

Correspondence

Psychological Medicine, 41 (2011).
doi:10.1017/S0033291711000717
First published online 6 May 2011

Letter to the Editor

The proposed 2/11 symptom algorithm for DSM-5 substance-use disorders is too lenient

Substance-use disorder (SUD) diagnoses are critically important for research and clinical practice. Unlike the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV), the draft diagnostic criteria for SUDs in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) define a single SUD for various substance classes – such as ‘alcohol-use disorder’ (AUD) – based upon 11 symptoms. This criterion set is composed of all seven DSM-IV substance dependence symptoms, three of the four DSM-IV substance-abuse symptoms (all but legal problems) and a craving symptom. AUD and other SUDs would be diagnosed if a person has at least 2/11 co-occurring symptoms. The 2/11 threshold was chosen, in large part, to emulate rates of any DSM-IV SUD (abuse plus dependence) (<http://www.DSM5.org>).

We believe that the 2/11 algorithm is too lenient and creates serious multiple problems. The proposed algorithm would diagnose many whose substance involvement has questionable clinical significance, leading SUD diagnosis away from mainstream neuro-behavioral theory regarding what constitutes a mental ‘disorder’ and ‘addiction’. Further, the algorithm allows so much heterogeneity that the clinical and research utility of the diagnostic category would be greatly compromised. We illustrate these points with past-year AUD symptom data from 29993 lifetime drinkers aged 21 years and above, from wave 2 of the National Epidemiological Survey of Alcohol Use and Related Conditions (NESARC; Grant *et al.* 2003) (analytic details available upon request). Unlike wave 1, wave 2 had a craving item allowing us to study proposed DSM-5 diagnoses with respect to AUD prevalence, severity and heterogeneity.

Prevalence

We used NESARC to estimate the past-year prevalence for DSM-IV alcohol dependence (5.02%) and any DSM-IV AUD (abuse plus dependence; 11.06%). Results for the proposed DSM-5 criterion set showed that a 2/11 threshold produced a past-year AUD rate of 12.40%, which dropped to 7.25% with a 3/11

threshold, 4.48% with a 4/11 threshold, and 2.93% with a 5/11 threshold. Although the prevalence estimates produced by the 2/11 and 3/11 thresholds were closest to that for any DSM-IV AUD, we do not believe that DSM-5 should be designed to emulate DSM-IV in this way. There are serious problems with the validity of the DSM-IV substance-abuse category, including over-diagnosis (Martin *et al.* 2008; Vergés *et al.* 2010), such that it should not be part of a target for SUD prevalence in DSM-5. DSM-IV alcohol dependence, a category with considerably more construct validity, represents a more reasonable comparator and is closest in prevalence to a 4/11 threshold.

Severity

Problems with the 2/11 algorithm must be understood in the context of the DSM-5 draft criteria, some of which are mild, not necessarily indicative of pathology, or are commonly misunderstood and over-endorsed, particularly among youth. Tolerance can present without significant harm, and alcohol tolerance is normative in adolescence and young adulthood (O’Neill & Sher, 2000). The symptom of hazardous use, usually given due to intoxicated driving, can reflect simple heedlessness rather than disorder. Although not problematic conceptually, the symptoms of using more or longer than intended (larger/longer) and persistent desire or repeated attempts to quit or cut down (quit/cut down) have proven difficult to operationalize, and standard wording is frequently misunderstood. These criteria are often incorrectly endorsed by drinkers for social reasons (e.g. conformity) rather than compulsion-based reasons (Chung & Martin, 2005; Caetano & Babor, 2006). Given the nature of these four symptoms, a 2/11 threshold means that many diagnosed cases may have mild levels of substance involvement and no meaningful pathology.

We examined past-year alcohol use in NESARC for those with exactly two, three, four and five past-year DSM-5 AUD symptoms. The number of standard drinks usually consumed on drinking days averaged a modest 3.79 (s.e. = 0.09) for those with two symptoms, and was 4.58 (s.e. = 0.19), 4.64 (s.e. = 0.16) and 5.33 (s.e. = 0.25) for those with three, four and five symptoms, respectively. Drinking five or more standard drinks per occasion at least weekly was reported by only 28.33% of the two-symptom group, but this percentage rose to 38.08, 47.22 and 54.14% among those with three, four and five symptoms, respectively.

These data suggest that the 2/11 algorithm would diagnose very mild cases that do not have the compulsive patterns of use that are thought to characterize 'addiction' in modern theory (Koob & Le Moal, 1997; Robinson & Berridge, 2003). This would be inconsistent with the traditional DSM definition of mental disorders as syndromes that reflect dysfunction in internal mechanisms (Wakefield & First, 2003), such as neuroadaptations that 'hijack' the brain's reward and incentive salience systems, leading to substance use.

The 2/11 algorithm allows tremendous variation in severity, with a more than five-fold difference in the number of symptoms a diagnosed case can have. This variation in severity is addressed, in part, by the draft criteria: the presence of 4/11 symptoms would define 'severe' rather than 'moderate' SUD. However, some persons with four symptoms can have a heavy and heedless substance-use pattern rather than a presentation that clearly reflects compulsive use, raising questions as to the meaning of 'severe'. The variability in severity among those with SUD, and even among those with 'severe' SUD, would greatly limit the utility of these categories.

Heterogeneity

The 2/11 algorithm allows a remarkable degree of heterogeneity for something described as a mental disorder, even among those with a similar number of symptoms. Although the polythetic nature of substance problems has long been recognized, the degree of variability in the symptom profile that the proposed algorithm allows makes it hard to recognize any sort of prototypic syndrome whatsoever. Among 1486 persons in NESARC with exactly two past-year DSM-5 symptoms, we observed 41 of the 55 two-symptom configurations that are possible with 11 items. No configuration characterized more than a fraction of this group. Among this two-symptom group, the most frequently endorsed individual criteria were larger/longer (55.30%), quit/cut down (39.10%), hazardous use (36.77%) and tolerance (24.81%). The 15 most common two-symptom configurations in this group (accounting for 88.8% of the total) contained at least one of these four criteria. Yet tolerance is mild and can present without problems, hazardous use can reflect heedless behavior rather than pathology, and larger/longer and quit/cut down are often misunderstood and over-endorsed.

We also characterized heterogeneity using the pairwise average number of shared symptoms. Those with exactly two symptoms shared, on average, only 0.70 symptoms (s.d. = 0.62). The average number of shared symptoms almost doubled to 1.36 (s.d. = 0.72) among those with exactly three symptoms, and rose to

2.14 (s.d. = 0.79) and 2.98 (s.d. = 0.84) among those with exactly four and five symptoms, respectively. Cases with two or more symptoms shared an average of 1.58 symptoms (s.d. = 1.19), but this number increased to 2.52 (s.d. = 1.37) in the 3+ group, 3.51 (s.d. = 1.48) in the 4+ group, and 4.51 (s.d. = 1.55) in the 5+ group. These results indicate that increasing the threshold for diagnosis with the DSM-5 draft criteria would substantially reduce the degree of heterogeneity among diagnosed cases.

Recommendations

If the diagnostic threshold of 2/11 symptoms is too lenient, what should the new threshold be? If all 11 draft criteria will be used in DSM-5, it is reasonable to consider diagnostic cut-points ranging from three to six symptoms. A 3/11 threshold would suffer from many of the limitations discussed herein, while a 6/11 threshold would probably prove to be too strict. However, we do not believe that the proposed 11-item criterion set is optimal. A preferable way to increase the diagnostic threshold for SUDs is to remove especially mild, poorly performing and theoretically compromised symptoms from the criterion set. Research is needed to contrast diagnostic algorithms with various permutations of criteria and thresholds, using relevant external criteria.

Because SUD symptom count shows a continuous and roughly linear association with other measures of substance involvement, a diagnostic threshold will not be determined by finding clear discontinuities in the association of symptom count with external validators. However, this does not mean that the choice of diagnostic threshold is arbitrary. Instead, the threshold should be explicitly designed to avoid what can be considered false-positive diagnostic assignments, and to reflect professional judgments of the importance of diagnosing substance problems at a certain level of severity and the costs of not doing so (Krueger *et al.* 2004). Choosing the threshold for an SUD diagnosis involves weighing the relative importance and cost of type I and type II errors. When a threshold is higher, there can be concerns about the availability of treatment for those who are subthreshold for a diagnosis, especially with regards to insurance coverage. However, a low threshold can lead to a highly inefficient allocation of scarce and expensive healthcare resources, and stigmatization of individuals lacking meaningful pathology (Room, 2006).

One partial solution to the issue of type II 'errors' is to think differently about subthreshold substance problems, which are an important public health issue. Subdiagnostic substance problems are key targets for prevention and intervention efforts, which are worthy

of financial and policy support. In this way, the field of addiction medicine can protect the critical concept of 'disorder' while broadening efforts to help those with substance problems.

Acknowledgments

Preparation of this letter was supported by National Institute on Alcohol Abuse and Alcoholism grants no. R01 AA13397 (to C.S.M.), no. K25 AA017456 (to D.L.S.) and no. R01 AA13987, no. R37 AA07231, no. T32 AA13526 and no. K05 AA017242 (to K.J.S.).

Declaration of Interest

None.

References

- Caetano R, Babor TF** (2006). Diagnosis of alcohol dependence in epidemiological surveys: an epidemic of youthful alcohol dependence or a case of measurement error? *Addiction* **101**, 111–114.
- Chung T, Martin CS** (2005). What were they thinking? Adolescents' interpretations of DSM-IV alcohol dependence symptom queries and implications for diagnostic validity. *Drug and Alcohol Dependence* **80**, 191–200.
- Grant BF, Moore TC, Kaplan KD** (2003). *Source and Accuracy Statement: Wave 1 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)*. National Institute on Alcohol Abuse and Alcoholism: Bethesda, MD.
- Koob GF, Le Moal M** (1997). Drug abuse: hedonic homeostatic dysfunction. *Science* **278**, 52–58.
- Krueger RF, Nichol PE, Hicks BM, Markon KE, Patrick CJ, Iacono WG, McGue M** (2004). Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychological Assessment* **16**, 107–119.
- Martin CS, Chung T, Langenbucher JW** (2008). How should we revise diagnostic criteria for substance use disorders in DSM-V? *Journal of Abnormal Psychology* **117**, 561–575.
- O'Neill S, Sher KJ** (2000). Tolerance and withdrawal during early adulthood: a longitudinal perspective. *Experimental and Clinical Psychopharmacology* **8**, 493–508.
- Robinson TE, Berridge KC** (2003). Addiction. *Annual Review of Psychology* **54**, 25–53.
- Room R** (2006). Taking account of cultural and societal influences on substance use diagnoses and criteria. *Addiction* **101** (Suppl. 1), 31–39.
- Vergés A, Steinley D, Trull TJ, Sher KJ** (2010). It's the algorithm! Why differential rates of chronicity and comorbidity are not evidence for the validity of the abuse–dependence distinction. *Journal of Abnormal Psychology* **119**, 650–661.
- Wakefield JC, First MB** (2003). Clarifying the distinction between disorder and nondisorder: confronting the overdiagnosis problem in DSM-V. In *Advancing DSM: Dilemmas in Psychiatric Diagnosis* (ed. K. A. Phillips,

M. B. First and H. A. Pincus), pp. 23–56. American Psychiatric Association: Washington, DC.

CHRISTOPHER S. MARTIN¹, DOUGLAS L. STEINLEY², ALVARO VERGÉS², KENNETH J. SHER²

¹ Department of Psychiatry, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

² Department of Psychological Sciences, University of Missouri-Columbia and Midwest Alcoholism Research Center, Columbia, MO, USA

Address for correspondence:

C. S. Martin

Department of Psychiatry,

University of Pittsburgh Medical Center,

3811 O'Hara Street, Pittsburgh, PA 15213, USA.

(Email: martincs@upmc.edu)

Psychological Medicine, **41** (2011).

doi:10.1017/S0033291711000729

First published online 13 May 2011

Letter to the Editor

Adiposity as a possible mediator of low testosterone salivary levels in adolescent boys in prodromal stages of psychosis

We have read with interest the article entitled 'Neuroendocrine markers of high risk of psychosis: salivary testosterone in adolescent boys with prodromal symptoms' by van Rijn *et al.* (2011), published in *Psychological Medicine*. The authors concluded that levels of testosterone were significantly lower in adolescents with prodromal symptoms as compared with non-clinical controls and that there were no statistically significant differences in oestradiol between groups.

A possible mechanism linking low testosterone and psychosis is overweight/obesity, which was not reported in the study. In fact, a robust association between the metabolic syndrome and low testosterone levels has been found (Kupelian *et al.* 2008), with hypogonadism being associated with obesity and type 2 diabetes mellitus (Dandona & Rosenberg, 2010). The relationship between schizophrenia and metabolic abnormalities is also a well-replicated finding, being reported even in the early stages of the disease (Verma *et al.* 2009). A possible explanation for the relationship between obesity and reduced testosterone is an increase in the activity of the enzyme aromatase, present in adipose tissue (Loves *et al.* 2008). Aromatase is a member of the cytochrome P450 superfamily that catalyses the conversion of C-19 androgens (testosterone and androstenedione) into C-18 oestrogens (oestradiol and oestrone) (Williams, 2010). Therefore, it is possible to hypothesize that adipose