Methodological Advances in the Study of Hidden Variables: A Demonstration on Clinical Alcohol Use Disorder Data

Erich Kummerfeld, PhD, Justin A. Anker, PhD, Alexander Rix, MS, Matt G. Kushner, PhD

University of Minnesota

November 16, 2016

Alcohol Use Disorder: Importance

- Alcohol use disorder (AUD) affects over 15 million people in the US alone (2015 National Survey on Drug Use and Health).

Alcohol Use Disorder: Importance

- Alcohol use disorder (AUD) affects over 15 million people in the US alone (2015 National Survey on Drug Use and Health).
- Alcohol misuse cost the United States $249.0 billion in 2010.

AUD: Observations That Sparked This Project

- 1 in 3 people with AUD also suffer from anxiety or depression (“internalizing” disorders).

AUD: Observations That Sparked This Project

- 1 in 3 people with AUD also suffer from anxiety or depression ("internalizing" disorders).
- Patients with both AUD and internalizing disorders are twice as likely to relapse.

AUD: Observations That Sparked This Project

- 1 in 3 people with AUD also suffer from anxiety or depression ("internalizing" disorders).
- Patients with both AUD and internalizing disorders are twice as likely to relapse.
- The mechanisms that produce and maintain these disorders are not well understood.

The Data

Measure	Mean (SD)	Range
Generalized anxiety	64.13 (11.59)	16–80
Depression	20.43 (17.30)	0–63
Social anxiety	32.43 (17.30)	0–80
Panic	10.99 (6.34)	0–28
Agoraphobia	31.59 (19.78)	0–100
Perceived stress	28.15 (5.50)	10–40
Self-efficacy	32.91 (10.91)	8–48
Drinking to cope	62.93 (12.15)	20–80
Drinking behavior	1608.76 (1271.51)	30–6840
Alcohol craving	2.67 (1.05)	0–4

Table: Variables measured from AUD patients, N = 362

Methods

- GLASSO
  - Popular method for learning undirected graphs

Methods

- GLASSO
  - Popular method for learning undirected graphs
- "Hillclimbing"
  - Popular method for learning directed acyclic graphs (DAGs)

Methods

- GLASSO
  - Popular method for learning undirected graphs
- "Hillclimbing"
  - Popular method for learning directed acyclic graphs (DAGs)
- Greedy Fast Causal Inference (GFCI)
  - Recent, untested method for learning partial ancestral graphs (PAGs)

Methods

- GLASSO
  - Popular method for learning undirected graphs
- "Hillclimbing"
  - Popular method for learning directed acyclic graphs (DAGs)
- Greedy Fast Causal Inference (GFCI)
  - Recent, untested method for learning partial ancestral graphs (PAGs)
- Factor Analysis
  - classic method for learning factor models

Methods

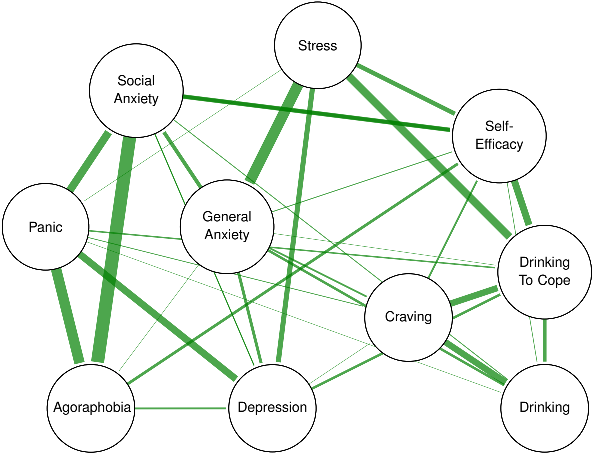
- GLASSO
  - Popular method for learning undirected graphs
- “Hillclimbing”
  - Popular method for learning directed acyclic graphs (DAGs)
- Greedy Fast Causal Inference (GFCI)
  - Recent, untested method for learning partial ancestral graphs (PAGs)
- Factor Analysis
  - classic method for learning factor models
- Find One Factor Clusters (FOFC)
  - Recent, untested method for learning latent variable models

Methods Comparison

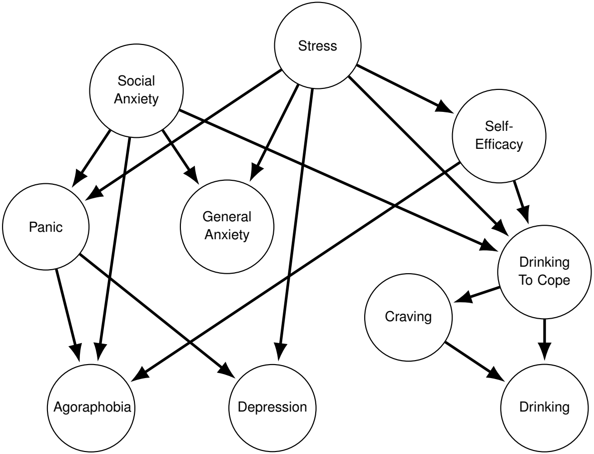
Method	Representation	Causal	Latents
GLASSO	Undirected graph	no	no
Hillclimbing	DAG	yes	no
GFCI	PAG	yes	allowed
Factor analysis	Factor model	no	modeled
FOFC	Latent variable model	yes	modeled

Table: Comparison of utilized learning methods

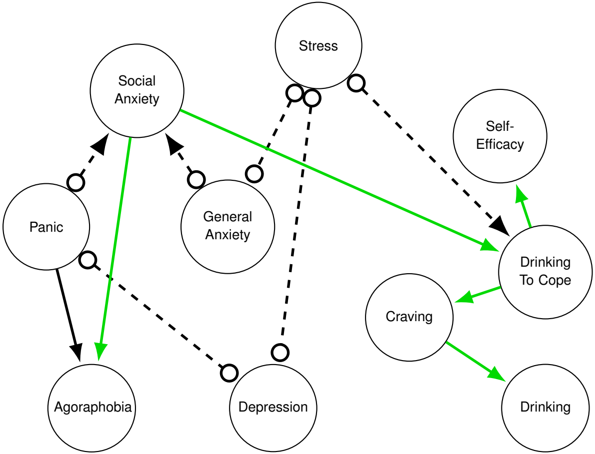
GLASSO


Stress
Social Anxiety
Self-Efficacy
Panic
General Anxiety
Drinking To Cope
Craving
Agoraphobia
Depression
Drinking

Hillclimbing



Greedy Fast Causal Inference (GFCI)



Factor Analysis

Variable	F1	F2	F3	F4
DEP1		0.47		
DEP2		0.56		
DEP3		0.55		
DEP4		0.56		
DEP5		0.47		
DEP6		0.50		
DEP7		**0.63**		
DEP8		**0.60**		
DEP9		0.44		
DEP10		0.41		
DEP11				
DEP12		**0.67**		
DEP13		**0.69**		
DEP14		0.53		
DEP15		**0.60**		
DEP16		0.48		
DEP17		0.57		
DEP18		0.51		
DEP19				
DEP20		0.39		
DEP21		0.30		
PAN1			**0.65**	
PAN2			**0.69**	
PAN3			**0.74**	
PAN4			**0.78**	
PAN5			**0.77**	
PAN6			**0.83**	
PAN7			**0.80**	

Variable	F1	F2	F3	F4
GAD1r				
GAD2	0.57			
GAD3r	0.31			
GAD4	**0.73**			
GAD5	**0.72**			
GAD6	**0.66**			
GAD7	**0.89**			
GAD8r	0.31			
GAD9	**0.72**			
GAD10r	0.33			
GAD11r	0.40			
GAD12	**0.69**			
GAD13	**0.67**			
GAD14	**0.83**			
GAD15	**0.90**			
GAD16	**0.70**			
STR1				
STR2				0.49
STR3				0.47
STR4r				**0.68**
STR5r				**0.74**
STR6				0.37
STR7r				0.32
STR8r				**0.66**
STR9				
STR10				0.60

Find One Factor Clusters (FOFC)

Variable	C1	C2	C3	C4
DEP1	0.43			
DEP2	0.48			
DEP3	0.30			
DEP4	0.60			
DEP5	0.37			
DEP6	**0.82**			
DEP7				
DEP8	**0.87**			
DEP9	**0.92**			
DEP10	**0.95**			
DEP11				
DEP12	**0.60**			
DEP13	**0.62**			
DEP14	**0.85**			
DEP15				
DEP16	**0.88**			
DEP17				
DEP18				
DEP19			0.43	
DEP20				
DEP21	**0.98**			
PAN1				
PAN2				
PAN3				
PAN4				
PAN5				
PAN6				
PAN7				

Variable	C1	C2	C3	C4
GAD1r		**0.68**		
GAD2		0.47		
GAD3r				
GAD4		0.30		
GAD5				
GAD6		0.49		
GAD7				
GAD8r				
GAD9				
GAD10r		0.33		
GAD11r		0.57		
GAD12		0.41		
GAD13		0.59		
GAD14				
GAD15				
GAD16				0.49
STR1				
STR2				
STR3				
STR4r				
STR5r				0.41
STR6				**0.67**
STR7r			0.42	
STR8r				0.46
STR9				**0.91**
STR10				

Future Work

- Investigate item-level graphical structures
- Determine why factor analysis and FOFC disagree about Panic items
- Apply methods to other data sets
- Continue to apply promising new methods to real data

Funding: This work was supported by NIAAA grant K01AA024805 (Anker), NIAAA grant R01-AA015069 (Kushner), NCATS grant UL1 TR002494 (Blazar), and NIH/NIMH grant 1R01MH116156-01A1 (Nielson).