

# Weight loss interventions on health-related quality of life in those with moderate to severe obesity: Findings from an individual patient data meta-analysis of randomized trials

John Buckell<sup>1</sup>  | Xue W. Mei<sup>2</sup> | Philip Clarke<sup>1</sup> | Paul Aveyard<sup>2</sup>  | Susan A. Jebb<sup>2</sup>

<sup>1</sup>Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>2</sup>Nuffield Department of Primary Health Care Sciences, University of Oxford, Oxford, UK

## Correspondence

John Buckell, University of Oxford, Richard Doll Building, Old Road Campus, Oxford OX3 7LF, UK.  
Email: john.buckell@ndph.ox.ac.uk

## Funding information

NIHR Oxford Biomedical Research Centre and Applied Research Collaboration

## Summary

The relationship between BMI and health-related quality of life (HRQoL) critically affects regulatory approval of interventions for weight loss, but evidence of the association is inconsistent. A higher standard of evidence than that available was sought with an IPD meta-analysis of 10,884 people enrolled in five randomized controlled trials of intentional weight loss interventions. Cross-sectional and longitudinal associations of BMI and HRQoL were estimated in mixed effects models specifying a latent variable for HRQoL. Spline regressions captured nonlinear associations across the range of BMI. In cross-sectional spline regressions, BMI was not associated with HRQoL for people with a BMI < 30 kg/m<sup>2</sup> but was for those with a higher BMI. In longitudinal spline regressions, decreases in BMI were positively associated with HRQoL for people with a BMI ≥ 25 kg/m<sup>2</sup>. The impact of change in BMI was larger for people with higher BMIs than for those with BMIs under 30 kg/m<sup>2</sup>. Lower BMI and decreases in BMI were related to higher HRQoL for people living with obesity but not in the population without excess weight. HRQoL gains from weight loss are greater for more severe obesity. Commissioners should use these estimates for future decision making.

## KEYWORDS

commissioning, health-related quality of life, IPD meta-analysis, weight loss

## 1 | INTRODUCTION

Excess weight is associated with substantial morbidity and premature mortality.<sup>1</sup> Excess weight reduces health-related quality of life (HRQoL)—that is, is an individual's or a group's perceived physical and mental health over time—through effects on incidence of weight-related disease but may also do so directly, compromising daily functioning through difficulty in performing everyday tasks and reducing mental well-being.<sup>2,3</sup> Interventions aiming at

weight loss have been shown to reduce the incidence of diabetes and premature mortality,<sup>4,5</sup> but the effect on HRQoL is uncertain.

Systematic reviews and meta-analyses indicate several issues with extant evidence on the relationship between BMI and HRQoL.<sup>2</sup> Within-study results are not always robust across different indicators of HRQoL.<sup>6</sup> There is some, though limited, evidence that the relationship between BMI and HRQoL is not linear.<sup>7,8</sup> Studies in specific patient groups may not apply to the general population of people with overweight. There are concerns that weight loss attempts may impair mental health.<sup>9</sup>

Paul Aveyard and Susan A. Jebb are of equal contribution.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Obesity Reviews* published by John Wiley & Sons Ltd on behalf of World Obesity Federation.

Across the world, health economic appraisals of behavioral interventions, pharmacotherapeutic interventions, and bariatric surgery have estimated the impact of the intervention on quality of life by assuming that cross-sectional differences between people of different BMIs represent the effect of *changing* weight on quality of life.<sup>10–15</sup> Studies have used estimates that differ markedly. Different values could alter decisions on whether to commission these interventions. In the United Kingdom, recent appraisal of liraglutide by the National Institute for Health and Care Excellence (NICE) used values from a cross-sectional, observational study.<sup>7</sup> Here, a nonlinear function of BMI was specified. HRQoL, expressed in utility, increased with BMI up to an inflection point (around BMI = 25 kg/m<sup>2</sup>); HRQoL declined thereafter. This is in stark contrast to prior approaches in the United Kingdom and internationally that have assumed a single value across the range of BMI. Another issue is that gains in HRQoL from weight loss are often assumed to be sustained, which may not be the case, particularly after non-surgical interventions. More robust estimates of the HRQoL-BMI relationship and the sustainability over time would enable consistent commissioning decisions.

In this paper, we estimate the association between BMI and HRQoL using longitudinal data from five large RCTs of behavioral weight loss interventions in an individual participant data (IPD) meta-analysis, where weight loss can be assumed to be intentional and not arising from disease. We assess whether the association varied over the distribution of BMI and whether the strength of association was different for the physical and mental domains of HRQoL. The latter in particular is of interest given the lack of existing evidence on this issue.

## 2 | METHODS

### 2.1 | Trial identification

Studies were identified using existing systematic reviews in obesity,<sup>2,16</sup> reference mining of systematic reviews, and consultation with experts in the field.

### 2.2 | Trial eligibility

(1) Adults (18 years or older) with overweight or obesity at baseline (BMI 25 kg/m<sup>2</sup> or above) and enrolled in a behavioral weight loss trial, indicating intention to lose weight. For trials that used behavioral and pharmacotherapeutic arms, only data from the behavioral arms were analyzed. Pharmacotherapeutic trials were excluded to avoid any additional impact on individuals' quality of life, for example, through side-effects. Surgical trials were excluded due to the higher weight of patients undergoing surgery, the substantially larger reduction associated with the intervention and potential for other effects on quality of life not directly related to weight loss. (2) Mean change in weight in the intervention groups  $\geq 2$  kg. (3) The main outcomes, participants'

BMI and HRQoL, collected at least twice. (4) Trials of more than 1,000 participants were enrolled to ensure that each trial has at least some bearing on the overall relationship of interest. Details of the five trials included—Look AHEAD, TOHP, DPP, WRAP, and DioGENES—are presented below and in detail elsewhere.<sup>17–21</sup>

#### 1 Diets with High or Low Protein Content and Glycemic Index for Weight-Loss Maintenance (DioGENES)

The population studied was overweight and obese adults in EU countries that had recently lost at least 8% of their bodyweight. A total of 1209 adults participated in the trial. Interventions were ad libitum diets in a two-by-two factorial design (low/high protein, low/high glycemic index); the control was a diet based on countries' general guidance without reference to glycemic index. The outcome was weight regain. The intervention lasted for a 26-week period; quality of life was collected up to this point. The HRQoL indicator used the obesity-specific Impact of Weight on Quality Of Life (IWQOL).

#### 2 Diabetes Prevention Programme (DPP)

The population studied was adults in the United States with elevated fasting and post-load plasma glucose. A total of 3234 adults were enrolled. Interventions were a weight loss-focused lifestyle intervention and metformin (we did not include this arm—see eligibility criteria); the control was a placebo. The outcome was the incidence of diabetes. Mean follow-up was 2.8 years. The HRQoL indicator used was the SF-36.

#### 3 Look AHEAD

The population studied was adults in the United States being between 55 and 76 years of age, being diagnosed with type II diabetes, and being overweight. A total of 5145 patients participated in the trial. Interventions were intensive lifestyle interventions (increased physical activity and reduced calorie intake); the control was diabetes support and education. Outcomes were adverse effects including death and AMI, stroke, or angina. Mean follow-up was 9.6 years. HRQoL indicators included SF-36, feelings thermometer, HUI2, and HUI3.

#### 4 Trials of Hypertension Prevention (TOHP)

The population studied was adults in the United States being between 30 and 54 years of age, with diastolic blood pressure between 80 and 89. A total of 2182 patients participated in the trial. Interventions were lifestyle interventions (weight reduction, sodium reduction, and stress management); the control was non-intervention. In addition, four nutritional supplements were compared, double-blinded, to a placebo. Outcomes were changes in blood pressure. Follow-up was 18 months. The HRQoL indicator used was the general well-being scale.

## 5 Extended and standard duration weight-loss program referrals for adults in primary care (WRAP)

The population studied was adults in the United Kingdom with BMI over  $28 \text{ kg m}^{-2}$  in primary care. Interventions were 12- and 52-week weight-management programs; control was brief advice and self-help material. The primary outcome was weight after 12 months of follow-up. Mean follow-up was 1.5 years. HRQoL indicators included the 3-level EuroQol Five Dimension (EQ-5D) and the EuroQol visual analogue scale (EQ-VAS).

## 2.3 | Outcomes

The main outcomes were indicators of HRQoL. These were the SF-36 pcs (physical component summary) and SF-36 mcs (mental component summary), Feelings Thermometer, IWQOL, EQ-5D, EQVAS, and general well-being schedule. All HRQoL outcomes are composite scores except for the EQVAS and the Feelings Thermometer. A description of these indicators is provided in Appendix A. The trials varied in the indicators of HRQoL collected, with no common measurement. LookAHEAD collected four indicators, WRAP two, and the remaining three trials collected one indicator each.

## 2.4 | Statistical analyses

HRQoL was treated as a latent variable within a system of equations. Measurement equations related HRQoL, as an independent variable, to the various indicators of HRQoL across the trials, each of which was separately treated as a dependent variable. Structural equations were specified with the latent variable for HRQoL as the dependent variable. Appendix B presents the rationale for this and presents the models. Mixed effects models, based on a pre-registered analysis plan, were developed.<sup>22</sup> First, HRQoL was regressed on BMI, the group (within-individual) mean of BMI, individual characteristics, trial fixed effects, and time-from-baseline fixed effects. (A model with intervention arm fixed effects was tested, but the main coefficient of interest was consistent; the reported specification was retained to make use of data of all individuals in all arms, all of whose BMI/HRQoL varied.) A random effect for individuals was also specified. Look AHEAD had longer follow up than all of the other trials. We therefore specified time fixed effects only for time periods in which multiple trials observed participants, to avoid attributing time series HRQoL variation in Look AHEAD to other trials. Second, a spline regression was estimated. The range of BMI was partitioned into five regions, each corresponding to one of five BMI categories: healthy weight, overweight, obesity category I, obesity category II, and obesity category III (underweight was omitted as there were no underweight participants enrolled in these trials). We examined the HRQoL-BMI relationship not only at different starting levels of BMI but also at different starting HRQoL levels (see Appendix H). In cross-sectional analyses, all available observations of all individuals in all time periods

were analyzed. Next, two similar regressions were estimated, but here, HRQoL was regressed on period-to-period BMI change to examine the relationship between HRQoL and longitudinal variation in BMI.

## 2.5 | Revisiting previous commissioning decisions

Our fitted model was used to repeat the cost-effectiveness of orlistat that informed NICE commissioning<sup>12</sup> using our updated HRQoL-BMI estimate. The VAS scale was used to derive utility as per the original study. All other metrics were taken from the original paper and held constant.

# 3 | RESULTS

Levels of missing HRQoL and BMI data varied across trials (Appendix E). An initial imputation exercise was conducted on each of the data sets separately to gauge the impact of missing data on the estimated relationship of interest. In all cases, the multiply imputed regression results were very similar to those without imputation (Appendix E).

## 3.1 | Descriptive statistics

After removing observations for which either BMI or HRQoL were missing, 58,723 observations of 10,884 individuals were available. The largest trial was Look AHEAD ( $n = 4900$ ; 45%), followed by TOHP ( $n = 2172$ ; 20%), DPP ( $n = 1750$ ; 16%), WRAP ( $n = 1241$ ; 11%), and DioGENES ( $n = 821$ ; 8%).

Characteristics of individuals at baseline are presented in Table 1. A total of 8881 individuals were analyzed at baseline. For DPP, data were available at baseline, but only in an aggregated form, which is not compatible with this analysis. Mean BMI at baseline was  $33.3 \text{ kg/m}^2$ . There were slightly more females than males (53%) with a mean age of 53 years. Individuals were predominantly white (76%) with around 13% black; this is broadly in line with census data from the trials' countries.

## 3.2 | Association of HRQoL and its indicators

The associations between HRQoL and all its indicators (SF-36 pcs, SF-36 mcs, EQ-5D, etc.) were positive and statistically significantly different from zero in the linear model and spline regressions (Table 2). Physical health quality of life (SF-36 pcs) had a stronger association with HRQoL than mental health quality of life (SF-36 mcs). In the BMI change (delta) models, neither obesity-specific quality of life (IWQOL) nor mental health quality of life (SF-36 mcs) were related to HRQoL. All other HRQoL indicators were significantly related to HRQoL.

**TABLE 1** Baseline descriptive statistics

	N	Mean	s.d.
SF-36 pcs	4834	47.99	7.90
SF-36 mcs	4834	54.15	8.00
Feelings Thermometer	4834	74.16	15.84
IWQoL	693	64.74	27.79
EQ VAS	1192	70.43	18.91
EQ-5D	1192	0.79	0.25
Wellbeing	2162	80.63	11.76
BMI	8881	33.30	6.31
Age	8860	53.00	10.71
	N	Count	Percentage
Female	8881	4740	53.4%
White	8881	6746	76.0%
Black	8881	1142	12.9%
Other	8881	245	2.8%
BMI category: Underweight	8881	4	0.1%
BMI category: Normal Weight	8881	615	6.9%
BMI category: Overweight	8881	2240	25.2%
BMI category: Obese Class I	8881	2992	33.7%
BMI category: Obese Class II	8881	1769	19.9%
BMI category: Obese Class III	8881	1261	14.2%

Abbreviations: EQ VAS, EuroQol Visual Analogue Scale; EQ-5D, EuroQol Five Dimension; IWQoL, Impact of Weight on Quality Of Life; SF-36 pcs, Short Form-36 physical component summary; SF-36 mcs, Short Form-36 mental component summary; Well-being, General Well-Being Schedule. Note: Continuous variables are listed in the upper panel; binary variables are listed in the lower panel. Varying numbers of observations for HRQoL indicators reflects that different trials collected different indicators of HRQoL or that data for some individuals were missing at baseline (e.g., for age). BMI categories (lower panel) follow the standard classification. While summarized, categorical data were available to replicate baseline statistics for the DPP trial, the raw data at baseline were unavailable and so absent here.

### 3.3 | Cross-sectional associations of HRQoL and BMI

In a linear model, BMI and HRQoL were negatively associated. A one-unit lower BMI was associated with a 0.13-standard deviation unit higher HRQoL (BMI =  $-0.13$ ; 95% CI:  $-0.14$  to  $-0.12$ ) (Table 3). In the spline regression, there was no evidence of association between HRQoL and BMI when BMI was below 30 kg/m<sup>2</sup>. At higher BMI, there was a significant inverse association. BMI in the range 30–34.9 kg/m<sup>2</sup> had the smallest association with HRQoL (BMI: 30–34.9 =  $-0.52$ ; 95% CI:  $-0.66$  to  $-0.37$ ). A stronger association was observed when BMI was 35–39.9 kg/m<sup>2</sup> (BMI: 35–39.9 =  $-1.44$ ; 95% CI:  $-1.70$  to  $-1.19$ ). The strongest association was observed when BMI  $\geq 40$  kg/m<sup>2</sup> (BMI:  $\geq 40$  =  $-2.95$ ; 95% CI:  $-3.39$  to  $-2.51$ ).

Taking the estimate from the linear model, one-unit lower in BMI would relate to higher HRQoL on its various indicators (listing only statistically significant associations): SF-36 pcs 0.38; SF-36 mcs 0.21; Feelings Thermometer 0.64; IWQoL 1.35; EQ-5D 0.01; EQVAS 0.01; well-being 0.52.

### 3.4 | Associations of change in BMI with HRQoL

In a model examining BMI change and HRQoL, there was a negative association. A one unit decrease in BMI was associated with a 0.09-standard deviation unit higher HRQoL (change in BMI =  $-0.09$ ; 95% CI:  $-0.10$  to  $-0.08$ ) (Table 4). In a delta spline regression, there was no evidence of association between HRQoL and changes in BMI when BMI was below 25. At higher BMI, there were significant inverse associations. The association was modest in BMI range 25–29.9 kg/m<sup>2</sup> at  $-0.09$  (95% CI:  $-0.13$  to  $-0.04$ ) and was progressively stronger in each 5-point higher grouping (Table 4). Above a BMI of 40 kg/m<sup>2</sup>, a one unit BMI increase was associated with a  $-0.15$  lower HRQoL; (95% CI:  $-0.23$  to  $-0.07$ ).

Here, a one-unit loss in BMI would relate to higher HRQoL on its various indicators (listing only statistically significant associations): SF-36 pcs 0.43; Feelings Thermometer 0.62; EQ-5D 0.01; EQVAS 0.01; well-being 0.35.

### 3.5 | Reconciling past commissioning decisions with updated BMI-HRQoL estimates

To demonstrate an application of our results, we revisit a cost-effectiveness evaluation of orlistat. The previous cost (£) per quality-adjusted life year (QALY) was £24,430.50 (Table 5). Two results from this study were then applied to the same calculation: cross-sectional results from the linear regression (cost per QALY: £80,499.86) and results from the delta regression (cost per QALY: £60,517.02). Hence, using the results of our study may have a sizable impact on whether weight reduction interventions such as orlistat are deemed cost-effective, particularly in countries such as England where thresholds of £20,000–£30,000 applied by decision making bodies such as NICE.

### 3.6 | Nonlinear HRQoL and QALYs

Nonlinear HRQoL change from spline regressions was converted into QALYs to demonstrate the heterogeneity of HRQoL across the range of BMI; QALYs ranged from 0 to 0.427 (Table 6). In both cases, the mean change in QALYs was different to that in each of the categories underscoring the importance of this heterogeneity. Stark differences between the linear model and the delta model reinforce the difference between BMI change and cross-sectional differences in BMI: cross-sectional models may be over-estimating QALY gains.

**TABLE 2** HRQoL association with HRQoL indicators in measurement equations from the linear model and spline regression

	Linear model			Spline regression			Delta model			Delta spline regression		
	Est	95% CI	p value	Est	95% CI	p value	Est	95% CI	p value	Est	95% CI	p value
SF-36 pcs	2.91	(2.761 to 3.054)	<0.01	1.56	(0.939 to 2.173)	<0.01	4.92	(4.552 to 5.293)	<0.01	3.73	(1.98 to 5.484)	<0.01
SF-36 mcs	1.57	(1.381 to 1.75)	<0.01	0.84	(0.512 to 1.174)	<0.01	-0.38	(-1.053 to 0.286)	0.29	-0.29	(-0.819 to 0.247)	0.32
Feelings Thermometer	4.81	(4.515 to 5.101)	<0.01	2.56	(1.558 to 3.554)	<0.01	7.10	(6.43 to 7.763)	<0.01	5.34	(2.879 to 7.796)	<0.01
IWQOL	10.22	(8.939 to 11.506)	<0.01	5.72	(3.515 to 7.918)	<0.01	2.60	(-1.407 to 6.615)	0.20	2.11	(-0.872 to 5.095)	0.31
EQ-5D	0.09	(0.082 to 0.1)	<0.01	0.05	(0.03 to 0.071)	<0.01	0.09	(0.067 to 0.117)	<0.01	0.07	(0.034 to 0.114)	<0.01
EQVAS	0.04	(0.034 to 0.043)	<0.01	0.02	(0.013 to 0.03)	<0.01	0.08	(0.064 to 0.092)	<0.01	0.06	(0.031 to 0.091)	<0.01
Well-being	3.93	(3.672 to 4.179)	<0.01	2.15	(1.307 to 2.984)	<0.01	3.97	(2.731 to 5.217)	<0.01	2.85	(1.545 to 4.155)	<0.01

Abbreviations: Est, estimate; EQ VAS, EuroQol Visual Analogue Scale; EQ-5D, EuroQol Five Dimension; IWQoL, Impact of Weight on Quality Of Life; SF-36 pcs, Short Form-36 physical component summary; SF-36 mcs, Short Form-36 mental component summary; Well-being, General Well-Being Schedule; 95% CI, 95% confidence interval.

Note: All of the indicators' ranges were 0–100, save for the EQ-5D, which had a range of less than 0 to 1 and General Well-Being Schedule that ranged from 0 to 110. NB - to replicate analyses of past commissioning decisions, the EQVAS was transformed to a utility as per Foxcroft et al.<sup>12</sup> NB - aside from the EQVAS, HRQoL scales are analyzed in their raw form without having preference tariffs applied.

## 4 | DISCUSSION

In a cross-sectional analysis of individual participant data from five intentional weight loss trials, HRQoL was negatively related to BMI, primarily because of an inverse association in those with BMI > 30, with no evidence of an association in the BMI range under 30 kg/m<sup>2</sup>. In longitudinal analysis, HRQoL was negatively related to change in BMI in people with a BMI > 25 kg/m<sup>2</sup>. The coefficients in the cross-sectional analyses for people with a high BMI were larger than those for change in BMI. The mean of the HRQoL indicators is close to those at the population mean. However, we were able to exploit BMI variation to examine how the HRQoL-BMI relationship varied across the range of BMI.

We used IPD from large randomized controlled trials in several countries to perform a meta-analysis to provide a robust synthesis of the available evidence. We were able to exploit both the cross-sectional and longitudinal aspects to provide robust evidence and new insights into the relationship between BMI and HRQoL. A natural limitation of the pooled data is that variables are not collected consistently across trials. We sought to overcome that and exploit the precision of pooled data through the use of latent variable modelling, but if the association between quality of life and each measure genuinely differs, this will only be partially successful. Incident or prevalent disease related to BMI could partly be responsible for the differences by BMI, but given the relatively short follow-ups, is less likely to cloud the data on change in BMI. A potential limitation is that one trial, LookAHEAD, provided 45% of the total participants and with the longest follow up. However, we found associations across all trials, and the results for each study were similar (Appendix E). It is not known whether these results can be generalized to other weight-loss interventions. A further issue is that we cannot rule out confounding if actions which resulted in weight loss directly impacted on HRQoL and not only through changes in BMI. The direction and magnitude of these effects are unknown. This would not be the case for control arm individuals (a model only on control arm individuals in Appendix G shows consistent results). However, beyond any specific effect caused by participation in a trial, it reflects what happens when people lose weight—in a world of other competing effects on HRQoL. Finally, direct comparison across the HRQoL indicators is inappropriate because they are all attempting to measure HRQoL in different ways, but our use of latent variable analysis overcame this to some extent. Indeed, we see our approach as a pragmatic solution to the vexing issue of modelling different HRQoL indicators together. However, the latent variable is not without potential theoretical limitations.<sup>23</sup> We have assumed within this approach that HRQoL is causing the indicators of it, but the opposite is of course possible. The latent variable is ultimately capturing shared variation among the separate HRQoL indicators; whether or not this constitutes a true HRQoL, if such a singular thing exists, is open to question. We reiterate that, aside from the EQVAS, HRQoL scales are analyzed in their raw form without having preference tariffs applied, and should be understood as such.

**TABLE 3** Cross-sectional estimates of associations between 1 kg/m<sup>2</sup> difference in BMI and standard deviation units of HRQoL from the structural equations from the linear model and spline regression

	Linear model			Spline regression		
	Est	95% CI	p value	Est	95% CI	p value
BMI	−0.13	(−0.139 to −0.124)	<0.01			
BMI: 18.5–24.9				0.00		
BMI: 25–29.9				0.00		
BMI: 30–34.9				−0.36	(−0.59 to −0.139)	<0.01
BMI: 35–39.9				−1.16	(−1.71 to −0.6)	<0.01
BMI: ≥40				−2.42	(−3.51 to −1.338)	<0.01

Abbreviations: Est, estimate; 95% CI, 95% confidence interval.

Note: HRQoL was in turn related to each of its indicators (see Table 2 for ranges of individual HRQoL indicators).

**TABLE 4** Longitudinal estimates of association between changes in BMI and standard deviation units of HRQoL from the structural equations from the delta model and delta spline regression

	Delta model			Delta spline regression		
	Est	95% CI	p value	Est	95% CI	p value
Change in BMI	−0.087848	(−0.099 to −0.075)	<0.01			
Change in BMI: 18.5–24.9				0.00		
Change in BMI: 25–29.9				−0.08522	(−0.131 to −0.039)	<0.01
Change in BMI: 30–34.9				−0.1048	(−0.162 to −0.046)	<0.01
Change in BMI: 35–39.9				−0.12335	(−0.193 to −0.052)	<0.01
Change in BMI: ≥40				−0.1524	(−0.231 to −0.073)	<0.01

Abbreviations: Est, estimate; 95% CI, 95% confidence interval.

Note: HRQoL was in turn related to each of its indicators (see Table 2 for ranges of individual HRQoL indicators).

**TABLE 5** NICE cost-effectiveness of orlistat using the original value and using our estimates

	Utility gain	BMI change	Change in QALYs	QALY change in the trial arm	QALYs/100	Cost/100	Cost per QALY
Foxcroft orlistat	0.017	6.862	0.117	1.493	0.927	22744.8	24,430.50
Buckell et al. static orlistat	0.005	6.862	0.036	0.455	0.283	22744.8	80,499.86
Buckell et al. delta orlistat	0.007	6.862	0.047	0.605	0.376	22744.8	60,517.02
Foxcroft Placebo	0.017	6.282	0.107	0.566			
Buckell et al. static Placebo	0.005	6.282	0.033	0.172			
Buckell et al. delta Placebo	0.007	6.282	0.043	0.229			

Abbreviation: QALY, quality adjusted life year.

Note: “Utility gain” is the gain in utility from a one unit decrease in BMI. This is derived from the EQVAS measure which is converted to a utility score based on Hakim et al.<sup>13</sup> “BMI change” is the total BMI change in the trial arm. “Change in QALYs” multiplies the first two columns. “QALY change in the trial arm” computes the QALY changes per 100 respondents in each trial arm. “QALYs/100” takes the difference in QALY changes between treatment and control. “Cost/100” is the cost of the treatment per 100 respondents. “Cost per QALY” is the cost per QALY of the intervention from the original study and using our revised estimates. Figures that are not computed in this paper can be found in the original study.<sup>12</sup>

While our findings generally confirm the direction of effect in previous estimates, we provide new data on how BMI change appears to influence HRQoL, and importantly, we show that the effect of this depends upon starting BMI. The association is stronger for physical HRQoL than mental HRQoL. Prior evidence of the

association between mental HRQoL and BMI is mixed, including mainly negative results, but some reports of positive associations or non-significant estimates.<sup>3,8,24</sup> One meta-analysis reported higher mental HRQoL in overweight than for healthy weight.<sup>8</sup> These inconsistent findings may be because the magnitude of the association is



**TABLE 6** Change in QALYs for non-linear changes in HRQoL from cross-sectional and delta models

Cross-sectional models	Utility gain	BMI change	QALYs	Delta models	Utility gain	BMI change	QALYs
BMI: 18.5–24.9	0.000	6.862	0.000	Change in BMI: 18.5–24.9	0.000	6.862	0.000
BMI: 25–29.9	0.000	6.862	0.000	Change in BMI: 25–29.9	0.007	6.862	0.046
BMI: 30–34.9	0.008	6.862	0.055	Change in BMI: 30–34.9	0.008	6.862	0.056
BMI: 35–39.9	0.025	6.862	0.174	Change in BMI: 35–39.9	0.010	6.862	0.066
BMI: ≥40	0.053	6.862	0.366	Change in BMI: ≥40	0.012	6.862	0.082
Mean	0.005	6.862	0.036	Mean	0.007	6.862	0.047

Abbreviation: QALY, quality adjusted life year.

Note: Utility gains are expressed in utils derived from a Time-Trade-Off-transformed visual analogue scale. BMI change is taken from previous NICE cost-effectiveness decision on orlistat.<sup>12</sup> QALYs shows the difference of in QALYs for each of the BMI categories using the estimates from the spline regressions. Mean shows the non-spline regression results for comparison.

small, and many studies are underpowered to detect effects. We found no evidence of any adverse effects of intentional weight loss on mental health, even after weight regain, providing some reassurance to those expressing concerns.<sup>9</sup>

Understanding HRQoL is key for commissioning decisions about treatment interventions and so critical to patient care. The present study offers the highest quality evidence to date on the relationship between BMI and health-related quality of life in the context of behavioral interventions. In updating prior cost-effectiveness calculations, a different commissioning decision would have been made if these results had been used (though we recognize that our results are derived from behavioral interventions and applied to a pharmacotherapeutic intervention). The present study highlights the stronger relationship between BMI and HRQoL for people with BMI ≥ 30 kg/m<sup>2</sup>, where treatments are most often targeted. Commissioners in the United Kingdom have previously used an estimate of HRQoL-BMI at the sample mean, disregarding the heterogeneity across the range of BMI. This implies gains in HRQoL used to determine the cost-effectiveness of interventions will be underestimated. This was applied in the case of *orlistat* when it was considered by NICE.<sup>12</sup> For recent appraisals, NICE have used estimates in which the value of HRQoL-BMI varies over BMI. However, these values rely on cross-sectional, observational evidence<sup>7</sup> which we show greatly overestimates the change observed with intervention as demonstrated in QALY conversions of our HRQoL estimates. Furthermore, HRQoL gains in appraisals have assumed to be sustained; our results indicate that this is not the case (Appendix D). Our results, which provide estimates specific to people with a BMI > 30 kg/m<sup>2</sup> and examine the change with behavioral interventions, are likely to provide a more appropriate estimate, especially since initial states of HRQoL vary between categories of BMI. Our results can be used for commissioning decisions involving any of the HRQoL indicators because we measured their association with HRQoL and in turn its association with BMI. Without a robust, well-accepted relationship, there is potential for inconsistencies to emerge in commissioning determinations. Similar analyses are warranted with regard to decisions on pharmacological agents and surgical interventions for weight management.

## ACKNOWLEDGMENTS

We thank Frauke Becker for providing STATA code for processing trials' data. We thank attendees at Health Economists' Study Group meeting and CERGAS at Bocconi University for useful comments when discussed. This work was supported by the NIHR Oxford Biomedical Research. PA and SAJ are NIHR senior investigators and funded by the NIHR Oxford Biomedical Research Centre and Applied Research Collaboration.

The Diabetes Prevention Program (DPP) was conducted by the DPP Research Group and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), the General Clinical Research Center Program, the National Institute of Child Health and Human Development (NICHD), the National Institute on Aging (NIA), the Office of Research on Women's Health, the Office of Research on Minority Health, the Centers for Disease Control and Prevention (CDC), and the American Diabetes Association. The data (and samples) from the DPP were supplied by the NIDDK Central Repositories. This manuscript was not prepared under the auspices of the DPP and does not represent analyses or conclusions of the DPP Research Group, the NIDDK Central Repositories, or the NIH.

Look AHEAD was conducted by the Look AHEAD Research Group and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); the National Heart, Lung, and Blood Institute (NHLBI); the National Institute of Nursing Research (NINR); the National Institute of Minority Health and Health Disparities (NIMHD); the Office of Research on Women's Health (ORWH); and the Centers for Disease Control and Prevention (CDC). The data (and samples) from Look AHEAD were supplied by the NIDDK Central Repositories. This manuscript was not prepared under the auspices of the Look AHEAD and does not represent analyses or conclusions of the Look AHEAD Research Group, the NIDDK Central Repositories, or the NIH.

This manuscript was prepared using TOHP Research Materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center and does not necessarily reflect the opinions or views of the TOHP or the NHLBI.

This manuscript was prepared using DioGENES Research Materials obtained from the DioGENES trial and does not necessarily reflect the opinions or views of the DioGENES trial. This manuscript was prepared using WRAP Research Materials obtained from the WRAP trial and does not necessarily reflect the opinions or views of the WRAP trial.

## CONFLICT OF INTEREST

There are no conflicts of interest to declare.

## ORCID

John Buckell  <https://orcid.org/0000-0002-4157-4217>

Paul Aveyard  <https://orcid.org/0000-0002-1802-4217>

## ENDNOTE

\* See the database of mapping studies, Dakin et al. (2018).

## REFERENCES

- Global Burden of Disease Obesity Collaborators. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *N Engl J Med*. 2017;377(1):13-27.
- Kolotkin RL, Andersen JR. A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life. *Clinical Obesity*. 2017;7(5):273-289. <https://doi.org/10.1111/cob.12203>
- Warkentin LM, Das D, Majumdar SR, Johnson JA, Padwal RS. The effect of weight loss on health-related quality of life: systematic review and meta-analysis of randomized trials. *Obes Rev*. 2014;15(3):169-182.
- DPP Research Group. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: the Diabetes Prevention Program Outcomes Study. *Lancet Diab Endocrinol*. 2015;3(11):866-875.
- Ma C, Avenell A, Bolland M, et al. Effects of weight loss interventions for adults who are obese on mortality, cardiovascular disease, and cancer: systematic review and meta-analysis. *BMJ*. 2017;359:1-4, j4849.
- McDonough CM, Grove MR, Tosteson TD, Lurie JD, Hilibrand AS, Tosteson ANA. Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. *Qual Life Res*. 2005;14(5):1321-1332.
- Søltøft F, Hammer M, Kragh N. The association of body mass index and health-related quality of life in the general population: data from the 2003 Health Survey of England. *Qual Life Res*. 2009;18(10):1293-1299.
- Ul-Haq Z, Mackay DF, Fenwick E, Pell JP. Meta-analysis of the association between body mass index and health-related quality of life among adults, assessed by the SF-36. *Obesity*. 2013;21(3):E322-E327.
- Jones RA, Lawlor ER, Griffin SJ, van Sluijs EMF, Ahern AL. Impact of adult weight management interventions on mental health: a systematic review and meta-analysis protocol. *BMJ Open*. 2020;10(1):e031857.
- Ananthapavan J, Sacks G, Brown V, et al. *Assessing Cost-Effectiveness of Obesity Prevention Policies in Australia 2018 (ACE-Obesity Policy)*. Melbourne: Deakin University; 2018.
- Chen G, Ratcliffe J, Olds T, Magarey A, Jones M, Leslie E. BMI, Health Behaviors, and Quality of Life in Children and Adolescents: A School-Based Study. *Pediatrics*. 2014;133(4):e868-e874.
- Foxcroft DR. Orlistat for the treatment of obesity: cost utility model. *Obes Rev*. 2005;6(4):323-328.
- Hakim Z, Wolf A, Garrison LP. Estimating the Effect of Changes in Body Mass Index on Health State Preferences. *Pharmacoeconomics*. 2002;20(6):393-404.
- Klarenbach S, Padwal R, Wiebe N, et al. *Bariatric Surgery for Severe Obesity: Systematic Review and Economic Evaluation [Internet]*. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2010 (Technology report; no. 129). [cited 2010-09-20]. Available from: <http://www.cadth.ca/index.php/en/hta/reportspublications/search?%26type=16>
- Lacey LA, Wolf A, O'Shea D, Erny S, Ruof J. Cost-effectiveness of orlistat for the treatment of overweight and obese patients in Ireland. *Int J Obes (Lond)*. 2005;29(8):975-982.
- Hartmann-Boyce J, Aveyard P, Hobbs R, et al. Systematic review of weight regain after behavioural weight management programmes and the impact on cardiometabolic risk factors, health outcomes, and quality of life. PROSPERO 2018 CRD42018105744 Available from: [https://www.crd.york.ac.uk/prosperto/display\\_record.php?ID=CRD42018105744](https://www.crd.york.ac.uk/prosperto/display_record.php?ID=CRD42018105744)
- Ahern AL, Wheeler GM, Aveyard P, et al. Extended and standard duration weight-loss programme referrals for adults in primary care (WRAP): a randomised controlled trial. *Lancet*. 2017;389(10085):2214-2225.
- DPP Research Group. Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin. *N Engl J Med*. 2002;346(6):393-403.
- Larsen TM, Dalskov S, Van Baak M, et al. The Diet, Obesity and Genes (Diogenes) Dietary Study in eight European countries – a comprehensive design for long-term intervention. *Obes Rev*. 2010;11(1):76-91.
- Look AHEAD Research Group. Cardiovascular Effects of Intensive Lifestyle Intervention in Type 2 Diabetes. *N Engl J Med*. 2013;369(2):145-154.
- Whelton PK, Appel L, Charleston J, et al. The Effects of Non-pharmacologic Interventions on Blood Pressure of Persons With High Normal Levels: Results of the Trials of Hypertension Prevention, Phase I. *JAMA*. 1992;267(9):1213-1220.
- Buckell J, Mei XW, Becker F, Clarke P, Aveyard P, Jebb S. The relationship between BMI and health-related quality of life (HRQoL) for overweight and obese individuals: evidence from weight loss trials. 2020. Retrieved from [osf.io/mqxtp](https://osf.io/mqxtp)
- Lee N, Cadogan JW. Problems with formative and higher-order reflective variables. *J Bus Res*. 2013;66(2):242-247.
- Kroes M, Osei-Assibey G, Baker-Searle R, Huang J. Impact of weight change on quality of life in adults with overweight/obesity in the United States: a systematic review. *Curr Med Res Opin*. 2016;32(3):485-508.

**How to cite this article:** Buckell J, Mei XW, Clarke P, Aveyard P, Jebb SA. Weight loss interventions on health-related quality of life in those with moderate to severe obesity: Findings from an individual patient data meta-analysis of randomized trials. *Obesity Reviews*. 2021;22(11):e13317. <https://doi.org/10.1111/obr.13317>



## APPENDIX A: HRQoL OUTCOMES

Main outcome(s): The main outcome of interest was HRQoL. Across the trials, various HRQoL indicators were used. Each trial included required at least one of the following indicators. Obesity-specific minimally clinically important distances (MCIDs) are reported (Warkentin et al., 2014) and more general MCIDs where obesity-specific values are not available (Zhang et al., 2021).

- i. Short Form-36 (SF-36, version 2; Brazier et al., 2002). General population-based quality of life measure. Comprises both physical (pcs) and mental (mcs) health. Physical health is measured using 21 items across four domains. Mental health is measured using 14 items across four domains. Items are summarized into overall pcs and mcs scores. These are *t* scores with a mean of 50 and standard deviation of 10, based on the U.S. population. The MCID for both the pcs and the mcs is 5.
- ii. Short-Form Impact of Weight on Quality Of Life (IWQOL-Lite; Kolotkin et al., 2002). An obesity-specific measure of HRQoL. There are 31 items over five scales (physical function, self-esteem, sexual life, public distress, and work). Items are summarized into overall scores ranging from 0 to 100. The MCID is 12.
- iii. EuroQol Five Dimension (EQ-5D; EuroQol, 1990). A general population, preference-based health status measure. It comprises five dimensions of health status (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). Items are summarized into overall scores ranging from 0 to 1. The MCID is 0.03.
- iv. EuroQol Visual Analogue Scale (EQVAS). A vertical scale ranging from 0 to 100. Zero is labelled “the worst imaginable health state” or “the worst health you can imagine.” and 100 is labelled “the best imaginable health state” or “the best health state you can imagine.” The MCID is 10.
- v. Feeling thermometer. A visual analogue scale of 0 to 100 that allows individuals to rate their quality of life on this scale. The MCID is between 0.061 and 0.074 (though this is not specific to obese individuals, see Zhang et al., 2021).
- vi. General Well-Being Schedule (Dupuy, 1977). A psychological scale measuring well-being and level of distress; 18 items cover six dimensions that are summarized in a scale of 0 to 110. Scores of 0–60 denote “severe” distress; 61–72 denote “moderate distress”; and 73 to 110 denote “positive well-being.”

HRQoL indicators collected varied on a trial-by-trial basis. Some trials collected multiple indicators, whereas other trials collected on a single indicator.

## APPENDIX B: LATENT VARIABLE FOR HRQoL

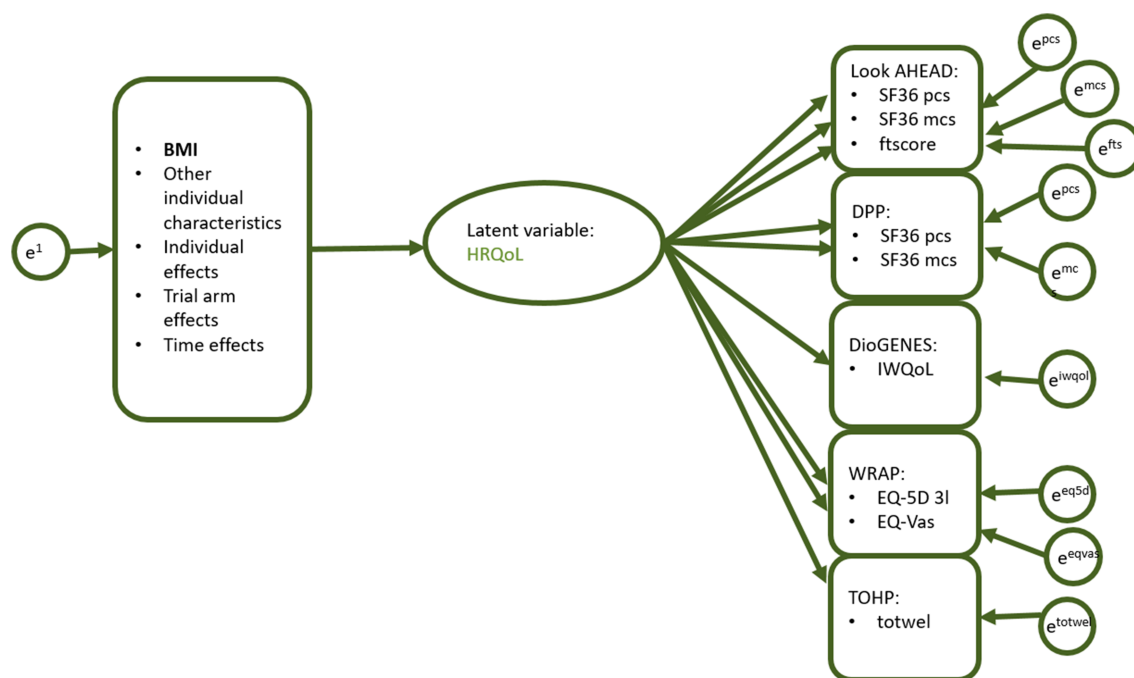
Health-related quality of life is unobservable. Indicator measures of HRQoL (i.e., SF-36, EQ-5D, etc.) serve, at best, as proxies for the underlying metric of interest. Treated as direct measures of HRQoL,

they are vulnerable to measurement error and reporting bias. Many indicators have been applied in studies of BMI: 12 meta-analyses studied in Kolotkin et al. (2017) reported 23 separate measures of HRQoL. Different indicators record and summarize multidimensional HRQoL differently. Thus, establishing which indicator best approximates it (i.e., minimizes measurement error) is not clear. The use of different HRQoL indicators across studies makes comparing HRQoL outcomes problematic. This presents vexing issues in this study where data are jointly modelled. Conventionally, this problem has been approached by mapping between HRQoL indicators (Devlin et al., 2020). This approach is less than ideal because it requires (a) an assumption that the indicator mapped to best approximates HRQoL; (b) that the mapping algorithm used does not induce error; and (c) that algorithms for mapping exist to do so.\* Elsewhere, the issue has been addressed by using contingency tables (Warkentin et al., 2014), but here much of the information in the data is lost and only comparable indices (such as generic indicators of HRQoL) are compared. In other meta-analyses, eligibility criteria was the use of the most common HRQoL indicator (SF-36; UI Haq et al., 2013), which limits the data to analyze because studies that use other indicators are discarded. Here, these issues are avoided by treating HRQoL as a latent variable, treating various HRQoL indicators as such rather than direct measures, and using a system of equations to estimate HRQoL-BMI (see, e.g., Feng et al., 2019, for a similar approach with EQ-5D). This approach relaxes the dependence on the HRQoL indicator to directly measure HRQoL. There is neither a need to assume a best-in-class indicator of HRQoL of those at hand, nor a need to map other HRQoL indicators to it. This approach allows us to dictate the specification of the model and the indicators used. We exploit this to examine BMI variation against both physical and mental HRQoL simultaneously. Further, since we no longer require a single dependent variable for HRQoL, it is possible to use all of the available indicators in the data, that is, multiple indicators for the same individual in a trial.

The latent variable is estimated using information on HRQoL that is contained in the data but not observed, determined in estimation by the fit of the model (under a set of assumptions, e.g., distributional assumptions on error terms). The latent variable uses all of the HRQoL data available. This is shown in Figure B1, where various trials use multiple and different indicators of HRQoL. Models within the system are jointly estimated. There are two types of equation here:

- Measurement equations: which use the indicator of HRQoL, for example, SF-36, as the dependent variable and measure the association of these indicators with the latent variable of HRQoL
- Structural equation: which treats the latent variable of HRQoL as the dependent variable and measures the association between HRQoL and individual characteristics such as BMI.

With these, we measure the relationship between HRQoL and BMI in the structural equation and then how BMI is related to the various indicators of HRQoL through the measurement equations.



**FIGURE B1** Schematic of system of equations for estimating HRQoL as a latent variable and its relationship with BMI. Square boxes are observed data; ellipses are unobserved variables; circles are noise

Structural equations:

$$HRQoL_{ijt} = \alpha_i + \gamma_{bmi} BMI_{ijt} + \gamma_{bmim} \overline{BMI}_{ijt} + \gamma' Z_{it} + \sum_{j=1}^4 \gamma'_j Trial_j + \sum_{t=1}^4 \gamma'_t Months\ since\ baseline_t + \varepsilon_{ijt} \quad (B1)$$

Where  $HRQoL_{it}$  is a latent variable of health-related quality of life.  $\gamma_{bmi}$  is the coefficient of interest, capturing the relationship between BMI and HRQoL. (We then in turn relate changes in BMI to changes in specific HRQoL indicators, using the  $\zeta$  coefficients in the measurement equations.)  $\overline{BMI}$  are the individual group means of BMI, so as to recover the within estimator from the random effects specification (Bell et al., 2019).  $Z$  is a vector of individual characteristics including age, gender and ethnicity;  $\gamma$  are to be estimated. Trial are trial-specific fixed effects; four parameters,  $\gamma_j$ , are estimated for 5 trials as one is set to zero to avoid linear dependency. Months since baseline are time periods (measured in months) since baseline; four parameters,  $\gamma_t$ , are estimated for 5 time periods as one is set to zero to avoid linear dependency.  $\alpha_i$  is an individual-specific random effect, assumed to be iid normally distributed, with zero mean and variance,  $\sigma_\alpha^2$ ; that is,  $\alpha \sim iid N(0, \sigma_\alpha^2)$ .

For spline regressions, BMI is partitioned into regions of its range, denoted BMI\_c according to the standard BMI classification, omitting underweight as no participants in the study were underweight.  $c = normal\ weight, overweight, obese\ category\ I, obese\ category\ II, obese\ category\ III$ . Boundary knots were placed at BMI=18.5 (as per the BMI categorization) and at BMI=55 (due to sparsity of data beyond BMI=55). Parameters for each region,  $\gamma_{bmi\_c}$ , are estimated. Similarly, the group means are specified for each individual and by grouping.

$$HRQoL_{ijt} = \alpha_i + \sum_{c=1}^5 \gamma_{bmi\_c} BMI\_c_{ijt} + \sum_{c=1}^5 \gamma_{bmim\_c} \overline{BMI\_c}_{ijt} + \gamma' Z_{it} + \sum_{j=1}^4 \gamma'_j Trial_j + \sum_{t=1}^4 \gamma'_t Months\ since\ baseline_t + \varepsilon_{ijt} \quad (B2)$$

Next, model (1) was re-specified to consider the relation of longitudinal changes in BMI and HRQoL. The period-to-period BMI changes are used in the linear model. We refer to this model as the delta model to refer to the change in BMI.

$$HRQoL_{ijt} = \alpha_i + \gamma_{bmi\_change} \Delta BMI_{ijt} + \gamma' Z_{it} + \sum_{j=1}^4 \gamma'_j Trial_j + \sum_{t=1}^4 \gamma'_t Months\ since\ baseline_t + \varepsilon_{ijt} \quad (B3)$$

Where  $\Delta BMI_{ijt} = BMI_{ijt} - BMI_{ijt-1}$ .

Finally, the longitudinal analogue of model (2) is specified. The period-to-period BMI changes are used in the spline regression. This is termed delta spline regression,

$$HRQoL_{ijt} = \alpha_i + \sum_{c=1}^5 \gamma_{bmi\_c} \Delta BMI\_c_{ijt} + \gamma' Z_{it} + \sum_{j=1}^4 \gamma'_j Trial_j + \sum_{t=1}^4 \gamma'_t Months\ since\ baseline_t + \varepsilon_{ijt} \quad (B4)$$

Where  $\Delta BMI\_c_{ijt} = BMI\_c_{ijt} - BMI\_c_{ijt-1}$ .

In all specifications, models are fully specified with all possible variables. Models are then iteratively refined by removing non-significant variables, or setting parameters to zero that do not influence the log-likelihood in pre-testing. Models are estimated using simulated log-likelihood (see e.g. Train, 2009). Modified Latin

Hypercube Sampling (MLHS) draws from a normal distribution are taken to construct the latent variable. Apollo software for R is used for estimation.

Measurement equations:

$$SF - 36(pcs)_{ijt} = \zeta_{pcs} HRQoL_{ijt} + \varepsilon_{ijt,pcs} \quad (B5)$$

$$SF - 36(mcs)_{ijt} = \zeta_{mcs} HRQoL_{ijt} + \varepsilon_{ijt,mcs} \quad (B6)$$

$$FT\ score_{ijt} = \zeta_{ft\ score} HRQoL_{ijt} + \varepsilon_{ijt,ft\ score} \quad (B7)$$

$$IWQoL_{ijt} = \zeta_{iwqol} HRQoL_{ijt} + \varepsilon_{ijt,iwqol} \quad (B8)$$

$$EQ - 5D_{ijt} = \zeta_{eq-5d} HRQoL_{ijt} + \varepsilon_{ijt,eq5d} \quad (B9)$$

$$EQVAS_{ijt} = \zeta_{eq-vas} HRQoL_{ijt} + \varepsilon_{ijt,vas} \quad (B10)$$

$$totwel_{ijt} = \zeta_{totwel} HRQoL_{ijt} + \varepsilon_{ijt,totwel} \quad (B11)$$

where indicators of HRQoL, for example,  $SF - 36(pcs)$ , are treated as dependent variables to be explained by the latent variable of HRQoL. Then, the  $\zeta$  capture the association between the indicators of HRQoL and the latent variable for HRQoL.

## APPENDIX C: ASSOCIATIONS OF HRQoL AND DEMOGRAPHIC VARIABLES

Females had lower HRQoL than males (estimate =  $-0.18$ ; 95% CI:  $-0.35$  to  $-0.01$ ); and black individuals had higher HRQoL than white individuals (estimate =  $0.43$ ; 95% CI:  $0.23$  to  $0.64$ ). HRQoL was negatively associated with age. Relative to under-65s, those between ages 65 and 70 had lower HRQoL (estimate =  $-0.17$ ; 95% CI:  $-0.24$  to  $-0.09$ ); and those over 70 had lower HRQoL (estimate =  $-0.66$ ; 95% CI:  $-0.78$  to  $-0.55$ ). HRQoL was lower in WRAP than LookAHEAD in the spline regression (estimate =  $-1.01$ ; 95% CI:  $-1.40$  to  $-0.62$ ) and higher in TOHP than LookAHEAD (estimate =  $0.87$ ; 95% CI:  $0.71$  to  $1.03$ ) (Table C1).

**TABLE C1** Associations of covariates and HRQoL

	Linear model			Spline regression		
	Est	95% CI	p value	Est	95% CI	p value
Gamma Female	-0.18286	(-0.351 to -0.014)	<0.01	-0.55191	(-0.943 to -0.16)	<0.01
Gamma Age category2	-0.16658	(-0.238 to -0.094)	<0.01	-0.28104	(-0.45 to -0.111)	<0.01
Gamma Age category3	-0.66396	(-0.78 to -0.547)	<0.01	-1.13937	(-1.621 to -0.656)	<0.01
Gamma Black	0.4338	(0.226 to 0.641)	<0.01	0.68008	(0.164 to 1.195)	<0.01
Gamma Other	-0.04807	(-0.335 to 0.239)		0.0451	(-0.486 to 0.576)	
Gamma DPP						
Gamma DioGENES						
Gamma WRAP	0.01315	(-0.17 to 0.196)		-0.95245	(-1.326 to -0.578)	<0.01
Gamma TOHP				0.8565	(0.514 to 1.198)	<0.01

Abbreviations: DPP, Diabetes Prevention Programme; Est, estimate; HRQoL, health-related quality of life; TOHP, Trials of Hypertension Prevention; 95% CI, 95% confidence interval.

Note: In the linear model, an additional specification interacts the time-from-baseline variables with trials' interventions.

## APPENDIX D: TIME-FROM-BASELINE CHANGE IN HRQoL

Time-from-baseline fixed effects indicate how HRQoL varied over time for all participants across all of the trials. We estimated this

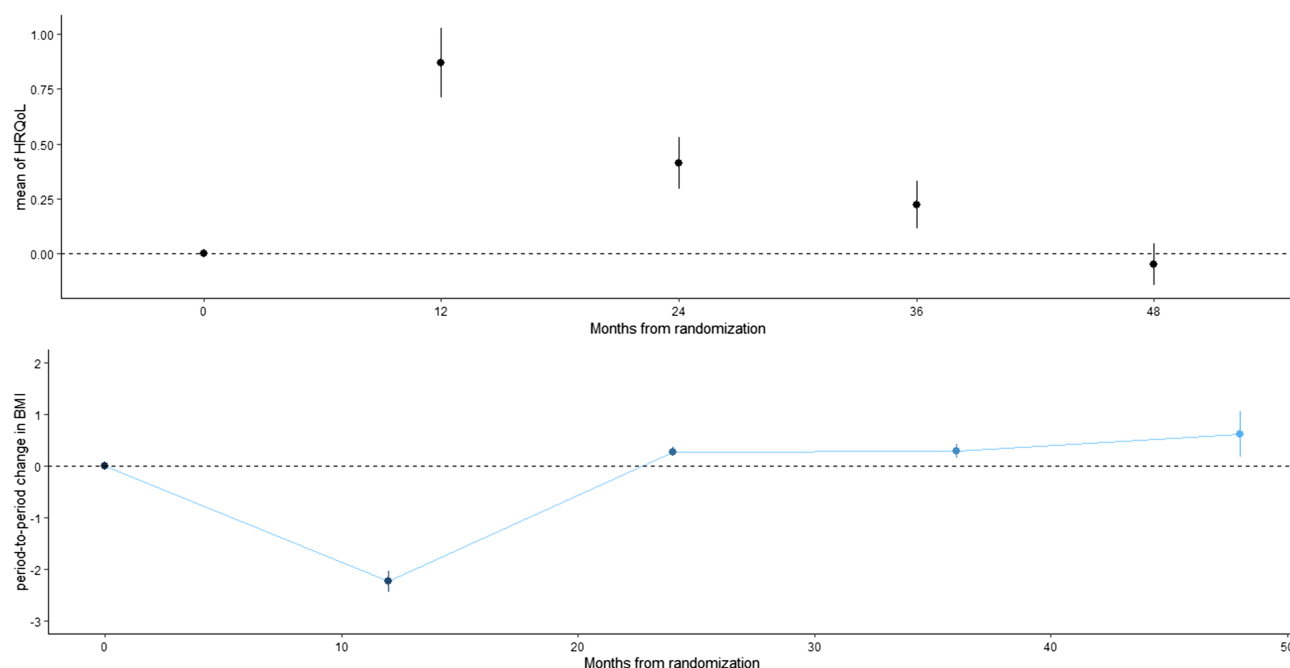
specification for the linear and the spline regressions. In both cases, HRQoL increased from baseline to 12 months and declined thereafter. By 48 months, the level of HRQoL had returned to that at baseline (Table D1 and Figure D1).

	Linear model			Spline regression		
	Est	95% CI	p value	Est	95% CI	p value
12 months	0.41	(0.348 to 0.473)	<0.01	0.86	(0.514 to 1.198)	<0.01
24 months	0.20	(0.141 to 0.251)	<0.01	0.40	(0.219 to 0.582)	<0.01
36 months	0.11	(0.059 to 0.166)	<0.01	0.22	(0.087 to 0.343)	<0.01
48 months	-0.02	(-0.069 to 0.029)	0.48	-0.05	(-0.142 to 0.048)	0.33

**TABLE D1** Estimates of longitudinal HRQoL change from the structural equations from the linear model and spline regression

Abbreviations: Est, estimate; HRQoL, health-related quality of life; 95% CI, 95% confidence interval.

Note: In the linear model, an additional specification interacts the time-from-baseline variables with trials' interventions.



**FIGURE D1** HRQoL and BMI change over time. Top panel: HRQoL at months after baseline. Lower panel: Period-to-period change in BMI

## APPENDIX E: MISSING DATA AND MODELS WITH MULTIPLE IMPUTATION

### Look AHEAD

Variable	Missing	Total	Percent missing
pcs	16,420	78,496	20.92
mcs	16,420	78,496	20.92
bmi	33,386	78,496	42.53
age_category	0	78,496	0.00
t	0	78,496	0.00
randarm	0	78,496	0.00
female	0	78,496	0.00
black	0	78,496	0.00
other	0	78,496	0.00
heduc	0	78,496	0.00
educ_missing	0	78,496	0.00
unemployed	0	78,496	0.00
unemployed~g	0	78,496	0.00
current_sm~r	0	78,496	0.00
past_smoker	0	78,496	0.00
cvdhistory	0	78,496	0.00
use_of_ins~n	0	78,496	0.00
hypertension	0	78,496	0.00

### DPP

Variable	Missing	Total	Percent missing
sfqal	8880	15,405	57.64
bmi	8204	15,405	53.26
female	0	15,405	0.00
black	0	15,405	0.00
hispanic	0	15,405	0.00
other	0	15,405	0.00
years_diab~s	0	15,405	0.00
heart_dise~e	0	15,405	0.00
stroke_bas~e	0	15,405	0.00
hours_sleep	4116	15,405	26.72
total_work	6322	15,405	41.04
total_hous~k	4312	15,405	27.99
total_recr~n	4163	15,405	27.02
alcohol_pe~k	6903	15,405	44.81
past_30_da~g	6903	15,405	44.81

### DioGENES

Variable	Missing	Total	Percent missing
iwqoltot_	1320	3363	39.25
bmi	1087	3363	32.32
female	0	3363	0.00
black	0	3363	0.00
other	0	3363	0.00
married	0	3363	0.00
age	543	3363	16.15

### WRAP

Variable	Missing	Total	Percent missing
eq5d	1448	5068	28.57
eq_vas	1398	5068	27.58
bmi_	1118	5068	22.06
female	0	5068	0.00
nonwhite	0	5068	0.00
age_category	0	5068	6.08
levelofedu~0	308	5068	6.08
household~e_	2056	5068	40.57
household~s_	2102	5068	41.48
bloodpress~_	2499	5068	49.31
lipidlower~_	2499	5068	49.31
diabetesmed_	2499	5068	49.31
obesitymed_	2499	5068	49.31
antidepres~_	2499	5068	49.31

### TOHP

Variable	Missing	Total	Percent missing
totwel	1676	6546	25.60
bmi	1621	6546	24.76
female	0	6546	0.00
black	0	6546	0.00
other	0	6546	0.00
college	0	6546	0.00
employed	0	6546	0.00
age	0	6546	0.00



Table E1 is as follows.

**TABLE E1** Comparison of non-imputed and imputed estimates for each of five trials

	Diogenes				DPP				Look AHEAD			
	OLS		FE		Spline		OLS		FE		Spline	
	raw	MI	raw	MI	raw	MI	raw	MI	raw	MI	raw	MI
BMI	-0.724 (-3.85)	-0.721 (-4.76)	-0.739 (-2.26)	-0.799 (-3.39)	0.000 (0.49)	0.000 (-0.06)	0.000 (0.49)	0.000 (-0.06)	0.002 (1.11)	0.000 (-0.06)	-0.435 (-25.79)	-0.432 (-37.92)
BMI1					0 (.)	0 (.)	0 (.)	0 (.)	0.006 (-2.51)	0.006 (-2.51)	-0.506 (-24.65)	-0.318 (-33.12)
BMI2					1.096 (0.23)	1.323 (-0.27)	1.096 (0.23)	1.323 (-0.27)	0.009 (1.35)	0.000 (-0.13)		
BMI3					-1.578 (-2.19)	-1.627 (-2.31)	-1.578 (-2.19)	-1.627 (-2.31)	0.004 (0.91)	0.000 (-0.49)		
BMI4					0.372 (0.85)	0.361 (-0.8)	0.372 (0.85)	0.361 (-0.8)	0.001 (0.15)	0.000 (-0.25)		
BMI5					-1.862 (-3.90)	-1.794 (-3.67)	-1.862 (-3.90)	-1.794 (-3.67)	-0.001 (-0.34)	0.000 (-0.08)		
BMI6					0 (.)	0 (.)	0 (.)	0 (.)	0 (.)	0 (.)		

Note: Three models are compared for each trial: OLS, fixed effects (FE) and spline regressions (spline). For each model, parameter estimates from the model on the raw data (raw) are presented next to those on the imputed data (MI) with robust t-ratios in parentheses.

**TABLE E1** Continued

	Look AHEAD				TOHP				WRAP			
	Spline		OLS		FE		Spline		OLS		FE	
	raw	MI	raw	MI	raw	MI	raw	MI	raw	MI	raw	MI
BMI			-0.281 (-4.28)	-0.283 (-4.81)	-0.683 (-5.03)	-0.303 (-4.80)			-0.009 (-6.69)	-0.010 (-8.55)	-0.008 (-3.96)	-0.009 (-4.58)
BMI1	0 (.)	0 (.)					2.178 (0.70)	2.127 (2.56)			0 (.)	
BMI2	0.119 (0.39)	0.153 (-0.64)					0.114 (0.54)	0.072 (0.32)			-0.0131 (-1.14)	-0.008 (-0.48)
BMI3	-0.3 (-3.54)	-0.308 (-5.06)					-0.505 (-3.82)	-0.483 (-3.54)			-0.00241 (-0.72)	-0.003 (-0.65)
BMI4	-0.322 (-5.26)	-0.308 (-6.74)					-0.345 (-1.85)	-0.371 (-1.83)			-0.011 (-4.46)	-0.016 (-4.77)
BMI5	-0.899 (-14.55)	-0.902 (-21.46)					-1.242 (-1.48)	-0.957 (-0.95)			-0.00493 (-1.20)	-0.003 (-0.61)
BMI6							0 (.)	0 (.)			-0.014 (-3.44)	-0.015 (-4.14)

Note: Three models are compared for each trial: OLS, fixed effects (FE) and spline regressions (spline). For each model, parameter estimates from the model on the raw data (raw) are presented next to those on the imputed data (MI) with robust t-ratios in parentheses.

## APPENDIX F: DIAGNOSTIC INFORMATION FROM THE ESTIMATED MODELS

Table F1 is as follows.

**TABLE F1** Model diagnostics

	Cross section		Delta	
	Linear	Spline	Linear	Spline
No. of individuals	10,884	10,884	8531	8531
No. of observations	58,719	58,719	27,715	27,715
Estimated parameters	29	33	23	26
Simulated log-likelihood (full fitted model)	−509,624.7	−509,944.4	−238,724.3	−238,722.4
AIC (full fitted model)	1,019,307	1,019,955	477,494.6	477,496.8
BIC (full fitted model)	1,019,595	1,020,282	477,705	477,734.6
Simulated log-likelihood (measurement equation: pcs)	−149,861	−149,997.1	−71,881.56	−71,870.63
Simulated log-likelihood (measurement equation: mcs)	−155,600.1	−155,668	−74,093.12	−74,093.65
Simulated log-likelihood (measurement equation: ft)	−172,332	−172,623.6	−82,508.44	−82,508.82
Simulated log-likelihood (measurement equation: iwqol)	−172,332	−8684.095	−2147.588	−2147.694
Simulated log-likelihood (measurement equation: eqvas)	2114.198	2104.347	719.3283	716.1988
Simulated log-likelihood (measurement equation: eq5d)	684.5828	647.1308	301.0425	300.6449
Simulated log-likelihood (measurement equation: well)	−18,063.81	−18,039.41	−9696.251	−9698.8

Abbreviations: AIC, Akaike information criteria; BIC, Bayesian information criteria; NB, different numbers of draws were required for the models to facilitate estimation, thus direct comparison of the simulated log-likelihood is not appropriate.

Note: Simulated log-likelihoods for the full model and for the measurement equations are presented—one for each of the HRQoL indicators.

## APPENDIX G: SENSITIVITY ANALYSIS USING ONLY PARTICIPANTS IN THE CONTROL ARMS OF THE TRIALS

Table G1 is as follows.

**TABLE G1** Estimated BMI coefficient and diagnostics from model using only individuals in the control arms of trials

	Estimate	95% CI	p value
BMI	−0.122	(−0.134 to −0.110)	<0.01
No. of Individuals	4258		
No. of Observations	26,556		
Estimated parameters	29		
Simulated log-likelihood (full fitted model)	−248,039.2		
AIC (full fitted model)	496,136.4		
BIC (full fitted model)	496,401.6		

Abbreviations: AIC, Akaike information criteria; BIC, Bayesian information criteria; Est, estimate; 95% CI, 95% confidence interval.

## APPENDIX H: BASELINE AVERAGE HRQoL VALUES AT BMI CATEGORIES

Table H1 is as follows.

**TABLE H1** Mean HRQoL scores at baseline by BMI category

	SF-36 pcs	SF-36 mcs	IWQoL	Feelings thermometer	EQ-5D	EQ VAS	Well-being
BMI category: Underweight							74.5
BMI category: Normal Weight	50.5	51.2	67.7	85.0			80.8
BMI category: Overweight	50.4	54.1	66.0	77.9	0.8	76.3	81.6
BMI category: Obese Class I	49.4	54.3	64.2	75.9	0.8	72.6	78.8
BMI category: Obese Class II	47.6	53.8	59.7	73.3	0.8	65.9	78.5
BMI category: Obese Class III	44.8	54.4	54.2	70.0	0.7	63.8	

## REFERENCES

- Brazier, J., Roberts, J., & Deverill, M. (2002). **The estimation of a preference-based measure of health from the SF-36.** *Journal of Health Economics*, 21(2), 271-292. [https://doi.org/10.1016/S0167-6296\(01\)00130-8](https://doi.org/10.1016/S0167-6296(01)00130-8)
- Bell, A., M. Fairbrother and K. Jones (2019). **"Fixed and random effects models: making an informed choice."** *Quality & Quantity* 53(2): 1051-1074.
- Devlin, N., D. Parkin and B. Janssen (2020). *Advanced topics. Methods for Analysing and Reporting EQ-5D Data.* N. Devlin, D. Parkin and B. Janssen. Cham, Springer International Publishing: 87-98.
- Dakin, H, Abel, L, Burns, R, Yang, Y. 2018. **Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement.** *Health and Quality of Life Outcomes* 16:31, 1.
- Dupuy, H.J. (1977). The general well-being schedule. In I. McDowell & C. Newell (Eds.), *Measuring Health: A Guide to Rating Scales and Questionnaire* (2nd ed) (pp. 206-213). USA: Oxford University Press.
- EuroQol. (1990). **"EuroQol—a new facility for the measurement of health-related quality of life."** *Health Policy* 16(3): 199-208.
- Feng, Y.-S., R. Jiang, T. Kohlmann and A. S. Pickard (2019). **"Exploring the internal structure of the EQ-5D using non-preference-based methods."** *Value in Health* 22(5): 527-536.
- Kolotkin, R. L., & Crosby, R. D. (2002). **Psychometric evaluation of the Impact of Weight on Quality of Life-Lite Questionnaire (IWQOL-lite) in a community sample.** *Quality of Life Research*, 11(2), 157-171.
- Kolotkin, R. L., & Andersen, J. R. (2017). **A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life.** *Clinical Obesity*, 7(5), 273-289. <https://doi.org/10.1111/cob.12203>
- Train, K. (2009). *Discrete Choice Methods With Simulation.* Cambridge, UK: Cambridge University Press.
- Ul-Haq, Z., D. F. Mackay, E. Fenwick and J. P. Pell (2013). **"Meta-analysis of the association between body mass index and health-related quality of life among adults, assessed by the SF-36."** *Obesity* 21(3): E322-E327.
- Warkentin, L. M., D. Das, S. R. Majumdar, J. A. Johnson and R. S. Padwal (2014). **"The effect of weight loss on health-related quality of life: systematic review and meta-analysis of randomized trials."** *Obesity Reviews* 15(3): 169-182.
- Warkentin, L. M., S. R. Majumdar, J. A. Johnson, C. B. Agborsangaya, C. F. Rueda-Clausen, A. M. Sharma, S. W. Klarenbach, S. Karmali, D. W. Birch and R. S. Padwal (2014). **"Weight loss required by the severely obese to achieve clinically important differences in health-related quality of life: two-year prospective cohort study."** *BMC Medicine* 12(1): 175.
- Zhang, P., K. M. Atkinson, G. A. Bray, H. Chen, J. M. Clark, M. Coday, G. R. Dutton, C. Egan, M. A. Espeland, M. Evans, J. P. Foreyt, F. L. Greenway, E. W. Gregg, H. P. Hazuda, J. O. Hill, E. S. Horton, V. S. Hubbard, P. J. Huckfeldt, S. D. Jackson, J. M. Jakicic, R. W. Jeffery, K. C. Johnson, S. E. Kahn, T. Killean, W. C. Knowler, M. Korytkowski, C. E. Lewis, N. M. Maruthur, S. Michaels, M. G. Montez, D. M. Nathan, J. Patricio, A. Peters, X. Pi-Sunyer, H. Pownall, B. Redmon, J. T. Rushing, H. Steinburg, T. A. Wadden, R. R. Wing, H. Wyatt and S. Z. Yanovski (2021). **"Within-trial cost-effectiveness of a structured lifestyle intervention in adults with overweight/obesity and type 2 diabetes: results from the Action for Health in Diabetes (Look AHEAD) study."** *Diabetes Care* 44(1): 67-74.