A framework evaluating the utility of multi-gene, multi-disease population-based panel testing that accounts for uncertainty in penetrance estimates

Jane W. Liang[1,2], Kurt D. Christensen[3-5], Robert C. Green[6,5,7,8], and Peter Kraft*[,9,10,1,5]

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[2]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA

[3]Center for Healthcare Research in Pediatrics, Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, Massachusetts

[4]Department of Population Medicine, Harvard Medical School, Boston, Massachusetts

[5]Broad Institute of MIT and Harvard, Cambridge, Massachusetts

[6]Mass General Brigham, Boston, MA, USA

[7]Ariadne Labs, Boston, MA, USA

[8]Harvard Medical School, Boston, MA, USA

[9]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[10]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

*phillip.kraft@nih.gov

Abstract

Panel germline testing allows for efficient detection of deleterious variants for multiple conditions, but the benefits and harms of identifying these variants are not always well-understood. We present a multi-gene, multi-disease aggregate utility formula that allows the user to consider adding or removing each gene in a panel based on variant frequency; estimated penetrances; and subjective disutilities for testing positive but not developing the disease and testing negative but developing the disease. We provide credible intervals for utility that reflect uncertainty in penetrance estimates. Rare, highly-penetrant deleterious variants tend to contribute positive net utilities for a wide variety of user-specified disutilities, even when accounting for parameter estimation uncertainty. However, the clinical utility of deleterious variants with moderate, uncertain penetrance depends more on assumed disutilities. The decision to include a gene on a panel depends on variant frequency, penetrance, and subjective utilities, and should account for uncertainties around these factors.

Introduction

Genetic screening for deleterious variants (DVs) in genes associated with monogenic hereditary conditions (typically pathogenic and likely pathogenic variants) can be a valuable component of risk management for opportunistic and population-based genomic screening[1,2]. Testing results can prompt heightened surveillance, prophylactic surgery, and other measures to enhance prevention or treatment. Technological advances such as next-generation sequencing have made simultaneous testing of multiple genes cheaper and more accurate than ever before[3,4]. Panel studies have led to many clinically significant findings that would have been missed by single-gene or single-syndrome testing[3–6]. However, the clinical utility of such comprehensive panel germline testing may not be universally appropriate for all contexts.

For some genes and diseases, published guidelines provide best practices about actions to take when deleterious variants are identified in the context of diagnostic testing. For other genes and settings (e.g. secondary findings or population screening), there is a lack of consensus on whether screening itself, or subsequent interventions based upon screening, should be recommended, often because the disease penetrance (the probability that carriers of deleterious variants will develop disease) is low or unknown. For example, while penetrance has been estimated through families with strong family histories, the penetrance estimates in population screening may still be uncertain[7]. If the benefits, risks, and guidelines are unclear for these genes, it could be harmful rather than beneficial to include them in a testing panel. Instead of mitigating risk and improving outcomes, testing may lead to unnecessary surveillance and overtreatment.

We consider a scenario in which the goal is to determine which genes should be included in a panel as part of non-diagnostic screening for a fixed group of diseases. This setting occurs at the stage when a panel is being built and is distinct from the problem of developing a clinician/patient-facing risk tool. While our approach is readily generalizable to other contexts, such as asymptomatic high-risk settings and the incorporation of variants of

uncertain significance, we will focus on population-based screening for deleterious variants in asymptomatic individuals as our motivating context. We propose an aggregate utility function that incorporates quantitative measures of genetic and disease characteristics (carrier prevalences and disease penetrances) and utility benefits and harms (harms and costs are sometimes termed “disutility”) for multiple diseases and germline tests. Positive utilities could include identifying individuals at high risk for disease who would benefit from intervention, who would remain unrecognized in the absence of testing. Disutilities could include anxiety or false reassurance in response to test results, unnecessary surveillance, and overtreatment. Utilities and disutilities can be individualized for specific diseases and tests, as well as patient and clinician concerns.

This approach generates a single net utility across all genes proposed for inclusion in a panel, but our construction also allows for the evaluation of each disease and gene combination on its own merits. We note that our notion of net utility and (dis)utilities is distinct from the *health utility* that is frequently used within decision science for valuing a disease state with respect to death and perfect health. This health utility is typically a value between 0 to 1, but our net utility may take on any positive or negative value, with positive net utilities indicating that it is beneficial to include the gene(s) and negative net utilities indicating that it is harmful.

Additionally, we incorporate credible intervals for disease penetrances that reflect our confidence in available penetrance estimates and propagate this uncertainty into the net utility calculation. Uncertainty may be due to lack of sufficient data to estimate prevalences or penetrances for certain DVs and diseases, as well as ancestral populations or other subgroups of interest. For sufficiently large penetrances, the net utility may provide evidence in favor of keeping the test even when the penetrance estimate is unreliable. For low or moderate penetrances, the net utility may point toward removing the gene from the panel or needing to improve the reliability of the penetrance estimate. This utility approach can be used to help formalize the decision-making process when designing a gene panel. Some approaches that

take uncertainty in risk estimates into account for other domains include Berry and Parmigiani[8] and Ding et al.[9] The former considers quantifying uncertainty for decision analysis of testing for BRCA1 and BRCA2 mutations, while the latter applies Bayesian methods to estimate the variance of an individual's polygenic risk score.

Some related methodology has been proposed for addressing how to select genes for panel germline testing, as well as the broader question of how to best leverage modern sequencing technology for clinical use. Most identify clinical interpretation and actionability as critical considerations; many focus on diagnostic applications, while we consider non-diagnostic applications (e.g. risk-stratified screening recommendations). Hall et al.[10] give an overview of the benefits and challenges in gene panel testing for inherited cancer risk assessment, highlighting ambiguous clinical utility as a potential disadvantage. Because genetic testing results may provide uncertainty rather than information for managing cancer risk, the authors recommend testing be used alongside professional consultation. Xue et al.[11] modify Shashi et al.[12]'s testing algorithm for evaluating which molecular diagnostic tool (single-gene tests, gene panels, or exome sequencing) to use for diagnostic yield, depending on the clinical setting. Finally, Mazzarotto et al.[13] develop a “diagnostic effectiveness” score for determining genes to include for hypertrophic cardiomyopathy genetic testing panels, based on variant classification and penetrance. They identify new genes to screen for but also suggest that panels beyond a limited size provide limited additional sensitivity.

An illustration of our approach to germline testing for deleterious variants on five genes (*ATM*, *BRCA1*, *BRCA2*, *CHEK2*, and *PALB2*) associated with increased risk of developing breast cancer follows in the Results section, along with a broad exploration of the effects of varying parameter inputs (true penetrance, uncertainty in penetrance estimates, variant frequencies, relative (dis)utilities). The Methods section details our formulation of general expressions for our proposed multi-gene, multi-disease aggregate utility. Our method is implemented in an R Shiny app, freely accessible at https://janewliang.shinyapps.io/agg_utility,

where users can enter parameter estimates and uncertainties for calculating their own net utilities.

Results

Female breast cancer application

We first consider a specific application for the aggregate utility approach that incorporates panel germline testing for *ATM*, *BRCA1*, *BRCA2*, *CHEK2*, and *PALB2* as part of risk assessment for female breast cancer, for hypothetical screening of a woman without a previous breast cancer diagnosis or breast cancer family history. These five genes are commonly included in risk panels for hereditary breast cancer. We chose this example for its familiarity and relevance, as well as the availability of empirical estimates of the lifetime risk of breast cancer in women for carriers of deleterious variants in these genes and their relative precisions. We stress that these results are largely presented for illustrative purposes. In particular, although our understanding of the absolute and relative uncertainty in the penetrance estimates for these genes is changing as more data become available in more diverse populations, the lifetime risk for carriers of DVs in these genes is relatively well known—the uncertainty in penetrances estimates for other diseases and other genes is often much greater[14,15]. Users are free to input their own prevalence and penetrance estimates, as well as the uncertainty in the penetrance estimates, into our R Shiny app to calculate their impact on likely net utility.

Deleterious variants in these genes have all been linked to breast cancer, but some are better studied than others. *BRCA1* and *BRCA2* DVs are highly penetrant with widely adopted guidelines for enhanced screening and other clinical interventions[16–19]. While DVs of *ATM*, *CHEK2*, and *PALB2* have been linked to breast cancer, the additional risk conferred is not as well understood[20–23], especially among individuals with non-European ancestries. In this section, quantities (prevalences and penetrances) involving *ATM*, *BRCA1*, *BRCA2*, and *PALB2* are

estimated for any deleterious variant in the given gene; for *CHEK2*, quantities are for the 1100delC variant only.

The carrier prevalences for DVs on *BRCA1* (0.00058) and *BRCA2* (0.00068) are calculated based on allele frequency estimates reported in Antoniou et al.[24] (see also Dullens et al.[25]; Krassuski et al.[26]). Those for *ATM* (0.0019), *CHEK2* (0.0026), and *PALB2* (0.00057) are calculated based on allele frequencies reported in Lee et al.[27] Cumulative lifetime penetrance estimates for female breast cancer are taken from the literature review performed by the All Syndromes Known to Man Evaluator[28–32]: 0.35 (*ATM*), 0.73 (*BRCA1*), 0.72 (*BRCA2*), 0.19 (*CHEK2*), and 0.38 (*PALB2*). This is the genotype-specific probability of developing breast cancer among females, prior to dying. In this illustration, we use parameters estimated from large meta-analyses, largely of studies from the USA, UK, Australia, or countries in Western Europe. Estimates for specific ancestries and subgroups may be used instead, to better reflect a different population of interest

To reflect our greater confidence in the penetrance estimates for *BRCA1* and *BRCA2*, we use a precision of $n_{D_i G_j} = 10,000$ to specify the parameters in their uncertainty distributions. $n_{D_i G_j}$ parameterizes the uncertainty in estimating the penetrance for disease *i* associated with deleterious variants in gene *j*; see Methods. Intuitively, the penetrance's uncertainty distribution can be thought of as a posterior distribution from a trial with $n_{D_i G_j}$ carriers, where larger values of $n_{D_i G_j}$ correspond to a greater degree of perceived certainty for the estimate. For *ATM*, *CHEK2*, and *PALB2*, we specify $n_{D_i G_j} = 100$, i.e. a smaller trial size of 100. (We chose these values to illustrate the impact of uncertainty on net utility calculations. They should not be taken as indicative of the absolute or relative strength of the available data on the penetrance of DVs in these genes.) Supplementary Figure 1 plots the uncertainty distributions of the five lifetime penetrance estimates. The wider spread for *ATM*, *CHEK2*, and *PALB2* reflects greater

uncertainty. Supplementary Table 1 summarizes the quantiles for these uncertainty distributions at 2.5%, 5%, 10%, 50%, 90%, 95%, and 97.5%.

We report the individual gene net utilities and aggregate utility for a multigene breast cancer panel testing for deleterious variants in all five of these genes. Individual gene utility is the net change in utility from including a gene in the panel relative to not screening for deleterious variants in that gene; the aggregate utility for a set of genes (and diseases) is the sum of individual gene (and disease) utilities (see Methods). The aggregate net utility Δ is a function of the frequency of deleterious variants, lifetime penetrances, and relative disutilities $\delta_{D_i=0,\, G_j=1}$, $\delta_{D_i=1,\, G_j=0}$, and K_{D_i,G_j} (indexed over gene i and disease j). Further detail is provided in Methods, but in brief, $\delta_{D_i=0,\, G_j=1}$ denotes the disutility for an individual who tests positive for the gene, but does not develop the phenotypic features of the associated disease (abbreviated G+D-); $\delta_{D_i=1,\, G_j=0}$ denotes the disutility for an individual who tests negative for the gene, but does develop phenotypic features of the associated disease (abbreviated G-D+); and K_{D_i,G_j} is the disutility of screening itself. Because hereditary predisposition for breast cancer drives only a portion of cases, most individuals in the general population who develop breast cancer over their lifetimes would fall in the G-D+ group. Since the deleterious variants considered are rare, G+D- individuals (those with hereditary cancer predisposition due to these genes) represent a small subgroup of those who never develop breast cancer. In general, the disutilities encompass a broad range of financial and non-financial harms and can vary across disease i and tested gene j. For ease of presentation, we assume that disutilities are the same across all five tests (denoted as $\delta_{D=0,\, G=1}$, $\delta_{D=1,\, G=0}$, and K), that $K = 0$, and (without loss of generality) that $\delta_{D=0,\, G=1} = 1$. We allow the utility ratio $\delta_{D=1,\, G=0}/\delta_{D=0,\, G=1}$ to vary from 0.1 to 10 in increments of $\log_{10}(0.1)$ (see Methods). Figure 1 plots the net utilities against this utility ratio for each of the individual genes, as well as the aggregate Δ for all five genes. Supplementary Figure 2 depicts

the same curves on a log-transformed x-axis, to help illustrate the behavior for small values. Supplementary Figure 3 presents a supplemental analysis where the disutilities for BRCA and BRCA2 differ from those specified for ATM, CHEK2, and PALB2.

As expected, the credible intervals for the *BRCA1* and *BRCA2* net utilities are very narrow and the credible intervals for *ATM*, *CHEK2*, and *PALB2* are wider, reflecting the widths of the credible intervals in their uncertainty penetrance distributions. The aggregate utility has the widest credible intervals of all, because it incorporates uncertainty from all five penetrance estimates.

We can consider interpreting the results in terms of utility thresholds, defined as the ratio of $\delta_{D_i=1,\, G_j=0}$ (disutility of G-D+) to $\delta_{D=0,\, G=1}$ (disutility of G+D-) such that the net utility is 0 (Table 1). Additional detail, including derivations, can be found in the Methods, but in general, a gene with a higher utility threshold (>1) can be interpreted as having a more limited range of subjective inputs $\delta_{D_i=1,\, G_j=0}$ and $\delta_{D=0,\, G=1}$ where it would still be beneficial to include it in the panel. These threshold values are quite low for *BRCA1* (0.37) and *BRCA2* (0.40), so even in a scenario where one is highly concerned with avoiding G+D- results, there is a wide range of possible disutilities that can be specified to result in a positive net utility. The net utility for testing these two genes is positive, except for some extreme cases when $\delta_{D=1,\, G=0}$ is very low compared to $\delta_{D=0,\, G=1}$. The curves for the probability of the net utility being positive resemble step functions, with the jump from being 0% positive to 100% positive occurring at a sharp, early point.

In contrast, the less-penetrant genes have utility thresholds above 1: *ATM* (1.9), *CHEK2* (4.1), and *PALB2* (1.6). In these cases, the G-D+ disutility needs to outweigh the G+D- disutility in order for it to be beneficial to keep the gene in the panel, sometimes by a considerable amount. Because of the greater uncertainty in the penetrance estimates, the lower bound of the utility threshold credible interval (Table 1) is noticeably even less favorable. The probability

curve for observing a positive net utility also bends toward 100% at a much more gradual incline. The aggregate utility threshold is somewhere intermediate (1.8), balancing between the larger and smaller effect sizes, as is the shape of the probability curve.

Heatmaps of the individual and aggregate net utilities while holding $K = 0$ and varying $\delta_{D=0,\, G=1}$ and $\delta_{D=1,\, G=0}$ from 1 to 100 in increments of 5 are depicted in Figure 2. Similar heatmaps for the probability of a positive net utility and the fifth percentile net utility are shown in Supplementary Figures 4 and 5. Utility thresholds based on the original penetrance estimates are drawn as solid black lines on all three sets of heatmaps. The dashed black reference lines have intercept 0 and slope 1, and correspond to cases where G+D- and G-D+ have equal disutilities. In Supplementary Figure 5, the additional solid blue lines indicate the utility thresholds based on the fifth percentiles of the penetrances' uncertainty distributions.

These heatmaps offer an alternative visualization as well as some additional insight on the behavior of the utilities under different G+D- and G-D+ disutility conditions. In Figure 2, the net utilities for *BRCA1* and *BRCA2* are positive (blue) for a much broader range of $\delta_{D=0,\, G=1}$ and $\delta_{D=1,\, G=0}$ values compared to the net utilities for the other genes, in concordance with the utility threshold discussion for Figure 1.

The utility threshold reference lines in the heatmaps for the probability of a positive net utility (Supplementary Figure 4) track with the regions where the probability of positive net utility transitions from 0 (white) to 1 (dark blue). The sharp transitions for *BRCA1* and *BRCA2* reflect their tight credible intervals, and the more gradual transitions for *ATM*, *CHEK2*, and *PALB2* reflect their wider credible intervals. The heatmaps for the fifth percentiles of the net utilities (Supplementary Figure 5) closely resemble the heatmaps for the net utilities based on the original penetrance estimates. The utility thresholds for *BRCA1* and *BRCA2* are quite similar; there is more variability in the utility thresholds for *ATM*, *CHEK2*, and *PALB2*. Again, this reflects the wider credible intervals and uncertainty distributions for these genes. Under a "near-worst

case scenario" interpretation (i.e. basing decisions about potential utility on the lowest 5th percentile of the utility distribution), *ATM*, *CHEK2*, and *PALB2* require the specification of an even larger G-D+ disutilities relative to their G+D- disutilities in order to result in positive net utilities for testing.

Net utility behavior as parameters vary

In order to explore the properties of our proposed net utility expression across a wider range of scenarios, we varied the parameters influencing the individual net utility for disease i and gene j, Δ_{D_i,G_j}, as follows. Let $D_i = \{0, 1\}$ be the indicator for developing disease i and $G_j = \{0, 1\}$ be the indicator for testing positive for carrying a deleterious variant on gene j (see Methods).

- G-D+ disutility $\delta_{D_i=1,\, G_j=0}$: Ranging from 0.1 to 10 in increments of $\log_{10}(0.1)$
- Cumulative lifetime disease penetrance $Pr(D_i = 1 \mid G_j = 1)$: $\{0.2, 0.4, 0.6, 0.8, 0.99\}$
- Carrier prevalence $Pr(G_j = 1)$: $\{0.001, 0.002, 0.003, 0.004\}$
- Precision n_{D_i,G_j} used to specify parameters in the uncertainty distribution:

  $\{10, 100, 1000, 10000\}$

The chosen penetrance and carrier prevalence values reflect those seen in clinical practice. Lifetime penetrances vary between 0.195 and 0.732 and carrier prevalences vary between 0.00114 and 0.00519 in the female breast cancer application. The ClinGen actionability reports[33] frequently list disease risks with broad ranges of possible values. For example, carriers of *STK11* deleterious variants have a 38-66% estimated risk of developing gastrointestinal cancer by age 60-70 and a 13-18% risk for gynecological cancer[34]. The penetrance of developing dopa-responsive dystonia among *GCH1* carriers is 87-100% for females and 35-55% for males[35]. Carriers of *MLH1*, *MSH2*, *MSH6* or *PMS2* have a 25-70% cumulative risk for colorectal

cancer by age 70 and 30-70% for endometrial cancer. Penetrances by age 70 for other cancers are generally lower in effect size and narrower in range of estimated values, including 1-9% for gastric, 2-16% for bladder, 6-14% for ovarian, 9-30% for prostate, and 5-14% for breast[36].

As in the previous subsection, we set the test disutility K_{D_i,G_j} to 0 and the G+D- disutility $\delta_{D_i=0,\,G_j=1}$ to 1, thereby normalizing the utility ratio to be $\delta_{D_i=1,\,G_j=0}\ /\delta_{D_i=0,\,G_j=1} = \delta_{D_i=1,\,G_j=0}$ and allowing it to range from 0.1 to 10 in increments of $\log_{10}(0.1)$. Figure 3 plots Δ_{D_i,G_j} against $\delta_{D_i=1,\,G_j=0}\ /\delta_{D_i=0,\,G_j=1}$ for each combination of parameters, with prevalences varying in the rows and penetrances varying in the columns. (Supplementary Figure 6 depicts the same curves on a log-transformed x-axis, to help illustrate the behavior for small values.) The colored shading represents 95% credible intervals for different values of n_{D_i,G_j}. Dashed reference lines are drawn to indicate the utility threshold in each scenario. Table 2 gives these threshold values with a 95% credible interval.

Overall, higher carrier prevalences and lower penetrances tend to correspond to wider credible intervals for all precision levels. However, if the carrier prevalence is very low, the credible intervals remain consistently narrow even when the penetrance is also very low, and similarly when the disease penetrance is very high in the presence of high prevalence. So the net utilities for rare, highly penetrant genes are more likely to have narrow credible intervals, independent of the amount of confidence we have about the penetrance estimates.

Interestingly, the credible intervals for net utilities at given prevalence and penetrance values grow wider as $\delta_{D_i=1,\,G_j=0}$ (disutility of a G-D+ result) increases relative to $\delta_{D_i=0,\,G_j=1}$ (disutility of a G+D- result). So, even supposing that one undervalues the relative disutility of G+D-, the uncertainty from the distribution of the penetrance allows the decision of which genes to keep in the panel to be less dependent on the exact choice of disutilities. Higher penetrances

correspond to lower utility thresholds (since we set $K_{D_i G_j} = 0$, the utility threshold does not depend on prevalence), which makes intuitive sense: high penetrance implies the proportion of carriers who do not develop the disease of interest is small, so interventions with smaller $\delta_{D_i=1,\, G_j=0} / \delta_{D_i=0,\, G_j=1}$ ratios can still have positive net utility.

Discussion

We have derived net utility expressions to aid in determining which genes add or detract utility from a genomic testing panel in the setting of population-based screening for deleterious variants in asymptomatic individuals. These expressions are functions of carrier prevalence and disease penetrance estimates, as well as user-specified disutilities for G+D-, G-D+, and testing. Our approach is flexible, and allows users to estimate impact in a variety of clinical contexts, from population-level applications to screening of high-risk populations. These expressions may provide a useful framework for determining which genes to include on a custom sequencing panel or which genes to include on a clinical report from whole exome sequencing, for asymptomatic individuals. One goal in this context is to prevent or mitigate the disease of interest by identifying high-risk individuals. The trade-offs in utility will therefore depend partly on the efficacy of the would-be-prescribed interventions for preventing poor health outcomes. We present utility thresholds, the probability of a positive utility, and lower bounds on the net utilities as summary values that can provide additional insight. As an illustration of our approach, we evaluated the net utility of population screening for deleterious variants in five breast cancer predisposition genes, as well as a hypothetical range of disease penetrances, carrier prevalences, and precision values (used for specifying the penetrance's uncertainty distribution).

Our work provides a needed approach for estimating the incremental utility or disutility of genetic screening for DVs in numerous genes and conditions simultaneously. Published estimates about the clinical benefits of genetic screening to date have focused on conditions

with reasonably developed evidence bases[37–40]. Yet, the ability of genomic sequencing to identify genetic variants associated with rare disorders, for which epidemiological evidence is typically limited, has been one of the most promising successes from advances in genetic testing capabilities[41–44]. Moreover, the American College of Medical Genetics and Genomics is integrating an increasing number of conditions into its recommendations for a minimal list for secondary findings disclosure, even when data about the penetrance of DVs in associated genes is limited[45–47]. The approach we have developed allows for better estimation of the benefits and harms of such recommendations, estimates that have been omitted from research to date[40]. Our tool provides a flexible approach that can accommodate varying measures of utility and disutility, including quality-adjusted life years, life years gained or lost, and death rates. Moreover, the tool can be easily tailored to accommodate utility and disutility for a variety of perspectives, from patient outcomes to societal impact[48].

Our approach does not directly account for potential challenges in curation accuracy for deleterious variants, and we generally do not distinguish between different DVs of the same gene (although our framework is easily modified to have several individual variants or classes of variants). We assume that modern germline testing technology detects DVs with near-perfect sensitivity and specificity. Here, we refer to sensitivity and specificity with respect to sequencing technology, not the problem of classifying variants as benign or deleterious on the basis of clinical sensitivity and specificity. We further assume that the carrier prevalences can be estimated with a high degree of accuracy, such that they do not contribute a significant amount of additional uncertainty to the net utilities. When formulating the net utility expression, we treat untested individuals as being equivalent to those who test negative for the DV(s) in question. In scenarios that deviate considerably from these conditions, we acknowledge that our approach may be of limited use.

Further work can explore challenges when building a utility that incorporates many more genes/variants or genes with unknown parameter estimates/variants of uncertain significance,

as well as accounting for the age of the person being tested or measured polygenic risk scores[47,49]. Accounting for the age of the person tested and treating penetrances as age-based distributions would allow us to model more complex relationships between DVs and diseases with incomplete penetrances. We can also conduct a more rigorous exploration of simplifying assumptions that reduce the number of disutility parameters that need to be specified. Nevertheless, our work provides a feasible approach to estimating the clinical benefits or harms of genetic screening. Tools such as ours are critically needed by policymakers and payers as they make decisions about how to regulate and reimburse the current generation of genomic tests.

Methods

Utility notation

Suppose that we are interested in risk assessment for some predetermined set of diseases, indexed $i = 1, \ldots, I$, and are considering the genes $j = 1, \ldots, J$ to be included in a panel for germline testing. We define an aggregate utility expression in terms of the following notation:

- $D_i = \{0, 1\}$ is the indicator for developing disease i.
- $G_j = \{0, 1\}$ is the indicator for testing positive for carrying a deleterious variant on gene j.
- $C_{D_i=1, G_j=1} > 0$ is the utility associated with the scenario where the individual tests positive for carrying a deleterious variant on gene j and does develop disease i (abbreviated G+D+).

- $C_{D_i=0,\, G_j=0} > 0$ is the utility associated with the scenario where the individual tests negative for carrying a deleterious variant for gene j and does not develop disease i (abbreviated G-D-).
- $C_{D_i=0,\, G_j=1} > 0$ is the utility associated with the scenario where the individual tests positive for carrying a deleterious variant on gene j but does not develop disease i (abbreviated G+D-). Assume that $C_{D_i=0,\, G_j=1} < C_{D_i=0,\, G_j=0}$.
- $C_{D_i=1,\, G_j=0} > 0$ is the utility associated with the scenario where the individual tests negative for carrying a deleterious variant for gene j but develops disease i (abbreviated G-D+). Assume that $C_{D_i=1,\, G_j=0} < C_{D_i=1,\, G_j=1}$.

We emphasize that the scenarios outlined by these definitions capture incomplete penetrance, as opposed to genotyping errors or misclassifying deleterious variants (DVs). In our notation, developing a disease (D+) refers to lifetime development of specific phenotypic features of a condition, and the converse (D-) refers to not developing those features. Let $\delta_{D_i=0,\, G_j=1} = C_{D_i=0,\, G_j=0} - C_{D_i=0,\, G_j=1}$ be the disutility associated with testing positive for gene j but not developing disease i (G+D-) or alternatively the utility benefit of testing negative for gene j and not developing disease i (G-D-), e.g. the disutility associated with unnecessary surveillance and over-treatment and possible anxiety due to a positive test. Similarly, define $\delta_{D_i=1,\, G_j=0} = C_{D_i=1,\, G_j=1} - C_{D_i=1,\, G_j=0}$ as the disutility associated with testing negative for gene j but developing disease i (G-D+) or alternatively the utility benefit of testing positive for gene j and developing disease i (G+D+), e.g. the disutility associated with default screening or preventive interventions relative to more intensive interventions along with false reassurance among those who go on to develop disease. We assume both $\delta_{D_i=0,\, G_j=1}$ and $\delta_{D_i=1,\, G_j=0}$ are greater than 0. In other words, we assume that if one does not develop the disease, testing negative for the

associated DV leads to more beneficial outcomes, and if one does develop the disease, testing positive leads to more beneficial outcomes. (We do not consider situations where either $\delta_{D_i=1,\, G_j=0} < 0$ or $\delta_{D_i=0,\, G_j=1} < 0$, although we note that these may exist; for example, where the G-D+ utility is larger than the G+D+ utility and "the cure is worse than the disease".) Finally, let K_{D_i,G_j} be the disutility (potentially including psychological or physical harms) associated with conducting the test for gene j in relation to disease i, independent of test results. Then the net utility for disease i in the setting where we test for gene j is

$$\begin{aligned} & C_{D_i=1,\, G_j=1}\, Pr(D_i = 1,\ G_j = 1) \ + C_{D_i=0,\, G_j=1}\, Pr(D_i = 0,\ G_j = 1) \\ & \qquad + \ C_{D_i=1,\, G_j=0}\, Pr(D_i = 1,\ G_j = 0) \ + C_{D_i=0,\, G_j=0}\, Pr(D_i = 0,\ G_j = 0) \ + K_{D_i,G_j} \\ & = Pr(D_i = 0)[(C_{D_i=0,\, G_j=0} - \delta_{D_i=0,\, G_j=1})Pr(G_j = 1 \mid D_i = 0) \ + C_{D_i=0,\, G_j=0}\, Pr(\, G_j = 0 \mid D_i = 0)] \\ & \quad + \ Pr(D_i = 1)[C_{D_i=1,\, G_j=1} Pr(G_j = 1 \mid D_i = 1) \ + \ (C_{D_i=1,\, G_j=1} - \delta_{D_i=1,\, G_j=0})Pr(\, G_j = 0 \mid D_i = 1)] \\ & + K_{D_i,G_j}. \end{aligned} \tag{1}$$

Assuming that the utility associated with developing disease i in the absence of testing information for gene j is equal to $C_{D_i=1,\, G_j=0}$, and assuming that the utility for not developing disease i in the absence of testing is equal to $C_{D_i=0,\, G_j=0}$, then the net utility for disease i in the scenario where we do not test for gene j is

$$\begin{aligned} & C_{D_i=1,\, G_j=0} Pr(D_i = 1) \ + C_{D_i=0,\, G_j=0} Pr(D_i = 0) \\ & = \ (C_{D_i=1,\, G_j=1} - \delta_{D_i=1,\, G_j=0})Pr(\, D_i = 1) \ + C_{D_i=0,\, G_j=0} Pr(D_i = 0). \end{aligned} \tag{2}$$

Of interest is the difference in utility for disease i when testing vs. not testing for gene j, which we define as the difference of Eq. (1) and Eq. (2):

$$\begin{aligned} & \Delta_{D_i,G_j} = Pr(D_i = 0)[- \delta_{D_i=0,\, G_j=1} Pr(G_j = 1 \mid D_i = 0)] \\ & + \ Pr(D_i = 1)[\delta_{D_i=1,\, G_j=0} Pr(G_j = 1 \mid D_i = 1)] \ + K_{D_i,G_j}. \end{aligned} \tag{3}$$

This difference in utility can be re-expressed in terms of $Pr(G_j = 1)$, which is the prevalence for DVs of gene j, and $Pr(D_i = 1 \mid G_j = 1)$, which is the cumulative lifetime risk or penetrance of developing disease i given that one is carries a DV in gene j (i.e., the penetrance):

$$\begin{aligned}\Delta_{D_i,G_j} &= [-\,\delta_{D_i=0,\,G_j=1}\,Pr(G_j = 1)\,Pr(D_i = 0 \mid G_j = 1)] \\ &\quad + [\delta_{D_i=1,\,G_j=0}\,Pr(G_j = 1)\,Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j} \\ &= [-\,\delta_{D_i=0,\,G_j=1}\,Pr(G_j = 1)\,(1 - Pr(D_i = 1 \mid G_j = 1))] \\ &\quad + [\delta_{D_i=1,\,G_j=0}\,Pr(G_j = 1)\,Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j} \\ &= Pr(G_j = 1)[-\,\delta_{D_i=0,\,G_j=1} + (\delta_{D_i=0,\,G_j=1} + \delta_{D_i=1,\,G_j=0})\,Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j}. \qquad (4)\end{aligned}$$

It is beneficial to test for gene j when $\Delta_{D_i,G_j} > 0$, which occurs when the utility for testing is greater than the utility for not testing. For simplicity, we will generally treat testing for a given gene as testing for a particular deleterious variant in the gene, but the framework readily extends to handle variant-specific tests, prevalences, and penetrances.

For multiple diseases (indexed by i) and tests (indexed by j), the aggregate utility Δ sums over all combinations of i and j. Doing so requires carrier prevalence and disease penetrance estimates for each gene and disease, as well as the specification of disutilities for testing positive for gene j but not developing disease i (G+D-), testing negative for gene j but developing disease i (G-D+), and testing itself. Δ provides a simple summary value while still allowing genes to be evaluated individually:

$$\begin{aligned}\Delta &= \sum_{(i,j)} \Delta_{D_i,G_j} \\ &= \sum_{(i,j)} \{\, Pr(D_i = 0)[-\,\delta_{D_i=0,\,G_j=1} Pr(G_j = 1 \mid D_i = 0)] \\ &\quad + Pr(D_i = 1)[\delta_{D_i=1,\,G_j=0} Pr(G_j = 1 \mid D_i = 1)] + K_{D_i,G_j}\} \\ &= \sum_{(i,j)} \{\, Pr(G_j = 1)[-\,\delta_{D_i=0,\,G_j=1} + (\delta_{D_i=0,\,G_j=1} + \delta_{D_i=1,\,G_j=0}) Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j}\}. \qquad (5)\end{aligned}$$

An additional K_{D_i,G_j} can be included for each (i, j) combination (with perhaps less weight given for each additional test), or a single overall disutility for all testing can be used. Since Δ is the sum of the net utilities of particular disease-gene pairs (i, j), the decision as to whether or not to include a given test on a multi-gene, multi-disease panel depends only on the individual net utility of that test.

The number of disutility parameters $\delta_{D_i=0,\, G_j=1}$, $\delta_{D_i=1,\, G_j=0}$, and K_{D_i,G_j} grows as the number of diseases and tests increases, but one can consider simplifications such as assuming the same (dis)utilities across diseases/tests or subgroups of diseases/tests. For example, it may be reasonable to assume that the disutility of each test for an additional gene j is negligible. Specification of these disutilities is largely subjective and should depend on the clinical setting and patient concerns.

Utility threshold

As a general guide for interpretation, if the user fixes the value of $\delta_{D_i=0,\, G_j=1}$, they can conceive of the value of $\delta_{D_i=1,\, G_j=0}$ as being a relative weight for the disutility of G-D+ vs the disutility of G+D-. More formally, for an individual test for disease i and gene j, a utility threshold can be defined as the value $b_{D_i,G_j} = \delta_{D_i=1,\, G_j=0}/\delta_{D_i=0,\, G_j=1}$ for which $\Delta_{D_i,G_j} = 0$:

$$0 = \Delta_{D_i,G_j}$$
$$= Pr(G_j = 1)[-\delta_{D_i=0,\, G_j=1} + (\delta_{D_i=0,\, G_j=1} + \delta_{D_i=1,\, G_j=0})Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j},$$

which implies

$$Pr(G_j = 1) - (K_{D_i,G_j}/\delta_{D_i=0,\, G_j=1}) = (1 + b_{D_i,G_j})Pr(D_i = 1 \mid G_j = 1)Pr(G_j = 1),$$

so

$$b_{D_i,G_j} = [1 - (K_{D_i,G_j}/\delta_{D_i=0,\, G_j=1})/Pr(G_j = 1)]/Pr(D_i = 1 \mid G_j = 1) - 1. \qquad (6)$$

Note that when $K_{D_i,G_j} = 0$, $b_{D_i,G_j} = 1/Pr(D_i = 1 \mid G_j = 1) - 1$ and depends only on the penetrance $Pr(D_i = 1 \mid G_j = 1)$. If the ratio of the disutility of G-D+ to the disutility of G+D- is greater than b_{D_i,G_j}, then including gene *j* to test for disease *i* has positive net utility. If the ratio needed to achieve a non-negative utility is unreasonable—e.g. in many settings ascribing a higher disutility to testing positive for the gene but not developing the disease compared to the disutility of testing negative for the gene but developing the disease would be inappropriate—then the test should not be kept as part of the panel. Basing analysis around a threshold ratio allows for an alternative interpretation that does not require upfront specification of the disutilities $\delta_{D_i=0,\, G_j=1}$, $\delta_{D_i=1,\, G_j=0}$, and K_{D_i,G_j}.

If one assumes that $\delta_{D_i=0,\, G_j=1} = \delta_{D=0,\, G=1}$ and $\delta_{D_i=1,\, G_j=0} = \delta_{D=1,\, G=0}$ for all values of i and j, then

$$\Delta = \sum_{(i,j)} \{ Pr(G_j = 1)[-\delta_{D=0,\, G=1} + (\delta_{D=0,\, G=1} + \delta_{D=1,\, G=0}) Pr(D_i = 1 \mid G_j = 1)] + K_{D_i,G_j} \}$$

$$= -\delta_{D=0,\, G=1} \sum_{(i,j)} Pr(G_j = 1) + (\delta_{D=0,\, G=1} + \delta_{D=1,\, G=0}) \sum_{(i,j)} Pr(G_j = 1)\, Pr(D_i = 1 \mid G_j = 1) + \sum_{(i,j)} K_{D_i,G_j}$$

$$= \delta_{D=0,\, G=1}[-\sum_{(i,j)} Pr(G_j = 1) + \sum_{(i,j)} (K_{D_i,G_j} / \delta_{D=0,\, G=1}) + (1 + \delta_{D=1,\, G=0}/\delta_{D=0,\, G=1}) \sum_{(i,j)} Pr(G_j = 1)\, Pr(D_i = 1 \mid G_j = 1)]$$

and the threshold $b = \delta_{D=1,\, G=0}/\delta_{D=0,\, G=1}$ when $\Delta = 0$ can be expressed as

$$b = \sum_{(i,j)} [Pr(G_j = 1) - (K_{D_i,G_j}/\delta_{D=0,\, G=1})] \,/\, \sum_{(i,j)} Pr(G_j = 1) Pr(D_i = 1 \mid G_j = 1)] - 1$$

$$= \sum_{(i,j)} Pr(G_j = 1) \,/\, \sum_{(i,j)} Pr(G_j = 1) Pr(D_i = 1 \mid G_j = 1)] - 1, \qquad (7)$$

where the last line holds when $K_{D_i,G_j} = 0$ for all i, j.

Uncertainty distribution for disease penetrance

Of additional interest is the incorporation of uncertainty in the penetrance estimates $Pr(D_i = 1 \mid G_j = 1)$. Denoting $p_{D_i,G_j} = Pr(D_i = 1 \mid G_j = 1)$ for a given disease i and gene j, we model the uncertainty in the penetrance p_{D_i,G_j} as a beta distribution Beta(α_{D_i,G_j}, β_{D_i,G_j}). One can motivate the choice of the parameters α_{D_i,G_j} and β_{D_i,G_j} by conceiving of the penetrance's uncertainty distribution as the posterior distribution from a trial of n_{D_i,G_j} carriers of deleterious variant j. Then, set $\alpha_{D_i,G_j} = n_{D_i,G_j} p_{D_i,G_j}$ to represent the expected number of cases of disease i in the trial and set $\beta_{D_i,G_j} = n_{D_i,G_j}(1 - p_{D_i,G_j})$ to represent the expected number of individuals who do not develop the disease. Through specification of the precision n_{D_i,G_j}, we can express our confidence level in the estimation of p_{D_i,G_j}, with larger values of n_{D_i,G_j} corresponding to a greater degree of certainty about the estimate and smaller ones indicating less confidence.

The uncertainty from p_{D_i,G_j} can then be propagated into a distribution and credible interval for the corresponding Δ_{D_i,G_j} and the aggregate Δ (assuming independence of p_{D_i,G_j} across all i, j), as well as additional summary values. We will assume that we are not concerned about incorporating uncertainty from estimating $Pr(G_j = 1)$. The probability that the individual net utility Δ_{D_i,G_j} is positive (i.e. adding the test for gene j makes an improvement) can be written as

$$\begin{aligned} Pr(\Delta_{D_i,G_j} > 0) &= Pr\{Pr(G_j = 1)[-\delta_{D_i=0,\,G_j=1} + (\delta_{D_i=0,\,G_j=1} + \delta_{D_i=1,\,G_j=0})p_{D_i,G_j}] + K_{D_i,G_j} > 0\} \\ &= Pr\{p_{D_i,G_j} > [\delta_{D_i=0,\,G_j=1} - K_{D_i,G_j}/Pr(G_j = 1)]/[\delta_{D_i=0,\,G_j=1} + \delta_{D_i=1,\,G_j=0}]\} \\ &= Pr\{p_{D_i,G_j} > \delta_{D_i=0,\,G_j=1}/[\delta_{D_i=0,\,G_j=1} + \delta_{D_i=1,\,G_j=0}]\}, \end{aligned} \tag{8}$$

where the last line holds if $K_{D_i,G_j} = 0$. $Pr(\Delta > 0)$ does not generally have a closed form, but can be calculated empirically from the sampling distributions of the p_{D_i,G_j}s. One can also derive a lower bound on the estimated Δ_{D_i,G_j}s that accounts for uncertainty by plugging in the fifth percentiles of the p_{D_i,G_j}s in the uncertainty distributions in place of $Pr(D_i = 1 \mid G_j = 1)$ in Equation (6). This fifth percentile represents a "near-worst case scenario" for the net utility in which the true disease penetrance is at the low end of its credible range.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Code Availability

Our net utility calculations are available in an R Shiny app, freely accessible at https://janewliang.shinyapps.io/agg_utility. Code is available at https://github.com/janewliang/agg_utility.

Acknowledgements

J.W.L. was supported by the National Cancer Institute at the National Institutes of Health (5T32CA009337). K.D.C. was supported by the National Human Genome Research Institute (K01HG009173).

Author Contributions

J.W.L (conceptualization, formal analysis, methodology, software, visualization, writing-original draft, writing-review & editing). K.D.C. (conceptualization, writing-review &

editing). R.C.G. (conceptualization, writing-review & editing). P.K (conceptualization, methodology, writing-review & editing).

Competing Interests

R.C.G. has received compensation for advising the following companies: AIA, Allelica, Atria, Fabric, Genome Web, Genomic Life, Grail, Verily, VinBigData; and is co-founder of Genome Medical and Nurture Genomics.

References

1. Murray MF, Evans JP, Khoury MJ. DNA-Based Population Screening: Potential Suitability and Important Knowledge Gaps. *JAMA*. 2020;323(4):307-308. doi:10.1001/jama.2019.18640
2. Shen EC, Srinivasan S, Passero LE, et al. Barriers and Facilitators for Population Genetic Screening in Healthy Populations: A Systematic Review. *Front Genet*. 2022;13:865384. doi:10.3389/fgene.2022.865384
3. Rosenthal ET, Bernhisel R, Brown K, Kidd J, Manley S. Clinical testing with a panel of 25 genes associated with increased cancer risk results in a significant increase in clinically significant findings across a broad range of cancer histories. *Cancer Genet*. 2017;218-219:58-68. doi:10.1016/j.cancergen.2017.09.003
4. Williams MS. Early Lessons from the Implementation of Genomic Medicine Programs. *Annu Rev Genomics Hum Genet*. 2019;20(1):389-411. doi:10.1146/annurev-genom-083118-014924
5. Buchanan AH, Lester Kirchner H, Schwartz MLB, et al. Clinical outcomes of a genomic screening program for actionable genetic conditions. *Genet Med*. 2020;22(11):1874-1882. doi:10.1038/s41436-020-0876-4
6. Blout Zawatsky CL, Shah N, Machini K, et al. Returning actionable genomic results in a research biobank: Analytic validity, clinical implementation, and resource utilization. *Am J Hum Genet*. 2021;108(12):2224-2237. doi:10.1016/j.ajhg.2021.10.005
7. Forrest IS, Chaudhary K, Vy HMT, et al. Population-Based Penetrance of Deleterious Clinical Variants. *JAMA*. 2022;327(4):350-359. doi:10.1001/jama.2021.23686
8. Berry DA, Parmigiani G. Assessing the benefits of testing for breast cancer susceptibility genes: a decision analysis. *Breast Dis*. 1998;10(1-2):115-125. doi:10.3233/bd-1998-101-213
9. Ding Y, Hou K, Burch KS, et al. Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification. *Nat Genet*. 2022;54(1):30-39. doi:10.1038/s41588-021-00961-5
10. Hall MJ, Forman AD, Pilarski R, Wiesner G, Giri VN. Gene Panel Testing for Inherited Cancer Risk. *J Natl Compr Canc Netw*. 2014;12(9):1339-1346. doi:10.6004/jnccn.2014.0128
11. Xue Y, Ankala A, Wilcox WR, Hegde MR. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genet Med*. 2015;17(6):444-451. doi:10.1038/gim.2014.122
12. Shashi V, McConkie-Rosell A, Rosell B, et al. The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genet Med*. 2014;16(2):176-182. doi:10.1038/gim.2013.99
13. Mazzarotto F, Girolami F, Boschi B, et al. Defining the diagnostic effectiveness of genes for inclusion in panels: the experience of two decades of genetic testing for hypertrophic cardiomyopathy at a single center. *Genet Med*. 2019;21(2):284-292. doi:10.1038/s41436-018-0046-0
14. Dorling L, Carvalho S, Allen J, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med*. 2021;384(5):428-439. doi:10.1056/nejmoa1913948
15. Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N Engl J Med*. 2021;384(5):440-451. doi:10.1056/NEJMoa2005936
16. Manahan ER, Kuerer HM, Sebastian M, et al. Consensus Guidelines on Genetic` Testing for Hereditary Breast Cancer from the American Society of Breast Surgeons. *Ann Surg Oncol*. 2019;26(10):3025-3031. doi:10.1245/s10434-019-07549-8

17. Elezaby M, Lees B, Maturen KE, et al. BRCA Mutation Carriers: Breast and Ovarian Cancer Screening Guidelines and Imaging Considerations. *Radiology*. 2019;291(3):554-569. doi:10.1148/radiol.2019181814
18. Schwartz MD, Isaacs C, Graves KD, et al. Long-term outcomes of BRCA1/BRCA2 testing: risk reduction and surveillance. *Cancer*. 2012;118(2):510-517. doi:10.1002/cncr.26294
19. Lee A, Moon BI, Kim TH. BRCA1/BRCA2 Pathogenic Variant Breast Cancer: Treatment and Prevention Strategies. *Ann Lab Med*. 2020;40(2):114-121. doi:10.3343/alm.2020.40.2.114
20. Bergstrom C, Pence C, Berg J, et al. Clinicopathological Features and Outcomes in Individuals with Breast Cancer and ATM, CHEK2, or PALB2 Mutations. *Ann Surg Oncol*. 2021;28(6):3383-3393. doi:10.1245/s10434-020-09158-2
21. Cragun D, Weidner A, Tezak A, Clouse K, Pal T. Cancer risk management among female BRCA1/2, PALB2, CHEK2, and ATM carriers. *Breast Cancer Res Treat*. 2020;182(2):421-428. doi:10.1007/s10549-020-05699-y
22. Filippini SE, Vega A. Breast cancer genes: beyond BRCA1 and BRCA2. *Front Biosci Landmark Ed*. 2013;18(4):1358-1372. doi:10.2741/4185
23. Byrnes GB, Southey MC, Hopper JL. Are the so-called low penetrance breast cancer genes, ATM, BRIP1, PALB2 and CHEK2, high risk for women with strong family histories? *Breast Cancer Res*. 2008;10(3):208. doi:10.1186/bcr2099
24. Antoniou AC, Pharoah PDP, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer*. 2002;86(1):76-83. doi:10.1038/sj.bjc.6600008
25. Dullens B, de Putter R, Lambertini M, et al. Cancer Surveillance in Healthy Carriers of Germline Pathogenic Variants in BRCA1/2: A Review of Secondary Prevention Guidelines. *J Oncol*. 2020;2020:9873954. doi:10.1155/2020/9873954
26. Krassuski L, Vennedey V, Stock S, Kautz-Freimuth S. Effectiveness of decision aids for female BRCA1 and BRCA2 mutation carriers: a systematic review. *BMC Med Inform Decis Mak*. 2019;19(1):154. doi:10.1186/s12911-019-0872-2
27. Lee AJ, Cunningham AP, Tischkowitz M, et al. Incorporating truncating variants in PALB2, CHEK2, and ATM into the BOADICEA breast cancer risk model. *Genet Med*. 2016;18(12):1190-1198. doi:10.1038/gim.2016.31
28. Braun D, Yang J, Griffin M, Parmigiani G, Hughes KS. A Clinical Decision Support Tool to Predict Cancer Risk for Commonly Tested Cancer-Related Germline Mutations. *J Genet Couns*. 2018;27(5):1187-1199. doi:10.1007/s10897-018-0238-4
29. Chen J, Bae E, Zhang L, et al. Penetrance of Breast and Ovarian Cancer in Women Who Carry a BRCA1/2 Mutation and Do Not Use Risk-Reducing Salpingo-Oophorectomy: An Updated Meta-Analysis. *JNCI Cancer Spectr*. 2020;4(4):pkaa029. doi:10.1093/jncics/pkaa029
30. Marabelli M, Cheng SC, Parmigiani G. Penetrance of ATM Gene Mutations in Breast Cancer: A Meta-Analysis of Different Measures of Risk. *Genet Epidemiol*. 2016;40(5):425-431. doi:10.1002/gepi.21971
31. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and Tumor Subtype–Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. *J Clin Oncol*. 2016;34(23):2750-2760. doi:10.1200/JCO.2016.66.5844
32. Antoniou AC, Casadei S, Heikkinen T, et al. Breast-Cancer Risk in Families with Mutations in PALB2. *N Engl J Med*. 2014;371(6):497-506. doi:10.1056/NEJMoa1400382
33. Rehm HL, Berg JS, Brooks LD, et al. ClinGen — The Clinical Genome Resource. *N Engl J Med*. 2015;372(23):2235-2242. doi:10.1056/NEJMsr1406261
34. van Lier MGF, Wagner A, Mathus-Vliegen EMH, Kuipers EJ, Steyerberg EW, van Leerdam ME. High Cancer Risk in Peutz–Jeghers Syndrome: A Systematic Review and

Surveillance Recommendations. *Off J Am Coll Gastroenterol ACG*. 2010;105(6):1258-1264. doi:10.1038/ajg.2009.725
35. Furukawa Y. *GTP Cyclohydrolase 1-Deficient Dopa-Responsive Dystonia*. University of Washington, Seattle, Seattle (WA); 1993. http://europepmc.org/books/NBK1508
36. Vasen HFA, Blanco I, Aktan-Collan K, et al. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. *Gut*. 2013;62(6):812-823. doi:10.1136/gutjnl-2012-304356
37. Yeh JM, Stout NK, Chaudhry A, et al. Universal newborn genetic screening for pediatric cancer predisposition syndromes: model-based insights. *Genet Med*. 2021;23(7):1366-1371. doi:10.1038/s41436-021-01124-x
38. Guzauskas GF, Garbett S, Zhou Z, et al. Cost-effectiveness of Population-Wide Genomic Screening for Hereditary Breast and Ovarian Cancer in the United States. *JAMA Netw Open*. 2020;3(10):e2022874. doi:10.1001/jamanetworkopen.2020.22874
39. Zhang L, Bao Y, Riaz M, et al. Population genomic screening of all young adults in a health-care system: a cost-effectiveness analysis. *Genet Med*. 2019;21(9):1958-1968. doi:10.1038/s41436-019-0457-6
40. Bennette CS, Gallego CJ, Burke W, Jarvik GP, Veenstra DL. The cost-effectiveness of returning incidental findings from next-generation genomic sequencing. *Genet Med Off J Am Coll Med Genet*. 2015;17(7):587-595. doi:10.1038/gim.2014.156
41. Lee H, Deignan JL, Dorrani N, et al. Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders. *JAMA*. 2014;312(18):1880-1887. doi:10.1001/jama.2014.14604
42. Zhu X, Petrovski S, Xie P, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med Off J Am Coll Med Genet*. 2015;17(10):774-781. doi:10.1038/gim.2014.191
43. Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med*. 2022;14(1):23. doi:10.1186/s13073-022-01026-w
44. Schuler BA, Nelson ET, Koziura M, Cogan JD, Hamid R, Phillips JA. Lessons learned: next-generation sequencing applied to undiagnosed genetic diseases. *J Clin Invest*. 2022;132(7):e154942. doi:10.1172/JCI154942
45. Miller DT, Lee K, Chung WK, et al. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(8):1381-1390. doi:10.1038/s41436-021-01172-3
46. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249-255. doi:10.1038/gim.2016.190
47. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565-574. doi:10.1038/gim.2013.73
48. Botkin JR, Teutsch SM, Kaye CI, et al. Outcomes of interest in evidence-based evaluations of genetic tests. *Genet Med*. 2010;12(4):228-235. doi:10.1097/GIM.0b013e3181cdde04
49. Gao C, Polley EC, Hart SN, et al. Risk of Breast Cancer Among Carriers of Pathogenic Variants in Breast Cancer Predisposition Genes Varies by Polygenic Risk Score. *J Clin Oncol*. 2021;39(23):2564-2573. doi:10.1200/JCO.20.01992

Figure Legends

Figure 1: Net utilities from the female breast cancer application plotted against the ratio $\delta_{D=1,\,G=0}/\delta_{D=0,\,G=1}$ for each of the individual genes, as well as the aggregate Δ for all five genes. $\delta_{D=1,\,G=0}$ is allowed to vary from 0.1 to 10 in increments of $\log_{10}(0.1)$, while fixing $K = 0$ and $\delta_{D=0,\,G=1} = 1$. 95% credible intervals are shaded in gray and represent uncertainty contributed by the penetrance estimates taken from a literature review (see text). Dashed reference lines are drawn to indicate the value of the ratio at which the net utility changes from negative to positive. The red dotted curve represents the probability of a positive net utility, with respect to the uncertainty distribution, at each value of $\delta_{D=1,\,G=0}/\delta_{D=0,\,G=1}$.

Figure 2: Heatmaps of the five individual female breast cancer net utilities and the aggregate utility ("All") while holding $K = 0$ and varying $\delta_{D=0,\,G=1}$ and $\delta_{D=1,\,G=0}$ from 1 to 100 in increments of 5. Positive net utilities are shaded blue and negative utilities are shaded red. Utility thresholds based on the penetrance estimates are drawn as solid black lines. The dashed black lines have intercept 0 and slope 1, and correspond to cases where the individual testing positive for the gene but not developing the disease and vice versa have equal disutilities.

Figure 3: Net utilities for a single gene and disease plotted against the utility ratio ($\delta_{D_i=1,\,G_j=0}/\delta_{D_i=0,\,G_j=1}$), while varying the G-D+ disutility, disease penetrance, carrier prevalence, and precision parameter. Test and G+D- disutilities are fixed at 0 and 1, respectively. The colored shading represents 95% credible intervals for different values of the precision. Dashed reference lines are drawn to indicate the utility threshold in each scenario. The dotted curves represents the probability of a positive net utility, with respect to the uncertainty distribution for each precision level and at each value of $\delta_{D_i=1,\,G_j=0}/\delta_{D_i=0,\,G_j=1}$.

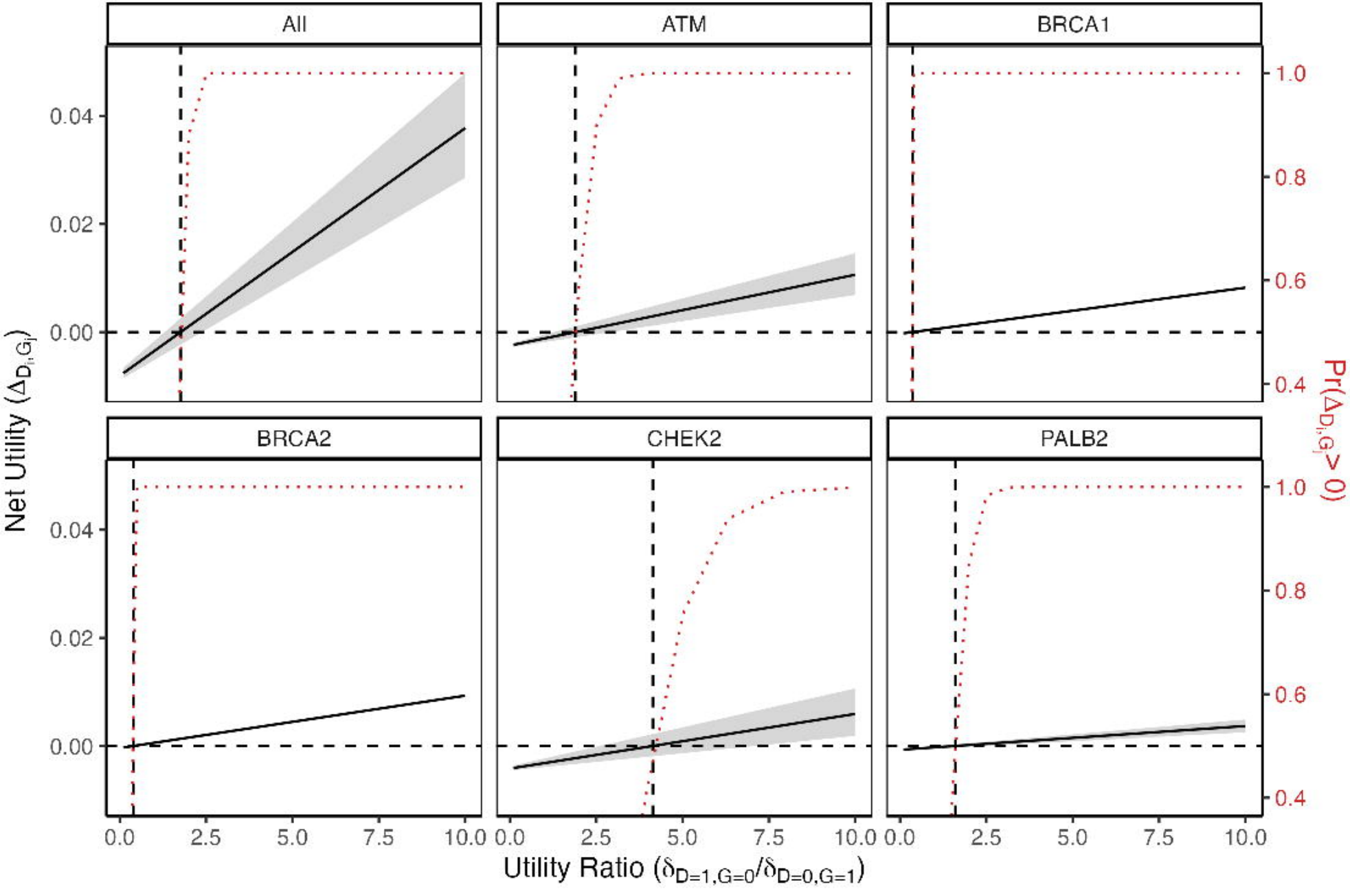
Tables

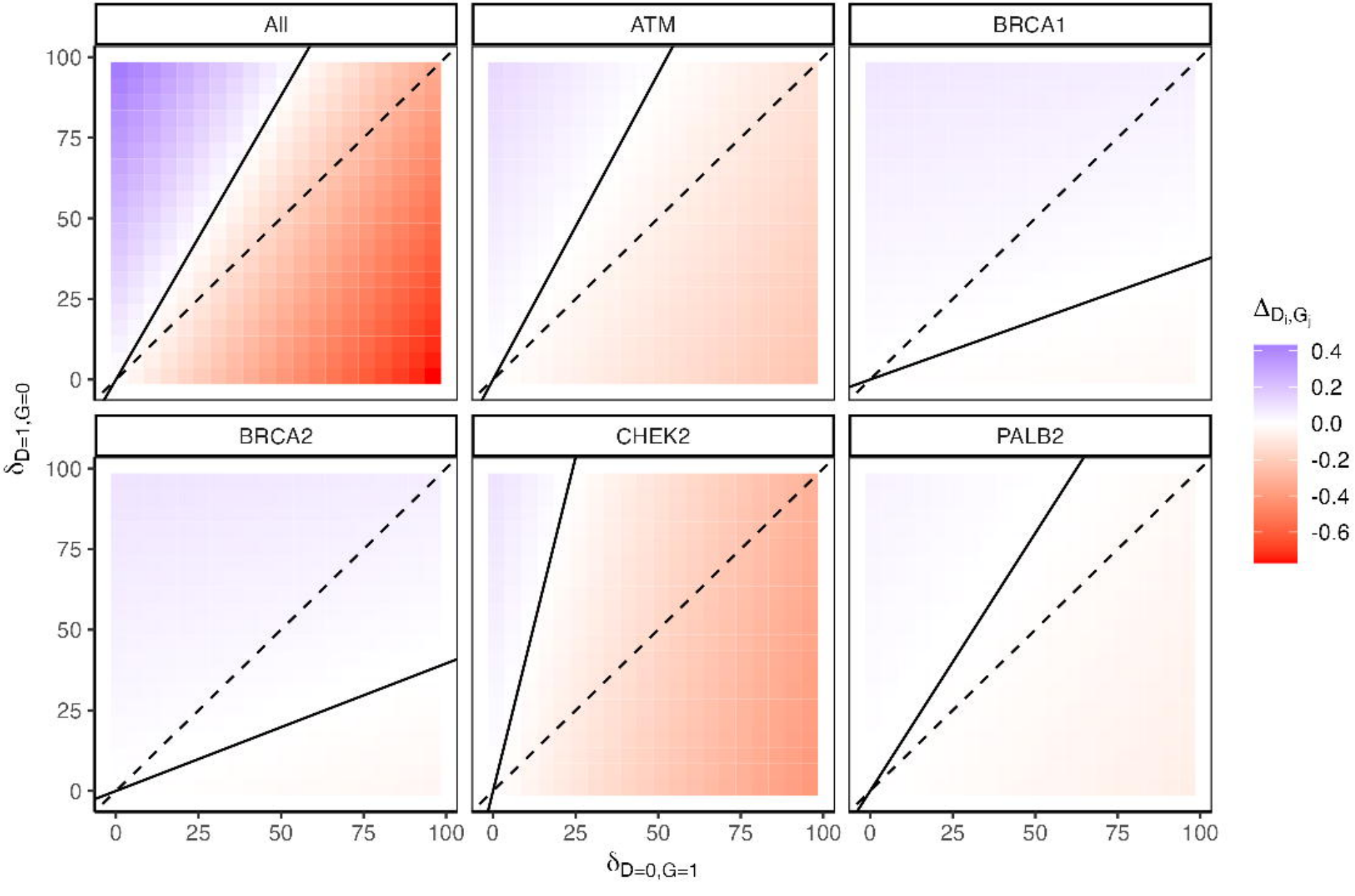
	Estimate	0.025	0.975
All	1.8	1.3	2.4
ATM	1.9	1.3	2.9
BRCA1	0.37	0.35	0.38
BRCA2	0.40	0.38	0.41
CHEK2	4.1	2.6	7.1
PALB2	1.6	1.08	2.42

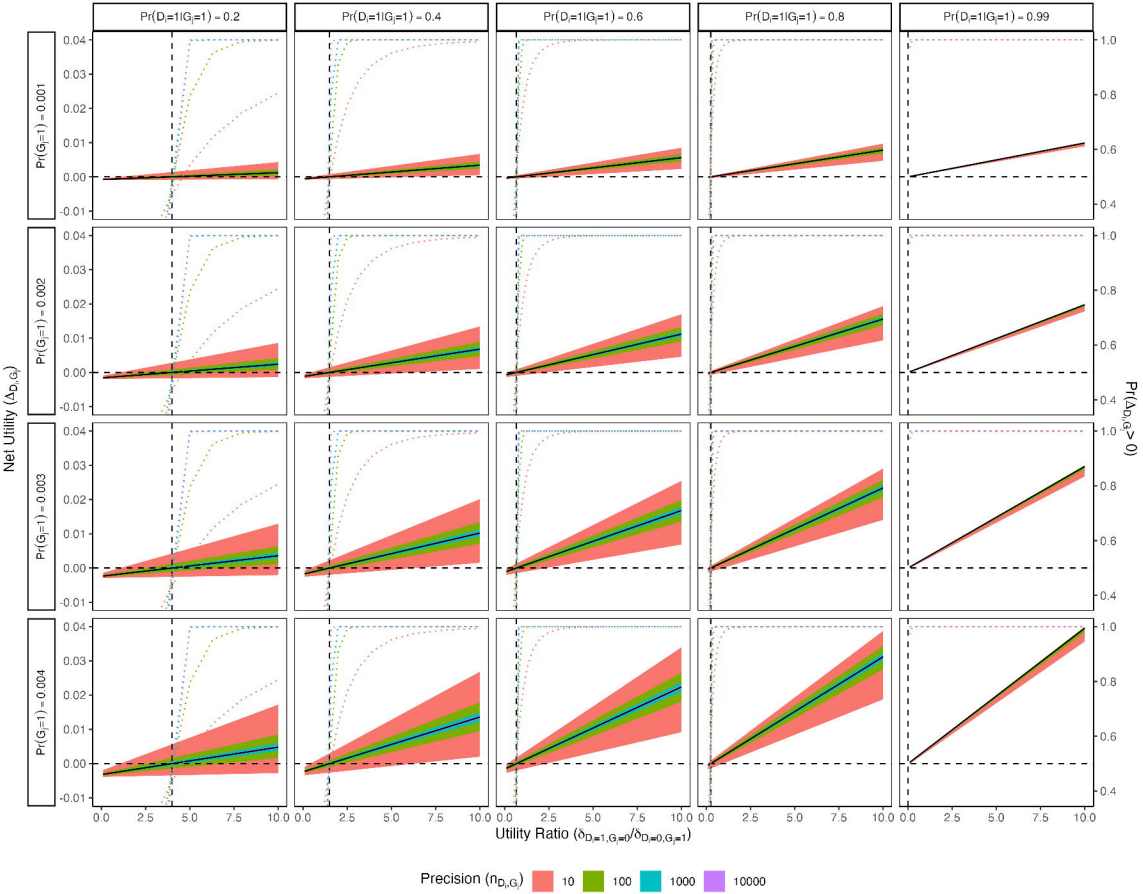
Table 1: Female breast cancer application utility thresholds $b_{D_i,G_j} = \delta_{D_i=1,\, G_j=0} / \delta_{D_i=0,\, G_j=1}$ at which the net utility is 0 for the five individual genes and overall ("All"). Lower threshold values (below 1) indicate that a positive net utility can be achieved with a lower G-D+ disutility, relative to the G+D- disutility. The upper and lower bounds for a 95% credible interval are also reported.

Penetrance	Precision	Estimate	0.025	0.975
0.2	**10**	4	1.1	35
	100		2.5	6.8
	1000		3.4	4.7
	10000		3.8	4.2
0.4	**10**	1.5	0.43	6.3
	100		1.0	2.3
	1000		1.3	1.7
	10000		1.4	1.6
0.6	**10**	0.67	0.16	2.3
	100		0.44	0.99
	1000		0.59	0.76
	10000		0.64	0.69
0.8	**10**	0.25	0.029	0.93
	100		0.15	0.40
	1000		0.21	0.29
	10000		0.24	0.26
0.99	**10**	0.01	0.000	0.11
	100		0.000	0.038
	1000		0.005	0.017
	10000		0.008	0.012

Table 2: Utility thresholds $b_{D_i,G_j} = \delta_{D_i=1,\, G_j=0} / \delta_{D_i=0,\, G_j=1}$ at which the net utility is 0 while varying the G-D+ disutility, disease penetrance, carrier prevalence, and precision parameter. Lower threshold values (below 1) indicate that a positive net utility can be achieved with a lower G-D+ disutility, relative to the G+D- disutility. The upper and lower bounds for a 95% credible interval are also reported.


All
ATM
BRCA1
BRCA2
CHEK2
PALB2
Net Utility (Δ_{D_i,G_j})
$Pr(\Delta_{D_i,G_j} > 0)$
0.04
0.02
0.00
1.0
0.8
0.6
0.4
0.0
2.5
5.0
7.5
10.0
Utility Ratio ($\delta_{D=1,G=0}/\delta_{D=0,G=1}$)


All
ATM
BRCA1
BRCA2
CHEK2
PALB2
100
75
50
25
0
$\delta_{D=1,G=0}$
$\delta_{D=0,G=1}$
Δ_{D_i,G_j}
0.4
0.2
0.0
-0.2
-0.4
-0.6


Pr(D_i=1|G_i=1) = 0.2
Pr(D_i=1|G_i=1) = 0.4
Pr(D_i=1|G_i=1) = 0.6
Pr(D_i=1|G_i=1) = 0.8
Pr(D_i=1|G_i=1) = 0.99
Pr(G_i=1) = 0.001
Pr(G_i=1) = 0.002
Pr(G_i=1) = 0.003
Pr(G_i=1) = 0.004
Net Utility (Δ_{D_i,G_i})
Pr(Δ_{D_i,G_i} > 0)
Utility Ratio (δ_{D_i=1,G_i=0}/δ_{D_i=0,G_i=1})
Precision (n_{D_i,G_i})
10
100
1000
10000