

## RESEARCH ARTICLE

# Genetic Factors Are Not the Major Causes of Chronic Diseases

Stephen M. Rappaport\*

School of Public Health, University of California, Berkeley, California, United States of America

\* [srappaport@berkeley.edu](mailto:srappaport@berkeley.edu)

## Abstract

The risk of acquiring a chronic disease is influenced by a person's genetics (G) and exposures received during life (the 'exposome', E) plus their interactions (GxE). Yet, investigators use genome-wide association studies (GWAS) to characterize G while relying on self-reported information to classify E. If E and GxE dominate disease risks, this imbalance obscures important causal factors. To estimate proportions of disease risk attributable to G (plus shared exposures), published data from Western European monozygotic (MZ) twins were used to estimate population attributable fractions (PAFs) for 28 chronic diseases. Genetic PAFs ranged from 3.4% for leukemia to 48.6% for asthma with a median value of 18.5%. Cancers had the lowest PAFs (median = 8.26%) while neurological (median = 26.1%) and lung (median = 33.6%) diseases had the highest PAFs. These PAFs were then linked with Western European mortality statistics to estimate deaths attributable to G for heart disease and nine cancer types. Of 1.53 million Western European deaths in 2000, 0.25 million (16.4%) could be attributed to genetics plus shared exposures. Given the modest influences of G-related factors on the risks of chronic diseases in MZ twins, the disparity in coverage of G and E in etiological research is problematic. To discover causes of disease, GWAS should be complemented with exposome-wide association studies (EWAS) that profile chemicals in biospecimens from incident disease cases and matched controls.



CrossMark  
click for updates

## OPEN ACCESS

**Citation:** Rappaport SM (2016) Genetic Factors Are Not the Major Causes of Chronic Diseases. PLoS ONE 11(4): e0154387. doi:10.1371/journal.pone.0154387

**Editor:** Rodney John Scott, University of Newcastle, AUSTRALIA

**Received:** February 26, 2016

**Accepted:** April 12, 2016

**Published:** April 22, 2016

**Copyright:** © 2016 Stephen M. Rappaport. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data were obtained from published sources and are fully tabulated in the manuscript.

**Funding:** Partial funding was received from grants P42ES05948 and P42ES0470518 from the National Institute for Environmental Health Sciences.

**Competing Interests:** The author has declared that no competing interests exist.

## Introduction

As the world's population ages, mortality increasingly reflects the ravages of complex chronic diseases, particularly cancer and heart disease [1]. A person's risk of succumbing to a chronic disease is linked to his or her genetics (G) and exposome (E, representing all exposures during life) plus GxE interactions. Although geneticists and epidemiologists have debated the importance of G and E as causes of chronic diseases, it is clear that both factors affect disease risks [2, 3]. However, most etiologic research has focused on genetic causes and has relegated exposures to secondary roles. For example, when queried on Feb. 6, 2016, there were 566,685 PubMed citations for the keywords "disease causes AND genetics" compared to 71,922 citations for "disease causes AND exposure", a ratio of about eight to one.

This genome-centric view of causation is motivated by the technologic ability to detect and manipulate genes, and fosters the notion that genetic factors are necessary determinants of disease that operate in a causal background of diverse exposures [4, 5]. Certainly, technologies

spawned by the human genome project led to stunningly comprehensive genome-wide association studies (GWAS) that investigated genomic variability across thousands of diseased and healthy subjects. Yet, because more than 2,000 GWAS rarely reported relative risks greater than 1.2 [6, 7], geneticists are turning to whole-genome sequencing in searches for ‘missing heritability’ [8, 9]. This motivation stems, at least in part, from calculations of heritability that do not differentiate disease variation arising from genetic factors and shared exposures [10].

In contrast to GWAS, the epidemiology of causal exposures still relies on self-reported and geographic information plus a few targeted measurements [11, 12], much as it did a century ago. Nonetheless, data from the World Health Organization (WHO) has attributed nearly half of global mortality to a handful of exposures (Table 1), mainly particulate air pollution (including indoor smoke and occupational exposure) (14% of all deaths), tobacco smoking and second-hand smoke (13%), high plasma levels of sodium (6%), and alcohol use (which is generally protective but can be harmful with high consumption) (5%) [13]. There is also strong epidemiologic evidence that genetically-stable populations experience profound alterations in cancer incidence across generations and with migration that logically reflect changing exposures [3, 14, 15]. Thus, the empirical evidence promotes the notion that exposures are necessary determinants of disease that operate in a causal background of genetic diversity. However, compared to GWAS, the universe of exposures that has been investigated for associations with chronic diseases essentially consists of airborne particulate matter plus a set of about 300 environmental chemicals and nutrients [16].

To investigate the global influence of genetic factors on chronic-disease risks, data from cohorts of monozygotic (MZ) twins in Western Europe were compiled to estimate population

**Table 1. Global deaths attributed to exposure-risk factors for chronic diseases.**

Risk factor	Attributed deaths	Percent of global deaths
Tobacco smoking	5,695,349	11.28
Indoor smoke	3,478,773	6.89
Ambient particulate pollution	3,223,540	6.38
Diet high in sodium	3,104,308	6.15
Alcohol use	2,735,511	5.42
Diet low in seafood omega-3 fatty acids	1,389,896	2.75
Lead exposure	674,038	1.33
Second-hand smoke	601,938	1.19
Diet low in polyunsaturated fatty acids	533,603	1.06
Diet high in trans fatty acids	515,260	1.02
Occupational chemicals	373,738	0.74
Drug use	157,805	0.31
Ambient ozone pollution	152,434	0.30
Diet low in calcium	125,594	0.25
Vitamin A deficiency	119,762	0.24
Iron deficiency	119,608	0.24
Residential radon	98,992	0.20
Zinc deficiency	97,330	0.19
<b>TOTAL</b>	<b>23,197,479</b>	<b>45.9</b>

Data were obtained from World Health Organization estimates of exposures affecting 50,506,784 global deaths in 2010 [13]. (Because of possible correlations across risk factors, attributed deaths and percentages may overestimate true values).

doi:10.1371/journal.pone.0154387.t001

attributable fractions (PAFs) for 28 chronic diseases, including prominent cancers, cardiovascular diseases, neurologic diseases, lung diseases, and autoimmune diseases. Because pairs of MZ twins have essentially identical genomes and also share many exposures [10], especially in early life, these PAFs estimate proportions of cases that would *theoretically* be prevented if interventions were able to remove particular combinations of genotypes and shared exposures [17, 18]. To further evaluate the impacts of G and E on the risks of chronic diseases, PAFs from MZ twins were linked with mortality statistics from Western Europe to estimate the numbers of deaths attributable to genetic factors and shared exposures for ischemic heart disease and prominent cancers.

## Materials and Methods

Data for estimation of PAFs were obtained from publications of disease phenotypes in large MZ-twin cohorts, i.e. [19–35], some of which had been curated by Roberts *et al.* [36]. Virtually all of the data were from Western European twins, primarily Swedish, Danish, and Finnish. The 28 diseases included nine types of cancer, cardiovascular diseases (heart disease and stroke), neurological diseases (Parkinson's disease, Alzheimer's disease, dementia, and migraine), lung diseases (chronic obstructive pulmonary disease and asthma), obesity-associated diseases (Type-2 diabetes and gallstone disease), autoimmune diseases (rheumatoid arthritis, Type-1 diabetes, and thyroid autoimmunity), genitourinary diseases (general dysuria, stress urinary incontinence, and pelvic organ prolapse), and three other syndromes (chronic fatigue, irritable bowel syndrome, and gastroesophageal reflux). The following data were extracted from each study: gender, number of MZ twin pairs ( $N_T$ ), number of concordant MZ pairs ( $N_C$ ), and number of discordant MZ pairs ( $N_D$ ). Each PAF (%) was estimated as  $P^*$  ( $RR-1)/RR$ , where  $P = 2N_C/(2N_C+N_D)*100$  represents the proportion (%) of case twins with an affected co-twin and  $RR = [2N_C/(2N_C+N_D)]/[N_D/(N_D+2N_T-2(N_C+N_D))]$  is the associated relative risk. If statistics were reported for both male and female twin pairs, then PAFs were estimated for the combined datasets. Since all of the studies were reported around the year 2000 for twins from Western Europe, mortality statistics for Western Europeans in that year were obtained for ischemic heart disease and relevant cancers from the WHO Global Burden of Disease Database [37, 38]. To estimate deaths attributable to genetics and shared exposures, the number of deaths for each disease type were multiplied by the corresponding PAF from MZ twins.

## Results and Discussion

Statistics from studies of MZ twins are summarized in Table 2. Estimated G-related PAFs ranged from 3.4% for leukemia to 48.6% for asthma with a median value of 18.5% and interquartile range of 9.9% to 24.2%. This indicates that fractions of cases attributable to genetics plus shared exposures tend to be modest, with three fourths of the phenotypes having PAFs less than 25%. In fact, G-related PAFs for only two phenotypes were greater than 40%, i.e. thyroid autoimmunity (42%) and asthma (49%).

Fig 1 displays the cumulative distribution for the 28 phenotypes with symbols representing disease categories. Although there was variability within a given category, cancers tended to have the lowest PAFs (median = 8.26%) while neurological (median = 26.1%) and lung (median = 33.6%) diseases had the highest PAFs. Although these are apparently the first estimates of PAFs derived exclusively from MZ twins, Hemminki and Czene reported familial PAFs for cancers in the Swedish-Family Cancer Database (10.2 million individuals) [18] that are consistent with these results.

**Table 2. Data sources and statistics for estimation of population attributable fractions (PAFs) in monozygotic twins.**

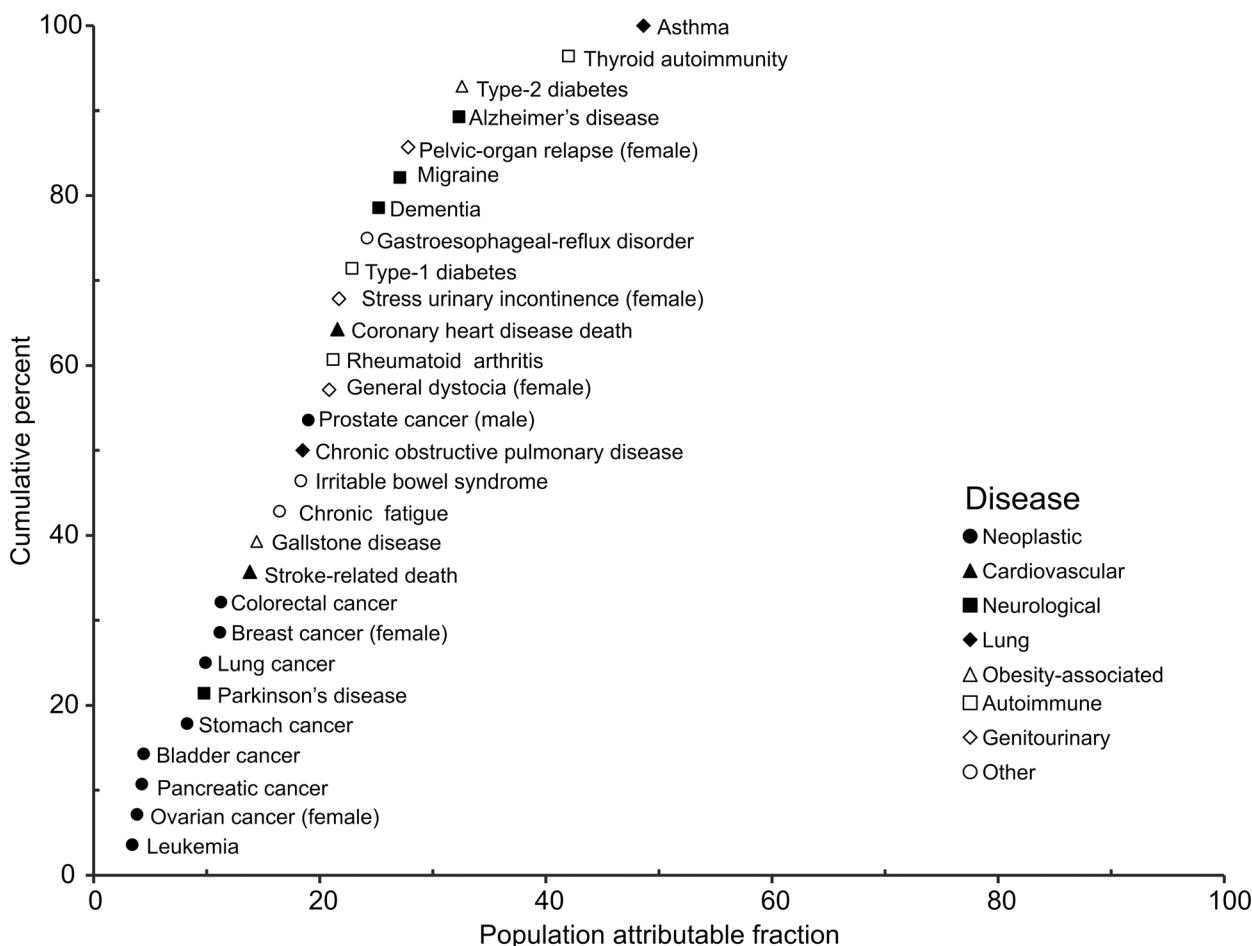
Disease	Country	N <sub>T</sub>	N <sub>C</sub>	N <sub>D</sub>	P (%)	RR	PAF (%)	Ref.
Bladder cancer	Sweden, Denmark, Finland	15,668	5	189	5.03	8.28	4.42	[19]
Breast cancer (female)	Sweden, Denmark, Finland	8,437	42	505	14.3	4.60	11.2	[19]
Colorectal cancer	Sweden, Denmark, Finland	15,668	30	416	12.6	9.35	11.3	[19]
Leukemia	Sweden, Denmark, Finland	15,668	2	103	3.74	11.3	3.41	[19]
Lung cancer	Sweden, Denmark, Finland	15,668	18	296	10.8	11.4	9.89	[19]
Ovarian cancer (female)	Sweden, Denmark, Finland	8,437	3	125	4.58	6.13	3.83	[19]
Pancreatic cancer	Sweden, Denmark, Finland	15,668	3	123	4.65	11.8	4.26	[19]
Prostate cancer (male)	Sweden, Denmark, Finland	7,231	40	299	21.1	9.94	19.0	[19]
Stomach cancer	Sweden, Denmark, Finland	15,668	11	223	8.98	12.5	8.26	[19]
Thyroid autoimmunity	Denmark	284	7	17	45.2	14.3	42.0	[20]
Type 1 diabetes	Finland	4,307	3	20	23.1	99.1	22.8	[21]
Type 2 diabetes	Finland	4,307	29	113	33.9	25.3	32.6	[21]
Gallstone disease	Sweden	11,073	112	956	19.0	4.16	14.4	[22]
Alzheimer's disease	Sweden	398	2	8	33.3	32.7	32.3	[23]
Dementia	Sweden	398	3	16	27.3	13.2	25.2	[23]
Chronic fatigue	Sweden	3,229	181	792	31.4	2.10	16.4	[24]
Gastroesophageal reflux disorder	Sweden	2178	95	370	33.9	3.48	24.2	[25]
Irritable bowel syndrome	Norway	1,252	14	97	22.4	5.49	18.3	[26]
Coronary heart disease death	Sweden	3,644	250	875	36.4	2.46	21.6	[27]
Stroke-related death	Denmark	3,852	35	316	18.1	4.20	13.8	[28]
General dystocia (female)	Sweden	928	40	173	31.6	2.93	20.8	[29]
Pelvic organ prolapse (female)	Sweden	3,376	34	157	30.2	12.6	27.8	[30]
Stress urinary incontinence (female)	Sweden	3,376	13	87	23.0	17.6	21.7	[30]
Migraine	European Union	9,077	382	1377	35.7	4.15	27.1	[31]
Rheumatoid arthritis	Finland	4,137	9	64	22.0	28.1	21.2	[32]
Asthma	Denmark	5,084	257	447	53.5	11.0	48.6	[33]
Parkinson disease	Sweden	8,590	9	151	10.6	12.0	9.76	[34]
Chronic obstructive pulmonary disease	Sweden, Denmark	7,747	18	149	19.5	20.0	18.5	[35]

Legend: N<sub>T</sub>, total twin pairs; N<sub>C</sub>, concordant twin pairs; N<sub>D</sub>, discordant twin pairs; P, proportion of case twins with an affected co-twin (%); RR, relative risk.

doi:10.1371/journal.pone.0154387.t002

Since heart disease and cancer are the two leading causes of mortality in Western Europe (and worldwide), the contributions of genetics plus shared exposures to incidence of these diseases were estimated as summarized in Fig 2. Assuming that the populations of MZ twins that were used to derive PAFs are reasonable surrogates for Western Europeans in the year 2000, then 0.25 million of the 1.53 million cancer and heart-disease deaths (16.4%) can be attributed to G-related factors.

Because comprehensive exposure data were not collected in the twin studies, it is not possible to directly estimate E-related PAFs or contributions to disease risks from G×E interactions. But given the modest values of G-related PAFs reported here, it is reasonable to infer that the combined effects of non-shared exposures (E) and G×E would be greater than those of G alone. This conjecture is supported by results of structural equation modeling by Lichtenstein *et al.* [19], who reported that non-shared exposures in monozygotic and dizygotic twins accounted for between 58% and 82% (median = 62%) of the variation in 12 types of cancer. Nonetheless, the hypothesized dominance of E and G×E on chronic-disease risks is at odds with a recent paper by Tomasetti and Vogelstein who found a strong correlation between cancer risks and total numbers of stem-cell divisions in various tissues, and concluded from this



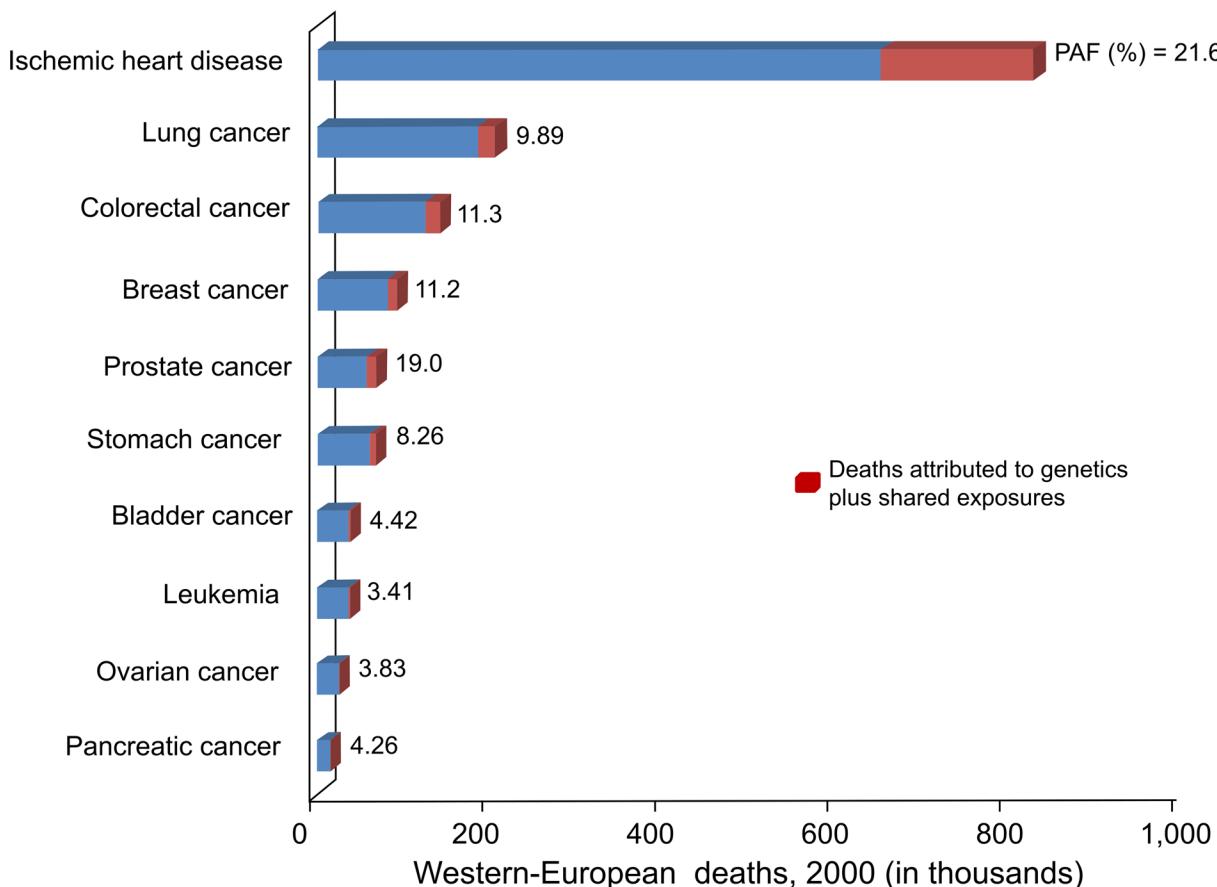
**Fig 1. Population attributable fractions (PAFs) for 28 disease phenotypes estimated from studies of monozygotic twins.** Sources of data and statistics are summarized in [Table 2](#).

doi:10.1371/journal.pone.0154387.g001

that 'bad luck' accounts for about two thirds of the variation in cancers [39]. In refuting this conclusion, Wu *et al.* pointed out that the bad-luck hypothesis is illogical because: it equates correlation with causation; it is inconsistent with the epidemiological evidence; and it requires that mutation signatures of cancers be correlated with age, which is rarely the case [15]. Recognizing that both intrinsic random errors and E-related factors can influence cancer risks, Wu *et al.* then applied models to the same data used by Tomasetti and Vogelstein [39], which allowed E-related factors to be estimated after adjustment for intrinsic random errors. Results indicated that E-related factors typically explained more than 90% of cancer risk, consistent with the small genetic PAFs observed for cancers in MZ twins (median = 8.26%).

## Conclusions

Because the human genome project planted the seeds for genome sequencing and large-scale omics technologies [5], it was inevitable that these methods would be used to search for causes of major diseases, and almost 2,000 GWAS have been reported [6]. Yet, the matrix of disease-associated genetic variants does not explain much heritability [7,9]. Indeed, Yang *et al.* predicted that between 20 and 50 causal genetic variants would be required to explain half the burden of a common disease, depending on the frequency of each variant and risk ratio of the



**Fig 2. Numbers of Western-European deaths in 2000 estimated for ischemic heart disease and nine cancer types (1.53 million total deaths from these causes).** The contributions attributed to genetics plus shared exposures are based on the population attributable fractions (PAFs) estimated from Western European monozygotic twins (Table 2).

doi:10.1371/journal.pone.0154387.g002

genotype [17]. The small genetic PAFs estimated here from studies of MZ twins (Table 2 and Fig 1) cast further doubt on the notion that our inherited genomes are the primary causes of chronic diseases. Nonetheless, the genome can influence disease outcomes through G×E interactions, and may also contribute through epistasis and heritable epigenetic effects that are as yet unknown. Thus, investigations of causes of chronic diseases should continue to consider genetic factors as part of a balanced strategy that characterizes both E and G with high resolution.

One avenue for discovering E-related risks would be to extend the data-driven approach embodied in GWAS and conduct exposome-wide association studies (EWAS) [40] via untargeted analyses of chemicals in blood (the ‘blood exposome’) [16]. Since disease processes can alter the blood exposome through dysregulation of systems biology, it is important that EWAS be conducted with archived biospecimens collected prior to diagnosis from incident cases and matched controls in prospective cohort studies. This makes it possible to distinguish chemical signatures of potentially causal exposures from those generated by progression of the disease (reverse causality) [40, 41].

A good example of this data-driven approach for EWAS is given by Wang *et al.* [42] who found 18 chemical features (out of more than 2000 detected) that were associated with cardiovascular disease in samples totaling only 75 incident cases and 75 matched controls. Three of

the features were identified as choline and its metabolites, betaine, and trimethylamine-*N*-oxide (TMAO), with TMAO exhibiting the strongest disease risk in follow-up studies [43, 44]. Since TMAO is a product of joint microbial and human metabolism of choline, the positive association between plasma TMAO and disease risk points to possible involvement of the gut microbiota in the etiology of cardiovascular disease. It is interesting that a study of colorectal cancer by Bae *et al.* [45] also found a positive association between plasma TMAO and disease risk, again suggesting involvement of the gut microbiota.

## Author Contributions

Conceived and designed the experiments: SMR. Performed the experiments: SMR. Analyzed the data: SMR. Contributed reagents/materials/analysis tools: SMR. Wrote the paper: SMR.

## References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death in 2007: a systematic analysis for the Global Burden of Disease Study 2007. *Lancet*. 2012; 380(9859):2095–128. Epub 2012/12/19. doi: S0140-6736(12)61728-0 [pii] doi: [10.1016/S0140-6736\(12\)61728-0](https://doi.org/10.1016/S0140-6736(12)61728-0) PMID: [23245604](https://pubmed.ncbi.nlm.nih.gov/23245604/).
2. Chakravarti A, Little P. Nature, nurture and human disease. *Nature*. 2003; 421(6921):412–4. Epub 2003/01/24. doi: [10.1038/nature01401](https://doi.org/10.1038/nature01401) [pii]. PMID: [12540911](https://pubmed.ncbi.nlm.nih.gov/12540911/).
3. Hemminki K, Lorenzo Bermejo J, Forsti A. The balance between heritable and environmental aetiology of human disease. *Nat Rev Genet*. 2006; 7(12):958–65. Epub 2006/12/02. doi: nrg2009 [pii] doi: [10.1038/nrg2009](https://doi.org/10.1038/nrg2009) PMID: [17139327](https://pubmed.ncbi.nlm.nih.gov/17139327/).
4. Gannett L. What's in a Cause?: The Pragmatic Dimensions of Genetic Explanations. *Biology and Philosophy*. 1999; 14:349–74.
5. Haga SB, Khoury MJ, Burke W. Genomic profiling to promote a healthy lifestyle: not ready for prime time. *Nat Genet*. 2003; 34(4):347–50. doi: [10.1038/ng0803-347](https://doi.org/10.1038/ng0803-347) PMID: [12923535](https://pubmed.ncbi.nlm.nih.gov/12923535/).
6. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014; 42(Database issue):D1001–6. doi: [10.1093/nar/gkt1229](https://doi.org/10.1093/nar/gkt1229) PMID: [24316577](https://pubmed.ncbi.nlm.nih.gov/24316577/); PubMed Central PMCID: [PMC3965119](https://pubmed.ncbi.nlm.nih.gov/PMC3965119/).
7. Chang CQ, Yesupriya A, Rowell JL, Pimentel CB, Clyne M, Gwinn M, et al. A systematic review of cancer GWAS and candidate gene meta-analyses reveals limited overlap but similar effect sizes. *European journal of human genetics: EJHG*. 2014; 22(3):402–8. doi: [10.1038/ejhg.2013.161](https://doi.org/10.1038/ejhg.2013.161) PMID: [23881057](https://pubmed.ncbi.nlm.nih.gov/23881057/); PubMed Central PMCID: [PMC3925284](https://pubmed.ncbi.nlm.nih.gov/PMC3925284/).
8. Goldstein DB. Common genetic variation and human traits. *N Engl J Med*. 2009; 360(17):1696–8. Epub 2009/04/17. doi: NEJMmp0806284 [pii] doi: [10.1056/NEJMmp0806284](https://doi.org/10.1056/NEJMmp0806284) PMID: [19369660](https://pubmed.ncbi.nlm.nih.gov/19369660/).
9. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461(7265):747–53. Epub 2009/10/09. doi: nature08494 [pii] doi: [10.1038/nature08494](https://doi.org/10.1038/nature08494) PMID: [19812666](https://pubmed.ncbi.nlm.nih.gov/19812666/); PubMed Central PMCID: [PMC2831613](https://pubmed.ncbi.nlm.nih.gov/PMC2831613/).
10. Vineis P. Misuse of genetic data in environmental epidemiology. *Ann N Y Acad Sci*. 2006; 1076:163–7. Epub 2006/11/23. doi: 1076/1/163 [pii] doi: [10.1196/annals.1371.060](https://doi.org/10.1196/annals.1371.060) PMID: [17119201](https://pubmed.ncbi.nlm.nih.gov/17119201/).
11. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev*. 2005; 14(8):1847–50. PMID: [16103423](https://pubmed.ncbi.nlm.nih.gov/16103423/).
12. Rappaport SM, Smith MT. Environment and disease risks. *Science*. 2010; 330(6003):460–1. Epub 2010/10/23. doi: 330/6003/460 [pii] doi: [10.1126/science.1192603](https://doi.org/10.1126/science.1192603) PMID: [20966241](https://pubmed.ncbi.nlm.nih.gov/20966241/).
13. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380(9859):2224–60. Epub 2012/12/19. doi: S0140-6736(12)61766-8 [pii] doi: [10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8) PMID: [23245609](https://pubmed.ncbi.nlm.nih.gov/23245609/).
14. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst*. 1981; 66(6):1191–308. Epub 1981/06/01. PMID: [7017215](https://pubmed.ncbi.nlm.nih.gov/7017215/).
15. Wu S, Powers S, Zhu W, Hannun YA. Substantial contribution of extrinsic risk factors to cancer development. *Nature*. 2016; 529(7584):43–7. doi: [10.1038/nature16166](https://doi.org/10.1038/nature16166) PMID: [26675728](https://pubmed.ncbi.nlm.nih.gov/26675728/).

16. Rappaport SM, Barupal DK, Wishart D, Vineis P, Scalbert A. The blood exposome and its role in discovering causes of disease. *Environ Health Perspect*. 2014; 122(8):769–74. doi: [10.1289/ehp.1308015](https://doi.org/10.1289/ehp.1308015) PMID: [24659601](https://pubmed.ncbi.nlm.nih.gov/24659601/).
17. Yang Q, Khouri MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? *Int J Epidemiol*. 2005; 34(5):1129–37. doi: [10.1093/ije/dyi130](https://doi.org/10.1093/ije/dyi130) PMID: [16043441](https://pubmed.ncbi.nlm.nih.gov/16043441/).
18. Hemminki K, Czene K. Attributable risks of familial cancer from the Family-Cancer Database. *Cancer Epidemiol Biomarkers Prev*. 2002; 11(12):1638–44. Epub 2002/12/24. PMID: [12496055](https://pubmed.ncbi.nlm.nih.gov/12496055/).
19. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000; 343(2):78–85. Epub 2000/07/13. PMID: [10891514](https://pubmed.ncbi.nlm.nih.gov/10891514/).
20. Hansen PS, Brix TH, Iachine I, Kyvik KO, Hegedus L. The relative importance of genetic and environmental effects for the early stages of thyroid autoimmunity: a study of healthy Danish twins. *European journal of endocrinology / European Federation of Endocrine Societies*. 2006; 154(1):29–38. doi: [10.1530/eje.1.02060](https://doi.org/10.1530/eje.1.02060) PMID: [16381988](https://pubmed.ncbi.nlm.nih.gov/16381988/).
21. Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia*. 1992; 35(11):1060–7. PMID: [1473616](https://pubmed.ncbi.nlm.nih.gov/1473616/).
22. Katsika D, Grjibovski A, Einarsson C, Lammert F, Lichtenstein P, Marschall HU. Genetic and environmental influences on symptomatic gallstone disease: a Swedish study of 43,141 twin pairs. *Hepatology*. 2005; 41(5):1138–43. doi: [10.1002/hep.20654](https://doi.org/10.1002/hep.20654) PMID: [15747383](https://pubmed.ncbi.nlm.nih.gov/15747383/).
23. Gatz M, Pedersen NL, Berg S, Johansson B, Johansson K, Mortimer JA, et al. Heritability for Alzheimer's disease: the study of dementia in Swedish twins. *The journals of gerontology Series A, Biological sciences and medical sciences*. 1997; 52(2):M117–25. PMID: [9060980](https://pubmed.ncbi.nlm.nih.gov/9060980/).
24. Sullivan PF, Evengard B, Jacks A, Pedersen NL. Twin analyses of chronic fatigue in a Swedish national sample. *Psychological medicine*. 2005; 35(9):1327–36. doi: [10.1017/S0033291705005222](https://doi.org/10.1017/S0033291705005222) PMID: [16168155](https://pubmed.ncbi.nlm.nih.gov/16168155/).
25. Cameron AJ, Lagergren J, Henriksson C, Nyren O, Locke GR 3rd, Pedersen NL. Gastroesophageal reflux disease in monozygotic and dizygotic twins. *Gastroenterology*. 2002; 122(1):55–9. PMID: [11781280](https://pubmed.ncbi.nlm.nih.gov/11781280/).
26. Bengtson MB, Ronning T, Vatn MH, Harris JR. Irritable bowel syndrome in twins: genes and environment. *Gut*. 2006; 55(12):1754–9. doi: [10.1136/gut.2006.097287](https://doi.org/10.1136/gut.2006.097287) PMID: [17008364](https://pubmed.ncbi.nlm.nih.gov/17008364/); PubMed Central PMCID: [PMC1856463](https://pubmed.ncbi.nlm.nih.gov/1856463/).
27. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med*. 2002; 252(3):247–54. PMID: [12270005](https://pubmed.ncbi.nlm.nih.gov/12270005/).
28. Bak S, Gaist D, Sindrup SH, Skytthe A, Christensen K. Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke; a journal of cerebral circulation*. 2002; 33(3):769–74. PMID: [11872902](https://pubmed.ncbi.nlm.nih.gov/11872902/).
29. Algovik M, Nilsson E, Cnattingius S, Lichtenstein P, Nordenskjold A, Westgren M. Genetic influence on dystocia. *Acta obstetricia et gynecologica Scandinavica*. 2004; 83(9):832–7. doi: [10.1111/j.0001-6349.2004.00544.x](https://doi.org/10.1111/j.0001-6349.2004.00544.x) PMID: [15315594](https://pubmed.ncbi.nlm.nih.gov/15315594/).
30. Altman D, Forsman M, Falconer C, Lichtenstein P. Genetic influence on stress urinary incontinence and pelvic organ prolapse. *European urology*. 2008; 54(4):918–22. doi: [10.1016/j.eururo.2007.12.004](https://doi.org/10.1016/j.eururo.2007.12.004) PMID: [18155350](https://pubmed.ncbi.nlm.nih.gov/18155350/).
31. Mulder EJ, Van Baal C, Gaist D, Kallela M, Kaprio J, Svensson DA, et al. Genetic and environmental influences on migraine: a twin study across six countries. *Twin research: the official journal of the International Society for Twin Studies*. 2003; 6(5):422–31. doi: [10.1375/136905203770326420](https://doi.org/10.1375/136905203770326420) PMID: [14624726](https://pubmed.ncbi.nlm.nih.gov/14624726/).
32. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis and rheumatism*. 2000; 43(1):30–7. doi: [10.1002/1529-0131\(200001\)43:1<30::AID-ANR5>3.0.CO;2-B](https://doi.org/10.1002/1529-0131(200001)43:1<30::AID-ANR5>3.0.CO;2-B) PMID: [10643697](https://pubmed.ncbi.nlm.nih.gov/10643697/).
33. Thomsen SF, van der Sluis S, Kyvik KO, Skytthe A, Backer V. Estimates of asthma heritability in a large twin sample. *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*. 2010; 40(7):1054–61. doi: [10.1111/j.1365-2222.2010.03525.x](https://doi.org/10.1111/j.1365-2222.2010.03525.x) PMID: [20528882](https://pubmed.ncbi.nlm.nih.gov/20528882/).
34. Wirdefeldt K, Gatz M, Reynolds CA, Prescott CA, Pedersen NL. Heritability of Parkinson disease in Swedish twins: a longitudinal study. *Neurobiology of aging*. 2011; 32(10):1923 e1–8. doi: [10.1016/j.neurobiolaging.2011.02.017](https://doi.org/10.1016/j.neurobiolaging.2011.02.017) PMID: [21482443](https://pubmed.ncbi.nlm.nih.gov/21482443/); PubMed Central PMCID: [PMC4452894](https://pubmed.ncbi.nlm.nih.gov/PMC4452894/).

35. Ingebrigtsen T, Thomsen SF, Vestbo J, van der Sluis S, Kyvik KO, Silverman EK, et al. Genetic influences on Chronic Obstructive Pulmonary Disease—a twin study. *Respiratory medicine*. 2010; 104(12):1890–5. doi: [10.1016/j.rmed.2010.05.004](https://doi.org/10.1016/j.rmed.2010.05.004) PMID: [20541380](https://pubmed.ncbi.nlm.nih.gov/20541380/).
36. Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The predictive capacity of personal genome sequencing. *Sci Transl Med*. 2012; 4(133):133ra58. doi: [10.1126/scitranslmed.3003380](https://doi.org/10.1126/scitranslmed.3003380) PMID: [22472521](https://pubmed.ncbi.nlm.nih.gov/22472521/); PubMed Central PMCID: [PMC3741669](https://pmid.ncbi.nlm.nih.gov/PMC3741669/).
37. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Murray CJ, et al. Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to 2010: the Global Burden of Disease 2010 study. *Circulation*. 2014; 129(14):1483–92. doi: [10.1161/CIRCULATIONAHA.113.004042](https://doi.org/10.1161/CIRCULATIONAHA.113.004042) PMID: [24573352](https://pubmed.ncbi.nlm.nih.gov/24573352/); PubMed Central PMCID: [PMC4181359](https://pmid.ncbi.nlm.nih.gov/PMC4181359/).
38. Shibuya K, Mathers CD, Boschi-Pinto C, Lopez AD, Murray CJ. Global and regional estimates of cancer mortality and incidence by site: II. Results for the global burden of disease 2000. *BMC cancer*. 2002; 2:37. PMID: [12502432](https://pubmed.ncbi.nlm.nih.gov/12502432/); PubMed Central PMCID: [PMC149364](https://pmid.ncbi.nlm.nih.gov/PMC149364/).
39. Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*. 2015; 347(6217):78–81. doi: [10.1126/science.1260825](https://doi.org/10.1126/science.1260825) PMID: [25554788](https://pubmed.ncbi.nlm.nih.gov/25554788/); PubMed Central PMCID: [PMCPMC4446723](https://pmid.ncbi.nlm.nih.gov/PMC4446723/).
40. Rappaport SM. Biomarkers intersect with the exposome. *Biomarkers*. 2012; 17(6):483–9. Epub 2012/06/08. doi: [10.3109/1354750X.2012.691553](https://doi.org/10.3109/1354750X.2012.691553) PMID: [22672124](https://pubmed.ncbi.nlm.nih.gov/22672124/).
41. Wild CP, Scalbert A, Herceg Z. Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk. *Environmental and molecular mutagenesis*. 2013; 54(7):480–99. doi: [10.1002/em.21777](https://doi.org/10.1002/em.21777) PMID: [23681765](https://pubmed.ncbi.nlm.nih.gov/23681765/).
42. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011; 472(7341):57–63. Epub 2011/04/09. doi: [nature09922](https://doi.org/10.1038/nature09922) [pii] doi: [10.1038/nature09922](https://doi.org/10.1038/nature09922) PMID: [21475195](https://pubmed.ncbi.nlm.nih.gov/21475195/); PubMed Central PMCID: [PMC3086762](https://pmid.ncbi.nlm.nih.gov/PMC3086762/).
43. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med*. 2013; 368(17):1575–84. Epub 2013/04/26. doi: [10.1056/NEJMoa1109400](https://doi.org/10.1056/NEJMoa1109400) PMID: [23614584](https://pubmed.ncbi.nlm.nih.gov/23614584/).
44. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med*. 2013; 19(5):576–85. Epub 2013/04/09. doi: [nm.3145](https://doi.org/10.1038/nm.3145) [pii] doi: [10.1038/nm.3145](https://doi.org/10.1038/nm.3145) PMID: [23563705](https://pubmed.ncbi.nlm.nih.gov/23563705/); PubMed Central PMCID: [PMC3650111](https://pmid.ncbi.nlm.nih.gov/PMC3650111/).
45. Bae S, Ulrich CM, Neuhouser ML, Malysheva O, Bailey LB, Xiao L, et al. Plasma choline metabolites and colorectal cancer risk in the Women's Health Initiative Observational Study. *Cancer research*. 2014; 74(24):7442–52. doi: [10.1158/0008-5472.CAN-14-1835](https://doi.org/10.1158/0008-5472.CAN-14-1835) PMID: [25336191](https://pubmed.ncbi.nlm.nih.gov/25336191/); PubMed Central PMCID: [PMC4268282](https://pmid.ncbi.nlm.nih.gov/PMC4268282/).