OPEN

How Do Early Weight Trajectories Explain Social Inequalities in Lung Function in Children With Cystic Fibrosis?

A Longitudinal Interventional Disparity Effects Analysis With Time-varying Mediators and Intermediate Confounders

Daniela K. Schlüter[a], Ruth H. Keogh[b], Rhian M. Daniel[c], Schadrac C. Agbla[d], and David Taylor-Robinson[a]

Abstract

Background: Children with cystic fibrosis (CF) from socioeconomically deprived areas have poorer growth, worse lung function, and shorter life expectancy than their less-deprived peers. While early growth is associated with lung function around age 6, it is unclear whether improving early growth in the most deprived children reduces inequalities in lung function.

Methods: We used data from the UK CF Registry, tracking children born 2000–2010 up to 2016. We extended the interventional disparity effects approach to the setting of a longitudinally measured mediator. Applying this approach, we estimated the association between socioeconomic deprivation (children in the least vs. most deprived population quintile; exposure) and lung function at first measurement (ages 6–8, outcome), and the role of early weight trajectories (ages 0–6) as mediators of this relationship. We adjusted for baseline confounding by sex, birthyear, and genotype and time-varying intermediate confounding by lung infection.

Results: The study included 853 children, with 165 children from the least and 172 from the most deprived quintiles. The average lung function difference between the least and most deprived quintiles was 4.5% of predicted forced expiratory volume in 1 second (95% confidence interval: 1.1-7.9). If the distribution of early weight trajectories in the most deprived children matched that in the least deprived children, this difference would reduce to 4% (95% confidence interval: 0.57- 7.4).

Conclusion: Socioeconomic deprivation has a strong negative association with lung function for children with CF. We estimate that improving early weight trajectories in the most deprived children would only marginally reduce these inequalities.

Keywords: Cystic fibrosis; Deprivation; Health disparities; Health inequalities; Interventional disparity measures; Interventional effects; Socioeconomic conditions

(Epidemiology 2025;36: 275–285)

Submitted December 14, 2021; accepted November 22, 2024

From the [a]Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, United Kingdom; [b]Department of Medical Statistics and Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, London, United Kingdom; [c]Division of Population Medicine, School of Medicine, Cardiff University, Cardiff, United Kingdom; and [d]Department of Health Data Science, University of Liverpool, Liverpool, United Kingdom.

D.K.S. and D.T.R were supported by the Strategic Research Centre "CF-EpiNet: Harnessing data to improve lives" funded by the Cystic Fibrosis Trust. D.T.R. is funded by the MRC on a Clinician Scientist Fellowship (MR/P008577/1). R.H.K. is supported by a UKRI Future Leaders Fellowship (MR/S017968/1). R.M.D. is supported by a Sir Henry Dale Fellowship from the Wellcome Trust and The Royal Society (107617/Z/15/Z).

Disclosure: The authors report no conflicts of interest.

The data used in this study are from the UK CF Registry and are available from the UK CF Registry on reasonable request. The computing code can be found in the Supplementary Material.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Daniela K. Schlüter, Department of Public Health, Policy and Systems, University of Liverpool, Liverpool L69 3GF, United Kingdom. E-mail: d.k.schlueter@liverpool.ac.uk.

ISSN: 1044-3983/25/362-275285
DOI: 10.1097/EDE.0000000000001826

Cystic fibrosis (CF) is a serious inherited condition affecting approximately one in 2500 babies in the UK and over 100,000 people worldwide.[1,2] CF is a genetic disease caused by variants of the CF transmembrane conductance regulator gene. Symptoms include the buildup of thick, sticky mucus in various organs, in particular those of the respiratory and digestive systems. There is no cure for CF and most people die prematurely from their disease through respiratory failure. The median life expectancy for a baby born with CF in the UK in 2022 is 56 years.[3]

Nutritional management is an essential part of the multidisciplinary care people with CF receive to preserve health and slow decline.[4] In the majority of people with CF, mucus

blocks the passage of enzymes from the pancreas to the small intestine. In addition, an imbalance of the secretion of bicarbonate and acid in the digestive system and reduced gastrointestinal motility all lead to malabsorption and maldigestion of fat, protein, and fat-soluble vitamins.[4–6] Thus, achieving and maintaining good nutritional status and, therefore, healthy growth are difficult for many people with CF and are further impacted by increased energy requirements due to infections and worsening lung function.[4]

Child growth and nutritional status in early life as measured by weight-for-age, height-for-age, and percent ideal body weight, are positively associated with lung function at age 6 for children with CF and it has been suggested that interventions aimed at improving child growth and nutrition in early life may improve later pulmonary function.[7] In children with CF diagnosed within 6 months after birth and who were malnourished at diagnosis, weight-for-age improvements to reach their birthweight percentile by age 2 is associated with improved lung function at age 6, while maintenance of weight-for-age, height-for-age and body mass index (BMI) percentiles from ages 2 to 6 has no strong additional effect on pulmonary function.[8]

Children with CF from socioeconomically deprived areas in the UK have worse outcomes than those from more affluent areas.[9–11] They have lower nutritional status as measured by weight-for-age, height-for-age, and BMI from infancy compared with children from less deprived areas. Lower lung function is also evident in children from more deprived areas as soon as this can be routinely measured at around age 6.[9] Although there is some evidence from other countries that these inequalities, the unfair differences in outcomes between children with CF from socioeconomically more and less deprived backgrounds, may widen with age,[12,13] no evidence has emerged that this is the case in the UK.[14] Therefore, understanding the pathways that lead to inequalities in lung function at age 6 is crucial to identifying potential targets for interventions that could reduce inequalities in CF across the life-course.

In this study, we were interested in examining how unequal growth and nutrition in the first 6 years of life can explain inequalities in lung function (when first measured around age 6) between groups of children from areas with different levels of socioeconomic deprivation in the UK CF population. More specifically, we estimated the association between socioeconomic deprivation at birth and lung function at first measurement between ages 6 and 8, and the extent to which this could be reduced if we were to intervene to improve the weight trajectories, as markers of growth and nutrition, in the most deprived children such that they have the same distribution as those in the least deprived children.

To estimate the effects of interest, we build on recent developments in the causal inference literature that have led to the definition of "interventional disparity effects"[15] and extend this approach to the setting with a longitudinally measured mediator.

METHODS

Data Source

We used data from the UK CF Registry, which has been collecting information since the 1990s and is estimated to cover 99% of the current UK CF population.[16] It records data from comprehensive annual reviews of individuals with CF, including assessments of lung function, nutritional status, and the microbiology of respiratory tract secretions since the previous review. Demographic details, including socioeconomic status, are also documented. Our dataset includes follow-up data up to the end of 2016.

Setting and Participants

We included children who were born after 2000, were diagnosed by newborn screening, had at least one lung function measure between ages 6 and 8 years and at least one weight measurement between birth and less than 7 years of age, i.e., before or at the same time as the lung function measure. We excluded individuals who did not have complete data on sex, genotype, year of birth, and socioeconomic conditions.

Outcome, Exposure, Mediator, and Confounders

Figure 1 shows the outcome, exposure, mediator and confounders, their assumed relationships, and when they are measured. Our outcome of interest was the first available lung function measurement taken between age 6 and up to but excluding age 9 (i.e., between ages 6 and 8) given by percent of predicted forced expiratory volume in one second (%FEV1) based on Global Lung Function Initiative reference equations, which adjust for sex, age, height, and ethnicity.[17] %FEV1 is treated as a continuous variable.

Our exposure of interest was socioeconomic conditions at birth. Measures of individual socioeconomic conditions are not available in the UK CF Registry. As a proxy for socioeconomic conditions at birth, we used the Index of Multiple Deprivation, a relative measure of deprivation at small-area level linked to first home-address postcode recorded in the registry. This index combines information on income, employment, health, education, access to services, crime/safety, and environment/housing.[18–21] Since the domains differ between England, Wales, Scotland, and Northern Ireland, we calculated country-specific Index of Multiple Deprivation z-scores[11] to obtain a comparable measure of deprivation across these regions. All children were included in the study, but our main interest was the difference in outcome between children in the highest and lowest quintile of deprivation (see eAppendix; http://links.lww.com/EDE/C210 for details). We, therefore, converted z-scores to quintiles.

Our mediator of interest was weight measured at annual review visits from birth up to but excluding age 7 as a proxy for growth and nutrition. We used weight-for-age z-scores based on the UK WHO reference[22] as a continuous variable. BMI percentile is the most appropriate measure of nutrition in

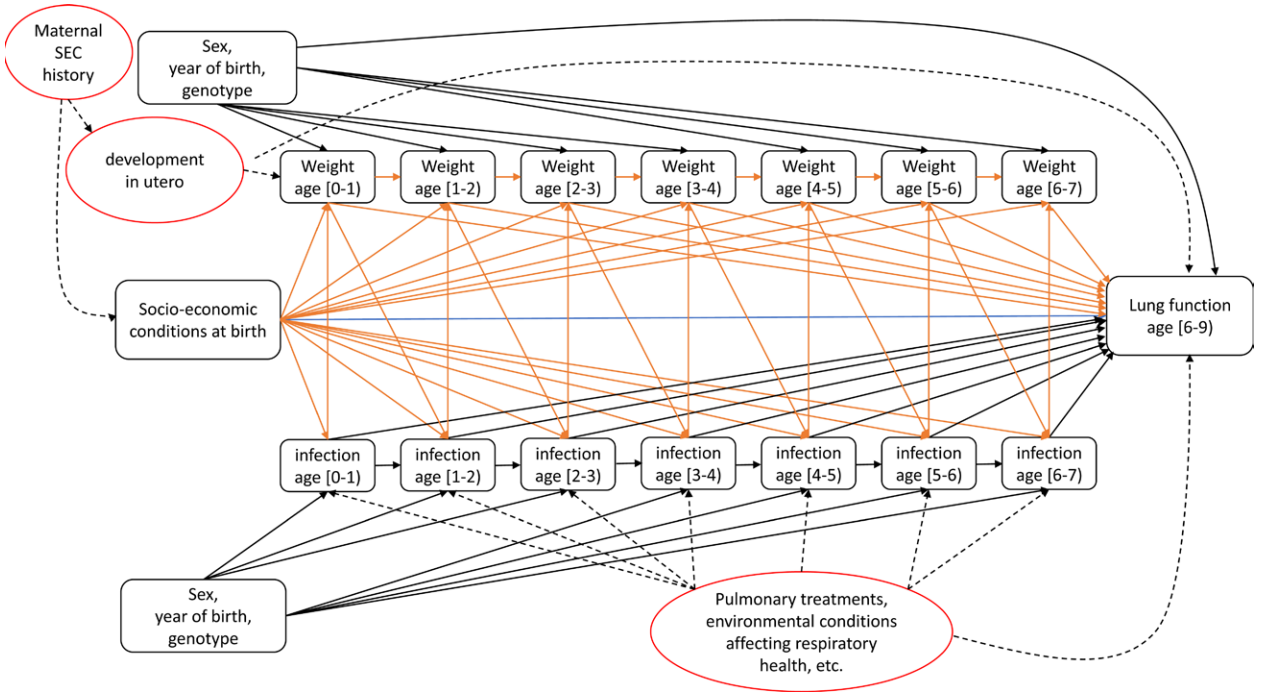

Figure 1. Assumed causal relationships between socioeconomic conditions at birth, early weight as marker of growth and nutrition, first lung function measurement and their common causes including baseline and time-varying confounders of the mediator-outcome relationship that are also affected by exposure, and latent variables, which are assumed present but not observed (ellipses). SEC indicates socio-economic conditions.

children with CF; however, it cannot be used in children under 2 years of age as no reference values are available in this age group.[23] BMI might also mask the effect of stunting because height is adjusted. For these reasons, we used weight-for-age z-scores in this study.

Confounders were identified based on our assumed causal relationships between socioeconomic conditions, weight, lung function, and their common causes (Figure 1). We included sex (male/female), birthyear (continuous), and genotype (categorical: number of F508del alleles – 0,1,2. F508del is the most common mutation in CF and people with two copies of F508del have generally the most severe and those with no F508del copy the least severe disease). Respiratory infection may be a time-varying confounder of the mediator–outcome relationship affected by the exposure.[9] We only considered infection by *Pseudomonas aeruginosa,* the most common cause of infection in the UK CF population under study.[24] At any annual review visit, it is recorded whether an infection with *P. aeruginosa* had been present since the last annual review visit.

Statistical Methods

Rationale and Effects of Interest

We were interested in estimating the association between socioeconomic conditions at birth (high vs. low) and lung function at first measurement (between ages 6 and 8) and the extent to which it can be explained by early weight trajectories. Figure 1 illustrates the assumed causal relationships between socioeconomic conditions, weight, first lung function measurement and their common causes, including baseline and time-varying confounders of the mediator–outcome relationship that are also affected by exposure (intermediate confounders), and latent variables, which are assumed present but not observed.

Deprivation at birth as measured by the Index of Multiple Deprivation is multifactorial and determined by a number of historical factors, and it is difficult to conceive of a single intervention (specifically one that would satisfy the consistency condition[25–27]) that could plausibly be implemented for subsets of the CF population and could shift this exposure. However, one can conceive of interventions that could be implemented to influence weight gain for children with CF, which is downstream of the effect of socioeconomic conditions at birth and also related to lung function at around age 6. This motivates defining our effects of interest in terms of interventional disparity effects, which are a variant of interventional effects.[28,29] They are based on the ideas of counterfactual disparity measures by VanderWeele and Robinson[30] and Naimi et al[31] and were formally defined in the context of multiple mediators by Micali et al.[15] Interventional disparity effect measures define the indirect effects in terms of the effect on the expected outcome in the exposed group if we could intervene on a mediator so as to shift the mediator distribution in the exposed to that in the unexposed;

direct effects are defined in terms of the association that would remain between exposure and expected outcome if we could intervene on a mediator so as to shift the mediator distribution in the exposed to that in the unexposed. In our context, this enables us to estimate how the association between socioeconomic conditions at birth and lung function at first measurement between ages 6 and 8 would change if we could intervene on weight trajectory up to age 6. In contrast to interventional effects, interventional disparity measures do not give a causal interpretation to the estimate of the effect of the exposure on the outcome. To ensure this is clear, we will refrain from using the term "direct effect" throughout the rest of the article and will rather refer to it as "residual association."

We further extend the approach of interventional disparity measures by considering a hypothetical intervention on a longitudinal mediator trajectory. This is similar to the extension of interventional effects to incorporate time-varying exposures and mediators.[32,33]

Let X be the exposure of interest, here socioeconomic conditions at birth. We are interested in the contrast between those born in the least and most deprived quintiles of the population. Let $X = 0$ indicate that children were born in the least deprived quintile and $X = 1$ indicate that children were born in the most deprived quintile. Let Y be the outcome (first lung function measurement at age [6–9]) and $\boldsymbol{C}$ the vector of baseline covariates (sex, year of birth, and genotype). Let M_t be a time-varying mediator (weight) and L_t a time-varying confounder of the mediator–outcome relationship affected by the exposure (infection), where t denotes age. The complete set of mediator measurements up to age T, when the outcome is measured, is denoted $\bar{M} = (M_1, \ldots, M_T)$ and similarly $\bar{L} = (L_1, \ldots, L_T)$. Let $\overline{M_c^x}$ denote a random draw from the distribution of mediator trajectories in those where $X = x$ and $\boldsymbol{C} = \boldsymbol{c}$. For justification of conditioning the draws on $\boldsymbol{C}$, see Supplementary Material; http://links.lww.com/EDE/C210. The interventional disparity measure residual association (*IDM-RA*) is the difference in the expected lung function between children from the least and most deprived quintiles, if the weight trajectories of all children were randomly drawn from the distribution of those observed in the least deprived children within strata of baseline confounders, standardized to the distribution of baseline confounders across both groups of children. Formally:

$$IDM - RA = \sum_c \left\{ E\left[Y\left(\overline{M_c^0} \right) \mid X = 0, C = c \right] - E[Y\left(\overline{M_c^0} \right) | X = 1, C = c)] \right\} P(C = c)$$

The interventional disparity measure indirect effect (*IDM-IE*) is the difference in expected lung function amongst the most deprived children if we were to randomly draw their weight trajectories from the distribution of these in least deprived children versus if we were to randomly draw their weight trajectories from the distribution of these in the most deprived children, again standardized to the distribution of baseline confounders across both groups of children. Formally:

$$IDM - IE = \sum_c \left\{ E\left[Y\left(\overline{M_c^0} \right) \mid X = 1, C = c \right] - E[Y\left(\overline{M_c^1} \right) | X = 1, C = c)] \right\} P(C = c)$$

Finally, the total adjusted association (*Adj-TA*) is the difference in expected lung function between the least and most deprived children standardized to the distribution of baseline confounders across the two groups of children. Formally:

$$\begin{aligned} Adj - TA &= \textstyle\sum_c \{E[Y \mid X = 0, C = c] - E[Y|X = 1, C = c]\} P(C = c) \\ &= IDM - RA + IDM - IE \end{aligned}$$

In all the estimands corresponding integrals and densities are used for continuous $\boldsymbol{C}$.

The *IDM-IE* captures the entire effect of socioeconomic deprivation on lung function via weight including any effect from deprivation to infection to weight to lung function. However, it does not include any effect of deprivation on lung function via infection if this is not also mediated by weight. For justification of this, see Supplementary Material; http://links.lww.com/EDE/C210.

Assumptions Required for Estimation Using Registry Data

We estimate the estimands specified above using UK CF Registry data. This requires several assumptions: (1) no interference, (2) consistency, (3) positivity, and (4) no unmeasured confounding of the mediator–outcome relationship. For explanations of assumptions 1–3 in this context, see eAppendix; http://links.lww.com/EDE/C210. The fourth assumption of "no unmeasured confounding of the mediator–outcome relationship" includes both baseline and time-dependent confounders as measured confounders. We include sex, genotype, and year of birth as baseline mediator–outcome confounders, and *P. aeruginosa* infection as a time-dependent confounder. However, we are not able to control for development in-utero, which we believe could confound the relationship between weight and lung function (Figure 1) but for which we do not have proxy measures in the registry data. We will assess the implications of this in the discussion.

As we do not aim to estimate the causal effect of the exposure on the outcome, the assumptions required for identifiability of these effects are weaker than for other causal mediation estimands.[15] In particular, we do not have to assume no unmeasured confounding of the exposure–outcome relationship.

Assuming all four conditions are met, we can identify the effects of interest from the observed data (eAppendix; http://links.lww.com/EDE/C210).

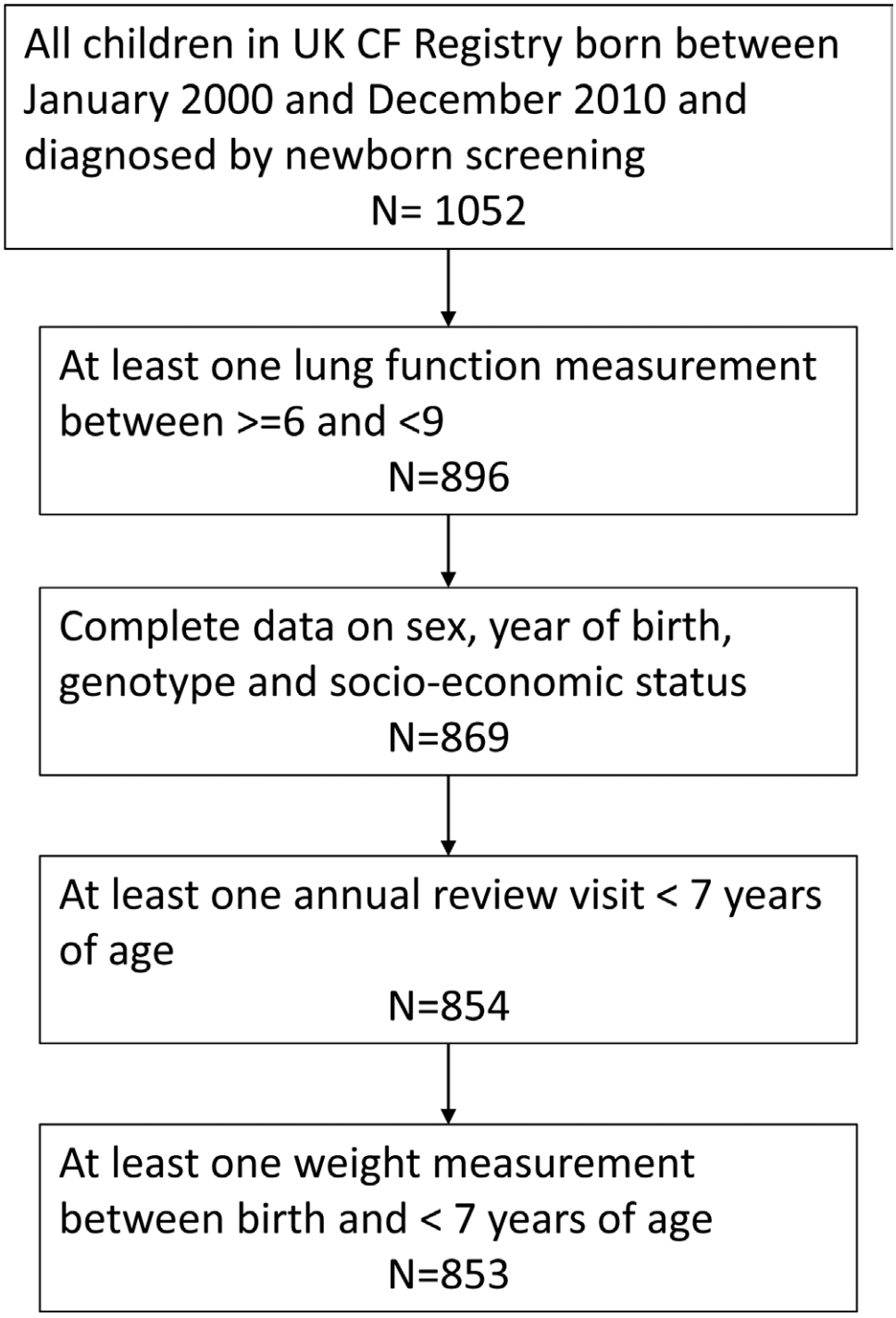

Figure 2. Derivation of the study population.

Estimation

Full details of the estimation procedure are given in the eAppendix; http://links.lww.com/EDE/C210. We provide an overview here. We discretized time and assumed that one weight and infection measure per year were recorded as is the guideline in the UK. Our discrete time intervals ran from one birthday up to but not including the next birthday. For example, measurements recorded after a child's sixth birthday, but before they turned 7, were labeled as measurements at age 6. We used Monte Carlo simulation to estimate the effects of interest.[15,29] This required a number of parametric models to be specified: (1) for the expected outcome Y conditional on exposure, a mediator at ages 0–6, baseline confounders and time-varying intermediate confounders at ages 0–6; (2) for

Table 1. Demographics and Clinical Characteristics of the Study Population Stratified by Deprivation Quintile

	1 = Least Deprived	2	3	4	5 = Most Deprived	Overall
n	165	164	180	172	172	853
Male (%)	95 (58)	76 (46)	89 (49)	79 (46)	94 (55)	433 (51)
F508 class (%)						
Heterozygous	72 (44)	74 (45)	72 (40)	67 (39)	67 (39)	352 (41)
Homozygous	77 (47)	74 (45)	91 (51)	90 (52)	86 (50)	418 (49)
Other	16 (10)	16 (10)	17 (9)	15 (9)	19 (11)	83 (10)
Cohort 2007–2010 (%)	107 (65)	110 (67)	120 (67)	115 (67)	115 (67)	567 (67)
First measured lung function mean (SD)	93 (14)	91 (14)	91 (15)	89 (17)	89 (18)	91 (16)
Age at first lung function measure (%)						
6	138 (83)	142 (86)	152 (84)	136 (79)	143 (83)	711 (83)
7	24 (15)	18 (11)	23 (13)	32 (19)	25 (15)	122 (14)
8	3 (2)	4 (2)	5 (3)	4 (2)	4 (2)	20 (2)
Ever diagnosed with pseudomonas (%)	98 (59)	105 (64)	108 (60)	101 (59)	95 (55)	507 (59)
Age at first pseudomonas diagnosis mean (SD)	2.8 (1.6)	3.2 (1.7)	3.0 (1.8)	3.0 (1.7)	2.9 (1.7)	3.0 (1.7)
Number of years with missing weight data (%)						
0	25 (15)	25 (15)	30 (17)	31 (18)	28 (16)	139 (16)
1	65 (39)	62 (38)	60 (33)	56 (33)	61 (36)	304 (36)
2	44 (27)	51 (31)	38 (21)	41 (24)	43 (25)	217 (25)
3	20 (12)	12 (7)	27 (15)	24 (14)	23 (13)	106 (12)
4–6	11 (7)	14 (9)	25 (14)	20 (12)	17 (10)	87 (10)
Weight z-scores mean (SD)						
At age (0–1)	−0.35 (1.2)	−0.54 (1.5)	−0.50 (1.5)	−0.45 (1.4)	−0.66 (1.4)	−0.49 (1.4)
At age (1–2)	0.32 (0.88)	0.40 (1.1)	0.40 (1.0)	0.27 (0.89)	0.08 (1.1)	0.30 (1.0)
At age (2–3)	0.10 (0.92)	0.26 (0.95)	0.37 (0.92)	0.17 (0.88)	0.14 (1.1)	0.21 (0.96)
At age (3–4)	0.11 (0.92)	0.11 (0.97)	0.18 (0.97)	0.09 (0.94)	−0.04 (1.1)	0.09 (0.97)
At age (4–5)	−0.17 (0.97)	−0.00 (1.1)	0.03 (1.0)	−0.07 (0.93)	−0.17 (1.1)	−0.08 (1.0)
At age (5–6)	−0.07 (0.95)	−0.08 (1.0)	−0.03 (0.98)	−0.15 (1.0)	−0.14 (1.1)	−0.09 (1.0)
At age (6–7)	−0.08 (0.96)	0.01 (1.1)	−0.00 (0.98)	−0.03 (1.0)	−0.18 (1.2)	−0.06 (1.1)

SD indicates standard deviation.

the distribution of the intermediate confounder L_t at each age 0–6 conditional on exposure, baseline confounders, mediator value at the previous age and previous intermediate confounder value; and (3) for the distribution of the mediator M_t at ages 0–6 conditional on exposure, previous mediator value, baseline confounders and intermediate confounder value at the same age. We fitted these models to data from all children in the study. See eAppendix; http://links.lww.com/EDE/C210 for the detailed model specification and goodness of model fit.

For each individual i in the least and most deprived groups, random values were drawn from the fitted distributions for the mediator M_t and the intermediate confounder L_t at each age $t = 0,\ldots, 6$ sequentially, conditional on baseline confounder values $\boldsymbol{c}_i$, previous mediator and intermediate confounder values, and $X = 0,1$. The expected outcome was then sampled for each individual based on the fitted distribution for Y, conditional on their baseline confounder values $\boldsymbol{c}_i$, $X = 0,1$, and the previously generated draws m_t and l_t for $t = 0–6$. The above process was repeated 1200 times to obtain 1200 draws of the mediators, intermediate confounders, and outcome Y (under $X = 0,1$) for each individual. We chose the number of repetitions (1,200) to render negligible the Monte Carlo error at two decimal places. We then averaged the conditional mean outcomes to give the standardized means of the potential outcomes, which we then inserted into the equations for the *IDM-RA, IDM-IE,* and *Adj-TA* to estimate the effects of interest. We used nonparametric bootstrap (1000 bootstrap samples) for variance estimation.

Duplicate Entries and Missing Data

In some cases, children had multiple measurements in a year, for example, if they moved between clinics. In these cases, we selected the data from the visit closest to the child's next birthday that had complete data on weight and infection status.

We used multiple multivariate imputations by chained equations[34] to impute data if weight and infection status were missing for any years (10 imputed datasets). See eAppendix; http://links.lww.com/EDE/C210 for the imputation model specification and goodness of fit. We applied the estimation algorithm to each imputed dataset and pooled point estimates across all 10 datasets. To estimate the 95% confidence intervals (CIs), we applied standard multiple imputation procedures[35] using the variance estimated across bootstrap samples

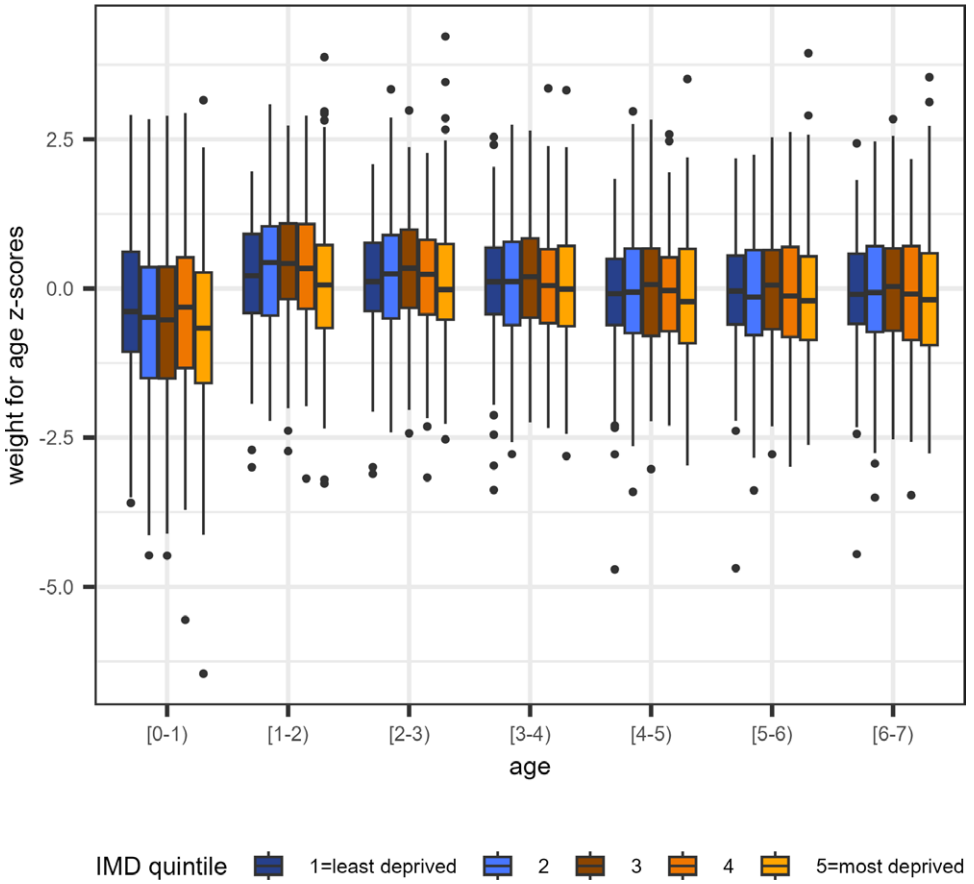

Figure 3. Weight for age z-scores by age and Index of Multiple Deprivation (IMD) quintile as given in Table 1.

as an approximation of the within imputation variance.[36] See eAppendix; http://links.lww.com/EDE/C210 for full details.

Sensitivity Analyses

Lung function is measured with considerable error, due to natural day-to-day variation and measurement error. In a sensitivity analysis, we used all available lung function data up to age 12 for individuals in the study population and fitted a linear mixed effects model with random intercept and random slope adjusted for sex, genotype, year of birth, and deprivation quintile. Based on this model, we predicted lung function at age 6 for all individuals which we used as an alternative outcome. This also explores the sensitivity of the results to the variability in the time point of outcome measurement.

Due to the relatively small sample size, the parametric models specified for the estimation procedure did not include interaction terms or higher order terms but we conducted a sensitivity analysis, where we included interaction terms between deprivation and all covariates in the parametric models (1)–(3) described above.

Ethical Considerations

NHS research ethics approval (Huntingdon Research Ethics Committee 07/Q0104/2) was granted for the collection of data into the UK CF Registry. The CF Registry committee approved the use of anonymized data in this study.

RESULTS

Study Population

In our population, 853 children met the eligibility criteria (Figure 2) including 165 in the least deprived quartile and 172 children in the most deprived quintiles (Table 1).

Only 139 individuals (16%) had complete weight data across all ages (15% and 16% in the least and most deprived quintiles, respectively) (Table 1). A total of 766 individuals (90%) had three or more measurements, including 154 (93%) and 155 (90%) individuals from the least and most deprived quintiles, respectively. Missing measurements were primarily due to missing review visits, although a small number of review visits (87) did not record weight data (18 and 17 in the least and most deprived quintiles of children, respectively). Among the observed data, children in the most deprived quintile were on average slightly lighter than those in the least deprived quintile across almost all ages 0–6 (Table 1 and Figure 3). The proportion of children ever diagnosed with *Pseudomonas* was similar across deprivation quintiles; the same was the case for the mean age at first infection.

Table 2. Estimated Total Adjusted Association and Interventional Disparity Measure Direct and Indirect Effects

	Main Analysis	Sensitivity Analysis 1	Sensitivity Analysis 2
	Point Estimate (95% Confidence Interval)		
Total adjusted association	4.5 (1.1, 7.9)	4.3 (1.7, 7)	4.4 (0.95, 7.9)
Interventional disparity measure indirect effect	0.53 (−0.64, 1.7)	0.47 (−0.46, 1.4)	0.23 (−1.6, 2.1)
Interventional disparity measure residual association	4 (0.57, 7.4)	3.9 (1.2, 6.5)	4.2 (0.43, 7.9)

In sensitivity analysis 1, we used predicted lung function at age 6 rather than observed lung function as outcome; in sensitivity analysis 2, we included interaction terms between deprivation and all covariates in the models for weight, infection, and lung function.

There was a social gradient in lung function at the first measurement. Mean lung function was 93%FEV1 (standard deviation: 14) and 89 %FEV1 (standard deviation: 18) in the least and most deprived quintiles, respectively (mean difference: 4.6, 95% CI: 1.2, 7.9%FEV1).

Estimated Total Adjusted Association and Interventional Disparity Measures

The estimated *Adj-TA*, given as the average difference in lung function between the least and most deprived quintile of children, was 4.5 percentage points (95% CI: 1.1, 7.9) (Table 2). The estimated *IDM-IE* via weight trajectories at ages 0–6 was 0.53 percentage points (95% CI: −0.64, 1.7). This is the estimated amount by which lung function would increase on average in children from the most deprived quintile of the population if their early weight could be improved to have the same distribution as that of children from the least deprived quintile of the population. The estimated *IDM-RA* was 4 percentage points (95% CI: 0.57, 7.4). The difference in lung function at first measurement between children from the least and most deprived quintiles of the population would, therefore, be on average 4 percentage points (95% CI: 0.57, 7.4) even if the weight at ages 0–6 was comparable between the least and the most deprived quintiles of children.

Sensitivity Analyses

The *Adj-TA, IDM-RA*, and *IDM-IE* from the sensitivity analyses are given in Table 2. The results are similar to the main results.

DISCUSSION

We assessed the relationship between socioeconomic conditions at birth and lung function at first measurement in children with CF in the UK, and the extent to which early inequalities in lung function can be explained by differences in weight trajectories from birth. We estimated that inequalities in early lung function would only be marginally reduced (4 %FEV1 [95% CI: 0.57, 7.4] vs. 4.5 %FEV1 [95% CI: 1.1, 7.9]) if early weight trajectories could be improved for children from the most socioeconomically deprived backgrounds.

To address this question, we extended the framework of interventional disparity effects to the setting of a longitudinal mediator and time-varying intermediate confounder. This allowed us to separate the association between socioeconomic conditions and lung function into the effects via distinct, longitudinally measured pathways, but without giving a causal interpretation to the estimate of the effect of the exposure. It, therefore, requires weaker assumptions than more commonly used mediation methods and may help gain insights in other areas of health inequalities research and social epidemiology where the consistency and no unmeasured confounding (of the exposure and mediator and or/outcome) assumptions are often not satisfied.[31,37]

Strength and Limitations

A strength of this study is that it uses the UK CF Registry data, which collects data on 99% of the current CF population. We excluded just under 200 children because of missing data; the majority of these (156) had no lung function measurement primarily because their first lung function measurement had not yet been taken at the end of our follow-up in 2016. We are, therefore, confident that the exclusion of individuals with missing data did not lead to bias and that our study population is representative of the UK CF population of children born from 2000 to 2010.

The data used in this study are recorded systematically at clinical review visits for the purposes of reporting and research and are, therefore, more reliable than many observational datasets. Measurement error in lung function is, however, nevertheless likely to be nonnegligible. This error is nondifferential and would, therefore, not be expected to lead to biased estimates.[38] We also conducted a sensitivity analysis where we used longitudinally collected lung function data to get an estimate of the error-free measure at age 6. As expected, the results agreed with the results from our main analysis.

The major limitation of this study is that our results rely on untestable assumptions, namely, "no interference," "consistency," and "no unmeasured confounding of the mediator–outcome relationship." In our setting, the "no interference" assumption states that the lung function of one child is not affected by the weight trajectory of another. This is plausible especially as children with CF are encouraged to not meet in person to prevent cross-infections. The second assumption, consistency, makes assumptions about the hypothetical intervention that would shift the weight distribution in the

most deprived children. We can imagine here a potentially complex intervention that could combine financial support with aspects of higher levels of training, engagement, and support from dieticians for children from socioeconomically deprived backgrounds and their families. The consistency assumption states that this intervention would not change the outcome for those children whose weight trajectory was set to the trajectory that was actually observed. However, it is conceivable that stronger engagement and more support for more deprived families would improve several other factors that may lead to improved lung function, such as clinic attendance and self-confidence in other aspects of CF care as well as reduced parental stress and improved mental health. Therefore, although we did not find any evidence that a hypothetical intervention on weight would substantially reduce inequalities in lung function, it may have positive effects beyond any direct effects on weight. Finally, we had to assume no unmeasured confounding of the mediator–outcome relationship. According to our assumed causal relationships (Figure 1), it is plausible that development in-utero impacts early weight and lung development, and there is some evidence that supports that birthweight (but not gestational age) is associated with lung function at age6.[39,40] Unfortunately, the UK CF registry has only recently started collecting data on birthweight and there was no proxy in the dataset that could be used to assess the impact of ignoring development in-utero in our analysis. However, as the effect of growth in-utero would have positive effects on early growth and lung function, then any hypothetical intervention on weight after birth would likely have even less of an impact than our estimated effect.

Our analysis further relies on the correct specification of the regression models used in the estimation procedure. Due to the small sample size, we did not include interactions and higher order terms in our models for the main analysis. However, we did conduct a sensitivity analysis in which we included interactions between our exposure and all covariates in the models. This did not substantially change our results although the uncertainty around estimates was increased.

A further potential limitation of the study is the use of the area-based Index of Multiple Deprivation rather than an individual measure of socioeconomic conditions. Some children living in the most deprived areas may be from affluent families and vice versa. Neighborhood-level deprivation is a relevant target for public health action. However, in further work, analyses should also be undertaken at individual level, for example, by linking family income or maternal education to the CF Registry data or collecting these data in the registry. Finally, there are differences in socioeconomic conditions between countries, for example, people in Wales are on average more deprived than those in England. Our current approach using within-country deprivation z-scores does not take this into account, and therefore, our analysis may be subject to exposure misclassification bias. However, approximately 75% of our study population is from England. Therefore, we expect this bias to have only a small effect if any. Other approaches that make the comparison of deprivation scores between countries possible, for example, by Abel et al,[41] require data that we did not have available in the UK CF Registry.

Comparison With Other Studies and Implications for Further Research and Clinical Practice

Weight in the first years of life has been shown to be strongly associated with later lung function in children with CF,[7,8,42–45] which has led to the development of CF clinical practice guidelines for nutrition management.[46] However, this is the first study that we are aware of that has looked at the role of early weight in the generation of socioeconomic inequalities in lung function in CF. Although optimal nutrition is undoubtedly important, we found that reducing inequalities in early weight would likely only have a marginal effect on inequalities in later lung function.

Other factors that may lie on the pathway from socioeconomic conditions to lung function have previously been studied, including the prescription of antibiotics,[9] exposure to environmental tobacco smoke,[47,48] and age at diagnosis.[11] However, it remains unclear what drives inequalities in early lung function and further research is needed. There is anecdotal evidence that the regularity of clinic attendance may display a social pattern due to barriers that more socioeconomically deprived families face regarding transportation and the ability to take time off work. Also, engagement with the clinical care team and health literacy of parents from more deprived backgrounds may be lower than in less deprived families and may have a role to play.[49,50] In addition, environmental factors and living conditions need further consideration.

In the absence of well-understood mediators of early inequalities, policy and practice should focus on addressing upstream drivers of child health inequalities.[51] Our results also highlight that it is important for the clinical care team of children with CF to be mindful of families' circumstances and focus on optimizing care and support for children with CF from the most deprived backgrounds. This involves paying special attention to the children's respiratory health independent of their growth and nutrition.

CONCLUSION

There is a strong association between socioeconomic conditions at birth and early lung function for children with CF, and we estimated that this would only be marginally reduced if early weight trajectories could be improved for children from the most deprived backgrounds.

The methods developed and applied here may also help to give insights into the generation of inequalities in other conditions.

ACKNOWLEDGMENTS

We thank the UK CF Registry team, and the UK CF centers and clinics for submitting data to the Registry. Special thanks to the people with cystic fibrosis and their families who have agreed for their UK CF Registry data to be used for research.

REFERENCES

1. Dodge JA, Morison S, Lewis PA, et al. Incidence, population, and survival of cystic fibrosis in the UK, 1968–95. *Arch Dis Child*. 1997;77:493–496.
2. Guo J, Garratt A, Hill A. Worldwide rates of diagnosis and effective treatment for cystic fibrosis. *J Cyst Fibros*. 2022;21:456–462.
3. UK Cystic Fibrosis Registry. 2022 annual data report. 2022.
4. Nutrition Working Group. *Nutritional management of cystic fibrosis*. 2016:60. Available at: https://www.cysticfibrosis.org.uk/sites/default/files/2020-12/Nutritional%20Management%20of%20cystic%20fibrosis%20Sep%2016.pdf. Accessed 26 May 2021.
5. Davis TA, Miller A, Hachem C, Velez C, Patel D. The current state of gastrointestinal motility evaluation in cystic fibrosis: a comprehensive literature review. *Transl Gastroenterol Hepatol*. 2023;9:10.
6. Sabharwal S. Gastrointestinal manifestations of cystic fibrosis. *Gastroenterol Hepatol*. 2016;12:43–47.
7. Konstan MW, Butler SM, Wohl MEB, et al; Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Growth and nutritional indexes in early life predict pulmonary function in cystic fibrosis. *J Pediatr*. 2003;142:624–630.
8. Lai HJ, Shoff SM, Farrell PM; Wisconsin Cystic Fibrosis Neonatal Screening Group. Recovery of birth weight z-score within two years of diagnosis is positively associated with pulmonary status at age six years in children with cystic fibrosis. *Pediatrics*. 2009;123:714–722.
9. Taylor-Robinson DC, Smyth RL, Diggle PJ, Whitehead M. The effect of social deprivation on clinical outcomes and the use of treatments in the UK cystic fibrosis population: a longitudinal study. *Lancet Respir Med*. 2013;1:121–128.
10. Barr HL, Britton J, Smyth AR, Fogarty AW. Association between socioeconomic status, sex, and age at death from cystic fibrosis in England and Wales (1959 to 2008): cross sectional study. *BMJ*. 2011;343:d4662.
11. Schlüter DK, Southern KW, Dryden C, Diggle P, Taylor-Robinson D. Impact of newborn screening on outcomes and social inequalities in cystic fibrosis: a UK CF registry-based study. *Thorax*. 2020;75:123–131.
12. Taylor-Robinson DC, Thielen K, Pressler T, et al. Low socioeconomic status is associated with worse lung function in the Danish cystic fibrosis population. *Eur Respir J*. 2014;44:1363–1366.
13. Schechter MS, Shelton BJ, Margolis PA, Fitzsimmons SC. The association of socioeconomic status with outcomes in cystic fibrosis patients in the United States. *Am J Respir Crit Care Med*. 2001;163:1331–1337.
14. Taylor-Robinson D, Diggle P, Smyth R, Whitehead M. Health trajectories in people with cystic fibrosis in the UK: exploring the effect of social deprivation. In: Burton-Jeangros C, Cullati S, Sacker A, Blane D, eds. *A Life Course Perspective on Health Trajectories and Transitions*. Springer; 2015. Available at: http://www.ncbi.nlm.nih.gov/books/NBK385365/. Accessed 25 May 2021.
15. Micali N, Daniel RM, Ploubidis GB, De Stavola BL. Maternal prepregnancy weight status and adolescent eating disorder behaviors: a longitudinal study of risk pathways. *Epidemiology*. 2018;29:579–589.
16. Taylor-Robinson D, Archangelidi O, Carr SB, et al; CF-EpinNet Collaboration. Data resource profile: the UK cystic fibrosis registry. *Int J Epidemiol*. 2018;47:9–10e.
17. Quanjer PH, Stanojevic S, Cole TJ, et al; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40:1324–1343.
18. English indices of deprivation. GOV.UK. https://www.gov.uk/government/collections/english-indices-of-deprivation. Accessed 27 May 2021.
19. Welsh index of multiple deprivation. Available at: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation. Accessed 27 May 2021.
20. SpatialData.gov.scot SG. Scottish index of multiple deprivation (SIMD) 2016. 2021. Available at: https://data.gov.uk/dataset/a448dd2a-9197-4ea0-8357-c2c9b3c29591/scottish-index-of-multiple-deprivation-simd-2016. Accessed 27 May 2021.
21. Northern Ireland multiple deprivation measures 2017. GOV.UK. Available at: https://www.gov.uk/government/statistics/northern-ireland-multiple-deprivation-measures-2017. Accessed 27 May 2021.
22. Wright CM, Williams AF, Elliman D, et al. Using the new UK-WHO growth charts. *BMJ*. 2010;340:c1140.
23. Cystic Fibrosis Trust. *Nutritional management of cystic fibrosis*; 2016. Available at: https://www.cysticfibrosis.org.uk/sites/default/files/2020-12/Nutritional%20Management%20of%20cystic%20fibrosis%20Sep%2016.pdf.
24. UK Cystic Fibrosis Trust. *Annual data report 2010*; 2011. Available at: https://www.cysticfibrosis.org.uk/sites/default/files/2020-12/2010%20Registry%20Annual%20Data%20Report.pdf. Accessed 6 July 2021.
25. Hernán MA, Robins JM. *Causal inference: what if?* Chapman & Hall/CRC; 2020.
26. Hernán MA. Does water kill? A call for less casual causal inferences. *Ann Epidemiol*. 2016;26:674–680.
27. Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes (Lond)*. 2008;32(Suppl 3):S8–14.
28. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator–outcome confounder. *Epidemiology*. 2014;25:300–306.
29. Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. *Epidemiology*. 2017;28:258–265.
30. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology*. 2014;25:473–484.
31. Naimi AI, Schnitzer ME, Moodie EEM, Bodnar LM. Mediation analysis for health disparities research. *Am J Epidemiol*. 2016;184:315–324.
32. VanderWeele T, Tchetgen ET. Mediation analysis with time-varying exposures and mediators. *Harv Univ Biostat Work Pap Ser*. 2014. Available at: https://biostats.bepress.com/harvardbiostat/paper168.
33. Lin SH, Young J, Logan R, Tchetgen Tchetgen EJ, VanderWeele TJ. Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators, and confounders. *Epidemiology*. 2017;28:266–274.
34. Buuren S van, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67.
35. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996;91:473–489.
36. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. *Stat Med*. 2018;37:2252–2266.
37. Rehkopf DH, Glymour MM, Osypuk TL. The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. *Curr Epidemiol Rep*. 2016;3:63–71.
38. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; 2015.
39. Darrah R, Nelson R, Damato EG, Decker M, Matthews A, Hodges CA. Growth deficiency in cystic fibrosis is observable at birth and predictive of early pulmonary function. *Biol Res Nurs*. 2016;18:498–504.
40. Atteih SE, Raraigh KS, Blackman SM, Cutting GR, Collaco JM. Predictive effects of low birth weight and small for gestational age status on respiratory and nutritional outcomes in cystic fibrosis. *J Cyst Fibros*. 2020;19:888–895.
41. Abel GA, Barclay ME, Payne RA. Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. *BMJ Open*. 2016;6:e012750.
42. Peterson ML, Jacobs DR, Milla CE. Longitudinal changes in growth parameters are correlated with changes in pulmonary function in children with cystic fibrosis. *Pediatrics*. 2003;112(3 Pt 1):588–592.
43. Pedreira CC, Robert RGD, Dalton V, et al. Association of body composition and lung function in children with cystic fibrosis. *Pediatr Pulmonol*. 2005;39:276–280.
44. Konstan MW, Morgan WJ, Butler SM, et al; Scientific Advisory Group and the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Risk factors for rate of decline in forced expiratory

volume in one second in children and adolescents with cystic fibrosis. *J Pediatr*. 2007;151:134–9, 139.e1.
45. Sanders DB, Fink A, Mayer-Hamblett N, et al. Early life growth trajectories in cystic fibrosis are associated with pulmonary function at age 6 years. *J Pediatr*. 2015;167:1081–1088.
46. Stallings VA, Stark LJ, Robinson KA, Feranchak AP, Quinton H; Clinical Practice Guidelines on Growth and Nutrition Subcommittee. Evidence-based practice recommendations for nutrition-related management of children and adults with cystic fibrosis and pancreatic insufficiency: results of a systematic review. *J Am Diet Assoc*. 2008;108:832–839.
47. Ong T, Schechter M, Yang J, et al; for the EPIC Study Group. Socioeconomic status, smoke exposure, and health outcomes in young children with cystic fibrosis. *Pediatrics*. 2017;139:e20162730. doi:10.1542/peds.2016-2730.
48. Oates GR, Baker E, Rowe SM, et al. Tobacco smoke exposure and socioeconomic factors are independent predictors of pulmonary decline in pediatric cystic fibrosis. *J Cyst Fibros*. 2020;19:783–790.
49. Jackson AD, Kirwan L, Gibney S, Jeleniewska P, Fletcher G, Doyle G. Associations between health literacy and patient outcomes in adolescents and young adults with cystic fibrosis. *Eur J Public Health*. 2020;30:112–118.
50. Mercer SW, Higgins M, Bikker AM, et al. General practitioners' empathy and health outcomes: a prospective observational study of consultations in areas of high and low deprivation. *Ann Fam Med*. 2016;14:117–124.
51. Pearce A, Dundas R, Whitehead M, Taylor-Robinson D. Pathways to inequalities in child health. *Arch Dis Child*. 2019;104:998–1003.