

## **Does household income predict health and educational outcomes in childhood better than neighbourhood deprivation?**

**Ieva Skarda<sup>1</sup>, Richard Cookson<sup>1</sup>, Ruth Gilbert<sup>2</sup>**

<sup>1</sup>Centre for Health Economics, University of York, York, YO10 5DD, UK

<sup>2</sup>Population Policy and Practice Research and Teaching Department, University College, London Great Ormond Street Institute of Child Health, London WC1N 1EH, UK

Address correspondence to Ieva Skarda, E-mail: [ieva.skarda@york.ac.uk](mailto:ieva.skarda@york.ac.uk)

Ieva Skarda (PhD), Research Fellow

Richard Andrew Cookson (PhD), Professor in Health Economics

Ruth Gilbert (MD), Professor of Clinical Epidemiology

All authors contributed to the conceptualisation, methods and writing of iterative drafts. Ieva Skarda conducted the analyses, with input from Richard Cookson and Ruth Gilbert.

## **Abstract**

### *Background*

Public health research and prevention policies often use the small area Index of Multiple Deprivation (IMD) at neighbourhood level to proxy individual socio-economic status because it is readily available. We investigated what household income adds to IMD in early childhood for predicting adverse health in adolescence.

### *Methods*

Using data from the Millennium Cohort Study, we analysed IMD and self-reported equivalised household income (ages 0-5) to predict outcomes at age 17: poor academic achievement, psychological distress, poor health, smoking, and obesity. Predictions were compared using IMD quintiles alone, household income alone, and both together.

### *Results*

Household income was a stronger and more consistent predictor of age 17 outcomes than IMD and revealed inequalities within neighbourhoods. Decreasing household income showed steep gradients in educational attainment and smoking across all IMD quintiles, and moderate gradients in obesity, psychological distress, and poor health in most quintiles. IMD did not predict smoking or psychological distress within any income group, or educational attainment within the poorest income group.

### *Conclusion*

Household income is associated with inequality gradients within all quintiles of neighbourhood IMD. Early childhood public health strategies should consider household income in combination with neighbourhood deprivation.

### **Keywords**

Public health, Socioeconomic factors, Children

## Introduction

The Index of Multiple Deprivation (IMD) has been widely used as a neighbourhood measure of deprivation over the last two decades to guide UK policy on health disparities (see Table A1 in the appendix for details). The success of IMD is due, in part, to its ready availability and routine linkage to administrative health and other datasets. Furthermore, IMD consistently shows a clear and monotonic social gradient in almost all indicators of health which underpins its use as a tool for identifying need and allocating services locally and nationally.<sup>[1]</sup>

Household income is an important indicator of socio-economic status that is strongly associated with health<sup>[2-9]</sup> yet is challenging to measure and link to routine health data.<sup>[10]</sup> Hence, public health research and prevention policies often use the IMD at the neighbourhood level to proxy individual socio-economic status. For example, NHS England statistics aimed at reducing health inequalities in children and young people currently focus on the most deprived 20% of neighbourhoods (in addition to vulnerable groups identified locally, such as ethnic minorities, those with learning disability, care leavers and those with five specific health conditions).<sup>[11]</sup>

However, IMD is a blunt instrument for guiding policies to reduce health inequalities. Firstly, 55% to 62% of the poorest households, based on household income, live outside the most deprived 20% of neighbourhoods (ranked using IMD), so are likely to be missed by such policies.<sup>[12-14]</sup> Secondly, focusing on the most deprived ignores the gradient of adverse health outcomes across the whole distribution of deprivation, particularly income deprivation. Thirdly, IMD does not capture the fact that household income is highly amenable to policy interventions directly affecting income, such as on wages, benefits, wealth, housing, and costs of other essentials.

This raises the question of whether household income, grouped in a useful way that avoids risk of disclosure, could add value to existing neighbourhood measures of deprivation.

Few studies have directly assessed the added value of individual-level measures of income for predicting adverse health outcomes – such as household income or personal earnings – compared with neighbourhood-level measures of deprivation. A recent scoping review concluded that individual-level measures of social disadvantage tended to identify stronger associations with adverse child health than neighbourhood-level measures.<sup>[15]</sup> More research is needed on the advantages of combining individual-level and neighbourhood-level measures, including to address intersectional inequalities and to counter the problem of the ecological fallacy inherent in neighbourhood-level measures.

Our aim was to address the evidence gap on whether household income adds important information, separately and in combination with neighbourhood IMD, that could be used by services and policy makers to identify groups in childhood with poor long-term health outcomes who might benefit from early intervention. We used the Millennium Cohort Study (MCS), which contains parent-reported household income and IMD in the early childhood years. We explored associations between quintiles of neighbourhood IMD and household income alone, and in combination, with five health-related outcomes at age 17: poor educational attainment, psychological distress, poor health, smoking, and obesity. All five outcomes have strong evidence of long-term impacts on health in adulthood.<sup>[16]</sup> We compared

associations between these outcomes and household income and neighbourhood deprivation separately and assessed the added value of both variables combined.

## Methods

### *Data and sample*

Data for this study were derived from waves 1 to 3 and wave 7 of the MCS, a nationally representative retrospective cohort study following individuals born in the UK between September 1, 2000, and January 11, 2002. The first survey in 2001–02 included 18,819 children, with subsequent surveys at ages 3, 5, 7, 11, 14, and 17. Additional details on MCS can be found elsewhere.<sup>[17]</sup> Ethics approval was obtained through the National Health Service Research Ethics Committee system, with written informed consent from parents up to age 14 and verbal consent from those aged 16 and older.

We addressed missing data primarily due to attrition using multiple imputation with chained equations. Imputation was carried out from the age 3 survey (sweep 2), our primary imputation sweep, which included 1,389 additional families not surveyed in sweep 1.<sup>[18], [19]</sup> Our primary analysis dataset includes observations on 15,367 children after imputation. Analyses were weighted using inverse probability weights to adjust for attrition and sampling design. Further details on the derivation of our primary analysis dataset are in the appendix (section A1 and Figure A1)."

### *Measures*

Data on neighbourhood deprivation was based on the IMD 2004 Overall Deciles (ranked separately for England, Scotland, Wales and Northern Ireland) which are geographically linked to the MCS households at the lower super output level (LSOA; see table A1, appendix, for details on all variables used in the analyses). The decile group numbers at the first three waves (9 months, 3 years, 5 years) were averaged over these waves, rounded and transformed into quintile groups.

Between 2001 and 2006, questions about household income in early childhood (after tax and other deductions but before housing costs) were asked during home interviews and reported by the main parent/ caregiver at ages 9 months, 3 years, and 5 years. Imputed banded responses were converted into continuous values and equivalised based on household size and composition.<sup>[20]</sup> Average household income across the first three waves was ranked into income quintiles. Health-related outcomes at age 17 are summarized in the appendix (Table A1) and detailed elsewhere.<sup>[16]</sup>

### *Statistical analyses*

We cross-tabulated the percentage prevalence of each of the five adverse health-related outcomes at age 17 across the five IMD quintiles and the five income quintiles in early childhood, as well as the twenty-five sub-groups defined by both IMD and income quintiles, using heatmap shading to visualize the degree of outcome prevalence (Table 1). We also report how the sample children were distributed across these subgroups (appendix Figure A2 and Figure A3).

We plotted the percentage prevalence of adverse outcomes across the sub-groups in two ways: (1) sorting first by IMD quintile and then by income quintile, allowing us to observe the inequality gradient within each IMD quintile as income decreases (Figure 1), and (2) sorting

first by income quintile and then by IMD quintile (Figure 2), allowing us to observe the inequality gradient within each income quintile as IMD decreases. The aim was to determine whether the inequality gradient changes when examining within-group variation and if this varies depending on the measure used.

For each adverse outcome, we computed the Inequality Gradient Slope Index separately for IMD quintiles, income quintiles, and both combined (appendix Table A3). This index quantifies the percentage point reduction in the likelihood of the adverse outcome when moving from the most deprived to the least deprived child, based on a linear model.<sup>[21]</sup>

Next, we ran modified Poisson regressions to estimate the risk ratios (with 95% confidence intervals) for the incidence of five adverse outcomes at age 17, given the IMD quintiles and income quintiles as predictors in separate models, as well as in a model with both measures together (Table 2). The aim was to compare how well each of the measures predicted the adverse outcomes and the inequality gradient in these outcomes.

In main regression analyses, we did not include any covariates as our aim is to inform government and researchers of the added value of household income in addition to IMD in the context of adverse adolescent health-related outcome prevalence, regardless of other measures available. Adjusting for covariates could potentially alter and likely underestimate any association with IMD or household income.<sup>[22], [23], [24]</sup> However, in sensitivity analyse, we analysed the extent to which other covariates (child's sex, eligibility for free school meals, single parent status, maternal age and number of siblings) routinely available in some administrative datasets such as ECHILD or birth registrations, capture the socioeconomic disadvantage reflected by the IMD and household income (appendix Table A4).

Finally, we plotted the receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) for the Poisson regression models predicting adverse outcomes at age 17, considering the IMD quintile group, the income quintile group, and both predictors together (appendix Figure A4, Figure A5). ROC curves quantify the diagnostic accuracy of prediction models, allowing comparison across multiple models.

The multiple imputation and regression analyses were done using STATA (version 18). The cross tabulation, figures and ROC curve plots were done using R (version R-4.4.0).

## Results

Of the 15,367 adolescents aged 17 years, 7,822 (50.9%) were male and 7,545 (49.1%) were female. The average weekly equivalised household income in nominal prices was £344.3. For around 29% of the sample in early childhood, their IMD quintile coincided with their income quintile. Of those with lowest 20% income, only 52% lived in the lowest IMD quintile, and 23% lived in the middle to richest areas (appendix Figure A3).

Overall, 5,655 (36.8%) of the adolescents achieved poor academic outcomes, 2,351 (15.3%) experienced psychological distress, 1,214 (7.9%) self-reported poor health, 1,583 (10.3%) were regular smokers, and 2,873 (18.7%) were obese (appendix Table A2).

Poor academic achievement was most prevalent in the lowest IMD quintile (54.2%, 95% CI 52.2-56.1; Table 1) and the poorest income quintile (62.6%, 95% CI 60.4-64.7; Table 1). A

steep inequality gradient was found with much lower prevalence of poor academic achievement in the least deprived IMD quintile (18.4%, 95% CI 16.4-20.5; Table 1) and richest income quintile (13.9%, 95% CI 12.3-15.4; Table 1).

The prevalence of all adverse outcomes was characterized by apparent inequality gradients in income— i.e., as one moves from the richest to the poorest quintile, the prevalence of adverse outcomes increases (Figure 1). In particular, poor academic achievement, followed by smoking, exhibited the steepest income inequality gradients consistently across all IMD quintile groups. Poor health also showed consistent inequality gradients in income. The income gradients for obesity and psychological distress were more variable, showing moderate gradients within most IMD quintile groups.

Adverse outcomes exhibited moderate to no inequality gradient in neighbourhood IMD within each income quintile (Figure 2). For example, the prevalence of poor academic achievement showed a moderate inequality gradient in IMD across the three middle-income quintiles, indicating that both income and neighbourhood deprivation contribute to poor academic achievement. However, children in the poorest income quintile demonstrated similarly poor attainment regardless of whether they resided in the least or most deprived neighbourhoods (61.9%, 95% CI 59.4-64.5 vs. 61.9%, 95% CI 45.1-78.7; Table 1). Those in the highest income quintile experienced the lowest rates of poor attainment in all IMD groups, with minimal variation within the high-income quintile according to neighbourhood deprivation levels.

The inequality gradients in IMD for psychological distress and poor health varied from weak to no gradient, particularly among those with the lowest 20% income. However, for obesity, neighbourhood IMD was associated with a moderate gradient of declining obesity among children in the three highest quintiles of household income who live in less deprived neighbourhoods. Smoking appears to relate to low income rather than neighbourhood quintile.

The slope index of inequality was higher across income quintiles than across IMD quintiles (appendix Table A3). For example, based on this index, we estimated a 61.1 percentage point reduction in the probability of poor academic achievement when moving from the poorest to the richest child in terms of household income; and a 44.1 percentage point reduction when moving from the most deprived to the least deprived child according to IMD. Similarly, we estimated an 8.3 percentage point reduction in the probability of psychological distress when moving from the poorest to the richest child; and a 4.4 percentage point reduction when moving from the most deprived to the least deprived child.

The regression analyses indicated that income was a stronger predictor of adverse outcomes than IMD, for all outcomes except for obesity (Table 2). The coefficients for income were higher than for IMD in models using each exposure alone and in the combined model. Coefficients for IMD were no longer significant at the 5% level after including income when predicting psychological distress and smoking, and were reduced, but still significant when predicting poor educational attainment, poor health and obesity. Income and IMD had effects of a similar magnitude for obesity, and IMD had a stronger effect in the combined model. These findings suggest independent effects of household income and neighbourhood IMD on poor educational attainment, poor health and obesity, reflecting the added value of including both measures. All models that included income consistently predicted more frequent adverse outcomes in the poorest subgroups.

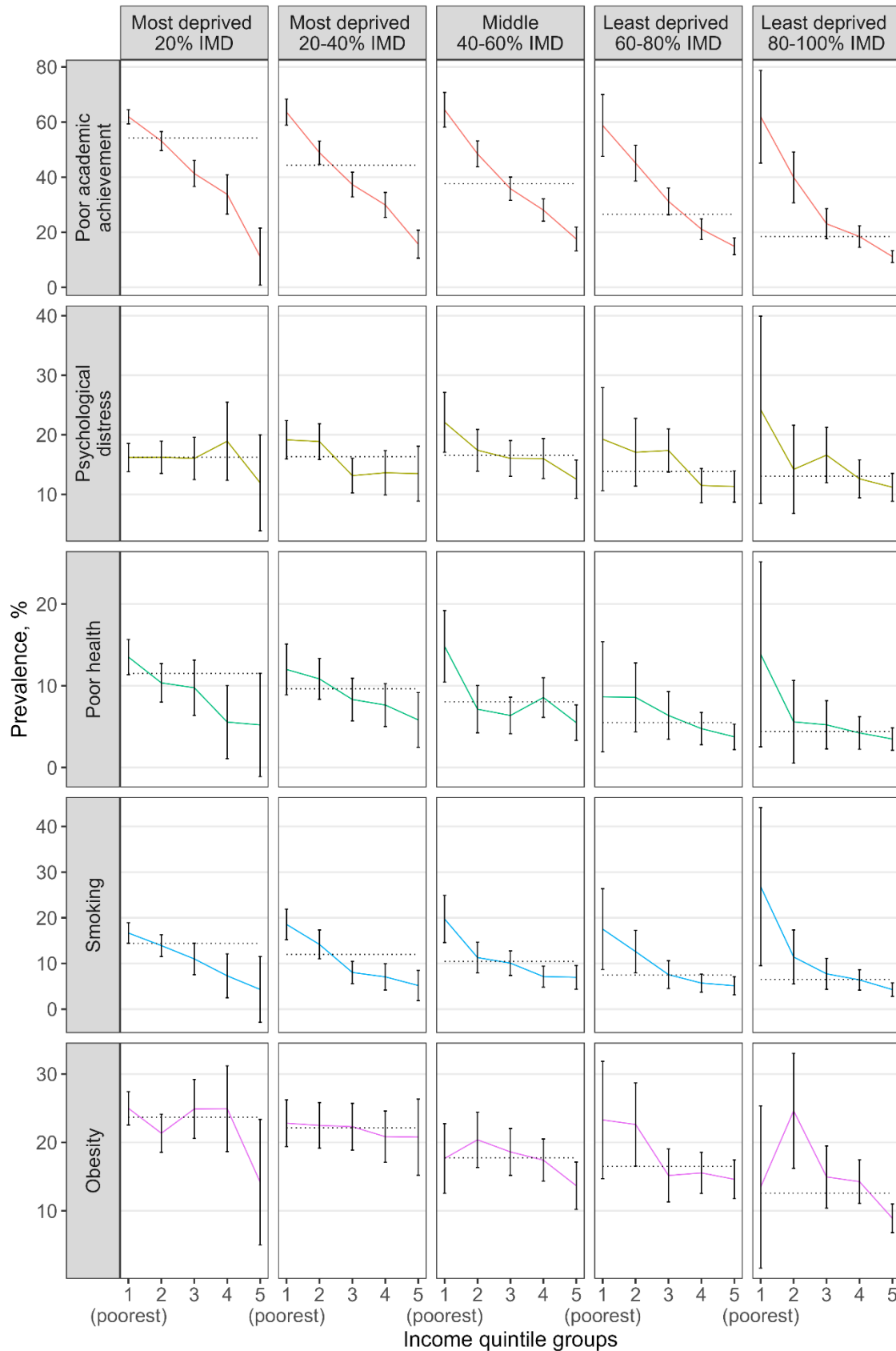
The finding of income being a stronger predictor than IMD for all outcomes except for obesity was supported by our sensitivity analyses.

First, the ROC curve analysis (appendix Figure A4) shows that the combined model (IMD & income) consistently returns the highest area under curve (AUC) when predicting the incidence of all five adverse outcomes, so both income and IMD have the highest predictive power when used together. However, income is the most important predictor: moving from a model with IMD quintiles as the sole predictors to a model with income quintiles as the predictors provides the highest increase in AUC (statistically significant at 95% level)<sup>[25]</sup>, particularly, when predicting the incidence of poor academic achievement, poor health, and smoking. This implies that using income instead of IMD as a predictor for these outcomes substantially increases the sensitivity while maintaining the specificity of the model.

Our regression analyses adjusted for the child's sex and socioeconomic characteristics (appendix Table A4), and the corresponding ROC curves (appendix Figure A4), found that household income during early childhood remained a relatively stronger predictor of adverse outcomes at age 17 and inequality in these outcomes.



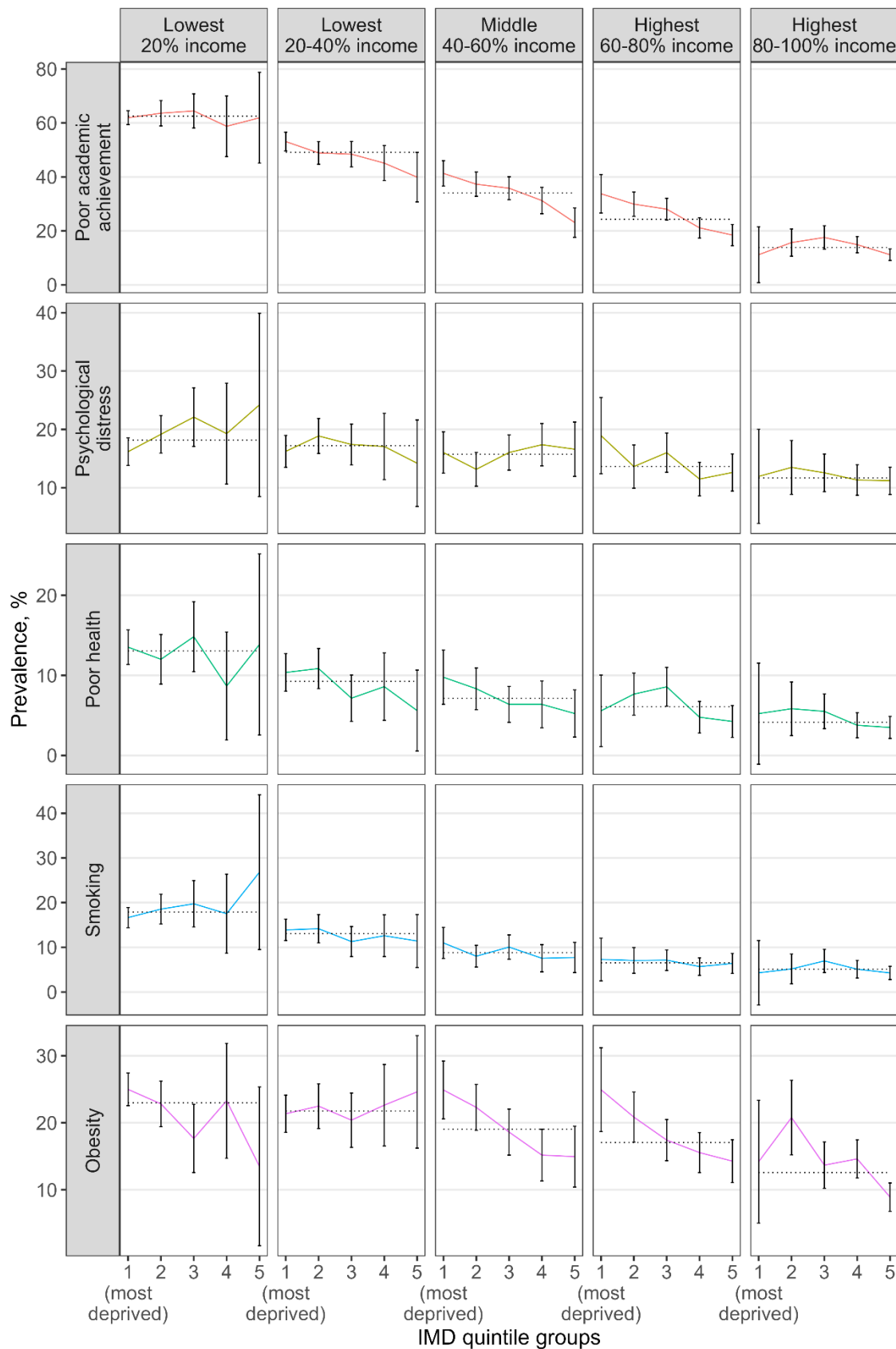
**Figure 1: Adverse outcome prevalence at age 17 in each of the IMD-income quintile subgroups**



*Note: The coloured lines depict outcome incidence across the subgroups defined by IMD and income quintile groups; the dashed lines represent outcome incidence in each IMD quintile group.*



**Figure 2: Adverse outcome prevalence at age 17 in each of the income-IMD quintile subgroups**



*Note: The coloured lines depict outcome incidence across the subgroups defined by income and IMD quintile groups; the dashed lines represent outcome incidence in each income quintile group.*

**Table 1. Crosstabulation heatmap of incidence of adverse adolescent outcomes at age 17 in the different IMD and income quintile subgroups**

	Most deprived 20% IMD	Most deprived 20-40% IMD	Middle 40-60% IMD	Least deprived 60-80% IMD	Least deprived 80-100% IMD	Total
<i>Poor academic achievement, %</i>						
Poorest 20% income	61.9 [59.4-64.5]	63.6 [58.9-68.3]	64.5 [58.2-70.8]	58.8 [47.5-70]	61.9 [45.1-78.7]	62.6 [60.4-64.7]
Poorest 20-40% income	53.1 [49.7-56.5]	48.9 [44.7-53.1]	48.4 [43.7-53.1]	45.1 [38.6-51.6]	39.9 [30.7-49.1]	49.1 [47-51.2]
Middle 40-60% income	41.3 [36.6-46]	37.3 [32.8-41.8]	35.8 [31.6-40.1]	31.2 [26.4-36.1]	23.1 [17.6-28.5]	34 [32-36.1]
Richest 60-80% income	33.8 [26.6-40.9]	29.9 [25.4-34.4]	28.1 [24-32.1]	21.1 [17.4-24.9]	18.5 [14.5-22.4]	24.3 [22.3-26.3]
Richest 80-100% income	11.2 [0.8-21.5]	15.7 [10.6-20.7]	17.6 [13.3-21.9]	14.9 [11.9-17.9]	11.2 [9-13.3]	13.9 [12.3-15.4]
<b>Total</b>	<b>54.2</b> [52.2-56.1]	<b>44.3</b> [42-46.6]	<b>37.6</b> [35.5-39.8]	<b>26.5</b> [24.4-28.6]	<b>18.4</b> [16.4-20.5]	<b>36.8</b> [35.8-37.8]
<i>Psychological distress, %</i>						
Poorest 20% income	16.2 [13.8-18.6]	19.2 [15.9-22.4]	22.1 [17.1-27.1]	19.3 [10.6-27.9]	24.2 [8.5-39.9]	18.2 [16.3-20.1]
Poorest 20-40% income	16.2 [13.5-19]	18.9 [15.9-21.9]	17.4 [13.9-20.9]	17.1 [11.4-22.8]	14.2 [6.8-21.6]	17.2 [15.6-18.9]
Middle 40-60% income	16 [12.5-19.6]	13.1 [10.3-16]	16 [13-19.1]	17.4 [13.8-21]	16.6 [11.9-21.3]	15.7 [14.2-17.3]
Richest 60-80% income	18.9 [12.4-25.5]	13.6 [9.9-17.3]	16 [12.6-19.4]	11.5 [8.6-14.4]	12.6 [9.4-15.8]	13.7 [12-15.3]
Richest 80-100% income	11.9 [3.9-20]	13.5 [8.9-18.1]	12.6 [9.3-15.8]	11.3 [8.7-13.9]	11.2 [8.9-13.5]	11.7 [10.2-13.1]
<b>Total</b>	<b>16.2</b> [14.7-17.8]	<b>16.3</b> [14.8-17.9]	<b>16.6</b> [15.1-18.1]	<b>13.9</b> [12.2-15.5]	<b>13</b> [11.2-14.9]	<b>15.3</b> [14.6-16]
<i>Poor health, %</i>						
Poorest 20% income	13.5 [11.4-15.7]	12 [8.9-15.1]	14.8 [10.5-19.2]	8.7 [1.9-15.4]	13.8 [2.5-25.2]	13.1 [11.3-14.8]
Poorest 20-40% income	10.4 [8-12.7]	10.8 [8.3-13.4]	7.1 [4.2-10]	8.6 [4.4-12.8]	5.6 [0.5-10.7]	9.2 [7.9-10.6]
Middle 40-60% income	9.8 [6.4-13.1]	8.3 [5.7-10.9]	6.4 [4.1-8.6]	6.4 [3.5-9.3]	5.2 [2.3-8.2]	7.1 [5.9-8.3]
Richest 60-80% income	5.6 [1.1-10]	7.7 [5-10.3]	8.6 [6.2-11]	4.8 [2.8-6.7]	4.2 [2.2-6.2]	6.1 [5-7.2]
Richest 80-100% income	5.2 [-1.1-11.5]	5.8 [2.5-9.2]	5.5 [3.3-7.7]	3.8 [2.2-5.3]	3.5 [2.1-4.9]	4.2 [3.2-5.1]
<b>Total</b>	<b>11.5</b> [10-13]	<b>9.6</b> [8.3-11]	<b>8.1</b> [6.8-9.3]	<b>5.5</b> [4.2-6.8]	<b>4.4</b> [3.4-5.4]	<b>7.9</b> [7.3-8.5]

*Smoking, %*

Poorest 20% income	16.7 [14.4-18.9]	18.6 [15.2-21.9]	19.7 [14.6-24.9]	17.5 [8.7-26.4]	26.8 [9.5-44.1]	17.9 [16.1-19.7]
Poorest 20-40% income	13.9 [11.5-16.3]	14.2 [11-17.3]	11.3 [7.9-14.7]	12.6 [7.9-17.3]	11.4 [5.5-17.3]	13.1 [11.6-14.5]
Middle 40-60% income	11 [7.5-14.4]	8 [5.6-10.5]	10.1 [7.4-12.8]	7.6 [4.5-10.6]	7.7 [4.3-11.1]	8.8 [7.5-10.2]
Richest 60-80% income	7.3 [2.5-12.1]	7.1 [4.2-9.9]	7.1 [4.8-9.4]	5.7 [3.7-7.7]	6.4 [4.2-8.6]	6.5 [5.4-7.6]
Richest 80-100% income	4.3 [-2.9-11.5]	5.2 [1.8-8.5]	7 [4.4-9.6]	5.1 [3.1-7.1]	4.3 [2.8-5.7]	5.1 [4-6.2]
<b>Total</b>	<b>14.4</b> <b>[12.8-15.9]</b>	<b>12.0</b> <b>[10.6-13.4]</b>	<b>10.5</b> <b>[9.1-11.9]</b>	<b>7.5</b> <b>[6.1-8.8]</b>	<b>6.5</b> <b>[5.3-7.8]</b>	<b>10.3</b> <b>[9.7-10.9]</b>

*Obesity, %*

Poorest 20% income	25 [22.5-27.4]	22.8 [19.4-26.2]	17.7 [12.6-22.7]	23.3 [14.7-31.9]	13.5 [1.6-25.3]	23.0 [21.2-24.7]
Poorest 20-40% income	21.3 [18.6-24.1]	22.5 [19.1-25.8]	20.4 [16.3-24.4]	22.6 [16.5-28.7]	24.6 [16.2-33]	21.8 [19.9-23.6]
Middle 40-60% income	24.9 [20.6-29.2]	22.3 [18.9-25.7]	18.6 [15.2-22]	15.2 [11.3-19]	15 [10.4-19.5]	19.1 [17.3-20.8]
Richest 60-80% income	24.9 [18.7-31.2]	20.8 [17.1-24.6]	17.4 [14.3-20.5]	15.5 [12.6-18.5]	14.3 [11.1-17.5]	17.1 [15.5-18.6]
Richest 80-100% income	14.2 [5.0-23.4]	20.8 [15.2-26.3]	13.7 [10.2-17.1]	14.6 [11.8-17.4]	8.9 [6.8-11]	12.6 [11.1-14]
<b>Total</b>	<b>23.7</b> <b>[22-25.4]</b>	<b>22.1</b> <b>[20.4-23.9]</b>	<b>17.8</b> <b>[16.1-19.4]</b>	<b>16.5</b> <b>[14.7-18.3]</b>	<b>12.6</b> <b>[11-14.2]</b>	<b>18.7</b> <b>[17.9-19.5]</b>

*Note: 95% confidence intervals reported in the square brackets. The heatmap shading visualises the magnitude of outcome prevalence, such that a darker shade corresponds to a higher prevalence and a lighter shade – to a lower prevalence.*

**Table 2. Risk ratios of inequalities in adverse outcomes at age 17 according to quintile of Index of Multiple Deprivation (IMD) or household income, or both**

	Model with IMD only		Model with income only		Model with IMD & income	
	RR	95% CI	RR	95% CI	RR	95% CI
<b>Poor academic achievement</b>						
Most deprived 20 % IMD	2.9***	2.6 - 3.3			1.5***	1.3 - 1.7
Most deprived 20-40 % IMD	2.4***	2.1 - 2.7			1.5***	1.3 - 1.7
Middle 40-60 % IMD	2.0***	1.8 - 2.3			1.5***	1.3 - 1.7
Least deprived 60-80 % IMD	1.4***	1.2 - 1.7			1.2**	1.1 - 1.4
Lowest 20 % income			4.5***	3.9 - 5.2	3.7***	3.2 - 4.4
Lowest 20-40 % income			3.4***	2.9 - 3.9	2.9***	2.5 - 3.4
Middle 40-60 % income			2.3***	2.0 - 2.7	2.1***	1.8 - 2.4
Highest 60-80 % income			1.7***	1.5 - 2.0	1.6***	1.4 - 1.9
<b>Psychological distress</b>						
Most deprived 20 % IMD	1.3*	1.0 - 1.5			1.0	0.8 - 1.3
Most deprived 20-40 % IMD	1.3*	1.0 - 1.5			1.1	0.9 - 1.3
Middle 40-60 % IMD	1.3**	1.1 - 1.5			1.2	0.9 - 1.4
Least deprived 60-80 % IMD	1.1	0.9 - 1.3			1.0	0.8 - 1.3
Lowest 20 % income			1.5***	1.3 - 1.8	1.5***	1.2 - 1.9
Lowest 20-40 % income			1.5***	1.2 - 1.8	1.4***	1.2 - 1.7
Middle 40-60 % income			1.3**	1.1 - 1.6	1.3**	1.1 - 1.6
Highest 60-80 % income			1.1	0.9 - 1.4	1.1	0.9 - 1.4
<b>Poor health</b>						
Most deprived 20 % IMD	2.5***	1.9 - 3.4			1.6*	1.1 - 2.2
Most deprived 20-40 % IMD	2.2***	1.6 - 2.9			1.5*	1.1 - 2.1
Middle 40-60 % IMD	1.8***	1.3 - 2.4			1.4*	1.0 - 1.9
Least deprived 60-80 % IMD	1.2	0.8 - 1.7			1.1	0.7 - 1.6
Lowest 20 % income			3.2***	2.4 - 4.3	2.4***	1.7 - 3.5
Lowest 20-40 % income			2.1***	1.6 - 2.9	1.7**	1.2 - 2.4
Middle 40-60 % income			1.8***	1.3 - 2.4	1.5*	1.1 - 2.2
Highest 60-80 % income			1.5*	1.1 - 2.1	1.4+	1.0 - 2.0
<b>Smoking</b>						
Most deprived 20 % IMD	2.2***	1.7 - 2.9			1.1	0.8 - 1.5
Most deprived 20-40 % IMD	1.8***	1.4 - 2.4			1.2	0.8 - 1.6
Middle 40-60 % IMD	1.6***	1.2 - 2.1			1.2	0.9 - 1.6
Least deprived 60-80 % IMD	1.2	0.9 - 1.5			1.0	0.7 - 1.4
Lowest 20 % income			3.8***	2.9 - 5.0	3.6***	2.5 - 5.0
Lowest 20-40 % income			2.5***	1.9 - 3.4	2.4***	1.7 - 3.3
Middle 40-60 % income			1.7**	1.2 - 2.3	1.6**	1.2 - 2.3
Highest 60-80 % income			1.4+	1.0 - 1.9	1.3+	1.0 - 1.8
<b>Obesity</b>						
Most deprived 20 % IMD	1.9***	1.6 - 2.2			1.6***	1.3 - 1.9
Most deprived 20-40 % IMD	1.8***	1.5 - 2.1			1.5***	1.3 - 1.8
Middle 40-60 % IMD	1.4***	1.2 - 1.7			1.3*	1.0 - 1.5
Least deprived 60-80 % IMD	1.3**	1.1 - 1.6			1.3**	1.1 - 1.6
Lowest 20 % income			1.7***	1.5 - 2.0	1.4***	1.2 - 1.7
Lowest 20-40 % income			1.7***	1.4 - 2.0	1.4***	1.2 - 1.7
Middle 40-60 % income			1.4***	1.2 - 1.7	1.3**	1.1 - 1.5
Highest 60-80 % income			1.3**	1.1 - 1.5	1.2*	1.0 - 1.4
Observations	15,367		15,367		15,367	

Note: \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , +  $p < 0.10$ ; the least deprived 80-100% IMD and highest 80-100% income are used as a reference. Regression results are based on a modified Poisson model (not adjusted for any confounders). The results from an adjusted regression are in the appendix (Table A4).

## Discussion

### *Main findings*

Household income in early childhood was a stronger and more consistent predictor than neighbourhood IMD for four of the five health-related outcomes at age 17. IMD was a slightly stronger predictor for obesity. Adverse outcomes increased consistently with decreasing household income, both in analyses based on income alone and in combination with IMD. In contrast, neighbourhood IMD was less consistently associated with adverse outcomes once income was considered.

Including both income and IMD in the analysis revealed patterns that would otherwise go unnoticed. Neighbourhood IMD was not associated with smoking or psychological distress in any of the income quintiles, nor with educational attainment in the poorest and richest income quintiles. Using IMD alone would miss the steep gradients of poorer educational attainment and increased smoking with decreasing household income within each IMD quintile.

### *What is already known on this topic?*

Given the lack of UK evidence highlighted in a recent systematic review<sup>[15]</sup>, it is not known which is a better proxy for early-years disadvantage: IMD or household income or both.

IMD and household income are different measures, leading to variations in their associations with adverse adolescent outcomes. The income deprivation component of IMD measures the proportion of low-income households meeting benefit thresholds, while household income is recorded for each child and grouped into quintiles across the income distribution (see appendix). IMD also captures multiple deprivation indicators, including low income, unemployment, and neighbourhood measures of education, health, crime, and access to amenities (listed in Table A1). Neighbourhood IMD is less sensitive and specific for detecting income-deprived households<sup>[12-14]</sup>, resulting in weaker inequality gradients compared to household-level measures. However, the multi-dimensionality of IMD may capture neighbourhood effects on young families regardless of income.<sup>[26]</sup>

### *What this study adds*

We found that household income in early childhood is a stronger predictor of adverse health outcomes in adolescence than neighbourhood IMD. However, combining income and IMD adds useful information for understanding subgroups at highest risk. Adding household income groupings to IMD could improve understanding of policies that affect household income, expose larger disparities in outcomes than is evident using IMD alone, and improve targeting of families who stand to benefit from early years interventions. Our findings suggest that the current core20plus5 approach advocated by NHS England<sup>[11]</sup>, is a poor predictor of children most at risk and does not consider how to address the gradient of disadvantage through policies that affect household income.<sup>[27]</sup>

Our findings strengthen arguments to widen the use and development of measure of individual or household income to address what has been called the inverse evidence law, whereby there is least evidence on the upstream determinants of health inequalities that are likely to have most impact on health but are most difficult to research.<sup>[10]</sup> Research on health inequalities from the Nordic countries, the Netherlands and North America, reports use of income alongside area-based deprivation indices.<sup>[26, 28-35]</sup> A recent US report describes a

comprehensive income dataset that incorporates nearly all taxable income which is used to address under-reporting in surveys and for tax administration, forecasting and research.<sup>[28, 29]</sup>

In the UK, improving information on costs and income is a priority for population studies on health and deprivation, as identified in a review and poverty experts.<sup>[36]</sup> Currently, the Office of National Statistics (ONS) derives household income measures, but allows their use only at the area level.<sup>[37, 38]</sup> However, recent ONS research, using employment data linked to health records, could serve as an exemplar for broader use of anonymized income data linked to administrative healthcare data for research.<sup>[39]</sup>

### *Limitations of this study*

First, the study sample was not large enough to draw statistical conclusions about the differences between the 25 subgroups combining household income and neighbourhood IMD. Second, missing data disproportionately affected disadvantaged groups but was addressed by multiple imputation<sup>[18, 19]</sup>, and by averaging income over three rounds of interviews in the early years. Third, we used parent-reported income as administrative data on household income is not yet available for England: differences between these sources need to be evaluated. Fourth, we used neighbourhood IMD, which includes population health outcomes, rather than restrict to the income deprivation domain of IMD. This was because IMD is a tool for policy and public health and restriction to income deprivation makes very little difference.<sup>[40]</sup>

### **Conclusion**

Household income in early childhood was a stronger predictor of health-related outcomes at age 17 than neighbourhood deprivation and was associated with steeper inequality gradients than currently recognised. Public health intervention in early childhood could be more effective if guided by household income as well as neighbourhood deprivation.

## **Acknowledgements**

We would like to thank our colleagues from the Children and Families Policy Research Unit, the Department for Health and Social Care, and the University of York for their useful discussions and feedback during the presentation of this work. We also express our gratitude to Aase Villadsen, who led the analyses of the cohort and variables for a study previously published in *The Lancet*, which we re-used and adapted for this study. Finally, we are grateful to the Centre for Longitudinal Studies (CLS), UCL Social Research Institute, for the use of the MCS data, and to the UK Data Service for making these data available. However, neither CLS nor the UK Data Service bear any responsibility for the analysis or interpretation of these data.

## **Data availability**

The datasets analysed during the current study are available through the UK Data Service [here](#). For further information on the Millennium Cohort Study, including data documentation and metadata, please visit the Centre for Longitudinal Studies [website](#).

## **Funding**

This independent research study was funded by the National Institute for Health and Care Research (NIHR) through the Children and Families Policy Research Unit as a responsive study (PR-PRU-1217-21301). Richard Cookson and Ieva Skarda also received support from the UK Prevention Research Partnership (ActEarly Programme, MR/S037527/1). Ruth Gilbert receives support from Health Data Research UK. The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care, or other funders.

## **Conflict of interest**

As already specified in the fundings section above, this is independent research supported by the NIHR through the Children and Families Policy Research Unit as a responsive study (PR-PRU-1217-21301). Richard Cookson and Ieva Skarda also received support from the Prevention Research Partnership (ActEarly Programme, MR/S037527/1). Ruth Gilbert receives support from Health Data Research UK. The authors have no other conflicts of interest to report.



## References

1. Cookson R, Asaria M, Ali S et al. Health equity indicators for the English NHS: a longitudinal whole-population study at the small-area level. *Health Services and Delivery Research* 2016. **4**(26). <https://doi.org/10.3310/hsdr04260>
2. Lai ET, Wickham S, Law C, Whitehead M, Barr B, Taylor-Robinson D. Poverty dynamics and health in late childhood in the UK: evidence from the Millennium Cohort Study. *Archives of Disease in Childhood* 2019. **104**(11): p. 1049-1055. <https://doi.org/10.1136/archdischild-2018-316702>
3. Straatmann VS, Lai ET, Lange T et al. How do early-life factors explain social inequalities in adolescent mental health? Findings from the UK Millennium Cohort Study. *Journal of Epidemiology & Community Health* 2019. **73**(11): p. 1049-1060. <https://doi.org/10.1136/jech-2019-212367>
4. Wickham S, Anwar E, Barr B, Law C, Taylor-Robinson D. Poverty and child health in the UK: using evidence for action. *Archives of Disease in Childhood* 2016. **101**(8): p. 759-66. <https://doi.org/10.1136/archdischild-2014-306746>
5. Marmot M. Social determinants of health inequalities. *Lancet* 2005. **365**(9464): p. 1099-104. [https://doi.org/10.1016/S0140-6736\(05\)71146-6](https://doi.org/10.1016/S0140-6736(05)71146-6)
6. Cooper K, Stewart K. Does household income affect children's outcomes? A systematic review of the evidence. *Child Indicators Research* 2021. **14**(3): p. 981-1005. <https://doi.org/10.1007/s12187-020-09782-0>
7. Troller-Renfree SV, Costanzo MA, Duncan GJ et al. The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences* 2022. **119**(5):e2115649119. <https://doi.org/10.1073/pnas.2115649119>
8. Goldfeld S, Downes M, Gray S et al. Household income supplements in early childhood to reduce inequities in children's development. *Social Science Medicine* 2024. **340**: p. 116430. <https://doi.org/10.1073/pnas.2115649119>
9. Almond D, Currie J, Duque V. Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature* 2018. **56**(4): p. 1360-1446. <https://doi.org/10.1257/jel.20171164>
10. Federation of Academies of Sciences and Humanities (ALLEA), the Federation of European Academies of Medicine (FEAM), *Health Inequalities Research. New Methods, Better Insights?* 2021. <https://doi.org/10.26356/health-inequalities>
11. NHS England. *Core20PLUS5 – An approach to reducing health inequalities for children and young people*. <https://www.england.nhs.uk/about/equality/equality-hub/national-healthcare-inequalities-improvement-programme/core20plus5/core20plus5-cyp/>, accessed 15 Jul 2024
12. McCartney G, Hoggett R, Walsh D, Lee D. How well do area-based deprivation indices identify income- and employment-deprived individuals across Great Britain today? *Public Health* 2023. **217**: p. 22-25. <https://doi.org/10.1016/j.puhe.2023.01.020>
13. McCartney G, Hoggett R. How well does the Scottish Index of Multiple Deprivation identify income and employment deprived individuals across the urban-rural spectrum and between local authorities? *Public Health* 2023. **217**: p. 26-32. <https://doi.org/10.1016/j.puhe.2023.01.009>
14. McLoone P. Targeting deprived areas within small areas in Scotland: population study. *BMJ* 2001. **323**(7309): p. 374-5. <https://doi.org/10.1136/bmj.323.7309.374>
15. Clery A, Grant C, Harron K, Bedford H, Woodman J. Measuring disadvantage in the early years in the UK: a systematic scoping review. *The Lancet* 2022. **400**: p. S32 [https://doi.org/10.1016/S0140-6736\(22\)02242-5](https://doi.org/10.1016/S0140-6736(22)02242-5)
16. Villadsen A, Asaria M, Skarda I et al. Clustering of adverse health and educational outcomes in adolescence following early childhood disadvantage: population-based retrospective UK cohort study. *Lancet Public Health* 2023. **8**(4): p. e286-e293. [https://doi.org/10.1016/S2468-2667\(23\)00029-4](https://doi.org/10.1016/S2468-2667(23)00029-4)
17. Joshi H, Fitzsimons E. The Millennium Cohort Study: the making of a multi-purpose resource for social science and policy. *Longitudinal and Life Course Studies* 2016. **7**(4): p. 409-430. <https://doi.org/10.14301/llics.v7i4.410>

18. Mostafa T, Narayanan M, Pongiglione B et al. Missing at random assumption made more plausible: evidence from the 1958 British birth cohort. *Journal of Clinical Epidemiology* 2021. **136**: p. 44-54. <https://doi.org/10.1016/j.jclinepi.2021.02.019>
19. Mishra S, Khare D. On comparative performance of multiple imputation methods for moderate to large proportions of missing data in clinical trials: a simulation study. *Journal of Medical Statistics and Informatics* 2014. **2**(1): p. 9. <https://doi.org/10.7243/2053-7662-2-9>
20. *Centre for Longitudinal Studies. Millennium Cohort Study: User Guide (Surveys 1–5), 9th edn.* London: University College London 2020. [https://doc.ukdataservice.ac.uk/doc/5795/mrdoc/pdf/mcs1-5\\_user\\_guide\\_ed9\\_2020-08-07.pdf](https://doc.ukdataservice.ac.uk/doc/5795/mrdoc/pdf/mcs1-5_user_guide_ed9_2020-08-07.pdf), accessed 23 Jul 2024
21. Cookson R, Griffin S, Norheim OF, Culyer AJ eds. *Distributional cost-effectiveness analysis: quantifying health equity impacts and trade-offs.* Oxford University Press 2020
22. Sariaslan A, Mikkonen J, Aaltonen M, Hiilamo H, Martikainen P, Fazel S. No causal associations between childhood family income and subsequent psychiatric disorders, substance misuse and violent crime arrests: a nationwide Finnish study of >650 000 individuals and their siblings. *International Journal of Epidemiology* 2021. **50**(5): p. 1628-1638. <https://doi.org/10.1093/ije/dyab099>
23. Rod NH, Lange T, Petersen AH. Do sibling comparisons answer the causal question? In response to 'No causal associations between childhood family income and subsequent psychiatric disorders, substance misuse and violent crime arrests'. *International Journal of Epidemiology* 2022. **51**(6): p. 2025-2026. <https://doi.org/10.1093/ije/dyab235>
24. Ledberg A, Rajaleid K, Modin B. Are there really no causal associations between childhood family income and subsequent outcomes? *International Journal of Epidemiology* 2022. **51**(6): p. 2027-2028. <https://doi.org/10.1093/ije/dyac016>
25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988: p. 837-845. <https://doi.org/10.2307/2531595>
26. Burgos Ochoa L, Bijlsma MJ, Steegers EA, Been JV, Bertens LC. Does neighborhood crime mediate the relationship between neighborhood socioeconomic status and birth outcomes? An application of the mediational G-formula. *American Journal of Epidemiology* 2023. **192**(6): p. 939-948. <https://doi.org/10.1093/aje/kwad037>
27. Marmot M. Health equity in England: the Marmot review 10 years on. *BMJ* 2020. **368**: p. m693. <https://doi.org/10.1136/bmj.m693>
28. Medalia C, Meyer BD, O'Hara AB, Wu D. Linking survey and administrative data to measure income, inequality, and mobility. *International Journal of Population Data Science* 2019. **4**(1). <https://doi.org/10.23889%2Fijpds.v4i1.939>
29. Corinth K, Meyer BD, Wu D. The Change in Poverty from 1995 to 2016 among Single-Parent Families. *AEA Papers and Proceedings* 2014. **112**: p. 345-350. <https://doi.org/10.1257/pandp.20221043>
30. Li X, Memarian E, Sundquist J, Zöller B, Sundquist K. Neighbourhood deprivation, individual-level familial and socio-demographic factors and diagnosed childhood obesity: a nationwide multilevel study from Sweden. *Obesity Facts* 2014. **7**(4): p. 253-263. <https://doi.org/10.1159/000365955>
31. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annual Review of Public Health* 2011. **32**: p. 91-108. <https://doi.org/10.1146/annurev-publhealth-031210-100700>
32. Thygesen LC, Ersbøll AK. Danish population-based registers for public health and health-related welfare research: introduction to the supplement. *Scandinavian Journal of Public Health* 2011. **39**(7 Suppl): p. 8-10. <https://doi.org/10.1177/1403494811409654>
33. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *European Journal of Epidemiology* 2014. **29**: p. 541-549. <https://doi.org/10.1007/s10654-014-9930-3>

34. Evensen M, Klitkou ST, Tollånes MC et al. Parental income gradients in adult health: a national cohort study. *BMC Medicine* 2021. **19**(1): p. 152. <https://doi.org/10.1186/s12916-021-02022-4>
35. Hynek KA, Abebe DS, Hollander AC, Liefbroer AC, Hauge LJ, Straiton ML. The association between persistent low parental income during preschool age and mental disorder in adolescence and early adulthood: a Norwegian register-based study of migrants and non-migrants. *BMC Psychiatry* 2022. **22**(1): p. 206. <https://doi.org/10.1186/s12888-022-03859-6>
36. Timpson K, McCartney G, Walsh D, Chabanis B. What is missing from how we measure and understand the experience of poverty and deprivation in population health analyses? *European Journal of Public Health* 2023. **33**(6): p. 974-980. <https://doi.org/10.1093/eurpub/ckad174>
37. Office for National Statistics. *Income estimates for small areas, England and Wales* 2023: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/smallareaincomeestimatesformiddlelayerssuperoutputareaseinglandandwales>, accessed 15 Jul 2024
38. Office for National Statistics Statistical Bulletin. Andrew Zellin. *Income estimates for small areas, England and Wales: financial year ending 2020. Small area model-based income estimates covering local areas called Middle layer Super Output Areas (MSOAs) in England and Wales* 2023 <https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/smallareamodelbasedincomeestimates/financialyearending2020>, accessed on 23 Jul 2024
39. HM Treasury. *£12.4 million to help change choices about work* 2023 <https://www.gov.uk/government/news/124-million-to-help-change-choices-about-work>, accessed 15 Jul 2024.
40. McCartney G, Hoggett R, Walsh D, Lee D. How important is it to avoid indices of deprivation that include health variables in analyses of health inequalities? *Public Health* 2023. **221**: p. 175-180. <https://doi.org/10.1016/j.puhe.2023.06.028>