Twins of Mistaken Zygosity (TOMZ): Evidence for Genetic Contributions to Dietary Patterns and Physiologic Traits

Erica P. Gunderson,[1] Ai-Lin Tsai,[1] Joe V. Selby[1], Bette Caan,[1] Elizabeth J. Mayer-Davis,[2] and Neil Risch[3]

[1] *Kaiser Permanente, Division of Research, California, United States of America*

[2] *Center for Research in Nutrition and Health Disparities, University of South Carolina, Arnold School of Public Health, Columbia, South Carolina, United States of America*

[3] *Institute for Human Genetics, University of California, San Francisco, California, United States of America*

Twin designs, comparing correlations in monozygotic (MZ) versus dizygotic (DZ) twins, have an extensive history. One major confounder in such studies is that MZ twins may share postnatal environmental influences more so than do DZ twins. To avoid such confounding, twins separated at or soon after birth have been studied, but their scarcity often makes this approach impractical. Another method has been to measure the degree of contact twins have maintained over time, and adjust the observed correlations. Here, we remove confounding by utilizing the discrepancy between biological and self-perceived zygosity to separate environmental from genetic sources of twin similarity. We analyzed dietary patterns and physiologic traits in 350 female twin pairs of the 1988 Kaiser Permanente Twin Registry. Among twin pairs, 175 were MZ by self-report and genetic testing (MZC), 136 were DZ by self-report and genetic testing (DZC), 30 were MZ by genetic testing but not by self-report (MZW), and 9 were DZ by genetic testing but not by self-report (DZW) but were excluded due to small sample size. For healthy food patterns, MZC and MZW intraclass correlations were similar and greater than for DZC, yielding positive and significant heritability estimates. For unhealthy food patterns, the MZC, MZW and DZC correlations were similar with no significant heritability. For physiologic traits, MZC and MZW correlations were similar and higher than those for DZC, indicating significant heritability, except for insulin for which MZW and DZC were similar and which showed modest heritability. Twins of mistaken zygosity (TOMZ) provides a useful approach to robust determination of heritability.

Comparisons of correlations for monozygotic (MZ) versus dizygotic (DZ) twin pairs have an extensive history in twins research. One major confounder in such studies is that the influences from sharing of postnatal environments may be greater for MZ than DZ twins. To avoid such confounding, one approach has been to study twins separated at or soon after birth, but their scarcity often makes this approach impractical. Another method has been to measure the degree of contact twins have maintained over time, and adjust observed correlations for degree of contact. However, the degree of contact may not account for all the difference in environmental similarity between MZ and DZ twins, especially early in life. Discordance in perceived versus true zygosity offers an alternative practical approach to control for environmental confounding in large studies that has rarely been utilized in twin research since first proposed in the 1960s.

In 1968, Scarr examined sources of environmental bias in twin studies, including the assumption of equal environmental variances for MZ and DZ co-twins depending on whether the parents are correct about their twins' zygosity (Scarr, 1968). Based on whether twin zygosity is perceived correctly, two important factors may be confounded: (1) the greater genetic differences in DZ twins, and (2) the greater differences in parental treatment of DZ pairs. In this early study where four MZ pairs and seven DZ pairs were misclassified, the data generally showed that parental similarity in the treatment of twins is determined by their perceived genetic relatedness (Scarr, 1968), although the sample sizes were quite small.

Studies comparing designation of zygosity by serologic markers or DNA typing versus questionnaires have reported misclassification rates ranging from 4% to 10% (Cohen et al., 1975; Hrubec & Neel, 1981; Kasriel & Eaves, 1976; King et al., 1980; Magnus et al., 1983; Nichols & Bilbro, Jr., 1966; Reed et al., 2005; Rietveld et al., 2000). Moreover, questionnaire response comparisons with serological markers in adult twins from a large United States

Received 16 January, 2006; accepted 6 April, 2006.

Address for correspondence: Erica P. Gunderson, Kaiser Permanente, Division of Research, 2000 Broadway, Oakland, CA 94612, USA. E-mail: epg@dor.kaiser.org

cohort indicated that a greater proportion of MZ pairs thought that they were DZ (7.6%; Jablon et al., 1967) than DZ pairs who thought they were MZ (2.2%; Hrubec & Neel, 1978).

In 1980, King and colleagues compared the diagnosis of twin zygosity by self-assessment with genetic analysis based on 20 independent genetic markers for 173 adult like-sex twin pairs from the Kaiser-Permanente Twin Registry (King et al., 1980). Concordance of the laboratory diagnosis with twins' self-assessment was about 90%. The error rate among laboratory diagnosed MZ twins was much higher than previous studies because of the greater number of markers used to establish zygosity in this study. The most predominant error was that 17% of MZ twins believed incorrectly that they were DZ, while only 2% of the DZ twins believed they were MZ. These investigators also suggested that twin pairs who believe that they are DZ, but who are MZ based on the genetic analysis, represent an important group to evaluate the genetic versus environmental influences on behavioral characteristics such as diet or smoking (King et al., 1980). The MZ twins who believe themselves DZ can be considered 'environmentally DZ, genetically MZ' twins, and might be used to study genetic and environmental influences separately. Comparability of the within-pair environmental discordance for these twins to the within-pair environmental discordance for other MZ twins may be helpful in determining the extent of genetic influences on twin pair shared environment as described by Scarr (1968).

If postnatal environmental sharing between twins is a function of their perceived zygosity, comparison of MZ correctly versus DZ incorrectly self-identified twin pairs should provide direct heritability estimates uncontaminated by excess environmental sharing of MZ twins; by contrast, comparison of correctly versus incorrectly self-identified pairs for MZ twins as well as for DZ twins should provide a direct estimate of MZ twin correlation due to excess environmental sharing of MZ compared to DZ twins.

Here we applied this method, twins of mistaken zygosity (TOMZ), as a distinctive approach to estimate the separate contributions of genetic and environmental influences on common traits. To demonstrate the utility of the TOMZ method, we selected two examples, dietary patterns and physiologic measures, which differ with respect to the degree of genetic and environmental influence. Genetic factors strongly influence levels of physiologic risk factors for coronary heart disease and insulin resistance (Acton et al., 2004; Austin et al., 1987; Edwards et al., 1997; Friedlander et al., 1997; Heller et al., 1993; Hewitt, 1997; Katoh et al., 2005; Selby et al., 1987). However, less is known about the influence of the shared environment on these risk factors for chronic disease. Dietary traits are influenced by environmental factors, and possibly to a lesser degree by genetic factors. Genetic differences in food preferences can be traced to biological variations in digestive enzymes in invertebrates. For example, amylase genotype has been correlated with food preference in amphipod crustraceans (Guarna & Borowsky, 1993). In humans, most studies have found evidence for genetic influences on overall energy and nutrient intakes (Collaku et al., 2004; de Castro, 2002; Reed et al., 1997), but evidence that meal patterns, eating behaviors (de Castro, 1993a, 1993b; Keller et al., 2002; Klump et al., 2000; Tholin et al., 2005; Segal, 2001) and specific food preferences are inherited is less consistent (Heitmann et al., 1999; Reed et al., 1997; van den Bree et al., 1999). Evidence from a study of MZ twins with discordant body weights indicates that preferences for fat, sweet foods, and alcohol may be acquired behaviors (Rissanen et al., 2002). Moreover, findings show that genetic influences are most likely to affect the overall pattern of interrelated food selections rather than individual food items or micronutrients (Reed et al., 1997; Rozin & Millman, 1987). For example, more healthful eating patterns may consist of multiple food items that affect future chronic disease and obesity, and the dietary patterns are likely to be influenced by both genetics and the environment.

Many studies have shown a link between overall energy and fat intake and genetic variation (Reed et al., 1997), yet, estimates of the contribution of the genetic and shared environmental components in determining 'healthy' and 'unhealthy' food patterns have not been well-characterized. In one study of over 4600 male and female twins, two independent eating patterns were identified: the first pattern included food items high in fat, salt and sugar, and the second was characteristic of a 'healthful' eating pattern (van den Bree et al., 1999). About 60% to 85% of the variability in eating patterns was associated with environmental factors.

The objective of this study was to demonstrate that TOMZ is a practical method for estimating the separate contributions of genetic and environmental influences on dietary patterns as well as physiologic risk factors for chronic disease. This method is based on comparison of the intraclass correlation estimates for twins whose zygosity had been classified correctly compared with twins whose zygosity had not.

A group of 350 female twin pairs from the Kaiser Permanente Twin Registry who participated in dietary interviews and physiologic assessments in 1988 and 1989 were included based on zygosity assessed by self-report and by laboratory diagnosis.

Materials and Methods

Study Subjects

The study cohort consisted of members of the Kaiser Permanente Twin Registry who agreed to visit the Kaiser Permanente Medical Center in Oakland, California in 1978 and 1979 for a study of coronary

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

heart disease risk factors, and health outcomes. This sample of 434 female twin pairs were born in 1960 or earlier and lived in the San Francisco Bay area. Recruitment procedures, sample selection, data collection study protocol, laboratory methodology and other basic features have been previously described in detail elsewhere (Austin, 1993; Austin et al., 1987; Edwards et al., 1994; Mayer et al., 1993). Briefly, study participants had a physical examination, a comprehensive laboratory test panel, zygosity assessment, and measurement of lipid and lipoprotein levels, and were administered a health questionnaire. The average age at the first examination was 42 years. Ninety per cent of the twin pairs reported themselves to be white, 7% were black and the remaining 3% were of other ethnicity. Thirty-nine per cent were college graduates, 34% had some college, 23% were high school graduates, and the remaining 4% had not completed high school. Sixty-seven per cent were living with a spouse or partner, 16% were divorced or separated and 8% were single.

The present analysis includes 350 adult female twin pairs (80.6% of the original cohort) who participated in the first examination of the Kaiser Permanente Women Twins Study in 1978 and 1979, and were reexamined at a second visit 11 years later in 1989 and 1990. The average age of the women at the second examination was 50 years (range 30 to 90 years), and similar to the original cohort, the majority (90%) were white, and more than 90% perceived their overall health status as good or excellent.

Data Collection

Dietary intake during the month before the clinical exam visit was assessed from the Health Habits and History Questionnaire, a self-administered, 100-item food frequency questionnaire (FFQ), developed at the National Cancer Institute (Block et al., 1986). The computer-scannable version includes nine frequency categories, and three portion size categories (small, medium and large) for each food item. The methods used to administer the FFQ to subjects have been previously described in detail (Mayer et al., 1993). Briefly, subjects reported the number of servings per day or per week and portion size consumed for each of the 100 food items within the past month.

In data analyses, the 100 food items were initially collapsed into 18 categories based on preselected major food types, protein, saturated fat and other nutrient content as follows: beef, pork, hot dogs, fish/poultry, citrus fruits, other fruits, tomatoes, carrots, yellow/green vegetables, salad, other vegetables, rice/potatoes, high fiber grains, cheese, butter/margarine, eggs, ice-cream, and sweets. Factor analysis was utilized to reduce the number of food categories further, and to identify distinct dietary patterns.

Physiologic measurements including anthropometry (body mass index [BMI], waist circumference), and fasting blood lipoproteins, triglycerides, glucose, and insulin were obtained at the examination in 1988 or 1989 under standardized methods which have been described in detail elsewhere (Friedlander et al., 1997; Mayer et al., 1993; Selby et al., 1994). Blood samples were obtained from subjects in the morning after an overnight fast of 9 hours or more and analyzed for plasma triglycerides, lipoproteins, glucose and insulin using methods previously reported. Height (m) and weight (kg) were obtained from subjects dressed in light clothing and without shoes. BMI was calculated as weight divided by height squared (kg/m^2). Waist circumference was measured at the natural indentation, or midway between the iliac crest and the lowermost extension of the rib cage in the midaxillary line. Blood samples were collected into EDTA-containing tubes. Plasma was separated within 2 hours and stored under refrigeration for a maximum of 72 hours before processing. Total cholesterol (TC), high-density lipoprotein cholesterol (HDL), and triglycerides were assessed using standard methods (Nagele et al., 1984; Warnick et al., 1985). Low density lipoprotein cholesterol (LDL) was estimated using the Friedewald formula (Friedewald et al., 1972). All assays were performed in the same laboratory (Selby et al., 1994). Plasma glucose levels were measured by the glucose oxidase method. Plasma insulin concentration was measured by radio-immunoassay at SmithKline Laboratories (Van Nuys, CA) using commercial kits (Selby et al., 1994). We excluded women self-reporting diabetes or taking diabetes medications from the sample for analyses of plasma glucose and insulin levels (n = 19).

Zygosity Assessment and Classification

Self-reported zygosity for each twin was based on her response to the question asked by the study interviewer of whether she was an identical or nonidentical twin or uncertain at the baseline interview in 1978. There were no additional questions about their perceived zygosity. Actual zygosity for each pair was determined by analysis of 20 polymorphic loci in 1979 (King et al., 1980). Based on the concordance or discordance of the self-reported and the laboratory zygosities, each twin pair was classified into one of four groups: correctly self-identified as MZ (MZC = MZ Correct), correctly self-identified as DZ (DZC = DZ Correct), those not correctly self-identified as DZ (MZW = MZ Wrong) and those not correctly self-identified as MZ (DZW = DZ Wrong). Pairs classified as MZC included pairs where both twins correctly reported their zygosity as MZ, or within a pair one of the twins correctly reported zygosity and the other was uncertain. The classification in the DZC group used the same criteria as it pertained to DZ instead of MZ zygosity. Pairs classified as MZW included pairs where at least one twin within the pair incorrectly classified herself as DZ or both twins were uncertain of their zygosity. The classification into the DZW group used the same criteria as it pertained to DZ instead of MZ status.

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

Statistical Methods

First, the dietary and physiologic variables were examined for normality. For those variables that were not normally distributed due to skewness, transformations were applied to approximate normality. For some variables the transformation required was square root, for others a log transformation was necessary, and for other extremely skewed variables, a double log transformation was needed.

Factor analysis with orthogonal rotation was applied to the 18 dietary items for data reduction purposes, and to explore patterns in the dietary data. The computer package SAS version 8 was used to produce two factors using PROC FACTOR, assuming varimax rotation and all prior communalities set to 1.0.

Estimation of Heritability

We first calculated intraclass correlations for the MZC, MZW and DZC groups for each variable to determine whether patterns were more consistent with a genetic, environmental, or mixed model. To estimate heritabilities, the computer program Mx, a structural equation modeling package, was used to fit structural equation models to observed twin covariance matrices and to obtain maximum likelihood estimates of components of variance (Neale et al., 2002). These components included variance due to genetic heritability (called 'A' or h^2), variance due to shared twin environment (called 'C' or c^2), and variance due to nonshared environmental effects (called 'E'). Heritability is estimated as the proportion of total variance due to A, while 'C' reflects the proportion of variance due to shared twin environmental exposures. The MZC and DZC groups, the two largest, were used to calculate heritability estimates.

Results

The zygosity of the 350 female adult twin pairs was determined by two methods: self-report and genetic analysis as shown in Table 1. In total, 175 pairs were MZ by self-report and genetic testing (MZC); 136 pairs were DZ by self-report and genetic testing (DZC); 30 were MZ by genetic testing but not by self-report (MZW); and nine pairs were DZ by genetic testing but not by self-report (DZW). The DZW pairs were excluded from further analysis due to small sample size.

The results of the principal components analysis gave evidence for two dietary factors accounting for 17% and 14% of the variance, respectively. These factors segregated the 18 food types (Table 2) into two distinct factors which are suggestive of 'healthy' and 'unhealthy' food patterns. The 'healthy' dietary pattern (Factor 1) was characterized by a low saturated fat, high fruit and vegetable foods intake (high fiber), and less concentrated sugars. The 'unhealthy' dietary pattern (Factor 2) was characterized by high saturated fat from animal food sources, refined carbohydrates and sugars (low fiber), and fatty meats intake. The remaining 16 factors explained between 3% and 7% of the variance, and it was clear in examining the distribution of variance explained that the first two factors were prominent while the remaining 16 formed a simple continuum. Therefore all our subsequent analyses focused only on the first two factors.

The food types that clustered together into Factor 1 (healthy food pattern) included fish or chicken, and vegetable food sources such as carrots, tomatoes, salad, green or yellow vegetables, fruits, high fiber grains, rice and potatoes. Food types that clustered together to form Factor 2 (unhealthy dietary pattern) included

Table 1

Kaiser Permanente Female Twin Pairs by Self-Reported and Genetically Determined Zygosity

	Genetic zygosity		
Self-report zygosity	MZ	DZ	Total
MZ/MZ	171[a]	5[b]	176
DZ/DZ	19[c]	129[d]	148
MZ/Unknown	4[a]	1[b]	5
DZ/Unknown	1[c]	7[d]	8
Unknown/Unknown	10[c]	3[b]	13
All	205	145	350

Note: [a]Defined as MZ-Correct (MZC; $n = 175$)
[b]Defined as DZ-Wrong (DZW; $n = 9$)
[c]Defined as MZ-Wrong (MZW; $n = 30$)
[d]Defined as DZ-Correct (DZC; $n = 136$)

Table 2

Factor Analysis Results (Weightings) for 18 Dietary Food Categories

Food Categories	Factor 1 ('healthy' dietary pattern)	Factor 2 ('unhealthy' dietary pattern)
Fish/chicken	.489	.154
Tomatoes	.443	.076
Carrots	.602	–.088
Salad	.612	–.044
Yellow/green vegetables	.618	–.089
Other vegetables	.594	.026
Citrus fruits	.498	–.019
Other fruits	.631	.046
High fiber grains	.553	.004
Rice/potatoes	.310	.461
Beef	–.051	.686
Pork	–.050	.440
Hot Dogs	–.056	.604
Eggs	.016	.517
Butter/margarine	.140	.578
Cheese	.053	.496
Ice-cream	.048	.374
Sweets/soda/desserts	–.211	.500

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

largely animal food sources, such as beef, pork, hot dogs, eggs, cheese, and higher fat and concentrated sweet foods including ice-cream, butter, margarine, soda, and desserts.

Intraclass correlations for the food types and the two factors signifying the two eating patterns were calculated for each of the zygosity groups, and estimates of heritability and shared environment influencing the dietary patterns were derived from the analysis using Mx (Table 3). Heritability is sometimes determined as twice the difference between the MZ and DZ twin intraclass correlations. For most variables in Table 3, an estimate based on this difference is similar to the Mx derived estimate of h^2, with the exception of LDL cholesterol, and to a lesser extent fasting glucose and insulin. The reason for the difference between the two methods of estimation is that the Mx analysis is variance based, and assumes the trait variance in MZ and DZ twins is the same, while the estimate based on the difference in intraclass correlations does not make this assumption. Therefore, when the variances are different between the two types of twins, the heritability estimates will differ. The largest difference in trait variance between MZ and DZ twins was for LDL cholesterol and less so for fasting insulin and glucose. However, since none of the variance differences between MZ and DZ twins were significant, we believe the Mx estimates of heritability to be the more robust.

The heritability based on the Mx analysis within the 'healthy eating pattern' was strongest for vegetable food types, which included carrots, tomatoes, salad, green or yellow vegetables, with h^2 estimates ranging from .33 to .43 for each food ($p < .05$); the overall heritability estimate for dietary Factor 1 was strong at .496 ($p < .01$; Table 3). None of the foods classified within the 'unhealthy' dietary pattern showed any statistically significant genetic influence on dietary choices, although several of the environmental components (c^2) were significant. The overall estimate for heritability for dietary Factor 2 was zero, whereas the environmental component c^2 was significant at .315 ($p < .05$; Table 3).

By contrast, twin pair intraclass correlations and heritability estimates for physiologic, anthropometric and metabolic characteristics were much stronger than for dietary patterns. Genetic influences on body weight and waist girth had the strongest effect with h^2 estimates above .75 ($p < .001$) in women twins. Genetic variance strongly influenced fasting serum lipoprotein and triglyceride levels with heritability estimates ranging from .50 to .74 ($p < .001$). Much weaker estimates of heritability were found for fasting glucose and insulin levels, .393 and .247, ($p < .05$), respectively. The influence of shared environment was double that of the genetic contribution for fasting insulin levels, an index of insulin resistance, with c^2 estimates of .499 ($p < .001$).

For most variables with significant heritability estimates and nonsignificant c^2 estimates from the Mx analysis (i.e., some of the healthy food types, Factor 1 and most of the physiologic variables), the intraclass correlations for the MZW group were very similar to the MZC group, and greater than for the DZC group. This observation reinforces the fact that the heritability estimate we derived from the MZC and DZC groups actually reflects genetic influences rather than environmental influences that are shared more by MZ than DZ twins. By contrast, those variables with a significant c^2 component, such as Factor 2 and fasting insulin, showed either that the three groups had similar intraclass correlations (Factor 2) or that the MZW and DZC groups had similar intraclass correlations, but differed (were less than) the MZC correlation (fasting insulin), reinforcing the evidence for shared environmental as opposed to genetic influences.

We also examined correlations between the dietary variables, including Factors 1 and 2, and the seven

Table 3

Twin Intraclass Correlations and Heritability Estimates for Dietary Factors and for Physiologic (Anthropometric and Metabolic) Characteristics

	Correlations by zygosity group				
Variable	MZC *(n)*	MZW *(n)*	DZC *(n)*	h^2	c^2
Dietary Factor 1 'healthy pattern'	.546 (175)	.498 (30)	.239 (136)	.496**	.030
Dietary Factor 2 'unhealthy pattern'	.287 (175)	.428 (30)	.318 (136)	.000	.315*
BMI	.763 (174)	.784 (30)	.390 (135)	.781***	.000
Waist circumference	.769 (174)	.797 (30)	.406 (135)	.759***	.022
Fasting glucose	.560 (162)	.587 (29)	.295 (122)	.393*	.156
Fasting insulin	.730 (165)	.641 (29)	.663 (126)	.247*	.499**
LDL cholesterol	.742 (172)	.746 (29)	.313 (130)	.640***	.076
HDL cholesterol	.697 (172)	.889 (29)	.296 (130)	.735***	.000
Triglycerides	.689 (172)	.769 (29)	.455 (130)	.500***	.210

Note: *$p < .05$

**$p < .01$

***$p < .001$

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

physiologic traits across all the twins. We found these correlations to be uniformly small and nonsignificant, ranging from –.07 to +.12 (median .025). These results indicate that the observed heritabilities for the physiological variables cannot be attributed to potential genetic influences underlying dietary choices, in particular the healthy diet variables that showed significant heritability.

Discussion

An underlying assumption of twin analysis is that with respect to the relevant nongenetic contributions, the environmental influences of DZ co-twins are as similar as those of MZ co-twins. If higher MZ versus DZ correlations are attributed entirely to genetic difference, whereas in fact DZ variations originate in greater differences in their environmental exposures, estimates of heritability will be biased in favor of genetic origins. It has been noted that MZ co-twins are more similar in a variety of behaviors, including sports activities, as well as in the treatment they receive from parents (Jones, 1955). Home environments of MZ co-twins tend to be more similar than those of DZ co-twins (Scarr, 1968), and hence, greater similarity in environmental exposures may be an important source of bias in twin studies where the twins are influenced by family members according to zygosity. The method evaluated in this study, TOMZ, is based on comparison of the correlations for twins who had been classified correctly versus those who had not been classified correctly by zygosity. TOMZ is useful for estimating the separate contributions of genetic and environmental influences including differences in the shared environment which would generally have been included in the influences attributed to genetics. For example, misclassified MZ twins are more likely to be treated differently by their parents, and misclassified DZ twins are more likely to be treated similarly. These parental environmental influences would result in a lessening of the differences between misclassified DZ pairs and exaggerating the differences between misclassified MZ pairs. The TOMZ method would allow for robust estimates of heritability that would minimize the environmental bias.

This method is applicable for estimating genetic and environmental influences for common behavioral or physiologic traits such as dietary patterns, physical activity, blood lipid profiles, and obesity where gene–environment interactions are most likely to occur.

Previous studies, including the Kaiser Permanente Women Twins Study, have found that genetics influence the levels of coronary heart disease risk factors, including total and central obesity, lipid profiles, and insulin resistance (Austin et al., 1997; Mayer et al., 1996; Rose et al., 1998; Selby et al., 1987). Genetic heritability estimates for anthropometric measurements ranged from .72 to .82 for waist circumference adjusted for age and BMI (Rose et al., 1998), and .55 for relative weight (weight for height; Austin et al., 1987). Consistent with our work, the estimates of genetic heritability are somewhat weaker for physiologic measurements such as fasting serum lipids, insulin and glucose levels adjusted for BMI and behavioral factors (Mayer et al., 1996; Selby et al., 1987).

Gene–environment interactions, including effects of dietary fat intake and genetic polymorphisms on plasma lipoprotein levels, have been widely studied. Evidence is less available for genetic determinants of dietary intake and eating behaviors. Genetic influences on food preferences have been demonstrated for only a limited number of food items (Falciglia & Norton, 1994; Heitmann et al., 1999; Krondl et al., 1983; Reed et al., 1997; Rozin & Millman, 1987). These studies have found evidence for genetic effects on the frequency of intake of individual food items such as flour and grain products, fruits and vegetables, citrus fruits, rice, bacon, cottage cheese, and chili pepper. For example, Rozin and Millman (1987) showed that correlations for individual foods are very weak for young adults and their parents ranging from zero to .4 (Rozin & Millman, 1987). One study found strong genetic influences on intakes of cottage cheese and orange juice (Falciglia & Norton, 1994), but another study identified influences on flour and grain products, fruits and vegetables, and rice (Heitmann et al., 1999). The inconsistency of the estimates and specific individual food items across studies may suggest that genetic influences are stronger for dietary patterns than on individual food items or absolute food intake.

Most twin studies have found evidence that eating behaviors have some genetic determination. A few studies have found significant genetic influences on meal frequencies and meal size that have been attributed to the heredity of overall macronutrient and caloric intake (de Castro, 1993a, 1993b). The separate genetic and environmental influences on dietary intake were examined in 66 MZ and 51 DZ twins who were reared apart and 30% of the variance in self-reported diet was attributable to genetic factors, and current family environment exerted only slight effects on dietary intake (Hur et al., 1998). In a very large twin study (n = 4640 pairs), genetic and environmental influences on eating patterns were examined in a sample of predominantly female, white twins over the age of 50 who self-reported zygosity (van den Bree et al., 1999). The study identified two independent eating patterns, the first included foods that had high amounts of fat, salt and sugar, and the second was consistent with healthful eating habits. For 'unhealthy' and 'healthy' eating patterns, respectively, about 53% and 48% of the variability was related to specific environmental factors, and about 17% and 12% of the variability was related to shared environment, which was not significant (van den Bree et al., 1999). These findings

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

are consistent with our findings for dietary patterns, in that genetic factors contributed to the two different eating patterns with heritability estimates of about 30% to 40% (van den Bree et al., 1999). Another similarity is that two distinct dietary patterns were identified, which were distinguished by the same attributes with respect to high fat, salt and refined sugars for the 'unhealthy' pattern.

In our study, the healthy dietary pattern accounted for a slightly higher proportion of the total variance in dietary intake than the unhealthy dietary pattern (17% vs. 14%). The reverse was true in the analysis by van den Bree et al. (1999). The estimates of heritability taking into account the misclassification of twins in our study were 50% for the healthy dietary pattern (Factor 1), and 0% for the unhealthy dietary pattern (Factor 2). The estimate for Factor 1 is comparable to that estimated by van den Bree et al. (1999; 40% heritability for the healthful eating pattern). Our estimate of the heritability of dietary Factor 2, the unhealthy eating pattern (0%), is lower than the 30% estimated by van den Bree et al. (1999). Our finding suggests that genetic influences are much greater for certain characteristics of dietary choices. It is unclear why our results for this factor differ from those of van den Bree et al. (1999).

In the van den Bree et al. (1999) study, zygosity was determined exclusively by self-report. To address the question of shared environmental bias, these authors examined twin correlations stratified by degree of contact of the twins in adulthood. They found little evidence that degree of contact influenced the twin correlations (i.e., those with more contact did not demonstrate higher correlations) for the dietary factor variables. This result is similar to ours in that we found no evidence for a difference in the MZ twin correlations based on whether they considered themselves to be MZ or not. Thus, the difference between our heritability estimates for the 'unhealthy' factor is unlikely due to different approaches to dealing with the confounding due to shared twin environment.

Our results show that biological zygosity determined the correlations for the anthropometric and physiologic measures. Earlier studies that did not use biological zygosity, but self-report of zygosity, are likely to have underestimated the true heritabilities for most variables. However, since the proportion of misclassified twins is small, the underestimation is likely not to be very large. Also, we note that among the 30 twin pairs included in the MZW group, 20 were incorrect about their zygosity while 10 were uncertain. These uncertain pairs may not represent the same degree of discordance in terms of environmental and genetic correlation as would be the case for those pairs who were incorrect. This could possibly also have attenuated differences between the MZC and MZW groups, if they existed.

Currently, there appear to be three different approaches for addressing the potential bias of differential environmental correlations in twin studies: (1) analysis of separated twins, (2) adjustment for degree of twin contact, and (3) separate analysis of twins of mistaken zygosity. Most likely the first paradigm is the most robust, in that twins are reared in separate (and presumably independent) postnatal environments, so that observed correlations should only reflect genetic influences, aside from those that occur prenatally or very early in life, prior to adoption. Of course, this approach has had limited application because of the scarcity of twins who have been independently adopted without subsequent contact.

The second and third paradigms provide an interesting and complementary approach. Measuring the degree of twin contact can serve as a surrogate for the degree of environmental correlation for a twin pair, and thus adjusting observed trait correlations for this degree of contact can eliminate environmental confounding. However, this will only be true to the extent that the correlations of relevant environmental exposures are also indirectly measured by the degree of contact. So, for example, if the relevant exposure is parental treatment, this adjustment will only be useful to the extent that similarity of parental treatment correlates with degree of time spent together by the twins. For example, one can imagine a scenario where parents treat their twins quite differently despite the fact that the twins spend a lot of time with each other; or, conversely, that parents treat the twins similarly even though they spend little time together. On the other hand, the TOMZ method is effective only to the extent that the correlation in environmental exposure tracks with perceived zygosity. For the example above, this would mean that parents treat their twins with a degree of similarity dependent entirely on their perception of their zygosity and not based on their true zygosity. It has been argued that twins may create a shared environment that is more dependent on their true zygosity than their perceived zygosity (Rutter et al., 2001). If that is the case, and the relevant environmental exposure is directly related to the environment that the twins create, then the perceived zygosity may be less relevant for addressing the environmental source of correlation than the actual zygosity. However, this scenario also portrays an interesting paradox, namely that the genes involved in determining the trait are not directly related to the trait itself, but to twin behaviors that lead to their environmental exposure. Possibly the best, or only, way to resolve this dilemma is to identify the relevant environmental exposures (rather than their surrogates) and directly measure their degree of correlation in the twins and their direct impact on the phenotype.

Most twin studies do not rely on self-report of zygosity, but rather on a questionnaire eliciting degree of physical similarity of the twins and how often they are mistaken by others. Presumably, there is some correspondence between questionnaire-diagnosed zygosity and self-belief of zygosity, but they are not necessarily the same. For example, some twins who

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

believe themselves to be DZ might be thought by others to be MZ and vice versa. On the other hand, some twins who are DZ but believe themselves to be MZ could dress alike and appear more similar to outside observers. One could also perform an analysis of twins discordant for zygosity between questionnaire results and DNA analysis (again, numbers of discordant pairs is unlikely to be large), and in this case one would be comparing biological MZ to biological DZ twins, controlling for degree of physical similarity. Again, the degree of potential environmental confounding this approach would eliminate depends on the extent to which the relevant environmental exposure tracks with physical characteristics of the twins.

For some traits, at least, it may be that self and parental perception of zygosity are the categories most allied with the relevant environmental exposures. In this regard, it is also pertinent to note that perception of zygosity may relate to the number of placentas found at birth. For example, if MZ twins are often believed at birth to be DZ twins because they are dichorionic, then the contrast between MZ twins of correct and mistaken zygosity will, at least to some extent, reflect potential prenatal differences associated with twins having a single versus multiple chorions.

While the optimal design of separated twins suffers from their infrequency, it is also a potential limitation of the TOMZ methodology that there are only relatively small numbers of twin pairs who misclassify themselves as to their biological zygosity, especially DZ twins if our estimates are accurate (although they are much more frequent than separated twins). Indeed, in our study, the number of incorrectly classified DZ twins was too small for further analysis. Thus, this method would not be appropriate for studies with small sample sizes, but in larger twin studies there should be adequate numbers of twins of mistaken identity and this method would be useful for identifying the contribution of shared environments to twin correlations. Furthermore, the reduced cost for genotyping and availability of DNA from saliva now make DNA zygosity determination feasible for even large twin samples.

Acknowledgments

Dr. Mary-Claire King and Dr. Beth Newman are acknowledged for their work establishing the Twin Registry at Kaiser Permanente and for collection of the epidemiologic and genetic data.

This study was supported by grants from the U.S. National Institutes of Health, the National Institute of Diabetes, Digestive and Kidney Diseases, Career Development Award, NIDDK Grant number K01 DK059944.

References

Acton, R. T., Go, R. C., & Roseman, J. M. (2004). Genetics and cardiovascular disease. *Ethnicity and Disease, 14,* S2–16.

Austin, M. A. (1993). The Kaiser-Permanente Women Twins Study data set. *Genetic Epidemiology, 10,* 519–522.

Austin, M. A., Friedlander, Y., Newman, B., Edwards, K., Mayer-Davis, E. J., & King, M. C. (1997). Genetic influences on changes in body mass index: A longitudinal analysis of women twins. *Obesity Research, 5,* 326–331.

Austin, M. A., King, M. C., Bawol, R. D., Hulley, S. B., & Friedman, G. D. (1987). Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *American Journal of Epidemiology, 125,* 308–318.

Block, G., Hartman, A. M., Dresser, C. M., Carroll, M. D., Gannon, J., & Gardner, L. (1986). A data-based approach to diet questionnaire design and testing. *American Journal of Epidemiology, 124,* 453–469.

Cohen, D. J., Dibble, E., Grawe, J. M., & Pollin, W. (1975). Reliably separating identical from fraternal twins. *Archives of General Psychiatry, 32,* 1371–1375.

Collaku, A., Rankinen, T., Rice, T., Leon, A. S., Rao, D. C., Skinner, J. S., Wilmore, J. H., & Bouchard, C. (2004). A genome-wide linkage scan for dietary energy and nutrient intakes: The Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study. *American Journal of Clinical Nutrition, 79,* 881–886.

de Castro, J. M. (1993a). Genetic influences on daily intake and meal patterns of humans. *Physiology and Behavior, 53,* 777–782.

de Castro, J. M. (1993b). Independence of genetic influences on body size, daily intake, and meal patterns of humans. *Physiology and Behavior, 54,* 633–639.

de Castro, J. M. (2002). Independence of heritable influences on the food intake of free-living humans. *Nutrition, 18,* 11–16.

Edwards, K. L., Austin, M. A., Newman, B., Mayer, E., Krauss, R. M., & Selby, J. V. (1994). Multivariate analysis of the insulin resistance syndrome in women. *Arteriosclerosis and Thrombosis, 14,* 1940–1945.

Edwards, K. L., Newman, B., Mayer, E., Selby, J. V., Krauss, R. M., & Austin, M. A. (1997). Heritability of factors of the insulin resistance syndrome in women twins. *Genetic Epidemiology, 14,* 241–253.

Falciglia, G. A., & Norton, P. A. (1994). Evidence for a genetic influence on preference for some foods. *Journal of the American Dietetic Association, 94,* 154–158.

Friedewald, W. T., Levy, R. I., & Fredrickson, D. S. (1972). Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clinical Chemistry, 18,* 499–502.

Friedlander, Y., Austin, M. A., Newman, B., Edwards, K., Mayer-Davis, E. I., & King, M. C. (1997). Heritability of longitudinal changes in coronary-heart-disease risk

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

factors in women twins. *American Journal of Human Genetics, 60,* 1502–1512.

Guarna, M. M., & Borowsky, R. L. (1993). Genetically controlled food preference: Biochemical mechanisms. *Proceedings of the National Academy Science USA, 90,* 5257–5261.

Heitmann, B. L., Harris, J. R., Lissner, L., & Pedersen, N. L. (1999). Genetic effects on weight change and food intake in Swedish adult twins. *American Journal of Clinical Nutrition, 69,* 597–602.

Heller, D. A., de Faire, U., Pedersen, N. L., Dahlen, G., & McClearn, G. E. (1993). Genetic and environmental influences on serum lipid levels in twins. *The New England Journal of Medicine, 328,* 1150–1156.

Hewitt, J. K. (1997). The genetics of obesity: What have genetic studies told us about the environment. *Behavior Genetics, 27,* 353–358.

Hrubec, Z., & Neel, J. V. (1978). The National Academy of Sciences — National Research Council Twin Registry: Ten years of operation. *Progress in Clinical and Biological Research, 24 Pt B,* 153–172.

Hrubec, Z., & Neel, J. V. (1981). Familial factors in early deaths: Twins followed 30 years to ages 51–61 in 1978. *Human Genetics, 59,* 39–46.

Hur, Y. M., Bouchard, T. J., Jr., & Eckert, E. (1998). Genetic and environmental influences on self-reported diet: A reared-apart twin study. *Physiology & Behavior, 64,* 629–636.

Jablon, S., Neel, J. V., Gershowitz, H., & Atkinson, G. F. (1967). The NAS-NRC twin panel: Methods of construction of the panel, zygosity diagnosis, and proposed use. *American Journal of Human Genetics, 19,* 133–161.

Jones, H. E. (1955). Perceived differences among twins. *Eugenics quarterly, 5,* 98–102.

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science, 8,* 263–266.

Katoh, S., Lehtovirta, M., Kaprio, J., Harjutsalo, V., Koskenvuo, M., Eriksson, J., Tajima, N., & Tuomilehto, J. (2005). Genetic and environmental effects on fasting and postchallenge plasma glucose and serum insulin values in Finnish twins. *Journal of Clinical Endocrinology and Metabolism, 90,* 2642–2647.

Keller, K. L., Pietrobelli, A., Must, S., & Faith, M. S. (2002). Genetics of eating and its relation to obesity. *Current Atherosclerosis Reports, 4,* 176–182.

King, M. C., Friedman, G. D., Lattanzio, D., Rodgers, G., Lewis, A. M., Dupuy, M. E., & Williams, H. (1980). Diagnosis of twin zygosity by self-assessment and by genetic analysis. *Acta Geneticae Medicae et Gemellologiae, 29,* 121–126.

Klump, K. L., McGue, M., & Iacono, W. G. (2000). Age differences in genetic and environmental influences on eating attitudes and behaviors in preadolescent and adolescent female twins. *Journal of Abnormal Psychology, 109,* 239–251.

Krondl, M., Coleman, P., Wade, J., & Milner, J. (1983). A twin study examining the genetic influence on food selection. *Human Nutrition. Applied Nutrition, 37 A,* 189–198.

Magnus, P., Berg, K., & Nance, W. E. (1983). Predicting zygosity in Norwegian twin pairs born 1915–1960. *Clinical Genetics, 24,* 103–112.

Mayer, E. J., Newman, B., Austin, M. A., Zhang, D., Quesenberry, C. P., Jr., Edwards, K., & Selby, J. V. (1996). Genetic and environmental influences on insulin levels and the insulin resistance syndrome: An analysis of women twins. *American Journal of Epidemiology, 143,* 323–332.

Mayer, E. J., Newman, B., Quesenberry, C. P., Jr., & Selby, J. V. (1993). Usual dietary fat intake and insulin concentrations in healthy women twins. *Diabetes Care, 16,* 1459–1469.

Nagele, U., Hagele, E. O., Sauer, G., Wiedemann, E., Lehmann, P., Wahlefeld, A. W., & Gruber, W. (1984). Reagent for the enzymatic determination of serum total triglycerides with improved lipolytic efficiency. *Journal of Clinical Chemistry and Clinical Biochemistry, 22,* 165–174.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2002). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Nichols, R. C., & Bilbro, W. C., Jr. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica, 16,* 265–275.

Reed, D. R., Bachmanov, A. A., Beauchamp, G. K., Tordoff, M. G., & Price, R. A. (1997). Heritable variation in food preferences and their contribution to obesity. *Behavior Genetics, 27,* 373–387.

Reed, T., Plassman, B. L., Tanner, C. M., Dick, D. M., Rinehart, S. A., & Nichols, W. C. (2005). Verification of self-report of zygosity determined via DNA testing in a subset of the NAS-NRC twin registry 40 years later. *Twin Research and Human Genetics, 8,* 362–367.

Rietveld, M. J., Der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research, 3,* 134–141.

Rissanen, A., Hakala, P., Lissner, L., Mattlar, C. E., Koskenvuo, M., & Ronnemaa, T. (2002). Acquired preference especially for dietary fat and obesity: A study of weight-discordant monozygotic twin pairs. *International Journal of Obesity and Related Metabolic Disorders, 26,* 973–977.

Rose, K. M., Newman, B., Mayer-Davis, E. J., & Selby, J. V. (1998). Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obesity Research, 6,* 383–392.

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press

Rozin, P., & Millman, L. (1987). Family environment, not heredity, accounts for family resemblances in food preferences and attitudes: A twin study. *Appetite, 8,* 125–134.

Rutter, M., Pickles, A., Murray, R., & Eaves, L. (2001). Testing hypotheses on specific environmental causal effects on behavior. *Psychological Bulletin, 127,* 291–324.

Scarr, S. (1968). Environmental biases in twin studies. In S. G. Vanderberg (Ed.), *Progress in human behavior genetics*. Baltimore: Johns Hopkins Press.

Segal, N. L. (2001). Twin studies of dietary behaviors: Why we eat when we do. *Twin Research, 4,* 478–480.

Selby, J. V., Austin, M. A., Sandholzer, C., Quesenberry, C. P., Jr., Zhang, D., Mayer, E., & Utermann, G. (1994). Environmental and behavioral influences on plasma lipoprotein(a) concentration in women twins. *Preventive Medicine, 23,* 345–353.

Selby, J. V., Newman, B., King, M. C., & Friedman, G. D. (1987). Environmental and behavioral determinants of fasting plasma glucose in women. A matched co-twin analysis. *American Journal of Epidemiology, 125,* 979–988.

Tholin, S., Rasmussen, F., Tynelius, P., & Karlsson, J. (2005). Genetic and environmental influences on eating behavior: The Swedish Young Male Twins Study. *American Journal of Clinical Nutrition, 81,* 564–569.

van den Bree, M. B., Eaves, L. J., & Dwyer, J. T. (1999). Genetic and environmental influences on eating patterns of twins aged ≥ =50 y. *American Journal of Clinical Nutrition, 70,* 456–465.

Warnick, G. R., Nguyen, T., & Albers, A. A. (1985). Comparison of improved precipitation methods for quantification of high-density lipoprotein cholesterol. *Clinical Chemistry, 31*, 217–222.

https://doi.org/10.1375/twin.9.4.540 Published online by Cambridge University Press