medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Early Life Domains as Predictors of Obesity and Hypertension Comorbidity: Findings from the 1970 British Cohort Study (BCS70)

S Stannard[1](ORCID: 0000-0002-6139-1020), RK Owen[2] (ORCID: 0000-0001-5977-376X), A Berrington[3] (ORCID: 0000-0002-1683-6668), N Ziauddeen[1] (ORCID: 0000-0002-8964-5029), SDS Fraser[1] (ORCID: 0000-0002-4172-4406), S Paranjothy[4], RB Hoyle[5] (ORCID: 0000-0002-1645-1071), N A Alwan[1,6,7] (ORCID: 0000-0002-4134-8463)

[1]School of Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[2]Population Data Science, Swansea University Medical School, Faculty of Medicine, Health & Life Science, Swansea University, Swansea, United Kingdom

[3] School of Economic, Social and Political Sciences, University of Southampton, Southampton, United Kingdom

[4]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen

[5]School of Mathematical Sciences, University of Southampton, Southampton, United Kingdom

[6]University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom

[7]NIHR Applied Research Collaboration Wessex, Southampton, United Kingdom

Corresponding Author:

Sebastian Stannard - S.J.Stannard@soton.ac.uk

University of Southampton, University Road, Southampton, SO17 1BJ, United Kingdom

+44(0)23 8059 5000

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Abstract

Background: Obesity and hypertension are major public health problems and are associated with adverse health outcomes. To model realistic prevention scenarios and inform policy, it may be helpful to conceptualise early lifecourse domains of risk and incorporate such information when predicting comorbidity outcomes. We identify exposures across five pre-hypothesised childhood domains and explore them as predictors of obesity and hypertension comorbidity in adulthood.

Methods: The analytical sample included 7858 participants in the 1970 British Cohort Study. The outcome was obesity (BMI of ≥30) and hypertension (blood pressure>140/90mm Hg or self-reported doctor's diagnosis) comorbidity at age 46. Early life domains included: 'prenatal, antenatal, neonatal and birth', 'developmental attributes and behaviour', 'child education and academic ability', 'socioeconomic factors' and 'parental and family environment'. We conducted prediction analysis of the outcome in three stages:(1) stepwise backward elimination to select variables for inclusion for each domain (2) calculation of predicted risk scores of obesity-hypertension for each cohort member within each domain (3) multivariable logistic regression analysis including domain-specific risk scores, sex and ethnicity to assess how well the outcome could be predicted. We additionally included potential adult predictors of obesity-hypertension comorbidity as sensitivity analysis.

Results: Including all domain-specific risk scores in the same model, all five domains were significant predictors of obesity-hypertension comorbidity. The predictive power of the model, measured by the area under the curve (AUC), was 0.63 (95%CI 0.61-0.65). Including adult predictors increase the AUC to 0.68 (95%CI 0.66-0.70), and three early life domains - the parental and family environment domain (OR 1.11 95%CI 1.05-1.17) the socioeconomic

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

factors domain (OR 1.09 95%CI 1.04-1.16), and the education and academic ability domain (OR 1.07 95%CI 1.02-1.13) remained predictors of obesity-hypertension comorbidity.

Conclusions: We found three robust domains for predicting obesity-hypertension comorbidity. Interventions that address these early life factors could reduce the burden of comorbidity.

Introduction

Obesity and hypertension are major public health problems [1-2]. In England, 26% of adults have obesity [3], and 30% of adults have hypertension [4]. Both conditions are associated with morbidities later in the lifecourse, including Type 2 diabetes, heart disease, kidney disease, renal disease, strokes, and some cancers, including breast and bowel cancer [5-8]. There is evidence suggesting that in England and Scotland since 2014, obesity and excess body fat have contributed to more deaths among people in middle- and old-age than smoking [9]. On a global level, in 2019 and across 204 countries, the leading Level 2 risk factor for attributable deaths was high systolic blood pressure, and between 2010 to 2019, one of the largest increases in risk exposure was for high body-mass index [10].

Obesity and hypertension are closely related and often co-occur [11], for example each condition occurs at higher frequency with the other than in a population free of either [12]. Research indicates that obesity accounts for 60-70% of hypertension, and individuals with obesity are 3.5 times more likely to have hypertension compared to normal weight individuals [13,14]. The combination of both conditions significantly increases the likelihood of adverse health outcomes such as cardiovascular disease, reduced sexual function, quality of life and mortality [12,15,16]. Previous literature has also identified obesity and

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

hypertension as common sentinel conditions, defined as the first long-term condition in the development of multiple long-term conditions [17-20].

A substantial body of evidence suggests that experiences in early life are crucial in determining outcomes such as obesity and hypertension. The aetiology of chronic disease has been strongly linked with environmental exposures in utero and early life [21]. Socioeconomic disadvantage in early life is also strongly related to obesity and hypertension [22]. Analyses of the Hertfordshire cohort study demonstrated that paternal social class was associated with future multimorbidity, including hypertension [23]. In the Aberdeen Children of the 1950s cohort, lower father's social class at birth was associated with early-onset multimorbidity, including hypertension [24], and in the 1970 British Cohort Study those with fathers from unskilled occupational groups (vs. professional) at birth had 43% higher risk of early-onset (age 46-48) multimorbidity including hypertension [25].

Many wider determinants acting in childhood are likely to increase the risk of disease in adulthood, and in previous research [26,27] we conceptualised exposures across 12 pre-defined childhood domains covering a range of social, economic, developmental, educational and environmental factors. Most previous research focuses on single exposure-outcome relationships, potentially to reduce statistical complexity, or to focus policy attention onto a specific aspect. However, focussing on single exposures does not reflect the reality and complexity of the early life course given children are likely to be exposed to combinations of intersectional factors across these domains, often simultaneously. We argue that analyses must begin to explore exposures as domains (i.e., a group of variables that represent an overarching theme) rather than the individual variables that form its components for three reasons. First, to provide a combined exposure measure that reflects multiple variables in the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

data rather than performing multiple statistical testing using all of the components in relation to the study outcomes. Second, to conceptualise the components within wider early life domains provides a better reflection of the childhood conditions in which people grow up. Third considering domains will better inform interventions and policy in childhood as incorporating information from multiple early life domains into the same analysis may help us understand the combined effects of different experiences across a range of early life domains on developing long-term conditions. This can provide actionable insights into developing complex multi-domain interventions in childhood that may help support people to live more healthily for longer across the lifecourse.

In this paper we look to address two research questions. Firstly, which early life domain(s) contribute to predicting the risk of obesity-hypertension comorbidity at mid-life? Secondly, which early life domain(s) are more important than others in explaining the variability in obesity-hypertension occurrence?

To achieve this, we aimed to predict the outcome of obesity-hypertension comorbidity using five early lifecourse domains. In order to do this while weighting the components of each domain, we also aimed to produce -as a first stage of the analysis- predicted risk scores of the outcome for each of the five pre-defined early life domains [26,27]. In a sensitivity analysis we included potential adult risk factors of obesity-hypertension comorbidity (number of days of exercise per week, highest educational qualification, weekly income, number of cigarettes smoked daily, hours spent on a weekday watching television, hours on a weekday spent on the internet (not for work related reasons), cohabitating with a partner, alcohol consumption and use of e-cigarettes) to explore which early life domains matter most for obesity-hypertension comorbidity taking account of such factors.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

This work forms part of a larger aim to model targeted multimorbidity prevention scenarios as part of the Multidisciplinary Ecosystem to study Lifecourse Determinants and Prevention of Early-onset Burdensome Multimorbidity (MELD-B) project [28].

Methods

Dataset

We used the 1970 British Cohort Study (BCS70) [29] that has followed 17196 cohort members in England, Scotland, Wales born in one week in 1970; to date, there have been 10 sweeps of data collection – 4 in childhood and 6 in adulthood. The comorbidity outcome of obesity and hypertension was measured at age 46 within a biomedical sweep with measurements conducted by a research nurse. All other variables were collected either at birth or age 10.

Outcome

The outcome was a combined obesity-hypertension phenotype at age 46. Blood pressure was measured via three systolic and diastolic blood pressure readings during a single appointment and administered by a research nurse. Hypertension was defined as an average blood pressure reading of over 140/90 mm Hg. We additionally classified hypertension if a participant reported (at age 46) that they had received a doctor's diagnosis of high blood pressure or hypertension, even if the blood pressure measurement was less than 140/90 mm Hg, since diagnosed hypertension may be accompanied by intake of antihypertensive medication, thus lowering blood pressure readings at the time of cohort measurement. Body mass index (BMI) was calculated via height and weight measurements taken during the same nurse appointment using the following formula: BMI = weight (kg) / height $(m)^2$. Obesity was defined as a BMI

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

of 30 or over. The obesity-hypertension comorbidity variable was considered as a binary (no/yes) variable.

Exposures (Five pre-hypothesised domains)

We previously conceptualised 12 domains of early life risk factors of future multimorbidity risk informed through a scoping literature/policy review and patient and public engagement [26]. In this paper we focus our analysis on 5 out of the 12 domains chosen because they showed unadjusted associations with the outcomes:

1. *Prenatal, antenatal, neonatal and birth domain* focused on the period from conception to the onset of labour, the circumstances and outcomes surrounding a birth, and the period immediately following birth.
2. *Developmental attributes and behaviour domain* focused on the developmental markers of children relating to cognition, coordination, personality types and behavioural traits.
3. *Child education and academic ability domain* related to the process of learning and educational achievement, especially in educational settings, and the knowledge an individual gains from these educational institutions.
4. *Socioeconomic factors domain* included factors relating to differences between individuals or groups of peoples caused mainly by their social and economic situation.
5. *Parental and family environment domain* incorporated the interactions between children and care givers, parenting styles, parental beliefs, attitudes and discipline, and wider family factors such as kin networks.

Supplementary Materials Table 1 includes all the variables that were initially considered for each domain. These variables were selected and categorised based on a previous data audit and PCA analysis [27] that identified early-life variables from multiple sweeps of

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

data that fitted into five early-life domains of future multimorbidity risk. This previous work reduced the dimensionality of the data and structured each of the five domains into mutually exclusive groups of variables based on similar characteristics [27].

Analytical sample

The analytical sample included all cohort members who had measured BMI and blood pressure at age 46 (n = 7858); this represented 45.7% of the original birth cohort. To preserve sample size and reduce bias in the estimates due to missing data we used multiple imputation. Multiple imputation was conducted by chained equations for missing observations at birth, age 10, and 46 [30]. 50 imputation cycles were constructed under the missing-at-random assumption [31-33], which has been found to be highly plausible in the British birth cohorts [34]. All variables were included in the imputation process. The outcome was included in the imputed models, but imputed outcome values were not used. For reference we include results based on complete case analysis in Supplementary Materials Tables 2 and 9.

Statistical analysis

Step 1: Stepwise backwards elimination to select variables for inclusion separately for each domain.

Firstly, stepwise backward elimination conducted on multiple imputed data, was used to select variables for inclusion separately for each domain. This method started with all potential variables identified for each domain as outlined in Supplementary Materials Table 1, then variables were removed sequentially based on a series of hypothesis tests. In backward elimination, variables were removed sequentially if the p-value for a variable exceeded the specified significance level which was set at 0.157. This level was chosen conservatively to reduce the risk of overfitting and is the equivalent to the Akaike

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

information criterion (AIC) [35,36]. Table 1 identifies the retained variables following stepwise backwards elimination for each domain in relation to the outcome of obesity-hypertension comorbidity at age 46, the descriptive statistics were based on the specific sample available at the sweep in which the variable was reported or measured.

Step 2: Predicted risk scores of obesity-hypertension comorbidity for each cohort member within each domain.

Secondly, logistic regression models then explored the relationship between retained variables within each domain following stepwise backwards elimination and odds of obesity-hypertension comorbidity. Based on this logistic regression modelling, and using the 'predict' function in STATA [37], predicted risk scores of the obesity-hypertension comorbidity outcome for each cohort member within each domain were calculated. In other words, each cohort member has five predicted score values in relation to the obesity-hypertension comorbidity outcome, one for each of the five domains. These predicted risk scores for each individual, and within each domain, were centred on the mean predicted risk score within that domain, and were bound between -1 and 1. A Pearson correlation matrix then explored the correlation between domain-specific risk scores. We generated predicted risk scores of obesity-hypertension for each of the five early life domains to account for the weight of each domain component in relation to the outcome rather than assuming each component contributes an equal weight to the outcome.

Step 3: A prediction model including five domain-specific risk scores.

The third step involved focussing on how well obesity-hypertension could be predicted including all five domain-specific risk scores produced in step 2 using a multivariable logistic regression model with the 'predict function', and adding sex (at birth) and ethnicity. We

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

produced the area under curve (AUC) statistic to assess the predictive performance of this model. Odds ratios and confidence intervals of the five domains within this model were used to identify the strongest domains that acted as predictors for obesity and hypertension comorbidity taking into account the effect of the other domains.

Step 4: A prediction model including five domain-specific risk scores and adult factors (sensitivity analyses)

The fourth step involved performing step 3 with the inclusion of adult factors that are potentially linked to both the exposures and the outcome. These were recorded at age 46 including number of days of exercise per week, highest educational qualification, occupational social class, weekly income, number of cigarettes smoked daily, hours spent on a weekday watching television, hours on a weekday spent on the internet (not for work related reasons), cohabitating with a partner, alcohol consumption and use of e-cigarettes.

Ethical considerations

Ethics approval for the MELD-B project has been obtained from the University of Southampton Faculty of Medicine Ethics committee (ERGO II Reference 66810).

Results

Step 1: Stepwise backwards elimination to select variables for inclusion in each domain.

Among the 7858 cohort members at age 46, 597 (7.6%) had obesity-hypertension comorbidity at age 46. Table 1 identifies the retained variables following stepwise backwards elimination for each domain in relation to the outcome of obesity-hypertension comorbidity at age 46. As shown in Table 1, in the socioeconomic factors domain, 9.2% of cohort members whose gross family income at age 10 (per week) was below £49 had obesity-

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

hypertension at age 46, compared to 4.2% of cohort members whose gross family income was above £250. In the prenatal, antenatal, neonatal domain, 8.8% of cohort members whose mothers smoked during pregnancy had obesity and hypertension at age 46 compared to a 6.7% prevalence in non-smokers. In the education and academic ability domain 10.1% of cohort members who reported some difficulty with reading at age 10 had obesity-hypertension at age 46 compared to a prevalence of 7.1% in those with no difficulty. Finally, for the parent and family environment domain of cohort members with a father by adoption (10.4%), stepfather (10.3%) or another father figure such as a grandparent (11.2%) at age 10 had obesity-hypertension at age 46 compared to 7.4% in those living with a biological father.

Table 1. Step 1: Variables retained in the five domains risk scores following stepwise backwards elimination, and the prevalence of obesity-hypertension comorbidity at age 46.

			Obesity and High Blood Pressure		
			No N (%)	Yes N (%)	*Total Sample*[1]
Prenatal, antenatal, neonatal and birth domain (variables recorded at birth)	Mother's parity at birth of cohort member	0	2849 (91.9%)	251 (8.1%)	*3100*
		1	2475 (93.1%)	183 (6.9%)	*2658*
		2	1127 (92.7%)	89 7.3%	*1216*
		3	460 (93.1%)	34 (6.9%)	*494*
		4	201 (89.3%)	24 (10.7%)	*225*
		5+	148 (89.7%)	17 (10.3%)	*165*
	Maternal smoking during pregnancy	Non-smoker	3190 (93.3%)	229 (6.7%)	*3419*
		Stopped pre/during	1272 (92.8%)	993 (7.2%)	*1371*

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

		Smoker	2758 (91.2%)	266 (8.8%)	*3024*
	Birthweight	No obesity-hypertension - mean (SD)		3318g (523g)	*7254*
		Yes obesity-hypertension - mean (SD)		3246g (483g)	*596*
	Maternal age at birth of cohort member	No obesity-hypertension - mean (SD)		26.0 (5.3)	*7220*
		Yes obesity-hypertension - mean (SD)		25.6 (5.6)	*594*
Education and academic ability domain (variables recorded at age 10)	Ability spelling – cohort member's assessment	Good ability	3210 (93.2%)	234 (6.9%)	*3444*
		Poor ability	2605 (91.0%)	258 (9.0%)	*2863*
	Difficulty writing – teacher's assessment	No difficulty	5173 (92.9%)	396 (7.1%)	*5569*
		Some difficulty	1050 (90.2%)	114 (9.8%)	*1164*
		Great difficulty	127 (94.8%)	7 (5.2%)	*134*
	Difficulty reading – teacher's assessment	No difficulty	5260 (92.9%)	402 (7.1%)	*5662*
		Some difficulty	950 (89.9%)	107 (10.1%)	*1057*
		Great difficulty	161 (93.6%)	11 (6.4%)	*172*
	Edinburgh Reading Test Scores	No obesity-hypertension - mean (SD)		42.5 (12.1)	*5776*
		Yes obesity-hypertension - mean (SD)		40.1 (12.2)	*493*
Parent and family environment domain (variables recorded at age	Father interest in education – teacher's assessment	Very interested	2200 (93.6%)	150 (6.4%)	*2350*
		Moderate interest	1136 (92.1%)	97 (7.8%)	*1233*
		Very little interest	178 (90.8%)	18 (9.2%)	*196*

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

10)		Uninterested	144 (92.3%)	12 (7.7%)	*156*
		Cannot say/no father	1658 (90.6%)	172 (9.5%)	*1830*
	Mother interest in education – teacher's assessment	Very interested	3058 (93.2%)	203 (6.8%)	*3281*
		Moderate interest	1694 (91.8%)	151 (8.2%)	*1845*
		Very little interest	242 (90.3%)	26 (9.7%)	*268*
		Uninterested	99 (90.8%)	10 (9.2%)	*109*
		Cannot say/no mother	636 (89.5%)	75 (10.6%)	*711*
	Family go on outings	Often	3570 (92.7%)	281 (7.3%)	*3851*
		Sometimes	2602 (91.9%)	229 (8.1%)	*2831*
		Rarely or never	216 (92.7%)	17 (7.3%)	*233*
	Father helps manage cohort member	Equal or similar amount	3346 (92.1%)	287 (7.9%)	*3633*
		Smaller part than mother	2118 (93.2%)	154 (6.8%)	*2272*
		Very small part	437 (92.4%)	36 (7.6%)	*473*
		Does not help	442 (91.1%)	43 (8.9%)	*485*
	Father Figure	Biological father	5584 (92.7%)	440 (7.4%)	*6024*
		Father by adoption	86 (89.6%)	10 (10.4%)	*96*
		Step father	235 (89.7%)	27 (10.3%)	*262*
		Other	111 (88.8%)	14 (11.2%)	*125*

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

		No father figure	415 (91.6%)	38 (8.4%)	*453*
Developmental attributes and behaviour domain (variables recorded at age 10)	Rutter behaviour	No obesity-hypertension - mean (SD)		424.0 (209.2)	*6047*
		Yes obesity-hypertension - mean (SD)		449.7 (196.8)	*497*
	Number of steps walking backwards	No obesity-hypertension - mean (SD)		15.7 (5.4)	*5975*
		Yes obesity-hypertension - mean (SD)		15.0 (5.6)	*488*
	Hand control rating	No obesity-hypertension - mean (SD)		30.3 (12.2)	*5815*
		Yes obesity-hypertension - mean (SD)		28.2 (12.2)	*496*
Socioeconomic factors domain (variables recorded at age 10)	Parental social class	I professional	439 (94.2%)	27 (5.8%)	*466*
		II managerial	1691 (94.1%)	106 (5.9%)	*1797*
		III non-manual	710 (92.8%)	55 (7.2%)	*765*
		III manual	2432 (91.5%)	226 (8.5%)	*2658*
		IV partly skilled	762 (89.5%)	89 (10.5%)	*851*
		V Unskilled	206 (91.9%)	18 (8.1%)	*224*
	Gross Income per week	£250+	408 (95.8%)	16 (4.2%)	*426*
		£200-£249	421 (92.3%)	35 (7.7%)	*456*
		£150-£199	1061 (94.0%)	68 (6.0%)	*1129*
		£100-£149	2072 (91.7%)	188 (8.3%)	*2260*
		£50-£99	1602 (91.0%)	158 (9.0%)	*1760*

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

		Under £49	334 (90.8%)	34 (9.2%)	*368*
	Housing tenure	Owned outright	737 (93.1%)	55 (6.9%)	*792*
		Private rent	191 (93.2%)	14 (6.8%)	*205*
		Being purchased (mortgage)	3585 (93.0%)	270 (7.0%)	*3855*
		Council rent	1685 (90.5%)	177 (9.5%)	*1862*
		Other	198 (92.1%)	17 (7.9%)	*215*

[1]*The total sample size was based on the specific sample available at the sweep in which the variable was reported or measured.*

Step 2: Generation of domain-specific predicted risk scores of obesity-hypertension comorbidity for each cohort member.

Supplementary Materials Figure 1. presents a histogram of the distribution of domain-specific predicted risk scores, generated in step 2, following logistic regression modelling that explored the relationship between retained variables following stepwise backwards elimination (step 1) and obesity-hypertension comorbidity for the five early life domains. In Supplementary Materials Tables 4-8, we include the regression coefficients of obesity-hypertension for these models, and for each domain separately. The domain risk scores were centred on the mean predicted risk score (for each domain) and bound between -1 and 1. As shown the largest range of domain predicted risk scores was for the prenatal, antenatal, neonatal and birth domain (range -0.05 – 0.14) and the smallest range was for the parental and environmental domain risk score (range -0.03 – 0.08).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Given there was likely to be correlation between domain risk scores, in Supplementary Materials Table 3 we present a Pearson correlation matrix exploring correlation across domains. The highest correlation was between both the developmental attributes and behaviour domain and the child education and academic ability domain (coefficient 0.29), and between the parental and family environment domain and the socioeconomic factors domain (coefficient 0.32), and therefore the correlation between predicted risk scores for each domain was low.

Step 3: Prediction modelling of obesity-hypertension including five domain-specific risk scores.

We calculated the area under curve (AUC) (included in Supplementary Materials 10) for the prediction model that included all five domain specific risk scores (produced in step 2), ethnicity and sex. Overall, the area under the curve was 0.63 (95%CI 0.61-0.65). Table 2 presents the odds ratios of obesity-hypertension at age 46 for the prediction model that included all five domain specific risk scores (produced in step 2), sex and ethnicity. As shown, all domain risk scores were significant predictors of obesity-hypertension comorbidity, with the strongest being the parental and family environment domain (OR 1.11 95%CI 1.05-1.17) and the socioeconomic factors domain (OR 1.11 95%CI 1.05-1.17). The weakest domain predictors included the prenatal, antenatal, neonatal and birth domain (OR 1.06 95%CI 1.01-1.12) and the education and academic ability domain (OR 1.07 95%CI 1.01-1.12).

Table 2. Step 3: Odds ratios of obesity-hypertension at age 46 in relation to domain-specific risk score of obesity-hypertension for five early life domains. Multiple imputed data (50 Imputations).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

	Odds ratios	95% Confidence interval
Developmental attributes and behaviour domain	**1.08**	**1.03 - 1.15**
Prenatal, antenatal, neonatal and birth domain	**1.06**	**1.01 - 1.12**
Parent and family environment domain	**1.11**	**1.05 - 1.17**
Education and academic ability domain	**1.07**	**1.01 - 1.12**
Socioeconomic factors domain	**1.11**	**1.05 - 1.17**
Child's sex: Women (base – men)	**0.77**	**0.64 - 0.91**
Mother ethnic group: Other (base – white)	1.05	0.69 - 1.61
Number of observations		7858

Step 4: Prediction modelling of obesity-hypertension comorbidity including five domain-specific risk scores and potential adult predictors (sensitivity analysis).

Table 3 presents the odds ratios of obesity-hypertension at age 46 for the prediction model that included all five domain specific risk scores (produced in step 2) ethnicity, sex and adult predictors. Overall the inclusion of adult predictors increased the area under the curve to 0.68 (95%CI 0.66-0.70) (Supplementary table 11). As shown, after the inclusion of adult predictors, the prenatal, antenatal, neonatal and birth domain (OR 1.05 95%CI 0.99-1.11) and the developmental attributes and behaviour domains (OR 1.07 95%CI 1.00-1.13) were no longer significant predictors of obesity-hypertension comorbidity. The other three domains remained significant predictors for obesity and hypertension comorbidity. These domains included the parental and family environment domain (OR 1.11 95%CI 1.05-1.17), the socioeconomic factors domain (OR 1.09 95%CI 1.04-1.16), and the education and academic abilities domain (OR 1.07 95%CI 1.02-1.13).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Table 3. Step 3: Odds ratios of obesity-hypertension at age 46 in relation to domain-specific risk score of obesity-hypertension for five early life domains including adult predictors. Multiple imputed data (50 Imputations).

	OR	95% Confidence Interval
Developmental attributes and behaviour domain	1.07	1.00 - 1.13
Prenatal, antenatal, neonatal and birth domain	1.05	0.99 - 1.11
Parent and family environment domain	**1.11**	**1.05 - 1.17**
Education and academic ability domain	**1.07**	**1.02 - 1.13**
Socioeconomic factors domain	**1.09**	**1.04 - 1.16**
Child's sex: Women (base – men)	**0.75**	**0.62 - 0.91**
Mother ethnic group: Other (base – white)	0.89	0.58 - 1.37
Number of cigarettes smoked a day	0.99	0.98 - 1.01
TV hours per weekday:1 to 2 hours (base - 1 hour or less)	**1.72**	**1.24 - 2.39**
2 to 3 hours	**1.46**	**1.03 - 2.06**
3 to 4 hours	**1.49**	**1.02 - 2.17**
More than 4 hours	**1.90**	**1.26 - 2.86**
Internet hours per weekday: 1 to 2 hours (base -1 hour or less)	1.12	0.91 - 1.38
2 to 3 hours	**1.37**	**1.04 - 1.80**
3 to 4 hours	1.37	0.95 - 1.98
More than 4 hours	**1.47**	**1.03 - 2.09**
Cohabitating: No (base – yes)	1.04	0.85 - 1.27
Alcohol consumption: Does not drink alcohol (base – normal)	**1.39**	**1.08 - 1.81**
High risk drinking	**1.23**	**1.00 - 1.51**
Social class: Lower managerial (base - higher managerial)	1.10	0.82 - 1.49
Intermediate occupations	1.12	0.77 - 1.63
Small employer	0.95	0.63 - 1.43
Lower supervisor	1.04	0.71 - 1.55
Semi-routine	0.89	0.59 - 1.35
Routine	0.82	0.52 - 1.28
Not stated	1.15	0.80 - 1.67
Not applicable	0.75	0.48 - 1.17
Days exercise per week: 1 (base – 0)	0.75	0.54 - 1.05
2	**0.74**	**0.55 - 0.99**

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

3	**0.74**	**0.55 - 0.99**
4	**0.67**	**0.47 - 0.97**
5	0.80	0.59 - 1.07
6	**0.56**	**0.32 - 0.98**
7	**0.54**	**0.41- 0.72**
Educational qualification: GCSE (base - no qualification)	1.32	**1.06 - 1.64**
A/AS level	1.24	0.81 - 1.91
Diploma	1.39	0.99 - 1.95
Degree	1.20	0.89 - 1.63
Higher degree	1.01	0.58 - 1.75
Weekly income	1.00	0.99 - 1.00
Managing financially: doing all right (base - doing well)	1.06	0.86 - 1.30
Just about getting by	1.24	0.95 - 1.61
Finding it quite difficult	**1.72**	**1.16 - 2.53**
Finding it very difficult	1.65	0.99 - 2.78
E-cigarettes: Previously smoked (base- never smoked)	0.97	0.62 - 1.39
Occasionally smoke	1.43	0.97 - 2.14
Smoke daily	0.82	0.56 - 1.21
Number of observations		7858

Discussion

Using the BCS70 data, early life domains showed associations with obesity-hypertension comorbidity at age 46 even after accounting for each other, with the strongest domain predictors for obesity-hypertension being the parental and family environment domain and the socioeconomic factors domain. After accounting for adult predictors including number of days of exercise per week, highest educational qualification, occupational social class, weekly income, number of cigarettes smoked daily, hours spent on a weekday watching television, hours on a weekday spent on the internet (not for work related reasons), cohabitating with a partner, alcohol consumption and use of e-cigarettes, associations with the outcome for the parental and family environment domain, the socioeconomic factors domain, and the education and academic abilities domain remained robust.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

In this paper we have quantified the effect of a combination of risk factors in early life in the shape of life domains on comorbidity outcomes. We recognise that prevention policy choices are rarely feasible or practical for one isolated risk factor and thinking in terms of domains may help addressing the early wider determinants of later health. Therefore, our motivation for this analysis stems from the observation that health research tends to investigate single exposure-outcome relationships, while children's experiences across a variety of early lifecourse behavioural, social, economic and transgenerational domains are intersecting, and therefore research needs to be able to incorporate information from multiple domains into the same analysis. However, it is also important that health research reflects on modifiable risk factors for ill health that focus on both direct and indirect factors including wider systemic and structural determinants of disease and health inequalities [38].

Previous research has indicated that early life factors, including environmental exposures in utero, socioeconomic disadvantage, and father's occupation are associated to the individual outcomes considered in this paper [21-25]. We extended this previous work to demonstrate that combinations of early-life risk factors represented through domains are also associated to obesity-hypertension comorbidity. In particular our results support research for different cohort studies including the Hertfordshire cohort study and the Aberdeen Children of the 1950s cohort suggesting that early life socioeconomic disadvantage including paternal social class and occupational social class are key in shaping multimorbidity [22-25]. We have demonstrated that this relationship remains even after considering the role of other early life domains and adult predictors.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

The finding that the parental and family environment domain remains important indicate that there may be lifelong health impacts stemming from the interaction between children and the primary care giver, parenting styles, parental beliefs, attitudes and discipline, and wider family factors such as kin networks in childhood.

Our results support recent policy directions such as those stated in the 2022 Public Heath Wales report and the 2021 The Best Start for Life report which highlight that the relationship between the parent and child, between the child's parents, and the family's relationships with their wider family, are key components of both children's wellbeing and early child development [39,40]. Additionally, we add to literature that has found parenting and family support, parenting warmth and parenting styles to be important determinants of psychiatric comorbidity [41]. These results highlight the importance of current UK policy interventions and research endeavours such as 'A Better Start', 'Family Hubs' and 'Start for Life' programmes that aim to give children and families the best start in life [42,43]. Identifying and supporting vulnerable families to develop parenting skills, ensuring children have a healthy and safe home environment are all interventions that may reduce multimorbidity in later life. In addition, continued efforts are needed to address the wider determinants of health such as income and equitable access to good quality housing and healthy food to support people to have healthy lives for longer.

Calculating domain-specific risk scores and using these within the same prediction model was our approach to reduce dimensionality of the individual variables within the domain, while weighing their components, preserving the structure of the key domains, and providing the opportunity to move beyond simple counts of binary variables as an indicator of risk.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Additionally, this method was developed building on previous work that derived each variable (within each domain) into a binary (yes/no) outcome, and then summed these binary variables to produce domain risk scores that described a count of adversity [44]. A limitation of this previous approach was that we were required to assume that all variables within each domain and all domain risk scores carried equal weight, and this was particularly problematic when we compared the relative importance of each domain adjusting for the other domains. A further issue was that by deriving variables into a binary indicator we disregarded information contained within the original data structure and we were required to implement arbitrary cut-off points. The methods developed here represent an alternative method that adjusted for the weighting of the individual domain risk scores (in the final prediction model), whilst maintaining the original data structure of the variables.

Strengths and limitations

Data from a large cohort study allowed us to capture a wide array of biological, social, environmental, behavioural and family variables in childhood to represent five early lifecourse domains. This depth of information would not have been available from most electronic health care records in either primary or secondary care. The data also afforded the opportunity to analyse objective measures of both obesity and hypertension.

However, the cohort is representative of births occurring in Britain in 1970 and as such lacks ethnic diversity. Additionally, no differentiation was made between the burdensome impact of different diseases or disease severity to the individual. For example, mild controlled hypertension or just making the threshold for obesity is unlikely to have the same impact on an individual's quality of life compared to unmanaged hypertension or clinically severe obesity (BMI over 40). Further, despite using measured rather than self-reported BMI, the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

BMI measurement continues to have a risk of overestimating body fat in those who have muscular builds.

It is also important to consider the possibility that we have over-adjusted within our models, resulting in overadjustment bias through adjusting for mediators or colliders [45]. This was unavoidable given our particular approach to the analysis given our domains were combined risk factors rather than individual variables. Further, children's experiences across a variety of early life course domains are intersecting and we argue that research needs to be able to incorporate information from multiple domains into the same analysis. We therefore felt it was important to explore multiple domains simultaneously to attempt to disentangle the strongest predictor for obesity-hypertension comorbidity. We considered taking a more traditional, epidemiological approach where we would consider the relationship between a single domain or a single variable within a domain and odds of obesity-hypertension, controlling for select individual variables from the other domains. However, the choice of domain would have been arbitrary, and this method would not have addressed this paper's research questions.

Finally, it is important to expand the methods presented here to consider the relationship to other multimorbidity clusters and outcomes that develop a more sophisticated understanding of multimorbidity, including focussing on burdensomeness and complexity. In previous research [27] we audited the early life variables available in two other cohort datasets. The data available in these datasets provide the opportunity to validate and compare our results across cohorts.

Conclusions

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We found the most robust domains for predicting obesity-hypertension were for the parental and family environment domain, the socioeconomic factors domain and the education and academic abilities domain. Developing methods for exploring the multidetermined nature of combined childhood risk factors for health such as the work presented here, can help to challenge existing understanding of the aetiology of health, develop new ideas and solutions, and facilitate improvements in developing and recognising health as a complex and multidetermined concept from a lifecourse perspective.

Conflict of Interest

R.O. is a member of the National Institute for Health and Care Excellence (NICE) Technology Appraisal Committee, member of the NICE Decision Support Unit (DSU), and associate member of the NICE Technical Support Unit (TSU). She has served as a paid consultant to the pharmaceutical industry and international reimbursement agencies, providing unrelated methodological advice. She reports teaching fees from the Association of British Pharmaceutical Industry (ABPI). R.H. is a member of the Scientific Board of the Smith Institute for Industrial Mathematics and System Engineering.

Author Contributions

S.F., N.A., R.H., S.P., R.O., S.S. and A.B. contributed to the conceptualisation of the MELD-B project. S.S., N.A., and S.F. obtained the datasets. All authors contributed to the conceptualisation of the paper. S.S., and N.A. led the design and planning of the paper. R.O. led the design of the statistical analysis. S.S., N.A., A.B., N.Z., R.H., and R.O. supported the design, planning and reviewing of the statistical analysis. S.S. performed the statistical analysis with support from N.Z. S.S. prepared all figures and graphs. S.S., and N.A. produced the initial draft of the manuscript. All authors were involved in editing and reviewing the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

manuscript, and approved the final manuscript. S.S., N.A., and S.F. take responsibility for the data and research governance.

Data Availability Statement

The BCS70 datasets generated and analysed in the current study are available from the UK Data Archive repository (available here: http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=795).

Acknowledgement

We would like to acknowledge all other members of the MELD-B Consortium, and we thank the participants of the BCS70 cohort study. We would also like to thank Jack Welch and our other PPIE colleagues.

References

1. National Health Service. Obesity Statistics. 2022. https://www.nhs.uk/conditions/obesity/.
2. Mills KT, Stefanescu A, & He J,T. Global epidemiology of hypertension. Nat. Rev. Nephrol. 2020; 16: 223–237.
3. House of Commons Library. Obesity Statistics. 2023. https://commonslibrary.parliament.uk/research-briefings/sn03336/#:~:text=Adult%20obesity%20in%20England,is%20classified%20as%20%27overweight%27.
4. National Health Service Digital. Health Survey for England – Part 2. 2021. https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2021-part-2/adult-health-hypertension

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

5. Wolin KY, Carson K, Colditz GA. Obesity and cancer. The Oncologist. 2010; 15: 556–565. pmid:20507889
6. Eckel RH, Kahn SE, Ferrannini E, Goldfine AB, Nathan DM, Schwartz MW, et al. Endocrine Society, American Diabetes Association, & European Association for the Study of Diabetes Obesity and type 2 diabetes: what can be unified and what needs to be individualized?. Diabetes Care. 2011; 34: 1424–1430.
7. Powell-Wiley TM, Poirier P, Burke LE, Després JP, Gordon-Larsen P, Lavie CJ, et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2021; 143: e984–e1010. pmid:33882682
8. Weldegiorgis, M., Woodward, M. The impact of hypertension on chronic kidney disease and end-stage renal disease is greater in men than women: a systematic review and meta-analysis. *BMC Nephrol* 21, 506 (2020).
9. Ho FK, Celis-Morales C, Petermann-Rocha F, Parra-Soto SL, Lewsey J, Mackay, D et al. Changes over 15 years in the contribution of adiposity and smoking to deaths in England and Scotland. *BMC Public Health. 2021;* **21**, 169. doi:10.1186/s12889-021-10167-3
10. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2019:396(10258): 1223 – 1249.
11. Leggio M, Lombardi M, Caldarone E, Severi P, D'Emidio S, Armeni M *et al*. The relationship between obesity and hypertension: an updated comprehensive overview on vicious twins. *Hypertens Res*. 2017;40: 947–963.
12. Frohlich ED, Messerli FH, Reisin E, Dunn FG. The problem of obesity and hypertension. Hypertension 1983;5:III71.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

13. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA. 2003;289(1):76–9.

14. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. JAMA. 1999;282(16):1523–9.

15. Natsis M, Antza C, Doundoulakis I, Stabouli S, Kotsis V. Hypertension in Obesity: Novel Insights. *Curr Hypertens Rev*. 2020;16(1):30-36.

16. Alidu H, Owiredu WKBA, Amidu N, Gyasi-Sarpong CK, Dapare PPM, Bawah AT, et al. Hypertension and obesity comorbidities increases coronary risk, affects domains of sexual function and sexual quality of life. Int J Impot Res. 2018;30(1):8-13.

17. Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. BMC Med. 2020;18(1).

18. Haug N, Deischinger C, Gyimesi M, Kautzky-Willer A, Thurner, S Klimek, P. High-risk multimorbidity patterns on the road to cardiovascular mortality. BMC Med. 2020; 18(1): 44.

19. Jørgensen IF, Aguayo-Orozco A, Lademann M, Brunal, S. Age-stratified longitudinal study of Alzheimer's and vascular dementia patients. Alzheimer's Dementia. 2020; 16(6): 908–917.

20. Strauss VY, Jones PW, Kadam UT, Jordan, KP. Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. J Clin Epidemiol. 2014; 67(10): 1163–1171.

21. Gluckman PD, Buklijas T, Hanson MA. The developmental origins of health and disease (DOHaD) concept: Past, present, and future. In Rosenfeld CS, editor. The Epigenome and Developmental Origins of Health and Disease; 2016. pp. 1-15.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

22. Harper S, Lynch J, Smith GD. Social determinants and the decline of cardiovascular diseases: understanding the links. Annual Review of Public Health. 2011;32:39-69.

23. Humphreys J, Jameson K, Cooper C, Dennison E. Early-life predictors of future multi-morbidity: results from the Hertfordshire Cohort. Age Ageing. 2018;47(3):474-478.

24. Johnston MC, Black C, Mercer SW, Prescott GJ, Crilly MA. Impact of educational attainment on the association between social class at birth and multimorbidity in middle age in the Aberdeen Children of the 1950s cohort study. BMJ Open. 2019;9(1):e024048.

25. Gondek D, Bann D, Brown M, Hamer M, Sullivan A, Ploubidis GB. Prevalence and early-life determinants of mid-life multimorbidity: evidence from the 1970 British birth cohort. BMC Public Health. 2021;21(1):1319. doi: 10.1186/s12889-021-11291-w.

26. Stannard S, Berrington A, Paranjothy S, Owen RK, Fraser SDS, Hoyle RB, et al. A conceptual framework for characterising lifecourse determinants of multiple long-term condition multimorbidity. Journal of Multimorbidity and Comorbidity. 2023;13.

27. Stannard S, Berrington A, Fraser SDS, Paranjothy S, Hoyle RB, Owen RK, et al. Mapping domains of early life determinants of future multimorbidity across three UK longitudinal cohort studies. Sci Rep. **14**, 21454 (2024).

28. Fraser SDS, Stannard S, Holland E, Boniface M, Hoyle RB, Wilkinson R, et al. Multidisciplinary ecosystem to study lifecourse determinants and prevention of early-onset burdensome multimorbidity (MELD-B) – protocol for a research collaboration. Journal of Multimorbidity and Comorbidity. 2023;13.

29. Sullivan A, Brown M, Hamer M, Ploubidis G. Cohort Profile Update: The 1970 British Cohort Study (BCS70). International J Epidemiol. 2022; dyac148.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

30. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2nd edition. Hoboken, NJ: Wiley; 2002.
31. Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of software packages for regression models with missing variables. American Statistical Association. 2001; 55: 244–54.
32. Jakobsen JC, Gluud C, Wetterslev J. When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts. BMC Medical Research Methodology. 2017; 17: 162.
33. Collins LM, Schafer JL, Kam CM. A Comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods. 2001; 6: 330–351.
34. Mostafa T, Narayanan M, Pongiglione B. Improving the plausibility of the missing at random assumption in the 1958 British birth cohort: A pragmatic data driven approach. CLS working paper number 2020/6. UCL Centre for Longitudinal Studies.
35. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. Biometrical Journal. 2018; *60*(3): 431–449.
36. Atkinson, A.C. A note on the generalized information criterion for choice of a model. Biometrika. 1980;67, 413-418.
37. STATA - predict function. 2024. https://www.stata.com/manuals13/p_predict.pdf
38. Alwan NA, Stannard S, Berrington A, Paranjothy S, Hoyle RB, Owen RK, et al. Risk factors for ill health: How do we specify what is 'modifiable'? PLOS Glob Public Health. 2024; 4(3): e0002887.
39. Public Health Wales. First 1000 Days, Developing a Public Health Approach to Supporting Parents Technical Report. 2022.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.13.24307277; this version posted October 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

https://phw.nhs.wales/publications/publications1/developing-a-public-health-approach-to-supporting-parents-technical-report/

40. HM Government. The Best Start for Life. 2021. https://assets.publishing.service.gov.uk/media/605c5e61d3bf7f2f0d94183a/The_best_start_for_life_a_vision_for_the_1_001_critical_days.pdf
41. Kalmijn M. Intergenerational transmission of health behaviors in a changing demographic context: The case of smoking and alcohol consumption. Soc Sci Med. 2022; 296: 114736.
42. The National Lottery Community Fund. A Better Start. 2015. A Better Start | The National Lottery Community Fund (tnlcommunityfund.org.uk)
43. Department of Health and Social Care and the Department for Education. Family Hubs and Start for Life Programme Guide. 2022. Family Hubs and Start for Life Programme Guide (publishing.service.gov.uk)
44. Abstract from the 2024 Society for Social Medicine and Poupoulation Health Annual Scientific Meeting. Stannard S, Berrington A, Fraser SDS, Hoyle RB, Paranjothy S, Owen RK, et al. P03 Domains of early-life risk and obesity-hypertension comorbidity: findings from two prospective birth cohorts. J Epidemiol Community Health 2024;**78:**A48.
45. Lu H, Cole S, Platt R, Schisterman E. Revisiting Overadjustment Bias. Epidemiology. 2021; 32(5):p e22-e23.