Risk factors relate to the variability of health outcomes as well as the mean: a GAMLSS tutorial

David Bann[1]*, Liam Wright[1], Tim J Cole[2]

[1]Centre for Longitudinal Studies, Social Research Institute, University College London, London, UK
[2]Great Ormond Street Institute of Child Health, University College London, London, UK

*corresponding author; david.bann@ucl.ac.uk, +44 020 7679 2000

Running head: risk factors and health outcome variability

Declarations of interest: none

Word count: 3581

Data availability: Available from the UK Data Archive:
https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200001

Background: Risk factors or interventions may affect the variability as well as the mean of health outcomes. Understanding this can aid aetiological understanding and public health translation, in that interventions which shift the outcome mean and reduce variability are typically preferable to those which affect only the mean. However, most commonly used statistical tools do not test for differences in variability. Tools that do have few epidemiological applications to date, and fewer applications still have attempted to explain their resulting findings. We thus provide a tutorial for investigating this using GAMLSS (Generalised Additive Models for Location, Scale and Shape).

Methods: The 1970 British birth cohort study was used, with body mass index (BMI; N=6,007) and mental wellbeing (Warwick-Edinburgh Mental Wellbeing Scale; N=7,104) measured in midlife (42-46 years) as outcomes. We used GAMLSS to investigate how multiple risk factors (sex, childhood social class and midlife physical inactivity) related to differences in health outcome mean and variability.

Results: Risk factors were related to sizable differences in outcome variability—for example males had marginally higher mean BMI yet 28% lower variability; lower social class and physical inactivity were each associated with higher mean and higher variability (6.1% and 13.5% higher variability, respectively). For mental wellbeing, gender was not associated with the mean while males had lower variability (-3.9%); lower social class and physical inactivity were each associated with lower mean yet higher variability (7.2% and 10.9% higher variability, respectively).

Conclusions: The results highlight how GAMLSS can be used to investigate how risk factors or interventions may influence the variability in health outcomes. This underutilised approach to the analysis of continuously distributed outcomes may have broader utility in epidemiologic, medical, and psychological sciences. A tutorial and replication syntax is provided online to facilitate this (https://osf.io/5tvz6/).

Funding: DB is supported by the Economic and Social Research Council (grant number ES/M001660/1), The Academy of Medical Sciences / Wellcome Trust ("Springboard Health of the Public in 2040" award: HOP001/1025); DB and LW are supported by the Medical Research Council (MR/V002147/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Key words: statistical methods, distributions, GAMLSS, quantile regression, body mass index, mental health, mental wellbeing

Introduction

What is health? Contrary to simplistic notions of its being defined as the absence of disease, it is now increasingly understood that most outcomes of public health significance are continuous in nature.[1] This applies to both physical and mental health outcomes.[2 3] The use of binary endpoints, while having utility in clinical applications, should not hinder investigation of the influences of health outcomes which are ultimately continuous. Further, analysing the determinants of health using continuous rather than binary outcomes is beneficial both practically (with more statistical power and less information loss) and substantively (greater aetiological understanding). Indeed, those at high risk of a developing an illness may comprise a minority of those who ultimately succumb.[4]

Studies into the effect on continuous outcomes of exposures, be they risk factors in observational studies or interventions in randomised trials,[3] typically focus on mean differences in the outcome, using linear regression. However linear regression assumes homoscedasticity, i.e. that the variability of the outcome is unrelated to the exposure, and often this is not the case. It is possible to extend regression analysis to model the variability as well as the mean, and this has benefits in terms of not only the model's fit but also its interpretation. If for example the intervention in a trial can be shown to reduce variability in the outcome, this could reasonably be viewed as evidence of intervention success[5] independent of the intervention's effect on the mean. Treatment for refractive vision errors—glasses, contact lenses, and/or corrective surgery—seeks to improve vision by shifting individuals towards a specified standard (e.g. 20/20 vision).[6] Successful treatments alter the mean refraction, but they are even more successful if they also reduce the substantial variability in refraction arising from the mix of short- and long-sighted individuals.

Similarly, obesity interventions aim to reduce body mass index (BMI) and shift treated individuals from overweight (25-30 kg/m^2), obese (>30 kg/m^2) or severely obese (>45 kg/m^2) to the normal range (20-25 kg/m^2). However here the effect of the intervention on variability is often to increase it. Even if not formally tested, visual comparisons of outcome distributions of some influential trials suggest that weight loss interventions increase rather than reduce BMI variability,[7] presumably since they are effective in some but not all participants.

Understanding if and how risk factors influence variability in health outcomes has aetiological significance, consistent with the goal of epidemiological science to understand the *distribution* of health.[8] Risk factors could feasibly affect outcome variability yet not affect the mean—for example, one study found that breastfeeding was not related to mean childhood BMI, yet was related to lower childhood BMI variability.[9] Similarly, sex may affect variability and/or average levels of an outcome—for instance, males may have greater variability than females in some cognitive traits[10] and brain structures.[11]

Identifying associations between risk factors and outcome variability may also be useful to identify the absence or presence of heterogeneity in susceptibility to interventions or risk factors and thus aid aetiological understanding. Indeed, the finding that substantial increases in mean BMI in recent decades have been matched by increases in BMI variability indicates that there may be differential susceptibility to the obesogenic environment.[12 13] In the context of randomised controlled trials, the finding of variability in treatment effects between individuals has been used to justify individualised approaches to treatment (personalised medicine). Reflecting the challenges of empirically testing this however, five separate meta-analyses have tested heterogeneity in response to antidepressant therapy; despite using the same dataset, different methods and divergent conclusions were drawn.[14]

Another advantage of modelling varability arises in common situations where the outcome under study is non-linearly related to other outcomes of interest. For instance, BMI influences mortality and morbidity rates, but the relationship between BMI and mortality is thought to be J-shaped[15]; compared with those in the normal range, mortality risks are greater for those who are under- or overweight. In this case, the total effect of an intervention to reduce BMI on these wider outcomes is not fully captured by its average BMI effect. Rather, understanding the total distributional effect on BMI is required.

Figure 1 shows three hypothetical scenarios for an intervention to affect the distribution of an outcome. In the first case (Panel A), the intervention has an impact that is consistent across the population: all individuals are affected and to the same extent. In the second case (Panel B), the intervention has the same mean impact, but variability is also increased: some are positively affected, others negatively. In the third case (Panel C), the mean is again increased, but so is skewness. There is heterogeneity in response, with some seeing more positive responses than others. The policy implications may be different in each case. In the second and third scenarios, efforts could be directed to identify those who are (more) positively impacted, so as to increase the net benefit or cost-effectiveness of the intervention. Indeed, in a choice between interventions, an intervention generating lower expected benefits but smaller variability in outcomes may be chosen, in so far as reducing inequalities is seen as a policy goal in itself.

Recent studies in biological,[16 17] environmental[18] and economic science[19-21] have begun to examine how risk factors relate to the distribution of the outcome of interest. However, there have been few epidemiological applications of this approach to date;[22] and fewer still that provide explanations for such findings, which are essential if such methods are to have utility. Indeed, one recent study which investigated the association between mental health symptoms and lower income explicitly avoided

interpretation of its findings on variability, focusing instead on issues relating to the application of such methods.[20]

Regression methods that allow variability to be modelled are uncommon. One particular method, Generalised Additive Models for Location, Scale and Shape (GAMLSS)[23] has become the standard for constructing growth reference centiles,[24] where the aim is to model the outcome's distribution as a function of age. It defines the distribution in terms of distribution moments, i.e. the mean, variance, and optionally skewness and kurtosis. This allows for factors influencing the higher moments to be identified in just the same way as for the mean, and it provides a simple and elegant interface for modelling variability in epidemiology.

Another arguably underutilised[25] and related statistical approach to investigating risk factors for continuous outcomes is quantile regression. Recent epidemiological studies using this method have found that risk factors for higher BMI—particularly lower social class and physical inactivity—have sizably larger effect sizes at higher BMI centiles.[26 27] This has potentially important policy implications—risk factors which have larger effects amongst those at highest health risk are likely to have a more favourable effect on population health than alternatives which do not.[26] However, the reason for this phenomenon is not yet understood—it is likely to be logically consistent with results of GAMLSS analyses in which risk factors influence outcome means, variability and/or skewness.

In this paper, we provide a worked example of the use and interpretation of GAMLSS. Accompanying this is an online tutorial and full replication syntax for running GAMLSS in R (https://osf.io/5tvz6/). We investigate whether and how several established risk factors—sex, childhood socioeconomic circumstances, and physical inactivity[28]—relate to differences in outcome mean and variability. We choose two different continuous outcomes, an indicator of adiposity (body mass index, BMI) and mental wellbeing. These are two weakly correlated health outcomes, each of independent importance to population health. Each risk factor-outcome combination is the subject of previous (separate) literature which focuses largely on mean differences only. For instance, low socioeconomic position in childhood has been repeatedly related to higher BMI[29 30] and worse mental wellbeing in adulthood;[31-33] greater physical activity has notable likely bi-directional links with lower BMI[34] and higher wellbeing;[35-37] while males and females seemingly have similar mean BMI and wellbeing,[33] this may mask differences in variability or skewness, as suggested in the sizable sex differences in overweight and obesity rates.[38]

The further investigation of differences in variability and skewness in these outcomes is therefore arguably of substantive interest, providing further motivation to the tutorial content. We highlight the

contribution of GAMLSS by contrasting results with the more commonly used linear regression and (less commonly used) quantile regression models.

Methods

Study sample

The 1970 British birth cohort study (1970c) consists of all 17,196 babies born in Britain during one week of March 1970, with 10 subsequent waves of follow-up from childhood to midlife.[39] At the most recent wave (46 years), 12,368 eligible participants (those alive and not lost to follow-up) were invited to be interviewed at home by trained research staff—8,581 participants provided at least some data in this wave. At all waves, informed consent was provided and ethical approval granted.

Health outcomes

We selected two outcomes in midlife which capture different dimensions of health and are continuously distributed: adiposity (BMI), and mental wellbeing (Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)). BMI was measured at 46 years, and wellbeing at 42 years.[31] WEMWBS consists of 14 positively worded items—such as "I've been feeling optimistic about the future" and "…feeling cheerful"—measured on a five-point Likert scale, which are summed to give a total well-being score ranging from 14 to 70 (highest well-being).[40]

Risk factors

We chose three risk factors across different domains—each of them likely to independently influence health outcomes.[28] They were coded as binary variables to simplify comparison of descriptive and GAMLSS results: sex (female/male), socioeconomic position (social class at birth; coded as non-manual/manual), and a behavioural risk factor (reported physical activity at 42 years; reported days in which the participant took part in exercise for 30 mins or more in a typical week 'working hard enough to raise your heart rate and break into a sweat', coded as active (≥1 days)/inactive (0 days)). We examined if the binary split of risk factors influenced the inferences drawn—additional analyses were conducted with them coded instead as categorical variables (social class in 6 categories and physical inactivity from 0-7 days).

Analytical strategy

To visually inspect the outcome distributions and their differences across risk factor groups, we first plotted separate kernel density estimates alongside relevant descriptive statistics (mean, standard deviation, and coefficient of variation (CoV = SD/mean)). This enables a descriptive depiction of variability, with unadjusted GAMLSS results corresponding to each descriptive statistic. We then used GAMLSS[23] separately with each outcome, to formally investigate whether risk factors were associated with 1) differences in mean outcome, 2) differences in outcome variability, 3) differences in outcome skewness. Linear regression analysis, in contrast, only enables mean differences in outcomes to be investigated.

GAMLSS is a form of regression analysis that estimates different 'moments' of the outcome distribution. The first moment is the location (see mean in Figure 1 panel a), the second is variance, which specifies the scale or spread (SD in Figure 1 panel b) the third is skewness which quantifies the relative size of the distribution tails (Figure 1 panel c). As in linear regression analyses covariates can optionally be included, and appropriate link functions can be chosen for use.

GAMLSS requires that the distribution is specified at the outset. In this tutorial we use two distributions which we recommend for use in epidemiological research of continuous outcomes. First, the normal distribution (called NO in GAMLSS), where location is measured by the mean and scale by the standard deviation (SD). The normal distribution has no 'shape' moments, as there is no skewness and kurtosis is fixed.

Second, a more complex distribution which enables skewness to be investigated: the Box-Cox Cole and Green (BCCG). Here location is the median, scale is the generalised coefficient of variation (CoV), which is calculated in the normal case as SD/mean, and shape is skewness as defined by the Box-Cox power required to transform the outcome distribution to normality. The transformation requires the outcome to be on the positive line, so zero or negative values are excluded. BCCG is effectively NO with added skewness, though parameterised differently. A Box-Cox power of 1 indicates that the distribution is normal, 0 is log-normal and -1 inverse normal, so a smaller (i.e. more negative) power corresponds to more right skewness.

After choosing a distribution, linear models are used to specify the relationship between the independent variables and the different moments of the outcome distribution. As with other regression models, GAMLSS provides a standard error for each estimated coefficient, from which 95% confidence intervals can be calculated. We note that more experienced users may wish to use alternative distributions which GAMLSS facilitates.[41]

In our primary analyses we used the NO and BCCG families. Differences in variability are modelled with a log link, and can be multiplied by 100 and interpreted as percentage differences in variability to aid interpretation.[42] Differences in the mean and median were also analysed as percentages, to aid comparability across outcomes and model estimates. To aid comparison of descriptive statistics and model estimation results, we first conducted analyses adjusting for each risk factor alone. We then adjusted for the risk factors jointly.

Separately we fitted conditional quantile regression models to estimate risk factor and BMI associations at the lower, middle and upper quartiles of the outcome distribution, i.e. the 25^{th}, 50^{th} and 75^{th} centiles. To aid comparison with methods more commonly used in the existing epidemiological

literature, we estimated generalised linear models which show the association between each risk factor and mean differences in outcomes.

All analyses were conducted using R v4.1.1. We used the *gamlss* package version 5.3-4 to produce gamlss models.[43] Syntax to replicate all analyses is presented online (https://osf.io/5tvz6/).

Results

6,007 participants had valid data for BMI and all risk factors, and 7,104 for WEMWBS. Mean BMI was 28.4 (SD = 5.5), and mean WEMWBS 49.2 (8.3). Higher BMI was weakly associated with lower wellbeing ($r = -0.07$, $p<0.01$). BMI was moderately right-skewed (Figure 2, left panel) and WEMWBS left-skewed (Figure 2, right panel). Visual and descriptive comparisons of the BMI and wellbeing distributions by risk factor suggest that differences in the outcome mean and variability are not always in the same direction.

GAMLSS results for the binary risk factors are shown in Tables 1 and 2, with the results using the extra risk factor categories in Supplementary Tables 1 and 2. Associations were similar in the unadjusted and mutually adjusted analyses, so the former are described below.

Body mass index

Males had higher mean BMI yet lower variability than females—see Figure 2 and Table 1. The SD for BMI was lower in males (4.6) than females (6.1) i.e., a 27.6% difference (difference in log(SD) *100). This matches the estimate obtained from GAMLSS—males had 27.6% (SE: 1.8%) less variability than females (Table 1).

In contrast, lower social class and physical inactivity were both associated with higher mean BMI and higher BMI variability (Figure 2 and Table 1). Those from lower social class households had 4% (SE 0.5%) higher mean BMI than those from non-manual classes, and 6.1% (1.9%) more variability. Physically inactive participants had 3.3% (0.6%) higher mean BMI and 13.5% (2.1%) more variability.

The GAMLSS results were similar with the BCCG distribution rather than NO (Table 1). That is, risk factors associated with higher mean BMI and higher SD were also associated with higher median BMI and higher CoV. Male sex and lower social class were both associated with less right skewness of the BMI distribution; the Box-Cox power was 0.5 (0.1) higher in males and 0.4 (0.1) higher for manual social class. Physical activity was not associated with outcome skewness.

Mental wellbeing – Warwick-Edinburgh Mental Wellbeing Scale

There was little evidence of sex differences in mean wellbeing, while males had marginally less variability than females by 4.0% (1.7%). Lower social class and physical inactivity were both associated with lower mean yet higher variability (Figure 2 and Table 2). Those from lower social class households had a 2.8% (0.4%) lower mean yet 7.2% (1.8%) higher variability. Physically inactive participants had 5.3% (0.5%) lower mean yet 10.9% (1.9%) higher variability. These findings were similar in mutually adjusted analyses (Table 2).

The results were similar with the BCCG distribution (Table 2). There was evidence suggesting that lower social class was associated with less skewness in the wellbeing distribution; sex and physical activity were not associated with outcome skewness.

Comparison with quantile regression findings

For BMI, the associations of lower social class and physical inactivity were stronger at upper quantiles (Table 3; e.g., manual social class had 3.7 (0.6) higher BMI at the the median, and 4.9 (0.7) at the 75^{th}); estimates at higher centiles were also estimated less precisely than at lower centiles (larger SE). In contrast sex differences were present at lower centiles but absent at the 75^{th} centile. These findings corresponded with those from GAMLSS using BCCG, with all BMI centiles plotted by risk factor group (Figure 3). This comparison highlights the utility of GAMLSS—risk factor differences in the mean, variability, and skewness can each be quantified and thus visually depicted.

For WEMWBS, the associations of lower social class and physical inactivity were also stronger at lower quantiles (Table 3), yet had larger standard errors. Sex was not associated with WEMWBS at any centile. These findings corresponded with those from GAMLSS (Figure 4).

Discussion

Using an underutilised analytical approach (GAMLSS), we present empirical evidence to support the idea that risk factors can relate to sizable differences in outcome variability, and even outcome skewness, in addition to differences in the outcome mean. Females had higher variability in BMI and mental wellbeing than males; lower social class and physical inactivity were each associated with higher variability in both BMI and mental wellbeing, despite having different directions of association with the mean (higher BMI yet lower mental wellbeing).

Our findings add to an emerging literature which has investigated associations between risk factors and outcome variability. Studies[11-17] have reported that risk factors associated with higher means are also associated with higher outcome variability. For example, Beyerlein et al (2008)[22] found that multiple risk factors for high childhood BMI (such as more frequent television viewing and greater rapid infant weight gain) were related to both higher mean BMI and greater variability in BMI. However, previous studies have not utilised multiple outcomes or nationally representative samples, and have not systematically considered explanations for such findings or their implications.

Our findings help to reconcile findings from GAMLSS with those using quantile regression[22 26 27] which have reported stronger effect sizes for BMI risk factors at higher BMI centiles. This finding is both consistent with and helps explain the GAMLSS findings. For instance, lower social class and physical inactivity are related to higher BMI mean and variability, yet less BMI skewness; the net result is higher effect estimates at upper centiles which are less precisely estimated, as seen in quantile regression. While both analytical approaches have merit, GAMLSS has a number of attractive features for use in aetiological research: it enables each distribution moment to be separately investigated, and uses predetermined distribution families which enable computation of sparsely distributed variables.

Why are risk factors associated with differences in outcome variability? There are multiple possible explanations. First, risk factors may not be sufficient for an outcome to occur but rather only have a causal effect in the presence of other factors, for instance as posited in models such as the *stress-diathesis* model of mental health.[44] Such additional factors could also operate as effect modifiers which increase the strength of the risk factor. Factors such as genetic propensity to weight gain may for example modify the effect on weight gain of exposure to adverse socioeconomic circumstances.[45] Other environmental factors could operate similarly—such that the association between lower social class and higher BMI is weaker amongst those living in a local environment which is less 'obesogenic' (i.e., less conducive to physical inactivity and lower energy intake).[46 47] The net result of such divergent effects would be increased variability since the effects would range from zero to the upper bound of the effect. This explanation may also apply to mental wellbeing, given evidence for

the myriad environmental[48 31] and genetic determinants[49 50] which could modify the effects observed in the current study.

Alternatively, between-person differences in confounding and/or measurement error may also lead to risk factors being associated with outcome variability. For example, in the present study physical activity was measured via a single item capturing reported activity of a moderate-vigorous intensity for at least 30 minutes per day; this is an imperfect reflection of the underlying exposure which may have a causal effect (e.g., total energy expenditure (across all intensities of activity) in the case of adiposity;[51] or time spent in specific activities conducive to wellbeing in the case of mental wellbeing[35]). The net result would be higher variability in those reporting higher physical activity levels. A related issue is the extent to which the exposure captures the same 'dose' across participants in a given study. The physical activity measure used here counted the number of days that bouts of activity lasted at least 30 minutes; this likely reflects substantial variability in the level of exercise actually undertaken, thus leading to greater differences in outcome variability. This could partly explain the associations of lower social class with greater outcome variability, since social class is one dimension of socioeconomic position, such that there may be substantial between-person variation in other dimensions (eg, parental education, income and/or wealth[38 39]) which may each influence outcomes, leading to greater variability.

The study highlights the fact that analyses by GAMLSS and quantile regression lead to similar results at the selected quantiles of the outcome distribution—see Figures 3 and 4. However GAMLSS, by analysing the whole distribution, can in some cases provide more efficient estimates of the quantiles. Compare for example the standard errors of the median as obtained by the BCCG distribution (Tables 2 and 3) and quantile regression (Table 4); for BMI the standard errors of around 0.5 are broadly similar the two ways, but for WEMWBS the GAMLSS standard errors are appreciably smaller.

Strengths and limitations

Strengths of this study include the analytical approach used (GAMLSS) to empirically investigate differences in outcome variability. While differences in variability can be informed by descriptive comparison (e.g., comparing standard deviations), GAMLSS additionally enables computation of estimates of precision and incorporates multivariable specifications (e.g., confounder or mediator adjustment; and inclusion of interaction terms). The use of the 1970 birth cohort data is an additional strength, enabling investigation of multiple risk factors and two largely orthogonal yet important continuous health outcomes. The national representation of this cohort is also advantageous—highly distorted sample selection can bias conventional epidemiological results (i.e. mean differences in outcomes),[54] and may also bias comparisons of outcome variability.

The study also has limitations. As in all observational studies, causal inference is challenging despite the use of longitudinal data. Associations of social class at birth with outcomes for example could be explained by unmeasured confounding—this may include factors such as parental mental health. This is challenging to falsify empirically owing to a lack of such data collected before birth. In contrast, sex is randomly assigned at birth, and thus its associations with outcomes are unlikely to be confounded. However, sex differences in reporting may bias associations with mental wellbeing. Physical activity and mental wellbeing were ascertained at broadly the same age, so that associations between the two could be explained by reverse causality; existing evidence appears to suggest bi-directionality of links between physical activity and both outcomes.[37 55] Finally, attrition led to lower power to precisely estimate smaller effect sizes (e.g. gender differences in mental wellbeing) or confirm null effects. Such attribution could potentially bias associations—those in worse health and adverse socioeconomic circumstances are disproportionately lost to follow-up.[56 57] The focus of principled approaches to handle missing data in epidemiology has been on the main parameter of interest—typically beta coefficients in linear regression models—and further empirical work is required to investigate the potential implications of (non-random) missingness for the variability and other moments of the outcome distribution.

Potential implications

This study used an underutilised approach to empirically investigate associations between risk factors and outcome variability in a single cohort study. Thus, our findings require replication and extension in other datasets across other risk factors and health outcomes. Future studies should also seek to explain their findings, and where possible falsify potential explanations. Understanding how risk factors relate to and/or cause differences in outcome variability is not a standard part of epidemiological training, and it entails additional analytical and conceptual complexity. Thus, with greater application of these tools an emerging consensus on best practice should develop. In the first instance we recommend both descriptive and formal investigation, and that analysts carefully consider the use of both absolute (e.g., SD) and relative (e.g., CoV) differences in variability. Since the CoV is fractional standard deviation (eg, SD/mean or log SD), its suitability of use depends on the *a priori* anticipated relationship between the mean and variance.

In the context of randomised controlled trials, the finding of variability in treatment effects between individuals has been used to justify individualised approaches to treatment (personalised medicine). It is beyond the scope of the current article to discuss the tractability of this for complex outcomes in which treatment effects are unpredictable.[58] Trials are designed typically to detect only mean differences in outcomes;[59] nevertheless, additionally presenting outcome variability before and after treatment would be helpful to better appraise intervention effects.[5] GAMLSS provides a useful framework with which to formally investigate this, even where the homoscedasticity assumption does

not hold (i.e., where risk factors or treatment groups differ in their outcome variance). Where there are multiple potential efficacious interventions, further studies could meta-analyse existing trials to identify the types of intervention which additionally reduce outcome variability.

Conclusion

We provide empirical support for the notion that risk factors or interventions can either reduce or increase variability in health outcomes. This finding is consistent with results from quantile regression analysis where a risk factor vs outcome association is stronger (or weaker) at higher outcome centiles. Such findings may be explained by heterogeneity in the causal effect of each exposure, by the influence of other (typically unmeasured) variables, and/or by measurement error. This underutilised approach to the analysis of continuously distributed outcomes may have broader utility in epidemiological, medical, and psychological sciences. Our tutorial and syntax content is designed to facilitate this.

Data availability

Available from the UK Data Archive:

https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200001

Legends

Figure 1. Simulated data for three interventions each having the same effect on the mean, but different effects on the variability (middle panel) and skewness (bottom panel).

Figure 1: Kernel density plots for body mass index and mental wellbeing, stratified by risk factor group. Note: CoV = coefficient of variation (SD/mean).

Figure 3. Association between risk factors and BMI by BMI centile. Plotted lines are calculated using GAMLSS estimation results of the entire outcome distribution; points at the 25^{th}, 50^{th}, and 75^{th} centiles are estimated using quantile regression models. Marginal effects show the differences in outcome between each risk group across the outcome distribution.

Figure 4. Association between risk factors and BMI by BMI centile. Plotted lines are calculated using GAMLSS estimation results of the entire outcome distribution; points at the 25^{th}, 50^{th}, and 75^{th} centiles are estimated using quantile regression models. Marginal effects show the differences in outcome between each risk group across the outcome distribution.

Supplementary File 1a. Risk factors in relation to body mass index (BMI): differences in mean, variability and skewness estimated by GAMLSS

Supplementary File 1b. Risk factors in relation to mental wellbeing (WEMWEBS): differences in mean, variability and skewness estimated by GAMLSS

453

454

455

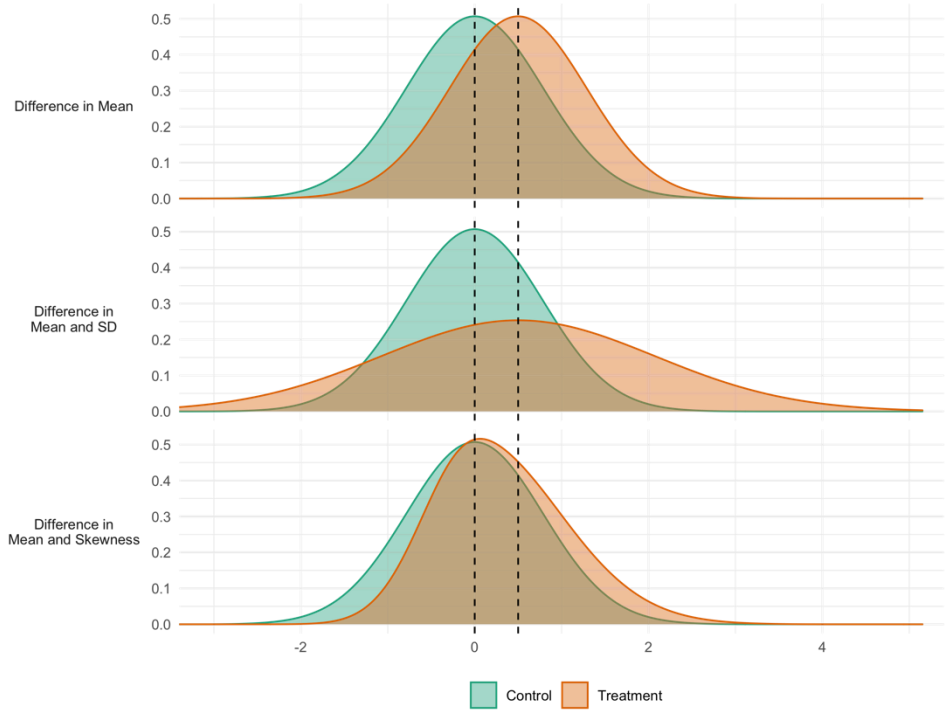


Figure 2. Simulated data for three interventions each having the same effect on the mean, but different effects on the variability (middle panel) and skewness (bottom panel)

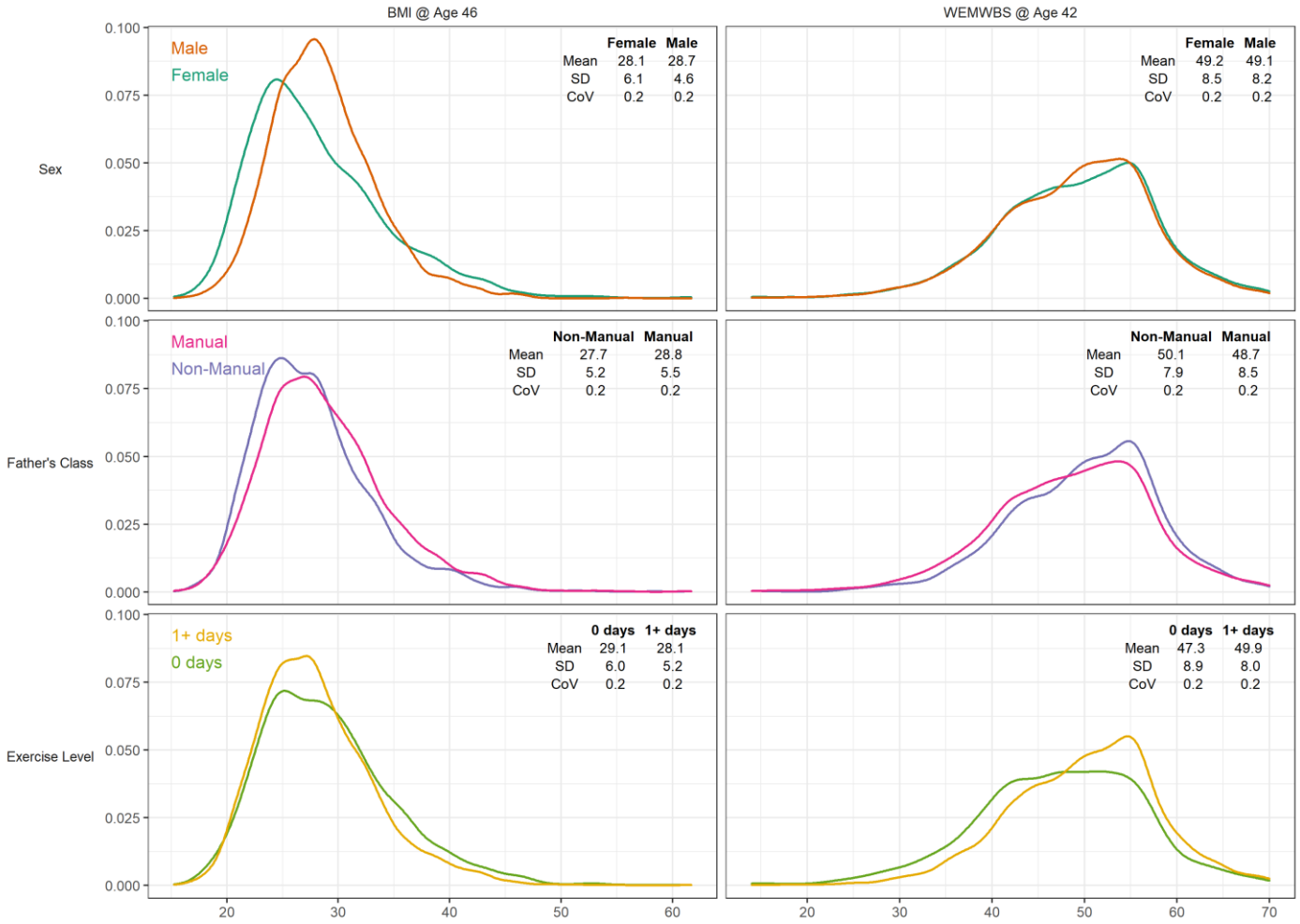


Figure 3: Kernel density plots for body mass index and mental wellbeing, stratified by risk factor group. Note: CoV = coefficient of variation (SD/mean)

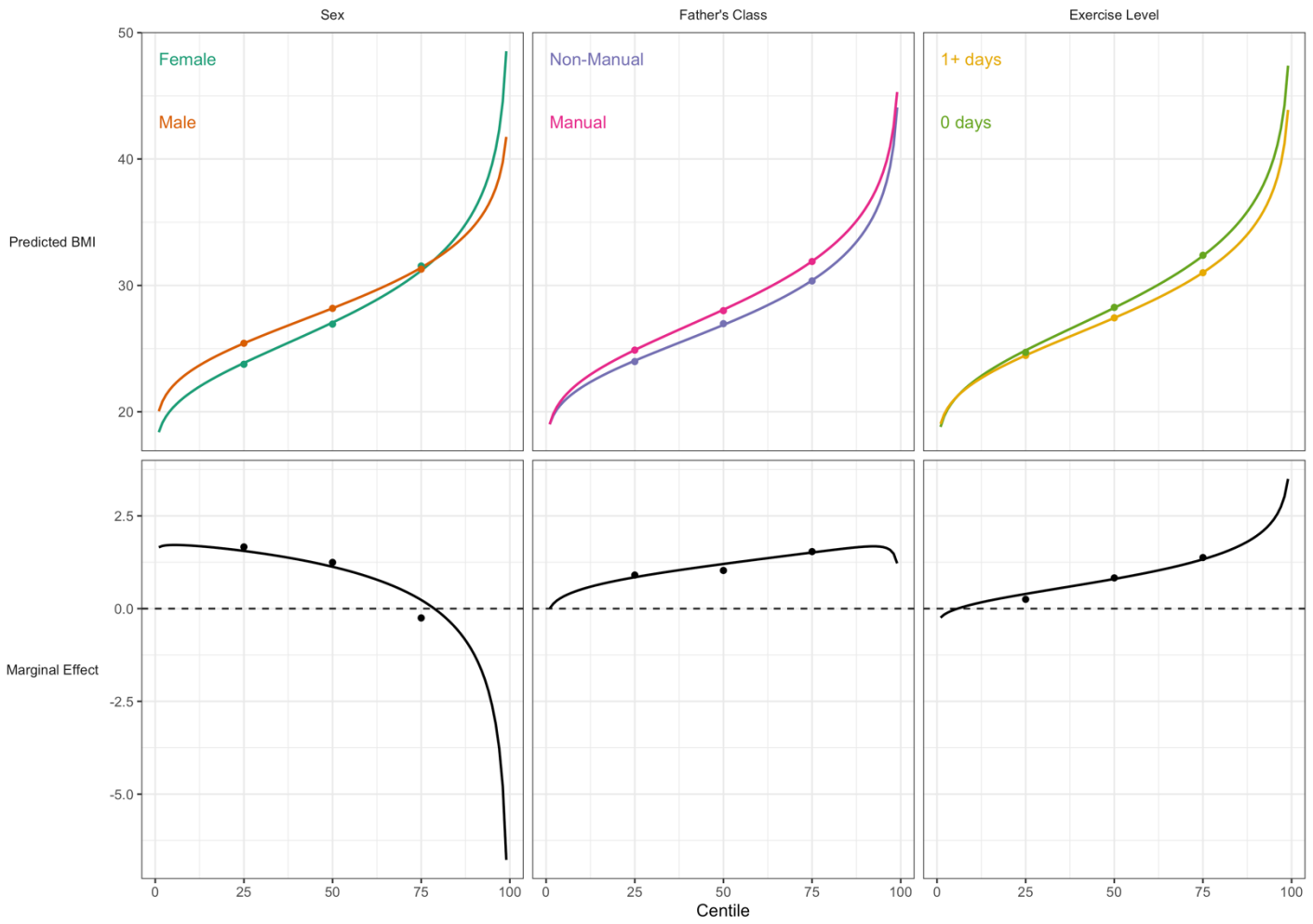


Figure 3. Association between risk factors and BMI by BMI centile. Plotted lines are calculated using GAMLSS estimation results of the entire outcome distribution; points at the 25th, 50th, and 75th centiles are estimated using quantile regression models. Marginal effects show the differences in outcome between each risk group across the outcome distribution.

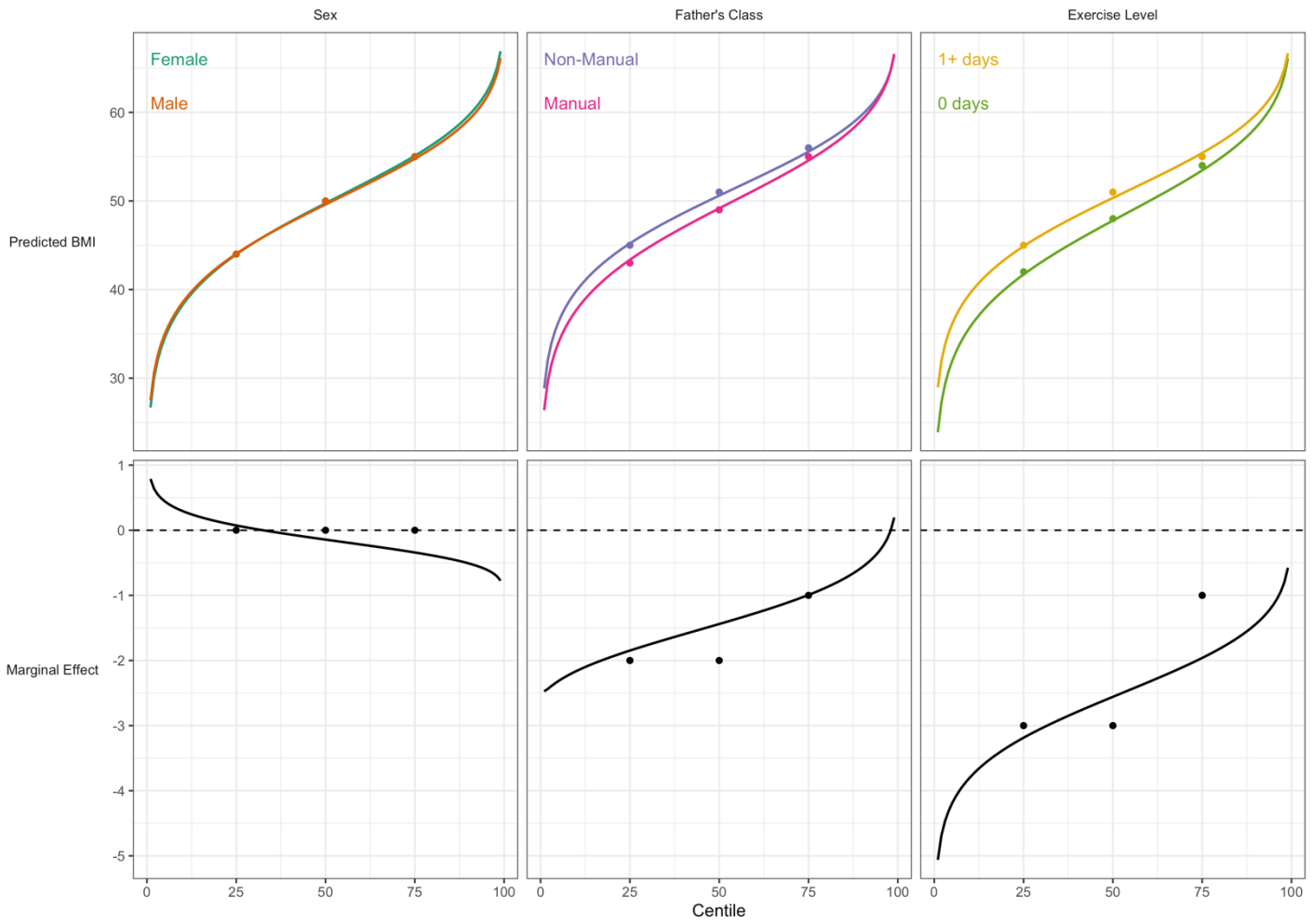


Figure 4. Association between risk factors and BMI by BMI centile. Plotted lines are calculated using GAMLSS estimation results of the entire outcome distribution; points at the 25th, 50th, and 75th centiles are estimated using quantile regression models. Marginal effects show the differences in outcome between each risk group across the outcome distribution.

Table 1. Risk factors in relation to body mass index: differences in mean, variability and skewness estimated by GAMLSS (n = 6,007)

		NO distribution		BCCG distribution		
Risk factor	%	Mean	SD	Median	CoV	Skewness*
Female (ref)	52.4%	28.1	6.1	26.9	0.22	1.10
Male	47.6%	28.7	4.6	28.2	0.16	0.75
Unadjusted difference, % (SE)		1.9 (0.5)	-27.6 (1.8)	4.1 (0.4)	-23 (1.8)	0.48 (0.11)
Adjusted[#] difference, % (SE)		2.2 (0.5)	-27.4 (1.8)	4.4 (0.4)	-22.6 (1.8)	0.54 (0.11)
Non-manual (ref)	36.3%	27.7	5.2	27	0.19	1.15
Manual social class	63.7%	28.8	5.5	28	0.19	0.90
Unadjusted difference, % (SE)		4.0 (0.5)	6.1 (1.9)	4.4 (0.5)	6 (1.9)	0.39 (0.11)
Adjusted[#] difference, % (SE)		3.8 (0.5)	5.5 (1.9)	4.3 (0.4)	5.6 (1.9)	0.40 (0.12)
Physically active (ref)	73%	28.1	5.2	27.4	0.19	0.97
Inactive	27%	29.1	6.0	28.3	0.21	0.94
Unadjusted difference, % (SE)		3.3 (0.6)	13.5 (2.1)	2.9 (0.5)	10.4 (2.1)	0.08 (0.12)
Adjusted[#] difference, % (SE)		3.3 (0.6)	12.1 (2.1)	3.1 (0.5)	9.3 (2.1)	0.12 (0.12)

[#]Estimates mutually adjusted for sex, social class and physical inactivity.

*Skewness is estimated as the Box-Cox power (that is, the power required to transform the outcome to a normal distribution); differences are the absolute difference in Box-Cox power in each subgroup estimated by GAMLSS. GAMLSS estimates multiple distribution moments simultaneously; thus, differences may not exactly correspond to descriptive comparisons reported above.

NO: normal distribution; BCCG: Box-Cox Cole and Green distribution: SD: standard deviation; CoV: coefficient of variation; GAMLSS: Generalized Additive Models for Location, Scale and Shape; SE, standard error.

Table 2. Risk factors in relation to mental wellbeing (WEMWBS): differences in mean, variability and skewness estimated by GAMLSS (n = 7,104)

		NO distribution		BCCG distribution		
Risk factor	%	Mean	SD	Median	COV	Skewness*
Female (ref)	52.8%	49.2	8.5	50	0.17	-0.41
Male	47.2%	49.1	8.2	50	0.17	-0.40
Unadjusted difference, % (SE)		-0.2 (0.4)	-3.9 (1.7)	-0.3 (0.4)	-3.5 (1.7)	0.02 (0.11)
Adjusted[#] difference, % (SE)		-0.6 (0.4)	-3.6 (1.7)	-0.7 (0.4)	-2.6 (1.7)	0.00 (0.11)
Non-manual (ref)	34.8%	50.1	7.9	51	0.16	-0.45
Manual social class	65.2%	48.7	8.5	49	0.17	-0.37
Unadjusted difference, % (SE)		-2.8 (0.4)	7.2 (1.8)	-2.9 (0.4)	10.9 (1.8)	-0.20 (0.12)
Adjusted[#] difference, % (SE)		-2.5 (0.4)	6.0 (1.8)	-2.7 (0.4)	9.8 (1.8)	-0.24 (0.12)
Physically active (ref)	72.4%	49.9	8.0	51	0.16	-0.38
Inactive	27.6%	47.3	8.9	48	0.19	-0.36
Unadjusted difference, % (SE)		-5.3 (0.5)	10.9 (1.9)	-5.2 (0.4)	16.2 (1.9)	-0.12 (0.12)
Adjusted[#] difference, % (SE)		-5.3 (0.5)	9.9 (1.9)	-5.1 (0.4)	15.2 (1.9)	-0.10 (0.12)

[#]Estimates mutually adjusted for sex, social class and physical inactivity.

*Skewness is estimated as the Box-Cox power (that is, the power required to transform the outcome to a normal distribution); differences are the absolute difference in Box-Cox power in each subgroup estimated by GAMLSS. GAMLSS estimates multiple distribution moments simultaneously; thus, differences may not exactly correspond to descriptive comparisons reported above.

NO: normal distribution; BCCG: Box-Cox Cole and Green distribution: SD: standard deviation; CoV: coefficient of variation; GAMLSS: Generalized Additive Models for Location, Scale and Shape; SE, standard error.

Table 3. Risk factors in relation to body mass index (BMI) and mental wellbeing (WEMWBS): percentage differences at multiple points of the outcome distribution estimated by quantile regression

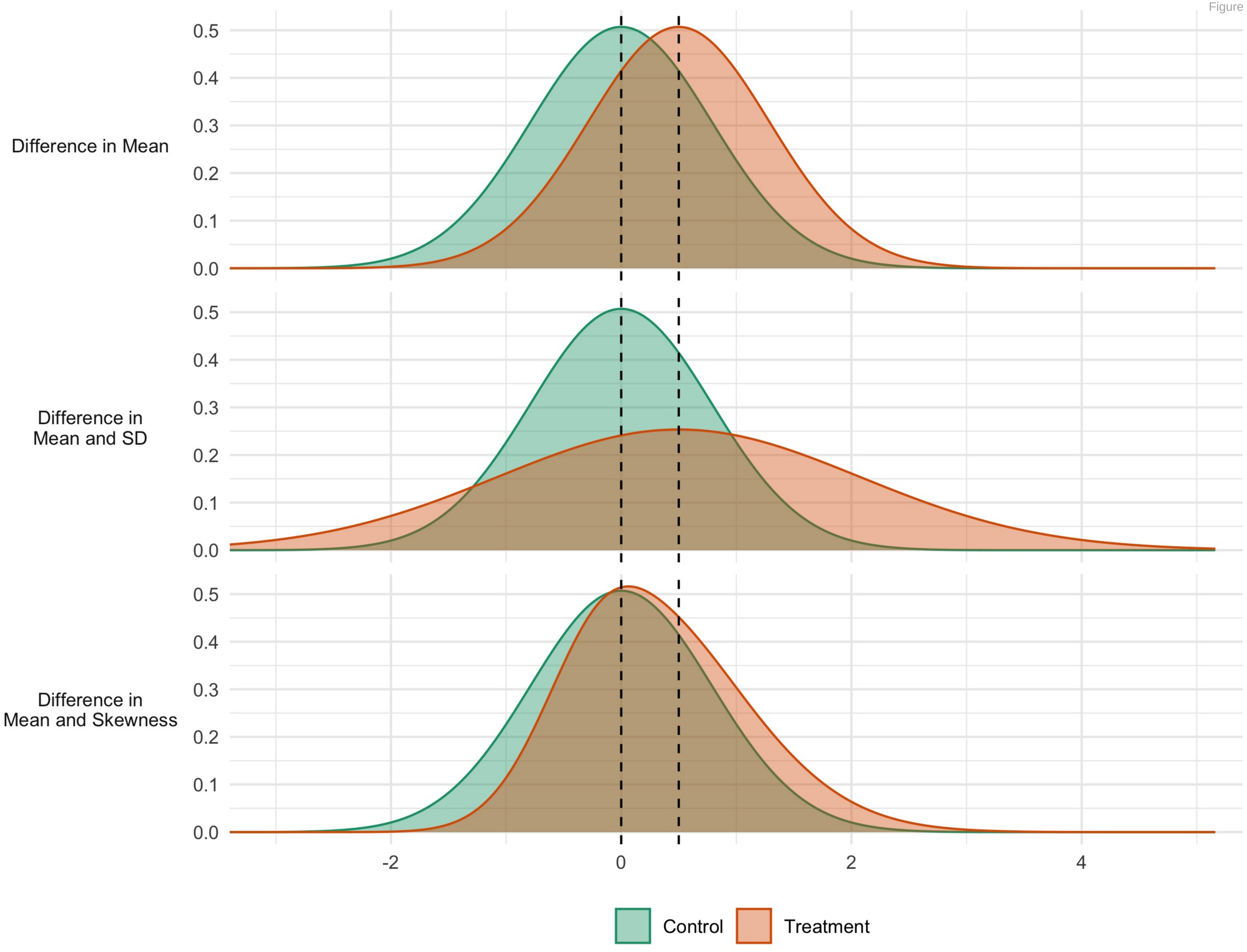
Outcome	Risk Factor	25th centile	50th centile	75th centile
BMI @ Age 46	Male vs female	6.8 (0.5)	4.5 (0.6)	-0.8 (0.7)
	Father's Class	3.7 (0.6)	3.7 (0.6)	4.9 (0.7)
	Exercise Level	1 (0.7)	3 (0.7)	4.3 (0.8)
WEMWBS @ Age 42	Sex	0 (0.7)	0 (0.5)	0 (0.3)
	Father's Class	-4.5 (0.7)	-4 (0.5)	-1.8 (0.3)
	Exercise Level	-6.9 (0.5)	-6.1 (0.5)	-1.8 (0.5)

Note: results show the percentage difference (log-transformed x 100) in BMI or mental wellbeing (WEMWEBS; standard errors in parenthesis) at different centiles of the outcome distribution; estimates are mutually adjusted.

References

1. Keyes KM, Galea S. Population health science: Oxford University Press 2016.
2. Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nature reviews genetics* 2009;10(12):872-78.
3. Keyes CL. The mental health continuum: From languishing to flourishing in life. *J Health Soc Behav* 2002:207-22.
4. Subramanian S, Kim R, Christakis NA. The" average" treatment effect: A construct ripe for retirement. A commentary on Deaton and Cartwright. *Soc Sci Med* 2018
5. Vitale S, Cotch MF, Sperduto RD. Prevalence of visual impairment in the United States. *JAMA* 2006;295(18):2158-63.
6. Truby H, Baic S, DeLooy A, et al. Randomised controlled trial of four commercial weight loss programmes in the UK: initial findings from the BBC “diet trials”. *BMJ* 2006;332(7553):1309-14.
7. Porta M. A dictionary of epidemiology. Oxford, UK: Oxford University Press 2008.
8. Beyerlein A, Toschke AM, Von Kries R. Breastfeeding and childhood obesity: shift of the entire BMI distribution or only the upper parts? *Obesity* 2008;16(12):2730-33.
9. Flegal KM, Troiano RP. Changes in the distribution of body mass index of adults and children in the US population. *Int J Obes* 2000;24(7):807.
10. Johnson W, Li L, Kuh D, et al. How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts. *PLoS Med* 2015;12(5):e1001828.
11. Sun J, Covaci A, Bustnes JO, et al. Temporal trends of legacy organochlorines in different white-tailed eagle (Haliaeetus albicilla) subpopulations: A retrospective investigation using archived feathers. *Environ Int* 2020;138:105618.
12. Nakagawa S, Poulin R, Mengersen K, et al. Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods Ecol Evol* 2015;6(2):143-52.
13. Pitt D, Trück S, van den Honert R, et al. Modeling risks from natural hazards with generalized additive models for location, scale and shape. *J Environ Manage* 2020;275:111075.
14. Hohberg M, Pütz P, Kneib T. Treatment effects beyond the mean using distributional regression: Methods and guidance. *PLoS One* 2020;15(2):e0226514.
15. Silbersdorff A, Schneider KS. Distributional Regression Techniques in Socioeconomic Research on the Inequality of Health with an Application on the Relationship between Mental Health and Income. *Int J Environ Res Public Health* 2019;16(20):4009.
16. Silbersdorff A, Lynch J, Klasen S, et al. Reconsidering the income-health relationship using distributional regression. *Health Econ* 2018;27(7):1074-88.
17. Beyerlein A, Fahrmeir L, Mansmann U, et al. Alternative regression models to assess increase in childhood BMI. *BMC Med Res Methodol* 2008;8(1):59.
18. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2005;54(3):507-54.
19. Cole T, Stanojevic S, Stocks J, et al. Age-and size-related reference ranges: A case study of spirometry through childhood and adulthood. *Stat Med* 2009;28(5):880-98.
20. Beyerlein A. Quantile regression—opportunities and challenges from a user's perspective. *Am J Epidemiol* 2014;180(3):330-31.
21. Bann D, Fitzsimons E, Johnson W. Determinants of the population health distribution: an illustration examining body mass index. *Int J Epidemiol* 2020;49(3):731–37.
22. Green MA, Rowe F. Explaining the widening distribution of Body Mass Index: A decomposition analysis of trends for England, 2002–2004 and 2012–2014. *Area* 2020
23. Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status and the 25× 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1· 7 million men and women. *Lancet* 2017;389(10075):1229-37.

24. Elliott J, Shepherd P. Cohort profile: 1970 British birth cohort (BCS70). *Int J Epidemiol* 2006;35(4):836-43.
25. Wood N, Hardy R, Bann D, et al. Childhood correlates of adult positive mental well-being in three British longitudinal studies. *J Epidemiol Community Health* 2021;75(2):177-84.
26. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health and Quality of life Outcomes* 2007;5(1):63.
27. Lewontin RC. On the measurement of relative variability. *Syst Zool* 1966;15(2):141-42.
28. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;11(10):1305-19.
29. Stasinopoulos DM, Rigby RA. Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* 2007;23(7):1-46.
30. Tyrrell J, Wood AR, Ames RM, et al. Gene–obesogenic environment interactions in the UK Biobank study. *Int J Epidemiol* 2017:dyw337.
31. Drewnowski A, Rehm CD, Solet D. Disparities in obesity rates: analysis by ZIP code area. *Soc Sci Med* 2007;65(12):2458-63.
32. Stafford M, Cummins S, Ellaway A, et al. Pathways to obesity: Identifying local, modifiable determinants of physical activity and diet. *Soc Sci Med* 2007;65(9):1882-97.
33. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. *Science* 2012;337(6101):1505-10.
34. Luciano M, Hagenaars SP, Davies G, et al. Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat Genet* 2018;50(1):6-11.
35. De Moor MH, Van Den Berg SM, Verweij KJ, et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA psychiatry* 2015;72(7):642-50.
36. Bann D, Kuh D, Wills AK, et al. Physical activity across adulthood in relation to fat and lean mass in early old age: findings from the MRC National Survey of Health and Development, 1946-2010. *Am J Epidemiol* 2014;179(10):1197-207.
37. Black SV, Cooper R, Martin KR, et al. Physical activity and mental well-being in a cohort aged 60–64 years. *Am J Prev Med* 2015;49(2):172-80.
38. Moulton V, Goodman A, Nasim B, et al. Parental Wealth and Children's Cognitive Ability, Mental, and Physical Health: Evidence From the UK Millennium Cohort Study. *Child Dev* 2021;92(1):115-23.
39. Galobardes B, Shaw M, Lawlor DA, et al. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006;60(1):7-12.
40. Munafò MR, Tilling K, Taylor AE, et al. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol* 2018;47(1):226-35.
41. Pereira SMP, Geoffroy M-C, Power C. Depressive symptoms and physical activity during 3 decades in adult life: bidirectional associations in a prospective cohort study. *JAMA psychiatry* 2014;71(12):1373-80.
42. Gibbs BB, Aaby D, Siddique J, et al. Bidirectional 10-year associations of accelerometer-measured sedentary behavior and activity categories with weight among middle-aged adults. *Int J Obes* 2020;44(3):559-67.
43. Mostafa T, Wiggins D. The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies. *Longitudinal and Life Course Studies* 2015;6(2):131-46.
44. Mostafa T, Narayanan M, Pongiglione B, et al. Improving the plausibility of the missing at random assumption in the 1958 British birth cohort: A pragmatic data driven approach. 2020
45. Smith GD. Epidemiology, epigenetics and the 'Gloomy Prospect': embracing randomness in population health research and practice. *Int J Epidemiol* 2011;40(3):537-62.
46. Senn S. Mastering variation: variance components and personalised medicine. *Stat Med* 2016;35(7):966-77.


Difference in Mean
Difference in Mean and SD
Difference in Mean and Skewness
0.5
0.4
0.3
0.2
0.1
0.0
-2
0
2
4
Control
Treatment

BMI @ Age 46

WEMWBS @ Age 42

Sex

Father's Class

Exercise Level

Male
Female

	Female	Male
Mean	28.1	28.7
SD	6.1	4.6
CoV	0.2	0.2

	Female	Male
Mean	49.2	49.1
SD	8.5	8.2
CoV	0.2	0.2

Manual
Non-Manual

	Non-Manual	Manual
Mean	27.7	28.8
SD	5.2	5.5
CoV	0.2	0.2

	Non-Manual	Manual
Mean	50.1	48.7
SD	7.9	8.5
CoV	0.2	0.2

1+ days
0 days

	0 days	1+ days
Mean	29.1	28.1
SD	6.0	5.2
CoV	0.2	0.2

	0 days	1+ days
Mean	47.3	49.9
SD	8.9	8.0
CoV	0.2	0.2

0.100
0.075
0.050
0.025
0.000

20
30
40
50
60
70

Sex

Father's Class

Exercise Level

Female

Male

Non-Manual

Manual

1+ days

0 days

Predicted BMI

50

40

30

20

2.5

0.0

-2.5

-5.0

Marginal Effect

0

25

50

75

100

Centile

Sex

Father's Class

Exercise Level

Female

Male

Non-Manual

Manual

1+ days

0 days

Predicted WEMWBS

60

50

40

30

Marginal Effect

1

0

-1

-2

-3

-4

-5

0

25

50

75

100

Centile