What explains increasing diabetes prevalence across cohorts? A causal mediation analysis in the 1958 and 1970 British birth cohort studies

Laura Gimeno[1]*, Alice Goisis[1], Jennifer B. Dowd[2] and George B. Ploubidis[1]

*Presenting author

[1]Centre for Longitudinal Studies, University College London, UK

[2]Leverhulme Centre for Demographic Studies and Nuffield College, University of Oxford, Oxford, UK

Abstract

The prevalence of diabetes has increased across successive generations, yet few studies have examined the factors explaining these cohort differences. Identifying these factors could inform efforts to reverse the Generational Health Drift in diabetes. We used data from the 1958 and 1970 British birth cohort studies to quantify cohort differences in observer-measured diabetes at age 44-48. A causal mediation approach was used to evaluate the contribution of six hypothesised mediators to this gap: obesity, psychological distress, housing tenure, relative poverty, breastfeeding, and maternal smoking during pregnancy. Interventional indirect and direct effects and the randomised analogue of the proportion mediated (rPM) were estimated using the parametric g formula. Diabetes prevalence was twice as high in the 1970 cohort compared to the 1958 cohort (6.5% vs. 3%). Obesity was the most important mediator of the cohort-diabetes association (rPM = 29.8%), with psychological distress (rPM = 6.4%) and housing tenure (rPM = 6.9%) also contributing. Relative poverty, maternal smoking and breastfeeding were not identified as significant mediators. Given that rates of obesity and mental ill-health have risen and rates of home ownership have declines across cohorts born since 1970, cohort differences may persist or widen as these more recently born cohorts enter midlife. Our results suggest that tackling obesity is key to reversing the Generational Health Drift in diabetes, but that there is also a need to address the impact of broader social determinants of health to fully achieve this goal.

Theoretical Focus

The prevalence of diabetes in the United Kingdom has increased over time and across cohorts, even when comparisons are made at the same age, a trend that we have previously referred to as a Generational Health Drift.[1,2] While diabetes and cardiovascular disease are still the leading cause of DALYs in the UK,[3] no existing studies have explored what lifecourse factors explaining cohort differences in diabetes. Existing work seeking to explain period trends between the early 1990s and 2010s has focused exclusively on the consequences of changing obesity prevalence, often using cross-sectional data, and reported that increases in adiposity explained 25-75% of the observed trend.[4–6] Research seeking to understand cohort differences in health more generally is limited,[7–10] and only one study has explored drivers of cohort differences in health in the UK, focusing on mental ill-health and early life risk factors.[11]

Asking what explains cohort differences in health is effectively a mediation question. To what extent is the association between birth year (exposure X) and diabetes (outcome Y), or indeed any other health outcome, *mediated* by a hypothesised set of variables (M)? This framing is useful given that in this case X cannot be intervened upon, but M variables theoretically. Only one paper to date has approached the question of cohort differences in health using a causal mediation approach.[11] Other papers have taken traditional approaches to mediation, which make a number of strong but implicit assumptions about the absence of unmeasured X-M, X-Y and M-Y confounding, and no intermediate confounding (M-Y confounding caused by X).[12] Additionally, the traditional approach is biased in the presence of X-M interaction.[12] Given the nature of X in this question, it is hard to think of variables that would impact the allocation of a given individual to a particular cohort, meaning that the assumption of no unmeasured X-M or X-Y confounding is plausible. However, the nature of X also means that any confounders of M-Y are likely to be intermediate confounding, and X-M interaction is highly likely. In this paper, we take a causal mediation approach to ask what factors might explain cohort differences in midlife diabetes prevalence between two cohorts born 12 years apart (Late Baby Boomers and Generation X), and use the parametric g formula to estimate interventional direct and indirect effects and approximate the proportion of the X-Y association mediated by a selected set of M, while appropriately accounting for intermediate confounding.[13]

Data

We used data from the 1958 National Child Development Study (1958c) and the 1970 British Cohort Study (1970c), two nationally representative birth cohort studies of people born in Britain (England, Wales and Scotland) in one week of 1958 and 1970. Both studies have collected rich data cross the lifecourse since birth, including on socioeconomic circumstances, early life factors, and health (including biomarkers). Detailed profiles of the studies have been published elsewhere.[14–16] We used data on respondents to biomedical sweeps at age 44-45 in the 1958c (n = 9377) and at age 46-48 in the 1970c (n = 8559), and leveraged information on mediators from sweeps at birth, age 5/7, and age 42.

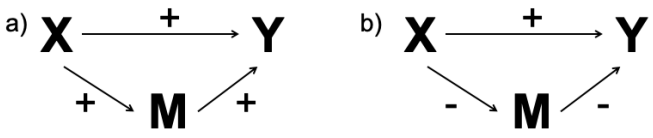
Methods

Diabetes at age 44-48 was defined using information of glycated haemoglobin (HbA1c) levels, an indicator of blood glucose levels over the previous 8-12 weeks,[17] and medication use. Individuals were classified as having diabetes if HbA1c ≥ 6.5% DCCT or if they were taking medication for diabetes.[18]

We considered the role of six hypothesised consistent mediators in explaining cohort-diabetes associations, which have changed in prevalence across cohorts and are associated with diabetes risk in the same direction (e.g., more prevalent in the 1970c and associated with higher diabetes risk, Figure 1).

Figure 5.3. Graphical depiction of consistent mediation

Note: Plus (+) signs indicate a positive association between variables, while minus (–) signs indicate a negative association. The association between X and Y is shown here as positive, since here the younger cohort is assumed to have a higher prevalence of the outcome, but could just as well be negative.



Inconsistent mediators were not considered in this study, as they act as regression suppressors, and can be thought of as factors without which the cohort gap would be larger than what was observed (e.g., educational attainment, smoking).

- **Obesity** at age 42, defined using Body Mass Index (BMI) ≥ 30 kg/m^2 based on harmonised self-reported height and weight.
- **Psychological distress** at age 42, using a cut-off score of ≥4 to the total 9-Item Malaise Inventory score, which captures symptoms of anxiety and depression and has been shown to exhibit measurement invariance across these two cohorts.[11]
- **Home ownership** at age 42, defined as owning home outright or buying with a mortgage, versus other.
- **Relative poverty** at age 42, defined as having an equivalised household income <60% of the cohort median.
- **Ever breastfed,** retrospectively reported by cohort members' mothers at age 5/7.
- **Maternal smoking,** reported by the nurse or midwife attending the cohort member's birth.

We first described the prevalence of diabetes and the different mediators in both cohorts, tested X-M and M-Y associations, and for X-M interaction using modified Poisson regressions with robust standard errors.[19] We then conducted causal mediation analysis using the parametric g formula in R (using the `CMAverse` package) to estimate interventional direct and indirect effects and the random analogue of the proportion mediation (rPM) while accounting for intermediate confounders and sex at birth.[13,20] Models were run for each mediator separately, allowing for X-M interaction, and then for all mediators simultaneously (not accounting for M-M interaction). The set of intermediate confounders included in each model is shown in **Table 1**. All analyses were conducted on imputed data among respondents to address item missingness, and observations were inverse probability weighted for non-response to account for attrition. Confidence intervals for the mediation analysis were obtained by bootstrapping.

Table 1. Intermediate confounders included in models.

Note: PD = psychological distress. Intermediate confounders are confounders of mediator-outcome associations caused by the exposure. The measurement of intermediate confounders had to temporally precede the measurement of the mediator. As adult mediators at age 42 are recognised to be a product of the lifecourse (i.e., they partially capture histories of exposure to obesity, mental ill-health, tenure and poverty), adjustments focused on childhood characteristics while adjustments for factors in adulthood were sparser to avoid controlling for factors along the causal pathway between the mediators and the outcome.

		Models						
Confounder	**Age**	**Obesity**	**PD**	**Tenure**	**Poverty**	**Breastfed**	**Mum smoked**	**All**
Gestational age	0					X		
Birthweight	0					X		
Father's social class	0	X	X	X	X	X	X	X
Mother's education	0	X	X	X	X	X	X	X
Parents' housing tenure	10/11	X	X	X	X			X
Household amenities	10/11	X	X	X	X			X
Childhood BMI	10/11	X	X	X	X			
Childhood mental health	10/11	X	X	X	X			
Childhood chronic conditions	10/11	X	X	X	X			
Childhood cognitive ability	10/11	X	X	X	X			X
Educational attainment	33/34	X	X	X	X			
Occupation	33/34	X	X	X	X			
Household income (log)	33/34	X	X					
Partnership status	33/34			X	X			

Findings

Cohort differences in the outcome and hypothesised mediators, and the association of mediators with diabetes are shown in **Table 2**. The risk of diabetes in early midlife was approximately twice as high in the 1970c as in the 1958c. This was due to both a higher proportion of cohort members with elevated HbA1c (RR = 1.60, 95% CI 1.26-2.05) and a higher proportion taking medication (RR = 2.14, 95% CI 1.68-2.74). Adjusting for sex, members of the 1970c were more likely to never have been breastfed, to have had mothers who smoked during pregnancy, and to be obese or experiencing psychological distress in adulthood. They were also less likely to be owner occupiers and more likely to be living relative poverty (despite absolute incomes being on average higher, because of growing income inequality). All hypothesised mediators were significantly associated with diabetes, adjusting for sex. Obesity was the mediator most strongly associated with the outcome, with cohort members who were obese at age 42 having a risk of diabetes that was more than 4 times higher than that of non-obese peers.

Table 2. Prevalence of diabetes and hypothesised consistent mediators in the 1958c and 1970c, cohort-mediator and mediator-outcome associations.

Note: RR = risk ratio. CI = confidence interval. Regression models adjust for sex at birth. Results are given for imputed and weighted data.

	1958c	1970c	Cohort assoc. (Ref: 1958c)	Diabetes assoc. (Ref: No diabetes)
	%	%	RR (95% CI)	RR (95% CI)
Outcome				
Diabetes (Hba1c + meds)	3.0	6.5	2.11 (1.71-2.61)	-
Mediators				
Obesity	16.0	22.6	1.43 (1.29-1.59)	4.49 (3.51-5.75)
Psych distress	13.8	23.1	1.65 (1.48-1.83)	1.75 (1.30-2.36)
Owns home	78.5	64.7	0.83 (0.81-0.86)	0.49 (0.39-0.62)
Relative poverty	18.9	23.3	1.26 (1.14-1.39)	1.60 (1.21-2.11)
Ever breastfed	68.6	38.4	0.55 (0.52-0.58)	0.72 (0.57-0.90)
Maternal smoking	39.3	46.4	1.13 (1.13-1.24)	1.36 (1.08-1.72)

Table 3 shows interventional effects from causal mediation analyses. Obesity was the most important mediator of the cohort-diabetes association, with changes in the prevalence of obesity across cohorts explaining nearly a third of the difference in diabetes prevalence (rPM = 29.8%). This was nearly 5 times the amount explained by changes in the prevalence of psychological distress and housing tenure, which also appeared to contribute to the gap (rPM = 6.4% and 6.9% respectively). Changes in the prevalence of relative poverty and maternal smoking did not appear to explain much of the cohort gap, and there was no evidence for an impact of changes in breastfeeding practices. When all mediators were considered simultaneously, they explained approximately 40% of the cohort difference in diabetes (rPM = 41.9%) suggesting that while obesity was the main driver, other consistent mediators also played a part.

Table 3. Interventional effects from causal mediation analysis using the parametric g formula.

Mediator	ITE (95% CI)	IDE (95% CI)	IIE (95% CI)	rPM (95% CI)
Obesity	2.27 (1.76, 2.85)	1.89 (1.47, 2.36)	1.20 (1.10, 1.39)	29.8 (16.2, 49.5)
Psychological distress	2.12 (1.69, 2.66)	2.05 (1.63, 2.57)	1.04 (1.00, 1.08)	6.4 (0.1, 14.5)
Housing tenure	2.03 (1.57, 2.57)	1.96 (1.52, 2.49)	1.04 (0.99, 1.57)	6.9 (-0.1, 13.6)
Relative poverty	2.10 (1.68, 2.62)	2.09 (1.67, 2.60)	1.01 (0.99, 1.03)	1.2 (-0.1, 5.0)
Breastfeeding	2.08 (1.64, 2.62)	2.04 (1.60, 2.60)	1.02 (0.93, 1.10)	3.2 (-14.4, 17.2)
Maternal smoking	2.08 (1.64, 2.63)	2.05 (1.62, 2.58)	1.02 (0.99, 1.04)	2.9 (-0.1, 8.2)
All mediators	2.21 (1.74, 2.80)	1.71 (1.33, 2.16)	1.30 (1.15, 1.51)	41.9 (24.2, 62.8)

Note: ITE = Interventional Total Effect. IDE = Interventional Direct Effect. IIE = Interventional Indirect Effect. rPM = randomised analogue of the proportion mediated, expressed in percentage form. CI = confidence interval. Models adjust for sex at birth and intermediate confounders specified in Table 1. Models for individual mediators allow for exposure-mediator interaction. The final model considering all mediators simultaneously did not account for possible mediator-mediator interaction.

Interpretation

Increases in the prevalence of obesity and psychological distress, and declines in home ownership contributed to the cohort gap in diabetes in midlife between the 1958c and 1970c. The prevalence of obesity continued to increase across cohorts born since 1970, and in more recently born generations, cohort differences in obesity and overweight prevalence have emerged in childhood and adolescence.[21] Similarly, evidence from more recently born British birth cohorts indicates that the prevalence of ill-health has continued to rise.[22] Rates of home ownership, an important indicator of wealth and material advantage, have continued to fall across cohorts born since 1970. For instance, only half of those born in 1989-1990 were owner occupiers at age 32, as opposed to 60% of those born in 1970 at age 30.[23,24] While relative poverty, breastfeeding and maternal smoking were not identified as important mediators of the cohort gap in diabetes, this does not mean that they are not important determinants of health to intervene upon, and they may still be important for other outcomes. For instance, because members of the 1970c had higher incomes on average, our measure of relative poverty would be expected to act on diabetes through purely psychosocial pathways, and it may be that these pathways are more important for other health outcomes (e.g., mental health) than for diabetes, where material factors may matter more (e.g., wealth captured in part by home ownership). Maternal smoking and breastfeeding have encouragingly improved over time,[25,26] while the relative poverty rate has remained stable over the last decade.[27] Overall, trends in mediators across more recent cohorts therefore suggest that the cohort gap in midlife diabetes prevalence will likely persist as these more recent generations enter older age. There is already some suggestion that this may be occurring as the number of people diagnosed with early-onset type 2 diabetes (<40 years) increased by 39% between 2016/17 and 2022/23.[28]

Whether diabetes prevalence in midlife does ultimately continue to increase across generations is still to play for and depends on how much the prevalence of these consistent mediators can be reduced, the extent to which the link between these mediators and diabetes can be weakened, or whether improvements in other risk factors can offset the impact of increasing mediator prevalence. Our findings suggest that achieving declines in obesity prevalence would likely help to close the cohort gap in diabetes, and the recent approval of GLP-1 agonists for weight management in the NHS may be one way to achieve this. However, our findings also suggest that the impact of wider determinants of health should not be forgotten.

References

1. Gimeno L, Goisis A, Dowd JB, Ploubidis GB. Cohort Differences in Physical Health and Disability in the United States and Europe. *J Gerontol B Psychol Sci Soc Sci*. 2024;**79**(8):gbae113.
2. Gimeno L, Moreno Agostino D, Danka M, et al. The Generational Health Drift: A Lifecourse Perspective. *Under Review*. 2024;
3. World Health Organization. Global Health Estimates 2021: Disease burden by cause, age, sex, by country and by region, 2000-2021 [Internet]. Geneva: World Health Organization; 2024. Available from: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys
4. Hardoon SL, Morris RW, Thomas MC, Wannamethee SG, Lennon LT, Whincup PH. Is the Recent Rise in Type 2 Diabetes Incidence From 1984 to 2007 Explained by the Trend in Increasing BMI?: Evidence from a prospective study of British men. *Diabetes Care*. 2010;**33**(7):1494–1496.
5. Stokes A, Preston SH. The contribution of rising adiposity to the increasing prevalence of diabetes in the United States. *Prev Med*. 2017;**101**:91–95.
6. Menke A, Rust KF, Fradkin J, Cheng YJ, Cowie CC. Associations between trends in race/ethnicity, aging, and body mass index with diabetes prevalence in the United States: a series of cross-sectional studies. *Ann Intern Med*. 2014;**161**(5):328–335.
7. Canizares M, Hogg-Johnson S, Gignac MAM, Glazier RH, Badley EM. Increasing Trajectories of Multimorbidity Over Time: Birth Cohort Differences and the Role of Changes in Obesity and Income. *J Gerontol B Psychol Sci Soc Sci*. 2018;**73**(7):1303–1314.
8. Badley E, Canizares M, Peruccio AV, Hogg-Johnson S, Gignac MAM. Benefits Gained, Benefits Lost: Comparing Baby Boomers to Other Generations in a Longitudinal Cohort Study of Self-Rated Health. *Milbank Q*. 2015;**93**(1):40–72.
9. Zheng H, Dirlam J, Choi Y, George L. Understanding the health decline of Americans in boomers to millennials. *Soc Sci Med*. 2023;**337**:116282.
10. Zheng H. A New Look at Cohort Trend and Underlying Mechanisms in Cognitive Functioning. *J Gerontol B Psychol Sci Soc Sci*. 2021;**76**(8):1652–1663.
11. Ploubidis GB, Sullivan A, Brown M, Goodman A. Psychological distress in mid-life: evidence from the 1958 and 1970 British birth cohorts. *Psychol Med*. 2017;**47**(2):291–303.
12. Daniel RM, De Stavola BL. Mediation Analysis for Life Course Studies. *Pathways to Health*. Dordrecht: Springer Nature; 2019. p. 1–40.
13. Daniel RM, De Stavola BL, Cousens SN. Gformula: Estimating Causal Effects in the Presence of Time-Varying Confounding or Mediation using the G-Computation Formula. *Stata J*. 2011;**11**(4):479–517.
14. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol*. 2006;**35**(1):34–41.
15. Elliott J, Shepherd P. Cohort Profile: 1970 British Birth Cohort (BCS70). *Int J Epidemiol*. 2006;**35**(4):836–843.

16. Sullivan A, Brown M, Hamer M, Ploubidis GB. Cohort Profile Update: The 1970 British Cohort Study (BCS70). *Int J Epidemiol*. 2023;**52**(3):e179–e186.
17. World Health Organization. Use of glycated haemoglobin (HbA1c) in diagnosis of diabetes mellitus [Internet]. Geneva: World Health Organization; 2011 Feb p. 25. Available from: https://www.who.int/publications/i/item/use-of-glycated-haemoglobin-(-hba1c)-in-diagnosis-of-diabetes-mellitus
18. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med*. 2005;**24**(19):2911–2935.
19. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;**159**(7):702–706.
20. Shi B, Choirat C, Coull BA, VanderWeele TJ, Valeri L. CMAverse: A Suite of Functions for Reproducible Causal Mediation Analyses. *Epidemiol*. 2021;**32**(5):e20–e22.
21. Johnson W, Li L, Kuh D, Hardy R. How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts. *PLOS Med*. 2015;**12**(5):e1001828.
22. McElroy E, Tibber M, Fearon P, Patalay P, Ploubidis GB. Socioeconomic and sex inequalities in parent-reported adolescent mental ill-health: time trends in four British birth cohorts. *J Child Psychol Psychiatry*. 2023;**64**(5):758–767.
23. Wu AF-W, Henderson M, Brown M, et al. Cohort Profile: Next Steps—the longitudinal study of people in England born in 1989–90. *Int J Epidemiol*. 2024;**53**(6):dyae152.
24. Smith K, Ferri E. Chapter 7: Housing. In: Ferri E, Bynner J, Wadsworth M, editors. *Changing Britain, Changing Lives: Three Generations At the Turn of the Century*. London: Institute of Education, University of London; 2003. p. 194–206.
25. Wright L, Kock L, Tattan-Birch H, Bann D. Cigarette Smoking Across Life from 1946 to 2018: Harmonisation of Four British Birth Cohort Studies [Internet]. medRxiv; 2024 [cited 2025 Sep 26]. p. 2024.12.06.24318606. Available from: https://www.medrxiv.org/content/10.1101/2024.12.06.24318606v1
26. Office for Health Imrpovement & Disparities. Official Statistics: Breastfeeding at 6 to 8 weeks, 2023 to 2024 statistical commentary [Internet]. *GOV.UK* 2024 [cited 2025 Oct 6]. Available from: https://www.gov.uk/government/statistics/breastfeeding-at-6-to-8-weeks-after-birth-annual-data-april-2023-to-march-2024/breastfeeding-at-6-to-8-weeks-2023-to-2024-statistical-commentary
27. Francis-Devine B. Poverty in the UK: Statistics [Internet]. London: House of Commons Library; 2025 Apr. Report No.: 7096. Available from: https://researchbriefings.files.parliament.uk/documents/SN07096/SN07096.pdf
28. Diabetes UK. Reverse the Trend: Reducing type 2 diabetes in young people [Internet]. London: Diabetes UK; 2024 May p. 19. Available from: https://www.diabetes.org.uk/sites/default/files/2024-06/Reverse%20the%20Trend%20report%20V2.pdf?