Pros and Cons of Minimal Phenotyping in Psychiatric Genetics

Patrick Sullivan, MD FRANZCP

UNC Chapel Hill

PGC Lead-PI

PGC Worldwide Lab Call Details

DATE: Friday, December 14, 2012
PRESENTER: Patrick F Sullivan
TITLE: "Pros and Cons of Minimal Phenotyping in Psych Genetics."
START: We will begin promptly on the hour.
1000 EST - US East Coast
0700 PST - US West Coast
1500 GMT - UK
1600 CET - Central Europe
0200 EDT – Australia (Sat, Dec 15th, 2012)
DURATION: 1 hour

TELEPHONE:
- US Toll free: 1 866 515.2912
- International direct: +1 617 399.5126
- Toll-free number? See http://www.btconferencing.com/globalaccess/?bid=75_public
- Operators will be on standby to assist with technical issues. "*0" will get you assistance.
- This conference line can handle up to 300 participants.

PASSCODE: 696 058 27

Lines are Muted

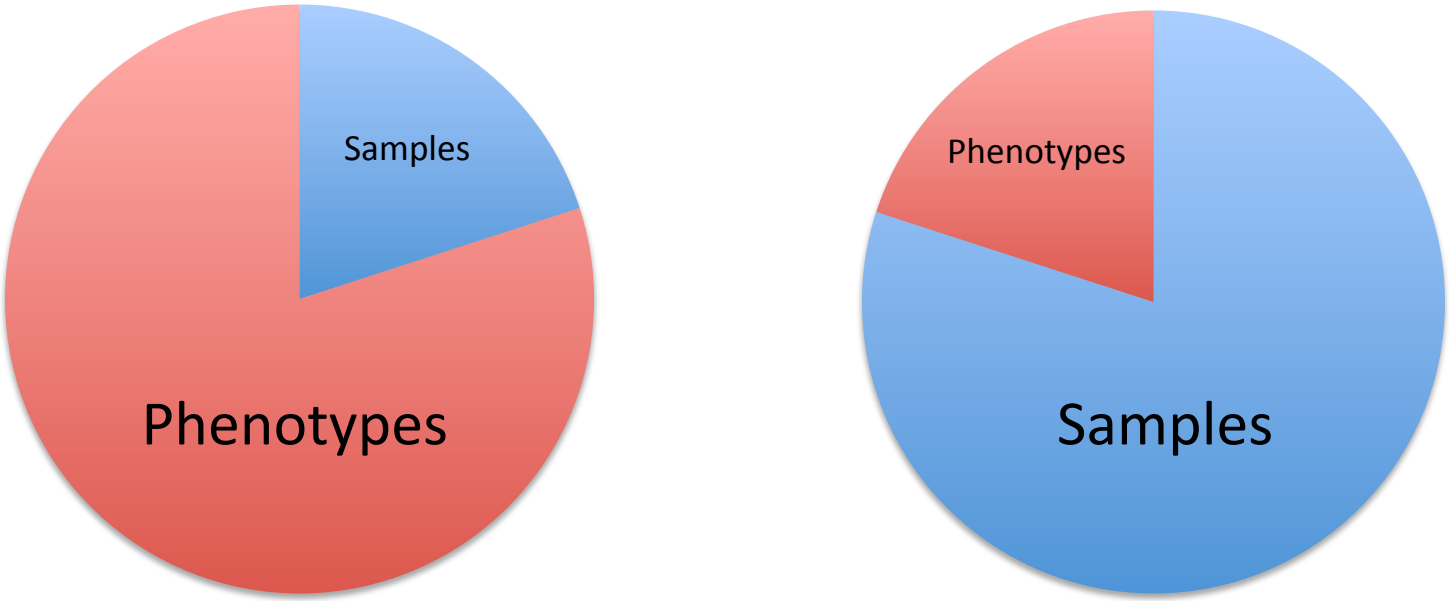
The operator has muted all lines. It is possible for just one person to ruin the call for everyone due to background noise, crying children, wind, typing, etc.

Operators announce calls one at a time during question and answer sessions.

Dial *1 if you would like to ask a question. Presenter will respond to calls as time allows.

Dial *0 if you need operator assistance at any time during the call.

If money is finite…what to do?



Or worst of both worlds? Middling numbers of phenotypes and samples?

Answer: “it depends”

It depends…on the problem (1)

- Phenotyping for **diagnostic clarity**
  - Our disorders have fuzzy boundaries, considerable comorbidity
  - Pro: increased info could help clarify dx
  - Con: Expense.
  - Con: Many practical issues – lifetime vs current symptoms, lifetime records not available, lack of standardized assessments, reporting bias of people with chronic mental illness, relevance of clinical-historical definitions to genome

It depends…on the problem (2)

- Endophenotyping to elucidate **heterogeneity**
  - Phenotypes not directly related to dx criteria, but might help identify clinical subsets
  - E.g., cognition, imaging, neurophys, biomarkers, etc.
  - Pro: continuous, more direct assessment of CNS etc.
  - Con: expense
  - Con: lifetime vs current, state vs trait, confounding
- Grade of evidence? Pure guess? Must have?

It depends…on the problem (3)

- If goal is **gene discovery**
  - Why has this been hard?
  - What lessons from the past?
  - Is it more phenotype imprecision or low power?

One view

- Assume:
  - Definite: low power for gene discovery
  - Possible: dx imprecision
  - One study can't do everything

- So, try it this way:
  - Step 1: increase N
  - Step 2: if find genes, then investigate clinical & endophenotype relevance

Maximize Sample Size

- Minimally adequate phenotyping
- Several ways to do
- All require assumptions
- 4 examples follow

Pause

Any questions so far?

CLOZUK

- Described:
  - Hampshere 2012 Mol Psych pmid 22614287
  - WCPG talk James Walters (talk link)
- Idea
  - UK clozapine clinics, treatment-resistant psychosis
  - 3rd party linked DNA-minimal phenotypes, anonymized. DNA blood after testing.
  - 1 year to get IRB approvals
  - 3 months to get >6000 samples
  - Costs: €20 (vs €1000 per case)
- *Verify:* genetic results similar to pgc-conventional

Electronic Medical Records

- Harvard i2b2
  - https://www.i2b2.org, pmid 21587298
  - Use of discarded clinical samples + text mining of electronic medical records

- Kaiser Permanente
  - Large HMO in California
  - gwas on 100K, 7K MDD (Neil Risch, in prep)

MDD

- MDD has problems – big N & no hits
  - Either we narrow phenotype (on-going)
  - Or we increase N (also on-going)
- 23andMe: has gwas on 180K plus "have you ever been dx'd with clinical depression", BIP and meds
- CHARGE & CESD (MDD sx, past 2 weeks)

Pheno experts skeptical, but willing to try. Verify.

Sweden

Thanks to Christina Hultman

Hospital Discharge Register

- Sweden pop 9M
- HDR – up to 8 ICD diagnoses
- >99% of all inpatient admissions
- Psych admissions 1973 to present

Operational definition of SCZ

- Assume >95% of people with SCZ are hospitalized
- Define SCZ as ≥ 2 inpatient admissions with discharge diagnosis consistent with SCZ
- ≥ 2 to avoid coding errors, etc
- ICD-8 295, ICD-9 295, ICD-10 F20 (exclude "latent SCZ", borderline PD)
- Manual dx refinement

(1) Is this definition SCZ?

General

- HDR basis of many peer-reviewed & highly cited papers on epidemiology of SCZ
- In general, HDR high agreement with direct med and psych diagnoses

(2) Is this definition SCZ?

Focus on SCZ diagnosis

- Nordic countries: SCZ more influenced by biological theories: “schizophrenia diagnosis has been given with great restriction in Swedish hospitals”)
- Ekholm 2005: HDR SCZ vs structured interview, 94% agreement
- Hultman: HDR vs chart review, 97% agree

(3) Is this definition SCZ?

- Epidemiology:
  - merge HDR with total population register
  - lifetime prev of SCZ in Sweden 0.407% (95% CI, 0.402-0.411%)
  - Saha (2005) meta-analysis 0.4% (pmid 15916472)

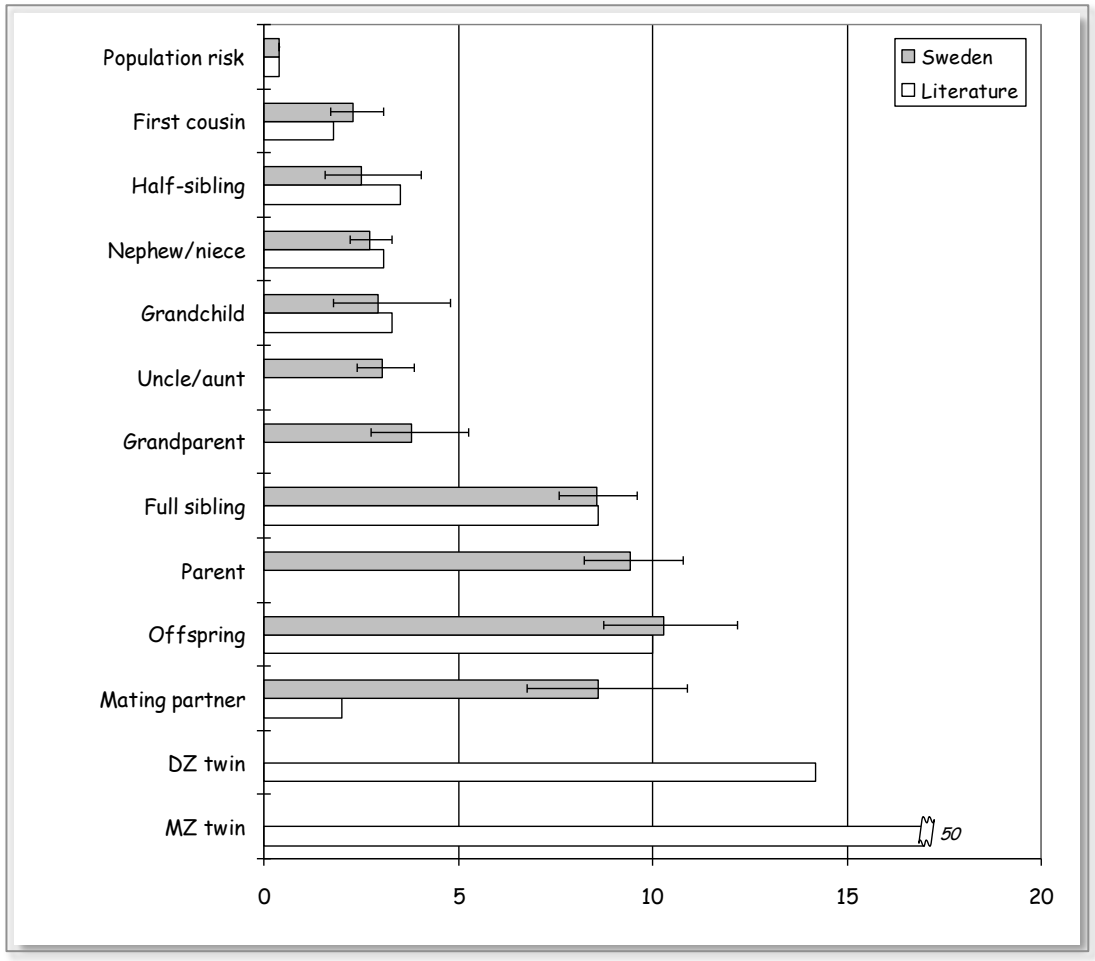
Effects of misclassification

- SCZ ≥2 will not be perfect
  - some cases will not be cases
  - for uncommon disorder, assume controls are correct
- Power calculations (Shaun Purcell)
  - Case misclassification rates 2.5%, 5%, and 10%
  - Across range of MAF and GRR
  - ratio of power with/without misclassification
  - 2.5% 0.98
  - 5% 0.95
  - 10% 0.91
- Conclude – misclassification ~5% acceptable
- Note that all studies have misclassification

(4) Is this definition SCZ?

- Genetic epidemiology:
  - N=32,536 met our ≥2 criterion for SCZ
  - merge HDR with Multi-Generation Register
  - 7,739,202 individuals, clustered into 3,664,856 family groups (1^{st} 2^{nd} 3^{rd} degree relatives)

(4, con't) Is this definition SCZ?



PMID 16863597, hazard ratios for risk of SCZ in relatives compared to Gottesman & Shields

(4, con't) Is this definition SCZ?

	Additive genetic effects (A)	Childhood shared environmental effects (C)	Non-shared environmental effects (E)
Non-hierarchical diagnoses			
Schizophrenia	64·3% (61·7%–67·5%)	4·5% (4·4%–7·4%)	31·1% (25·1%–33·9%)
Bipolar disorder	58·6% (56·4%–61·8%)	3·4% (2·3%–6·2%)	38·0% (32·0%–41·2%)
Comorbidity	63·4% (62·0%–64·9%)	5·9% (4·0%–6·8%)	30·6% (28·7%–32·3%)

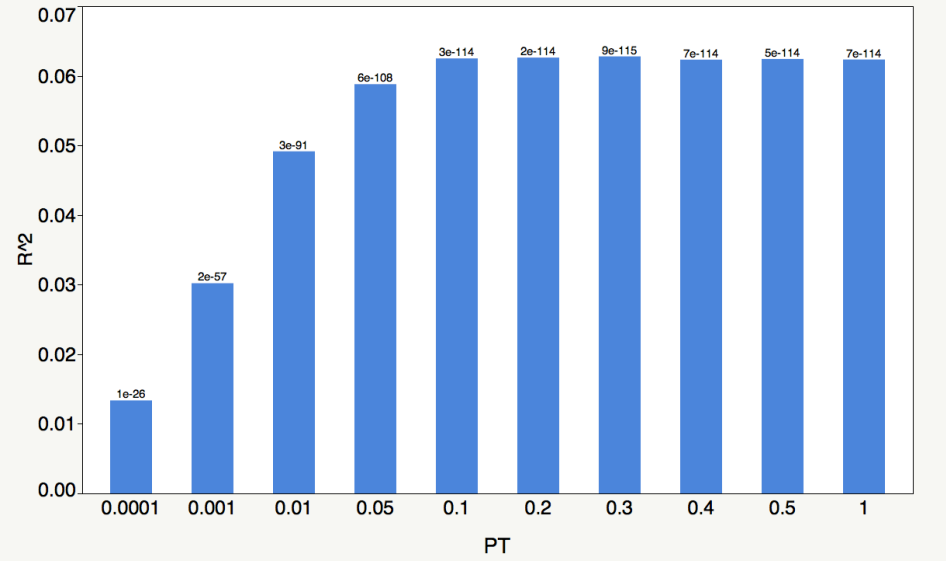
Heritability of SCZ in Sweden, pmid 19150704

Similar to recent Danish paper (Wray & Gottesman)

Less than my twin meta-analysis (0.81, pmid 14662550)

(5) Is this definition SCZ?

- Genetic results:
  - PGC SCZ 1 conventional diagnosis, 8832 cases
  - Sweden, 5001 cases
- PGC vs Sweden:
  - Sign test, $2x10^{-22}$
  - Risk profile scores (130K SNPs)



Sweden

- HDR ≥2 SCZ discharge diagnoses
  - As a group, this maps well onto “schizophrenia”
  - Based on epi, gen epi, genetics

Implications

For the goal of increasing N

- Minimal phenotyping can work
  - Must be done with care & thoughtfulness
  - Might work badly for some disorders
  - No panacea, for goal of gene identification

Relevance for PGC

- Psychchip being developed
- Looks to be gwas + exome + 20K custom SNPs for ~ 100 $US
- PGC can pay for 100K samples
- Chip available for anyone (PGC will +++ support any grant applications)

We need samples: can you get 1000+ cases fast?

Write a grant!

Thanks

My lab

Sweden colleagues
- Christina Hultman
- Patrik Magnusson
- Anna Kahler
- Paul Lichtenstein

Mt Sinai
- Pamela Sklar
- Shaun Purcell

Broad
- Ed Scolnick
- Steve McCarroll
- Jennifer Moran
- Stephan Ripke
- Ben Neale

Funders
- NIMH
- Stanley Center

The PGC !

Sweden papers: GWAS submitted (N=11K), in prep ex chip (N=11K) & ex seq (N=5K). Planned – CNV, GxE, etc

Questions?