Discover Social Science and Health

BRIEF COMMUNICATION Open Access

Self-rated health and mortality: family background and genetic precursors

Brian Karl Finch[1,2*], Deborah Finkel[1,3], Margaret Gatz[1], Nancy L. Pedersen[4], Chandra A. Reynolds[5], Perminder Sachdev[6], Amanda Selwood[6], Nick Martin[7] and Miriam A. Mosing[4,8]

*Correspondence:
Brian Karl Finch
brian.finch@usc.edu

Full list of author information is available at the end of the article

Abstract

Background Self-rated health (SRH) strongly predicts future mortality, even after controlling for various confounding factors. This study investigates two potential confounders of the SRH/mortality relationship—shared family background and genetics.

Methods We analyze a dataset of several harmonized twin studies from the integrating genes and environment from multiple studies consortium. Utilizing a within-between twin methodology, we assess whether the SRH/mortality relationship can be explained by social and genetic inheritance.

Results Our within-twin estimates are notably lower than the observational estimates, although the difference is statistically non-significant, indicating no substantial confounding from family background. Additionally, we find no significant interaction effects by zygosity, suggesting no confounding from shared genetic factors.

Conclusions SRH has been shown to be robust to multiple sources of variation, including demographic sub-groups and contemporary controls for clinical assessments. This study reaffirms the resilience of the SRH/mortality relationship against confounding from shared family background and genetic factors.

1 Introduction

Measures of self-rated health (SRH) have been validated by a large number of studies. Most of these studies find that poorer self-ratings of health have predictive value for future mortality [1], even after controlling for a wide variety of measures that cover medical, physical, cognitive, emotional, and social status domains [1]. These risks are similar for most major racial/ethnic and gender groups in the United States [2], the sole exception being that less acculturated and foreign-born Latinx are known to have self-reports that are discordant with objective measures and are less predictive of future mortality [3]. Studies of health among adults have shown that while physician assessments are predictive of future morbidity/mortality, they are often out-performed by individual subjective assessments, including SRH [4]. Finally, research has found the SRH/mortality

relationship to be robust to regional variation in the United States [5], and SRH is valid in multiple countries and languages [6].

Such a robust and highly predictive single item measure seems to belie most of what we know about psychometric research and begs the question as to why this measure is so consistently predictive over time and across cultural milieu. A summary of the literature argues that this paradoxical measure is predictive for four major reasons [1]. First, the measure is *inclusive* in that captures a holistic domain of health that is not reducible to the presence/absence of disease and seems to capture both clinical and sub-clinical disease stages [1]. Second, SRH is a *dynamic* measure that is based on a continuous monitoring of bodies and health that physicians are largely unaware of; in addition, this measure is valid for assessing both current states of health and future health trajectories [1]. Third, SRH has been found to *influence preventive health behaviors* (physical activity, e.g.), and better SRH ratings are associated with less risky behaviors and less abnormal behavior during illness [1]. Fourth, SRH is found to *reflect* a host of economic, psychological, and social *resources*, including: education, financial well-being, social support, optimism, perceived control, and self-efficacy [1]. Although accounting for each of these factors can reduce the relationship between SRH and various health outcomes, none of them fully account for the predictive validity of the measure.

Virtually all studies to date have attempted to account for the predictive validity of SRH by controlling for contemporaneous measures of health, and ignore both social and genetic inheritance. Recent advances in genomics and statistical methods that can account for genetics and shared childhood background have proven these two factors to be strong confounders of many well-established relationships in public health research. For example, the oft-replicated relationship between Type-2 diabetes and dementia is now thought to be spurious due in large part to genetic pleiotropy—that is, common genetic factors that predict both Type-2 diabetes and dementia [7]. In addition, the virtually iron-clad relationship between education and health is now known to be at least partially attenuated by both family-level characteristics (socio-economic status, e.g.) and common genetic factors that predict both educational attainment and long-term health [8].

Although there are no molecular studies of genetic pleiotropy for both SRH and mortality, separate studies show that SRH ($h^2 = 0.40$) [9] and longevity [10] are highly heritable, and exhibit underlying genetic influences. Given large genetic overlap in many determinants of health, common genetic variants may confound the SRH/mortality relationship. Similarly, given the rich literature on the social determinants of health, it is also plausible that shared family background experience—such as socio-economic status during childhood—might also confound this robust relationship. Thus, it is our intention to test the hypothesis that the SRH/mortality relationship may be due to both shared family background and genetic influences. The proposed within-twin methodology and the comparison of both DZ and MZ twins will provide direct evidence of potential confounding by both of these sources.

2 Methods

In order to test whether the SRH/mortality relationship is robust to confounding from common genetic variants and family background, we use a large harmonized sample (n = 5,743) of both monozygotic (MZ n = 1,944) and both opposite and same-sex

dizygotic twins (DZ n = 3,799) from the Interplay of Genes and Environment across Multiple Studies (IGEMS) Consortium [11]—using samples from five twin cohorts in both Sweden and Australia. While the entire IGEMS sample includes more than 120,000 twins from 21 international cohorts in 5 countries, our analysis requires us to use only twins with a mortality follow-up, which restricts our sample size. This sub-sample includes the following 7 twin cohorts: the Swedish Adoption/Twin Study of Aging [12], Study of Origins of Variance in the Oldest-Old [13], Aging in Women and Men [14], Longitudinal Study of Aging Danish Twins [15], Middle Age Danish Twin study [16], Australian Over 50's study [17], and Older Australian Twin Study [18].

We harmonize several different SRH measures and create standardized t-scores ($\overline{X} = 50$, s.d. = 10) within each twin cohort which are then merged for a pooled analysis. In the pooled sample, t-scores range from 32 to 89, with higher ratings indicating better health. A further description of both the general harmonization approach [19] and the approach for self-rated health [20] are described elsewhere. Month/year of birth and month/year of death or last measurement wave for right-censored twins are used to create a time-to-event variable with dead/not dead indicating the event of interest (dead n = 3,179, 55%). We employ *between-within* methods for twin research, which, in the absence of measurement error and substantial non-shared confounding, can produce a specific type of causal estimate that is purged of both genetic and family background confounds [21]. This approach requires *centering* measures of SRH within twin pairs by computing an average SRH for each pair (the between effect) and then differencing the individual twins from their average (the within effect; see Eq. 1). If an association between SRH and Mortality is partially attenuated in the intra-pair analyses of DZ twins and fully attenuated in MZ twins, the association may be attributable to genetic confounding. If an association is attenuated in the intra-pair analyses of both DZ and MZ twins, the association may be attributable to confounding from shared childhood environment. If an association persists in both sets of twin analyses, it would suggest an effect of SRH not due to confounding from shared familial factors nor common genetics.

First, we specify a Cox proportional hazards model for the i_{th} twin in the j_{th} twin-pair that regresses time to death on SRH (both between (B) and within (W) twin pairs), and a set of statistical controls, including: dummies for twin study (DZ = 0, MZ = 1), sex, age, birth cohort (pre-Great Depression, pre-WWII, and post-WWII), educational attainment (International Standard Classification of Education), and zygosity. This hazard models treats the hazard function as the dependent variable, which represents the instantaneous risk of death occurring at a given time, given that the event has not yet occurred. For model specification purposes, we include an event indicator that specifies dead (1) vs. alive/right-censored (0) and then include a duration variable which is months until death/right-censoring for those still alive. A second model includes interactions between both within/between SRH estimates and zygosity, in order to test for additional genetic confounds among MZ twins (see Eq. 1). Standard errors are estimated using Huber/White/sandwich robust estimation in Stata.

$$\begin{aligned} Mortality_{ij} = {} & \alpha_0 + \beta_B\left(\overline{SRH}_j\right) + \beta_W\left(SRH_{ij} - \overline{SRH}_j\right) + \beta_1\left(Study_{ij}\right) \\ & + \beta_2\left(Sex_{ij}\right) + \beta_3\left(Age_{ij}\right) + \beta_4\left(BirthCohort_{ij}\right) + \beta_5\left(ISCED_{ij}\right) + \beta_6\left(Zygosity_{ij}\right) \\ & + \beta_7\left(\overline{SRH}_j\right)\left(Zygosity_{ij}\right) + \beta_8\left(SRH_{ij} - \overline{SRH}_j\right)\left(Zygosity_{ij}\right) + \epsilon_{ij} \end{aligned} \quad (1)$$

Table 1 Cox proportional hazards survival model for time to death: "between-within twin" SRH estimates

Variable	Model 1			Model 2		
	B	SE	95% CI	B	SE	95% CI
SRH_B	−.0267	.0027	−.0320, −.0215	−0253	.0024	−.0300, −.0207
SRH_W	−.0181	.0038	−.0255, −.0107	−.0124	.0042	−.0206, −.0042
SRH_B * Zygosity				.0027	.0021	−.0014, .0068
SRH_W * Zygosity				−.0056	.0066	−.0185, .0074

Controls for sex*, age*, education*, cohort*, zygosity, and twin study*(2/4)

*$p < .05$ for control variables

3 Results

Results are presented in Table 1. Both the within and between SRH estimates in Model 1 are statistically significant, indicating that SRH predicts mortality, as expected. The within-twin coefficients estimate an 18% decrease for every 1 standard deviation increase in SRH. However, although 23% of the between-twin effect is attenuated in the within-twin estimates, the confidence intervals overlap indicating no statistically significant attenuation due to shared family background. It is important to note that our comparisons used a population-based observational model to estimate the overall effect, which was then compared to our within-twin effect, following recommendations in the literature [22, 23]. Model 2 specifies an interaction with zygosity (MZ = 1) and finds that the MZ and DZ twin SRH estimates are similar in magnitude as there is no significant interaction with the between-twin estimate, and most importantly, finds no statistically significant genetic confounding as the within-estimate by zygosity interaction is non-significant.

As a test for robustness, we also sex-stratified our models but found no significant substantive differences. While SRH was more strongly predictive of the male sample, the confounding estimates were similar for males and females and the interaction with zygosity remained non-significant. Further, since controlling for educational attainment might over-control for family socio-economic background, we re-specified models with and without this variable with virtually identical results. In addition, we considered interactions [24] for zygosity by age, zygosity by sex, and zygosity by EA3 as potential confounders, but none were significant, nor did they affect the substantive results reported above [25].

4 Discussion and conclusions

Despite being only a single-item measure, SRH has shown to be both highly predictive of future mortality in multiple research settings, and particularly robust to common sources of confounding, even comprehensive clinical assessment. At the same time, many long-standing relationships between various health measures have been shown to be at least partially attenuated, if not fully explained, by common sources of variation. Two common sources of confounding—shared family background and genetics—have proven to at least partially, if not fully account for several epidemiological relationships. However, our research shows that SRH is robust to these confounds and remains a simple, yet strong predictor of current health and future health trajectories. While 23% of the initial relationship is explained by familial background, this effect was not statistically significant. Although we are using one of the largest twin samples with this specific combination of variables that we know of, we still may be lacking in power to uncover

relationships that may exist when pooling even larger samples. As such, our conclusions should be tempered, while also recognizing the frequent robustness of SRH to confounding. Finally, although our results did not indicate statistically significant effect modification by zygosity, there was some attenuation of our estimates in the MZ models suggesting marginal confounding by genetics. This is consistent with the literature on genetic pleiotropy, which suggests that the genetic predictors for SRH are also correlated with a host of mental and physical health outcomes that are clearly related to premature death [26].

Authors contributions
BKF developed the initial analysis plan, ran the analyses with assistance from MM, and wrote the first draft. All authors contributed to the analytical design and methodology. BKF, DF, MG, and NP all contributed to the formation of the current IGEMS research consortium and made data available from individual twin cohorts. CR assisted with statistical design and interpretation. PS, AS, NM, and MM all harmonized data and assisted with merging to the IGEMS data consortium. BKF and MM conducted statistical analysis and all authors assisted with revisions and finalization of the draft.

Funding
This work was supported in part by The National Institute of Aging grants R01AG059329, R01AG081248, and by the Center for Economic and Social Research at the University of Southern California.

Data availability
Data Availability. Data are restricted and available only via contractual agreement with the IGEMS Consortium. Code Availability. Programming code will be made available on the USC IGEMS website upon publication and will also be available from the corresponding author upon request.

Code availability
Programming code will be made available on the USC IGEMS website upon publication and will also be available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate
Not Applicable.

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA
[2]Department of Sociology and Spatial Sciences, University of Southern California, Los Angeles, CA, USA
[3]Department of Psychology, Jönköping University, Jönköping, Sweden
[4]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[5]Department of Psychology and Neuroscience, Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, CO, USA
[6]Centre for Healthy Brain Ageing, School of Clinical Medicine, University of New South Wales, Sydney, Australia
[7]Genetic Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Herston, Australia
[8]Department of Cognitive Neuropsychology, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

Received: 8 October 2024 / Accepted: 27 August 2025
Published online: 06 October 2025

References

1. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997;38(1):21–37.
2. McGee DL, et al. Self-reported health status and mortality in a multiethnic US cohort. Am J Epidemiol. 1999;149(1):41–6.
3. Finch BK, et al. Validity of self-rated health among Latino(a)s. Am J Epidemiol. 2002;155(8):755–9.
4. Ferraro KF, Farmer MM. Utility of health data from social surveys: is there a gold standard for measuring morbidity? Am Sociol Rev. 1999;64(2):303–15.
5. Berchick ER, Lynch SM. Regional variation in the predictive validity of self-rated health for mortality. SSM Popul Health. 2017;3:275–82.
6. Pérez-Zepeda MU, et al. Assessing the validity of self-rated health with the short physical performance battery: a cross-sectional analysis of the International Mobility in Aging Study. PLoS ONE. 2016;11(4):e0153855.
7. Thomassen JQ, et al. Type-2 diabetes and risk of dementia: observational and Mendelian randomisation studies in 1 million individuals. Epidemiol Psychiatr Sci. 2020;29:e118.

8. Fujiwara T, Kawachi I. Is education causally related to better health? A twin fixed-effect study in the USA. Int J Epidemiol. 2009;38(5):1310–22.
9. Svedberg P, et al. Self-rated health in a longitudinal perspective: a 9-year follow-up twin study. J Gerontol B Psychol Sci Soc Sci. 2005;60(6):S331-s340.
10. Tesi N, et al. Polygenic risk score of longevity predicts longer survival across an age continuum. J Gerontol A Biol Sci Med Sci. 2021;76(5):750–9.
11. Pedersen NL et al. IGEMS: the consortium on interplay of genes and environment across multiple studies—an update. Twin Res Human Gene. 2019;1–8.
12. Pedersen NL, et al. The Swedish adoption twin study of aging: an update. Acta Genet Med Gemellol Twin Res. 1991;40(1):7–20.
13. McClearn GE, et al. Substantial genetic influence on cognitive abilities in twins 80 or more years old. Science. 1997;276:1560–3.
14. Gold CH, et al. Gender and health: a study of older unlike-sex twins. J Gerontol B Psychol Sci Soc Sci. 2002;57B:S168–76.
15. Christensen K, et al. A Danish population-based twin study on general health in the elderly. J Aging Health. 1999;11(1):49–64.
16. Osler M, et al. Marital status and twins' health and behavior: an analysis of middle-aged Danish twins. Psychosom Med. 2008;70(4):482–7.
17. Hopper JL, et al. Australian twin registry: 30 years of progress. Twin Res Hum Genet. 2013;16(1):34–42.
18. Sachdev PS, et al. A comprehensive neuropsychiatric study of elderly twins: the Older Australian Twins Study. Twin Res Hum Genet. 2009;12(6):573–82.
19. Gatz M, et al. Data harmonization in aging research: not so fast. Exp Aging Res. 2015;41(5):475–95.
20. Franz CE, et al. Facets of subjective health from early adulthood to old age. J Aging Health. 2017;29(1):149–71.
21. Sjölander A, Frisell T, Öberg S. Causal interpretation of between-within models for twin research. Epidemiol Methods. 2012;1:217–37.
22. Demange PA, et al. Educational attainment and psychiatric diagnoses: a national registry data and two-sample Mendelian randomization study. Nat Ment Health. 2024;2(6):668–79.
23. Carlin JB, et al. Regression models for twin studies: a critical review. Int J Epidemiol. 2005;34(5):1089–99.
24. Keller MC. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. Biol Psychiatry. 2014;75(1):18–24.
25. Yzerbyt VY, Muller D, Judd CM. Adjusting researchers' approach to adjustment: on the use of covariates when testing interactions. J Exp Soc Psychol. 2004;40(3):424–31.
26. Gale CR, et al. Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. Transl Psychiatry. 2016;6(4):e791.

Publisher's Note