Genomic Screening of the General Population

Michael Adams[1], James Evans[1], Gail Henderson[2], Jonathan S. Berg[1]

1. Department of Genetics, UNC-Chapel Hill 2. Department of Social Medicine, UNC-Chapel Hill

Introduction

- Genomic screening of the general population for preventable, monogenic disease has potential to decrease morbidity and mortality
- The selection of which genetic variants to return has tremendous impact on the specificity and positive predictive value of the test, which in turn has important downstream consequences for the success of any such endeavor
- We selected 17 genes for 11 conditions that are among the most medically actionable of the Mendelian disorders for genomic screening
- We screened 478 exome sequences for potentially pathogenic variants in these genes with 5 variant selection algorithms, and show the false positive rate of these algorithms

Methods

- Variants from 478 exomes from a diagnostic sequencing study (NCGENES) were loaded into a PostgresSQL database (v.9.0.3) for annotation and facilitation of queries.
- Population allele frequency estimates were determined using the Exome Aggregation Consortium (ExAC), a resource composed of 63,358 unrelated individuals sequenced through a variety of studies

The specificity, false positive rate, and number of variants returned per 1000 people screened was calculated for each of five variant selection algorithms:

- VSA-1 includes rare variants classified as “Pathogenic” in ClinVar. **This is the least sensitive algorithm**
- VSA-2 adds rare predicted truncating variants (nonsense, frameshift, canonical splice-site)
- VSA-3 adds variants classified as “Likely Pathogenic” in ClinVar and/or as a “Disease Mutation (DM)” in HGMD
- VSA-4 adds rare missense variants with CADD scores >13 that are located within a conserved functional domain
- VSA-5 adds all rare missense variants, regardless of CADD score or location. **This is the most sensitive algorithm**

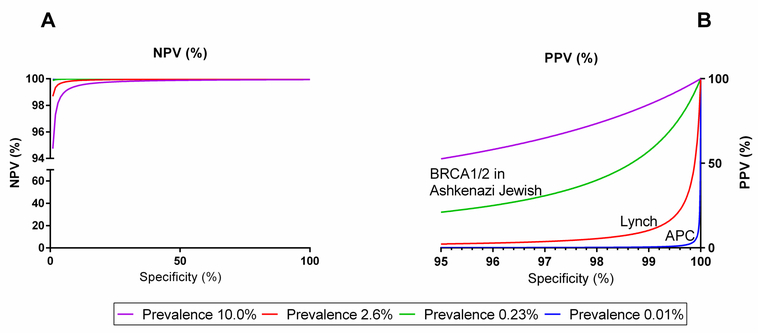
The medical literature was reviewed in order to estimate the clinical sensitivity of diagnostic testing (corrected for locus heterogeneity), and the NNS based on the minimal sensitivity.

Table 1: minimum sensitivities of screening 17 genes for known variants (corresponding to VSA-3)

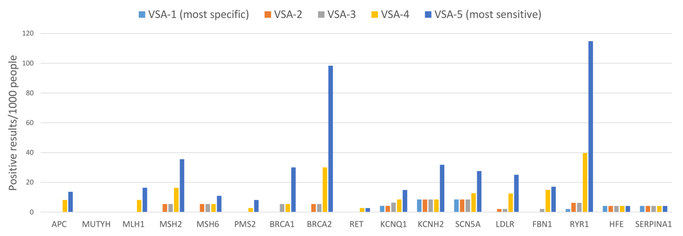
Category	Phenotype	Gene	Population prevalence	Clinical sensitivity	Cases due to known variants	Minimum clinical sensitivity	NNS	Notes
Cancer	Polyposis	*APC*	1/10,000	70%	84%	**59%**	14,000	This calculation is for classical polyposis (>100 polyps)
		MUTYH	1/20,000	99%	>73%	**72%**	20,000	MUTYH also contributes to a fraction of cases of attenuated polyposis; clinical sensitivity is much lower in that case
	Lynch syndrome	*MLH1*	1/910	33%	35%	**11%**	3,000	
		MSH2	1/1,100	31%				
		MSH6	1/6,300	8%				
		PMS2	1/14,000	4%				
	Hereditary Breast/Ovarian cancer	*BRCA1*	1/300 whites; 1/88 Ashkenazi Jewish	67%	84%	**56%**	450	
		BRCA2	1/140 whites; 1/66 Ashkenazi Jewish	70%			208	
	MEN 2A/2B	*RET*	1/34,000	95%	>99%	**94%**	36,000	
Cardiovascular	Long QT Syndrome	*KCNQ1*	1/5,800	32%	46%	**15%**	6900	
	Long QT Syndrome	*KCNH2*	1/7,800	28%	70%	**20%**		
	Long QT Syndrome	*SCN5A*	1/19,000	25%	63%	**16%**		
	Brugada syndrome		1/2,000	20%			10,000	
	Familial Hyperlipidemia	*LDLR*	1/200 Europe	80%			250	
	Marfan Syndrome	*FBN1*	1/5,000	91%			5500	FBN1 is also implicated in non-syndromic aortopathies
Other	Malignant Hyperthermia	*RYR1*	1/3300	86%	30%	**26%**	3800	
	Hereditary Hemochromatosis	*HFE*	1/230	90%	94%	**85%**	250	
	Alpha-1-Antitrypsin	*SERPINA1*	1/2000	95%	95%	**90%**	2100	

Results

- The NPV (A) and PPV (B) of genetic screening is demonstrated for prevalence values ranging across four log scales (from 10% to 0.01% prevalence).
- For any rare disease, the NPV is less influenced by the test characteristics while PPV has extreme dependence on specificity.



- The yield of potentially pathogenic variants using 5 variant selection algorithms with varying sensitivity is shown below. The number of people who would screen positive per 1000 individuals screened is displayed on the vertical axis



The false positive rate (1-Specificity) of 5 variant selection algorithms across all 17 genes is shown below:

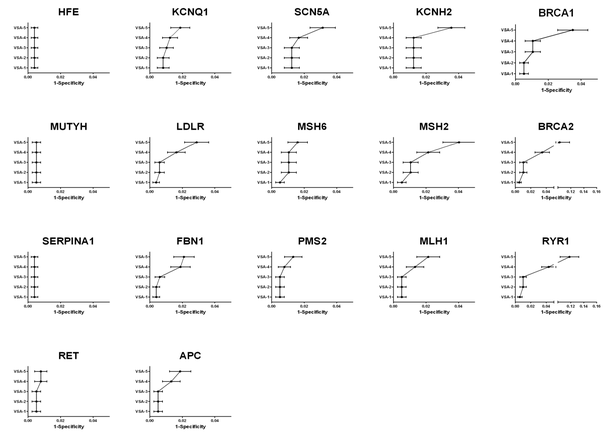


Table 2 shows the consequences of a positive screen for each gene and the suggested VSA to achieve the desired false positive rate

Category	Gene	Interventions	Suggested False Positive Standard	Mutational Spectrum (derived from HGMD and Clinvar)	Variant selection algorithm
Cancer	APC	Colonoscopy, endoscopy screening, thyroid ultrasound, surgery	Low tolerance	81% truncating; 19% missense	Strict Known Pathogenic
	MUTYH	Colonoscopy, endoscopy	Low tolerance	32% truncating; 68% missense	Strict Known Pathogenic
	MLH1	Colonoscopy, endoscopy, endometrial biopsy, possible surgery (prophylactic hysterectomy and salpingo-oophorectomy)	Low tolerance	50% truncating; 50% missense	Strict Known Pathogenic
	MSH2			65% truncating; 35% missense	
	MSH6			59% truncating; 41% missense	
	PMS2			65% truncating; 35% missense	
	BRCA1	Breast imaging, prophylactic mastectomy and/or salpingo-oophorectomy	Very low tolerance	60% truncating; 40% missense	Strict Known Pathogenic
	BRCA2			53% truncating; 47% missense	
	RET	Prophylactic thyroidectomy, serum metanephrine blood test	Very low tolerance	11% truncating; 89% missense	Strict Known Pathogenic
Cardiovascular	KCNQ1	Cardiology consultation, electrocardiogram (ECG), beta-blocker medication if ECG positive; implantable cardioverter-defibrillator if symptomatic	Low tolerance	13% truncating; 87% missense	Strict Known Pathogenic
	KCNH2			36% truncating; 64% missense	
	SCN5A			34% truncating; 66% missense	
	LDLR	Lipid biochemical screening, pharmacotherapy if needed	High tolerance	21% truncating; 79% missense	Known and Stringent Missense
	FBN1	Echocardiography, ophthalmologic screening	Low tolerance	25% truncating; 75% missense	Strict Known and Novel
Other	RYR1	Avoidance of specific anesthetics	High tolerance	7% truncating; 93% missense	Known and Stringent Missense
	HFE	Ferritin biochemical screening, phlebotomy	High tolerance	24% truncating; 76% missense	Known and Stringent Missense
	SERPINA1	Avoidance of smoke exposure	High tolerance	19% truncating; 81% missense	Known and Stringent Missense

Conclusions

- To optimize public health benefits from screening the genome for preventable, rare disease, the highest specificity must be ensured
- Disease-specific false positive rates can be chosen, and different variant selection algorithms may be pursued depending on the presence of a confirmatory test and other downstream consequences of screening

This work was supported by NHGRI 2P50 HG004488 (“Genescreen”), NHGRI 1U01HG006487-01 (“NCGENES”), and the Howard Holderness Distinguished Medical Scholars Foundation