The Positive Predictive Value of Genetic Screening Tests

Jill Hagenkord,[1] Birgit Funke,[2] Emily Qian,[2] Madhuri Hegde,[3] Kevin B Jacobs,[4] Matthew Ferber,[5] Matthew Lebo,[6] Adam Buchanan,[7] David Bick[8]

[1] MDisrupt, San Jose, CA, USA
[2]Veritas Genetics, Danvers, MA, USA
[3] Perkin Elmer Genomics, Branford, CT, USA
[4] Progenity, Ann Arbor, MI, USA
[5]Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, MN, USA
[6]Laboratory for Molecular Medicine at Partners HealthCare, Boston, MA, USA
[7]Geisinger Genomic Medicine Institute, Danville, PA, USA
[8]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA

Corresponding Author: Jill Hagenkord, MD[1]
Email address: jill@mdisrupt.com

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Abstract

As the cost of DNA sequencing continues to fall, more individuals are opting to undergo genomic testing in the absence of a clinical indication. Testing asymptomatic individuals for unsuspected conditions is not new to the medical and public health communities. However, the application of screening principles to inherited diseases with unknown prevalence in an unselected (low risk) test population raises unique challenges. This paper examines the effect of disease prevalence on the positive predictive value of a test result. Many inherited conditions have very low prevalence in an unselected population, which increases the probability that some likely pathogenic variants may be false positives. In situations where the adverse impact of a false positive result is significant, laboratories should address this issue by either increasing the interpretive specificity of the test, performing a clinical confirmatory test to establish the presence of disease, or restricting the test to a population with increased disease prevalence. Here, we review the statistical concepts relevant to screening tests, apply these concepts to genetic disease screening, create a model to estimate prevalence and positive predictive value, and provide a framework for further discussion.

Key words: genetic screening, epidemiologic methods, clinical validity, positive predictive value, clinical test validation

Introduction

Imagine that you are a clinician and a patient brings you a genetic screening report for inherited conditions. The individual has a positive result for an inherited cancer syndrome but does not have a personal or family history of this disorder. How should you advise this patient?

Clinical DNA sequencing for inherited diseases has typically been performed in specialized genetics laboratories that focus primarily on sequence-based tests intended for individuals suspected of having a genetic condition. In the scenario above, the patient has had a screening test. Screening tests are intended to identify the presence of an as-yet-undiagnosed disease in individuals without signs or symptoms. (Maxim et al., 2014; Trevethan, 2017) The prior probability of disease in this setting is often quite low and, therefore, the possibility of false results may be quite high.

We are now entering an era in which DNA sequencing is carried out in different populations for different purposes. (Figure 1) This commentary explains the statistical origin of potential false positive results when screening asymptomatic individuals for rare conditions, and why it is necessary to know the prevalence of the condition and the positive predictive value (PPV) of a test *in the intended target population*. We apply these general concepts to genetic diseases, propose a model to estimate PPV in the absence of concrete prevalence data, and offer options to mitigate the risk of false positive results in genetic screening tests.

General Principles of Clinical Test Design

The design of laboratory tests begins with an intended use statement, which describes what the assay is testing, the testing technology, why the test is performed, acceptable sample types, and who is or is not an appropriate test subject. (Jennings et al., 2009) The intended use guides the

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

laboratory director's choice of technologies and informs the sensitivity/specificity trade-off used to determine the most appropriate cut-off between a "positive" and "negative" result. (Supplemental Figure 1) In addition, the intended use helps the physician order the appropriate test for a particular clinical question. During test development, a clinical validation study is typically conducted by performing the new test on samples from patients confirmed to have the disease by a "gold standard" or index test and patients confirmed to be free of the disease. Ideally, the new test will discriminate between these cohorts perfectly; in practice there will be false positive and false negative results. Although it is critical to establish the *analytical* validity of a test, the more salient metric for test optimization is *clinical* validity. (Table 1) Clinical validity reflects the test's ability to properly categorize those with the disease as positive (clinical sensitivity) and those without the disease as negative (clinical specificity). (Maxim et al., 2014) In this commentary, we focus on clinical validity. Furthermore, when we use the term "confirmatory test" we are referring to a "gold standard" or index test that confirms the presence or absence of the condition (Supplemental Table 1), not the presence or absence of a genetic variant (orthogonal technical confirmation).

As mentioned above, clinical sensitivity and specificity are determined using a reference cohort during test development. The predictive values, in contrast, provide the post-test probability of disease *for an individual*. (Screening (medicine) - Wikipedia 2005) Predictive values address the probability that a person with a positive result has the condition (positive predictive value or PPV) or that the person with a negative result does not have the condition (negative predictive value or NPV). (Trevethan, 2017) Unlike sensitivity and specificity, PPV and NPV will vary depending on the prevalence of the disease in the test population. (Figure 2) In the clinic, the PPV and NPV are more useful than sensitivity and specificity since in practice the presence/absence of disease is unknown in the individual prior to testing. (Akobeng, 2007)

Estimates of PPV and NPV can be used during test development to determine the most appropriate cut-off threshold between a positive result and negative result as appropriate for the test's intended use. Sliding the cut off towards higher specificity will result in fewer false positives, while sliding the cut-off toward higher sensitivity will decrease the possibility of false negatives. (Supplemental Figure 1) The optimal cut-off will vary based on the intended use. When designing a screening test, the extent to which true positive and true negative results are medically desirable and the extent to which false positive and false negative results are tolerable or even acceptable must be weighed. (Trevethan, 2017) Considerations include the immediate and long-term burden on the healthcare system, the treatability of the condition, psychosocial effects, and the potential over-utilization of diagnostic procedures or surveillance. (Maxim et al., 2014; Trevethan, 2017)

If a screening test produces a significant number of false positives, the cost of managing the false positives may outweigh the potential benefits of identifying the true positives. For example, recommending a lifetime of increased surveillance for breast cancer has both cost and risk. The more mammograms a woman undergoes, the more likely she will have a false positive result. This will result in unnecessary invasive follow-up tests. The chance of a false positive result after one mammogram is 10%, depending on age. Younger women are more likely to have a false positive result than older women. After 10 yearly mammograms, the chance of having a false positive is 50-60%.("Mammogram Accuracy - Accuracy of Mammograms") Without data to

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

support the safety and efficacy of that intervention, the screening may do more harm than good. There are many examples where interventions assumed to have benefit resulted in considerable harm.(Maxim et al., 2014; Trevethan, 2017) False negatives, on the other hand, could result in a missed opportunity for early detection.

To appreciate the fact that a screening test with high sensitivity and specificity can still have a low PPV *if the prevalence of the disease in the test population is sufficiently low* (Figure 2), consider the following examples:

- HIV screening test in sub-Saharan Africa →Most positives are true positives because the prevalence of HIV is high.
- HIV screening test in rural Canada →Most positives are false positives because the prevalence of HIV is low.
- Fecal occult blood test in people over 80 years old → Most positives are true positives because the prevalence of colon cancer is high.
- Fecal occult blood test in people under 20 years old → Most positives are false positives because the prevalence of colon cancer is low.

If a test has a 100% clinical sensitivity and 95% clinical specificity, then:

- For disease prevalence of 10%, the positive predictive value is 69%.
- For disease prevalence of 1.0%, the positive predictive value is 17%.
- For disease prevalence of 0.1%, the positive predictive value is 2%.

There are three ways to minimize false positives and thereby improve the PPV when designing a screening test: 1) increase the specificity of the screening test, 2) pair the screening test with a clinical confirmatory test, 3) employ the screening test in a population with a higher disease prevalence. Raising the "positive" cut off (i.e., increasing the stringency) will result in fewer positive results.(Eisenberg, 1995) (Figure 1 and Supplemental Figure 1) If the test was intended for use in a higher-prevalence setting, one would then have to consider the impact of increased specificity on the number of false negatives that will result. However, when the prevalence of disease is very low, increasing the specificity will have only a small effect on sensitivity.

Genetic Screening

The principles of screening were first published In 1968 by the World Health Organization (Wilson, Jungner & Health Organization, 1968) and adapted to DNA-based preventive screening in 2003. (McCabe & McCabe, 2003) Not all screening tests for inherited diseases available today were designed using these criteria. Nevertheless, genetic testing in the absence of a clinical indication is occurring more frequently, such as issuance of secondary reports in the context of a diagnostic exome (opportunistic screening), elective sequencing paid for by a curious individual, population screening offered by health systems and employers, and research studies involving return of results. (Brothers, Vassy & Green, 2019; Lu et al., 2019)

When screening an asymptomatic population for inherited diseases, the clinical implications of the results are confounded by uncertainty surrounding both the PPV of the test and the penetrance of the condition. PPV and penetrance can be confused because they are similar concepts. An example of the meaning of PPV in the context of a genetic disease is provided in

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Table 1. PPV provides information about FALSE POSITIVE tests. In contrast, penetrance is a feature of the condition (not the test) and is only applicable to TRUE POSITIVE genetic tests.

When designing a screening test, it is important to quantify the possibility that a result is a false positive. Outside of genetics, screening tests are intentionally designed to permit some false positives to avoid missing true positives. These false positives are tolerated because the standard practice is to follow a positive screening test with a confirmatory diagnostic test. The diagnostic confirmatory test is designed to be more specific so that it can identify the false positives detected by the screen. (Maxim et al., 2014) For genetic tests that do not have a clinical confirmatory test, this two-step process is not possible, and therefore the specificity of the screening test itself must be high.

There are two sources of false positives in genetic testing: analytical and clinical. (Table 1) This commentary focuses on the clinical sensitivity and specificity: How well does the test detect the *condition* when it is present and produce a negative result when the condition is not present *as compared to a gold standard diagnostic test*? Ideally, in the test development phase we would have 1000 samples from individuals known to have the condition and 1000 who do not have the condition. We could then determine the number of positive results identified by the genetic test that were false positives at different levels of specificity. Designing a test for inherited disease syndromes is challenging because a gold standard diagnostic test does not exist for many diseases. As a result, the clinical sensitivity and specificity cannot be definitively determined, although reasonable estimates are possible.

The clinical specificity of DNA sequencing tests for inherited diseases can be derived from the level of certainty that the identified variants cause disease. (Adams et al., 2016) When a variant is classified as "likely pathogenic," there is a possibility that it *is not* a cause of disease. The ACMG/AMP variant classification guideline advises that the cut off between positive test result (ie, clinically actionable) and negative test result is between variant of uncertain significance (VUS) and likely pathogenic (LP) when testing high-risk individuals. (Richards et al., 2015) In other words, a clinician should act on pathogenic and likely pathogenic results. They should not introduce or withhold a therapy based on variants that score below LP. The authors of the guideline suggested that the level of confidence in LP should be at least 90%, in which case, 1 in 10 likely pathogenic variants may be false positives, a medically appropriate cut-off when testing high-risk patients. However, in the absence of empirical data from a gold standard comparator set, the authors had to *estimate* the confidence of a likely pathogenic variant truly causing disease. This confidence level is an estimate of the clinical specificity. (Adams et al., 2016) (Figure 1) The guideline warns that applying this same cut-off to an asymptomatic test population may result in an unacceptable number of false positives. (Richards et al., 2015)

A Model for Estimating the PPV of Genetic Diseases

To use the results of a genetic test effectively the clinician needs to know the PPV of the test, not just the specificity. Determining the PPV of a test results requires knowledge of the prevalence of the disease in the test population. For many genetic diseases, we do not know the prevalence of the condition (or the prevalence of pathogenic variants) in an unselected population because most genetic studies have been carried out in a population with prior evidence of the disease.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Nevertheless, it is possible to estimate the expected prevalence of a condition, and thereby estimate the probability that a "likely pathogenic" variant is a false positive.

We applied the principles described above to the screening of asymptomatic individuals in an unselected (low risk) population for autosomal dominant monogenic inherited diseases in order to approximate the PPV across various disease prevalences. Assumptions in the model:

- 100% analytical test sensitivity.
- 90% confidence in (the specificity of) a "likely pathogenic" classification using the ACMG guidelines. In practice, the actual confidence will vary by variant type and gene/condition.
- 100% confidence in (the specificity of) a "pathogenic" classification.
- ⅓ of all positive results are "likely pathogenic" variants and ⅔ of all positive results are "pathogenic". Together, these compose the "overall positive rate"
- Prevalence of the CDC Tier 1 conditions, which include Hereditary Breast and Ovarian Cancer Syndrome (HBOC), Lynch syndrome (LS) and Familial hypercholesterolemia (FH), range from 1/500 to 1/200 in an unselected population. (Murray, M. F.,et al, 2018)
- The overall positive rate used for CDC Tier 1 (Abul-Husn et al., 2016; Manickam et al., 2018) and ACMG Secondary Findings V2.0 (ie, ACMG59™) is 1.5% and 3.0%, respectively, with the reported range for overall positive rates of ACMG59™ between 1% and 8.5%. (Dorschner et al., 2013; Olfson et al., 2015; Jamuar et al., 2016; Rego et al., 2018) (Table 2)
- Prevalence of ACMG59™ and other inherited diseases in an unselected population is not known but this model assumes a range between 1/10,000 and 1/500. Some may be much rarer. (Tables 2 and 3)

The specificity of the test depends on the overall positive rate of the test, the type of variant, the strength of the gene-disease association, and knowledge of the specific gene/disorder. Adams et al. assert that the clinical specificity of a "likely pathogenic" variant is 90-95% assuming a pre-clinical (test validation) cohort study that included an equal number of known positive and negative samples. (Adams et al., 2016) When the overall positive rate in practice is not 50%, the clinical specificity should be recalculated. Supplemental Figure 2 provides an example of how overall positive rate is used in determining specificity.

Like other models, ours assumes a 90% confidence that a likely pathogenic variant is a true positive. (Richards et al., 2015; Adams et al., 2016; Tavtigian et al., 2018) In practice there is a range of specificity that will vary by gene/disease pair as discussed and by variant type. For a monogenic disorder with a dominant inheritance pattern caused primarily by loss of gene function from simple variants, the specificity of variant classification will generally decrease for each of the following functional categories: known pathogenic variants, loss-of-function variants in relevant gene domains (nonsense and frameshift variants), canonical splice altering variants and non-canonical splice variants with some functional evidence, missense variants with credible functional evidence. This trend toward decreasing specificity is due to the inherent complexity of interpretation and the rate of errors that can occur in applying evidence toward a classification. Specificity (confidence) will be 100% for well-known pathogenic variants. This will decrease to below 90% in categories near the end of the list. As with other types of clinical tests, the positive

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

rate will increase as more genes and variant types are tested or when less stringent criteria are used to specify a variant as "positive." The false positive rate will increase as the overall positive rate increases.

The model examines two sets of genetic disorders, the CDC Tier 1 Conditions and the ACMG59™, which are currently used to screen asymptomatic individuals. The CDC Tier 1 Conditions were proposed by Murray, et al as conditions for which population screening could be implemented, provided outcomes were being measured. (Murray, M. F., et al., 2018) Importantly, the three CDC tier 1 conditions (HBOC, LS, FH) have been studied in large unselected populations so that disease prevalences (and penetrance) are established. As a result, a PPV can be determined. The ACMG has issued recommendations for reporting secondary findings in clinical exome and genome sequencing. They enumerate 27 conditions (59 genes) appropriate for inclusion in an opportunistic screen. (Kalia et al., 2017) The general population prevalence of disease associated with a number of these genes is unknown. While the ACMG specifies that this set of genes is not appropriate for general screening, several entities are testing asymptomatic individuals for these 59 genes.

The estimated PPV for a test with 100% sensitivity and a given specificity at different disease prevalences that are representative of inherited genetic conditions in the CDC Tier 1 and ACMG59™ lists are presented in Tables 2 and 3. Note how small changes in specificity can have a significant impact on PPV when testing for diseases with a low prevalence. The PPV for the CDC Tier 1 conditions (prevalence 1/200 to 1/500) is in the 80-90% range. (Murray, M. F., et al., 2018) The prevalence of several conditions included in ACMG59™ is less frequent than 1/25,000.("OMIM Entry - # 145600 - MALIGNANT HYPERTHERMIA, SUSCEPTIBILITY TO, 1; MHS1", "OMIM Entry - # 180200 - RETINOBLASTOMA; RB1", "OMIM Entry - # 193300 - VON HIPPEL-LINDAU SYNDROME; VHL", "OMIM Entry - # 158350 - COWDEN SYNDROME 1; CWS1") Several conditions are so rare that firm epidemiological estimates are not available. Our model provides PPV estimates for diseases with a prevalence as low as 1/10,000. These conditions require screening 10,000 people in order to detect a single true positive result. In such low prevalence conditions, a specificity of 99.94% results in a PPV of only 10%. (Adams et al., 2016)

When the prevalence for a condition in a population is sufficiently low, the false positive rate for a test is greater than the prevalence of the condition (the False Positive Paradox).(Ndase et al., 2015; Wikipedia contributors, 2019) In this scenario, more disease-free individuals will test positive than diseased individuals. As a result, the clinician who is accustomed to evaluating a positive test result drawn from a high-risk population may erroneously conclude that a positive test means that the individual is affected, when in fact a false positive is far more likely. At 99.9% specificity, the False Positive Paradox becomes a consideration at a disease prevalence of 1/1000. At this disease prevalence there will be one false positive for each true positive. If the specificity is reduced to 99.5%, testing would yield 5 false positives for each true positive. For a disease prevalence of 1/10,000 a 99.99% specificity results in one false positive for each true positive. ("Diagnostic Test Calculator") Only tests with 100% specificity escape this paradox.

Recommendations to Minimize False Positives in Genetic Screening

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

False positives in genetic screening can be minimized by adjusting the "abnormal" cut-off to increase the specificity and thus maximize PPV. Options include:

- Increase the specificity by reporting only known pathogenic variants (100% specificity)
- Increase the specificity by reporting known pathogenic variants and high confidence likely pathogenic variants (high specificity)
- Titrate clinical specificity based on clinical implications and availability of a confirmatory functional test. (Adams et al., 2016)

Hereditary breast and ovarian cancer syndrome (HBOC) is an example of a disorder where titration of clinical specificity might be considered. A very low false positive rate is required because consideration of risk-reducing surgery is an important management step, increased surveillance has adverse consequences and high cost, and no confirmatory tests exist to ensure that a genetic variant is indeed disease-causing. Thus, only carefully curated pathogenic variants and, perhaps, high confidence truncating variants should be reported in a screening context. In other cases, inclusion of rare, potentially damaging missense variants may be acceptable, depending on the specificity of variant selection algorithms for the gene, the spectrum of pathogenic mutations observed, and the false positive tolerance based on clinical implications of a positive screening result. (Adams et al., 2016)

When increasing the stringency, we recommend that laboratories adjust the cut-off between positive and negative interpretation of results (Lu et al., 2019) rather than adjusting variant classification, as ACMG (Richards et al., 2015) and others have recommended.(Adams et al., 2016; Murray, 2016) For example, reporting pathogenic and high confidence likely pathogenic variants as "positive" and lower confidence likely pathogenic variants as "negative" would accomplish this goal. (Figure 1). Increasing the cut-off adjusts the meaning of the variant for this particular patient based on their pretest probability, but does not change the classification of the variant itself.(Lu et al., 2019) And, the risk of confusion in the clinic would be lower if genetic laboratories adopted the traditional practice of adjusting the positive/negative cut-off when adjusting test sensitivity and specificity. Note that the impact of this approach on clinical sensitivity will be negligible given the low prevalence of these conditions, therefore there is minimal benefit to including the lower confidence novel variants that inherently reduce the specificity of the screening test. (Eisenberg, 1995; Adams et al., 2016) In addition to increasing the clinical specificity of the test, there are other approaches that minimize the risk of false positives. The screening test could be paired with a confirmatory test that has close to 100% specificity. Alternatively, the laboratory's intended use statement could articulate a particular subpopulation that has a higher prior probability of disease.

Some laboratories and programs have adjusted their stringency when pre-test probability of disease is low, while others have not. In the MyCode Community Health Initiative, Geisinger filters out lower confidence likely pathogenic variants and only reports pathogenic and high confidence likely pathogenic variants (together called 'expected pathogenic' variants) as positive results. They designate a variant as expected pathogenic on the basis of 1) classification in ClinVar with a *2 or *3 status, indicating strong evidence for pathogenicity; 2) predicted loss of function in genes in which loss of function is the mechanism of disease; or 3) both.(Manickam et al., 2018) Mayo Clinic's GeneGuide™ only reports a limited list of known pathogenic variants

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

as a positive result, despite their ability to detect novel variants. ("Mayo Clinic GeneGuide - Results," 2018) Other laboratories have not adjusted their interpretation for an unselected test population. As a result, there is a possibility of receiving a positive result or a negative result for the same variant depending on the laboratory performing the test. This underscores the need for consistent application of screening principles and clear statements of intended use by genetic laboratories.

Lastly, we considered whether PPV could be extrapolated from intra-laboratory reclassification rates, published reclassification rates from high volume laboratories, the reclassification rate in public variant databases, or from genetic conditions that have a well-established functional confirmatory test that can serve as the "gold standard." While reclassification rates for different laboratories is an important consideration when a physician makes management decisions, we reasoned that the reclassification rate is not a replacement for a "gold standard" derived test performance metrics such as PPV. Although there are some genetic conditions that have that have a "gold standard" comparator, such as hemoglobinopathies or coagulopathies, full gene sequencing studies in an unselected population have not been published with phenotypic data for these conditions. Performance metrics obtained for these gene/condition pairs cannot be confidently extrapolated to other genes, irrespective of the data. Datasets from high volume laboratories indicate that reclassification rates are laboratory dependent, based on high-risk populations, and fluctuate with differences in test volume in each year, the rate at which new data appears that affect classes of variants, and how many genes were available for testing each year.(Macklin et al., 2018; Mersch et al., 2018) In contrast, the PPV for an individual gene/disease pair is dependent on test specificity and the prevalence of the disease in the test population. Thus, specificity will decrease with an increasing number of likely pathogenic calls and increasing panel size and decrease with disease prevalence in the test population.

Conclusion

In summary, it is important for laboratories to optimize tests according to their intended use. A test employed outside of its intended use may result in an unacceptable number of false results. In some genetic screening delivery models, there may be a communication separation between the patient, the ordering physician, and the patient's primary care provider. For example, some laboratories allow an independent third-party network of physicians to review an individual's health history and order a genetic screening test. By clearly articulating the following in the pre-participation informational collateral, consent, reports, and physician education material, laboratories can minimize miscommunications:

- The "intended use" or "purpose of" the test, including whether it is designed for screening, diagnosis, or monitoring; the specific analyte or condition of interest; the target test population; the technology used; and how the results should be used.
- The PPV for the gene/condition pair in question in the test population.
- If the test is intended to be used as a screening test, the availability or absence of a clinical confirmatory test should be noted.
- The spectrum of clinical features of the syndrome and the penetrance of each feature in a low-risk population, if known, or clearly state that it is not known.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

- A reference to patient management guidelines that have been proven safe and effective in an asymptomatic population, if available, or clearly state that none exist.
- Participants should be aware that post-test preventive care may not be covered by insurance. Coverage will depend on the level of evidence for the utility of the intervention and the type of insurance. For most commercial payers, the presence of a positive genetic result is sufficient for diagnosis and coverage of the CDC Tier 1 conditions, but the same is not necessarily true for original Medicare or conditions outside CDC Tier 1. (Health Leaders Media, 2013)
- Laboratories should consider conducting comprehension testing on these limitations and/or require positive results to be communicated to the patient by a genetic counselor.

Let's return to the clinician at the beginning of this review whose patient arrives with an unsolicited genetic screening test report that is positive. Here are some examples suggesting how the provider might advise a patient:

- **Result: HBOC**, **pathogenic** variant in *BRCA1*. The PPV is 100% in both low risk (unselected population) and high-risk (early onset breast/ovarian families) population. The risks of breast and ovarian cancer in an unselected cohort appears to be only slightly less than that observed in high-risk cohorts.(Gabai-Kapara et al., 2014; Akbari, Gojska & Narod, 2017; Metcalfe et al., 2018) The interventions proven safe and effective for high-risk patients can be applied to individuals from an unselected population. (Manickam et al., 2018) **Clinical Action:** It is appropriate to follow NCCN (National Comprehensive Cancer Network) management guidelines. The presence of a positive genetic test result alone meets criteria for most commercial insurance coverage policies for the management of HBOC. Original Medicare may not cover preventive care.
- **Result: HBOC**, **likely pathogenic** variant in *BRCA1*. The PPV is 80-90% in an unselected population. A clinical confirmatory test is not available to determine if this is a true positive. **Clinical Action:** The appropriate action is undetermined.
- **Result: Peutz-Jeghers syndrome (PJS), likely pathogenic** variant in *STK11*. The prevalence of PJS is less than 1/10,000 and the PPV is under 10% in an unselected population. A clinical confirmatory test is not available to determine if this is a true positive. If it is a true positive, the penetrance has not been studied in an unselected population. **Clinical Action:** You cannot advise on likelihood of disease in this patient. Additional surveillance has not been demonstrated to be useful or cost-effective. Insurance coverage for increased surveillance may be challenged.
- **Result: Familial Adenomatous Polyposis (FAP)**, **pathogenic** variant in *APC*. The PPV is 100%. The person has FAP or attenuated FAP. However, the penetrance of polyposis for individuals from an unselected population appears to be very low, although extracolonic features may occur. (Rocha, et al. 2019) **Clinical Action:** The NCCN guidelines for managing FAP patients is not appropriate since polyposis is unlikely. No guidelines exist for how to best manage this patient. Increased surveillance has not been demonstrated to be useful or cost-effective. Insurance coverage for increased surveillance may be challenged.
- **Result: Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC), likely pathogenic** variant in *DSP*. The prevalence is approximately 1/5000(Romero et al., 2013) and the PPV is estimated to be 17%. The penetrance in individuals from an unselected

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

population appears to be very low. (Haggerty et al., 2017) **Clinical Action:** There is no definitive diagnostic standard. No guidelines exist for how to best manage or counsel this patient. Increased surveillance has not been demonstrated to be useful or cost-effective. Insurance coverage for increased surveillance may be challenged. In addition, one of the key risk management recommendations - limiting exercise - runs counter to physical activity recommendations known to be beneficial for a wide range of health conditions.

- **Familial Hypercholesterolemia (FH) likely pathogenic** variant in *LDLR*. The prevalence of FH in an unselected population is approximately 1/220.(Khera et al., 2016; Abul-Husn et al., 2016; Akioyamen et al., 2017) The estimated PPV is approximately 90%. **Clinical Action:** A functional study, such as LDL cholesterol level, can help adjudicate whether this is a true positive or a false positive. Since the penetrance of FH has been determined to be approximately the same in high risk and unselected cohorts, it is appropriate to manage confirmed cases according to FH guidelines. (Reiner, 2015) The presence of a positive genetic test result alone meets criteria for most commercial insurance coverage policies for the preventive management of FH.

Despite the absence of guidance, genetic testing of low-risk individuals is occurring more frequently and is likely to increase significantly in the near future. In the absence of penetrance and prevalence data, with few confirmatory tests, and few clinical utility guidelines describing preventive interventions, some would argue that genetic testing of low-risk individuals should not be conducted or restricted to certain circumstances, such as opportunistic screening in the context of a genetically literate care team. However, due to consumer demand, genetic testing of low-risk individuals will likely proceed. Laboratories can mitigate the risk by modeling estimates of the missing data, by designing screening tests to minimize potential harms from false positives, and by providing consumers and physicians abundantly clear limitations to the clinical utility of the results.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

References

Abul-Husn NS, Manickam K, Jones LK, Wright EA, Hartzel DN, Gonzaga-Jauregui C, O'Dushlaine C, Leader JB, Lester Kirchner H, Lindbuchler DM, Barr ML, Giovanni MA, Ritchie MD, Overton JD, Reid JG, Metpally RPR, Wardeh AH, Borecki IB, Yancopoulos GD, Baras A, Shuldiner AR, Gottesman O, Ledbetter DH, Carey DJ, Dewey FE, Murray MF. 2016. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science* 354. DOI: 10.1126/science.aaf7000.

Adams MC, GeneScreen Investigators, Evans JP, Henderson GE, Berg JS. 2016. The promise and peril of genomic screening in the general population. *Genetics in Medicine* 18:593–599. DOI: 10.1038/gim.2015.136.

Akbari MR, Gojska N, Narod SA. 2017. Coming of age in Canada: a study of population-based genetic testing for breast and ovarian cancer. *Current oncology* 24:282–283.

Akioyamen LE, Genest J, Shan SD, Reel RL, Albaum JM, Chu A, Tu JV. 2017. Estimating the prevalence of heterozygous familial hypercholesterolaemia: a systematic review and meta-analysis. *BMJ open* 7:e016461.

Akobeng AK. 2007. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatrica* 96:338–341. DOI: 10.1111/j.1651-2227.2006.00180.x.

Brothers KB, Vassy JL, Green RC. 2019. Reconciling Opportunistic and Population Screening in Clinical Genomics. *Mayo Clinic proceedings. Mayo Clinic* 94:103–109.

Contributors to Wikimedia projects. 2005.Screening (medicine) - Wikipedia. *Available at https://en.wikipedia.org/wiki/Screening_(medicine)* (accessed June 28, 2019).

Diagnostic Test Calculator. *Available at http://araw.mede.uic.edu/cgi-bin/testcalc.pl* (accessed June 28, 2019).

Dorschner MO, Amendola LM, Turner EH, Robertson PD, Shirts BH, Gallego CJ, Bennett RL, Jones KL, Tokita MJ, Bennett JT, Kim JH, Rosenthal EA, Kim DS, National Heart, Lung, and Blood Institute Grand Opportunity Exome Sequencing Project, Tabor HK, Bamshad MJ, Motulsky AG, Scott CR, Pritchard CC, Walsh T, Burke W, Raskind WH, Byers P, Hisama FM, Nickerson DA, Jarvik GP. 2013. Actionable, pathogenic incidental findings in 1,000 participants' exomes. *American journal of human genetics* 93:631–640.

Eisenberg MJ. 1995. Accuracy and predictive values in clinical decision-making. *Cleveland Clinic Journal of Medicine* 62:311–316. DOI: 10.3949/ccjm.62.5.311.

Gabai-Kapara E, Lahad A, Kaufman B, Friedman E, Segev S, Renbaum P, Beeri R, Gal M, Grinshpun-Cohen J,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Djemal K, Mandell JB, Lee MK, Beller U, Catane R, King M-C, Levy-Lahad E. 2014. Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2. *Proceedings of the National Academy of Sciences of the United States of America* 111:14205–14210.

Haggerty CM, James CA, Calkins H, Tichnell C, Leader JB, Hartzel DN, Nevius CD, Pendergrass SA, Person TN, Schwartz M, Ritchie MD, Carey DJ, Ledbetter DH, Williams MS, Dewey FE, Lopez A, Penn J, Overton JD, Reid JG, Lebo M, Mason-Suares H, Austin-Tse C, Rehm HL, Delisle BP, Makowski DJ, Mehra VC, Murray MF, Fornwalt BK. 2017. Electronic health record phenotype in subjects with genetic variants associated with arrhythmogenic right ventricular cardiomyopathy: a study of 30,716 subjects with exome sequencing. *Genetics in medicine: official journal of the American College of Medical Genetics* 19:1245–1252.

Jamuar SS, Kuan JL, Brett M, Tiang Z, Tan WLW, Lim JY, Liew WKM, Javed A, Liew WK, Law HY, Tan ES, Lai A, Ng I, Teo YY, Venkatesh B, Reversade B, Tan EC, Foo R. 2016. Incidentalome from Genomic Sequencing: A Barrier to Personalized Medicine? *EBioMedicine* 5:211–216.

Jennings L, Van Deerlin VM, Gulley ML, College of American Pathologists Molecular Pathology Resource Committee. 2009. Recommended principles and practices for validating clinical molecular pathology tests. *Archives of pathology & laboratory medicine* 133:743–755.

Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, Others. 2017. 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. *Genetics in medicine: official journal of the American College of Medical Genetics* 19:249.

Khera AV, Won H-H, Peloso GM, Lawson KS, Bartz TM, Deng X, van Leeuwen EM, Natarajan P, Emdin CA, Bick AG, Morrison AC, Brody JA, Gupta N, Nomura A, Kessler T, Duga S, Bis JC, van Duijn CM, Cupples LA, Psaty B, Rader DJ, Danesh J, Schunkert H, McPherson R, Farrall M, Watkins H, Lander E, Wilson JG, Correa A, Boerwinkle E, Merlini PA, Ardissino D, Saleheen D, Gabriel S, Kathiresan S. 2016. Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. *Journal of the American College of Cardiology* 67:2578–2589.

Lu JT, Ferber M, Hagenkord J, Levin E, South S, Kang HP, Strong KA, Bick DP. 2019. Evaluation for Genetic Disorders in the Absence of a Clinical Indication for Testing: Elective Genomic Testing. *The Journal of molecular diagnostics: JMD* 21:3–12.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Macklin S, Durand N, Atwal P, Hines S. 2018. Observed frequency and challenges of variant reclassification in a hereditary cancer clinic. *Genetics in medicine: official journal of the American College of Medical Genetics* 20:346–350.

Mammogram Accuracy - Accuracy of Mammograms. *Available at https://ww5.komen.org/BreastCancer/AccuracyofMammograms.html* (accessed June 28, 2019).

Manickam K, Buchanan AH, Schwartz MLB, Hallquist MLG, Williams JL, Rahm AK, Rocha H, Savatt JM, Evans AE, Butry LM, Lazzeri AL, Lindbuchler DM, Flansburg CN, Leeming R, Vogel VG, Lebo MS, Mason-Suares HM, Hoskinson DC, Abul-Husn NS, Dewey FE, Overton JD, Reid JG, Baras A, Willard HF, McCormick CZ, Krishnamurthy SB, Hartzel DN, Kost KA, Lavage DR, Sturm AC, Frisbie LR, Nate Person T, Metpally RP, Giovanni MA, Lowry LE, Leader JB, Ritchie MD, Carey DJ, Justice AE, Lester Kirchner H, Andrew Faucett W, Williams MS, Ledbetter DH, Murray MF. 2018. Exome Sequencing–Based Screening for BRCA1/2 Expected Pathogenic Variants Among Adult Biobank Participants. *JAMA Network Open* 1:e182140–e182140.

Maxim LD, Daniel Maxim L, Niebo R, Utell MJ. 2014. Screening tests: a review with examples. *Inhalation Toxicology* 26:811–828. DOI: 10.3109/08958378.2014.955932.

Mayo Clinic GeneGuide - Results. 2018. *Available at https://www.mayoclinic.org/mayoclinic-geneguide/results* (accessed June 28, 2019).

McCabe LL, McCabe ERB. 2003. Population screening in the age of genomic medicine. *Journal of medicine*.

Media H.Fortunately, Angelina Jolie Isn't On Medicare. *Available at https://www.healthleadersmedia.com/finance/fortunately-angelina-jolie-isnt-medicare* (accessed August 17, 2019).

Mersch J, Brown N, Pirzadeh-Miller S, Mundt E, Cox HC, Brown K, Aston M, Esterling L, Manley S, Ross T. 2018. Prevalence of Variant Reclassification Following Hereditary Cancer Genetic Testing. *JAMA: the journal of the American Medical Association* 320:1266–1274.

Metcalfe KA, Lubinski J, Gronwald J, Huzarski T, McCuaig J, Lynch HT, Karlan B, Foulkes WD, Singer CF, Neuhausen SL, Senter L, Eisen A, Sun P, Narod SA, Hereditary Breast Cancer Clinical Study Group. 2018. The risk of breast cancer in BRCA1 and BRCA2 mutation carriers without a first-degree relative with breast cancer. *Clinical genetics* 93:1063–1068.

Murray MF. 2016. Your DNA is not your diagnosis: getting diagnoses right following secondary genomic findings.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Genetics in medicine: official journal of the American College of Medical Genetics 18:765–767.

Murray, M. F., J. P. Evans, M. Angrist, K. Chan, W. Uhlmann, D. L. Doyle, S. M. Fullerton, T. Ganiats, J. Hagenkord, S. Imhof, S. H. Rim, L. Ortmann, N. Aziz, W. D. Dotson, E. Matloff, K. Young, K. Kaphingst, A. Bradbury, J. Scott, C. Wang, A. Zauber, M. Levine, B. Korf, D. Leonard, C. Wicklund, G. Isham, and M. J. Khoury. 2018. A Proposed Approach for Implementing Genomics-Based Screening Programs for Healthy Adults. *NAM Perspectives, National Academy of Medicine, Washington, DC*. DOI: 10.31478/201812a.

Ndase P, Celum C, Kidoguchi L, Ronald A, Fife KH, Bukusi E, Donnell D, Baeten JM, Partners PrEP Study Team. 2015. Frequency of false positive rapid HIV serologic tests in African men and women receiving PrEP for HIV prevention: implications for programmatic roll-out of biomedical interventions. *PloS one* 10:e0123005.

Olfson E, Cottrell CE, Davidson NO, Gurnett CA, Heusel JW, Stitziel NO, Chen L-S, Hartz S, Nagarajan R, Saccone NL, Bierut LJ. 2015. Identification of Medically Actionable Secondary Findings in the 1000 Genomes. *PloS one* 10:e0135193.

OMIM Entry - # 145600 - MALIGNANT HYPERTHERMIA, SUSCEPTIBILITY TO, 1; MHS1. *Available at https://omim.org/entry/145600* (accessed June 28, 2019).

OMIM Entry - # 158350 - COWDEN SYNDROME 1; CWS1. *Available at https://omim.org/entry/158350* (accessed June 28, 2019).

OMIM Entry - # 180200 - RETINOBLASTOMA; RB1. *Available at https://omim.org/entry/180200* (accessed June 28, 2019).

OMIM Entry - # 193300 - VON HIPPEL-LINDAU SYNDROME; VHL. *Available at https://omim.org/entry/193300* (accessed June 28, 2019).

Rao G. 2004. Remembering the meanings of sensitivity, specificity, and predictive values. *The Journal of family practice* 53:53.

Rego S, Dagan-Rosenfeld O, Zhou W, Sailani MR, Limcaoco P, Colbert E, Avina M, Wheeler J, Craig C, Salins D, Röst HL, Dunn J, McLaughlin T, Steinmetz LM, Bernstein JA, Snyder MP. 2018. High-frequency actionable pathogenic exome variants in an average-risk cohort. *Cold Spring Harbor molecular case studies* 4. DOI: 10.1101/mcs.a003178.

Reiner Ž. 2015. Management of patients with familial hypercholesterolaemia. *Nature reviews. Cardiology* 12:565–575.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

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, ACMG Laboratory Quality Assurance Committee. 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. *Genetics in medicine: official journal of the American College of Medical Genetics* 17:405–424.

Rocha H. APC Variant Identification in an Unselected Patient Population: Where are the Polyps?

Romero-Brufau S, Huddleston JM, Escobar GJ, Liebow M. 2015. Why the C -statistic is not informative to evaluate early warning scores and what metrics to use. *Critical care* 19:1–6.

Romero J, Mejia-Lopez E, Manrique C, Lucariello R. 2013. Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC/D): A Systematic Literature Review. *Clinical Medicine Insights. Cardiology* 7:97–114.

Tavtigian SV, on behalf of the ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI), Greenblatt MS, Harrison SM, Nussbaum RL, Prabhu SA, Boucher KM, Biesecker LG. 2018. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genetics in Medicine* 20:1054–1060. DOI: 10.1038/gim.2017.210.

Tenny S, Hoffman MR. 2019. Prevalence. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing,.

Trevethan R. 2017. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Frontiers in public health* 5:307.

Wikipedia contributors. 2019.Base rate fallacy. *Available at https://en.wikipedia.org/w/index.php?title=Base_rate_fallacy&oldid=901526725* (accessed June 28, 2019).

Wilson JMG, Jungner G, Health Organization W. 1968. Principles and practice of screening for disease.

TABLES AND FIGURES

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

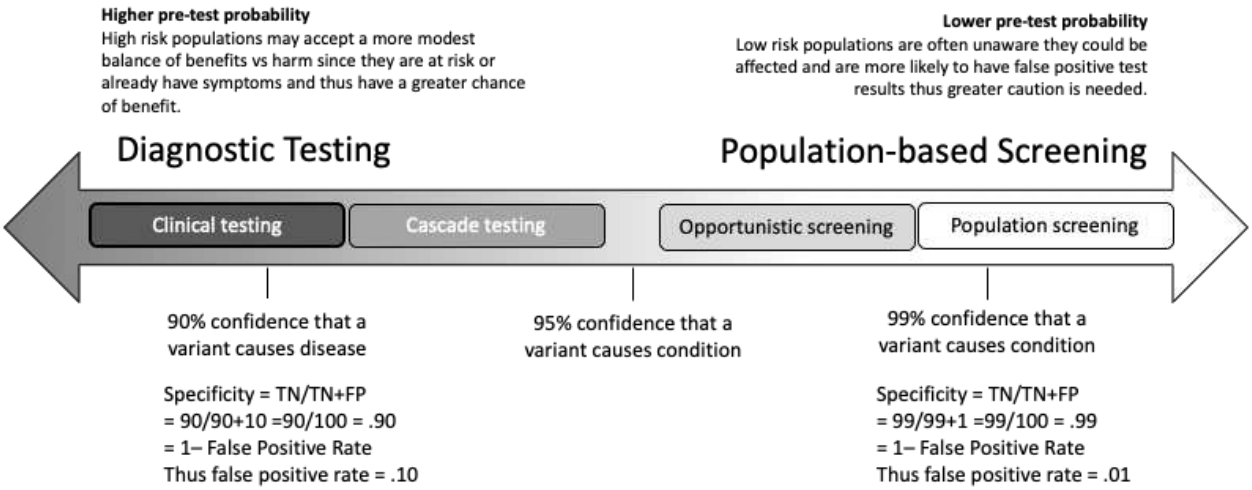


Figure 1. Intended Use or Purpose of the Test. The clinical implications of false results must be considered when determining the appropriate sensitivity and specificity for the intended use. Tests intended for higher risk populations (diagnostic testing) may tolerate more false positives than tests intended for low risk populations (screening). The false positive rate is reflected in the test's clinical specificity.

	Generic definitions	**Genetics example**
Analytical validity	The test accurately detects the analyte when it is present (analytical sensitivity) and does not detect it when it is absent (analytical specificity). Confirm by orthogonal technology during development phase.	The test accurately detects a sequence variant in *BRCA1* when it is present and does not detect it when it is absent. Confirm by orthogonal technology until performance metrics are well-established.
Clinical validity	The test accurately detects the disease when it is known to be present and does not detect it when it is known to be absent. A "gold standard" or index test is used to identify samples from individuals known to have the condition and those known not to have the condition. During test development, these samples are processed using the new test. The results are compared to those of the index test to determine the test's sensitivity and specificity.	The test accurately identifies samples from individuals known to have HBOC Syndrome and it does not identify HBOC Syndrome in samples from individuals known not to have HBOC Syndrome. A "gold standard" test is not available to establish a truth set for use during development.
Clinical Sensitivity	The ability of the test to correctly identify those patients with the disease when present. (Note: This is different than diagnostic yield.)	"I know this person has HBOC. What is the chance that the test will show that this person has it?"(Rao, 2004)

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Clinical Specificity	The ability of the test to correctly identify those patients without the disease when absent.	"I know this person doesn't have HBOC. What is the chance that the test will show that this person doesn't have it?"(Rao, 2004)
Positive Predictive Value	The probability that a person with a positive result has the condition. Dependent on the prevalence of the disease in the test population.	"I just received a positive HBOC test result for my patient. What is the chance that my patient actually has the disease?"(Rao, 2004) Dependent on the prevalence of the disease in the test population.
Negative Predictive Value	The probability that a person with a negative result does not have the condition. Dependent on the prevalence of the disease in the test population.	"I just received a negative HBOC test result for my patient. What is the chance that my patient actually doesn't have the disease?"(Rao, 2004) (Assume 100% analytical sensitivity) Dependent on the prevalence of the disease in the test population.
Penetrance	Not applicable.	The proportion of individuals with an inherited genetic syndrome (eg, HBOC) who exhibit clinical symptoms (eg, breast cancer) over time (eg, lifetime).
Classification	Evidence-based scoring system for determining likelihood of disease. Example Pap Smear: Cancer, Cervical Intraepithelial Neoplasia (3 levels), Atypical Squamous Cells of Uncertain Significance, Benign	Evidence-based scoring system for determining whether a variant is likely to cause disease. Example DNA Sequencing: Pathogenic, Likely pathogenic, Variant of uncertain significance, Likely Benign, Benign
Interpretation	Positive/Negative (Positive results are clinically actionable)	Positive/Negative (Positive results are clinically actionable)

Table 1. Generic test development definitions and examples of their application to genetic testing.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

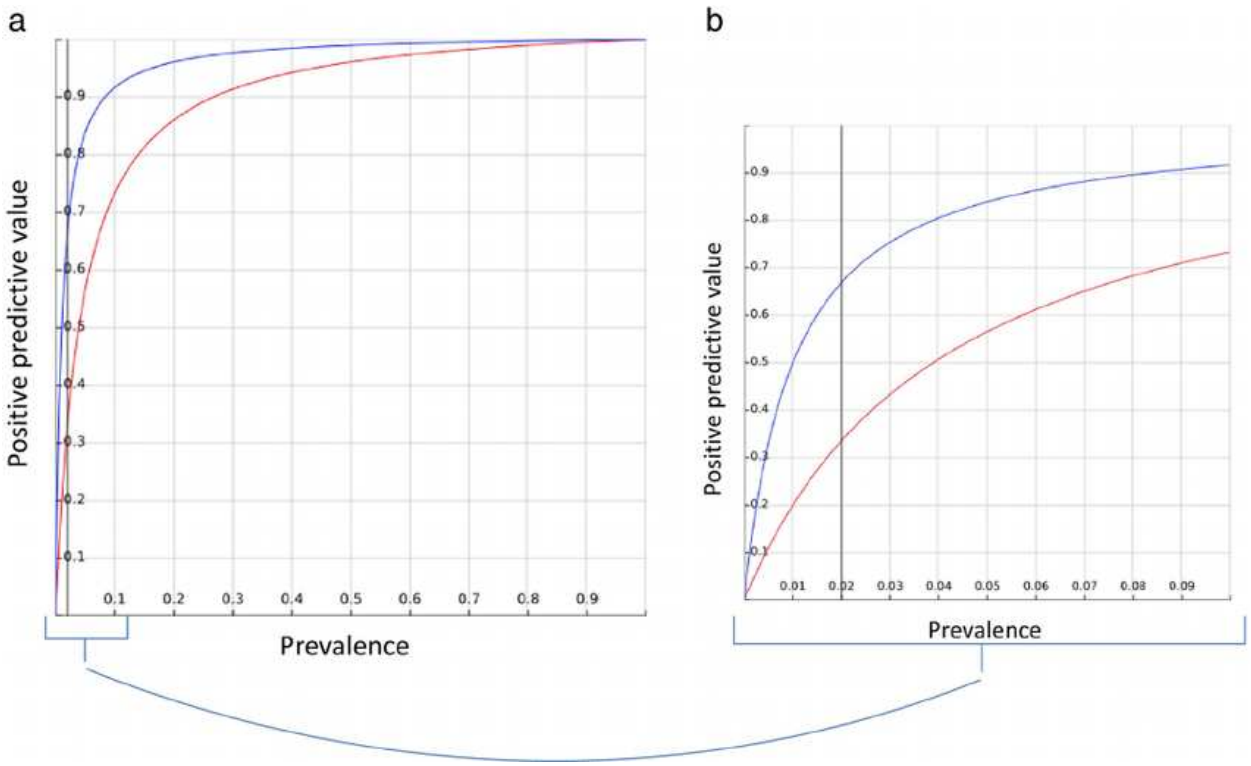


Figure 2. Impact of prevalence on PPV. When the disease prevalence is low in the test population, small changes in the specificity can have a large impact on the positive predictive value of the test. PPV as a function of prevalence for two tests: Test A (blue), with a sensitivity of 99% and a specificity of 99%; and Test B (red), with a sensitivity of 99% and a specificity of 96%. a) Full range of possible PPV and prevalence, from 0 to 1. b) Magnified region of prevalence <0.1, a gray line to show an example prevalence of 0.02. A decrease of only 3% in specificity can mean a 50% decrease in PPV: from 0.33 to 0.66. Adapted with permission from Romero-Brufau, et al. (Romero-Brufau et al., 2015) The prevalence of the disease is equal to the *a priori* probability that a subject selected at random from the test population has the condition.

Table 2: Estimating the Test Specificity

Example	Overall Positive rate	LP rate	Specificity	Prevalence	PPV Range
CDC Tier 1	1.50%	0.50%	99.95%	1/200-1/500	80% - 91%
ACMG59™	3.00%	1.00%	99.90%	1/200-1/10,000*	9% - 83%
ACMG59™ (at 6% overall positive rate)	6.00%	2.00%	99.79%	1/200-1/10,000*	5% - 72%

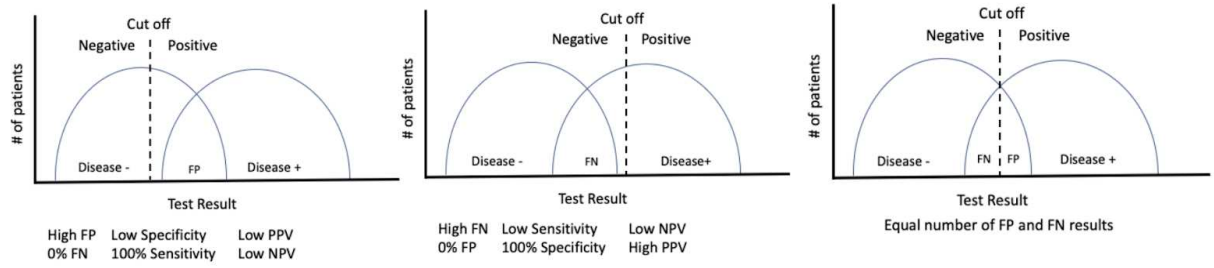
Table 2. Estimating the Test Specificity. Test specificity is estimated by assuming ⅓ of the overall positive rate is due to likely pathogenic variants (LP) and then calculating the specificity as described in Supplemental Figure 2 (Specificity = TN/(TN+FP) = (1 - Positive Rate)/(1 - 29/30 x Positive Rate). Note how specificity changes with overall positive rate (orange). *Some conditions on ACMG59™ have a prevalence less than 1/10,000.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Table 3: Computing the PPV (post-test probability) from prevalence and specificity						
		Disease Prevalence (pre-test probability)				
		1/10,000	1/1,000	1/500	1/200	1/50
Specificity	99.00%	0.99%	9.10%	16.69%	33.44%	67.11%
	99.50%	1.96%	16.68%	28.61%	50.13%	80.32%
	99.80%	4.76%	33.36%	50.05%	71.53%	91.07%
	99.90%	9.09%	50.03%	66.71%	83.40%	95.33%
	99.95%	16.66%	66.68%	80.03%	90.95%	97.60%
	99.995%	66.67%	95.24%	97.57%	99.01%	99.76%

Table 3. Computing the PPV from prevalence and specificity. Small decreases in specificity can have a significant impact on the PPV of likely pathogenic variants over a range of prevalences representative of monogenic inherited diseases. The PPV calculations for CDC Tier 1 (yellow) and ACMG59 (orange) conditions are as follows: PPV = sensitivity x prevalence / [sensitivity x prevalence + (1 - specificity) x (1 - prevalence)].(Tenny & Hoffman, 2019) The specificity estimates for CDC Tier 1 (yellow) and ACMG59™ (orange) conditions are from Table 2. The model is intended to provide estimates and show trends. In practice, each condition should be considered individually.

Supplemental Content

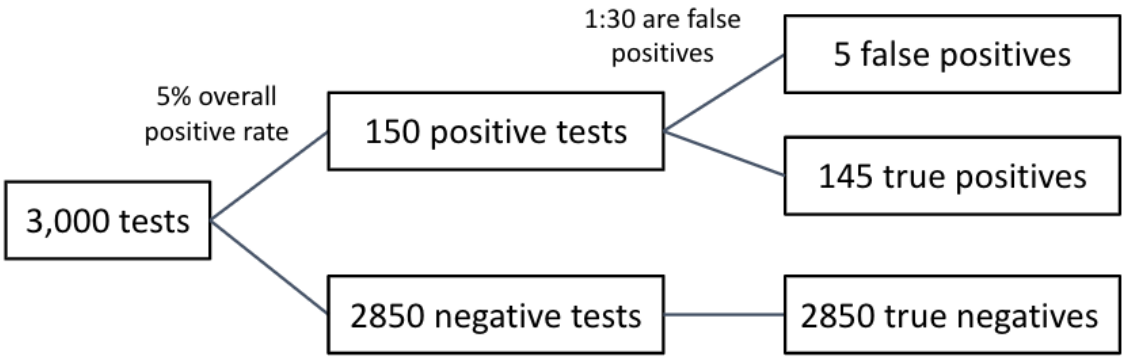


Supplemental Figure 1. The sensitivity and specificity of a test can be adjusted during the test development phase to allow for more or less false positives or false negatives, as appropriate for the test's intended use. The graphs show the results obtained by the new test when it tested a cohort of known positive samples and known negative samples, as determined by an existing diagnostic test. Ideally, the new test could perfectly discriminate between positive and negative samples. In practice this is rarely the case, so thoughtful trade-offs between desired sensitivity and specificity are necessary. FP = false positive, FN = false negative, PPV = positive predictive value, NPV = negative predictive value.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019

Screening test	Confirmatory test
Pap smear	Colposcopy
SickleDex or Thalassemia by genetic screen	Hemoglobin electrophoresis
HIV by ELISA	HIV by Western blot
Non-invasive prenatal testing (NIPT)	Karyotype/FISH
Genetic Factor XIII Deficiency or Glanzmann's	Mixing studies or Platelet Activation Test
Newborn screening by MS/MS	Genetic testing

Supplemental Table 1. Screening test and confirmatory diagnostic test pairs. Examples of screening tests and confirmatory tests. Confirmatory tests can serve as the "gold standard" or index test comparator during test validation as well as a follow up test for positive screening results in practice.



Specificity = TN/(TN+FP) = 2850 / (2850 + 5) = 99.82%

Supplemental Figure 2. Relationship between overall positive rate and specificity. In this example, the test has 100% analytical sensitivity and specificity, as determined during the test validation. After processing 3,000 samples, the lab determined that the positive rate in their setting is 5%. This means they have 150 positive results and 2850 negative results. The positive/negative cut off has been set to allow for 1/10 LP positives to be a false positive; thus 1/30 positives are false positives. Therefore, 5 of the positive results are false positives. and test specificity = 2850/(2850+5) = 99.82%. This method is used to calculate specificities for Tables 2 and 3 for varying positive test rates.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27922v1 | CC BY 4.0 Open Access | rec: 28 Aug 2019, publ: 28 Aug 2019