Obesity, Social And Demographic Context, And The Human Dopamine Receptor D4 (DRD4) Gene

Guang Guo, Kari North, and Seulki Choi
University of North Carolina, Chapel Hill

The prevalence of overweight (defined as body mass index (BMI) of 25-29.9 kg/m^2) and obesity (BMI of 30 or more) has increased dramatically in the US population over the past 20-30 years. Currently, over 65% of adults are considered overweight and 31% obese (Manson and Bassuk 2003). This health problem is becoming common in children; in 1999-2000, 16% of US children aged 12-19 years were classified as obese (Ogden et al. 2002).

The understanding of genetic effects on obesity has increased enormously in recent years (Comuzzie 2002; Comuzzie and Allison 1998; Chagnon et al. 2003). Studies have consistently shown that 40-70% of the variation in obesity-related measures is heritable. Recent reports of genome scans on BMI have found linkage to regions on several chromosomes using diverse samples (e.g., Lee et al. 1999; Norman et al. 1997; Hanson et al. 1998; Comuzzie et al. 1997; Mitchell et al. 1999; Hsueh et al. 2001; Ohman et al. 2000; Hager et al. 1998; Hinney et al. 2000; Price et al. 2001; Kissebah et al. 2000). The questions are not whether genes matter for obesity, but how they matter and how genetic effects are moderated by environmental influences across the life course.

Since genes are unlikely to have changed substantially over the previously two decades, the increase in overweight and obesity in the same period seems to be due to changes in environment. However, not all people become overweight or obese; individuals are subject to different sets of environmental factors. Obesity-related measures have been demonstrated to be associated with demographic factors such as race, gender, and age (Jacobson and Rowe 1998; Casas et al. 2001; Fabsitz et al. 1992; Gordon-Larsen et al. 1997), with social environmental measures such as socio-economic status (Kimm et al. 1996) and concentrated poverty, and with lifestyle measures such as physical activity and diet (Epstein et al. 1991; Epstein et al. 1995; Gordon-Larsen et al. 1999).

Obesity is a complex health problem influenced by both genes and environment, and most likely an interaction between the two. In the proposed study, we will investigate which demographic, social, and other environmental factors may boost or suppress the expression of heritability for obesity among adolescents in the US population using the genetically informative sibling data from the three waves of the National Longitudinal Study of Adolescent Health (Add Health) collected in 1994, 1995, and 2001.

Molecular genetic studies have identified a large number of genetic loci that might be responsible for obesity (Chagnon 2003 et al.; Comuzzie and Allison 1998). Any candidate gene study focusing on one or a few of these genes is likely to deal with only a small portion of the total phenotypic variance. Tremendous effort is necessary to examine each of these genes' impact on obesity and their interactions with potential environmental

factors. Our proposed investigation will serve as an exploratory analysis and give clues to which demographic, social, and other environmental factors might be interacting with genes to influence adolescent obesity. It will pave the way for molecular genetic studies that attempt to sort out the effects of specific genes on the development of obesity. While each of these molecular studies is likely to focus on a small fraction of the total phenotypic variance due to genetic factors, our sibling-based investigation will provide estimates related to the total variance for obesity.

The Add Health Study. The data source for the proposed analysis is the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a school-based study of the health-related behaviors of adolescents in grades 7-12 in 1994 in the United States (Bearman, Jones, and Udry 1997). The Study has been designed to explore the causes of these health behaviors. It postulates that various levels of social context such as families, friends, schools, neighborhoods, and communities play roles in shaping adolescents' health behaviors. Data are collected from adolescents themselves, from their parents, siblings, friends, peers, romantic partners, fellow students, and from school administrators. Already existing national data sources about neighborhoods and communities have been incorporated. The Add Health is longitudinal with adolescents interviewed three times during a 7-year period from 1994 to 2001.

In 1994-5, the Wave I in-school questionnaire was completed by more than 90,000 adolescents from 134 schools. The school sample was stratified by region, ethnic mix, size, urbanicity (urban/suburban/rural), and school type (public/private/parochial). All students who completed an in-school questionnaire, as well as those who did not complete a questionnaire but who were listed on a school roster, were eligible for selection into the in-home sample. The Wave II in-home survey interviewed a total of 20,754 adolescents from April through August of 1996. The in-home sample includes a core sample of 12,105 individuals representative of adolescents in grades 7 to 12 during the 1994-1995 school year in the United States. The in-home sample also includes special over-samples of several minority groups, adolescents from saturated schools in which all enrolled students were eligible for the in-home interview, disabled adolescents, and genetically related adolescents. At the time of the Wave II in-home interviews, a parental questionnaire was completed by one of the adolescent's parents (usually the mother) or guardians. Wave III interviewed the original Wave I respondents between August 2001 and April 2002. The respondents were between 18 and 26 years old in Wave III. Interviews with 15,197 original respondents were completed in Wave III.

The Sibling Sample. The data for our analysis come from the sibling sample within the Add Health Study, which has deliberately incorporated the behavior-genetic designs as components in an otherwise traditional survey. The sibling sample is composed of six groups: monozygotic twins, dizygotic twins, full biological siblings, half biological siblings, cousins, and biologically unrelated adolescents living in the same household. The result is an unprecedented genetic sample for a national study of this magnitude. The genetic sample in Wave I consists of more than 3,000 pairs of adolescents who are genetically related in varying degrees including 768 pairs of twins, 1,249 pairs of full siblings, 424 pairs of half siblings, and 657 pairs of biologically unrelated adolescents

living in the same household. These data represent pairs of adolescents who took the exact the same questionnaire and who share the same home environment, and in most cases the same school and same neighborhood. This design creates a precious opportunity to explore the relative contributions of genetics and environment to health and health behaviors.

In Waves I and II, the classification of the twins into monozygotic pairs and dizygotic pairs was based primarily on self-reports of confusability of appearance (Rowe et al.1999). Recently in Wave III, the zygosity of the twins were re-determined through a comparison of their match on DNA markers. Some corrections have been made over the original zygosity determination based on a questionnaire on confusability of appearance.

Obesity Measures. Add Health measures adiposity by the standard body mass index (BMI) calculated by BMI= (weight in kilos) / (height in meters ^{2}). Body weight was measured with a balance beam scale calibrated to zero and recorded to the nearest pound. Height was measured to the nearest centimeter with a fixed stadiometer, with the participant's head in the Frankfort horizontal plane. Add Health has collected three waves of obesity data in 1994-5 (Wave I), 1996 (Wave II), and 2001-2 (Wave III) spanning 7-8 years from when the respondents were aged 12-18 to when they are 19-25. The height and weight measures were self-reported in Wave I and both self-reported and measured in Waves II and III. The correlation between self-reports and measures are 0.93 for the males and 0.92 for the females in Wave II indicating the accuracy of self-reports. The longitudinal measure of obesity enables us to examine the possible changes in the relative contribution of genetics and environment to obesity during the study years when the prevalence of obesity is on the rise.

Alternatively, obesity may be measured by a binary variable based on BMI. Following the recommendations made by the World Health Organization (WHO) and developed by Must and others based on the US NHANES I data collected in 1971-1975, we use sex- and age- specific BMI cutoff points to define adolescent obesity and overweight (World Health Organization Expert Committee 1995; Must et al. 1991a; Must et al. 1991b). Specifically, the WHO recommends using the BMI 85th percentile to define overweight and this definition has been widely used in the US and other countries and referred to as the 'WHO/NCHS reference'. Although the WHO Expert Committee recommends using both the BMI 95th percentile and the triceps skinfold thickness 90th percentile to define adolescent obesity, the latter has been used infrequently due to the difficulty of measuring triceps skinfold thickness in large population-based studies. Instead, the BMI 95th percentile has been commonly used to define child and adolescent obesity (Must et al. 1991a; Must et al. 1991b; Himes and Dietz 1994; Kuczmarski et al. 2000). A recent panel of experts suggests that the BMI is the most appropriate index for assessing overweight during adolescence (Himes and Dietz 1994).

The human dopamine receptor D4 (DRD4) gene, located near the telomere of hromosome 11p, exhibits an unusual amount of expressed polymorphism (Lichter et

al. 1993; Ding et al. 2002; Grady et al. 2003). Much of this variation is the result of length and of SNP changes in a 48-bp tandem repeat (VNTR) in exon 3, encoding the third intracellular loop of this D2-like receptor. Alleles containing 2-11 repeats (2R-11R) have been found, with over 67 different haplotype variants uncovered to date (Ding et al. 2002; Grady et al. 2003). The three most common variants—2R, 4R, and 7R—represent 190% of the observed population allelic diversity. In most geographical locations, the 4R allele is the most common, whereas 2R and 7R allele frequencies vary widely (Chang et al. 1996; Ding et al. 2002).

Objective Of The Analysis and Preliminary Results. The purpose of the project is (1) to investigate whether there is an association between obesity and the 7R variant of the DRD4 gene and (2) once the association is established in the general population, to see if the association varies by social and demographic context. We attach twp tables that describe our preliminary results. Table 1 provides the descriptive statistics. Tables 2 and 3 display the regression results showing a consistent protective effect of the homozygote of the 7R. This result holds for raw BMI and BMI percentile score.

Table 1. Descriptive Statistics

Variable	Means	s.e.
BMI		
Wave 1	22.39	4.50
Wave 2, self reported	22.92	4.68
Wave 2, measured	22.94	4.94
Wave 3, self reported	25.66	5.85
Wave 3, measured	26.31	6.34
Combined sample	24.09	5.57
BMI Percentile		
Wave 1	58.48	28.22
Wave 2, self reported	57.73	28.86
Wave 2, measured	56.72	30.11
Wave 3, self reported	63.75	29.00
Wave 3, measured	66.21	29.15
Combined sample	60.67	29.31
Age		
Wave 1	16.03	1.64
Wave 2, self reported	17.00	1.65
Wave 2, measured	16.98	1.65
Wave 3, self reported	21.89	1.70
Wave 3, measured	21.87	1.71
Combined sample	18.82	3.08
Genotype		
1 if None	.62	
1 if Heterozygote	.34	
1 if Homozygote	.05	
Race		
1 if White	.64	
1 if Black	.20	
1 if Hispanic	.16	
Sex 1 if male	.47	
Number of Observations	10778	
Number of Individuals	2284	

Table 2. Regression analysis of BMI on Genotypes

	All-ethnicities	White	Black	Hispanic
Gene (Homozygote)	-1.18 *	-.48	-2.34 *	-1.77
	(.51)	(.65)	(1.10)	(1.11)
Age	.61 ***	.58 ***	.67 ***	.67 ***
	(.01)	(.01)	(.02)	(.02)
Sex (Male)	-.09	.22	-2.15 ***	1.16 *
	(.21)	(.25)	(.52)	(.51)
Race[a]				
Black	1.04 ***			
	(.30)			
Hispanic	1.10 ***			
	(.33)			
Intercept	12.28 ***	12.71 ***	13.28 ***	11.70 ***
	(.22)	(.27)	(.54)	(.58)
-2 log L	54504.6	34019.1	11214.2	9086.0
N. of Observations	10778	6894	2136	1748
N. of individuals	2284	1448	464	372

[a] White is the reference category
* p<.05; ** p<.01; *** p<.001

Table 3. Regression analysis of BMI Percentile on Genotypes

	All-ethnicities	White	Black	Hispanic
Gene (Homozygote)	-7.87 **	-4.33	-13.38 *	-12.31 *
	(2.71)	(3.67)	(5.24)	(5.88)
Age	1.34 ***	1.28 ***	1.10 ***	1.85 ***
	(.05)	(.06)	(.12)	(.12)
Race[a]				
Black	5.15 **			
	(1.62)			
Hispanic	7.40 ***			
	(1.76)			
Intercept	33.80 ***	34.67 ***	43.66 ***	31.76 ***
	(1.26)	(1.44)	(2.67)	(2.83)
-2 log L	92848.0	59180.7	18581.6	15013.0
N. of Observations	10778	6894	2136	1748
N. of individuals	2284	1448	464	372

[a] White is the reference category
* p<.05; ** p<.01; *** p<.001