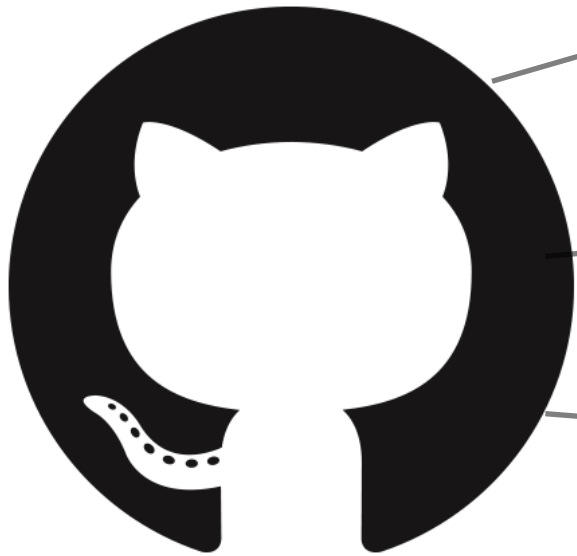# Leverage proxy-cases in your data!

Beta version available!
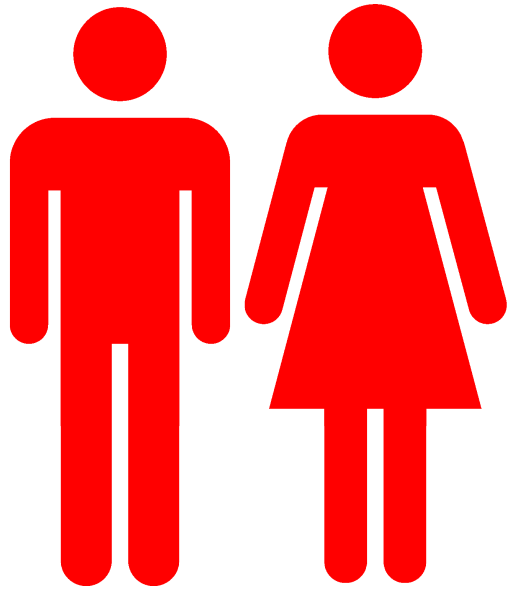
Identify unaffected relatives of cases in your data

Peform meta-analysis using GWAX and GWAS with shared controls

Create phenotype files for BOLT-LMM for multiple models

https://github.com/bnwolford/proxyPower
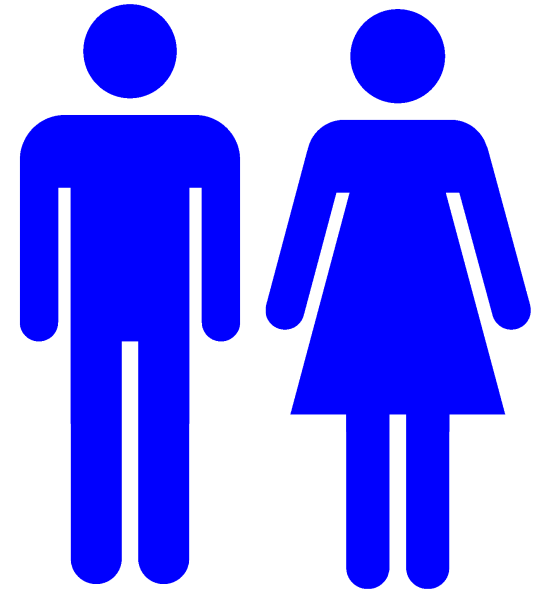
Proxy-cases are unaffected first degree relatives of cases

Affected

Unaffected first degree relative of affected

Unaffected

Case
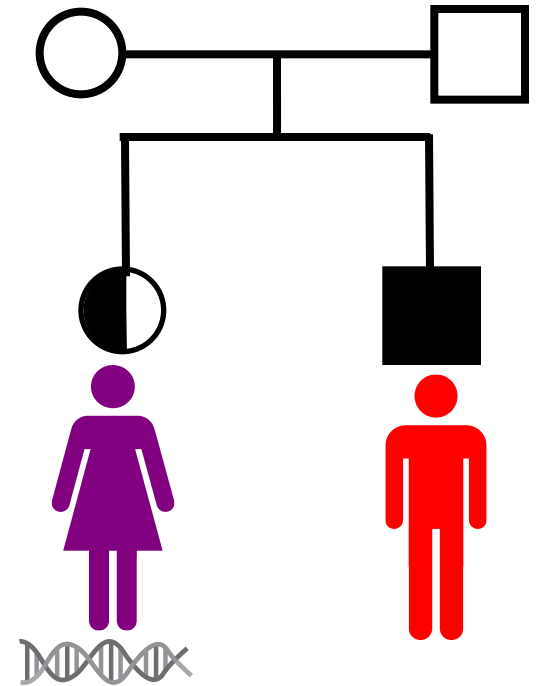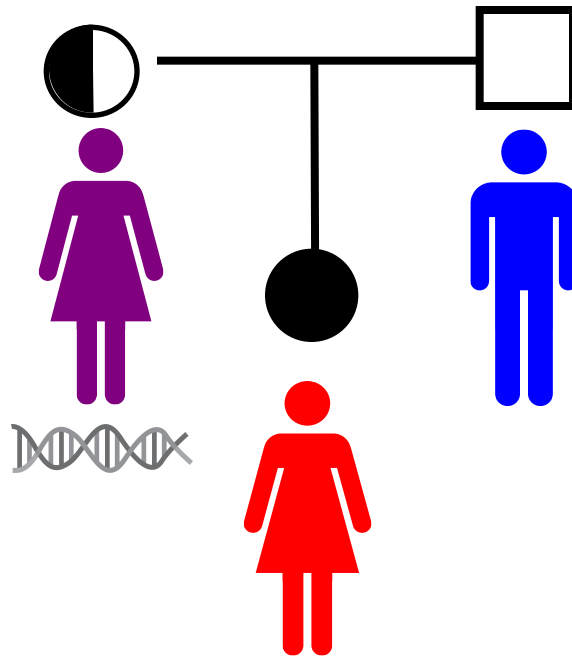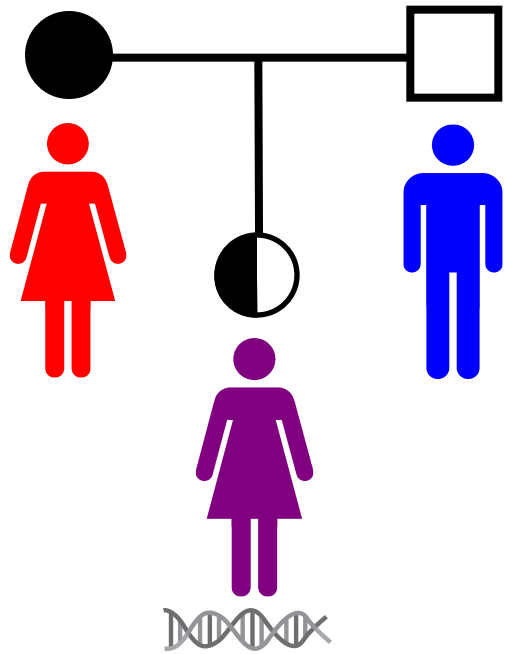
Proxy-case

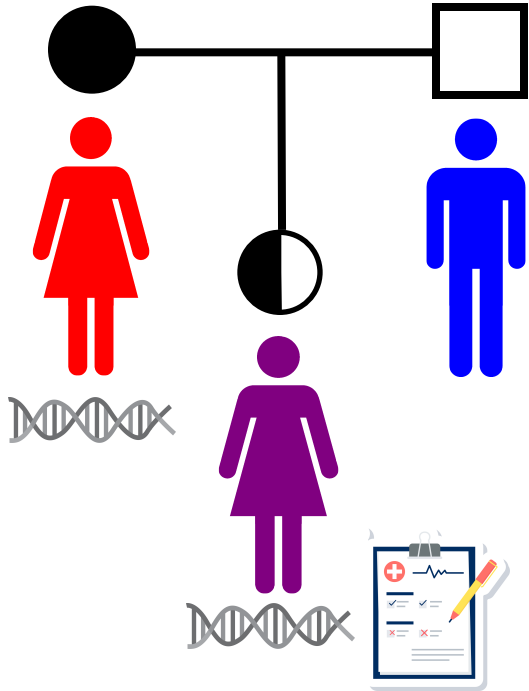Control

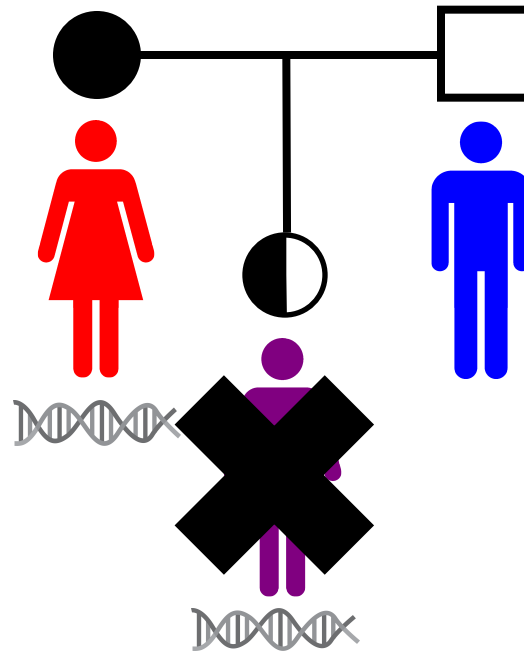Proxy-case has some combination of affected mother, father, child, sibling depending on epidemiological survey wording

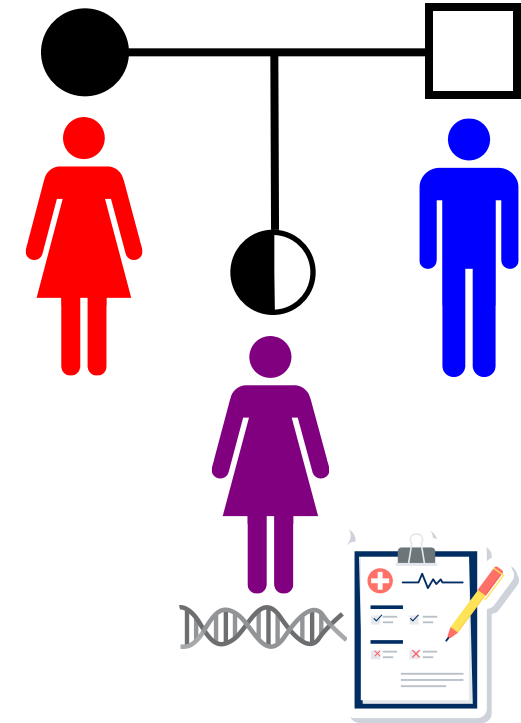# Proxy-case definition

*double counting genetic info of cases

Self report only:
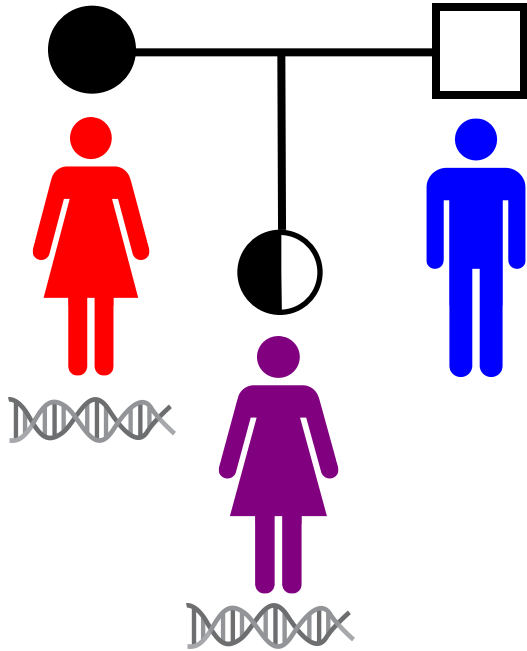Identify proxy-cases who report an affected relative on a epidemiological survey*

Self report minus kinship:
Exclude proxy-cases whose affected relative is in dataset, use only proxy-cases whose affected relative is not in dataset
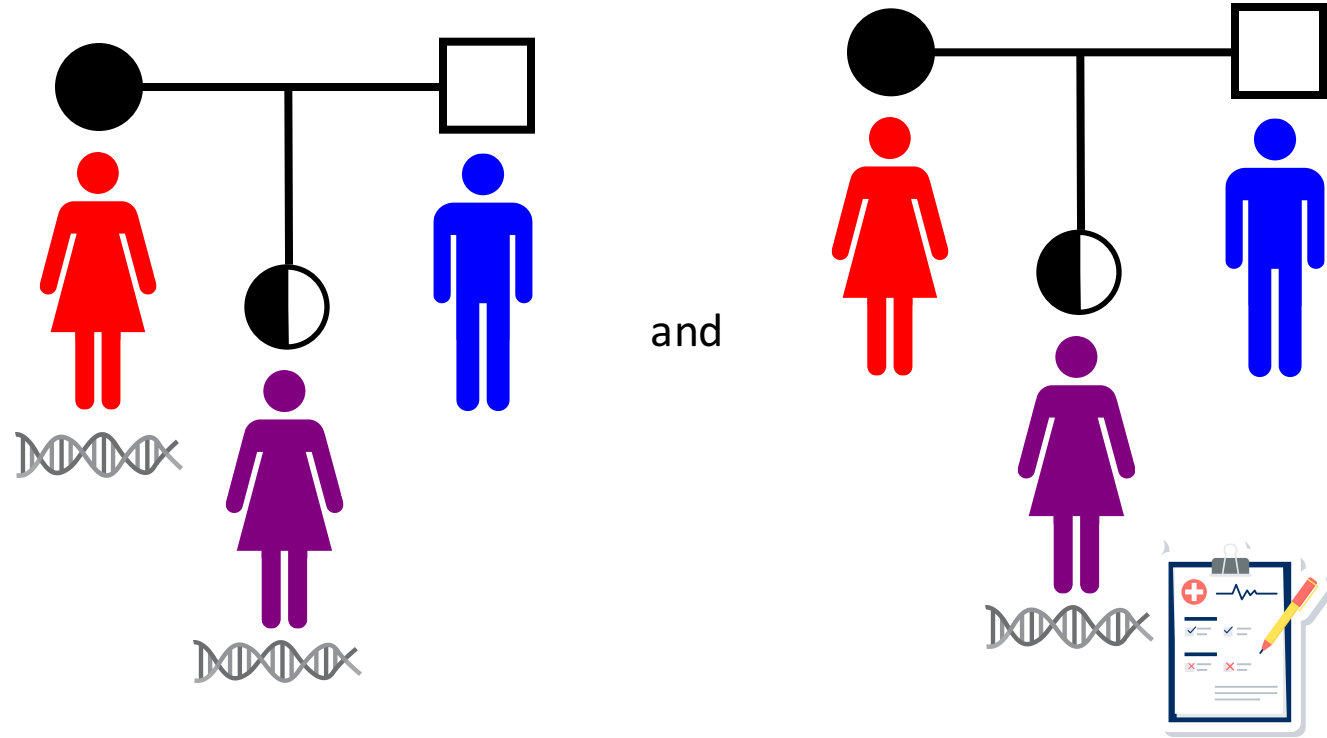
Proxy-case definition

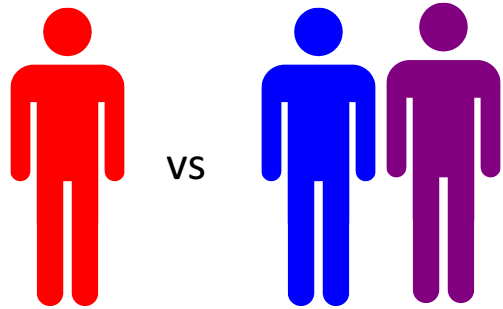*double counting genetic info of cases

Kinship only:
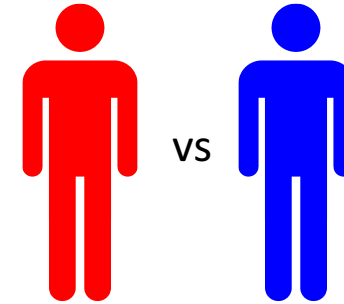Identify proxy-cases whose affected relative is in the dataset using kinship matrix

and

Self report plus kinship:
Identify proxy-cases whose affected relative is in dataset even if the sample did not report having an affected relative in the epidemiological survey*
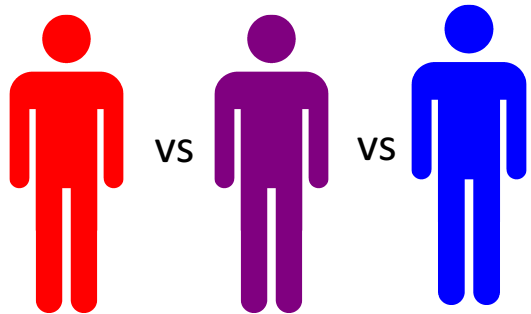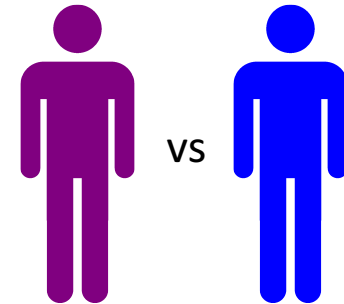
Study designs in cohort-based GWAS

Study designs in cohort-based GWAS

Group cases and proxy-cases

Your cohort

Control

Case

Self reported
relative of a case

Relative of case
in cohort via
kinship matrix

d **c** **b**

e

a
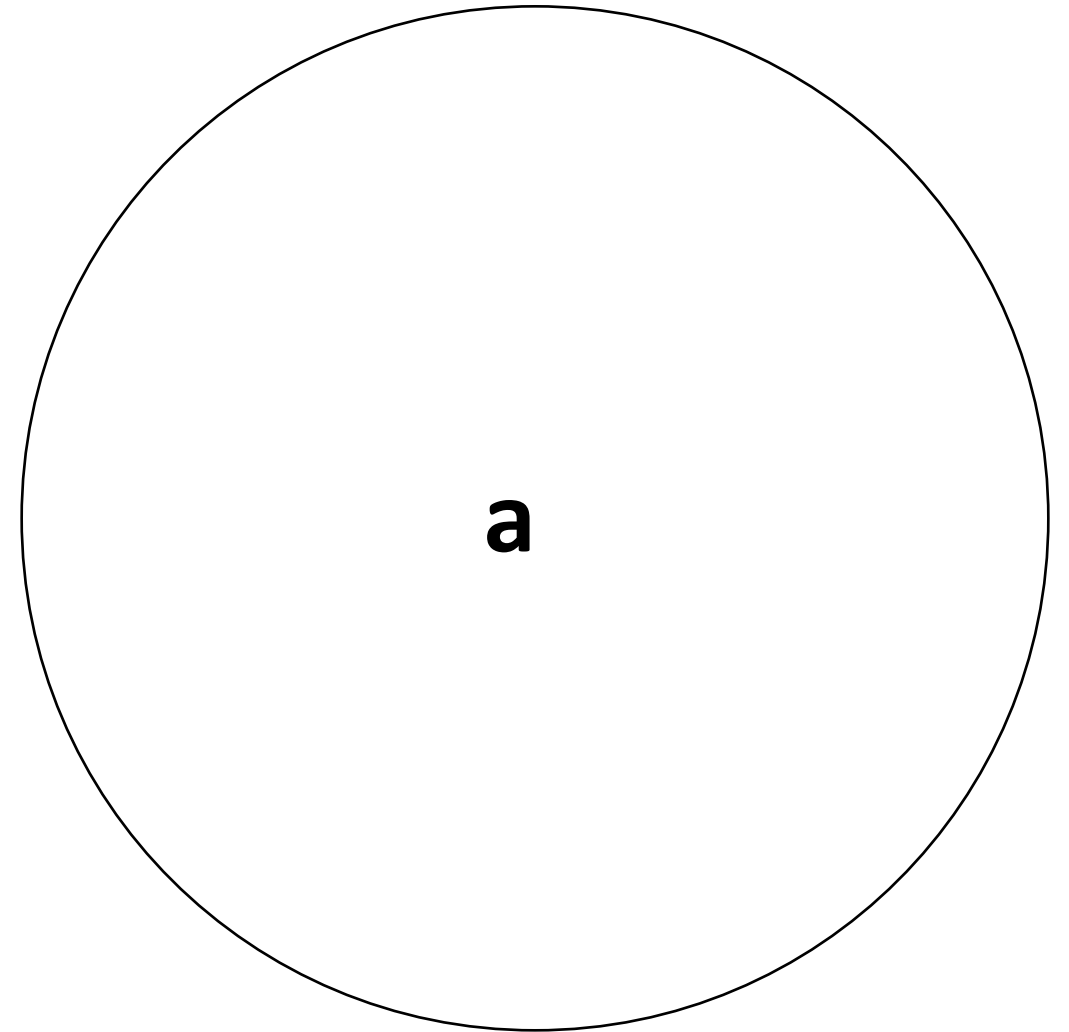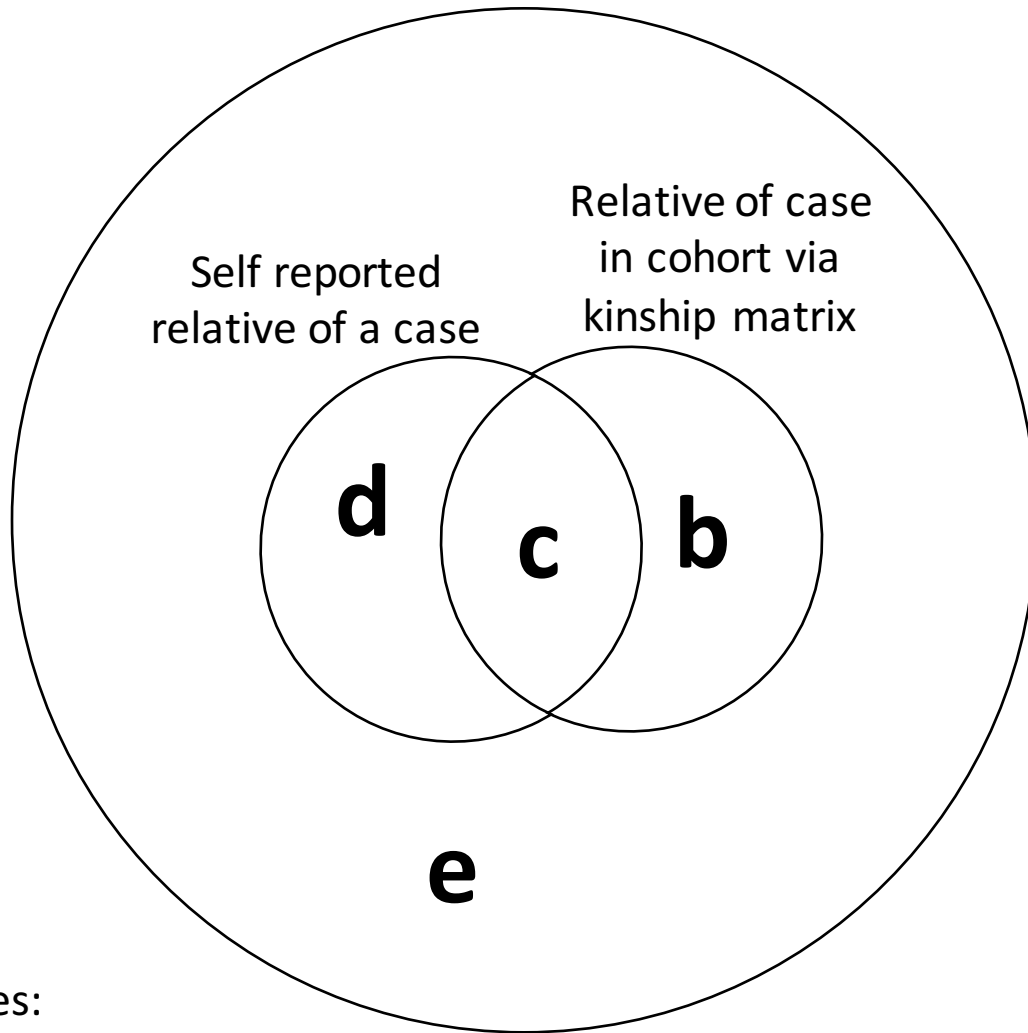
Notes:
Using first degree relative information, ignoring second degree relatives
Non conservative control
Group c is self reported proxy-cases who are relatives of cases in the cohort

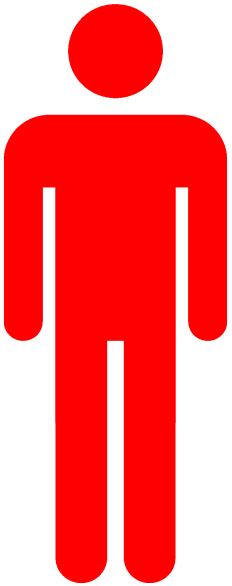|		Model			
Proxy-case definition		Proxy-case	GWAS (1)	Exclude proxy-cases (2)	GWAX (3)	Model proxy-cases (4)
Self report only (SR)	cd	a vs bcde	a vs be	cd vs be	a vs cd vs be
Self report plus kinship (SPK)	bcd	a vs bcde	a vs e	bcd vs e	a vs bcd vs e
Self report minus kinship (SMK)	d	a vs bcde	a vs e	d vs e	a vs d vs e
Kinship only (K)	bc	a vs bcde	a vs de	bc vs cd	a vs bc vs de

Notes:

Group a is NA for Model 3

Proxy-case group is NA for Model 2

For SMK definition, group bc is NA for Models 2, 3, 4

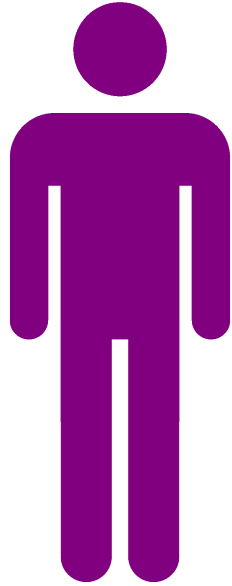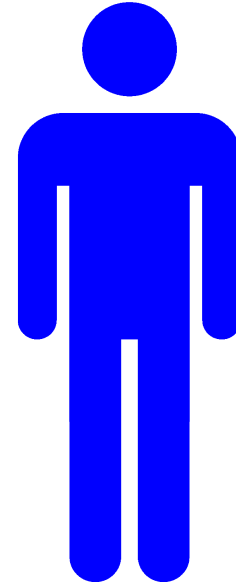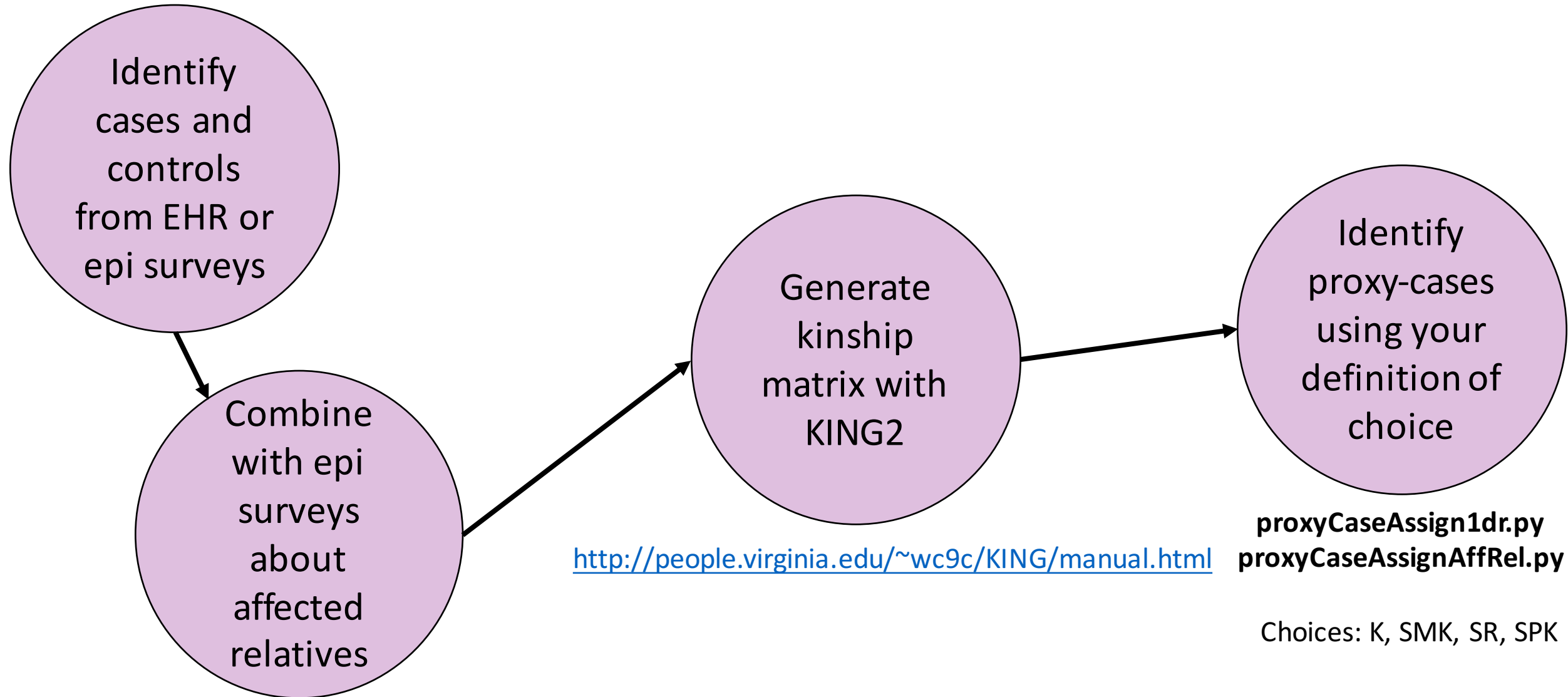# Modeling coefficient of relationship to a case

$F = 1$ vs $F = 0.5$ vs $F = 0$
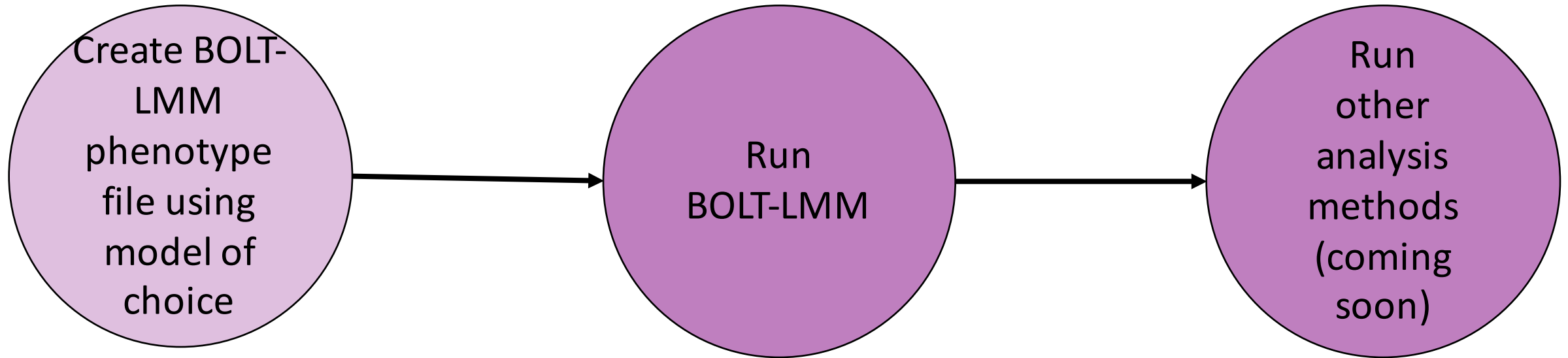
proxyPower workflow

Identify cases and controls from EHR or epi surveys

Combine with epi surveys about affected relatives

Generate kinship matrix with KING2

Identify proxy-cases using your definition of choice

http://people.virginia.edu/~wc9c/KING/manual.html

proxyCaseAssign1dr.py
proxyCaseAssignAffRel.py

Choices: K, SMK, SR, SPK

proxyPower workflow

proxyModel.py

https://data.broadinstitute.org/alkesgroup/BOLT-LMM/

Choices: 1, 2, 3, 4, 5