasonably stable when compared across the different time intervals. Change in the EQ-5D-5L index value tended to correlate better with PROMIS-29 change scores than the change in VAS did; for example, the correlation between the change in EQ-5D-5L index value between baseline and month 12 correlated with the PROMIS-29 PCS at 0.57 and with the MCS at 0.50, whereas correlations with the EQ-5D-5L VAS score were 0.30 and 0.37, respectively. The strongest correlation at each time interval was between the change in EQ-5D-5L index value and change in PROMIS-29 PCS, except for baseline to month 3, when this was comparable with the correlation with change in the MCS.

Between baseline and month 12, 21.8% of responders (at both time points) reported an increase in EQ-5D-5L utility index score that equated to 0.06 or more, which is the MCID (see [Table 44](#)).

Receiver operating characteristic curves indicated that both the PCS and MCS were able to discriminate well between participants who achieved the EQ-5D-5L index value MCID of an increase of 0.06 or more between baseline and month 12 (AUC 0.74 and 0.70, respectively) (see [Figure 10](#)).

## Discussion

The PROMIS-29 demonstrated reasonable concurrent validity relative to the EQ-5D-5L in community-dwelling adults aged 65 years and above with multimorbidity. Correlations between the PROMIS-29 and EQ-5D-5L were all in the expected direction. The utility index scores and the VAS correlated well with the PCS and MCS at all time points ( $\rho \geq 0.58$ ); however, the strength of correlations between the five individual domains of the EQ-5D-5L and the seven subscale T-scores of the PROMIS-29 varied. For instance, at baseline, correlations of 0.6 and over were observed between (order always EQ-5D-5L domain, then PROMIS-29 subscale): mobility and physical functioning; usual activities and physical functioning; pain/discomfort and pain interference; and anxiety/depression and depression. Other correlations were weaker.



**TABLE 41** Correlations between the PROMIS-29 and EQ-5D-5L scores

EQ-5D-5L	PROMIS-29								
	PF	Anx	Dep	Fatigue	SD	SP	PI	PCS	MCS
<b>Baseline</b>									
<i>Index value</i>	0.71	-0.35	-0.44	-0.54	-0.42	0.57	-0.69	0.74	0.66
<i>VAS score</i>	0.59	-0.26	-0.38	-0.51	-0.35	0.55	-0.54	0.61	0.59
<i>Mobility</i>	-0.62	0.07	0.15	0.27	0.19	-0.38	0.50	-0.59	-0.34
<i>Self-care</i>	-0.40	0.17	0.23	0.26	0.21	-0.31	0.32	-0.38	-0.30
<i>Usual activities</i>	-0.60	0.20	0.28	0.42	0.28	-0.53	0.52	-0.60	-0.49
<i>Pain/discomfort</i>	-0.46	0.14	0.19	0.34	0.29	-0.35	0.64	-0.50	-0.43
<i>Anxiety/depression</i>	-0.20	0.53	0.60	0.40	0.24	-0.29	0.18	-0.22	-0.42
<b>Month 3</b>									
<i>Index value</i>	0.76	-0.36	-0.45	-0.63	-0.39	0.63	-0.72	0.79	0.72
<i>VAS score</i>	0.62	-0.33	-0.40	-0.61	-0.42	0.61	-0.60	0.65	0.68
<i>Mobility</i>	-0.66	0.07	0.13	0.32	0.15	-0.42	0.47	-0.62	-0.36
<i>Self-care</i>	-0.45	0.09	0.19	0.32	0.21	-0.33	0.34	-0.42	-0.32
<i>Usual activities</i>	-0.69	0.13	0.19	0.40	0.21	-0.53	0.57	-0.67	-0.47
<i>Pain/discomfort</i>	-0.57	0.19	0.21	0.43	0.27	-0.43	0.67	-0.60	-0.50
<i>Anxiety/depression</i>	-0.15	0.59	0.63	0.38	0.26	-0.28	0.19	-0.19	-0.43
<b>Month 6</b>									
<i>Index value</i>	0.75	-0.42	-0.50	-0.59	-0.35	0.63	-0.74	0.79	0.72
<i>VAS score</i>	0.57	-0.36	-0.41	-0.56	-0.28	0.49	-0.53	0.60	0.61
<i>Mobility</i>	-0.66	0.17	0.25	0.34	0.16	-0.43	0.48	-0.63	-0.40
<i>Self-care</i>	-0.46	0.20	0.27	0.28	0.18	-0.35	0.38	-0.44	-0.34

EQ-5D-5L	PROMIS-29								
	PF	Anx	Dep	Fatigue	SD	SP	PI	PCS	MCS
<i>Usual activities</i>	-0.68	0.25	0.30	0.43	0.20	-0.52	0.54	-0.67	-0.50
<i>Pain/discomfort</i>	-0.55	0.21	0.26	0.40	0.24	-0.40	0.69	-0.58	-0.48
<i>Anxiety/depression</i>	-0.25	0.55	0.62	0.41	0.26	-0.38	0.28	-0.28	-0.50
<b>Month 12</b>									
<i>Index value</i>	0.74	-0.44	-0.47	-0.59	-0.37	0.68	-0.73	0.78	0.72
<i>VAS score</i>	0.54	-0.41	-0.44	-0.57	-0.36	0.56	-0.59	0.58	0.64
<i>Mobility</i>	-0.66	0.18	0.21	0.34	0.15	-0.46	0.50	-0.63	-0.39
<i>Self-care</i>	-0.49	0.20	0.20	0.33	0.14	-0.41	0.35	-0.47	-0.36
<i>Usual activities</i>	-0.66	0.23	0.28	0.40	0.17	-0.55	0.54	-0.64	-0.46
<i>Pain/discomfort</i>	-0.51	0.21	0.23	0.38	0.27	-0.45	0.65	-0.55	-0.47
<i>Anxiety/depression</i>	-0.19	0.65	0.66	0.44	0.30	-0.35	0.25	-0.23	-0.49

Anx, anxiety; Dep, depression; PF, physical function; PI, pain interference; SD, sleep disturbance; SP, social participation.

**TABLE 42** Mean change score (paired *t*-test) and responsiveness estimates of SRM and Cohen's *d* effect size for each measure and time interval

Time interval	Change metric	EQ-5D-5L index value	EQ-5D-5L VAS	PROMIS-29 PCS	PROMIS-29 MCS
3M - Baseline	<i>n</i>	413	413	406	406
	Mean change score (95% CI), <i>p</i> -value	-0.002 (-0.015 to 0.011), 0.75	0.33 (-1.15 to 1.81), 0.66	-0.59 (-1.14 to -0.04), 0.04	-1.16 (-1.72 to -0.60), < 0.001
	SRM (95% CI), <i>p</i> -value	-0.016 (-0.111 to 0.081), 0.75	0.02 (-0.08 to 0.12), 0.67	-0.10 (-0.20 to -0.01), 0.03	-0.20 (-0.29 to -0.11), < 0.001
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.012 (-0.090 to 0.065), 0.76	0.02 (-0.07 to 0.10), 0.67	-0.07 (-0.13 to -0.01), 0.03	-0.14 (-0.21 to -0.07), < 0.001
6M - Baseline	<i>n</i>	397	395	384	384
	Mean change score (95% CI), <i>p</i> -value	-0.022 (-0.039 to -0.005), 0.01	-1.92 (-3.65 to -0.19), 0.03	-0.82 (-1.45 to -0.19), 0.01	-1.35 (-1.95 to -0.75), < 0.001
	SRM (95% CI), <i>p</i> -value	-0.131 (-0.229 to -0.032), = 0.01	-0.11 (-0.21 to -0.01), = 0.03	-0.13 (-0.23 to -0.03), = 0.01	-0.22 (-0.32 to -0.13), < 0.001
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.130 (-0.229 to -0.031), 0.01	-0.11 (-0.21 to -0.01), 0.03	-0.09 (-0.16 to -0.02), 0.01	-0.17 (-0.23 to -0.09), < 0.001
12M - Baseline	<i>n</i>	394	394	377	377
	Mean change score (95% CI), <i>p</i> -value	-0.032 (-0.050 to -0.014), < 0.001	-2.87 (-4.73 to -1.01), 0.003	-1.49 (-2.18 to -0.80), < 0.001	-1.59 (-2.24 to -0.94), < 0.001
	SRM (95% CI), <i>p</i> -value	-0.178 (-0.277 to -0.078), < 0.001	-0.15 (-0.25 to -0.05), 0.003	-0.22 (-0.32 to -0.12), < 0.001	-0.25 (-0.34 to -0.15), < 0.001
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.189 (-0.293 to -0.085), < 0.001	-0.16 (-0.26 to -0.06), 0.002	-0.17 (-0.25 to -0.09), < 0.001	-0.19 (-0.27 to -0.12), < 0.001
6M - 3M	<i>n</i>	392	388	376	376
	Mean change score (95% CI), <i>p</i> -value	-0.017 (-0.030 to -0.004), 0.01	-1.90 (-3.38 to -0.42), 0.01	-0.16 (-0.68 to 0.36), 0.54	-0.17 (-0.65 to 0.31), 0.48
	SRM (95% CI), <i>p</i> -value	-0.126 (-0.228 to -0.025), 0.01	-0.13 (-0.23 to -0.02), 0.02	-0.03 (-0.13 to 0.07), 0.53	-0.04 (-0.14 to 0.06), 0.48
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.092 (-0.163 to -0.020), 0.01	-0.11 (-0.20 to -0.03), 0.01	-0.02 (-0.08 to 0.04), 0.53	-0.02 (-0.08 to 0.04), 0.48

Time interval	Change metric	EQ-5D-5L index value	EQ-5D-5L VAS	PROMIS-29 PCS	PROMIS-29 MCS
12M – 3M	<i>n</i>	385	358	368	368
	Mean change score (95% CI), <i>p</i> -value	-0.030 (-0.045 to -0.015), < 0.001	-3.07 (-4.71 to -1.43), < 0.001	-0.84 (-1.43 to -0.25), 0.01	-0.36 (-0.93 to 0.21), 0.22
	SRM (95% CI), <i>p</i> -value	-0.201 (-0.304 to -0.098), < 0.001	-0.19 (-0.29 to -0.08), < 0.001	-0.15 (-0.25 to -0.04), 0.01	-0.06 (-0.17 to 0.04), 0.23
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.163 (-0.242 to -0.083), < 0.001	-0.18 (-0.28 to -0.08), < 0.001	-0.09 (-0.16 to -0.03), 0.01	-0.04 (-0.11 to 0.03), 0.23
12M – 6M	<i>n</i>	386	382	365	365
	Mean change score (95% CI), <i>p</i> -value	-0.011 (-0.025 to 0.003), 0.12	-0.88 (-2.45 to 0.69), 0.27	-0.72 (-1.32 to -0.12), 0.02	-0.28 (-0.79 to 0.23), 0.29
	SRM (95% CI), <i>p</i> -value	-0.079 (-0.176 to 0.017), 0.11	-0.06 (-0.16 to 0.04), 0.27	-0.12 (-0.23 to -0.02), 0.02	-0.06 (-0.16 to 0.05), 0.30
	Cohen's <i>d</i> (95% CI), <i>p</i> -value	-0.052 (-0.117 to 0.013), 0.12	-0.05 (-0.13 to 0.04), 0.28	-0.08 (-0.15 to -0.01), 0.02	-0.03 (-0.09 to 0.03), 0.30

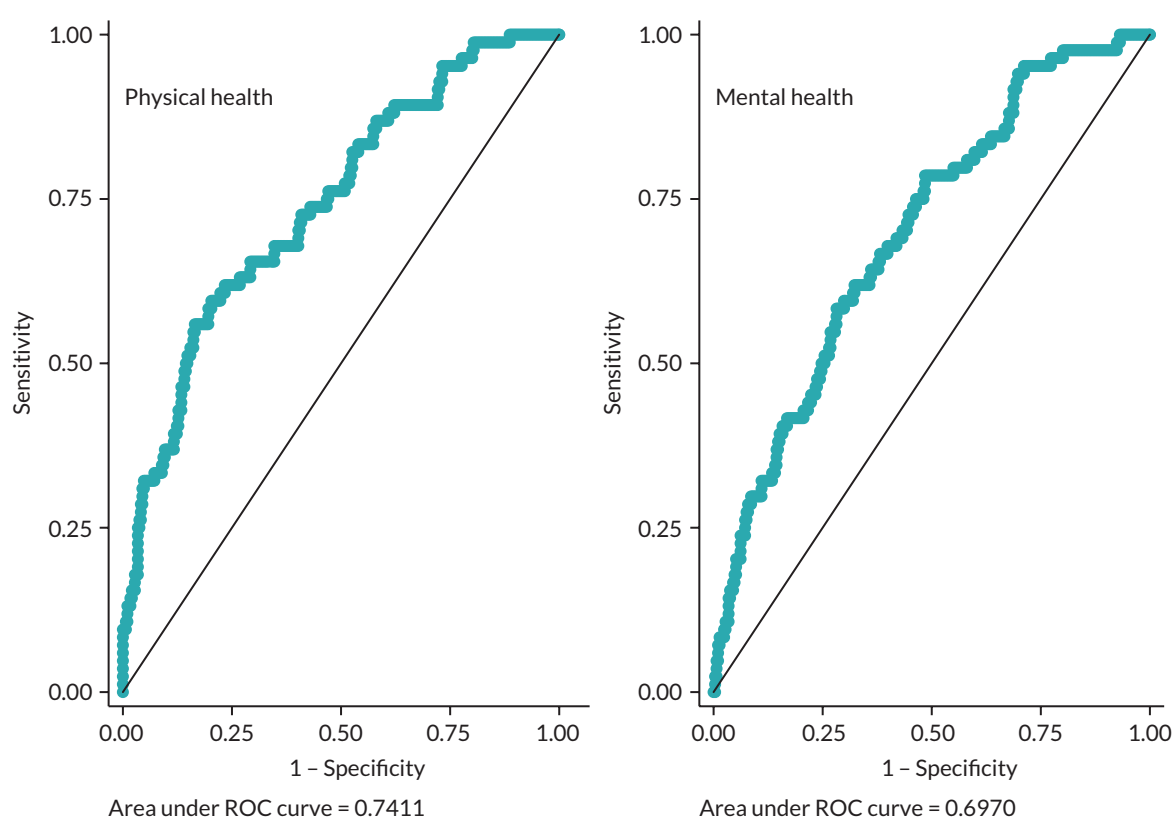
**TABLE 43** Correlation of change scores of EQ-5D-5L with change scores for PROMIS-29

<b>M3 vs. baseline</b>		<b>N</b>	<b>Correlation (95% CI)</b>	<b>p-value</b>
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	405	0.45 (0.37 to 0.52)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	405	0.46 (0.38 to 0.53)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	405	0.27 (0.18 to 0.36)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	405	0.31 (0.22 to 0.39)	< 0.001
<b>M6 vs. baseline</b>				
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	383	0.51 (0.43 to 0.58)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	383	0.45 (0.36 to 0.52)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	382	0.34 (0.25 to 0.43)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	382	0.31 (0.21 to 0.39)	< 0.001
<b>M12 vs. baseline</b>				
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	375	0.57 (0.50 to 0.64)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	375	0.50 (0.42 to 0.58)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	375	0.30 (0.20 to 0.39)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	375	0.37 (0.28 to 0.46)	< 0.001
<b>M6 vs. M3</b>				
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	376	0.49 (0.41 to 0.56)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	376	0.42 (0.33 to 0.50)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	373	0.34 (0.24 to 0.42)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	373	0.31 (0.22 to 0.40)	< 0.001
<b>M12 vs. M3</b>				
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	367	0.52 (0.44 to 0.59)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	367	0.50 (0.42 to 0.57)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	365	0.25 (0.15 to 0.35)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	365	0.40 (0.31 to 0.49)	< 0.001
<b>M12 vs. M6</b>				
<i>Change in EQ-5D-5L index value</i>	<i>Change in PROMIS-29 PCS</i>	364	0.49 (0.40 to 0.56)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	364	0.39 (0.30 to 0.47)	< 0.001
<i>Change in EQ-5D-5L VAS</i>	<i>Change in PROMIS-29 PCS</i>	361	0.30 (0.20 to 0.39)	< 0.001
	<i>Change in PROMIS-29 MCS</i>	361	0.37 (0.27 to 0.45)	< 0.001

The PROMIS-29 and EQ-5D-5L showed similar internal responsiveness by comparing the SRM and Cohen’s *d* effect size for the differences between scores at the varying time intervals. Between baseline and month 12, Cohen’s *d* for the change in the utility index score, VAS, PCS and MCS fell in the small range between –0.19 and –0.16. This demonstrates that there was a small deterioration in HRQoL over the 12-month follow-up of the trial, and this was detected by all measures. These changes were statistically significant but unlikely to be clinically meaningful. Only a fifth of participants who provided valid EQ-5D-5L data at baseline and 12 months had an increase in utility index score of 0.06 or greater, which is suggestive of a clinically meaningful improvement in HRQoL.

**TABLE 44** Participants achieving minimum clinically important difference, based on an increase of 0.06 in EQ-5D-5L index value between time points

Time interval	EQ-5D-5L index value increased by MCID Yes, n (%)
Baseline-M3	97 (23.5)
Baseline-M6	91 (22.9)
Baseline-M12	86 (21.8)
M3-M6	75 (19.1)
M3-M12	74 (19.1)
M6-M12	79 (20.5)



**FIGURE 10** Receiver operating characteristic curves for the PROMIS-29 physical and mental health scores against the external criterion of an improvement equating to the minimum clinically important difference of 0.06 in the EQ-5D-5L index value between baseline and month 12.

Both the PROMIS-29 PCS and MCS demonstrated reasonable external responsiveness in that they were able to discriminate between patients who achieved the EQ-5D-5L utility index score MCID of an increase of 0.06 or more between baseline and month 12 (AUC 0.74 and 0.70, respectively).

We were unable to find any other studies that evaluated the validity of the PROMIS-29 against the EQ-5D-5L. A recent review by Pan *et al.*<sup>85</sup> collated the published evidence based on the relationship between these two instruments and identified six studies. Three main types of relationships were examined in the studies: (1) comparing PROMIS-29 and EQ-5D as descriptive systems; (2) mapping PROMIS-29 domains to EQ-5D utilities; and (3) comparing and transforming PROMIS-29 utilities

to EQ-5D utilities. None considered the validity of the PROMIS-29 relative to the EQ-5D-5L as a measure of HRQoL. Shim and Hamilton<sup>87</sup> compared the responsiveness of the EQ-5D-3L and PROMIS-10 with the Oxford Knee Score in patients undergoing total knee replacements. They confirmed good responsiveness of the PROMIS-10, but comparability of these results to ours is low due to the alternative versions of the instruments used and the different patient populations.

A strength of our study is the large sample size (data from at least 361 participants could be included in each element of the analysis for this substudy) from a multicentre RCT, suggesting broad generalisability.

The decision about which PROM to use to assess HRQoL should consider the patient population and responsiveness to change of the PROM in this population and the specific elements of HRQoL relevant to the population and expected to be impacted by any intervention offered. While the EQ-5D-5L utility index score and VAS and the PROMIS-29 MCS and PCS were well correlated with each other and appeared to have similar responsiveness to change in this study, researchers may still choose to include both in their studies as their individual domains/subscales appear to measure distinct elements of HRQoL. For example, the self-care domain of the EQ-5D-5L and the fatigue and sleep disturbance subscales of the PROMIS-29 do not correlate well with any of the other measures, suggesting they capture distinct domains.

Recent developments have allowed a utility score to be calculated from the PROMIS-29, but value sets are, to date, only available for the USA.<sup>88-90</sup> Utility index scores were not calculated from the PROMIS-29 in this study. Future research could compare the validity of the PROMIS-29 and EQ-5D-5L utility index scores in this population.

## Substudy 2: Does including £5 and/or a pen in the recruitment pack enhance recruitment?

This abstract has been reproduced from Fairhurst *et al.*<sup>91</sup>

The abstract is from 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 work is properly cited. The full text is presented in [Report Supplementary Material 5](#) and includes minor additions and formatting changes to the original text.

### Background

Monetary and other incentives may increase recruitment to RCTs.

### Methods

2 × 2 factorial 'SWAT' of including a pen and/or £5 (Great British pounds) in cash with a postal recruitment pack to increase the number of participants randomised into the host trial ('Gentle Years Yoga') for older adults with multimorbidity. Secondary outcomes: return, and time to return, of screening form and the cost per additional participant randomised. Binary data were analysed using logistic regression and time to return using Cox proportional hazards regression.

### Results

Eight hundred and eighteen potential host trial participants were included. Between those sent a pen ( $n = 409$ ) and not sent a pen ( $n = 409$ ), there was no evidence of a difference in the proportion of participants randomised [15 (3.7%) vs. 11 (2.7%); OR 1.38, 95% CI 0.63 to 3.04], in returning a screening form [66 (16.1%) vs. 61 (14.9%); OR 1.10, 95% CI 0.75 to 1.61] nor in time to return the screening form [Hazard Ratio (HR) 1.09, 95% CI 0.77 to 1.55]. Between those sent £5 ( $n = 409$ ) and not sent £5

( $n = 409$ ), there was no evidence of increased randomisation [14 (3.4%) vs. 12 (2.9%); OR 1.18, 95% CI 0.54 to 2.57], but more screening forms were returned [77 (18.8%) vs. 50 (12.2%); OR 1.67, 95% CI 1.13 to 2.45], and there was decreased time to return screening forms (HR 1.56, 95% CI 1.09 to 2.22). No significant interaction between the interventions was observed. The cost per additional participant randomised was £32 and £1000 for the pen and £5, respectively.

### Conclusion

A small, monetary incentive did not result in more participants being randomised into the host trial but did encourage increased and faster response to the recruitment invitation. Since it is relatively costly, we do not recommend this intervention for use to increase recruitment in this population. Pens were cheaper but did not provide evidence of benefit.

### Substudy 3: Does sending a pen with a follow-up questionnaire enhance return rates?

This abstract has been reproduced from Fairhurst *et al.*<sup>92</sup>

The abstract is from 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 work is properly cited. The full text is presented in [Report Supplementary Material 6](#) and includes minor additions and formatting changes to the original text.

### Background

Poor response rates to follow-up questionnaires can adversely affect the progress of a RCT and the validity of its results. This embedded 'SWAT' aimed to investigate the impact of including a pen with the postal 3-month questionnaire completed by the trial participants on the response rates to this questionnaire.

### Methods

This study was a two-armed RCT nested in the GYY trial. Participants in the intervention group of the GYY trial were allocated 1 : 1 using simple randomisation to either receive a pen (intervention) or no pen with their 3-month questionnaire (control). The primary outcome was the proportion of participants sent a 3-month questionnaire who returned it. Secondary outcomes were time taken to return the questionnaire, proportion of participants who sent a reminder to return the questionnaire and completeness of the questionnaire. Binary outcomes were analysed using logistic regression, time to return by Cox Proportional hazards regression and number of items completed by linear regression.

### Results

There were 111 participants randomised to the pen group and 118 to the no pen group who were sent a 3-month questionnaire. There was no evidence of a difference in return rates between the two groups [pen 107 (96.4%), no pen 117 (99.2%); OR 0.23, 95% CI 0.02 to 2.19;  $p = 0.20$ ]. Furthermore, there was no evidence of a difference between the two groups in terms of time to return the questionnaire (HR 0.90, 95% CI 0.69 to 1.18,  $p = 0.47$ ), the proportion of participants sent a reminder (OR 0.85, 95% CI 0.48 to 1.53,  $p = 0.60$ ) nor the number of items completed (MD 0.51, 95% CI -0.04 to 1.06,  $p = 0.07$ ).

### Conclusion

The inclusion of a pen with the postal 3-month follow-up questionnaire did not have a statistically significant effect on response rate.



## Substudy 4: Does offering a free yoga session to control participants after the 12-month follow-up assessment enhance retention and reduce contamination?

### Introduction

In many yoga trials, participants are randomised to either a yoga intervention or a usual care (i.e. no yoga) control group. Given the nature of the intervention, blinding of the participants to their allocation is not possible, which can then create undesirable consequences in these trials because of disappointment among members of the control group who do not receive the yoga intervention. Recruitment of participants for the trial may be compromised because eligible patients decline because they do not want to be randomised to the control group.<sup>93</sup> Further, patients who enrol may be highly motivated to undertake yoga and, therefore, those randomised to the control group may begin yoga or increase their physical activity levels independently.<sup>94</sup> This non-compliance by controls may lead to a decrease of power to detect a clinically important intervention effect. Last, participants randomised to the control group may experience resentful demoralisation and withdraw from the trial.

Various trial designs have been used to address these issues with control group participants: (1) instructions given before the start of the intervention (e.g. 'Please don't change your level of exercise/physical activity during the course of the study'<sup>95</sup>), (2) offering an alternative intervention to control patients, such as education about exercise, stretches, etc. and (3) offering the intervention to control patients after the intervention (e.g. in wait-list or cross-over study design).<sup>96</sup> The latter strategy was evaluated in this SWAT.

The systematic review of Bisschop *et al.*<sup>95</sup> provided an overview of these different types of control groups in exercise-oncology trials and explored the influence on contamination and dropout rates.<sup>95</sup> The lowest contamination and dropout rates were observed in studies with control groups that were offered an intervention after the intervention period (contamination in 7.1% of studies, excess dropout rate  $-4.7 \pm 9.2\%$ ), but randomised trials are needed to clarify the effects of delayed interventions.

The objective of this methodological SWAT was to evaluate the effects of offering a free yoga class after the 12-month follow-up assessment versus no offer on rates of retention and contamination in the usual care group participants.

### Methods

#### Study design

This SWAT was a two-armed RCT embedded in the GYY trial.<sup>1</sup> The SWAT was registered with the Northern Ireland Network for Trials Methodology Research SWAT repository on 1 April 2019 (SWAT93; [www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/SWATSWARInformation/Repositories/SWATStore/](http://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/SWATSWARInformation/Repositories/SWATStore/)).

#### Participants

This study included participants allocated to the usual care arm of the GYY trial. Participants in the intervention arm were included in a different retention SWAT, namely the inclusion or not of a pen with the 3-month follow-up questionnaire (described in the previous section).

After randomisation to the main trial, participants allocated to the usual care group were immediately randomised again to receive: the offer of a one-off group yoga class which would take place after their 12-month follow-up is completed, or no offer. Participants randomised to receive the offer of a one-off class were informed immediately after randomisation.

Participants were informed in advance, via the main trial PIS, that some participants in the usual care group would be randomised to receive a one-off yoga class. Specific consent for the SWAT was not obtained; this was approved by the REC as it was considered low risk. Written informed consent for the GYY main trial was obtained from all participants who took part.

## Intervention

### Intervention

Offer of a one-off GYY class which took place after the final (12-month) follow-up was completed.

### Comparator

No offer of a yoga class.

### Outcome measures

The primary outcome was the proportion of participants in each group who returned at least one of the follow-up questionnaires (3, 6 or 12 months). Secondary outcomes included the proportion of participants who returned all three follow-up questionnaires and the proportion of usual care participants who reported use of non-GYY yoga throughout the trial follow-up.

### Sample size

All usual care participants in the host trial were randomised into this embedded trial. The host trial recruited 454 participants (240 intervention; 214 usual care). A sample size of 214 gave us 80% power to detect an increase in the percentage of participants returning at least one follow-up questionnaire from 85% in the no-offer arm to 96% in the offer arm.

### Randomisation

Simple 1 : 1 randomisation was used to allocate usual care participants to receive or not to receive, the offer of the free yoga class.

### Statistical analysis

Analyses were conducted under the principles of ITT using two-tailed tests at the 5% significance level. Analyses were conducted in Stata version 17.<sup>34</sup> Binary data were compared using logistic regression. The treatment effect is presented as an OR with an associated 95% CI and *p*-value.

## Results

In total, 214 participants were randomised into the usual care group of the GYY trial, of which 111 (51.9%) were allocated to receive the offer of the free yoga class and 103 (48.1%) to receive no offer.

Usual care participants offered a free yoga class were slightly more likely to return at least one of the post-randomisation follow-up questionnaires ( $n = 104$ , 93.7%) than usual care participants not offered a free yoga class ( $n = 91$ , 88.4%), though this difference was not statistically significant (OR 1.96, 95% CI 0.74 to 5.19;  $p = 0.18$ ).

There was no evidence of a difference in the likelihood of participants returning all three post-randomisation follow-up questionnaires (offer of a free yoga class  $n = 87$ , 78.4% vs. no offer  $n = 82$ , 79.6%; OR 0.90, 95% CI 0.46 to 1.77;  $p = 0.76$ ).

There was no evidence of a difference in the likelihood of participants reporting use of non-GYY yoga throughout the trial follow-up (offer of a free yoga class  $n = 15$ , 13.5% vs. no offer  $n = 15$ , 14.6%; OR 0.92, 95% CI 0.42 to 1.98;  $p = 0.83$ ).

## Discussion

The results of this trial do not indicate any demonstrable benefit of offering a free yoga class versus no offer after the 12-month follow-up assessment on rates of retention and contamination in the usual care group participants in the GYY trial. Although usual care participants offered a yoga class were slightly more likely to return at least one of the post-randomisation follow-up questionnaires compared to 'no offer' participants, the difference was not statistically significant. However, this trial was underpowered to detect a 5 percentage point difference, as was observed.

To our knowledge, this is the first SWAT investigating the offer of a single free intervention session after study completion on retention and contamination within a control group of a trial. Our study does not support the findings of Bisschop *et al.*,<sup>95</sup> where the lowest contamination and dropout rates were found in control groups offered an intervention after the intervention period; however, this is not entirely comparable to our study as it combines different intervention offers and not just a single free intervention session. Our study supports the finding of Courneya *et al.*<sup>97</sup> in terms of contamination, who suggest that offering the intervention to control participants at the end of the study was not sufficient to eliminate the contamination. In the study by Courneya *et al.*,<sup>97</sup> however, participants were asked not to initiate any structured exercise over the course of the intervention, so again not entirely comparable to our study.

The strength of this study was that it was a randomised trial; however, since it was conducted in a population of older adults with multimorbidity, particularly during the COVID-19 pandemic, findings may not be generalisable to other populations or contexts.

The GYY trial provided information on the importance of the control group in the PIS at the start of the study that may have lessened the potential benefit of the offer of a free yoga class at the end of the study on contamination. It was pointed out by Courneya *et al.* that a suggested solution to contamination is to ensure that participants understand the implications of random assignment and what is expected of them in each group before joining the study.<sup>97</sup> Alternatively, participants in the control arm may have under-reported yoga participation because they know it is important for the control group to abstain from yoga practice. It did not, however, state in the GYY trial PIS that control participants should not practise yoga or take part in other exercises.

Being assigned to a control group may be disappointing to participants and, consequently, some may continue or begin an exercise programme despite their group assignment.<sup>98</sup> Additional physical activity may introduce unintended confounders. Participants not offered the free yoga class may have therefore been more likely to take up another non-yoga physical activity, knowing they were not going to be offered a yoga session at the end of the trial; however, we were unable to monitor this as participants did not record other physical activity they were taking part in, only yoga. Researchers rarely monitor contamination rates, presumably because it is assumed that participants in the control group will not exercise.<sup>97</sup> Measures of contamination with not only the exercise intervention but other physical activities in control groups should therefore be considered in future studies.

The GYY trial implemented additional retention strategies including sending a text message to participants a few days before their postal questionnaire arrived, an unconditional £5 'thank you' payment and reminder questionnaires and phone calls. All of these may have lessened the potential benefit of the offer of a free yoga class at the end of the study on retention. Future studies should therefore consider other retention strategies they are implementing if they want to test the true effects of offering the intervention to control participants once the study has been completed.

The knowledge about the free yoga class was given at the beginning of the trial only. Perhaps a reminder at each follow-up would have increased the effects of the offer. Participants also had to wait 12 months to be offered the free yoga class, which could potentially be too long a time. Future studies should look at adding reminders of the incentive at various time points in the study, such as at each follow-up, which may improve follow-up return rates.

The offer in this study was for one free yoga class only. A longer free yoga course might have increased the effects of the offer. Future studies should therefore look at the length of a free delayed intervention offered.

It would be valuable for future studies to evaluate control group experiences of the design of the delayed intervention incentive.

Offering a free yoga class versus no offer of a free yoga class after the 12-month follow-up assessment did not have any significant effects on rates of retention and contamination in the usual care group participants in the GYY trial. Future studies should consider using reminders about the offer, other retention and contamination strategies used to ensure they are not diluting the effects of the offer and the length of free sessions they are offering, for example, more than one free session. It would also be beneficial for future studies to evaluate control group experiences of the design of the delayed intervention incentive.



## Chapter 7 Discussion

In this chapter, we summarise and discuss the main findings from the GYY trial, compare these with findings from previous studies and discuss the strengths and limitations of our work.

### Summary of main findings

In this study, we have assessed the clinical and cost-effectiveness of offering the GYY programme to people aged 65 years or older with multimorbidity.

A primary goal of the yoga programme was to improve participants' HRQoL. The results show no statistically or clinically significant effect from offering the GYY programme in respect of HRQoL measured using the EQ-5D-5L utility index score, which was the primary outcome. Another measure of HRQoL, the PROMIS-29, showed similar findings; that is, all of the PROMIS outcomes showed no evidence of effect except for pain interference and pain intensity, which showed small improvements associated with the intervention. There were no statistically significant differences in the secondary outcomes of depression, anxiety, loneliness or falls. No serious, related AEs were reported.

The economic evaluation showed that the intervention was associated with additional costs of £80.85 per participant and generated an additional 0.0178 QALYs per participant, on average, compared with usual care. The combined effect was that the GYY programme was likely to be cost-effective at the usual thresholds for WTP.

The fidelity checks indicated that the group classes were mostly delivered as planned. Deviations from the standardised protocol were few, minor and resolved with feedback. Class attendance rates were good overall (mean 8.8/12 sessions attended). Adherence to home yoga practice was also good at the intervention end point but reduced at post-intervention follow-up.

The process evaluation interviews highlighted that participants joined the trial for a range of reasons, including personal benefit, curiosity and altruism. Some participants favoured online delivery and some favoured face to face, with no clear delivery preference overall. Participants viewed GYY as a suitable and safe activity for older people with varying health issues, independent of delivery format. The perceived impact of the GYY programme ranged from minimal to transformative. For some participants, there was no impact on their health or lifestyle. For others, yoga became an integral part of their lives, and they felt it generated a broad range of benefits, including improvements in physical function, joint pain and stiffness and mental well-being.

In summary, the GYY programme was not associated with any statistically significant benefits in terms of HRQoL, mental health, loneliness or falls. However, the intervention was safe, acceptable to most participants and highly valued by some participants. The economic evaluation suggests that the intervention could be cost-effective.

### Interpretation of findings

There were no statistically significant differences in any of the key outcome measures. Furthermore, for the primary outcome of HRQoL measured using the EQ-5D-5L utility index score, the 95% CI excluded a clinically meaningful benefit (which we defined as an improvement of at least 0.06). Although these findings indicate that the intervention is not clinically effective, we should reflect on what part measurement and fidelity of implementing the intervention had on these findings. The potential role of each of these factors will be discussed and comparisons made with relevant previous studies.

### ***Intervention fidelity***

The fidelity checks indicated that the yoga teachers received adequate training and delivered the intervention mainly as intended. Indeed, deviations from the standardised protocol were few, minor and easily resolved with feedback. As is the case in routine practice, yoga teachers sometimes adapted the postures and activities so that they were accessible for participants with specific health problems.

This was a pragmatic trial testing the clinical and cost-effectiveness of offering the GYY programme, which comprised 12 weekly group-based classes plus encouragement to practice at home on non-class days. Class attendance rates were good. Of the 240 participants randomised to yoga, 192 (80%) met our pre-specified definition of 'fully compliant'; that is, they attended at least three of the first six sessions and at least three other sessions. In a previous RCT evaluating yoga for low back pain, only 60% of participants met this level of attendance.<sup>99</sup> Adherence to home yoga practice was also good at the intervention end point (3 months: 83% of intervention participants reporting a median of 4 weekly sessions) but reduced at post-intervention follow-up (12 months: 48% of intervention participants reporting a median of 3 weekly sessions). A similar pattern was observed in the yoga for low back pain trial.<sup>99</sup> The process evaluation interview data provide some insight into factors influencing home practice. An important facilitator of sustained practice was the biopsychosocial benefit gained from practice, whereas barriers included perceived lack of benefit, lack of interest in yoga and a preference for doing pre-existing physical exercise routines. Although adherence to the GYY programme varied between participants, this reflects a real-world implementation of the intervention in the context of current service provision for this population. Finally, the self-reported data on yoga practice at follow-up show little evidence of contamination in the usual care group.

In summary, the intervention was delivered largely as intended, so poor fidelity of implementing the intervention does not appear to be a major factor explaining the clinical effectiveness results.

### ***Sensitivity of outcome measures***

A key consideration is whether or not the outcome measures, particularly the EQ-5D-5L, which was the primary outcome, are sensitive to change in such populations.<sup>100</sup> This point was discussed at length both within the research team and with the funding body when the study was first designed. It is well recognised that the EQ-5D measure has limited sensitivity to change,<sup>101</sup> although responsiveness has been demonstrated in older adults with multimorbidity.<sup>102</sup> Our decision to use the EQ-5D-5L was based on several considerations. First, improving HRQoL was the most important aim of our intervention, and it was, therefore, appropriate to choose it as the primary outcome measure despite the known measurement difficulties. Second, the EQ-5D is the 'gold-standard' HRQoL measure recommended by NICE for comparing the benefits of different interventions. Third, we used the five-level version, which was designed to be more responsive than the older three-level version.<sup>103</sup> Finally, data from an earlier pilot study of the GYY programme suggested that the EQ-5D-5L was sensitive to intervention effects in a similar population after 3 months' follow-up.<sup>19</sup> The uncertainty about the appropriateness of using the EQ-5D-5L prompted us to include a second measure of HRQoL, the PROMIS-29, and the consistent results across all key outcomes add confidence to our interpretation of the findings.

The possibility of a ceiling effect owing to the participants being relatively healthy at baseline is also worthy of consideration. Potential participants were selected on the basis of their having two or more chronic conditions rather than having poor HRQoL. This approach would reduce the capacity of the trial to show benefit in patients who had few problems at baseline. The low recruitment rate (3.5% of those sent an invitation pack were randomised) also raises the possibility of participation bias, a common feature in physical activity research whereby study participants are generally fitter and healthier than those who do not volunteer.<sup>104</sup> The mean EQ-5D-5L utility index score at baseline was not markedly different from that previously reported for a nationally representative sample of older adults with multimorbidity (0.739 vs. 0.713)<sup>105</sup> and shows that there was room for improvement in HRQoL overall. However, most of the participant interviewees viewed their health as good, prompting some to question why they had been invited to the trial. One might therefore question if the GYY programme would be



more effective if it was targeted at people with greater health needs. These might be identified through more sophisticated selection algorithms (e.g. people with substantial multimorbidity and polypharmacy or using an index such as the electronic Frailty Index)<sup>106</sup> or by allowing GPs to use their clinical judgement to offer it to the patients whom they felt would benefit from it.

### **Not clinically effective**

As indicated above, some of the lack of effect of the intervention could be due to the heterogeneous study population, in which some participants were relatively healthy at baseline. Another possibility is that the GYY programme, as offered in this trial, is simply not effective at improving HRQoL or the key secondary outcomes in this population. The lack of meaningful change across all key outcomes would support this interpretation. Furthermore, although the exploratory CACE analysis provided some evidence that increased compliance was associated with greater effectiveness, the difference in effect was still small and probably not clinically important.

One interpretation could be that the 'dose' of yoga was insufficient to have a perceivable or measurable impact. In the absence of definitive guidance on programme duration, frequency or intensity, we were guided by previous studies and routine practice, as well as an intention to test a pragmatic and scalable version of the intervention. Twelve weeks is a commonly used intervention duration in yoga trials,<sup>107,108</sup> and the mixture of weekly group classes with encouragement to perform regular home practice reflected how GYY is typically delivered outside of the trial. Regarding 'intensity', GYY activities are intentionally gentle to make them accessible to older people with a wide range of health problems. Interview data support that this was the case, with the majority of participant interviewees viewing GYY as a form of gentle and unchallenging exercise. In some respect, it may be unrealistic to expect much change in health outcomes resulting from a 12-week programme of gentle yoga. More challenging forms of yoga may be effective, but they may also be less accessible to this population. Longer programmes with more frequent classes may also have more potential to improve health outcomes but would require more resources to deliver and might be less acceptable to participants.

The pivot to online delivery part way through this trial may have also influenced the clinical effectiveness results. A concern was that the after-class social time would be less meaningful in online versus face-to-face classes. However, an exploratory subgroup analysis showed no evidence of a differential effect on the primary outcome between delivery modes; but it is acknowledged that this analysis is underpowered. The process evaluation highlighted that a majority of interviewees had little interest in the social component of classes, and yoga teachers and participants noted that attendance for this component was low across several classes.

### **Comparison with previous studies**

This is the first RCT, of which the authors are aware, to investigate the clinical and cost-effectiveness of yoga for older people with multimorbidity. However, there have been RCTs of yoga for older people and RCTs of exercise or other patient-orientated interventions for people with multimorbidity. Therefore, the comparisons presented here have limitations in their applicability.

### **Effects of yoga on HRQoL and mental well-being in older people**

Tulloch *et al.* published a systematic review in 2018 of 'physical' yoga interventions for improving outcomes in people aged 60 + years.<sup>109</sup> Twelve trials with 752 participants were included. The yoga interventions varied in terms of style (including Hatha, Iyengar and Viniyoga), frequency (1–3 classes per week, often with additional regular home practice) and duration (8–24 weeks). Meta-analyses demonstrated that, compared with inactive control, yoga produced a medium effect on HRQoL (SMD 0.51, 95% CI 0.25 to 0.76) and a small effect on mental well-being (SMD 0.38, 95% CI 0.15 to 0.62). The trials were judged to be of high methodological quality.

The original pilot trial of GYY was not included in this review because it was published after the final search date of January 2017. This study of 52 participants, of which 25 were randomised to a 10-week



GY programme and 27 to wait-list control, found a statistically significant improvement associated with the intervention (adjusted MD 0.12, 95% CI 0.03 to 0.21).<sup>19</sup> The participants in the pilot trial had a similar mean age to those in the current trial (74.8 vs. 73.5 years) but included a higher proportion of women (90% vs. 61%). The pilot trial also had a much lower proportion of participants who had multimorbidity (54% vs. 100%) because it included physically inactive adults aged 60+ years and not specifically people with multimorbidity. This suggests that the current trial included a less healthy population; however, the baseline EQ-5D scores were similar.

### **Effects of exercise therapy on HRQoL, physical function and mental health in people with multimorbidity**

Bricca *et al.* published a systematic review in 2020 of exercise therapy for improving outcomes in people with multimorbidity.<sup>110</sup> Twenty-three RCTs with 3363 participants, testing an exercise therapy intervention (mean duration 13.0 weeks, typically aerobic exercise or combined aerobic and resistance exercise), showed that exercise therapy improved HRQoL (SMD 0.37, 95% CI 0.14 to 0.61) and objectively measured physical function (SMD 0.33, 95% CI 0.17 to 0.49) and reduced depression symptoms (SMD -0.80, 95% CI -1.21 to -0.40) and anxiety symptoms (SMD -0.49, 95% CI -0.99 to 0.01). Meta-regression showed that increasing age was associated with lower effect sizes for HRQoL, and greater baseline depression severity was associated with greater reduction of depression symptoms. The overall quality of evidence for all the outcomes was rated as low, mainly due to risk of bias, inconsistency and indirectness.

### **Effects of patient-orientated interventions in people with multimorbidity**

Smith *et al.* published an updated Cochrane review in 2021 of trials of interventions for improving outcomes in people with multimorbidity in primary care and community settings.<sup>8</sup> This review included 17 RCTs, five of which evaluated predominantly patient-oriented interventions, for example self-management support groups. A pooled effect size was not calculated for HRQoL due to substantial heterogeneity. Nevertheless, the findings were mixed and did not suggest that patient-orientated interventions are generally effective. The review authors suggested that interventions with a specific focus on functional capacity and activity participation may possibly be effective. However, the evidence to support this conclusion was quite limited, and such focused interventions are not likely to address the wide range of problems experienced by people with multimorbidity. The main conclusion was that there were only a limited number of studies, with considerable heterogeneity in their findings; therefore, further high-quality pragmatic trials are needed.

### **Summary**

Our findings regarding HRQoL seem consistent with the literature on patient-orientated interventions for people with multimorbidity (i.e. little or no benefit), but inconsistent with previous findings on yoga for older people and exercise therapy for people with multimorbidity (where small to medium beneficial effects are apparent). The reasons behind this 'mixed picture' are unclear; however, the outcome of the current trial provides the best estimate of the 'true' effects of offering GYY to older people with multimorbidity, specifically in the context of the UK healthcare system. The effects of yoga in this specific population should be further explored through meta-analysis once additional combinable studies have been performed.

### **Cost-effectiveness**

The results of the economic evaluation conducted alongside this trial suggest that offering the GYY intervention in addition to usual care could be a cost-effective option in terms of QALYs gained calculated using the EQ-5D-5L. Over the 12-month follow-up, the average cost per participant was £1886 (95% CI £1796 to £1976) and £1965 (95% CI £1882 to £2048) in the usual care and intervention groups, respectively. Compared with the usual care group, the intervention group incurred an additional mean cost of £81 (95% CI £77 to £85) but generated more QALYs (MD 0.0178, 95% CI 0.0175 to 0.0180). The ICER for the base-case analysis involving the multiply imputed data set was found to be £4546 per additional QALY, and the probability of the intervention being cost-effective at

a WTP threshold of £20,000 per QALY gained was 79%. The results appear robust across a range of assumptions that were explored using sensitivity analyses.

A limitation of the economic analysis was the presence of missing data. We had complete data on intervention costs for all participants and high completion rates for healthcare resource use at baseline. However, many participants had missing responses for healthcare usage items at later time points. Because of the cumulative nature of cost data, one missing observation renders all other data for that participant unusable in a CCA. The presence of missing follow-up healthcare usage data therefore reduced the sample size for the CCA to  $n = 192$  (42%). A test of the missing data mechanism indicated that data were MAR, indicating that multiple imputation was an appropriate method for handling missing data in the base-case analysis.

An important consideration is why the intervention group experienced higher costs than the usual care group over the 12-month follow-up (MD £81). Key cost drivers of the analysis included hospital-based services, specifically inpatient stays, accident and emergency visits (including overnight stays), day-case visits, outpatient attendances and medications. Also, as expected, the cost of the intervention itself was a prominent cost at £188 in the base case. Exploration of the subcomponents of health service use cost collected in this study indicated that the intervention group had higher costs in four components (GP visit at GP practice/home, GP phone consultation, mental health services, day-case hospital visit), and the usual care group had higher costs in two components (inpatient hospital nights, medications). Overall, it appears that the increased resources associated with delivering the GYY courses were partially offset by a slight reduction in the overall costs of healthcare usage over the 12-month follow-up. Another limitation of the economic evaluation is that it does not account for any differences in costs and QALYs that may be expected over the longer term (> 12 months post randomisation).

The cost of the intervention (£175 per participant for face to face, £196 per participant for online) was less expensive than other yoga or exercise-based therapies that have been used to treat long-term conditions. For example, a yoga programme for treating chronic low back pain, for which there is evidence of effectiveness, costs £293 per participant.<sup>111</sup> Elsewhere, a group-based supervised exercise programme for intermittent claudication has been estimated to cost £288 per participant,<sup>112</sup> a group- and hospital-based pulmonary rehabilitation programme cost £718 per participant<sup>113</sup> and a facilitated self-care and home-based cardiac rehabilitation programme for heart failure cost £418 per participant.<sup>114</sup>

## Strengths and limitations

This trial has several strengths. It is the first adequately powered RCT of yoga for older people with multimorbidity and was rigorously undertaken in line with recommended standards for individually randomised trials. The trial was prospectively registered, and the protocol was published. External validity was enhanced by using broad eligibility criteria and recruiting from a range of general practices across England and Wales. Randomisation was conducted by a secure, remote, web-based system with concealed allocation. The intervention was standardised and delivered by 12 experienced teachers who all held a regulated qualification in teaching GYY. Class attendance rates were good, as was adherence to home yoga practice during the intervention period. Recruitment was completed within the initial intended time period, the number of participants randomised provided sufficient power as per our sample size assumptions and there were high rates of participant follow-up over 12 months. We collected data on a range of outcomes, several of which feature in a core outcome set for multimorbidity trials.<sup>115</sup> The two randomised groups were comparable in almost all of the baseline characteristics. We also conducted economic and process evaluations, both of which have been lacking in most previous studies of yoga or interventions for multimorbidity. The trial was reported in line with CONSORT and other relevant guidelines.<sup>116,117</sup> Finally, an independent TSC helped ensure that the trial was conducted as planned and that participant safety issues were considered.

The trial also has some limitations. First, adherence to home yoga practice decreased over time past the intervention end point, which may have diluted any effects; however, the trial was highly pragmatic and reflects the effectiveness of the intervention in real-world implementation. Second, only 3.5% of invited patients were recruited. This rate of recruitment is typical of trials using this type of intervention and recruitment strategy,<sup>118</sup> but it raises the possibility of recruitment bias. The characteristics of the trial participants are considered further in the section on *Equality, Diversity and Inclusion* below. Third, there was some chance of imbalance in gender at baseline, but this was adjusted for in a SA and did not change the conclusions. Fourth, the COVID-19 pandemic required us to change our processes for recruitment, follow-up and intervention delivery part way through the trial. Regarding intervention delivery, BWY continues to offer a mixture of face-to-face and online GYY classes, so the mixture of course types included in this trial reflects current practice. Fifth, the use of subjective, participant self-reported outcome measures, as opposed to objective measures, has the potential to introduce reporting bias since participants were not blinded to their allocation in this trial. Finally, we specified a large number of secondary outcomes, which increases the risk of false-positive findings due to multiple comparisons.<sup>24</sup> However, in pragmatic trials such as this, it is important to collect data for a broad range of outcome measures relevant to stakeholders (patients, clinicians, policy-makers, etc.).<sup>119</sup> The consistent findings across all key outcomes add confidence to our interpretation of the results.

### Equality, diversity and inclusion

The trial enrolled 454 participants from 15 general practices in England and Wales. The mean age was 73.5 years (SD = 6.2), 60.6% of participants were female and 2% were from non-white ethnic groups. Participants had a median of three conditions, the mean EQ-5D-5L utility index score at baseline was 0.739 (SD = 0.169) and the mean IMD decile was 7.54 (SD = 2.65).

Data from the nationally representative Clinical Practice Research Datalink-GOLD database showed that, in the UK in 2012, 55% of multimorbid adults aged 65 years or older were female.<sup>120</sup> Data for England and Wales from the 2011 Census indicate that 4.5% of people aged 65 years or older were from non-white ethnic groups.<sup>121</sup> [Appendix 10](#) (see [Table 51](#)) shows expected versus actual ethnicity percentages for each of the trial sites. Data from the nationally representative General Practice Patient Survey showed that, in the UK in 2012, multimorbid adults aged 55 years or older had a mean EQ-5D-5L utility index score of 0.713 (SD = 0.219).<sup>105</sup>

Collectively, the above data indicate that our trial participants were reasonably representative of the wider multimorbid older adult population. However, there may have been a slight under-representation of males, non-white ethnic groups and people with lower socioeconomic/health status.

Steps that we took to optimise participation of relevant people included using broad eligibility criteria, identifying potential participants from GP databases, recruiting from a range of general practices and regions in England and Wales and using carefully worded recruitment materials. However, inclusiveness for the target population may have been limited by some aspects of the study. For example, we excluded people who were unable to read or speak English. However, this was done because the GYY teachers who were available to work on the trial all spoke English; thus, it may have been difficult for them to effectively communicate with non-English-speaking participants. We also could possibly have sought to work with a more diverse range of sites; however, site selection was largely determined by where the teachers were based (at least for the face-to-face courses).

The research staff at the Universities of York and Northumbria were appointed to ensure equality and diversity. The research team included a mixture of genders and a range of skills, experience and expertise. The project offered Professor Tew the opportunity to lead a full-scale multisite trial for the first time under the mentorship of Professor Hewitt. It also provided the two yoga consultants (LB, JH) with the experience of being involved in a large research trial. Within YU, three early-career

researchers gained experience taking lead roles in managing the project and developing strong team working. Our data administration team developed new ways of working to ensure the trial continued during the pandemic. Through working on this project, our Trial Support Officer (administrator grade) demonstrated her skills and knowledge and was promoted to a Trial Coordinator/Research Fellow. Members of the public who were representative of the study population were also involved in all aspects of the study, as described in the next section. Our TMG member, Valerie Mount, gained insight into trial methodology.

## Patient and public involvement

The design and conduct of this trial benefited from patient and public involvement (PPI). The aim of PPI in all aspects of the study was to ensure that the voices of older adults (with multimorbidity) were woven through the research, such that the results would be of direct benefit to them.

We worked with PPI representatives from grant preparation through to dissemination. Because the study linked to a themed call (complex health and care needs in older people), we were aware that older adults had been involved through the NICE guideline committee and NIHR prioritisation work, but we additionally worked with seven PPI representatives, all of whom were aged 65 years or older with multimorbidity and experience of GYY participation. This involvement extended across considerations around research design, development and iteration of participant information resources, research management and troubleshooting (as members of the TMG and TSC), interpretation of the data and writing and dissemination of the findings.

An example of how PPI shaped the research included consideration of how to limit the dropout rate in the usual care group. The PPI representatives expressed concerns that the dropout rate would be higher in the usual care group if there were no benefits to being in the study and suggested that some interaction would be required to make them feel involved. Some of their ideas that were implemented as SWATs included financial incentives and providing a free yoga class and details of suitable yoga classes at the end of the study.

Another example is related to the practicalities of intervention delivery. Consistent with the PPI representatives' feedback, we only recruited participants from general practices that were located close to the yoga venue (for face-to-face courses); we only used yoga venues that had good transport links, and we made sure that all trial classes were delivered between the hours of 1000 and 1500.

The PPI representatives also helped shape the participant-facing study documents. For example, we made sure that the questionnaires used a suitably large font size in case participants had sight problems. Key words and phrases were capitalised and bolded for emphasis. The colour of the paper was carefully considered to ensure the text could be seen clearly. We also included a contact number for a researcher who could help them complete the questionnaire.

## Implications for health care

### *Does this trial provide sufficient evidence to justify implementing Gentle Years Yoga courses as a funded NHS service?*

No. In this trial, the GYY courses were offered as if they were a funded NHS service, with referrals initiated in primary care and the programme provided free of charge to patients in the community. If shown to provide benefit relative to usual care alone, the intervention could potentially be made available to primary care professionals to offer to eligible patients. The findings lead us towards the somewhat paradoxical conclusion of 'not clinically effective but probably cost-effective'. The cost-effectiveness data alone may imply that the intervention should be adopted; however, it has been

argued that only exceptionally should a single trial provide grounds for implementation.<sup>122</sup> NICE and similar decision-making bodies specify that, rather than being based on a single trial, economic evaluation should generally be based on the totality of the evidence established by systematic review and meta-analysis.

### ***Should self-funded Gentle Years Yoga classes be recommended to older people with multimorbidity?***

Possibly. Outside of the trial setting, GYY classes are available to attend on a self-pay basis, either online or face to face, in many parts of the UK. Our findings indicate that the intervention is safe, acceptable and, in some cases, highly valued in this population. Healthcare professionals or social prescribing link workers could therefore consider recommending self-funded GYY classes where they appear to be a 'good fit' with an individual's needs and preferences. Routinely recommending GYY to older adults with multimorbidity would be unlikely to improve HRQoL at the population level, but a more targeted approach may provide various benefits to individuals. Given its gentle nature, GYY might be best targeted towards older adults who are frail or have a greater disease burden. For such individuals, the classes might be sufficiently stimulating to provide benefit and act as a gateway to more challenging forms of yoga or other forms of physical activity. From a different perspective, GYY classes could be done to contribute towards achieving a healthy amount of physical activity.<sup>123</sup>

Whether or not GYY classes are recommended may be influenced by several factors. A prerequisite to training for the GYY teaching qualification is for a yoga teacher to be insured on the basis of having a Level 3 qualification in Teaching Yoga or equivalent, any tradition. Gentle Years Yoga should only be delivered by insured teachers who hold the regulated BWYQ Level 4 Certificate in Teaching GYY or by insured teachers who are under the supervision of their GYY trainer and have completed the first observation/assessment of GYY teaching during unit 2, the practical unit. The 42-hour theory unit in 'Understanding the Principles of Adapting Yoga for Older Adults' is a prerequisite for unit 2.

As of January 2023, 66 yoga teachers across England, Scotland and Wales have the BWYQ Level 4 Certificate in Teaching GYY, and another 63 are teaching GYY classes as part of the assessment to receive the Ofqual-regulated qualification in 2023. This would give a total of 129 fully qualified GYY teachers in the locations shown on the map in [Figure 11](#).

Another important factor will be the awareness of healthcare professionals and link workers of what GYY is, its 'target audience' and evidence base and what classes are available to the people they are supporting. If there is a demand for GYY teachers and marketing support, it is realistic for BWY and other BWYQ-recognised centres to be able to train around 450 teachers per year to receive the regulated BWYQ Level 4 Teaching GYY qualification.

### ***A word of caution***

We tested a specific yoga programme in a specific population, that is, a 12-week programme of GYY in people aged 65 years or older with multimorbidity. We caution against extrapolating our findings to other contexts (e.g. other styles of yoga, other populations). When considering the role of yoga for a specific, single condition (e.g. chronic low back pain), other more directly applicable literature should be consulted.

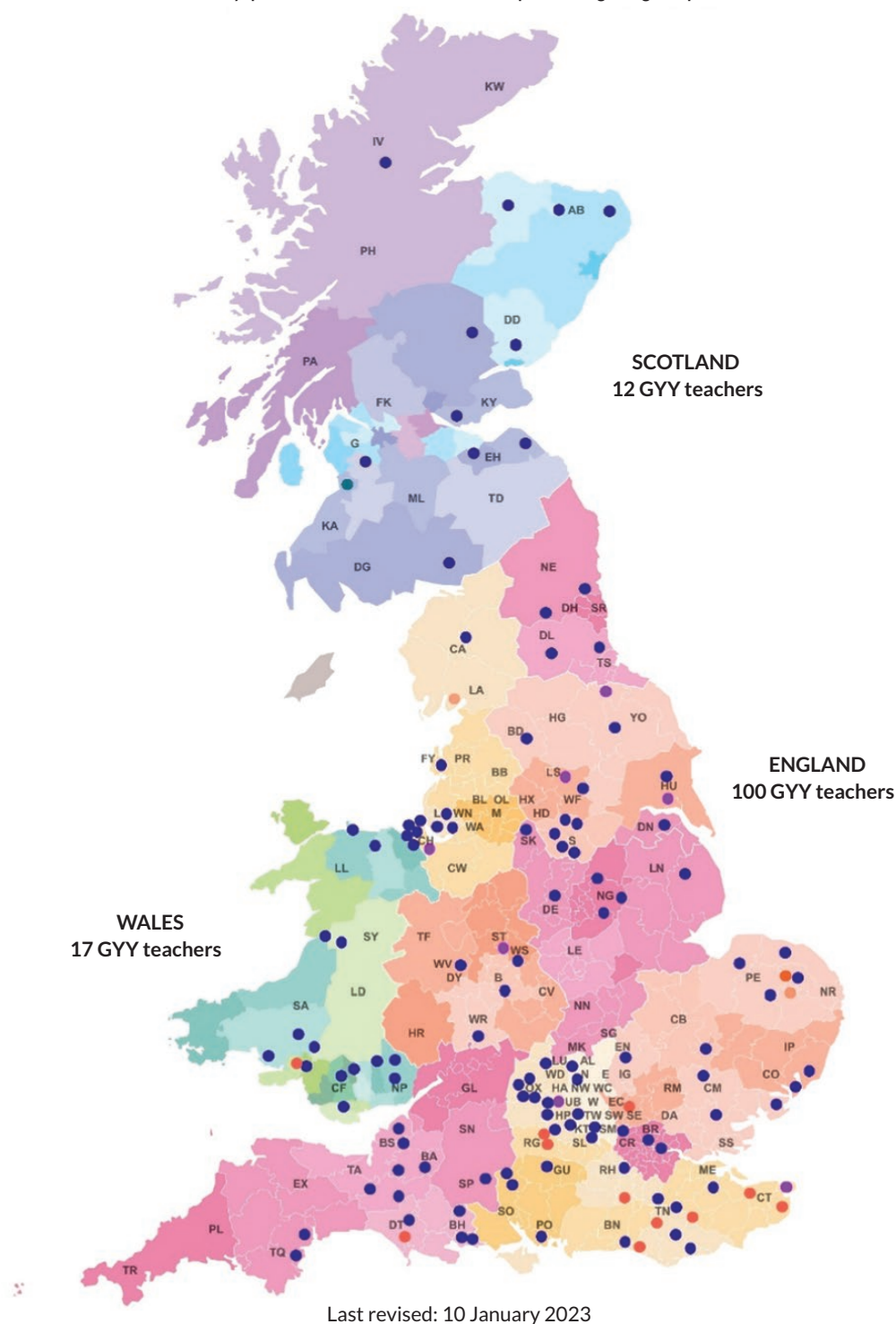
## **Recommendations for research**

We have identified several potential areas of future research, listed here in no specific order. We suggest that a more formal process of scoping and prioritisation be undertaken to determine how these are best taken forward in the wider context of yoga and multimorbidity research and ongoing work in these areas.



### BWY Gentle Years Yoga (GY) postcode map

Locations where the GYY programme is already available to older adults with either a fully qualified GYY teacher or a trainee in process of getting the qualification in 2023



**FIGURE 11** Location of GYY-trained yoga teachers.

- Further research could be done to try and identify subgroups of people who are most likely to benefit from this type of intervention.
- Further research is needed to explore the impact of yoga programmes on health inequalities and to identify best practices for widening access, maintaining engagement and tailoring support. For

example, it would be useful to understand the barriers and facilitators to broaden the diversity within those training as yoga teachers as well as for potential participants.

- Further research could be done to investigate the optimal type and dose of yoga that might be required to produce measurable and perceivable benefits to health outcomes.
- Further work is needed to help build a consensus about the most appropriate eligibility criteria to use in multimorbidity intervention trials.
- Further research is needed to explore the plausibility of our seemingly paradoxical finding of the intervention being not clinically effective but probably cost-effective. This could include longer-term cost-effectiveness modelling.

## Conclusion

The offer of a 12-week programme of GYY was not associated with any statistically significant benefits in terms of HRQoL, mental health, loneliness or falls in older adults with multimorbidity. However, the intervention was safe, acceptable to most participants and highly valued by some participants. The economic evaluation suggests that the intervention could be cost-effective.

# Additional information

## Acknowledgements

The authors wish to thank the participants for taking part in the trial. We would like to acknowledge the support of the NIHR Clinical Research Network. We would specifically like to thank Matthew Bailey, Maddy Elliot, Emma Filby, David Goodge and Val Wadsworth at YTU, University of York, for their contribution to data collection and data management and their assistance with the day-to-day running of the study. Additionally, we would like to thank the NIHR HTA Programme for funding the GYY trial.

We would like to thank the yoga teachers: Rosemary Bennett, Audrey Blow, Stephanie Braysmith, Sheree Cox, Eve Douglas, Clare Gardner, Celia Grieve, Ann Haggard, Emma Middleton, Lina Newstead, Paul Smith and Steve Smith.

Trial Sponsor: Northumbria University.

We would like to thank the independent members of the TSC for their support and guidance throughout the trial: Professor Andrew Judge (Chairperson; Professor of Translational Statistics, University of Bristol), Professor Andrew Clegg (Professor of Geriatric Medicine, University of Leeds), Professor Charlie Foster (Professor of Physical Activity & Public Health, University of Bristol), Dr Françoise Freedman (Senior Yoga Teacher and Lecturer of Social Anthropology, University of Cambridge), Mavis Giles (PPI representative), Karen Sherlock (PPI representative), Dr Teik Goh (GP, The Garth Surgery, Guisborough), Professor Barbara Hanratty (Professor of Primary Care and Public Health, Newcastle University), Professor Stewart Mercer (Professor of Primary Care and Multimorbidity, University of Edinburgh) and Professor Dawn Skelton (Professor in Ageing and Health, Glasgow Caledonian University).

Our thanks are also given to the following individuals for their input to this project: Clare Cook (PPI Representative), Belen Corbacho (Research Fellow, University of York), Pauline Fleming (BWY representative), Dorothy Hosein (Chief Executive Officer, BWY), Gillian Osborne (BWY Representative), Shirley-Anne Paul (Trial Coordinator, University of York), Jenny Roche (Trainee Statistician, University of York), Helen Tilbrook (Trial Coordinator, University of York) and Ian Watt (Professor, University of York).

We would like to thank staff from the GP practices involved in identifying and inviting participants to the study and also those who helped provide participant prescription data: Amherst Medical Practice (Kent), Banbury Cross Health Centre (Banbury), Bellevue Group Practice (Newport), Bridge View Medical (Bristol), Church Street Practice (Wantage), Diadem Medical Practice (Hull), Moss Healthcare (Harrogate), Eastgate Medical Group (Harrogate), Falldon Way Medical Centre (Bristol), Heswall & Pensby Group Practice (Wirral), Paxton Medical Practice (Wirral), St Bartholomew's Medical Centre (Oxford), Summertown Health Centre (Oxford), Wellspring Surgery (Bristol) and Westbury on Trym Primary Care Centre (Bristol).

## Contributions of authors

**Garry Alan Tew** (<https://orcid.org/0000-0002-8610-0613>) (Professor of Clinical Exercise Science, York St John University) was the chief investigator for the GYY trial. He had overall responsibility for the design and implementation of the study and writing of the report, with final approval of the report submission.



**Laura Wiley** (<https://orcid.org/0000-0001-9619-4484>) (Research Fellow) was a trial co-ordinator who assisted with the day-to-day management of the study and was involved in writing the initial version of the report.

**Lesley Ward** (<https://orcid.org/0000-0001-7285-2032>) (Research Fellow) undertook the process evaluation interviews, observations and qualitative analysis and contributed to writing the final report.

**Jessica Grace Hugill-Jones** (<https://orcid.org/0000-0002-9915-3941>) (Research Fellow) was a trial co-ordinator and assisted with the day-to-day management of the study.

**Camila Sofia Maturana** (<https://orcid.org/0000-0002-4946-8515>) (Research Fellow) was a trial co-ordinator and assisted with the day-to-day management of the study.

**Caroline Marie Fairhurst** (<https://orcid.org/0000-0003-0547-462X>) (Senior Research Fellow, Statistician) was a co-applicant, contributed to the grant application and trial protocol, wrote the statistical analysis plan, conducted the statistical analysis and was involved in writing the initial version of the report.

**Kerry Jane Bell** (<https://orcid.org/0000-0001-5124-138X>) (Research Fellow) contributed to the health economics analysis and was involved in writing the health economics section of the report.

**Laura Bissell** (<https://orcid.org/0000-0002-4784-2745>) (Yoga Consultant) was a co-applicant, helped standardise and assess intervention delivery and contributed to writing the report.

**Alison Booth** (<https://orcid.org/0000-0001-7138-6295>) (Senior Research Fellow) was the trial manager and contributed to writing the report.

**Jenny Howsam** (<https://orcid.org/0000-0002-5055-7072>) (Yoga Consultant) was a co-applicant, helped standardise and assess intervention delivery and contributed to writing the report.

**Valerie Mount** (<https://orcid.org/0000-0001-7089-6007>) was a PPI Representative and contributed to writing the plain English summary.

**Sarah Jane Ronaldson** (<https://orcid.org/0000-0001-8321-786X>) (Research Fellow, Health Economics) wrote the health economics analysis plan, conducted the health economics analysis and wrote the health economics section of the report.

**Fiona Rose** (<https://orcid.org/0000-0003-0587-683X>) (Research Fellow) was a trial support officer and then a trial co-ordinator, assisted with the day-to-day management of the study and contributed to writing the final report.

**Tim Rapley** (<https://orcid.org/0000-0003-4836-4279>) (Professor of Applied Health Care Research) was a co-applicant, contributed to the grant application and trial protocol, oversaw the process evaluation work and wrote the process evaluation section of this report.

**David John Torgerson** (<https://orcid.org/0000-0002-1667-4275>) (Professor, YTU) was a co-applicant, contributed to the grant application and trial protocol and contributed to writing the report.

**David Yates** (<https://orcid.org/0000-0003-4696-262X>) (Consultant in Anaesthesia and Critical Care Medicine) provided clinical oversight to the project and contributed to writing the final report.

**Catherine Elizabeth Hewitt** (<https://orcid.org/0000-0002-0415-3536>) (Professor, Director of YTU) was a co-applicant, contributed to the grant application and trial protocol, was a mentor to the chief investigator and contributed to writing the report.

All authors read and gave approval to the final manuscript.

## Disclosure of interests

**Full disclosure of interests:** Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/KPGN4216>.

**Primary conflicts of interest:** Laura Bissell and Jenny Howsam cocreated the British Wheel of Yoga (BWY) Gentle Years Yoga programme. Laura Bissell is a teacher trainer and external quality assurer who serves as a trustee-director and Chair of British Wheel of Yoga Qualifications (BWYQ), a separate company/registered charity that operates as an Ofqual-recognised awarding organisation for multiple training centres. Jenny Howsam is the BWYQ operations co-ordinator in charge of the awarding organisation's External Quality Assurance Department. Catherine Elizabeth Hewitt is Deputy Chair of the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) Programme commissioning board, on the NIHR Clinical Trials Unit Standing Advisory Committee, the NIHR HTA Post-Funding Committee teleconference and the NIHR HTA Funding Committee Policy Group (formerly CSG). David John Torgerson is the Director of a CTU which receives funding from the NIHR.

## Patient data statement

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety and plan NHS services. Patient data should be kept safe and secure to protect everyone's privacy, and it is important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

## Data-sharing statement

All trial data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review. The qualitative process evaluation data generated is not suitable for sharing beyond that contained within the report. Further information can be obtained from the corresponding author.

## Ethics statement

Ethics approval for the study was obtained from the North East – York Research Ethics Committee (REC reference number 19/NE/0072) on 24 April 2019.

## Information governance statement

YTU is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679.

Under the Data Protection legislation, YTU is the Data Processor, Northumbria University is the Data Controller and we process personal data in accordance with their instructions. You can find out more

about how we handle personal data, including how to exercise your individual rights, and the contact details for YU's Data Protection Officer here [www.york.ac.uk/records-management/dp/](http://www.york.ac.uk/records-management/dp/).

## **Publications**

Tew GA, Bissell L, Corbacho B, Fairhurst C, Howsam J, Hugill-Jones J, *et al.* Yoga for older adults with multimorbidity (the Gentle Years Yoga Trial): study protocol for a randomised controlled trial. *Trials* 2021;**22**:269.

Fairhurst C, Parkinson G, Hewitt C, Maturana C, Wiley L, Rose F, *et al.* Enclosing a pen in a postal questionnaire follow-up to increase response rate: a study within a trial. [version 2; peer review: 2 approved]. *NIHR Open Res* 2023;**2**:53.

Fairhurst C, Roche J, Bissel L, Hewitt C, Hugill-Jones J, Howsam J, *et al.* A 2 × 2 randomised factorial SWAT of the use of a pen and small, financial incentive to improve recruitment in a randomised controlled trial of yoga for older adults with multimorbidity. *F1000 Res* 2022;**10**:326.

## References

1. Tew GA, Bissell L, Corbacho B, Fairhurst C, Howsam J, Hugill-Jones J, *et al*. Yoga for older adults with multimorbidity (the Gentle Years Yoga Trial): study protocol for a randomised controlled trial. *Trials* 2021;**22**(1):269.
2. Johnston MC, Crilly M, Black C, Prescott GJ, Mercer SW. Defining and measuring multimorbidity: a systematic review of systematic reviews. *Eur J Public Health* 2019;**29**(1):182–9.
3. Kingston A, Robinson L, Booth H, Knapp M, Jagger C; MODEM project. Projections of multimorbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. *Age Ageing* 2018;**47**(3):374–80.
4. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Arch Intern Med* 2002;**162**(20):2269–76.
5. Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;**61**(582):e12–21.
6. McPhail SM. Multimorbidity in chronic disease: impact on health care resources and costs. *Risk Manag Healthc Policy* 2016;**9**:143–56.
7. Eckardt M, Bretschneider C, van den Bussche H, König HH; MultiCare Study Group. Analysis of health care costs in elderly patients with multiple chronic conditions using a finite mixture of generalized linear models. *Health Econ* 2017;**26**(5):582–99.
8. Smith SM, Wallace E, Clyne B, Boland F, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community setting: a systematic review. *Syst Rev* 2021;**10**(1):271.
9. McCall MC, Ward A, Roberts NW, Heneghan C. Overview of systematic reviews: yoga as a therapeutic intervention for adults with acute and chronic health conditions. *Evid Based Complement Alternat Med* 2013;**2013**:945895.
10. Sivaramakrishnan D, Fitzsimons C, Kelly P, Ludwig K, Mutrie N, Saunders DH, Baker G. The effects of yoga compared to active and inactive controls on physical function and health related quality of life in older adults – systematic review and meta-analysis of randomised controlled trials. *Int J Behav Nutr Phys Act* 2019;**16**(1):33.
11. Wieland LS, Skoetz N, Pilkington K, Vempati R, D’Adamo CR, Berman BM. Yoga treatment for chronic non-specific low back pain. *Cochrane Database Syst Rev* 2017;**1**(1):CD010671.
12. Yang ZY, Zhong HB, Mao C, Yuan JQ, Huang YF, Wu XY, *et al*. Yoga for asthma. *Sao Paulo Med J* 2016;**134**(4):368.
13. Cramer H, Lauche R, Klose P, Lange S, Langhorst J, Dobos GJ. Yoga for improving health-related quality of life, mental health and cancer-related symptoms in women diagnosed with breast cancer. *Cochrane Database Syst Rev* 2017;**1**(1):CD010802.
14. Hartley L, Dyakova M, Holmes J, Clarke A, Lee MS, Ernst E, Rees K. Yoga for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev* 2014;**2014**(5):CD010072.
15. Patel NK, Newstead AH, Ferrer RL. The effects of yoga on physical functioning and health related quality of life in older adults: a systematic review and meta-analysis. *J Altern Complement Med* 2012;**18**(10):902–17.
16. Gothe NP, McAuley E. Yoga is as good as stretching-strengthening exercises in improving functional fitness outcomes: results from a randomized controlled trial. *J Gerontol A Biol Sci Med Sci* 2016;**71**(3):406–11.

17. Youkhana S, Dean CM, Wolff M, Sherrington C, Tiedemann A. Yoga-based exercise improves balance and mobility in people aged 60 and over: a systematic review and meta-analysis. *Age Ageing* 2016;**45**(1):21–9.
18. Andronis L, Kinghorn P, Qiao S, Whitehurst DG, Durrell S, McLeod H. Cost-effectiveness of non-invasive and non-pharmacological interventions for low back pain: a systematic literature review. *Appl Health Econ Health Policy* 2017;**15**(2):173–201.
19. Tew GA, Howsam J, Hardy M, Bissell L. Adapted yoga to improve physical function and health-related quality of life in physically-inactive older adults: a randomised controlled pilot trial. *BMC Geriatr* 2017;**17**(1):131.
20. Smelt AF, van der Weele GM, Blom JW, Gussekloo J, Assendelft WJ. How usual is usual care in pragmatic intervention studies in primary care? An overview of recent trials. *Br J Gen Pract* 2010;**60**(576):e305–18.
21. 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**(10):1727–36.
22. National Institute for Health and Care Excellence. *Position Statement on Use of the EQ-5D-5L Valuation Set for England* (updated October 2019). 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 22 June 2023).
23. 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**(5):708–15.
24. Salisbury C, Man MS, Bower P, Guthrie B, Chaplin K, Gaunt DM, *et al.* Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach. *Lancet* 2018;**392**(10141):41–50.
25. Fischer F, Gibbons C, Coste J, Valderas JM, Rose M, Lepage A. Measurement invariance and general population reference values of the PROMIS Profile 29 in the UK, France, and Germany. *Qual Life Res* 2018;**27**(4):999–1014.
26. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009;**114**(1–3):163–73.
27. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;**166**(10):1092–7.
28. Hughes ME, Waite LJ, Hawkey LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. *Res Aging* 2004;**26**(6):655–72.
29. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A* 2013;**110**(15):5797–801.
30. Mattsson M, Sandqvist G, Hesselstrand R, Nordin A, Bostrom C. Validity and reliability of the Patient Health Questionnaire-8 in Swedish for individuals with systemic sclerosis. *Rheumatol Int* 2020;**40**(10):1675–87.
31. Bayliss EA, Ellis JL, Steiner JF. Seniors' self-reported multimorbidity captured biopsychosocial factors not incorporated into two other data-based morbidity measures. *J Clin Epidemiol* 2009;**62**(5):550–7.e1.
32. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005;**14**(6):1523–32.
33. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. *Value Health* 2017;**20**(4):644–50.

34. StataCorp. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC; 2021.
35. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
36. EuroQol Research Foundation. *EQ-5D-5L User Guide: Basic Information on How to Use the EQ-5D-5L Instrument*. 2019. URL: <https://euroqol.org/publications/user-guides> (accessed 22 June 2023).
37. Billingham L, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Assess* 1998;**3**(10):1–152.
38. Department of Health. *NHS Reference Costs 2020/21*. Department of Health; 2020. URL: [www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/](http://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/) (accessed 22 June 2023).
39. Jones K, Burns A. *Unit Costs of Health and Social Care 2021*. Canterbury: University of Kent, Personal Social Services Research Unit; 2021. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/) (accessed 22 June 2023).
40. British Medical Association, Royal Pharmaceutical Society. *British National Formulary (online)*. London: BMJ Group and Pharmaceutical Press; 2022. URL: <https://bnf.nice.org.uk/> (accessed 22 June 2023).
41. NHS Business Services Authority, NHS Prescription Services. *NHS Electronic Drug Tariff (online)*. URL: [www.drugtariff.nhsbsa.nhs.uk2022](http://www.drugtariff.nhsbsa.nhs.uk2022) (accessed 22 June 2023).
42. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: University of Kent, Personal Social Services Research Unit; 2015. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2015/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2015/) (accessed 22 June 2023).
43. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: University of Kent, Personal Social Services Research Unit; 2014. URL: [www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2014/](http://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2014/) (accessed 22 June 2023).
44. Department of Health. *NHS Reference Costs 2017/18*. Department of Health; 2017. URL: [www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#archive](http://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/#archive) (accessed 22 June 2023).
45. Bupa. *Private GP Appointments*. 2022. URL: [www.Bupa.co.uk/health/payg/gp-services](http://www.Bupa.co.uk/health/payg/gp-services) (accessed 22 June 2023).
46. myTribe Health Insurance. *What Is the Cost of Private Physiotherapy?* 2022. URL: [www.mytribein-surance.co.uk/treatment/what-is-the-cost-of-private-physiotherapy](http://www.mytribein-surance.co.uk/treatment/what-is-the-cost-of-private-physiotherapy) (accessed 22 June 2023).
47. Psychiatry-UK. *Psychiatry Fees – Explained*. 2022. URL: [www.Psychiatry-uk.com/fees](http://www.Psychiatry-uk.com/fees) (accessed 22 June 2023).
48. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40.
49. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**(4):377–99.
50. Rubin DB, Wiley I. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley; 1987.
51. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making* 1998;**18**(2 Suppl.):S68–80.
52. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Econ* 2004;**13**(5):405–15.
53. Bell K, Corbacho B, Ronaldson S, Richardson G, Hood K, Sanders J, *et al.*, Building Blocks Trial Group. Costs and consequences of the Family Nurse Partnership (FNP) programme in England: evidence from the Building Blocks trial. *F1000Research* 2019;**8**:1640.



54. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *PharmacoEconomics* 2014;**32**(12):1157–70.
55. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res* 2017; **27**(4):591–608.
56. Patton MQ. *Qualitative Research & Evaluation Methods*. 3rd edn. Thousand Oaks, CA: SAGE Publications Ltd; 2002.
57. Rapley T. Some pragmatics of data analysis. In Silverman D, editor. *Qualitative Research*. 5th edn. London: SAGE Publications Ltd; 2021. pp. 341–56.
58. Layder D. *Sociological Practice: Linking Theory and Social Research*. London: SAGE Publications Ltd; 1998.
59. Glaser BG. The constant comparative method of qualitative analysis. *Soc Probl* 1965;**12**(4):436–45.
60. Charmaz K. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: SAGE Publications Ltd; 2006.
61. Seale C. *The Quality of Qualitative Research*. London: SAGE Publications Ltd; 1999.
62. Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In Ritchie J, Lewis J, editors. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd; 2003. pp. 219–62.
63. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci* 2015;**10**:21.
64. McCann SK, Campbell MK, Entwistle VA. Reasons for participating in randomised controlled trials: conditional altruism and considerations for self. *Trials* 2010;**11**:31.
65. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;**6**:42.
66. Brosnan P, Nauphal M, Tompson MC. Acceptability and feasibility of the online delivery of hatha yoga: a systematic review of the literature. *Complement Ther Med* 2021;**60**:102742.
67. Haynes A, Gilchrist H, Oliveira JS, Sherrington C, Tiedemann A. 'I wouldn't have joined if it wasn't online': understanding older people's engagement with teleyoga classes for fall prevention. *BMC Complement Med Ther* 2022;**22**(1):283.
68. Huler G, Macdonald B. Rethinking social relationships in old age: digitalization and the social lives of older adults. *Am Psychol* 2020;**75**(4):554–66.
69. Keohane S, Swarbrick C, Helal. *Barriers and Facilitators to Technology Among Older Adults During COVID-19 Pandemic: A Systematic Review Using Thematic Analysis – International Conference on Human-Computer Interaction*. Cham: Springer; 2022.
70. Vereijken B, Helbostad JL. The potential for technology to enhance physical activity among older people. In Nyman SR, Barker A, Haines T, Horton K, Musselwhite C, Peeters G, et al., *The Palgrave Handbook of Ageing and Physical Activity Promotion*. Cham: Palgrave Macmillan; 2018. pp. 713–31.
71. Haynes A, Gilchrist H, Oliveira JS, Grunseit A, Sherrington C, Lord S, Tiedemann A. What helps older people persevere with yoga classes? A realist process evaluation of a COVID-19-affected yoga program for fall prevention. *BMC Publ Health* 2022;**22**(1):463.
72. Gilchrist H, Haynes A, Oliveira JS, Grunseit A, Sherrington C, Bauman A, et al. The value of mind-body connection in physical activity for older people. *J Aging Phys Act* 2022;**1**:8.

73. Butzer B, LoRusso AM, Windsor R, Riley F, Frame K, Khalsa SBS, Conboy L. A qualitative examination of yoga for middle school adolescents. *Adv Sch Ment Health Promot* 2017;**10**(3):195–219.
74. Cramer H, Lauche R, Haller H, Langhorst J, Dobos G, Berger B. 'I'm more in balance': a qualitative study of yoga for patients with chronic neck pain. *J Altern Complement Med* 2013;**19**(6):536–42.
75. Valderas JM. Multimorbidity, not a health condition or complexity by another name. *Eur J Gen Pract* 2015;**21**(4):213–4.
76. Ford JC, Ford JA. Multimorbidity: will it stand the test of time? *Age Ageing* 2018;**47**(1):6–8.
77. Cheshire A, Richards R, Cartwright T. 'Joining a group was inspiring': a qualitative study of service users' experiences of yoga on social prescription. *BMC Complement Med Ther* 2022;**22**(1):67.
78. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Per Med* 2011;**8**(2):161–73.
79. Schork NJ. Personalized medicine: time for one-person trials. *Nature* 2015;**520**(7549):609–11.
80. Stenman U, Hakama M, Knekt P, Aromaa A, Teppo L, Leinonen J, *et al.* Measurement and modeling of health-related quality of life. *Epidem Demog Public Health* 2010;**195**:130–5.
81. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, *et al.*, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 2010;**63**(11):1179–94.
82. *HealthMeasures*. PROMIS; 2022 URL: [www.healthmeasures.net/explore-measurement-systems/promis](http://www.healthmeasures.net/explore-measurement-systems/promis). (accessed 22 June 2023).
83. Huang W, Rose AJ, Bayliss E, Baseman L, Butcher E, Garcia RE, Edelen MO. Adapting summary scores for the PROMIS-29 v2.0 for use among older adults with multiple chronic conditions. *Qual Life Res* 2019;**28**(1):199–210.
84. Hanmer J, Dewitt B, Yu L, Tsevat J, Roberts M, Revicki D, *et al.* Cross-sectional validation of the PROMIS: preference scoring system. *PLOS ONE* 2018;**13**(7):e0201093.
85. Pan T, Mulhern B, Viney R, Norman R, Tran-Duy A, Hanmer J, Devlin N. Evidence on the relationship between PROMIS-29 and EQ-5D: a literature review. *Qual Life Res* 2021;**31**:79–89.
86. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York: Routledge; 1988.
87. Shim J, Hamilton DF. Comparative responsiveness of the PROMIS-10 Global Health and EQ-5D questionnaires in patients undergoing total knee arthroplasty. *Bone Joint J* 2019;**101-B**(7):832–7.
88. Dewitt B, Feeny D, Fischhoff B, Cella D, Hays RD, Hess R, *et al.* Estimation of a preference-based summary score for the patient-reported outcomes measurement information system: the PROMIS®-Preference (PROPr) scoring system. *Med Decis Making* 2018;**38**(6):683–98.
89. Craig BM, Reeve BB, Brown PM, Cella D, Hays RD, Lipscomb J, *et al.* US valuation of health outcomes measured using the PROMIS-29. *Value Health* 2014;**17**(8):846–53.
90. Dewitt B, Jalal H, Hanmer J. Computing PROPr utility scores for PROMIS® profile instruments. *Value Health* 2020;**23**(3):370–8.
91. Fairhurst C, Roche J, Bissell L, Hewitt C, Hugill-Jones J, Howsam J, *et al.* A 2x2 randomised factorial SWAT of the use of a pen and small, financial incentive to improve recruitment in a randomised controlled trial of yoga for older adults with multimorbidity. *F1000Res* 2021;**10**:326.



92. Fairhurst C, Parkinson G, Hewitt C, Maturana C, Wiley L, Rose F, *et al.* Enclosing a pen in a postal questionnaire follow-up to increase response rate: a study within a trial. *NIHR Open Res* 2022;**2**:53.
93. Courneya KS, Forbes CC, Trinh L, Sellar CM, Friedenreich CM, Reiman T. Patient satisfaction with participation in a randomized exercise trial: effects of randomization and a usual care posttrial exercise program. *Clin Trials* 2013;**10**(6):959–66.
94. Hertogh EM, Schuit AJ, Peeters PHM, Monninkhof EM. Noncompliance in lifestyle intervention studies: the instrumental variable method provides insight into the bias. *J Clin Epidemiol* 2010;**63**(8):900–6.
95. Bisschop CNS, Courneya KS, Velthuis MJ, Monninkhof EM, Jones LW, Friedenreich C, *et al.* Control group design, contamination and drop-out in exercise oncology trials: a systematic review. *PLOS ONE* 2015;**10**(3):e0120996.
96. Waters L, Reeves M, Fjeldsoe B, Eakin E. Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. *J Phys Act Health* 2012;**9**(6):884–95.
97. Courneya KS, Friedenreich CM, Quinney HA, Fields AL, Jones LW, Fairey AS. A randomized trial of exercise and quality of life in colorectal cancer survivors. *Eur J Cancer Care (Engl)* 2003;**12**(4):347–57.
98. Mock V, Pickett M, Ropka ME, Muscari Lin E, Stewart KJ, Rhodes VA, *et al.* Fatigue and quality of life outcomes of exercise during cancer treatment. *Cancer Pract* 2001;**9**(3):119–27.
99. Tilbrook HE, Hewitt CE, Aplin JD, Semlyen A, Trehwela A, Watt I, Torgerson DJ. Compliance effects in a randomised controlled trial of yoga for chronic low back pain: a methodological study. *Physiotherapy* 2014;**100**(3):256–62.
100. Salisbury C, Man MS, Chaplin K, Mann C, Bower P, Brookes S, *et al.* A patient-centred intervention to improve the management of multimorbidity in general practice: the 3D RCT. *Health Serv Del Res* 2019;**7**:1–238.
101. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. *Qual Life Res* 2005;**14**(7):1651–68.
102. Bhadhuri A, Kind P, Salari P, Jungo KT, Boland B, Byrne S, *et al.* Measurement properties of EQ-5D-3L and EQ-5D-5L in recording self-reported health status in older patients with substantial multimorbidity and polypharmacy. *Health Qual Life Outcomes* 2020;**18**(1):317.
103. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, *et al.* Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;**22**(7):1717–27.
104. de Souto Barreto P, Ferrandez AM, Saliba-Serre B. Are older adults who volunteer to participate in an exercise study fitter and healthier than nonvolunteers? The participation bias of the study population. *J Phys Act Health* 2013;**10**(3):359–67.
105. Thompson AJ, Turner AJ. A comparison of the EQ-5D-3L and EQ-5D-5L. *Pharmacoeconomics* 2020;**38**(6):575–91.
106. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, *et al.* Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing* 2016;**45**(3):353–60.
107. Ward L, Stebbings S, Cherkin D, Baxter GD. Components and reporting of yoga interventions for musculoskeletal conditions: a systematic review of randomised controlled trials. *Complement Ther Med* 2014;**22**(5):909–19.

108. Elwy AR, Groessl EJ, Eisen SV, Riley KE, Maiya M, Lee JP, *et al.* A systematic scoping review of yoga intervention components and study quality. *Am J Prev Med* 2014;**47**(2):220–32.
109. Tulloch A, Bombell H, Dean C, Tiedemann A. Yoga-based exercise improves health-related quality of life and mental well-being in older people: a systematic review of randomised controlled trials. *Age Ageing* 2018;**47**(4):537–44.
110. Bricca A, Harris LK, Jager M, Smith SM, Juhl CB, Skou ST. Benefits and harms of exercise therapy in people with multimorbidity: a systematic review and meta-analysis of randomised controlled trials. *Ageing Res Rev* 2020;**63**:101166.
111. Chuang LH, Soares MO, Tilbrook H, Cox H, Hewitt CE, Aplin J, *et al.* A pragmatic multicentered randomized controlled trial of yoga for chronic low back pain: economic evaluation. *Spine (Phila Pa 1976)* 2012;**37**(18):1593–601.
112. Bermingham SL, Sparrow K, Mullis R, Fox M, Shearman C, Bradbury A, Michaels J. The cost-effectiveness of supervised exercise for the treatment of intermittent claudication. *Eur J Vasc Endovasc Surg* 2013;**46**(6):707–14.
113. Griffiths TL, Phillips CJ, Davies S, Burr ML, Campbell IA. Cost effectiveness of an outpatient multidisciplinary pulmonary rehabilitation programme. *Thorax* 2001;**56**(10):779–84.
114. Dalal HM, Taylor RS, Jolly K, Davis RC, Doherty P, Miles J, *et al.* The effects and costs of home-based rehabilitation for heart failure with reduced ejection fraction: the REACH-HF multicentre randomized controlled trial. *Eur J Prev Cardiol* 2019;**26**(3):262–72.
115. Smith SM, Wallace E, Salisbury C, Sasseville M, Bayliss E, Fortin M. A Core Outcome Set for multimorbidity research (COSmm). *Ann Fam Med* 2018;**16**(2):132–8.
116. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Obstet Gynecol* 2010;**115**(5):1063–70.
117. Moonaz S, Nault D, Cramer H, Ward L. CLARIFY 2021: explanation and elaboration of the Delphi-based guidelines for the reporting of yoga research. *BMJ Open* 2021;**11**(8):e045812.
118. Tilbrook HE, Cox H, Hewitt CE, Kang'ombe AR, Chuang LH, Jayakody S, *et al.* Yoga for chronic low back pain: a randomized trial. *Ann Intern Med* 2011;**155**(9):569–78.
119. Ware JH, Hamel MB. Pragmatic trials—guides to better patient care? *N Engl J Med* 2011;**364**(18):1685–7.
120. Zhu Y, Edwards D, Mant J, Payne RA, Kiddle S. Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study. *BMC Med* 2020;**18**(1):78.
121. Office for National Statistics. 2011 Census. 2011. URL: [www.ons.gov.uk/census/2011census](http://www.ons.gov.uk/census/2011census) (accessed 22 June 2023).
122. Raftery J, Williams HC, Clarke A, Thornton J, Norrie J, Snooks H, Stein K. 'Not clinically effective but cost-effective' – paradoxical conclusions in randomised controlled trials with 'doubly null' results: a cross-sectional study. *BMJ Open* 2020;**10**(1):e029596.
123. Foster C, Armstrong M, Hillsdon M, Skelton D, Mavroei A, Cavill N, *et al.* *Muscle and Bone Strengthening and Balance Activities for General Health Benefits in Adults and Older Adults: Summary of a Rapid Evidence Review for the UK Chief Medical Officers' Update of the Physical Activity Guidelines*. London: UK Public Health England; 2018. pp. 1–31.



# Appendix 1 Research Ethics Committee approved amendments to the study protocol

## Recruitment

Due to the COVID-19 pandemic, various amendments were required to ensure that recruitment could continue. Alternative consent and questionnaire data collection processes were added, including the electronic completion of the consent form by participants via an online form, the collection of participant screening and baseline questionnaire data via telephone and the entry of this data electronically via an online form by study investigators. Circumstances that would make a participant unable to participate in an online yoga class were added to the exclusion criteria. These included no internet access, unfamiliar or unable to use the internet, no suitable device, insufficient space at home and no sturdy chair for use in the classes. It was also specified that potential participants needed to be available to attend 9 of the 12 classes. The method of contacting a patient to inform them of their ineligibility or eligibility (and therefore group allocation) was updated to say that this will either be done via telephone, e-mail or letter. Another amendment associated with recruitment challenges due to the COVID-19 pandemic was reconsidering the sample size in October 2021. This meant that recruitment closed with a total of 454 participants randomised (214 usual care/240 intervention). Other minor amendments to recruitment included updating an error to the estimated number of GP practices to be recruited, which was corrected to 36, and information about the use of Docmail used by GP practices for mailing out invitations to patients was added.

## Process evaluation

Amendments were made to clarify the processes (including consent) for interviews with study participants and yoga teachers and adding that trial staff may also be interviewed. Due to the COVID-19 pandemic, it was indicated that electronic consent may be used for the interviews and class observations. Also, the option of using video conferencing for interviews was added. The number of yoga teacher interviews at pilot phase sites (4–8) and main phase sites (6–12) was stated. Stakeholders were added to the list of groups to be interviewed. The sample size of trial participant interviews was later increased to 35.

## Intervention

Details were added to clarify the intervention processes, including that YTU would send participants the yoga teacher's health questionnaires to complete before their first yoga class, and that yoga teachers would alert YTU of any AEs. Due to the COVID-19 pandemic, it was also detailed that classes would be conducted either face to face in the community or online via video conferencing. Class sizes were amended to the increased range of 8–15 participants instead of 10–15. This was due to the classes being held online. In addition, it was clarified that the GYY certification is a level 4 regulated qualification, and more detail was provided to explain what the GYY training involves and that teachers will also receive training in trial processes. A further amendment was made clarifying that the participants will be provided with advice for continuing yoga at home and in suitable yoga classes in the community that they may want to attend on a self-pay basis.

## Follow-up

Due to the COVID-19 pandemic, amendments were required to follow-up processes including that the collection of follow-up questionnaire data may be done via telephone and the entry of this data would be done electronically via an online form by study investigators. It was also added that a £5 shopping voucher may be sent to participants instead of £5 cash.

## Other

Additional amendments made to the protocol include detailing the sending out of regular newsletters to trial participants, the procedure for recording planned hospitalisations and elective surgery unrelated to the intervention, a statement to say that the allocation will be revealed to a participant's GP in response to an adverse health event if necessary and to indicate that the internal pilot phase will cover the period of the two waves of courses at the pilot sites. Updates to wording throughout the protocol were also made for consistency, that is 'instructors' was updated to 'teachers'.



**TABLE 45** Mean and SD of EQ-5D-5L index value by time point and trial site

Anonymised site	Size	Baseline	Month 3	Month 6	Month 12
1	N = 18	0.777 (0.104)	0.738 (0.131)	0.725 (0.114)	0.757 (0.091)
2	N = 25	0.673 (0.201)	0.696 (0.207)	0.703 (0.192)	0.659 (0.184)
3	N = 26	0.741 (0.120)	0.711 (0.125)	0.709 (0.130)	0.739 (0.120)
4	N = 27	0.745 (0.155)	0.729 (0.217)	0.699 (0.216)	0.740 (0.188)
5	N = 25	0.754 (0.171)	0.730 (0.179)	0.723 (0.206)	0.682 (0.279)
6	N = 24	0.716 (0.204)	0.765 (0.117)	0.756 (0.275)	0.718 (0.220)
7	N = 20	0.701 (0.190)	0.733 (0.185)	0.688 (0.186)	0.711 (0.195)
8	N = 16	0.698 (0.216)	0.656 (0.278)	0.609 (0.271)	0.719 (0.241)
9	N = 24	0.760 (0.160)	0.759 (0.167)	0.739 (0.214)	0.705 (0.228)
10	N = 25	0.728 (0.153)	0.683 (0.256)	0.682 (0.267)	0.653 (0.295)
11	N = 28	0.646 (0.204)	0.705 (0.177)	0.669 (0.238)	0.611 (0.240)
12	N = 28	0.761 (0.122)	0.737 (0.202)	0.733 (0.209)	0.712 (0.208)
13	N = 24	0.819 (0.106)	0.828 (0.121)	0.851 (0.136)	0.766 (0.233)
14	N = 16	0.804 (0.154)	0.826 (0.149)	0.777 (0.194)	0.754 (0.186)
15	N = 35	0.756 (0.156)	0.725 (0.217)	0.712 (0.257)	0.712 (0.197)
16	N = 24	0.752 (0.157)	0.768 (0.144)	0.723 (0.172)	0.739 (0.129)
17	N = 24	0.774 (0.201)	0.774 (0.197)	0.782 (0.199)	0.762 (0.201)
18	N = 29	0.788 (0.139)	0.777 (0.145)	0.745 (0.201)	0.735 (0.183)
19	N = 16	0.609 (0.186)	0.662 (0.157)	0.643 (0.185)	0.552 (0.346)
Total	N = 454	0.739 (0.169)	0.737 (0.184)	0.720 (0.212)	0.707 (0.214)

## Appendix 3 Number and proportion of participants with complete case data by group

TABLE 46 Number and proportion of participants with complete case data by group

Time point	Intervention (N = 240), n (%)	Usual care (N = 214), n (%)
Baseline	175 (72.9)	165 (77.1)
3 months	152 (63.3)	128 (59.8)
6 months	142 (59.2)	124 (57.9)
12 months	170 (70.8)	145 (67.8)
Total trial duration	105 (43.8)	87 (40.7)





## Appendix 4 Further missing data information

TABLE 47 Logistic regression for indicators of missing cost and QALY data on baseline variables

Logistic regression for indicators of missing cost and QALY data on baseline variables	OR (95% CI)	
	Missing cost data at 12 months	Missing QALY data at 12 months
BASELINE VARIABLES:		
Gender (female)	1.53 (0.99 to 2.38)	1.80 <sup>a</sup> (1.03 to 3.16)
Age	1.07 <sup>a</sup> (1.03 to 1.11)	1.05 <sup>a</sup> (1.01 to 1.10)
Site	0.85 <sup>a</sup> (0.82 to 0.88)	0.95 <sup>a</sup> (0.92 to 0.99)
EQ-5D-5L at baseline	0.49 (0.13 to 1.80)	0.81 (0.18 to 3.75)
Treatment allocation (intervention vs. usual care)	0.62 <sup>a</sup> (0.40 to 0.96)	0.53 <sup>a</sup> (0.31 to 0.89)
a Statistically significant at the 5% level.		



## Appendix 5 Completion and missingness of EQ-5D-5L questionnaires

TABLE 48 Completion and missingness of EQ-5D-5L questionnaires by group and time point

Follow-up	Completed EQ-5D-5L		Missing EQ-5D-5L ( $\geq 1$ dimension missing)	
	Intervention ( $n = 240$ ) (%)	Usual care ( $n = 214$ ) (%)	Intervention ( $n = 240$ ) (%)	Usual care ( $n = 214$ ) (%)
Baseline	240 (100.0)	213 (99.5)	0 (0)	1 (0.5)
3 months	224 (93.3)	189 (88.3)	16 (6.7)	25 (11.7)
6 months	216 (90.0)	179 (83.6)	24 (10.0)	35 (16.4)
12 months	211 (87.9)	178 (83.2)	29 (12.1)	36 (16.8)



## Appendix 6 Number of missing dimensions for invalid EQ-5D-5L questionnaires

TABLE 49 Number of missing dimensions for invalid EQ-5D-5L questionnaires by group and time point

EQ-5D-5L	Intervention: number of missing dimensions					Usual care: number of missing dimensions				
	1	2	3	4	5	1	2	3	4	5
Follow-up	1	2	3	4	5	1	2	3	4	5
Baseline	0	0	0	0	0	1	0	0	0	0
3 months	0	0	0	0	16	0	0	0	0	25
6 months	0	0	0	0	24	0	0	0	0	35
12 months	1	0	0	0	28	0	0	0	0	36



## Appendix 7 Proportion reporting EQ-5D-5L levels 1–5 by dimension



**TABLE 50** Proportion reporting EQ-5D-5L levels 1–5 by dimension, group and time point

EQ-5D-5L scale	Health state severity <sup>a</sup>	Baseline		3 months		6 months		12 months	
		Intervention	Usual care	Intervention	Usual care	Intervention	Usual care	Intervention	Usual care
Mobility	Level 1	98 (40.8%)	88 (41.1%)	112 (46.7%)	86 (40.2%)	107 (44.6%)	76 (35.5%)	97 (40.4%)	74 (34.6%)
	Level 2	79 (32.9%)	64 (29.9%)	66 (27.5%)	50 (23.4%)	70 (29.2%)	52 (24.3%)	66 (27.5%)	49 (22.9%)
	Level 3	53 (22.1%)	46 (21.5%)	36 (15.0%)	40 (18.7%)	24 (10.0%)	33 (15.4%)	35 (14.6%)	38 (17.8%)
	Level 4	9 (3.8%)	16 (7.5%)	9 (3.8%)	13 (6.1%)	14 (5.8%)	17 (7.9%)	12 (5.0%)	17 (7.9%)
	Level 5	1 (0.4%)	0 (0%)	1 (0.4%)	0 (0%)	1 (0.4%)	1 (0.5%)	1 (0.4%)	0 (0%)
	Missing	0 (0%)	0 (0%)	16 (6.7%)	25 (11.7%)	24 (10.0%)	35 (16.4%)	29 (12.1%)	36 (16.8%)
Self-care	Level 1	207 (86.3%)	180 (84.1%)	191 (79.6%)	153 (71.5%)	183 (76.3%)	137 (64.0%)	177 (73.8%)	138 (64.5%)
	Level 2	28 (11.7%)	22 (10.3%)	24 (10.0%)	20 (9.3%)	21 (8.8%)	25 (11.7%)	26 (10.8%)	28 (13.1%)
	Level 3	5 (2.1%)	10 (4.7%)	8 (3.3%)	15 (7.0%)	11 (4.6%)	13 (6.1%)	7 (2.9%)	11 (5.1%)
	Level 4	0 (0%)	2 (0.9%)	1 (0.4%)	1 (0.5%)	1 (0.4%)	3 (1.4%)	1 (0.4%)	1 (0.5%)
	Level 5	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.5%)	1 (0.4%)	0 (0%)
	Missing	0 (0%)	0 (0%)	16 (6.7%)	25 (11.7%)	24 (10.0%)	35 (16.4%)	28 (11.7%)	36 (16.8%)
Usual activities	Level 1	125 (52.1%)	109 (50.9%)	108 (45.0%)	94 (43.9%)	95 (39.6%)	86 (40.2%)	89 (37.1%)	83 (38.8%)
	Level 2	65 (27.1%)	61 (28.5%)	80 (33.3%)	55 (25.7%)	83 (34.6%)	55 (25.7%)	82 (34.2%)	55 (25.7%)
	Level 3	45 (18.8%)	39 (18.2%)	31 (12.9%)	31 (14.5%)	28 (11.7%)	24 (11.2%)	35 (14.6%)	27 (12.6%)
	Level 4	5 (2.1%)	5 (2.3%)	8 (3.3%)	8 (3.7%)	9 (3.8%)	14 (6.5%)	5 (2.1%)	11 (5.1%)
	Level 5	0 (0%)	0 (0%)	1 (0.4%)	1 (0.5%)	1 (0.4%)	0 (0%)	1 (0.4%)	2 (0.9%)
	Missing	0 (0%)	0 (0%)	16 (6.7%)	25 (11.7%)	24 (10.0%)	35 (16.4%)	28 (11.7%)	36 (16.8%)

EQ-5D-5L scale	Health state severity <sup>a</sup>	Baseline		3 months		6 months		12 months	
		Intervention	Usual care	Intervention	Usual care	Intervention	Usual care	Intervention	Usual care
Pain/ discomfort	Level 1	63 (26.3%)	47 (22.0%)	55	43	53 (22.1%)	41 (19.2%)	53 (22.1%)	31 (14.5%)
	Level 2	93 (38.8%)	89 (41.6%)	99 (41.3%)	82 (38.3%)	94 (39.2%)	68 (31.8%)	95 (39.6%)	74 (34.6%)
	Level 3	67 (27.9%)	64 (29.9%)	58 (24.2%)	51 (23.8%)	54 (22.5%)	55 (25.7%)	47 (19.6%)	56 (26.2%)
	Level 4	17 (7.1%)	14 (6.5%)	12 (5.0%)	10 (4.7%)	13 (5.4%)	14 (6.5%)	14 (5.8%)	16 (7.5%)
	Level 5	0 (0%)	0 (0%)	0 (0%)	3 (1.4%)	2 (0.8%)	1 (0.5%)	3 (1.3%)	1 (0.5%)
	Missing	0 (0%)	0 (0%)	16 (6.7%)	25 (11.7%)	24 (10.0%)	35 (16.4%)	28 (11.7%)	36 (16.8%)
Anxiety/ depression	Level 1	153 (63.8%)	137 (64.0%)	132 (55.0%)	116 (54.2%)	130 (54.2%)	104 (48.6%)	119 (49.6%)	109 (50.9%)
	Level 2	62 (25.8%)	51 (23.8%)	69 (28.8%)	51 (23.8%)	60 (25.0%)	51 (23.8%)	63 (26.3%)	46 (21.5%)
	Level 3	20 (8.3%)	23 (10.7%)	22 (9.2%)	18 (8.4%)	24 (10.0%)	20 (9.3%)	28 (11.7%)	19 (8.9%)
	Level 4	4 (1.7%)	2 (0.9%)	1 (0.4%)	4 (1.9%)	2 (0.8%)	3 (1.4%)	2 (0.8%)	4 (1.9%)
	Level 5	1 (0.4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.5%)	0 (0%)	0 (0%)
	Missing	0 (0%)	1 (0.5%)	16 (6.7%)	25 (11.7%)	24 (10.0%)	35 (16.4%)	28 (11.7%)	36 (16.8%)

<sup>a</sup> Level 1 – no problems; level 2 – slight problems; level 3 – moderate problems; level 4 – severe problems; level 5 – extreme problems.



## Appendix 8 Cost-effectiveness planes for sensitivity analyses 1–6

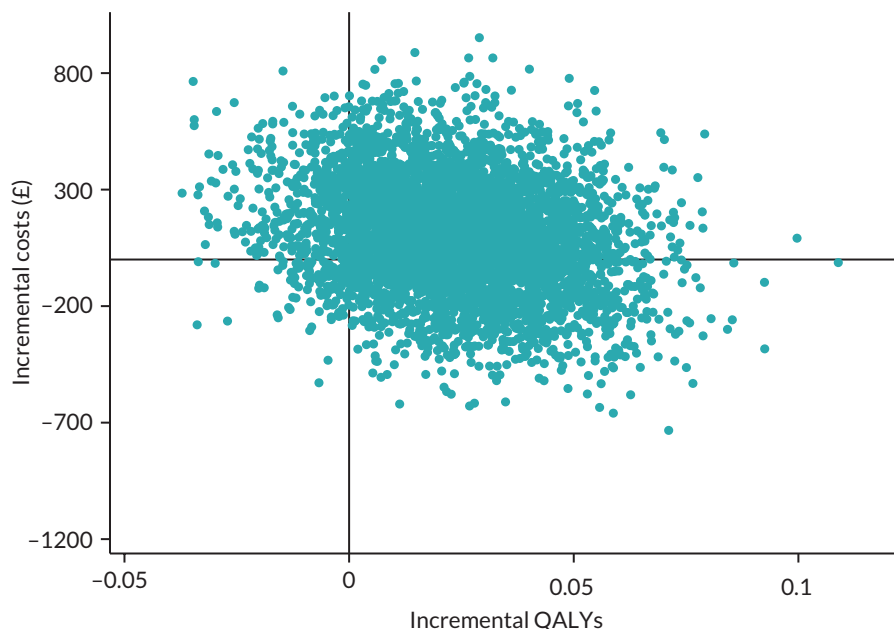


FIGURE 12 Cost-effectiveness plane for SA1: complete-case analysis.

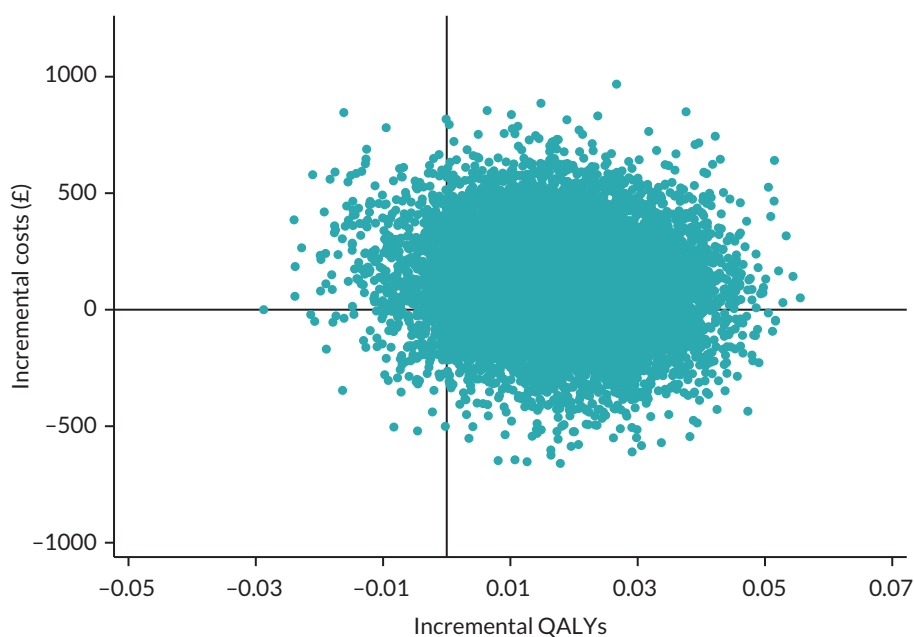
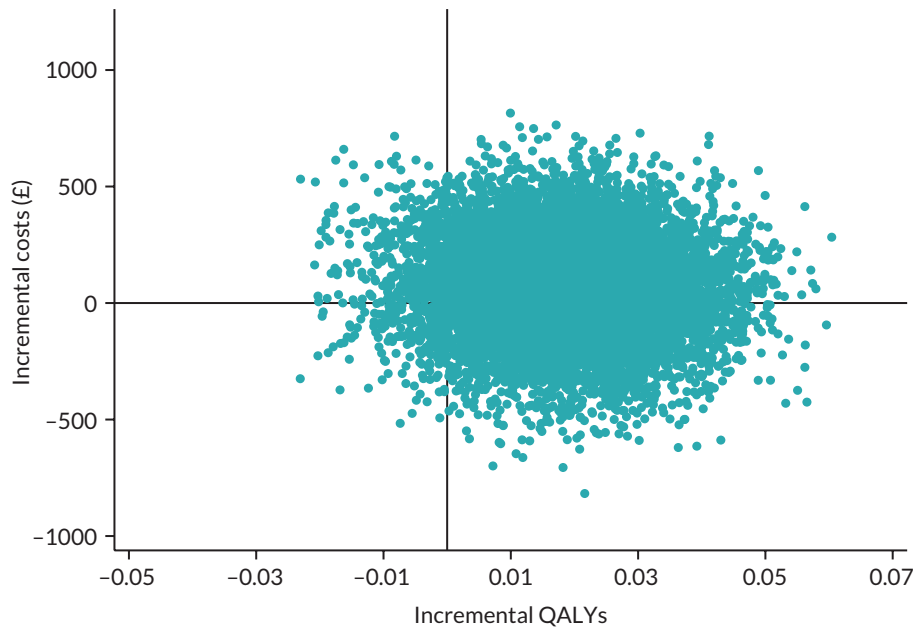
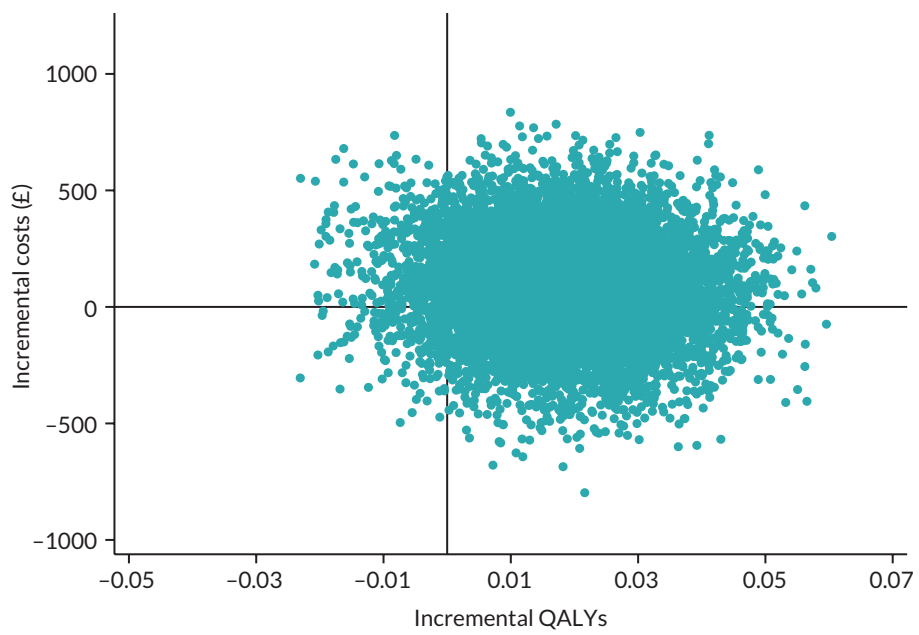


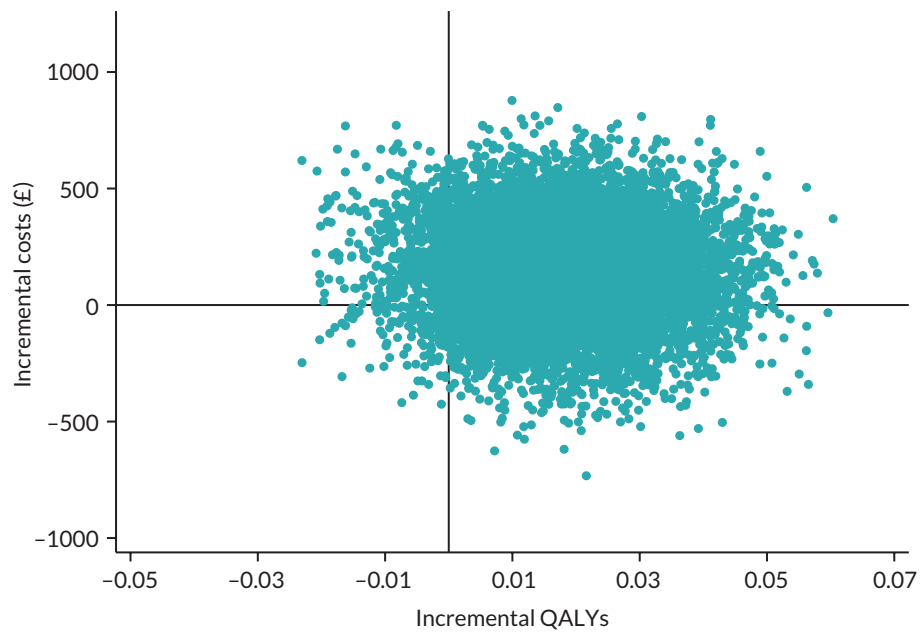
FIGURE 13 Cost-effectiveness plane for SA2: personal costs.



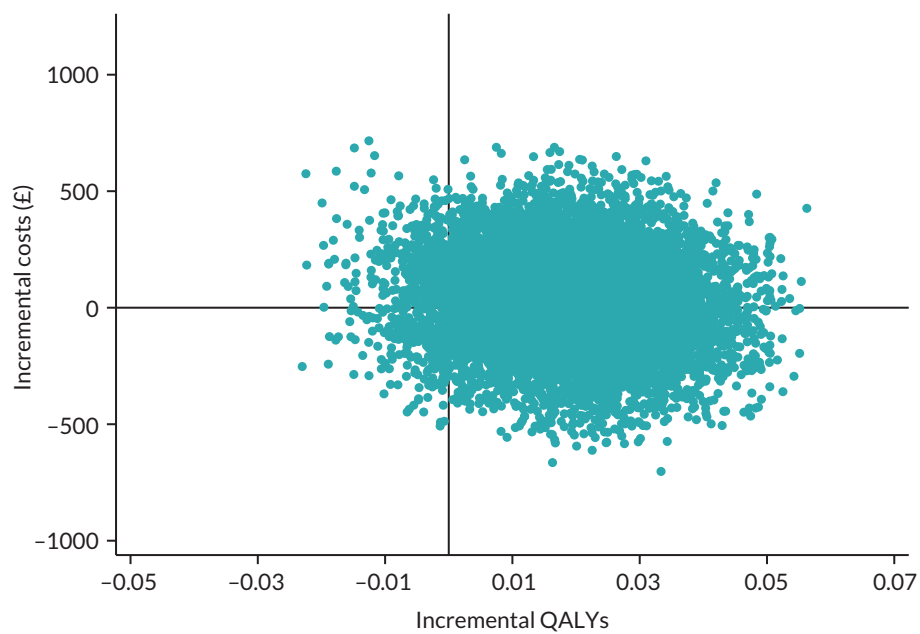
**FIGURE 14** Cost-effectiveness plane for SA3: face-to-face yoga classes only.



**FIGURE 15** Cost-effectiveness plane for SA4: online yoga classes only.



**FIGURE 16** Cost-effectiveness plane for SA5: medication costs removed.



**FIGURE 17** Cost-effectiveness plane for SA6: age and gender removed.



## Appendix 9 Cost-effectiveness acceptability curves for sensitivity analyses 1–6

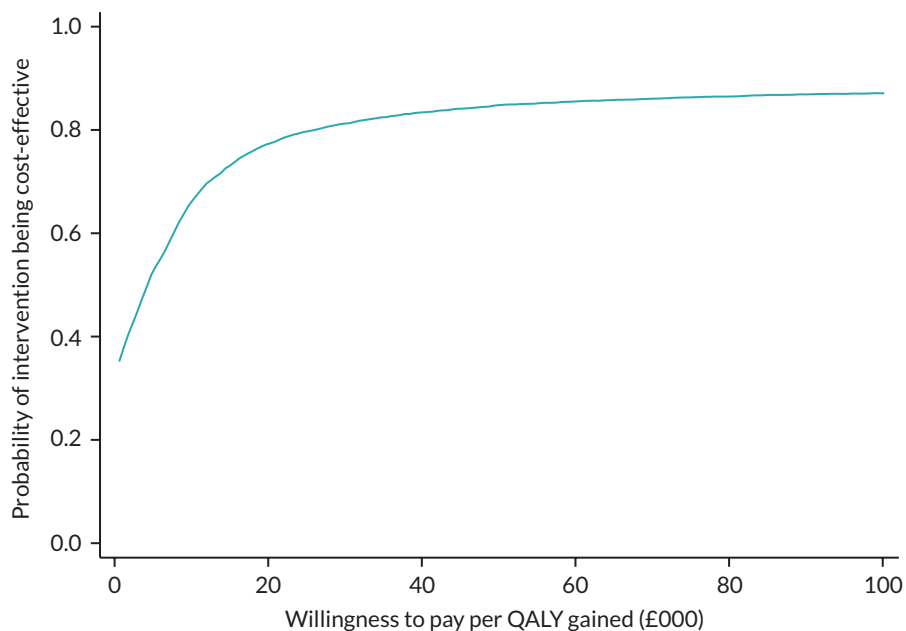


FIGURE 18 Cost-effectiveness acceptability curve for SA1: CCA.

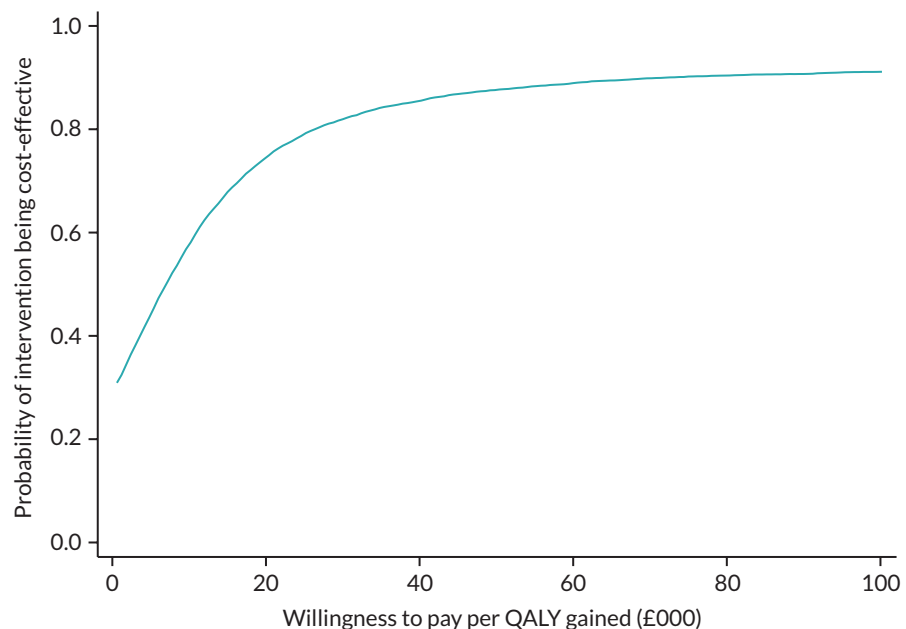
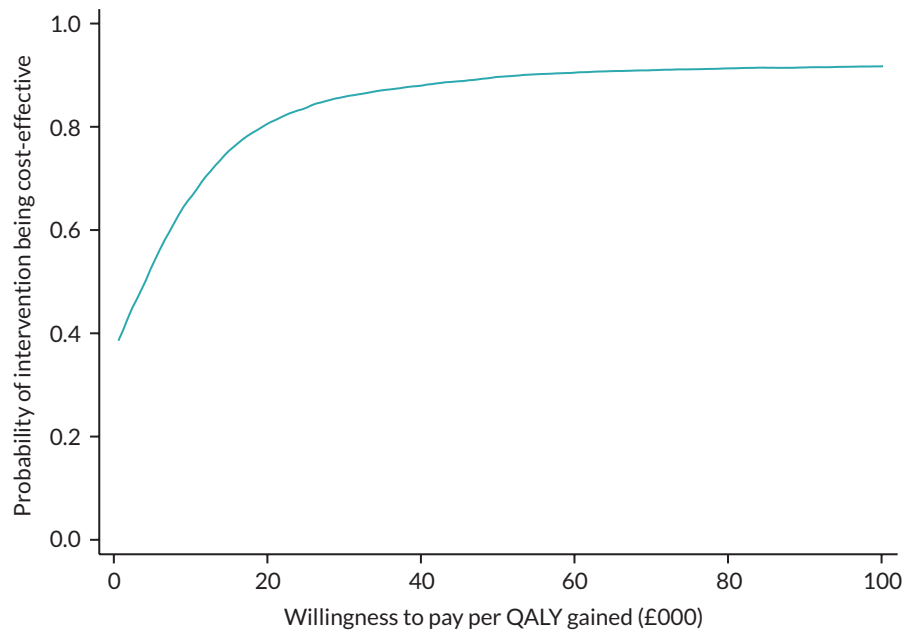
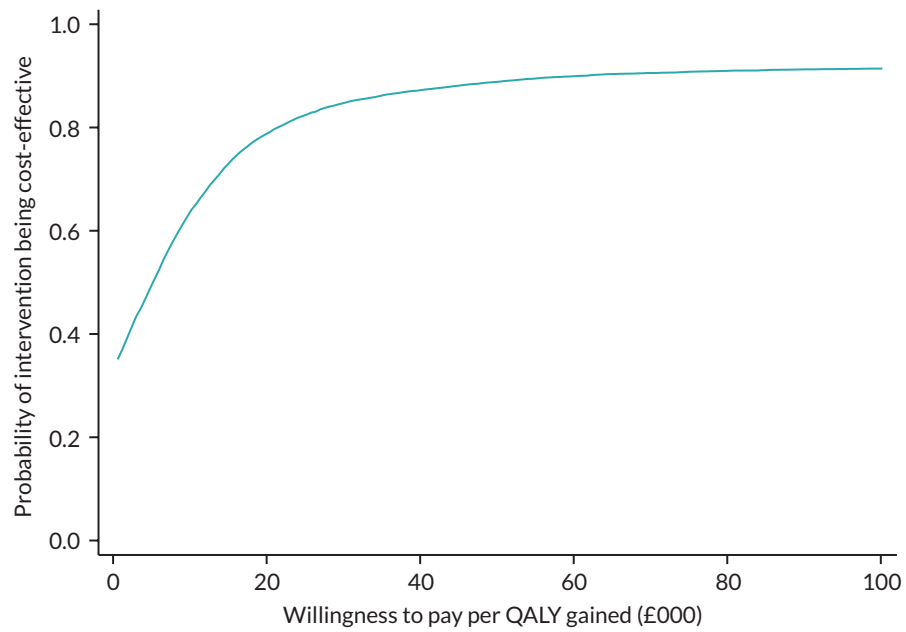


FIGURE 19 Cost-effectiveness acceptability curve for SA2: personal cost.

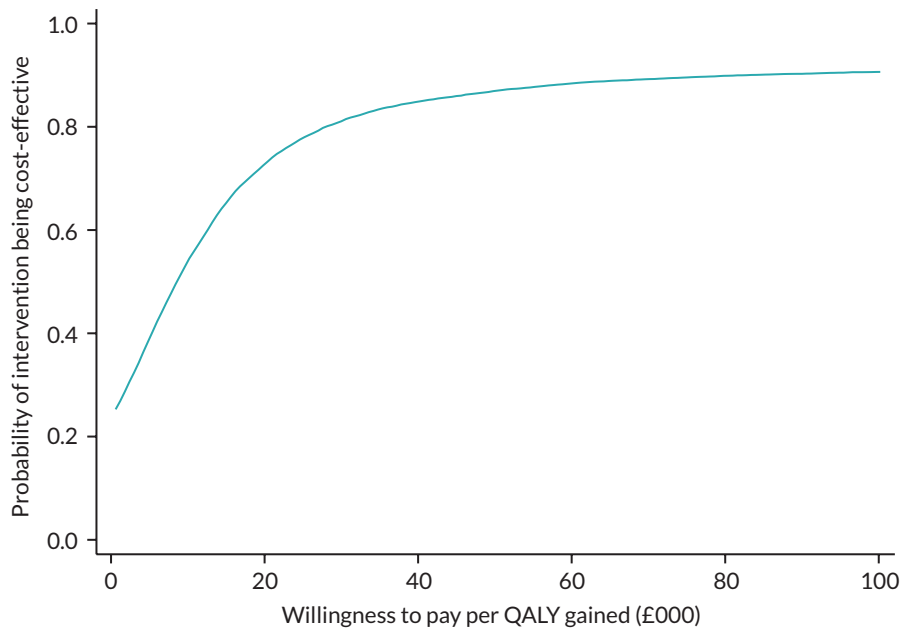




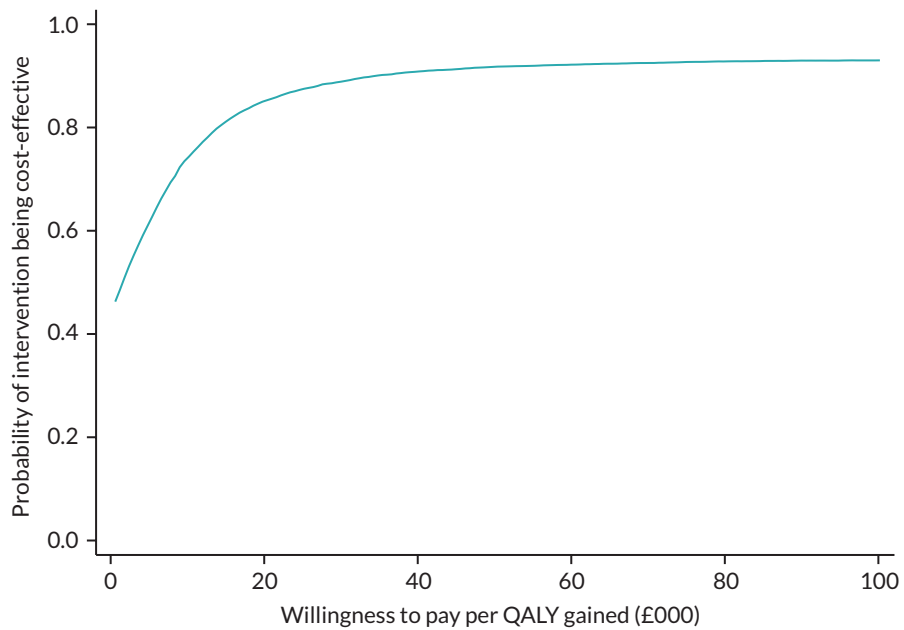
**FIGURE 20** Cost-effectiveness acceptability curve for SA3: face-to-face yoga classes only.



**FIGURE 21** Cost-effectiveness acceptability curve for SA4: online yoga classes only.



**FIGURE 22** Cost-effectiveness acceptability curve for SA5: medication costs removed.



**FIGURE 23** Cost-effectiveness acceptability curve for SA6: age and gender removed.



## Appendix 10 Expected and actual site-level ethnicity (percentages)

**TABLE 51** Expected and actual site-level ethnicity (percentages)

Trial site (anonymised)	White British/Irish/other		Black African/Caribbean/ other		Asian Indian/Pakistani/ Bangladeshi/other/Chinese		Mixed/multiple groups		Other	
	Expected	Actual	Expected	Actual	Expected	Actual	Expected	Actual	Expected	Actual
1	95.6	96.6	0.6	0.0	2.1	3.4	1.4	0.0	0.3	0.0
2	95.6	100.0	0.6	0.0	2.1	0.0	1.4	0.0	0.3	0.0
3	95.5	100.0	0.6	0.0	2.1	0.0	1.4	0.0	0.3	0.0
4	96.0	100.0	0.7	0.0	1.7	0.0	1.1	0.0	0.4	0.0
5	95.6	100.0	0.8	0.0	1.9	0.0	1.1	0.0	0.6	0.0
6	97.2	100.0	0.3	0.0	1.4	0.0	0.9	0.0	0.2	0.0
7	90.5	100.0	1.8	0.0	5.0	0.0	2.1	0.0	0.5	0.0
8	90.6	96.0	3.0	0.0	3.5	0.0	2.3	4.0	0.5	0.0
9	97.3	100.0	0.2	0.0	1.4	0.0	0.9	0.0	0.2	0.0
10	97.3	100.0	0.2	0.0	1.4	0.0	0.9	0.0	0.2	0.0
11	97.3	96.0	0.2	0.0	1.4	0.0	0.9	0.0	0.2	0.0
12	90.6	100.0	2.6	0.0	4.1	0.0	2.2	0.0	0.5	0.0
13	90.8	100.0	2.6	0.0	3.9	0.0	2.2	0.0	0.5	0.0
14	90.6	91.7	1.9	0.0	4.9	8.3	2.1	0.0	0.5	0.0
15	96.2	100.0	0.8	0.0	1.5	0.0	1.2	0.0	0.3	0.0
16	90.6	95.8	2.6	0.0	4.1	0.0	2.2	0.0	0.5	0.0
17	90.6	94.4	3.0	5.6	3.5	0.0	2.3	0.0	0.5	0.0
18	90.6	96.4	2.4	0.0	4.3	3.6	2.2	0.0	0.5	0.0
19	97.2	93.8	0.3	0.0	1.4	0.0	0.9	6.3	0.2	0.0



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***