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

Erich Kummerfeld, PhD[1], Justin A. Anker, PhD[2], Alexander Rix, MS[1], Matt G. Kushner, PhD[2]
[1]**University of Minnesota – Institute for Health Informatics, Minneapolis, Minnesota**
[2]**University of Minnesota – Department of Psychiatry, Minneapolis, Minnesota**

Abstract *Research in the domain of psychopathology has been hindered by hidden variables—variables that are important to understanding and treating psychopathological illnesses but are unmeasured. Recent methodological advances in machine learning have culminated in the ability to discover and identify the influence of hidden variables that confound the observed relationships among measured variables. We apply a combination of traditional methods and more recent advances to a data set of alcohol use disorder patients with comorbid internalizing disorders, and find that the increasingly advanced methods produce increasingly informative and reliable results. These results include novel findings evaluated positively by our psychopathologists, as well as findings validated with knowledge from existing literature. We also find that advanced graph discovery methods can guide the use of latent variable modeling procedures, which can in turn explain the output of the graph discovery methods, resulting in a synergistic relationship between two seemingly distinct classes of methods.*

Introduction

According to the 2015 National Survey on Drug Use and Health, alcohol use disorder (AUD) affects over 15 million people in the US alone, and in 2010 it was estimated that alcohol misuse cost the United States $249.0 billion[1]. Approximately one third of that population also suffers from anxiety or depression ("internalizing") disorders, and following treatment, patients who suffer from both AUD and internalizing disorders are twice as likely to relapse[2–5]. As in many psychopathology domains, the mechanisms that produce and maintain these disorders are not well understood, so there is a critical need for discoveries that inform the prevention and treatment of AUD.

To better understand the mechanisms underlying the high rate of co-occurrence between AUD and internalizing disorders, we analyzed a high quality clinical data set containing psychiatric measurements of a cohort of alcohol use disorder patients with a secondary internalizing disorder diagnosis. We used a combination of traditional methods for studying the structure of psychopathology data and more recently developed methods for studying structural hidden variables. Specifically, we used the Graphical Least Absolute Shrinkage and Selection Operator (GLASSO)[6], Scutari's version of Hillclimbing (SHC)[7], Greedy Fast Causal Inference (GFCI)[8], factor analysis (FA)[9], and Find One Factor Clusters (FOFC)[10].

Methodologically we found that a combination of GFCI and FOFC offered the most reliable and informative knowledge regarding the structure of co-occurring alcohol use disorder and internalizing disorders. With these methods we discovered a prominent cause of alcohol consumption that has been previously conjectured in the literature to play a special role in AUD, and identified an unmeasured common cause influencing a mixture of anxiety and stress items.

Data

Data was collected from a 21-day community-based residential chemical dependency treatment program, and a subset of patients ($N = 362$) were selected with primary alcohol use disorder and a secondary anxiety disorder. Measures of anxiety and depression ("internalizing") symptoms, stress and coping abilities, drinking behaviors, and alcohol craving were collected on every patient, with no skip questions and very few missing values, leading to a high quality data set. Variables were constructed from individual items based on standard scales for the various internalizing disorders.

Methods

The first method we consider is the graphical lasso (GLASSO), a popular tool used for discovering unoriented graphs from observational data, including data related to psychopathology. GLASSO estimates the inverse covariance matrix using an L_1 penalty, which is a well-understood and studied statistical object. The primary distinction between

Measure	Mean (SD)	Range	Description
Generalized anxiety	64.13 (11.59)	16-80	The total score on the Penn State Worry Questionnaire.[11]
Depression	20.43 (17.30)	0-63	The total score on the Beck Depression Inventor[12].
Social anxiety	32.43 (17.30)	0-80	The total score on the Social Phobia Scale[13].
Panic	10.99 (6.34)	0-28	The total score on the Panic Disorder Severity Scale[14].
Agoraphobia	31.59 (19.78)	0-100	The summed score from the Mobility Inventory for Agoraphobia[15].
Perceived stress	28.15 (5.50)	10-40	The total score on the Perceived Stress Scale[16].
Self-efficacy	32.91 (10.91)	8-48	The total score on the negative affect subscale of the Situational Confidence Questionnaire[17].
Drinking to cope	62.93 (12.15)	20-80	The Unpleasant Emotions subscale of the Inventory of Drinking Situations[18].
Drinking behavior	1608.76 (1271.51)	30-6840	The total drinks consumed during the 4 months prior to residential treatment entry assessed with the Timeline Follow-Back Interview[19].
Alcohol craving	2.67 (1.05)	0 to 4	The frequency of alcohol craving during the 30 days prior to treatment assessed with an item from the Obsessive Compulsive Drinking Scale[20].

Table 1: Measured variables in the clinical data set, N = 362

GLASSO and the other graph-learning algorithms we employed is that GLASSO learns an undirected graph that does not encode any causal information, so it serves as a point of comparison for the causal methods. Undirected graphs can be difficult to interpret, especially as the number of variables and edges increases. Because of this, interpretation of these graphs is typically done at a relatively high level: the graph is fed into an analysis method which evaluates the various nodes in the graph according to various graphical metrics, such as "centrality" and "connectedness". The nodes that rank highly on these metrics are identified as being important nodes in the network, and are often conjectured to be important targets for treatment or for further investigation. Well-connected groups of nodes ("clusters") can also be identified as collections of variables that seem to be categorically similar.

Scutari's version of Hillclimbing[7] has recently been utilized in some psychopathology publications[21]. We utilized the version of this method implemented in the R package $bnlearn$. This method attempts to learn the causal structure of the variables by optimizing a complexity-penalized likelihood score, typically the Bayesian Information Criterion (BIC). It has not been proven to be correct in the infinite data limit, but can perform well in simulations. It outputs a directed acyclic graph (DAG) in which nodes may be connected to each other by arrows. The directed edges in these graphs are frequently interpreted causally, such that the edge $A \longrightarrow B$ is interpreted to mean that the variable A causally influences the variable B, as utilized by Pearl[22] and Spirtes[23], and explicated by Woodward[24]. There are numerous other methods that produce directed, or partially directed, graphical models such as PC[23] and GES[25]; we selected SHC due to its use in prior psychopathology research.

Unlike the previous methods, GFCI[8] learns the relationships among the variables without assuming that there are no hidden common causes. We used the implementation of this method found in the $Tetrad$ software package. It has been proven correct in the infinite data limit, and while benchmark simulations of its performance on finite sample sizes are as-yet limited in scope, the benchmarking that has been done so far is promising[8]. In terms of scalability to large numbers of variables, these more complex algorithms are naturally slower than methods like GLASSO, however there are many data sets, such as the one covered in this paper, which are well within GFCI's feasibility bounds. As part of being able to tolerate the possible existence of hidden common causes, GFCI outputs a partial ancestral graph (PAG)[23], a graphical representation that encodes the possibility of latent confounders. PAGs use a rich set of edge types to encode a large amount of information, including whether a given relationship is definitely, possibly, or definitely not confounded, as well as whether a variable definitely, possible, or definitely does not cause another variable. In the

typical representation, the inclusion of a circle at either end of the edge indicates the possibility that a latent variable may be responsible for part (or all) of the statistical signal between those variables, while the inclusion of arrowheads at both ends of the edge indicates that the relationship is definitely due to a latent variable causing them both. The total set of possible edge types of a PAG is quite large and beyond the scope of this paper, however we include a reference Table 3 for the edges that are relevant for our demonstration below.

For analyzing unmeasured common causes, the number of available methods is limited. There are two primary approaches: Factor Analysis (FA)[9] is the traditional technique, while graphical latent effect estimation (GLEE)[10,26,27] is a recent development. We used the $factanal$ function in R, and the implementation of FOFC found in the $Tetrad$ software package. For our purposes, the primary distinction between FA and GLEE is the presence or absence of correctness proofs for detecting unmeasured common causes: no FA method has such a proof, while several GLEE methods do. As such the interpretation of FA results can be limited, since there is no theorem stating what they discover under what circumstances. On the other hand, GLEE methods have been proven to identify unmeasured common causes under very broad assumptions, making the interpretation of their results simple and reliable.

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

Table 2: Comparison of utilized learning methods

Application to Clinical Data

We first consider the undirected graph learned by GLASSO[28], shown in Figure 1. The graph shows dense connections amongst the various internalizing disorders, and also has the strongest connections from the alcohol focused variables (Drinking and Craving) to the drinking to cope (DTC) variable. DTC also has the highest "centrality" in the graph, and is between Craving and Drinking and the rest of the graph. These heuristics indicate that DTC could play an important role in the common co-occurrence of drinking problems and internalizing disorders, but they don't offer concrete causal information.

A lack of causal information is an implicit limitation of any method, like GLASSO, that produces undirected networks: causation is inherently a directed relationship. To begin investigating the causal relationships among the variables in this data set we used SHC[7] to learn a directed acyclic graph (DAG), shown in Figure 2. The DAG encodes a variety of descriptive statistics and causal information that is absent from the GLASSO model. For example, according to this DAG, conditioning or controlling for DTC makes Drinking statistically independent of Depression. For causal information, the DAG implies that effectively treating someone's depression would not affect their drinking, but effectively treating their social anxiety or stress would. Some of this information is already known in the literature, e.g. that DTC causes Craving, and that Craving causes Drinking; however some of this information is novel, e.g. that Social anxiety might be partly responsible for the co-occurrence of depression and AUD. If we can be confident that this DAG represents the true causal processes among the measured variables, then some critical information regarding the treatment of AUD leaps out of the graph: DTC, Stress, and Social anxiety are identified as the only treatment targets that would make any impact on actual drinking behavior.

Therein lies a critical problem with methods that only produce DAGs: how can we be confident that the resulting graph represents the truth? In complex and difficult domains such as psychopathology, researchers are rarely in a position where they believe that their data contains all of the important causal factors. DAG-learning methods are widely known to make errors in the presence of hidden common causes[23], so if we believe that our data is, or might be, subject to such confounding, then this casts doubt on the truth of the DAGs we learn.

Taking these concerns seriously, we applied GFCI[8], which produced the PAG shown in Figure 3. PAGs utilize a rich set of possible edge types, and covering them all is outside the scope of this paper, but Table 3 provides the interpretation of the edges found in this PAG. This graph confirms some of the information contained in the DAG,

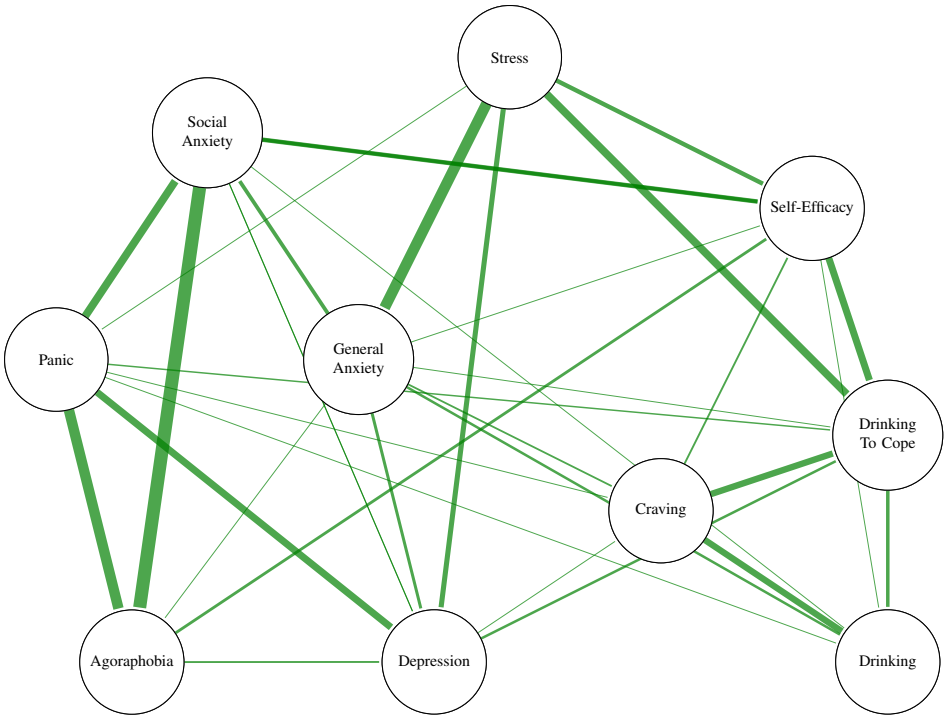

Figure 1: Visualization of GLASSO network

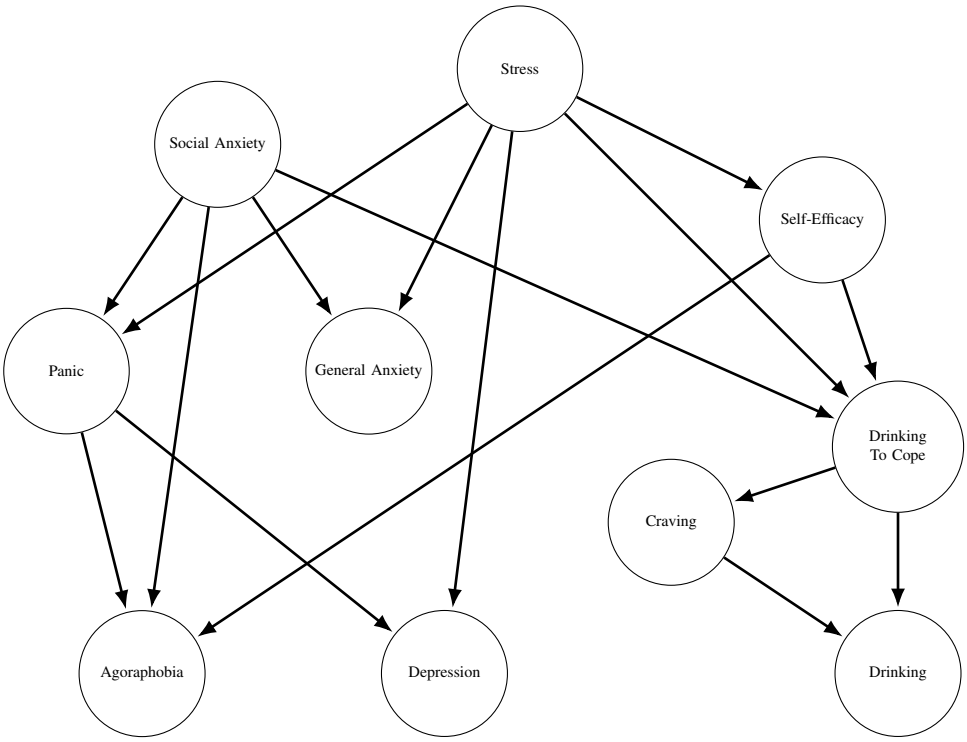

Figure 2: Visualization of SHC DAG

such as the causal chain DTC $\longrightarrow$ Craving $\longrightarrow$ Drinking, but disagrees with the DAG in other areas, such as whether

Table 3: Edge types in a Partial Ancestral Graph (PAG)

Edge Type	Meaning
A ○- - - -○ B	Precisely one of the following is true: 1. A causes B 2. B causes A 3. A and B are confounded 4. both 1. and 3. 5. both 2. and 3.
A ○- - - -▸ B	Precisely one of the following is true: 1. A causes B 2. A and B are confounded. 3. both 1. and 2.
A ——▸ B	All of the following are true: 1. A is a direct or indirect cause of B. 2. A and B are not confounded. 3. B is not a cause of A.
A ——▸ B	All of the following are true: 1. A is a direct cause of B. 2. A and B are not confounded. 3. B is not a cause of A.

DTC causes self-efficacy or self-efficacy causes DTC. All of the variables that are adjacent to each other in the PAG are also adjacent in the DAG, but the DAG contains a small number of adjacencies that are not present in the PAG. It's possible that these differences are errors induced by the presence of confounding variables, but these differences could also be due to differences in the parameter settings we used for SHC and GFCI: the two methods have distinct sets of parameters with different interpretations and meanings, so it is difficult to make sure that they are equivalently calibrated.

There are a few striking similarities between the DAG and the PAG. DTC is causally upstream of Craving and Drinking in both graphs, and GFCI goes a step further by confirming that these relationships are not confounded by unmeasured variables. GFCI also discovers that the causal effect of Social anxiety on DTC is not confounded, but the causal relationship between Stress and DTC might be: in fact it allows for the possibility that Stress is not a cause of DTC at all, but rather their correlation could be due entirely to an unmeasured common cause.

GFCI also identifies numerous other places in the graph where unmeasured variables could be influencing the observed relationships: half of the causal relationships in the PAG are possibly confounded. This casts doubt on the orientations that SHC gave to those edges. The possibility of confounding is especially present within the distress domain (Depression and General anxiety), as all of the causal relationships connecting these variables to other variables in the graph

are possibly confounded. There is also a small number of orientation disagreements between the GFCI PAG and the DAG learned by SHC, but these are simply low confidence graph features.

Our psychopathologists were surprised at the prominence of Social anxiety within the causal networks over other possible causes of Drinking. Both models indicated that Social anxiety served as an initiator of the causal chain of conditions that terminate with Drinking, and GFCI confirmed that this entire causal pathway is unconfounded. The social anxiety $\longrightarrow$ DTC $\longrightarrow$ Craving $\longrightarrow$ Drinking causal chain is supported by findings from several other studies. In one study, social anxiety was highly correlated with endorsement of drinking in unpleasant emotions and, in fact, DTC mediated the relationship between social anxiety and problematic alcohol consumption[29]. In another study of a community sample of individuals with co-occurring social anxiety and alcohol use disorder, DTC with social anxiety mediated the relationship between social anxiety and drinking problems[30]. In still another study of individuals with AUD, drinking to cope with social anxiety mediated the relationship between social anxiety and drinking problems[31].

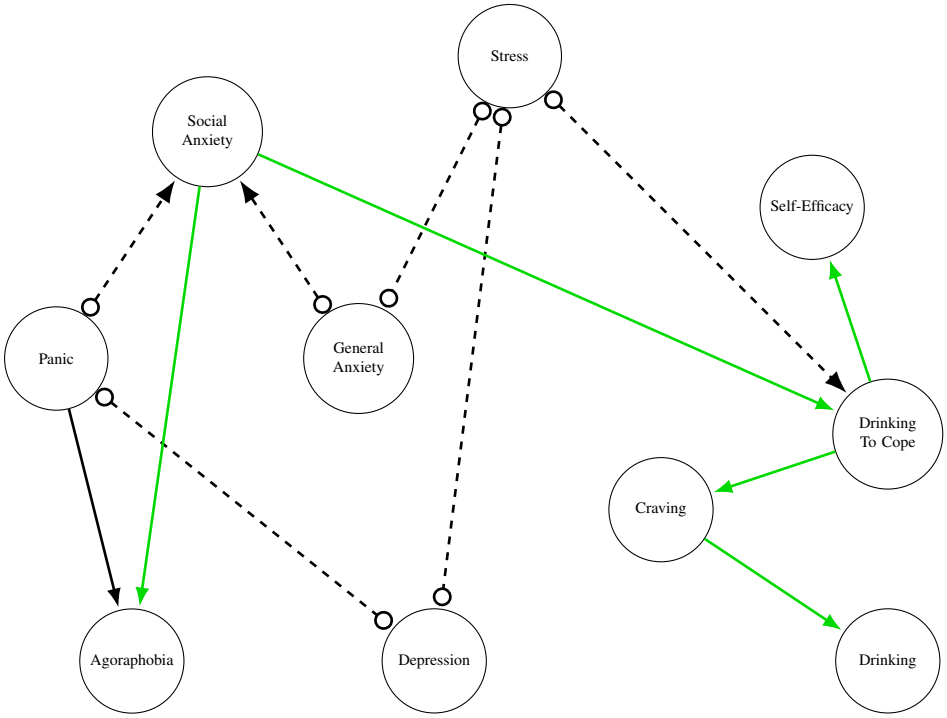

Figure 3: Visualization of GFCI PAG

Since GFCI identified several pairs of variables which might be confounded, we decided to investigate the possible confounders by using methods for investigating unmeasured common causes. Aside from Drinking and Craving, the variables are scores calculated from multiple survey items, so we decided to investigate the hypothesis that some survey items from different scores were actually caused by a single unmeasured variable. To test this hypothesis, we analyzed the item-level data for the Stress, General anxiety, Depression, and Panic scores, as the PAG indicated that all of the relationships among these scores might be confounded. This created a data set of 54 items: 10 for Stress, 21 for Depression, 7 for Panic, and 16 for General anxiety.

We applied factor analysis to the item-level data, using four factors since we know they come from four scores. This was augmented with the oblique promax rotation, since the scores are correlated (as shown by the undirected graph, the DAG, and the PAG). The results are shown in Table 4. The factor analysis does not identify any cross-loadings, even at a low cutoff threshold of 0.3, which would imply that the items are all measuring only the factors/scores that they are supposed to.

The lack of findings from factor analysis is contrasted by the results of a graphical latent effect estimation (GLEE) method, FOFC[10], that we also used on the item-level data to test for item-level confounding. While FOFC has been

Table 4: Factor loadings for STR, DEP, GAD, and PAN items. Loadings were calculated using the factanal function in R, with 4 factors and the oblique promax rotation. A cutoff value of 0.3 is used, and values above 0.6 are bolded.

Factor loadings for DEP and PAN items

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**	

Factor loadings for GAD and STR items

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

Table 5: FOFC output for STR, DEP, GAD, and PAN items, after being aggregated with K-means for k=4 with 50 random restarts. Centroid values are shown with a cutoff value of 0.3 and values above 0.6 in bold.

Centroid values for DEP and PAN items

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				

Centroid values for GAD and STR items

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				

proven correct, its output on finite sample sizes can depend on the variable order, so we ran FOFC 100 times on random variable orderings and stored all of the output factors that it identified. Output factors that loaded onto only a small number of items (≤ 4) were dropped since it is known that such small factors are less reliable than larger factors for this method. K-means clustering (with $k = 4$ selected to mirror the choice of 4 factors) was used to aggregate the output from the 100 FOFC runs. The cluster centroids are shown in Table 5. Notably, FOFC finds that some cross-score items share a latent common cause. In particular, centroid 4 shows one General anxiety item sharing a hidden cause with 4 Stress items. Our psychopathology experts inspected these test items and were excited to find that they were related to feelings of being overwhelmed by the daily obligations and routines of life, content which is largely missing from the other General anxiety and Stress items and is very relevant to the lives of patients with AUD. Further work is required to investigate this possible source of item-level cross-score confounding, but this could ultimately lead to better estimates of the existing scores, and even to identifying novel important psychopathological concepts that may require unique psychopathological attention.

Conclusion

In this paper we applied several statistical methods of varying levels of sophistication to a clinical data set of patients with alcohol use disorder. The methods vary significantly in terms of what can be learned from them, in particular in terms of causal information and the ability to detect and identify unmeasured common causes. In doing so we demonstrated how recent advances are enabling the discovery of novel and important knowledge in the difficult domain of psychopathology. Such knowledge included not only clinical insights about the treatment of patients with alcohol use disorder, but also potential problems with the tests being used to identify and measure concepts such as general anxiety and stress. In the future, we hope to explore these potential cross-score latent common causes by analyzing other data sets that use the same items, or potentially by constructing new sets of items intended specifically to target these latent variables. Importantly, none of these methods are tailored to the domain of psychopathology: they could easily be applied to other clinical domains, as well as to more biological domains such as protein signalling[32] or gene expression. Many such fields would benefit from adopting these cutting-edge representations and statistical methods in their investigations.

Acknowledgements

This work was supported by NIAAA grant K01AA024805 awarded to the Justin Anker, and NIAAA grant R01-AA015069 awarded to Matt Kushner. The authors acknowledge the Center for Causal Discovery for the development of the statistical software (Tetrad) used by Erich Kummerfeld and Alexander Rix to conduct the GFCI and FOFC analyses. The authors thank Constantin Aliferis for helpful discussions and feedback on an earlier version of this manuscript.

References

[1] Sacks JJ, Gonzales KR, Bouchery EE, Tomedi LE, Brewer RD. 2010 national and state costs of excessive alcohol consumption. American journal of preventive medicine. 2015;49(5):e73–e79.

[2] Greenfield SF, Weiss RD, Muenz LR, Vagge LM, Kelly JF, Bello LR, et al. The effect of depression on return to drinking: a prospective study. Archives of general psychiatry. 1998;55(3):259–265.

[3] Kaufmann CN, Chen LY, Crum RM, Mojtabai R. Treatment seeking and barriers to treatment for alcohol use in persons with alcohol use disorders and comorbid mood or anxiety disorders. Social psychiatry and psychiatric epidemiology. 2014;49(9):1489–1499.

[4] Kushner MG, Abrams K, Thuras P, Hanson KL, Brekke M, Sletten S. Follow-up study of anxiety disorder and alcohol dependence in comorbid alcoholism treatment patients. Alcoholism: Clinical and Experimental Research. 2005;29(8):1432–1443.

[5] Schellekens A, de Jong C, Buitelaar J, Verkes R. Co-morbid anxiety disorders predict early relapse after inpatient alcohol treatment. European Psychiatry. 2015;30(1):128–136.

[6] Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics. 2008;9(3):432–441.

[7] Scutari M. Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software. 2010;35(3):1–22.

[8] Ogarrio JM, Spirtes P, Ramsey J. A hybrid causal search algorithm for latent variable models. In: Conference on Probabilistic Graphical Models; 2016. p. 368–379.

[9] Krueger RF. The structure of common mental disorders. Archives of General Psychiatry. 1999;56(10):921–926.

[10] Kummerfeld E, Ramsey J. Causal clustering for 1-factor measurement models. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM; 2016. p. 1655–1664.

[11] Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the penn state worry questionnaire. Behaviour research and therapy. 1990;28(6):487–495.

[12] Beck AT, Steer RA, Brown GK. Beck depression inventory-II. San Antonio. 1996;78(2):490–8.

[13] Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety1. Behaviour research and therapy. 1998;36(4):455–470.

[14] Houck PR, Spiegel DA, Shear MK, Rucci P. Reliability of the self-report version of the panic disorder severity scale. Depression and anxiety. 2002;15(4):183–185.

[15] Chambless DL, Caputo GC, Jasin SE, Gracely EJ, Williams C. The mobility inventory for agoraphobia. Behaviour research and therapy. 1985;23(1):35–44.

[16] Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. Journal of health and social behavior. 1983;p. 385–396.

[17] Annis H, Graham JM, of Ontario ARF, Foundation OAR. Situational Confidence Questionnaire (SCQ): User's Guide. Addiction Research Foundation; 1988.

[18] Annis HM, Graham JM. Profile types on the Inventory of Drinking Situations: Implications for relapse prevention counseling. Psychology of Addictive Behaviors. 1995;9(3):176.

[19] Sobell LC, Sobell MB. Timeline follow-back. In: Measuring alcohol consumption. Springer; 1992. p. 41–72.

[20] Anton RF, Moak DH, Latham PK. The obsessive compulsive drinking scale: a new method of assessing outcome in alcoholism treatment studies. Archives of general psychiatry. 1996;53(3):225–231.

[21] McNally R, Mair P, Mugno B, Riemann B. Co-morbid obsessive–compulsive disorder and depression: A Bayesian network approach. Psychological medicine. 2017;47(7):1204–1214.

[22] Pearl J. Causality. Cambridge university press; 2009.

[23] Spirtes P, Glymour CN, Scheines R. Causation, prediction, and search. MIT press; 2000.

[24] Woodward J. Making things happen: A theory of causal explanation. Oxford university press; 2005.

[25] Chickering DM. Optimal structure identification with greedy search. Journal of machine learning research. 2002;3(Nov):507–554.

[26] Kummerfeld E, Ramsey J, Yang R, Spirtes P, Scheines R. Causal clustering for 2-factor measurement models. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer; 2014. p. 34–49.

[27] Silva R, Scheine R, Glymour C, Spirtes P. Learning the structure of linear latent variable models. Journal of Machine Learning Research. 2006;7(Feb):191–246.

[28] Anker JJ, Forbes MK, Almquist ZW, Menk JS, Thuras P, Unruh AS, et al. A network approach to modeling comorbid internalizing and alcohol use disorders. Journal of abnormal psychology. 2017;126(3):325.

[29] Buckner JD, Eggleston AM, Schmidt NB. Social anxiety and problematic alcohol consumption: The mediating role of drinking motives and situations. Behavior therapy. 2006;37(4):381–391.

[30] Clerkin EM, Sarfan LD, Parsons EM, Magee JC. Mindfulness Facets, Social Anxiety, and Drinking to Cope with Social Anxiety: Testing Mediators of Drinking Problems. Mindfulness. 2017;8(1):159–170.

[31] Cooper R, Hildebrandt S, Gerlach A. Drinking motives in alcohol use disorder patients with and without social anxiety disorder. Anxiety, Stress & Coping. 2014;27(1):113–122.

[32] Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal protein-signaling networks derived from multiparameter single-cell data. Science. 2005;308(5721):523–529.