



On the relationship between the heritability and the attributable fraction

Elisabeth Dahlqvist¹ · Patrik K. E. Magnusson¹ · Yudi Pawitan¹ · Arvid Sjölander¹

Received: 3 January 2019 / Accepted: 22 March 2019 / Published online: 2 April 2019
© The Author(s) 2019

Abstract

Heritability is the most commonly used measure of genetic contribution to disease outcomes. Being the fraction of the variance of latent trait liability attributable to genetic factors, heritability of binary traits is a difficult technical concept that is sometimes misinterpreted as the more-easily understandable concept of attributable fraction. In this paper we use the liability threshold model to describe the analytical relationship between heritability and attributable fraction. Towards this end, we consider a hypothetical intervention that is aimed to reduce the genetic risk of the disease for a specified target group of the population. We show how the relation between the heritability and the attributable fraction depends on the disease prevalence, the intervention effect and the size of the target group. We use two real examples to illustrate the practical implications of our theoretical results.

Introduction

Measuring the genetic influence on human diseases is one of the most important topics in medicine and epidemiology. This is commonly done in terms of the heritability, routinely reported using studies of siblings and twins (Lichtenstein et al. 2000; Mucci et al. 2016). For continuous traits, heritability is defined as the proportion of phenotypic variation that can be attributed to genetic variation. This definition can also be used for binary traits; however, for binary traits it is more common to define the heritability as the proportion of variance on the latent liability scale attributed to genetic variation, which is conceptually more difficult and often misinterpreted (Visscher et al. 2008; Witte et al. 2014). For example, in attempts to assess how much of a disease burden can be attributed to genetic factors, heritability is sometimes misinterpreted as an attributable fraction (AF) (Mucci et al. 2016). The AF for a particular exposure is defined as the proportion of disease cases that would be prevented if the exposure was eliminated from the population (Levin 1953). This definition differs from the heritability in two important aspects. First, in contrast to the AF, the heritability does not measure the effect of an intervention. Second, while the

AF refers to a specific and well-defined exposure, which (in principle) can be eliminated, the heritability captures, in a loose sense, the aggregated impact of variation over the whole genome on the disease.

Despite differences between the concepts, it is intuitively reasonable that the heritability conveys something meaningful about the impact of genetic interventions. At one extreme, when the heritability is equal to 1, we would expect genetic interventions to have a large potential for reducing the disease prevalence. At the other extreme, when it is zero, genetic interventions will not have any impact on the disease prevalence. However, no formal analysis of the relationship between the overall heritability and the AF exists. Previous studies have either restricted the interest to the heritability attributed to a limited set of SNPs (Wang et al. 2018) or addressed the overall heritability but lacked a general formalization of this relationship (Ramakrishnan and Thacker 2012). In this work, we derive a formal link between the overall heritability and the AF by using the liability threshold model (Falconer 1965) and an extension of the AF which allows for continuous exposures (Morgenstern and Bursic 1982; Taguri et al. 2012).

The outline of this paper is as follows. First, we review the theory behind the liability threshold model and use the model to derive the relation between the AF and the heritability. Next, we illustrate the practical implications of this relationship with two real examples: one concerns the prevention of cardiovascular events by medication, and the

✉ Elisabeth Dahlqvist
elisabeth.dahlqvist@ki.se

¹ Karolinska Institute, Nobels väg 12A, 171 77 Stockholm, Sweden

other a comparison between two strategies for breast cancer prevention.

Theory

The liability threshold model

Falconer (1965) developed the liability threshold model to explain how a large number of environmental and genetic factors result in a dichotomous disease outbreak. The model assumes that the binary observed disease status Y can be described by a continuous latent (i.e. unobserved) liability L , which captures all genetic and environmental risk factors that influence the disease risk. Let G and E be scalar measures that summarize these genetic and environmental risk factors, respectively; we hereafter refer to G and E as the genetic and environmental ‘risks’. The liability model assumes that

$$L = G + E. \quad (1)$$

Moreover, it assumes that the disease occurs when the subject’s liability L exceeds a threshold β , so that the disease prevalence $p(Y = 1)$ is equal to $p(L > \beta)$.

The epidemiological literature distinguishes between the point prevalence (i.e. the proportion of the population that has the disease at a given point in time) and the lifetime prevalence (i.e. the proportion of the population that develops the disease at some point during life). Here, we are interested in ever experiencing the disease and we are thus interested in the lifetime prevalence.

Before proceeding, we emphasize an important feature of the liability threshold model. Because the model is additive, it is, for any given value $E < \infty$, possible to find values of G for which the liability falls below the threshold β . In particular, this happens when $G = -\infty$. The model thus assumes that there exists an ‘optimal’ genetic composition that will prevent the disease from occurring, regardless of what environment the subject is exposed to. For multifactorial diseases, this assumption may be reasonable as an approximation to reality, at least when the environment is not too extreme. We note, though, that even for multifactorial diseases one can easily conceive of extreme environments that would cause the disease, regardless of the subject’s genetic composition (e.g. living inside a nuclear reactor will always cause cancer).

In line with the literature (Falconer and Mackay 1996), we assume that G and E are independent and normally distributed with mean zero and variances σ_G^2 and σ_E^2 , respectively. By assuming that G and E are linearly related to L as in (1), it follows that the liability is normally distributed with mean zero and variance $\sigma_G^2 + \sigma_E^2$. Moreover, the joint

distribution of the liability L and the genetic risk G is bivariate normal:

$$\begin{pmatrix} G \\ L \end{pmatrix} \sim N \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_G^2 & \sigma_G^2 \\ \sigma_G^2 & \sigma_G^2 + \sigma_E^2 \end{pmatrix} \right\}. \quad (2)$$

The heritability as a parameter in the liability threshold model

The heritability (h^2) is defined as the ratio of genetic variation and phenotypic variation, or variation in liability,

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2}. \quad (3)$$

Hence, the correlation between the liability and the genetic risk is the square root of the heritability,

$$\begin{aligned} \text{Corr}(L, G) &= \frac{\text{Cov}(L, G)}{\sqrt{\text{Var}(L)\text{Var}(G)}} = \frac{\sigma_G^2}{\sigma_G \sqrt{\sigma_G^2 + \sigma_E^2}} \\ &= \frac{\sigma_G}{\sqrt{\sigma_G^2 + \sigma_E^2}} = h. \end{aligned} \quad (4)$$

Thus, the joint distribution of the genetic risk and the liability is a function of the heritability. In the next section, we show how this can be used to link the AF with the heritability.

The attributable fraction

The AF is a commonly used measure in epidemiology for quantifying the impact of an exposure on an outcome (Levin 1953). The AF is defined as

$$\text{AF} = \frac{p - p^*}{p}, \quad (5)$$

where p is the proportion of subjects that will ever get the disease in the factual situation, and p^* is the counterfactual proportion that will ever get the disease under an intervention that eliminates the exposure from the population.

The AF measures the (net) proportion of disease cases prevented by the intervention. For instance, suppose that the factual prevalence is 5%, and that the intervention reduces the prevalence to 1%. The proportion of prevented disease cases is then equal to $(0.05 - 0.01)/0.05 = 80\%$.

In our context, the exposure of interest is the genetic risk G in the liability model. This is supposed to capture the aggregated impact of variation over the whole genome on the disease, which cannot be ‘eliminated’ in any meaningful sense. We can, however, think of hypothetical interventions that aim at manipulating the genetic risk in other ways, e.g.

by changing its distribution or shifting it by a fixed constant. Thus, we use the standard definition of the AF in (5), but allow for p^* to represent the counterfactual disease prevalence under such ‘generalized’ interventions. This generalization of the AF is sometimes referred to as the generalized impact fraction (Morgenstern and Bursic 1982; Taguri et al. 2012).

In practice, interventions may be targeted towards subgroups of the population due to, for instance, considerations of cost-effectiveness. For example, even though preventive medication against high blood pressure and cholesterol have been shown to reduce the risk of cardiovascular events (Yusuf et al. 2016), it might not be possible to implement this intervention on the whole population, since those at low risk will not have a sufficiently high benefit to motivate bearing the cost of the medication. Instead, it is more efficient to target the intervention to those at the highest genetic risk of a cardiovascular event (Tada et al. 2016).

We allow for targeted interventions that only apply to subjects with particularly high genetic risks, e.g. those who have a genetic risk above a certain quantile in the genetic risk distribution. We thus define the target group as those subjects for whom $G > b\sigma_G$, where b is a fixed constant that corresponds to a quantile in the standard normal distribution. For instance, $b = 1.64$ and $b = 1.96$ correspond to 0.95 and 0.975 quantiles, respectively, and setting b to $-\infty$ implies that we include the whole population in the target group. In practice, a high genetic risk group could be identified by using familial risk or a genetic risk score (Yoon et al. 2002; Belsky et al. 2013; Khera et al. 2016; Tada et al. 2016).

We consider interventions that reduce the genetic risk G with an amount $k\sigma_G$, where k is a fixed positive constant. Defining L^* as the liability under the intervention we thus have that

$$L^* = G - k\sigma_G + E.$$

If the intervention is targeted to a subgroup of the population, the counterfactual disease prevalence p^* in (5) can be divided into two components. For subjects who belong to the target group, the genetic risk is reduced to $G - k\sigma_G$, whereas for subjects outside the target group the genetic risk is unchanged. The counterfactual (joint) probability of developing the disease and being in the target group is thus $p(L^* > \beta, G > b\sigma_G) = p(L > \beta + k\sigma_G, G > b\sigma_G)$, where

$$\begin{aligned} p(L > \beta + k\sigma_G, G > b\sigma_G) &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}} + \frac{k\sigma_G}{\sqrt{\sigma_G^2 + \sigma_E^2}}, \frac{G}{\sigma_G} > b\right) \\ &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}} + kh, \frac{G}{\sigma_G} > b\right). \end{aligned} \tag{6}$$

Since the intervention is only given to subjects with genetic risk G above the threshold $b\sigma_G$, the difference between the factual (observed) prevalence p and counterfactual (under the intervention) prevalence p^* arises only because of shifted liability levels within this target group. Thus, without loss of generalization we can express the numerator in (5) as

$$\begin{aligned} p - p^* &= p(L > \beta, G > b\sigma_G) - p(L - k\sigma_G > \beta, G > b\sigma_G) \\ &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}}, \frac{G}{\sigma_G} > b\right) \\ &\quad - p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}} + kh, \frac{G}{\sigma_G} > b\right). \end{aligned} \tag{7}$$

Furthermore, we have that

$$\begin{aligned} p &= p(L > \beta) \\ &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}}\right) \\ &= \Phi\left(-\frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}}\right), \end{aligned} \tag{8}$$

where $\Phi(\cdot)$ denotes the standard (i.e. mean 0, variance 1) normal distribution function. The bivariate standard normal distribution function with correlation coefficient h is denoted as $\Phi(\cdot, \cdot; h)$. We can then express the difference in prevalence as

$$p - p^* = \Phi\{\Phi^{-1}(p), -b; h\} - \Phi\{\Phi^{-1}(p) - kh, -b; h\}. \tag{9}$$

The expression in (9) shows how the disease prevalence is modified by the intervention, k , the size of the target group, b , and the square root of the heritability, h . We refer to Appendix A for a detailed derivation of (9).

From (9), the AF can be written as

$$\text{AF}(p, b, k, h) = \frac{\Phi\{\Phi^{-1}(p), -b; h\} - \Phi\{\Phi^{-1}(p) - kh, -b; h\}}{p}. \tag{10}$$

In practice, we may set the constant k depending on what reduction in disease prevalence we imagine a given intervention would result in for a fixed heritability and size of the target group. For example, suppose that the intervention is given to the whole population ($b = -\infty$), so that the expression in (9) simplifies to $\Phi\{\Phi^{-1}(p)\} - \Phi\{\Phi^{-1}(p) - kh\}$.

Suppose further that the disease prevalence is 5% before the intervention and 2.5% after the intervention. The reduction Δ is then the difference between the 0.95 and 0.975 quantiles of the standard normal distribution, i.e. $\Delta = 1.96 - 1.64 = 0.32$. For a particular heritability, say $h^2 = 0.64$, this implies that $k = \frac{0.32}{0.8} = 0.4$. Thus, a 2.5 percentage point reduction in disease prevalence for a disease with a heritability of 64% corresponds to $k = 0.4$. Moreover, setting k to ∞ reduces G to $-\infty$, thus optimizing the genetic composition so that the disease is guaranteed to be prevented.

Properties of the attributable fraction

The expression in (10) depends on four parameters: the intervention effect, k , the size of the target group, b , the heritability h^2 , and the disease prevalence, p . When k increases, the second term in the numerator increases as well. Thus, the AF increases monotonically with k , which is intuitively reasonable. At the one extreme $k = 0$ (no intervention), the two terms in the numerator become equal, so that the AF becomes equal to 0. At the other extreme $k = \infty$ (the genetic composition is optimized for those in the target group) the second term in the numerator equals 0, so that the AF simplifies to

$$AF(p, b, k = \infty, h^2) = \frac{\Phi\{\Phi^{-1}(p), -b; h\}}{p} \tag{11}$$

It is not immediately obvious from the expression in (10) how the AF depends on b , but it can be shown (see Appendix B) that the AF decreases monotonically with b . This is intuitively reasonable; the smaller the target group, the smaller is the impact of the intervention. At the one extreme $b = \infty$ (no one is targeted by the intervention), both terms in the numerator of (10) equal 0, so that the AF equals 0 as well. At the other extreme $b = -\infty$ (the whole population is included in the target group), both terms in the numerator of (10) simplify to univariate distribution functions, so that the AF simplifies to

$$\begin{aligned} AF(p, b = -\infty, k, h^2) &= \frac{\Phi\{\Phi^{-1}(p)\} - \Phi\{\Phi^{-1}(p) - kh\}}{p} \\ &= \frac{p - \Phi\{\Phi^{-1}(p) - kh\}}{p} \end{aligned} \tag{12}$$

When both $k = \infty$ and $b = -\infty$, it can be seen from either (11) or (12) that the AF equals 1. This makes intuitive sense; if the genetic composition is optimized for everybody in the population, then 100% of all disease cases would be prevented.

One would perhaps expect that the AF increases monotonically with h^2 ; i.e. the more heritable the disease, the larger is the impact of genetic interventions. However, the AF is not (necessarily) a monotone function of h^2 . Figure 1 shows the AF as a function of h^2 for $k = 1, p = 0.5$ and with the target group of the intervention ranging between 1 and 30% of those at the highest genetic risk. For small target groups, the AF increases up to h^2 between 0.6 and 0.8, and then starts to decrease. However, we see that for large target groups (25% or larger), the relationship between AF and h^2 is monotone. In general, if the intervention is given to the whole population and $b = -\infty$, we observe from the expression in (12) that the AF does indeed increase monotonically with h^2 .

We note that the hypothetical scenario in Fig. 1 is unrealistic, in that the disease prevalence is unusually high.

The relationship between the AF and the disease prevalence is rather intricate, as the prevalence appears in both numerator and denominator of the expression in (10). When $p = 1$ (everybody gets the disease), both terms in the numerator of (10) simplify to the univariate distribution function $\Phi(-b)$, so that the AF equals 0. When p approaches 0, it can be shown that the AF approaches 1 (see Appendix B). We have made an extensive grid search over k, b, h^2 and p . Based on this grid search, we conjecture that the AF decreases monotonically with p ; however, we have not been able to prove this analytically.

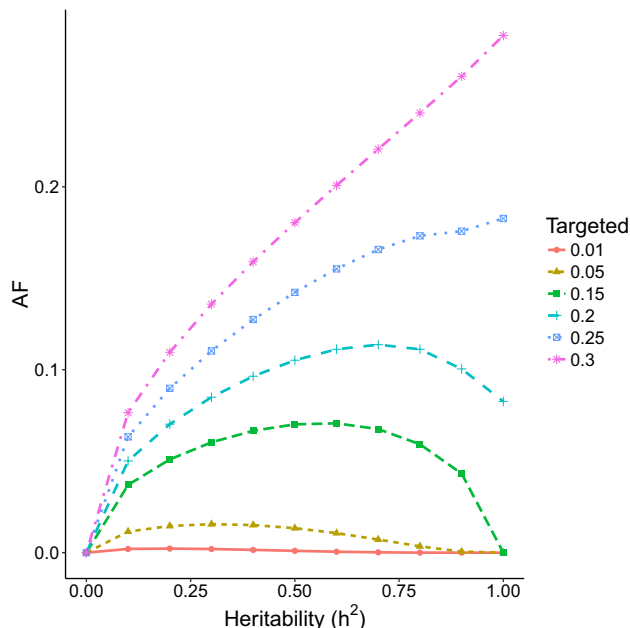


Fig. 1 The AF as a function of h^2 for $k = 1, p = 0.5$ and target group of between 1 and 30% of those at the highest genetic risk

Examples

In the previous section, we have derived the theoretical relationship between the heritability and the AF, by means of a hypothetical intervention that reduces the genetic risk of a disease. In this section, we illustrate the implications of this relationship through practical examples. In particular, we show how we can use the relationship to investigate the population impact of various intervention strategies.

Example 1: blood pressure and cholesterol-lowering medication

High blood pressure and cholesterol levels are well-known risk factors for cardiovascular events such as acute myocardial infarction (AMI) and stroke (Mozaffarian et al. 2015; Khera et al. 2016; Yusuf et al. 2016). These risk factors both have strong genetic components (Weissglas-Volkov and Pajukanta 2010; van Rijn et al. 2007) and can be lowered by preventive medical treatment with statins (cholesterol lowering) and angiotensin II receptor antagonists (blood pressure lowering) (Yusuf et al. 2016).

In a Swedish study, the heritability of AMI was estimated to be 36% and the prevalence of AMI in this cohort was approximately 6% (Zdravkovic et al. 2007). In a Danish study, the heritability for stroke was estimated to be 17% and the estimated prevalence of stroke in this cohort was around 4% (Bak et al. 2002).

We will now use these examples to investigate how the differences in heritability and prevalence between AMI and stroke impact the AF. We compare how the population impact differs between the diseases depending on if the intervention is given to those at 1% or 5% highest genetic risk of AMI and stroke, i.e. $b = 2.33$ or $b = 1.65$, respectively. The target groups could possibly be identified by a genetic risk score developed for cardiovascular disease (Thanassoulis et al. 2012).

If the intervention is given to 5% at the highest genetic risk, we suppose that this intervention can reduce the prevalence by 1.1 percentage points (i.e. from 6% to 4.9%) for AMI and 0.4 percentage points (i.e. from 4% to 3.6%) for stroke. Such a genetic risk reduction corresponds to an intervention effect of $k = 1$ for both diseases.

Figures 2 and 3 illustrate the AF as a function of heritability and prevalence. The prevalences are 2, 4, 6, 8 and 10% and the intervention effect is fixed at $k = 1$. In Fig. 2, the intervention is targeted at the 1% at highest genetic risk and in Fig. 3 it is given to the 5% highest genetic risk.

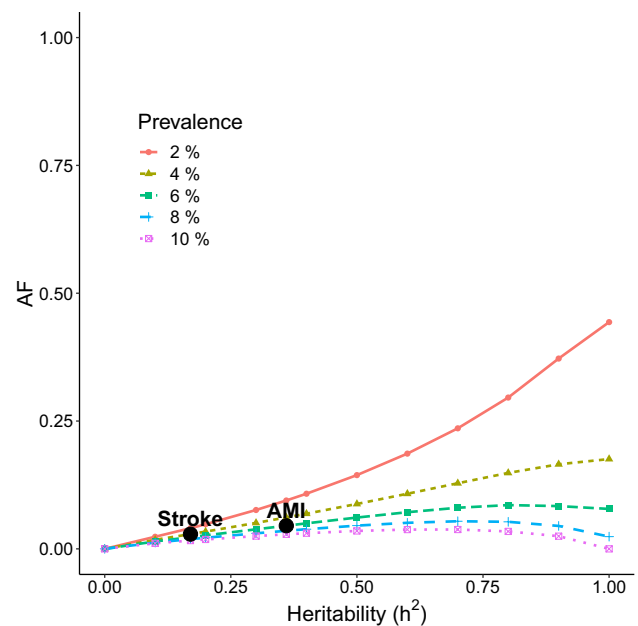


Fig. 2 The AF as a function of heritability for intervention given to the 1% at the highest genetic risk. The intervention effect is assumed to be $k = 1$

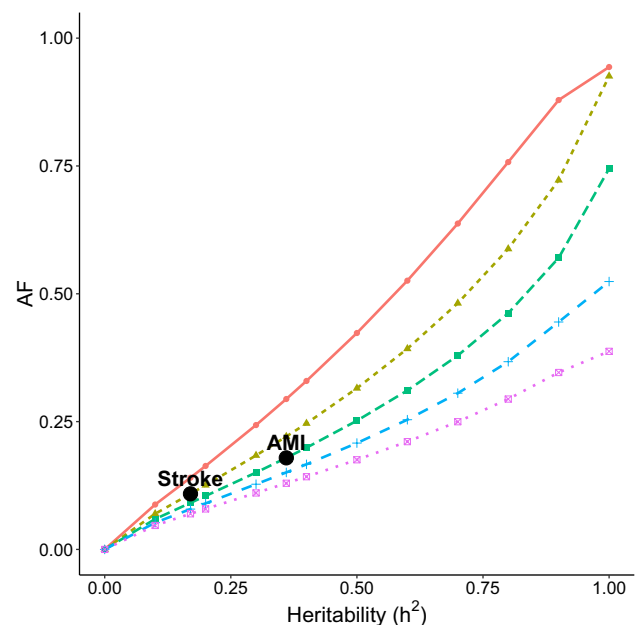


Fig. 3 The AF as a function of heritability for intervention given to the 5% at the highest genetic risk. The intervention effect is assumed to be $k = 1$

The examples of AMI and stroke are marked in Figs. 2 and 3. If the intervention is given to 1% of the population, the AF is 4.5% for AMI and 2.9% for stroke. When the

intervention is given to 5% of the population, the AF is 17.9% for AMI and 10.9% for stroke. From these two examples we observe how an intervention with larger coverage may drastically increase the population impact of the intervention. We note, however, that the benefit of increasing the target group eventually levels off, as the intervention also covers subjects with small genetic risk who do not benefit from the intervention. In general, we also observe that the smaller the prevalence, the larger is the AF for a fixed heritability, target group and intervention effect. These, and other features of the relationships between the heritability, intervention effect, target group size and disease prevalence can be investigated using our Shiny app ‘afteritability’ (Dahlqwist et al. 2018).

Example 2: prevention of breast cancer

Breast cancer is one of the most common cancers in women, with a lifetime prevalence of 8.1% and a heritability of 31%, estimated from Swedish twin data (Möller et al. 2016). Prevention strategies for breast cancer include screening programs and specialized treatments for those women at the highest genetic risk of breast cancer (Moyer and U.S. Preventive Services Task Force 2014). For example, women who are carriers of the BRCA1/2 mutations are often offered surgical removal of all breast tissue susceptible to cancer, bilateral prophylactic mastectomy (BPM) (Rebbeck et al. 2004). Another prevention strategy is treatment with tamoxifen (Moyer and US Preventive Services Task Force 2013, 2014). However, none of these interventions are used on a large scale since tamoxifen have severe side effects and may increase the risk of other adverse health outcomes (Nichols et al. 2015), and BPM is an invasive surgical procedure that may lead to additional complications (Moyer and US Preventive Services Task Force 2013). In this example we are interested in the difference in population impact of these two interventions for a fixed heritability and prevalence of breast cancer.

Studies have shown that BPM almost eliminates the risk of breast cancer (Rebbeck et al. 2004). We thus assume that the intervention effect of the BPM intervention almost eliminates breast cancer within the target group, e.g. $k = 10$. Moreover, preventive treatment with tamoxifen has been shown to almost halve the cumulative rate of invasive breast cancer (Fisher et al. 2005). Thus, based on these studies we assume that preventive tamoxifen treatment can reduce the lifetime prevalence within the target group by around 50%, which approximately corresponds to $k = 1$.

The target groups for these interventions can be chosen based on a genetic risk score for breast cancer (Shieh et al.

2016). We assume that BPM will only be given to women at the highest 1% genetic risk of breast cancer, i.e. $b = 2.33$. Tamoxifen treatment is not as invasive as the BPM intervention. However, due to the adverse effects of tamoxifen (Fisher et al. 1994; van Leeuwen et al. 1994), it is only recommended to women with a high genetic risk of breast cancer (Moyer and US Preventive Services Task Force 2013). Therefore we assume that that this intervention can be given to, at most, women at the highest 5% genetic risk of breast cancer, i.e. $b = 1.64$. We compare the AF for tamoxifen given to those at the 1% versus 5% highest genetic risk.

Figure 4 illustrates the AF as a function of the intervention effect k , for heritability 31% and target group sizes equal to 1% and 5%. We observe that the AF is 6.7% for the BPM intervention given to the 1% at the highest genetic risk, 3.1% for the tamoxifen intervention given to the 1% at the highest genetic risk and 13% for the tamoxifen intervention given to the 5% at the highest genetic risk. Thus, even though BPM almost eliminates breast cancer within the target group, it has a smaller impact than the less efficient tamoxifen intervention given to 5%, but larger impact than tamoxifen intervention given to 1%. This example illustrates that a large effect of a prevention strategy may not have a large population impact if the intervention is limited to a small part of those at risk.

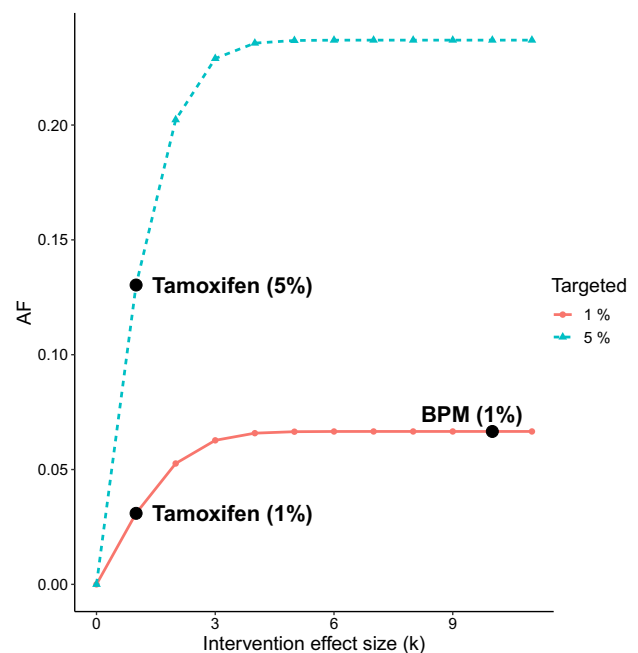


Fig. 4 The AF for breast cancer with $h^2 = 31\%$ and $p = 8.1\%$. Comparison of BPM given to a target group of 1% with $k = 10$ and tamoxifen given to target group of 1% versus 5% with $k = 1$

Discussion

Heritability is a central concept in genetic epidemiology. Yet, because it is defined in terms of proportion of variance of latent disease liability, it is difficult to interpret the population implications of a particular value of the heritability (Witte et al. 2014). Attempts to do so sometimes interprets heritability as an attributable fraction (AF) (Mucci et al. 2016). In this article, we have shown how the relationship between the heritability and the AF can be formalized and how these results can be used to understand the population implications of a particular value of the heritability.

The relationship between the heritability and the AF is rather intricate, since it depends on several parameters in a non-linear way. Both the disease prevalence and the effect of the genetic intervention modify the impact of the heritability on the AF. In reality, interventions may not always be targeted to the whole population. We have accounted for this by adding the possibility to consider situations where the intervention is only targeted at those at the highest genetic risk. Intuitively, one would expect that the AF increases monotonically with the heritability. However, we have shown by examples that this is not necessarily the case. In particular, we have shown that, if the prevalence is high and the target group is small, then the AF may increase with the heritability up to a certain point, after which it will start to decrease.

Two examples have been used to illustrate how our results can be used to understand the relationship between the heritability and the AF for different scenarios. The first example is an intervention with blood pressure and cholesterol-lowering medication to prevent AMI and stroke. In this example we have considered the target group, b , and the intervention effect, k , to be fixed and we have compared the AF separately for AMI and stroke. For the same intervention, the AF is larger for AMI compared to stroke due to the larger heritability of AMI. In the second example, we have compared two interventions to prevent breast cancer, bilateral prophylactic mastectomy and tamoxifen with a fixed heritability, h^2 , and prevalence, p . Bilateral prophylactic mastectomy (BPM) has a large intervention effect, but can only be given to a limited target group. Compared to BPM, tamoxifen has a smaller intervention effect but can be given to a larger target group. In this example, we observed how the AF is larger for the tamoxifen intervention compared to the BPM intervention despite its lower intervention effect since the target group is larger.

We are not the first to use the AF to measure the effect of genes. However, virtually all applications that we are aware of define the exposure as a single SNP/gene (Claus et al.

1996; Witte et al. 2014; Khoury et al. 2005) or a limited set of SNPs (Wang et al. 2018), thus measuring the effect of that single SNP and not the aggregated impact of the whole genome. A notable exception is Ramakrishnan and Thacker (2012), who used the AF for twin data and defined the exposure for a given twin as the disease status in the co-twin. The authors claimed that, by using this exposure definition, their attributable fraction would measure the proportion of disease cases that are ‘due to heritability’. However, they did not formally motivate this claim, and it is not obvious whether the term ‘due to heritability’ is empirically meaningful.

To derive the relation between the heritability and the AF, we have used the liability threshold model, which relies on several important assumptions. Specifically, it assumes that the genetic and environmental risks for the disease are normally distributed, the effect of genes and environment add up to the liability (i.e. that there is no additive statistical interaction between the risk factors) and the genetic and environmental risks are independent. Regarding the first assumption, it is reasonable to assume that the genetic and environmental risks are normally distributed, since we are considering complex traits that depend on the accumulated small contributions from many different genetic and environmental factors. The second assumption of mainly additive effects from the genetic and environmental factors has been debated to a great extent (Hill et al. 2008). However, there is not much evidence of statistical interaction effects between genes and environment (Hill et al. 2008; Hunter 2005). The third assumption of no gene–environment correlation is violated if genes affect the environment or vice versa. Gene–environment correlation might occur due to various reasons (Jaffee and Price 2008) and should be carefully considered in each particular application.

In this article, we have conceptualized ‘genetic interventions’ as all interventions that modify the genetic risk. These can be pure genetic interventions, such as gene therapy, or interventions that target the mechanisms by which the genes exert their effect, such as bilateral prophylactic mastectomy. Thus, by ‘genetic’ we do not necessarily mean that the genes are modified per se, but rather that the intervention modifies the circumstances that allow the genetic variants to manifest.

Throughout, we have defined the target group as those at the highest genetic risk. As suggested by one of the reviewers, target groups may in practice also be defined in terms of environmental risk for disease. In Appendix C we derive the relation between the AF and the heritability, when the target groups are defined as those subjects for which the total (genetic and environmental) liability

exceeds a certain threshold. We show that, in fact, this alternative definition of the target group leads to simpler calculations, which only involves univariate normal distribution functions.

Acknowledgements We would like acknowledge the biostatistics group at MEB, Zoltán Kutalik, Ralf Kuja-Halkola and Patrick Sullivan for valuable discussions and comments on this work. We acknowledge the financial support from the Swedish Research Council through the Swedish Initiative for Research on Microdata in the Social and Medical Sciences (SIMSAM) framework grant no. 340-2013-5867.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Proof of (9)

We have that

$$\begin{aligned}
 p - p^* &= p(L > \beta) - p(L^* > \beta) \\
 &= \{p(L > \beta, G > b\sigma_G) + p(L > \beta, G \leq b\sigma_G)\} \\
 &\quad - \{p(L^* > \beta, G > b\sigma_G) + p(L^* > \beta, G \leq b\sigma_G)\} \\
 &= p(L > \beta, G > b\sigma_G) - p(L^* > \beta, G > b\sigma_G) \\
 &= p(L > \beta, G > b\sigma_G) - p(L - k\sigma_G > \beta, G > b\sigma_G) \\
 &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}}, \frac{G}{\sigma_G} > b\right) \\
 &\quad - p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \frac{\beta + k\sigma_G}{\sqrt{\sigma_G^2 + \sigma_E^2}}, \frac{G}{\sigma_G} > b\right) \\
 &= p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > -\Phi^{-1}(p), \frac{G}{\sigma_G} > b\right) \\
 &\quad - p\left(\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > -\Phi^{-1}(p) + kh, \frac{G}{\sigma_G} > b\right) \\
 &= \Phi\{\Phi^{-1}(p), -b; h\} - \Phi\{\Phi^{-1}(p) - kh, -b; h\},
 \end{aligned}$$

where the second last equality follows from (8), and the last equality follows from (2) and (4).

Appendix B: Proof of relations between the AF and k, b, h^2 and p

Let X and Y be bivariate standard normal with correlation h . We then have the following:

- The AF is decreasing in b . *Proof* From (10) we have that

$$\begin{aligned}
 \frac{\partial \text{AF}}{\partial b} &= \frac{p\{X < \Phi^{-1}(p), Y = -b\} - p\{X < \Phi^{-1}(p) - kh, Y = -b\}}{p} \\
 &= -p(Y = -b) \frac{p\{X < \Phi^{-1}(p) | Y = -b\} - p\{X < \Phi^{-1}(p) - kh | Y = -b\}}{p} \\
 &= -\phi(b) \frac{\Phi\left\{\frac{\Phi^{-1}(p) + hb}{\sqrt{1-h^2}}\right\} - \Phi\left\{\frac{\Phi^{-1}(p) + h(b-k)}{\sqrt{1-h^2}}\right\}}{p} < 0,
 \end{aligned}$$

where the third equality follows from the conditional distribution of X , given Y , and k being positive.

- $\lim_{p \rightarrow 0} \text{AF} = 1$. *Proof:* Define $q = \Phi^{-1}(p)$. We have that

$$\begin{aligned}
 \lim_{p \rightarrow 0} \text{AF} &= \lim_{q \rightarrow -\infty} \text{AF} \\
 &= \lim_{q \rightarrow -\infty} \frac{\partial\{p(X < q, Y < -b) - p(X < q - kh, Y < -b)\}/\partial q}{\partial\Phi(q)/\partial q} \\
 &= \lim_{q \rightarrow -\infty} \frac{p(X = q, Y < -b) - p(X = q - kh, Y < -b)}{\phi(q)} \\
 &= \lim_{q \rightarrow -\infty} \frac{p(X = q)p(Y < -b | X = q) - p(X = q - kh)p(Y < -b | X = q - kh)}{\phi(q)} \\
 &= \lim_{q \rightarrow -\infty} \frac{(2\pi)^{-1/2} \left[\exp(-q^2/2) \Phi\left(\frac{-b-qh}{\sqrt{1-h^2}}\right) - \exp\{-(q-kh)^2/2\} \Phi\left(\frac{-b-(q-kh)h}{\sqrt{1-h^2}}\right) \right]}{(2\pi)^{-1/2} \exp(-q^2/2)} \\
 &= \lim_{q \rightarrow -\infty} \Phi\left(\frac{-b-qh}{\sqrt{1-h^2}}\right) - \exp\{2qkh - (kh)^2\}/2 \Phi\left(\frac{-b-(q-kh)h}{\sqrt{1-h^2}}\right) \\
 &= 1
 \end{aligned}$$

where the second equality follows from L'Hospital's rule, and the last equality follows by $\lim_{q \rightarrow -\infty} \Phi\left(\frac{-b-qh}{\sqrt{1-h^2}}\right) =$

$$\lim_{q \rightarrow -\infty} \Phi\left(\frac{-b-(q-kh)h}{\sqrt{1-h^2}}\right) = 1 \text{ and } \lim_{q \rightarrow -\infty} \exp\{2qkh - (kh)^2\}/2 = 0.$$

Appendix C: Relationship when the intervention is based on overall liability

To accommodate both genetic and environmental factors, we may target the intervention to subjects with $L > b\sqrt{\sigma_G^2 + \sigma_E^2}$. We then have that

$$\begin{aligned}
 p - p^* &= p(L > \beta) - p(L^* > \beta) \\
 &= \left\{ p\left(L > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right. \\
 &\quad \left. + p\left(L > \beta, L \leq b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right\} \\
 &\quad - \left\{ p\left(L^* > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right. \\
 &\quad \left. + p\left(L^* > \beta, L \leq b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right\} \\
 &= p\left(L > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \\
 &\quad - p\left(L^* > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \\
 &= p\left(L > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \\
 &\quad - p\left(L - k\sigma_G > \beta, L > b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \\
 &= p\left\{ L > \max\left(\beta, b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right\} \\
 &\quad - p\left\{ L > \max\left(\beta + k\sigma_G, b\sqrt{\sigma_G^2 + \sigma_E^2}\right) \right\} \\
 &= p\left\{ \frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \max\left(\frac{\beta}{\sqrt{\sigma_G^2 + \sigma_E^2}}, b\right) \right\} \\
 &\quad - p\left\{ \frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \max\left(\frac{\beta + k\sigma_G}{\sqrt{\sigma_G^2 + \sigma_E^2}}, b\right) \right\} \\
 &= p\left[\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \max\{-\Phi^{-1}(p), b\} \right] \\
 &\quad - p\left[\frac{L}{\sqrt{\sigma_G^2 + \sigma_E^2}} > \max\{-\Phi^{-1}(p) + kh, b\} \right] \\
 &= \Phi[\min\{\Phi^{-1}(p), -b\}] \\
 &\quad - \Phi[\min\{\Phi^{-1}(p) - kh, -b\}],
 \end{aligned}$$

where the second last equality follows from (3) and (8), and the last equality follows from (2). Thus, AF can be written as

$$AF(p, b, k, h) = \frac{\Phi[\min\{\Phi^{-1}(p), -b\}] - \Phi[\min\{\Phi^{-1}(p) - kh, -b\}]}{p}$$

We observe that, for this alternative definition of the target group, the AF becomes a relatively simple function of (p, b, k, h) , since it only involves univariate normal distribution functions. Below, we investigate how the AF depends on (p, b, k, h) .

We first note that

$$\Phi^{-1}(p) - kh > -b \Rightarrow AF = 0, \tag{13}$$

so in this case, AF does not depend on (p, b, k, h) .

Dependency on k :

- $k < \frac{\Phi^{-1}(p) + b}{h} \Rightarrow AF = 0$. Follows from (13).
- $k > \frac{\Phi^{-1}(p) + b}{h}$: AF increases monotonically with k .
- $\lim_{k \rightarrow \infty} = \frac{h\Phi[\min\{\Phi^{-1}(p), -b\}]}{p}$.

Dependency on h :

- $h < \frac{\Phi^{-1}(p) + b}{k} \Rightarrow AF = 0$. Follows from (13).
- $h > \frac{\Phi^{-1}(p) + b}{k}$: AF increases monotonically with h .
- $h = 1 \Rightarrow AF = \frac{\Phi[\min\{\Phi^{-1}(p), -b\}] - \Phi[\min\{\Phi^{-1}(p) - k, -b\}]}{p}$.

Dependency on b :

- $b > -\Phi^{-1}(p) + kh \Rightarrow AF = 0$. Follows from (13).
- $-\Phi^{-1}(p) < b < -\Phi^{-1}(p) + kh \Rightarrow AF = \frac{\Phi(-b) - \Phi\{\Phi^{-1}(p) - kh\}}{p}$, which decreases monotonically with b .
- $b < -\Phi^{-1}(p) \Rightarrow AF = 1 - \frac{\Phi\{\Phi^{-1}(p) - kh\}}{p}$; doesn't depend on b .

Dependency on p :

- $p > \Phi(kh - b) \Rightarrow AF = 0$. Follows from (13).
- $\Phi(-b) < p < \Phi(kh - b) \Rightarrow AF = \frac{\Phi(-b) - \Phi\{\Phi^{-1}(p) - kh\}}{p}$, which decreases monotonically with p . Proof: Define $q = \Phi^{-1}(p)$. We then have that

$$AF = \frac{\Phi(-b)}{\Phi(q)} - \frac{\Phi(q - kh)}{\Phi(q)}. \tag{14}$$

For the first term, we have that

$$\frac{\partial}{\partial q} \left\{ \frac{\Phi(-b)}{\Phi(q)} \right\} < 0.$$

For the second term, we have that

$$\frac{\partial}{\partial q} \left\{ -\frac{\Phi(q - kh)}{\Phi(q)} \right\} = \frac{\phi(q)\Phi(q - kh) - \Phi(q)\phi(q - kh)}{\Phi^2(q)}.$$

This expression is negative if and only if

$$\begin{aligned} \phi(q)\Phi(q - kh) - \Phi(q)\phi(q - kh) < 0 &\Leftrightarrow \\ -\frac{\phi(q)}{\Phi(q)} > -\frac{\phi(q - kh)}{\Phi(q - kh)} > 0, \end{aligned}$$

Which is true, since $-\frac{\phi(q)}{\Phi(q)} = E(X|X < q) > E(X|X < q - kh) = -\frac{\phi(q - kh)}{\Phi(q - kh)}$, where X has a standard normal distribution.

Thus, $\frac{\partial AF}{\partial q} < 0 \Rightarrow \frac{\partial AF}{\partial p} < 0$, which concludes the proof.

- $p < \Phi(-b) \Rightarrow AF = 1 - \frac{\Phi\{\Phi^{-1}(p) - kh\}}{p}$, which decreases monotonically with p , and similar proof as above.
- $\lim_{p \rightarrow 0} AF = 1$. Proof:

$$\begin{aligned} \lim_{p \rightarrow 0} AF &= \lim_{q \rightarrow -\infty} AF \\ &= \lim_{q \rightarrow -\infty} 1 - \frac{\Phi(q - kh)}{\Phi(q)} \\ &= \lim_{q \rightarrow -\infty} 1 - \frac{\partial\{\Phi(q - kh)/\partial q\}}{\partial\Phi(q)/\partial q} \\ &= \lim_{q \rightarrow -\infty} 1 - \frac{\phi(q - kh)}{\phi(q)} \\ &= \lim_{q \rightarrow -\infty} 1 - \frac{(2\pi)^{-1/2} \exp\{-(q - kh)^2/2\}}{(2\pi)^{-1/2} \exp\{-q^2/2\}} \\ &= \lim_{q \rightarrow -\infty} 1 - \exp\{[2qkh - (kh)^2]/2\} \\ &= 1, \end{aligned}$$

where the third equality follows from L'Hospital's rule.

References

Bak S, Gaist D, Sindrup SH, Skytthe A, Christensen K (2002) Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke* 33(3):769–774

Belsky DW, Moffitt TE, Sugden K, Williams B, Houts R, McCarthy J, Caspi A (2013) Development and evaluation of a genetic risk score for obesity. *Biodemography Soc Biol* 59(1):85–100

Claus EB, Schildkraut JM, Thompson WD, Risch NJ (1996) The genetic attributable risk of breast and ovarian cancer. *Cancer* 77(11):2318–2324

Dahlquist E, Magnusson P, Pawitan Y, Sjölander A (2018) Afheritability: a tool to visualize the relationship between the attributable fraction and the heritability. <https://afheritability.shinyapps.io/afheritability/>. Accessed 19 Jan 2018

Falconer DS (1965) The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 29(1):51–76

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman, London

Fisher B, Costantino JP, Redmond CK, Fisher ER, Wickerham DL, Cronin WM (1994) Endometrial cancer in Tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. *JNCI: J Nat Cancer Inst* 86(7):527–537

Fisher B, Costantino JP, Wickerham DL, Cecchini RS, Cronin WM, Robidoux A, Bevers TB, Kavanah MT, Atkins JN, Margolese RG, Runowicz CD, James JM, Ford LG, Wolmark N (2005) Tamoxifen for the prevention of breast cancer: current status of the national surgical adjuvant breast and bowel project P-1 study. *JNCI: J Nat Cancer Inst* 97(22):1652–1662

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLOS Genet* 4(2):e1000008

Hunter DJ (2005) Gene-environment interactions in human diseases. *Nat Rev Genet* 6(4):287–298

Jaffee SR, Price TS (2008) Genotype-environment correlations: implications for determining the relationship between environmental exposures and psychiatric illness. *Psychiatry* 71(2):496–499

Khera AV, Emdin CA, Drake I, Natarajan P, Bick AG, Cook NR, Chasman DI, Baber U, Mehran R, Rader DJ, Fuster V, Boerwinkle E, Melander O, Orho-Melander M, Ridker PM, Kathiresan S (2016) Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med* 375(24):2349–2358

Khoury MJ, Davis R, Gwinn M, Lindgren ML, Yoon P (2005) Do we need genomic research for the prevention of common diseases with environmental causes? *Am J Epidemiol* 161(9):799–805

van Leeuwen FE, van den Belt-Dusebout AW, van Leeuwen FE, Benraadt J, Diepenhorst FW, van Tinteren H, Coebergh JWW, Kiemeny LALM, Gimbirre CHF, Otter R, Schouten LJ, Damhuis RAM, Benraadt J, Bontenbal M (1994) Risk of endometrial cancer after tamoxifen treatment of breast cancer. *The Lancet* 343(8895):448–452

Levin ML (1953) The occurrence of lung cancer in man. *Acta-Unio Internationalis Contra Cancrum* 9(3):531–541

Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K (2000) Environmental and heritable factors in the causation of cancer analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343(2):78–85

Möller S, Mucci LA, Harris JR, Scheike T, Holst K, Halekoh U, Adami HO, Czene K, Christensen K, Holm NV, Pukkala E, Skytthe A, Kaprio J, Hjelmborg JB (2016) The heritability of breast cancer among women in the Nordic twin study of cancer. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 25(1):145–150

Morgenstern H, Bursic ES (1982) A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population. *J Community Health* 7(4):292–309

Moyer VA, US Preventive Services Task Force (2013) Medications to decrease the risk for breast cancer in women: recommendations from the US Preventive Services Task Force recommendation statement. *Ann Intern Med* 159(10):698–708

Moyer VA, US Preventive Services Task Force (2014) Risk assessment, genetic counseling, and genetic testing for BRCA-related cancer in women: US Preventive Services Task Force recommendation statement. *Ann Internal Med* 160(4):271–281

Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, de Ferranti S, Desprs JP, Fullerton HJ, Howard VJ, Huffman MD, Judd SE, Kissela BM, Lackland DT, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Matchar DB, McGuire DK, Mohler ER, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Sorlie

- PD, Stein J, Towfighi A, Turan TN, Virani SS, Willey JZ, Woo D, Yeh RW, Turner MB, American Heart Association Statistics Committee and Stroke Statistics Subcommittee (2015) Heart disease and stroke statistics-2015 update: a report from the American Heart Association. *Circulation* 131(4):e29–322
- Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, Graff RE, Holst K, Möller S, Unger RH, McIntosh C, Nuttall E, Brandt I, Penney KL, Hartman M, Kraft P, Parmigiani G, Christensen K, Koskenvuo M, Holm NV, Heikkilä K, Pukkala E, Skytthe A, Adami HO, Kaprio J, Nordic Twin Study of Cancer (NorTwinCan) Collaboration (2016) familial risk and heritability of cancer among twins in Nordic countries. *JAMA* 315(1):68–76
- Nichols HB, DeRoo LA, Scharf DR, Sandler DP (2015) Risk-benefit profiles of women using tamoxifen for chemoprevention. *JNCI: J Nat Cancer Inst* 107(1). <https://doi.org/10.1093/jnci/dju354>
- Ramakrishnan V, Thacker LR (2012) Population attributable fraction as a measure of heritability in dichotomous twin data. *Commun Stat: Simul Comput* 41(3):405–418
- Rebbeck TR, Friebel T, Lynch HT, Neuhausen SL, van 't Veer L, Garber JE, Evans GR, Narod SA, Isaacs C, Matloff E, Daly MB, Olopade OI, Weber BL (2004) Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE study group. *J Clin Oncol Off J Am Soc Clin Oncol* 22(6):1055–1062
- van Rijn MJE, Schut AFC, Aulchenko YS, Deinum J, Sayed-Tabatabaei FA, Yazdanpanah M, Isaacs A, Axenovich TI, Zorkoltseva IV, Zillikens MC, Pols HAP, Witteman JCM, Oostra BA, van Duijn CM (2007) Heritability of blood pressure traits and the genetic contribution to blood pressure variance explained by four blood-pressure-related genes. *J Hypertens* 25(3):565–570
- Shieh Y, Hu D, Ma L, Huntsman S, Gard CC, Leung JWT, Tice JA, Vachon CM, Cummings SR, Kerlikowske K, Ziv E (2016) Breast cancer risk prediction using a clinical risk model and polygenic risk score. *Breast Cancer Res Treat* 159(3):513–525
- Tada H, Melander O, Louie JZ, Catanese JJ, Rowland CM, Devlin JJ, Kathiresan S, Shiffman D (2016) Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur Heart J* 37(6):561–567
- Taguri M, Matsuyama Y, Ohashi Y, Harada A, Ueshima H (2012) Doubly robust estimation of the generalized impact fraction. *Biostatistics* 13(3):455–467
- Thanassoulis G, Peloso GM, Pencina MJ, Hoffmann U, Fox CS, Cupples LA, Levy D, D'Agostino RB, Hwang SJ, O'Donnell CJ (2012) A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium—the Framingham heart study. *Circ: Cardiovasc Genet p CIRCGENETICS*.111.961342
- Visscher PM, Hill WG, Wray NR (2008) Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 9(4):255–266
- Wang T, Hosgood HD, Lan Q, Xue X (2018) The relationship between population attributable fraction and heritability in genetic studies. *Front. Genet.* 9:352
- Weissglas-Volkov D, Pajukanta P (2010) Genetic causes of high and low serum HDL-cholesterol. *J Lipid Res* 51(8):2032–2057
- Witte JS, Visscher PM, Wray NR (2014) The contribution of genetic variants to disease depends on the ruler. *Nat Rev Genet* 15(11):765–776
- Yoon PW, Scheuner MT, Peterson-Oehlke KL, Gwinn M, Faucett A, Khoury MJ (2002) Can family history be used as a tool for public health and preventive medicine? *Genet Med* 4(4):304–310
- Yusuf S, Lonn E, Pais P, Bosch J, López-Jaramillo P, Zhu J, Xavier D, Avezum A, Leiter LA, Piegas LS, Parkhomenko A, Keltai M, Keltai K, Sliwa K, Chazova I, Peters RJG, Held C, Yusuf K, Lewis BS, Jansky P, Khunti K, Toff WD, Reid CM, Varigos J, Accini JL, McKelvie R, Pogue J, Jung H, Liu L, Diaz R, Dans A, Dagenais G, Investigators H (2016) Blood-pressure and cholesterol lowering in persons without cardiovascular disease. *N Engl J Med* 374(21):2032–2043
- Zdravkovic S, Wienke A, Pedersen NL, de Faire U (2007) Genetic susceptibility of myocardial infarction. *Twin Res Hum Genet* 10(6):848–852

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.