Research Article

Cancer Epidemiology, Biomarkers & Prevention

A Multilevel Model of Postmenopausal Breast Cancer Incidence

Robert A. Hiatt[1], Travis C. Porco[1,2], Fengchen Liu[2], Kaya Balke[1], Allan Balmain[3], Janice Barlow[4], Dejana Braithwaite[1], Ana V. Diez-Roux[5], Lawrence H. Kushi[6], Mark M. Moasser[7], Zena Werb[8], Gayle C. Windham[9], and David H. Rehkopf[10]

Abstract

Background: Breast cancer has a complex etiology that includes genetic, biologic, behavioral, environmental, and social factors. Etiologic factors are frequently studied in isolation with adjustment for confounding, mediating, and moderating effects of other factors. A complex systems model approach may present a more comprehensive picture of the multifactorial etiology of breast cancer.

Methods: We took a transdisciplinary approach with experts from relevant fields to develop a conceptual model of the etiology of postmenopausal breast cancer. The model incorporated evidence of both the strength of association and the quality of the evidence. We operationalized this conceptual model through a mathematical simulation model with a subset of variables, namely, age, race/ethnicity, age at menarche, age at first birth, age at menopause, obesity, alcohol consumption, income, tobacco use, use of hormone therapy (HT), and BRCA1/2 genotype.

Results: In simulating incidence for California in 2000, the separate impact of individual variables was modest, but reduction in HT, increase in the age at menarche, and to a lesser extent reduction in excess BMI >30 kg/m^2 were more substantial.

Conclusions: Complex systems models can yield new insights on the etiologic factors involved in postmenopausal breast cancer. Modification of factors at a population level may only modestly affect risk estimates, while still having an important impact on the absolute number of women affected.

Impact: This novel effort highlighted the complexity of breast cancer etiology, revealed areas of challenge in the methodology of developing complex systems models, and suggested additional areas for further study. *Cancer Epidemiol Biomarkers Prev; 23(10); 2078–92. ©2014 AACR.*

Introduction

Factors at multiple levels of organization ranging from the biologic to societal influence most diseases. Breast cancer is an example of a disease with such a complex etiology and various social, physical, individual, and biologic factors have been identified as being associated with breast cancer incidence (1). But as in many fields of scientific inquiry, these factors have generally been treated as distinct and unrelated to each other. The common tendency to focus on single causal agents obscures the reality of a complex web of causation (2). Individuals in the general public still talk about "the cause" of breast cancer, but the reality is not so simple.

The interaction of causal factors across biologic, social, and environmental determinants is poorly understood, but this understanding is necessary for developing prevention strategies (3). The lack of integration among risk factors at different levels, as well as the failure to consider feedbacks and dependencies undermines a full understanding of the causes of breast cancer and prospects for its prevention. We report here on the development of a multilevel model of the causes of postmenopausal breast cancer incidence.

We had four primary objectives in mind. The first was to create a conceptual model that included relevant and important known and potential causes within major domains, specifically sociocultural, behavioral and lifestyle physical/chemical and biologic domains. A second

[1]Department of Epidemiology and Biostatistics, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California. [2]Francis I. Proctor Foundation, University of California San Francisco, San Francisco, California. [3]Department of Biochemistry and Biophysics, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California. [4]Zero Breast Cancer, San Rafael, California. [5]Department of Epidemiology, University of Michigan, Ann Arbor, Michigan. [6]Division of Research, Kaiser Permanente, Oakland, California. [7]Department of Medicine, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California. [8]Department of Anatomy, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California. [9]Division of Environmental and Occupational Disease Control, California Department of Public Health, Richmond, California. [10]Department of Medicine, Stanford University, Stanford, California.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Robert A. Hiatt, University of California San Francisco, 185 Berry Street, Suite 5700, San Francisco, CA 94107-0560. Phone: 415-514-8113; Fax: 415-514-8150; E-mail: rhiatt@epi.ucsf.edu

doi: 10.1158/1055-9965.EPI-14-0403

objective was to develop a mathematical simulation model that could provide a quantitative framework for exploration of "what if?" scenarios of how interventions could influence long-term breast cancer incidence rates. The use of a model is critical because there are currently no datasets that offer all of the variables included in the complex system we envision.

Third, we wanted this model to be adaptable so that as new scientific evidence accumulates, we want to be able to integrate it into the model. The fourth objective was to create a model that could be used by multiple audiences, including scientists and the general public, to foster understanding of the range of factors that contribute to causing breast cancer.

Materials and Methods

The model was the result of a systematic process and based on input from a range of experts across multiple disciplines. The strategy for development was considered in four parts: (i) the selection of the experts, (ii) a series of meetings with thoughtful, unrushed, and respectful interaction among the experts to develop the components of the model, (iii) refinement of the model based on input from a larger range of potential model users, and (iv) the development and application of a mathematical model that used a selected subset of factors from the full model.

Process for creating the conceptual model

We first convened experts in the fields of genetics (A. Balmain), cell biology (Z. Werb), clinical oncology (M. Moasser), nutrition (L.H. Kushi), environmental health (G.C. Windham), epidemiology (D. Braithwaite, A.V. Diez-Roux, L.H. Kushi, R.A. Hiatt, D.H. Rehkopf, and G.C. Windham), breast cancer advocacy (J. Barlow), complex systems (T.C. Porco and A.V. Diez-Roux), and mathematical modeling (A.V. Diez-Roux and T.C. Porco). Over the course of one year, the group met for three extended meetings to develop the model, with a core group working to implement the group consensus. The first meeting was focused on listening to and reflecting on each individual perspective on the causes of breast cancer incidence, which differed substantially between scientific disciplines. In the second meeting, participants worked to integrate their perspectives into a single model. In the third meeting, participants worked to refine the final draft version of the model that had been constructed in the interim. Finally, the model was circulated to these experts and a wider range of knowledgeable individuals for feedback on both content and format and then a final review by all coauthors. The two end products were an annotated conceptual model and a smaller mathematical model that drew largely modifiable factors from the parent conceptual model.

Deciding on the factors to be included in the conceptual model

As a first step, we created a broad working model by collecting proposed causal risk factors from each panel member. The second step was for each panel member to select the best evidence from published peer-reviewed literature (i.e, 2–3 papers) that supported these causal relationships. The third step was for each panel member to interpret the strength of association and quality of evidence for each factor based on standard criteria (see below). The fourth step was identifying, from the broader literature, how each of the factors was related to each other. The fifth step was interpreting the strength of association and strength of evidence from the literature for connections between factors in the models.

Inclusion and exclusion of factors

To create a practical and useful model, we adopted a number of constraints. We decided that the model would be specific to invasive breast cancer incidence, rather than mortality. This was due to our focus on causes of breast cancer for prevention and not wanting to include factors in diagnosis and treatment that influence mortality. We also chose to focus our model on postmenopausal breast cancer incidence because the causes of breast cancer differ by menopausal status. Roughly 78% of breast cancer in California, our chosen geographic area of study, occurs in women over the age of 50 years (4). Also, more genetic determinants of pre- versus postmenopausal breast cancer have been found, suggesting a relatively greater importance of environmental causes of postmenopausal breast cancer.

We did not include screening as a factor in this model for two reasons. First, our intent was to highlight factors important in the etiology, not early detection of breast cancer. Second, the impact of screening on the incidence of invasive cancer (the outcome of this model) was probably small by 2000 when the rate of screening had reached a steady or slightly decreasing state in California (5). Thus, although it is clear that breast cancer screening will influence incidence rates, our model is meant to include only truly etiologic factors that could be modified through primary prevention.

The model is not specific to particular subtypes of breast cancer [e.g., estrogen receptor (ER) status, basal cell type] because there was not yet sufficient evidence on differential causation by subtype. We also decided that to create a more comprehensive framework, we should include factors even if the types of evidence for causes for these factors differed. For example, for some factors there are neither population-based studies (e.g., inflammation and immune function) nor quantitative data (e.g., ancestry) available. Finally, we could not include all possible factors and decided to limit the number of factors to what the expert committee judged to be major factors in each domain.

Strength of association

To illustrate the strength of association between factors in the model based on the best evidence available, we derived our categories of association based on the range of relative risks of studies of breast cancer. Our three

categories were: category 1, RR > 3.0 (strong); category 2, RR 1.8–3.0 (modest); category 3, RR 1.1–1.8 (weak). These specific categories are based on frequently used levels similar to those used by the Harvard Cancer Risk Index (6) and allow for a reasonably good differentiation of the strength of association for breast cancer-related risk factors.

Quality of the data

The expert panel felt that is was important to express the overall quality of the evidence, even though the type of evidence differed across factors included in our model. This approach allowed factors to be included in the model that were representative of etiologic factors of interest (e.g., environmental factors) but for which little human evidence currently exists.

The U.S. Preventive Services Task force (USPSTF) grades the quality of the overall evidence on a 3-point scale (good, fair, poor): where *good* means "evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes," *fair* means "evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes," and *poor* is appropriate when "evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes" (7). It is difficult to assess the level of evidence that relates a concept like country of birth to other factors in the model, or some of the immunologic evidence that has not been examined in human studies, but generally our three categories—1, strong; 2, moderate; and 3, weak—are analogous to the three USPSTF levels of evidence and should be considered relative measures within the range of factors that we examined.

Mathematical model

We used the structure of the conceptual model as the basis for representing the processes that lead to population level postmenopausal breast cancer incidence rates. For tractability, the simulation model presented here used a subset of 11 factors that are included in the overall conceptual model. These factors were age, self-identified race/ethnicity (classified as "Black," "Hispanic or Latina," "White," or "Other"), age at menarche, age at first birth, age at menopause, obesity, alcohol consumption, income, tobacco use, use of hormone therapy (HT), and BRCA1/2 genotype. These factors were chosen because they include those with the strongest empirical support for a causal role in breast cancer incidence and because they include factors that are generally modifiable (i.e., income, alcohol, tobacco, and HT use). For each predictive factor, we first derived a marginal distribution for each ethnic group using census data, National Health and Nutrition Examination Survey (NHANES) data or other sources. We then derived the ethnic group-specific correlation matrix for these predictive factors from NHANES data. We generated a simulated population matching the racial/ethnic composition and age distribution of California women in 2000, using this correlation structure. We then applied a function to predict the cancer incidence based on these risk factors. All simulations were conducted using R v 3.0 for MacIntosh (www.r-project.org).

This approach allows for the incorporation of results from various types of studies about the relations between variables. Data from published studies can be used to inform various rules (or parameters) encoded in the model. In addition, model output (e.g., predictions) can be contrasted with various types of population data. The distributions of variables can also be based on existing population data. The covariance and distributions of the 11 factors included in our model were calculated from the female population age of 50 years and older in NHANES III, which is a population-based nationally representative source that was surveyed between 1988–1994 (8).

California Department of Finance data for the 2000 census were used to first establish the fraction of the population in each of four race/ethnicity classifications: Black, Hispanic, White, and Other (the latter category includes mixed race, Asian and Pacific Islander, Native American and Native Hawaiian; ref. 9). We then used California Department of Finance demographic data to yield the age structure for women of the age 55 years and older for each ethnic classification based on the year 2000 census. The distribution of income by ethnic group was derived from analysis of the 2000 census (10).

For each of the four ethnic classifications, the marginal distribution of the following variables was derived from NHANES data: HT (HT-percent of the population), body mass index (BMI), alcohol use (g/day), age at menopause, age at menarche, age at first birth more than 30 years, income, and tobacco use (percentage of the population). We approximated these marginal distributions by using a piecewise linear cumulative distribution function whose deciles match the ethnic group-specific marginal distribution of each risk factor based on NHANES.

We used the normal-to-anything transformation to generate samples from the prescribed marginal distributions and the given correlation structure (11). Calibration was conducted to insure convergence with tolerance ≤ 0.01 for each correlation. Simulating cancer incidence from the joint distribution of the predictive factors was conducted by simply multiplying the relative hazards for each risk factor given the other risk factors and age of each individual.

Results

Conceptual model

The factors shown in Fig. 1, supported by the literature shown in Table 1, were the final results of our model building process. We divided selected factors into the four domains of interest: sociocultural, behavioral, physical-

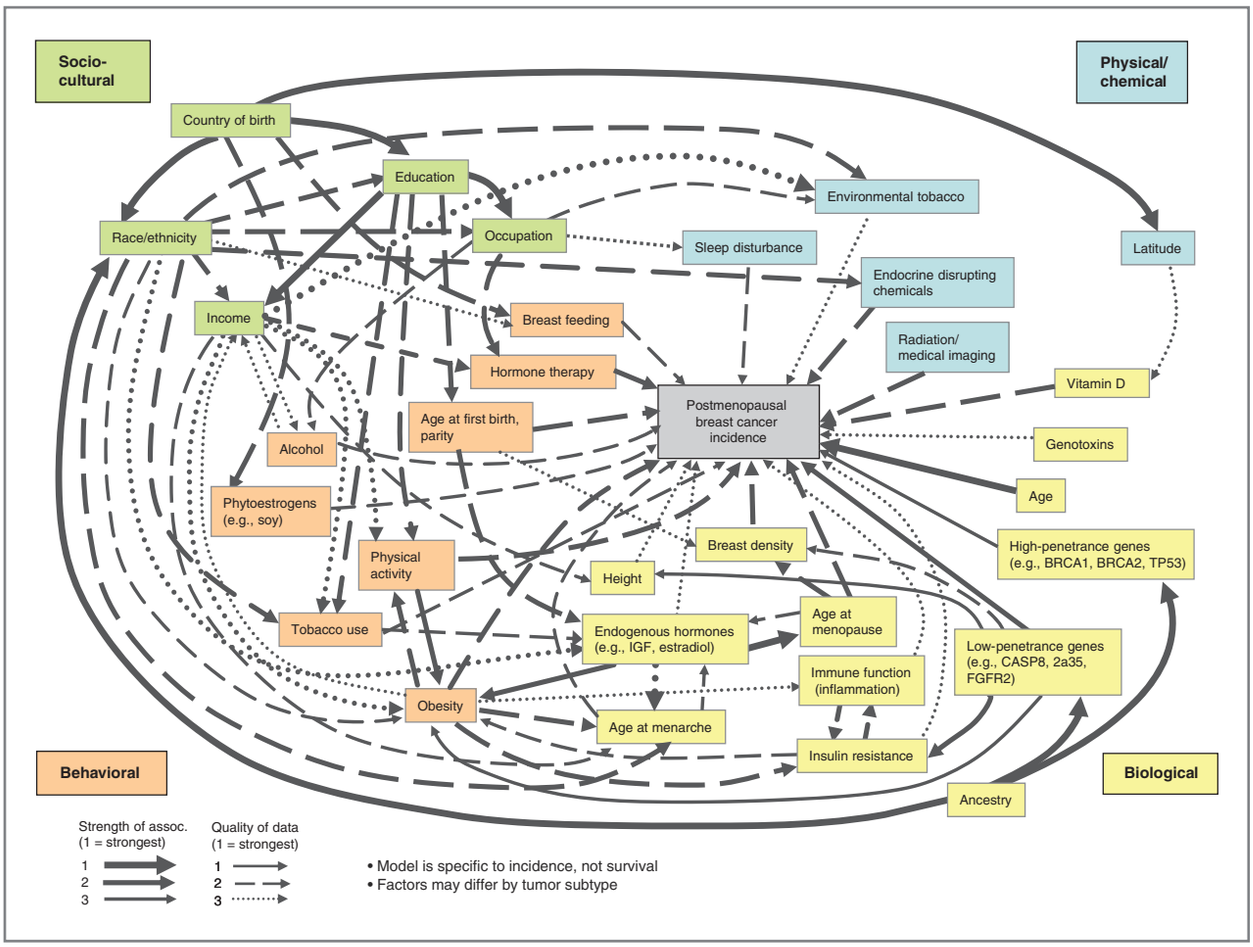

Figure 1. Conceptual model of the etiology of postmenopausal breast cancer. This model depicts factors related to postmenopausal breast cancer from studies in women in four domains: sociocultural, physical-chemical, behavioral, and biologic. The strength of the relationship is depicted by the weight of line of the arrow, with the thickest arrows indicating the strongest associations. The quality of the data is depicted by solid, broken, or dotted arrows, with the highest quality evidence in the solid arrows. References used to provide the evidence for the strength and quality of the associations depicted are given in Table 1. The model is specific to incidence and not mortality from breast cancer. There was insufficient evidence to differentiate between subtypes of breast cancer.

chemical, and biologic. The table provides a list of each of the relationships (arrows) that are in the model, indicating the starting node, the ending node, the rank of strength of association, and the rank of the quality of evidence supporting this arrow along with the supporting citations. The direction of the influence, increasing or decreasing incidence, is not specified by the model itself, but provided by the supporting literature. The references provided are not exhaustive, but meant to provide a source for the evidence that can be explored or updated further by the interested reader.

The complexity of the model (Fig. 1) not only illustrates the many factors thought to have a causal role in breast cancer etiology, but may also lead readers to consider absent or interesting areas for further clarification or research. Although many of the well-documented risk factors for postmenopausal breast cancer incidence show arrows going directly to this outcome in our model, there are also a number of other factors that mediate or moderate their relation to breast cancer. For example, many of the factors in the social environment domain are mediated by behavioral factors that affect postmenopausal breast cancer incidence. There are also a few instances where arrows go in both directions between factors in the model, indicating that there may be bidirectional influences. Although we sought to be as inclusive as possible in documented interacting connections between factors, two were left out of the picture for clarity of presentation: age and country of birth. Both of these factors influence the majority of other factors in our model, but the addition of this many arrows would obscure the other connections presented.

The nature of the four domains is described in Table 2 and they should best be considered to be heuristic categories, rather than absolute. Indeed, one of our primary motivations for the development of this model was that factors are interdependent and may not neatly fit into one category.

Table 1. Literature used to support the conceptual model

Arrow start	Arrow end	Strength	Quality	Reference
Socio/cultural				
Country of birth	Breast feeding	2	2	(12, 13)
Country of birth	Education	1	1	(14, 15)
Country of birth	Latitude	1	1	(16)
Country of birth	Phytoestrogens	2	2	(17, 18)
Country of birth	Race/ethnicity	1	1	(19, 20)
Education	Age at first birth, parity	2	2	(21)
Education	Income	1	1	(22, 23)
Education	Occupation	1	1	(24, 25)
Education	Physical activity	2	2	(26)
Education	Tobacco use	2	2	(27, 28)
Income	Age at menarche	3	2	(29)
Income	Alcohol	3	3	(30, 31)
Income	Environmental tobacco	1	3	(32, 33)
Income	HT	2	2	(34)
Income	Obesity	2	3	(35, 36)
Income	Physical activity	2	3	(26, 37)
Income	Tobacco use	2	3	(38)
Income	Height	3	2	(39, 40)
Occupation	Alcohol	3	2	(41)
Occupation	Environmental tobacco	3	2	(42)
Occupation	HT	2	2	(34)
Occupation	Sleep disturbance	3	3	(43)
Race/ethnicity	Age at menarche	2	2	(29)
Race/ethnicity	Breast feeding	3	3	(13, 44)
Race/ethnicity	Education	2	2	(45, 46)
Race/ethnicity	Endocrine disrupting chemicals	2	2	(47, 48)
Race/ethnicity	Endogenous hormones	2	3	(49–51)
Race/ethnicity	Environmental tobacco	2	2	(52, 53)
Race/ethnicity	Income	2	2	(46)
Race/ethnicity	Obesity	3	2	(35)
Race/ethnicity	Occupation	2	2	(46, 54)
Race/ethnicity	Tobacco use	2	2	(27)
Physical				
Endocrine disrupting chemicals	Incidence	2	2	(55, 56)
Environmental tobacco	Incidence	3	3	(42, 57)
Latitude	Vitamin D	3	3	(58, 59)
Radiation	Incidence	2	2	(60, 61)
Sleep disturbance	Incidence	3	2	(62, 63)
Behavioral				
Age at first birth, parity	Breast density	3	3	(64)
Age at first birth, parity	Endogenous hormones	2	2	(65)
Age at first birth, parity	Incidence	2	2	(66, 67)
Alcohol	Incidence	3	2	(68–71)
Alcohol	Income	3	3	(72)
Breast feeding	Incidence	3	2	(73)
HT	Incidence	2	1	(74, 75)
Obesity	Age at menarche	2	2	(76)
Obesity	Immune function	3	3	(77, 78)
Obesity	Incidence	2	2	(68, 69, 79, 80)
Obesity	Income	3	3	(36, 81)
Obesity	Insulin resistance	2	2	(82)

(Continued on the following page)

Table 1. Literature used to support the conceptual model (Cont'd)

Arrow start	Arrow end	Strength	Quality	Reference
Obesity	Physical activity	2	2	(83)
Physical activity	Incidence	2	2	(68, 69)
Physical activity	Obesity	2	1	(84, 85)
Phytoestrogens	Incidence	3	2	(86–88)
Tobacco use	Endogenous hormones	3	2	(89)
Tobacco use	Incidence	3	2	(90, 91)
Biologic				
Age	Incidence	1	1	(66)
Age at menarche	Endogenous hormones	3	2	(65, 66)
Age at menarche	Incidence	3	2	(66, 92)
Age at menopause	Breast density	2	2	(93, 94)
Age at menopause	Endogenous hormones	3	2	(65)
Age at menopause	Incidence	2	2	(66)
Ancestry	High-penetrance genes	1	1	(95–98)
Ancestry	Low-penetrance genes	1	1	(99–104)
Ancestry	Race/ethnicity	1	1	(19, 105)
Breast density	Incidence	2	2	(106–109)
Endogenous hormones	Age at menarche	1	3	(110–112)
Endogenous hormones	Age at menopause	1	1	(113)
Endogenous hormones	Obesity	2	1	(113)
Endogenous hormones	Incidence	3	3	(65)
Genotoxins	Incidence	3	3	(56, 114)
Height	Incidence	3	3	(115, 116)
High-penetrance genes	Incidence	3	1	(100, 117, 118)
Immune function	Incidence	3	3	(119)
Immune function	Insulin resistance	2	2	(120, 121)
Insulin resistance	Immune function	2	2	(120, 121)
Insulin resistance	Incidence	3	3	(122, 123)
Insulin resistance	Obesity	3	2	(124, 125)
Low-penetrance genes	Breast density	3	2	(126, 127)
Low-penetrance genes	Height	3	1	(128, 129)
Low-penetrance genes	Incidence	2	1	(99, 100, 102, 103, 130, 131)
Low-penetrance genes	Insulin resistance	2	1	(132, 133)
Low-penetrance genes	Obesity	3	1	(134–136)
Vitamin D	Incidence	2	2	(137, 138)

NOTE: The starting point for an arrow is the independent variable and the ending point is the dependent variable. Strength of the relationship is categorized as (1) strong (RR > 3.0), (2) modest (RR >1.8–3.0), or (3) weak (1.1–1.8). Quality of the evidence is categorized as (1) strong, (2) moderate, or (3) weak.

Mathematical model

Figure 2 shows the relationships between the subset of 11 factors that are included in the mathematical model. We applied relative hazard estimates for each of these factors, derived from systematic reviews and high-quality studies (66, 67, 70, 71, 74, 75, 79, 80, 91, 92, 100, 117, 118, 139), together with race/ethnic group-specific estimates of the joint distribution of these factors derived from NHANES III, to the population of California excluding sparsely populated areas (Table 3). We then applied a base rate of incidence of breast cancer to the lowest risk group and simulated the progression of cancer in each ethnic group over time. Our results produce incidence rates that closely approximate the observed rates in California in the time period under consideration (i.e., ~2000). We also reproduce the differences observed for self-reported race and ethnicity. Established "nonmodifiable" risk factors affect risks in the direction and with the magnitude expected.

Table 4 presents the estimated impact of changes at a population level for five of these selected factors considered one at a time in categories of age and race/ethnicity, but adjusting for the effects of the other variables. Overall, the impact of interventions or change in modifiable risk factors is modest. However, the model suggests that the greatest impact in 2000 would be from reductions in HT

Table 2. Explanations of interrelated domains

Domain	Description
Physical-chemical environment	Factors that fit most with the physical world, both the natural world and the physical world as created by humans. Factors in this category range from the latitude of residence, to physical and chemical environmental hazards that are located near work and housing. In our conceptual model, most of these factors are influenced at least in part by the social environment.
Sociocultural environment	Social and cultural factors that are dependent on geography, social networks, and social class that affect access to resources and networks. Factors in this category range from country of birth to occupation.
Individual behavioral	Individual choices that are made on diet, physical activity, smoking, HT, and child bearing. In our conceptual model, most of these factors are influenced by the social environment and in turn have effects on biologic pathways.
Biologic	Biologic states of individuals that range from morphologic (e.g., breast density) to genetic (e.g., BRCA1). In our conceptual model, many of these factors are impacted by individual behaviors and the social and physical environment.

use and an increase in the age at menarche followed by reductions in excess BMI. Reductions in alcohol or tobacco use have only small effects as would be expected from the levels of attributable risk estimated from prospective cohort studies (70, 90). Interestingly, if the age at menarche, which has been dropping over the last century (144, 145), were to be reversed (by a year or 18 months), breast cancer incidence would be reduced by 5.5% overall.

Incidence rates by the three major race/ethnic categories were, as expected, highest in Whites, lowest in Latinas, and intermediate in Blacks. There were no substantial differences to the influences of change in any of the risk factors by race/ethnic category.

We assessed the absolute risk reductions associated with modifications in risk factors and breast cancer incidence in this model. A 50% reduction in the proportion of

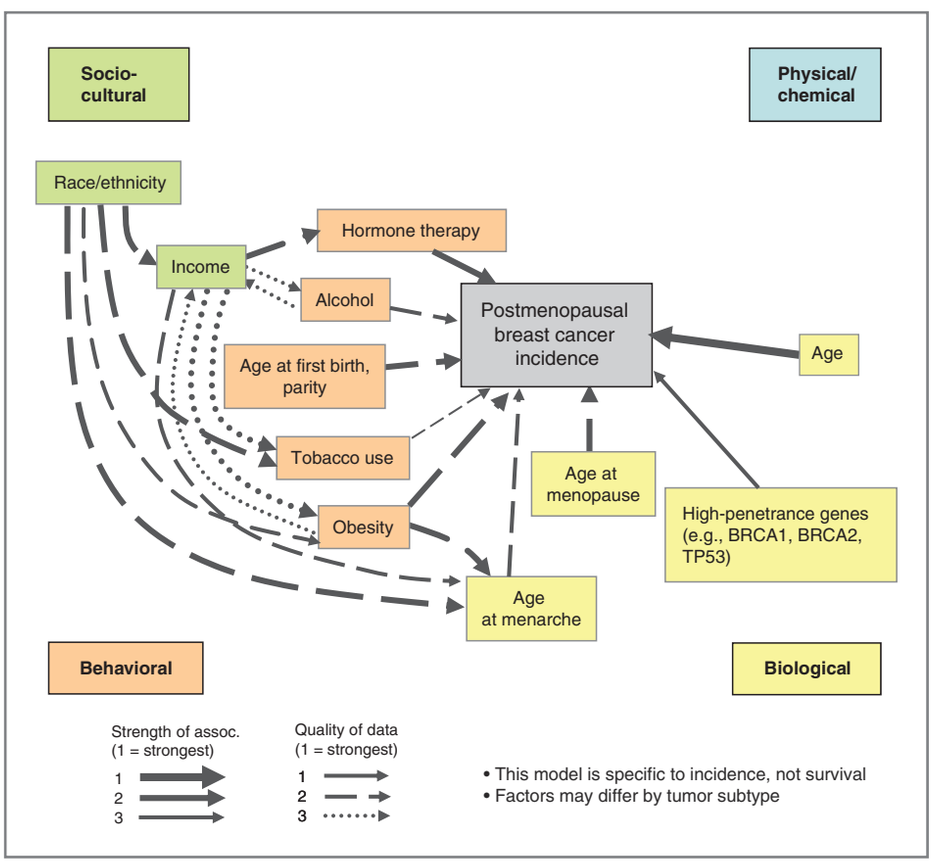

Figure 2. Etiologic factors included in mathematical model. Selected elements from conceptual model (Fig. 1) that were used in a mathematical model demonstrating the magnitude of effects from each element while adjusting for all the other elements shown. The strength of the relationship is depicted by the weight of line of the arrow, with the thickest arrows indicating the strongest associations. The quality of the data is depicted by solid, broken, or dotted arrows, with the highest quality evidence in the solid arrows (see the text for results).

Table 3. Predictive factors for the incidence model of breast cancer are shown

Name	Marginal distribution (source)	Meta-analysis estimated relative hazard and CI		Meta-analysis reference (if available and applicable)
Race-age (y)	California census[a]	See Supplementary Table S1 for matrix of risk estimates.		N/A
Age at first birth (years, all among parous women)	NHANES[a]	Category	RR	(140)
		<20 y	1.00	
		20–24	1.04 (0.91–1.18)	
		25–29	1.17 (1.02–1.34)	
		30–34	1.30 (1.11–1.51)	
		35+	1.40 (1.15–1.70)	
Age at menarche (y)	NHANES[a]	Age	RR	(139)
		<11	1.19 (1.13–1.25)	
		11	1.09 (1.06–1.12)	
		12	1.07 (1.05–1.09)	
		13	1.00 (0.98–1.02)	
		14	0.98 (0.96–1.00)	
		15	0.92 (0.89–0.95)	
		≥16	0.82 (0.79–0.85)	
Age at menopause (y)	NHANES[a]	Age	RR	(139)
		<40	0.67 (0.62–0.73)	
		40–44	0.73 (0.70–0.77)	
		45–49	0.86 (0.84–0.89)	
		50–54	1.00 (0.98–1.02)	
		≥55	1.12 (1.07–1.17)	
Alcohol use (g/day)	NHANES[a]	g/d	RR (SE)	(141)
		0	1.00 (0.015)	
		<5	1.01 (0.020)	
		5–14	1.01 (0.023)	
		15–24	1.19 (0.048)	
		25–34	1.22 (0.056)	
		35–44	1.18 (0.093)	
		≥45	1.49 (0.110)	
		OR increase in relative risk per 10 g/day: 7.1% (1.3%)		
BRCA[b]	0.1%	5.0 (4–6) estimated		
Race/ethnicity	California census			N/A
HT	NHANES[a]	Current combined vs. never use: 2.14 (2.04–2.24) Current estrogen-only vs. never use: 1.32 (1.24–1.39)		(142)
Obesity	NHANES[a]	Risk ratio per 5 U of BMI for postmenopausal breast 1.12 (1.08–1.16)		(143)

(Continued on the following page)

Table 3. Predictive factors for the incidence model of breast cancer are shown (Cont'd)

Name	Marginal distribution (source)	Meta-analysis estimated relative hazard and CI	Meta-analysis reference (if available and applicable)
Tobacco use	NHANES[a]	Relative risk of ever vs. never smokers, for those who went through natural menopause at: Age <45 y RR = 1.11 SE = 0.15 Age 45–49 RR = 0.98 SE = 0.08 Age ≥50 years RR = 1.12 SE = 0.06	(141)

[a]Race/ethnic group–specific marginal distributions were derived from NHANES data; see Table 2 for details.
[b]BRCA was assumed to have a population frequency of 0.1% in all race/ethnic groups independent of all other risk factors.

the population with a BMI ≥30 kg/m^2 would result in a decrease in the simulated rate of invasive breast cancer from 393/100,000 (no intervention) to 384.4/100,000 for an estimated total of 386 fewer cases of breast cancer per year in the state of California in 2010 with a census population of 4,486,843 women over 55 years of age. A reduction in the use of HT by 50%, which in 2000 was still very common, resulted in a rate of 288.3/100,000 or 4,697 fewer breast cancer cases.

Discussion

Our conceptual model (Fig. 1) is a graphical depiction of the relationships between factors that play a role in determining the etiology of postmenopausal breast cancer incidence. This model indicates a wide range of factors at different levels of organization that impact each other to influence breast cancer incidence rates. The findings from the mathematical model begin to illustrate how accounting for multiple factors can be used to integrate

Table 4. Rates of invasive postmenopausal breast cancer incidence with SDs by age category and race/ethnicity for risk factors in mathematical model for women ≥ 55 years of age and estimated impact of a change (degree change) at the population level of selected modifiable risk factors on incidence per 100,000 women by age at diagnosis and race/ethnic group, California, 2009

Predictive factor	Degree change	Total (White, Black, Latino)	SD	55–64 y	SD	65–74 y	SD	75+ y	SD	White	SD	Black	SD	Latina	SD
Total observed		379		314		451		423		430		379		254	
Total simulated		393.0	0.8	306.9	1.3	452.7	1.8	450.9	1.6	431.7	1.1	364.1	2.5	245.4	1.4
Excess BMI	50% decrease	384.4	0.8	300.4	1.2	442.6	1.7	440.9	1.5	423.3	1.0	349.4	2.4	238.3	1.4
	100% decrease	375.8	0.8	293.8	1.2	432.5	1.6	430.9	1.5	414.9	1.0	334.7	2.2	231.2	1.3
Alcohol consumption	25% decrease	391.9	0.8	305.9	1.3	451.4	1.7	449.7	1.6	430.5	1.0	363.5	2.5	244.3	1.4
	50% decrease	389.5	0.8	303.9	1.3	448.7	1.7	447.1	1.6	427.5	1.0	362.5	2.5	243.6	1.4
Tobacco use: % of population	25% decrease	392.0	0.8	305.8	1.3	451.5	1.8	450.1	1.6	430.5	1.1	362.9	2.5	244.8	1.4
	50% decrease	390.9	0.8	304.6	1.3	450.3	1.7	449.4	1.6	429.3	1.1	361.8	2.5	244.3	1.4
Age at menarche	1 y increase	377.4	0.8	294.3	1.2	434.5	1.7	433.5	1.5	415.3	1.0	346.9	2.4	233.5	1.4
	1.5 y increase	371.7	0.8	289.8	1.2	428.0	1.7	427.0	1.5	409.1	1.0	341.4	2.4	229.8	1.3
HT: % of population	50% decrease	288.3	0.7	225.2	1.1	332.1	1.6	330.7	1.4	316.7	1.0	267.1	2.3	180.0	1.3
	100% decrease	183.7	0.4	143.4	0.6	211.5	0.8	210.7	0.7	201.7	0.5	170.1	1.2	114.7	0.7

NOTE: Rates were simulated from 100,000 persons with 800 iterations, and were age adjusted to the 2,000 U.S. Standard Population (19 age groups - Census P25-1130: http://www.census.gov/prod/1/pop/p25-1130/p251130.pdf). The simulated incidence rates were from one parameter set using the average value in Table 3.

various types of evidence and to estimate the impact of various types of manipulations or interventions. Importantly, the exercise of specifying these models highlights gaps in our understanding and suggests new areas for data collection (e.g., endocrine disrupting chemicals, medical radiation exposure).

Existing models for breast cancer causation have been very successful in some ways, but not in others. For the past three decades, one of the most recognized and well-used conceptual framework of cancer causation overall comes from the analysis of Doll and Peto in 1981, frequently referred to as a model, which divides up the proportionate domain of causes for contributing to cancer (e.g., environment, diet; ref. 146). This approach has been more recently updated for the United Kingdom (147) and for the United States (148). These perspectives have been helpful in identifying which factors might be further investigated and that there are multiple causes. However, these causal factors have been viewed as relatively separate and mutually exclusive. For example, the impact of the environment has been regarded as separate from impact of individual level behavior although these domains frequently interact in cancer causation.

There are also a number of specific prediction models of cancer; the most well recognized and used for breast cancer is the Gail model (149). This model has proved to be useful for clinical applications in predicting individual future risk of breast cancer, based on the consideration of a number of biologic and behavioral factors, including the number of biopsies taken.

Mathematical models have also been used to explore the role of risk factors in breast cancer. For example, the Pike model has been used to explain the age–incidence curve in breast cancer (150). Modification of this model based on parameters from cohort data has resulted in further refinements and accuracy focused on individual level reproductive risk factors. Rosner and Colditz extended the Pike model using a log-incidence model to show the effects of reproductive risk factors on breast cancer incidence using data from the Nurses Health Study (151). Our model builds on this prior work by considering potentially causal variables in multiple domains and levels potentially amenable to intervention.

More recently, there has been more interest in complex systems modeling in health-related work. A notable example is the Foresight model, a conceptual model that depicts the causes of obesity that was developed by an expert commission in the United Kingdom (152). The development of this model was not constrained by parsimony, and the model includes more than 100 causal factors. Another model developed by Galea and colleagues depicts how a complex system can shed light on how measures of social class and neighborhood can influence health and disease across time (153). Auchincloss and colleagues have developed agent-based models to examine neighborhood influences on diet and walking (154).

Our model, which focuses only on postmenopausal breast cancer, has a number of strengths. First, it allows for the consideration of multiple levels of causes across multiple domains that are not typically a part of particular research domains (e.g., country of origin, sleep disturbance). Second, it is designed to focus consideration on the prevention of breast cancer, as compared with prior models that were more focused on treatment. Third, it is adaptable as new evidence comes to light. Finally, it can help highlight gaps where, because there is no data or weak evidence, further exploration and research are needed.

There are also a number of limitations and self-imposed constraints in our model. First, many of the relationships (i.e., arrows) in our conceptual model relied on evidence that may not estimate true causes, but instead be the result of confounding or bias. Another limitation is that although there are more risk factors that could be included (e.g., metals, mother's age at menarche), it was the view of the expert committee that inclusion of more risk factors with a lower level of supporting evidence would diminish the interpretability of the model. We also understand that risk factors may be different for pre- and postmenopausal breast cancer, by ER status and by breast cancer subtype (e.g., basal cell type). Future modifications of the model could expand to include subtypes as more information on risk factors emerges. Finally, breast cancer is a neoplasm with risk factors that come into play all along the life course (155). Although we included breast feeding, age at menarche, height, age at first birth, and age at menopause, we intend to explore better ways to display life course etiologic factors in future versions of the conceptual model.

There were also a number of limitations to the implementation of the mathematical model. First, at this stage of development, we did not include feedback in this model (e.g., the effect of alcohol consumption on income) although this is a characteristic of complexity models and we plan to consider this for subsequent versions. We also assumed linear relationships between factors in the models, which is almost certainly not correct. Unfortunately, very little data exist on more detailed specifications of relationships between variables and we did not include interactions between factors in this model. Finally, there are interdependencies between individuals, such as have been demonstrated for obesity and tobacco use in studies of social networks (156, 157), which will require more development in the future. Feedbacks and dependencies are two hallmarks of complex systems and are undoubtedly present in the system that results in breast cancer.

Our conceptual model and the complex system it depicts can be used in a number of ways by researchers, policymakers, and the public. (158) Researchers may be aided in seeing the "big picture" of breast cancer causation and where new transdisciplinary research is needed. This model should help make the connections across domains and areas of disciplinary expertise. It also highlights gaps in knowledge for further research such as a better

understanding of the role of environmental toxicants and their interactions with social factors, biologic pathways for environmental and behavioral factors, and strengthening evidence where the quality of data is weak. For policymakers, the model can help identify the multiple avenues for primary prevention and the need for resources and funding in specific areas. It may make clearer the need for considering multiple domains for interventions within their jurisdictions and areas of influence. The public, breast cancer advocates, and other lay stakeholders can use the model to understand the complexities in breast cancer causation (i.e., there is not "a cause" of breast cancer), and that these causes extend beyond genetic susceptibility, traditional reproductive and lifestyle risk factors, and potential environmental carcinogens.

In conclusion, we hope the growing understanding of the etiology of breast cancer will allow us to modify the relationships and pathways in this newly developed model and make it a continuously updated resource for scientists and the public. The model can and should be modified with the results of new knowledge and research. Equally important, although our model is focused on the causes of postmenopausal breast cancer, this approach can serve as a framework to be applied to other diseases and outcomes. By doing so, it can lead to better appreciation of the complex interplay of causal factors at multiple organizational levels, while simultaneously providing greater clarity on avenues for prevention. Complex systems approaches are increasingly called for in population health research yet specific applications to research questions beyond infectious disease remain rare. We have developed a conceptual model of breast cancer and operationalized it through a mathematical model. The advantages of this approach are illustrated in terms of (i) formulating dynamic conceptual models of breast cancer etiology that are explicit and can be understood and debated by various stakeholders, (ii) integrating various types of data, (iii) identifying the possible impact of various interventions and policies, and (iv) identifying gaps in knowledge where new data are needed.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: R.A. Hiatt, T.C. Porco, J. Barlow, A.V. Diaz-Roux, L.H. Kushi, M.M. Moasser, D.H. Rehkopf
Development of methodology: R.A. Hiatt, T.C. Porco, J. Barlow, L.H. Kushi, Z. Werb, G.C. Windham, D.H. Rehkopf
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R.A. Hiatt, L.H. Kushi, Z. Werb, G.C. Windham, D.H. Rehkopf
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): R.A. Hiatt, T.C. Porco, F. Liu, D. Braithwaite, L.H. Kushi, M.M. Moasser, Z. Werb, D.H. Rehkopf
Writing, review, and/or revision of the manuscript: R.A. Hiatt, A. Balmain, J. Barlow, D. Braithwaite, A.V. Diaz-Roux, L.H. Kushi, M.M. Moasser, Z. Werb, G.C. Windham, D.H. Rehkopf, K. Balke
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R.A. Hiatt, K. Balke, D.H. Rehkopf
Study supervision: R.A. Hiatt, D.H. Rehkopf

Acknowledgments

The authors thank Elad Ziv, Peggy Reynolds, Laura van t Veer, and Jeanne Rizzo for their comments on an early draft of the article.

Grant Support

This work was supported by the California Breast Cancer Research Program (15QB-8301; to R.A Hiatt, T.C. Porco, K. Balke, A. Balmain, J. Barlow, A.V. Diez-Roux, L.H. Kushi, M.M. Moasser, Z. Werb, G.C. Windham, and D.H. Rehkopf).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 14, 2014; revised June 10, 2014; accepted July 1, 2014; published OnlineFirst July 13, 2014.

References

1. Kelsey JL, Bernstein L. Epidemiology and prevention of breast cancer. Annu Rev Public Health 1996;17:47–67.
2. Krieger N. Epidemiology and the web of causation: has anyone seen the spider? Soc Sci Med 1994;39:887–903.
3. Colditz GA, Wei EK. Preventability of cancer: the relative contributions of biologic and social and physical environmental determinants of cancer mortality. Annu Rev Public Health 2012;33:137–56.
4. California Department of Public Health, Cancer Surveillance and Research Branch,. California Cancer Registry, Cancer Inquiry System, 2005–2009. 2011 [cited 2012 10/28]. Available from: www.cancer-rates.info/ca/index.php.
5. National Cancer Institute, NIH, DHHS,. Cancer Trends Progress Report – 2011/2012 Update. Bethesda, MD; 2012.
6. Colditz GA, Atwood KA, Emmons K, Monson RR, Willett WC, Trichopoulos D, et al. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. Risk Index Working Group, Harvard Center for Cancer Prevention. Cancer Causes Control 2000;11:477–88.
7. U.S. Preventive Services Task Force. Grade Definitions 2008. Available from: http://www.uspreventiveservicestaskforce.org/uspstf/grades.htm.
8. Third National Health and Nutrition Examination Survey (NHANES III), 1988–94. Catalog Number 76200. NHANES III Examination Data File Documentation. Ages two months and older. December 1996.
9. California Department of Finance DRU. 2013 [cited 2012 March 1]. Available from: http://www.dof.ca.gov/research/demographic/.
10. Lopez A. Race and Income in California: Census 2000 Profiles. Report. Stanford, CA: Stanford University, Ethnicity CfCSiRa; 2003 13.
11. Cario MC, Nelson BL. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical Report. Evanston, IL: Northwestern University, Sciences DoIEaM; 1997.
12. Dettwyler KA. Breastfeeding: Biocultural Perspectives.Stuart-Macadam P, Dettwyler KA, editors. Hawthorne, NY: Aldine de Gruyter; 1995.
13. Dettwyler KA. When to wean: biological versus cultural perspectives. Clin Obstet Gynecol 2004;47:712–23.
14. Chiswick BR, DebBurman N. Educational attainment: analysis by immigrant generation. Econ Educ Rev 2004;23:361–79.
15. Betts JR, Lofstrom M. The educational attainment of immigrants: trends and implications. NBER Work Pap Ser 1998:45:38.
16. Goode JP. Goode's World Atlas. Chicago: Rand McNally; 1978.

17. Morton MS, Arisaka O, Miyake N, Morgan LD, Evans BA. Phytoestrogen concentrations in serum from Japanese men and women over forty years of age. J Nutr 2002;132:3168–71.
18. Wu AH, Ziegler RG, Horn-Ross PL, Nomura AM, West DW, Kolonel LN, et al. Tofu and risk of breast cancer in Asian-Americans. Cancer Epidemiol Biomarkers Prev 1996;5:901–6.
19. Bhopal R, Donaldson L. White, European, Western, Caucasian, or what? Inappropriate labeling in research on race, ethnicity, and health. Am J Public Health 1998;88:1303–7.
20. Krieger N. Refiguring "race": epidemiology, racialized biology, and biological expressions of race relations. Int J Health Serv 2000;30: 211–6.
21. Rindfuss RR, Morgan SP, Offutt K. Education and the changing age pattern of American fertility: 1963–1989. Demography 1996;33: 277–90.
22. Card D. The causal effect of education on earnings. In:Ashenfelter O, Card D, editors. Handbook of Labor Economics. 3A. Amsterdam: Elsevier North Holland; 1999.
23. Psacharopoulos G, Patrinos HA. Returns to investment in education: a further update. Educ Econ 2004;12:111–34.
24. Sewell WH, Hauser R. Education, Occupation and Earnings: Achievement in the Early Career. New York: Academic Press; 1975.
25. Kitagawa EM, Hauser R. Differential mortality in the United States: A study in socioeconomic epidemiology. Cambridge: Harvard Press; 1973.
26. Ham SA, Ainsworth BE. Disparities in data on Healthy People 2010 physical activity objectives collected by accelerometry and self-report. Am J Public Health 2010;100 Suppl 1:S263–8.
27. Escobedo LG, Anda RF, Smith PF, Remington PL, Mast EE. Sociodemographic characteristics of cigarette smoking initiation in the United States. Implications for smoking prevention policy. JAMA 1990;264:1550–5.
28. Escobedo LG, Peddicord JP. Smoking prevalence in US birth cohorts: the influence of gender and education. Am J Public Health 1996;86:231–6.
29. Braithwaite D, Moore DH, Lustig RH, Epel ES, Ong KK, Rehkopf DH, et al. Socioeconomic status in relation to early menarche among black and white girls. Cancer Causes Control 2009;20:713–20.
30. Dawson DA, Grant BF, Chou SP, Pickering RP. Subgroup variation in U.S. drinking patterns: results of the 1992 national longitudinal alcohol epidemiologic study. J Subst Abuse 1995;7:331–44.
31. Babor TF, Mendelson JH, Greenberg I, Kuehnle J. Experimental analysis of the "happy hour": effects of purchase price on alcohol consumption. Psychopharmacology (Berl) 1978;58:35–41.
32. Singh GK, Siahpush M, Kogan MD. Disparities in children's exposure to environmental tobacco smoke in the United States, 2007. Pediatrics 2010;126:4–13.
33. Bolte G, Fromme H. Socioeconomic determinants of children's environmental tobacco smoke exposure and family's home smoking policy. Eur J Public Health 2009;19:52–8.
34. Marks NF, Shinberg DS. Socioeconomic status differences in hormone therapy. Am J Epidemiol 1998;148:581–93.
35. Wang Y, Beydoun MA. The obesity epidemic in the United States–gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. Epidemiol Rev 2007;29:6–28.
36. Cawley J. The impact of obesity on wages. J Hum Resour 2004;2:451–74.
37. Marshall SJ, Jones DA, Ainsworth BE, Reis JP, Levy SS, Macera CA. Race/ethnicity, social class, and leisure-time physical inactivity. Med Sci Sports Exerc 2007;39:44–51.
38. Farrelly MC, Nonnemaker JM, Watson KA. The consequences of high cigarette excise taxes for low-income smokers. PLoS ONE 2012;7: e43838.
39. Yip R, Scanlon K, Trowbridge F. Trends and patterns in height and weight status of low-income U.S. children. Crit Rev Food Sci Nutr 1993;33:409–21.
40. Komlos J, Breitfelder A. Height of US-born non-Hispanic children and adolescents ages 2–19, born 1942–2002 in the NHANES samples. Am J Hum Biol 2008;20:66–71.
41. Diala CC, Muntaner C, Walrath C. Gender, occupational, and socioeconomic correlates of alcohol and drug abuse among U.S. rural, metropolitan, and urban residents. Am J Drug Alcohol Abuse 2004;30:409–28.
42. CDC, Office on Smoking and Health. The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General. Atlanta GA: US DHHS; 2006.
43. Virtanen M, Ferrie JE, Gimeno D, Vahtera J, Elovainio M, Singh-Manoux A, et al. Long working hours and sleep disturbances: the Whitehall II prospective cohort study. Sleep 2009;32:737–45.
44. Dettwyler KA. Beauty and the Breast: The Cultural Context of Breastfeeding in the United States. In:Stuart-Macadam P, Dettwyler KA, editors. Breastfeeding: Biocultural Perspectives. Hawthorne, NY: Aldine de Gruyter; 1995.
45. Williams DR, Mohammed SA, Leavell J, Collins C. Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. Ann N Y Acad Sci 2010;1186:69–101.
46. LaVeist TA. Disentangling race and socioeconomic status: a key to understanding health inequalities. J Urban Health 2005;82:iii26–34.
47. James RA, Hertz-Picciotto I, Willman E, Keller JA, Charles MJ. Determinants of serum polychlorinated biphenyls and organochlorine pesticides measured in women from the child health and development study cohort, 1963–1967. Environ Health Perspect 2002; 110:617–24.
48. Windham GC, Pinney SM, Sjodin A, Lum R, Jones RS, Needham LL, et al. Body burdens of brominated flame retardants and other persistent organo-halogenated compounds and their descriptors in US girls. Environ Res 2010;110:251–7.
49. Setiawan VW, Haiman CA, Stanczyk FZ, Le Marchand L, Henderson BE. Racial/ethnic differences in postmenopausal endogenous hormones: the multiethnic cohort study. Cancer Epidemiol Biomarkers Prev 2006;15:1849–55.
50. Pinheiro SP, Holmes MD, Pollak MN, Barbieri RL, Hankinson SE. Racial differences in premenopausal endogenous hormones. Cancer Epidemiol Biomarkers Prev 2005;14:2147–53.
51. Woolcott CG, Shvetsov YB, Stanczyk FZ, Wilkens LR, White KK, Caberto C, et al. Plasma sex hormone concentrations and breast cancer risk in an ethnically diverse population of postmenopausal women: the Multiethnic Cohort Study. Endocr Relat Cancer 2010;17:125–34.
52. Iribarren C, Friedman GD, Klatsky AL, Eisner MD. Exposure to environmental tobacco smoke: association with personal characteristics and self reported health conditions. J Epidemiol Community Health 2001;55:721–8.
53. Pirkle JL, Flegal KM, Bernert JT, Brody DJ, Etzel RA, Maurer KR. Exposure of the US population to environmental tobacco smoke: the Third National Health and Nutrition Examination Survey, 1988 to 1991. JAMA 1996;275:1233–40.
54. Williams DR. Race/ethnicity and socioeconomic status: measurement and methodological issues. Int J Health Serv 1996;26:483–505.
55. Aronson KJ, Miller AB, Woolcott CG, Sterns EE, McCready DR, Lickley LA, et al. Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2000;9:55–63.
56. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. Environmental pollutants and breast cancer: epidemiologic studies. Cancer 2007;109:2667–711.
57. Miller MD, Marty MA, Broadwin R, Johnson KC, Salmon AG, Winder B, et al. The association between exposure to environmental tobacco smoke and breast cancer: a review by the California Environmental Protection Agency. Prev Med 2007;44:93–106.
58. Hanley DA, Davison KS. Vitamin D insufficiency in North America. J Nutr 2005;135:332–7.
59. Hagenau T, Vest R, Gissel TN, Poulsen CS, Erlandsen M, Mosekilde L, et al. Global vitamin D levels in relation to age, gender, skin pigmentation and latitude: an ecologic meta-regression analysis. Osteoporos Int 2009;20:133–40.
60. Aisenberg AC, Finkelstein DM, Doppke KP, Koerner FC, Boivin JF, Willett CG. High risk of breast carcinoma after irradiation of young women with Hodgkin's disease. Cancer 1997;79:1203–10.

61. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. Breast Cancer Res 2005;7:21–32.
62. Megdal SP, Kroenke CH, Laden F, Pukkala E, Schernhammer ES. Night work and breast cancer risk: a systematic review and meta-analysis. Eur J Cancer 2005;41:2023–32.
63. Schernhammer ES, Kroenke CH, Laden F, Hankinson SE. Night work and risk of breast cancer. Epidemiology 2006;17:108–11.
64. El-Bastawissi AY, White E, Mandelson MT, Taplin SH. Reproductive and hormonal factors associated with mammographic breast density by age (United States). Cancer Causes Control 2000;11:955–63.
65. Bernstein L, Ross RK. Endogenous hormones and breast cancer risk. Epidemiol Rev 1993;15:48–65.
66. Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. N Engl J Med 2000;342:564–71.
67. Trapido EJ. Age at first birth, parity, and breast cancer risk. Cancer 1983;51:946–8.
68. World Cancer Research Fund/American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR, 2007.
69. World Cancer Research Fund/American Institute for Cancer Research Systematic Literature Review Continuous Update Report. The associations between food, nutrition, and physical activity and the risk of breast cancer. Imperial College London. Continuous Update Team Members. November 2008. www.dietandcancerreport.org/cancer_resource_center/downloads/cu/Breast-Cancer-SLR-2008.pdf.
70. Lew JQ, Freedman ND, Leitzmann MF, Brinton LA, Hoover RN, Hollenbeck AR, et al. Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study. Am J Epidemiol 2009;170:308–17.
71. Smith-Warner SA, Spiegelman D, Yaun SS, van den Brandt PA, Folsom AR, Goldbohm RA, et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. JAMA 1998;279:535–40.
72. Lye J, Hirschberg J. Alcohol consumption and human capital: a retrospective study of the literature. J Econ Surv 2010;24:309–38.
73. Lipworth L, Bailey LR, Trichopoulos D. History of breast-feeding in relation to breast cancer risk: a review of the epidemiologic literature. J Natl Cancer Inst 2000;92:302–12.
74. Glass AG, Lacey JV Jr, Carreon JD, Hoover RN. Breast cancer incidence, 1980–2006: combined roles of menopausal hormone therapy, screening mammography, and estrogen receptor status. J Natl Cancer Inst 2007;99:1152–61.
75. Krieger N, Chen JT, Waterman PD. Decline in US breast cancer rates after the Women's Health Initiative: socioeconomic and racial/ethnic differentials. Am J Public Health 2010;100 Suppl 1:S132–9.
76. Anderson SE, Dallal GE, Must A. Relative weight and race influence average age at menarche: results from two nationally representative surveys of US girls studied 25 years apart. Pediatrics 2003;111:844–50.
77. Marti A, Marcos A, Martinez JA. Obesity and immune function relationships. Obes Rev 2001;2:131–40.
78. Nieman DC, Henson DA, Nehlsen-Cannarella SL, Ekkens M, Utter AC, Butterworth DE, et al. Influence of obesity on immune function. J Am Diet Assoc 1999;99:294–9.
79. Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, Manson JE, et al. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997;278:1407–11.
80. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003;348:1625–38.
81. Pagan JA, Davila A. Obesity, occupational attainment and earnings: Consequences of Obesity. Soc Sci Q 1997;78:756–70.
82. Seidell JC. Obesity, insulin resistance and diabetes–a worldwide epidemic. Br J Nutr 2000;83 Suppl 1:S5–8.
83. Dishman RK, Sallis JF, Orenstein DR. The determinants of physical activity and exercise. Public Health Rep 1985;100:158–71.
84. Eisenmann JC, Bartee RT, Wang MQ. Physical activity, TV viewing, and weight in U.S. youth: 1999 Youth Risk Behavior Survey. Obes Res 2002;10:379–85.
85. Weinsier RL, Hunter GR, Heini AF, Goran MI, Sell SM. The etiology of obesity: relative contribution of metabolic factors, diet, and physical activity. Am J Med 1998;105:145–50.
86. Thanos J, Cotterchio M, Boucher BA, Kreiger N, Thompson LU. Adolescent dietary phytoestrogen intake and breast cancer risk (Canada). Cancer Causes Control 2006;17:1253–61.
87. Wu AH, Yu MC, Tseng CC, Pike MC. Epidemiology of soy exposures and breast cancer risk. Br J Cancer 2008;98:9–14.
88. Peeters PH, Keinan-Boker L, van der Schouw YT, Grobbee DE. Phytoestrogens and breast cancer risk. Review of the epidemiological evidence. Breast Cancer Res Treat 2003;77:171–83.
89. Windham GC, Mitchell P, Anderson M, Lasley BL. Cigarette smoking and effects on hormone function in premenopausal women. Environ Health Perspect 2005;113:1285–90.
90. Ambrosone CB, Kropp S, Yang J, Yao S, Shields PG, Chang-Claude J. Cigarette smoking, N-acetyltransferase 2 genotypes, and breast cancer risk: pooled analysis and meta-analysis. Cancer Epidemiol Biomarkers Prev 2008;17:15–26.
91. Hiatt RA, Fireman BH. Smoking, menopause, and breast cancer. J Natl Cancer Inst 1986;76:833–8.
92. Li CI, Malone KE, Daling JR, Potter JD, Bernstein L, Marchbanks PA, et al. Timing of menarche and first full-term birth in relation to breast cancer risk. Am J Epidemiol 2008;167:230–9.
93. Verheus M, Peeters PH, Kaaks R, van Noord PA, Grobbee DE, van Gils CH. Premenopausal insulin-like growth factor-I serum levels and changes in breast density over menopause. Cancer Epidemiol Biomarkers Prev 2007;16:451–7.
94. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control 2000;11:653–62.
95. Nanda R, Schumm LP, Cummings S, Fackenthal JD, Sveen L, Ademuyiwa F, et al. Genetic testing in an ethnically diverse cohort of high-risk women: a comparative analysis of BRCA1 and BRCA2 mutations in American families of European and African ancestry. JAMA 2005;294:1925–33.
96. Schubert EL, Mefford HC, Dann JL, Argonza RH, Hull J, King MC. BRCA1 and BRCA2 mutations in Ashkenazi Jewish families with breast and ovarian cancer. Genet Test 1997;1:41–6.
97. Tonin P, Weber B, Offit K, Couch F, Rebbeck TR, Neuhausen SL, et al. Frequency of recurrent BRCA1 and BRCA2 mutations in Ashkenazi Jewish breast cancer families. Nat Med 1996;2:1179–83.
98. Varley JM, Evans DG, Birch JM. Li-Fraumeni syndrome–a molecular and clinical review. Br J Cancer 1997;76:1–14.
99. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, et al. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 2007;39:352–8.
100. Garcia-Closas M, Chanock S. Genetic susceptibility loci for breast cancer by estrogen receptor status. Clin Cancer Res 2008;14:8000–9.
101. Sun T, Gao Y, Tan W, Ma S, Shi Y, Yao J, et al. A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. Nat Genet 2007;39:605–13.
102. Zhang J, Qiu LX, Wang ZH, Leaw SJ, Wang BY, Wang JL, et al. Current evidence on the relationship between three polymorphisms in the FGFR2 gene and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010;124:419–24.
103. Rebbeck TR, DeMichele A, Tran TV, Panossian S, Bunin GR, Troxel AB, et al. Hormone-dependent effects of FGFR2 and MAP3K1 in breast cancer susceptibility in a population-based sample of postmenopausal African-American and European-American women. Carcinogenesis 2009;30:269–74.
104. Raskin L, Pinchev M, Arad C, Lejbkowicz F, Tamir A, Rennert HS, et al. FGFR2 is a breast cancer susceptibility gene in Jewish and Arab Israeli populations. Cancer Epidemiol Biomarkers Prev 2008;17:1060–5.
105. Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. Genome Biol 2002;3:comment2007.

106. Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004;13:715–22.
107. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15:1159–69.
108. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007;356:227–36.
109. Kerlikowske K, Cook AJ, Buist DS, Cummings SR, Vachon C, Vacek P, et al. Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. J Clin Oncol 2010;28: 3830–7.
110. Apter D, Reinila M, Vihko R. Some endocrine characteristics of early menarche, a risk factor for breast cancer, are preserved into adulthood. Int J Cancer 1989;44:783–7.
111. Bernstein L, Pike MC, Ross RK, Henderson BE. Age at menarche and estrogen concentrations of adult women. Cancer Causes Control 1991;2:221–5.
112. Emaus A, Espetvedt S, Veierod MB, Ballard-Barbash R, Furberg AS, Ellison PT, et al. 17-beta-estradiol in relation to age at menarche and adult obesity in premenopausal women. Hum Reprod 2008;23: 919–27.
113. Key TJ, Appleby PN, Reeves GK, Roddam AW, Helzlsouer KJ, Alberg AJ, et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. Br J Cancer 2011;105:709–22.
114. Zheng W, Lee SA. Well-done meat intake, heterocyclic amine exposure, and cancer risk. Nutr Cancer 2009;61:437–46.
115. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol 2000; 152:514–27.
116. Li CI, Littman AJ, White E. Relationship between age maximum height is attained, age at menarche, and age at first full-term birth and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2007;16:2144–9.
117. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat Rev Cancer 2007;7:937–48.
118. Malkin D, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, Kim DH, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250:1233–8.
119. Harris RE, Chlebowski RT, Jackson RD, Frid DJ, Ascenseo JL, Anderson G, et al. Breast cancer and nonsteroidal anti-inflammatory drugs: prospective results from the Women's Health Initiative. Cancer Res 2003;63:6096–101.
120. Greenfield JR, Campbell LV. Relationship between inflammation, insulin resistance and type 2 diabetes: 'cause or effect'? Curr Diabetes Rev 2006;2:195–211.
121. Kalupahana NS, Moustaid-Moussa N, Claycombe KJ. Immunity as a link between obesity and insulin resistance. Mol Aspects Med 2012;33:26–34.
122. Larsson SC, Mantzoros CS, Wolk A. Diabetes mellitus and risk of breast cancer: a meta-analysis. Int J Cancer 2007;121:856–62.
123. Healy LA, Ryan AM, Carroll P, Ennis D, Crowley V, Boyle T, et al. Metabolic syndrome, central obesity and insulin resistance are associated with adverse pathological features in postmenopausal breast cancer. Clin Oncol 2010;22:281–8.
124. Kahn BB, Flier JS. Obesity and insulin resistance. J Clin Invest 2000;106:473–81.
125. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature 2006;444:840–6.
126. Lindstrom S, Vachon CM, Li J, Varghese J, Thompson D, Warren R, et al. Common variants in ZNF365 are associated with both mammographic density and breast cancer risk. Nat Genet 2011; 43:185–7.
127. Odefrey F, Stone J, Gurrin LC, Byrnes GB, Apicella C, Dite GS, et al. Common genetic variants associated with breast cancer and mammographic density measures that predict disease. Cancer Res 2010;70:1449–58.
128. Lettre G. Recent progress in the study of the genetics of height. Hum Genet 2011;129:465–72.
129. Visscher PM, McEvoy B, Yang J. From Galton to GWAS: quantitative genetics of human height. Genet Res 2010;92:371–9.
130. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet 2008;4:e1000054.
131. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.
132. Zhan S, Ho SC. Meta-analysis of the association of the Trp64Arg polymorphism in the beta3 adrenergic receptor with insulin resistance. Obes Res 2005;13:1709–19.
133. Zhao T, Zhao J, Yang W. Association of the fatty acid-binding protein 2 gene Ala54Thr polymorphism with insulin resistance and blood glucose: a meta-analysis in 13451 subjects. Diabetes Metab Res Rev 2010;26:357–64.
134. Saunders CL, Chiodini BD, Sham P, Lewis CM, Abkevich V, Adeyemo AA, et al. Meta-analysis of genome-wide linkage studies in BMI and obesity. Obesity 2007;15:2263–75.
135. Bouchard C, Perusse L. Current status of the human obesity gene map. Obes Res 1996;4:81–90.
136. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010; 42:937–48.
137. Garland CF, Gorham ED, Mohr SB, Grant WB, Giovannucci EL, Lipkin M, et al. Vitamin D and prevention of breast cancer: pooled analysis. J Steroid Biochem Mol Biol 2007;103:708–11.
138. John EM, Schwartz GG, Dreon DM, Koo J. Vitamin D and breast cancer risk: the NHANES I Epidemiologic follow-up study, 1971–1975 to 1992. National Health and Nutrition Examination Survey. Cancer Epidemiol Biomarkers Prev 1999;8: 399–406.
139. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol 2012;13: 1141–51.
140. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, et al. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries. Int J Cancer 1990;46: 597–603.
141. Collaborative Group on Hormonal Factors in Breast Cancer. Alcohol, tobacco and breast cancer–collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. Br J Cancer 2002; 87:1234–45.
142. Reeves GK, Beral V, Green J, Gathani T, Bull D. Hormonal therapy for menopause and breast-cancer risk by histological type: a cohort study and meta-analysis. Lancet Oncol 2006;7:910–8.
143. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008;371: 569–78.
144. Euling SY, Herman-Giddens ME, Lee PA, Selevan SG, Juul A, Sorensen TI, et al. Examination of US puberty-timing data from 1940 to 1994 for secular trends: panel findings. Pediatrics 2008;121 Suppl 3: S172–91.
145. Biro FM, Galvez MP, Greenspan LC, Succop PA, Vangeepuram N, Pinney SM, et al. Pubertal assessment method and baseline characteristics in a mixed longitudinal study of girls. Pediatrics 2010;126: e583–90.
146. 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;66:1191–308.
147. Peto R. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. Br J Cancer 2011;105:S1.

148. Schottenfeld D, Beebe-Dimmer JL, Buffler PA, Omenn GS. Current perspective on the global and United States cancer burden attributable to lifestyle and environmental risk factors. Annu Rev Public Health 2013;34:97–117.
149. Gail MH, Greene MH. Gail model and breast cancer. Lancet 2000; 355:1017.
150. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature 1983;303:767–70.
151. Rosner B, Colditz GA. Nurses' health study: log-incidence mathematical model of breast cancer incidence. J Natl Cancer Inst 1996;88:359–64.
152. Foresight Programme. Tackling Obesities: Future Choices - Building the Obesity System Map. Project Report. Government Office for Science, UK; 2007 October 2007. Report No.: DIUS/PUB/2K/10/07/NP.
153. Galea S, Hall C, Kaplan GA. Social epidemiology and complex system dynamic modelling as applied to health behaviour and drug use research. Int J Drug Policy 2009;20:209–16.
154. Auchincloss AH, Riolo RL, Brown DG, Cook J, Diez Roux AV. An agent-based model of income inequalities in diet in the context of residential segregation. Am J Prev Med 2011;40: 303–11.
155. Institute of Medicine (IOM). Breast cancer and the environment: A life course approach. Washington, DC: The National Academies Press; 2012.
156. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007;357:370–9.
157. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008;358:2249–58.
158. Homer JB, Hirsch GB. System dynamics modeling for public health: background and opportunities. Am J Public Health 2006;96:452–8.