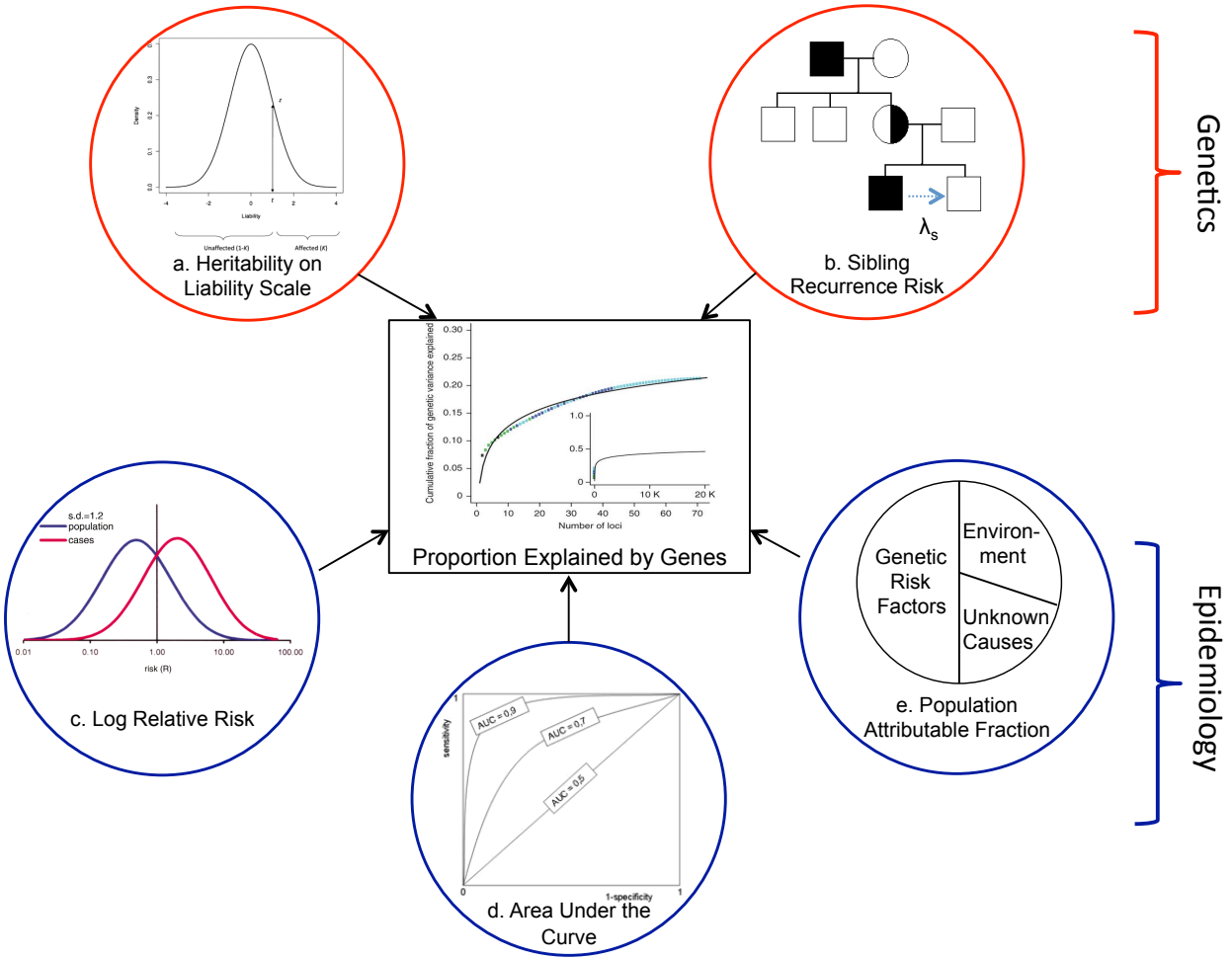
2017 SISG Module 10: Statistical & Quantitative Genetics of Disease

Converging fields of genetics, epidemiology & genetic epidemiology



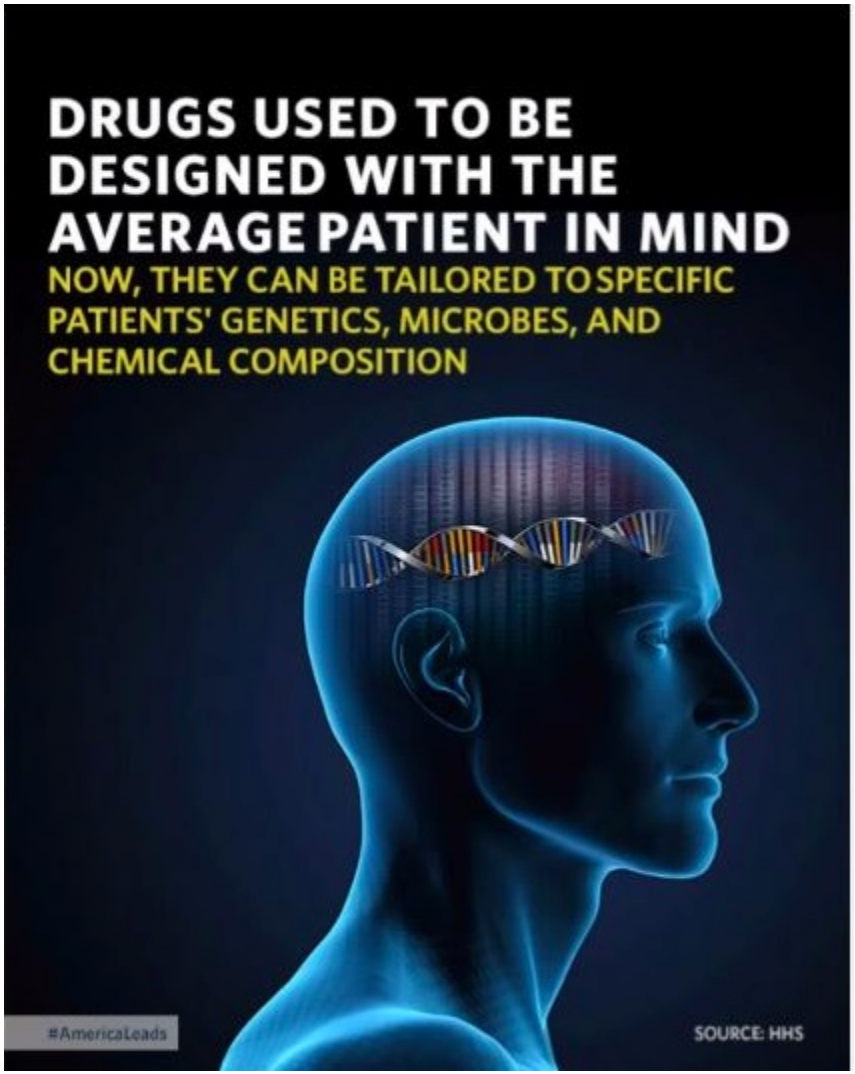
Naomi Wray
Psychiatric genetics

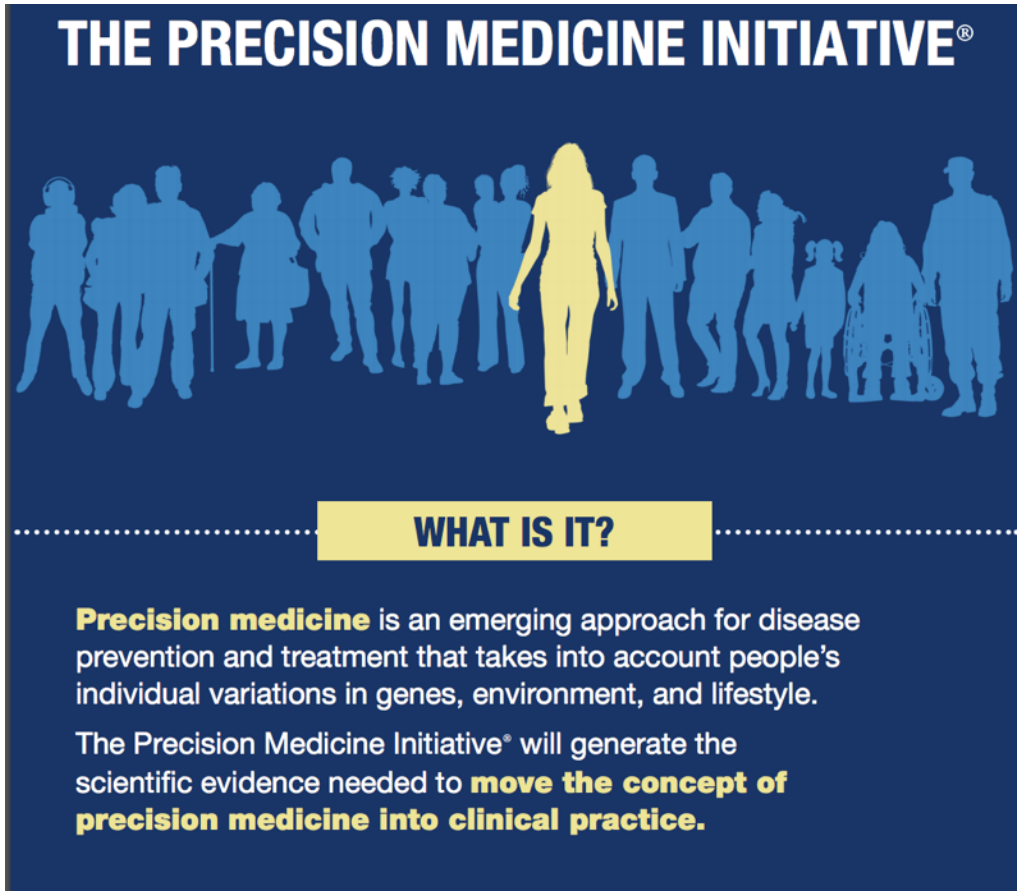
John Witte
Cancer genetics

Motivation for this module

- To unite the language of quantitative genetics (QG) and epidemiology
- Quantitative genetics of disease is often a tack on to QG of quantitative traits –here we make it the focus
- The new era of genomics bring QG of genetics of disease back into the foreground – a renewed relevance
- Understanding of prediction of disease risk in the precision medicine era

Precision Medicine Initiatives





http://syndication.nih.gov/multimedia/pmi/infographics/pmi-infographic.pdf

LONGER-TERM GOALS

Create a research cohort of **> 1 million American volunteers** who will share genetic data, biological samples, and diet/lifestyle information, all linked to their electronic health records if they choose.

Pioneer **a new model for doing science** that emphasizes **engaged participants, responsible data sharing, and privacy protection.**

Research based upon the cohort data will:

- Advance **pharmacogenomics,** the right drug for the right patient at the right dose
- Identify new targets for **treatment and prevention**
- Test whether **mobile devices** can encourage healthy behaviors
- Lay **scientific foundation** for precision medicine for **many diseases**

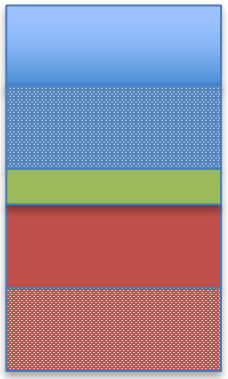
Course Outline

Thursday morning
- Lecture 1: Genetic epidemiology of disease; Heritability of liability (Naomi)
- Lecture 2: Single locus disease analysis (John)

Thursday afternoon
- Lecture 3: Single locus disease model; Power calculation for disease model (Naomi)
- Lecture 4: Modeling interactions: gene-environment, epistasis (John)

Friday morning
- Lecture 5:Multi-locus disease model (Naomi)
- Lecture 6: Modeling interactions: gene-environment, epistasis (John)

Friday afternoon
- Lecture 7: Risk Prediction (Naomi)
- Lecture 8: Rare variants (John)



Naomi lecture — More quantitative genetics theory
practical
Coffee

John lecture — More statistics/data analysis
practical

2017 SISG Brisbane Module 10: Statistical & Quantitative Genetics of Disease

Lecture 1
Quantifying the genetic contribution to disease

Naomi Wray

Aims of Lecture 1

If a disease affects 1% of the population and has heritability 80%

We will show why these statements are consistent :

If an individual is affected ~8% of his/her siblings affected

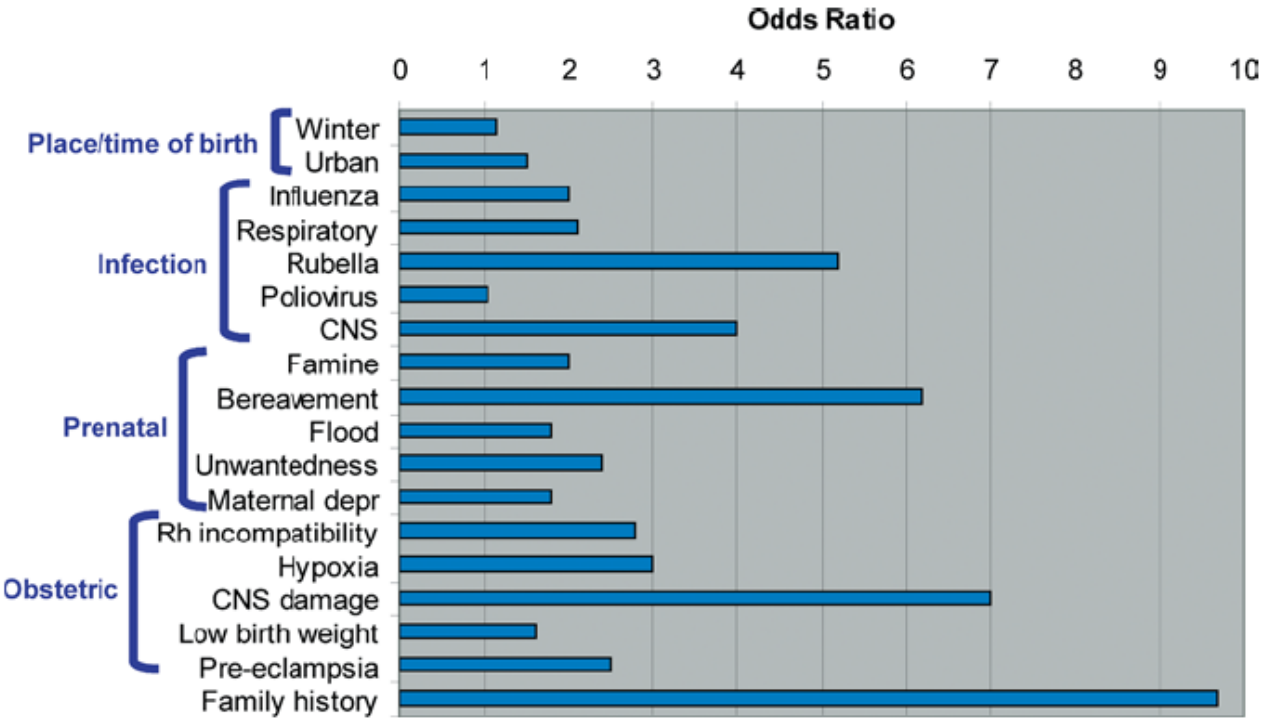
If an MZ twin is affected ~50% of their co-twins are affected

If an individual is affected > 60% will have no known family history

Bringing together genetic epidemiology and quantitative genetics

- The key papers were published 40 and 70 years ago……

Risk Factors for Schizophrenia



DOI: 10.1371/journal.pmed.0020212.g001

Figure 1. Comparison of a Selected Set of Relatively Well-Established Risk Factors for Schizophrenia, Focusing Mainly on Pre- and Antenatal Factors [6] (abbreviations: CNS, central nervous system; depr, depression; Rh, Rhesus)

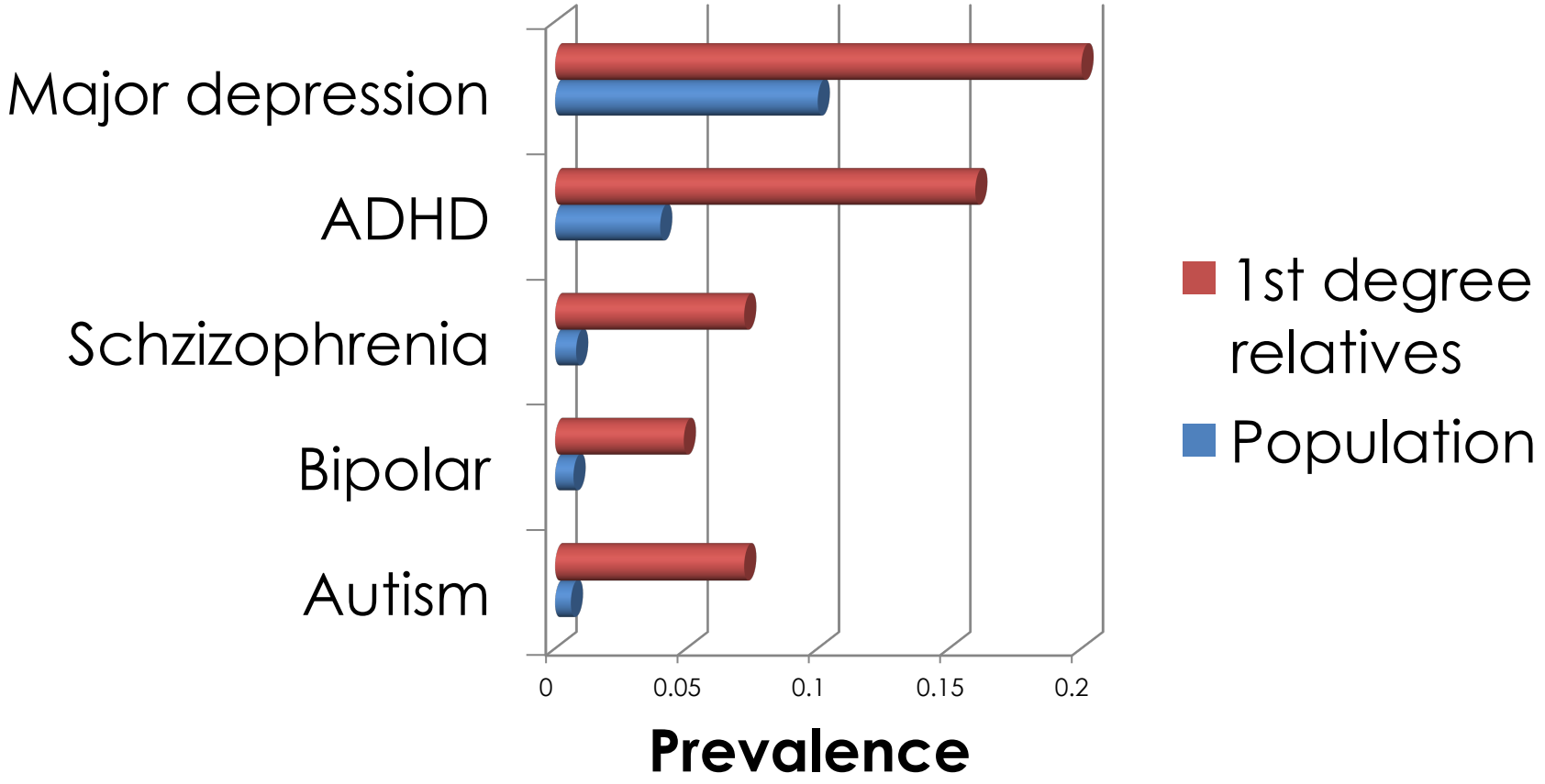
Sullivan, PLoS Med 05

Complex genetic diseases

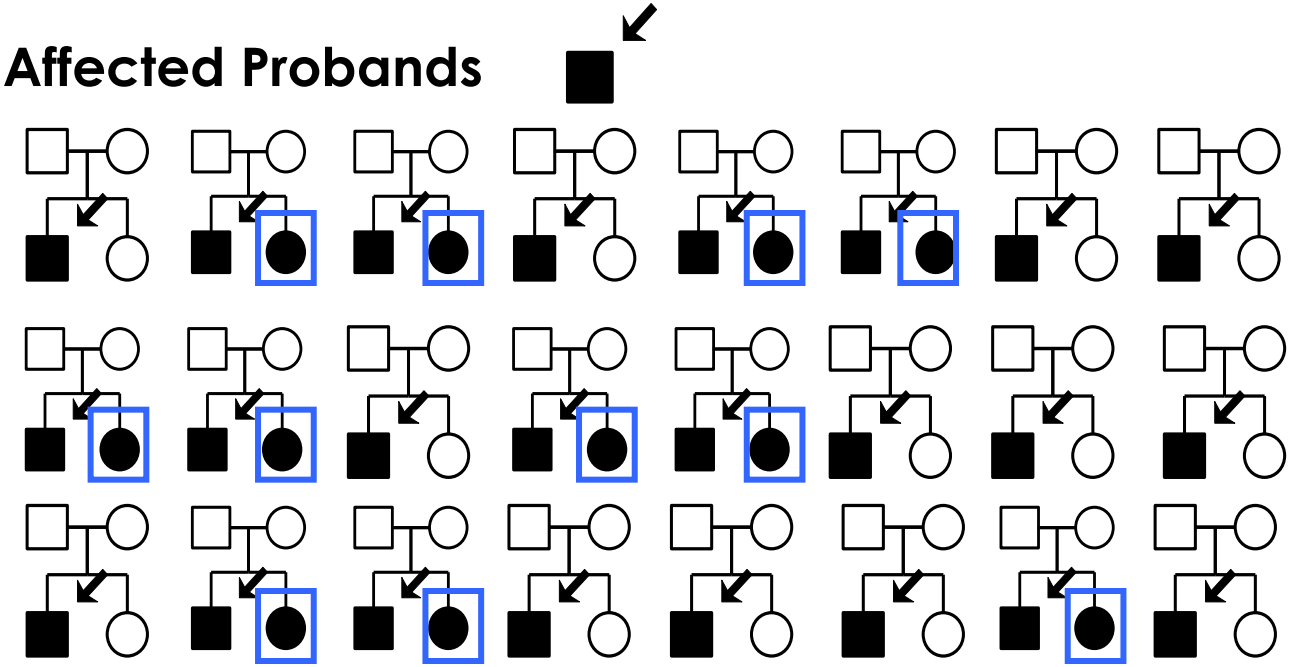
- Unlike Mendelian disorders, there is no clear pattern of inheritance
- Tend to "run" in families
- Few large pedigrees of multiply affected individuals
- Most people have no known family history

What can we learn from genetic epidemiology about genetic architecture?

Evidence for a genetic contribution comes from risks to relatives



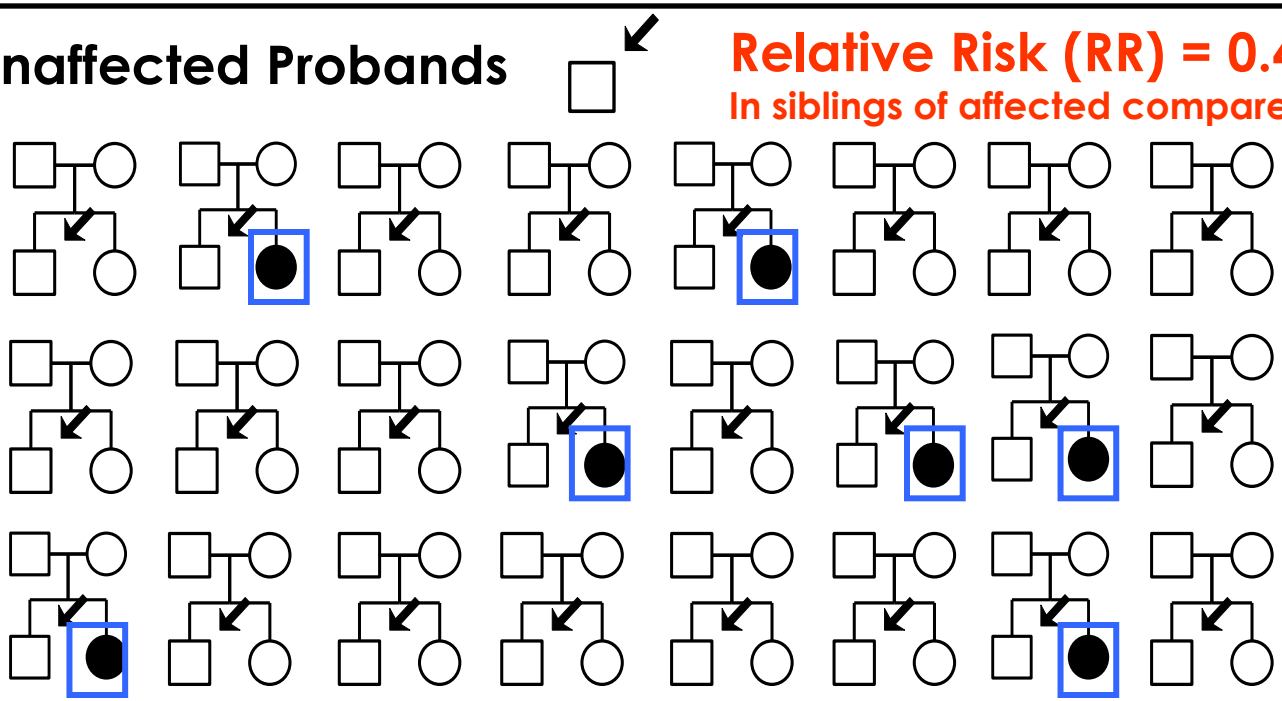
Affected Probands



13/30 are affected;

Risk = 0.433

Unaffected Probands



Relative Risk (RR) = 0.433 / 0.267 = 1.63

In siblings of affected compared to unaffected probands

8/30 are affected;

Risk = 0.267

Slide credit: Dale Nyholt

Relative risk to relatives
Recurrence risk to relatives

How much more likely are you to be diseased if your relative is affected compared to a person selected randomly from the population?

$$\text{Relative risk to relatives } (\lambda_R) = \frac{p(\text{affected} \mid \text{relative affected})}{p(\text{affected in population})} = \frac{K_R}{K}$$

How to estimate p(affected | relative affected) ?

- Collect population samples – cases infrequent
- Collect samples of case families and assess family members

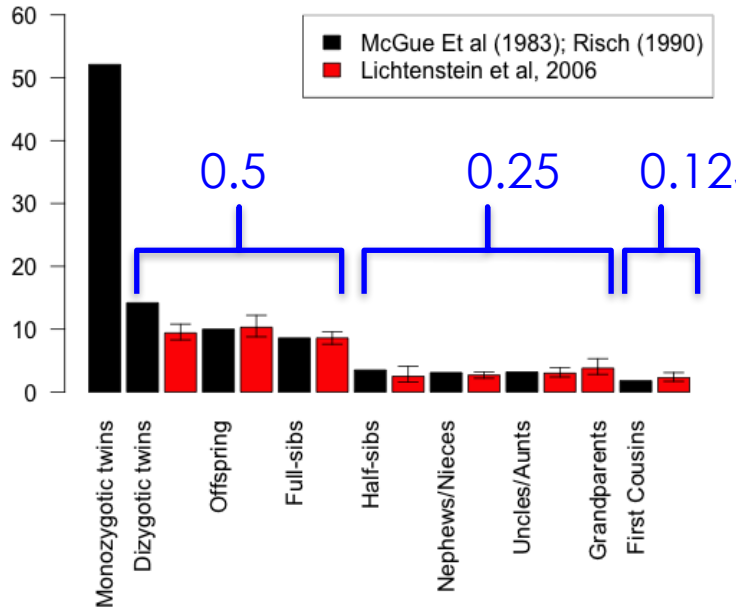
How to estimate p(affected in population) ?

- Census or national health statistics
  - Is definition of affected same in population sample as family sample
- Collect control families and assess family members

$$\text{If disease is not common} \quad \lambda_R = \frac{p(\text{sibling affected} \mid \text{case family})}{p(\text{sibling affected} \mid \text{control family})}$$

Schizophrenia risks to relatives

Relatives	Coefficient of relationship	Risch McGue et al	Lichtenstein et al Estimate	95% CI
Monozygotic twins	1	**52.1**		
Dizygotic twins	½	**14.2**		
Parent	½		**9.4**	**8.3 - 10.8**
Offspring	½	**10.0**	**10.3**	**8.8 - 12.2**
Full-sibs	½	**8.6**	**8.6**	**7.6 - 9.6**
Half-sibs	¼	**3.5**	**2.5**	**1.6 - 4.1**
Nephews/Nieces	¼	**3.1**	**2.7**	**2.2 - 3.2**
Uncles/Aunts	¼	**3.2**	**3.0**	**2.4 - 3.9**
Grandparents	¼		**3.8**	**2.8 - 5.3**
First Cousins	1/8	**1.8**	**2.3**	**1.7 – 3.1**
Offspring of 2 affected parents	½ but ascertained		**89**	**19 - 672**



Baseline risk, K = 0.85% McGue et al
= 0.407% Lichtenstein et al

Risch(1990) Linkage Strategies for Genetically Complex Traits AJHG
McGue et al (1983) Genetic Epidemiology 2: 99
Lichtenstein et al (2006) Recurrence risks for schizophrenia in a Swedish National Cohort.Psychological Medicine

James (1971) relationship between K and K_R

Y = scores of disease yes/no for individuals

Y_R = scores of disease yes/no in relatives of X

K proportion of the population affected

$E(Y) = E(Y_R) = K$

$K_R = E(Y_R \mid Y=1)$

Probability that both X and Y = 1: $E(YY_R) = K*K_R$

$Cov(Y,Y_R) = E(YY_R) – E(Y)*E(Y_R) = K*K_R – K^2$

$= (K_R – K)K = (\lambda_R - 1)K^2 = Cov_R$

This covariance is measurable based on observation, but what underpins this covariance?

James (1971) Frequency in relatives for an all-or-non trait Ann Hum Genet 35 47

Derivation from Risch (1990) Linkage strategies for genetically complex traits. I Multi-locus models. AJHG

Covariance between relatives

Basic quantitative genetics model:

$Y = G + \varepsilon$

$Y = A + D + I + \varepsilon$

$Cov_R = Cov(Y, Y_R) =$

$Cov(G + \varepsilon, G_R + \varepsilon_R) = Cov(G, G_R)$

$= Cov(A + D + I, A_R + D_R + I_R)$

$= Cov(A, A_R) + Cov(D, D_R) + Cov(I, I_R)$

$= a_R V(A) + u_R V(D) + a_R^2 V(AA) + \ldots$

General covariance between relatives

cov_R = covariance between relatives on the disease scale

$$cov_R = a_R V_{Ao} + u_R V_{Do} + a_R^2 V_{AAo} + a_R u_R V_{ADo} + \cdots$$

	V_A	V_D	V_{AA}	V_{AD}	V_{DD}
Offspring-parent	½	0	¼	0	0
Half-sib	¼	0	$^1/_{16}$	0	0
Full-sib	½	¼	¼	$^1/_8$	$^1/_{16}$
MZ twin	1	1	1	1	1
General	a_R	u_R	a_R^2	$a_R\, u_R$	u_R^2

$cov_R = (K_R - K)K = (\lambda_R - 1)K^2$ $\quad V_P = K(1-K)$ (from a few slides back!)

An estimate of narrow sense (additive) heritability on the disease scale is

$$\widehat{h_o^2} = \frac{(\lambda_R - 1)K^2}{a_R K(1-K)} = \frac{(\lambda_R - 1)K}{a_R(1-K)}$$

But covR contains non-additive genetic terms.
We don't know if non-additive genetic effects exist - What to do?

Estimate $\widehat{h_o^2}$ from different types of relatives to see if the estimates are consistent

James (1971) Frequency in relatives for an all-or-non trait Ann Hum Genet 35 47

James (1971) genetic variance on the disease scale

$$\widehat{h_o^2} = \frac{(\lambda_R - 1)K^2}{a_R K(1-K)} = \frac{(\lambda_R - 1)K}{a_R(1-K)}$$

K = 0.0085

λ_{OP}= 10 a_R= ½ $\quad \widehat{h_o^2} = \dfrac{(10-1)0.0085}{\frac{1}{2}(1-0.0085)} = 0.154$

λ_{HS} = 3 a_R= ¼ $\quad \widehat{h_o^2} = 0.069$

λ_{FS} = 8.6 a_R= ½ $\quad \widehat{h_o^2} = 0.130$

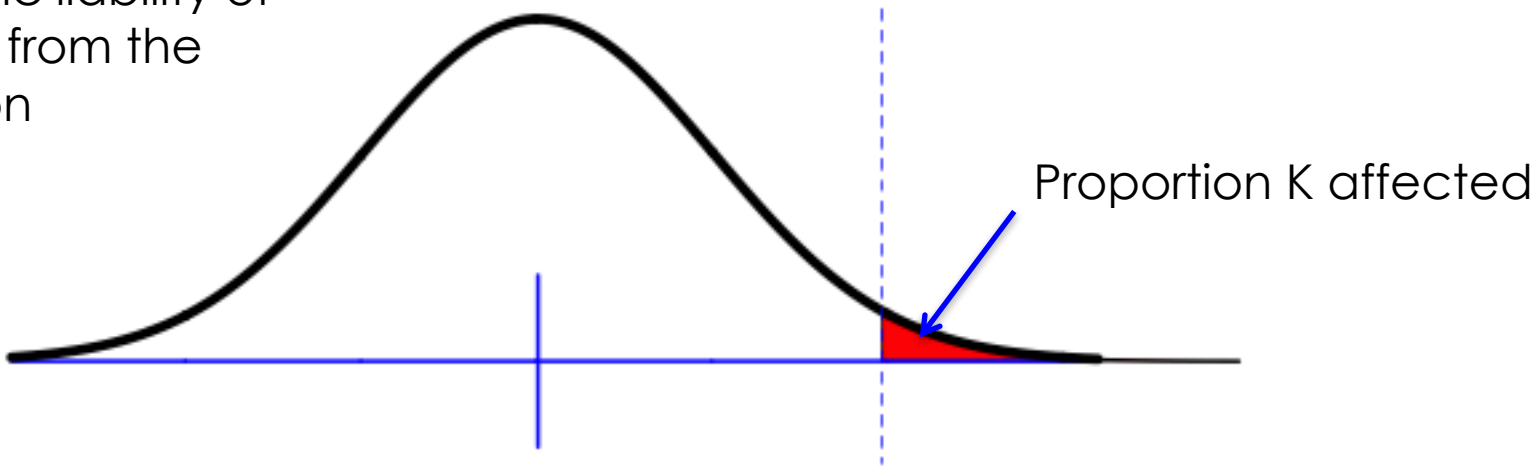
λ_{MZ}= 52 a_R= 1 $\quad \widehat{h_o^2} = 0.438$

The estimates of $\widehat{h_o^2}$ are very different (even if sampling variance is taken into account)

Implies that the estimates of $\widehat{h_o^2}$ are contaminated by non-additive variance on this scale of measurement

Liability threshold model

Phenotypic liability of a sample from the population



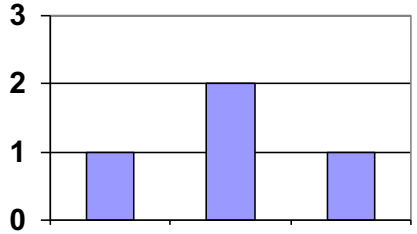
Assumption of normality

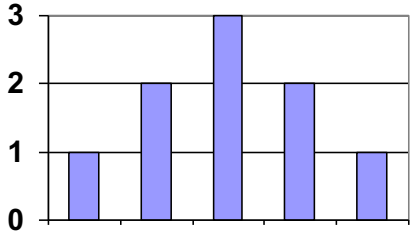
- Only appropriate for multifactorial disease
- i.e. more than a few genes but doesn't have to be highly polygenic
- Key – unimodal

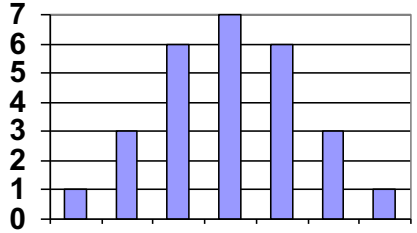
Does an undrlying normality assumption make sense?

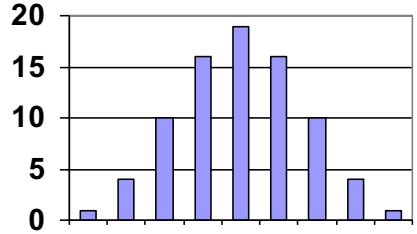
Assumes approximately normal distribution of liability
Makes sense for many genetic variants and environmental/noise factors

1 Locus	2 Locus	3 Locus	4 Locus
→ 3 Genotypes	→ 9 Genotypes	→ 27 Genotypes	→ 81 Genotypes
→ 3 Classes	→ 5 Classes	→ 7 Classes	→ 9 Classes



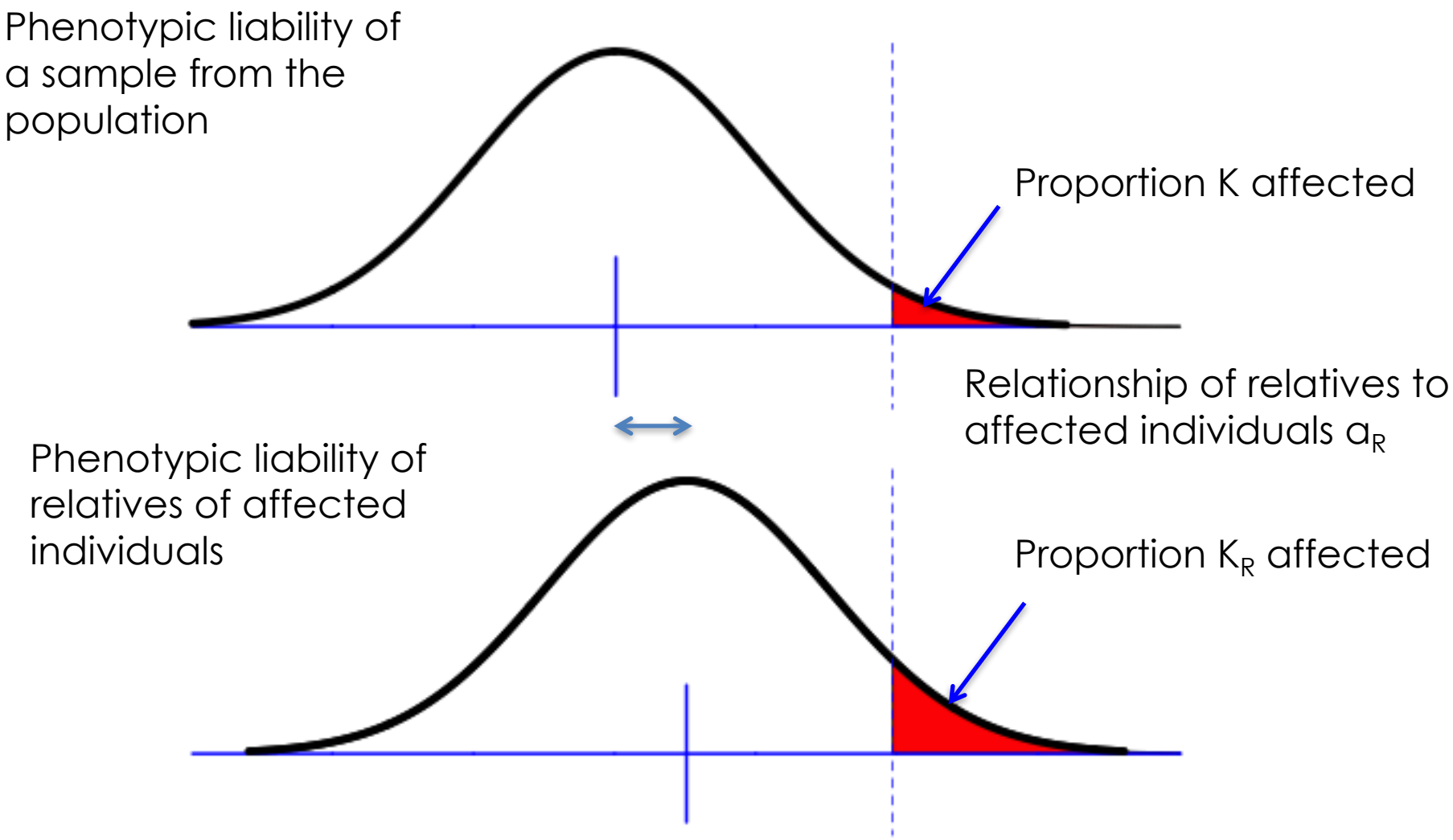






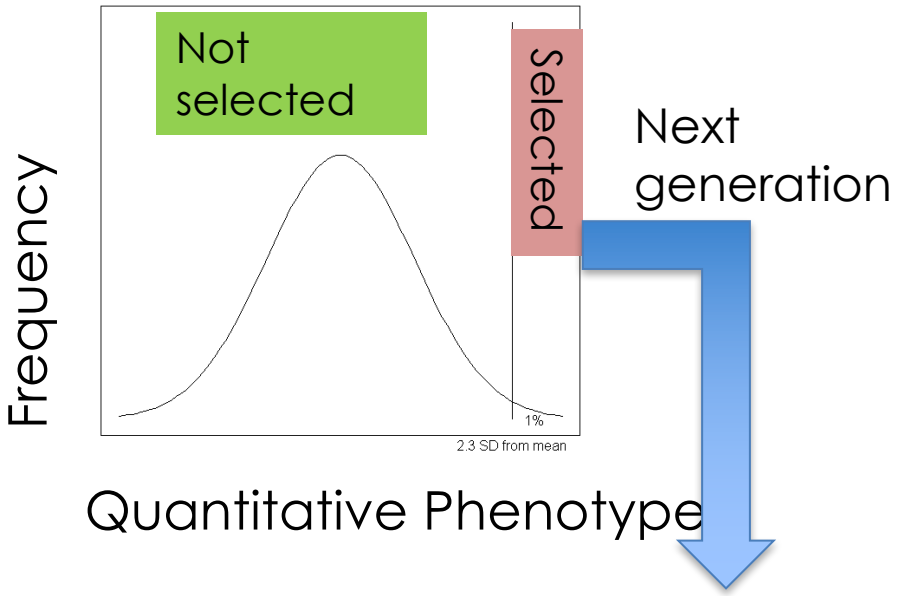
Each Locus has alleles R and r, R = risk alleles.
Each class has a different count of number of risk alleles

Falconer (1965)

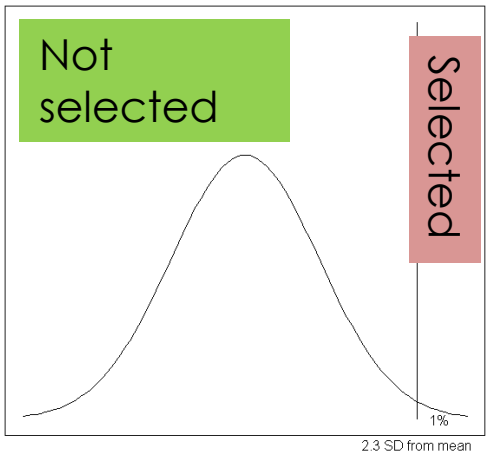


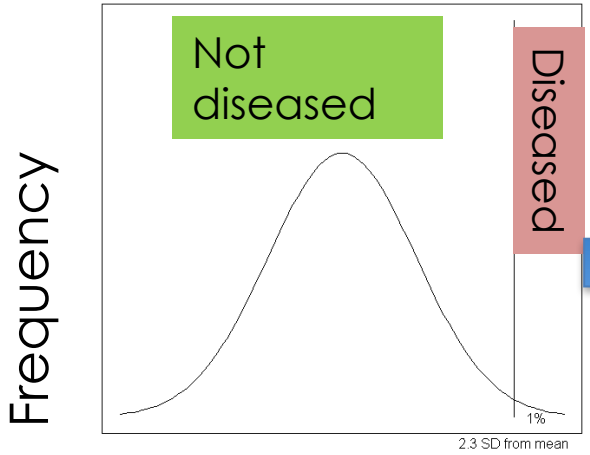
Using normal distribution theory what percentage of the variance in liability is attributale to genetic factors given K, K_R and a_R

Prediction of response to selection and rates of inbreeding under directional selection

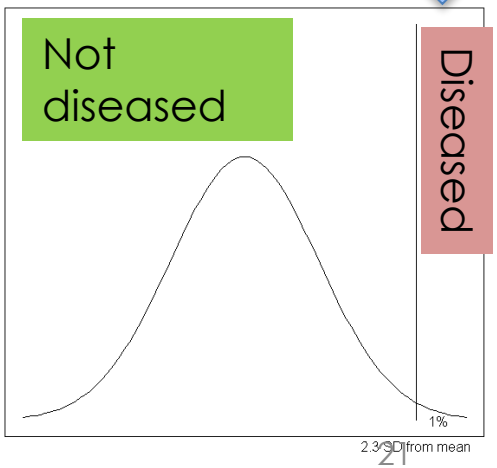


Next generation



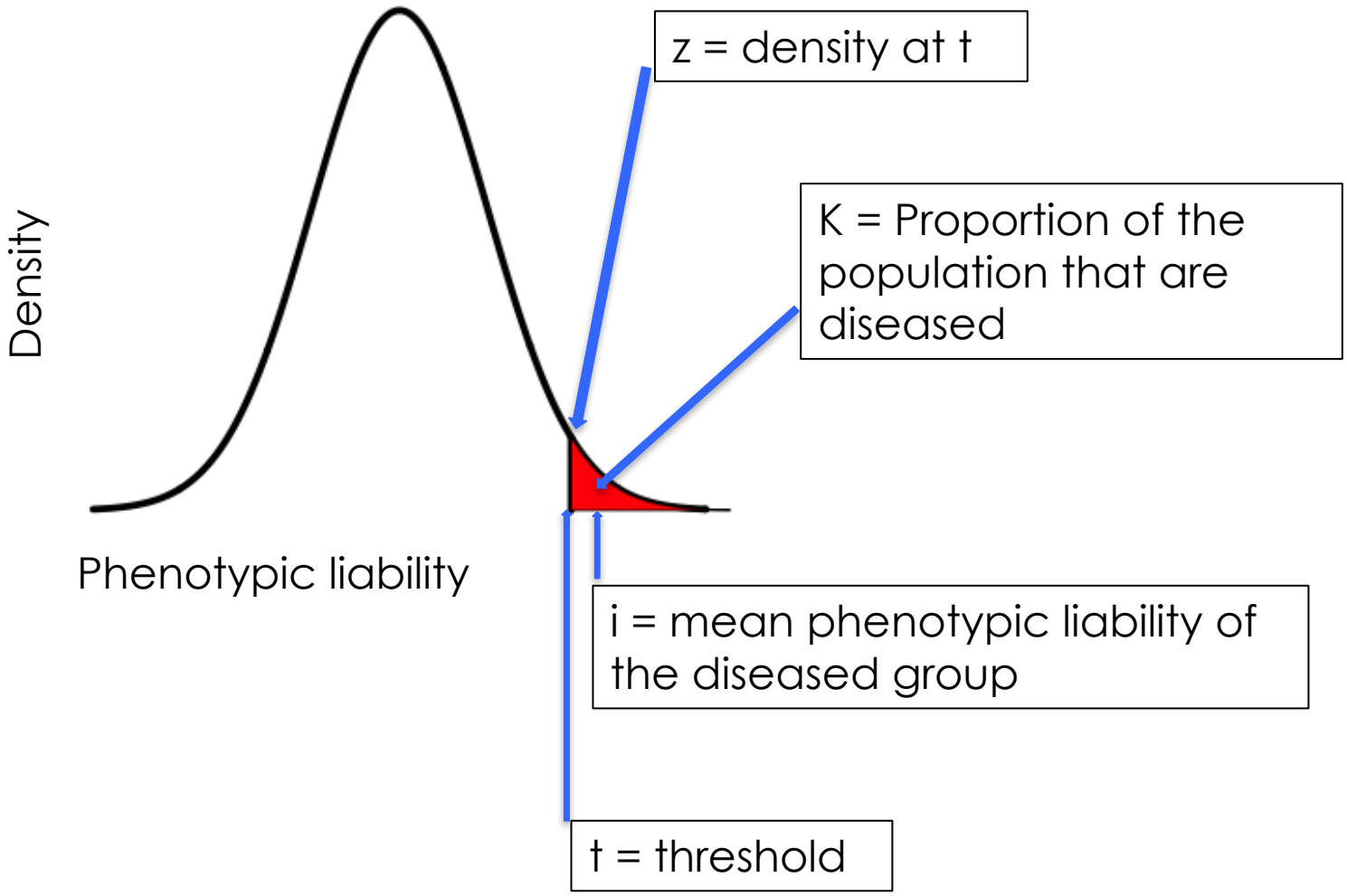


Relatives



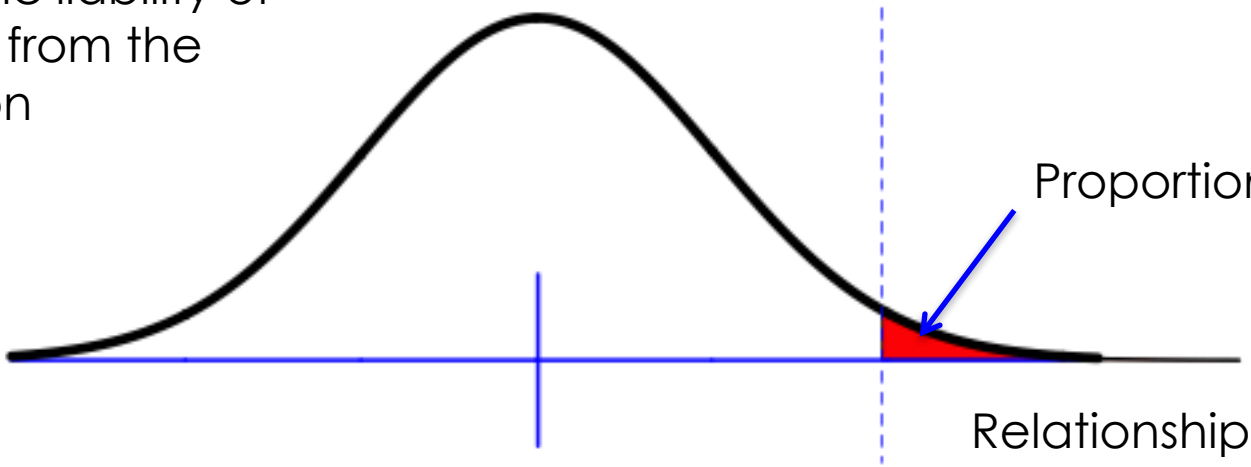
Strong parallels to quantitative genetics of disease

Definitions



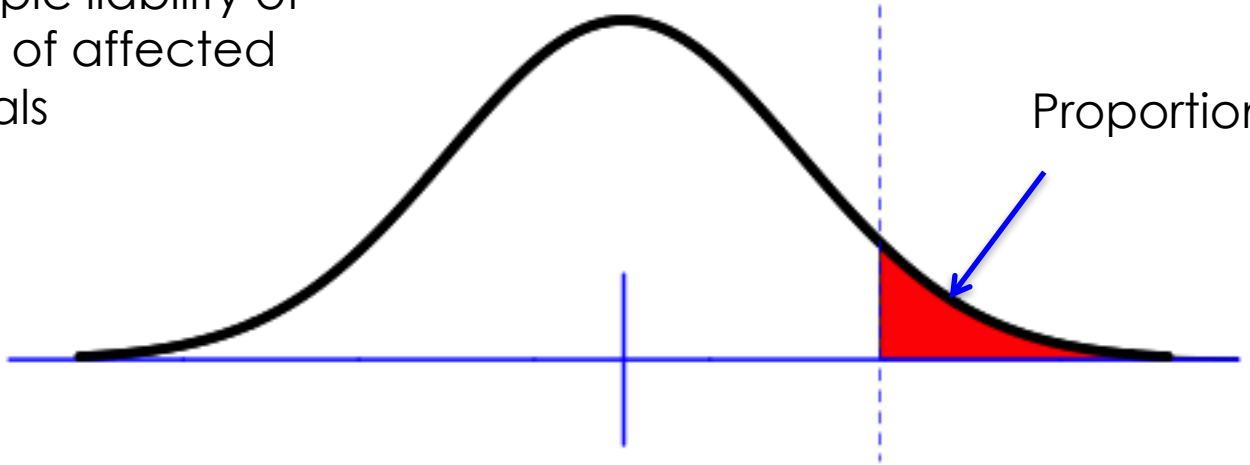
How to get from observed risks to relatives to heritability?- Falconer (1965)

Phenotypic liability of a sample from the population



Relationship of relatives to affected individuals r

Phenotypic liability of relatives of affected individuals



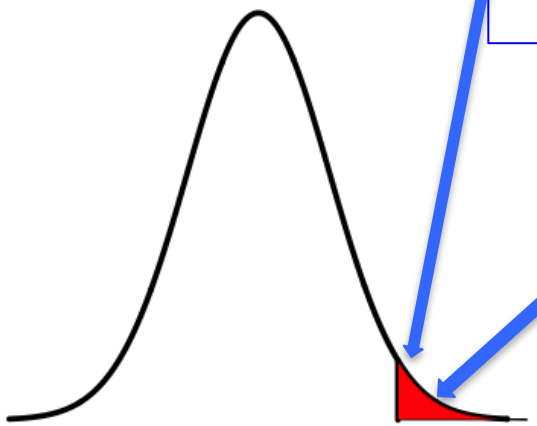
Using normal distribution theory what percentage of the variance in liability is attributale to genetic factors given K, K_R and r

Liability Threshold Model –truncated normal distribution theory

Φ(x) =cumulative density until liability x
standard normal distribution function
φ (x) = probability density at x
Phi

z = density at t

z = φ (t) $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$ = dnorm(t)



K = Proportion of the population that are diseased

K= 1-Φ(t) = 1-pnorm(t)

Standard Deviation =1
$\sigma_p = 1$

i = mean phenotypic liability of the diseased group

i= z/K "selection intensity"

Variance in liability amongst the diseased individuals
= σ_p^2 (1-k), where k = i(i-t)

t = threshold

t= $\Phi^{-1}(1-K)$ = qnorm(1-K)

Inverse standard normal distribution (probit) function

Mean of diseased group

- Pearson & Lee (1908) On the generalized probable error in normal correlation. Biometrika
- Lee (1915) Table of Gaussian tail functions..Biometrika
- Fisher (1941) Properties and application of Hh functions. Introduction to mathematical tables
- Cohen (1949) On estimating the mean and standard deviation of truncated normal distributions Am Stat Association
- Cohen & Woodward (1953)Pearson-Lee-Fisher Functions of singly truncated normal distributions. Biometrics

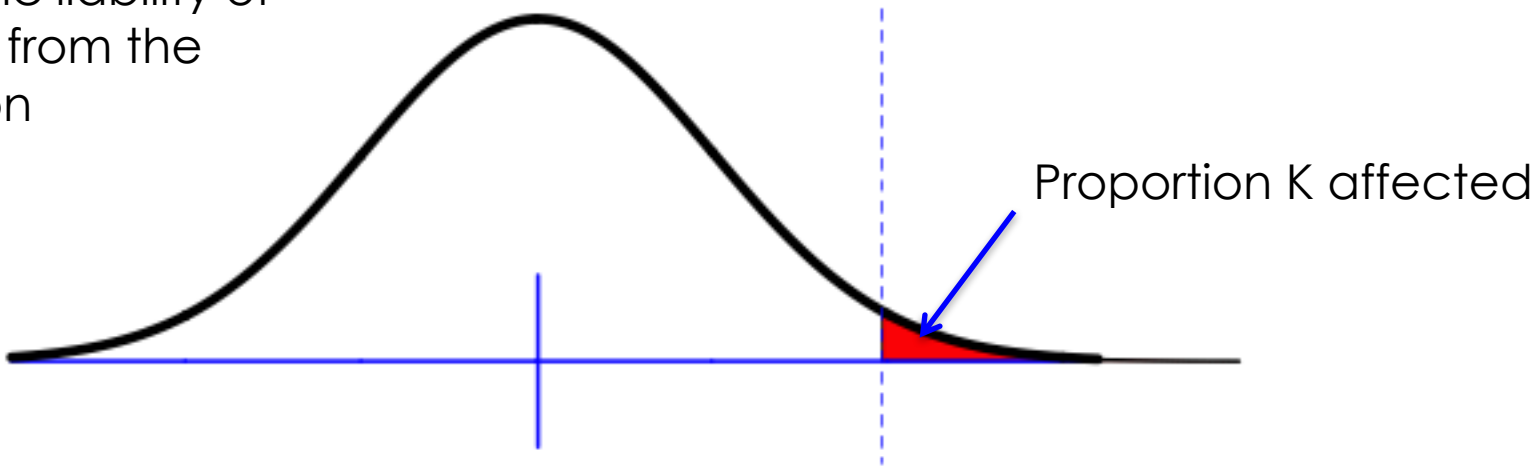
Mean (i): = sum(x * freq of x)
The phenotype frequencies must sum to 1, hence the denominator

$$i = \frac{\int_t^\infty x\phi(x)dx}{\int_t^\infty \phi(x)dx} = \frac{\int_t^\infty x\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx}{K} = \frac{\phi(t)}{K} = \frac{z}{K}$$

Lynch and Walsh equations 2.13 and 2.14; variance equation 2.15

Falconer (1965)

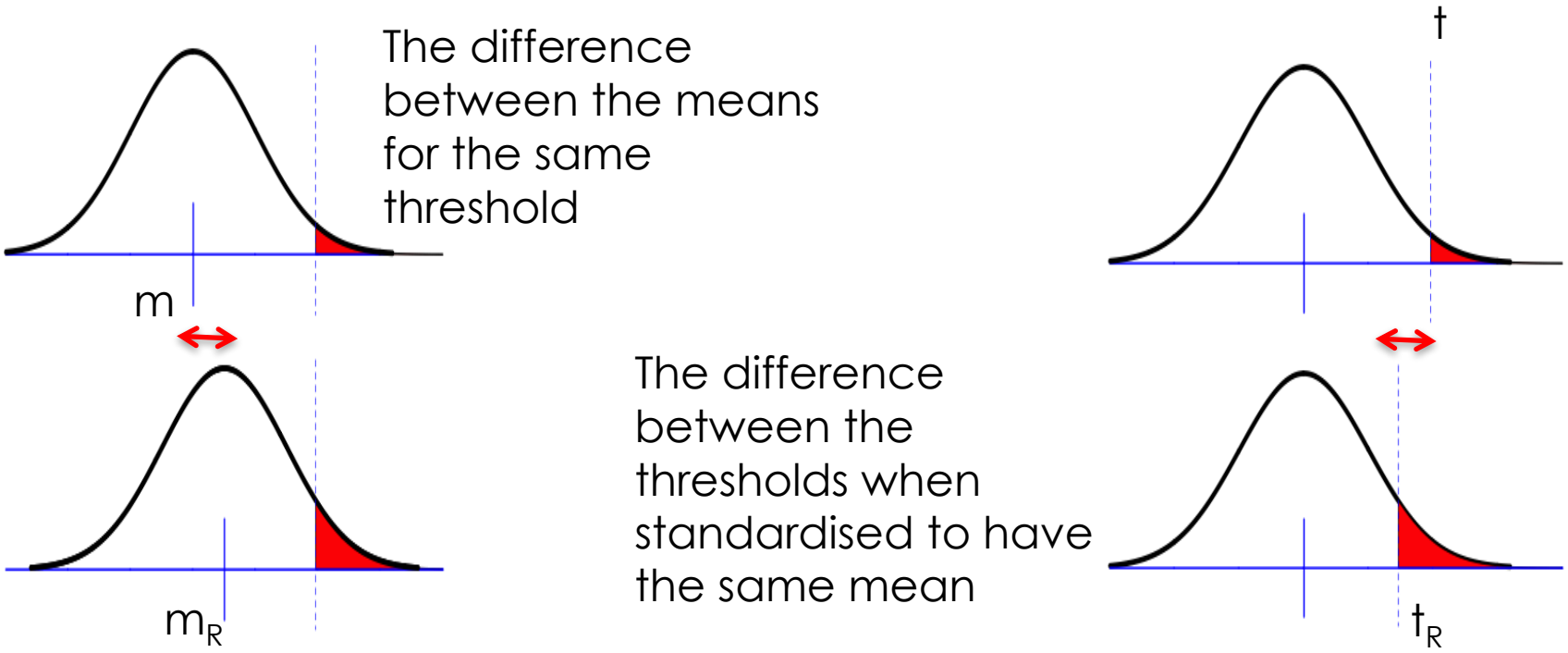
Phenotypic liability of a sample from the population



Assumption of normality

- Only appropriate for multifactorial disease
- i.e. more than a few genes but doesn't have to be highly polygenic
- Key – unimodal

Falconer (1965)



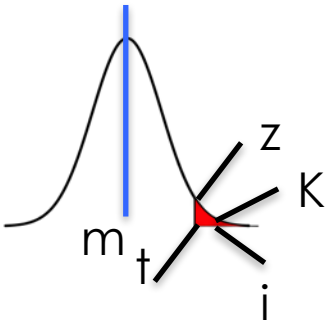
$$m_R - m = t - t_R$$

Given the difference in thresholds, and given known additive genetic relationship between relatives, what proportion of the total variance must be due to genetic factors

Falconer (1965) The inheritance of liability to certain diseases, estimated from incidences in relatives, Ann. Hum Genet. 29 51

Crittenden (1961) an interpretation o familial aggregation based on multiple genetic and environmental factors 27 Ann NY Acad Sci 91 769

Calculate heritability of liability using regression theory



X = phenotypic liability for individuals

Y = phenotypic liability for relatives of X

$E(X) = E(Y) = m = 0$

Relationship between X and Y is linear

$Y = \mu_Y + b_{Y.X}(X-\mu_x)+ \varepsilon$

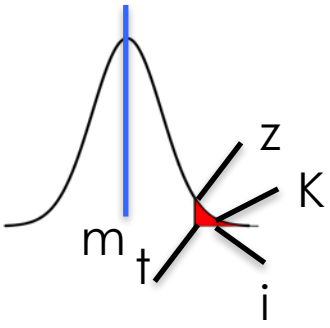
$= m + \frac{cov(A_R,A)}{Var(X)}(X-m) + \varepsilon$, since $m = 0$

$= \frac{a_R\sigma_a^2}{\sigma_p^2} X + \varepsilon = a_R h^2 X + \varepsilon$

Falconer (1965) The inheritance of liability to certain diseases, estimated from incidences in relatives, Ann. Hum Genet. 29 51

Crittenden (1961) an interpretation o familial aggregation based on multiple genetic and environmental factors Ann NY Acad Sci. 91 769

Calculate heritability of liability using regression theory



Y = phenotypic liability for individuals

Y_R = phenotypic liability for relatives of X

$Y_R = a_R h^2 Y + \varepsilon$

For affected individuals Y = i

Expected phenotypic liability of relatives of those affected

$E(Y \mid Y>t) = m_R - m = t - t_R$

Substitute $t - t_R = a_R h^2 i$

Rearrange $h^2 = (t - t_R)/i a_R$

Falconer (1965) The inheritance of liability to certain diseases, estimated from incidences in relatives, Ann. Hum Genet. 29 51

Crittenden (1961) an interpretation o familial aggregation based on multiple genetic and environmental factors Ann NY Acad Sci. 91 769

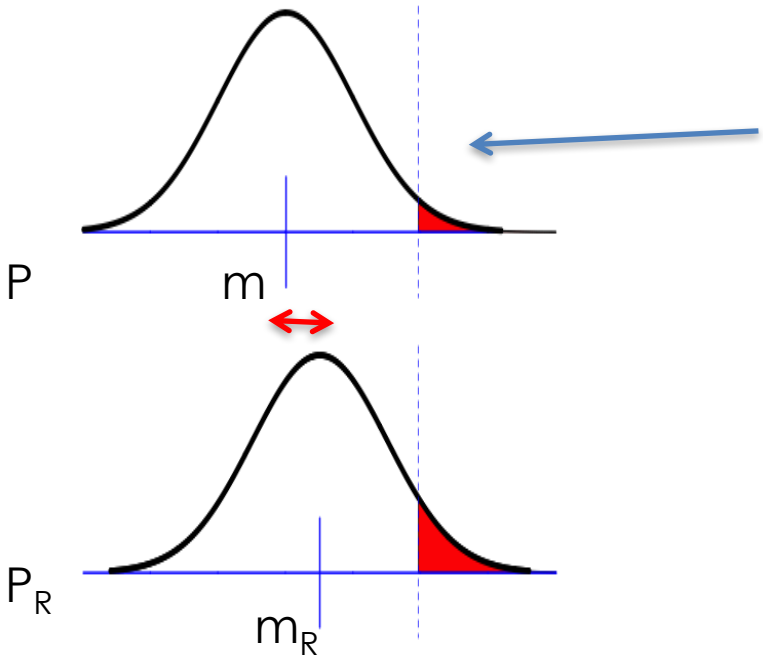
Assumptions made by Falconer (1965)

Assumption: Covariance between relatives reflects only shared additive genetic effects

Check: Use different types of relatives with different a_R and different u_R(dominance coefficient) and different shared environment to see consistency of estimates of h^2

Assumption: Phenotypic variance in relatives is unaffected by ascertainment on affected probands

Accounting for reduction in variance in relatives as a result of ascertainment on affected individuals



Variance in liability amongst the diseased individuals
= σ_p^2 (1-k), where k = i(i-t)

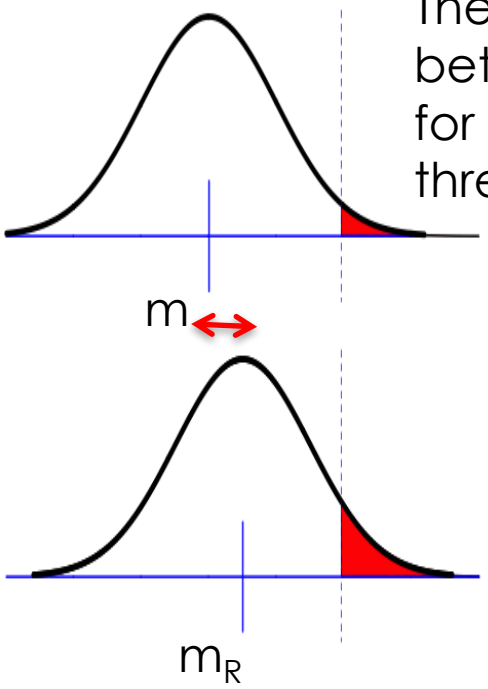
Variance in liability amongst relatives the diseased individuals

$V(P_R \mid P>t) = V(P_R) - kCov(P_R,P)^2$

$= 1 - k(a_R h^2)^2 = 1 - ka_R^2 h^4$

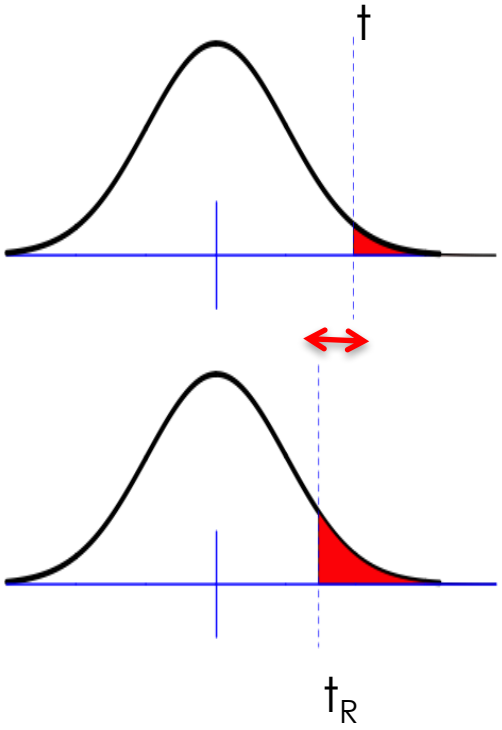
Reich, James, Morris (1972) The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. Ann Hum Gen 36: 163.

Reich et al: heritability of liability



The difference between the means for the same threshold

The difference between the thresholds when standardised to have the mean 0 and variance 1



$$m_R - m = t - t_R \sqrt{1 - ka_R^2 h^4}$$

Reich, James, Morris (1972) The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. Ann Hum Gen 36: 163.

Reich et al: heritability of liability

Y = phenotypic liability for individuals

Y_R = phenotypic liability for relatives of those with Y

$Y_R = a_R h^2 Y + \varepsilon$

For affected individuals Y = i

Expected phenotypic liability of relatives of those affected

$E(Y_R \mid Y>t) = m_R - m = t - t_R\sqrt{1 - ka_R^2 h^4}$

Substitute $t - t_R\sqrt{1 - ka_R^2 h^4} = a_R h^2 i$

Rearrange

$$h^2 = \frac{t - t_R\sqrt{1 - (1 - t/i)(t^2 - t_R^2)}}{a_R(i + (i - t)t_R^2)}$$

Also useful – calculation of t_R when K and h^2 are known

$$t_R = \frac{t - a_R i h^2}{\sqrt{1 - a_R^2 h^4 k}}$$

Practical

Uses simulation to give understanding to the theory.

How to calculate heritability of liability from risks to relatives.

Feel for sample size and sampling variation

Relationship between narrow sense heritability on disease and liability scales

Simulate P = A+ E

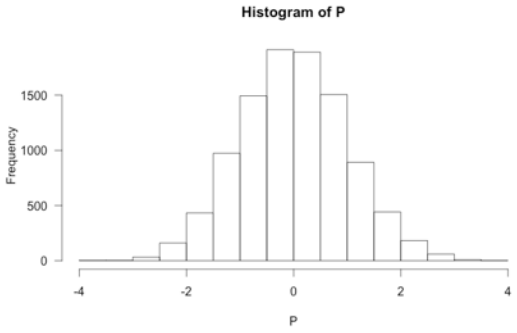
```
> ###################################
> # simulate P = A +  E
> ###################################
> N = 1e4                    #number of people
> h2 = 0.8                   #heritability
> sdA=sqrt(h2)               #genetic standard deviations
> sdE=sqrt(1-h2)             #residual standard deviations
> A = rnorm(N,0,sdA)         #additive genetic values drawn from a normal distribution with mean 0 and std dev sqrt(h2)
> E = rnorm(N,0,sdE)         #"everything else" values drawn from a normal distribution with mean 0 and std dev sqrt(1-h2)
> P = A + E                  #phenotypes
```

```
> dat=data.frame(A,E,P)
> head(dat)
           A           E           P
1 -0.0086812  0.25262196  0.2439408
2 -0.1636576  0.34454330  0.1808857
3  1.4371881 -0.03332416  1.4038639
4  0.6690834 -1.22955515 -0.5604718
5  0.2508860  0.14259020  0.3934762
6 -1.8710844  0.94433320 -0.9267512
```

```
> mean(A)
[1] -0.004441688
> mean(E)
[1] 0.008263525
> mean(P)
[1] 0.003821838
```

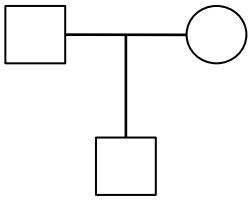
```
> var(A)
[1] 0.8059709
> var(E)
[1] 0.2002046
> var(P)
[1] 0.9998455
```

```
> var(dat)
           A           E         P
A  0.8059709 -0.00316500 0.8028059
E -0.0031650  0.20020460 0.1970396
P  0.8028059  0.19703959 0.9998455
```



If we only measure P how do we estimate heritability?

Need relatives



$$P = A + E$$

$$P_{dad} = A_{dad} + E_{dad}$$
$$P_{mum} = A_{mum} + E_{mum}$$

$$P_{child} = A_{child} + E_{child}$$

$$A_{child} = 0.5*A_{mum} + 0.5*A_{dad} + A_w$$

genetic segregation
unique to the child

What is the variance of A_w?

$$Var(A_{child}) = 0.25*Var(A_{mum}) + 0.25*Var(A_{dad}) + Var(A_w)$$

$$Var(A) = 0.25*Var(A) + 0.25*Var(A) + Var(A_w)$$

$$Var(A_w) = 0.5*Var(A)$$

Half of the genetic variance in a population is within family variance

Segregation Variation

2 parents

$(2^{23})^2 = 7 \times 10^{13}$

23 pairs of chromosomes

Choose 1 from the pair

Ignoring recombination which will make the # combinations even bigger

Half the genetic variation in a population is generated by the sampling of genetic material within families

Simulate families

```
> ####################################
> #Simulate parents and a child under a polygenic model
> ####################################
> N = 1e5                    #number of families
> h2 = 0.8                   #heritability
> sdA = sqrt(h2)               #genetic standard deviations
> sdE = sqrt(1-h2)             #residual standard deviations
> sdW = sqrt(0.5*h2)
>
> # Dads
> A_dad = rnorm(N,0,sdA)  #fathers' additive genetic values
> E_dad = rnorm(N,0,sdE)   #fathers' residual values
> P_dad = A_dad + E_dad       #fathers' phenotypic liability values
>
> # Mums
> A_mum = rnorm(N,0,sdA)    #mothers' additive genetic values
> E_mum = rnorm(N,0,sdE)    #mothers' residual values
> P_mum = A_mum + E_mum     #mothers' phenotypic liability values
>
> # Children
> A_child  =  0.5*A_dad + 0.5*A_mum  + rnorm(N,0,sdW) #childrens' genetic liability is the mid-parent value plus a withinfamily deviation
> E_child  = rnorm(N, mean=0, sdE) # childrens' residual values
> P_child  = A_child + E_child       # childrens' phenotypic liability values
>
```

```
> datA = data.frame(A_dad, A_mum, A_child)
> head(datA)
      A_dad      A_mum    A_child
1 0.2816500  0.8070750  0.7774883
2 0.9239563  0.6403608  0.2447249
3 1.2305561  1.3768674  0.6054434
4 0.7557954  1.9800166  1.8447414
5 1.0707037  1.7275696  0.8774031
6 1.1953478 -1.3135453 -1.0321547
```

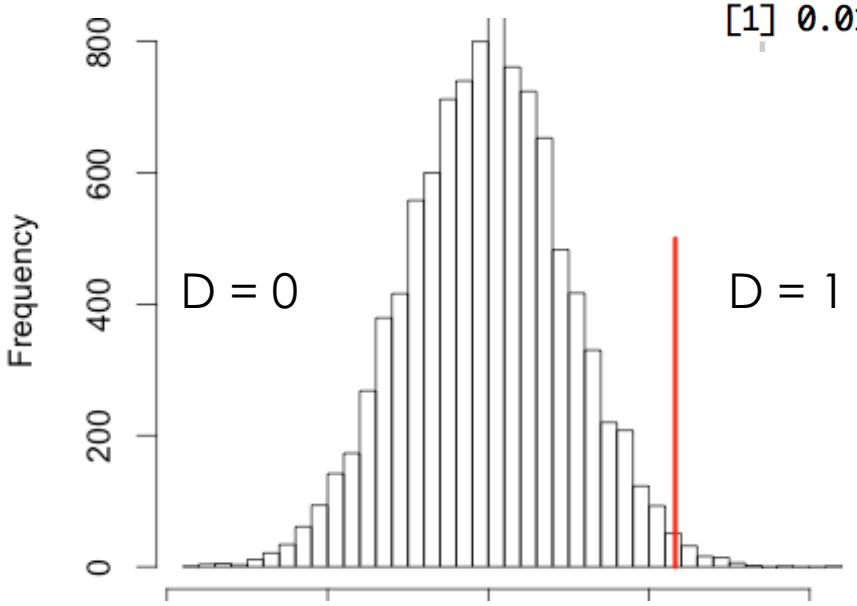
```
> var(datA)
              A_dad       A_mum   A_child
A_dad   0.796370183 0.000156606 0.3961720
A_mum   0.000156606 0.798393999 0.4010472
A_child 0.396171970 0.401047172 0.7961398
```

```
> datP = data.frame(P_dad,P_mum,P_child)
> var(datP)
               P_dad         P_mum   P_child
P_dad    0.9975084811 -0.0006500908 0.3987118
P_mum   -0.0006500908  0.9993500833 0.4009034
P_child  0.3987118131  0.4009033881 1.0009886
>
```

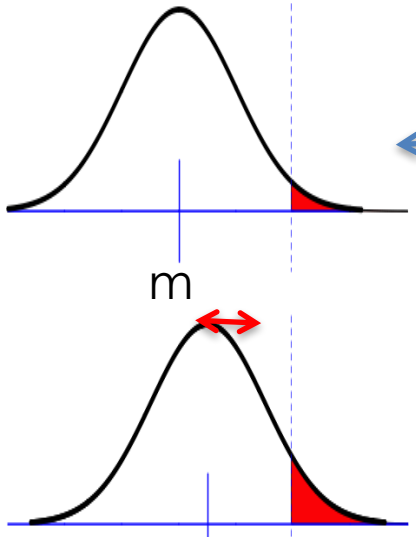
Simulation, phenotype is now liability

```
> ################################################
> # Make a disease trait
> ################################################
> K = 0.01              # lifetime prevalence of disease
> t=qnorm(1-K,0,1)      # liability threshold from normal distribution theory
>
> D = c(rep(0,N))       # first set all individuals to be non-diseased = 0
> D[P>t] = 1            # those with liability higher than the threshold are diseased = 1
```

```
> sum(D)/N    # observed value of K, lifetime disease risk
[1] 0.0102
```



Accounting for reduction in variance in relatives as a result of ascertainment on affected individuals



Variance in liability amongst the diseased individuals
= ((1-i(i-t)) = (1-k)

Variance in liability amongst relatives the diseased individuals = 1- $i(i-t)(a_R h^2)^2$

$$t_R = \frac{t - a_R i h^2}{\sqrt{1 - a_R^2 h^4 k}}$$

$$h^2 = \frac{t - t_R\sqrt{1 - (1 - t/i)(t^2 - t_R^2)}}{a_R(i + (i - t)t_R^2)}$$

```
> h2l=function(t,tR,i,aR){(t-tR*sqrt(1-(1-t/i)*(t^2-tR^2)))/(aR*(i+(i-t)*tR^2))} # heritability of liability with Reich et al correct
ion **use this one
> (h2l_est=h2l(t_est,t_dad,i_est,0.5))
[1] 0.7857835
> (h2l_est=h2l(t_est,t_MZ,i_est,1))
[1] 0.7985478
```

Reich, James, Morris (1972) The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. Ann Hum Gen 36: 163.

Practical

1. **Polygenic models generate a normal distribution of genetic values.**

a) Simulate a population of N=10,000 for 10 loci of frequency p

- Binomial distribution of genotypes
- G1, G2..G10=rbinom(N,2,p), set p =0.5
- Make a count of risk alleles across 1,2,..10 loci
- R1=G1, R2=G1+G2, …R10 = G1+G2…+G10
- Plot histogram of R1…R10

b) repeat for allele freq p = 0.1

c) set p randomly eg uniform c(runif(10,0,1))

d) a-c demonstrate normal distribution of risk allele count.
If the effect size for the risk locus at SNP i is a_i then what is the distribution of variance of risk allele. Draw the ai from different distributions.
Skip this come back if there is time

2. Using simulation to explore the liability threshold model.

Section 2a-2e. Already programmed.

2a. Run the section – generates sliders (make plot window as big as possible) – Not so important

2b-2e Run line by line

2b. Simulates phenotypic liability and disease status of parents and children

2c. Some graphs and calculates risks to relatives

2d. Compare simulated values with normal distribution theory

2e. Estimate heritability from recurrence risks to relatives

2f. Complete table to feel sampling variation

Regression of offspring quantitative phenotype on mid parent value.

$$\frac{Cov(Y_O,(Y_M+Y_D)/2)}{Var((Y_M+Y_D)/2)} = \frac{2*0.5*V(A)/2}{2*V(P)/4} = \frac{V(A)}{V(P)} = h^2$$

2g. Extend the simulation to include different types of relatives

##########################

Add to the simulation a Monozygotic twin of the child

Add to the simulation a full-sibling of the child

Add to the simulation a paternal half-sibling of the child

Calculate lambdaMZ, lambdaFS, and lambdaHS

Estimate heritability of liability from lambdaMZ, lambdaFS, and lambdaHS