Prevention, Use of Health Services, and Genes: Implications of Genetics for Policy Formation

George L. Wehby
Benjamin W. Domingue
Jason D. Boardman

Abstract

We evaluate the hypothesis that genetic factors influence the use of health services and prevention behaviors in a national sample of adult twins in the United States. The analysis compares the correlation of these outcomes between identical twins, who share all their genes, to the correlation between nonidentical twins, who share, on average, only one-half of their genes. Because the environmental similarities of twins are assumed to be the same for identical and nonidentical twin pairs, researchers can partition the variance in behavioral outcomes that are due to genetic and environmental factors. Using established methods in this field, we find evidence of significant genetic influences on preferences toward prevention, overall prevention effort, routine checkups, and prescription drug use. Use of curative services does not appear to be influenced by genes. Our findings offer several implications for policymakers and researchers and suggest that genetics could be informative for health services and policy research.

INTRODUCTION

There is extensive debate about how much recent national health care reform will actually affect the use of health services. At the heart of this discussion are questions about how the Patient Protection and Affordable Care Act (PPACA) mandates that individuals obtain insurance coverage and insurers fully cover preventive services will change use of preventive care and medical services. Individuals may respond very differently to the PPACA provisions aimed at improving access to health care and expanding the use of preventive services, and some of these differences may have genetic origins.

This paper examines the degree to which differences in personal prevention effort and use of health services across a population can be explained by genetic variation in the population. The analysis compares the correlation of these outcomes between identical twins, who share all their genes, to the correlation between nonidentical twins, who share, on average, only one-half of their genes. Because the environmental similarities of twins are assumed to be the same for identical and nonidentical

Journal of Policy Analysis and Management, Vol. 34, No. 3, 519–536 (2015)

Published by Wiley Periodicals, Inc. View this article online at wileyonlinelibrary.com/journal/pam
DOI:10.1002/pam.21835

twin pairs, researchers can partition the variance in behavioral outcomes that are due to genetic and environmental factors. Differences in these correlations measure the extent to which heritable factors may determine population differences in prevention behaviors and use of health services, at least under the policies that existed before the PPACA.

This work also indicates the potential of genes to moderate policy effects on these outcomes; the greater the outcome variation that is explained by genetics, the more likely it is that genes can moderate policy effects. Genetic moderation of policy effects has already been observed in other contexts. For example, Fletcher (2012) reported that individuals carrying a specific genetic risk variant that may predispose to smoking did not change their smoking behaviors in response to tobacco taxation, whereas individuals carrying the protective variant associated with reduced smoking did change. A comparable result was shown by Boardman (2009) using a twin design similar to that of our study. Such interactions between policy and genetics could be relevant for other health outcomes and behaviors. For example, there is some concern that tying participation in employer-sponsored wellness programs such as weight-reduction initiatives to insurance premium rebates (expanded under the PPACA by up to 30 percent of total premiums) may penalize those with genetic predisposition for obesity or low physical activity (Downey, 2014).

Understanding the influence of genes on prevention behaviors and health services use has the potential to offer important directions for future research on health services and policy. If genes are found to have an effect on these outcomes, future work could begin to unravel the specific genes involved. Identifying the physiological pathways could shed light on specific genetic pathways that may moderate policy effects causing policies to exacerbate or reduce genetically influenced differences in health and health services use. Furthermore, it could ultimately help develop targeted interventions to improve efficiency in health care demand.

We employ a national sample of U.S. twins interviewed in 1995 and 1996 to evaluate this possibility empirically. In addition to use of preventive and curative health services, we examine measures of preferences toward prevention, other personal beliefs about health care, and overall prevention effort, all of which are relevant to a person's overall relationship with the health services sector and this relationship's effect on health. We find evidence that prevention preferences and effort, routine checkups, and prescription drug use are influenced by genes. Use of other types of health care services including hospitalizations, outpatient treatment for physical illness, mental health visits, and urgent care appear to be much less related to genes. These findings indicate that genetic factors played an important role in preventive care use at a time before PPACA (1995 and 1996), despite several policies and organizational initiatives that aimed to improve prevention and access to preventive care (including those related to managed care). These findings also raise the possibility that genetic differences may modify policy effects on use of preventive care and prescription drugs.

A very small body of work (summarized below) has considered a genetic basis for health services use, and these studies have focused on care seeking for specific health conditions and excluded more generic measures such as physician visits, hospitalizations, and emergency room visits. Furthermore, these studies do not include measures of preferences for prevention and health services use or measures of overall prevention effort. We expand the literature by evaluating a wide array of generic measures related to health services use and prevention preferences and effort.

LITERATURE REVIEW

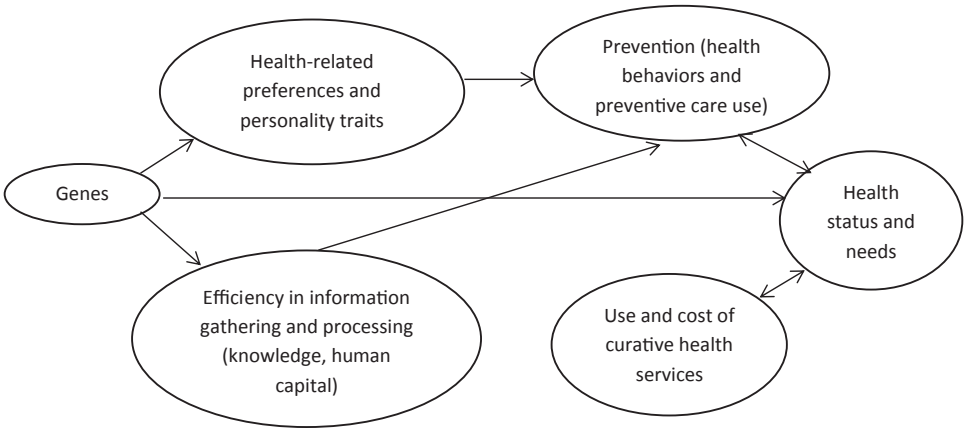
The few studies to examine the contribution of genetic variation to overall variation in health care use have compared identical and nonidentical twins, an approach commonly used in social and behavioral sciences to estimate how much of the variation in an outcome can be explained by genes. Under certain assumptions discussed below, the difference in correlation of an outcome between identical twins reared together, who share all of their genes, and that between nonidentical twins (also reared together), who share one-half their genes on average, can be easily shown to be half of the outcome variance explained by genetic variation.

The extant studies have produced important findings. True et al. (1997) studied four health conditions (high blood pressure, mental health, joint problems, and hearing problems) and treatment seeking related to these conditions in 1987, in a sample of 3,600 male twins from the Vietnam Era Twin study (VET). They show that a considerable fraction of the variation in the health conditions (24 to 52 percent) and an even higher fraction for the treatment-seeking behaviors (42 to 56 percent) can be explained by genetic factors. This difference suggests an important role for genetics in explaining variation in health services use in a population. Another study on treatment seeking for alcoholism using the same data source and telephone survey data collected in 1992 for about 3,350 male twin pairs estimated that 41 percent of the variance in this outcome is explained by genetic differences (True et al., 1996).

Less work has been done on the use of preventive care. To our knowledge, only one study has investigated the extent of genetic influence on a direct measure of use of preventive services. Treloar, McDonald, & Martin (1999) studied women's participation in cancer screening programs using data from an Australian twin sample and reported that a sizeable fraction (between 37 and 66 percent) of the variation in service seeking could have been influenced by genes (lowest genetic influence on breast self-examination and highest for Pap smear). Other studies have evaluated genetic influence on other prevention activities outside of the health care market. For example, exercise has been reported to be genetically influenced, with genes explaining up to 42 to 71 percent of the variation (Moor et al., 2011; Stubbe et al., 2006).

Some studies have evaluated generic health outcomes, including general health status and well-being, which are related to prevention and health services use. Romeis et al. (2000) studied self-rated health among 4,600 twin pairs from the 1987 VET study survey and reported that close to 40 percent of the variation could be attributed to genes. When the researchers controlled for the presence of major health problems, the explained variation declined slightly, to 32.5 percent, indicating that genetic factors influence self-reported health status beyond the direct effect of major health problems. Similarly, Romeis et al. (2005) reported that genes explained 17 to 33 percent of the variation in the SF-36—a commonly used measure of health status—in a sample of about 2,900 male twin pairs from the VET registry with survey data from 1992 and 1995. These results are not limited to twin studies. Using data on genetic variants across the genome from the Health and Retirement Study, Boardman, Domingue, and Daw (2015) reported that 22 percent of the variance of self-rated health was accounted for by genetic influences. Taken together, all these results are important because it is clear that genetic influences on health care use and prevention may be relatively larger than genetic influences on health, per se.

Little is known, however, about whether genes influence generic measures of health services use such as routine checkups, emergency room visits, hospitalizations, and prescription drug use, independently of health problems. Also, not much is known about how measures of preferences toward prevention and overall prevention effort are influenced by genes.



Notes: The figure shows a general conceptual model for various relationships through which genes can influence prevention effort, health care use and cost, and health status. Bidirectional arrows reflect two-way relationships where two domains affect each other.

Figure 1. Conceptual Framework for Potential Pathways Linking Genes to Prevention, Use of Health Services, and Health.

A CONCEPTUAL FRAMEWORK FOR GENETIC INFLUENCES

Theoretically, genes could contribute to differences in prevention effort and use of health services through multiple pathways. Of particular significance are genetic influences on disease risk and health needs, preferences, and information, which are all well-recognized determinants of health behaviors and health care demand. These potential pathways are highlighted in Figure 1 and discussed below. Our purpose here is not to develop a theoretical model to be tested in this paper but rather to illustrate several possible relationships between genes and the outcomes we study.

Genes may contribute to variation in use of health services simply through effects on disease, disability, and aging, which are major drivers of health services use and costs. Indeed, the distribution of total health care spending across the population has been heavily skewed, with more than half of expenditures incurred by the top 5 percent of users (Berk & Monheit, 2001), and expenditures significantly increasing with age, by up to four times more at ages 65 plus than for younger adults (Hartman et al., 2008). There is ample evidence that genes modify risks for disease (both prevalence and severity), especially for important chronic and complex conditions such as cardiovascular disease, diabetes, cancer, and mental health problems. Numerous twin studies have shown that much of the variation in several health conditions and measures can be explained by genetic factors, including as much as 40 percent for self-reported health status, 30 to 40 percent for major depression, 30 percent for hypertension, 50 to 80 percent for obesity and body mass index, over 50 percent for type-2 diabetes, 40 to 90 percent for asthma, and 30 percent for strokes (Agarwal, Williams, & Fisher, 2005; Carlsson et al., 2013; Kendler et al., 2006; Romeis et al., 2000; Thomsen et al., 2010). Genes may also affect the pace of aging and contribute to variation in health outcomes and activity limitations related to aging (Kenyon, 2010).

Another channel through which genes may influence prevention and use of health services is their effects on health-related preferences and personality traits. There is evidence for genetic influence on a range of general behavior preferences including risk taking, future discounting, self-efficacy, and trust in others (Benjamin, Ebstein,

& Belmaker, 2008). These preferences are conceptually relevant for personal investments in health (Grossman, 1972). For example, individuals who are more present oriented and discount future utility at a greater rate may be less likely to invest in their health through use of preventive services, since future returns from these investments are less valuable to them. As much as 20 percent of the variation in risk taking may be explained by genes (Cesarini et al., 2009). Risky behaviors such as smoking persistence and alcohol dependence have been shown to have an important genetic component (some estimates exceed 50 percent; Boardman, Blalock, & Pampel, 2010; Maes et al., 2004; Stacey, Clarke, & Schumann, 2009). Variation in how much individuals discount future outcomes may also be partly explained by genes, by as much as 30 to 50 percent (Anokhin et al., 2011). The tendency to trust others is also relevant since trust in health care providers has been shown to enhance the patient–provider relationship and continuity of care (Mainous et al., 2001). Close to 10 to 30 percent of the variation in trust may have genetic origins (Cesarini et al., 2008; Oskarsson et al., 2012), although estimates as high as 60 percent have also been reported (Sturgis et al., 2010).

Personal beliefs and knowledge about one's own health needs and the availability, quality/effectiveness, and cost of health services can have important effects on health behaviors and health care demand. These beliefs and knowledge are related to cognitive ability and human capital, which in turn are partly influenced by genes. For example, close to half of the variation in educational attainment (40 to 70 percent) may be explained by genes (Branigan, McCallum, & Freese, 2013). Therefore, genes could affect prevention and use of health services through a secondary pathway related to personal efficacy in information gathering and processing.

Of course, environmental effects such as wage rates, social capital, and access to quality health care play an important role in these pathways. The relative importance of genetic versus environmental effects could also vary between different outcomes. We aim at decomposing these two sources of variation for a broad range of outcomes related to prevention and use of health services. Ultimately, the extent to which genes influence prevention and use of health services is an empirical question.

DATA

Sample

We use data on a national sample of twins from the National Survey of Midlife Development in the United States (MIDUS I) conducted in 1995 and 1996. MIDUS provides a broad set of measures on prevention and health services use that are of interest for our study. The survey collected health and socioeconomic data from adults aged 25 to 74. The total MIDUS I sample included 957 twin pairs. Of those, 907 pairs (1,814 individuals) had data on whether the twins were identical and consistent data on self-reported birth year/age between twins. Of the 907 pairs, 349 pairs were identical, 321 same-sex nonidentical, and 237 different-sex nonidentical twins. The sample for our analysis ranges from 742 to 795 twin pairs with both twins having complete data, depending on the outcome evaluated.[1]

[1] The MIDUS I sample was followed up between 2004 and 2006 under MIDUS II to collect updated data. We use only MIDUS I data in this paper because the MIDUS II sample is smaller and less powerful than MIDUS I. The follow-up included 742 pairs; of those 604 had data on identical/nonidentical twin status and consistent data on self-reported birth year/age between twins. The MIDUS II sample with complete data on the outcomes we analyze for both twins ranges from 362 to 410 pairs.

Measures

MIDUS provides data on multiple measures of prevention effort, health services use, and related individual characteristics such as preferences toward prevention and health care providers. The outcomes we evaluate can be conceptually grouped into six general areas: prevention preferences, overall prevention effort, other health-related personality traits (including self-efficacy, personal health knowledge, and belief in healthcare effectiveness), preventive care use, use of curative health services, and general health status. We describe these below and include detailed definitions and summary statistics in Table 1.

Prevention Preferences

Prevention preferences are represented by responses to questions about one's beliefs in the effectiveness of prevention. One question requires the respondent to express agreement or disagreement, on a seven-category scale from "strongly disagree" to "strongly agree," with the following statement: "Keeping healthy depends on things that I can do." Two other questions use a similar scale to ask about effects of personal prevention on reducing heart attack and cancer risks. For these questions and others measured on a similar scale, we generally categorize respondents into two groups: those who strongly agreed versus those who less than strongly agreed or disagreed with the statement. This simplification is necessary since the models are not appropriate for ordinal data.[2]

In addition to evaluating each of these three measures separately, we generate an index for prevention preferences by aggregating them using principal component analysis (PCA).[3] PCA is more reliable for aggregating multiple variables than arbitrarily assigning equal or different weights to the various variables being aggregated. Furthermore, using PCA is intuitive, as the first component captures the single most important source of variation shared between these variables. We estimate the principal components with the original ordinal scales in order to capture as much of the variation in the variables as possible when generating the aggregate index.[4] The first principal component explains 72 percent of the variation in the aggregated variables. By design, the index has a mean around zero and includes both positive and negative values; an increase indicates stronger preference toward prevention.

Overall Prevention Effort

We measure overall prevention effort by two questions. The first asks the respondent to agree or disagree, again on the seven-category "strongly agree" to "strongly disagree" scale, with the statement "I work hard at trying to stay healthy." We again dichotomize this measure into whether or not the respondent indicated "strongly agree." The second question asks the respondent to rank, on a 0 to 10 scale, the thought and effort he or she puts into health. We analyze this measure as a continuous outcome.

[2] Considering these ordinal measures as continuous is not ideal given their skewed distributions and their lack of interval data properties.
[3] The index is generated by multiplying the weights from the first principal component by the measures being aggregated and summing the product terms.
[4] The correlations between the latent variables underlying the observed measures are estimated using maximum likelihood (Kolenikov & Angeles, 2004).

Table 1. Study measures and summary statistics.

Measure	MIDUS question/variable description	Percent or mean (SD)	No. of twin pairs
Prevention preferences			
Responsible*	Keeping healthy depends on things that I can do	60.8	784
Heart risk prevention*	There are certain things I can do for myself to reduce the risk of a heart attack	76.5	785
Cancer risk prevention*	There are certain things I can do for myself to reduce the risk of getting cancer	40.6	778
Prevention preference index	Scores from first principal component from a PCA of above three measures	−0.10 (1.14)	775
Overall prevention effort			
Works hard*	I work hard at trying to stay healthy	23.1	769
Thought/effort	Using a 0 to 10 scale where 0 means "no thought or effort" and 10 means "very much thought and effort," how much thought and effort do you put into your health these days?	7.3 (2.0)	790
Other health-related personality traits			
Health-related self-efficacy	Using a 0 to 10 scale where 0 means "no control at all" and 10 means "very much control," how would you rate the amount of control you have over your health these days?	7.9 (1.8)	789
Knowledge about own health†	I am often aware of various things happening within my body	23.5	784
Belief in health care effectiveness*	When I am sick, getting better is in the doctor's hands	28.2	775
Preventive care	Any visit to doctor/hospital/clinic for a routine physical checkup or gynecological exam in past 12 months	69.1	757
Use of curative services			
Hospitalization	At least one hospitalization in past 12 months	10.6	761
Outpatient treatment for physical health	At least one visit to doctor/hospital/clinic/orthodontist/ophthalmologist for scheduled treatment/surgery in past 12 months	28.9	742
Mental health care	At least one visit to psychiatrist/general practitioner/other MD for emotional/mental/personal problem in past 12 months	17.9	767
Any urgent care	At least one visit to doctor/ER/clinic for urgent treatment in past 12 months	30.3	753
Two or more urgent visits	At least two visits to doctor/ER/clinic for urgent treatment in past 12 months	9.9	753

Table 1. Continued.

Measure	MIDUS question/variable description	Percent or mean (SD)	No. of twin pairs
Prescription drug use	Used any prescription medicine in past 30 days	42.8	791
Health status			
Self-rated health	Using a scale from 0 to 10 where 0 means "the worst possible health" and 10 means "the best possible health," how would you rate your health these days? (10 *minus* response on this scale)	2.3 (1.6)	792
Number of chronic conditions	Sum of chronic health conditions (based on a list of questions about 29 conditions)	2.2 (2.3)	795
Health status index	Scores from first principal component from a PCA of above two measures under health status	0.00 (1.19)	791

Notes: The table reports the study variables, the exact question from MIDUS or description of variable construction, descriptive statistics (percent for dichotomous variables and means and SD for continuous variables), and number of twin pairs with complete data on each variable. The summary statistics are based on the total number of twins (individuals) in the total analytical sample of 907 pairs with complete data.

*Question is on a seven-category scale from "strongly disagree" to "strongly agree." The dichotomous variable is constructed to indicate those who strongly agreed with the statement versus other responders.

†Question is on a four-category scale from "not at all true" to "extremely true." The dichotomous variable is constructed to indicate those who responded "extremely true" versus other responders.

Other Health-Related Personality Traits

We evaluate three other personality traits that are conceptually related to prevention and use of health services. Health-related self-efficacy is captured by a measure of one's perception of personal control over health on a 0 to 10 scale. Knowledge about personal health needs is measured by a question on one's awareness of body changes; responses use a four-category scale and are dichotomized to capture those who are highly aware of things happening in their body. Personal beliefs about health care effectiveness are captured by the extent of the respondent's agreement with a statement that getting better when ill is in the doctor's hands.

Use of Preventive and Curative Services

We employ multiple measures of health services use. Use of preventive care is measured by whether a respondent visited health care providers for a routine physical checkup or gynecological exam in the past 12 months, the only such measure of preventive services in MIDUS I. In contrast, the data set provides several measures of health services use for curative care during the past 12 months that we evaluate, including visiting health providers for scheduled treatments (separately for physical and mental health care), an indicator for any hospitalization, and use of urgent/emergency care, which we evaluate separately as (1) any use and (2) having two or more urgent visits. Questions regarding scheduled treatments and urgent care use do not distinguish between provider types (e.g., dentist vs. doctor) or care settings (e.g., physician's office vs. emergency room). Nonetheless, these measures are useful for an exploratory look into potential genetic effects on their variation.

Health Status

We also evaluate genetic influences on generic measures of health status that are commonly studied in conjunction with health services use and are known to have a strong genetic component. We measure health status by two variables, own rating of health on a 0 to10 scale, and the total number of chronic conditions the individual has. Self-reported health is one of the most commonly used measures of health status and well-being in health services and policy research and is a strong predictor of multiple aspects of physical and cognitive functioning, limitations in daily activities, overall well-being, and mortality risk (Idler & Benyamini, 1997). As we note above, Romeis et al. (2000) have reported that genes explain a sizeable fraction of the variation in self-rated health. Similarly, the number of chronic conditions is an important indicator of health and well-being and a predictor of quality of life and functioning. We analyze these two health measures separately and then aggregate them into a health status index using the first component from a PCA (the first component explains 70 percent of the variation in these two variables). An increase in the index represents worsening health.

EMPIRICAL MODELS

The main goal of our empirical analysis is to assess how much of the variation in the measures of prevention and health services use described above can be explained by genetic variation between individuals. The standard approach involves comparing, for each outcome, the correlation between identical twins to that between nonidentical twins.[5] This approach decomposes the variance in an outcome into genetic influences, environmental factors that twins share with one another, and unique environmental factors that twins do not share with one another.[6] In its simplest form, the model requires the following assumptions: identical twins do not share a more similar home environment than nonidentical twins;[7] parents are not genetically related (mating is random); genetic effects can be modeled as additive (each copy of the genetic risk variant adds linearly to the outcome); and there are no gene-by-environment interactions. Some assumptions can be relaxed under certain extensions of the basic model.[8] We focus on estimating the basic model under the above assumptions in order to minimize the number of tests and the

[5] The twin model has been employed in several studies to estimate the influence of genes on multiple economic and social traits related to health, such as income and economic preferences, including risk aversion, time discounting, and trust (see detailed references in Benjamin et al., 2012).

[6] This model, commonly referred to as the ACE model, decomposes the outcome variance into additive genetic influences (A), environmental factors shared by twins (C), and unique environmental factors (E). The fraction of the variance explained by genes is a measure of how heritable the outcome is.

[7] We estimated a logistic regression to evaluate if the following variables predict whether twins are identical or nonidentical: race (white vs. nonwhite), indicators for financial status of the family when child was growing up relative to other families, number of times the family moved to a new neighborhood during childhood, and indicators for mother's educational level. These variables were jointly insignificant ($P = 0.53$). Detailed regression results are in Appendix Table A1. This analysis does not represent a full test of the equal environment assumption, but indicates that there are no differences in these specific indicators between the identical and nonidentical twin groups in the study sample. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[8] Issues related to nonrandom mating (Domingue et al., 2014) and other aspects of traditional twin-based models have met with renewed criticism in recent years (Burt & Simons, 2014; Charney & English, 2013). However, other scholars strongly argue for the merit of the basic twin model and robustness of its results despite potential limitations (Barnes et al., 2014).

magnitude of type-1 error.[9] We recognize that the different models may yield different estimates;[10] however, the basic model is appropriate for the main goal of this paper, which is to evaluate the extent to which genes may explain a broad range of measures related to prevention and health services use and implications for health policy.

As a secondary analysis, we explore the extent to which prevention preferences represented in the PCA-generated prevention index are correlated with the health status index (combining self-reported health status and number of chronic conditions) because of shared genetic effects (i.e., because of genes that are related to both outcomes). This analysis elucidates the behavioral mechanisms that are genetically influenced and that link preferences and health. Relying on assumptions similar to those described above, the approach examines the covariance across twins and across the two outcomes (prevention preferences and health status). Within twin pairs, the cross-twin cross-outcome analysis evaluates the extent to which one could predict the health status of the first twin as a function of prevention preferences of the second twin, and how this varies between identical and nonidentical twins (a stronger relationship between identical twins supports the hypothesis that the two outcomes may be linked by shared genetic influences). We provide the technical details of this model in Appendix B.[11] We do not evaluate shared genetic influences between health status and the other outcomes studied, such as prevention effort or use of preventive and curative health services, as these other outcomes can be either causes or effects of health status.[12]

RESULTS

Table 2 reports results from the twin analysis of the difference in correlations between identical and nonidentical twins, which decomposes the outcome variance into three fractions explained, respectively, by genes, environment shared between twins, and environment unique to each twin. Of the six categories of outcomes that we examine, four show evidence for genetic influences: prevention preferences, overall prevention effort, preventive care use, and health status. In contrast, we find no evidence for genetic influences on the other health-related personality traits including self-efficacy, personal health knowledge, and belief in health care effectiveness, and on use of curative care for physical or mental health problems, except

[9] We estimate a threshold model for the binary outcomes and a linear model for the continuous outcomes using the OpenMx suite of behavioral genetic models that is available through the computational program R (Boker et al., 2010). Mx is a statistical software package for structural equation modeling that is akin to commercial packages such as LISREL, EQS, and Amos (Neale et al., 2006). Because of the flexibility of the modeling, Mx has been the software most widely used for analyzing pairs of siblings and twins to provide estimates of outcome variance due to genetic and environmental factors. More recently, the Mx structure has been modified to work within the R programming environment through the package OpenMx (Boker et al., 2011).

[10] For example, one could test models that assume no effects from environmental factors shared between twins or models that assume dominant or interactive genetic effects in addition to the additive effects.

[11] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[12] Individuals may modify their use of preventive and curative care because of acute illnesses or chronic health problems. Similarly, personal health knowledge, control over health, and belief in health care effectiveness may be influenced by health status. These relationships are depicted using bidirectional arrows in Figure 1. In such cases, evaluating shared genetic influence in this framework would not be meaningful unless the reverse effects from health were appropriately modeled. In contrast, the measured preferences toward prevention are likely more exogenous and less affected by health status than these other measures.

Table 2. Decomposition of outcome variance into genetic and environmental components.

Outcome	Proportion of variance due to			*P*-value for genetic effects
	Genetic effects	Environmental effects shared by twins	Environmental effects unique to each twin	
Prevention preferences				
Responsible	0.286	0.0	0.714	0.157
Heart risk prevention	0.360	0.0	0.641	0.037
Cancer risk prevention	0.400	0.0	0.600	0.036
Prevention preference index	0.324	0.0	0.676	0.014
Overall prevention effort				
Works hard	0.540	0.0	0.460	0.002
Thought/effort	0.297	0.036	0.666	0.028
Other health-related personality traits				
Health-related self-efficacy	0.053	0.134	0.813	0.729
Knowledge about own health	0.137	0.061	0.803	0.610
Belief in health care effectiveness	0.0	0.269	0.731	0.999
Preventive care	0.409	0.0	0.591	0.039
Use of curative services				
Hospitalization	0.0	0.128	0.873	0.999
Outpatient treatment for physical health	0.170	0.020	0.810	0.517
Mental health care	0.269	0.0	0.732	0.124
Any urgent care	0.047	0.017	0.936	0.862
Two or more urgent visits	0.0	0.0	1.0	0.999
Prescription drug use	0.439	0.0	0.561	0.038
Health status				
Self-rated health	0.290	0.047	0.663	0.038
Number of chronic conditions	0.495	0.0	0.505	<0.0001
Health status index	0.484	0.0	0.516	<0.0001

Notes: The *P*-values for the significance of the genetic effects (i.e., proportion of variance explained by genes) are based on likelihood ratio test comparing the model to one without genetic effects.

prescription drug use. Environmental factors shared between twins do not appear to be relevant in explaining the variance of several outcomes, including prevention preferences, preventive visits, and prescription use. This finding is consistent with those of several twin studies on economic preferences (Benjamin et al., 2012). We focus below on summarizing the main genetic influences that we observe.

Prevention Preferences

We observe moderate to strong genetic influence on the three measures of prevention preferences, with genes explaining between 29 and 40 percent of the outcome variances. These genetic effects are statistically significant, except for the dichotomous indicator (*Responsible*) for strongly thinking that health depends on personal effort. Genetic differences explain over one-third of the variance in expressing strong beliefs that personal effort reduces heart attacks and cancer. Similarly,

genes account for about one-third of the variance of the PCA-generated index combining the three measures of prevention preferences.

Overall Prevention Effort

A strong genetic influence is observed for the two measures of overall prevention effort, consistent with the results described above for prevention preferences. Genes explain over one-half of the variance in asserting strongly that one works hard to stay healthy. Similarly, genetic variation accounts for about one-third of the variance in how much thought and effort the respondent puts into health. Together, the results for prevention preferences and effort suggest an important genetic influence on personal decisions about prevention.

Preventive Care

The strong genetic effects on prevention preferences and effort are also observed for actual use of preventive care. Genetic factors explain close to 40 percent of the variance in whether the person had a routine checkup in the past 12 months. The consistency in results across the three conceptually related outcome categories of prevention preferences, prevention effort, and preventive care use supports the validity of the analytical model.

Curative Services

The evidence of genetic relevance is much weaker for use of curative services than for the prevention measures. Of all the measures for curative care, only use of prescription drugs (in the past 30 days) has a strong genetic component; over 40 percent of this outcome variance is explained by genetic differences. In contrast, there is no evidence that genes are important for explaining other measures of curative services, including outpatient visits for physical health problems, hospitalizations, and urgent care. The only noteworthy exception is having a mental health visit; close to one-quarter of this outcome variance could be accounted for by genetic differences, a finding that is consistent with the evidence of genetic effects on mental health problems in the medical literature. However, the estimated genetic component for this outcome is statistically insignificant in our analysis, perhaps because of statistical power issues.

Health Status

In accord with the literature, we find important genetic influences on both self-reported health and number of chronic conditions. As much as one-third of the variance of self-reported health status is explained by genes. Similarly, close to half of the variance in the number of chronic conditions is accounted for by genetic differences. When we combine both measures into a health status index using PCA, about half of the variance of this index is explained by genes.

Shared Genetic Effects between Prevention Preferences and Health Status

We also evaluate the possibility that the genes influencing prevention preferences may be the same as those that influence health status. That is, some of the association between these two outcomes may be due to common genetic influences. We use the aggregated measures of prevention preferences and health status in this analysis. This evaluation is further supported by the evidence for genetic influence on prevention preferences and health status described above, as well as the correlation (albeit relatively small) between the prevention preference and health status indices

($r = -0.19$).[13] The findings are somewhat mixed on the existence of genetic influences shared between these two outcomes (details of this analysis are in Appendix B[14]). On the one hand, there is a stronger cross-twin cross-outcome correlation for identical than nonidentical twins (−0.09 vs. −0.05), suggesting some shared genetic effects between these two outcomes. On the other hand, the estimate of the shared genetic component is statistically insignificant. Therefore, this exploratory analysis provides only very weak evidence of the possibility of a shared genetic mechanism linking prevention preferences and health status.

CONCLUSIONS

Our findings suggest important genetic effects on prevention preferences, overall prevention effort, use of routine checkups, and prescription drug use. There is also an important genetic influence on self-reported health status and number of chronic conditions, both of which are commonly used health measures linked to prevention and use of health services. In contrast, we find no evidence of genetic influences on visits to medical professionals to treat physical health problems, hospitalization, and urgent care use. This finding is more surprising for outpatient visits than for urgent care and to some extent hospitalizations. Both urgent care and hospitalization are more likely to be influenced by events that are partly outside the individual's control, and less likely to be related to genes than outpatient care use, which involves more personal discretion. It is possible that genes play a greater role in the onset of chronic conditions and their "permanent" health care consequences such as use of prescription drugs than in the actual prognoses of chronic conditions and in the incidence of acute health problems, which are more strongly related to seeking outpatient or inpatient treatment and may be more sensitive to the environment. However, there is some indication that genes could still be relevant to seeking treatment for mental health problems. It is also important to note that the shared environment is less prominent than unique environmental factors for most of the outcomes evaluated, with a few exceptions (self-efficacy, belief in health care effectiveness, hospitalizations), a finding consistent with other studies of related outcomes such as economic preferences (Benjamin et al., 2012).

The study findings have important implications for policymaking.[15] One interpretation of the results is that the health care policy environment in the mid-1990s left substantial room for genetic factors to contribute to differences in prevention effort and preventive care use, despite several health care reform initiatives to improve access to preventive care. This era had significant expansions in preventive care use, with the extensive penetration of managed care and health maintenance organizations into health care delivery systems. This finding could suggest that reducing differences in use of preventive services may require additional interventions beyond slight modifications in contextual factors (e.g., very small changes in the supply of primary health care providers).

It remains unclear whether and how the observed genetic influence on use of preventive services before the PPACA would change with the PPACA provisions that expand insurance coverage and eliminate out-of-pocket cost for preventive care.

[13] The negative correlation is due to the opposite directions of the indices given the loading coefficients on the first principal component: an increase in the prevention index reflects stronger preferences for prevention; an increase in the health status index represents worsening health.

[14] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[15] Even if a large percentage of the variation in an outcome is explained by genetic differences, such a finding does not imply that this outcome variation is not amenable to policy interventions.

The fact that the majority of the population had such coverage at relatively low out-of-pocket cost before the PPACA suggests that a fairly small proportion of the population would be directly and dramatically affected by these provisions. Also, we find a strong genetic influence on prevention preferences and personal prevention effort outside the health care market, which are less targeted by the PPACA than use of preventive services. This genetic influence on choice of prevention independent of affordability raises the possibility that genes could modify the effects of the PPACA and other health policies on use of preventive services. Whether the PPACA will mute or exacerbate genetic differences in preventive care use is an open question that could be explored when post-PPACA data become available. For example, our approach could be employed in the future to evaluate how expansions of state Medicaid programs under the PPACA modify genetic influences on prevention and health services use and related outcomes.

Our results imply that the use of curative services is much more related to environmental factors than is the use of preventive care. Fisher and Wennberg (2003) show strong relationships between environmental factors (such as provider supply) and inpatient care use or outpatient visits to specialists, but not use of preventive services. Our results suggest that perhaps most of the variation in U.S. health care expenditures (mostly allocated to inpatient and outpatient treatments) is potentially driven by environmental factors and not by genetic differences. Reining in the costs of the U.S. health care system requires understanding the root causes of the population variation in health services use. Individual-level heterogeneity in unobservables is a major obstacle to causal inference on the determinants of health care demand. Our findings indicate that genetic heterogeneity is of relatively low concern in estimating demand for most curative inpatient and outpatient services, with the potential exception of mental health services. In contrast, our results highlight the need to consider genetic heterogeneity when studying prevention, use of prescription drugs, health status, and chronic conditions.[16]

Our results are generally consistent with those from previous studies reviewed above that investigated genetic influences on related outcomes, even though the outcome measures used in these studies were not the same as ours. Our estimates for the extent of genetic influences on prevention preferences (including for preventing cancer) and effort are consistent with those that Treloar, McDonald, and Martin (1999) reported for participation in cancer screening by an Australian sample of women. Similarly, our estimates for prescription drug use and mental health visits are in the range of those that True et al. (1997) found for seeking treatment for selective chronic conditions (such as high blood pressure and mental health problems) and for seeking treatment for alcoholism (1996), with both studies involving male twins from the VET registry. Also, our estimates of genetic influence on the health status measures are close to those reported by Romeis et al. (2000, 2005), again for the VET sample.[17]

[16] Our analysis of the possibility that the same genes may contribute to both prevention preferences and health status suggests that genetic factors may be a source of confounding in observational studies of the association between prevention and health that do not account for individual-level unobservables such as genetic factors. However, this analysis deserves replication in future research using a larger twin sample with more power to detect genetic effects shared between these outcomes.

[17] Finding results that are generally consistent with those from the VET samples limited to male twins does necessarily mean that there is no heterogeneity in genetic influences by gender. Unfortunately, our sample is not large enough to evaluate heterogeneity by gender in any meaningful way. We would have only 163 identical male twin pairs and 123 nonidentical male twin pairs, and 186 identical female twin pairs and 198 nonidentical female twin pairs. Clearly, these sample sizes are very small for a useful comparison by gender. Exploring heterogeneity by gender could be considered in future research with sufficient sample size.

Even though the MIDUS provides a national sample of twins, we do not know whether the findings can be generalized to subpopulations of particular interest to policymakers such as the uninsured, individuals of low socioeconomic status, and the elderly. The majority of people in the sample had health insurance coverage (over 91 percent). Also, even though the twin sample covered a wide age range (25 to 75 years), the majority (91.4 percent) were younger than 65. Use of curative health services is substantially higher among older adults (Hartman et al., 2008). It is possible that genetic influences on use of health services intensify with age owing to greater onset of chronic conditions, which we find to be strongly related to genes. Here it is important to note that both members of a twin pair must still be alive in order to examine age-related changes in the relative contributions of genes to prevention and health services use. If some of the outcomes that we study are associated with mortality, then we are not able to properly evaluate the age gradient because of these selective processes. Moreover, we cannot well evaluate heterogeneity in genetic influences by age in this data set because of the small size of age-stratified samples and the limited power to test subgroup differences.[18] Another limitation on the generalizability of our results derives from organizational changes in the health care system over time.[19] Finally, our data on use of health services combine multiple provider specialties and health care delivery settings, have a relatively short reporting period (the past 12 months), and potentially suffer from some error or bias. These limitations could have artificially lowered the genetic effects on use of curative services.

Our work echoes the conclusions from previous studies that genetics can bring useful information and tools into health services and policy research. Genetic models can be used to explain differences in prevention effort and demand for services and studies of genetic mechanisms may identify relevant behavioral pathways that can subsequently be targeted to improve health and well-being. For example, if the genetic effects on prevention preferences operate mainly through risk tolerance and fear of monetary loss, insurers or employers could require a minimum use of preventive services (such as at least one annual visit for routine checkups), with a financial penalty for failure to comply in the form of higher coinsurance for curative services. In the somewhat near future, it is possible that individuals could be offered genetic screening by their employers or insurance programs (e.g., in exchange for insurance rebates). Such screening could not only assess disease risks but also predict patterns of health care use in order to provide individualized interventions and counseling to optimize use of preventive services.

For researchers, investigating genetic mechanisms for prevention and health services use has promise, but also multiple challenges such as those previously described for other disciplines such as economics (Benjamin et al., 2012). Since genetic variants can be relatively easily measured, they can serve as indicators for latent

[18] We generally observe close differences in correlations between identical and nonidentical twins for most of the outcomes with a significant genetic component (from Table 2) when comparing the full sample (age range of 25 to 74 years) with two age subgroups: 25 to 64 and 25 to 44 years (Appendix Table C1). These results suggest that genetic influences are relatively stable across age for the majority of people in the sample, who are younger than 65 years, but do not provide much information on the estimates for individuals 65 years and older. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[19] There have been multiple organizational changes in the health care system since the mid-1990s, a period characterized by large HMO penetration, low insurance premium growth, gatekeeping, and extensive use of controls on health care use introduced by managed care plans. Since then, some of these controls have changed for the majority of privately covered people, while other use-management techniques, such as care management programs for chronic conditions, have been introduced. The health care workforce is also changing, with a decline in the supply of primary care physicians.

and difficult-to-measure variables such as preferences and personality traits and be included as control variables to improve explanatory power. There are, however, major challenges to be addressed before such work could yield useful information. The complexity of human behavior suggests that multiple genes could be at play, each having a very small effect. Identifying these effects requires very large samples and well-measured outcomes.[20] Furthermore, understanding how these genes function in shaping human behavior is complex. It will be long before we can fully realize the value of unravelling the molecular genetics of outcomes such as prevention behaviors and health services use. In the short term, however, discovering genes that contribute to these outcomes could enable us to evaluate potential differences in health policy consequences for groups bearing different genetic risks.

GEORGE L. WEHBY is a Research Associate at the NBER and an Associate Professor in the Departments of Health Management and Policy, Economics, Preventive & Community Dentistry, and the Public Policy Center at the University of Iowa, 145 N. Riverside Drive, 100 College of Public Health Building N248, Iowa City, IA 52242 (e-mail: george-wehby@uiowa.edu).

BENJAMIN W. DOMINGUE is a Research Associate at the Institute of Behavioral Science at the University of Colorado, 1440 15th Street, Boulder, CO 80302 (e-mail: Benjamin.domingue@colorado.edu).

JASON D. BOARDMAN is a Professor in the Department of Sociology & the Institute of Behavioral Science at the University of Colorado, 1440 15th Street, Boulder, CO 80302 (e-mail: boardman@colorado.edu).

REFERENCES

Agarwal, A., Williams, G. H., & Fisher, N. D. (2005). Genetics of human hypertension. Trends in Endocrinology and Metabolism, 16, 127–133.

Anokhin, A. P., Golosheykin, S., Grant, J. D., & Heath, A. C. (2011). Heritability of delay discounting in adolescence: A longitudinal twin study. Behavior Genetics, 41, 175–183.

Barnes, J. C., Wright, J. P., Boutwell, B. B., Schwartz, J. A., Connolly, E. J., Nedelec, J. L., & Beaver, K. M. (2014). Demonstrating the validity of twin research in criminology. Criminology, 52, 588–626.

Benjamin, D. J., Cesarini, D., Chabriset, C. F., Glaeser, E. L., Laibson, D. I., Guðnason, V., & Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics. Annual Review of Economics, 4, 627–662.

Benjamin, J., Ebstein, R., & Belmaker, R. (Eds.). (2008). Molecular genetics and the human personality. American Psychiatric Publishing, Washington, D.C., London, England.

Berk, M. L., & Monheit, A. C. (2001). The concentration of health care expenditures, revisited. Health Affairs (Millwood), 20, 9–18.

Boardman, J. D. (2009). State-level moderation of genetic tendencies to smoke. American Journal of Public Health, 99, 480–486.

Boardman, J. D., Blalock, C. L., & Pampel, F. C. (2010). Trends in the genetic influences on smoking. Journal of Health and Social Behavior, 51, 108–123.

[20] A recent study of educational attainment combining data on over one million genetic variants and 100,000 individuals illustrates this challenge; only three variants explaining a very small fraction (0.02 percent) of the variation in educational attainment were statistically significant after researchers corrected for multiple testing (Rietveld et al., 2013).

Boardman, J. D., Domingue, B. W., & Daw, J. (2015). What can genes tell us about the relationship between education and health? Social Science & Medicine, 127, 171–80.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., & Fox, J. (2010). OpenMx: Multipurpose software for statistical modeling (Version R package version 1.0.4). Virginia.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., & Fox, J. (2011). OpenMx: An open source extended structural equation modeling framework. Psychometrika, 76, 306–317.

Branigan, A. R., McCallum, K. J., & Freese, J. (2013). Variation in the heritability of educational attainment: An international meta-analysis. Social Forces, 92, 109–140.

Burt, C. H., & Simons, R. (2014). Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. Criminology, 52, 223–262.

Carlsson, S., Ahlbom, A., Lichtenstein, P., & Andersson, T. (2013). Shared genetic influence of BMI, physical activity and type 2 diabetes: A twin study. Diabetologia, 56, 1031–1035.

Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P., & Wallace, B. (2008). Heritability of cooperative behavior in the trust game. Proceedings of the National Academy of Sciences USA, 105, 3721–3726.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Genetic variation in preferences for giving and risk taking. Quarterly Journal of Economics, 124, 809–842.

Charney, E., & English, W. (2013). Genopolitics and the science of genetics. American Political Science Review, 107, 382–395.

Domingue, B. W., Fletcher, J., Conley, D., & Boardman, J. D. (2014). Genetic and educational assortative mating among US adults. Proceedings of the National Academy of Sciences USA, 111, 7996–8000.

Downey, M. (2014). Response to Dr. Cawley. Journal of Policy Analysis and Management, 33, 832–834.

Fisher, E. S., & Wennberg, J. E. (2003). Health care quality, geographic variations, and the challenge of supply-sensitive care. Perspectives in Biology and Medicine, 46, 69–79.

Fletcher, J. M. (2012). Why have tobacco control policies stalled? Using genetic moderation to examine policy impacts. PLoS One, 7, e50576.

Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political Economy, 80, 223–255.

Hartman, M., Catlin, A., Lassman, D., Cylus, J., & Heffler, S. (2008). U.S. health spending by age, selected years through 2004. Health Affairs (Millwood), 27, w1–w12.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. Journal of Health and Social Behavior, 38, 21–37.

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). A Swedish national twin study of lifetime major depression. American Journal of Psychiatry, 163, 109–114.

Kenyon, C. J. (2010). Nature, Mar 25;464(7288), 504–512. doi: 10.1038/nature08980. Review. Erratum in: Nature, 2010 Sep 30;467(7315), 622.

Kolenikov, S., & Angeles, G. (2004). The use of discrete data in principal component analysis with applications to socio-economic indices. CPC/MEASURE Working Paper.

Maes, H. H., Sullivan, P. F., Bulik, C. M., Neale, M. C., Prescott, C. A., Eaves, L. J., & Kendler, K. S. (2004). A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. Psychological Medicine, 34, 1251–1261.

Mainous, A. G. III, Baker, R., Love, M. M., Gray, D. P., & Gill, J. M. (2001). Continuity of care and trust in one's physician: Evidence from primary care in the United States and the United Kingdom. Family Medicine, 33, 22–27.

Moor, M. M., Willemsen, G., Rebollo-Mesa, I., Stubbe, J., Geus, E. C., & Boomsma, D. (2011). Exercise participation in adolescents and their parents: Evidence for genetic and generation specific environmental effects. Behavior Genetics, 41, 211–222.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. (2006). Mx: Statistical modeling (7th ed.). Richmond: Department of Psychiatry, Virginia Commonwealth University.

Oskarsson, S., Dawes, C., Johannesson, M., & Magnusson, P. K. (2012). The genetic origins of the relationship between psychological traits and social trust. Twin Research and Human Genetics, 15, 21–33.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., & Koellinger, P. D. (2013). GWAS of 126559 individuals identifies genetic variants associated with educational attainment. Science, 340, 1467–1471.

Romeis, J., Scherrer, J., Xian, H., Eisen, S., Bucholz, K., Heath, A., & True, W. R. (2000). Heritability of self-reported health status. Health Services Research, 35, 995–1010.

Romeis, J. C., Heath, A. C., Xian, H., Eisen, S. A., Scherrer, J. F., Pedersen, N. L., & True, W. R. (2005). Heritability of SF-36 among middle-age, middle-class, male-male twins. Medical Care, 43, 1147–1154.

Stacey, D., Clarke, T. K., & Schumann, G. (2009). The genetics of alcoholism. Current Psychiatry Reports, 11, 364–369.

Stubbe, J. H., Boomsma, D. I., Vink, J. M., Cornes, B. K., Martin, N. G., Skytthe, A., & de Geus, E. J. (2006). Genetic influences on exercise participation in 37051 twin pairs from seven countries. PLoS One, 1, e22.

Sturgis, P., Read, S., Hatemi, P. K., Zhu, G., Trull, T., Wright, M. J., & Martin, N. G. (2010). A genetic basis for social trust? Political Behavior, 32, 205–230.

Thomsen, S. F., van der Sluis, S., Kyvik, K. O., Skytthe, A., & Backer, V. (2010). Estimates of asthma heritability in a large twin sample. Clinical and Experimental Allergy, 40, 1054–1061.

Treloar, S. A., McDonald, C. A., & Martin, N. G. (1999). Genetics of early cancer detection behaviours in Australian female twins. Twin Research, 2, 33–42.

True, W. R., Heath, A. C., Bucholz, K., Slutske, W., Romeis, J. C., Scherrer, J. F., & Tsuang, M. T. (1996). Models of treatment seeking for alcoholism: The role of genes and environment. Alcoholism, Clinical and Experimental Research, 20, 1577–1581.

True, W. R., Romeis, J. C., Heath, A. C., Flick, L. H., Shaw, L., Eisen, S. A., Goldberg, J., & Lyons, M. J. (1997). Genetic and environmental contributions to healthcare need and utilization: A twin analysis. Health Services Research, 32, 37–53.

APPENDIX A

Table A1. Coefficients of logistic regression of identical versus nonidentical twin status on family background indicators.

Variable	Coefficient	Std. Err.	*P*-value
White vs. nonwhite	−0.131	0.296	0.659
Family financial background*			
A lot better off	−0.192	0.434	0.657
Somewhat better off	−0.016	0.201	0.937
A little better off	0.426	0.173	0.014
A little worse off	−0.115	0.167	0.493
Somewhat worse off	−0.018	0.208	0.93
A lot worse off	−0.139	0.322	0.666
Number of times moved to new neighborhood	−0.001	0.021	0.96
Mother's educational level†			
No school/some grade school	−0.003	0.271	0.992
Eighth grade/junior high school	−0.240	0.208	0.25
Some high school	0.018	0.186	0.922
Some college	0.064	0.201	0.75
College graduate	−0.188	0.232	0.417
Intercept	−0.278	0.306	0.363

Notes: All covariates refer to childhood period.
*Reference category = same as average family.
†Reference category = high school graduate.
Model estimated at individual twin level (1,548 twins with complete data on all variables); standard errors clustered at family level.

APPENDIX B

Evaluating Shared Genetic Effects

We employ a bivariate Cholesky model to estimate the genetic effects shared between prevention preferences and health status. This model relies on the same assumptions stated above—that identical and nonidentical twins share environments to the same extent and that nonidentical twins share, on average, one-half of their genes—but it examines the covariance across twins and across traits as is shown graphically in Appendix Figure B1. The Cholesky model estimates the parameters needed to describe the genetic correlation as is detailed in Appendix Figure B1. This bivariate model estimates three latent factors (A, C, E as defined above) for each manifest variable. The important difference is the inclusion of a path between A_1 and the prevention preference index and A_1 and the health status index. Evidence that trait 1 (prevention preferences) and trait 2 (health status) may be correlated because of common genetic influence is assessed by testing path a_{21}, which is the residual genetic influence from A_1 (genetic influence on trait 1) that is also affecting trait 2. The paths described in Appendix Figure B1 can be used with equation (B.1) to estimate the genetic correlation (r_g) coefficient as:

$$r_g = \frac{a_{21}a_{11}}{\sqrt{a_{11}^2\left(a_{21}^2 + a_{22}^2\right)}}. \tag{B.1}$$

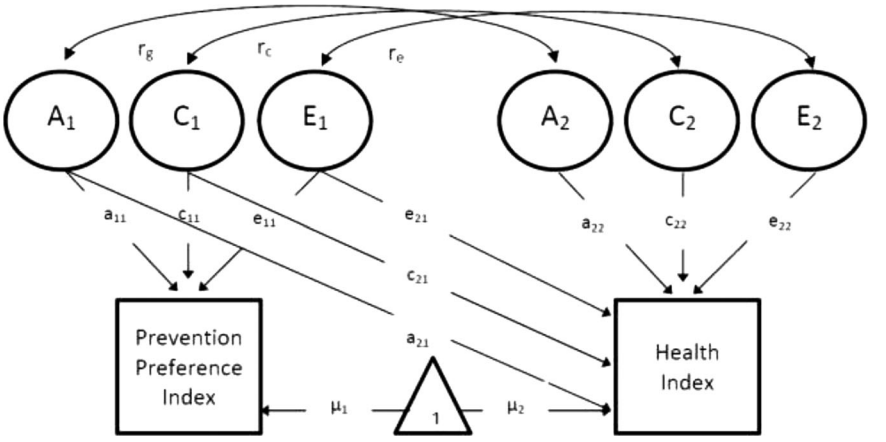


Figure B1. Path Parameters from the Bivariate Cholesky Models: Genetic and Environmental Correlations for Prevention Preference and Health Status Indices.

Table B1. Cross-twin, cross-trait, and cross-twin cross-trait correlations for prevention preference and health status indices.

	Identical twins ($n = 349$)	Nonidentical twins ($n = 558$)
Cross-twin		
r (prevention1, prevention2)	0.2933	0.1560
	[0.0000]	[0.0007]
r (health1, health2)	0.4428	0.2381
	[0.0000]	[0.0000]
Cross-outcome		
r (prevention1, health1)	−0.2055	−0.1752
	[0.0002]	[0.0001]
r (prevention2, health2)	−0.1274	−0.1901
	[0.0208]	[0.0000]
Cross-twin cross-outcome		
r (prevention1, health2)	−0.0858	−0.0796
	[0.1315]	[0.0848]
r (prevention2, health1)	−0.1005	−0.0199
	[0.0749]	[0.6669]

Notes: *P*-values of correlations are in brackets.

This correlation is characterized by the curved arrow at the top of Appendix Figure B1 describing the correlation between the two latent factors (A_1 and A_2). It is important to note, however, that the Cholesky model is best fit for large samples and strong correlations between the traits. Therefore, applying the Cholesky model is naturally limited in our case by the relatively small sample size and the small correlation between the prevention preference and health status indices (correlation coefficient = −0.19).

We first show in Appendix Table B1 descriptive analyses of the cross-twin, cross-outcome, and cross-twin cross-outcome correlations. The first entries echo those of the analyses exploring genetic effects on each outcome separately. Namely, the intrapair correlations between identical twin pairs for prevention ($r = 0.2933$) and health ($r = 0.4428$) are stronger than the correlations between nonidentical twins ($r = 0.1560$ and 0.2381, respectively). These comparisons support the notion that

Table B2. Bivariate Cholesky parameter estimates.

	Estimate	SE	*t*	*P*-value
a_{11}	−0.652	0.050	−12.976	0.000
a_{21}	0.248	0.073	3.388	0.001
a_{22}	0.796	0.044	17.957	0.000
c_{11}	0.000	1.045	0.000	1.000
c_{21}	0.000	1.114	0.000	1.000
c_{22}	0.000	0.583	0.000	1.000
e_{11}	−0.938	0.033	−28.780	0.000
e_{21}	0.098	0.044	2.220	0.027
e_{22}	0.855	0.031	27.291	0.000
$\bar{u}_1$	−0.100	0.031	−3.201	0.001
$\bar{u}_2$	0.016	0.034	0.462	0.644

Note: See Appendix Figure B1 for meaning of specific paths.

these two outcomes are partly explained by genetic variation. There is somewhat mixed evidence regarding the cross-outcome correlations when we compare identical and nonidentical twins; the average correlation between the two outcomes ($r = -0.19$) is fairly similar for identical and nonidentical twins. The most important statistics in this table are the cross-twin cross-outcome correlations. These estimates describe the degree to which one is able to predict the second outcome of the first twin as a function of the value of the first outcome of the second twin. That is, does the prevention preference of the first twin correlate with the health status of the second twin? Most importantly, if these estimates are higher for identical twins, then there is evidence that shared genes may affect both outcomes. The output provides some (albeit fairly weak) evidence for this notion. The average cross-twin cross-outcome correlation for identical twins is −0.09, compared to an average of −0.05 for nonidentical twins.

Appendix Table B2 presents the parameter estimates (paths) from the Cholesky model described above and standard errors. From these estimates, we calculate a genetic correlation (rG) of −0.299. The most important path in this specification is the a_{21} ($b = 0.248$, SE = 0.073, $t = 3.388$), which provides some evidence that the relatively small correlation between prevention preferences and health is due, in part, to genes associated with both outcomes. All three additive genetic paths are statistically significant, but gauging the statistical significance of genetic correlation is more complicated to assess. There are two approaches to assessing significance. First, one can compare the full saturated ACE model with one in which the r_g estimate is fixed to 0. This approach provided a likelihood ratio test of 2.34 with one degree of freedom, which corresponds to a *P*-value of 0.126. Second, one can bootstrap these models and describe the bounds of the empirical confidence interval. Bootstrapping with 1,000 replications yielded a 95 percent confidence interval of (−0.999, 0.174) and a 90 percent confidence interval of (−0.860, 0.048). The model fit and bootstrap methods do not provide enough evidence to reject the null hypothesis that the r_g estimate is equal to zero in the population. However, the direction and magnitude of these associations provide some evidence that these two outcomes may be influenced very weakly by common sets of genes. The lack of significance of the r_g estimate may be partly due to the small sample size and the relatively small correlation between the prevention preference and health status indices.

APPENDIX C

Table C1. Twin correlations by age.

	25 to 74 years ($n = 907$)		25 to 64 years ($n = 829$)		25 to 44 years ($n = 474$)	
Outcome	Identical twins ($n = 349$)	Nonidentical twins ($n = 558$)	Identical twins ($n = 326$)	Nonidentical twins ($n = 503$)	Identical twins ($n = 189$)	Nonidentical twins ($n = 285$)
Heart risk prevention	0.232	0.056	0.239	0.056	0.280	0.041
Cancer risk prevention	0.264	0.094	0.234	0.085	0.288	0.107
Prevention preference index	0.293	0.156	0.305	0.170	0.290	0.134
Works hard	0.367	0.080	0.385	0.076	0.295	0.127
Thought/effort	0.330	0.182	0.316	0.199	0.351	0.172
Preventive care	0.262	0.102	0.259	0.090	0.337	0.139
Prescription drug use	0.282	0.133	0.268	0.161	0.221	0.145
Self-rated health	0.296	0.205	0.274	0.211	0.379	0.247
Number of chronic conditions	0.457	0.242	0.449	0.260	0.475	0.152
Health status index	0.443	0.238	0.423	0.255	0.514	0.236

Notes: Twin correlations on outcomes with significant genetic effects (Table 2) between identical and nonidentical twins by age group. The total number of twin pairs (n) is in parentheses; the number of twin pairs with complete data varied by outcome.