Incorporating calibrated functional assay data into the *BRCA1* Ex-UV database

Bryony A. Thompson,[1,2] Russell Bell,[1] Bryan E. Welm,[3] John Burn,[4] Sean V. Tavtigian[1]

[1] Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah

[2] Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Melbourne, Australia

[3] Department of Surgery, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah

4 Institute of Genetic Medicine, International Centre for Life, Newcastle University, Newcastle upon Tyne, United Kingdom

Correspondence:

Sean V. Tavtigian, PhD

Professor of Oncological Science

Co-Leader, Cancer Control and Population Sciences Program

Huntsman Cancer Institute at the University of Utah

2000 Circle of Hope, Room 3262, Salt Lake City, UT 84112-5550

Email: Sean.Tavtigian@hci.utah.edu

Acknowledgements: BAT is a National Health and Medical Research Council CJ Martin Early Career Fellow. SVT is supported in part by the US NCI Cancer Center Support Grant P30 CA042014. Support was also provided by Astra Zeneca via Newcastle University, UK, as part of the BRCA Challenge, a flagship project of the Global Alliance for Genomics and Health (GA4GH).

Disclosure statement: The authors have no conflicts of interest to declare.

Abstract:

Driven by massively parallel sequencing and allied technologies, the scale of genetic predisposition testing is on a dramatic uptrend. While many patients are found to carry clinically actionable pathogenic sequence variants, testing also reveals enormous numbers of Unclassified Variants (UV), or Variants of Uncertain Significance (VUS), most of which are rare missense substitutions. Following IARC variant classification guidelines, quantitative methods have been developed to integrate multiple data types for clinical UV evaluation in *BRCA1/2*; results from these analyses are recorded in the BRCA gene Ex-UV database (hci-exlovd.hci.utah.edu). In variant classification the rate-limiting step is often accumulation of patient observational data. Recently, functional assays evaluating *BRCA1* RING domain and C-terminal substitutions have been calibrated, enabling variant classification through a two-component combination of sequence analysis-based predictions with functional assay results. This two-component classification was embedded in a decision tree with safeguards to avoid misclassification. For the two-component analysis, sensitivity is 89.5%, specificity is 100%, and the error rate 0.0%. Classification of UV as likely pathogenic or likely neutral does not require certainty – the probabilistic definitions of the categories imply an error rate. Combining sequence analysis with functional assay data in two-component analysis added 154 *BRCA1* variants to the Ex-UV database.

Introduction

Inactivating mutations in *BRCA1* (MIM #113705) are associated with an increased risk of breast and ovarian cancer. While many patients who undergo clinical mutation testing are found to carry clinically actionable pathogenic sequence variants, testing also reveals large numbers of Unclassified Variants (UV), or Variants of Uncertain Significance (VUS), most of which are rare missense substitutions. All known pathogenic non-spliceogenic *BRCA1* missense substitutions are found in the N-terminal RING domain and C-terminal BRCT domains (Easton et al. 2007; Lindor et al. 2012; Vallée et al. 2012). The RING domain binds BARD1 to act as an E3 ubiquitin ligase (Densham et al. 2016), and the BRCT domains bind a number of proteins involved in the DNA damage response (reviewed in Mermershtain and Glover 2013). Many functional assays have been developed to test the impact of missense substitutions on these functionally important domains, and other regions of the BRCA1 protein.

The BRCA gene Ex-UV database (hci-exlovd.hci.utah.edu) contains *BRCA1* and *BRCA2* UV clinically classified or reclassified using a quantitative Integrated Evaluation (QIE) (Vallée et al. 2012). The classifications are based on calibrated sequence analysis to generate a prior probability of pathogenicity (Prior P), patient observational data to generate odds in favor of pathogenicity, and Bayes rule to combine these data into a posterior probability of pathogenicity (Posterior P). For purposes of clinical translation and patient counselling/management, the Posterior P is translated through the IARC 5-category classifier (Plon et al. 2008). Currently the database includes fields for sequence analysis-based Prior Ps for missense analysis (Tavtigian et al. 2008), and predicted effects on splicing (Vallée et al. 2016). It also includes odds in favor of pathogenicity in the form of likelihood ratios (LR) for co-segregation with cancer phenotype (Thompson et al. 2003), strength of personal and family cancer history (Easton et al. 2007),

tumor pathology (Chenevix-Trench et al. 2006; Spurdle et al. 2008, 2014), and co-occurrence with known pathogenic variants (Goldgar et al. 2004); the product of these four LRs is the observational LR in favor of pathogenicity (Obs LR).

The rate-limiting step in UV classification has been accumulation of sufficient Obs LR from patient observational data to modify the Prior P into a Posterior P >0.95 (likely pathogenic) or <0.05 (likely not pathogenic/likely neutral). Functional assay estimates of damage to protein function can be used as a proxy for pathogenicity if the assay outputs are empirically calibrated (either against known pathogenic and know neutral variants or against patient observational data) to read out a variable that is more directly related to pathogenicity e.g., a functional likelihood ratio in favor of pathogenicity (Fxnl LR). Once such a calibration is in place and the functional assay(s) are considered sufficiently accurate for use in clinical variant classification, the assay(s) can become directly useful for medical genetics related evaluation of missense substitutions and other UVs. In principle, the QIE enables classification through a two-component combination of sequence-based *in silico* analyses and a calibrated functional assay. In the companion article [Paquette et al., manuscript in preparation] we describe the calibration of a mammalian two-hybrid assay testing BARD1 binding capabilities of BRCA1 RING domain substitutions. In addition, Woods et al. (Woods et al. 2016) recently described the calibration of mammalian transcriptional activation assays for evaluation of missense substitutions in the BRCA1 C-terminal region. Here we report an update of the *BRCA1* Ex-UV database to incorporate a Fxnl LR based on these calibrated functional assays into the Integrated Evaluation of *BRCA1* sequence variants. This has led to the addition of 154 variants to the database.

Methods

Database updates

Three new fields have been added to the LOVD *BRCA1* Ex-UV database (hci-exlovd.hci.utah.edu, under the gene database BRCA1fx): Functional LR, Functional Assay Reference, and Clinical Severity. The Clinical Severity field will be used to differentiate pathogenic variants (i.e. confers medically actionable increased risk) associated with different levels of risk. The values used to populate Clinical Severity are high-risk, moderate-risk, and not applicable. The Clinical Severity category is assigned to a variant based on the estimated cumulative risks to age 80: high-risk >32%; moderate-risk 18-32% (Easton et al. 2015). Variants assigned to Class 5 or Class 4 by QIE are assumed to be high-risk unless patient observational data in the form of co-segregation with the cancer phenotype, strength of personal and family cancer history, or a case-control odds ratio are indicative of moderate-risk. At this time, p.R1699Q is the only proven moderate-risk *BRCA1* variant (Spurdle et al. 2012). Clinical Severity for variants assigned to Class 3, Class 2, or Class 1 by QIE are marked not applicable. The other functional assay-related LOVD fields are populated based on the functional assay data available in the 'datasets' section described below.

Datasets

Only missense substitutions that have been observed in cases were added to the LOVD database, due to assumed contribution of a minimal observational datum (personal history of cancer) in the derivation of the sequence analysis Prior P and Fxnl LR. For the BRCA1 RING domain (aa 2–103), there were 11 natural (i.e. originally identified in cancer cases) already classified missense substitutions recorded in the Ex-UV database. However, because only two were neutral, the assay calibration described in the companion research article [Paquette et al., manuscript in preparation] included three human Class 3 substitutions with Obs LR <0.50, plus

four cross-species multiple sequence alignment neutral missense substitutions (i.e. the alternate amino acid is present in at least two mammals). To avoid circularities, these were not classified in Paquette et al [manuscript in preparation] and are not included here. There are 242 'natural' *BRCA1* missense substitutions between amino acids 1,315-1,863 evaluated in Woods et al (Woods et al. 2016). These include 13 in the coiled-coil domain (aa 1,392-1,424); 140 over the BRCT1, linker, and BRCT2 domains (aa 1,646-1,855); and 89 in the interval between the coiled-coil and BRCT1. All 242 of these are included here.

The missense analysis Prior P and splicing Prior P for each variant are taken from the HCI Breast Cancer Genes Prior Probabilities database (hci-priors.hci.utah.edu/PRIORS) (Vallée et al. 2016). The derivation of the functional Odds Path for the BRCA1 RING domain missense variants is described in the companion article [Paquette et al, manuscript in preparation]. For the missense variants in Woods *et al*. (Woods et al. 2016) the supplementary materials provided the data used to derive the Fxnl LR. The Woods *et al*. team used a Bayesian hierarchical model that takes into account experimental heterogeneity in the *in cellulo* functional assay to estimate log odds in favor of pathogenicity for each variant tested (Iversen et al. 2011; Woods et al. 2016). The Odds Path was then calculated by exponentiation of the reported log odds (also referred to as a Bayes factor).

Statistical analyses

Sensitivity was calculated as (# true pathogenic variants detected by a classifier / total # true pathogenic variants). Specificity was calculated as (# true neutral variants detected by a classifier / total # true neutral variants). Error rate was calculated as [(# false pathogenic variants + # false neutral variants detected by a classifier) / (total # true pathogenic variants + # true

neutral variants)]. 95% confidence intervals on these proportions were estimated using the binomial confidence interval calculator in STATA 13.1 (StataCorp).

Results and Discussion

Embedding quantitative integrated evaluation, with inclusion of calibrated functional assays, in a decision tree.

Sequence analysis-based Prior Ps are available for all missense substitutions in *BRCA1* that can be reached by a single nucleotide substitution. Any of these substitutions can be subjected to QIE if one or more of four types of patient observational data are available: personal and family cancer history, co-segregation with cancer phenotype, co-occurrence between UVs and clearly pathogenic variants, and/or tumor immunohistochemistry & grade (Lindor et al. 2012; Vallée et al. 2012). Here, we add calibrated function assays to the mix. To guide the analysis of individual substitutions while providing some safeguards against misclassification, we embed the QIE in a decision tree (Figure 1). Logic behind key nodes of the tree is explained in the next few paragraphs. The discussion proceeds from the assumption that valid functional assay data exists (node 1).

Inclusion of patient observational data (node 2). Recently, Vallee et al. asked the philosophical question of how much observational data are required in order to perform a *bona fide* integrated evaluation, and concluded that Obs LR of ≤0.5 or ≥2 are reasonable inner boundaries for the magnitude of observational data required (Vallée et al. 2016). Noting that some variants, especially moderate-risk variants such as *BRCA1* p.R1699Q, can be resistant to classification because the observational data can be conflicting, a more refined criterion would be that a multicomponent QIE with observational data requires either that the Obs LR criterion

of ≤0.5 or ≥2 is met, or else at least enough observational data to expect an Obs LR ≥2 if the variant were pathogenic. As examples of the expected Obs LR, tumor estrogen receptor (ER) status is an extremely effective data type because ~75% of breast tumors from *BRCA1* mutation carriers are ER-, and estrogen receptor negativity provides an Obs LR of 2.6 in favor of pathogenicity (so long as the case was not ascertained because of their tumor's immunohistochemical profile) (Spurdle et al. 2014). Therefore, data from just one tumor is enough to meet the expected Obs LR criterion. At the other extreme, analysis of co-observation between UVs and clearly pathogenic variants is relatively ineffective for classification as pathogenic because a specific *BRCA1* UV has to be observed about 17 times in the absence of a pathogenic variants in order to generate an Obs LR of 2.0 in favor of pathogenicity (Goldgar et al. 2004; Tavtigian et al. 2006). If the observational/ expected observational LR criterion is met, multicomponent QIE is applied; otherwise, analysis moves to node 3 and a two-component evaluation as described in the next paragraph.

Currently there are two functional assays evaluating the impact of *BRCA1* missense substitutions that have been calibrated. These are a mammalian two-hybrid assay testing the RING domain [Paquette et al., manuscript in preparation], and a transcriptional activation assay used to test many reported missense substitutions in the BRCA1 C-terminal region (Iversen et al. 2011; Woods et al. 2016). At the first step of two-component classification (node 3), a simple safeguard against error could be that if the missense analysis Prior P and the Fxnl LR are discordant, then classification is not permitted. Accordingly, consider a variant where the missense analysis Prior P is 0.49 and the Fxnl LR is 1.04. One might say that these data are discordant because they are on opposite sides of neutrality. But both results are saying the same thing… “we don’t know, Uncertain”. So it would be perfectly reasonable to combine these into

a Posterior P, which would be 0.50. A more interesting situation would be if the missense analysis Prior P gave 2-fold evidence against pathogenicity (naively, Prior P = 0.33) and functional assay gave 2-fold evidence in favor (Fxnl LR = 2.0). Here, we think that it's fair to say that they "disagree". So, we propose that if the missense analysis Prior P and Fxnl LR are on opposite sides of a factor-of-2 window around neutrality, we will NOT combine them via a two-component QIE, and the missense substitution remains Unclassified. For *BRCA1*, a small complication is that the background Prior P for an individual rare missense substitution is approximately 0.1, not 0.5 (Abkevich et al. 2004; Goldgar et al. 2004). Thus the definitions for non-discordant combinations are missense substitution sequence analysis Prior P <0.05 and Fxnl LR ≤2.00, or Prior P >0.18 and Fxnl LR ≥0.50 (Figure 1).

In general, there are two ways for a missense substitution to be pathogenic: the missense substitution can damage a key function of the protein, or the underlying nucleotide substitution can alter mRNA splicing in a way that results in either a non-productive transcript or damage to a key function of the protein. The missense analysis Prior P and Fxnl LR only address damage to a key function of the protein. Accordingly, consider a missense substitution where the Bayesian combination of the missense analysis Prior P and Fxnl LR naively results in a probability of pathogenicity <0.05, but the mRNA splicing analysis Prior P (Splice Prior P) (Vallee et al. 2016) is >0.05 (Figure 1, node 4). Since the functional assays evaluated here do not assess spliceogenicity, the overall probability of pathogenicity remains >0.05 and is given by the Splice Prior P. In this case, the missense substitution remains Unclassified. Note that this issue is explicitly asymmetric. If the Bayesian combination of the missense analysis Prior P and Fxnl LR results in a probability of pathogenicity >0.95, but the Splice Prior P <0.05, the substitution has Post P >0.95 and can be considered IARC Class 4, Likely Pathogenic. For

missense substitutions that pass node 4, the Post P is used to place the substitution on a truncated 3 Class version of the IARC classifier.

Class 4 - Likely pathogenic: Posterior P >0.95

Class 3 - Uncertain: Posterior P ≥0.05 & ≤0.95

Class 2 - Likely not pathogenic: Posterior P <0.05

Moving back to node 2, if the observational/ expected Obs LR criterion is met, multicomponent QIE can be applied and the analysis moves to node 5. Here, the question is much the same as at node 4: which is higher, the Bayesian combination of the missense analysis Prior P and Fxnl LR, or the Splice Prior P? The Obs LR from patient data is applied to the higher of these two probabilities to calculate the missense substitution's Post P. The Post P is then used to place the substitution in the standard 5 Class IARC classifier (Plon et al. 2008).

Performance on previously classified variants with patient observational data.

Between Paquette et al [manuscript in preparation] and Woods et al (2016), there were 56 previously classified missense substitutions with valid functional assay data that met the observational/ expected Obs LR criterion. Twenty four of 24 Class 5 substitutions remained in Class 5. Of two substitutions previously placed in Class 4, one remained in Class 4 and the other moved up to Class 5. Of three substitutions previously placed in Class 3, two moved up to Class 5 and one moved up to Class 4. Both of the substitutions previously placed in Class 2 moved down to Class 1, and all 25 previously placed in Class 1 remained in Class 1.

To get some idea of the performance of classification based on sequence analysis Prior Ps and Fxnl LRs alone, one can ask what the fate of these 56 variants would have been if they were subjected to the two-component evaluation. Fifteen of these were explicitly used as internal standards during assay calibration; and were excluded them from our analyses to avoid

circularities. Two of the known neutral substitutions (p.M1652T and p.G1706A) would have been remanded for further analysis because their missense Prior P and Fxnl LR were discordant (Table 1). If two-component QIE were applied to the remaining 41, two known pathogenic substitutions (p.R1495M and p.A1623G) would have been remanded for further analysis because their splicing Prior P was >0.05 and their splicing Prior P was greater than the Bayesian combination of the missense Prior P and the Functional LR (Table 1). The remaining 39 substitutions – 17 pathogenic and 22 neutral – would all have been classified correctly. Including the two remanded substitutions, sensitivity was 89.5% (95% CI: 67%- 99%), specificity was 100.0% (85% - 100%), and the classification rate was 95.1% (83% - 99%). The error rate – classification of known neutral variants as pathogenic or vice versa – was 0.0%.

Application to variants without patient observational data.

From Woods et al (2016), an additional 197 UV missense substitutions had been subjected to a functional assay and were available for two-component QIE. Of these, 30 failed the concordance test and were remanded for further analyses (Table 2). An additional 14 had splicing Prior Ps >0.05 and their splicing Prior P was greater than the Bayesian combination of the missense Prior P and the Functional LR; these were also remanded for further analyses (Table 2). Of the remaining 153 missense substitutions, 41 had two-component QIE Posterior Ps >0.95 and were classified as Class 4, Likely Pathogenic; 4 had Posterior P ≥0.05 and <0.95 and were classified as Class 3, Uncertain, and 108 had Posterior Ps <0.05 and were classified as Class 2, Likely Not Pathogenic.

Since there was no evidence for error in the two-component QIE analysis of known neutral or pathogenic missense substitutions, we explored how sensitive two-component QIE would be to misspecification of either the missense analysis Prior Ps or the Fxnl LRs. For this

analysis, we assumed that the total number of misclassified variants could be estimated from their Posterior Ps, if these were accurate. For groups of pathogenic variants, the estimated number or errors would be given by equation 1, and for groups of neutral variants by equation 2:

Eq1 $\sum_{1}^{n} 1 - PP_n$ Eq2 $\sum_{1}^{n} PP_n$

Where ***n*** is an index number for each variant in the group and ***PP*** is the Posterior P for each such variant. The misclassification rate would simply be the estimated number of misclassified variants divided by ***n***, with the understanding that a rate >0.05 would violate the IARC guideline for the likely neutral or likely pathogenic categories.

For the 41 likely pathogenic substitutions, the estimated error rate was 0.0026. Testing the consequence of the possibility that the missense Prior Ps are over-estimated, lowering the Align-GVGD prior probabilities from their published point-estimates (0.81, 0.66, and 0.29 for C65, C55 to C35, and C25 & C15, respectively) (Tavtigian et al. 2008) to the published 95% lower bound of their confidence intervals (0.61, 0.31, and 0.09, in the same order) increased the expected error rate to 0.0082. As a second alternative, the Fxnl LRs could be inflated. Testing the consequence of this possibility, we estimated the error rate as a function of Fxnl LR deflation (Figure 2). If the Fxnl LR were systematically inflated by a factor of 11, then the expected error rate would be ~0.025, and if it were inflated by a factor of 32, then the expected error rate would exceed 0.050 (Figure 2A). This latter estimate is equivalent to saying that if the *reported* Fxnl LR was 32.0, but the *actual* LR is ~1.0; and if the *reported* LR is 1,000, but the *actual* LR is ~ 31.3; then the error rate would exceed 0.05. As a third alternative, the missense Prior P and Fxnl LR could be partially non-independent, so that combining their full magnitude via Bayes rule over-estimates the Post P. Since the sequence analysis Prior Ps were calibrated several years before the functional assays, this was tested by reducing the magnitude of the Fxnl LR used to

calculate the Post P. The linear deflation described above was one test. As the the Fxnl LRs were reported in Woods et al (2016) as log(LR), we also tested linear deflation of the log(LR), which is described as "root deflation of the Fxnl LR" in Figure 2B. The Fxnl LRs would have to have been inflated to the 2.8th power to produce an error rate of 0.025 and to the 4.3rd power for the error rate to exceed 0.05.

The 108 likely neutral substitutions break into two sub-groups: 72 falling upstream of the BRCT-linker-BRCT2 key domain, and 36 falling within the key domain (see Figure 3). The 72 upstream substitutions fall in an interval where there is no evidence for pathogenic missense substitutions except for those that damage mRNA splicing, so we did not perform any analysis. The 36 in-domain substitutions have an estimated error rate of 0.0028. These are all Align-GVGD C0 substitutions with missense Prior Ps of 0.03. Testing the possibility that this missense Prior P is under-estimated, increasing it to the published 95% upper bound of the confidence interval (0.05) increased the expected error rate to 0.0047; if the actual Prior P was 0.24, then the estimated error rate would be ~0.025 (Figure 2C), and if the prior was 0.45 the error rate would be ~0.05. Alternatively, the functional LRs could be deflated. Testing the consequence of this possibility, we estimated the error rate as a function of Fxnl LR re-flation (Figure 3) If the Fxnl LR were systematically deflated by a factor of 11, then the expected error rate would be 0.025 (Figure 2A), and if it were deflated by a factor of 27, then the expected error rate would exceed 0.050. This latter estimate is equivalent to saying that if the *reported* Fxnl LR was 0.037, but the *actual* LR is ~1.0; and if the *reported* LR is 0.001, but the *actual* LR is ~ 0.027; then the error rate would exceed 0.05. Since the missense Prior P point estimate for these substitutions is below 0.05, and almost all of the relevant Fxnl LRs are <1.0, linear inflation of the log(LR),

simulating non-independence between the missense Prior P and the Fxnl LR does not model a false classification rate.

Safeguards against UV misclassification built into the decision tree and quantitative integrated evaluation with functional assay likelihood ratios.

As integrated within the decision tree of Figure 1, QIE with Fxnl LRs includes four specific safeguards designed to reduce the risk of UV misclassification. Each will also result in some UVs that would have been correctly classified instead resting unclassified pending accumulation of further data.

1. The Fxnl LR for variants with severe loss of function is capped at 1,000, and the Fxnl LR for a variant with full function is capped at 0.01. These caps prevent stand-alone classification by the Fxnl LR in a multi-component QIE. For example, a key domain Align-GVGD C65 missense substitution has a missense Prior P of 0.81. If such a variant had a minimally concordant Obs LR of 2.0, it would have a Posterior P of 0.895 – trending towards but not quite reaching Likely Pathogenic. But a Fxnl LR of ≤0.0061 (strongly indicative of neutrality) would result in a new Posterior P of <0.050, reversing the combination of Prior P and observational data to a classification of Likely Neutral. A Fxnl LR cap at 0.01 prevents this stand-alone reversal. Similarly, a key domain Align-GVGD C0 missense substitution has a missense Prior P of 0.03. If such a variant had an Obs LR of 0.5, it would be classified as Class 2 Likely Not Pathogenic with a Posterior P of 0.0152. But a Fxnl LR of 1,230 would drive the Posterior P above 0.950 and reverse the classification to Likely Pathogenic. So a Fxnl LR cap at 1,000 prevents stand-alone reversal of the combined prior and observational data. These caps are somewhat *ad hoc*, tuned to the actual missense Priors Ps and minimal Obs LR data inclusion rules already embedded in the QIE (Tavtigian et al.

2008; Vallee et al. 2016). Nonetheless, these caps also correspond to the stand-alone classification criteria from the original framework for quantitative classification of *BRCA1/2* UVs (Goldgar et al. 2004).

2. Unclassified Variants are remanded for further analysis if their missense Prior P and Fxnl LR are discordant. For the missense Prior P, there are clear instances of substitutions with low Prior Ps that are nonetheless pathogenic because of missense dysfunction, and clear instances of substitutions with high missense Prior Ps that are clearly neutral or nearly so. On the Fxnl LR side, *BRCA1* p.R1699Q has a higher Fxnl LR than does p.R1699W. Yet evidence from personal and family cancer histories are consistent with p.R1699Q conferring a moderate-risk phenotype whereas p.R1699W is consistent with a high-risk phenotype (Spurdle et al 2012; Easton et al. 2007). Since it is therefore clear that the Fxnl LR is not a perfect predictor of disease risk, the discordance rule is a prudent precaution, albeit with a somewhat arbitrary definition.
3. The *BRCA1* missense substitutions p.L1407P and p.M1411T present a more interesting issue: discordance between the missense Prior P and the Fxnl LR because of a key domain specification issue. Both substitutions fall in the BRCA1 coiled-coil domain (Figure 3) that mediates interaction between BRCA1 and PALB2. Both are predicted to be deleterious (align GVGD score = C65), and demonstrate a damaging effect in functional assays (Fxnl LR >6). Both variants abrogate PALB2 binding to the coiled-coil domain, and there is evidence that damaging this interaction compromises repair of DNA double strand breaks by homologous recombination (Sy et al. 2009; Woods et al. 2016). Nonetheless, it is not currently known whether missense substitutions that disrupt the function of this coiled coil domain are associated with an increased cancer risk, and missense substitutions in this

domain are assigned missense Prior Ps of 0.02 (Easton et al. 2007; Tavtigian et al. 2008). Patient observational data could be used to answer the question, if data were available from several pedigrees that segregate these variants – which is exactly the rate-limiting problem for variant classification. Alternatively, specific mouse knock-ins of this class of missense substitution, e.g. Align-GVGD C65 substitutions that also show clear damage in a functional assay, could be used to estimate penetrance compared to known reference sequence and pathogenic controls (Shakya et al. 2011). The same issue arises for missense substitutions in the *BRCA2* (exon 2-3) PALB2 interaction domain and (exon 27) RAD51 disassembly domain (Xia et al. 2006; Davies and Pelligrini 2007).

4. Unclassified Variants are also remanded for further analysis if their splicing Prior P is ≥0.05 and their splicing Prior P is > the Bayesian combination of the missense Prior P and the Fxnl LR. Potentially spliceogenic missense substitutions that could damage a splice donor, damage a splice acceptor, or create a *de novo* donor are caught by this safeguard, specifically reducing the probability of false negative classification through QIE.

Towards en masse classification

None of the 81 BRCA1 missense substitutions between the coiled-coil and BRCT domains (aa 1,425-1,645) demonstrated loss of function in the transcriptional activation assay (see Figure 3). Using personal and summary family history data, it has previously been shown that non-spliceogenic missense substitutions falling outside of the key functional domains are associated with <2% probability of pathogenicity (Easton et al. 2007; Tavtigian et al. 2008; Vallée et al. 2016). Those results, which were based on patient observational data, are further supported by the functional assay data. Thus one could begin to argue that all 1,247 non-spliceogenic (splicing Prior P <0.05) missense substitutions reachable by a single nucleotide

substitution within this 220 amino acid segment could be *a priori* classified as Class 2 – likely not pathogenic.

In closing, classification of sequence variants for purposes of clinical cancer genetics and patient management does not require recently deceased political commentator John McGlaughlin's notion of "metaphysical certitude". Accepting a "Likely Pathogenic" threshold of 0.95 implies that ~2.5% of variants so classified will turn out to have been wrongly classified. The ACMG Likely Pathogenic threshold of 0.90 (Richards et al. 2015) implies that ~5% of variants so classified will turn out to have been wrongly classified. If we look back on our results 10 years from now and find that >2.5% of Likely Pathogenic variants were actually neutral, it will mean that these methods were mis-calibrated and insufficiently stringent. But if we look back 10 years from now and find that <1.0% of Likely Pathogenic variants were actually neutral, it will mean that the classification was too stringent – in fact equivalent to the Clearly Pathogenic criterion. It will also mean that fewer UVs were classified than could have been, and that fewer patients and their at-risk relatives benefited from genotype-based genetic counseling than could have been the case. The classification error rate observed when two-component QIE was applied to known neutral and known pathogenic *BRCA1* missense substitutions was 0.00. In addition, the two-component classification model appears to be reasonably robust to systematic errors in the form of (1) mis-calibration of the sequence analysis-based prior probabilities of pathogenicity, (2) overly optimistic calibration of the functional assays to generate their Functional LRs, or (3) partial non-independence between the sequence analysis-based prior probabilities and the Functional LRs. Therefore, it appears that we are poised to move bravely into a world of higher efficiency *BRCA1* sequence variant classification.

References:

Abkevich V, Zharkikh A, Deffenbaugh AM, Frank D, Chen Y, Shattuck D, Skolnick MH, Gutin A, Tavtigian S V. 2004. Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation. J Med Genet 41:492–507.

Caux-Moncoutier V, Castéra L, Tirapo C, Michaux D, Rémon M-A, Laugé A, Rouleau E, Pauw A De, Buecher B, Gauthier-Villars M, Viovy J-L, Stoppa-Lyonnet D, et al. 2011. EMMA, a cost- and time-effective diagnostic method for simultaneous detection of point mutations and large-scale genomic rearrangements: application to BRCA1 and BRCA2 in 1,525 patients. Hum Mutat 32:325–334.

Chenevix-Trench G, Healey S, Lakhani S, Waring P, Cummings M, Brinkworth R, Deffenbaugh AM, Burbidge LA, Pruss D, Judkins T, Scholl T, Bekessy A, et al. 2006. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. Cancer Res 66:2019–27.

Colombo M, Vecchi G De, Caleca L, Foglia C, Ripamonti CB, Ficarazzi F, Barile M, Varesco L, Peissel B, Manoukian S, Radice P, Stratton M, et al. 2013. Comparative In Vitro and In Silico Analyses of Variants in Splicing Regions of BRCA1 and BRCA2 Genes and Characterization of Novel Pathogenic Mutations. PLoS One 8:e57173.

Davies R and Pelligrini L. 2007. Interaction with the BRCA2 C-terminus Protects RAD51–DNA Filaments from Disassembly by BRC Repeats. Nature Structural and Molecular Biology 14:475-483.

Densham RM, Garvin AJ, Stone HR, Strachan J, Baldock RA, Daza-Martin M, Fletcher A, Blair-Reid S, Beesley J, Johal B, Pearl LH, Neely R, et al. 2016. Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection. Nat Struct Mol Biol 23:647–55.

Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian S V., Monteiro ANA, Iversen ES, Couch FJ, Goldgar DE. 2007. A Systematic Genetic Assessment of 1,433 Sequence Variants of Unknown Clinical Significance in the BRCA1 and BRCA2 Breast Cancer–Predisposition Genes. Am J Hum Genet 81:873–883.

Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian S V, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, et al. 2015. Gene-panel sequencing

and the prediction of breast-cancer risk. N Engl J Med 372:2243–2257.

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian S V., Couch FJ. 2004. Integrated Evaluation of DNA Sequence Variants of Unknown Clinical Significance: Application to BRCA1 and BRCA2. Am J Hum Genet 75:535–544.

Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Léone M, et al. 2012. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. Hum Mutat 33:1228–1238.

Iversen ES, Couch FJ, Goldgar DE, Tavtigian S V., Monteiro ANA. 2011. A Computational Method to Classify Variants of Uncertain Significance Using Functional Assay Data with Application to BRCA1. Cancer Epidemiol Biomarkers Prev 20:1078–1088.

Lindor NM, Guidugli L, Wang X, Vallée MP, Monteiro ANA, Tavtigian S, Goldgar DE, Couch FJ. 2012. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). Hum Mutat 33:8–21.

Mermershtain I, Glover JNM. 2013. Structural mechanisms underlying signaling in the cellular response to DNA double strand breaks. Mutat Res Mol Mech Mutagen 750:15–22.

Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FBL, Hoogerbrugge N, Spurdle AB, Tavtigian S V. 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat 29:1282–1291.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17:405–424.

Shakya R, Reid LJ, Reczek CR, Cole F, Egli D, Lin C-S, deRooij DG, Hirsch S, Ravi K, Hicks JB, Szabolcs M, Jasin M, et al. 2011. BRCA1 Tumor Suppression Depends on BRCT Phosphoprotein Binding, But Not Its E3 Ligase Activity. Science 334:525-528.

Spurdle AB, Couch FJ, Parsons MT, McGuffog L, Barrowdale D, Bolla MK, Wang Q, Healey S, Schmutzler RK, Wappenschmidt B, Rhiem K, Hahnen E, et al. 2014. Refined histopathological predictors of BRCA1 and BRCA2mutation status: a large-scale analysis of breast cancer

characteristics from the BCAC, CIMBA, and ENIGMA consortia. Breast Cancer Res 16:3419.

Spurdle AB, Lakhani SR, Healey S, Parry S, Silva LM Da, Brinkworth R, Hopper JL, Brown MA, Babikyan D, Chenevix-Trench G, Tavtigian S V, Goldgar DE, et al. 2008. Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis--a report from the kConFab Investigators. J Clin Oncol 26:1657–63.

Spurdle AB, Whiley PJ, Thompson B, Feng B, Healey S, Brown MA, Pettigrew C, kConFab, Asperen CJ Van, Ausems MGEM, Kattentidt-Mouravieva AA, Ouweland AMW van den, et al. 2012. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. J Med Genet 49:525–32.

Sy SMH, Huen MSY, Chen J. 2009. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. Proc Natl Acad Sci 106:7155–7160.

Tavtigian S V., Byrnes GB, Goldgar DE, Thomas A. 2008. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 29:1342–1354.

Tavtigian S V, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, Silva D de, Zharkikh A, Thomas A. 2006. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43:295–305.

Thompson D, Easton DF, Goldgar DE. 2003. A full-likelihood method for the evaluation of causality of sequence variants from family data. Am J Hum Genet 73(3):652-655.

Vallée MP, Francy TC, Judkins MK, Babikyan D, Lesueur F, Gammon A, Goldgar DE, Couch FJ, Tavtigian S V. 2012. Classification of missense substitutions in the BRCA genes: A database dedicated to Ex-UVs. Hum Mutat 33:22–28.

Vallée MP, Sera TL Di, Nix DA, Paquette AM, Parsons MT, Bell R, Hoffman A, Hogervorst FBL, Goldgar DE, Spurdle AB, Tavtigian S V. 2016. Adding In Silico Assessment of Potential Splice Aberration to the Integrated Evaluation of BRCA Gene Unclassified Variants. Hum Mutat 37:627–639.

Walker LC, Whiley PJ, Couch FJ, Farrugia DJ, Healey S, Eccles DM, Lin F, Butler SA, Goff SA, Thompson BA, Lakhani SR, Silva LM Da, et al. 2010. Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications

for prediction of pathogenicity. Hum Mutat 31:E1484–E1505.

Woods NT, Baskin R, Golubeva V, Jhuraney A, De-Gregoriis G, Vaclova T, Goldgar DE, Couch FJ, Carvalho MA, Iversen ES, Monteiro AN. 2016. Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance. npj Genomic Med 1:16001.

Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, Christ N, Liu X, Jasin M, Couch FJ, Livingston DM. 2006. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. Mol Cell 22:719-729.

Yang Y, Swaminathan S, Martin BK, Sharan SK. 2003. Aberrant splicing induced by missense mutations in BRCA1: clues from a humanized mouse model. Hum Mol Genet 12:2121–31.

Table 1. Quantitative integrative evaluation of BRCA1 missense substitutions with discordant Prior Ps

Variant	Protein	Functional Odds Path	Prior P Type	Prior P	Observational Odds Path	Posterior P
Discordant sequence analysis Prior P and functional assay result (n=2)						
c.4955T>C	p.M1652T	~0.025*	Missense	0.29	$2.7x10^{-4}$	$2.76x10^{-6}$
c.5117G>A	p.G1706A	~0.04*	Missense	0.66	$2.55x10^{-5}$	$1.98x10^{-6}$
Splicing prediction overrules functional assay (n=2)						
c.4484G>T	p.R1495M	<0.01	Missense	0.02		$2.04x10^{-4}$
			Splicing	**0.34**	$1.23x10^{6}$	1.000
c.4868C>G	p.A1623G	<0.01	Missense	0.02		$2.04x10^{-4}$
			Splicing	**0.64**	389.1	0.999

Note: Observational Odds Path is the product of observational data LRs reported on the BRCA Ex-UV database. The bolded Prior P contributed to classification of the variant. *Odds Path is estimated from variants with a similar level of assay activity because variant is a known neutral control. BF – Bayes factor; LR – likelihood ratio; Missense – sequence analysis; P – probability of pathogenicity.

Table 2. Two-component evaluation of BRCA1 missense substitutions with discordant Prior Ps

Variant	Protein	Functional Odds Path	Prior P Type	Prior P	Posterior P
Discordant sequence analysis Prior P and functional assay result (n=30)					
c.4220T>C	p.L1407P	11.98	Missense	0.02	0.197
c.4232T>C	p.M1411T	6.41	Missense	0.02	0.116
c.4957G>A	p.V1653M	>1,000	Missense	0.03	0.969
c.4993G>A	p.V1665M	0.275	Missense	0.29	0.101
c.5074G>A	p.D1692N	0.215	Missense	0.29	0.081
c.5086G>C	p.V1696L	0.157	Missense	0.29	0.060
c.5096G>C	p.R1699P	0.055	Missense	0.81	0.190
c.5201T>C	p.F1734S	0.048	Missense	0.66	0.085
c.5254G>A	p.A1752T	>1,000	Missense	0.03	0.969
c.5254G>C	p.A1752P	>1,000	Missense	0.03	0.969
c.5255C>T	p.A1752V	77.37	Missense	0.03	0.705
c.5288G>T	p.G1763V	10.17	Missense	0.03	0.239
c.5312C>G	p.P1771R	0.040	Missense	0.29	0.016
c.5312C>T	p.P1771L	0.085	Missense	0.29	0.034
c.5317A>T	p.T1773S	0.043	Missense	0.66	0.077
c.5318C>T	p.T1773I	0.121	Missense	0.81	0.340
c.5332G>T	p.D1778Y	0.059	Missense	0.29	0.024
c.5348T>C	p.M1783T	0.247	Missense	0.66	0.324
c.5359T>A	p.C1787S	<0.01	Missense	0.81	0.041
c.5363G>A	p.G1788D	0.133	Missense	0.81	0.362
c.5365G>T	p.A1789S	0.026	Missense	0.81	0.099
c.5369C>A	p.S1790Y	<0.01	Missense	0.29	0.004

c.5408G>C	p.G1803A	>1,000	Missense	0.03	0.969
c.5423T>C	p.V1808A	0.049	Missense	0.29	0.019
c.5425G>T	p.V1809F	>1,000	Missense	0.03	0.969
c.5432A>G	p.Q1811R	>1,000	Missense	0.03	0.969
c.5497G>A	p.V1833M	33.04	Missense	0.03	0.505
c.5504G>C	p.R1835P	<0.01	Missense	0.29	0.004
c.5506G>A	p.E1836K	117.4	Missense	0.03	0.784
c.5527G>C	p.A1843P	149.9	Missense	0.03	0.823
Splicing prediction overrules functional assay (n=14)					
c.4115G>A	p.C1372Y	0.010	Missense	0.02	$2.14x10^{-4}$
			Splicing	0.3	
c.4184A>G	p.Q1395R	0.021	Missense	0.02	$4.37x10^{-4}$
			Splicing	0.34	
c.4205A>G	p.H1402R	0.028	Missense	0.02	0.001
			Splicing	0.3	
c.4213A>G	p.I1405V	<0.01	Missense	0.02	$2.04x10^{-4}$
			Splicing	0.3	
c.4484G>A	p.R1495K	<0.01	Missense	0.02	$2.04x10^{-4}$
			Splicing	0.34	
c.4489T>G	p.S1497A	<0.01	Missense	0.02	$2.04x10^{-4}$
			Splicing	0.3	
c.4675G>A	p.E1559K	0.024	Missense	0.02	$4.99x10^{-4}$
			Splicing	0.97	
c.4675G>C	p.E1559Q	<0.01	Missense	0.02	$2.04x10^{-4}$
			Splicing	0.97	
c.4892G>A	p.S1631N	<0.01	Missense	0.02	$2.04x10^{-4}$
			Splicing	0.3	

c.4987A>C	p.M1663L	0.080	Missense	0.03	0.002
			Splicing	0.34	
c.4988T>A	p.M1663K	0.026	Missense	0.03	0.001
			Splicing	0.34	
c.5332G>A	p.D1778N	0.044	Missense	0.03	0.001
			Splicing	0.34	
c.5347A>C	p.M1783L	0.012	Missense	0.03	3.81×10^{-4}
			Splicing	0.3	
c.5467G>A	p.A1823T	<0.01	Missense	0.03	2.04×10^{-4}
			Splicing	0.34	

BF – Bayes factor; LR – likelihood ratio; Missense – sequence analysis; P – probability of pathogenicity.

Figure Legends

Figure 1. Functional assay decision tree. N – the number of BRCA1 missense variants with functional Odds Path [Paquette et al., manuscript in preparation] (Woods et al. 2016). P – probability of pathogenicity. LR – likelihood ratio. *There must be enough subjects/observations that if the variant were pathogenic, then the expected LR (ELR) would be >2, or if the variant were neutral, the ELR would be <0.5. †Overall sequence analysis prior for BRCA1/2 is 0.1 (Abkevich et al. 2004; Goldgar et al. 2004), thus 2-fold sequence analysis-based evidence in favor of neutrality implies a prior of 0.053 and 2-fold sequence analysis-based evidence in favor of pathogenicity implies a prior of 0.18. §Observational LRs currently include: segregation; tumor pathology; summary family history; and co-occurrence. ¶Functional Odds Path was not used in calculation of Posterior P for 9 variants tested with BRCA1 RING domain assay and 3 variants tested with BRCA1 transactivation assay, because these missense substitutions were used in calibration of the assays. #Variants are not predicted to affect canonical splice acceptor site or create a *de novo* acceptor site.

Figure 2. **Sensitivity of two-component classification to miscalibration of the sequence analysis-based prior probabilities or functional assay likelihood ratios**. A: Sensitivity to overestimated Fxnl LRs for key domain missense substitutions if the missense Prior Ps for the substitutions are >0.18, and sensitivity to underestimation of the Fxnl Lrs for key domain missense substitutions if the missense Prior Ps for the substitutions are <0.05. B: Sensitivity to

non-independence between the missense Prior Ps and Fxnl LRs. C: Sensitivity to underestimated missense Prior Ps for key domain C0 missense substitutions. In each panel, the horizontal dashed grey line marks a false classification rate of 0.025.

Figure 3. Map of BRCA1 missense variants tested in functional assays. A schematic of the missense variants tested in the RING domain and the C-terminal of the BRCA1 protein color-coded by the 5-category classifier. The concordant variants were classified using two-component and multicomponent quantitative integrated evaluation, as outlined in Figure 1. The discordant variants include missense substitutions that could not be classified using functional assay Odds Path, color-coded to what the assay results indicated based on two-component evaluation. The dashed lines are the variants where the splicing predictions overruled the results of two-component evaluation. Locations of domains: RING – aa 2-103; CC, coiled-coil – aa 1,392-1,424; BRCT1 – aa 1,646-1,736; BRCT2 – aa 1,760-1,855.

Figure 1

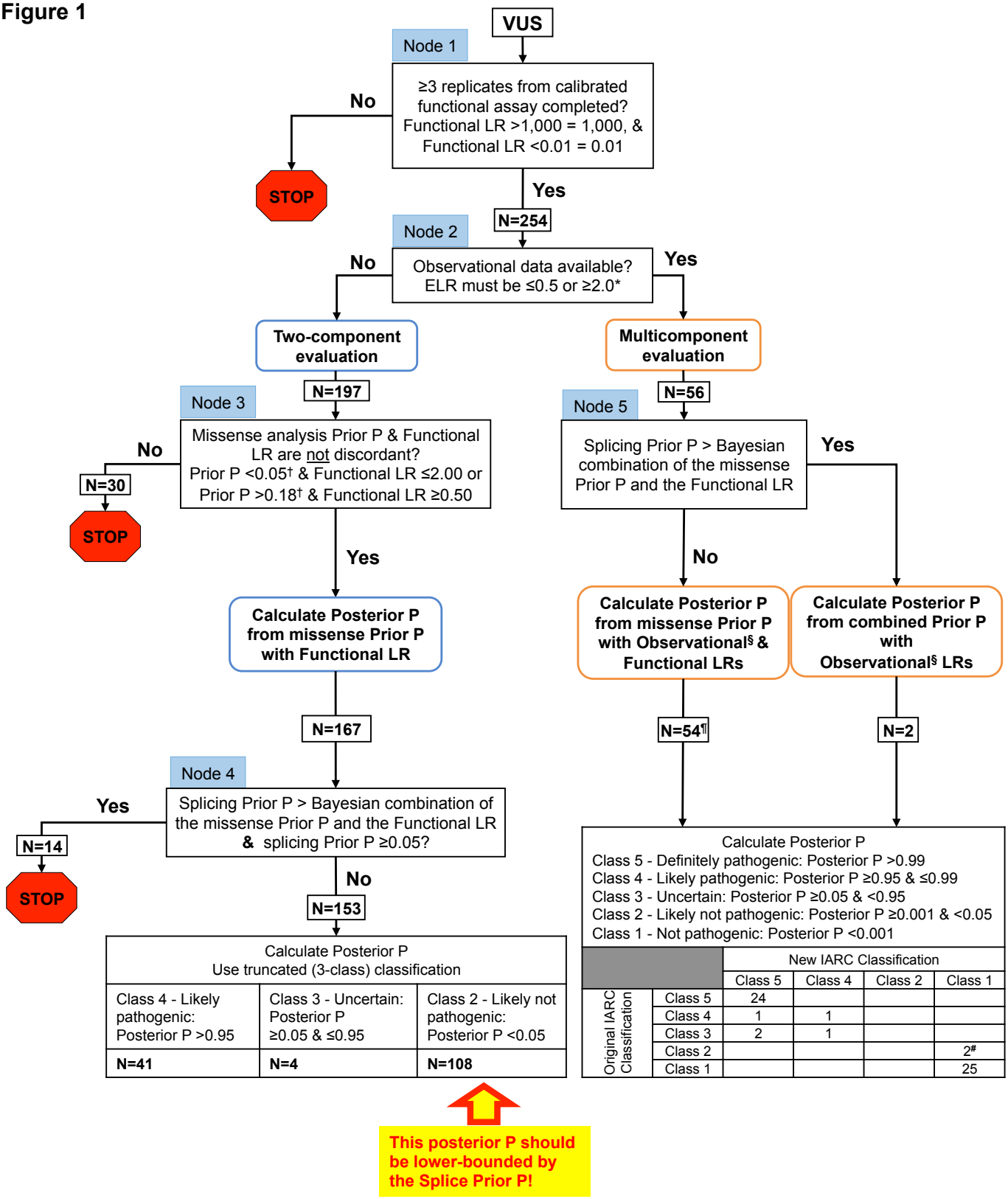

VUS
Node 1
≥3 replicates from calibrated functional assay completed? Functional LR >1,000 = 1,000, & Functional LR <0.01 = 0.01
No
STOP
Yes
N=254
Node 2
Observational data available? ELR must be ≤0.5 or ≥2.0*
No
Yes
Two-component evaluation
Multicomponent evaluation
N=197
N=56
Node 3
Missense analysis Prior P & Functional LR are not discordant? Prior P <0.05† & Functional LR ≤2.00 or Prior P >0.18† & Functional LR ≥0.50
No
N=30
STOP
Yes
Node 5
Splicing Prior P > Bayesian combination of the missense Prior P and the Functional LR
Yes
No
Calculate Posterior P from missense Prior P with Functional LR
Calculate Posterior P from missense Prior P with Observational§ & Functional LRs
Calculate Posterior P from combined Prior P with Observational§ LRs
N=167
N=54¶
N=2
Node 4
Splicing Prior P > Bayesian combination of the missense Prior P and the Functional LR & splicing Prior P ≥0.05?
Yes
N=14
STOP
No
N=153
Calculate Posterior P
Class 5 - Definitely pathogenic: Posterior P >0.99
Class 4 - Likely pathogenic: Posterior P ≥0.95 & ≤0.99
Class 3 - Uncertain: Posterior P ≥0.05 & <0.95
Class 2 - Likely not pathogenic: Posterior P ≥0.001 & <0.05
Class 1 - Not pathogenic: Posterior P <0.001
New IARC Classification
Class 5 | Class 4 | Class 2 | Class 1
Original IARC Classification
Class 5 | 24
Class 4 | 1 | 1
Class 3 | 2 | 1
Class 2 | 2#
Class 1 | 25
Calculate Posterior P
Use truncated (3-class) classification
Class 4 - Likely pathogenic: Posterior P >0.95
Class 3 - Uncertain: Posterior P ≥0.05 & ≤0.95
Class 2 - Likely not pathogenic: Posterior P <0.05
N=41
N=4
N=108
This posterior P should be lower-bounded by the Splice Prior P!

Figure 2

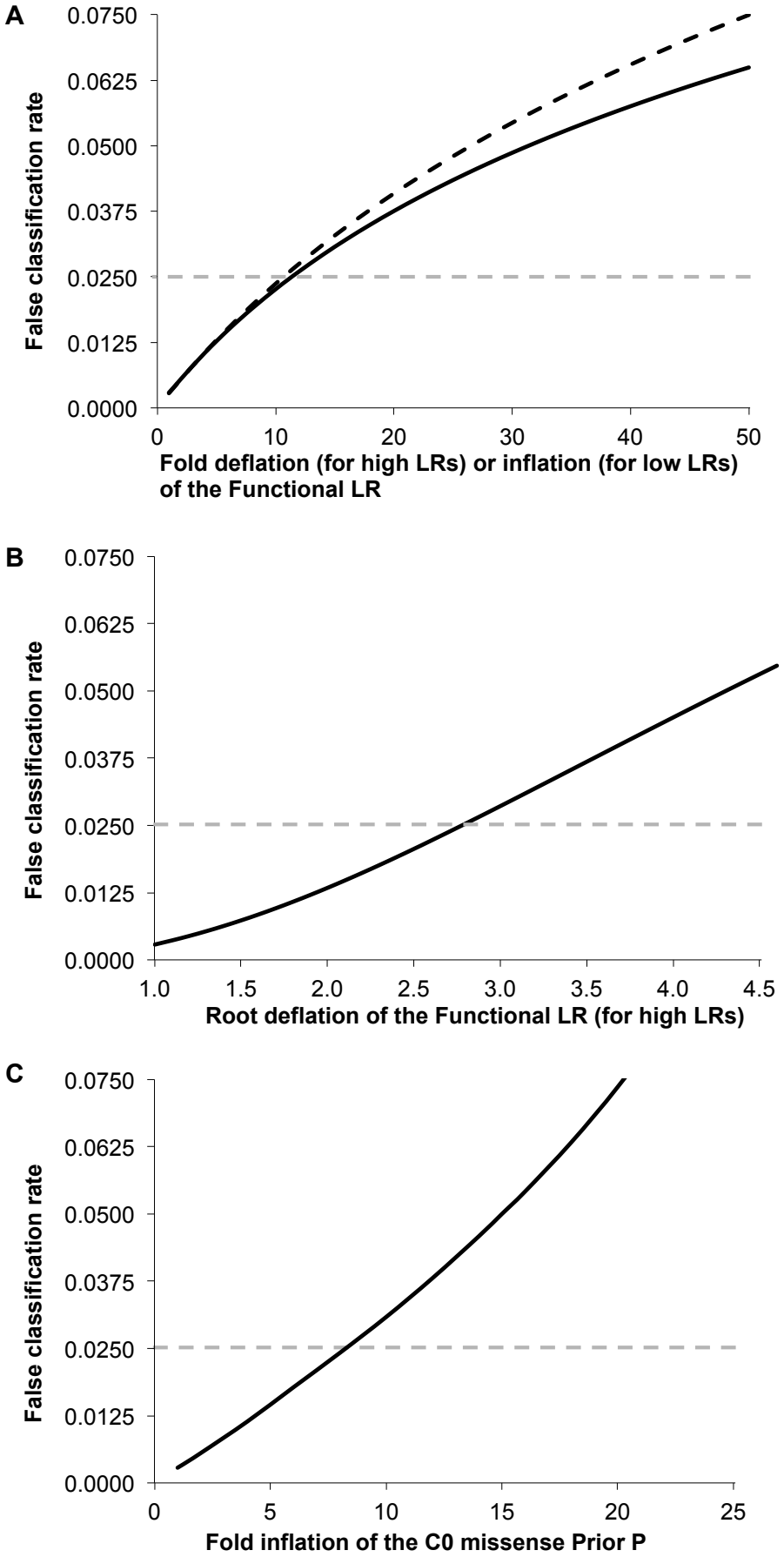
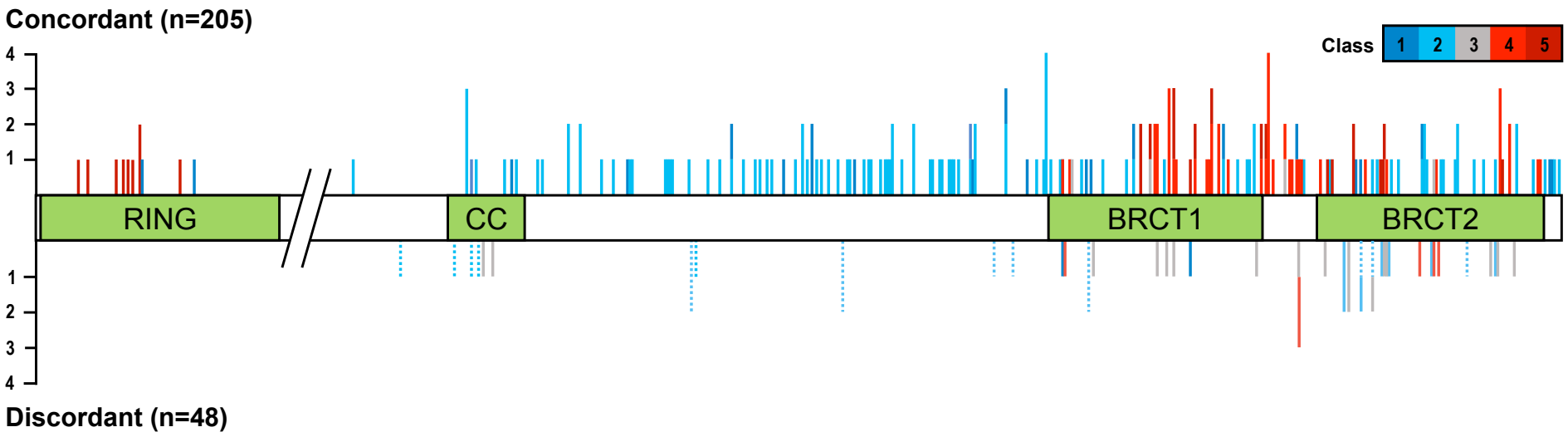

A
0.0750
0.0625
0.0500
0.0375
0.0250
0.0125
0.0000
False classification rate
0
10
20
30
40
50
Fold deflation (for high LRs) or inflation (for low LRs) of the Functional LR
B
0.0750
0.0625
0.0500
0.0375
0.0250
0.0125
0.0000
False classification rate
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Root deflation of the Functional LR (for high LRs)
C
0.0750
0.0625
0.0500
0.0375
0.0250
0.0125
0.0000
False classification rate
0
5
10
15
20
25
Fold inflation of the C0 missense Prior P

Figure 3


Concordant (n=205)
Class
1
2
3
4
5
4
3
2
1
RING
CC
BRCT1
BRCT2
1
2
3
4
Discordant (n=48)