Calculating the “Number Needed” in Genomic Screening

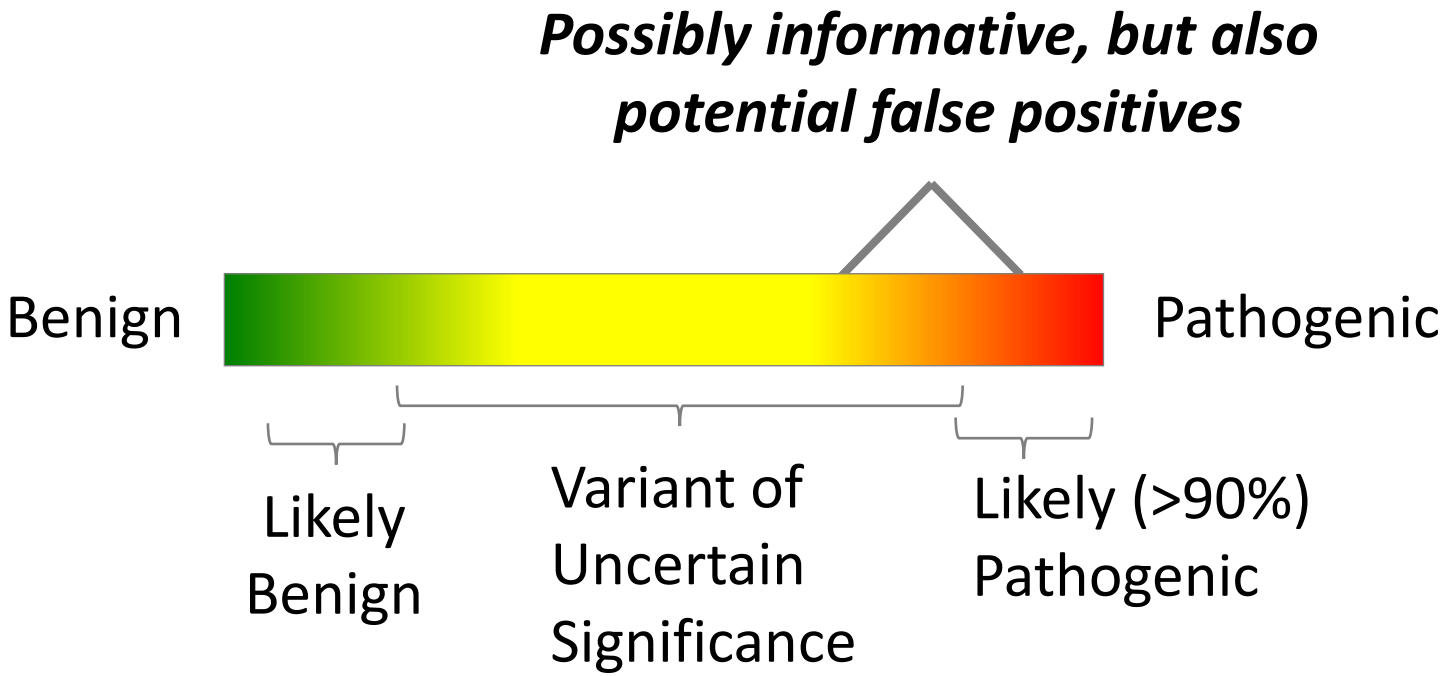
Jonathan S. Berg, MD/PhD
Department of Genetics
UNC Chapel Hill

Commonly used metrics for diagnostics

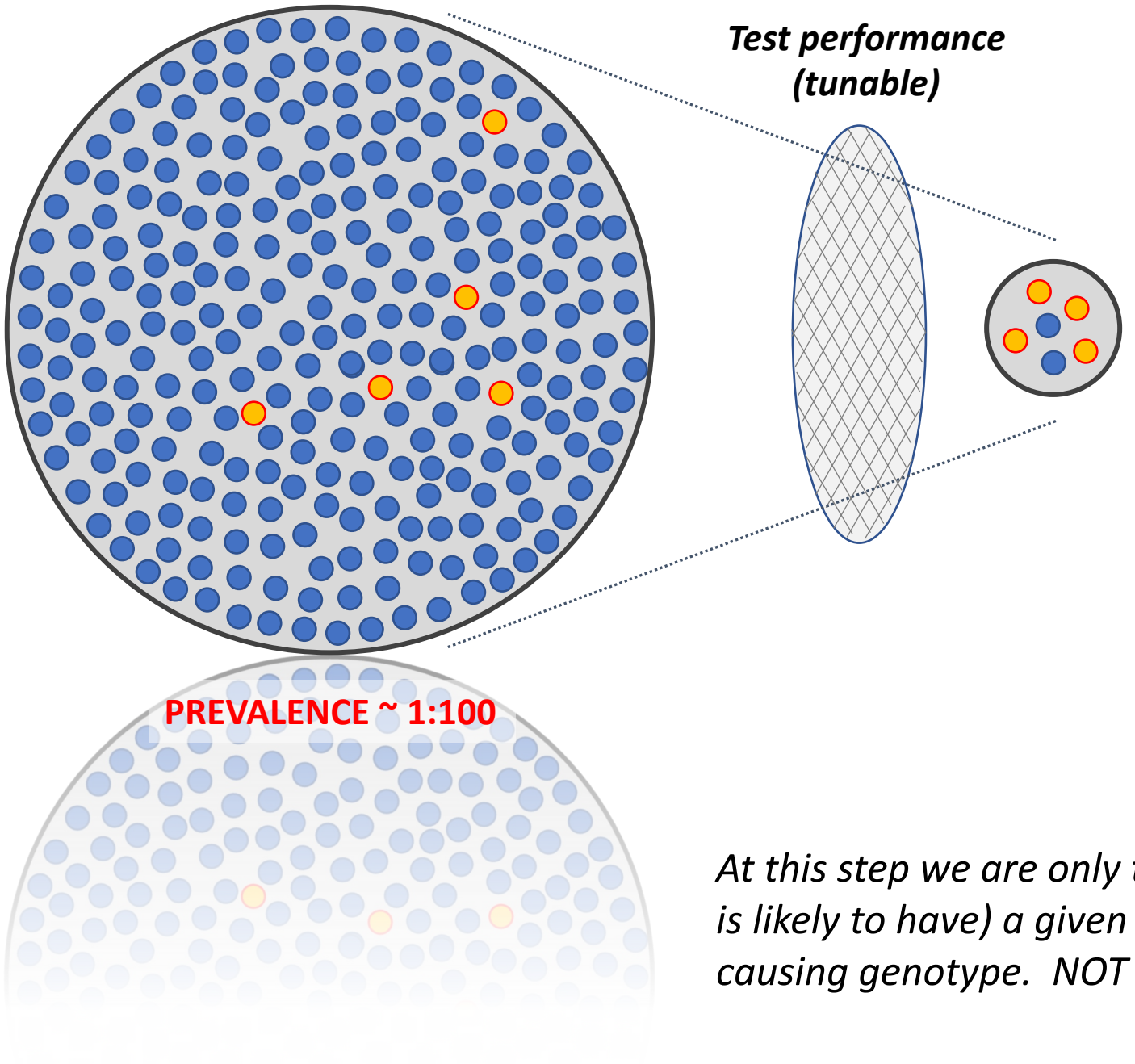
- NNT = Number of people that need to be treated to prevent one adverse event (Rembold, 1996)
  - Starts with individuals already identified as having risk factors
- NNS = Number of people that need to be screened to prevent one death or adverse event (Rembold, 1998)
  - Extends the NNT concept to screening for the underlying risk factors
- These concepts involve *risk for poor health outcomes* and the interventions that would be used to prevent them
- Can we use similar logic to examine genomic screening for monogenic diseases that convey high risk?

Part 1:
Test performance and population prevalence

Genomic screening and molecular diagnosis



- By definition, clinical variant classification defines a likelihood of pathogenicity
- High prior probability of monogenic disease in the clinical diagnostic setting generally means that P/LP variants are treated the same
- In screening, a molecular diagnosis consisting of LP variant(s) could represent a false positive result



	DISEASE	NO DISEASE
TEST +	TRUE POSITIVE	FALSE POSITIVE
TEST -	FALSE NEGATIVE	TRUE NEGATIVE

SENSITIVITY 80%

SPECIFICITY 99%

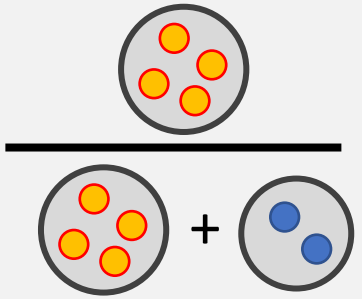
At this step we are only trying to determine whether an individual has (or is likely to have) a given monogenic disease, based on a putative disease-causing genotype. NOT whether they will manifest symptoms of disease.

	DISEASE	NO DISEASE
TEST +	TRUE POSITIVE	FALSE POSITIVE
TEST -	FALSE NEGATIVE	TRUE NEGATIVE

SENSITIVITY
80%

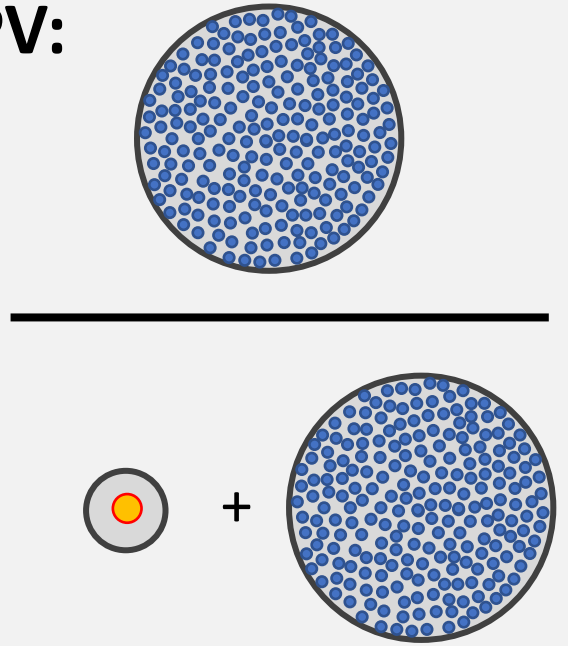
SPECIFICITY
99%

PPV:



= 67%

NPV:



>99.9%

The “number needed” for genomic screening (pt. 1)

- The number needed to diagnose (NNDx) one TP depends on sensitivity
- The positive predictive value (PPV) reflects the number of FP per TP
- Decisions about where to set the threshold for a “positive” screening result can be made based on the prevalence of conditions being screened and the tolerance for FP

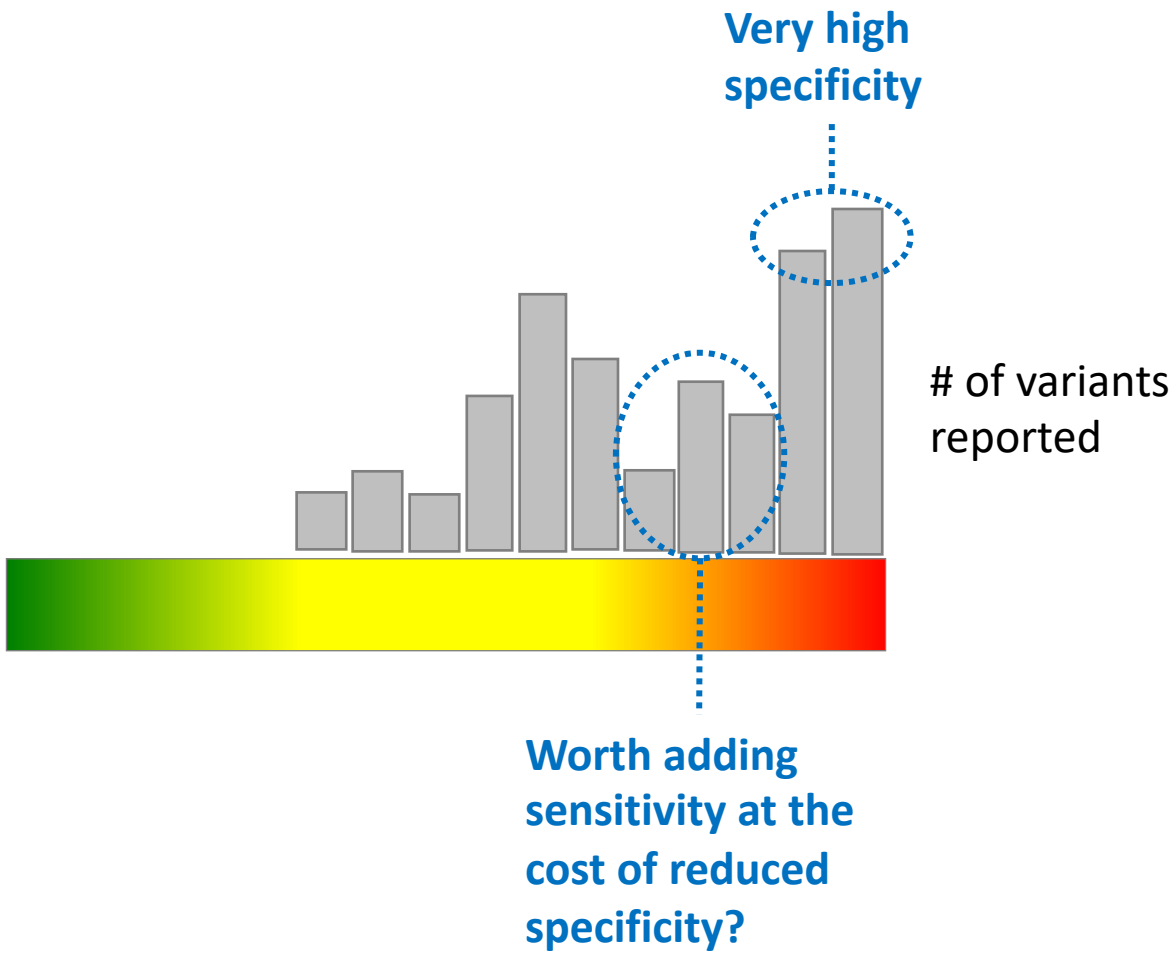
Prevalence	Clinical sensitivity	Clinical specificity	NNDx (per 1 TP)	FP (per 1 TP)	PPV
1/250	100% 90% 75%	75% 99% 100%	250 278 333	62 3 0	0.016 0.265 1
1/5000	100% 90% 75%	75% 99% 100%	5000 5556 6667	1250 56 0	0.0008 0.0177 1
1/80,000	100% 90% 75%	75% 99% 100%	80,000 88,889 106,667	20,000 889 0	<0.0001 0.0011 1
1/1,000,000	100% 90% 75%	75% 99% 100%	1,000,000 1,111,111 1,333,333	250,000 11,111 0	<0.0001 <0.0001 1

Very low prevalence characteristic of most rare diseases could create massive numbers of FP for every TP unless only well-established pathogenic variants are included as positive screening results

Critical values to know for the “number needed”

- *Prevalence* of each monogenic disease considered for screening
  - Most estimates are very broad and not evidence-based
  - How do we ascertain? What are the diagnostic criteria?
    - Clinical diagnosis overestimates penetrance and underestimates population prevalence if unaffected individuals are not ascertained; complicated by locus heterogeneity
    - Molecular diagnosis in affected individuals solves the locus heterogeneity problem but still biased due to ascertainment of affected individuals and most clinical studies do not have a true “denominator”
    - Population frequency of Pathogenic variants (plus/minus “expected pathogenic”) could provide a reasonable lower bound estimate

Thresholding clinical performance of genomic tests

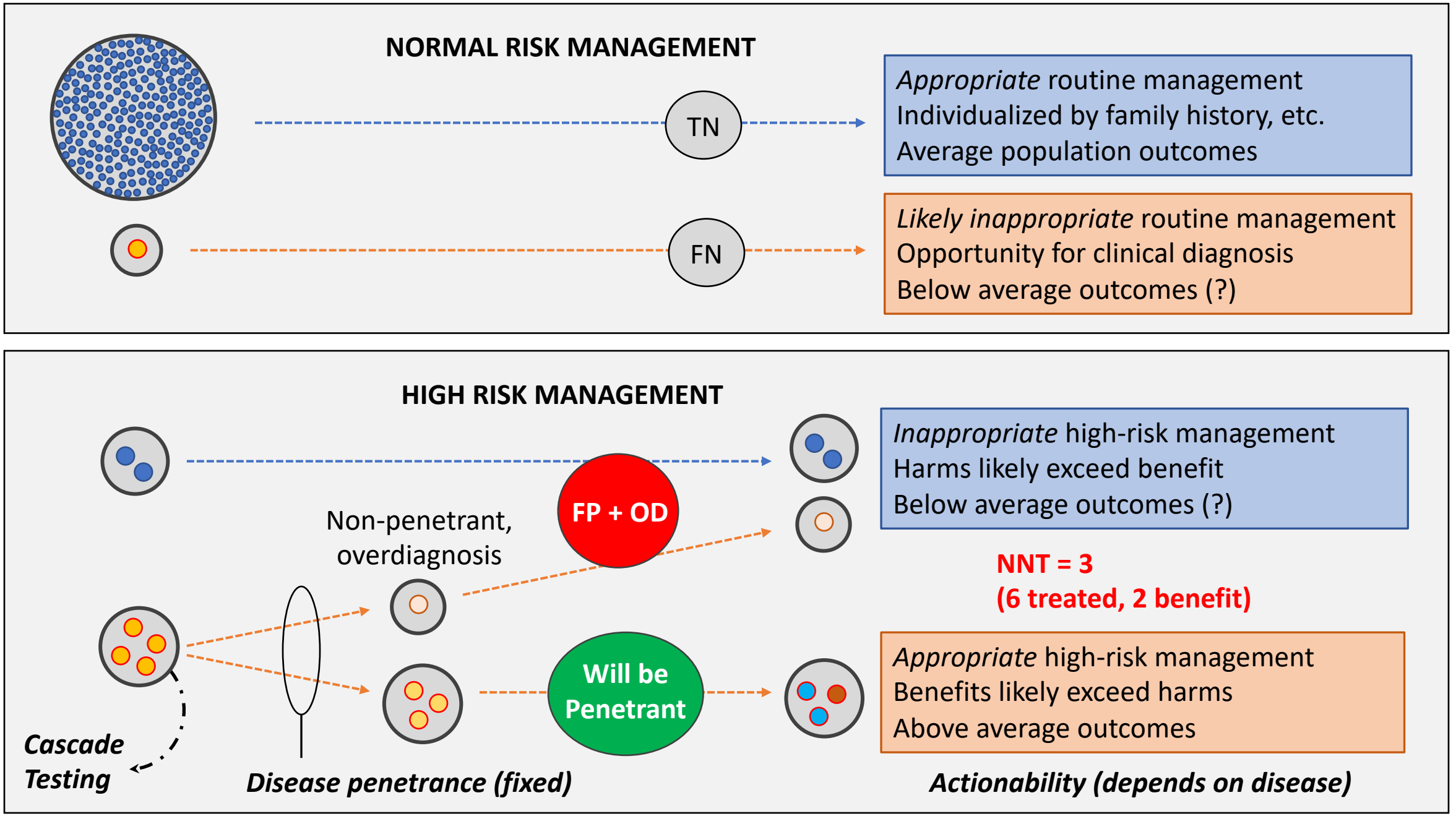


- Spectrum and frequency of different variant types reported in a genomic test
- Clinical labs could share these data to help with decisions about positive screening thresholds (rather than just all P/LP variants)
- Conditions with a *very well understood catalog of pathogenic and benign variants* will naturally have better performance [ClinGen]

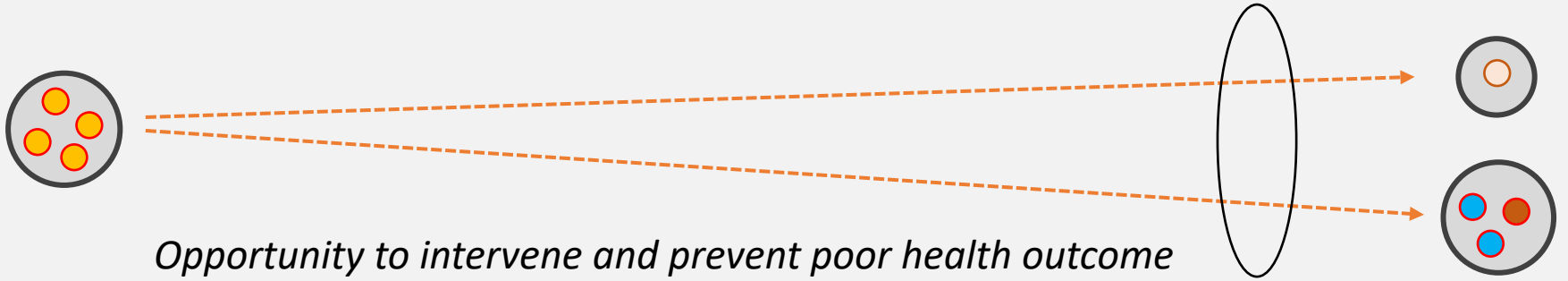
Critical values to know for the “number needed”

- *Clinical performance* of genomic screening test(s) for each monogenic disease considered for screening
  - Best data
  - If one variant is responsible for all cases = 100% sensitive, 100% specific
  - Most diseases have a much more complex mixture of variants identified in clinical testing
    - If P/LP/VUS are identified in 90% of cases and negative results found in 10%, upper bound of sensitivity = 90%
    - Specificity could be estimated for a given variant threshold, as the weighted average of the “probability of pathogenicity” for the collection of variants reported
    - Slightly more complicated for recessive conditions (compound heterozygosity P/LP with VUS) given the very low chance of identifying two rare variants in any individual

Part 2:
Penetrance, actionability, and preventing poor health outcomes

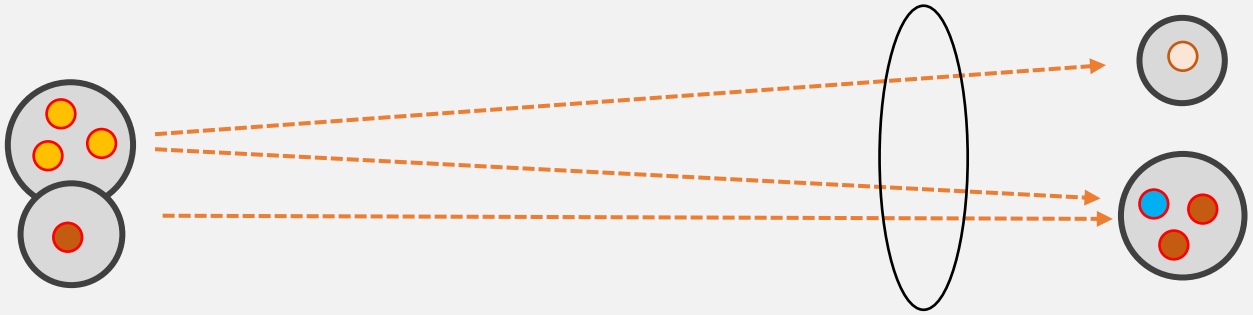

NORMAL RISK MANAGEMENT
TN
FN
Appropriate routine management
Individualized by family history, etc.
Average population outcomes
Likely inappropriate routine management
Opportunity for clinical diagnosis
Below average outcomes (?)
HIGH RISK MANAGEMENT
Non-penetrant,
overdiagnosis
FP + OD
Will be
Penetrant
Cascade
Testing
Disease penetrance (fixed)
Inappropriate high-risk management
Harms likely exceed benefit
Below average outcomes (?)
NNT = 3
(6 treated, 2 benefit)
Appropriate high-risk management
Benefits likely exceed harms
Above average outcomes
Actionability (depends on disease)

Screening prior to symptom onset



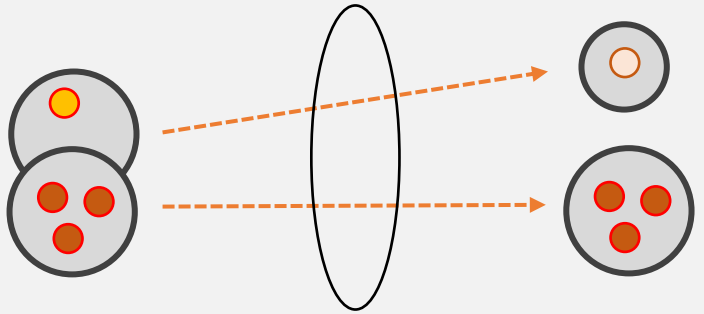
Opportunity to intervene and prevent poor health outcome

Screening coincident with symptom onset



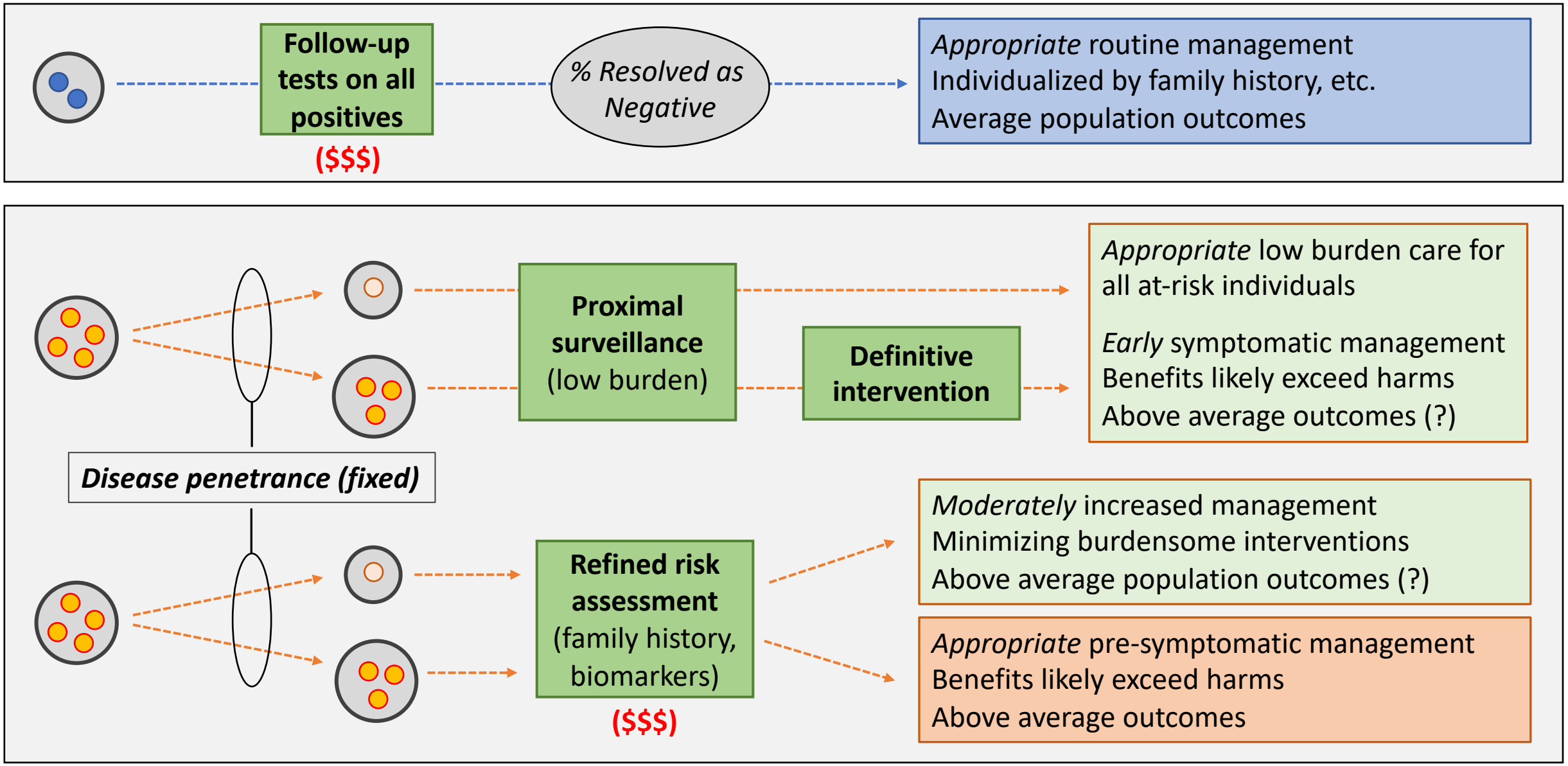
Understanding the natural history of disease and age-based penetrance will help to better tailor the timing of genomic screening and maximize the ability to reduce poor health outcomes

Screening after with symptom onset



Critical values to know for the “number needed”

- *Penetrance* of each monogenic disease considered for screening
  - Estimates from affected cohorts are subject to ascertainment bias, population-based ascertainment may reveal lower penetrance
    - ➢ This will increase the NNT since a greater proportion of individuals with a molecular diagnosis will not benefit from high-risk intervention
  - Better characterization of age-based onset of symptoms/natural history is necessary to determine optimal timing for screening and intervention
    - ➢ Ongoing surveillance and intervention contribute substantially to cost of the screening program, initiation too early may reduce cost effectiveness
    - ➢ Overdiagnosis and overtreatment of individuals who will be non-penetrant decreases the overall effectiveness of the screening program



Strategies to reduce false positives and mitigate the harms of overdiagnosis (determined for each condition) will be required to enable tailoring of test performance thresholds to maximize case finding

Critical values to know for the “number needed”

- *Quantitative actionability* of each monogenic disease considered for screening
  - How much reduction in morbidity/mortality can be expected among those who undergo high-risk management strategies?
  - How effective are strategies to reduce false positives and mitigate overdiagnosis?
  - In the absence of controlled trials or long-term follow-up, we will need credible estimates of NNT to reduce poor health outcomes

Conclusions

- Well-calibrated implementation of genomic screening will require key evidence for each monogenic disease considered:
  - Monogenic disease prevalence
  - Clinical test performance
  - Natural history / age of onset / penetrance
  - Quantitative actionability estimates
- Determine the variant threshold that qualifies for a “positive screen” based on PPV, ability to reduce FP, and strategies to mitigate overdiagnosis
- Consider how to incorporate these concepts into cost effectiveness analyses, varying the age at which screening is conducted based on natural history of disease