Practice of Epidemiology

Life-Course Analysis of a Fat Mass and Obesity-Associated (*FTO*) Gene Variant and Body Mass Index in the Northern Finland Birth Cohort 1966 Using Structural Equation Modeling

Marika Kaakinen, Esa Läärä, Anneli Pouta, Anna-Liisa Hartikainen, Jaana Laitinen, Tuija H. Tammelin, Karl-Heinz Herzig, Ulla Sovio, Amanda J. Bennett, Leena Peltonen†, Mark I. McCarthy, Paul Elliott, Bianca De Stavola, and Marjo-Riitta Järvelin*

* Correspondence to Dr. Marjo-Riitta Jarvelin, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, Norfolk Place, London W2 1PG, United Kingdom (e-mail: m.jarvelin@imperial.ac.uk).

† Deceased.

The association between variation in the fat mass and obesity-associated (*FTO*) gene and adulthood body mass index (BMI; weight (kg)/height (m)^2) is well-replicated. More thorough analyses utilizing phenotypic data over the life course may deepen our understanding of the development of BMI and thus help in the prevention of obesity. The authors used a structural equation modeling approach to explore the network of variables associated with BMI from the prenatal period to age 31 years (1965–1997) in 4,435 subjects from the Northern Finland Birth Cohort 1966. The use of structural equation modeling permitted the easy inclusion of variables with missing values in the analyses without separate imputation steps, as well as differentiation between direct and indirect effects. There was an association between the *FTO* single nucleotide polymorphism rs9939609 and BMI at age 31 years that persisted after controlling for several relevant factors during the life course. The total effect of the *FTO* variant on adult BMI was mostly composed of the direct effect, but a notable part was also arising indirectly via its effects on earlier BMI development. In addition to well-established genetic determinants, many life-course factors such as physical activity, in spite of not showing mediation or interaction, had a strong independent effect on BMI.

body mass index; molecular epidemiology; structural equation model

Abbreviations: BMI, body mass index; CI, confidence interval; *FTO*, fat mass and obesity-associated; MET, metabolic equivalent; SD, standard deviation; SEM, structural equation modeling.

The prevalence of obesity is rapidly increasing in both developed and developing countries. Obesity predisposes people to many chronic diseases, such as the metabolic syndrome, type 2 diabetes, and cardiovascular disease (1). Recent progress in genome-wide association studies has led to the discovery of novel genetic variants associated with body mass index (BMI; weight (kg)/height (m)^2) and increased risk of obesity (2–5). The strongest signals discovered to date are located in the fat mass and obesity-associated (*FTO*) gene, which was originally found within a study on type 2 diabetes genes, but the association was mediated by BMI (6). Since then the association between *FTO* and BMI has been replicated in several studies (3, 4,

7, 8). The association between *FTO* and BMI growth throughout the life course is still somewhat unclear, but some studies suggest that the effect starts to show at least as early as approximately age 7 years (9–11).

Genetic variants discovered so far explain only a small proportion of the variability in body weight. For instance, in the Northern Finland Birth Cohort 1966, variants in the *FTO* and melanocortin 4 receptor (*MC4R*) genes explain only 0.55% of the total variation in adult BMI (12). The heritability of BMI has been estimated to be moderate-to-high (40%–80%) (13, 14), so there are probably many common single nucleotide polymorphisms of comparable effect sizes yet to be identified and obviously stronger

Am J Epidemiol 2010;172:653–665

underlying rare variants that wait to be discovered (4). Meanwhile, it is important to study the interplay between life-course factors and the genetic variants discovered so far.

It has been suggested that the *FTO* gene plays a role in appetite regulation (15) and that it is associated with energy expenditure (16), energy intake (17, 18), and diminished satiety (19), whereas 2 recent studies found no evidence for an association between nutrition and *FTO* (20, 21). In several studies, investigators have also reported a significant effect of interaction between *FTO* and physical activity on BMI (22–26). However, in 1 relatively large study, Jonsson et al. (27) found no evidence for interaction between *FTO* and physical activity.

To identify mediators or modifiers of the genetic associations, carefully characterized cohorts are needed (28), as well as appropriate statistical methods for dealing with complex relations. Multiple regression analysis has often been used as a standard method, yet a model with several terms may produce biased and unstable estimates because of sparse data and multicollinearity (29). Standard multiple regression also ignores the presumed causal and temporal ordering of exposure variables and their interrelations (30) and thus can provide information only on direct effects conditional on all of the other variables in the model (31), whereas an appropriate path analysis can provide deeper insight into the interrelations of the variables, that is, indirect effects and mediation.

Structural equation modeling (SEM), which includes path analysis and latent variables (32), can be used to study associations between variables thought to be causally ordered along the life course. Within this modeling approach, variables are subdivided into background factors, such as parental characteristics; intermediate outcomes, such as BMI at birth and BMI during childhood; and distal outcomes, such as adult BMI. Formally, relations among these factors are specified via simultaneous equations, and then the covariance structure of the assumed model is estimated.

We used SEM to examine the effects of the *FTO* rs9939609 variant on adult BMI in a large, prospectively followed birth cohort on which we had data from the prenatal period to adulthood. Our aim was to obtain more information on the networks around the variables along the life course and a better understanding of the mechanisms by which the *FTO* polymorphism plays a role in weight gain, together with other factors. We also hypothesized that the effect of the *FTO* variant might be mediated or modified through behavioral factors that have been suggested to be associated with *FTO*, such as diet and physical activity.

MATERIALS AND METHODS

Participants

The study population consisted of persons belonging to the Northern Finland Birth Cohort 1966. Initially, all mothers with expected delivery dates in 1966 who were living in the 2 northernmost provinces of Finland, Oulu and Lapland, were invited; over 96% participated (12,055 mothers with 12,058 liveborn children) (33). Data on the

prenatal and perinatal period were collected via questionnaires administered by local midwives in the antenatal clinics. In 1980, at the age of 14 years, all living cohort members with known addresses received postal questionnaires containing questions on their growth, health habits, and family situation. An abridged version of the questionnaire was sent to the parents in cases where the adolescent did not respond or to the school health nurse if neither the child nor the parent responded. The postal questionnaire was returned by 94% ($n = 11,010$) of the adolescents, 52% ($n = 389$) of the parents, and 97% ($n = 354$) of the nurses. In 1997, at the age of 31 years, the subjects received a postal questionnaire including questions on their health, lifestyle, and occupation, and it was returned by 75% ($n = 8,767$). At the same time, those subjects living in the original target area (northern Finland) or in the capital area (Helsinki) were invited to undergo a clinical examination, and 71% ($n = 6,033$) of those invited participated (34). At this point, blood samples were drawn and DNA was extracted for 5,753 subjects. For the present study, we included persons who had both genotype data and measured BMI data available at age 31 years. After exclusion of multiple births, 4,435 persons (2,137 men and 2,298 women) remained for analysis.

All participants gave their written, informed consent when DNA was taken at age 31 years. The University of Oulu ethics committee approved the study.

Genotyping

Genotyping of the samples was performed using the TaqMan single nucleotide polymorphism genotyping assay (Applied Biosystems, Warrington, United Kingdom) according to the manufacturer's protocol. Genotyping was carried out for 5,365 samples from the cohort. The allele frequencies of the single nucleotide polymorphism rs9939609 were not observed to deviate essentially from Hardy-Weinberg equilibrium ($P = 0.33$). The duplicate concordance rate was 99.9%, and the genotype success rate in the sample was 88% ($n = 4,701$).

Phenotypic, behavioral, and environmental variables

Data on variables related to the prenatal period or birth were obtained from the questionnaire targeted toward the mothers during pregnancy and supplemented after delivery. Maternal prepregnancy BMI and gestational age were treated as continuous variables in the analyses. The categorizations of parity, maternal smoking after the second month of pregnancy, familial socioeconomic status (based on the father's occupation, or on the mother's if single), and maternal hypertension during pregnancy are shown in Table 1.

Regarding variables assessed at age 14 years, we calculated BMI from self-reported height and weight and classified the adolescent's smoking status into 1) nonsmoker (never smoked or had tried smoking), 2) occasional smoker (smoked occasionally or about twice a week), and 3) regular smoker (smoked daily) (35). Alcohol consumption was classified into 1) nonconsumer (never drank alcohol, had merely tasted it, or had consumed alcohol occasionally) and 2) regular consumer (drank monthly or more

Table 1. Characteristics of the Northern Finland Birth Cohort 1966 Study Sample, 1965–1997

Characteristic	Total	No.	%	Mean (SD)
FTO rs9939609 genotype	4,435			
TT		1,678	37.8	
AT		2,068	46.6	
AA		689	15.5	
Prenatal factors				
Maternal BMI[a]	4,052			23.2 (3.2)
Maternal age, years	4,427			28.2 (6.6)
Parity	4,424			
0		1,384	31.3	
1–3		2,193	49.6	
≥4		847	19.2	
Maternal smoking during second month of pregnancy	4,325			
Nonsmoker		3,715	85.9	
1–10 cigarettes/day		510	11.8	
>10 cigarettes/day		100	2.3	
Blood pressure during pregnancy	4,344			
Normotensive		2,418	55.7	
Gestational hypertension[b]		779	17.9	
Elevated systolic blood pressure		361	8.3	
Elevated diastolic blood pressure		338	7.8	
Not determined/not known		448	10.3	
Family SES	4,403			
I + II (professional)		1,027	23.3	
III (skilled worker)		1,476	33.5	
IV (unskilled worker)		951	21.6	
V (farmer)		949	21.6	
Characteristics at birth				
Sex	4,435			
Male		2,137	48.2	
Female		2,298	51.8	
BMI	4,404			13.8 (1.3)
Gestational age, weeks	4,278			40.1 (1.8)
Characteristics at age 14 years				
BMI	3,957			19.4 (2.5)
Frequency of participation in sports	4,191			
Daily		738	17.6	
Every other day		821	19.6	
Twice a week		934	22.3	
Once a week		688	16.4	
Less than once a week		1,010	24.1	

Table continues

Table 1. Continued

Characteristic	Total	No.	%	Mean (SD)
Alcohol consumption	4,238			
Nonregular intake		4,125	97.3	
Regular intake		113	2.7	
Smoking	4,243			
Nonsmoker		3,515	82.4	
Occasional smoker		463	10.9	
Regular smoker		265	6.3	
Family SES	4,260			
I + II (professional)		1,223	28.7	
III (skilled worker)		1,463	34.3	
IV (unskilled worker)		935	22.0	
V (farmer)		639	15.0	
Characteristics at age 31 years				
BMI	4,435			24.7 (4.3)
Quartile of physical activity, MET-hours per week	4,022			14.8 (14.6)
0–3.7		1,106	27.5	
3.8–10.8		1,155	28.7	
10.9–20.5		1,080	26.9	
≥20.6		681	16.9	
Tertile of alcohol consumption, g/day	4,296			9.2 (15.8)
Abstainer		405	9.4	
0.1–2.5		1,283	29.9	
2.6–8.7		1,298	30.2	
≥8.8		1,310	30.5	
Smoking	4,162			
Nonsmoker		2,385	57.3	
1–10 cigarettes/day		887	21.3	
>10 cigarettes/day		890	21.4	
Unhealthy diet score[c]	4,395			
0–1		1,380	31.4	
2–3		2,495	56.8	
4–5		520	11.8	
SES	4,387			
I + II (professional)		1,039	23.7	
III (skilled worker)		1,354	30.9	
IV (unskilled worker)		1,113	25.4	
V (farmer)		159	3.6	
VI (other)		722	16.5	

Abbreviations: BMI, body mass index; FTO, fat mass and obesity-associated; MET, metabolic equivalent; SD, standard deviation; SES, socioeconomic status.

[a] Weight (kg)/height (m)^2.

[b] Includes gestational hypertension, chronic hypertension, preeclampsia, and superimposed preeclampsia.

[c] An unhealthy diet was defined as daily or almost daily consumption of sausage and less frequent (twice a week or less often) consumption of rye bread or crisp bread, fresh vegetables and salads, and berries or fruit. One point was given for each of these counts, and scores could range from 0 to 5.
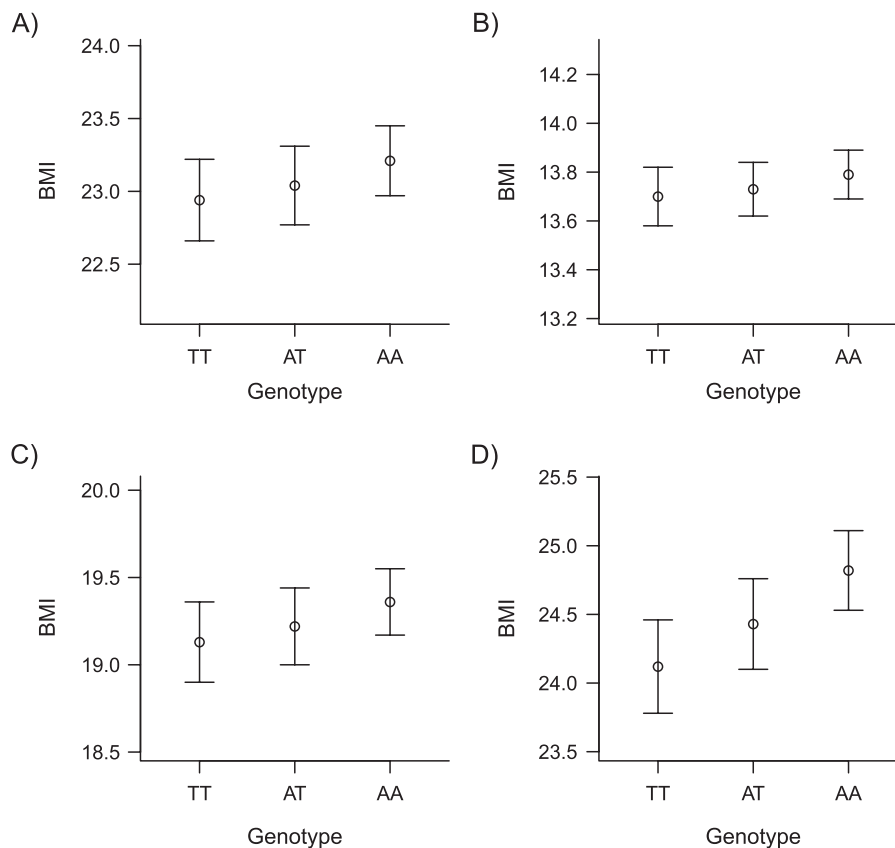
Figure 1. Relation between the fat mass and obesity-associated (*FTO*) rs9939609 genotype and A) maternal body mass index (BMI; weight (kg)/height (m)^2); B) the subject's own BMI at birth; C) BMI at age 14 years; and D) BMI at age 31 years in the Northern Finland Birth Cohort 1966, 1965–1997. Data are presented as geometric mean values (circles) with back-transformed 95% confidence intervals (bars). Results were adjusted for sex and gestational age (for birth BMI) or sex only (for BMI at ages 14 and 31 years).

often) (36). The subjects were asked how often they participated in sports after school hours and were classified into 5 physical activity groups: 1) daily, 2) every other day, 3) twice a week, 4) once a week, and 5) less than once a week (37). Familial socioeconomic status was coded similarly to prepregnancy socioeconomic status.

Weight and height at age 31 years were measured at the clinical examination, and BMI was calculated from those measurements. Data on other background variables were obtained from the postal questionnaire filled in at the same age. Smoking during the past year was categorized as 1) nonsmoker, 2) 1–10 cigarettes/day, or 3) >10 cigarettes/day. Alcohol consumption was measured with several questions on type, amount, and frequency of alcohol consumption, and the information was validated against 7-day food diaries. It was transformed into daily intake (g/day) (34) and was further classified into abstainers and consumers divided according to tertile of daily intake (Table 1). Frequencies of consumption of various types of food were also ascertained, and an unhealthy diet was defined as daily or almost daily consumption of sausage and less frequent (twice a week or less often) consumption of rye bread or crisp bread, fresh vegetables and salads, and berries or fruit. One point was given for each of these counts (38), and the unhealthy diet

score, ranging from 0 to 5, was used in the analyses. The subjects were also asked about the frequency and duration of light and brisk physical activities. These data were transferred into metabolic equivalent (MET)-hours per week (39) and further classified into quartiles. In the calculations, an intensity value of 3 METs was used for light physical activity and 5 METs was used for brisk physical activity. The subject's own socioeconomic status was based on occupation and employment data. It was classified from I (high) to IV (low), plus farmers and others (student, pensioner, long-term unemployed, or not defined).

Statistical analyses

Associations between *FTO* rs9939609 and maternal prepregnancy BMI, BMI at birth, BMI at age 14 years, and BMI at age 31 years, respectively, were first analyzed with multiple linear regression models assuming an additive model for genotype. All of the outcome variables were natural logarithm-transformed to reduce skewness, and the analyses were adjusted for sex (BMI at ages 14 and 31 years) or sex and gestational age (BMI at birth). Results are presented as geometric mean values and 95% confidence intervals.
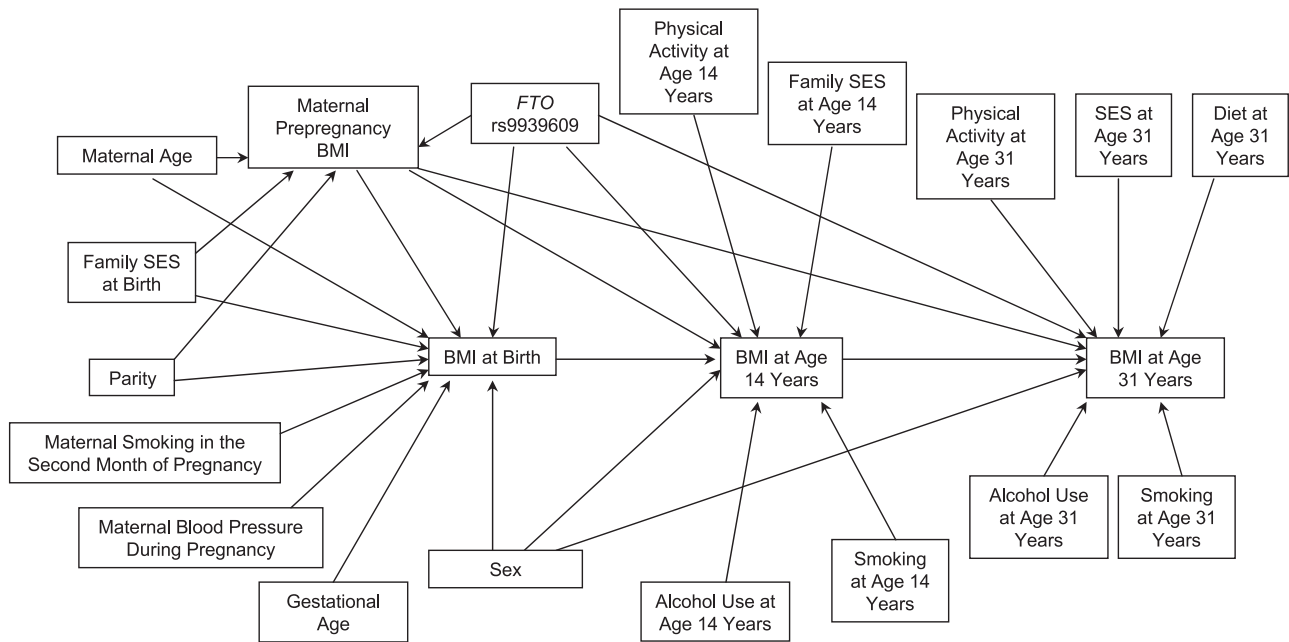
Figure 2. Structural equation model of relations between the fat mass and obesity-associated (*FTO*) rs9939609 genotype and body mass index (BMI; weight (kg)/height (m)²) fitted to data from the Northern Finland Birth Cohort 1966, 1965–1997. Four simultaneous equations were fitted using maternal BMI, BMI at birth, BMI at age 14 years, and BMI at age 31 years as the outcome variables. SES, socioeconomic status.

A SEM approach was used to model the assumed underlying relations between the studied variables. Four simultaneous equations were fitted with maternal prepregnancy BMI, BMI at birth, BMI at age 14 years, and BMI at age 31 years as the outcome variables and assuming additive models for the effect of *FTO* rs9939609. All data on the outcome variables were natural logarithm-transformed; hence, the results are presented as percentage changes with 95% confidence intervals. Estimation of the parameters was carried out by the method of maximum likelihood, assuming normal distributions for the outcome variables. The expectation-maximization algorithm (40) implemented in Mplus (41) was used in the analyses of incomplete data. The expectation-maximization algorithm relies on the assumption that data are missing at random (42). For comparison, we also estimated model parameters for complete cases only (data not shown). Model evaluations were carried out on the basis of the following goodness-of-fit indices: the comparative fit index and the root mean square error of approximation. The comparative fit index is a fit index that is independent of sample size, and the model is thought to have a good fit if the index exceeds 0.95 (43, 44). The root mean square error of approximation measures the discrepancy between the model and the observed covariance matrix and is expressed per degree of freedom, thus taking into account the complexity of the model. A value less than 0.05 indicates a good fit (45). Modification indices (46) were used in detecting model misspecifications. All of the analyses were conducted with SAS, version 9.1.3 (SAS Institute Inc., Cary, North Carolina) and Mplus, version 3.12 (41). All *P* values reported are 2-sided.

RESULTS

The distributions of *FTO* rs9939609, BMI, and background factors are presented in Table 1. Of the 4,435 study subjects, 46.6% were heterozygous for the risk A allele and 15.5% were AA homozygous. Data on maternal BMI were available for 4,052 subjects, with a mean of 23.2 (standard deviation (SD), 3.2). The subjects' mean BMIs were 13.8 (SD, 1.3) at birth, 19.4 (SD, 2.5) at age 14 years, and 24.7 (SD, 4.3) at age 31 years.

Figure 1 shows the geometric mean BMIs and back-transformed 95% confidence intervals according to *FTO* rs9939609 genotype. Because the Northern Finland Birth Cohort 1966 was part of the database used in the replication of the original *FTO* finding, these associations between BMI at ages 14 years and 31 years were reported previously by Frayling et al. (6). We additionally included maternal BMI and BMI at birth in the analyses. Carrying the risk A allele was associated with a 1.40% (95% confidence interval (CI): 0.72, 2.09) higher BMI at age 31 years (per A-allele change from an adjusted additive model corresponding to a 0.34-unit (95% CI: 0.18, 0.51) higher BMI, $P = 5.1 \times 10^{-5}$) and a 0.58% (95% CI: 0.00, 1.16) higher BMI at age 14 years (0.11-unit higher BMI (95% CI: 0.00, 0.22), $P = 0.05$). Weaker evidence of effects pointing in the same direction on maternal BMI and BMI at birth were also observed (maternal BMI: 0.55% (95% CI: −0.05, 1.14), corresponding to 0.11 units (95% CI: −0.01, 0.22), $P = 0.07$; BMI at birth: 0.32% (95% CI: −0.09, 0.72), corresponding to 0.04 units (95% CI: −0.01, 0.10), $P = 0.12$).

The SEM fitted to the data is depicted in Figure 2. It shows the relations we assumed to be underlying among the available variables in our study sample, based on

Table 2.   Results From Structural Equation Modeling of Relations Between the Fat Mass and Obesity-Associated (*FTO*) rs9939609 Genotype and Life-Course Data in the Northern Finland Birth Cohort 1966, 1965–1997

Model	β (×10⁻³)	Standard Error (×10⁻³)	Standardized β	P Value	Change, ‰	95% CI	Corresponding Change in Mean BMI, g/m²	95% CI
Maternal BMI[a,b]								
FTO rs9939609 (additive model)	6.00	2.73	0.031	0.03	6.18	0.64, 12.0	143	14.9, 279
Maternal age, years	6.53	0.39	0.325	<0.001	6.75	5.94, 7.56	157	138, 176
Parity								
0	Referent							
1–3	20.1	4.34	0.150	<0.001	22.2	12.3, 33.1	516	285, 768
≥4	50.2	7.86	0.376	<0.001	65.2	41.6, 92.7	1,512	965, 2,150
Family SES at birth								
I + II (professional)	Referent							
III (skilled worker)	4.46	4.77	0.034	0.35	4.56	−4.78, 14.8	106	−111, 344
IV (unskilled worker)	11.8	5.76	0.088	0.04	12.5	0.51, 26.0	290	11.9, 602
V (farmer)	25.8	5.79	0.192	<0.001	29.4	15.5, 44.9	682	360, 1,042
BMI at birth[b]								
FTO rs9939609 (additive model)	2.56	1.91	0.018	0.18	2.60	−1.18, 6.52	35.8	−16.3, 90.0
Sex (female vs. male)	−6.00	2.70	−0.062	0.03	−5.83	−10.7, −0.71	−80.4	−147, −9.80
Gestational age, weeks	15.2	0.89	0.288	<0.001	16.4	14.4, 18.5	227	199, 255
Log maternal BMI	0.11	0.01	0.155	<0.001	0.11	0.09, 0.14	1.60	1.20, 1.90
Maternal age, years	−0.51	0.31	−0.035	0.09	−0.51	−1.10, 0.09	−7.00	−15.2, 1.20
Parity								
0	Referent							
1–3	30.8	3.42	0.316	<0.001	36.0	27.2, 45.4	497	375, 627
≥4	43.6	5.53	0.447	<0.001	54.6	38.7, 72.3	753	534, 997
Maternal smoking								
Nonsmoker	Referent							
1–10 cigarettes/day	−11.9	4.45	−0.124	0.01	−11.2	−18.7, −3.16	−155	−257, −43.6
>10 cigarettes/day	−31.0	8.99	−0.317	0.001	−26.6	−38.5, −12.5	−367	−531, −172
Maternal blood pressure								
Normotensive	Referent							
Gestational hypertension[c]	−18.1	4.04	−0.185	<0.001	−16.5	−22.9, −9.63	−228	−316, −133
Elevated systolic blood pressure	8.11	4.98	0.083	0.10	8.44	−1.64, 19.6	117	−22.7, 270
Elevated diastolic blood pressure	−1.68	5.52	−0.019	0.76	−1.66	−11.8, 9.57	−23.0	−162, 132
Not determined/not known	−3.57	4.81	−0.036	0.46	−3.51	−12.2, 6.03	−48.4	−168, 83.1

Table continues

previous knowledge of the associations (38, 47) and the correlation structure of the variables. Note that we also specified a relation between child's genotype and maternal BMI, because half of the child's genotype is inherited from the mother, and thus it partly represents the mother's genotype. The initial model was modified by removing nonsignificant associations whose inclusion would have worsened the overall model fit considerably and by adding new paths based on modification indices. Although we hypothesized that the *FTO* effect may be mediated through diet, unfortunately it was not possible to examine this adequately with our rather crude diet measurement, which showed no association with the *FTO* variant. Including a mediating path through diet would also have considerably worsened the overall model fit in terms of the comparative fit index (a drop from 0.92 to 0.74). Additionally, we tested for an interaction between the *FTO* variant and physical activity, but no evidence for it was observed ($P > 0.20$ for all interaction terms); thus, the terms were omitted from the final model. All of the mediating paths through BMI measurements were

Table 2.   Continued

Model	β (×10⁻³)	Standard Error (×10⁻³)	Standardized β	*P* Value	Change, ‰	95% CI	Corresponding Change in Mean BMI, g/m²	95% CI
Family SES at birth								
I + II (professional)	Referent							
III (skilled worker)	−10.3	3.65	−0.106	0.01	−9.82	−16.1, −3.12	−136	−222, −43.1
IV (unskilled worker)	−15.6	4.19	−0.161	<0.001	−14.5	−21.2, −7.13	−199	−293, −98.3
V (farmer)	−15.3	4.29	−0.156	<0.001	−14.1	−21.1, −6.61	−195	−291, −91.2
BMI at age 14 years[b]								
FTO rs9939609 (additive model)	4.83	2.88	0.026	0.09	4.95	−0.82, 11.1	96.1	−15.8, 214
Sex (female vs. male)	8.07	4.11	0.062	0.05	8.40	0.02, 17.5	163	0.40, 339
Log birth BMI	0.11	0.02	0.080	<0.001	0.11	0.06, 0.15	2.10	1.20, 2.90
Log maternal BMI	0.17	0.02	0.175	<0.001	0.17	0.14, 0.20	3.30	2.60, 3.90
Frequency of participation in sports at age 14 years								
Daily	Referent							
Every other day	1.60	6.47	0.013	0.80	1.61	−10.5, 15.4	31.3	−204, 298
Twice a week	0.99	6.21	0.007	0.87	1.00	−10.6, 14.1	19.3	−205, 273
Once a week	5.55	7.14	0.043	0.44	5.71	−8.11, 21.6	111	−157, 419
Less than once a week	−1.61	6.75	−0.012	0.81	−1.60	−13.8, 12.3	−31.0	−267, 239
Alcohol consumption at age 14 years								
Nonregular intake	Referent							
Regular intake	31.6	12.6	0.248	0.01	37.2	7.18, 75.6	721	139, 1,467
Smoking at age 14 years								
Nonsmoker	Referent							
Occasional smoker	18.0	6.27	0.138	0.004	19.7	5.86, 35.4	382	114, 686
Regular smoker	7.38	9.06	0.058	0.42	7.66	−9.85, 28.6	149	−191, 554
Family SES at age 14 years								
I + II (professional)	Referent							
III (skilled worker)	2.08	4.86	0.017	0.67	2.10	−7.18, 12.3	40.8	−139, 239
IV (unskilled worker)	−5.21	5.86	−0.041	0.37	−5.07	−15.4, 6.49	−98.4	−298, 126
V (farmer)	−1.76	6.13	−0.014	0.77	−1.75	−12.9, 10.8	−33.9	−250, 210
BMI at age 31 years[b]								
FTO rs9939609 (additive model)	9.22	2.88	0.040	0.001	9.66	3.63, 16.0	239	89.6, 396
Sex (female vs. male)	−44.8	4.77	−0.278	<0.001	−36.1	−41.8, −29.8	−891	−1,032, −737

Table continues

left in the model, although some of these paths showed only weak evidence of a direct association (the effect of *FTO* rs9939609 on BMI at birth and BMI at age 14 years).

Because of the categorical nature of several variables in our analysis, we did not allow for correlations between variables and did not specify associations between sex, alcohol, smoking, and physical activity, for instance. When we conducted the analyses separately for men and women (data not shown), we observed sex differences in the estimated effects of these variables on BMI, but these differences did not influence the estimated effect of the *FTO* variant on BMI, which was our

main interest in this study. Thus, we report results from the analysis conducted for men and women together.

For cross-validation of our results, we randomly assigned study subjects to a training sample and a validation sample (48); we first conducted the analyses in the training sample and then validated them in the other sample. Because the results did not differ substantially between the samples (data not shown), we merged the samples and report the results for the whole sample.

The model gave a good fit in terms of the root mean square error of approximation (0.025) and an adequate fit

Table 2.   Continued

Model	β (×10⁻³)	Standard Error (×10⁻³)	Standardized β	P Value	Change, ‰	95% CI	Corresponding Change in Mean BMI, g/m²	95% CI
Log maternal BMI	0.11	0.02	0.095	<0.001	0.11	0.08, 0.15	2.80	2.00, 3.60
Log BMI at age 14 years	0.66	0.03	0.529	<0.001	0.66	0.60, 0.72	16.4	14.9, 17.9
Quartile of physical activity, MET-hours per week								
0–3.7	26.5	6.83	0.165	0.001	30.4	14.0, 49.0	750	347, 1,211
3.8–10.8	15.1	5.98	0.093	0.01	16.3	3.41, 30.7	402	84.1, 759
10.9–20.5	2.39	5.59	0.015	0.67	2.42	−8.21, 14.3	59.8	−203, 353
≥20.6	Referent							
Tertile of alcohol consumption, g/day								
Abstainer	Referent							
0.1–2.5	−3.70	8.39	−0.022	0.66	−3.63	−18.2, 13.6	−89.7	−451, 336
2.6–8.7	−7.56	8.16	−0.048	0.35	−7.29	−21.0, 8.80	−180	−519, 217
≥8.8	8.93	8.43	0.054	0.29	9.34	−7.32, 29.0	231	−181, 716
Smoking at age 31 years								
Nonsmoker	Referent							
1–10 cigarettes/day	3.86	5.53	0.024	0.49	3.93	−6.74, 15.8	97.1	−167, 391
>10 cigarettes/day	−9.52	5.94	−0.059	0.11	−9.08	−19.1, 2.14	−224	−471, 52.8
Unhealthy diet score[d]	4.65	2.01	0.033	0.02	4.76	0.70, 8.97	118	17.3, 222
SES at age 31 years								
I + II (professional)	Referent							
III (skilled worker)	15.4	5.67	0.094	0.01	16.6	4.34, 30.3	410	107, 749
IV (unskilled worker)	11.1	5.21	0.069	0.04	11.8	0.90, 23.8	290	22.3, 587
V (farmer)	18.7	11.3	0.118	0.10	20.5	−3.48, 50.5	507	−86.0, 1,248
VI (other)	19.8	6.88	0.124	0.004	21.9	6.52, 39.5	541	161, 976

Abbreviations: BMI, body mass index; CI, confidence interval; FTO, fat mass and obesity-associated; MET, metabolic equivalent; SES, socioeconomic status.

[a] Weight (kg)/height (m)².

[b] The outcome variables maternal BMI and BMI at birth, age 14 years, and age 31 years were natural logarithm-transformed.

[c] Includes gestational hypertension, chronic hypertension, preeclampsia, and superimposed preeclampsia.

[d] An unhealthy diet was defined as daily or almost daily consumption of sausage and less frequent (twice a week or less often) consumption of rye bread or crisp bread, fresh vegetables and salads, and berries or fruit. One point was given for each of these counts, and scores could range from 0 to 5.

in terms of the comparative fit index (0.92). The regression coefficients conditional on all of the variables in the SEM analysis are shown in Table 2. Note that because of some very small effect sizes, the estimates are presented in 10^{-3} scale, the changes are presented in per mils (per 1,000), and the corresponding changes in mean BMI are presented in g/m² instead of the conventional kg/m². The standardized β coefficients are used to compare the relative importance of the independent variables, since they describe the change in the outcome variable in SD units per a 1-SD change in the continuous predictor and per the change from 0 to 1 in a binary predictor. These standardized coefficients suggest that FTO rs9939609 would have a modest effect on BMI in comparison with some of the early background exposures, which show a much stronger independent effect (e.g., the standardized regression coefficients for maternal BMI were 0.031 SD units for the FTO variant and 0.325 SD units for

maternal age). However, for outcomes in adulthood, the estimated effect of the FTO variant on BMI was comparable with that of smoking and socioeconomic status (for example). The model including both genetic and life-course factors explained 20% of the total variation in maternal BMI and explained 16%, 5%, and 34% of the variation in BMI at birth, age 14 years, and age 31 years, respectively.

Table 3 shows the estimated indirect, direct, and total effects of the FTO variant on BMI, calculated assuming that the relations depicted in Figure 2 are correct. The total effects of FTO rs9939609 on maternal BMI, BMI at birth, and BMI at age 14 years were strengthened in comparison with the cross-sectional explorative analyses shown in Figure 1. For BMI at age 31 years, the evidence for the association remained strong ($P = 5.0 \times 10^{-5}$). The estimated direct effect of a per-A-allele change (conditional on all of the other variables in the model) on adult BMI

Table 3. Direct, Indirect, and Total Effects of Fat Mass and Obesity-Associated (*FTO*) Genotype rs9939609 on Body Mass Index During the Life Course in the Northern Finland Birth Cohort 1966, 1965–1997

Model	β (×10⁻³)	Standard Error (×10⁻³)	Standardized β	P Value	Change, ‰	95% CI	Corresponding Change in Mean BMI, g/m²	95% CI
Maternal BMI[a]								
Total effect	6.00	2.73	0.031	0.03	6.18	0.64, 12.0	143	14.9, 279
BMI at birth								
Total effect	3.24	1.94	0.023	0.09	3.30	−0.56, 7.30	45.5	−7.71, 101
Total indirect effect	0.68	0.32	0.005	0.03	0.68	0.06, 1.31	9.42	0.84, 18.1
1) *FTO*–maternal BMI–birth BMI	0.68	0.32	0.005	0.03	0.68	0.06, 1.31	9.42	0.84, 18.1
Direct effect	2.56	1.91	0.018	0.18	2.60	−1.18, 6.52	35.8	−16.3, 90.0
BMI at age 14 years								
Total effect	6.19	2.95	0.033	0.04	6.39	0.41, 12.7	124	7.88, 23.5
Total indirect effect	1.36	0.55	0.007	0.01	1.37	0.28, 2.47	26.6	5.52, 47.8
1) *FTO*–maternal BMI–BMI at 14	1.02	0.48	0.005	0.03	1.02	0.08, 1.97	19.8	1.55, 38.3
2) *FTO*–birth BMI–BMI at 14	0.27	0.21	0.001	0.19	0.27	−0.14, 0.68	5.28	−2.67, 13.3
3) *FTO*–maternal BMI–birth BMI–BMI at 14	0.07	0.04	0.000	0.05	0.07	0.00, 0.14	1.40	−0.01, 2.81
Direct effect	4.83	2.88	0.026	0.09	4.95	−0.82, 11.1	92.1	−15.8, 214
BMI at age 31 years								
Total effect	14.0	3.45	0.060	5.0×10^{-5}	15.0	7.50, 23.1	371	185, 570
Total indirect effect	4.78	2.04	0.021	0.02	4.90	0.79, 9.17	121	19.6, 227
1) *FTO*–maternal BMI–BMI at 31	0.69	0.33	0.003	0.04	0.69	0.04, 1.34	17.0	1.02, 33.1
2) *FTO*–BMI at 14–BMI at 31	3.20	1.90	0.014	0.09	3.25	−0.53, 7.17	80.2	−13.0, 177
3) *FTO*–maternal BMI–BMI at 14–BMI at 31	0.67	0.32	0.003	0.03	0.67	0.05, 1.30	16.7	1.20, 32.2
4) *FTO*–birth BMI–BMI at 14–BMI at 31	0.18	0.14	0.001	0.19	0.18	−0.09, 0.45	4.45	−2.28, 11.2
5) *FTO*–maternal BMI–birth BMI–BMI at 14–BMI at 31	0.05	0.02	0.000	0.05	0.05	0.00, 0.10	1.19	0.02, 2.35
Direct effect	9.22	2.88	0.040	0.001	9.66	3.63, 16.0	239	89.6, 396

Abbreviations: BMI, body mass index; CI, confidence interval; FTO, fat mass and obesity-associated.

[a] Weight (kg)/height (m)².

was 0.97% (95% CI: 0.36, 1.60), which corresponds to an increase of 0.24 units (95% CI: 0.09, 0.40) in mean BMI. Indirect effects of the *FTO* variant were observed through maternal BMI, BMI at birth, and BMI at age 14 years (Figure 2, Table 3). The effects through BMI at birth were modest, since the association between the *FTO* variant and BMI at birth was of small magnitude (0.26%, 95% CI: −0.12, 0.65). Adding all of the indirect effects together, an increase of 0.49% (95% CI: 0.08, 0.92) in adult BMI was observed, which is equivalent to 0.12 units (95% CI: 0.02, 0.23). The total effect, which is the sum of the indirect and direct effects, was then 1.50% (95% CI: 0.75, 2.30), corresponding to a 0.37-unit (95% CI: 0.19, 0.57) increase in mean BMI.

Attrition and missing data

The subset of Northern Finland Birth Cohort 1966 subjects who participated in the clinical examination at age 31 years has been shown to be well-representative of the original study population (49). We further compared the distributions of all of the variables used in the present study between subjects who had complete data on all of the selected variables (*n* = 2,761) and those who had missing information on at least 1 of the variables (*n* = 1,674). With regard to the maternal characteristics, subjects with missing values on any of the variables used in the analyses were more likely to have a slightly older mother (mean age at delivery = 28.5 years vs. 28.0 years), a mother with more

children (mean parity = 3.2 vs. 2.8), a mother belonging to a lower socioeconomic status group (proportion of unskilled workers = 25% vs. 20%), and a mother who was a heavy smoker during pregnancy (3.6% smoking >10 cigarettes/day vs. 1.6%) in comparison with subjects with complete data. The subjects with missing data on any of the variables used in the analyses were themselves more likely to be male (52% vs. 47%), to be physically inactive at age 14 years (28% vs. 22%), to be a regular smoker both at age 14 years (7.2% vs. 5.8%) and at age 31 years (28% vs. 18%), and to come from a lower socioeconomic status group both at age 14 years (proportion of unskilled workers = 25% vs. 20%) and at age 31 years (30% vs. 23%).

We fitted a SEM for complete cases only ($n = 2,761$; data not shown) and compared the estimates with those obtained from analysis including all cases. In general, all of the estimates pointed in the same direction and were approximately of the same magnitude as in the all-cases analysis, but the effect of the *FTO* variant on BMI at age 14 years was attenuated in the complete-case analyses ($\beta = 1.22 \times 10^{-3}$, 95% CI: -7.46×10^{-3}, 9.89×10^{-3}) as compared with the all-cases analyses ($\beta = 4.83 \times 10^{-3}$, 95% CI: -2.60×10^{-3}, 12.3×10^{-3}). Note that the variables identified as influencing the missingness mechanisms were included in the model used for the all-cases analysis, and therefore imbalances between completers and noncompleters were implicitly taken into account in the all-cases analyses.

DISCUSSION

We analyzed the effects of *FTO* rs9939609 on BMI in a large sample with good representation of the general population of northern Finland using SEM, taking advantage of data on a large selection of nongenetic exposures associated with BMI during the life course. This study provided positive evidence for an association between the *FTO* variant and adult BMI despite control for several factors during the life course, and weaker evidence for associations with BMI at age 14 years and maternal BMI.

In this study, we were not able to observe any mediation or modification of the genetic effect through potentially relevant health behavioral exposures. However, the study showed that nongenetic life-course factors are important determinants of BMI development in addition to genetic factors, since many of these factors had a strong independent effect on BMI. We identified mediation through earlier BMI development, which is in line with previous findings that the *FTO* polymorphism would start affecting BMI by the time of adolescence (6, 9–11). However, the effect of the *FTO* variant on adult BMI was not fully mediated via earlier BMI development, indicating that the variant continues to function actively over the later life course as well. This is an important observation, since age-varying associations may cause failure to replicate a genetic-association finding (50, 51).

Our inability to find evidence for any mediation or modification through behavioral variables may be due to the low precision of some of the variables available in the present study. For instance, our rather crude measure of unhealthy diet, which is a surrogate for total energy intake, was in-

adequate in the attempt to shed light on the contradictory findings about the association between nutrition and *FTO*. The quality of data seems to play a huge role in the detection of gene-environment interaction analyses (52). In addition, we acknowledge a need for bigger sample sizes, since complicated models with several parameters require quite substantial sample sizes, as do gene-environment interaction analyses (53, 54).

The estimation of indirect and direct effects assumes that the specified model is correct (55). We aimed to build a model that would be a good approximation of the reality by including several relevant variables in the model. An indication of successful selection and model-fitting, in spite of some inaccuracies in variable measurement, was the fact that (for instance) the assumed model explained 34% of the total variation in adult BMI. However, one drawback of these models is that they are difficult to replicate in other cohorts as such, because not many cohorts have similar data. Therefore, we conducted cross-validation within the study population itself as a sensitivity check of the estimates.

Our analyses were conducted using observations with missing values via the expectation-maximization algorithm in maximum likelihood estimation. This allowed us to use the data to their full potential, with a noteworthy increase in statistical power. This is important in studies utilizing life-course data, since attrition and missing values are common in data collected over a long period of time. The expectation-maximization algorithm relies on the assumption that data are missing at random, which cannot be tested in practice (56). However, we conducted the analysis for complete cases only, and the estimates obtained were of a similar direction and, for most of the parameters, a similar magnitude as those from the analysis using observations with missing values. Only the estimated effect of the *FTO* variant on BMI at age 14 years was notably attenuated in the complete-case analyses as compared with the all-cases analyses. However, the 95% confidence intervals for the parameter estimates from the analyses overlapped.

The other advantage of using SEM is that it deals with the collinearity problem efficiently. We had repeated measurements of several variables, and putting them into a single equation in the multiple regression analysis could have produced problems in estimation due to collinearity, as demonstrated in a similar kind of study including repeated BMI measurements by Gamborg et al. (57). By using SEM, we avoided this problem and also obtained interpretable estimates of indirect and direct effects.

Recently, Mendelian randomization (58) has been widely used to study the mediating effect of variables in genetic epidemiology, and SEM has been very rarely used to address causal questions. The commentary by Tu (59) also highlighted the underutilization of SEM in epidemiology. Tu concluded that SEM might be a step in the right direction in the field of epidemiology. Especially in the genetic field, the advantages of SEM have been utilized in systems biology, quantitative trait loci analysis, twin studies, animal models, and linkage analyses (e.g., see Stein et al. (60)), but its use in genetic epidemiology is still very limited. However, some promising studies that have investigated gene-environment

interactions using this method have already been published (61, 62).

In conclusion, we estimated the effects of the *FTO* rs9939609 variant on BMI measurements taken over the life course in the largest study so far to collect extensive data from early pregnancy to adulthood, using a SEM approach, and we showed that the associations remain robust despite controlling for several relevant factors during the life course. Mediation of the *FTO* effect on adult BMI was observed via body mass development but not via an interaction with physical activity. Evidence for mediation or effect modification by diet could not be evaluated. More analyses of this kind should be carried out in large cohort studies with adequate statistical power that have carefully collected information over the life course. SEM proved to be an efficient analytical tool for modeling the complex networks around genetic and nongenetic variables.

REFERENCES

1. Smith SC Jr. Multiple risk factors for cardiovascular disease and diabetes mellitus. *Am J Med.* 2007;120(3 suppl 1):S3–S11.
2. Loos RJ, Lindgren CM, Li S, et al. Common variants near *MC4R* are associated with fat mass, weight and risk of obesity. *Nat Genet.* 2008;40(6):768–775.
3. Thorleifsson G, Walters GB, Gudbjartsson DF, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet.* 2009; 41(1):18–24.
4. Willer CJ, Speliotes EK, Loos RJ, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet.* 2009;41(1):25–34.
5. Meyre D, Delplanque J, Chèvre JC, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet.* 2009;41(2):157–159.
6. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science.* 2007; 316(5826):889–894.
7. Dina C, Meyre D, Gallina S, et al. Variation in *FTO* contributes to childhood obesity and severe adult obesity. *Nat Genet.* 2007;39(6):724–726.
8. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the *FTO* gene are associated with obesity-related traits. *PLoS Genet.* 2007;3(7):e115. (doi: 10.1371/journal.pgen.0030115).
9. Hakanen M, Raitakari OT, Lehtimäki T, et al. *FTO* genotype is associated with body mass index after the age of seven years

but not with energy intake or leisure-time physical activity. *J Clin Endocrinol Metab.* 2009;94(4):1281–1287.

10. Haworth CM, Carnell S, Meaburn EL, et al. Increasing heritability of BMI and stronger associations with the *FTO* gene over childhood. *Obesity (Silver Spring).* 2008;16(12):2663–2668.

11. Jess T, Zimmermann E, Kring SI, et al. Impact on weight dynamics and general growth of the common *FTO* rs9939609: a longitudinal Danish cohort study. *Int J Obes (Lond).* 2008;32(9):1388–1394.

12. Sabatti C, Service SK, Hartikainen AL, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet.* 2009;41(1):35–46.

13. Hjelmborg JB, Fagnani C, Silventoinen K, et al. Genetic influences on growth traits of BMI: a longitudinal study of adult twins. *Obesity (Silver Spring).* 2008;16(4):847–852.

14. Yang W, Kelly T, He J. Genetic epidemiology of obesity. *Epidemiol Rev.* 2007;29(1):49–61.

15. Fredriksson R, Hägglund M, Olszewski PK, et al. The obesity gene, *FTO*, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. *Endocrinology.* 2008;149(5):2062–2071.

16. Do R, Bailey SD, Desbiens K, et al. Genetic variants of *FTO* influence adiposity, insulin sensitivity, leptin levels, and resting metabolic rate in the Quebec Family Study. *Diabetes.* 2008;57(4):1147–1150.

17. Speakman JR, Rance KA, Johnstone AM. Polymorphisms of the *FTO* gene are associated with variation in energy intake, but not energy expenditure. *Obesity (Silver Spring).* 2008;16(8):1961–1965.

18. Timpson NJ, Emmett PM, Frayling TM, et al. The fat mass- and obesity-associated locus and dietary intake in children. *Am J Clin Nutr.* 2008;88(4):971–978.

19. Wardle J, Carnell S, Haworth CM, et al. Obesity associated genetic variation in *FTO* is associated with diminished satiety. *J Clin Endocrinol Metab.* 2008;93(9):3640–3643.

20. Stutzmann F, Cauchi S, Durand E, et al. Common genetic variation near *MC4R* is associated with eating behaviour patterns in European populations. *Int J Obes (Lond).* 2009;33(3):373–378.

21. Johnson L, van Jaarsveld CH, Emmett PM, et al. Dietary energy density affects fat mass in early adolescence and is not modified by *FTO* variants. *PLoS One.* 2009;4(3):e4594. (doi:10.1371/journal.pone.0004594).

22. Andreasen CH, Stender-Petersen KL, Mogensen MS, et al. Low physical activity accentuates the effect of the *FTO* rs9939609 polymorphism on body fat accumulation. *Diabetes.* 2008;57(1):95–101.

23. Cauchi S, Stutzmann F, Cavalcanti-Proença C, et al. Combined effects of *MC4R* and *FTO* common genetic variants on obesity in European general populations. *J Mol Med.* 2009;87(5):537–546.

24. Rampersaud E, Mitchell BD, Pollin TI, et al. Physical activity and the association of common *FTO* gene variants with body mass index and obesity. *Arch Intern Med.* 2008;168(16):1791–1797.

25. Vimaleswaran KS, Li S, Zhao JH, et al. Physical activity attenuates the body mass index-increasing influence of genetic variation in the *FTO* gene. *Am J Clin Nutr.* 2009;90(2):425–428.

26. Franks PW, Jablonski KA, Delahanty LM, et al. Assessing gene-treatment interactions at the *FTO* and *INSIG2* loci on obesity-related traits in the Diabetes Prevention Program. *Diabetologia.* 2008;51(12):2214–2223.

27. Jonsson A, Renström F, Lyssenko V, et al. Assessing the effect of interaction between an *FTO* variant (rs9939609) and

28. physical activity on obesity in 15,925 Swedish and 2,511 Finnish adults. *Diabetologia.* 2009;52(7):1334–1338.

28. Manolio TA. Cohort studies and the genetics of complex disease. *Nat Genet.* 2009;41(1):5–6.

29. Robins JM, Greenland S. The role of model selection in causal inference from nonexperimental data. *Am J Epidemiol.* 1986;123(3):392–402.

30. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives [editorial]. *Int J Epidemiol.* 2002;31(2):285–293.

31. De Stavola BL, Nitsch D, dos Santos Silva I, et al. Statistical issues in life course epidemiology. *Am J Epidemiol.* 2006;163(1):84–96.

32. Bollen KA. *Structural Equations With Latent Variables.* New York, NY: John Wiley & Sons, Inc; 1989.

33. Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. *Acta Paediatr Scand.* 1969;193(suppl):1–71.

34. Järvelin MR, Sovio U, King V, et al. Early life factors and blood pressure at age 31 years in the 1966 Northern Finland Birth Cohort. *Hypertension.* 2004;44(6):838–846.

35. Isohanni I, Järvelin MR, Rantakallio P, et al. Juvenile and early adulthood smoking and adult educational achievements—a 31-year follow-up of the Northern Finland 1966 Birth Cohort. *Scand J Public Health.* 2001;29(2):87–95.

36. Rantakallio P. Family background to and personal characteristics underlying teenage smoking. Background to teenage smoking. *Scand J Soc Med.* 1983;11(1):17–22.

37. Tammelin T, Laitinen J, Näyhä S. Change in the level of physical activity from adolescence into adulthood and obesity at the age of 31 years. *Int J Obes Relat Metab Disord.* 2004;28(6):775–782.

38. Laitinen J, Pietiläinen K, Wadsworth M, et al. Predictors of abdominal obesity among 31-y-old men and women born in Northern Finland in 1966. *Eur J Clin Nutr.* 2004;58(1):180–190.

39. US Department of Health and Human Services. *2008 Physical Activity Guidelines for Americans.* Washington, DC: US Department of Health and Human Services; 2008. (http://www.health.gov/paguidelines/pdf/paguide.pdf). (Accessed November 23, 2009).

40. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Stat Methodol.* 1977;39(1):1–38.

41. Muthén L, Muthén B. *Mplus: Statistical Analysis With Latent Variables.* Los Angeles, CA: Muthén and Muthén; 1998.

42. Little R, Rubin D. *Statistical Analysis With Missing Data.* New York, NY: John Wiley & Sons, Inc; 2002.

43. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull.* 1990;107(2):238–246.

44. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling.* 1999;6(1):1–55.

45. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models.* Newbury Park, CA: Sage Publications; 1993:445–455.

46. Jöreskog KG, Sörbom D. *LISREL 7: A Guide to the Program and Applications.* Chicago, IL: SPSS, Inc; 1988.

47. Järvelin MR, Elliott P, Kleinschmidt I, et al. Ecological and individual predictors of birthweight in a Northern Finland Birth Cohort 1986. *Paediatr Perinat Epidemiol.* 1997;11(3):298–312.

48. Cudeck R, Browne MW. Cross-validation of covariance structures. *Multivariate Behav Res.* 1983;18:147–167.

49. Sovio U, King V, Miettunen J, et al. Cloninger's temperament dimensions, socio-economic and lifestyle factors and metabolic syndrome markers at age 31 years in the Northern Finland Birth Cohort 1966. *J Health Psychol.* 2007;12(2): 371–382.

50. Lasky-Su J, Lyon HN, Emilsson V, et al. On the replication of genetic associations: timing can be everything! *Am J Hum Genet.* 2008;82(4):849–858.

51. Lyon HN, Emilsson V, Hinney A, et al. The association of a SNP upstream of *INSIG2* with body mass index is reproduced in several but not all cohorts. *PLoS Genet.* 2007;3(4):e61. (doi: 10.1371/journal.pgen.0030061).

52. Wong MY, Day NE, Luan JA, et al. The detection of gene-environment interaction for continuous traits: should we deal with measurement error by bigger studies or better measurement? *Int J Epidemiol.* 2003;32(1):51–57.

53. Dempfle A, Scherag A, Hein R, et al. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet.* 2008;16(10): 1164–1172.

54. Freathy RM, Timpson NJ, Lawlor DA, et al. Common variation in the *FTO* gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. *Diabetes.* 2008; 57(5):1419–1426.

55. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol.* 2002;31(1):163–165.

56. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods.* 2002;7(2):147–177.

57. Gamborg M, Andersen PK, Baker JL, et al. Life course path analysis of birth weight, childhood growth, and adult systolic blood pressure. *Am J Epidemiol.* 2009;169(10):1167–1178.

58. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22.

59. Tu YK. Commentary: is structural equation modelling a step forward for epidemiologists? *Int J Epidemiol.* 2009;38(2): 549–551.

60. Stein CM, Song Y, Elston RC, et al. Structural equation model-based genome scan for the metabolic syndrome. *BMC Genet.* 2003;31(4 suppl 1):S99. (doi: 10.1186/1471-2156-4-S1-S99).

61. Song XY, Lee SY, Ma RC, et al. Phenotype-genotype interactions on renal function in type 2 diabetes: an analysis using structural equation modelling. *Diabetologia.* 2009;52(8): 1543–1553.

62. Nock NL, Larkin EK, Morris NJ, et al. Modeling the complex gene × environment interplay in the simulated rheumatoid arthritis GAW15 data using latent variable structural equation modeling. *BMC Proc.* 2007;1(suppl 1):S118. (doi: 10.1186/1753-6561-1-s1-s118).