STATE OF LIFE

Fit For Purpose

The Social Value of Physical Activity Referral Programmes delivered by GM Active

Technical Report

State of Life
London
1 July 2025

Acknowledgements & Credits

Report Author: Jasmine Kazantzis

State of Life Working Group: Lizzie Trotter, Rose Fawcett, Will Watt, Allan Little

GM Active Working Group: Michelle Childs - Chair, GM Active Health Sub-Group; Jon Keating - Head of Business Operations, GM Active; John Oxley - CEO, Life Leisure Stockport; Adam Blezard - Health and Community Lead, Wigan Council; Carly Heselwood - Health and Wellbeing Manager, Trafford Leisure; Chelsea Mears - Live Active Manager, Active Tameside; Elanor Reynolds - Health Manager, Be Well Wigan; Lauren Connis - Health and Fitness Brand Manager, Oldham Community Leisure; Lynsey Johnson - Head of Wellbeing, Wigan Council; Michael Tuson - Strategic Policy and Partnerships Lead, Salford Community Leisure; Nicky Winkley - Sport and Leisure Manager, Salford Community Leisure; Paul Gardner - Community Development Manager, Your Trust Rochdale; Shelley Caulfield - Wellness Service Manager, Bury Live Well; Jackie Veal - Head of Wellness, Bury Council; Shaun Higgins - Director of Health and Social Care, Active Tameside; Simon Blair - Deputy CEO, Oldham Community Leisure; Stuart Lockwood - CEO, Oldham Community Leisure; Louise Sword - Senior Active Lifestyles Officer, Salford Community Leisure; Nicola Woodall - Exercise Referral Lead, Your Trust Rochdale

With thanks to: Greater Manchester Moving, for their ongoing strategic leadership and support; All GM Active delivery staff and survey participants who made this research possible

We would also like to acknowledge the formative role of the early work carried out with Life Leisure to produce a first-order estimate of the social value of the Physical Activity Referral in Stockport (PARiS) programme in Brinnington Park Leisure Centre. The pilot study laid the foundation for this full-scale evaluation. It offered critical proof of concept and highlighted the potential for wider measurement of wellbeing and health outcomes in exercise referral schemes.

Suggested citation: Kazantzis, Jasmine, Lizzie Trotter, Rose Fawcett, Will Watt and Allan Little (2025) *Fit For Purpose: The Social Value of Physical Activity Referral Programmes delivered by GM Active.* London: State of Life and GM Active

STATE OF LIFE

Executive Summary

- **GM Active's physical activity referral programmes are delivering powerful annual benefits per person, valued at £21,800 in wellbeing gains based on HM Treasury's WELLBY measure, and £5,600 in health improvements using the NHS QALY measure.**
- This evaluation is based on survey data from over 2,000 participants and a control group of 1,000 non-participants drawn from waiting lists.
- Taking part in the programme is associated with significant improvements in personal and mental wellbeing, community cohesion, individual development, physical health, and reduced reliance on NHS services.
- Using the **NICE-endorsed EQ-5D index**, we estimate an average health improvement equivalent to **0.08 QALYs per person per year,** which is above the estimated minimal clinically important difference for England's EQ-5D index score of 0.037, and **valued at £5,600 per person, per year.**
- **Self reported GP consultations fell by 19% among participants**, further signalling the programme's potential to relieve pressure on NHS frontline services.
- Programme participants report an average increase of **1.32 points in life satisfaction** - a substantial boost on a 0-10 scale.
- Using HM Treasury-approved wellbeing valuation methods, this converts into a **social value of £21,800 per person per year**, placing GM Active's impact well above most comparable health or physical activity interventions.
- These effects are especially pronounced for those with typically lower baseline wellbeing: those who are inactive, living with disabilities, or from deprived areas.
- Frequent, supervised, and sustained participation is most strongly associated with health and wellbeing gains; at least 4 sessions a week is optimal.
- Key benefits are not limited to a particular moment in the programme lifecycle; they appear early and extend beyond the period of active participation.
- Fixed effects analysis supports the programme's impact on wellbeing and community cohesion, though weaker results for physical health highlight the need for further longitudinal research with larger samples to strengthen causal inference.
- These findings make the case for expanding and embedding exercise referral schemes as a core component of preventative healthcare. They also support GM's wider ambition to create a population health system that reduces inequality, boosts wellbeing, and reduces pressure on public services.

STATE OF LIFE

STATE OF LIFE

Glossary

EQ-5D-5L Questionnaire: A standardised, widely used instrument for measuring health-related quality of life. It assesses five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.

Exercise referral schemes: programmes that include a mix of supervised sessions (e.g. structured group activities or one-to-one support with an instructor), non-supervised sessions (e.g. open gym access or self-guided activity), and specialist classes tailored for people with specific health conditions (e.g. cardiac rehabilitation or pain management).

HM Treasury's Green Book: Official guidance from the UK government on how to appraise policies, programmes, and projects, including guidance on valuing wellbeing.

IMD (Index of Multiple Deprivation): A UK government statistical release that provides a relative measure of deprivation across small areas in England.

NICE (National Institute for Health and Care Excellence): An organisation in the UK, sponsored by the Department of Health and Social Care, that provides guidance and advice on public health and social care.

Pooled OLS (Ordinary Least Squares) Regression: A statistical method used to analyse the relationship between a dependent variable and one or more independent variables, combining data from individuals across multiple time points. It can account for the influence of factors we're not directly interested in, like sociodemographics.

Ordered Logit Regression: A statistical method used when the dependent variable has categories with a natural order, like "inactive," "fairly active," and "active"). It models the probability of being in a particular category based on independent variables.

FE (Fixed Effects) Regression: A statistical method that controls for unchanging individual or group characteristics in panel data, allowing to isolate the effects of variables that change over time.

QALY (Quality-Adjusted Life Year): A measure of health outcome that combines both the length and quality of life. One QALY is equal to one year of life in perfect health.

SALS (Sport England's Short Active Lives Survey): A survey used to measure and classify physical activity levels.

SWEMWBS (Short Warwick-Edinburgh Mental Wellbeing Scale): A scale used to measure mental wellbeing and psychological functioning.

WELLBY (Wellbeing-Adjusted Life Year): A measure of wellbeing that quantifies the impact of an intervention on life satisfaction, typically valued in monetary terms.

1. Introduction

GM Active is a collective of twelve organisations that function across all ten boroughs of Greater Manchester (GM). Together, they deliver a wide spectrum of services spanning sport, physical activity, health, wellbeing, and culture.

All organisations provide a range of 'health intervention programmes' not restricted to but including exercise referral schemes. Some of these interventions are externally commissioned while others are delivered independently, but all are designed for those who need them most: people living with disability, long-term health conditions, and/or chronic inactivity that could lead to more serious health issues.

The exercise referral schemes vary in format but typically include a mix of supervised sessions (e.g. structured group activities or one-to-one support with an instructor), non-supervised sessions (e.g. open gym access or self-guided activity), and specialist classes tailored for people with specific health conditions (e.g. cardiac rehabilitation or pain management). Some participants also engage in activities independently, supported by initial referrals or guidance. This flexible delivery model is intended to meet people where they are, in terms of their health, confidence, and preferences, and to provide pathways into sustained physical activity.

This research study was commissioned to evaluate and demonstrate credible evidence of the social impact and value of GM Active exercise referral schemes, with the aim of informing a broader, long-term understanding of how leisure facilities can enhance quality of life and be re-positioned in making a valuable contribution to improved population health in the context of an integrated care system. Specifically, it assesses how these schemes improve health and wellbeing using valuation approaches aligned with HM Treasury's Green Book Supplementary Guidance: Wellbeing Guidance for Appraisal[1]. As stated in the supplementary guidance (p.5):

"The appraisal of social value, also known as public value, is based on the principles and ideas of welfare economics and concerns overall social welfare efficiency, not simply economic market efficiency. Social or public value therefore includes all significant costs and benefits that affect the welfare and wellbeing of the population, not just market effects."

[1] https://assets.publishing.service.gov.uk/media/60fa9169d3bf7f0448719daf/Wellbeing_guidance_for_appraisal_-_supplementary_Green_Book_guidance.pdf

At the heart of this analysis is the premise that improvements in health and wellbeing among participants are the main socially desirable outcome of GM Active's referral programmes.

To quantify this, we collected data via an individual-level survey of participants and waitlisters (non-participants) at two points in time. We applied statistical and econometric techniques, including regression analysis, to estimate the causal impact of the programme on personal wellbeing, community cohesion, health, and other outcomes. These impacts were then monetised using the latest UK Government guidance on valuing health and wellbeing for policy evaluation, using the Quality-adjusted Life Year (QALY) and Wellbeing-adjusted Life Year (WELLBY) approaches, respectively.

The remainder of the report outlines the research questions and methodology used, presents the results of the survey and estimation of impacts, details the valuation of health and wellbeing gains, and concludes with a discussion of the study's limitations, implications, and areas for further research.

2. Research Questions

This study aims to address the following research questions:

1. Is participation in GM Active physical activity referral programmes associated with improvements in key outcomes, including:
    a. Personal wellbeing
    b. Mental wellbeing
    c. Physical health
    d. Physical activity level
    e. Community cohesion
    f. Individual development
2. Does the duration, intensity and/or type of programme participation play a role in key wellbeing outcomes?
3. Do wellbeing outcomes vary across key demographic characteristics or referral sources?
4. Is participation in the programme associated with reduced NHS service utilisation?
5. What is the monetary value of any health and wellbeing improvements?
6. Do programme participants intend to remain physically active over the next six months, and if so, as members of the leisure centre?

3. Methodology and Data

To measure the effect of GM Active's exercise referral schemes on participants' health and wellbeing, State of Life developed a bespoke survey which was administered by eight GM Active partner organisations to programme participants and waitlisters. The participating GM Active organisations and their respective schemes are outlined in the following table.

Table 1: Participating GM Active organisations and their respective schemes

GM Active organisation	Exercise Referral Scheme
1. Bury Council	Bury Live Well Service
2. Wigan Council	Be Well
3. Life Leisure (Stockport)	Physical Activity Referral in Stockport (PARiS)
4. Trafford Leisure	Physical Activity Referral
5. Active Tameside	Live Active
6. Salford Community Leisure	Active Lifestyles
7. Oldham Active (Oldham Community Leisure)	Exercise Referral Scheme
8. Your Trust (Rochdale)	Exercise Referral Scheme

3.1 Research design, sample and control

GM Active operates with a waiting list for individuals eligible to join the exercise referral scheme.[2] Schemes accept participants from their waiting list mostly on a first-come, first-served basis, however some prioritise clients referred from a specialist NHS programme (e.g. cardiac, pulmonary or neuro rehab team) first. In this study, we leveraged this waiting list as a comparison group, forming a quasi-experimental design. This approach allowed us to compare people with similar characteristics - some who received the intervention and others who had not yet received it - so we could estimate the intervention's impact. By using the waiting list, we avoid ethical concerns associated with fully randomised selection, while also accounting for other influencing factors. Although randomised controlled trials (RCTs) are considered the gold standard for establishing cause and effect, they are often expensive and

[2] Some examples of health conditions included in the selection criteria for the programmes are previous heart attacks or heart disease, high blood pressure, high cholesterol, overweight or obese, diabetes type 2, depression, joint pain and COPD.

may not reflect real-world conditions. Our method improves external validity (i.e. the extent to which the results are likely to apply to broader, real-world settings) while remaining cost-effective and practical.

3.2 Data collection

Survey responses were collected in two waves. Wave 1 was conducted between 7th August and 18th September 2024 and wave 2 followed between 4th December 2024 and 7th February 2025. Following data cleaning to remove invalid responses, 3,181 responses were retained for analysis. 290 individuals were present in both rounds, 32 of which stayed on the waitlist between rounds, 55 moved from the waitlist to the programme, and 203 stayed in the current/past participant group.

Table 2: The sample, following data cleaning

	Non-participants	Participants	All
Wave 1 (7 Aug -18th Sept 2024)	577	1,027	1,604
Wave 2 (4th Dec 2024 - 7th Feb 2025)	482	1,095	1,577
Total	1,059	2,122	3,181

3.3 Analytical approach

We begin by presenting descriptive statistics that summarise average outcomes for both the participant group (treatment) and the waitlist group (control). This initial comparison provides a broad understanding of trends and differences across a range of key wellbeing, health, and social outcomes before applying deeper statistical methods.

The survey captured the following self-reported outcomes:

- **Personal wellbeing**: ONS-4 - life satisfaction, happiness, anxiety, and worthwhileness[3]
- **Mental wellbeing**: Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS)[4]
- **Physical health**: Health-related quality of life, using the EQ-5D-5L questionnaire[5], and NHS service use (GP consultations and emergency services)

[3] https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingsurveyuserguide
[4] https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/strengths/
[5] https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/; cleaned using EuroQol's published STATA syntax code

- **Physical activity levels**: Sport England's Short Active Lives Survey (SALS) activity level classifications: Inactive / Fairly active / Active, based on moderate-equivalent minutes per week from all activities[6]
- **Community cohesion**: loneliness, sense of belonging, sense of trust[7]
- **Individual development**: self-efficacy, measuring agreement with the statement - 'I can achieve most of the goals I set myself'.

Full detail on the questions and instruments used to capture these outcomes can be found in appendix 1.

Following the descriptive stage, we perform pooled ordinary least squares (OLS) regression analysis to explore the relationships between participation in the exercise referral schemes and the outcome measures, while controlling for potentially confounding variables. This model is cross-sectional in nature but provides a more robust estimate of associations between participation and outcomes than simple comparisons, by accounting for differences in observable characteristics. This method enables us to isolate the effect of programme participation from other factors known to influence wellbeing, accounting for individual differences that may otherwise bias the estimated effect of the intervention. Specifically, we control for a comprehensive set of demographic and background characteristics, including: gender, age group, ethnicity, marital status, highest educational qualification, employment status, religion, number of children, presence of a long-term health impairment, and local area deprivation. We also control for the wave in which the data was collected, survey response mode (i.e. self-completion or via interview), and the organisation/ leisure centre offering the programme to the participant.

The primary independent variable of interest is programme participation status. In wave 1 of data collection, respondents were given the option to report as either on the waitlist or having started or completed a programme. In wave 2, the survey was refined to distinguish between those who had started and those who had completed the programme, allowing for more nuanced analysis of engagement levels.

In addition to estimating overall programme effects, we explore the impact of duration, intensity and type of participation. We also carry out disaggregated analyses to examine

[6] https://evaluationframework.sportengland.org/media/1357/short-active-lives-survey-what-it-is-and-how-to-use-it-1.pdf
[7] Loneliness and sense of belonging questions from Understanding Society, and sense of trust and self-efficacy from Active Lives

differential effects across key subgroups, including gender, age group, area deprivation levels, presence of a long-term health condition and referral source.

To test the robustness of our findings, we run several sensitivity checks, including the addition of self-reported general health as a control variable, panel regression analysis with fixed effects, and a wave 2-only analysis, to assess the varying impacts of current versus completed participation compared to being on the waitlist. The fixed effects models allow us to account for all stable, unobserved individual differences, making stronger causal inference possible. This model is used only where we have repeated measures for the same individuals, and thus represents the most sophisticated level of analysis applied in this evaluation. These checks help ensure that the estimated effects are not driven by model specification or sampling differences across waves.

Finally, to estimate the social value of GM Active's exercise referral schemes in monetary terms, we apply HM Treasury-endorsed valuation methods. Wellbeing improvements are converted into Wellbeing-Adjusted Life Years (WELLBYs), and physical health gains are converted into Quality-Adjusted Life Years (QALYs).

4. Findings

4.1 Summary statistics

In this section, we examine the characteristics of the sample and present the summary statistics for the outcomes of interest. This provides a clear picture of the sample's profile and the outcome measures we are evaluating before moving on to regression analysis.

4.1.1 About the sample

To support the validity of this quasi-experimental approach, it is important to demonstrate that the treatment group (participants) and control group (waitlisters/non-participants) are broadly similar in their key demographic characteristics. This similarity helps ensure that any differences in outcomes can be more confidently attributed to the intervention itself, rather than to pre-existing differences between the groups. However, where differences do exist, they can be reasonably accounted for in the regression analysis by controlling for participant characteristics, helping to reduce potential bias.

Based on the demographic data in Table 3, participants and non-participants are broadly similar across key characteristics. Participants are slightly older than non-participants on average, which is to be expected given that the participant group includes individuals who completed the programme anywhere from a few months to several years ago, allowing more time for age differences to emerge. They are also more likely to be married and/or retired, life circumstances that usually correlate with older age.

A difference is observed in the proportion reporting a long-term physical or mental health condition, with 65% of participants compared to 73% of non-participants. Given that such conditions are a key eligibility criterion for GM Active's physical activity referral scheme, this is noteworthy. However, further breakdowns of participants in wave 2 suggest that the proportion with a long-term condition is similar for those on the waiting list (72%) and for those currently on the programme (68%), and slightly lower for previous programme participants (64%). This may reflect positive changes in participants' health over time, rather than differences at the point of entry.

Overall, the demographic similarities between groups support their comparability, with the observed difference in health conditions potentially pointing to programme impact rather than bias in sample composition.

Table 3: Demographics of participants compared to non-participants

	Non-participants	Participants
Number of respondents[8]	972	1,905
Demographics		
Gender		
Male	32.2%	33.6%
Female	65.7%	64.7%
Other	0.3%	0.3%
Age, in 6 categories		
16-34	10.8%	4.6%
35-44	11.7%	8.6%
45-54	17.9%	13.9%
55-64	25.9%	26.7%
65-74	20.5%	27.5%
75+	8.8%	14.3%
Ethnicity		
White	88.5%	90.8%
Asian or Asian British	4.7%	3.0%
Black or Black British	1.7%	1.7%
Mixed	1.1%	1.4%
Other	1.2%	1.0%
Marital status		

[8] For the 304 individuals who responded to the survey in both waves, we present their most recent (wave 2) demographic information in Table 3.

	Non-participants	Participants
Number of respondents[8]	972	1,905
Demographics		
Single	22.5%	15.6%
Living as a couple (but not legally married)	11.3%	9.1%
Married	45.9%	54.4%
Divorced	6.9%	7.3%
Separated	2.7%	2.0%
Widowed	6.8%	8.4%
Highest educational qualification[9]		
Level 1 (GCSE D-G) and below	6.3%	4.0%
Level 2 (GCSE A*-C) and equivalents	17.8%	16.1%
Level 3 (A-levels) and equivalents	13.4%	12.1%
Level 4 (CHE) and equivalents, or higher	27.8%	34.8%
No qualification	12.6%	10.6%
Other	8.7%	8.3%
Employment status		
In-paid employment (full-time or part-time)	34.3%	30.8%
Long-term sick or disabled	14.3%	9.2%
Looking after family or home	3.6%	1.9%
Retired	32.0%	45.4%
Self-employed	2.8%	4.1%
Student (full-/part-time)	0.9%	0.5%
Unemployed	7.8%	4.8%

[9] In wave 1, we identified a substantial amount of responses in the "other" category that could have fallen into one of the other 4 education level categories, therefore we expanded the options in wave 2 to 7 levels, but later collapsed the additional levels into "Level 4 or above".

	Non-participants	Participants
Number of respondents[8]	972	1,905
Demographics		
Other	1.5%	0.7%
Religion		
Christian	57.3%	65.4%
Muslim	4.7%	2.9%
No religion	28.8%	24.1%
Other[10]	4.3%	3.8%
Number of children		
None	78.5%	84.9%
1	10.1%	6.7%
2	6.1%	3.2%
3 or more	2.3%	1.8%
Has a long-term physical or mental health condition		
No	20.1%	28.8%
Yes	73.3%	65.0%
Local Area Deprivation (IMD[11]) 2024, in 3 categories		
High deprivation (1-3)	45.7%	36.4%
Medium deprivation (4-7)	27.7%	32.8%
Low deprivation (8-10)	21.8%	25.6%

4.1.2 Outcome measures

Next, we compared average outcomes between participants and non-participants. These are simple comparisons that show raw differences, without adjusting for other sociodemographic

[10] Jewish, Hindu, Buddhist and Sikh were given as options in the question.
[11] Index of Multiple Deprivation

factors, and **should not be used to infer the impact of the programme**. As shown in Table 4, participants reported markedly better outcomes across a wide range of wellbeing, health, and social indicators. On average, they scored higher on personal wellbeing measures (e.g. life satisfaction: 6.6 vs. 4.9 on a 0-10 scale), mental wellbeing (SWEMWBS: 24.6 vs. 21.9 on a 7-35 scale), and physical health, with a notably higher average EQ-5D index (0.70 vs. 0.58 on a -0.285-1 scale). Participants also reported fewer GP visits and emergency service uses over the past month, on average, suggesting lower healthcare utilisation. As expected, rates of physical activity were notably higher among participants, with only 18% classified as inactive versus 42% among non-participants. In addition, participants reported greater social connectedness - feeling less lonely, more trusting, and a greater sense of belonging - alongside higher self-efficacy. Taken together, these descriptive statistics indicate a consistent association between programme participation and improved health and wellbeing outcomes, which is further explored in the regression analysis in section 4.2.

Table 4: Comparing outcomes of participants and non-participants

	Non-participants	Participants	All
Number of responses[12]	1,059	2,122	3,181
Outcomes (response scale)			
Personal Wellbeing			
Life Satisfaction (0-10)	4.85 (1047)[13]	6.58 (2111)	6.00 (3158)
Happiness (0-10)	5.33 (1040)	6.84 (2106)	6.34 (3146)
Anxiety (0-10)	4.54 (1040)	3.77 (2101)	4.03 (3141)
Worthwhileness (0-10)	5.67 (1038)	7.17 (2109)	6.68 (3147)
Mental Wellbeing			
SWEMBWBS (7-35)	21.93 (1034)	24.56 (2084)	23.69 (3118)
Physical Health			
Health-related QoL (EQ-5D Index, -0.285-1)	0.58 (1020)	0.70 (2063)	0.66 (3083)

[12] For the 304 individuals who responded to the survey in both waves, we include both of their responses to key outcomes in Table 3

[13] Numbers in brackets indicate the number of valid responses for each outcome measure. These may be lower than the total sample size due to item non-response or partial survey completion.

	Non-participants	Participants	All
Number of responses[12]	1,059	2,122	3,181
Outcomes (response scale)			
GP visits in the past months	1.26 (1024)	0.94 (2074)	1.05 (3098)
Emergency service use in the past month	0.15 (1037)	0.10 (2086)	0.12 (3123)
Physical Activity (PA) Classifications			
Inactive - fewer than 30 mins of PA a week	42.0% (445/1059)	18.4% (391/2122)	26.3% (836/3181)
Fairly Active - 30-149 mins of PA a week	12.8% (136/1059)	17.1% (362/2122)	15.7% (498/3181)
Active - at least 150 mins of PA a week	21.7% (230/1059)	38.5% (817/2122)	32.9% (1047/3181)
Community Cohesion			
Loneliness (1-5)	3.15 (1047)	2.70 (2111)	2.85 (3158)
Belonging (1-5)	3.34 (1044)	3.57 (2113)	3.49 (3157)
Trust (1-5)	3.13 (1043)	3.37 (2105)	3.29 (3148)
Individual Development			
Self-efficacy	3.11 (1046)	3.46 (2111)	3.35 (3157)

4.2 Regression Results

While the descriptive statistics are useful for illustrating broad patterns, they cannot tell us whether the observed differences are caused by the programme. Participants and non-participants may differ in many ways unrelated to the intervention, such as age, gender, socioeconomic status, or pre-existing motivation and health. To account for these differences, we use pooled OLS regression analysis, which adjusts for a range of confounding variables. This allows us to better isolate the relationship between participation in GM Active's exercise referral schemes and a range of wellbeing, health, and social outcomes, except for the three-level activity classification variable (inactive, fairly active, active), where we estimated

an ordered logit regression model.[14] Results are shown below and full regression tables are available on request[15].

Each row in the tables below represents a separate regression model, where the dependent variable (listed in the first column) is one of the key outcomes of interest. The second column shows the coefficient for our main treatment variable: participation in GM Active's exercise referral programme (coded as 1 for current and previous participants, 0 for non-participants). This coefficient represents the average difference in the outcome between participants and non-participants, after controlling for other variables.

Because ordered logit coefficients are less intuitive to interpret directly, we report marginal effects for physical activity classification which show the average change in the probability of being in each activity level associated with programme participation.

All models control for a consistent set of background characteristics as outlined in section 3.3. For clarity and space, only the treatment effect is reported here, but full regression outputs are available on request.

Table 5: Regression coefficients indicating impact of participation on key outcomes

Outcome variable (response scale)	Coefficient
Personal Wellbeing	
Life Satisfaction (0-10)	1.32***
Happiness (0-10)	1.11***
Anxiety (0-10)	-0.44***
Worthwhileness (0-10)	1.09***
Mental Wellbeing	
SWEMBWBS (7-35)	1.46***
Physical Health	
Health-related QoL (EQ-5D Index, -0.285-1 scale)	0.08***
GP visits in the past month	-0.22***

[14] An ordered logit model is used for physical activity classification as it accounts for the non-linear, ordinal nature of the outcome without assuming equal spacing between categories, unlike OLS.
[15] To access full regression tables, please contact hello@stateoflife.org.

Outcome variable (response scale)	Coefficient
Emergency service use in the past month	-0.04
Community Cohesion	
Loneliness (1 - never to 5-often/always)	-0.21***
Belonging (1 - strongly disagree to 5 - strongly agree)	0.11***
Trust (1 - strongly disagree to 5 - strongly agree)	0.10***
Individual Development	
Self-efficacy (1 - strongly disagree to 5 - strongly agree)	0.21***
Physical Activity (PA) Classifications	Marginal effect[16]
Inactive - fewer than 30 mins of PA a week	-0.21***
Fairly Active - 30-149 mins of PA a week	-0.01
Active - at least 150 mins of PA a week	0.22***

Stars indicate statistical significance levels: * p<10%, ** p<5%, *** p<1%.

4.2.1 Personal wellbeing

As shown in Table 5, participation in GM Active's exercise referral schemes is associated with statistically significant improvements in personal wellbeing. On average, life satisfaction scores are 1.32 points higher (on a 0-10 scale) among participants than non-participants, while happiness and worthwhileness scores are higher by 1.11 and 1.09 points, respectively. These are sizeable effects when benchmarked against population-level changes in wellbeing: analysis of national data shows moving from unemployment to employment leads to a 0.46-point increase in life satisfaction.[17] Participation in GM Active's exercise referral schemes is associated with an almost 3x times greater difference. Anxiety scores are also significantly 0.44 points lower, indicating reduced anxiety. These collective findings suggest that the scheme improves all aspects of personal wellbeing, pointing to a robust and multidimensional uplift in subjective wellbeing.

[16] Marginal effects show the average change in the probability of being in each physical activity category (Inactive, Fairly Active, Active) associated with participating in the programme, holding other variables constant. These can be interpreted as percentage point changes.

[17] https://assets.publishing.service.gov.uk/media/60fa9169d3bf7f0448719daf/Wellbeing_guidance_for_appraisal_-_supplementary_Green_Book_guidance.pdf, pg. 62

4.2.2 Mental wellbeing

In addition to improvements in subjective wellbeing, participants also report better mental wellbeing as measured by the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS). On average, scores are 1.46 points higher among participants than non-participants (7–35 scale) - a difference that is both statistically significant and practically meaningful. According to the University of Warwick, the minimal detectable change on the SWEMWBS at the individual level is estimated to be between 1 and 3 points, depending on the calculation method used, suggesting that participation in the scheme is likely contributing to real, perceptible improvements in mental wellbeing, rather than merely statistical variation.[18] This suggests that the programme is not only improving how participants feel but is also likely contributing to better overall psychological functioning.[19]

4.2.3 Physical health

Participants in the scheme report significantly greater physical health outcomes, as measured by both a preference-based approach (EQ-5D) and healthcare usage (GP consultations and emergency service use). The average EQ-5D index score, used widely in health economics to quantify health-related quality of life, is 0.08 points higher among participants. This is notable on a scale where even a 0.037 change is often regarded as clinically important in England.[20] In terms of service use, participants report 0.22 fewer GP visits per month, a 19% difference.[21] These reductions point to meaningful demand-side savings for the health system and signal the preventive health benefits of structured physical activity support.

4.2.4 Community cohesion

Participation is also associated with stronger social connectedness. Participants report lower levels of loneliness (-0.21 points) and a modestly greater sense of belonging (+0.11) and trust in others (+0.10), all statistically significant at the 1% level. While the effects are smaller in magnitude than wellbeing gains (even when considering differences in measurement scales) these effects are consistent and reinforce the role of physical activity schemes in supporting community cohesion. These outcomes also suggest the programme may generate “spillover” social benefits that go beyond individual health and wellbeing.

---

[18] https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/wemwbsvsswemwbs/

[19] https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/strengths/

[20] The minimally important difference (MID) for England’s scoring algorithm has been estimated to be 0.037

[21] This percentage change was calculated using the marginal effects from the regression analysis, dividing the difference in the predicted number of GP visits between participants and waitlisters (0.22) by predicted GP visits of waitlisters (1.18).

4.2.5 Individual development

Participation in the exercise referral programme is also associated with greater self-efficacy, that is, the belief in one's ability to manage and take control of life situations. Participants scored 0.21 points higher on the self-efficacy scale (1 to 5) than non-participants, a significant difference. Increased self-efficacy is particularly valuable in the context of behaviour change: people who feel more capable are more likely to sustain physical activity, manage long-term health conditions, and engage with wider health-promoting behaviours.[22] This finding strengthens the case for exercise referral schemes as tools for longer-term empowerment and resilience.

4.2.6 Physical activity

The ordered logit analysis of physical activity classifications reveals a strong positive association between programme participation and being 'active', defined as engaging in at least 150 minutes of activity per week. Participation is associated with a 22 percentage point (pp) higher probability of being active, and a 21pp lower probability of being inactive. These effects suggest a significant behavioural shift driven by the scheme. The net difference in the "fairly active" group was not statistically significant, likely because participants are transitioning both from inactivity to fairly active, and from fairly active to fully active. This pattern indicates that the programme is helping individuals move progressively up the physical activity spectrum.

4.2.7 Duration, intensity and type of participation (on life satisfaction)

Having established the overall association between programme participation and life satisfaction, we next examined whether this relationship varies by how participants engage, specifically, in terms of duration, intensity, and type of activity. Understanding these gradients can help identify what forms of engagement are most effective and inform programme design to maximise wellbeing impact.

To explore this, we estimated a series of regression models where the main treatment variable, programme participation, was replaced by sub-categories representing duration, intensity, and type of engagement. This allows us to estimate the difference in life satisfaction

[22] https://pmc.ncbi.nlm.nih.gov/articles/PMC6003667/

for each subgroup compared to those not yet participating, helping us understand which aspects of engagement are most strongly associated with wellbeing gains.

The results presented in Table 6 reveal a clear pattern: life satisfaction is positively associated with both longer duration and greater intensity of participation. Even minimal engagement (0-3 weeks) is associated with a 0.78-point higher life satisfaction score, while the largest difference is seen among those participating for 4-6 weeks (+1.61). All duration bands from 4 weeks onwards show consistent and significant gains of around 1.4 points or more, suggesting that the benefits appear early and are sustained over time. See Figure 1 in section 4.4 for a graphical representation of duration on the programme.

The intensity of engagement similarly shows a positive association that resembles a dose-response relationship. Participants attending fewer than three sessions per week still report life satisfaction scores that are 1.16-points higher, on average, but this association rises to 1.87 points among those attending 4-6 sessions a week, the highest of any subgroup, and remains high at 1.83 points for 7+ sessions. These differences suggest that the frequency of attendance plays a crucial role in shaping the programme's impact, and that at least 4 sessions a week is optimal.

When examining the type of activity, both supervised and non-supervised sessions are positively associated with life satisfaction, but supervised activity yields a larger effect: 0.56 points higher life satisfaction for every additional supervised session, compared to 0.30 points higher for every additional non-supervised session. This may reflect the added value of structured support, social interaction, or professional guidance. Attending a specialist class, designed to support condition-specific needs, is also linked to a more substantial difference (+1.60) compared to those not attending such sessions (+1.22), highlighting the potential added benefit of sessions designed for specific health conditions.

Taken together, these findings underscore that while any engagement is better than none, more frequent, longer-lasting, and structured forms of participation are most strongly associated with wellbeing gains. This has important implications for how referral schemes are delivered and incentivised. Encouraging consistent attendance, offering tailored sessions, and ensuring access to supervised support may amplify the programme's effectiveness, particularly for individuals who stand to benefit most.

Table 6: Regression coefficients indicating impact of duration, intensity and type of participation on life satisfaction

Treatment variable sub-category	Life Satisfaction Coefficient
Current or past participant	1.32***
Duration of participation	
0-3 weeks	0.78***
4-6 weeks	1.61***
7-9 weeks	1.38***
10-12 weeks	1.43***
12+ weeks (completers)	1.39***
Intensity of participation	
Fewer than 3 sessions a week	1.16***
4-6 sessions a week	1.87***
More than 7 sessions a week	1.83***
Type of participation	
Number of weekly supervised sessions	0.56***
Number of weekly non-supervised sessions	0.30***
Attends a weekly specialist class (e.g. cardiac rehab class)	1.60***
Doesn't attend a specialist class	1.22***

Stars indicate statistical significance levels: * p<10%, ** p<5%, *** p<1%.

4.2.8 Disaggregation - variation in life satisfaction by different groups

To better understand how different groups experience the wellbeing benefits of participation, we disaggregated life satisfaction effects by key participant characteristics, including gender, age, area deprivation (IMD), long-term health conditions, and referral source.[23] This analysis helps highlight where the programme may be having the greatest impact and for whom. This

[23] A disaggregation by ethnicity would have been included, however the sample is ~90% white, therefore variation is too low to accurately estimate varying effects.

could inform targeted approaches in future delivery. Disaggregations are reported in Table 7 below.

By gender, the results show virtually no difference: both men and women report equally higher life satisfaction scores following participation. Due to small sample sizes at the youngest and oldest age bands, we collapsed age into two categories. Participants aged 16–64 show a stronger association with higher life satisfaction (+1.41) than those aged 65 and over (+1.13). This pattern may reflect differences in baseline wellbeing or the ways different age groups engage with the programme.

Disaggregating by deprivation level reveals that individuals living in more deprived areas (IMD decile 1-3) tend to report the largest differences in life satisfaction, whilst those from medium and low deprivation areas also show positive associations, though the magnitude is marginally smaller. This suggests the programme is having its greatest relative impact where baseline wellbeing tends to be lower, supporting its potential role in reducing geographic inequalities in health and wellbeing.

Health status also plays a role: those with a long-term health condition report nearly double the difference in life satisfaction (+1.46) compared to those without (+0.75). This highlights the programme's particular value for people managing ongoing physical or mental health challenges - a key priority group for the intervention.

Finally, disaggregation by referral source shows variation in life satisfaction scores, with self-referrals associated with the highest levels (+1.74), likely reflecting higher baseline motivation or a stronger personal commitment to change. Those referred by GPs, nurses, and physiotherapists all show meaningful differences (+1.21 to +1.34), while those from rehabilitation departments report slightly smaller differences (+1.01), possibly due to more complex health needs or differing expectations.

In sum, these disaggregated findings reinforce the broad and inclusive value of the programme, with particularly strong impacts for those with long term physical or mental health conditions and people living in more deprived communities - the very groups who may need these wellbeing gains the most.

Table 7: Life satisfaction regression coefficients disaggregated by participant characteristics

Treatment variable sub-category	Life Satisfaction Coefficient
All current or past participants	1.32***
By gender	
Male	1.32***
Female	1.31***
By age, in 2 categories	
16-64	1.41***
65+	1.13***
By local deprivation (IMD decile 2024), in 3 categories	
High deprivation (1-3)	1.42***
Medium deprivation (4-7)	1.36***
Low deprivation (8-10)	1.18***
By long-term health condition	
No	0.75***
Yes	1.46***
By referral source	
GP	1.21***
Physiotherapist	1.21***
Self-referral	1.74***
Nurse	1.34***
Rehabilitation department	1.01***

Stars indicate statistical significance levels: * p<10%, ** p<5%, *** p<1%.

4.3 Sensitivity checks

This section presents results from a series of sensitivity checks designed to test the stability of the main findings under alternative modelling assumptions. We explored how the estimated

effects change when taking into account self-reported general health and fixed effects (i.e. adjusting for individual characteristics that don't change over time, such as personality or unmeasurable background factors), and focusing solely on the wave 2 sample where more granular participation data are available. Tables 8 and 9 summarise the coefficients across these different specifications.

Table 8: Regression coefficients from models under various sensitivity checks

Outcome variable (response scale)	Coefficient		
	Main model	+ self-reported general health	+ fixed effects
Personal Wellbeing			
Life Satisfaction (0-10)	1.32***	1.02***	1.17**
Happiness (0-10)	1.11***	0.84***	0.70
Anxiety (0-10)	-0.44***	-0.31***	-0.25
Worthwhileness (0-10)	1.09***	0.82***	0.60
Mental Wellbeing			
SWEMBWBS (7-35)	1.46***	0.90***	1.39
Physical Health			
Health-related QoL (EQ-5D Index, -0.285-1 scale)	0.08***	N/A[24]	-0.00
GP visits in the past month	-0.22***	N/A	-0.45
Emergency service use in the past month	-0.04	N/A	0.06
Community Cohesion			
Loneliness (1 - never to 5-often/always)	-0.21***	-0.14***	-0.27**
Belonging (1 - strongly disagree to 5 - strongly agree)	0.11***	0.06	0.22
Trust (1 - strongly disagree to 5 - strongly agree)	0.10***	0.06	0.12
Individual Development			
Self-efficacy (1 - strongly disagree to 5 - strongly agree)	0.21***	0.10**	0.01

[24] We do not include self-reported health as a control in the regressions on health related outcomes.

Outcome variable (response scale)	Coefficient		
Physical Activity (PA) Classifications	Marginal effect		
Inactive - fewer than 30 mins of PA a week	-0.21***	-0.17***	N/A[25]
Fairly Active - 30-149 mins of PA a week	-0.01	-0.01**	N/A
Active - at least 150 mins of PA a week	0.22***	0.18***	N/A

Stars indicate statistical significance levels: * p<10%, ** p<5%, *** p<1%.

4.3.1 Self-reported health as a control

In our main specification, we excluded self-reported general health from the control variables to avoid over-controlling for a potential outcome of the intervention (as shown in column 1 of Table 8). However, as a sensitivity test, we re-ran the regressions with self-reported general health included to see whether it "explained away" the treatment effects.

As shown in column (2) of Table 8, the inclusion of health reduces the magnitude of most coefficients but does not eliminate them. For example, the effect of participation on life satisfaction drops from 1.32 to 1.02, and happiness falls from 1.11 to 0.84 (both still significant at the 1% level). Anxiety and worthwhileness also remain significant, but at a smaller magnitude. This suggests that while part of the observed difference in wellbeing may be mediated through better health, there remains a direct association between participation and wellbeing outcomes that is not explained by self-reported health.

In the domain of physical activity, the effects also attenuate slightly but remain statistically significant, particularly for the likelihood of being "active" (+0.18) and less likely to be "inactive" (-0.17). For self-efficacy, the coefficient drops from 0.21 to 0.10 (significant at the 5% level), indicating that some of the observed association may operate through better perceived or actual health. Overall, including general health as a control reduces but does not undermine the core findings, strengthening our confidence that the effects are not solely due to underlying differences in baseline health.

4.3.2 Fixed effects regression

To further assess the robustness and mitigate the risk of unobserved individual-level bias, we estimated fixed effects (FE) models. This specification controls for all time-invariant individual

[25] The fixed effects ordered logit model cannot be accurately estimated in the software package used (i.e. STATA).

characteristics, whether observed or not, as well as wave-level fixed effects to account for shared shocks across time.

Unlike the main cross-sectional regressions, the FE model estimates the within-person change in outcomes associated with programme participation, providing a stricter test of causal inference. However, the trade-off is a significantly smaller analytic sample (290 individuals), limited to those with data from both waves and non-constant outcome values. This reduction in sample size lowers statistical power, which helps explain why several coefficients, while directionally similar, lose statistical significance in the FE model, particularly for some wellbeing and healthcare use outcomes.

The results from the FE models are shown in column (3). While the statistical power is somewhat reduced, the direction of most coefficients remains consistent. Life satisfaction, for example, continues to show a significant association of 1.17 points (significant at the 5% level), and loneliness is still significantly lower (-0.27). Other wellbeing outcomes retain positive associations, although lose statistical significance, likely due to more conservative estimates and reduced sample efficiency.

Notably, for physical health outcomes, the FE models show no statistically significant results. The earlier association between the programme and higher health-related quality of life (measured by the EQ-5D index) disappears once we account for unchanging individual characteristics that we can't control for, as they aren't directly observed or measured, such as long-standing health behaviours, attitudes, or personal motivation. This suggests that the cross-sectional association, even after adjusting for observed characteristics, may still have been influenced by who chooses to take part in the programme. Similarly, the reduction in GP service use seen previously is not observed in the FE models. This points to the possibility that the pooled OLS regressions reflected selection effects or other unobserved personal factors, rather than a direct impact of the intervention itself.

However, the consistency of positive and significant results for key wellbeing and community cohesion outcomes provides reassurance that the programme's core benefits are not simply artefacts of local context, timing, or stable individual differences. At the same time, the loss of significance in several outcomes, particularly those related to physical health, underlines the importance of interpreting these results with caution. These findings highlight both the strength and limits of the available data, suggesting that while the programme appears to

deliver robust wellbeing benefits, further longitudinal analysis with larger samples will be critical to strengthen causal claims in other domains.

4.3.3 Wave 2-only regression

As a final sensitivity check, we re-estimated the core outcomes using only data from wave 2, the only survey wave where we could distinguish between current and past participants. This allowed us to test whether results are consistent when isolating outcomes at a single time point and whether the wellbeing benefits differ across current participants and completers.

Table 9 presents the results. Both current and past participants show statistically significant associations with better outcomes across nearly all domains, relative to non-participants. While the differences are modest, past participants tend to report slightly stronger effects, particularly for life satisfaction (+1.47 vs. +1.27), happiness (+1.22 vs. +0.93), and self-efficacy (+0.32 vs. +0.21). Mental wellbeing scores (SWEMWBS) are also marginally higher for past participants (+1.61 vs. +1.36), suggesting that the benefits of the programme may persist or even grow following completion. Reductions in loneliness are also more pronounced among past participants (-0.23) than current ones (-0.14), potentially indicating that social or emotional improvements consolidate over time. Physical activity patterns and GP visit reductions remain consistent across both groups. Unlike the main pooled OLS results, the effects on belonging and trust in the wave 2-only regression are smaller and statistically insignificant at the 10% level. However, the coefficients remain positive and directionally consistent, with magnitudes similar to the full-sample results. This suggests that the lack of statistical significance may be due to reduced statistical power in the smaller wave 2-only sample, rather than a true absence of effect.

Taken together, these results strengthen confidence in the programme's sustained effects, showing that key benefits are evident both during and after participation and are not limited to a particular moment in the programme lifecycle.

Table 9: Regression coefficients split by participation type

Outcome variable (response scale)	Coefficient		
	Current **or** past participant	Current participant	Past participant
Personal Wellbeing			
Life Satisfaction (0-10)	1.40***	1.27***	1.47***
Happiness (0-10)	1.12***	0.93***	1.22***
Anxiety (0-10)	-0.39**	-0.45**	-0.36*
Worthwhileness (0-10)	1.12***	0.98***	1.19***
Mental Wellbeing			
SWEMBWBS (7-35)	1.52***	1.36***	1.61***
Physical Health			
Health-related QoL (EQ-5D Index, -0.285-1 scale)	0.09***	0.08***	0.10***
GP visits in the past month	-0.28***	-0.25**	-0.30***
Emergency service use in the past month	-0.04	-0.02	-0.05
Community Cohesion			
Loneliness (1 - never to 5-often/always)	-0.20***	-0.14	-0.23***
Belonging (1 - strongly disagree to 5 - strongly agree)	0.08	0.06	0.08
Trust (1 - strongly disagree to 5 - strongly agree)	0.06	0.01	0.09
Individual Development			
Self-efficacy (1 - strongly disagree to 5 - strongly agree)	0.28***	0.21***	0.32***
Physical Activity (PA) Classifications	Marginal effect		
Inactive - fewer than 30 mins of PA a week	-0.21***	-0.22***	-0.20***
Fairly Active - 30-149 mins of PA a week	0.00	0.00	0.00
Active - at least 150 mins of PA a week	0.21***	0.22***	0.20***

Stars indicate statistical significance levels: * p<10%, ** p<5%, *** p<1%.

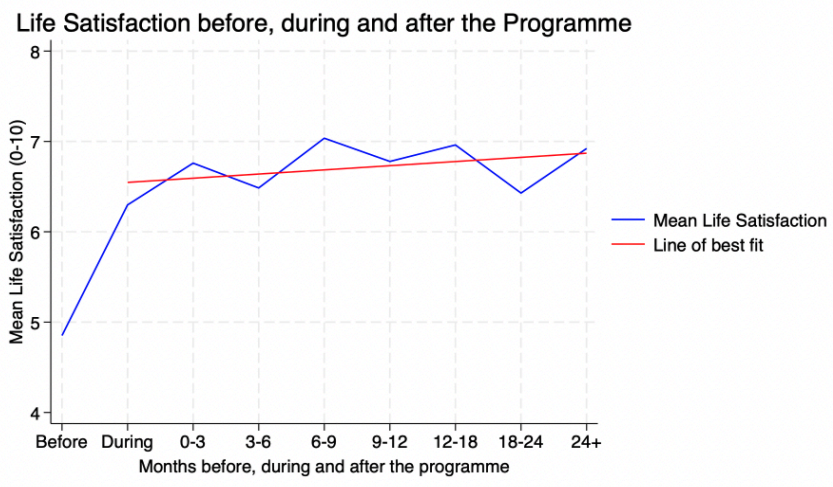
4.4 Wellbeing value

To estimate the social value of GM Active's exercise referral schemes, we use a wellbeing valuation approach based on Wellbeing-Adjusted Life Years (WELLBYs), defined as a one-point change in life satisfaction (on a scale from 0 to 10) experienced by one person for one year. This metric allows us to translate subjective wellbeing gains into a common monetary unit that can be compared across programmes, sectors, or interventions.

The coefficient of programme participation on life satisfaction - 1.32 points - forms the basis of this valuation. This estimate, derived from our main regression model (Table 5), reflects the average difference in life satisfaction between participants and non-participants. While we cannot make strong causal claims due to the observational nature of the data, the model includes adjustments for a wide range of demographic and socioeconomic characteristics known to influence wellbeing. As such, the estimate represents a relatively robust comparison between those who have participated in the programme and those who have not yet begun.

In order to claim a per person yearly WELLBY of 1.32, we must make the assumption that the life satisfaction impact lasts for one year. In support of this assumption, the following chart plots the mean life satisfaction before, during and after the programme, indicating that the life satisfaction increase sustains well past 1 year.[26]

Figure 1: Chart plotting life satisfaction before, during and after the programme



[26] "Before" refers to the waitlist group, "during" refers to those on the programme/reporting a duration between 0-12 weeks, and "after" refers to those who have completed/ reporting a duration beyond 12 weeks, in 3-month increments (e.g. "0-3" = up to 3 months after completion).

To assign a monetary value to this life satisfaction gain, we apply the recommended valuation rate for a WELLBY from HM Treasury's Green Book supplementary guidance on wellbeing (2021): £13,000 in 2019 prices. Adjusting this for inflation to 2025[27] gives a value of £16,500 per WELLBY. When multiplied by the estimated effect of programme participation (1.32 WELLBYs), this results in a wellbeing value of approximately £21,800 per participant, per year.

This estimated per-person wellbeing value of £21,800 is high, but not implausible, when considered in the context of other evidence on the value of physical activity and preventative health and wellbeing programmes - especially those targeting populations with lower baseline wellbeing.[28]

For example, in our work for Sport England, we found that the wellbeing impact of physical activity is considerably higher for those who are inactive or living with disabilities and long-term health issues[29]. Being physically active generates an annual wellbeing value of £2,500 (in 2023 prices) per person, whilst for those who are disabled or living with long-term health conditions, being physically active can generate a wellbeing value of up to £5,100 per person per year. Programmes targeted at young people in disadvantaged areas, such as Active Row, have shown WELLBY values of up to £6,000 per participant per year, reflecting the greater potential for wellbeing uplift among those starting from a lower baseline[30]. And previous studies valuing similar interventions to GM Active's, like the Essex County's Prevention and Enablement Model, which aims to improve the health and wellbeing of older adults through targeted community support, is associated with a wellbeing impact of 1.71 life satisfaction points, corresponding to a value of £22,230 per person per year (in 2019 prices)[31]. This suggests that the GM Active programme's value is in line with other successful programmes focused on improving health and wellbeing in populations with varying levels of initial disadvantage.

It is also important to note that this valuation reflects only the subjective wellbeing benefit, as captured through life satisfaction. It does not account for the potential downstream fiscal benefits, such as reduced demand on NHS services, improved productivity, or lower reliance

[27] https://assets.publishing.service.gov.uk/media/60fa9169d3bf7f0448719daf/Wellbeing_guidance_for_appraisal_-_supplementary_Green_Book_guidance.pdf , page 57

[28] There is strong evidence that people living in more deprived areas, those with disabilities or long-term health conditions, and individuals who are physically inactive tend to report lower life satisfaction and mental wellbeing compared to the general population (e.g., What Works Centre for Wellbeing 2017; ONS, 2023; Sport England Active Lives Survey, 2025).

[29] https://www.sportengland.org/news-and-inspiration/sport-and-physical-activity-generates-over-100-billion-social-value

[30] https://heyzine.com/flip-book/907baf2538.html#page/1

[31] https://www.sportforconfidence.com/our-services/prevention-enablement-model/

on employment and welfare support, that might reasonably follow from improved physical and mental health, greater social connectedness, and increased activity levels. These additional societal and economic impacts, which were evidenced earlier in our analysis through reduced GP visits and improved physical health and mental wellbeing, are likely to enhance the overall case for investment even further. The following sections explore some of these health-related outcomes and potential cost savings in more detail.

4.5 Health value

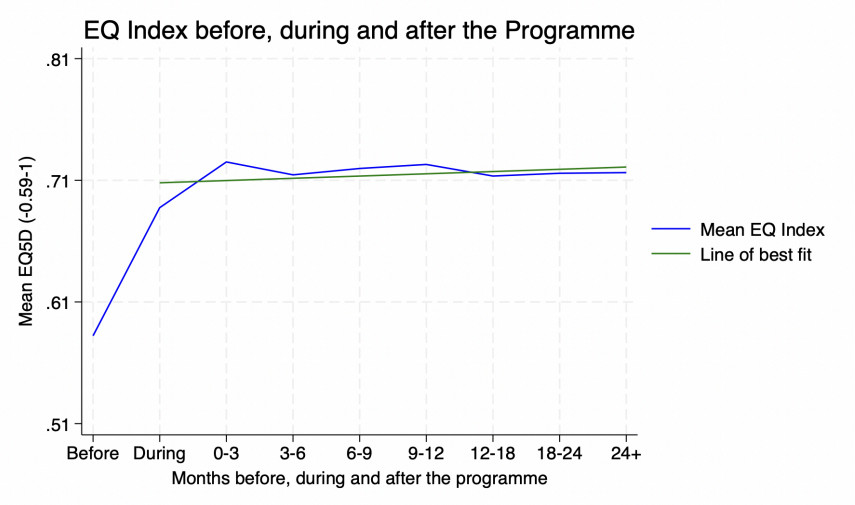
As a complement to the wellbeing-based valuation presented in the previous section, we also apply a health economic approach to estimate the health value of GM Active's exercise referral programme using Quality-Adjusted Life Years (QALYs). This provides an estimate of the programme's impact on health-related quality of life, using a metric widely adopted by the National Institute for Health and Care Excellence (NICE) to evaluate the cost-effectiveness of health interventions.

A QALY represents one year of life lived in perfect health, and it incorporates both the length and quality of life. Health-related quality of life can be measured using the EQ-5D-5L questionnaire, a validated instrument included in our survey. The resulting index was constructed using the English Devlin value set[32] and EuroQol's published STATA syntax code[33]. This index ranges from -0.285 (worst health state) to 1 (perfect health). In our analysis, participants in the GM Active referral programme experienced an average increase in EQ-5D index score of 0.08 (Table 5), suggesting an average QALY gain of 0.08 per person per year.

We need to once more make the assumption that this impact lasts for one year. The following chart plots the mean EQ-5D index score before, during and after the programme, indicating that the health-related quality of life increase sustains well past 1 year.

[32] https://onlinelibrary.wiley.com/doi/10.1002/hec.3564
[33] https://euroqol.org/wp-content/uploads/2024/01/ENG_value-set_STATA.txt

Figure 2: Chart plotting EQ Index before, during and after the programme



To assign a monetary value, we follow the Department of Health and Social Care's valuation of a QALY at £70,000,[34] reflecting the estimated total economic value to society, beyond just NHS savings. This valuation yields an estimated health benefit of £5,600 per participant per year (0.08 QALYs × £70,000). While our data suggest the health benefits of participation are likely to persist beyond one year - potentially up to two years or more - we adopt a conservative approach and report a one-year valuation only. Further longitudinal data would be needed to robustly quantify longer-term effects.

This QALY gain can also be interpreted in relation to NICE's cost-effectiveness thresholds, which guide whether a health intervention represents good value for money. NICE typically considers interventions costing £20,000 to £30,000 per QALY gained to be cost-effective[35]. Based on our estimated 0.08 QALYs gained per participant, the GM Active programme would meet NICE's £20,000 threshold if the cost per participant were less than £1,600 per year (0.08 × £20,000).

If we apply a more conservative NHS production cost estimate of £15,000 per QALY, the programme would still represent good value if delivered at a cost of under £1,200 per participant, per year. Whilst we acknowledge more work needs to be done to understand the

[34] https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2020
[35] https://www.nice.org.uk/news/blogs/should-nice-s-cost-effectiveness-thresholds-change-

full costs of the programmes and conduct a thorough cost utility analysis[36], early estimations place the cost of delivering the programme at roughly £240 per person per year (see appendix 2). This is well below NICE's cost-effectiveness threshold, suggesting the programme represents strong value for money from a health system perspective.

Beyond the measured improvement in quality of life, there is evidence of reduced health service use, which carries real economic implications. Participants in the GM Active programme reported 19% fewer GP consultations in the past month compared to waitlisted individuals. Given that a single GP appointment costs the NHS approximately £40,[37] and assuming an average of 1.2 visits per person per month, as in the control group, this reduction could represent a potential annual saving of around £110 per person. Extrapolated across thousands of participants, these savings could add up significantly, especially when combined with broader gains in health, wellbeing, and reduced long-term care needs. While these are first-order estimates, they point to an important additional mechanism through which exercise referral schemes relieve pressure on primary care services.

4.6 Market value - a sustainable business model for leisure centres

Finally, we look at a key and more traditional measure of economics - market value, defined here as the potential for sustained consumer engagement and economic return following initial public investment. Once again, we see encouraging results.

In the wave 2 survey, we asked participants whether they intend to remain physically active over the next six months, and whether as a member of their current centre or elsewhere. 80% of participants reported an intention to remain active, with nearly two-thirds (66%) intending to do so as customers of their GM Active leisure centre. This suggests that many participants transition from being publicly funded referrals to self-sustaining members, contributing directly to the financial viability of the service providers.

This behavioural shift aligns with the findings above that the wellbeing and health benefits of participation can persist well beyond the initial 12-week intervention, lasting many years for some participants. These enduring benefits likely point to ongoing engagement, as individuals continue to invest in physical activity due to their experience of it improving their wellbeing and health-related quality of life.

[36] https://www.gov.uk/guidance/cost-utility-analysis-health-economic-studies
[37] https://kar.kent.ac.uk/109563/ p.73

These findings point to a double benefit model: public funding initiates health improvements in inactive individuals, while also catalysing long-term lifestyle change that supports a more sustainable business model for leisure centres. This dynamic exemplifies how preventative health investment can yield both social value and market value, aligning public health goals with economic sustainability.

5. Discussion

This report has presented a detailed evaluation of GM Active's exercise referral programmes, using data from over 3,000 participants across two waves and applying robust analytical techniques to estimate the impact on health and wellbeing. By combining regression analysis with both WELLBY and QALY valuation methods, the analysis offers a comprehensive assessment of the social and health value generated by the intervention. This final section summarises the key conclusions, reflects on the limitations, considers the broader implications for policy and practice, and identifies priorities for further research.

5.1 Conclusions

The evidence strongly suggests that GM Active's physical activity referral programmes are associated with meaningful improvements in personal wellbeing, physical health, and other quality of life outcomes, as well as reductions in health service usage. Participants in the programme reported, on average, a 1.32-point increase in life satisfaction compared to non-participants (waitlisters), and more frequent, longer-lasting, and structured forms of participation in a GM Active programme augment these changes. These statistically significant wellbeing effects are not only robust when controlling for a wide range of sociodemographics and to various sensitivity tests, but also translate into substantial social value when assessed using the HM Treasury endorsed WELLBY-based valuation approach.

The estimated value of £21,800 per person per year places the programme's wellbeing impact at the higher end of the spectrum compared to other evaluated public health and sport-based interventions, particularly those targeting disadvantaged groups. Similarly, using the EQ-5D index as a health-related quality of life measure, the average 0.08-point increase equates to a QALY value of £5,600 per person per year.

Importantly, the wellbeing effects are most pronounced among those with typically lower initial wellbeing: those living with long-term health conditions or from more deprived areas. The greater magnitude of benefits observed among these participants underscores the potential for such programmes to help reduce health inequalities and contribute to levelling up wellbeing outcomes across Greater Manchester.

5.2 Limitations of the study

Despite the strength of the findings, several limitations must be acknowledged. Most crucially, the analysis is not based on a randomised controlled trial, meaning causal inference is limited. While the regression models control for a wide range of confounders, some unobservable differences - especially time-variant factors not picked up in the fixed-effects regression - between participants and non-participants may still bias the results.

The use of a waitlist as a comparison group and a variable for duration of participation strengthens the analysis by approximating a dose-response relationship, which, under reasonable assumptions, brings this evaluation in line with Level 3 of the Nesta Standards of Evidence.[38] This is consistent with a quasi-experimental design that meets the expectations of the UK Government's Green and Magenta Books for robust evaluation practice. However, it still falls short of causal certainty. Establishing definitive causality would require a more rigorous and significantly more expensive research design, such as a randomised controlled trial or a longitudinal study following individuals over time. While such designs are often considered the gold standard, they can be expensive, impractical and less likely to reflect real-world service delivery contexts. Therefore, this study strikes a balance between methodological rigour and feasibility.

Another limitation relates to attrition, which can range from 25% to 55% depending on the programme.[39] This dropout can introduce bias if the individuals who disengage differ systematically from those who remain, for instance, in terms of motivation, wellbeing, or circumstances. Although fixed-effects models mitigate bias from time-invariant characteristics, they cannot fully account for attrition-related differences, particularly if disengagement is linked to unmeasured, time-varying factors.

Selection bias is also a concern, both in terms of entry into the programme and participation in the survey. Programmes that accept self-referrals or rely on partner referrals may attract individuals who are more motivated or who already have higher baseline wellbeing, health literacy, or engagement with support services. Similarly, individuals who choose to complete follow-up surveys may not be representative of all participants, particularly if those who

[38] https://media.nesta.org.uk/documents/standards_of_evidence.pdf

[39] This information comes from a survey asked to GM Active partner organisations, reporting on the delivery and characteristics of their respective scheme.

experienced less positive outcomes were less likely to respond. These biases could lead to overestimation of average effects.

A further source of potential bias arises from the self-reported nature of key outcomes, such as GP visits and physical activity levels. These measures are inherently vulnerable to recall error, where participants may misremember or misreport the frequency of healthcare use or the amount of physical activity undertaken. Additionally, social desirability bias may lead individuals to overstate positive behaviours or underreport negative experiences, consciously or unconsciously presenting themselves in a more favourable light.

Lastly, the monetary valuations of wellbeing and health benefits rely on standard assumptions regarding the persistence of effects and the value of WELLBYs and QALYs. In this report, we assume that the observed life satisfaction and health improvements last for at least one year. While some evidence from past participant information supports this assumption that benefits endure for 24+ months, longitudinal tracking would be required to confirm the duration and trajectory of impacts over time.

5.3 Implications

The results of this evaluation have important implications for policy, practice, and future investment in preventative health and wellbeing interventions. The observed wellbeing and health benefits strongly support the case for continued and potentially expanded funding for GM Active's exercise referral schemes. These programmes are not only improving individual outcomes but are also likely reducing demand on NHS services and contributing to broader public health goals.

Importantly, 80% of participants reported an intention to remain physically active over the next six months, with nearly two-thirds (66%) indicating they plan to continue as customers of their GM Active centre. This suggests a sustainable model where initial patients become ongoing customers of the leisure centres delivering the service.

The findings also highlight the value of physical activity as a lifestyle-based intervention that may be preferable to pharmaceutical treatments for some individuals. Rather than relying on medication, participants are engaging in an approach that can be empowering, non-invasive, and conducive to long-term habit change. The fact that these interventions are delivered in local community settings - leisure centres that are more familiar, accessible, and less clinical

than hospitals - also enhances their convenience, especially for those who might be deterred by more traditional healthcare settings.

Coupled with the finding that more structured and sustained engagement yields the greatest benefits, this suggests that the programme has the potential for sustained impact, with a high likelihood of continued physical activity even after completion. In light of these findings, policymakers should consider how programme designs can further support long-term adherence, through mechanisms such as follow-up coaching, peer support, or flexible re-engagement options.

The study also found particularly strong effects among participants with lower baseline wellbeing, indicating that targeted interventions for these groups may offer the highest returns on investment in terms of equity and effectiveness. This highlights the importance of prioritising resources and efforts for those who stand to benefit most, such as individuals living with long-term health conditions or those from more deprived areas.

While we did not assess the value-for-money of the programme directly due to the lack of cost data, the dual application of wellbeing and health economic valuation methods offers a valuable framework for evaluating other complex interventions that intersect health, community, and social policy domains. This integrated approach provides a richer understanding of impact than traditional metrics alone, offering a model for future evaluations.

5.4 Areas for further study

While this study provides strong evidence of benefit, further research is needed to deepen and validate these findings. Longer-term tracking of participants would help establish the durability of wellbeing and health improvements beyond the initial year. A deeper understanding of whether the gains observed continue, plateau, or decline after programme completion is essential for assessing the long-term return on investment.

Future studies should also explore the feasibility of implementing more rigorous quasi-experimental or randomised designs to strengthen causal inference. Methods such as propensity score matching, regression discontinuity, or difference-in-differences could offer a more practical solution to the resource and ethical constraints of a randomised controlled trial.

In addition, the collection of detailed cost data, both in terms of programme delivery and participant engagement (attrition and drop out rates), would enable a more comprehensive cost-benefit analysis, particularly if combined with administrative data on healthcare usage. This would allow for a clearer assessment of the programme's value for money and understanding whether the health and wellbeing value, in addition to the healthcare cost relief, outweigh the programme costs would be crucial for justifying its continued funding and expansion.

Given the relatively high attrition rate in some programmes, further work is also needed to explore how engagement can be optimised across different population groups. This includes understanding the barriers to sustained participation, such as motivation, accessibility, or perceived value of the programme, and could include testing strategies like tailored follow-up, flexible session formats, and additional support services.

Finally, the potential for linking survey data with NHS health records or other administrative datasets offers a promising avenue for validating self-reported impacts and exploring broader system-level benefits. These next steps would enhance the evidence base for exercise referral schemes and support more informed decision-making about their role in improving population health and wellbeing.

References

Devlin, N., Shah, K., Feng, Y., Mulhern, B., & van Hout, B. (2017). Valuing health-related quality of life: An EQ-5D-5L value set for England. *Value in Health*, 27(1), 7–22. https://doi.org/10.1002/hec.3564

EuroQol Research Foundation. (2024). *ENG value set STATA file*. https://euroqol.org/wp-content/uploads/2024/01/ENG_value-set_STATA.txt

EuroQol Research Foundation. (2025). *EQ-5D-5L instrument overview*. https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/

HM Treasury. (2020). *The Green Book: Appraisal and evaluation in central government*. https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2020

HM Treasury. (2021). *Wellbeing guidance for appraisal: Supplementary Green Book guidance*. https://assets.publishing.service.gov.uk/media/60fa9169d3bf7f0448719daf/Wellbeing_guidance_for_appraisal_-_supplementary_Green_Book_guidance.pdf

NICE. (2024). *Should NICE's cost-effectiveness thresholds change?* https://www.nice.org.uk/news/blogs/should-nice-s-cost-effectiveness-thresholds-change-

Office for National Statistics. (2023). *Measuring National Wellbeing: April 2022 to March 2023*. https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/measuringnationalwellbeing/april2022tomarch2023

Office for National Statistics. (2025). *Personal wellbeing survey user guide*. https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingsurveyuserguide

Jones, K., et al. (2025). *Unit Costs of Health and Social Care 2024 Manual*. University of Kent. https://kar.kent.ac.uk/109563/

Sport England. (2025). *Active Lives Adult Survey, Nov 2023–24*. https://sportengland-production-files.s3.eu-west-2.amazonaws.com/s3fs-public/2025-04/ActiveLivesAdult-Nov23-24_V9-23-04-25-10-03-03-02.pdf

Sport England. (2024). *Sport and physical activity generates over £100 billion in social value*. https://www.sportengland.org/news-and-inspiration/sport-and-physical-activity-generates-over-100-billion-social-value

State of Life & Sport for Confidence. (2023). *Prevention and Enablement Model Evaluation Report*. https://www.sportforconfidence.com/our-services/prevention-enablement-model/

State of Life & London Youth Rowing. (2023). *Active Row social impact report*. https://heyzine.com/flip-book/907baf2538.html#page/1

UK Government. (2020). *Cost-utility analysis: Health economic studies guidance*. https://www.gov.uk/guidance/cost-utility-analysis-health-economic-studies

Warwick Medical School. (2020). *Why use WEMWBS?* https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/strengths/

Warwick Medical School. (2020). *WEMWBS: 14-item vs 7-item scale* https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/wemwbsvsswemwbs/

What Works Wellbeing. (2017). *Drivers of wellbeing inequality*. https://whatworkswellbeing.org/resources/drivers-of-wellbeing-inequality/

Nesta. (2013). *Standards of evidence: An approach that balances the need for evidence with innovation* https://media.nesta.org.uk/documents/standards_of_evidence.pdf

Appendix 1: List of Outcomes

Personal Wellbeing

We used the ONS-4 personal wellbeing questions developed by the Office for National Statistics (ONS), a widely used national measure of subjective wellbeing in the UK. Respondents answered the following questions on a 0-10 scale (0 = “not at all”, 10 = “completely”):

- Overall, how satisfied are you with your life nowadays?
- Overall, how happy did you feel yesterday?
- Overall, how anxious did you feel yesterday?
- Overall, to what extent do you feel the things you do in your life are worthwhile?

Mental Wellbeing

Mental wellbeing was assessed using the 7-item Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS). Responses are provided on a 5-point scale (1 = “none of the time”, 5 = “all of the time”), with total scores transformed according to established guidelines to provide a final score ranging from 7 to 35.

SWEMWBS statements:

1. I've been feeling optimistic about the future
2. I've been feeling useful
3. I've been feeling relaxed
4. I've been dealing with problems well
5. I've been thinking clearly
6. I've been feeling close to other people
7. I've been able to make up my own mind about things

Physical Health

Health-related quality of life was measured using the EQ-5D-5L instrument, developed by the EuroQol Group and recommended by NICE. It includes five dimensions:

- Mobility
- Self-care
- Usual activities
- Pain/discomfort
- Anxiety/depression

Each dimension has five response levels (no problems to extreme problems), and responses are converted into a single index value using the English Devlin value set.

Participants also reported their use of:

- **GP services**: “How many times in the last month have you consulted with your GP?”
- **Emergency services**: “How many times in the last month have you accessed urgent or emergency care services (A&E or ambulance)?”

Physical Activity Levels

Physical activity was measured using Sport England’s Short Active Lives Survey (SALS). Respondents reported the amount of moderate and vigorous activity they engaged in during the previous week across all activities (e.g. walking, sports, exercise classes).

This was used to classify participants as:

- **Inactive**: Fewer than 30 minutes per week
- **Fairly active**: 30 to 149 minutes per week
- **Active**: 150+ minutes per week

These categories align with UK Chief Medical Officer guidelines and Sport England classification standards.

Community Cohesion and Individual Development

To assess social connectedness and cohesion, we asked participants to rate their agreement or frequency of the following:

- **Loneliness**: “How often do you feel lonely?” (1 = “never”, 5 = “often/always”)
- **Sense of belonging**: “I feel like I belong to my local area” (1 = “strongly disagree”, 5 = “strongly agree”)

- **Trust in others**: “Generally speaking, most people can be trusted” (1 = “strongly disagree”, 5 = “strongly agree”)
- **Self-efficacy**: “I can achieve most of the goals I set myself” (1 = “strongly disagree”, 5 = “strongly agree”)

Appendix 2: Cost-effectiveness case study

To get a rough idea of the cost-effectiveness of GM Active's physical activity referral schemes, we collated indicative programme cost data from several of the partner organisations. While not derived from a full economic costing study, these figures offer a useful benchmark for initial value-for-money assessment.

Based on available submissions, the **average cost per participant is approximately £240**, with a range from £140 to £290 depending on programme structure, delivery model, and inclusion of additional services (e.g. falls prevention).

For example, in 2023-24:

- One organisation reported total programme delivery costs of £205,000 (excluding falls prevention), with 1,900 referrals, an 85% uptake rate, and a 45% completion rate. Adjusting for these rates[40], the cost per completer is approximately **£280**.
- Another reported costs of £265,000 for 1,166 completers, resulting in an estimated cost per participant of **£230**.
- A third reported the costs of their exercise on referral programme to be £250,000 for 1,500 referrals, but with an uptake rate of 75% and a completion rate of 65%. Adjusting for these rates, the cost per completed participant is approximately **£260**.

These figures compare favourably to NICE's cost-effectiveness thresholds, which typically consider interventions good value for money if they deliver a QALY (Quality-Adjusted Life Year) for less than £20,000 to £30,000. Based on this evaluation, GM Active's average health-related quality of life improvement is 0.08 QALYs. To meet the lower NICE threshold of £20,000 per QALY, the programme would need to cost less than £1,600 per participant per year. At an estimated average cost of £240 per person, GM Active's schemes appear to offer exceptional value, potentially making them **at least 6 times more cost-effective than the NICE / NHS benchmark** in delivering outcomes.

While these are early estimates, they strongly support the case for a more detailed economic evaluation and suggest that GM Active's community-based approach to prevention is not only impactful, but highly cost-effective.

[40] To estimate the number of successful completions, the number of referrals are first multiplied by the uptake rate, and this number is then multiplied by the completion rate.