



Published in final edited form as:

*Drug Alcohol Depend.* 2014 August 1; 141: 153–158. doi:10.1016/j.drugalcdep.2014.04.029.

## Data Compatibility in the Addiction Sciences: An Examination of Measure Commonality\*

Kevin P. Conway<sup>1</sup>, Genevieve C. Vullo<sup>1,2</sup>, Ashley P. Kennedy<sup>3</sup>, Matthew S. Finger<sup>1</sup>, Arpana Agrawal<sup>4</sup>, James M. Bjork<sup>5</sup>, Lindsay A. Farrer<sup>6</sup>, Dana B. Hancock<sup>7</sup>, Andrea Hussong<sup>8</sup>, Paul Wakim<sup>9</sup>, Wayne Huggins<sup>7</sup>, Tabitha Hendershot<sup>7</sup>, Destiney S. Nettles<sup>7</sup>, Joseph Pratt<sup>7</sup>, Deborah Maiese<sup>7</sup>, Heather A. Junkins<sup>10</sup>, Erin M. Ramos<sup>10</sup>, Lisa C. Strader<sup>7</sup>, Carol M. Hamilton<sup>7</sup>, and Kenneth J. Sher<sup>11</sup>

<sup>1</sup>Division of Epidemiology, Services, and Prevention Research, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, Bethesda, MD

<sup>2</sup>Kelly Government Solutions, Bethesda, MD

<sup>3</sup>Intramural Research Program, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services

<sup>4</sup>Washington University in St. Louis, St. Louis, MO

<sup>5</sup>Division of Clinical Neuroscience and Behavioral Research, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, Bethesda, MD

<sup>6</sup>Boston University, Boston, MA

<sup>7</sup>RTI International, Research Triangle Park, NC

<sup>8</sup>University of North Carolina at Chapel Hill, Chapel Hill, NC

<sup>9</sup>Center for the Clinical Trials Network, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, Bethesda, MD

<sup>10</sup>National Human Genome Research Institute, Bethesda, MD

<sup>11</sup>University of Missouri, Columbia, MO

### Abstract

\*Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

Corresponding Author: Kevin P. Conway, Deputy Director, Division of Epidemiology, Services, and Prevention Research (DESPR), National Institute on Drug Abuse, 6001 Executive Boulevard, Suite 5185, Bethesda, Maryland 20892-9589, tel. (301) 443-6504, fax. (301) 443-2636, [kconway@nida.nih.gov](mailto:kconway@nida.nih.gov).

#### Contributors

Authors Conway and Vullo designed the study, conducted the initial analysis and, along with Finger and Kennedy, drafted the initial manuscript. All authors contributed meaningfully to the drafting, reviewing, and revising of the manuscript.

#### Conflict of Interest

No conflict declared.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The need for comprehensive analysis to compare and combine data across multiple studies in order to validate and extend results is widely recognized. This paper aims to assess the extent of data compatibility in the substance abuse and addiction (SAA) sciences through an examination of measure commonality, defined as the use of similar measures, across grants funded by the National Institute on Drug Abuse (NIDA) and the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Data were extracted from applications of funded, active grants involving human-subjects research in four scientific areas (epidemiology, prevention, services, and treatment) and six frequently assessed scientific domains. A total of 548 distinct measures were cited across 141 randomly sampled applications. Commonality, as assessed by density (range of 0–1) of shared measurement, was examined. Results showed that commonality was low and varied by domain/area. Commonality was most prominent for (1) diagnostic interviews (structured and semi-structured) for substance use disorders and psychopathology (density of 0.88), followed by (2) scales to assess dimensions of substance use problems and disorders (0.70), (3) scales to assess dimensions of affect and psychopathology (0.69), (4) measures of substance use quantity and frequency (0.62), (5) measures of personality traits (0.40), and (6) assessments of cognitive/neurologic ability (0.22). The areas of prevention (density of 0.41) and treatment (0.42) had greater commonality than epidemiology (0.36) and services (0.32). To address the lack of measure commonality, NIDA and its scientific partners recommend and provide common measures for SAA researchers within the PhenX Toolkit.

## Keywords

Measure commonality; data harmonization; standard measures; gene-environment interactions; substance use; abuse; addiction

## 1. INTRODUCTION

The need for comprehensive analysis to compare and combine data across multiple studies in order to validate and extend results is widely recognized. When a cross-study analysis includes multiple data sets that lack measure commonality, data harmonization is needed in order to encompass the data from related, but not identical, measures and protocols. Data harmonization efforts require thoughtful consideration of what data can/should be combined, and typically include a time-consuming re-coding step prior to carrying out the cross-study analysis. While procedures for combining data sets with nonidentical measures are under active development and represent an important tool (Hussong et al. 2013; Fortier et al., 2010), greater confidence in construct equivalence (the extent to which a concept assessed across studies has the same meaning across studies) may result from using common measures during the data-collection phase. Consequently, greater use of common measures facilitates pooling of data across multiple independent studies. Such pooling enables researchers to undertake analyses addressing important research questions that the individual studies themselves are illequipped to address.

In the addiction sciences, the ability to pool data and perform integrative data analysis is particularly important and timely. The identification and verification of relatively weak associations and more complex relations, such as gene-gene and gene-environment interactions related to addictions, is vastly improved via analysis of larger samples sizes than

any single study contains or is fiscally feasible to ascertain anew (Agrawal et al., 2012; Duncan and Keller, 2011; Heath et al., 2011; Sher et al., 2010). The advantages of measure commonality also apply to intervention, epidemiological, and developmentally informative studies, including accelerated cohort designs that examine trajectories over a broad age span during a study (or studies) of short duration (Miyazaki and Raudenbush, 2000).

While it is generally recognized that low levels of measure commonality is problematic in the addiction sciences, we are aware of no empirical examination of this topic. This paper examines measure commonality across grants funded by the National Institute on Drug Abuse (NIDA) and the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Analysis of measures within grants funded by these two institutes – collectively the world's largest funding source of research in the field of substance use, abuse, and addiction – provides a comprehensive and representative picture of the extent of measure commonality in the field. Following the empirical examination, this paper discusses current efforts to increase measurement commonality and accelerate scientific discovery, and highlights important resources for substance abuse and addiction research.

## 2. METHODS

To estimate the commonality of measures related to substance use, abuse and addiction, a portfolio analysis was conducted across grants funded by NIDA and NIAAA to identify occurrences where similar measures were proposed by two or more applications. Data were extracted from applications of funded, active grants involving human-subjects research in the areas of epidemiology, prevention, services, and treatment. These areas were targeted because we assumed that grant applications within each of these areas would study related topics, and that there would be a tendency for similarity in the constructs and measures. Furthermore, investigators in these areas increasingly recognize the importance of data comparability, interoperability, and integration across multiple studies (Curran and Hussong, 2009), particularly in the area of gene-environment interactions (Bennett et al., 2011; Bierut, 2011; Cornelis et al., 2010; Duncan and Keller, 2011), a research portfolio that is largely housed within the epidemiology domain.

### 2.1 Sampling and procedure

Our intention was to sample a reasonably sized and representative number of R01 grants that were active as of late March/early April, 2011 (N=829). This date corresponds with the initiation of the NIDA supplement to the PhenX Toolkit to create the Substance Abuse and Addiction (SAA) collection, described later in this manuscript, and is intended to serve as a baseline assessment of measure commonality. After numbering each of the 829 grants, we sampled 141 using a random number generator program in Microsoft Excel; this sample size was sufficient to give a 7.5% margin of error at the 95% confidence level.

### 2.2 Coding and data analysis

The 141 grant applications were manually examined to code the proposed measures of all constructs related to substance use, abuse, and addiction. For estimating reliability, a sample of applications (n=14) was randomly selected for coding by a second independent

researcher, resulting in a total of 190 observations of measures used. These observations were compared for agreement between two coders. A total of 162 observations matched exactly, reflecting 85.3% agreement.

The number of applications proposing to use each measure was calculated separately by scientific domain and scientific area. For the scientific domains, we focused on the six most frequently assessed in SAA research: (1) diagnostic interviews (structured and semi-structured) for substance use disorders and psychopathology, (2) scales to assess dimensions of substance use problems and disorders, (3) scales to assess dimensions of affect and psychopathology, (4) measures of substance use quantity and frequency, (5) measures of personality traits, and (6) assessments of cognitive/neurologic ability. There were four scientific areas defined by administrative branch or program unit: epidemiology at NIDA (n=28) and NIAAA (n=11), prevention at NIDA (n=23) and NIAAA (n=19), services at NIDA (n=17) and NIAAA (n=5), and treatment at NIDA (n=22) and NIAAA (n=16).

The use of common measures was then examined within domain and area. Density of measurement commonality was calculated and used as a unit of comparison across domain/area. To calculate this value, the number of distinct measures was divided by the frequency of uses of measures by applications, then standardized on a scale from 0–1, with 1 being highly common ( $1 - [(total\ n\ distinct\ measures)/(n\ uses\ of\ all\ measures)]$ ).

### 3. RESULTS

A total of 548 distinct measures were cited across the 141 grant applications. This included a broad range of measures such as the Structured Clinical Interview for DSM Disorders (SCID; First et al., 1996; Spitzer et al., 1992), Fagerstrom Test for Nicotine Dependence (FTND; Heatherton et al. 1991), the Multidimensional Personality Questionnaire (MPQ; Tellegen and Waller, 2008), the Beck Depression Inventory (BDI; Beck et al., 1961), and the Trail Making Test, parts A and B (TMT; Reitan, 1994). Figure 1 shows the top 10 most common measures in each of the six most commonly assessed domains. The domain of diagnostic interviews of substance use disorders and psychopathology had the greatest commonality (108 uses of 13 unique measures, density of 0.88), followed by the domain of scales to assess dimensions substance use problems and disorders (136 uses of 41 unique measures, density of 0.70). More moderate degrees of commonality were observed for the domain of scales to assess dimensions of affect and psychopathology (190 uses of 59 unique measures, density of 0.69) and the domain of measures of substance use quantity and frequency (124 uses of 47 unique measures, density of 0.62). Commonality was more limited in the personality domain (40 uses of 24 unique measures, density of 0.40) and the cognitive/neurologic domain (67 uses of 52 unique measures, density of 0.22).

Similarly, there was limited commonality within each area of science (density: 0.32–0.42). The areas of prevention (251 uses of 149 unique measures, density of 0.41) and treatment (429 uses of 250 unique measures, density of 0.42) had greater commonality than epidemiology (343 uses of 221 unique measures, density of 0.36) and services (165 uses of 113 unique measures, density of 0.32). Additional post-hoc analyses revealed high levels of commonality (density: 0.70–0.81) *within* certain domains within these scientific areas. In the

areas of epidemiology (0.81), prevention (0.70), and treatment (0.70), the diagnostic measures of substance use disorders or psychopathology had greater commonality compared to the other domains. In the services area, the substance use quantity and frequency domain had the greatest commonality (0.60).

#### 4. DISCUSSION

The analysis revealed a dearth of measure commonality across multiple related areas of SAA science. Across the areas of epidemiology, prevention, treatment, and services research funded by NIDA or NIAAA, the use of common measures was rare. An exception was observed for the assessment of substance use and psychiatric disorders, perhaps because of the limited number of measures to choose from in this domain. Additional analyses (results not shown) replicated this general lack of measure commonality in other areas of SAA (e.g., clinical neuroscience), further suggesting its pervasiveness. This lack of commonality is not without consequences. The absence of common measures in SAA impedes cross-study comparability, replication, and large-scale data integration, the latter of which is paramount to advancing high-priority areas that require very large sample sizes, including addiction genomics (Ripke et al., 2013).

Our findings should be considered in light of several study limitations. First, the data were drawn from applications of funded grants, and there was no attempt to contact investigators to inquire about the use of measures not mentioned in the applications and/or changes in the use of proposed measures following funding. Second, measures used by studies not funded by NIDA or NIAAA were not captured. Third, the sample comprised NIH grants across related areas of addiction science (epidemiology, treatment, services, and prevention), suggesting a potential for overestimating measurement commonality in the addiction field. Fourth, the classification of measures into domains was somewhat crude, as some measures (e.g., Self-Ratings of the Effects of Alcohol) do not fit neatly into one of the broad domains and, within a domain, serve different purposes (e.g., unidimensional versus multidimensional assessments of depression) or target different developmental groups (e.g., children vs. adults). Finally, the definition of measure commonality focused on similarity of *measures* rather than items/questions within measures; this method would have missed occasions of commonality when studies using different measures used some of the same items/questions. Future work will assess and track item-level commonality, including the conditions under which it occurs (i.e., different studies using all items from the same measure, different studies using some items from the same measure, different studies using the same items albeit from different measures). This analytic approach would directly assess and monitor item-level commonality, and build upon findings from the current paper. Overall, these limitations suggest that caution be taken when generalizing and interpreting the results.

To address the lack of measure commonality documented herein, NIDA and its scientific partners launched an effort that provides investigators with, via the PhenX Toolkit (<https://www.phenxtoolkit.org/>), a common set of tools and resources that allow their work to span the manifold of SAA areas. As described elsewhere (Hamilton et al., 2011a, 2011b; Maiese et al., 2013), the PhenX Toolkit is a catalog of well-established, standard measures of

phenotypes and exposures recommended for human-subjects studies across 21 fields of research, and its development followed a rigorous consensus process led by domain experts and vetted by the scientific community. To expand the depth and breadth of measures in the PhenX Toolkit, the SAA project was launched in early 2011, with support from NIDA, to identify a “Core” collection and six “Specialty” collections recommended for human-subjects studies in the addiction sciences (see <http://grants.nih.gov/grants/guide/notice-files/NOT-DA-12-008.html>). A total of 43 new measures in support of SAA research were added to the Toolkit in February 2012 (Table 1).

Importantly, we note that pursuit of measure commonality is not to replace specific or specialized measures crucial to an investigator’s core hypothesis, but to create a common item set across studies to allow for the comparing and pooling of data. The PhenX Toolkit and other measure-commonality and data-harmonization efforts and tools will not only increase research productivity, but will also foment scientific discovery within and beyond the addiction sciences. There are several other existing tools and initiatives aimed at data-harmonization and measure-commonality, located in the e-supplementary materials. Collectively, these efforts reflect an embracing of a “big data” culture of large-scale data integration, and underscore a growing emphasis on sharing and creating opportunities to create multi-study linkages. Hand-in-hand with increasing use of common measures is the need to develop effective mechanisms for data sharing. To date at NIH, a number of data-sharing efforts have been initiated, which have been included in the Supplementary Materials<sup>1</sup>.

The NIH will continue to play a central role in fostering a culture of cooperation, data sharing, and collaboration. In this sense, measure commonality must be viewed as an essential component of a larger strategy to marshal investigators to more effectively reap the benefits of the considerable Federal investment in biomedical and health research to advance science and improve the nation’s health.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Role of Funding Source

Funding for the PhenX Toolkit was primarily provided by the National Human Genome Research Institute (NHGRI) of the National Institutes of Health (NIH) Cooperative Agreement U01-HG004597 with supplements from the Office of Behavioral and Social Sciences Research (OBSSR) and the National Institute on Drug Abuse (NIDA). The views and opinions expressed in this report are those of the authors and should not be construed to represent the views of NIDA, NHGRI, NIAAA, or any of the sponsoring organizations, agencies, or the U.S. Government.

The authors thank Dr. Marcia S. Scott of NIAAA, the PhenX Steering Committee (<https://www.phenx.org/Default.aspx?tabid=57>), the Substance Abuse Addiction Scientific Panel (<https://www.phenx.org/Default.aspx?tabid=689>), the NIH liaisons (<https://www.phenx.org/Default.aspx?tabid=702>), and the Working Group members (<https://www.phenxtoolkit.org/index.php?pageLink=saa.wg123>) for their many contributions.

---

<sup>1</sup>Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

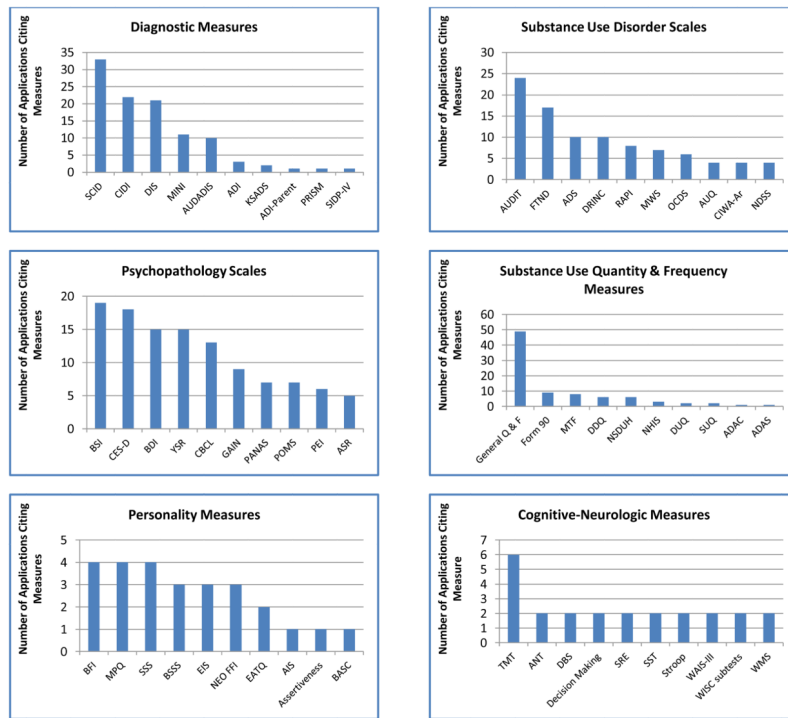


## References

- Agrawal A, Freedman ND, Cheng YC, Lin P, Shaffer JR, Sun Q, Taylor K, Yaspan B, Cole JW, Cornelis MC, DeSensi RS, Fitzpatrick A, Heiss G, Kang JH, O'Connell J, Bennett S, Bookman E, Bucholz KK, Caporaso N, Crout R, Dick DM, Edenberg HJ, Goate A, Hesselbrock V, Kittner S, Kramer J, Nurnberger JI Jr, Qi L, Rice JP, Schuckit M, van Dam RM, Boerwinkle E, Hu F, Levy S, Marazita M, Mitchell BD, Pasquale LR, Bierut LJ. Measuring alcohol consumption for genomic meta-analyses of alcohol intake: opportunities and challenges. *Am J Clin Nutr.* 2012; 95:539–547. [PubMed: 22301922]
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry.* 1961; 4:561–571. [PubMed: 13688369]
- Bennett SN, Caporaso N, Fitzpatrick AL, Agrawal A, Barnes K, Boyd HA, Cornelis MC, Hansel NN, Heiss G, Heit JA, Kang JH, Kittner SJ, Kraft P, Lowe W, Marazita ML, Monroe KR, Pasquale LR, Ramos EM, van Dam RM, Udren J, Williams K. Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience. *Genet Epidemiol.* 2011; 35:159–173. [PubMed: 21284036]
- Bierut LJ. Genetic vulnerability and susceptibility to substance dependence. *Neuron.* 2011; 69:618–627. [PubMed: 21338875]
- Cornelis MC, Agrawal A, Cole JW, Hansel NN, Barnes KC, Beaty TH, Bennett SN, Bierut LJ, Boerwinkle E, Doheny KF, Feenstra B, Feingold E, Fornage M, Haiman CA, Harris EL, Hayes MG, Heit JA, Hu FB, Kang JH, Laurie CC, Ling H, Manolio TA, Marazita ML, Mathias RA, Mirel DB, Paschall J, Pasquale LR, Pugh EW, Rice JP, Udren J, van Dam RM, Wang X, Wiggs JL, Williams K, Yu K. The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Genet Epidemiol.* 2010; 34:364–372. [PubMed: 20091798]
- Maiese, DR.; Hendershot, TP.; Strader, LC.; Wagener, DK.; Hammond, JA.; Huggins, W.; Kwok, RK.; Hancock, DB.; Whitehead, NS.; Nettles, DS.; Pratt, JG.; Scott, MS.; Conway, KP.; Junkins, HA.; Ramos, EM.; Hamilton, CM. RTI Press Publication No. MR-0027-1310. RTI Press; Research Triangle Park, NC: 2013. PhenX—Establishing a Consensus Process to Select Common Measures for Collaborative Research.
- Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry.* 2011; 168:1041–1049. [PubMed: 21890791]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. Structured Clinical Interview for the DSM-IV Axis I Disorders. American Psychiatric Press, Inc; Washington, DC: 1996.
- Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, Hammond JA, Huggins W, Jackman D, Pan H, Nettles DS, Beaty TH, Farrer LA, Kraft P, Marazita ML, Ordovas JM, Pato CN, Spitz MR, Wagener D, Williams M, Junkins HA, Harlan WR, Ramos EM, Haines J. Hamilton et al. Respond to “Consolidating Data Harmonization”. *Am J Epidemiol.* 2011b; 174:265–266.
- Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, Hammond JA, Huggins W, Jackman D, Pan H, Nettles DS, Beaty TH, Farrer LA, Kraft P, Marazita ML, Ordovas JM, Pato CN, Spitz MR, Wagener D, Williams M, Junkins HA, Harlan WR, Ramos EM, Haines J. The PhenX Toolkit: get the most from your measures. *Am J Epidemiol.* 2011a; 174:253–260. [PubMed: 21749974]
- Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA, McEvoy BP, Schrage AJ, Grant JD, Chou YL, Zhu R, Henders AK, Medland SE, Gordon SD, Nelson EC, Agrawal A, Nyholt DR, Bucholz KK, Madden PA, Montgomery GW. A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry.* 2011; 70:513–518. [PubMed: 21529783]
- Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict.* 1991; 86:1119–1127. [PubMed: 1932883]
- Hussong AM, Curran PJ, Bauer DJ. Integrative data analysis in clinical psychology research. *Annu Rev Clin Psychol.* 2013; 9:61–89. [PubMed: 23394226]

- Office of Extramural Research at the National Institutes of Health. NIH Research Project Grant Program (R01). US Department of Health and Human Services; 2010. 3-9-2012 <http://grants.nih.gov/grants/funding/r01.htm> [accessed on April 28, 2014]
- Miyazaki Y, Raudenbush SW. Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychol Methods*. 2000; 5:44–63. [PubMed: 10937322]
- Reitan, R. Halstead-Reitan Neuropsychological Battery. Reitan Neuropsychology Laboratories; Tempe, AZ: 1994. Trail Making Test.
- Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, Sanchez N, Stahl EA, Williams S, Wray NR, Xia K, Bettella F, Borglum AD, Bulik-Sullivan BK, Cormican P, Craddock N, de Leeuw C, Durmishi N, Gill M, Golimbet V, Hamshere ML, Holmans P, Hougaard DM, Kendler KS, Lin K, Morris DW, Mors O, Mortensen PB, Neale BM, O'Neill FA, Owen MJ, Milovancevic MP, Posthuma D, Powell J, Richards AL, Riley BP, Ruderfer D, Rujescu D, Sigurdsson E, Silagadze T, Smit AB, Stefansson H, Steinberg S, Suvisaari J, Tosato S, Verhage M, Walters JT, Levinson DF, Gejman PV, Kendler KS, Laurent C, Mowry BJ, O'Donovan MC, Owen MJ, Pulver AE, Riley BP, Schwab SG, Wildenauer DB, Dudbridge F, Holmans P, Shi J, Albus M, Alexander M, Campion D, Cohen D, Dikeos D, Duan J, Eichhammer P, Godard S, Hansen M, Lerer FB, Liang KY, Maier W, Mallet J, Nertney DA, Nestadt G, Norton N, O'Neill FA, Papadimitriou GN, Ribble R, Sanders AR, Silverman JM, Walsh D, Williams NM, Wormley B, Arranz MJ, Bakker S, Bender S, Bramon E, Collier D, Crespo-Facorro B, Hall J, Iyegbe C, Jablensky A, Kahn RS, Kalaydjieva L, Lawrie S, Lewis CM, Lin K, Linszen DH, Mata I, McIntosh A, Murray RM, Ophoff RA, Powell J, Rujescu D, Van Os J, Walshe M, Weisbrod M, Wiersma D, Donnelly P, Barroso I, Blackwell JM, Bramon E, Brown MA, Casas JP, Corvin AP, Deloukas P, Duncanson A, Jankowski J, Markus HS, Mathew CG, Palmer CN, Plomin R, Rautanen A, Sawcer SJ, Trembath RC, Viswanathan AC, Wood NW, Spencer CC, Band G, Bellenguez C, Freeman C, Hellenthal G, Giannoulatou E, Pirinen M, Pearson RD, Strange A, Su Z, Vukcevic D, Donnelly P, Langford C, Hunt SE, Edkins S, Gwilliam R, Blackburn H, Bumpstead SJ, Dronov S, Gillman M, Gray E, Hammond N, Jayakumar A, McCann OT, Liddle J, Potter SC, Ravindrarajah R, Ricketts M, Tashakkori-Ghanbaria A, Waller MJ, Weston P, Widaa S, Whittaker P, Barroso I, Deloukas P, Mathew CG, Blackwell JM, Brown MA, Corvin AP, McCarthy MI, Spencer CC, Bramon E, Corvin AP, O'Donovan MC, Stefansson K, Scolnick E, Purcell S, McCarroll SA, Sklar P, Hultman CM, Sullivan PF. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*. 2013; 45:1150–1159. [PubMed: 23974872]
- Sher KJ, Dick DM, Crabbe JC, Hutchison KE, O'Malley SS, Heath AC. Consilient research approaches in studying gene x environment interactions in alcohol research. *Addict Biol*. 2010; 15:200–216. [PubMed: 20148780]
- Spitzer RL, Williams JW, Gibbon M, First MB. The structured clinical interview FOR DSM-III-R (SCID): I: history, rationale, and description. *Arch Gen Psychiatry*. 1992; 49:624–629. [PubMed: 1637252]
- Tellegen, A.; Waller, NG. Exploring personality through test construction: development of the Multidimensional Personality Questionnaire. In: Boyle, GJ.; Matthews, G.; Saklofske, DH., editors. *The SAGE Handbook of Personality Theory and Assessment, Vol 2: Personality Measurement and Testing*. Sage Publications, Inc; Thousand Oaks, CA, US: 2008. p. 261-292.





**Figure 1.**

Top 10 Most Common Measures in Six Frequently Assessed Domains

**Diagnostic Measures:** SCID = Structured Clinical Interview for DSM-IV; CIDI = University of Michigan Composite International Diagnostic; DIS = Diagnostic Interview Schedule – based instruments; MINI = Mini-International Neuropsychiatric Interview; AUDADIS = Alcohol Use Disorder and Associated Disabilities Interview Schedule; ADI = Adolescent Diagnostic Interview; KSADS = Kiddie Schedule for Affective Disorders and Schizophrenia; ADI-Parent = Adolescent Diagnostic Interview-Parent; PRISM = Psychiatric Research Interview for Substance and Mental Disorders; SIDP-IV Structured Interview for DSM-IV Personality.

**Substance Use Disorder Scales:** AUDIT = Alcohol Use Disorders Identification Test; FTND = Fagerström Test for Nicotine Dependence; ADS = Alcohol Dependence Scale; DRINC = Drinker Inventory of Consequences; RAPI = Rutgers Alcohol Problem Index; MWS = Minnesota Withdrawal Scale; OCD5 = Obsessive Compulsive Drinking Scale; AUQ = Alcohol Urge Questionnaire; CIWA-Ar = Clinical Institute Withdrawal Assessment for Alcohol – Revised; NDSS = Nicotine Dependence Syndrome Scale.

**Psychopathology Scales:** BSI = Brief Symptom Inventory; CES-D = Center for Epidemiologic Studies Depression Scale; BDI = Beck Depression Inventory; YSR = Youth Self-Report of the Child Behavior Checklist; CBCL = Child Behavior Checklist; GAIN = Global Appraisal of Individual Needs; PANAS = Positive and Negative Affect Schedule; POMS = Profile of Mood States; PEI = Personal Experiences Inventory; ASR = Adult Self-Report, Adult behavior checklist.

**Substance Use Quantity and Frequency Measures:** General Q and F = General Quantity and Frequency Questions; Form 90 = Form 90 from Project Match; MTF = Monitoring the Future Quantity and Frequency of Substance Use Questions; DDQ = Daily Drinking

Questionnaire; NSDUH = National Survey of Drug Use and Health Quantity and Frequency of Substance Use Questions; NHIS = National Health Interview Survey; DUQ = Drug Use Questionnaire; SUQ = Substance Use Questionnaire; ADAC = Adapted Daily Alcohol consumption; ADAS = American Drug and Alcohol Survey

**Personality Measures:** BFI = Big Five Inventory; MPQ = Multidimensional Personality Questionnaire; SSS = Sensation Seeking Scale (Zuckerman); BSSS = Brief Sensation Seeking Scale; EIS = Eysenck Impulsivity Scale; NEO FFI = NEO Five Factor Inventory; EATQ = Early Adolescent Temperament Questionnaire – Revised; AIS = Avoidance and Inflexibility Scale; Assertiveness = Assertiveness (Gambrill and Richey); BASC = Behavioral Assessment System for Children.

**Cognitive-Neurologic Measures:** TMT = Trail Making Test A and B; ANT = Attention Network Task; DBS = Decisional Balance Scale; Decision Making = Decision making (5-item scale developed by Botvin, Baker); SRE = Self-Ratings of the Effects of Alcohol; SST = Stop-signal task; Stroop = Stroop Color and Word Test; WAIS-III = Wechsler Adult Intelligence Scale III Comprehension Subscale; WISC subtests = Wechsler Intelligence Scale for Children subtests; WMS = Wechsler Memory Scale.

**Table 1**

## Substance Abuse and Addiction Collection in the PhenX Toolkit

<a href="https://www.phenxtoolkit.org/index.php?pageLink=browse.conceptualgroupsandid=2169andbreadcrumbs=2169">https://www.phenxtoolkit.org/index.php?pageLink=browse.conceptualgroupsandid=2169andbreadcrumbs=2169</a>		
	Measure Count	Measures
<b>Core</b>		
Tier 1	19	Current Age; Ethnicity; Gender; Race; Current Educational Attainment; Current Employment Status; Current Marital Status; Body Mass Index; Current Quality of Life; Self-report of Human Immunodeficiency Virus (HIV) Testing; Tobacco - Smoking Status; Tobacco - 30-day Quantity and Frequency; Tobacco - Age of Initiation of Use; Alcohol - Lifetime Use; Alcohol - 30 Day Quantity and Frequency; Alcohol - Age at First Use; Substances - Lifetime Use; Substances - 30 Day Frequency; Substances - Age at First Use
Tier 2	8	Annual Family Income; Child-Reported Parental Educational Attainment; Household Roster-Relationships; Occupation/Occupational History; Family History of Substance Use Problems; Internalizing, Externalizing, and Substance Use Disorders Screener; Peer/Partner Substance Use and Tolerance of Substance Use; Social Networks
<b>Specialty Areas</b>		
Assessment of Substance Use and Substance Use Disorders	7	Alcohol Breathalyzer; Family History of Substance Use Problems; Lab-based Urine Drug Analysis; On-site Urine Drug Screen; Patterns of Substance Use; Screening and Severity of Substance Use Problems; Substance Abuse and Dependence - Past Year
Substance Use-related Co-morbidities and Health-related Outcomes	8	Behavioral Health Screener; Comprehensive HIV Risk Assessment; Duration of Abstinence; Health Quality Index for Economic Evaluations; Self-Help Involvement; Service Utilization and Cost; Substance Abuse Treatment, Participation, and Involvement; Victimization Screener
Substance Use-related Community Factors	8	Cigarette Price; Disapproval of Substance Use; Perceived Availability of Illegal Drugs and Alcohol; Perceived Harm of Substance Use; Point of Sale Environment for Alcohol and Tobacco; Substance Abuse Treatment Quality; Treatment Availability and Level of Care; Unmet Need for Alcohol or Drug Treatment
Substance Use-related Neurobehavioral and Cognitive Risk Factors	7	Cognitive Flexibility (Dimensional Change Card Sort); Decision Making (Iowa Gambling Task); Delayed Reward Discounting (Monetary Choice Questionnaire); Inhibitory Control (Stop Signal Paradigm); Laboratory Test of Risk Taking (Balloon Analogue Risk Task); Motor and Attentional Impulsivity (Immediate and Delayed Memory Task); Response Inhibition (Go/NoGo Task)
Substance Use-related Psychosocial Risk Factors	6	Community Risk and Protective Factors; Externalizing Problems; Family Risk and Protective Factors; Peer/Partner Substance Use and Tolerance of Substance Use; School Risk and Protective Factors; Sensitivity to Reward and Punishment; Temperament
Substance-specific Intermediate Phenotypes	5	Acute Subjective Response to Substances - Current; Acute Subjective Responses to Substances - Retrospective; Expectancies - Alcohol, Tobacco, and Other Substances; Motives

<a href="https://www.phenxtoolkit.org/index.php?pageLink=browse.conceptualgroupsandid=2169andbreadcrumbs=2169">https://www.phenxtoolkit.org/index.php?pageLink=browse.conceptualgroupsandid=2169andbreadcrumbs=2169</a>		
	Measure Count	Measures
		- Alcohol, Tobacco, and Other Substances; Self-reported Craving